Curvas Elípticas na Criptografia

A Troca de Chaves de Diffie-Hellman por Curvas Elípticas

Enno Nagel

ICMC da Universidade Estadual de São Paulo, São Carlos,
Junho 2018

A criptografia estuda a transformação de um

texto inteligível

$\Downarrow$

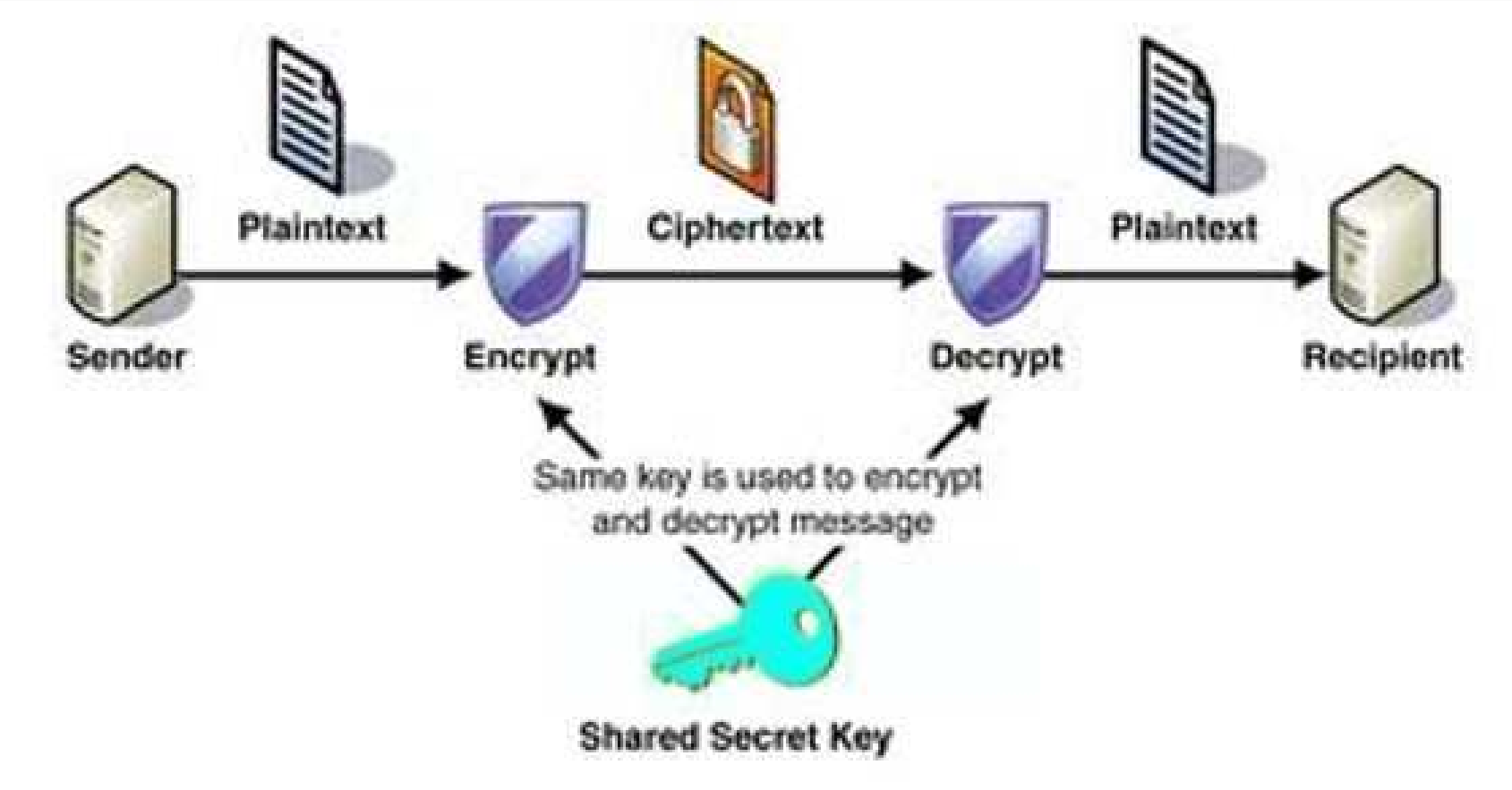
texto *in*inteligível

tal que só uma informação adicional secreta, a *chave*, permite desfazê-la.

Criptografia Simétrica

A criptografia é **simétrica** se

chave para cifrar = chave para decifrar.



A criptografia simétrica já era usada pelos egípcios 2000 anos antes de Cristo, e todos os exemplos históricos são simétricos.

Exemplos

Recordamos:

- as *Substituições do Alfabeto*,
  - por traslado,

$$\mathrm{A} \mapsto \mathrm{D}, \mathrm{B} \mapsto \mathrm{E}, \mathrm{C} \mapsto \mathrm{F}, \ldots, \mathrm{W} \mapsto \mathrm{Z}, \mathrm{X} \mapsto \mathrm{A}, \ldots, \mathrm{Z} \mapsto \mathrm{C}.$$

  - por permutação;

A	B	…	Y	Z
↓	↓	…	↓	↓
E	Z	…	G	A

- a *Transposição do Texto* (claro) pela cítala, que torna linhas em colunas; por exemplo,

l	u	a
m	e	l

transforma-se em

l	m
u	e
a	l

Figura 1: A cítala enrolada por um cinto de couro

Criptografia Moderna

A criptografia estuda a transformação de

dados inteligíveis

$$\Downarrow$$

dados *in*inteligíveis

tal que só uma informação adicional secreta, a **chave**, permite desfazê-la.

Dados

Outrora (na época análoga):

dados = textos.

Hoje (na época digital):

dados = arquivo digital (de texto, imagem, som, vídeo, ...)
= sequência de bits (= 0,1)
= sequência de bytes (= 00, 01, ..., FE, FF)
= número (= 0,1,2,3 ...).

Codificação de Textos

Texto inglês (= sequência de letras latinos)

A codificação **ASCII** cobre todas as letras inglesas (além dos símbolos de pontuação, por exemplo) e envia um símbolo a um byte. Por exemplo,

$$\mathrm{A} \mapsto 65, ..., \mathrm{Z} \mapsto 90; a \mapsto 97, ..., z \mapsto 122.$$

Texto Internacional (= sequência de símbolos)

