Criptografía sobre curvas elípticas

José Galaviz Casas

Departamento de Matemáticas,
Facultad de Ciencias,
Universidad Nacional Autónoma de México.

Contenido

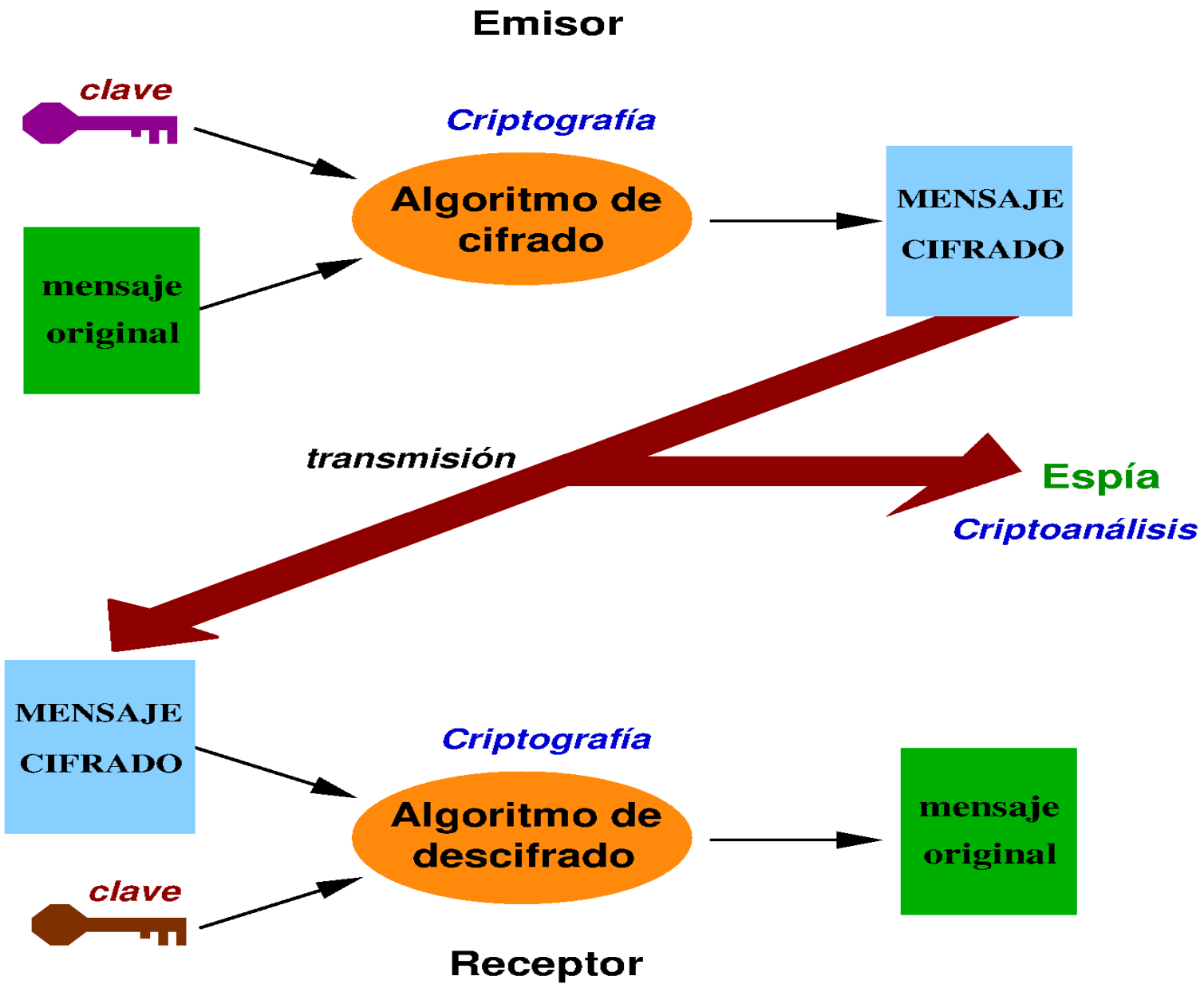
Sistema criptográfico (elementos 1):

- Un **emisor** del mensaje secreto.
- Un **receptor** a quien va dirigido el mensaje y que está autorizado a conocer los secretos contenidos en él.
- Una entidad que pretende, sin autorización, conocer los secretos: el **enemigo**.
- Un **canal de comunicación** inseguro por el que viaja el mensaje entre el emisor y el receptor y que suponemos intervenido por el enemigo.

Sistema criptográfico (elementos 2):

- El **mensaje claro** cuyo significado se desea hacer saber al receptor.
- El **mensaje cifrado** que viajará por el canal y cuyo contenido secreto es el mismo que el del mensaje claro.
- Un par de elementos, llamados **claves** o **llaves**, que permiten obtener el mensaje cifrado a partir del mensaje claro y viceversa.
- Un **algoritmo de cifrado** que recibe como entrada el texto claro y la clave de cifrado y obtiene el texto cifrado como salida.
- Un **algoritmo de descifrado** que recibe como entrada el texto cifrado y la clave de descifrado y obtiene el texto claro original.

Esquema general de un sistema criptográfico



Cifrado simétrico Vs. asimétrico

- Si la llave de cifrado y de descifrado son la misma el sistema es **simétrico**.
- En caso contrario se denomina **asimétrico**.

Cifrados simétricos vigentes

- AES (*Advanced Encryption Standard*, Rijndael).
- ChaCha.
- Camellia.
- ARIA.
- 3DES.

Sistemas criptográficos simétricos: ventajas

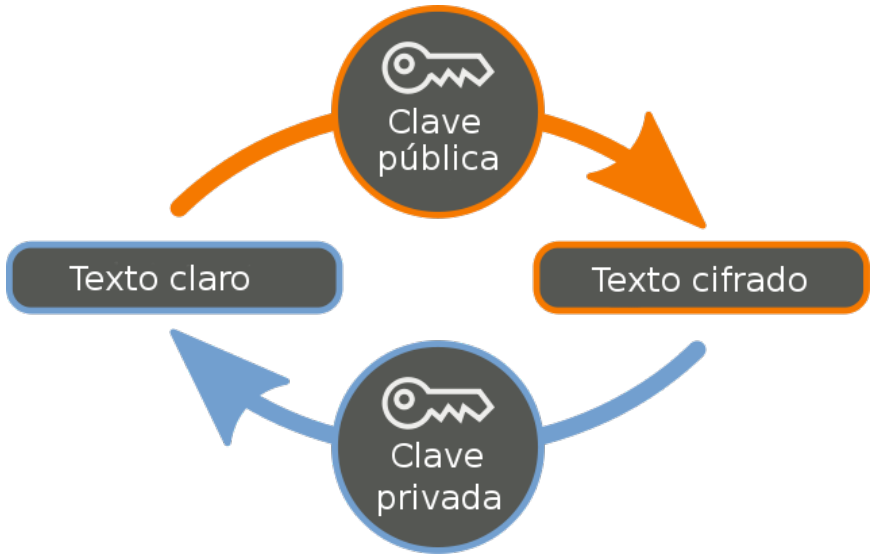
- Versatilidad. Se pueden implementar eficientemente con facilidad, tanto en hardware como en software: sólo se requiere de operaciones sencillas de manejo de bits, permutaciones, corrimientos, xor, etc.
- Alto rendimiento, chip AES con clave de 256 bits: 494 MB/s para cifrar y 2.6 GB/s para descifrar [Liv14]. En software dos ordenes de magnitud más lentos pero aún muy rápidos.
- Claves relativamente cortas respecto a la longitud del texto claro: cientos de bits.
- Se conocen muy bien (en términos generales) sus fortalezas y debilidades.

