Universidade Federal de Uberlândia - UFU
Faculdade de Matemática - FAMAT
Coordenação dos Cursos de Bacharelado e Licenciatura em Matemática

Trabalho de Conclusão de Curso

Curvas Elípticas e Criptografia

Aluno: Tiago Aprigio Bezerra Meireles

Orientador: Prof. Dr. Alonso Sepúlveda Castellanos

Tiago Aprigio Bezerra Meireles

Curvas Elípticas e Criptografia

Trabalho apresentado à Faculdade de Matemática, como parte dos requisitos para obtenção do título de Bacharel em matemática

Universidade Federal de Uberlândia – UFU

Faculdade de Matemática

Orientador: Prof. Dr. Alonso Sepúlveda Castellanos

Uberlândia-MG

2020

Universidade Federal de Uberlândia - UFU
Faculdade de Matemática - FAMAT
Coordenação dos Cursos de Bacharelado e Licenciatura em Matemática

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Matemática.

Uberlândia, de dezembro de 2020.

BANCA EXAMINADORA

Prof. Dr. Alonso Sepúlveda Castellanos

Prof. Dr. Cícero Fernandes de Carvalho

Prof. Dr. Neiton Pereira da Silva

Uberlândia-MG
2020

Agradecimentos

Agradeço primeiramente à Deus, por ter me dado força, persistência, coragem e me abençoado muito durante toda a minha graduação.

Aos meus pais, Joaquim e Lourdes, que são minhas maiores referências aqui na Terra. Sem eles eu não teria conseguido chegar até aqui.

Aos meus irmãos, Daniel e Marcos, pela companhia e cumplicidade.

À minha namorada Julia, que foi um dos melhores presentes que ganhei da universidade. Ela me deu muito apoio, principalmente na reta final do curso.

À minha sogra, Solange, e ao meu sogro, Jõao, pelos cuidados que tiveram comigo.

Ao meu grande amigo Haniel Josué, pela grande amizade que temos a bastante tempo.

Ao meu orientador e amigo Alonso Sepúlveda, pelas orientações, conselhos de vida e pela amizade.

Aos professores da faculdade que tive o prazer de conhecer, em especial, aos professores, Cícero Fernandes, Elisa Regina, Marcio Fenille, Luciana Alves, Marcus Bronzi e Fábio Bertoloto.

Aos meus amigos do PET e do curso, Gabriela, Alef, Ana Laura, Dhara, Leonardo, Lucas Gomes, Fernanda, Gabriel, Aloísio e Lucas Santos.

"Mas, buscai primeiro o reino de Deus, e a sua justiça,
e todas estas coisas vos serão acrescentadas."
Mateus 6:33

Resumo

O objetivo deste trabalho é estudar algumas propriedades, resultados e aplicações criptográficas das curvas elípticas definidas sobre corpos finitos de característica diferente de 2 e 3. Mostramos de forma algébrica e geométrica que os pontos de uma curva elíptica formam um grupo abeliano. Além disso, foi mostrado o famoso Teorema de Hasse, que determina uma cota para o número de pontos de curvas elípticas sobre corpos finitos. Também, estudamos o problema do logaritmo discreto, cuja dificuldade de resolução torna os criptossistemas sobre curvas elípticas mais seguros.

Palavras-chave: Curvas Elípticas, Logaritmo discreto, Teorema de Hasse, Aplicações criptográficas.

Abstract

The objective of this work is to study some properties, results and cryptographic applications of the elliptic curves defined on finite fields of different characteristics of 2 and 3. We show in an algebraic and geometric way that the points of an elliptic curve form an abelian group. In addition, the famous Hasse's Theorem was shown, which determines a dimension for the number of points of elliptic curves on finite fields. We also studied the discrete logarithm problem, which, the difficulty in solving it, makes cryptosystems on elliptic curves safer.

Keywords: Elliptic Curves, Discrete Logarithm , Hasse's Theorem, Cryptographic Applications.

Sumário

Introdução

O estudo de curvas elípticas surgiu a partir dos problemas da teoria dos números. A teoria das equações Diofantinas é uma corrente da teoria que trata de soluções de equações polinomiais contidas nos números inteiros ou nos números racionais. Existem muitos problemas famosos em equações diofantinas. Um dos mais famosos problemas na história da matemática e talvez um dos que mais inspirou o desenvolvimento de novas teorias é o chamado Último Teorema de Fermat.

As curvas elípticas e suas propriedades foram estudadas em matemática de forma teórica por muito tempo, desde o segundo ou terceiro século a.C., e com o desenvolvimento tecnólogico suas aplicações tem sido de grande importância na área de criptografia, com o objetivo de trazer maior segurança na transmissão de dados.

Além disso, diversos métodos são utilizados no estudo de curvas elípticas e aparecem naturalmente em diversas áreas de matemática, de teoria dos números à análise, e de criptografia à física matemática. A demonstração do último teorema de Fermat [10], integrais e funções elípticas, fatoração de inteiros grandes [9], formas modulares e resolução da equação do pêndulo são exemplos dessa presença. Isso tudo e a facilidade de serem definidas, sem necessidade de muitos pré-requisitos, as tornam objetos muito interessantes a serem estudadas.

Vale salientar que curvas elípticas não são elipses, uma vez que elipses são seções cônicas e seções cônicas são dadas por equações do segundo grau. Estas curvas denominam-se elípticas porque surgem no estudo de uma classe específica de funções complexas chamadas funções elípticas.

Um dos problemas que encontramos na teoria de curvas elípticas é calcular seu número de pontos racionais sobre um corpo finito e encontrar sua estrutura de grupo. Neste sentido, temos um resultado que nos dá um limitante inferior e superior que é o famoso Teorema de Hasse (Teorema 4.10). Esse resultado foi primeiramente conjecturado por Emil Artin no ano de 1924 em sua tese e foi provado pelo alemão Helmut Hasse em 1933.

Através dos anos a criptografia vem sendo usada para enviar mensagens ocultas com o objetivo de serem lidas somente por pessoas autorizadas. Entre os problemas que a criptografia resolve, referentes à segurança em uma comunicação estão a privacidade, integridade, autenticidade e o não repúdio das mensagens.

Desde o descobrimento da criptografia de chave pública em 1975 por Diffie e Hellman [8], tem surgido uma variedade de grupos finitos cíclicos para a implementação destes

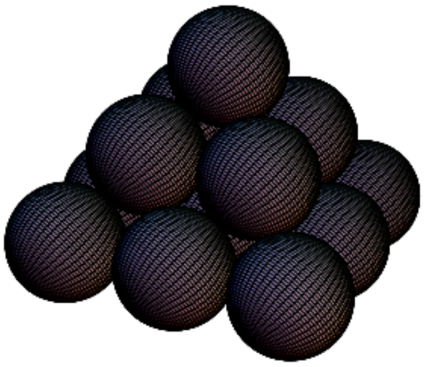
criptossistemas. Em 1985, Victor Miller e Neal Koblitz propuseram, independentemente, os criptossistemas baseados em Curvas Elípticas baseando sua segurança na intratabilidade do problema do Logaritmo Discreto no grupo de pontos da curva [2]. Desta forma, se amplia a visão para o uso de curvas algébricas como ferramentas úteis na construção de grupos cíclicos finitos.

No capítulo 1 são dadas algumas motivações para o estudo das curvas elípticas. No capítulo 2, foram definidos planos projetivos, equações de Weierstrass e a soma de pontos racionais de uma curva elíptica, e mostramos que com essa soma o conjunto dos pontos racionais forma um grupo. No capítulo 3, foram definidos Pontos de Torsão e foi feito um breve estudo do emparelhamento de Weil, pois, a partir disso, foi possível demosntrar o importante Teorema de Hasse no capítulo 4 . No capitulo 5, foi definido o Logaritmo Discreto, cuja dificuldade de cálculo está ligada com a segurança dos criptossistemas apresentados no capítulo 6.

Foi utilizado o software online MAGMA [11] para fazer várias contas deste trabalho.

1 Motivação para estudar Curvas Elípticas

Suponha que uma coleção de esferas está empilhada na forma de uma pirâmide quadrada com uma esfera na camada do topo, 2^2 na segunda camada, 3^2 na terceira camada, etc. Se a pilha desmoronar, é possível rearranjar as esferas para que fique na forma de um quadrado? (Ou seja, a quantidade de esferas é um quadrado perfeito?)



Se a pirâmide tem 3 camadas, então isso não pode ser feito, pois, haverá $1+4+9 = 14$ esferas, que não é um quadrado perfeito. Claro que, se tiver apenas uma esfera, formará uma pirâmide de altura 1 e também um quadrado um por um. Se não há esferas, temos uma pirâmide de altura 0 e um quadrado zero por zero. Além desses casos triviais, existem outros? Propomos a encontrar outro caso usando o método de Diophantus.

Se a pirâmide tem x camadas, então há

$$1^2 + 2^2 + 3^3 + \cdots + x^2 = \frac{x(x+1)(2x+1)}{6}$$

esferas. Queremos que essa expressão seja um quadrado perfeito, ou seja, queremos encontrar a solução de

$$y^2 = \frac{x(x+1)(2x+1)}{6}$$

para inteiros positivos x, y. Uma equação desse tipo representa uma **curva elíptica**. O gráfico em $\mathbb{R}^2$ é dado na figura 1.

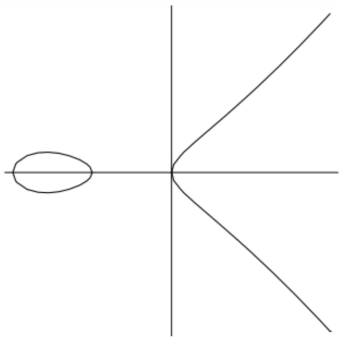


Figura 1 – $y^2 = \dfrac{x(x+1)(2x+1)}{6}$

O método de Diophantus usa os pontos que já conhecemos para produzir novos pontos. Vamos começar com os pontos $(0,0)$ e $(1,1)$. A reta que passa por esses dois pontos é $y = x$, e intersectando com a curva nos dá a equação

$$x^2 = \frac{x(x+1)(2x+1)}{6} = \frac{1}{3}x^3 + \frac{1}{2}x^2 + \frac{1}{6}x$$

E, portanto

$$x^3 - \frac{3}{2}x^2 + \frac{1}{2}x = 0.$$

Felizmente já conhecemos duas raízes dessa equação: $x = 0$ e $x = 1$. Isso é porque as raízes são as abscissas das interseções entre a reta e a curva. Antes de encontrarmos a terceira raiz, observe que

$$(x-a)(x-b)(x-c) = x^3 - (a+b+c)x^2 + (ab+ac+bc)x - abc.$$

Portanto, quando o coeficiente de x^3 é 1, o negativo do coeficiente de x^2 é a soma das raízes.

No nosso caso, temos as raízes $0, 1$, e x, então $0 + 1 + x = \dfrac{3}{2}$ Portanto, $x = 1/2$. Como a reta foi $y = x$, temos $y = 1/2$ também. É difícil ver o que isso significa em termos da pilha de esferas, mas pelo menos encontramos outro ponto sobre a curva. De fato, automaticamente temos mais um ponto, a saber, $(1/2, -1/2)$, por causa da simetria da curva em relação ao eixo x.

Vamos repetir o procedimento acima usando os pontos $(1/2, -1/2)$ e $(1,1)$. Por que usamos esses pontos? Estamos procurando por um ponto da interseção de algum lugar no primeiro quadrante, e a reta que passa por esses dois pontos parece ser a melhor escolha. A reta encontrada é $y = 3x - 2$. Intersectando com a curva nos dá

$$(3x-2)^2 = \frac{x(x+1)(2x+1)}{6}$$

Reorganizando a equação, temos $x^3 - \frac{51}{2}x^2 + \cdots = 0$. Já conhecemos as raízes $1/2$ e 1, então obtemos $\frac{1}{2} + 1 + x = \frac{51}{2}$, Portanto, $x = 24$. Como $y = 3x - 2$, temos que $y = 70$. Isso quer dizer que

$$1^2 + 2^2 + 3^2 + \cdots + 24^2 = 70^2$$

Se nós temos 4900 esferas podemos organizá-las em uma pirâmide de altura 24, ou colocá-las na forma de um quadrado 70 por 70. Se continuássemos repetindo o procedimento acima, por exemplo, usando os pontos já encontrados como um dos nossos pontos, vamos obter infinitas soluções racionais para nossa equação. No entanto, pode ser mostrado que $(24, 70)$ é a única solução em inteiros positivos para o nosso problema, além da soluçao trivial $x = 1$. Isso exige técnicas mais sofisticadas.

Finalmente, vamos considerar a equação quártica de Fermat. Nós queremos mostrar que

$$a^4 + b^4 = c^4 \tag{1.1}$$

não possui soluções, sendo que $a,\ b,\ c$ são números inteiros diferentes de zero. Essa equação representa o caso mais fácil do Último Teorema de Fermat, que afirma que a soma de duas potências n-ésimas de números inteiros diferentes de zero não pode ser igual a uma potência n-ésima diferente de zero quando $n \geq 3$. Esse resultado geral foi provado por Wiles (usando trabalhos de Frey, Ribet, Serre, Mazur, Taylor, $\cdots$) em 1994 usando propriedades de curvas elípticas.

Suponha $a^4 + b^4 = c^4$ com $a \neq 0$. Tome

$$x = 2\frac{b^2 + c^2}{a^2}, \quad y = 4\frac{b(b^2 + c^2)}{a^3} \tag{1.2}$$

Um cálculo direto mostra que $y^2 = x^3 - 4x$.

Nos estudos das curvas elípticas sobre os racionais podemos mostrar que as únicas soluções dessa equação são $(x, y) = (0, 0), (2, 0), (-2, 0)$. Todos esses pontos substituídos em (1.2) nos dão $b = 0$, logo, não existem soluções inteiras não-triviais de (1.1).

A equação cúbica de Fermat também pode ser transformada na de uma curva elíptica. Suponha que $a^3 + b^3 = c^3$ e $abc \neq 0$. Como $a^3 + b^3 = (a + b)(a^2 - ab + b^2)$, devemos ter $a + b \neq 0$. Tome $x = 12\frac{c}{a+b}$ e $y = 36\frac{a-b}{a+b}$. Então, $y^2 = x^3 - 432$. Pode ser mostrado (mas isso não é fácil) que as únicas soluções racionais dessa equação são $(x, y) = (12, \pm 36)$ O caso $y = 36$ nos dá $a - b = a + b$, então $b = 0$. Analogamente, $y = -36$ nos dá $a = 0$. Portanto não existem soluções para $a^3 + b^3 = c^3$ quando $abc \neq 0$.

2 Conceitos básicos

2.1 Plano Projetivo e o Ponto no Infinito

Antes de começarmos a definir plano projetivo, faremos a seguinte pergunta: existem retas paralelas que se encontram em algum ponto? O plano projetivo nos permite dar sentido e resposta a essa pergunta.

Definição 2.1. Seja $\mathbb{K}$ um corpo. No conjunto $\mathbb{K}^3 \setminus \{(0,0,0)\}$, definimos a relação de equivalência $(x_1, y_1, z_1) \sim (x_2, y_2, z_2) \Leftrightarrow$ existe $\lambda \in \mathbb{K}$, com $\lambda \neq 0$, tal que $(x_1, y_1, z_1) = \lambda(x_2, y_2, z_2)$. Considere o conjunto quociente dessa relação de equivalência

$$\mathbb{P}^2_{\mathbb{K}} = \{(x_1, x_2, x_3) \in \mathbb{K}^3 : (x_1, x_2, x_3) \neq (0,0,0)\}/\sim .$$

Chamaremos $\mathbb{P}^2_{\mathbb{K}}$ de **plano projetivo** e seus elementos de **pontos projetivos**. Se (x_1, x_2, x_3) é um ponto de $\mathbb{K}^3$, $(x_1, x_2, x_3) \neq (0,0,0)$, sua classe de equivalência é denotada por $(x_1 : x_2 : x_3)$.

Definição 2.2. Se $(x : y : z)$ é um ponto com $z \neq 0$, então $(x : y : z) = (x/z : y/z : 1)$. Estes são definidos como os **pontos finitos** (vale salientar que esta definição não significa que há uma quantidade finita desses pontos) em $\mathbb{P}^2_{\mathbb{K}}$. No entanto, se $z = 0$, os pontos $(x : y : 0)$ são chamados **pontos no infinito** em $\mathbb{P}^2_{\mathbb{K}}$.

Definição 2.3. O **plano afim bi-dimensional** sobre $\mathbb{K}$ é frequentemente denotado

$$\mathbb{A}^2_{\mathbb{K}} = \{(x, y) \in \mathbb{K} \times \mathbb{K}\}$$

E, temos a inclusão $\mathbb{A}^2_{\mathbb{K}} \hookrightarrow \mathbb{P}^2_{\mathbb{K}}$ dada por $(x, y) \mapsto (x : y : 1)$.

Dessa maneira, o plano afim é identificado com os pontos finitos em $\mathbb{P}^2_{\mathbb{K}}$, e podemos ver que $\mathbb{P}^2_{\mathbb{K}} = \mathbb{A}^2_{\mathbb{K}} \cup \{\text{pontos no infinito}\}$.

Definição 2.4. Um polinômio é **homogêneo** em $\mathbb{K}[x, y, z]$ de grau n se é uma soma de termos da forma $ax^i y^j z^k$ com $a \in \mathbb{K}$ e $i + j + k = n$.

Exemplo 2.5. O polinômio $F(x, y, z) = 2x^3 - 5xyz + 7yz^2$ é homogêneo de grau 3.

Se um polinômio F é homogêneo de grau n, então $F(\lambda x, \lambda y, \lambda z) = \lambda^n F(x, y, z)$ para todo $\lambda \in \mathbb{K}$. Segue que se F é homogêneo de algum grau e $(x_1, y_1, z_1) = \lambda(x_2, y_2, z_2)$, então $F(x_1, y_1, z_1) = 0$ se, e somente se, $F(x_2, y_2, z_2) = 0$. Portanto, um zero de F em $\mathbb{P}^2_{\mathbb{K}}$ não depende da escolha do representante da classe de equivalência, então o conjunto de zeros de F em $\mathbb{P}^2_{\mathbb{K}}$ está bem definido.

Se $F(x, y, z)$ é um polinômio arbitrário em (x, y, z), então não podemos falar sobre um ponto em $\mathbb{P}^2_{\mathbb{K}}$ onde $F(x, y, z) = 0$, desde que esse dependa do representante (x, y, z) da classe de equivalência.

Exemplo 2.6. Seja $F(x, y, z) = x^2 + 2y - 3z$. Então, $F(1, 1, 1) = 0$, logo, devemos ter o cuidado de dizer que F se anula em $(1 : 1 : 1)$. Mas, $F(2, 2, 2) = 2$ e $(1 : 1 : 1) = (2 : 2 : 2)$.

Para evitar esse problema, precisamos trabalhar com polinômios homogêneos.

Se $f(x, y)$ é um polinômio em x e y, então podemos torná-lo homogêneo inserindo potências apropriadas de z.

Exemplo 2.7. Seja $f(x, y) = y^2 - x^3 - Ax - B$, então obtemos o polinômio homogêneo $F(x, y, z) = y^2z - x^3 - Axz^2 - Bz^3$.

Se F é homogêneo de grau n, então $F(x, y, z) = z^n f\left(\frac{x}{z}, \frac{y}{z}\right)$, onde $f(x, y) = F(x, y, 1)$.

Agora podemos ver o que significa duas retas paralelas se encontrar no infinito.

Sejam

$$y = mx + b_1, \qquad y = mx + b_2 \quad \in \mathbb{K}[x, y]$$

duas retas paralelas não verticais com $b_1 \neq b_2$.

Homogenizando as duas retas, temos

$$y = mx + b_1z, \qquad y = mx + b_2z \quad \in \mathbb{K}[x, y, z]$$

Resolvendo as duas equações simultaneamente, temos

$$mx + b_1z = mx + b_2z \Rightarrow b_1z = b_2z \Rightarrow z = 0$$

Agora, substituindo em uma das retas, obtemos a interseção

$$z = 0 \quad \text{e} \quad y = mx$$

Como não podemos ter x, y, z sendo 0 simultaneamente, devemos ter $x \neq 0$.

Sendo assim, a interseção das duas retas paralelas(não verticais) é

$$(x : mx : 0) = (1 : m : 0)$$

que é um ponto no infinito em $\mathbb{P}^2_{\mathbb{K}}$.

Se $x = c_1$ e $x = c_2$ são duas retas verticais, basta homegenizá-las como feito a cima, ou seja, $x = c_1z$ e $x = c_2z$. Assim, de maneira análoga encontraremos a interseção das duas retas, que é o ponto

$$(0 : y : 0) = (0 : 1 : 0)$$

que é também um ponto no infinito em $\mathbb{P}^2_{\mathbb{K}}$

No entanto, neste trabalho usaremos na maioria das vezes coordenadas afins(não-projetivas) e tratamos o ponto no infinito como um caso especial quando preciso. Uma exceção é a prova da associatividade da lei de grupo sobre os pontos de uma curva elíptica, onde será conveniente usar as coordenadas projetivas.

2.2 Equações de Weierstrass

Uma forma de descrever uma curva elíptica é a partir da definição de uma equação de Weiertrass. A presente seção visa definir tal equação que será muito utilizada neste trabalho.

Definição 2.8. Uma **curva elíptica** E sobre um corpo $\mathbb{K}$ é o conjunto dos pontos $(X : Y : Z) \in \mathbb{P}^2_{\mathbb{K}}$ que satisfazem a equação

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3, \tag{2.1}$$

onde $a_1, a_2, a_3, a_4, a_6 \in \mathbb{K}$, são tais que

$$\Delta = \Delta(E) := -b_1^2b_4 - 8b_2^3 - 27b_3^2 + 9b_1b_2b_3 \neq 0,$$

com

$$b_1 = a_1^2 + 4a_2, \qquad b_2 = a_1a_3 + 2a_4$$
$$b_3 = a_3^2 + 4a_6, \qquad b_4 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2.$$

Temos que a curva E é não singular (em outras palavras, as derivadas parciais de (2.1) não se anulam simultaneamente), pois Δ (chamado o discriminante da curva) é diferente de zero. Essa equação será definida como **Equação generalizada de Weierstrass**.

Vejamos que uma curva elíptica possui somente um ponto no infinito. Tomando $Z = 0$, temos que o ponto no infinito da curva elíptica é $(0 : Y : 0) = (0 : 1 : 0)$.

Agora, quando $Z \neq 0$, considere as funções racionais $x := \dfrac{X}{Z}$ e $y := \dfrac{Y}{Z}$ restritas a E, temos que de (2.1) podemos tomar E como sendo simplesmente o ponto $(0 : 1 : 0)$ e o conjunto dos pontos em $\mathbb{A}^2_{\mathbb{K}}$ que satisfazem a equação afim

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

onde os $a_i's$ são tomados como em (2.1).

Vamos supor que x, y pertencem a alguma extensão $\mathbb{L} \supseteq \mathbb{K}$, então escreveremos $E(\mathbb{L})$ para denotar o conjunto de pontos que satisfazem a Equação generalizada de Weierstrass. Portanto, dizemos que

$$E(\mathbb{L}) = \{(0 : 1 : 0)\} \cup \{(x, y) \in \mathbb{L} \times \mathbb{L} \mid y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\}.$$

Agora, suponha que a característica do corpo não é 2, então podemos dividir a Equação generalizada de Weierstrass por 2 e completar o quadrado, conforme mostrado a seguir

$$\begin{aligned} & y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \\ \Rightarrow\ & y^2 + 2y\left(\frac{a_1x + a_3}{2}\right) + \left(\frac{a_1x + a_3}{2}\right)^2 = x^3 + a_2x^2 + a_4x + a_6 + \left(\frac{a_1x + a_3}{2}\right)^2 \\ \Rightarrow\ & \left(y + \frac{a_1x + a_3}{2}\right)^2 = x^3 + \left(a_2 + \frac{a_1^2}{4}\right)x^2 + \left(a_4 + \frac{a_1a_3}{2}\right)x + \left(\frac{a_3^2}{4} + a_6\right), \end{aligned}$$

que pode ser escrito como

$$y_1^2 = x^3 + a_2'x^2 + a_4'x + a_6',$$

com $y_1 = y + \dfrac{a_1x + a_3}{2}$ e com algumas constantes $a_2', a_4', a_6' \in \mathbb{K}$. Se a característica também não é 3, então podemos tomar $x_1 = x + \dfrac{a_2'}{3}$ e obtemos

$$\begin{aligned} y_1^2 &= (x_1 - \frac{a_2'}{3})^3 + a_2'(x_1 - \frac{a_2'}{3})^2 + a_4'(x_1 - \frac{a_2'}{3}) + a_6' \\ \Rightarrow y_1^2 &= x_1^3 + Ax_1 + B, \end{aligned}$$

para algumas contantes $A, B \in \mathbb{K}$.