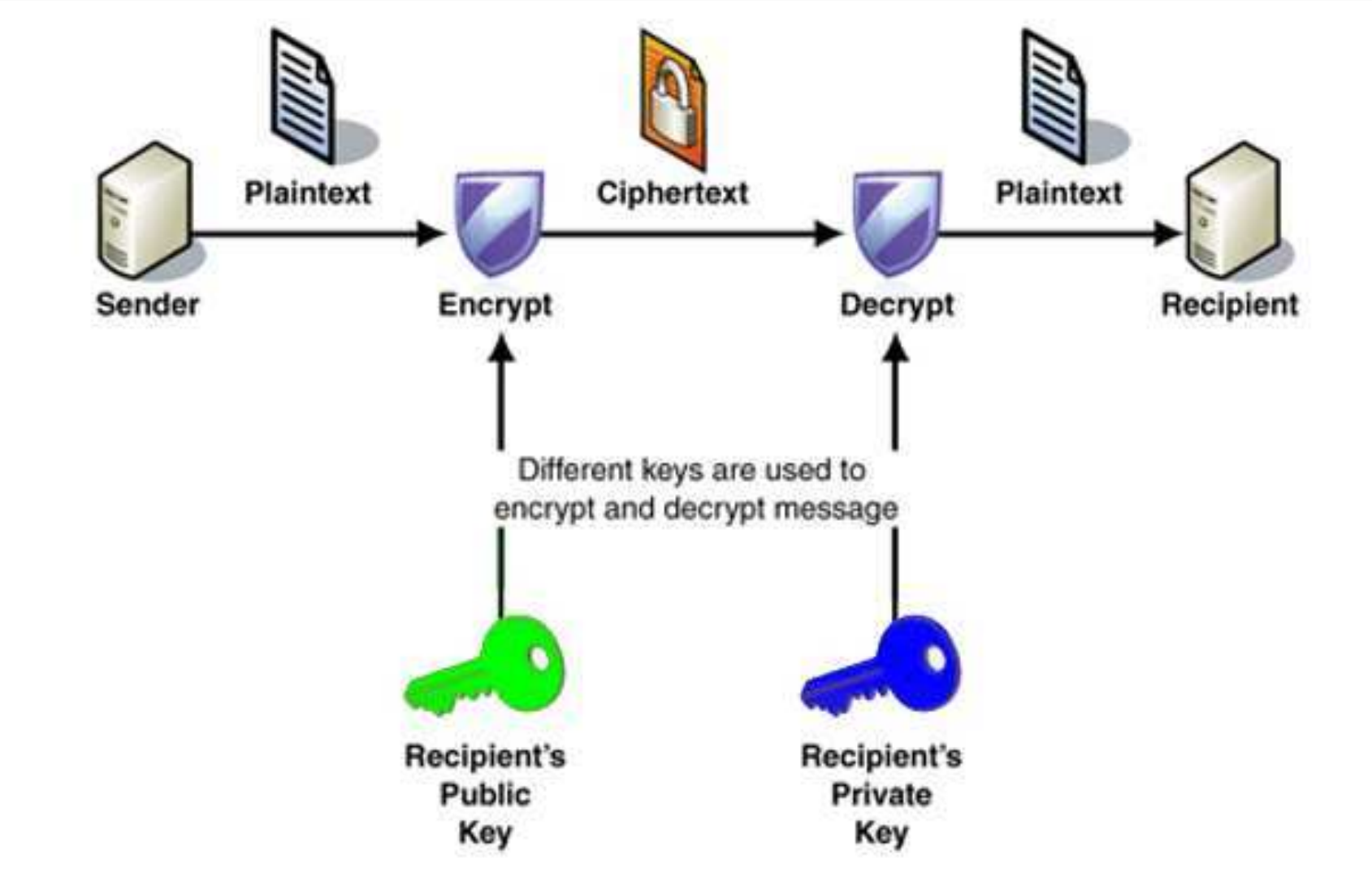
A codificação **UTF-8** inclui a ASCII e cobre todas as letras de todos os idiomas (por exemplo, chinês, coreano, ... e, além disso, por exemplo todos os símbolos matemáticos).
Ela envia um símbolo a $1, 2, 3$ ou até 4 bytes, onde

$$\#\{\text{bytes}\} = \#\{\text{dígitos bin. consecutivos iniciais iguais a } 1\} + 1$$

Criptografia Assimétrica

A criptografia é **assimétrica** se

chave para cifrar ≠ chave para decifrar.



Comparação Criptografia Simétrica e Assimétrica

Como vantagem em comparação a criptografia simétrica, a criptografia assimétrica

- evita o risco de comprometimento da chave secreta na troca da chave com o cifrador.

Por outro lado,

- a um nível de segurança comparável, os algoritmos simétricos são mais rápidos que os assimétricos.

$\Longrightarrow$ na prática, comunica-se

1. com **chaves assimétricas** (para trocar uma chave simétrica)
2. com **uma chave simétrica**.

Inventores da Ideia da Criptografia Assimétrica

A criptografia assimétrica foi sugerida pela primeira vez, publicamente, em Diffie e Hellman (1976). Antes (uns 20 anos), somente os serviços secretos tiveram consciência desta técnica.

Figura 2: Diffie e Hellman

O Anel dos Números Inteiros

Denotenão

$$\mathbf{Z} = \{..., -2, -1, 0, 1, 2, ...\} \quad \text{e} \quad \mathbf{R}$$

os *anéis* dos números inteiros e dos números reais (= a reta).

Anel

Um **anel** (comutativo com 1) é

- um conjunto que contém
  - $\mathbf{0}$ (= o elemento neutro da adição), e
  - $\mathbf{1}$ (= o elemento neutro da multiplicação);
- sobre o qual operam
  - a **adição** $+$ (e o seu *inverso* $-$), e
  - a **multiplicação** $\cdot$,

  que satisfazem a lei *associativa*, *comutativa* e *distributiva*.

Funções sobre Conjuntos Discretos

Na criptografia assimétrica,

- a *facilidade* de cifrar (um número), e
- a *dificuldade* de decifrar (um número)

baseiam-se em uma função invertível tal que

- ela é **facilmente** computável, mas
- a sua função inversa é **dificilmente** computável.

Exemplos para tais funções são

- a *exponenciação* $x \mapsto g^x$ (no algoritmo `Diffie-Hellman`), e
- a *potenciação* $x \mapsto x^e$ (no algoritmo `RSA`)

mas **não** definidas sobre $\mathbf{Z}$, porque $\mathbf{Z}$ é incluso em $\mathbf{R}$ e ambas as funções, exponenciação e potenciação, são **contínuas** sobre $\mathbf{R}$.