Sistemas criptográficos simétricos: desventajas

- La clave debe permanecer secreta y la poseen tanto el emisor como el receptor, así que debe haber doble garantía.
- Acordar la clave debe hacerse fuera del sistema.
- Hay que cambiar claves con periodicidad para dificultar el criptoanálisis por volumen de datos (lineal, p.ej.).

Imaginemos:

- Un sistema en el que cada usuario U posee dos claves:
  - Una clave de cifrado que le dice a todo mundo. Pública.
  - Una clave de descifrado que sólo él conoce. Privada.
- Cualquiera puede enviar mensajes cifrados a U.
- Sólo U puede descifrar los mensajes cifrados que le envían.

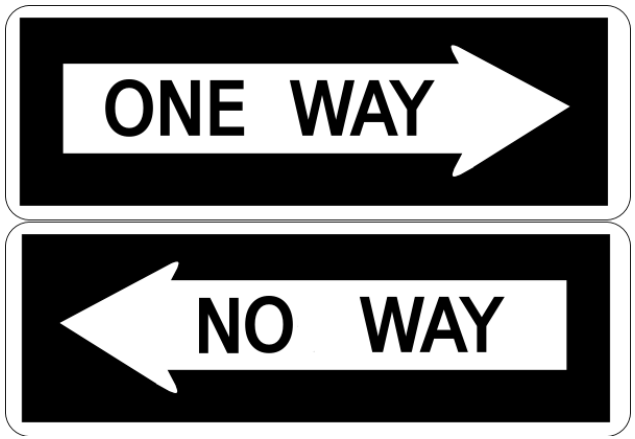


¿Qué se necesita?

- Necesitamos un sistema asimétrico.
- En el que conocer la clave de cifrado no sea muy útil para calcular la de descifrado.

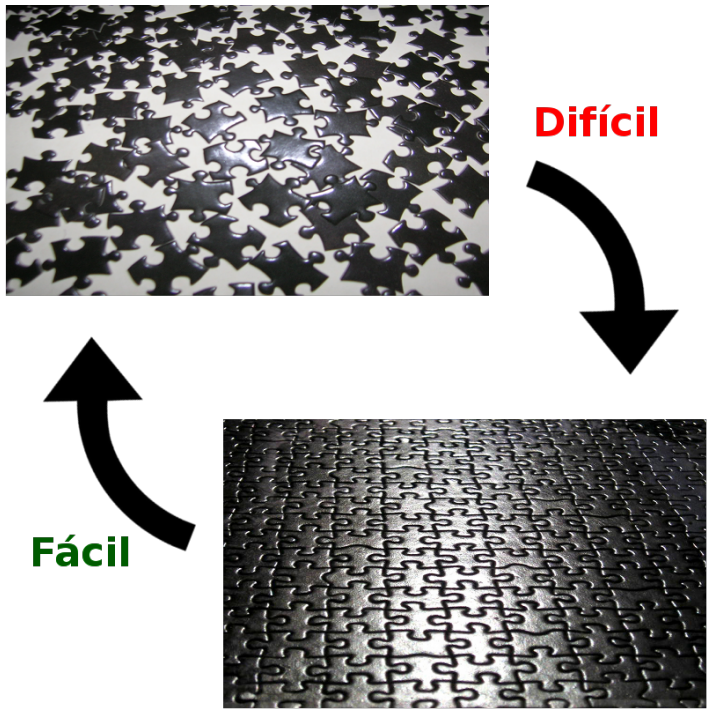
Dos puntos de vista

- Se podría, si podemos encontrar funciones en las que "ir" en un sentido es fácil e "ir" en sentido contrario no: fácil aplicar f, difícil aplicar f^{-1}.
- Es un concepto computacional, se refiere a la dificultad (complejidad) general de calcular el valor de la inversa.



Funciones de un sólo sentido:

- Ejemplo: rompecabezas.
- Es difícil armarlo (pregúntenle a mi esposa).
- Es muy fácil desarmarlo (pregúntenle a mi gato).



Pero no basta:

- Idea: escribo un mensaje atrás de un *Ravensburger*, armado, de $\mathbf{N} \times \mathbf{10^3}$ piezas que me regaló un cuate; a quien envío luego de regreso su rompecabezas desarmado, con la intención de que lea el mensaje.
- Un espía atrapa el rompecabezas en tránsito. Es difícil (inútil) que lo arme.
- Pero también para el destinatario.
- Necesitamos algo más...

Funciones de puerta de trampa:

- ... necesitamos un elemento extra que, si es conocido por alguien, haga que la inversa también sea fácil de calcular.
- Si mi cuate me regaló el rompecabezas con las piezas numeradas en un patrón peculiar y sólo él tiene ese patrón ¡listo!.
- La puerta de trampa sólo es fácil de abrir si se conoce el truco.
- La clave para descifrar es el patrón.

Infraestructura

- Los mensajes se codifican como números.
- Cifrar y descifrar consistirá en aplicar operaciones a números.
- Los que estarán representados en binario en la memoria de una computadora.
- Con un cierto tamaño establecido.

Ejemplo de funciones de un sólo sentido

Logaritmo discreto en un grupo cíclico o en un campo finito:

- Es fácil calcular $\mathbf{n} = \mathbf{r}^{\mathbf{k}}$, dados $\mathbf{r}$ y $\mathbf{k}$.
- Pero en general no es fácil calcular $\mathbf{k}$ tal que $\mathbf{n} = \mathbf{r}^{\mathbf{k}}$ dados $\mathbf{n}$ y $\mathbf{r}$.
- La solución puede no existir.