Definição 2.9. Se a característica do corpo $\mathbb{K}$ for diferente de 2 e 3, uma **curva elíptica** E sobre um corpo $\mathbb{K}$ é o conjunto dos pontos $(X : Y : Z) \in \mathbb{P}^2_{\mathbb{K}}$ que satisfazem a equação

$$Y^2Z = X^3 + AXZ^2 + BZ^3, \qquad A, B \in \mathbb{K}$$

onde

$$\Delta = -(4A^3 + 27B^2) \neq 0.$$

Essa equação vai ser referida simplesmente como **Equação de Weierstrass** para uma curva elíptica E.

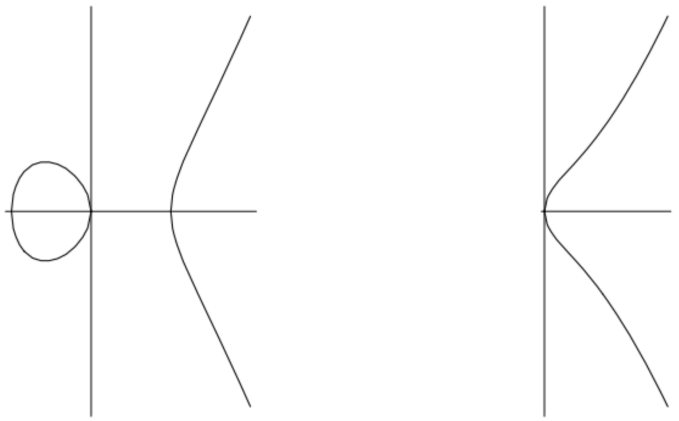
Da mesma forma que foi feita na equação generalizada de Weierstrass, se queremos considerar pontos com coordenadas em alguma extensão $\mathbb{L} \supseteq \mathbb{K}(\ \mathrm{char}(\mathbb{K} \neq 2, 3)\ )$, escrevemos $E(\mathbb{L})$ para denotar o conjunto dos pontos que satisfaz a Equação de Weierstrass. Portanto dizemos que

$$E(\mathbb{L}) = \{(0 : 1 : 0)\} \cup \{(x, y) \in \mathbb{L} \times \mathbb{L} \mid y^2 = x^3 + Ax + B\}$$

Vale a pena observar que, conforme visto na seção 2.1, $(0 : 1 : 0)$ pertence a todas as retas verticais, então toda reta vertical intersecta E no ponto no infinito. Além do mais, como $(0 : 1 : 0) = (0 : -1 : 0)$, o "topo"e o "extremo inferior"do eixo y são os mesmos.

A partir de agora quando referirmos a uma curva elíptica, estaremos supondo que o corpo a qual a curva está definida seja de característica diferente de 2 e 3, a não ser que digamos o contrário.

Não é possível desenhar figuras significativas de curvas elípticas sobre corpos finitos. No entanto, para intuição, é útil pensar em termos de gráfico sobre os números reais. Estes têm duas formas básicas, como mostramos a seguir:



(a) $y^2 = x^3 - x$ (b) $y^2 = x^3 + x$

A cúbica $y^2 = x^3 - x$ no primeiro caso tem três raízes reais distintas. No segundo caso, a cúbica $y^2 = x^3 + x$ tem uma única raiz real.

O que acontece se existe uma raiz múltipla? Nesse trabalho, não estudaremos este caso, pois estamos assumindo que $4A^3 + 27B^2 \neq 0$.

Se as raízes da equação cúbica da curva elíptica são r_1, r_2, r_3, então pode ser mostrado que o discriminante da curva é:

$$((r_1 - r_2)(r_1 - r_3)(r_2 - r_3))^2 = -(4A^3 + 27B^2).$$

Portanto, as raízes devem ser distintas.

Na maior parte desse trabalho, vamos desenvolver a teoria usando a equação de Weierstrass, ocasionalmente apontando quais modificações precisam ser feitas em características 2 e 3.

Finalmente, suponha que começamos com uma equação

$$cy^2 = dx^3 + ax + b$$

com $c, d \neq 0$. Multiplicando ambos os lados da equação por c^3d^2 para obter

$$(c^2dy)^2 = (cdx)^3 + (ac^2d)(cdx) + (bc^3d^2).$$

A mudança de variáveis

$$y_1 = c^2dy, \qquad x_1 = cdx$$

nos dá uma equação na forma de Weierstrass.

2.3 A Lei de Grupo

Nesta seção, será definida uma soma sobre os pontos racionais de uma curva elíptica, e mostramos que com essa soma o conjunto dos pontos racionais forma um grupo abeliano. Lembramos que um grupo abeliano é um conjunto de elementos associados a uma operação que combina dois elementos quaisquer para formar um terceiro, essa operação deve satisfazer as seguintes propriedades: comutatividade, associatividade, existência de um elemento neutro e cada elemento deve ter um elemento inverso.

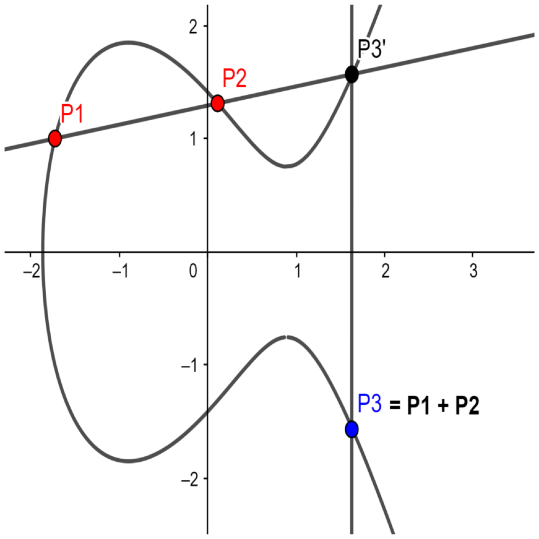


Figura 2 – Soma de pontos sobre uma curva elíptica.

Sejam os pontos distintos $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ sobre uma curva elíptica E dada pela equação $y^2 = x^3 + Ax + B$. Defina um novo ponto P_3 como se segue. Seja a reta L passando por P_1 e P_2. Vamos ver abaixo que L intersecta E em um terceito ponto P_3'. Depois, refletindo P_3' no eixo-x (i.e., mude o sinal da coordenada-y) para obter P_3. Assim, definimos a soma de pontos como:

$$P_1 + P_2 = P_3$$

Exemplos abaixo vão mostrar que isso não é o mesmo que somar as coordenadas dos pontos. Talvez, poderia ser melhor denotar essa operação por $P_1 +_E P_2$, mas optamos por uma notação mais simples, pois não vamos somar os pontos pela soma de coordenadas.

Suponha primeiro que $P_1 \neq P_2$ e que nenhum ponto é $(0 : 1 : 0)$. Seja a reta L que passa por P_1 e P_2. Sua inclinação é

$$m = \frac{y_2 - y_1}{x_2 - x_1},$$

Se $x_1 = x_2$, então L é vertical. Vamos tratar esse caso depois, então vamos assumir que $x_1 \neq x_2$. A equação de L é então

$$y = m(x - x_1) + y_1.$$

Para encontrar a intersecção com E, substitua na equação da curva elíptica para obter

$$(m(x - x_1) + y_1)^2 = x^3 + Ax + B.$$

Essa equação pode ser rearranjada para a forma $0 = x^3 - m^2x^2 + \cdots$. As três raízes dessa cúbica correspondem aos três pontos de intersecção de L com E. Portanto, poderíamos fatorar a cúbica para obter o terceiro valor de x. Mas há uma maneira mais simples. Sabemos que um polinômio cúbico $x^3 = ax^2 + bx + c$ com raízes r, s, t, então

$$x^3 + ax^2 + bx + c = (x - r)(x - s)(x - t) = x^3 - (r + s + t)x^2 + \cdots .$$

Portanto, $r + s + t = -a$. Se conhecemos duas raízes r, s, então podemos descobrir a terceira, que é $t = -a - r - s$

Em nosso caso, obtemos $x = m^2 - x_1 - x_2$ e $y = m(x - x_1) + y_1$. Agora, reflita através do eixo-x para obter $P_3 = (x_3, y_3)$:

$$x_3 = m^2 - x_1 - x_2, \qquad y_3 = -(m(x_3 - x_1) + y_1) = m(x_1 - x_3) - y_1.$$

No caso que $x_1 = x_2$, mas $y_1 \neq y_2$, a reta que passa por P_1 e P_2 é uma reta vertical, que portanto intersecta E em $(0 : 1 : 0)$. Refletindo $(0 : 1 : 0)$ através do eixo-x obtemos o mesmo ponto $((0 : -1 : 0) = (0 : 1 : 0))$. Portanto, nesse caso $P_1 + P_2 = (0 : 1 : 0)$.

Agora consideramos o caso onde $P_1 = P_2 = (x_1, y_1)$. Quando dois pontos sobre a curva estão muito pertos um do outro, a reta que passa por eles se aproxima de uma reta tangente. Portanto, quando os dois pontos coincidem, tomamos a reta L que passa por eles como sendo a reta tangente. Diferenciação implícita nos permite encontrar a inclinação m de L:

$$2y\frac{dy}{dx} = 3x^2 + A, \quad \text{então} \quad m = \frac{dy}{dx} = \frac{3x_1^2 + A}{2y_1}.$$

Se $y_1 = 0$ então a reta é vertical e temos $P_1 + P_2 = 2P_1 = (0 : 1 : 0)$, como antes. Portanto, assumimos que $y_1 \neq 0$. A equação de L é $y = m(x - x_1) + y_1$, como antes. Obtemos a equação cúbica $0 = x^3 - m^2x^2 + \cdots$. Desta vez, conhecemos somente uma raiz, a saber, x_1, mas é uma raiz dupla, pois L é tangente à E em P_1. Portanto, procedendo como antes, obtemos

$$x_3 = m^2 - 2x_1, \qquad y_3 = m(x_1 - x_3) - y_1.$$

Finalmente, suponha que $P_2 = (0 : 1 : 0)$. A reta que passa por P_1 e $(0 : 1 : 0)$ é uma reta vertical que intersecta E no ponto P_1' que é a reflexão de P_1 através do eixo-x. Quando refletimos P_1' através do eixo-x obtemos $P_3 = P_1 + P_2$, voltamos em P_1. Portanto,

$$P_1 + (0 : 1 : 0) = P_1$$

para todos pontos P_1 em E. Em particular vale para $P_1 = (0 : 1 : 0)$, logo, $(0 : 1 : 0) + (0 : 1 : 0) = (0 : 1 : 0)$.

LEI DE GRUPO

Seja E uma curva elíptica definida por $y^2 = x^3 + Ax + B$. Sejam $P_1 = (x_1, y_1)$ e $P_2 = (x_2, y_2)$ pontos em E com $P_1, P_2 \neq (0 : 1 : 0)$. Defina $P_1 + P_2 = P_3 = (x_3, y_3)$ como se segue:

1. Se $x_1 \neq x_2$, então

$$x_3 = m^2 - x_1 - x_2, \qquad y_3 = m(x_1 - x_3) - y_1, \qquad \text{onde } m = \frac{y_2 - y_1}{x_2 - x_1}$$

2. Se $x_1 = x_2$ mas $y_1 \neq y_2$, então $P_1 + P_2 = (0 : 1 : 0)$.

3. Se $P_1 = P_2$ e $y_1 \neq 0$, então

$$x_3 = m^2 - 2x_1, \qquad y_3 = m(x_1 - x_3) - y_1, \qquad \text{onde } m = \frac{3x_1^2 + A}{2y_1}.$$

4. Se $P_1 = P_2$ e $y_1 = 0$, então $P_1 + P_2 = (0 : 1 : 0)$

   Além disso, defina

$$P + (0 : 1 : 0) = P$$

   para todos pontos P em E.

Observe pelas fórmulas que, quando P_1 e P_2 têm coordenadas em um corpo $\mathbb{L}$ que contém os coeficientes A e B, então $P_1 + P_2$ também têm coordenadas em $\mathbb{L}$. Portanto, $E(\mathbb{L})$ é fechado em relação a operação adição de pontos feito acima.

Teorema 2.10. *A soma de pontos sobre uma curva elíptica E satisfaz as seguintes propriedades:*

1. *$P_1 + P_2 = P_2 + P_1$ para todos P_1, P_2 em E (comutatividade).*

2. *$P + (0 : 1 : 0) = P$ para todos pontos P em E (existência de identidade).*

3. *Dado P em E, existe P' em E com $P + P' = (0 : 1 : 0)$. Esse ponto P' será geralmente denotado por $-P$. (existência de inverso).*

4. *$(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3)$ para todos P_1, P_2, P_3 em E (associatividade).*

   Em outras palavras, os pontos sobre E formam um grupo abeliano aditivo com $(0 : 1 : 0)$ como o elemento identidade.

Demonstração.

- A comutatividade é óbvia, quer seja pelas fórmulas, quer seja pelo fato de que a reta que passa por P_1 e P_2 é a mesma que a reta que passa por P_2 e P_1

- A propriedade de identidade de $(0 : 1 : 0)$ vale por definição.

- Para a existência de inversos, seja P' a reflexão de P através do eixo-x. Então $P + P' = (0 : 1 : 0)$.

- Finalmente, precisamos provar a associatividade. Essa é de longe a propriedade mais sutil e não óbvia da soma de pontos sobre E. É possível definir muitas leis de composição satisfazendo $(1), (2), (3)$ para pontos sobre E, ou mais simples ou mais complicadas do que essa que foi considerada. Mas é muito improvável que tal lei irá ser associativa. Afinal, começamos com dois pontos P_1 e P_2 e realizamos um determinado procedimento para obter um terceiro ponto $P_1 + P_2 = P_3$. Então repetimos o procedimento com $P_1 + P_2$ e P_3 para obter $(P_1 + P_2) + P_3$. Se começarmos somando P_2 e P_3, então calculamos $P_1 + (P_2 + P_3)$, parece não haver razão óbvia que essa soma daria o mesmo ponto que o outro calculado. A lei de associatividade pode ser verificada por cálculos usando as fórmulas. Há vários casos, dependendo ou não se $P_1 = P_2$, se $P_3 = (P_1 + P_2)$, etc., e isso torna a prova muito complicada. No entanto, preferimos uma abordagem diferente que será dada na próxima seção. □

Observação 2.11. Para a equação de Weierstrass, se $P = (x, y)$, então $-P = (x, -y)$. Para a equação generalizada de Weierstrass (2.1), não é o mesmo caso. Se $P = (x, y)$ está sobre a curva descrita por (2.1), então

$$-P = (x, -a_1x - a_3 - y).$$

De fato, sejam $P_1 = (x, y)$ e $P_2 = (x, y_1)$ pontos sobre a curva descrita pela equação generalizada de Weierstrass. Então, temos que $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ e $y_1^2 + a_1xy_1 + a_3y_1 = x^3 + a_2x^2 + a_4x + a_6$ (ou seja, y e y_1 são as raízes do polinômio $p(t) = t^2 + a_1xt + a_3t - x^3 - a_2x^2 - a_4x - a_6$). Pelo mesmo método feito acima com a equação cúbica, temos que a soma das raízes dessa equação quadrada em termos de t é o coeficiente negativo que acompanha o termo t. Portanto, temos que $y + y_1 = -a_1x - a_3$, logo, $y_1 = -a_1x - a_3 - y)$ como desejado.

Exemplo 2.12. Os cálculos usados no método de Diophantus feitos no Capítulo 1 podem agora serem interpretados como soma de pontos sobre curvas elípticas. Sobre a curva

$$y^2 = \frac{x(x+1)(2x+1)}{6},$$

temos

$$(0,0) + (1,1) = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \left(\frac{1}{2}, \frac{1}{2}\right) + (1,1) = (24, -70)$$

Se P é um ponto sobre uma curva elíptica e k é um inteiro positivo, então kP denota $P+P+\cdots+P$ (com k parcelas). Se $k<0$, então $kP=(-P)+(-P)+\cdots+(-P)$, com $|k|$ parcelas. Para calcular kP para um inteiro grande k, é ineficiente somar P com si próprio repetidamente. É mais rápido usar a duplicação sucessiva. Por exemplo, para calcular $19P$, calculamos

$$2P,\quad 4P=2P+2P,\quad 8P=4P+4P,\quad 16P=8P+8P,\quad 19P=16P+2P+P.$$

Esse método nos permite calcular kP para k muito grande, digamos de cem dígitos, muito rápido. A única dificuldade é que o tamanho das coordenadas dos pontos aumentam rapidamente se estamos trabalhando sobre números racionais. No entanto, quando estamos trabalhando sobre um corpo finito, por exemplo $\mathbb{F}_p$, isso não é um problema porque podemos continuar reduzindo mod p e portanto, manter os números relativamente pequenos. Observe que a associatividade é a propriedade que nos permite fazer esses cálculos sem se preocupar com a ordem que usamos para combinar as somas.

Por outro lado, se estamos trabalhando sobre corpos finitos grandes e são dados P e kP, é muito difícil determinar o valor de k. Esse é chamado o **problema do logaritmo discreto** para curvas elípticas e é a base para as aplicações criptográficas que serão discutidas no capítulo 6.

2.4 Demonstração da Associatividade

Agora, vamos mostrar que a propriedade associativa é satisfeita pelo conjunto de pontos racionais de uma curva elíptica E. A ideia básica da demonstração é a seguinte. Sejam P,Q,R pontos racionais sobre E. Para calcular $-((P+Q)+R)$ precisamos de formar as retas $\ell_1=\overline{PQ}$, $m_2=\overline{(0:1:0),P+Q}$, e $\ell_3=\overline{R,P+Q}$, e veja onde elas intersectam E. Para calcular $-(P+(Q+R))$ precisamos formar as retas $m_1=\overline{QR}$, $\ell_2=\overline{(0:1:0),Q+R}$, e $m_3=\overline{P,Q+R}$. É fácil ver que os pontos $P_{ij}=\ell_i\cap m_j$ estão em E, exceto possivelmente P_{33}.

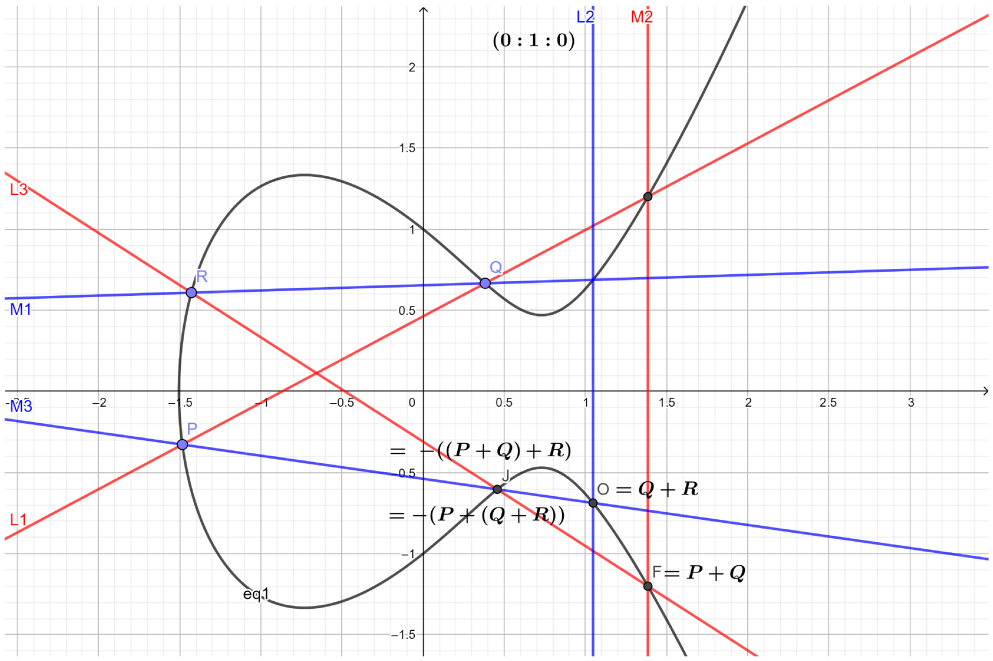


Figura 3 – Ideia da demonstração da associtatividade.

Vamos mostrar no Teorema 2.23 que ter os oitos pontos $P_{ij} \neq P_{33}$ sobre E implica P_{33} estar sobre a curva E. Como ℓ_3 intersecta E nos pontos $R, P+Q, -((P+Q)+R)$, devemos ter $-((P+Q)+R) = P_{33}$. Analogamente, $-(P+(Q+R)) = P_{33}$, então

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

que implica a propriedade associativa.

Há três aspectos principais técnicos que devem ser tratados. Primeiro, alguns dos pontos P_{ij} poderiam estar no infinito, então precisamos usar coordenadas projetivas. Segundo, uma reta poderia ser tangente à E, que significa que dois pontos P_{ij} poderiam ser iguais. Portanto, precisamos de uma definiçao cuidadosa da ordem em que uma reta intersecta uma curva. Terceiro, duas das retas poderiam ser iguais. Lidar com esses aspectos técnicos ocupa a maior parte da nossa atenção durante a prova.

Primeiro, precisamos discutir retas em $\mathbb{P}^2_{\mathbb{K}}$.

Definição 2.13. Uma **reta em** $\mathbb{P}^2_{\mathbb{K}}$ é descrita pela equação linear: $ax + by + cz = 0$. Algumas vezes é útil dar uma descrição paramétrica:

$$\begin{aligned} x &= a_1u + b_1v \\ y &= a_2u + b_2v \\ z &= a_3u + b_3v \end{aligned} \tag{2.2}$$

onde u, v percorre $\mathbb{K}$, e pelo menos um dos u, v é não-nulo.

Exemplo 2.14. Se $a \neq 0$, a reta $ax+by+cz = 0$ pode ser descrita por $x = -\frac{b}{a}u - \frac{c}{a}v, y = u, z = v$.

Agora, suponha que todos os pontos (a_i, b_i) são múltiplos um do outro, digamos $(a_i, b_i) = \lambda_i(a_1, b_1)$. Então,

$$\begin{aligned} x &= a_1u + b_1v \\ y &= \lambda_2(a_1u + b_1v) = \lambda_2 x \\ z &= \lambda_3(a_1u + b_1v) = \lambda_3 x \end{aligned}$$

Portanto, $(x, y, z) = x(1, \lambda_2, \lambda_3)$ para todos u, v tais que $x \neq 0$. Então conseguimos um ponto, ao invés de uma reta, no espaço projetivo. Portanto, precisamos de uma condição sobre os coeficientes $a_1, \cdots, b_3$ de (2.2) para garantirmos que realmente temos uma reta. Não é difícil ver que devemos exigir que a matriz

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$$

tenha posto 2.

Se $(u_1, v_1) = \lambda(u_2, v_2)$ para algum $\lambda \in \mathbb{K}^\times$, então (u_1, v_1) e (u_2, v_2) produz triplas equivalentes (x, y, z). Portanto, podemos considerar (u, v) percorrendo pontos $(u : v)$ no espaço projetivo unidimensional $\mathbb{P}^1_{\mathbb{K}}$. Consequentemente, uma reta corresponde a uma cópia da reta projetiva $\mathbb{P}^1_{\mathbb{K}}$ incorporada no plano projetivo.

Precisamos quantificar a ordem em que uma reta intersecta uma curva em um ponto.

Lema 2.15. *Seja $G(u, v) \in K[u, v]$ um polinômio homogêneo não-nulo e tome $(u_0 : v_0) \in \mathbb{P}^1_{\mathbb{K}}$. Então existe um inteiro $k \geq 0$ e um polinômio $H(u, v)$ com $H(u_0, v_0) \neq 0$ tal que*

$$G(u, v) = (v_0u - u_0v)^k H(u, v).$$

Demonstração. Suponha que $v_0 \neq 0$. Seja m o grau de G. Tome $g(u) = G(u, v_0)$. Fatorando a maior potência possível de $u - u_0$, podemos escrever $g(u) = (u - u_0)^k h(u)$ para algum k e para algum polinômio h de grau $m - k$ com $h(u_0) \neq 0$. Defina $H(u, v) = \left(\frac{v^{m-k}}{v_0^m}\right) h\left(\frac{uv_0}{v}\right)$, então $H(u, v)$ é homogêneo de grau $m - k$.