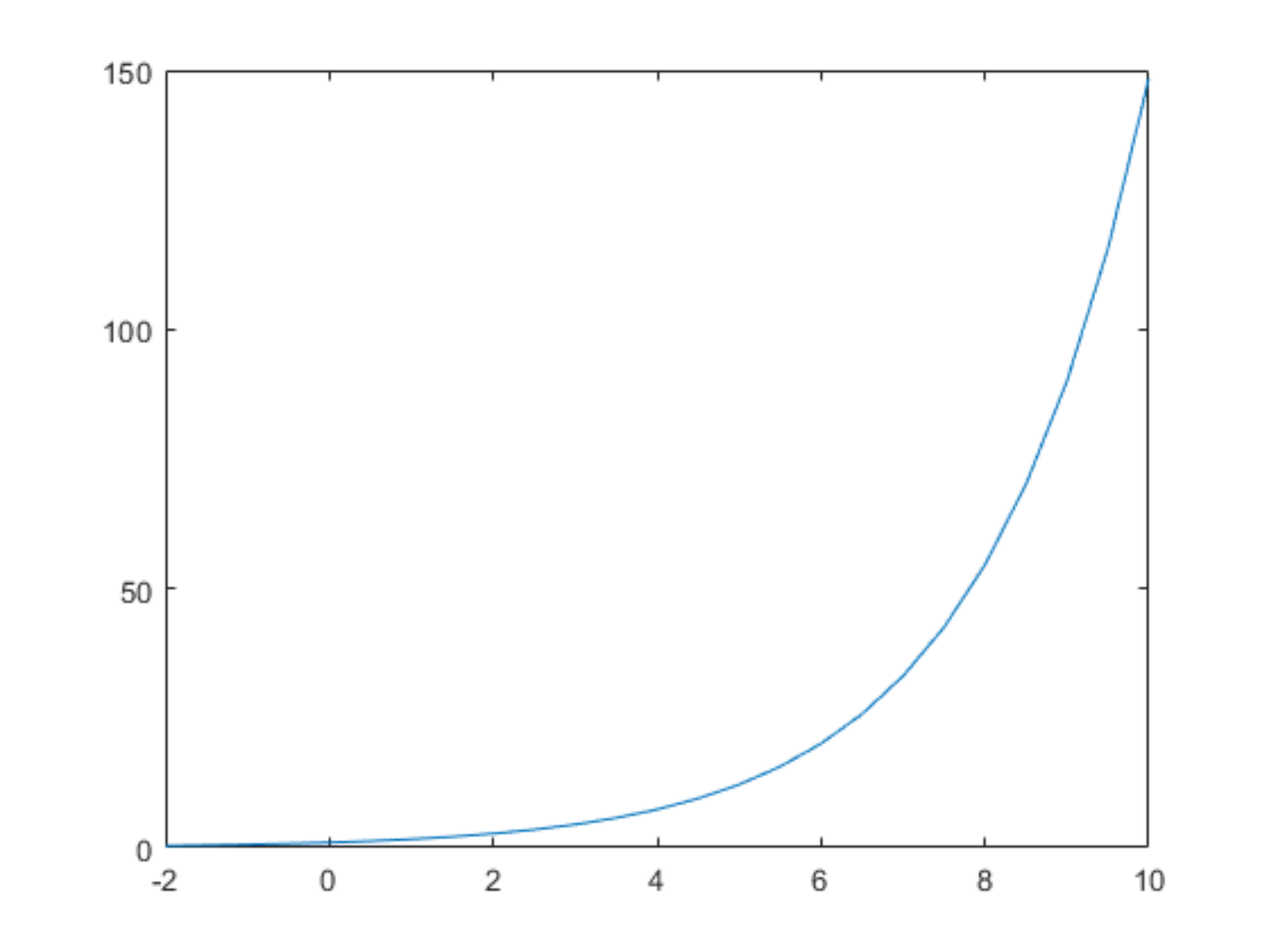


Figura 3: Exponencial $x \mapsto e^x$

Se estas funções fossem definidas sobre o anel $\mathbf{Z}$ e o qual é incluso em $\mathbf{R}$, os seus **inversos** podiam ser **aproximados** sobre $\mathbf{R}$, por exemplo, pelo **Método da Bisseção**: Dado y_0, encontrar um x_0 tal que $f(x) = y_0$ equivale encontrar um **zero** x_0 da função

$$x \mapsto f(x) - y_0.$$

1. (*Inicialização*) Escolha um intervalo $[x', x'']$ tal que

$$f(x') < 0 \quad \text{e} \quad f(x'') > 0.$$

2. (*Recalibração*) Calcule o seu meio $x''' := \frac{x'+x''}{2}$. Vale
   - ou $f(x''') = 0$, então $x''' = x_0$,
   - ou $f(x''') < 0$, então substituta o bordo esquerdo x' por x''',
   - ou $f(x''') > 0$, então substituta o bordo direito x'' por x'''.

   e itere com os novos bordos do intervalo.

Pelo **Teorema do Valor Intermediário**, o zero é garantido de estar no intervalo, que a cada passo diminui e converge à interseção.

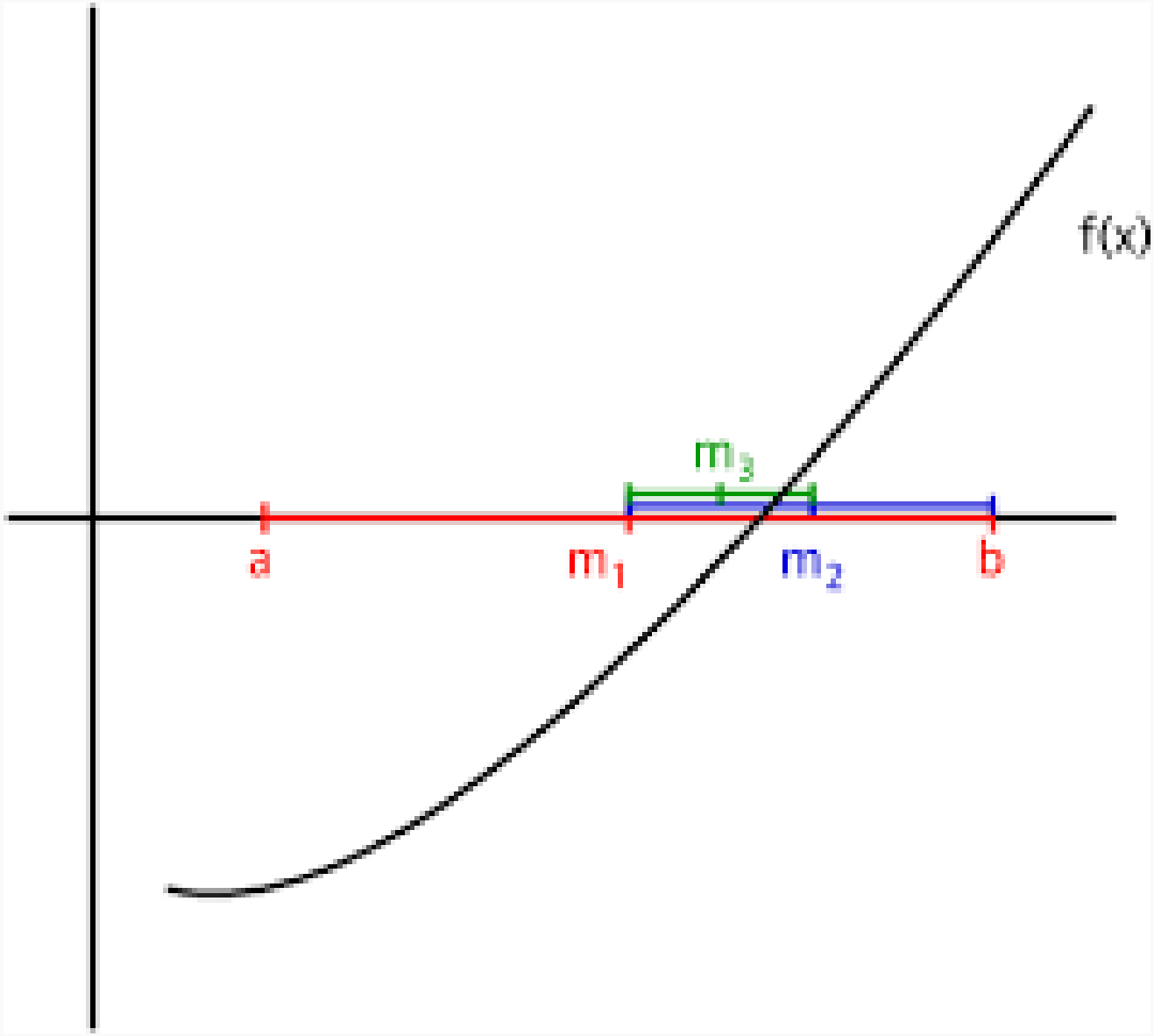


Figura 4: Bisseção de uma função contínua

Anéis Finitos

Para evitar a iterada aproximação de zeros e assim **dificultar** a computação do inverso (além de facilitar a computação da função), a criptografia assimétrica desenrola-se sobre os **anéis finitos**

$$\mathbf{Z}/m\mathbf{Z} = \{0, 1, ..., m-1\}.$$

Neles, vale $m = 1 + \cdots + 1 = 0$ e **toda adição (e logo toda multiplicação e toda potenciação) tem resultado** $< m$ e assim $\mathbf{Z}/m\mathbf{Z}$ é como anel **não** incluso em $\mathbf{R}$. Por exemplo, para $m = 7$, vale

$$2^2 = 2 \cdot 2 = 4 \quad \text{e} \quad 3^2 = 3 \cdot 3 = 7 + 2 = 0 + 2 = 2.$$

Introduzamos estes anéis finitos primeiro pelo exemplo $\mathbf{Z}/12\mathbf{Z}$ (= o anel dos números das horas do relógio) e depois em geral.