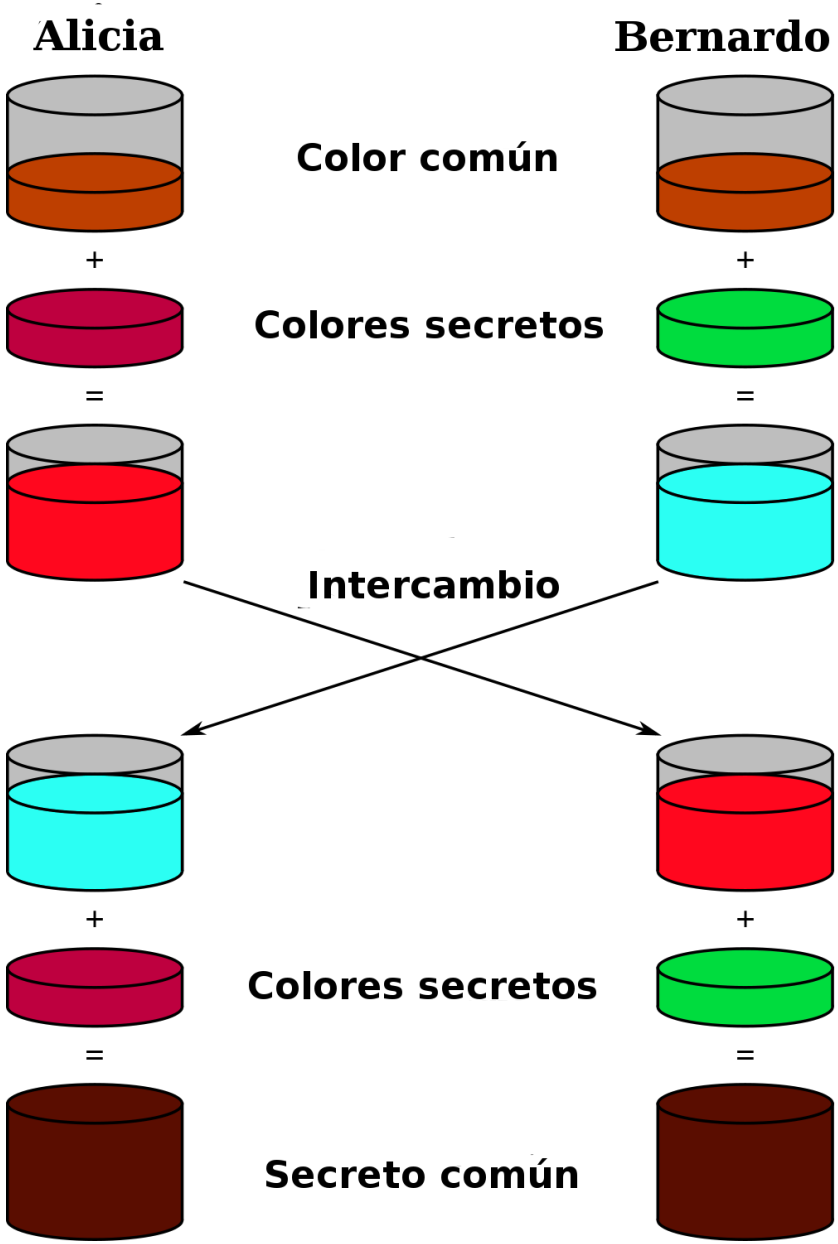
Los pioneros...

...oficiales, de este lado del Atlántico:
Whitfield Diffie, Martin Hellman, Ralph Merkle, Stanford, 1976.

Pinturas

¿Han visto un catálogo de pinturas?

Diffie-Hellman con botes de pintura



Protocolo de intercambio de llave de Diffie-Hellman: inicio

Alicia y Bernardo quieren ponerse de acuerdo en una clave secreta, sin que nadie se entere, usando el inseguro canal para comunicarse.

- Alicia y Bernardo se ponen de acuerdo en un par de números:
  - Un primo grande $\mathbf{p}$.
  - Un generador $\mathbf{g} \in \mathbb{Z}_{\mathbf{p}}^* = \mathbb{Z}_{\mathbf{p}} \setminus \{\mathbf{0}\}$.
- No importa si los escucha el enemigo.

Protocolo de intercambio de llave de Diffie-Hellman: ejecución

1. Alicia elige un entero aleatorio grande α, con $\mathbf{0} < \alpha < \mathbf{p} - \mathbf{1}$, y le envía a Bernardo: $\mathbf{X} = \mathbf{g}^{\alpha} \pmod{\mathbf{p}}$.
2. Bernardo elige un entero aleatorio grande β, con $\mathbf{0} < \beta < \mathbf{p} - \mathbf{1}$, y le envía a Alicia: $\mathbf{Y} = \mathbf{g}^{\beta} \pmod{\mathbf{p}}$.
3. Alicia calcula $\mathbf{K} = \mathbf{Y}^{\alpha} \pmod{\mathbf{p}}$.
4. Bernardo calcula $\mathbf{K}' = \mathbf{X}^{\beta} \pmod{\mathbf{p}}$.

$$\mathbf{K} = \mathbf{K}' = \mathbf{g}^{\alpha\beta} \pmod{\mathbf{p}}$$

El criptoanálisis

Un espía conoce **p**, **g**, **X**, y **Y**. Para calcular **K** tendría que hacer una de dos cosas:

1. Obtener el logaritmo discreto de **X** ó **Y** en base **g** módulo **p** para obtener α ó β, respectivamente, y poder calcular $\mathbf{g}^{\alpha\beta} \pmod{\mathbf{p}}$.
2. Calcular $\mathbf{g}^{\alpha\beta} \pmod{\mathbf{p}}$ de alguna manera diferente a la opción anterior.

Conjetura DH

La conjetura de Diffie-Hellman es que la segunda opción no es posible.
Así que la seguridad del protocolo estriba en que:

- Creemos que calcular el logaritmo discreto es difícil.
- Creemos que es el único medio para conocer **K**.

Ventajas

Con el protocolo de Diffie-Hellman se resuelve el problema de acordar la clave para usar un sistema simétrico rápido y eficiente. El número secreto común puede usarse como clave del sistema simétrico.

¡Así funciona el comercio en Internet!

Sistemas criptográficos de llave pública: ventajas

- Sólo la llave privada debe permanecer secreta y sólo la posee una persona.
- Se puede pensar fácilmente en una red de usuarios con directorio público de claves.
- Es fácil implementar un sistema de firmas digitales.
- En un sistema con muchos usuarios se requiere de relativamente pocas llaves.