Como $G(u, v)$ é homogêneo de grau m, temos

$$G\left(\frac{u}{v}, 1\right) = G\left(\frac{u}{v}, \frac{v}{v}\right) = \frac{1}{v^m} G(u, v). \tag{2.3}$$

Assim,

$$g\left(\frac{uv_0}{v}\right) = G\left(\frac{uv_0}{v}, v_0\right) = v_0^m G\left(\frac{u}{v}, 1\right) \tag{2.4}$$

Portanto, pelas equações (2.3) e (2.4) temos

$$G(u,v) = v^m G\left(\frac{u}{v}, 1\right) = \frac{v^m}{v_0^m} g\left(\frac{uv_0}{v}\right) = \frac{v^m}{v_0^m}\left(\frac{uv_0}{v} - u_0\right)^k h\left(\frac{uv_0}{v}\right) =$$

$$= \frac{v^{m-k}}{v_0^m}(uv_0 - u_0v)^k h\left(\frac{uv_0}{v}\right) = (uv_0 - u_0v)^k H(u,v)$$

como desejado.

Se $v_0 = 0$, então $u_0 \neq 0$. Invertendo as regras de u e v nos dá a prova nesse caso. □

Definição 2.16. Seja $f(x,y) = 0$ (onde f é um polinômio) que descreve uma curva C no plano afim $\mathbb{A}^2_{\mathbb{K}}$ e seja

$$x = a_1t + b_1, y = a_2t + b_2$$

uma reta escrita em termos do parâmetro t. Defina

$$\tilde{f}(t) = f(a_1t + b_1, a_2t + b_2).$$

Então L intersecta C quando $t = t_0$ se $\tilde{f}(t_0) = 0$. Se $(t - t_0)^2$ divide $\tilde{f}(t)$, então L é tangente a C (se o ponto correspondente a t_0 é não-singular). Mais geralmente, dizemos que **a ordem em que** L **intersecta** C **é** n **no ponto** (x,y) **correspondente a** $t = t_0$ se $(t - t_0)^n$ é a maior potência de $(t - t_0)$ que divide $\tilde{f}(t)$.

A versão homogênea da definição acima é a seguinte.

Definição 2.17. Seja $F(x,y,z)$ um polinômio homogêneo, então $F = 0$ descreve uma curva C em $\mathbb{P}^2_{\mathbb{K}}$. Seja L uma reta dada parametricamente por (2.2) e defina

$$\tilde{F}(u,v) = F(a_1u + b_1v, a_2u + b_2v, a_3u + b_3v).$$

Dizemos que **a ordem em que** L **intersecta** C **é** n **no ponto** $P = (x_0 : y_0 : z_0)$ **correspondente a** $(u : v) = (u_0 : v_0)$ se $(v_0u - u_0v)^n$ é a maior potência de $(v_0u - u_0v)$ dividindo $\tilde{F}(u,v)$. Denotamos isso por

$$\mathrm{ord}_{L,P}(F) = n.$$

Se $\tilde{F}$ é identicamente 0, então denotamos $\mathrm{ord}_{L,P}(F) = \infty$. Não é difícil mostrar que $\mathrm{ord}_{L,P}(F)$ é independente da escolha de parametrização da reta L. Observe que $v = v_0 = 1$ corresponde à situação não homogênea acima, e as definições coincidem (pelo menos quando $z \neq 0$). A vantagem da formulação homogênea é que nos permite a tratar pontos no infinito junto com os pontos finitos de uma maneira uniforme.

Lema 2.18. *Sejam L_1 e L_2 retas intersectando em um ponto P, e para $i = 1, 2$, seja $L_i(x,y,z)$ seu polinômio linear, respectivamente. Então $ord_{L_1,P}(L_2) = 1$ a menos que $L_1(x,y,z) = \alpha L_2(x,y,z)$ para alguma constante α, nesse caso $ord_{L_1,P}(L_2) = \infty$.*

Demonstração. Quando substituímos a parametrização de L_1 em $L_2(x, y, z)$, obtemos $\tilde{L}_2$, que é uma expressão linear em termos de u e v. Seja P o ponto da hipótese correspondente a $(u_0 : v_0)$. Como $\tilde{L}_2(u_0, v_0) = 0$, segue do lema 2.15 que $\tilde{L}_2(u, v) = \beta(v_0 u - u_0 v)$ para alguma constante β. Se $\beta \neq 0$, então $\text{ord}_{L_1,P}(L_2) = 1$. Se $\beta = 0$, então todos os pontos sobre L_1 estão em L_2. Como dois pontos em $\mathbb{P}^2_{\mathbb{K}}$ determinam uma reta, e L_1 tem pelo menos três pontos ($\mathbb{P}^1_{\mathbb{K}}$ sempre contém os pontos $(1 : 0), (0 : 1), (1 : 1)$), segue que L_1 e L_2 são a mesma reta. Portanto $L_1(x, y, z)$ é proporcional a $L_2(x, y, z)$. □

Geralmente, uma reta que intersecta uma curva em ordem de pelo menos 2 é tangente a curva. No entanto, considere a curva C definida por

$$F(x, y, z) = y^2 z - x^3 = 0$$

Seja

$$x = au, \quad y = bu, \quad z = v$$

uma reta que passa por $P = (0 : 0 : 1)$. Observe que P corresponde a $(u : v) = (0 : 1)$. Temos $\tilde{F}(u, v) = u^2(b^2 v - a^3 u)$, então toda reta que passa por P intersecta C em ordem pelo menos 2. A reta com $b = 0$, que é a melhor escolha para a tangente em P, intersecta C em ordem 3. A parte afim de C é a curva $y^2 = x^3$. O ponto $(0, 0)$ é uma singularidade da curva, e é por isso que as intersecções em P têm ordens mais altas do que o esperado. Essa é uma situação que geralmente queremos evitar.

Definição 2.19. Uma curva C em $\mathbb{P}^2_{\mathbb{K}}$ definida por $F(x, y, z) = 0$ é dita ser **não singular** se pelo menos uma das derivadas parciais F_x, F_y, F_z é não nula em P.

Exemplo 2.20. Considere uma curva elíptica definida por $F(x, y, z) = y^2 z - x^3 - Axz^2 - Bz^3 = 0$, e suponha que a característica do nosso corpo $\mathbb{K}$ é diferente de 2 e 3. Temos

$$F_x = -3x^2 - Az^2, \quad F_y = 2yz, \quad F_z = y^2 - 2Axz - 3Bz^2$$

Suponha que $P = (x : y : z)$ é um ponto singular. Se $z = 0$, então $F_x = 0$ implica que $x = 0$ e $F_z = 0$ implica que $y = 0$, então $P = (0 : 0 : 0)$, que é impossível. Portanto $z \neq 0$, então podemos pegar $z = 1$(e portanto o ignoramos). Se $F_y = 0$, então $y = 0$. Como $(x : y : 1)$ está sobre a curva, x deve satisfazer $x^3 + Ax + B = 0$. Se $F_x = -(3^2 + A) = 0$, então x é uma raiz de um polinômio e uma raiz de sua derivada, consequentemente uma raiz dupla. Como assumimos que o polinômio cúbico não tem raízes múltiplas, teremos uma contradição. Contudo, uma curva elíptica não apresenta pontos singulares. Observe que isto é verdade mesmo se estamos considerando pontos com coordenadas em $\overline{\mathbb{K}}$ (= fecho algébrico de $\mathbb{K}$). Em geral, por uma **curva não singular**, significa que a curva não tem pontos singulares em $\overline{\mathbb{K}}$.

Se permitimos o polinômio cúbico $x^3 + Ax + B$ ter uma raíz múltipla x, então é fácil de ver que a curva tem uma singularidade em $(x : 0 : 1)$.

Se P é um ponto não singular da curva $F(x, y, z) = 0$, então a reta tangente em P

$$F_x(P)x + F_y(P)y + F_z(P)z = 0 \tag{2.5}$$

Exemplo 2.21. Se $F(x, y, z) = y^2z - x^3 - Axz^2 - Bz^3 = 0$, então obtemos a **reta tangente** em $(x_0 : y_0 : z_0)$ é

$$(-3x_0^2 - Az_0^2)x + 2y_0z_0y + (y_0^2 - 2Ax_0z_0 - 3Bz_0^2)z = 0.$$

Se adotamos $z_0 = z = 1$, então obtemos

$$(-3x_0^2 - A)x + 2y_0y + (y_0^2 - 2Ax_0 - 3B) = 0.$$

Usando o fato que $y_0^2 = x_0^3 + Ax_0 + B$, podemos reescrever como

$$(-3x_0^2 - A)(x - x_0) + 2y_0(y - y_0) = 0.$$

Lema 2.22. *Seja $F(x, y, z) = 0$, tal que descreva uma curva C. Se P é um ponto não singular de C, então existe exatamente uma reta em $\mathbb{P}^2_{\mathbb{K}}$ que intersecta C em ordem de pelo menos 2, e é tangente de C em P.*

Demonstração. Seja L uma reta intersectando C em ordem $k \geq 1$. Parametrize L por (2.2) e substitua em F. Isso nos dá $\tilde{F}(u, v)$. Seja $(u_0 : v_0)$ correspondente a P. Então $\tilde{F} = (v_0u - u_0v)^k H(u, v)$ para algum $H(u, v)$ com $H(u_0, v_0) \neq 0$. Portanto,

$$\tilde{F}_u(u, v) = kv_0(v_0u - u_0v)^{k-1}H(u, v) + (v_0u - u_0v)^k H_u(u, v)$$

e

$$\tilde{F}_v(u, v) = -kv_0(v_0u - u_0v)^{k-1}H(u, v) + (v_0u - u_0v)^k H_v(u, v).$$

Segue que $k \geq 2$ se, e só se, $\tilde{F}_u = \tilde{F}_v(u_0, v_0) = 0$.

Suponha que $k \geq 2$. Pela regra da cadeia temos

$$\tilde{F}_u = a_1F_x + a_2F_y + a_3F_z = 0, \quad \tilde{F}_v = b_1F_x + b_2F_y + b_3F_z = 0 \tag{2.6}$$

em P. Lembre-se que como a parametrização (2.2) dá uma reta, os vetores (a_1, a_2, a_3) e (b_1, b_2, b_3) devem ser linearmente independente.

Suponha que L' é outra reta que intersecta C em ordem de pelo menos 2. Então obtemos outro conjunto de equações

$$a_1'F_x + a_2'F_y + a_3'F_z = 0, \quad b_1'F_x + b_2'F_y + b_3'F_z = 0$$

em P.

Se os vetores $\mathbf{a'} = (a_1', a_2', a_3')$ e $\mathbf{b'} = (b_1', b_2', b_3')$ geram o mesmo plano em $\mathbb{K}^3$ como $\mathbf{a} = (a_1, a_2, a_3)$ e $\mathbf{b} = (b_1, b_2, b_3)$, então

$$\mathbf{a'} = \alpha\mathbf{a} + \beta\mathbf{b}, \quad \mathbf{b'} = \gamma\mathbf{a} + \delta\mathbf{b}$$

para alguma matriz invertível

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

. Portanto,

$$u\mathbf{a'} + v\mathbf{b'} = (u\alpha + v\gamma)\mathbf{a} + (u\beta + v\delta)\mathbf{b} = u_1\mathbf{a} + v_1\mathbf{b}$$

para uma nova escolha de parâmetros u_1, v_1. Isso significa que L e L' são as mesmas retas.

De fato, se L e L' são retas diferentes, então $\mathbf{a}, \mathbf{b}$ e $\mathbf{a}', \mathbf{b}'$ geram planos diferentes, então os vetores $\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}'$ devem gerar todos de $\mathbb{K}^3$. Como (F_x, F_y, F_z) tem produto escalar igual a 0 com esses vetores, esse deve ser o vetor 0. Isso significa que P é um ponto singular, contrariando a nossa hipótese.

Finalmente, precisamos mostrar que a reta tangente intersecta a curva em ordem de pelo menos 2. Suponha, por exemplo, que $F_x \neq 0$ em P. Os casos onde $F_x \neq 0$ e $F_z \neq 0$ são semelhantes. A reta tangente dada pela equação (2.5) pode ser parametrizada por

$$x = -\left(\frac{F_y}{F_x}\right)u - \left(\frac{F_z}{F_x}\right)v, \qquad y = u, \qquad z = v,$$

então

$$a_1 = -\frac{F_y}{F_x}, \quad b_1 = -\frac{F_z}{F_x}, \quad a_2 = 1, \quad b_2 = 0, \quad a_3 = 0, \quad b_3 = 1$$

na notação de (2.2). Substitua em (2.6) para obter

$$\tilde{F}_u = \left(\frac{-F_y}{F_x}\right)F_x + F_y = 0, \qquad \tilde{F}_v = \left(\frac{-F_z}{F_x}\right)F_x + F_z = 0$$

Pela discussão no início da demonstração, isso significa que a reta tangente intersecta a curva em ordem $k \geq 2$. □

A associatividade da soma dos pontos sobre a curva elíptica seguirá facilmente do próximo resultado. A demonstração pode ser simplificada se os pontos P_{ij} são assumidos serem distintos. Os casos onde pontos são iguais correspondem a situações onde retas tangentes são usadas na definição da lei do grupo. Correspondentemente, esse caso é onde é mais difícil verificar a associatividade pelo cálculo direto com as fórmulas da lei de grupo.

Teorema 2.23. *Seja $C(x, y, z)$ um polinômio cúbico homogêneo, e seja C a curva em $\mathbb{P}^2_{\mathbb{K}}$ descrita por $C(x, y, z) = 0$. Sejam ℓ_1, ℓ_2, ℓ_3 e m_1, m_2, m_3 retas em $\mathbb{P}^2_{\mathbb{K}}$ tais que $\ell_i \neq m_j$ para todo i, j. Seja P_{ij} o ponto de interseção de ℓ_i e m_j. Suponha P_{ij} um ponto não singular sobre a curva C para todos $(i, j) \neq (3, 3)$. Além disso, exigimos que se, para algum i, existe $k \geq 2$ dos pontos P_{i1}, P_{i2}, P_{i3} igual ao mesmo ponto, então ℓ_i intersecta C em ordem de pelo menos k nesse ponto. Além disso, se, para algum j, existe $j \geq 2$ de pontos P_{1j}, P_{2j}, P_{3j} igual ao mesmo ponto, então m_j intersecta C em ordem de pelo menos k nesse ponto. Então P_{33} também está sobre a curva C.*

Demonstração. Expresse ℓ_1 na forma paramétrica (2.2). Então $C(x, y, z)$ torna $\tilde{C}(u, v)$. A reta ℓ_1 passa por P_{11}, P_{12}, P_{13}. Sejam $(u_1 : v_1), (u_2 : v_2), (u_3 : v_3)$ os parâmetros em ℓ_1 para esses pontos. Como esses pontos estão sobre C, temos $\tilde{C}(u_i, v_i) = 0$ para $i = 1, 2, 3$.

Seja $m_j(x, y, z) = a_j x + b_j y + c_j z = 0$ a equação de m_j. Substituindo a parametrização para ℓ_1 se tem $\tilde{m}_j(u, v)$. Como P_{ij} está em m_j, temos $\tilde{m}_j(u_j, v_j) = 0$ para $j = 1, 2, 3$. Como $\ell_1 \neq m_j$ e como os zeros de $\tilde{m}_j$ nos dão as intersecções de ℓ_1 e m_j, a função $\tilde{m}_j(u, v)$ se anula apenas em P_{1j}, então a forma linear de $\tilde{m}_j$ não é nula. Portanto, o produto $\tilde{m}_1(u, v)\tilde{m}_2(u, v)\tilde{m}_3(u, v)$ é um polinômio homogêneo cúbico não nulo. Precisamos relacionar esse produto com $\tilde{C}$. Para isso, iremos provar o seguinte lema.

Lema 2.24. *Sejam $R(u, v)$ e $S(u, v)$ polinômios homogêneos de grau 3 com $S(u, v)$ não identicamente 0, e suponha que exista três pontos $(u_i : v_i), i = 1, 2, 3$, em que R e S se anulam. Além do mais, se k desses pontos são iguais ao mesmo ponto, exigimos que R e S se anulem em ordem de pelo menos k nesse ponto (isto é, $(v_i u - u_i v)^k$ divide R e S). Então existe uma constante $\alpha \in \mathbb{K}$ tal que $R = \alpha S$.*

Demonstração. Primeiro, observe que um polinômio homogêneo cúbico não nulo $S(u, v)$ pode ter no máximo 3 zeros $(u : v)$ em $\mathbb{P}^1_{\mathbb{K}}$ (contando multiplicidades). De fato, fatore a maior potência possível de v, digamos v^k. Então, $S(u, v)$ se anula em ordem k em $(1 : 0)$, e $S(u, v) = v^k S_0(u, v)$ com $S_0(1, 0) \neq 0$. Como $S_0(u, 1)$ é um polinômio de grau $3 - k$, o polinômio $S_0(u, 1)$ pode ter no máximo $3 - k$ zeros, contando multiplicidades (tem exatamente $3 - k$ se $\mathbb{K}$ é algebricamente fechado). Todos pontos $(u : v) \neq (1 : 0)$ pode ser escrito na forma $(u : 1)$, então $S_0(u, v)$ tem no máximo $3 - k$ zeros. Portanto, $S(u, v)$ tem no máximo $k + (3 - k) = 3$ zeros em $\mathbb{P}^1_{\mathbb{K}}$.

Seja $(u_0 : v_0)$ qualquer ponto em $\mathbb{P}^1_{\mathbb{K}}$ não é igual a qualquer dos $(u_i : v_i)$. (Ponto técnico: Se $\mathbb{K}$ tem apenas dois elementos em $\mathbb{P}^1_{\mathbb{K}}$, então $\mathbb{P}^1_{\mathbb{K}}$ tem apenas três elementos. Nesse caso, extenda $\mathbb{K}$ à $G\mathbb{F}(4)$ (corpo finito de 4 elementos e característica 2). O α que obtemos é forçado estar em $\mathbb{K}$ desde que seja a razão de um coeficiente de R e um coeficiente de S, dos quais ambos estão em $\mathbb{K}$.) Como S pode ter no máximo três zeros, $S(u_0, v_0) \neq 0$. Defina $\alpha = \dfrac{R(u_0, v_0)}{S(u_0, v_0)}$. Então, $R(u, v) - \alpha S(u, v)$ é um polinômio homogêneo que se anula nos quatros pontos $(u_i : v_i), i = 0, 1, 2, 3$. Portanto, $R - \alpha S$ deve ser identicamente nula. □

Voltando para a demonstração do teorema, observamos que $\tilde{C}$ e $\tilde{m}_1\tilde{m}_2\tilde{m}_3$ se anulam nos pontos $(u_i : v_i)$, para $i = 1, 2, 3$. Além do mais, se k dos pontos P_{1j} são os mesmos pontos, então k das funções lineares se anulam nesse ponto, então o produto $\tilde{m}_1(u, v)\tilde{m}_2(u, v)\tilde{m}_3(u, v)$ se anula em ordem pelo menos k. Pela hipótese, $\tilde{C}$ se anula em ordem pelo menos k nessa situação. Pelo lema, existe uma constante α, tal que

$$\tilde{C} = \alpha\tilde{m}_1\tilde{m}_2\tilde{m}_3.$$

Defina

$$C_1(x,y,z) = C(x,y,z) - \alpha\tilde{m}_1(x,y,z)\tilde{m}_2(x,y,z)\tilde{m}_3(x,y,z).$$

A reta ℓ_1 pode ser descrita por uma equação linear $\ell_1(x,y,z) = ax + by + cz = 0$. Pelo menos um coeficiente é diferente de 0, então vamos assumir $a \neq 0$. Os outros casos são análogos. A parametrização da reta ℓ_1 pode ser tomada da seguinte forma

$$x = -(\frac{b}{a})u - (\frac{c}{a})v, \quad y = u, \quad z = v. \tag{2.7}$$

Então, $\tilde{C}_1(u,v) = C_1\left(-\left(\frac{b}{a}\right)u - \left(\frac{c}{a}\right)v, u, v\right)$. Escreva $C_1(x,y,z)$ como um polinômio em x com polinômios em y, z como coeficientes. Escrevendo

$$x^n = \left(\frac{1}{a^n}\right)((ax+by+cz) - (by+cz))^n = \frac{1}{a^n}((ax+bx+cz)^n + \cdots),$$

podemos rearranjar $C_1(x,y,z)$ na forma de um polinômio em $ax+by+cz$, cujo coeficientes são polinômios em y, z:

$$C_1(x,y,z) = a_3(y,z)(ax+by+cz)^3 + \cdots + a_0(y,z) \tag{2.8}$$

Substituindo (2.7) em (2.8), temos

$$0 = \tilde{C}_1(u,v) = a_0(u,v),$$

pois $ax + bx + cz$ se anula identicamente quando x, y, z são escritos em termos de u, v. Portant, $a_0(y,z) = a_0(u,v)$ é um polinômio nulo. Segue de (2.8) que $C_1(x,y,z)$ é um múltiplo de $\ell_1(x,y,z) = ax + bx + cz$.

Analogamente, existe uma constante β tal que $C(x,y,z) - \beta\ell_1\ell_2\ell_3$ é um múltiplo de m_1. Defina

$$D(x,y,z) = C - \alpha m_1 m_2 m_3 - \beta\ell_1\ell_2\ell_3.$$

Então, $D(x,y,z)$ é um múltiplo de ℓ_1 e um múltiplo de m_1.

Lema 2.25. *$D(x,y,z)$ é um múltiplo de $\ell_1(x,y,z)m_1(x,y,z)$.*

Demonstração. Escreva $D = m_1 D_1$. Precisamos mostrar que ℓ_1 divide D_1. Parametrize a reta ℓ_1 como em (2.7) (de novo, estamos considerando o caso $a \neq 0$). Substituindo essa na relação $D = m_1 D_1$ temos $\tilde{D} = \tilde{m}_1\tilde{D}_1$. Como ℓ_1 divide D, temos $\tilde{D} = 0$. Como $m_1 \neq \ell_1$, temos $\tilde{m}_1 \neq 0$. Portanto, $\tilde{D}_1(u,v)$ é o polinômio nulo. Como a cima, isso implica que $D_1(x,y,z)$ é um múltiplo de ℓ_1, como desejado. □

Pelo lema, $D(x,y,z) = \ell_1 m_1 \ell$, onde $l(x,y,z)$ é linear. Por hipótese, $C = 0$ em P_{22}, P_{23}, P_{32}. Também, $\ell_1\ell_2\ell_3$ e $m_1m_2m_3$ se anulam nesses pontos. Portanto, $D(x,y,z)$ se anulam nesses pontos. Nosso objetivo é mostrar que D é identicamente 0.

Lema 2.26. $\ell(P_{22}) = \ell(P_{23}) = \ell(P_{32}) = 0$

Demonstração. Primeiro, suponha que $P_{13} \neq P_{23}$. Se $\ell_1(P_{23}) = 0$, então P_{23} está na reta ℓ_1 e também em ℓ_2 e m_3 por definição. Portanto, P_{23} é igual a intersecção P_{13} de ℓ_1 e m_3. Como P_{23} e P_{13} no momento são assumidos como distintos, então isso é uma contradição. Portanto, $\ell_1(P_{23}) \neq 0$. Como $D(P_{23}) = 0$, segue que $m_1(P_{23})\ell(P_{23}) = 0$.

Suponha agora que $P_{13} = P_{23}$. Então, pela hipótese do teorema, m_3 é tangente à C em P_{23}, então $\text{ord}_{m_3,P_{23}}(C) \geq 2$. Como $P_{13} = P_{23}$ e P_{23} pertence à m_3, temos que $\text{ord}_{m_3,P_{23}}(\ell_1) = \text{ord}_{m_3,P_{23}}(\ell_2) = 1$. Portanto, $\text{ord}_{m_3,P_{23}}(\alpha\ell_1\ell_2\ell_3) \geq 2$. Além disso, $\text{ord}_{m_3,P_{23}}(\beta m_1 m_2 m_3) = (0:1:0)$. Portanto, $\text{ord}_{m_3,P_{23}}(D) \geq 2$, como D é uma soma de termos, cada um deles se anulam em ordem pelo menos 2. Mas, $\text{ord}_{m_3,P_{23}}(\ell_1) = 1$, então temos

$$\text{ord}_{m_3,P_{23}}(m_1\ell) = \text{ord}_{m_3,P_{23}}(D) - \text{ord}m_3, P_{23}(\ell_1) \geq 1.$$

Portanto, $m_1(P_{23})\ell(P_{23}) = 0$. Em ambos os casos, temos $m_1(P_{23})\ell(P_{23}) = 0$. Se $m_1(P_{23}) \neq 0$, então $\ell_1(P_{23}) = 0$, como desejado.