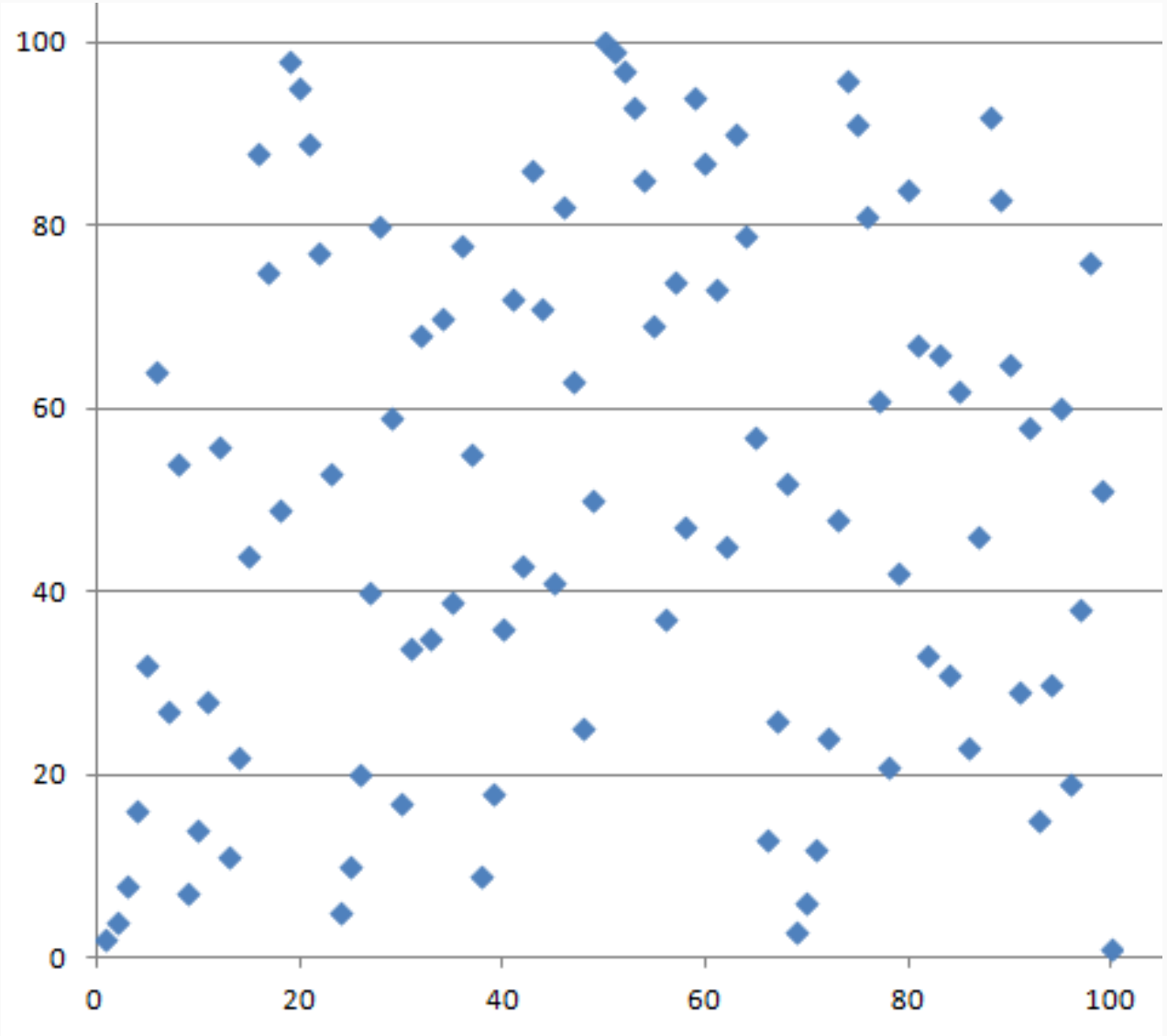


Figura 5: O Exponencial com base 2 em $\mathbf{Z}/101\mathbf{Z}$

Relógio = Anel dos números $1, 2, ..., 11, 12 = 0$

Figura 6: Relógio como Anel dos números $1, 2, ..., 11, 12 = 0$

Aritmética do Relógio

O exemplo prototípico de aritmética modular é a aritmética do relógio, em que vale a equação

$$12 = 0,$$

e que implica, entre outros, as equações

$$9 + 4 = 1 \quad \text{e} \quad 1 - 2 = 11;$$

Quer dizer,

- 4 horas depois das 9 horas é 1 hora, e
- 2 horas antes da 1 hora são 11 horas.

Formalmente, derivamos estas equações das igualdades

$$9+4 = 13 = 12+1 = 0+1 = 1 \quad \text{e} \quad 1-2 = -1+0 = -1+12 = 11.$$

Podemos ir mais longe: $9 + 24 = 9$, quer dizer se agora são 9 horas, então 24 horas mais tarde também. Formalmente,

$$9 + 24 = 9 + 2 \cdot 12 = 9 + 2 \cdot 0 = 9.$$

Em geral, para quaisquer a e x em $\mathbf{Z}$,

$$a + 12 \cdot x = a$$

ou, equivalentemente, para quaisquer a e b em $\mathbf{Z}$,

$$a = b \quad \text{se } 12 \mid a - b$$

Congruência Modular

Seja $m \geq 1$ um inteiro. Os números inteiros a e b são ***congruentes módulo*** *m ou, em fórmulas,*

$$a \equiv b \mod m.$$

se $m \mid a - b$, isto é se a sua diferença $a - b$ é divisível por m. O número m é o ***módulo***.