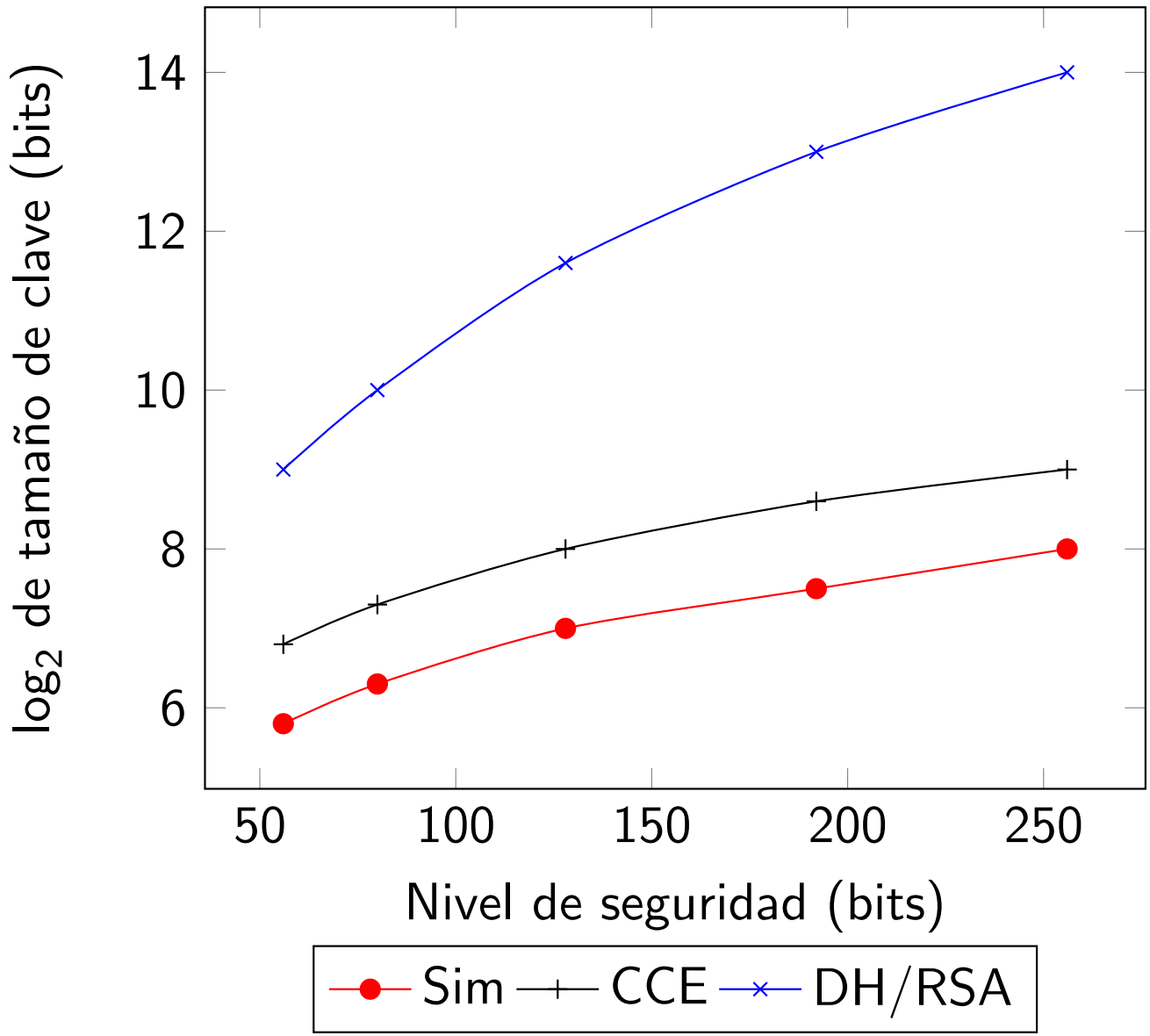
Sistemas criptográficos de llave pública: desventajas

- La implementación requiere de mucha infraestructura adicional. Bibliotecas de manejo de enterotes con aritmética modular, pruebas de primalidad.
- Mucho más lentos que los simétricos: 16 Gbits/seg [Fre15].
- Sabemos menos de ellos, la seguridad se basa en el supuesto de que hay problemas difíciles y de que no hay modo de hacer las cosas más que por la ruta difícil.

Sistemas criptográficos de llave pública: desventajas

- Hay más recursos teóricos para obtener atajos en el criptoanálisis (Vs. ataque de fuerza bruta en simétricos).
- El tamaño de las llaves es mucho mayor (un orden de magnitud) al requerido por los sistemas simétricos para obtener una seguridad comparable.

Simétricos	DH/RSA	CCE
80	1024	163
128	3072	283
192	7680	409
256	15360	571


14
12
10
8
6
$\log_2$ de tamaño de clave (bits)
50
100
150
200
250
Nivel de seguridad (bits)
Sim
CCE
DH/RSA