Se $m_1(P_{23}) = 0$, então P_{23} está em m_1, e também em ℓ_2 e m_3, por definição. Portanto, $P_{23} = P_{21}$, pois ℓ_2 e m_1 intersecta em um único ponto. Pela hipótese do teorema, ℓ_2 é portanto, tangente à C em P_{23}. Assim, $\text{ord}_{\ell_2,P_{23}}(D) \geq 2$, logo

$$\text{ord}_{\ell_2,P_{23}}(\ell_1\ell) \geq 1.$$

Se nesse caso tivéssemos $\ell_1(P_{23}) = 0$, então P_{23} está em ℓ_1, ℓ_2, m_3. Portanto, $P_{13} = P_{23}$. Pela hipótese, a reta m_3 é tangente à C em P_{23}. Como P_{23} é um ponto não singular de C, pelo lema 4.9 temos que $\ell_2 = m_3$, contrariando a hipótese do teorema. Sendo assim, temos $\ell_1(P_{23}) \neq 0$, então $\ell(P_{23}) = 0$. Os casos, $\ell(P_{22}) = \ell(P_{32}) = 0$ são análogos. □

Se $\ell(x, y, z)$ é identicamente 0, então D é identicamente 0. Portanto, suponha que $\ell(x, y, z)$ é diferente de 0 e então, define uma reta ℓ.

Primeiro, suponha que P_{23}, P_{22}, P_{32} são distintos. Então, ℓ e ℓ_2 são retas passando por P_{23} e P_{22}. Assim, $\ell = \ell_2$. Consequentemente, $\ell = m_2$. Portanto, $\ell_2 = m_2$, que é uma contradição.

Agora suponha que $P_{32} = P_{22}$. Então, m_2 é tangente à C em P_{22}. Como antes,

$$\text{ord}_{m_2,P_{22}}(\ell_1 m_1 \ell) \geq 2.$$

Queremos mostrar que isso força ℓ ser a mesma reta que m_2.

Se $m_1(P_{22}) = 0$, então P_{22} está em m_1, m_2, ℓ_2. Portanto, $P_{21} = P_{22}$. Isso significa que ℓ_2 é tangente à C em P_{22}. Pelo lema 4.9, $\ell_2 = m_2$, que é uma contradição. Logo, $m_1(P_{22}) \neq 0$.

Se $\ell_1(P_{22}) \neq 0$, então $\text{ord}_{m_2,P_{22}}(\ell) \geq 2$. Isso significa que ℓ é a mesma reta que m_2.

Se $\ell_1(P_{22}) = 0$, então $P_{22} = P_{32}$ está em $\ell_1, \ell_2, \ell_3, m_3$, então $P_{12} = P_{22} = P_{32}$. Portanto $\text{ord}_{m_2,P_{22}}(C) \geq 3$. Pelo mesmo raciocínio acima, temos agora $\text{ord}_{m_2,P_{22}}(\ell_1 m_1 \ell) \geq 3$. Como já provamos que $m_1(P_{22}) \neq 0$, temos $\text{ord}_{m_2,P_{22}}(\ell) \geq 2$. Isso significa que ℓ é a mesma reta que m_2.

Portanto, agora provamos, assumindo que $P_{32} = P_{22}$, que ℓ é a mesma reta que m_2. Pelo lema 2.26, P_{23} está em ℓ, e portanto em m_2. Além disso, o ponto P_{23} está também em ℓ_2 e m_3. Portanto, $P_{22} = P_{23}$. Isso significa que ℓ_2 é tangente à C em P_{22}. Como $P_{32} = P_{22}$ significa que m_2 é também tangente à C em P_{22}, sendo assim, temos $\ell_2 = m_2$, que é uma contradição. Portanto, $P_{32} \neq P_{22}$ (sob a hipótese que $\ell \neq 0$).

Prova-se de modo análogo que $P_{23} \neq P_{22}$.

Finalmente, suponha que $P_{23} = P_{32}$. Então P_{23} está em ℓ_2, ℓ_3, m_2, m_3. Isso força $P_{22} = P_{32}$, que já mostramos que é impossível.

Portanto, todas as possibilidades levaram à contradições. Segue que $\ell(x, y, z)$ deve ser identicamente 0. Portanto, $D = 0$, então

$$C = \alpha \ell_1 \ell_2 \ell_3 + \beta m_1 m_2 m_3.$$

Como ℓ_3 e m_3 se anulam em P_{33}, temos $C(P_{33}) = 0$, como desejado. Isso completa a prova do Teorema 2.23. □

Podemos provar agora a associatividade da soma para uma curva elíptica. Sejam P, Q, R pontos sobre E. Defina as retas

$$\ell_1 = \overline{PQ}, \quad \ell_2 = \overline{(0:1:0), Q+R} \quad \ell_3 = \overline{R, P+Q}$$

$$m_1 = \overline{QR}, \quad m_2 = \overline{(0:1:0), P+Q} \quad m_3 = \overline{P, Q+R}.$$

Temos as seguintes interseções:

	ℓ_1	ℓ_2	ℓ_3
m_1	Q	$-(Q+R)$	R
m_2	$-(P+Q)$	$(0:1:0)$	$(P+Q)$
m_3	P	$Q+R$	X

Assuma no momento que as hipóteses do teorema são satisfeitas. Então todos os pontos na tabela, incluindo X, estão na curva E. A reta ℓ_3 tem três pontos de intersecção com E, a saber, $R, P+Q$, e X. Pela definição da soma de pontos sobre uma curva elíptica, $X = -((P+Q)+R)$. Analogamente, m_3, intersecta C em 3 pontos, que significa que $X = -(P+(Q+R))$. Portanto, após refletir através do eixo x, obtemos $(P+Q)+R = P+(Q+R)$, como desejado.

Resta verificar as hipóteses do teorema, ou seja que as ordens da intersecção estão correctas e que as retas ℓ_i são distintas das retas m_j.

Primeiro, queremos dispensar os casos onde $(0:1:0)$ ocorre. Como apontado anteriormente, a reta tangente em $(0:1:0)$ intersecta a curva somente em $(0:1:0)$. Segue que se duas das entradas em uma linha ou coluna da tabela acima de intersecções são iguais a $(0:1:0)$, então a terceira entrada também é $(0:1:0)$, e a reta intersecta a curva em ordem 3. Portanto, essa hipótese é satisfeita.

Também é possível tratar diretamente os casos onde alguns pontos de intersecção $P, Q, R \pm (P+Q), \pm(Q+R)$ são $(0:1:0)$. Nos casos onde pelo menos um dos pontos P, Q, R é $(0:1:0)$ a associatividade é trivial. De fato, fazendo para o caso quando $P = (0:1:0)$, temos $P + (Q+R) = (Q+R) = (Q+R) + P$.

Se $P + Q = (0:1:0)$, então $(P+Q) + R = (0:1:0) + R = R$. Por outro lado, a soma $Q + R$ é computada desenhando primeiro a reta L que passa por Q e R, que intersecta E em $-(Q+R)$. Como $P + Q = (0:1:0)$, a reflexão de Q através do eixo x é P. Portanto, a reflexão L' de L passa por $-Q = P, -R$ e $Q + R$. A soma $P + (Q+R)$ é encontrada por desenhar a reta que passa por P e $Q + R$, que é L'. No entanto, já observamos que o terceiro ponto da intersecção de L' com E é $-R$. Sendo assim, refletindo $-R$ temos $P + (Q+R) = R$, logo a associatividade vale nesse caso.

O caso onde $Q + R = (0:1:0)$ é análogo ao caso anterior.

Finalmente, precisamos considerar o que acontece se alguma reta ℓ_i for igual a alguma reta m_j, uma vez que o Teorema 2.23 não se aplica.

Primeiro, observe que se P, Q, R são colineares, então a associatividade é verificada facilmente. De fato, como os três pontos estão na mesma reta, logo $P + (Q+R) = P - P = (0:1:0)$ e $(P+Q) + R = -R + R = (0:1:0)$.

Segundo, suponha que $P, Q, Q+R$ são colineares. Então $P + (Q+R) = -Q$. Além disso, $P + Q = -(Q+R)$, então $(P+Q) + R = -(Q+R) + R$. O próximo Lemma nos permite afirmar que $-(Q+R) + R = -Q$, o que completa a demonstração desse caso.

Lema 2.27. *Sejam P_1, P_2 pontos sobre uma curva elíptica. Então $(P_1 + P_2) - P_2 = P_1$ e $-(P_1 + P_2) + P_2 = -P_1$.*

Demonstração. As duas relações são reflexões uma da outra, logo é suficiente provar a segunda. A reta L que passa por P_1 e P_2 intersecta a curva elíptica em $-(P_1 + P_2)$. Considerando L como a reta que passa por $-(P_1+P_2)$ e P_2 temos que $-(P_1+P_2)+P_2 = -P_1$, como afirmado. □

Suponha que $\ell_i = m_j$ para algum i, j. Consideramos vários casos. Pela discussão acima, podemos assumir que todos pontos na tabela de intersecções são finitos, exceto

para $(0:1:0)$ e possivelmente X. Observe que cada ℓ_i e cada m_j encontra E em três pontos (contando multiplicidade), um dos quais é P_{ij}. Se as duas retas coincidem, então os outros dois pontos devem coincidir em mesma ordem.

Segue as 9 possibilidades possíveis

1 $\ell_1 = m_1$: Então P, Q, R são colineares, e a associatividade segue.

2 $\ell_1 = m_2$: Nesse caso, $P, Q, (0:1:0)$, então $P + Q = (0:1:0)$ e a associatividade segue pelo cálculo direto feito acima.

3 $\ell_2 = m_1$: análogo ao caso anterior, pois $Q + R = (0:1:0)$

4 $\ell_1 = m_3$: Então $P, Q, Q + R$ são colineares e associatividade nesse caso foi provada acima.

5 $\ell_3 = m_1$: Análogo ao caso anterior, pois nesse caso $Q, R, P + Q$ são colineares.

6 $\ell_2 = m_2$: Então $P + Q$ devem ser $\pm(Q + R)$. Se $P + Q = Q + R$, então por comutatividade mais o Lema acima temos

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

Portanto,

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

Se $P + Q = -(Q + R)$, então

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

e

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

Logo, a associatividade vale

7 $\ell_2 = m_3$: Nesse caso, a reta m_3 passa por P e $(Q + R)$ intersecta E em $(0:1:0)$, então $P = -(Q + R)$. Como $-(Q + R), Q, R$ são colineares, temos que P, Q, R são colineares e a associatividade vale.

8 $\ell_3 = m_2$: Análogo ao caso anterior, mas desta vez teremos $-(P + Q) = R$.

9 $\ell_3 = m_3$: Como ℓ_3 não intersecta E em 4 pontos (contando multiplicidades), é fácil ver que $P = R$ ou $P = P + Q$ ou $Q + R = P + Q$ ou $Q + R = R$. O caso $P = R$ foi trado no caso $\ell_2 = m_2$. Suponha que $P = P + Q$. Somando $-P$ e aplicando o Lema 2.27 temos $(0:1:0) = Q$, nesse caso a associatividade segue imediatamente. O caso $Q + R = R$ é análogo. Se $Q + R = P + Q$, então somando $-Q$ e aplicando o Lema 2.27 temos $P = R$, logo vale a associatividade.

Se $\ell_i \neq m_j$ para todos i, j, então as hipóteses do teorema são satisfeitas, então a soma de pontos sobre uma curva elíptica é associativa, como provamos acima. Isso completa a prova da associatividade.

2.5 Endomorfismos

O principal objetivo dessa seção é provar a Proposição 2.32, que será usada na demontração do importante Teorema de Hasse no capítulo 4. Também provaremos alguns resultados técnicos à respeito dos endomorfismos.

Definição 2.28. Um homomorfismo $\alpha : E(\overline{\mathbb{K}}) \to E(\overline{\mathbb{K}})$ que é dado por funções racionais é denominado **endomorfismo** de E. Em outras palavras, $\alpha(P_1 + P_2) = \alpha(P_1) + \alpha(P_2)$, e há funções racionais (quocientes de polinômios) $R_1(x, y), R_2(x, y)$ com coeficientes em $\overline{\mathbb{K}}$, tais que

$$\alpha(x, y) = (R_1(x, y), R_2(x, y))$$

para todos $(x, y) \in E(\overline{\mathbb{K}})$.

Há alguns detalhes técnicos quando as funções racionais não estão definidas em um ponto. Esses serão tratados abaixo. Desde que α é um homomorfismo, temos $\alpha((0 : 1 : 0)) = (0 : 1 : 0)$. Vamos também assumir que α não é trivial; isto é, existe algum (x, y), tal que $\alpha(x, y) \neq (0 : 1 : 0)$. O endomorfismo trivial que mapeia todo ponto à $(0 : 1 : 0)$ será denotado por 0.

Exemplo 2.29. Seja E dado por $y^2 = x^3 + Ax + B$ e tome $\alpha(P) = 2P$. Então α é um homomorfimo e

$$\alpha(x, y) = (R_1(x, y), R_2(x, y)),$$

onde

$$R_1(x, y) = \left(\frac{3x^2 + A}{2y}\right)^2 - 2x$$

$$R_2(x, y) = \left(\frac{3x^2 + A}{2y}\right)\left(3x - \left(\frac{3x^2 + A}{2y}\right)^2\right) - y$$

Como α é um homomorfismo dado por funções racionais, então α é um endomorfismo de E.

Irá ser útil ter uma forma padrão para as funções racionais que descrevem um endomorfismo. Para simplicidade, assumimos que nossa curva elíptica é dada na forma de Weierstrass. Seja $R(x, y)$ uma função racional qualquer. Desde que $y^2 = x^3 + Ax + B$ para todo $(x, y) \in E(\overline{\mathbb{K}})$, conseguimos substituir qualquer potência par de y por um polinômio em x e substituir qualquer potência ímpar de y por y vezes um polinômio em x e obtenha uma função racional que dê a mesma função que $R(x, y)$ sobre os pontos em $E(\overline{\mathbb{K}})$.

Portanto, podemos assumir que

$$R(x,y) = \frac{p_1(x) + p_2(x)y}{p_3(x) + p_4(x)y}$$

Além disso, podemos racionalizar o denominador multiplicando o numerador e o denominador por $p_3 - p_4 y$ e então substituir y^2 por $x^3 + Ax + B$. Isso nos dá

$$R(x,y) = \frac{p_1(x)p_3(x) + p_2(x)p_3(x) - p_2(x)p_4(x)y^2 - p_1(x)p_4(x)y}{p_3(x)^2 - p_a(x)^2 y^2} = \frac{q_1(x) + q_2(x)y}{q_3(x)}. \tag{2.9}$$

Considere um endomorfismo dado por

$$\alpha(x,y) = (R_1(x,y), R_2(x,y)),$$

como acima. Desde que α é um homomorfismo, temos que dado $P = (x,y) \in E(\overline{\mathbb{K}})$

$$(0:1:0) = \alpha(P - P) = \alpha(P) + \alpha(-P), \quad \text{ou seja} \quad \alpha(-(x,y)) = -\alpha(x,y)$$

Portanto,

$$\alpha(x,-y) = \alpha(-(x,y)) = -\alpha(x,y).$$

Isso significa que

$$R_1(x,-y) = R_1(x,y) \quad \text{e} \quad R_2(x,-y) = -R_2(x,y),$$

pois $\alpha(x,-y) = (R_1(x,-y), R_2(x,-y)) = -(R_1(x,y), R_2(x,y)) = (R_1(x,y), -R_2(x,y))$.

Portanto, se R_1 é escrito da forma 2.9, temos

$$R_1(x,-y) = \frac{q_1(x) + q_2(x)(-y)}{q_3(x)} = \frac{q_1(x) + q_2(x)y}{q_3(x)} = R_1(x,y) \Rightarrow q_2(x) = 0.$$

e se R_2 é escrito da forma 2.9, temos

$$R_2(x,-y) = \frac{q_1(x) + q_2(x)(-y)}{q_3(x)} = -\frac{q_1(x) + q_2(x)y}{q_3(x)} = -R_2(x,y) \Rightarrow q_1(x) = 0.$$

Portanto, podemos supor que

$$\alpha(x,y) = (r_1(x), r_2(x)y)$$

com funções racionais $r_1(x), r_2(x)$.

Podemos agora dizer o que acontece quando uma das funções racionais não está definida em um ponto. Escreva

$$r_1(x) = \frac{p(x)}{q(x)}$$

com polinômios $p(x)$ e $q(x)$ que não tenham um favor comum. Se $q(x) = 0$ para algum ponto (x,y), então assumimos que $\alpha(x,y) = (0:1:0)$. Se $q(x) \neq 0$, então $r_2(x)$ está bem definido. De fato, se $\alpha(p(x)/q(x), r_2(x)y)$, então, $r_2(x)^2y^2 = ((p(x)/q(x))^3 + Ap(x)/q(x) + B = u(x)/q(x)^3$; consequentemente as funções racionais que definem α estão bem definidas.

Definição 2.30. Seja α um endomorfismo de E da forma $\alpha(x, y) = (r_1(x), r_2(x)y)$, assim, definimos o **grau** de α como sendo

$$deg(\alpha) = max\{deg\ p(x), deg\ q(x)\}$$

se α não é trivial. Quando $\alpha = 0$, definimos $deg(0) = 0$.

Se $\alpha \neq 0$ dizemos que é um endomorfismo **separável** se a derivada $r_1'(x)$ não é identicamente nula.

Um importante exemplo de um endomorfismo é o **mapa de Frobenius**. Suponha que a curva E é definida sobre o corpo finito $\mathbb{F}_q$. Defina $\phi_q(x, y) = (x^q, y^q)$. O mapa de Frobenius ϕ_q desempenha um papel crucial na teorial de curvas elípticas sobre $\mathbb{F}_q$.

Lema 2.31. *Seja E definida sobre $\mathbb{F}_q$. Então ϕ_q é um endomorfismo de E de grau q, e ϕ_q não é separável.*

Demonstração. Desde que $\phi_q(x, y) = (x^q, y^q)$, o mapa é dado por funções racionais (de fato, por polinômios) e o grau é q. Portanto, só nos resta provar que $\phi_q : E(\overline{\mathbb{F}}_q) \to E(\overline{\mathbb{F}}_q)$ é homomorfismo e não é separável. Sejam $(x_1, y_1), (x_2, y_2) \in E(\overline{\mathbb{F}}_q)$ com $x_1 \neq x_2$. A soma desses dois pontos é (x_3, y_3), com

$$x_3 = m^2 - x_1 - x_2, \quad y_3 = m(x_1 - x_3) - y_1, \ \text{ onde } m = \frac{y_2 - y_1}{x_2 - x_1}$$

(estamos trabalhando com a forma de Weierstrass aqui; a prova para a forma generalizada de Weierstrass é essencialmente a mesma). Eleve tudo à q-ésima potência para obter

$$x_3^q = m'^2 - x_1^q - x_2^q, \quad y_3 = m'(x_1^q - x_3^q) - y_1^q, \ \text{ onde } m' = \frac{y_2^q - y_1^q}{x_2^q - x_1^q}$$

Isso significa que

$$\phi_q((x_1, y_1) + (x_2, y_2)) = \phi_q((x_3, y_3)) = (x_3^q, y_3^q) = \phi_q(x_1, y_1) + \phi_q(x_2, y_2)$$

Os casos onde $x_1 = x_2$ ou onde um dos pontos é $(0 : 1 : 0)$ são verificados de modo semelhante, usando as fórmulas de suas respectivas somas.
No entanto, há uma sutileza que surge quando somamos um ponto à ele mesmo. A fórmula diz que $2(x_1, y_1) = (x_3, y_3)$, com

$$x_3 = m^2 - 2x_1, \quad y_3 = m(x_1 - x_3) - y_1, \quad \text{onde } m = \frac{3x_1^2 + A}{2y_1}.$$

Quando isso é elevado à $q-$ésima potência, obtemos

$$x_3^q = m'^2 - 2x_1^q, \quad y_3^q = m'(x_1^q - x_3^q) - y_1^q \quad \text{onde } m' = \frac{3^q(x_1^q)^2 + A^q}{2^q y_1^q}.$$

Desde que $2, 3, A \in \mathbb{F}_q$, temos $2^q = 2, 3^q = 3, A^q = A$. Isso significa que obtemos a fórmula para dobrar o ponto (x_1^q, y_1^q) sobre E (se A^q não é igual A, estaríamos trabalhando sobre uma nova curva elíptica com A^q no lugar de A). Como ϕ_q é um homomorfismo dado por funções racionais, é um endomorfismo de E. Desde que, $q = 0$ em $\mathbb{F}_q$, a derivada de x^q é identicamente nula. Portanto, ϕ_q não é separável. □

O seguinte resultado vai ser crucial na prova do teorema de Hasse no capítulo 4 e na prova do Teorema 4.10.

Proposição 2.32. *Seja $\alpha \neq 0$ um endomorfismo separável de uma curva elíptica E. Então*

$$deg\ \alpha = \#Ker(\alpha)$$

onde $Ker(\alpha)$ é o núcleo do endomorfismo $\alpha : E(\overline{\mathbb{K}}) \to E(\overline{\mathbb{K}})$.

Se $\alpha \neq 0$ não é separável, então

$$deg\ \alpha > \#Ker(\alpha)$$

Demonstração. Escreva $\alpha(x, y) = (r_1(x), yr_2(x))$ com $r_1(x) = \dfrac{p(x)}{q(x)}$, como acima. Se α é separável, então $r_1' \neq 0$, logo $p'q - pq'$ não é o polinômio nulo.

Seja S o conjunto dos $x \in \overline{\mathbb{K}}$, tais que $(pq' - p'q)(x)q(x) = 0$. Seja $(a, b) \in E(\overline{\mathbb{K}})$, tal que

1. $a \neq 0, b \neq 0, \quad (a\, b) \neq (0 : 1 : 0)$,
2. $deg(p(x) - aq(x)) = max\{deg(p), deg(q)\} = deg(\alpha)$,
3. $a \notin r_1(S)$, e
4. $(a, b) \in \alpha(E(\overline{\mathbb{K}}))$.

   Desde que $pq' - p'q$ é diferente do polinômio nulo, S é um conjunto finito, consequentemente sua imagem sob α é finito. A função $r_1(x)$ é facilmente vista como tendo infinito valores distintos à medida que x percorre em $\overline{\mathbb{K}}$. Como, para cada x, existe um ponto $(x, y) \in E(\overline{\mathbb{K}})$, vemos que $\alpha(E(\overline{\mathbb{K}}))$ é um conjunto infinito. Portanto, tal (a, b) existe.

   Afirmamos que há exatamente $deg(\alpha)$ pontos $(x_1, y_1) \in E(\overline{\mathbb{K}})$ tal que $\alpha(x_1, y_1) = (a, b)$. Para tal ponto, temos

$$\frac{p(x_1)}{q(x_1)} = a, \quad y_1 r_2(x_1) = b.$$

   Desde que $(a, b) \neq (0 : 1 : 0)$, devemos ter $q(x_1) \neq 0$. Assim, por uma observação feita após definirmos endomorfismo separável, temos que $r_2(x_1)$ está bem definido.

Como $b \neq 0$ e $y_1 r_2(x_1) = b$, devemos ter $y_1 = \dfrac{b}{r_2(x_1)}$. Portanto, x_1 determina y_1 nesse caso, então precisamos somente contar os valores de x_1.

Pela hipótese (2), $p(x) - aq(x) = 0$ tem $deg(\alpha)$ raízes, contando multiplicidades. Portanto, devemos mostrar que $p - aq$ não tem raízes múltiplas. Suponha que x_0 é uma raiz múltipla. Então,

$$p(x_0) - aq(x_0) = 0 \quad \text{e} \quad p'(x_0) - aq'(x_0) = 0.$$

Multiplicando as equações $p = aq$ e $aq' = p'$ temos

$$ap(x_0)q'(x_0) = ap'(x_0)q(x_0).$$

Como $a \neq 0$, isso implica que x_0 é uma raiz de $pq' - p'q$, então $x_0 \in S$. Portanto, $a = r_1(x_0) \in r_1(S)$, contrariando a hipótese. Segue que $p - aq$ não tem raízes múltiplas, e portanto tem $deg(\alpha)$ raízes distintas.