O Anel Quociente

Dado um número inteiro m, construiremos de duas vias, primeiro pela teórica e depois pela prática,

- o **menor** anel (= conjunto que contém 0 e 1 e sobre o qual $+$ e $\cdot$ operam), denotado por $\mathbf{Z}/m\mathbf{Z}$,
- com **uma aplicação** $\bar{\cdot} \colon \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$ denotada

$$x \mapsto \bar{x}$$

que **satisfaz**

$$\overline{x+y} = \overline{x} + \overline{y}$$

(e portanto $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$, em particular, $\bar{0} = 0$ e $\bar{1} = 1$), e

- tal que

$$x \mapsto 0 \iff m \mid x \qquad (\dagger)$$

ou, equivalentemente,

$$x \equiv y \bmod m \iff \bar{x} = \bar{y} \text{ em } \mathbf{Z}/m\mathbf{Z}.$$

O Anel $\mathbf{Z}/7\mathbf{Z}$ para $m = 7$

Figura 7: Os comprimidos semanais

O Anel $\mathbf{Z}/m\mathbf{Z}$ para $m = 15$

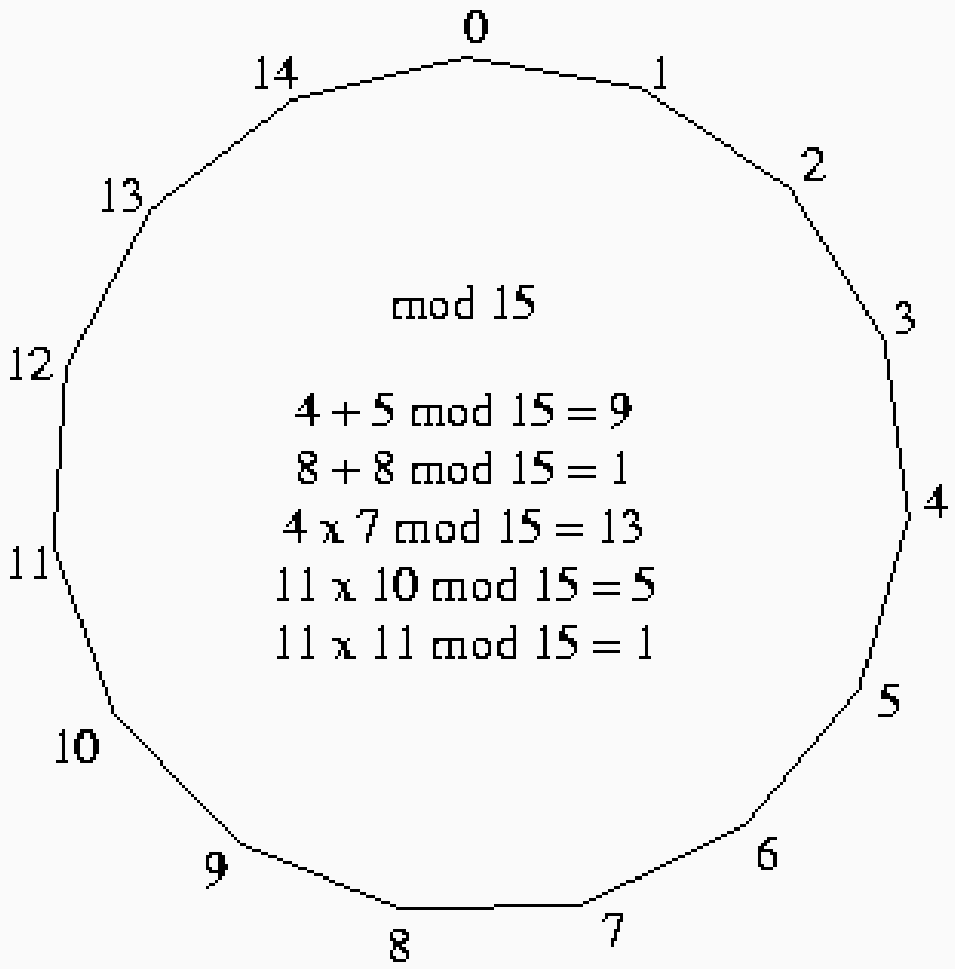


Figura 8: O relógio com 15 horas

Construção *Prática*

- como **conjunto**

$$\mathbf{Z}/m\mathbf{Z} := \{0, ..., m-1\};$$

e como **aplicação** $\bar{\cdot}\colon \mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$

$$x \mapsto r \quad \text{onde } x = qm + r \quad \text{com } r \text{ em } \{0, ..., m-1\};$$

- como **elementos neutros** da adição e multiplicação 0 e 1,
- como **operações** $+$ e $\cdot$

$$x+y := r \quad \text{onde } x+y = qm+r \quad \text{com } r \text{ em } \{0, ..., m-1\}, \text{ e}$$

$$x \cdot y = r \quad \text{onde } x \cdot y = qm + r \quad \text{com } r \text{ em } \{0, ..., m-1\}.$$

O anel $\mathbf{Z}/p\mathbf{Z}$ para p primo é um *corpo*, isto é, nele podemos dividir por todos os números (exceto 0); é denotado $\mathbf{F}_p$.

Tabelas de Adição e Multiplicação para $m = 4$

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

*	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Troca de Chaves de Diffie-Hellman

Para Alice e Bob construírem uma chave secreta através de um canal inseguro, combinam em primeiro lugar

- um número primo p *apropriado*, e
- um número natural g *apropriado*.

1. Alice, pra gerar **uma metade** da chave, escolhe um número a,
   - calcula $\mathrm{A} \equiv g^a \bmod p$, e
   - transmite A ao Bob.
2. Bob, pra gerar **outra metade** da chave, escolhe um número b,
   - calcula $\mathrm{B} \equiv g^b \bmod p$, e
   - transmite B à Alice.
3. A chave secreta **mútua** entre Alice e Bob é

$$c := \mathrm{A}^b = (g^a)^b = g^{ab} = g^{ba} = (g^b)^a = \mathrm{B}^a \bmod p.$$

Dificuldade = Computar o Logaritmo módulo p

Um olheiro obteria a chave secreta $c = \mathrm{A}^b = \mathrm{B}^a$ a partir de A e B, se pudesse computar

$$a = \log_g \mathrm{A} \quad \text{ou} \quad b = \log_g \mathrm{B} \mod p;$$

isto é, o *logaritmo* $\log_g$ inverte a *potenciação* $x \mapsto g^x = y$,

$$\log_g y = x \quad \text{com } x \text{ tal que } g^x = y.$$

Enquanto a potenciação é facilmente computável, o **logaritmo é dificilmente computável** para escolhas de p e g **apropriadas**:

- o número primo p
  - seja **grande** e
  - exista um número primo **grande** tem de dividir $p - 1$.
- as potências da base g gerem um **grande** conjunto (finito)

$$\{g, g^2, g^3, \ldots\}.$$

Números Apropriados

O Teorema de Euclides garante que existam números primos arbitrariamente **grandes** (> 1024 bits),

$$\#\{ \text{números primos} \} = \infty$$

e quase todo número primo p satisfaz que

- exista um primo **grande** (> 768 bits) que divide $p - 1$.

O Teorema da *Raiz Primitiva* garante que sempre exista g com

$$\mathbf{F}_p^* = \{1, 2, 3, ..., p-1\} = \{g, g^2, g^3, ..., g^{p-1}\},$$

Em particular, se p é grande, então o conjunto gerado pelas potências da base g é **grande** (= $p - 1 \geq 1024$ bits).

Na prática, p e g são adotados de uma fonte confiável.

Curvas Elípticas

Uma curva E sobre um corpo (de característica $\neq 2, 3$) é *elíptica* se dada por uma equação

$$y^2 = x^3 + ax + b$$

para coeficientes a e b tais que a curva não seja *singular*, isto é, que a sua *discriminante* não desvaneça, $4a^3 + 27b^2 \neq 0$.

Após escolha de um domínio (por exemplo, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ ou $\mathbf{F}_p$ para um número primo p) os pontos (x, y) que solvem esta equação formam uma curva no plano sobre ele.