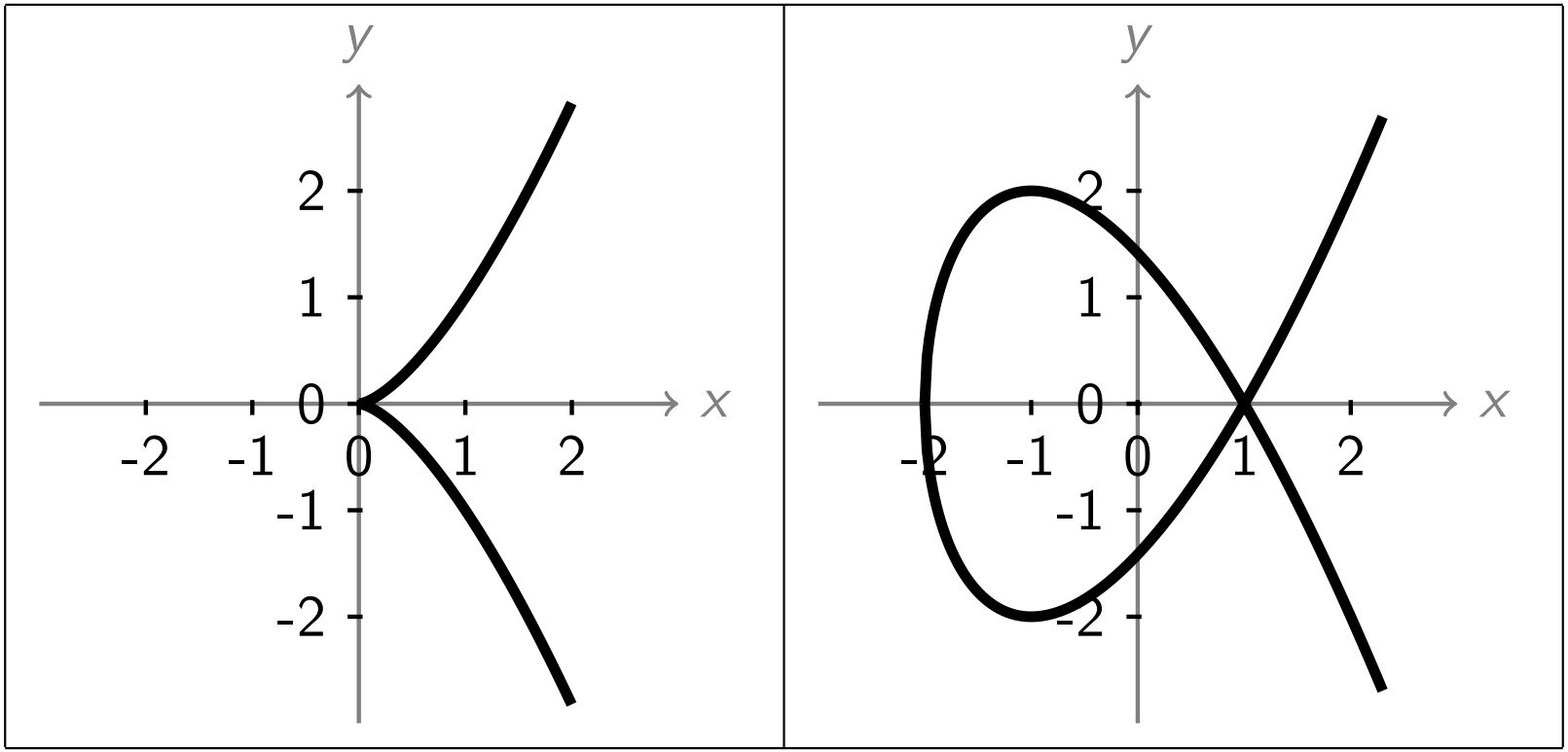
Curva elíptica

$$E: \ y^2 = x^3 + ax + b \tag{1}$$

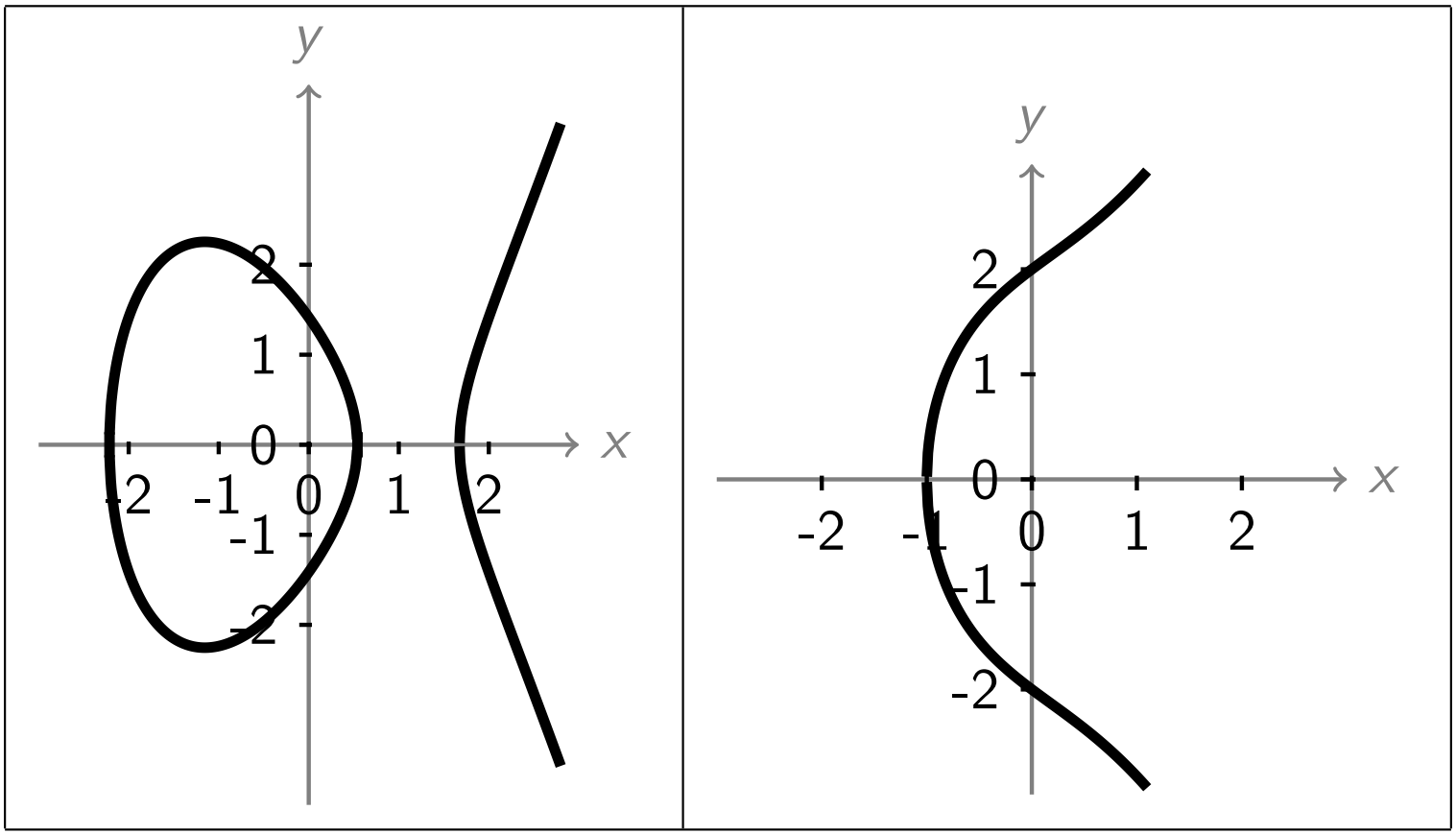
Sujeta a:

$$4a^3 + 27b^2 > 0 \tag{2}$$

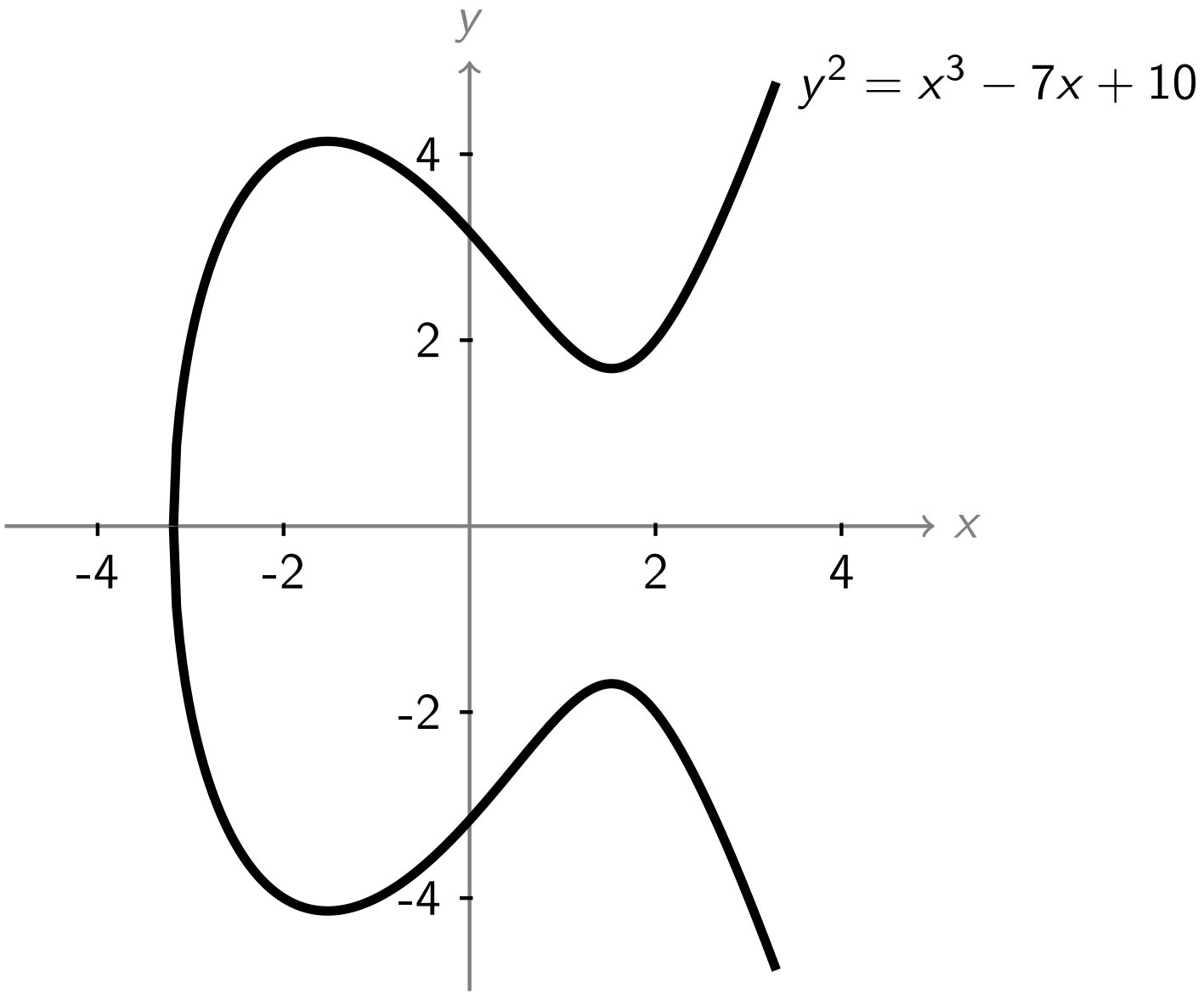
Así se ven



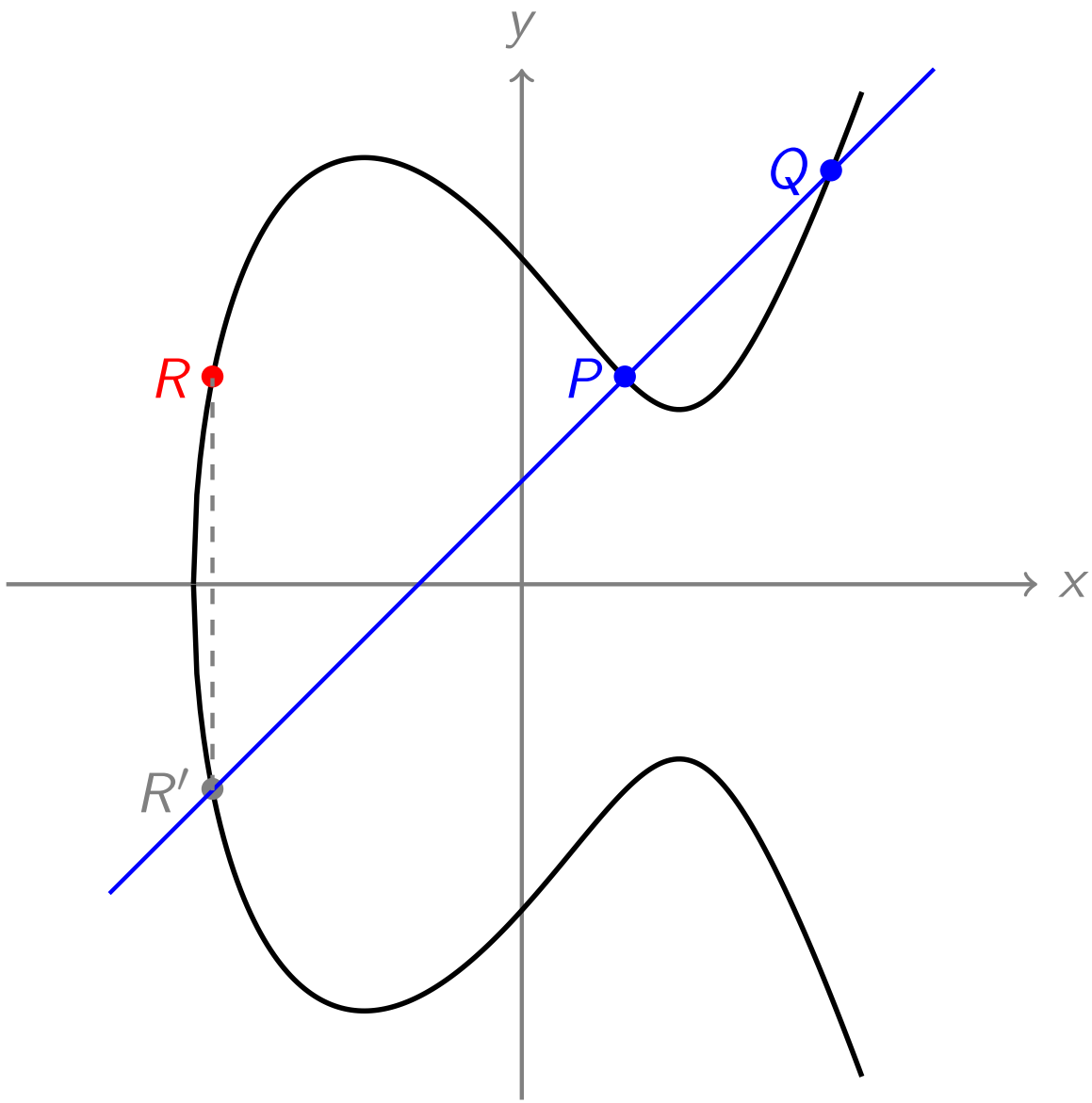
Así se ven



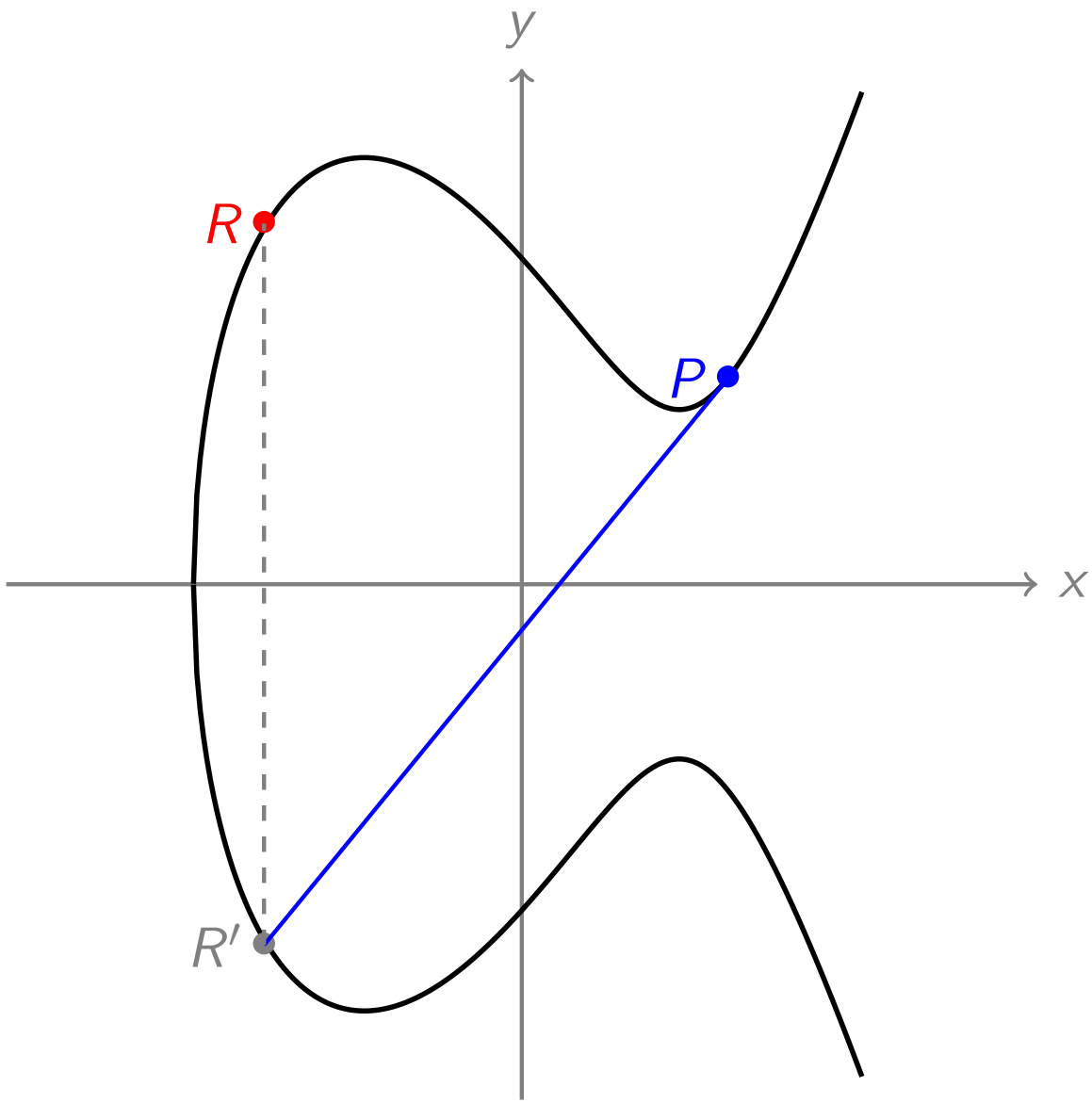
Ejemplo



Suma de puntos



Duplicación de un punto



Conjunto finito

Una curva elíptica sobre $\mathbb{Z}_p$ ($p > 3$ primo) es el conjunto de parejas $(x, y) \in \mathbb{Z}_p^2$ que satisfacen:

$$E: \ y^2 \equiv x^3 + ax + b \pmod{p} \tag{3}$$

con $a, b \in \mathbb{Z}_p$, sujetas a la restricción:

$$4a^3 + 27b^2 \not\equiv 0 \pmod{p} \tag{4}$$

junto con un punto imaginario llamado *punto al infinito* denotado con $\mathcal{O}$.

La operación

Sean $P = (x_1, y_1)$, $Q = (x_2, y_2)$ dos puntos sobre una curva elíptica de acuerdo a la definición y $R = (x_3, y_3)$ el punto sobre la curva que corresponde a la suma de P y Q, entonces:

- Si $P \neq Q$ pero $x_1 = x_2$, entonces $R = P + Q = \mathcal{O}$.
- Si $P = Q$ y $y_1 = y_2 = 0$, entonces $R = P + Q = 2P = \mathcal{O}$.
- En otro caso:

$$x_3 \equiv s^2 - x_1 - x_2 \pmod{p} \tag{5}$$

$$y_3 \equiv s(x_1 - x_3) - y_1 \pmod{p} \tag{6}$$

