



# Clase 23



# Factorización de Enteros

- Como vimos cuando estudiamos el RSA, la seguridad en algunos criptosistemas de clave pública se basa en la dificultad de factorizar enteros que son producto de dos primos grandes. En efecto, no se conoce algoritmo eficiente (es decir, polinomial en el número de bits de  $n$ ) para factorizar un  $n$  de la forma  $n = p \cdot q$  ( $p$  y  $q$  primos), en general.
- Ahora bien, en casos particulares, aunque  $p$  y  $q$  sean muy grandes, hay algoritmos eficientes que permiten factorizar  $n$ , y los veremos a continuación.




# Método de Factorización de Fermat


- Sea  $n = p \cdot q$ ,  $p$  y  $q$  primos,  $p > q$ .
- Veamos que si  $p - q$  es pequeño (del orden de una potencia de  $\log n$ ) se tiene una manera eficiente de factorizar  $n$ .
- Proponemos un cambio de variables que permita escribir la factorización de  $n$  como una diferencia de cuadrados, es decir, en lugar de hallar directamente  $p$  y  $q$ , proponemos encontrar antes  $a$  y  $b$  tales que:
  - $$\begin{cases} p = a + b \\ q = a - b \end{cases}$$
- De donde la igualdad:  $n = p \cdot q$  se transforma en:
- $n = (a + b) \cdot (a - b) = a^2 - b^2$




Pierre de Fermat

- De aquí se sigue que factorizar  $n$  es equivalente a escribirlo como diferencia de cuadrados cuyas bases  $a$  y  $b$  no son consecutivas, de modo que  $a - b \neq 1$  y obtenemos la factorización no trivial  $n = (a + b) \cdot (a - b)$ .
- Observación: las incógnitas  $a$  y  $b$  son enteras, como se ve de resolver el sistema  $2 \times 2$  que las define, obteniendo:
  - $a = (p + q)/2$ ,  $b = (p - q)/2$ , pues  $p$  y  $q$  son ambos impares.
- Si suponemos que  $p - q$  es pequeño, por ejemplo:  $p - q < \log n$ , podemos hallar  $a$  y  $b$  en una cantidad pequeña de pasos. En efecto:  $b = (p - q)/2 < (\log n)/2$  y  $a = (p + q)/2$  cumple:  $a^2 - n = b^2 < \left(\frac{\log n}{2}\right)^2 \Rightarrow a^2 < n + \left(\frac{\log n}{2}\right)^2$
- Además:  $a^2 - n = b^2 > 0 \Rightarrow a^2 > n \Rightarrow a > \sqrt{n}$ . Por lo tanto:
- $\sqrt{n} < a < \sqrt{n + \left(\frac{\log n}{2}\right)^2}$ .

- 
- Para hallar  $a$  se prueba uno por uno los enteros de este intervalo. Veamos que la cantidad de enteros en este intervalo es como mucho  $(\log n)/2$ . Si llamo  $U$  a esta cantidad tenemos que  $U < \sqrt{n + \left(\frac{\log n}{2}\right)^2} - \sqrt{n}$ .
  - Sea  $V = \sqrt{n + \left(\frac{\log n}{2}\right)^2} + \sqrt{n}$ . Vemos que  $U < V$  y que:
  - $U \cdot V = n + \left(\frac{\log n}{2}\right)^2 - n = \left(\frac{\log n}{2}\right)^2$ . Por lo tanto  $U < \frac{\log n}{2}$ .