Além dos pontos no plano, existe também o ponto no infinito (ou ponto *ideal*) que é denotado 0. Resumimos que, como conjunto, a curva elíptica é dada por,

$$\mathrm{E} := \{(x, y) : \mathrm{E}(x, y) = 0\} \cup \{0\}$$

Sobre um corpo finito $\mathbf{F}_q$, o número dos pontos $\#\mathrm{E}$ é limitado por $q + 1 - t$ onde $t \leq 2\sqrt{q}$, isto é, assimptoticamente igual a $\mathbf{F}_q^* = q - 1$. (A computação de $\#\mathrm{E}$ leva pelo algoritmo de Schoof $\mathrm{O}(n^5)$ operações com $n = \log_2 q$ o número de dígitos binários de q.)

Para o domínio $\mathbf{R}$, as curvas assumem as seguintes formas no plano real ao a e b variarem:

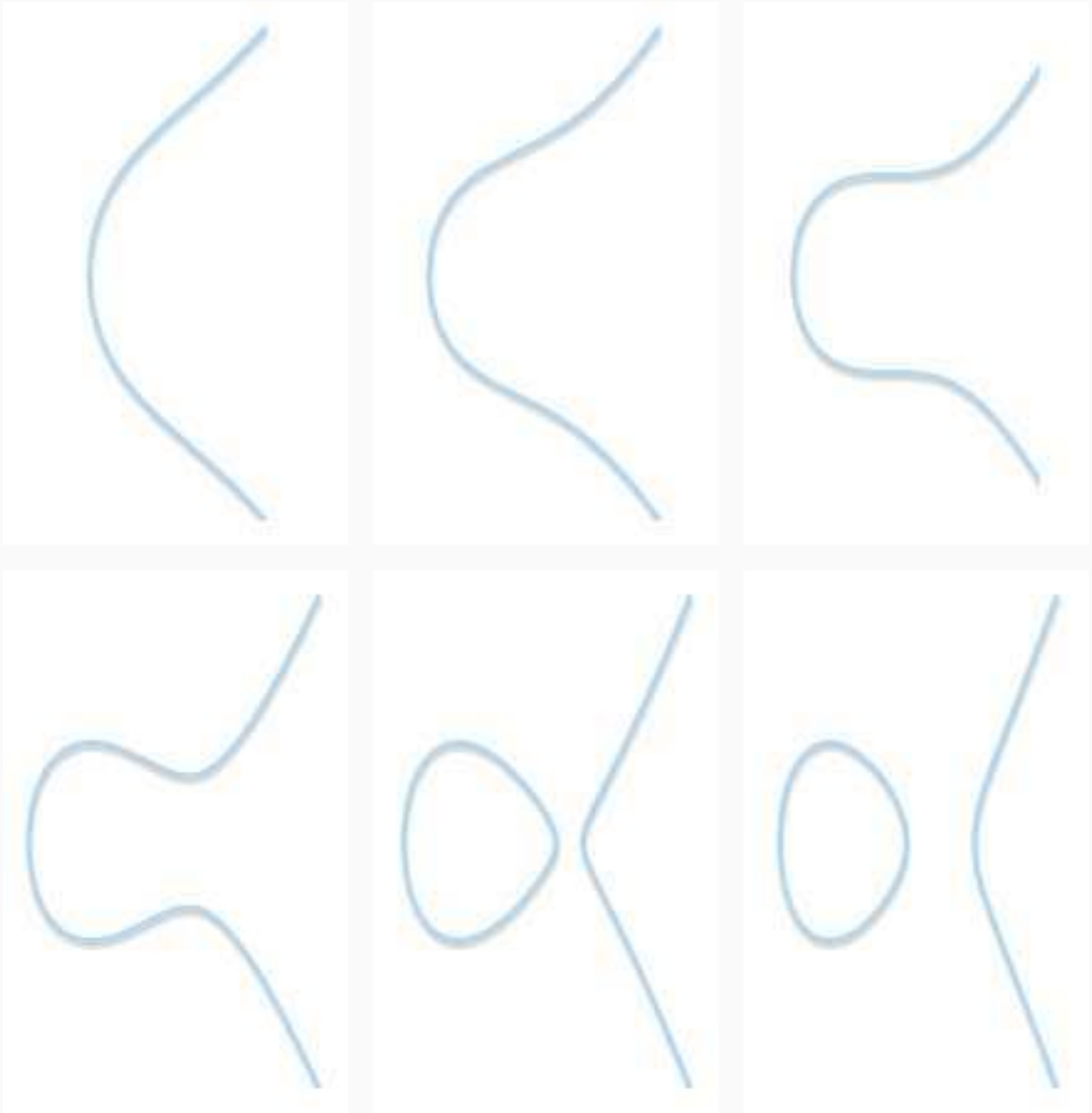


Figura 9: Curvas elípticas reais

Enquanto sobre os corpos finitos, obtemos um conjunto discreto de pontos (simétrico em volta do eixo horizontal médio).