Donde

$$s \equiv \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & (\text{mód } p) \quad \text{si } P \neq Q \\ \\ \frac{3x_1^2 + a}{2y_1} & (\text{mód } p) \quad \text{si } P = Q \end{cases} \tag{7}$$

El punto que corresponde con el inverso aditivo de P, denotado como $-P$ es:

$$-P \equiv (x_1, -y_1) \pmod{p} = (x_1, p - y_1)$$

Propiedades

Dados una curva elíptica E sobre $\mathbb{Z}_p$ ($p > 3$ primo) y los puntos cualesquiera $P, P_1, P_2, P_3 \in E$:

- Cerradura. $P_1 + P_2 \in E$.
- Asociatividad. $P_1 + (P_2 + P_3) = (P_1 + P_2) + P_3$.
- Existencia del neutro aditivo. Existe $\mathcal{O}$ (el punto al infinito) tal que $P + \mathcal{O} = \mathcal{O} + P = P$.
- Existencia del inverso aditivo. Para toda $P \in E$ existe un elemento distinguido denotado $-P$ tal que $P + (-P) = \mathcal{O}$.
- Conmutatividad. $P_1 + P_2 = P_2 + P_1$.

Grupo cíclico

El conjunto de puntos sobre una curva elíptica módulo p, es un grupo cíclico.
Un elemento $G \in E$ es llamado un *generador* de E si, para cualquier otro elemento $P \in E$ existe una k tal que $P = k\,G$.