Como há exatamente $deg(\alpha)$ pontos (x_1, y_1) com $\alpha(x_1, y_1) = (a, b)$, o núcleo de α tem $deg(\alpha)$ elementos.

É claro que, desde que α é um homomorfismo, para cada $(a, b) \in \alpha(E(\overline{\mathbb{K}}))$, há exatamente $deg(\alpha)$ pontos (x_1, y_1) com $\alpha(x_1, y_1) = (a, b)$. As hipóteses sobre (a, b) foram feitas durante a prova para obter esse resultado para pelo menos um ponto, que é suficiente.

Se α não é separável, então os passos da prova acima são válidas, exceto que $p' - aq'$ é sempre o poninômio nulo, então $p(x) - aq(x) = 0$ sempre tem raízes multíplas e portanto tem menor do que $deg(\alpha)$ soluções.

□

Teorema 2.33. *Seja E uma curva elíptica definida sobre um corpo $\mathbb{K}$. Seja $\alpha \neq 0$ um endomorfismo de E. Então $\alpha : E(\overline{\mathbb{K}}) \to E(\overline{\mathbb{K}})$ é sobrejetora.*

Demonstração. Seja $(a, b) \in E(\overline{\mathbb{K}})$. Podemos assumir que $(a, b) \neq (0 : 1 : 0)$, pois temos $\alpha(0 : 1 : 0) = (0 : 1 : 0)$. Seja $r_1(x) = \dfrac{p(x)}{q(x)}$ como acima. Se $p(x) - aq(x)$ não é um polinômio constante, então tal polinômio tem uma raiz x_0. Desde que p e q não tenha raízes comuns, $q(x_0) \neq 0$. Escolha $y_0 \in \overline{\mathbb{K}}$ para ser a raiz quadrada de $x_0^3 + Ax_0 + B$. Então $\alpha(x_0, y_0)$ está bem definido e é igual a (a, b') para algum b'. Como $b'^2 = a^3 + Aa + B = b^2$, temos $b = \pm b'$. Se $b' = b$, acabou. Se $b' = -b$, então $\alpha(x_0, -y_0) = (a, -b') = (a, b)$.

Precisamos considerar agora o caso quando $p - aq$ é constante. Desde que $E(\overline{\mathbb{K}})$ é infinito e o núcleo de α é finito, apenas um número finito de pontos de $E(\overline{\mathbb{K}})$ pode ser mapeado para um ponto dada uma determinada coordenada x. Portanto, ou $p(x)$ ou $q(x)$ não é constante. Se p e q são dois polinômios não constantes, então há no máximo

uma constante a tal que $p - aq$ é constante. De fato, se a' é outro número tal que $p - a'q$ é constante, então $(a' - a)q = (p - aq) - (p - a'q)$ é constante e $(a' - a)p = a'(p - aq) - a(p - a'q)$ é constante, que implica que p e q são constantes. Portanto, no máximo dois pontos, (a, b) e $(a, -b)$ para algum b, que não são a imagem de α. Seja (a_1, b_1) qualquer outro ponto. Então $\alpha(P_1) = (a_1, b_1)$ para algum P_1. Conseguimos escolher (a_1, b_1) tal que $(a_1, b_1) + (a, b) \neq (a, \pm b)$, logo existe P_2 com $\alpha(P_2) = (a_1, b_1) + (a, b)$. Então $\alpha(P_2 - P_1) = (a, b)$, e $\alpha(P_1 - P_2) = (a, -b)$. Portanto, α é sobrejetora.

□

Para aplicações posteriores, precisamos de um critério conveniente de separabilidade. Se (x, y) é um ponto variável em $y^2 = x^2 + Ax + B$, então conseguimos diferenciar y com respeito a x:

$$2yy' = 3x^2 + A$$

Analogamente, podemos diferenciar uma função racional $f(x, y)$ com respeito a x:

$$\frac{d}{dx} f(x, y) = f_x(x, y) + f_y(x, y)y',$$

onde f_x e f_y denota as derivadas parciais.

Lema 2.34. *Seja E a curva elíptica $y^2 = x^3 + Ax + B$. Fixe um ponto (u, v) em E. Escreva*

$$(x, y) + (u, v) = (f(x, y), g(x, y)),$$

onde $f(x, y)$ e $g(x, y)$ são funções racionais de x, y (os coeficientes dependem de (u, v)) e y é considerado como uma função de x satisfazendo $\dfrac{dy}{dx} = \dfrac{3x^2 + A}{2y}$. Então

$$\frac{\frac{d}{dx} f(x, y)}{g(x, y)} = \frac{1}{y}.$$

Demonstração. As fórmulas 2.3 da soma de dois pontos de uma curva elíptica nos dá

$$f(x, y) = \left(\frac{y - v}{x - u}\right)^2 - x - u$$

$$g(x, y) = \frac{-(y - v)^3 + x(y - v)(x - u)^2 + 2u(y - v)(x - u)^2 - v(x - u)^3}{(x - u)^3}$$

$$\frac{d}{dx} f(x, y) = \frac{2y'(y - v)(x - u) - 2(y - v)^2 - (x - u)^3}{(x - u)^3}$$

Agora, usando o fato de que $2yy' = 3x^2 + A$ e fazendo alguns cálculos temos

$$(x - u)^3 \left(y \frac{d}{dx} f(x, y) - g(x, y)\right) =$$

$$v(Au + u^3 - v^2 - Ax - x^3 + y^2) + y(-Au - u^3 + v^2 + Ax + x^3 - y^2).$$

Desde que (u, v) e (x, y) estão em E, temos $v^2 = u^3 + Au + B$ e $y^2 = x^3 + Ax + B$. Logo, a expressão acima ficará

$$v(-B + B) + y(B - B) = 0$$

Portanto, $y\dfrac{d}{dx}f(x, y) = g(x, y)$. □

Lema 2.35. *Sejam $\alpha_1, \alpha_2, \alpha_3$ endormorfismos não nulos de uma curva elíptica E com $\alpha_1 + \alpha_2 = \alpha_3$. Escreva*

$$\alpha_j(x, y) = (R_{\alpha_j}(x), yS_{\alpha_j}(x)).$$

Suponha que existem constantes $c_{\alpha_1}, c_{\alpha_2}$ tais que

$$\frac{R'_{\alpha_1}(x)}{S_{\alpha_1}(x)} = c_{\alpha_1}, \qquad \frac{R'_{\alpha_2}(x)}{S_{\alpha_2}(x)} = c_{\alpha_2}.$$

Então

$$\frac{R'_{\alpha_3}(x)}{S_{\alpha_3}(x)} = c_{\alpha_1} + c_{\alpha_2}.$$

Demonstração. Sejam (x_1, y_1) e (x_2, y_2) pontos em E. Escreva

$$(x_3, y_3) = (x_1, y_1) + (x_2, y_2)$$

onde

$$(x_1, y_1) = \alpha_1(x, y), \qquad (x_2, y_2) = \alpha_2(x, y).$$

Então x_3 e y_3 são funções racionais de x_1, y_1, x_2, y_2, que por sua vez são funções racionais de x, y. Pelo lema 2.34, com $(u, v) = (x_2, y_2)$,

$$\frac{\partial x_3}{\partial x_1} + \frac{\partial x_3}{\partial y_1}\frac{dy_1}{dx_1} = \frac{y_3}{y_1}.$$

Analogamente,

$$\frac{\partial x_3}{\partial x_2} + \frac{\partial x_3}{\partial y_2}\frac{dy_2}{dx_2} = \frac{y_3}{y_2}$$

Por hipótese,

$$\frac{dx_j}{dx} = c_{\alpha_j}\frac{y_j}{y}, \qquad \text{onde } \; x_j = R_{\alpha_j}(x) \;\text{ e }\; y_j = yS_{\alpha_j}(x)$$

para $j = 1, 2$. Pela regra da cadeia,

$$\frac{dx_3}{dx} = \frac{\partial x_3}{\partial x_1}\frac{dx_1}{dx} + \frac{\partial x_3}{\partial y_1}\frac{dy_1}{dx_1}\frac{dx_1}{dx} + \frac{\partial x_3}{\partial x_2}\frac{dx_2}{dx} + \frac{\partial x_3}{\partial y_2}\frac{dy_2}{dx_2}\frac{dx_2}{dx} =$$

$$= \frac{y_3}{y_1}\frac{y_1}{y}c_{\alpha_1} + \frac{y_3}{y_2}\frac{y_2}{y}c_{\alpha_2} = (c_{\alpha_1} + c_{\alpha_2})\frac{y_3}{y}.$$

Dividindo por $\dfrac{y_3}{y}$ o resultado segue. □

Proposição 2.36. *Seja E uma curva elíptica definida sobre um corpo $\mathbb{K}$, e seja n um inteiro não nulo. Suponha que a multiplicação por n em E é dada por*

$$n(x, y) = (R_n(x), yS_n(x))$$

para todo $(x, y) \in E(\overline{\mathbb{K}})$, onde R_n e S_n são funções racionais. Então

$$\frac{R'_n(x)}{S_n(x)} = n.$$

Portanto, a multiplicação por n é separável se, e só se, n não é um múltiplo da característica p do corpo.

Demonstração. Primeiro, observe que $-n(x, y) = (R_{-n}(x), yS_{-n}(x)) = (R_n, -yS_n)$. Então, $R_{-n} = R_n$ e $S_{-n} = -S_n$. Assim, temos $\dfrac{R'_{-n}}{S_{-n}} = -\dfrac{R'_n}{S_n}$. Portanto, o resultado para n positivo implica o resultado para n negativo.

Note que a primeira parte da proposição é trivialmente verdade para $n = 1$, pois $(x, y) = (R_1(x), yS_1(x))$, onde $R_1(x) = x$ e $S_1(x) = 1$, logo, $\dfrac{R'_1(x)}{S_1(x)} = 1$. Agora, suponha verdade para n, então o lema 2.35 implica que, $\dfrac{R'_{n+1}(x)}{S_{n+1}(x)} = \dfrac{R'_n(x)}{S_n(x)} + \dfrac{R'_1(x)}{S_1(x)} = n + 1$. Portanto, $\dfrac{R'_n}{S_n} = n$ para todo n.

Temos $R'_n(x) \neq 0$ se, e só se, $n = \dfrac{R'_n(x)}{S_n(x)} \neq 0$, que é equivalente a p não dividir n. Como a definição de separabilidade é que $R'_n \neq 0$, isso prova a segunda parte da proposição. □

Finalmente, usaremos o lema 2.35 para provar um resultado que será necessário em seções posteriores. Seja E uma curva elíptica definida sobre um corpo finito $\mathbb{F}_q$. O endomorfismo de Frobenius ϕ_q é definido por $\phi_q(x, y) = (x^q, y^q)$. Pelo lema 2.31 temos que ϕ_q é um endomorfismo de E.

Proposição 2.37. *Seja E uma curva elíptica definida sobre $\mathbb{F}_q$, onde q é uma potência do primo p. Sejam r e s inteiros, ambos não nulos. O endomorfismo $r\phi_q + s$ é separável se, e só se, $p \nmid s$.*

Demonstração. Escreva o endormorfismo da multiplicação por r como

$$r(x, y) = (R_r(x), yS_r(x)).$$

Então

$$(R_{r\phi_q}(x), yS_{r\phi_q}(x)) = (r\phi_q)(x, y) = (\phi_q r)(x, y) = (R_r^q(x), y^q S_r^q(x)) =$$
$$= \left(R_r^q(x), y(x^3 + Ax + B)^{(q-1)/2} S_r^q(x)\right).$$

Portanto,

$$c_{r\phi_q} = \frac{R'_{r\phi_q}}{S_{r\phi_q}} = \frac{qR_r^{q-1}R'_r}{S_{r\phi_q}} = 0.$$

Além disso, $c_s = \dfrac{R'_s}{S_s} = s$ pela proposição 2.36. Pelo lema 2.35,

$$\frac{R'_{r\phi_q+s}}{S_{r\phi_q+s}} = c_{r\phi_q+s} = c_{r\phi_q} + c_s = 0 + s = s.$$

Portanto, $R'_{r\phi_q+s} \neq 0$ se, e só se, $p \nmid s$. □

3 Pontos de Torsão

Neste capítulo, estudaremos os pontos racionais de ordem finita, chamados de pontos de torsão, que desempenham um papel importante no estudo de curvas elípticas. Vamos ver no Capítulo 4 de curvas elípticas sobre corpos finitos que todos os pontos são pontos de torsão. Primeiro consideramos os casos elementares de pontos racionais de ordem de torsão 2 e 3, e então determinar a situação geral. Finalmente, discutimos o importante emparelhamento de Weil.

3.1 Pontos de Torsão

Seja E uma curva elíptica definida sobre um corpo $\mathbb{K}$. Seja n um inteiro positivo (lembre que $\overline{\mathbb{K}}$ = fecho algébrico de $\mathbb{K}$). Estamos interessado no conjunto

$$E[n] = \{P \in E(\overline{\mathbb{K}}) \mid nP = (0:1:0))$$

Enfatizamos que $E[n]$ contém pontos com coordenadas em $\overline{\mathbb{K}}$, e não somente em $\mathbb{K}$.

Quando a característica de $\mathbb{K}$ é diferente de 2, a equação geral de E pode ser transformada numa equação da forma $y^2 =$ cúbica, e é fácil determinar $E[2]$. De fato,

$$y^2 = (x - e_1)(x - e_2)(x - e_3),$$

com $e_1, e_2, e_3 \in \overline{\mathbb{K}}$. Um ponto $P = (x, y)$ satisfaz $2P = (0:1:0)$ se, e só se, a reta tangente em P é vertical. É fácil ver que isso significa que $y = 0$ (pois $2P = (0:1:0) \Leftrightarrow P = (x, y) = (x, -y) = -P \Leftrightarrow y = 0$), então

$$E[2] = \{(0:1:0), (e_1, 0), (e_2, 0), (e_3, 0)\}.$$

Como um grupo abstrato, esse é isomorfo a $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. A seguir, vamos ver que em característica 2 é mais sutil. Em [[1], p.48] é mostrado que E pode ser assumida ter uma das seguintes formas

$$(I)\ \ y^2 + xy + x^3 + a_2x^2 + a_6 = 0 \ \text{ ou } \ (II)\ \ y^2 + a_3y + x^3 + a_4x + a_6 = 0.$$

No primeiro caso, $a_6 \neq 0$ e no segundo caso, $a_3 \neq 0$ (caso contrário as curvas seriam singulares). Se $P = (x, y)$ é um ponto de ordem 2, então a reta tangente em P deve ser vertical, que significa que a derivada parcial com respeito a y deve se anular. No caso I, isso significa que $x = 0$, pois em característica 2 temos que $2y = 0$. Substitua $x = 0$ em I para obter $0 = y^2 + a_6 = (y + \sqrt{a_6})^2$. Portanto, $(0, \sqrt{a_6})$ é o único ponto de ordem 2 (raízes quadradas são únicas em característica 2), então

$$E[2] = \{(0:1:0), (0, \sqrt{a_6})\}.$$

Como um grupo abstrato, esse é isomorfo a $\mathbb{Z}_2$.

No caso II, a derivada parcial com respeito a y é $a_3 \neq 0$. Portanto, não existe nenhum ponto de ordem 2, então

$$E[2] = \{(0:1:0)\}.$$

Resumimos a discussão precedente a seguir.

Proposição 3.1. *Seja E uma curva elíptica sobre um corpo $\mathbb{K}$. Se a característica de $\mathbb{K}$ é diferente de 2, então*

$$E[2] \simeq \mathbb{Z}_2 \oplus \mathbb{Z}_2.$$

Se a característica de $\mathbb{K}$ é 2, então

$$E[2] \simeq \{0\} \quad ou \quad E[2] \simeq \mathbb{Z}_2.$$

Agora vamos olhar para o conjunto $E[3]$. Primeiro, suponha que a característica de $\mathbb{K}$ é diferente de 2 ou 3, então E pode ser dada pela equação $y^2 = x^3 + Ax + B$. Um ponto P satisfaz $3P = (0:1:0)$ se, e só se, $2P = -P$. Isso significa que a coordenada x de $2P$ é igual a coordenada x de P (a coordenada y difere no sinal; claro que, se elas são iguais, então $2P = P$, logo, $P = (0:1:0)$. Nas fórmulas de soma de pontos sobre uma curva elíptica mostrada na seção 3, temos que a coordenada x de $2P$ é $m^2 - 2x$, logo

$$m^2 - 2x = x, \quad \text{onde} \quad m = \frac{3x^2 + A}{2y}.$$

Usando o fato que $y^2 = x^3 + Ax + B$, temos que

$$\frac{(3x^2 + A)^2}{(2y)^2} = 3x \Rightarrow (3x^2 + A)^2 = 12x(x^3 + Ax + B).$$

Isso simplifica a

$$3x^4 + 6Ax^2 + 12Bx - A^2 = 0.$$

O discriminante desse polinômio é $-6912(4A^3 + 27B^2)^2$, que é diferente de zero (pois $4A^3 + 27B^2 \neq 0$). Portanto, o polinômio não tem raízes múltiplas. Existem 4 valores distintos de x (em $\overline{\mathbb{K}}$), e cada x nos dá 2 valores de y, então temos oito pontos de ordem 3. Desde que $(0:1:0)$ também está em $E[3]$, podemos ver que $E[3]$ é um grupo de ordem 9 em que todo elemento tem ordem 3. Segue que

$$E[3] \simeq \mathbb{Z}_3 \oplus \mathbb{Z}_3.$$

Agora vamos olhar quando $\mathbb{K}$ tem característica 2. Podemos assumir que E tem a forma $y^2 = x^3 + a_2x^2 + a_4x + a_6$. Novamente, queremos que a coordenada x de $2P$ seja igual a coordenada x de P. Calculamos a coordenada x de $2P$ pelo procedimento usual e igualamos

a coordenada x de P. Alguns termos desaparecem, porque $3 = 0$. Observe que derivando implicitamente a equação da curva elíptica obtemos $y' = \dfrac{2a_2x + a_4}{2y} = m$. Da mesma forma feita na seção 3, temos que a coordenada x de $2P$ é $m^2 - a_2 - 2x$. Assim, obtemos

$$\left(\frac{2a_2x + a_4}{2y}\right)^2 - a_2 - 2x = x \Rightarrow \left(\frac{2a_2x + a_4}{2y}\right)^2 - a_2 = 3x = 0.$$

Isso simplifica a $a_2x^3 + a_2a_6 - a_4^2 = 0$. Note que não podemos ter $a_2 = a_4 = 0$, pois, então $x^3 + a_6 = (x + a_6^{1/3})^3$ tem raízes múltiplas, então a_2 ou a_4 é diferente de zero.

Se $a_2 = 0$, então temos $-a_4^2 = 0$, que não pode acontecer, então não há valores de x. Portanto, $E[3] = \{(0 : 1 : 0)\}$ nesse caso.

Se $a_2 \neq 0$, então obtemos uma equação da forma $a_2(x^3 + a) = 0$, que tem uma única raiz tripla em característica 3. Portanto, há um valor de x, e dois valores correspondentes de y. Isso nos dá 2 pontos de ordem 3. Desde que $(0 : 1 : 0)$ também está em $E[3]$, então, temos que $E[3]$ tem ordem 3, portanto $E[3] \simeq \mathbb{Z}_3$ como grupo abstrato.

A situação geral é dada a seguir.

Teorema 3.2. *Seja E uma curva elíptica sobre um corpo $\mathbb{K}$ e seja n um positivo inteiro. Se a característica de $\mathbb{K}$ não divide n, ou é 0, então*

$$E[n] \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n.$$

Se a característica de $\mathbb{K}$ é $p > 0$ e $p|n$, escreva $n = p^r n'$ com $p \nmid n'$. Então

$$E[n] \simeq \mathbb{Z}_{n'} \oplus \mathbb{Z}_{n'} \quad ou \quad \mathbb{Z}_n \oplus Z_{n'}.$$

A prova desse teorema pode ser obtido em [[1], p. 79].

Seja n um inteiro positivo não divisível pela característica de $\mathbb{K}$. Escolha uma **base** $\{\beta_1, \beta_2\}$ de $E[n] \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n$. Isso significa que todo elemento de $E[n]$ é escrito da forma $m_1\beta_1 + m_2\beta_2$ com inteiros m_1, m_2. Observe que m_1, m_2 são unicamente determinados mod n. Seja $\alpha : E(\overline{\mathbb{K}}) \to E(\overline{\mathbb{K}})$ um homomorfismo. Então α leva $E[n]$ em $E[n]$. De fato, seja $P \in E[n]$, então temos $(0 : 1 : 0) = \alpha((0 : 1 : 0)) = \alpha(nP) = n\alpha(P)$, logo, $\alpha(P) \in E[n]$. Portanto, existem $a, b, c, d \in \mathbb{Z}_n$, tais que

$$\alpha(\beta_1) = a\beta_1 + c\beta_2, \qquad \alpha(\beta_2) = b\beta_1 + d\beta_2.$$

Portanto, cada homomorfismo $\alpha : E(\overline{\mathbb{K}}) \to E(\overline{\mathbb{K}})$ é representado por uma matriz 2×2

$$\alpha_{\mathbf{n}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Composição de homomorfismos corresponde a multiplicação das matrizes correspondentes.

Exemplo 3.3. Seja E uma curva elíptica definida sobre $\mathbb{R}$ dada pela equação $y^2 = x^3 - 2$, e considere $n = 2$. Então

$$E[2] = \{(0:1:0), (2^{1/3}, 0), (\zeta 2^{1/3}, 0), (\zeta^2 2^{1/3}, 0)\},$$

onde ζ é uma raiz cúbica não trivial da unidade. Defina $\beta_1 = (2^{1/3}, 0), \beta_2 = (\zeta 2^{1/3}, 0)$. Então $\{\beta_1, \beta_2\}$ é uma base de $E[2]$, pois $\beta_3 = (\zeta^2 2^{1/3}, 0) = \beta_1 + \beta_2$.

Seja $\alpha : E(\mathbb{C}) \to E(\mathbb{C})$ uma conjugação complexa: $\alpha(x, y) = (\overline{x}, \overline{y})$, onde a barra denota conjugação complexa. É fácil verificar que α é um homomorfismo. De fato, desde que todos os coeficientes das fórmulas da lei de grupo tenha coeficientes reais, temos $\overline{P_1} + \overline{P_2} = \overline{P_1 + P_2}$. Isso é o mesmo que $\alpha(P_1) + \alpha(P_2) = \alpha(P_1 + P_2)$. Temos

$$\alpha(\beta_1) = 1 \cdot \beta_1 + 0 \cdot \beta_2, \quad \alpha(\beta_2) = \beta_3 = 1 \cdot \beta_1 + 1 \cdot \beta_2.$$

Portanto, obtemos a matriz

$$\alpha_{\mathbf{2}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Observe que $\alpha \circ \alpha$ é a identidade, que corresponde ao fato de que α_2^2 é a matriz identidade mod 2.

3.2 O emparelhamento de Weil

O emparelhamento de Weil aplicado em pontos racionais de ordem de torsão n de uma curva elíptica é uma ferramenta bastante importante para o estudo destas curvas. Por exemplo, será usado no capítulo 4 para provar o importante Teorema de Hasse quanto ao número de pontos em uma curva elíptica definida sobre um corpo finito.

Seja E uma curva elíptica sobre um corpo $\mathbb{K}$ e seja n um inteiro não divisível pela característica de $\mathbb{K}$. Então $E[n] \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n$. Seja $\mu_n = \{x \in \overline{\mathbb{K}} | x^n = 1\}$ o grupo das $n-$ésimas raízes da unidade em $\overline{\mathbb{K}}$. Desde que a característica de $\mathbb{K}$ não divide n, a equação $x^n = 1$ não tem raízes múltiplas, logo, tem n raízes em $\overline{\mathbb{K}}$. Portanto, μ_n é um grupo cíclico de ordem n. Qualquer gerador ζ de μ_n é chamado de **raiz primitiva n-ésima da unidade**. Isto é equivalente a dizer que $\zeta^k = 1$ se, e só se, n divide k.