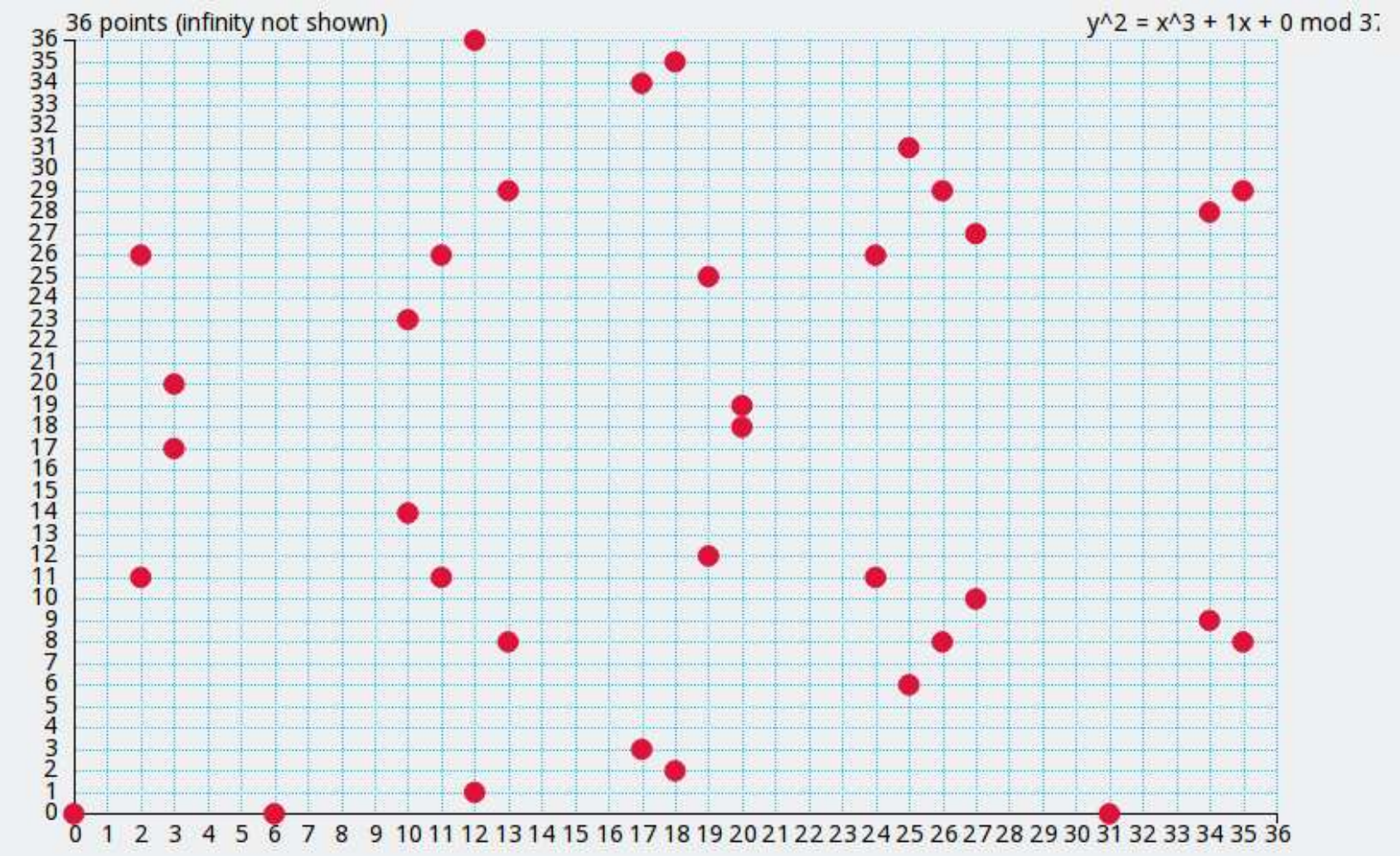


Figura 10: Curva elíptica modular

Exemplo usado na Criptografia

A Curve25519 com

$$y^2 = x^3 + 486662x^2 + x$$

sobre $\mathbf{F}_p$ com $p = 2^{255} - 19$ (de onde o seu nome); tem ordem

$$\#\mathrm{E}/8 = 2^{252} + 27742317777372353535851937790883648493.$$

Adição Geométrica

Entre todas as curvas, a graça das elípticas (dadas por uma equação $y^2 = x^3 + ax + b$) é que permitem somar pontos ($p + q + r = 0$ se uma reta passa por p, q e r). O grupo dado por uma curva elíptica é obtido assim:

- os elementos são os seus pontos, isto é, os pares (x, y) com entradas no corpo que satisfazem a equação;
- o elemento neutro 0 é (geralmente) o ponto infinito $(0, 0)$;
- a adição é dada, geometricamente por $p + q + r = 0$ se os pontos p, q e r são colineares.

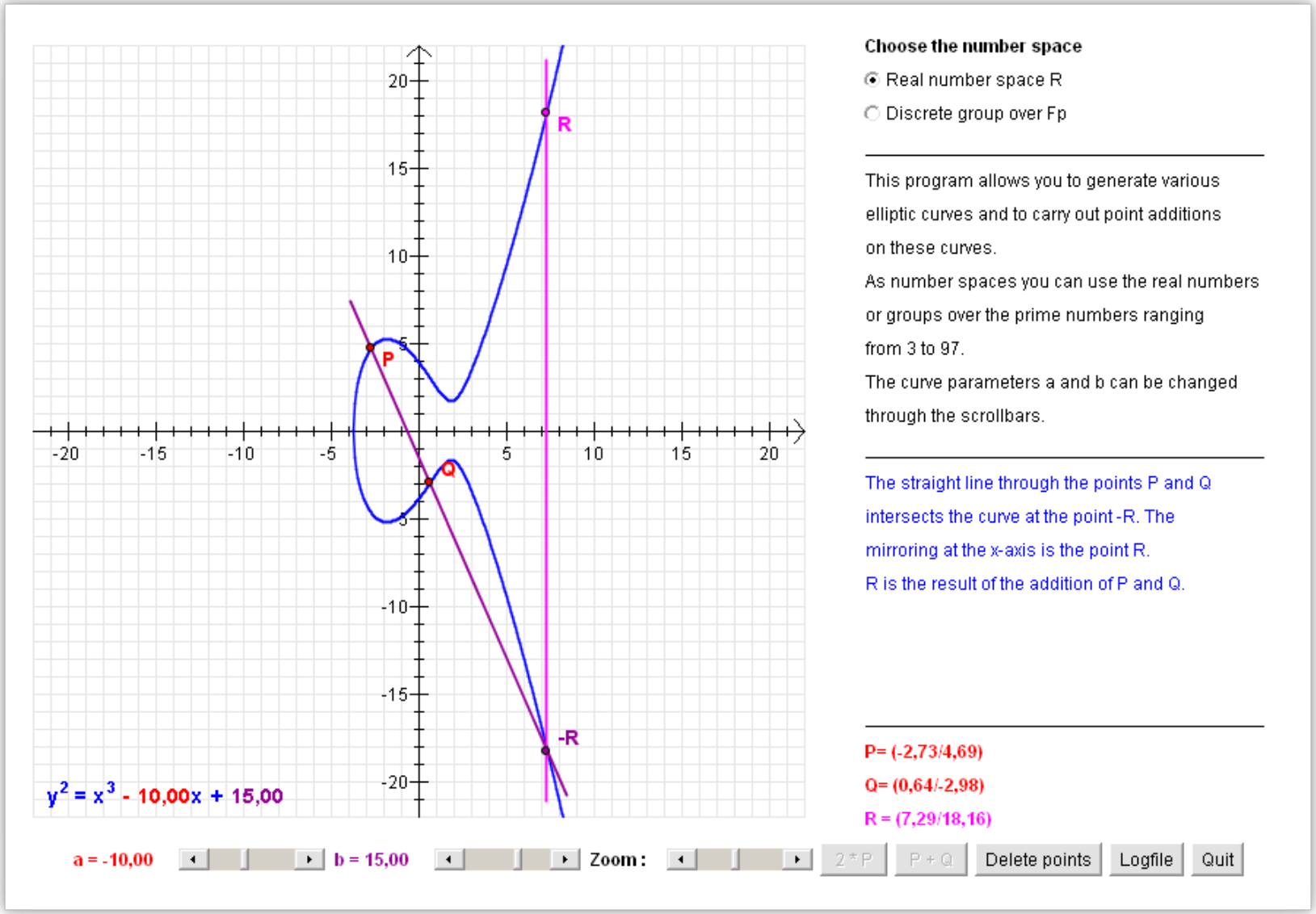


Figura 11: Adição de dois pontos sobre os números reais no CrypTool 1

Adição Algébrica

A adição de dois pontos de uma curva elíptica é dada por uma fórmula algébrica, isto é, envolve unicamente as operações básicas da adição, multiplicação e potenciação: Denote

$$\mathrm{P} + \mathrm{Q} = \mathrm{R} \quad \text{e} \quad (x_p, y_p) + (x_q, y_q) = (x_r, y_r).$$

Se a curve E é dada por $x^3 + ax + b$, então

$$x_r = \lambda^2 - x_p - x_q \quad \text{e} \quad y_r = \lambda(x_p - x_r) - y_p \quad (*)$$

onde

$$\lambda = \frac{y_q - y_p}{x_q - x_p} \quad \text{caso } x_q \neq x_p, \quad \text{e} \quad \lambda = \frac{3x_p^2 + a}{2y_p} \quad \text{caso } x_q = x_p.$$