Ejemplo

Sea

$$E : y^2 \equiv x^3 - 7x + 10 \pmod{19} \tag{8}$$

recordemos que todas las operaciones deben hacerse módulo 19.
El punto $P_1 = (9, 7)$ genera los siguientes elementos de E:

$$\{P_1 = (9,7), 2P_1 = (18,15), 3P_1 = (1,17), 4P_1 = (7,0),$$
$$5P_1 = (1,2), 6P_1 = (18,4), 7P_1 = (9,12), 8P_1 = \mathcal{O}\}$$

El grupo

x	y	Orden	x	y	Orden
1	2	8	1	17	8
2	2	24	2	17	24
3	4	24	3	15	24
5	9	12	5	10	12
7	0	2	9	7	8
9	12	8	10	3	24
10	16	24	12	1	3
12	18	3	13	8	12
13	11	12	16	2	6
16	17	6	17	4	24
17	15	24	18	4	4
18	15	4	$\mathcal{O}$		0

El problema del logaritmo discreto

Dada una curva elíptica módulo un número primo p, sean G un elemento generador y P otro punto cualquiera, con $G, P \in < E, p, + >$, el *Problema del Logaritmo Discreto* consiste en encontrar un número $k \in \mathbb{N}$ tal que:

$$kG \equiv P \pmod{p}$$

Diffie-Hellman

1. Elegir un primo p grande y los valores de a y b en la expresión

$$E : y^2 \equiv x^3 + a \cdot x + b \pmod{p}$$

2. Elegir un elemento G, generador de los puntos de E.
3. A elige $k_A <| E |$ (el orden de $< E, p, + >$).
4. A calcula $Q_A = k_A \cdot G$. Nótese que este no es un escalar, es un punto de E.
5. B elige $k_B <| E |$.
6. B calcula $Q_B = k_B \cdot G$.
7. A envía a B el punto Q_A.
8. B envía a A el punto Q_B.
9. A calcula $R_A = k_A \cdot Q_B$
10. B calcula $R_B = k_B \cdot Q_A$

Funciona

Por supuesto, tanto A como B llegan al mismo punto, dado que:

$$\begin{aligned} R_B &= k_B \cdot Q_A = k_B \cdot (k_A \cdot G) \\ &= (k_B \cdot k_A) \cdot G = (k_A \cdot k_B) \cdot G \\ &= k_A \cdot (k_B \cdot G) = k_A \cdot Q_B \\ &= R_A \end{aligned}$$

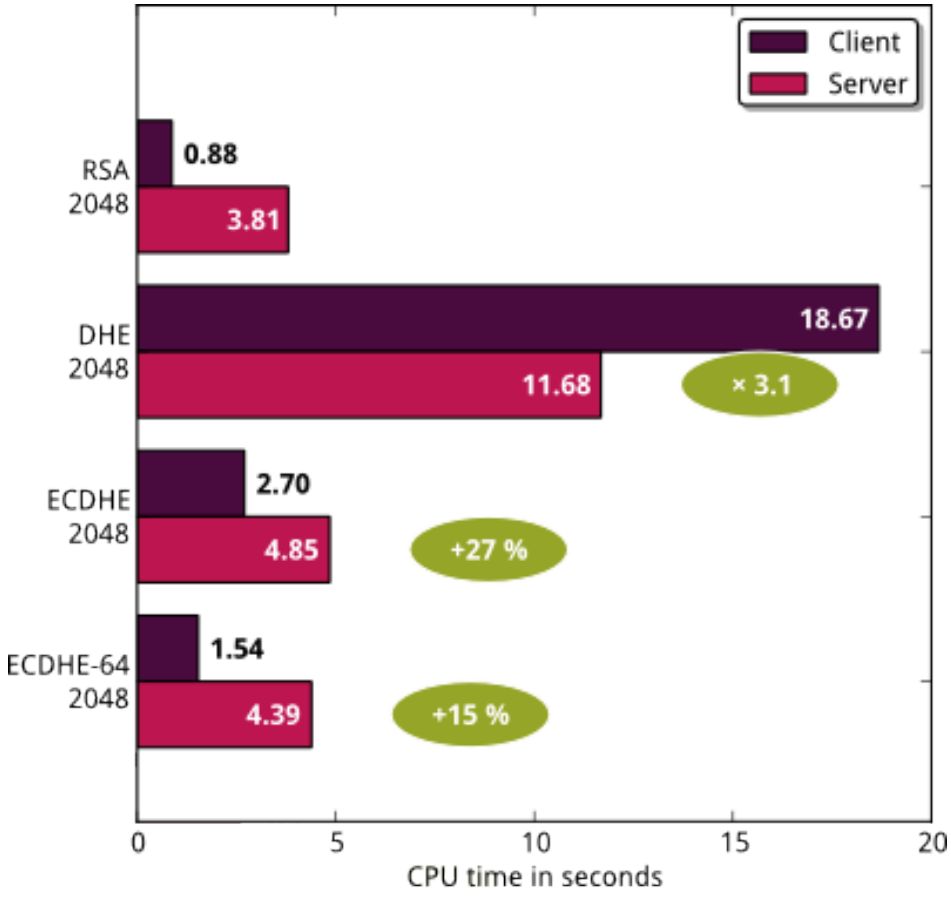
¿Qué tan bien?

Hoy en día (TLS, V1.2 + actualizaciones hasta 2011) todo servidor debe tener implementadas 25 curvas elípticas diferentes: 11 sobre un campo primo y 14 sobre un campo finito de caraterística 2.

En TLS V1.3 permanecerán 3 de ellas (sobre campos primos). Las demás o no se usaban o, si eran mal usadas, podían ser inseguras.

Todos los sistemas operativos y actualmente todos los navegadores de internet.

Desempeño (1000 ejecuciones)



¿Qué sigue?

- Llegaron para quedarse.
- El furor inicial se ha ido atemperando.
- Buena relación costo-beneficio.
- NSA: “Desafortunadamente, el crecimiento del uso de curvas elípticas ha mermado el progreso en la investigación sobre la computación cuántica, lo que exige una reevaluación de nuestra estrategia de cifrado”.

Referencias

[Liv14] Liviero, Belinda, *Intel® AES-NI Performance Enhancements: HyTrust DataControl Case Study*, septiembre 2014.
https://software.intel.com/en-us/articles/intel-aes-ni-performance-enhancements-hytrust-datacontrol-case-study

[Fre15] Freescale Semiconductor, *Freescale C29x Crypto Coprocessor Family Product Brief*, julio 2015.
http://www.nxp.com/files/32bit/doc/prod_brief/C29xPB.pdf

[Ver11] Vernat, Vincent, *FSSL/TLS & Perfect Forward Secrecy*, noviembre 2011.
https://vincent.bernat.im/en/blog/2011-ssl-perfect-forward-s