Teorema 3.4. *Seja E uma curva elíptica sobre um corpo $\mathbb{K}$ e seja n um inteiro positivo. Suponha que a característica de $\mathbb{K}$ não divida n. Então existe uma aplicação*

$$e_n : E[n] \times E[n] \to \mu_n,$$

*chamada o **emparelhamento de Weil**, que satisfaz as seguintes propriedades*

1. *e_n é bilinear em cada variável, ou seja*

$$e_n(S_1 + S_2, T) = e_n(S_1, T) e_n(S_2, T)$$

e

$$e_n(S, T_1 + T_2) = e_n(S, T_1)e_n(S, T_2)$$

para todos $S, S_1, S_2, T, T_1, T_2 \in E[n]$.

2. e_n *é não degenerado em cada variável. Isto é, se* $e_n(S, T) = 1$ *para todo* $T \in E[n]$, *então* $S = (0 : 1 : 0)$, *e também se* $e_n(S, T) = 1$ *para todo* $S \in E[n]$, *então* $T = (0 : 1 : 0)$.

3. $e_n(T, T) = 1$ *para todo* $T \in E[n]$.

4. $e_n(T, S) = e_n(S, T)^{-1}$ *para todo* $S, T \in E[n]$.

5. $e_n(\sigma S, \sigma T) = \sigma(e_n(S, T))$ *para todos automorfismos* σ *de* $\overline{\mathbb{K}}$, *tais que* σ *é o mapa identidade sobre os coeficientes de* E *(se* E *está na forma de Weierstrass, isto significa que* $\sigma(A) = A$ *e* $\sigma(B) = B$*).*

6. $e_n(\alpha(S), \alpha(T)) = e_n(S, T)^{deg(\alpha)}$ *para todos endomorfismos separáveis* α *de* E. *Se os coeficientes de* E *pertencem a um corpo finito* $\mathbb{F}_q$, *então a igualdade também vale quando* α *é o endomorfismo de Frobenius* ϕ_q.

A prova do teorema pode ser encontrado em [[1], p. 350]. Nessa seção, vamos mostrar algumas consequências do emparelhamento de Weil.

Corolário 3.5. *Seja* $\{T_1, T_2\}$ *uma base de* $E[n]$. *Então* $e_n(T_1, T_2)$ *é uma raiz primitiva* $n-$*ésima da unidade.*

Demonstração. Suponha que $e_n(T_1, T_2) = \zeta$ com $\zeta^d = 1$. Então $e_n(T_1, dT_2) = e_n(T_1, T_2 + \cdots + T_2) = e_n(T_1, T_2) \cdots e_n(T_1, T_2) = \zeta^d = 1$. Do mesmo modo, $e_n(T_2, dT_2) = e_n(T_2, T_2)^d = 1$ (de(1) e (3)). Seja $S \in E[n]$. Então, $S = aT_1 + bT_2$ para alguns inteiros a, b. Portanto,

$$e_n(S, dT_2) = e_n(aT_1 + bT_2, dT_2) = e_n(T_1, dT_2)^a e_n(T_2, dT_2)^b = 1 \cdot 1 = 1.$$

Como isso vale para todo S, (2) implica que $dT_2 = (0 : 1 : 0)$. E como $dT_2 = (0 : 1 : 0)$ se, e só se, $n|d$ (pois $T_2 \in E[n]$), segue que ζ é uma raiz primitiva $n-$ésima da unidade. □

Corolário 3.6. *Se* $E[n] \subseteq E(\mathbb{K})$, *então* $\mu_n \subset \mathbb{K}$.

Demonstração. Seja σ qualquer automorfismo de $\overline{\mathbb{K}}$ tal que σ é a identidade em $\mathbb{K}$. Sejam T_1, T_2 uma base de $E[n]$. Como $E[n] \subseteq E(\mathbb{K})$, então T_1 e T_2 têm coordenadas em $\mathbb{K}$, logo, $\sigma T_1 = T_1$ e $\sigma T_2 = T_2$. Por (5),

$$\zeta = e_n(T_1, T_2) = e_n(\sigma T_1, \sigma T_2) = \sigma(e_n(T_1, T_2)) = \sigma(\zeta).$$

O teorema fundamental da teoria de Galois diz que se um elemento $x \in \overline{\mathbb{K}}$ é fixado por todos esses automorfismos σ, então $x \in \mathbb{K}$. Portanto, $\zeta \in \mathbb{K}$. Desde que ζ é uma raiz primitiva $n-$ésima da unidade pelo corolário anterior, segue que $\mu_n \subset \mathbb{K}$. □

Agora, usamos o emparelhamento de Weil para deduzir duas proposições que serão usadas na demonstração do teorema de Hasse. Lembre que se α é um endomorfismo de E, então representaremos a matriz de α por

$$\alpha_n = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

com entradas em $\mathbb{Z}_n$, descrevendo a ação de α em relação a uma base $\{T_1, T_2\}$ de $E[n]$.

Proposição 3.7. *Seja α um endomorfismo de uma curva elíptica E definida sobre um corpo $\mathbb{K}$. Seja n um inteiro não divisível pela característica de $\mathbb{K}$. Então $det(\alpha_n) \equiv deg(\alpha)\ (mod\ n)$.*

Demonstração. Pelo corolário 3.5, $\zeta = e_n(T_1, T_2)$ é uma raiz primitiva $n-$ésima da unidade. Pela parte (6) do Teorema 3.4, temos

$$\begin{aligned} \zeta^{deg(\alpha)} &= e_n(\alpha(T_1), \alpha(T_2)) = e_n(aT_1 + cT_2, bT_1 + dT_2) \\ &= e_n(T_1, T_1)^{ab} e_n(T_1, T_2)^{ad} e_n(T_2, T_1)^{cb} e_n(T_2, T_2)^{cd} \\ &= \zeta^{ad-bc}, \end{aligned}$$

pelas propriedades do emparelhamento de Weil. Como ζ é uma raiz primitiva $n-$ésima da unidade, $deg(\alpha) \equiv ad - bc\ (mod\ n)$. □

Sejam α e β endomorfismos de E e sejam a, b inteiros. O endomorfismo $a\alpha + b\beta$ é definido por

$$(a\alpha + b\beta)(P) = a\alpha(P) + b\beta(P).$$

Aqui $a\alpha(P)$ significa multiplicação em E de $\alpha(P)$ pelo inteiro a. O resultado é então adicionado em E a $b\beta(P)$. Todo esse processo pode ser descrito por funções racionais. Portanto, $a\alpha + b\beta$ é um endomorfismo.

Proposição 3.8. $deg(a\alpha + b\beta) = a^2 deg(\alpha) + b^2 deg(\beta) + ab(deg(\alpha + \beta) - deg(\alpha) - deg(\beta))$.

Demonstração. Seja n qualquer inteiro não divisível pela característica de $\mathbb{K}$. Represente α e β por matrizes α_n e β_n (com respeito a alguma base de $E[n]$). Então $a\alpha_n + b\beta_n$ dá a ação de $a\alpha + b\beta$ em $E[n]$. Defina $\alpha_n = \begin{pmatrix} a' & b' \\ c & d \end{pmatrix}$ e $\beta_n = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$, assim, temos

$$det\ (a\alpha_n + b\beta_n) = a^2 det\ \alpha_n + b^2 det\ \beta_n + ab(det\ (\alpha_n + \beta_n) - det\ \alpha_n - det\ \beta_n)$$

para quaisquer matrizes α_n e β_n. Portanto, pela proposição anterior

$$\begin{aligned} &deg\ (a\alpha + b\beta) \equiv \\ &a^2 deg\ \alpha + b^2 deg\ \beta + ab(deg\ (\alpha + \beta) - deg\ \alpha - deg\ \beta)\ (mod\ n). \end{aligned}$$

Desde que esta equivalência é verdade para n tão grande quanto queramos, então

$$deg\ (a\alpha + b\beta) = a^2 deg\ \alpha + b^2 deg\ \beta + ab(deg\ (\alpha + \beta) - deg\ \alpha - deg\ \beta).$$

□

4 Curvas Elípticas sobre Corpos Finitos

Seja $\mathbb{F}$ um corpo finito e seja E uma curva elíptica definida sobre $\mathbb{F}$. Desde que existem somente finitos pares (x, y) com $x, y \in \mathbb{F}$, o grupo $E(\mathbb{F})$ é finito. Nesse capítulo, demonstraremos o famoso Teorema de Hasse e alguns resultados interessantes quando se trata de curva elíptica definida sobre corpos finitos. Os resultados, por si só, não são apenas interessantes, mas também é o ponto de partida para aplicações criptográficas.

4.1 Exemplos

Primeiramente, vamos considerar alguns exemplos.

Exemplo 4.1. Seja E a curva $y^2 = x^2 + x + 1$ sobre $\mathbb{F}_5$. Para contar os pontos em E, fazemos uma lista de possíveis valores para x, e então calculamos $x^3 + x + 1$ $(mod\ 5)$ e as raízes quadradas y de $x^3 + x + 1$ $(mod\ 5)$. Isto nos dá os pontos em E.

x	$x^3 + x + 1$	y	Pontos
0	1	± 1	$(0, 1), (0, 4)$
1	3	–	–
2	1	± 1	$(2, 1), (2, 4)$
3	1	± 1	$(3, 1), (3, 4)$
4	4	± 2	$(4, 2), (4, 3)$
$(0 : 1 : 0)$		$(0 : 1 : 0)$	$(0 : 1 : 0)$

Portanto, $E(\mathbb{F}_5)$ tem ordem 9.

Vamos computar $(3, 1) + (2, 4)$ em E. A inclinação da reta que passa pelos dois pontos é

$$\frac{4 - 1}{2 - 3} \equiv 2\ (mod\ 5).$$

Portanto, a reta é $y = 2(x - 3) + 1 \equiv 2x\ (mod\ 5)$. Substituindo isto em $y^2 = x^3 + x + 1$ e rearranjando, temos que $0 = x^3 - 4x^2 + x + 1$.

A soma das raízes é 4, e conhecemos as raízes 3 e 2. Portanto a raiz restante é $x = 4$. Desde que $y = 2x$, temos que $y \equiv 3$. Refletindo através do eixo x temos que a soma $(3 + 1) + (2, 4) = (4, -3) = (4, 2)$.

É claro que, podíamos ter usado as fórmulas da Seção 3 diretamente. Com poucos cálculos podemos mostrar que $E(\mathbb{F}_5)$ é cíclico, gerado por $(0, 1)$.

Exemplo 4.2. Seja E a curva elíptica $y^2 = x^3 + 2$ sobre $\mathbb{F}_7$. Então

$$E(\mathbb{F}_7) = \{(0 : 1 : 0), (0, 3), (0, 4), (3, 1), (3, 6), (5, 1)(5, 6), (6, 1), (6, 6)\}.$$

Uma conta simples mostra que todos estes pontos P satisfazem $3P = (0 : 1 : 0)$, então o grupo é isomorfo a $\mathbb{Z}_3 \oplus \mathbb{Z}_3$.

Exemplo 4.3. Vamos considerar a curva elíptica E dada por $y^2 + xy = x^3 + 1$ definida sobre $\mathbb{F}_2$. Podemos achar os pontos como antes e obter

$$E(\mathbb{F}_2) = \{\infty, (0,1), (1,0), (1,1)\}.$$

Este é um grupo cíclico de ordem 4. Os pontos $(1,0), (1,1)$ tem ordem 4 e o ponto $(0,1)$ tem ordem 2. Agora vamos olhar em $E(\mathbb{F}_4)$. Relembre que $\mathbb{F}_4$ é o corpo finito com 4 elementos. Nós podemos escrevê-lo como $\mathbb{F}_4 = 0, 1, \omega, \omega^2$, com a relação $\omega^2 + \omega + 1 = 0$ (que implica, depois de multiplicar por $\omega + 1$, que $\omega^3 = 1$). Vamos listar os elementos de $E(\mathbb{F}_4)$.

$$\begin{aligned} x &= 0 \Rightarrow y^2 = 1 \Rightarrow y = 1 \\ x &= 1 \Rightarrow y^2 + y = 0 \Rightarrow y = 0, 1 \\ x &= \omega \Rightarrow y^2 + \omega y = 0 \Rightarrow y = 0, \omega \\ x &= \omega^2 \Rightarrow y^2 + \omega^2 y = 0, \omega^2. \end{aligned}$$

Portanto,

$$E(\mathbb{F}_4) = \{(0:1:0), (0,1), (1,0), (1,1), (\omega, 0), (\omega, \omega), (\omega^2, 0), (\omega^2, \omega^2)\}.$$

Como $E(\mathbb{F}_4)$ tem ordem 8, então $E(\mathbb{F}_4)$ é isormorfo a $\mathbb{Z}_8, \mathbb{Z}_4 \oplus \mathbb{Z}_2$ ou a $\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$. Como estamos em característica 2, existe no máximo um ponto de ordem 2 (vide Proposição 3.1). E Temos que $(0,1)$ tem ordem 2. Portanto, $E(\mathbb{F}_4)$ é isomorfo a $\mathbb{Z}_8$, ou seja, $E(\mathbb{F}_4)$ é cíclico de ordem 8. Qualquer um dos 4 pontos que contém ω ou ω^2 é um gerador. Isto pode ser verificado pelo cálculo direto, ou observando que esses pontos não pertencem ao subgrupo $E(\mathbb{F}_2)$ de ordem 4. Seja $\phi_2(x, y) = (x^2, y^2)$ o mapa de Frobenius. É fácil ver que ϕ_2 permuta os elementos de $E(\mathbb{F}_4)$, e

$$E(\mathbb{F}_2) = \{(x, y) \in E(\mathbb{F}_4) \mid \phi_2(x, y) = (x, y)\}.$$

Em geral, para qualquer curva E definida sobre $\mathbb{F}_q$ e qualquer extensão $\mathbb{F}$ de $\mathbb{F}_q$, o mapa de Frobenius ϕ_q permuta os elementos de $E(\mathbb{F})$ e é a identidade sobre o grupo $E(\mathbb{F}_q)$.(Veja o Lema 4.6).

Teorema 4.4. *Seja E uma curva elíptica sobre um corpo finito $\mathbb{F}_q$. Então*

$$E(\mathbb{F}_q) \simeq \mathbb{Z}_n \quad \textit{ou} \quad E(\mathbb{F}_q) \simeq \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2}$$

para algum inteiro $n \geq 1$, ou para alguns inteiros $n_1, n_2 \geq 1$ com n_1 dividindo n_2.

Demonstração. Um resultado básico na teoria de grupos diz que um grupo abeliano finito é isomorfo a uma soma direta de grupos cíclicos

$$\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_r},$$

com $n_i | n_{i+1}$ para $1 \leq i \leq r-1$. Como, para cada i, o grupo $\mathbb{Z}_{n_i}$ tem n_1 elementos de ordem que divide n_1, pois $n_1 | n_i$, logo existe $x \in \mathbb{Z}_{n_i}$, tal que $|x| = n_1$ e todos os elementos de $< x >$ têm ordem igual a um divisor de n_1), assim, $E(\mathbb{F}_q)$ tem n_1^r elementos de ordem que divide n_1. Pelo Teorema 3.2, existe no máximo n_1^2 desses pontos (mesmo que se estivéssemos considerando coordenadas em $\overline{\mathbb{F}}_q$). Portanto, $r \leq 2$. Como queríamos demonstrar. O grupo é trivial se $r = 0$; este caso está coberto por $n = 1$ no teorema. □

4.2 Teorema de Hasse e Alguns Resultados

A seguir mostraremos uma proposição sobre corpos finitos que será usado para provar o Lema 4.6.

Proposição 4.5. *Seja $\mathbb{F}_q$ um corpo finito com característica p , $q = p^n$ e $n \in \mathbb{N}$. Dados dois elementos a, b de $\overline{\mathbb{F}}_q$ temos que*

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

e

$$\mathbb{F}_q = \{\alpha \in \overline{\mathbb{F}}_p | \alpha^q = \alpha\},$$

onde $\overline{\mathbb{F}}_p$ é o fecho algébrico de $\mathbb{F}_p$

Demonstração. Se $1 \leq j \leq p-1$, o coeficiente binomial $\binom{p}{j} = \dfrac{p!}{j!(p-j)!}$ tem um fator de p em seu numerador que não é cancelado pelo denominador, então

$$\binom{p}{j} \equiv 0 \ (mod\ p).$$

Portanto,

$$(a+b)^p = a^p + \binom{p}{1} a^{p-1}b + \binom{p}{2} a^{p-2}b^2 + \cdots + b^p = a^p + b^p$$

pois estamos trabalhando em característica p. Portanto, por indução segue que

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

para todos $x, y \in \overline{\mathbb{F}}_q$.

Agora, seja $\alpha \in \mathbb{F}_q$, sabemos que o grupo $\mathbb{F}_q^\times$ de elementos não nulos de $\mathbb{F}_q$ forma um grupo de ordem $q-1$, então $\alpha^{q-1} \equiv 1 \ (mod\ p)$ quando $0 \neq \alpha \in \mathbb{F}_q$. Portanto, $\alpha^q \equiv \alpha \ (mod\ p)$ para todo $\alpha \in \mathbb{F}_q$.

Lembre-se que o polinômio $g(X)$ tem múltiplas raízes se, e só se, $g(X)$ e $g'(X)$ têm uma raiz comum. Como

$$\frac{d}{dX}(X^q - X) = qX^{q-1} - 1 = -1,$$

pois $q = p^n = 0$ em $\mathbb{F}_p$, o polinômio não tem raízes múltiplas. Portanto, há q distintos $\alpha \in \overline{\mathbb{F}}_p$, tal que $\alpha^q = \alpha$. Uma vez que ambos os conjuntos na declaração do teorema têm q elementos e um está contido no outro, eles são iguais. □

Seja $\mathbb{F}_q$ um corpo finito com fecho algébrico $\overline{\mathbb{F}}_q$ e seja

$$\begin{aligned} \phi_q : \overline{\mathbb{F}}_q &\longrightarrow \overline{\mathbb{F}}_q, \\ x &\mapsto x^q \end{aligned}$$

o mapa de Frobenius para $\mathbb{F}_q$. Seja E uma curva elíptica definida sobre $\mathbb{F}_q$. Então ϕ_q age nas coordenadas dos pontos em $E(\overline{\mathbb{F}}_q)$:

$$\phi_q(x, y) = (x^q, y^q), \quad \phi_q((0 : 1 : 0)) = (0 : 1 : 0).$$

Lema 4.6. *Seja E definida sobre $\mathbb{F}_q$, e seja $(x, y) \in E(\overline{\mathbb{F}}_q)$.*

1. $\phi_q(x, y) \in E(\overline{\mathbb{F}}_q)$
2. $(x, y) \in E(\mathbb{F}_q)$ *se, e somente se,* $\phi_q(x, y) = (x, y)$.

Demonstração. Como a prova é a mesma para as equações de Weierstrass e as equações generalizadas de Weierstrass, trabalhamos com a forma geral. Temos

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

com $a_i \in \mathbb{F}_q$. Elevando a equação à q-ésima potência e usando a Proposição 4.5, obtemos

$$(y^q)^2 + a_1(x^qy^q) + a_3(y^q) = (x^q)^3 + a_2(x^q)^2 + a_4(x^q) + a_6.$$

Isso significa que (x^q, y^q) está em E, que prova (1). Para (2), lembre-se que $x \in \mathbb{F}_q$ se, e só se, $\phi_q(x) = x$ e analogamente para y. Lembremos que $(x, y) \in E(\overline{\mathbb{F}}_q)$, portanto

$$(x, y) \in E(\mathbb{F}_q) \Leftrightarrow x, y \in \mathbb{F}_q \Leftrightarrow \phi_q(x) = x \text{ e } \phi_q(y) = y \Leftrightarrow \phi(x, y) = (x, y)$$

□

Lema 4.7. *Seja E uma curva elíptica definida sobre $\mathbb{F}_q$. Então ϕ_q é um endomorfismo de E de grau q, e ϕ_q não é separável.*

Este é o mesmo Lema 2.31. Observe que o núcleo do endomorfismo ϕ_q é trivial. Isto está relacionado com o fato de que ϕ_q não é separável. Veja proposição 2.32.

O seguinte resultado é a chave para contagem de pontos das curvas elípticas sobre corpos finitos. Desde que ϕ_q é um endomorfismo de E, assim $\phi_q^2 = \phi_q \circ \phi_q$ e também $\phi_q^n = \phi_q \circ \phi_q \circ \circ \cdots \circ \phi_q$ para todo $n \geq 1$. Como a multiplicação por -1 é também um endomorfismo, a soma $\phi_q^n - 1$ é um endomorfismo de E.

O seguinte resultado vai ser crucial na prova do teorema de Hasse.

Proposição 4.8. *Seja E definida sobre $\mathbb{F}_q$ e seja $n \geq 1$*

1 $Ker(\phi_q^n - 1) = E(\mathbb{F}_{q^n})$.

2 Se $\phi_q^n - 1$ é um endomorfismo separável, então $\#E(\mathbb{F}_{q^n}) = deg(\phi_q^n - 1)$.

Demonstração. Como ϕ_q^n é o mapa de Frobenius do corpo $\mathbb{F}_{q^n}$, a parte (1) é apenas a reafirmação do Lema 4.6. O fato de que ϕ_q^n é separável foi provado na Proposição 2.37 . Portanto, (2) segue da Proposição 2.32. □

Lema 4.9. *Sejam r, s inteiros com $mdc(s, q) = 1$. Então $deg(r\phi_q - s) = r^2q + s^2 - rsa$, onde $a = q + 1 - \#E(\mathbb{F}_q)$.*

Demonstração. A Proposição 3.8 implica que

$$deg(r\phi_q - s) = r^2 deg(\phi_q) + s^2 deg(-1) + rs(deg(\phi_q - 1) - deg(\phi_q) - deg(-1)).$$

Como $deg(\phi_q) = q$, $deg(-1) = 1$ e pela Proposição 4.8, temos

$$deg(r\phi_q - s) = r^2q + s^2 - rsa.$$

□

Teorema 4.10. ***Teorema de Hasse.*** *Seja E uma curva elíptica sobre o corpo finito $\mathbb{F}_q$. Então a ordem de $E(\mathbb{F}_q)$ satisfaz*

$$|q + 1 - \#E(\mathbb{F}_q)| \leq 2\sqrt{q}$$

Demonstração. Como $deg(r\phi_q - s) \geq 0$, o Lema 4.9 implica que $q(\frac{r}{s})^2 - a(\frac{r}{s}) + 1 \geq 0$ para todos r, s com $mdc(s, q) = 1$. Afirmamos que o conjunto dos números racionais $\frac{r}{s}$, tal que $mdc(s, q) = 1$ é denso em $\mathbb{R}$. De fato, se $q = p^n$, onde p é um primo ímpar, tome $s = 2^m$, sendo assim, seja um intervalo $(a, b) \subset \mathbb{R}$ com $a < b$. Pela propriedade Arquimediana, existe $m \in \mathbb{N}$ tal que, $0 < \frac{1}{m} < b - a$, logo, $0 < \frac{1}{2^m} < \frac{1}{m} < b - a$, daí segue que $1 < 2^m(b - a) = 2^m b - 2^m a$, portanto, existe $r \in \mathbb{N}$, tal que $2^m a < r < 2^m b$, ou seja, $a < \frac{r}{2^m} < b$. No entanto, quando p é 2, tome $s = 3^m$ e o resultado segue analogamente.

Sendo assim, temos $qx^2 - ax + 1 \geq 0$ para todos números reais x. Portanto, o discriminante do polinômio é negativo ou 0, que significa que $a^2 - 4q \leq 0$, logo, $|a = q+1-\#E(\mathbb{F}_q)| \leq 2\sqrt{q}$, como queríamos demonstrar. □

Seja $P \in E(\mathbb{F}_q)$. A **ordem** de P é o menor inteiro positivo k tal que $kP = (0 : 1 : 0)$. Um resultado fundamental da teoria de grupos (um corolário do Teorema de Lagrange) é que a ordem de um ponto sempre divide a ordem do grupo $E(\mathbb{F}_q)$. Além disso, para um inteiro n temos $nP = (0 : 1 : 0)$ se, e só se, a ordem de P divide n. Pelo Teorema de Hasse, $\#E(\mathbb{F}_q)$ pertence a um intervalo de tamanho $4\sqrt{q}$. Portanto, se encontrarmos um ponto de ordem maior do que $4\sqrt{q}$, pode haver apenas um múltiplo dessa ordem no intervalo correto, e deve ser $\#E(\mathbb{F}_q)$. Mesmo se a ordem do ponto é menor do que $4\sqrt{q}$, obtemos uma lista pequena de possibilidades para $\#E(\mathbb{F}_q)$.