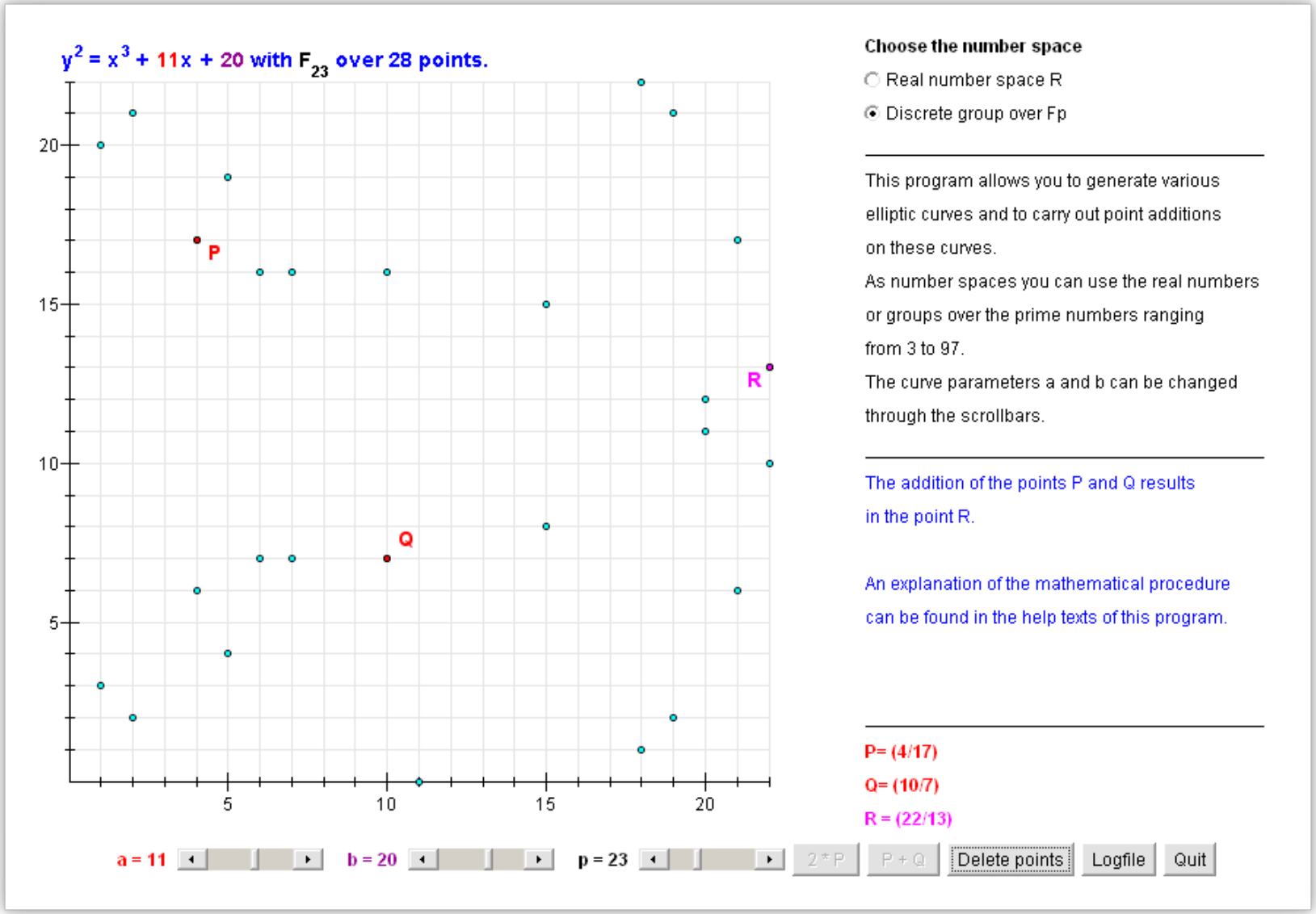


Figura 12: Adição de dois pontos sobre um corpo finito no CrypTool 1

Diffie-Hellman

Ao restringirmos às soluções (x, y) em $\mathbf{F}_p \times \mathbf{F}_p$ para um grande primo p e fixarmos um tal ponto P,

- enquanto, dado um número n, é fácil computar $\mathrm{Q} = n\mathrm{P} = \mathrm{P} + \ldots + \mathrm{P}$,
- em contraste, dado $\mathrm{Q} = \mathrm{P} + \ldots + \mathrm{P}$, é dificílimo computar quantas vezes P foi adicionado, isto é, computar o escalar n tal que $\mathrm{Q} = n\mathrm{P}$.

A criptografia por curvas elípticas `ECC` (Elliptic Curve Criptography) é uma variação do protocolo de `Diffie-Hellman`: Em vez de multiplicarmos repetidamente (n vezes) a base g, isto é, calcular $g^n = g \cdots g$, *adicionamos* repetidamente (n vezes) um ponto G, isto é, $n \cdot \mathrm{G} = \mathrm{G} + \cdots + \mathrm{G}$.

Vantagem

A vantagem de usar

- o logaritmo sobre uma curva elíptica sobre $\mathbf{F}_p$ (isto é, a função que para dados pontos G e Y determina o escalar x em $\mathbf{N}$ tal que $\mathrm{Y} = x\mathrm{G}$)

em vez do

- logaritmo sobre o grupo multiplicativo sobre $\mathbf{F}_p$ (isto é, a função que para dados números g e y determina o escalar x em $\mathbf{N}$ tal que $y = g^x$)

é que o tempo para computar do logaritmo aumenta em dependência do número dos dígitos binários de p,

- *linearmente* para o logaritmo sobre uma curva elíptica, enquanto
- *logaritmicamente* para o logaritmo multiplicativo.

Quanto maior o número de bits, maior este fator como mostra esta tabela:

Chave Simétrica	Chave Asimétrica Comum	Chave Elíptica
80	1024	160
112	2048	224
128	3072	256
192	7680	384
256	15360	512

Troca de Chaves

No protocolo ECDH (Elliptic Curve Diffie-Hellman), para Alice e Bob construírem uma chave secreta, combinam

- um número primo p *apropriado*,
- uma curva elíptica E *apropriada* sobre $\mathbf{F}_p$, e
- um ponto G *apropriado* em E.

1. Alice, pra gerar **uma metade** da chave, escolhe a inteiro,
   - calcula $\mathrm{A} \equiv a\mathrm{G}$, e
   - transmite A ao Bob.
2. Bob, pra gerar **outra metade** da chave, escolhe um b inteiro,
   - calcula $\mathrm{B} \equiv b\mathrm{G}$, e
   - transmite B à Alice.
3. A chave secreta **mútua** entre Alice e Bob é

$$c := b\mathrm{A} = ba\mathrm{G} = ab\mathrm{G} = ba\mathrm{G} = ab\mathrm{G} = a\mathrm{B}$$

Os algoritmos usando ECC

A criptografia por curvas elípticas ECC (Elliptic Curve Criptography) usa a troca de Chaves de Diffie-Hellman para

1. estabelecer uma chave secreta, para
2. transformá-la por um *hash* criptográfico, para
3. usá-la para cifrar a comunicação por um algoritmo criptográfico simétrica.

É padronizado pelo ECIES (Elliptic Curve Integrated Encryption Scheme), um procedimento *híbrido* (mistura criptografia assimétrica com criptografia simétrica).

Isto é, Uma vez a chave secreta mútua c estabelecida, Alice e Bob usam-na para cifras simétricas como AES or 3DES.

A cifra simétrica AES no dia-a-dia

Figura 13: Cifração de uma rede sem fio pelo AES

Referências

Diffie, Whitfield, e Martin Hellman. 1976. «New directions in cryptography». *IEEE transactions on Information Theory* 22 (6). IEEE: 644–54.

1. Aritmética Modular
2. Troca Multiplicativa
3. Troca Elíptica