Agora, vamos aplicar o Teorema de Hasse.

Exemplo 4.11. Seja E a curva $y^2 = x^3 + 7x + 1$ sobre $\mathbb{F}_{101}$. É possível mostrar que o ponto $(0, 1)$ tem ordem 116 (utilizando [11]), então $\#E(\mathbb{F}_{101})$ é um múltiplo de 116. O Teorema de Hasse (Teorema 4.10) diz que

$$101 + 1 - 2\sqrt{101} \leq \#E(\mathbb{F}_{101}) \leq 101 + 1 + 2\sqrt{101},$$

que significa que $82 \leq \#E(\mathbb{F}_{101}) \leq 122$. O único múltiplo de 116 nesse intervalo é 116, logo, $\#E(\mathbb{F}_{101}) = 116$. Como um corolário, encontramos que o grupo de pontos é cíclico de ordem 116, gerado por $(0, 1)$.

Exemplo 4.12. Seja E a curva elíptica $y^2 = x^3 - 10x + 21$ sobre $\mathbb{F}_{557}$. O ponto $(2, 3)$ tem ordem 189. O Teorema de Hasse implica que $511 \leq \#E(\mathbb{F}_{557}) \leq 605$. O único múltiplo de 189 nesse intervalo é 567. Portanto, $\#E(\mathbb{F}_{557}) = 567$.

Exemplo 4.13. Seja E a curva elíptica $y^2 = x^3 + 7x + 12$ sobre $\mathbb{F}_{103}$. O ponto $(-1, 2)$ tem ordem 13 e o ponto $(19, 0)$ tem ordem 2. Portanto, $\#E(\mathbb{F}_{103})$ é um múltiplo de 26. O Teorema de Hasse implica que $84 \leq \#E(\mathbb{F}_{103}) \leq 124$. O único múltiplo de 26 nesse intervalo é 104, então $\#E(\mathbb{F}_{103}) = 104$.

Se tivermos que $E(\mathbb{F}_q) \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n$, então fica mais difícil de encontrar a ordem do grupo de pontos, mas esse isomorfismo é bastante raro, como mostra o próximo resultado.

Proposição 4.14. *Seja E uma curva elíptica sobre $\mathbb{F}_q$ e suponha que*

$$E(\mathbb{F}_q) \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n$$

para algum inteiro n. Então, ou $q = n^2 + 1$ ou $q = n^2 \pm n + 1$ ou $q = (n \pm q)^2$.

Demonstração. Supondo que $E(\mathbb{F}_q) \simeq \mathbb{Z}_n \oplus \mathbb{Z}_n$, então $\#E(\mathbb{F}_q) = n^2$. Pelo Teorema de Hasse, $n^2 = q + 1 - a$, com $|a| \leq 2\sqrt{q}$. Para provar a proposição, vamos usar o seguinte lema, que coloca uma restrição forte sobre a.

Lema 4.15. $a \equiv 2 \ (mod\ n)$.

Demonstração. Seja p a característica de $\mathbb{F}_q$. Então, $p \nmid n$; caso contrário, existiria p^2 pontos em $E[p]$, que é impossível em característica p pelo Teorema 3.2.

Como $E[n] \subseteq E(\mathbb{F}_q)$, o Corolário 3.6 implica que as n-ésimas raízes da unidade estão em $\mathbb{F}_q$, logo, $q-1$ deve ser um múltiplo de n (pois μ_n é um subgrupo de $\mathbb{F}_q^\times$ e a ordem de μ_n é n). Portanto, $a = q+1-n^2 = (q-1+2)-n^2 = nk+2-n^2 \equiv 2 \ (mod\ n)$.

□

Escreva $a = 2+kn$ para algum inteiro k. Então

$$n^2 = q+1-a = q-1 = kn, \qquad \text{portanto,} \quad q = n^2+kn+1.$$

Pelo Teorema de Hasse, $|2+kn| \leq 2\sqrt{q}$. Elevando ao quadrado essa última inequação, temos

$$4+4kn+k^2n^2 \leq 4q = 4(n^2+kn+1) \Rightarrow k^2 \leq 4.$$

Portanto, $|k| \leq 2$. As possibilidades $k = 0, \pm 1, \pm 2$ dá os valores de q listados na proposição. Isto completa a prova da Proposição 4.14. □

A maioria dos valores de q não são da forma dada na proposição, e mesmo para tais q a maioria das curvas elípticas $E(\mathbb{F}_q)$ não são isomorfas a $\mathbb{Z}_n \oplus \mathbb{Z}_n$ (somente uma pequena fração de curvas elípticas tem ordem n^2), portanto, podemos considerar $\mathbb{Z}_n \oplus \mathbb{Z}_n$ como um caso raro.

Teorema 4.16. *Seja E uma curva elíptica definida sobre $\mathbb{F}_q$. Seja $a = q+1-\#E(\mathbb{F}_q)$. Então, $\phi_q^2 - a\phi_q + q = 0$ como endomorfismo de E, e a é o único inteiro que satisfaz tal equação. Em outras palavras, se $(x,y) \in \overline{\mathbb{F}}_q$, então*

$$\left(x^{q^2}, y^{q^2}\right) - a(x^q, y^q) + q(x,y) = (0:1:0),$$

e a é o único inteiro tal que essa relação vale para todo $(x,y) \in \overline{\mathbb{F}_q}$. Além disso, a é o único inteiro que satisfaz $a \equiv Traço((\phi_q)_m) \ (mod\ m)$ para todo m com $mdc(m,q) = 1$.

Demonstração. Primeiro, note que, se $\phi_q^2 - a\phi_q + q$ não é o endomorfismo nulo, então seu núcleo é finito (Proposição 2.32). Vamos mostrar que se o núcleo é infinito, consequentemente o endomorfismo é 0.

Seja $m \geq 1$ um inteiro com $mdc(m,q) = 1$. Lembre que ϕ_q induz uma matriz $(\phi_q)_m$ que descreve a ação de ϕ_q em $E[m]$. Considere

$$(\phi_q)_m = \begin{pmatrix} s & t \\ u & v \end{pmatrix}$$

Como $\phi_q - 1$ é separável pela Proposição 2.37, então pelas Proposições 2.32 e 3.7 temos que

$$\begin{aligned} \#Ker(\phi_q - 1) = deg(\phi_q - 1) &\equiv det((\phi_q)_m - I) \\ &= sv - tu - (s + v) + 1 \ (mod\ m). \end{aligned}$$

Pela Proposição 3.7, $sv - tu = det((\phi_q)_m) \equiv deg(\phi_q) = q \ (mod\ m)$. Pela Proposição 4.8, $a = q + 1 - \#E(\mathbb{F}_q) = q + 1 - deg(\phi_q - 1)$, logo, $\#Ker(\phi_q - 1) = q + 1 - a$. Portanto,

$$\begin{aligned} &sv - tu - (s + v) + 1 \equiv q + 1 - a \ (mod\ m) \Rightarrow q - (s + v) + 1 \equiv q + 1 - a \ (mod\ m) \\ \Rightarrow &Traço((\phi_q)_m) = s + v \equiv a \ (mod\ m). \end{aligned}$$

Pelo Teorema de Cayley-Hamilton da álgebra linear, temos

$$(\phi_q)_m^2 - a(\phi_q)_m + qI \equiv 0 \ (mod\ m),$$

onde I é a matriz identidade 2×2. (Observe que $X^2 - aX + q$ é o polinômio característico de $(\phi_q)_m$.) Isso significa que o endomorfismo $\phi_q^2 - a\phi_q + q$ é identicamente nulo em $E[m]$. Como existem infinitas escolhas de m, o núcleo de $\phi_q^2 - a\phi_q + q$ é infinito, então o endomorfismo é 0.

Suponha $a_1 \neq 0$ que satisfaz $\phi_q^2 - a_1\phi_q + q = 0$. Então

$$(a - a_1)\phi_q = (\phi_q^2 - a_1\phi_q + q) - (\phi_q^2 - a\phi_q + q) = 0.$$

Pelo Teorema 2.33, $\phi_q : E(\overline{\mathbb{F}}_q) \to E(\overline{\mathbb{F}}_q)$ é sobrejetora. Portanto, $(a - a_1)P = (0 : 1 : 0)$, para todo $P \in E(\overline{\mathbb{F}_q})$. Com efeito, seja $P \in E(\overline{\mathbb{F}}_q)$, então existe $Q \in E(\overline{\mathbb{F}_q})$, tal que $\phi_q(Q) = P$, assim, $(a - a_1)P = (a - a_1)\phi_q(Q) = 0$. Em particular, $(a - a_1)P = (0 : 1 : 0)$ com $P \in E[m]$, para todo $m \geq 1$. Como existem pontos em $E[m]$ de ordem m quando $mdc(m, q) = 1$, temos que $a - a_1 \equiv 0 \ (mod\ m)$ para tais m. De fato, seja $P \in E[m]$, tal que a ordem de P é m, então, $kmP = 0$, com $k \in \mathbb{Z}$, logo $(a - a_1)P = Pkm = 0$, portanto, $a - a_1 = km \equiv 0 \ (mod\ m)$. Sendo assim, como a escolha de m é arbitrária, concluímos que $a - a_1 = 0$, logo, a é único. □

Destacamos o seguinte resultado, que podemos obter da prova do teorema anterior.

Proposição 4.17. *Seja E uma curva elíptica sobre $\mathbb{F}_q$ e seja $(\phi_q)_m$ a matriz dada pela ação do mapa de Frobenius ϕ_q em $E[m]$. Defina $a = q + 1 - \#E(\mathbb{F}_q)$. Então*

$$Traço((\phi_q)_m) \equiv a \ (mod\ m), \qquad det((\phi_q)_m) \equiv q \ (mod\ m).$$

O polinômio $X^2 - aX + q$ é frequentemente chamado o ***polinômio característico de Frobenius.***

Alguma vezes temos uma curva elíptica definida sobre um corpo finito $\mathbb{F}_q$ de ordem pequena e queremos conhecer a ordem de $E(\mathbb{F}_{q^n})$ para algum n. Conseguimos determinar a ordem de $E(\mathbb{F}_{q^n})$ quando $n = 1$ listando os pontos ou por algum outro procedimento elementar. O fato surpreendente é que isso nos permite determinar a ordem para todo n.

Teorema 4.18. *Considere* $\#E(\mathbb{F}_q) = q + 1 - a$. *Escreva* $X^2 - aX + q = (X - \alpha)(X - \beta)$. *Então*

$$\#E(\mathbb{F}_{q^n}) = q^n + 1 - (\alpha^n + \beta^n)$$

para todo $n \geq 1$.

Demonstração. Primeiro, precisamos mostrar que $\alpha^n + \beta^n$ é um inteiro.

Lema 4.19. *Considere* $s_n = \alpha^n + \beta^n$. *Então,* $s_0 = 2, s_1 = a,$ *e* $s_{n+1} = as_n - qs_{n-1}$ *para todo* $n \geq 1$.

Demonstração. Multiplique as relações $\alpha^2 - a\alpha + q = 0$ e $\beta^2 - a\beta + q$ por α^{n-1} e por β^{n-1}, respectivamente, para obter $\alpha^{n+1} = a\alpha^n - q\alpha^{n-1}$ e $\beta^{n+1} = a\beta^n - q\beta^{n-1}$. Agora, bastar somar as duas relações para obter o lema. □

Segue imediatamente do lema que $\alpha^n + \beta^n$ é um inteiro para todo $n \geq 0$.

Defina

$$f(X) = (X^n - \alpha^n)(X^n - \beta^n) = X^{2n} - (\alpha^n + \beta^n)X^n + q^n.$$

Então $X^2 - aX + q = (X - \alpha)(X - \beta)$ divide $f(X)$. Segue imediatamente do algoritmo da divisão de Euclides para polinômios que o quociente é um polinômio $Q(X)$ com coeficientes inteiros (os pontos principais são que o coeficiente líder de $X^2 - ax + q$ é 1 e que esse polinômio e $f(X)$ têm coeficientes inteiros.) Portanto,

$$(\phi_q^n)^2 - (\alpha^n + \beta^n)\phi_q^n + q^n = f(\phi_q) = Q(\phi_q) = Q(\phi_q)(\phi_q^2 - a\phi_q + q) = 0,$$

como endomorfismos de E, pelo Teorema 4.16. Note que $\phi_q^n = \phi_{q^n}$. Pelo Teorema 4.16, existe um único inteiro k tal que $\phi_{q^n}^2 - k\phi_{q^n} + q^n = 0$, e tal k é determinado por $k = q^n + 1 = \#E(\mathbb{F}_{q^n})$. Portanto,

$$\alpha^n - \beta^n = q^n + 1 - \#E(\mathbb{F}_{q^n}).$$

Isto completa a prova do Teorema 4.18. □

Exemplo 4.20. No Exemplo 4.3, mostramos que a curva elíptica E dada por $y^2 + xy = x^3 + 1$ sobre $\mathbb{F}_2$ satisfaz $\#E(\mathbb{F}_2) = 4$. Portanto, $a = 2 + 1 - 4 = -1$, e obtemos o polinômio

$$X^2 + X + 2 = \left(X - \frac{-1 + \sqrt{-7}}{2}\right)\left(X - \frac{-1 - \sqrt{-7}}{2}\right).$$

O Teorema 4.18 diz que

$$\#E(\mathbb{F}_4) = 4 + 1 - \left(\frac{-1+\sqrt{-7}}{2}\right)^2 - \left(\frac{-1-\sqrt{-7}}{2}\right)^2.$$

Ao invés de calcularmos a última expressão diretamente, podemos usar a recorrência do Lema 4.19

$$s_2 = as_1 - 2s_0 = -(-1) - 2(2) = -3.$$

Segue que $\#E(\mathbb{F}_4) = 4 + 1 - (-3) = 8$, que é o mesmo resultado quando calculamos por listagem de pontos.

Analogamente, usando a recorrência, temos

$$\left(\frac{-1+\sqrt{-7}}{2}\right)^{101} + \left(\frac{-1-\sqrt{-7}}{2}\right)^{101} = 2969292210605269.$$

Portanto,

$$\begin{aligned}\#E(\mathbb{F}_{2^{101}}) &= 2^{101} + 1 - 2969292210605269 \\ &= 2535301200456455833701195805484.\end{aligned}$$

A vantagem do Teorema 4.18 é que nos permite determinar a ordem do grupo para certas curvas muito rápido. A desvantagem é que exige que a curva seja definida sobre um corpo finito de ordem pequena.

5 O Logaritmo Discreto

Seja p um primo e seja a, b inteiros que são não-nulos $mod\ p$. Suponha que sabemos que existe um inteiro k tal que

$$a^k \equiv b\ (mod\ p)$$

O clássico **problema do logaritmo discreto** é achar k. Desde que $k + (p-1)$ é também uma solução, a resposta k deve ser considerada como sendo definida $mod\ p-1$, ou mod um divisor d de $p-1$ se $a^d \equiv 1(mod\ p)$.

Mais geralmente, seja G qualquer grupo, escrito multiplicativamente para o momento, e seja $a, b \in G$. Suponha que sabemos que $a^k = b$ para algum inteiro k. Neste contexto, o problema do logaritmo discreto é novamente achar k. Por exemplo, G poderia ser o grupo multiplicativo $\mathbb{F}_q^\times$ no corpo finito. Também, G poderia ser $E(\mathbb{F}_q)$ para alguma curva elíptíca, nesse caso a e b são pontos em E e nós estamos tentando achar um inteiro k com $ka = b$.

No Capítulo 6 iremos conhecer algumas aplicações criptográficas do problema do logaritmo discreto. A segurança dos criptossistemas vão depender da dificuldade de resolver o problema do logaritmo discreto.

Uma maneira de atacar o problema do logaritmo discreto é simplesmente por força bruta: tente todos os valores possíveis de k até que um funcione. Isto é impraticável quando a resposta k pode ser um inteiro de inúmeras centenas de digítos, que é um tamanho típico usado em criptografia. Portanto, técnicas melhores são necessárias.

5.1 Ataques Gerais em Logaritmos Discretos

Nesta seção, discutimos ataques que funcionam para grupos arbitrários. Primeiro, na subseção 5.1.1 é mostrado um algoritmo melhor (ou seja, que gasta menos tempo para ser executado) para calcular logaritmos discretos sobre $\mathbb{Z}_p^*$ (veja mais em [3]) e na subseção 5.1.2 como nosso foco principal são curvas elípticas, escrevemos nosso grupo G aditivamente. Portanto, são dados $P, Q \in G$ e estamos tentando resolver $kP = Q$ (sempre assumimos que k existe). Seja N a ordem de G. Geralmente, nós assumimos que N é conhecido. Por simplicidade, é geralmente assumido que P gera G.

5.1.1 Um algoritmo melhor para calcular logaritmos discretos sobre $\mathbb{Z}_p^*$

Na subseção 5.1.2 mostraremos um algoritmo que calcula logaritmos discretos que é executado no tempo $O(\sqrt{p})$. Nesta subseção iremos mostrar um algoritmo melhor que requer uma complexidade de $O(\log_2 p)$ no tempo se $p-1$ tem somente fatores primos pequenos.

A partir de agora, iremos considerar o seguinte par de funções inversas:

$$y \equiv \alpha^x (\text{mod } p)$$

$$x \equiv \log_\alpha y \text{ sobre } \mathbb{Z}_p^*$$

que são referidas como exponenciais e funções logaritmicas de base α, modulo p, onde p é primo, e α é um elemento primitivo ou gerador fixado de $\mathbb{Z}_p^*$.

5.1.1.1 Um algoritmo para $p = 2^n + 1$

Nos é dados $\alpha,\ p$, e y, com α um elemento primitivo de $\mathbb{Z}_p^*$, e devemos encontrar x tal que $y \equiv \alpha^x (\text{mod } p)$.

Podemos assumir $0 \leq x \leq p-2$, ou seja, $x \ (\text{mod } p-1) = x (\text{mod } 2^n)$.

Quando $p = 2^n + 1$, x é facilmente determinado por encontrar a expansão binária $\{b_0, \cdots, b_{n-1}\}$ de x, i.e,

$$x = \sum_{i=0}^{n-1} b_i 2^i$$

O bit menos significante b_0 de x é determinado por elevar y à potência $(p-1)/2 = 2^{n-1}$ e aplicar a regra

$$y^{(p-1)/2} (\text{mod } p) \equiv \begin{cases} +1, & b_0 = 0 \\ -1, & b_0 = 1 \end{cases}$$

Esse fato é estabelecido por observar que, como α é um elemento primitivo, então $|\alpha| = p-1$, logo, $\alpha^{p-1} \equiv \alpha^{(p-1)/2} \cdot \alpha^{(p-1)/2} \equiv 1 (\text{mod } p)$.

No entanto, $\alpha^{(p-1)/2} \equiv 1 (\text{mod } p)$ ou $\alpha^{(p-1)/2} \equiv -1 (\text{mod } p)$. Como $|\alpha| = p-1$, concluímos que $\alpha^{(p-1)/2} \equiv -1 (\text{mod } p)$, e portanto, $y^{(p-1)/2} \equiv (\alpha^x)^{(p-1)/2} \equiv (-1)^x (\text{mod } p)$.

Então, se $b_0 = 0$, temos, $y^{(p-1)/2} \equiv (-1)^x \equiv (-1)^{0+2(b_1+b_2 2+\cdots+b_{n-1}2^{n-2})} \equiv 1 (\text{mod } p)$. Da mesma forma, se $b_0 = 1$, temos, $y^{(p-1)/2} \equiv (-1)^x \equiv -1 (\text{mod } p)$.

O próximo bit na expansão de x é então determinado tomando $z \equiv y\alpha^{-b_0} \equiv \alpha^x \alpha^{-b_0} \equiv \alpha^{x_1} (\text{mod } p)$, onde $x_1 = \sum_{i=1}^{n-1} b_i 2^i$.

Claramente, x_1 é um múltiplo de quatro se, e somente se, $b_1 = 0$. Se $b_1 = 1$, então x_1 é divisível por dois, mas não por quatro.

Portanto, se $b_1 = 0$, temos, $z^{(p-1)/4} \equiv \alpha^{(x_1)(p-1)/4} \equiv \alpha^{4k(p-1)/4} \equiv \alpha^{k(p-1)} \equiv 1(\text{mod } p)$.

Da mesma forma, se $b_1 = 1$, temos, $z^{(p-1)/4} \equiv \alpha^{(x_1)(p-1)/4} \equiv \alpha^{2(1+2t)(p-1)/4} \equiv \alpha^{(1+2t)(p-1)/2} \equiv \alpha^{(p-1)/2}\alpha^{(p-1)t} \equiv -1(\text{mod } p)$

Assim, temos

$$z^{(p-1)/4}(\text{mod } p) \equiv \begin{cases} +1, & b_1 = 0 \\ -1, & b_1 = 1 \end{cases}$$

Os bits remanescentes de x são determinados de uma maneira similar, como será mostrado a seguir. Para calcular o i-ésimo bit, consideraremos $m = (p-1)/2^{i+1}$ e $z \equiv y\alpha^{-b_0-2b_1-\ldots-2^{i-1}b_{i-1}} \equiv \alpha^{x_i}(\text{mod } p)$, onde $x_i = \sum_{j=i}^{n-1} b_j 2^j$.

Assim, elevando z à m-ésima potência, temos

$$z^m \equiv \alpha^{(x_i m)} \equiv \alpha^{[(p-1)/2]\cdot(x_i/2^i)} \equiv (-1)^{x_i/2^i} \equiv$$

$$(-1)^{b_i+2(b_{i+1}+\cdots+2^{(n-1)-(1+i)}b_{n-1})} \equiv (-1)^{b_i}(\text{mod } p),$$

portanto, $z^m \equiv 1(\text{mod } p)$ se, e somente se, $b_i = 0$, e $z^m \equiv -1(\text{mod } p)$ se, e somente se, $b_i = 1$. Logo, foi encontrado a expansão binária de x.

5.1.1.2 Um algoritmo para primos arbitrários

O algoritmo mostrado acima é limitado para primos da forma $p = 2^n + 1$, porém, o maior primo conhecido dessa forma é $2^{16} + 1$. Desta forma, é necessário uma generalização desse algoritmo para primos arbitrários p.

Seja $p - 1 = p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}$, $p_i < p_{i+1}$ a fatoração prima de $p - 1$, onde os p_i's são primos distintos e os n_i's são inteiros positivos. x O valor de $x \pmod{p_i^{n_i}}$ irá ser determinado por $i = 1, \ldots, k$, e os resultados serão combinados com o Teorema Chinês do Resto para obter $x \pmod{\prod_{i=1}^{k} p_i^{n_i}} = x(\text{mod } p-1) = x$ com $0 \leq x \leq p-2$.

Considere a seguinte expansão de $x \pmod{p_i^{n_i}}$: $x \pmod{p_i^{n_i}} = \sum_{j=0}^{n_i-1} b_j p_i^j$, onde $0 \leq b_j \leq p_i - 1$.

O bit menos significante, b_0, é determinado elevando y à potência $(p-1)/p_i$,

$$y^{(p-1)/p_i} \equiv \alpha^{x(p-1)/p_i} \equiv \alpha^{(b_0+b_1p_i+\ldots+b_{n_i-1}p_i^{n_i-1})(p-1)/p_i} \equiv$$

$$\alpha^{b_0(p-1)/p_i}\alpha^{(b_1+\ldots+b_{n_i-1}p_i^{n_i-1-p_i})(p-1)} \equiv (\gamma_i)^{b_0}(\text{mod } p)$$

onde $\gamma_i = \alpha^{(p-1)/p_i}$ é uma raiz p_i-ésima primitiva da unidade. No entanto, há somente p_i valores possíveis para $y^{(p-1)/p_i} \pmod{p}$, e o valor resultante determina exclusivamente b_0.

O próximo dígito b_1 na expansão da base p_i de $x \pmod{p_i}$ é determinado de maneira similar do algoritmo anterior.

Seja $z \equiv y \cdot \alpha^{-b_0} \equiv \alpha^{x_1} (\text{mod } p)$, onde $x_1 = \sum_{j=1}^{n_i-1} b_j p_i^j$. Agora, elevando z à potência $(p-1)/p_i^2$, temos

$$z^{(p-1)/p_i^2} \equiv \alpha^{(p-1)\cdot x_1/p_i^2} \equiv (\gamma_i)^{x_1/p_i} \equiv (\gamma_i)^{b_1} (\text{mod } p)$$

Denovo, há somente p_i valores possíveis de $z^{(p-1)/p_i^2}$, e esse valor determina b_1. Esse processo cotinua para determinar todos coeficientes b_j.

Exemplo 5.1. Seja o grupo $\mathbb{Z}_{271}^* = \langle 6 \rangle$. Queremos encontrar x, tal que:

$$6^x \equiv 239 (\text{mod } 271)$$

Temos que $270 = 2 \cdot 3^3 \cdot 5$, assim, iremos fazer o processo do algoritmo anterior para resolver o logaritmo discreto.

Para $p_1 = 3^3$: então $x (\text{mod } 3^3) = \sum_{i=0}^{2} b_i 3^i$, com $0 \leq b_0 \leq 2$.

Agora, elevando 239 à potência $(271-1)/3 = 90$, temos:

$$239^{90} \equiv 28 (\text{mod } 271)$$

Também temos, $(\gamma_1)^{b_0} = (6^{90})^{b_0} \equiv (242)^{b_0} \equiv 28 (\text{mod } 271)$, ou seja, $b_0 = 2$.

Para encontrar b_1, consideraremos$z \equiv 239 \cdot 6^{-b_0=-2} \equiv 239 \cdot 226^2 \equiv 240 (\text{mod } 271)$

Agora, elevando 240 à potência $(271-1)/3^2 = 30$, temos:

$$240^{30} \equiv 242 (\text{mod } 271)$$

Também temos, $(\gamma_1)^{b_1} \equiv 242^{b_1} \equiv 242 (\text{mod } 271)$, ou seja $b_1 = 1$

Por fim, considerando $z_1 \equiv 239 \cdot 6^{-b_0=-2} \cdot 6^{-3b_1=-3} \equiv 240 \cdot 226^3 \equiv 242 (\text{mod } 271)$, de modo análogo aos bits encontrados anteriormente, concluímos que $b_2 = 1$.

Portanto, $x \equiv 2 + 3 + 3^2 \equiv 14 (\text{mod } 3^3)$

Para $p_2 = 2$ e $p_3 = 5$: de maneira análoga ao processo anterior temos que, $x \equiv 1 (\text{mod } 2)$ e $x \equiv 0 (\text{mod } 5)$

Combinando esses resultados encontrados com o Teorema Chinês do resto, temos

$$\log_6 239 \equiv 95 (\text{mod } 271).$$

5.1.2 Baby Step, Giant Step

Este método, desenvolvido por D. Shanks [4], requer aproximadamente $\sqrt{N}$ passos e por volta de $\sqrt{N}$ armazenamento. Portanto, somente funciona bem para tamanho moderado N. O procedimento é como se segue.

1. Fixe um inteiro $m \geq \sqrt{N}$ e compute mP.
2. Faça e armazene uma lista de iP para $0 \leq i < m$.
3. Compute os pontos $Q - jmP$ para $j = 0, 1, \cdots m - 1$ até que um seja igual a um elemento da lista armazenada.
4. Se $iP = Q - jmP$, temos $Q = kP$ com $k \equiv i + jm(mod\ N)$.

Por que isto funciona? Desde que $m^2 > N$, podemos assumir que a resposta k satisfaz $0 \leq k < m^2$. Escreva $k = k_0 + mk_1$ com $k_0 \equiv k\ (mod\ m)$ e $0 \leq k_0 < m$ e seja $k_1 = (k - k_0)/m$. Então $0 \leq k_1 < m$. Quando $i = k_0$ e $j = k_1$, nós temos

$$Q - k_1 mP = kP - k_1 mP = k_0 P,$$

então há uma combinação.

O ponto iP é calculado adicionando P (um "**baby step**") para $(i-1)P$. O ponto $Q - jmP$ é computado adicionando $-mP$ (um "**giant step**") para $Q - (j-1)mP$. O método foi desenvolvido por Shanks para computações na teoria dos números algébricos.

Note que não precisamos saber a ordem exata N de G. Apenas exigimos um limite superior para N. Portanto, para curvas elípticas sobre $\mathbb{F}_q$, podíamos usar este método com $m^2 \geq q + 1 + 2\sqrt{q}$, pelo Teorema de Hasse.

Exemplo 5.2. Seja $G = E(\mathbb{F}_{41})$, onde E é dado por $y^2 = x^3 + 2x + 1$. Seja $P = (0, 1)$ e $Q = (30, 40)$. Pelo teorema de Hasse's, sabemos que a ordem de G é no máximo 54, então tomamos $m = 8$. Os pontos iP para $1 \leq i \leq 7$ são

$$(0, 1), (1, 39), (8, 23), (38, 38), (23, 23), (20, 28), (26, 9).$$

Calculamos $Q - jmP$ para $j = 0, 1, 2$ e obtemos $(30, 40), (9, 25), (26, 9)$, paramos nesse terceiro ponto, pois esse ponto é igual a $7P$. Desde que $j = 2$ conseguimos a combinação, assim, temos $(30, 40) = (7 + 2 \cdot 8)P = 23P$. Portanto, $k = 23$.

6 Aplicações criptográficas

Neste capítulo, discutiremos alguns criptossistemas baseado em curvas elípticas, especialmente no problema do logaritmo discreto para curvas elípticas.

Alguém pode se perguntar porque curvas elípticas são usadas em situações de criptografia. A razão é que curvas elípticas providenciam segurança equivalente para sistemas clássicos enquanto usa-se menos bits. Por exemplo é estimado em [5] que uma chave de tamanho 4096 bits do RSA (veja [7]) dá o mesmo nível de segurança que 313 bits em um sistema de curva elíptica. Isto significa que implementações de criptossistemas de uma curva elíptica requer um chip de tamanho pequeno, menos consumo do computador, etc. Daswani e Bonech [6] realizou experimentos usando 3Com's PalmPilot, que é um dispositivo pequeno portátil, que é maior que um cartão inteligente, mas menor do que um computador laptop. Eles descobriram que gerar uma chave 512−bit RSA leva 3.4 minutos, enquanto que gerar 163-bit ECC-DSA leva 0.597 segundos. Embora certos procedimentos, tal como verificações de assinatura, tenham sido mais rápido para RSA, os métodos de curva elíptica, tal como ECC-DSA, claramente oferece um alto aumento na velocidade em várias situações.

6.1 A Configuração Básica

Alice quer mandar uma mensagem, normalmente chamada **texto simples**, para Bob. Para evitar que a bisbilhoteira Eve leia a mensagem, ela a criptografa para obter o **texto cifrado**. Quando Bob recebe o texto cifrado, ele a descriptografa e lê a mensagem. Para criptografar a mensagem, Alice usa uma **chave de criptografia**. Bob usa a **chave de descriptografia** para descriptografar o texto cifrado. Claramente, a chave de criptografia deve ser mantida em segredo para Eve.

Há dois tipos básicos de criptografia. Na **criptografia simétrica**, a chave de criptografia e de descriptografia são as mesmas, ou uma pode ser facilmente deduzida da outra. Métodos populares de criptografia simétrica são a Data Encryption Standard (DES) e a Advanced Encryption Standard (AES, normalmente referente pelo seu nome original Rijadael). Neste caso, Alice e Bob precisam ter alguma maneira para estabelecer uma chave. Por exemplo, Bob poderia enviar um mensageiro para Alice com diversos dias de antecedência. Então, quando for a hora de enviar a mensagem, ambos terão a chave. Claramente isto é imprático em várias situações.

O outro tipo de criptografia é a **criptografia de chave pública**, ou criptografia assimétrica. Neste caso, Alice e Bob não precisam ter contato prévio. Bob publica uma

chave pública de criptografia, no qual Alice usa. Ele também têm uma chave privada de criptografia que o permite descriptografar textos cifrados. Como todo mundo conhece a chave de criptografia, deveria ser inviável deduzir a chave de descriptografia a partir da chave de criptografia. O sistema mais famoso de chave pública é conhecida como RSA [7] e é baseada na dificuldade de fatoração de inteiros grandes em fatores primos. Outro sistema bem conhecido é devido ao ElGamal e é baseado na dificuldade do problema do logaritmo discreto.

Geralmente, sistemas de chave pública são mais lentos do que bons sistemas simétricos. Além disso, é comum, usar um sistema de chave pública para estabelecer uma chave, que então é usada no sistema simétrico. O melhoramento na velocidade é importante quando grandes quantidades de dados estão sendo transmitidos.

6.2 Troca de Chave Diffie-Hellman

Alice e Bob querem concordar com uma chave em comum que eles possam usar para trocas de dados via um esquema simétrico de criptografia, tal como DES ou AES. Por exemplo, Alice e Bob poderiam ser bancos que queriam transmitir dados financeiros. É impráticavel e consome tempo para usar um mensageiro para entregar a chave. Além do mais, assumimos que Alice e Bob não tenham tido contato prévio, e portanto, os únicos canais de comunicação entre eles são públicos. Um jeito de estabelecer uma chave secreta é o método seguinte, devido aos criptógrafos estadunidenses Bailey Whitfield Diffie e Martin Edward Hellman [8] (na verdade, eles usaram grupos multiplicativos de corpos finitos).

1. Alice e Bob concordam com uma curva elíptica E sobre um corpo finito $\mathbb{F}_q$ tal que o problema do logaritmo discreto é difícil em $E(\mathbb{F}_q)$. Eles também concordam com um ponto $P \in E(\mathbb{F}_q)$, tal que, o subgrupo gerado por P tem uma ordem grande (geralmente, a curva e o ponto são escolhidos de tal forma que a ordem é um primo grande).

2. Alice escolhe um inteiro secreto a, computa $P_a = aP$, e envia P_a para Bob.

3. Bob escolhe um inteiro secreto b, computa $P_b = bP$, e envia P_b para Alice.

4. Alice computa $aP_b = abP$.

5. Bob computa $bP_a = baP$.

6. Alice e Bob usa algum método aceito publicamente para extrair uma chave de abP. Por exemplo, eles poderiam usar os últimos 256 bits da x-coordenada de abP como a chave. Ou eles podiam avaliar uma função hash na x-coordenada.

A única informação que a bisbilhoteira Eve vê é a curva E, o corpo finito $\mathbb{F}_q$, e os pontos P, aP e bP. Portanto, ela precisa resolver o seguinte

PROBLEMA DIFFIE-HELLMAN

Dados P, aP, e bP em $E(\mathbb{F}_q)$, calcular abP.

Se Eve conseguir resolver logaritmos discretos em $E(\mathbb{F}_q)$, então ela pode usar P e aP para encontrar a. Assim, ela consegue computar abP. No entanto, não é conhecido uma maneira de computar abP sem que primeiro resolva um problema do logaritmo discreto.

6.3 Criptossistema Massey-Omura

Alice deseja enviar uma mensagem para Bob utilizando canais públicos. Uma maneira de fazer isso é o seguinte. Alice coloca sua mensagem em uma caixa e coloca um cadeado nela. Ela envia a caixa para Bob. Bob coloca seu cadeado na caixa e envia de volta para Alice. Então, Alice abre o cadeado dela e envia a caixa de volta para Bob. Assim, Bob remove seu cadeado, abre a caixa, e lê a mensagem.

Esse procedimento pode ser implementado matematicamente como se segue.

1. Alice e Bob escolhem uma curva elíptica E sobre um corpo finito $\mathbb{F}_q$, tal que o problema do logaritmo discreto é difícil em $E(\mathbb{F}_q)$. Defina $N = \#E(\mathbb{F}_q)$.
2. Alice representa sua mensagem como um ponto $M \in E(\mathbb{F}_q)$. (Vamos discutir como fazer isso abaixo.)
3. Alice escolhe um inteiro secreto m_A com $mdc(m_A, N) = 1$, computa $M_1 = m_A M$, e envia M_1 à Bob.
4. Bob escolhe um inteiro secreto m_B com $mdc(m_B, N) = 1$, computa $M_2 = m_B M_1$, e envia M_2 à Alice.
5. Alice computa $m_A^{-1} \in \mathbb{Z}_N$. Ela computa $M_3 = m_A^{-1} M_2$ e envia M_3 à Bob.
6. Bob computa $m_B^{-1} \in \mathbb{Z}_N$. Ele computa $M_4 = m_B^{-1} M_3$. Então, $M_4 = M$ é a mensagem.

Vamos mostrar que M_4 é a mensagem original M. Formalmente, temos

$$M_4 = m_B^{-1} m_A^{-1} m_B m_A M = M,$$

mas precisamos justificar o fato que m_A^{-1}, que é um inteiro representando o inverso de m_A^{-1} mod N, e m_A cancelam um ao outro. Temos $m_A^{-1} \equiv 1 \ (mod\ N)$, então, $m_A^{-1} m_A = 1 + kN$ para algum k. O grupo $E(\mathbb{F}_q)$ tem ordem N, logo, o Teorema de Lagrange implica que $NR = (0 : 1 : 0)$ para todo $R \in E(\mathbb{F}_q)$. Portanto,

$$m_A^{-1} m_A R = (1 + kN)R = R + (0 : 1 : 0) = R.$$

Aplicando isso a $R = m_B M$, encontramos que

$$M_3 = m_A^{-1} m_B m_A M = m_B M.$$

Analogamente, m_B^{-1} e m_B se cancelam, então

$$M_4 = m_B^{-1} M_3 = m_B^{-1} m_B M = M.$$

A bisbilhoteira Eve conhece $E(\mathbb{F}_q)$ e os pontos $m_A M, m_B m_A M$, e $m_B M$. Defina $a = m_A^{-1}, b = m_B^{-1}, P = m_A m_B M$. Então, vimos que *Eve* conhece P, bP, aP e quer encontrar abP. Isto é o Problema Diffie-Hellman.

Resta mostrar como representar uma mensagem como um ponto sobre uma curva elíptica. Usamos um método proposto por Koblitz [2]. Para corpos de característica diferente de 2 e 3, a curva elíptica é dada pela equação $y = x^3 + Ax + B$. Suponhamos que as mensagens básicas $\mathbf{m}$ são números naturais tais que $0 \leq \mathbf{m} \leq R$ com $R \in \mathbb{Z}$. Fixamos t um número natural e escolhemos um corpo $\mathbb{F}_q$, tais que $q > Rt$. Os inteiros entre 1 e Rt podem ser escritos da forma $\mathbf{m} \cdot t + j$ com $j = 1, \cdots, t-1$, assim podemos identificá-los como elementos de $\mathbb{F}_q$. Dado uma mensagem $\mathbf{m}$, tomemos $mt = x \in \mathbb{F}_q$ e calculamos o valor de $f(x) = x^3 + Ax + B$. Se $f(x)$ for um quadrado em $\mathbb{F}_q$, então tomamos $P_m = (x, y)$ com $y = \sqrt{f(x)}$; caso contrário tomamos $x = mt + 1$ e calculamos $f(x)$ e determinamos se é um quadrado em $\mathbb{F}_q$. Em geral, para algum valor $j < t$, temos que $f(x = mt + j)$ é um quadrado, então tomamos $P_m = (x, \sqrt{f(x)})$. Para recuperar a mensagem original $\mathbf{m}$ aplicamos a fórmula $\lfloor x/t \rfloor$. A probabilidade de falhar este procedimento para codificar é aproximadamente de $1/2^t$.

A seguir mostramos uma forma de implementar o criptossistema Massey-Omura, com o objetivo de dar uma ideia do uso das curvas elípticas na criptografia.

Exemplo 6.1. Consideramos a curva elíptica $E(\mathbb{F}_{751})$ definida por $y^2 = x^3 - 7x + 2$. Para codificar a mensagem identificamos as letras do alfabeto com o conjunto $\{10, 11, \cdots, 35\}$, respectivamente. O usuário **A** deseja enviar a mensagem **m= APROVADONOTCC** para o usuário **B**, assim, usando o método acima, codifica as letras como pontos da curva elíptica. Assim, $0 \leq \mathbf{m} \leq 35 = R$, escolhemos $t = 20$, então $q = 751 \geq Rt = 700$. Primeiro, mostraremos o processo feito para encontrar a codificação da letra A. Temos que A = 10, logo, $10 \cdot 20 = 200 = x$, portanto, $f(x) = 200^3 - 7(200) + 2 \equiv 452 \ (mod\ 751)$. No entanto, a equação $y^2 \equiv 452 \ (mod\ 751)$ não tem solução. Portanto, tomemos $x + 1 = 201$, calculemos $f(201) \equiv 135 \ (mod\ 751)$, mas 135 também não é um quadrado em $\mathbb{F}_{751}$. Assim, tomemos $x + 2 = 202$, temos $f(202) \equiv 273 \ (mod\ 751)$ e nesse caso, a equação $y^2 \equiv 273 \ (mod\ 751)$ tem solução, a saber, $y = 32$ ou $y = 719$. Escolhemos $y = 32$ e, portanto, $P_A = (202, 32)$. Fazendo o mesmo processo para as outras letras, obtemos $P_P = (501, 53), P_R = (541, 104), P_O = (485, 263), P_V = (621, 324), P_D = (261, 110)$, $P_N = (460, 247), P_T = (581, 283), P_C = (241, 372)$.

O usuário **A** escolhe seu inteiro secreto $m_A = 5$ ($mdc(5, \#E(\mathbb{F}_{751}) = 764) = 1$) e computa os pontos $m_AP_A = 5(202, 32) = (553, 147)$, $m_AP_P = (535, 214)$, $m_AP_R = (232, 634)$, $m_AP_O = (49, 553)$, $m_AP_V = (98, 16)$, $m_AP_D = (627, 163)$, $m_AP_N = (529, 475)$, $m_AP_T = (128, 233)$, $m_AP_C = (398, 743)$. Logo, o usuário **A** envia a seguinte mensagem criptografada (M.C.) para o usuário **B**

$M.C. = \{(553, 147), (535, 214), (232, 634), (49, 553), (98, 16), (553, 147), (627, 163), (49, 553), (529, 475), (49, 553), (128, 233), (398, 743), (398, 743)\}$.

O usuário **B** recebe a mensagem, multiplica cada ponto de (M.C.) pelo seu inteiro secreto $m_B = 3$ ($mdc(3, 764) = 1$), respectivamente, e envia a seguinte mensagem para o usuário **A**

$\{(263, 658), (627, 163), (50, 509), (111, 745), (691, 589), (263, 658), (569, 88), (111, 745)$
$(238, 456), (111, 745), (331, 64), (55, 696)(55, 696)\}$.

O usuário **A** recebe a mensagem, multiplica cada ponto por $m_A^{-1} = 153$, respectivamente, e envia a seguinte mensagem para o usuário **B**

$\{(324, 143), (261, 110), (10, 593), (237, 187), (140, 362), (324, 143), (714, 725), (237, 187)$
$(471, 374), (237, 187), (173, 109), (353, 428), (353, 428)\}$.

Agora, o usuário **B** consegue recuperar a mensagem original multiplicando esses pontos por $m_B^{-1} = 255$, respectivamente, e, por último o usuário **B** decodifica esta mensagem como foi dito anteriormente, por exemplo, para decodificar $P_A = (202, 32)$, calculemos $\lfloor 202/20 \rfloor = 10 = A$.

6.4 Criptossistema ElGamal

Alice deseja enviar uma messagem à Bob. Primeiro, Bob estabelece sua chave pública como a seguir. Ele escolhe uma curva elíptica E sobre um corpo finito $\mathbb{F}_q$ tal que o problema do logaritmo discreto é difícil em $E(\mathbb{F}_q)$. Também, ele escolhe um ponto P em E (geralmente, é escolhido P de tal forma que sua ordem seja um primo grande). Ele escolhe um inteiro secreto s e computa $Q = sP$. A curva elíptica E, o corpo finito $\mathbb{F}_q$, e os pontos P e Q são chaves públicas de Bob. Elas se tornam públicas. A chave privada de Bob é o inteiro s.

Para enviar uma mensagem à Bob, Alice faz o seguinte

1. Faz o download da chave pública de Bob.
2. Expressa sua mensagem como um ponto $M \in E(\mathbb{F}_q)$.

3. Escolhe um inteiro secreto k e computa $M_1 = kP$.

4. Computa $M_2 = M + kQ$.

5. Envia M_1, M_2 à Bob.

   Bob decifra calculando

   $$M = M_2 - sM_1$$

   Essa decifragem funciona porque

   $$M_2 - sM_1 = (M + kQ) - s(kP) = M + k(sP) - skP = M.$$

A intrusa Eve conhece as informações públicas de Bob e os pontos M_1 e M_2. Se ela conseguir calcular logaritmos discretos, então ela consegue usar P e Q para encontrar s, assim, ela consegue decifrar a mensagem calculando $M_2 - sM_1$. Além disso, ela poderia usar P e M_1 para encontrar k. E, então, ela consegue calcular $M = M_2 - kQ$. Se ela não conseguir calcular logaritmos discretos, não parece haver uma maneira de encontrar M.

É importante Alice usar um k aleatório diferente cada vez que ela envia uma mensagem para Bob. Suponha que Alice usa o mesmo k para ambos M e M'. Assim, teria $M_1 = M_1'$. Então, Eve computa $M_2' - M_2 = M' - M$. Suponha que M é um anúncio de vendas que se torna público um dia depois. Então, Eve descobre M, logo, ela calcula $M' = M - M_2 + M_2'$. Portanto, o conhecimento de M permite Eve deduzir M' nesse caso.

Segue um exemplo de uma implementação do criptossistema ElGamal.

Exemplo 6.2. Alice deseja enviar a mensagem **m = JESUS** à Bob. Primeiro, ela faz o download da chave pública de Bob, que nesse caso será, $E : y^2 = x^3 + 5x - 3$ sobre o corpo $\mathbb{F}_{839}$, $P = (154, 5)$ e $Q = 7P = (343, 44)$, depois, ela codifica a mensagem da mesma forma como foi feito no Exemplo 6.1 (escolhendo $t = 23$) e obtém

$$\mathbf{M} = \{P_J = (437, 354), P_E = (323, 38), P_S = (647, 384), P_U = (690, 243)P_S = (647, 384))\}.$$

Alice escolhe um inteiro secreto $k = 9$ e calcula $kP = 9(154, 5) = (107, 744)$. Defina $M_1 = (107, 744)$. Agora, ela soma cada ponto de $\mathbf{M}$ com $kQ = 9(343, 44) = (325, 556)$, respectivamente, e obtém-se

$$\mathbf{M}_2 = \{b_1 = (637, 606), b_2 = (152, 623), b_3 = (355, 101), b_4 = (216, 70), b_5 = (355, 101)\}.$$

E, por fim, Alice envia à Bob

$$\mathbf{M.C.} = \{(M_1, b_1), (M_1, b_2), (M_1, b_3), (M_1, b_4), (M_1, b_5)\}.$$

Bob decifra **M.C.** calculando $c_i = b_i - sM_1$, assim, ele obtém

$$\mathbf{M} = \{P_J = c_1, P_E = c_2, P_S = c_3, P_U = c_4, P_S = c_5\}.$$

Referências

[1] Washington, Lawrence. *Elliptic Curves - Number Theory and Cryptography*, vol.2 , 2008.

[2] Koblitz, Neal. *Elliptic curve cryptosystems*, Mathematics of computation 48.177 (1987): 203-209.

[3] Pohlig, Stephen, and Martin Hellman. *An improved algorithm for computing logarithms over GF (p) and its cryptographic significance (Corresp.)*, IEEE Transactions on information Theory 24.1 (1978): 106-110.

[4] Shanks, Daniel. *Class number, a theory of factorization, and genera*, Proc. of Symp. Math. Soc., 1971. Vol. 20. (1971): 415-440.

[5] Blake, Ian, et al. Elliptic curves in cryptography. Vol. 265. Cambridge university press, 1999.

[6] Daswani, Neil, and Dan Boneh. *Experimenting with electronic commerce on the PalmPilot*, International Conference on Financial Cryptography. Springer, Berlin, Heidelberg, (1999): 1-16.

[7] Coutinho, Severino Colier. *Números inteiros e criptografia RSA*, IMPA, (1997): 181-193.

[8] Diffie, Whitfield, and Martin Hellman. *New directions in cryptography*, IEEE transactions on Information Theory 22.6 (1976): 644-654.

[9] Lenstra Jr, Hendrik W. *Factoring integers with elliptic curves*, Annals of mathematics (1987): 649-673.

[10] Wiles, Andrew. *Modular elliptic curves and Fermat's last theorem*, Annals of mathematics 141.3 (1995): 443-551.

[11] http://magma.maths.usyd.edu.au/calc/