- 
- Por lo tanto, probamos con los enteros de este intervalo hasta hallar un  $a$  tal que  $a^2 - n = x$  sea un cuadrado perfecto.
  - Calculamos, partiendo de una aproximación a  $\sqrt{n}$ , el entero  $[\sqrt{n}]$ , es decir, el mayor entero  $z$  tal que  $z \leq \sqrt{n}$  y a partir de este valor los valores de  $a$ :
  - $a = [\sqrt{n}] + 1, [\sqrt{n}] + 2, \dots$ , y para cada uno de ellos calculamos  $a^2 - n = x$  y comprobamos si el valor obtenido es un cuadrado perfecto. Una manera fácil de comprobar esto último es calcular aproximadamente  $\sqrt{x}$ , coger el entero  $b$  más cercano a este valor y verificar si se cumple o no  $b^2 = x$ .
  - Una vez hallado el  $a$  tal que  $a^2 - n = x = b^2$  para  $b$  entero ya tenemos:
  - $n = a^2 - b^2 = (a + b) \cdot (a - b) = p \cdot q$ , es decir, tenemos la factorización de  $n$ .

- 
- ¿Y si esta factorización fuera trivial, es decir, si fuera  $a - b = 1$ ?
  - En ese caso  $a = b + 1 \Rightarrow n = a + b = 2 \cdot b + 1 \Rightarrow b = (n-1)/2$  y  $a = (n+1)/2$ , pero como este valor de  $a$  no pertenece al intervalo  $[\sqrt{n}, \sqrt{n + \left(\frac{\log n}{2}\right)^2}]$  vemos que nunca nos topáramos con este caso!! Esto tiene sentido porque para esta factorización trivial los dos factores son 1 y  $n$ , que NO ESTÁN CERCA uno del otro.
  - El método que hemos visto es rápido bajo la hipótesis de que  $p - q$  sea pequeño, pues en este caso el intervalo en el que buscar  $a$  tiene longitud pequeña. Si se parte de suponer que  $p - q < W \cdot (\log n)^C$  para  $W$  y  $C$  constantes positivas, también se deduce que el algoritmo acabará rápido, es decir, la cantidad de pasos será de un orden polinomial en  $\log n$ .







# Método de Factorización p-1 de Pollard

- El siguiente método permite dar una factorización eficiente (rápida) de un entero  $n = p \cdot q$ , donde  $p$  y  $q$  son primos grandes, suponiendo que el primo  $p$  (o  $q$ , o ambos) es tal que  $p-1$  tiene "una propiedad especial"
- **Definición:** Dada una cota  $B > 0$ , decimos que un entero  $x$  es B-suave si todos los factores primos de  $x$  son menores que  $B$ . Decimos que un entero  $x$  es B-suave en potencias si, al descomponer  $x$  en producto de primos:
- $x = p_1^{a_1} \cdot p_2^{a_2} \cdot \dots \cdot p_r^{a_r}$ , se cumple que  $p_i^{a_i} \leq B$ ,  $\forall i = 1, 2, \dots, r$ .
- En la aplicación a la factorización de un entero  $n$ , fijaremos una cota  $B$  pequeña, del orden de  $\log n$ , y para este valor de  $B$  nos interesará que cierto valor sea B-suave en potencias.



Professor John Pollard


- 
- Dada una tal cota  $B$ , comenzamos por precalcular el **mínimo común múltiplo de todos los enteros menores o iguales que  $B$** :
  - (1) Criba de Eratóstenes: Calcular el conjunto  $P$  de todos los primos  $p \leq B$ .
  - (2) Calcular el producto:  $m = \prod_{p \in P} p^{\lfloor \log_p B \rfloor}$
  - De este modo hemos formado  $m$  con todos los primos hasta  $B$ , cada uno de ellos con el mayor exponente  $w$  tal que  $p^w \leq B$ , puesto que  $p^{\log_p B} = B$ . Esto prueba que  $m$  es el mínimo común múltiplo buscado.
  - Este valor lo calculamos para una cota  $B$  del orden de  $\log n$  (la idea es coger  $B$  lo más grande posible, siempre que el cálculo previo pueda hacerse en un tiempo razonable, y se pueda almacenar el resultado, es decir,  $m$ , en el ordenador).


- 
- Una vez calculado este valor, el método de Pollard intenta factorizar  $n = p \cdot q$  y funciona bajo la condición de que al menos uno de los dos factores primos de  $n$ , por ejemplo  $p$ , tenga la propiedad de que  $p - 1$  sea B-suave en potencias.




# Clase 24

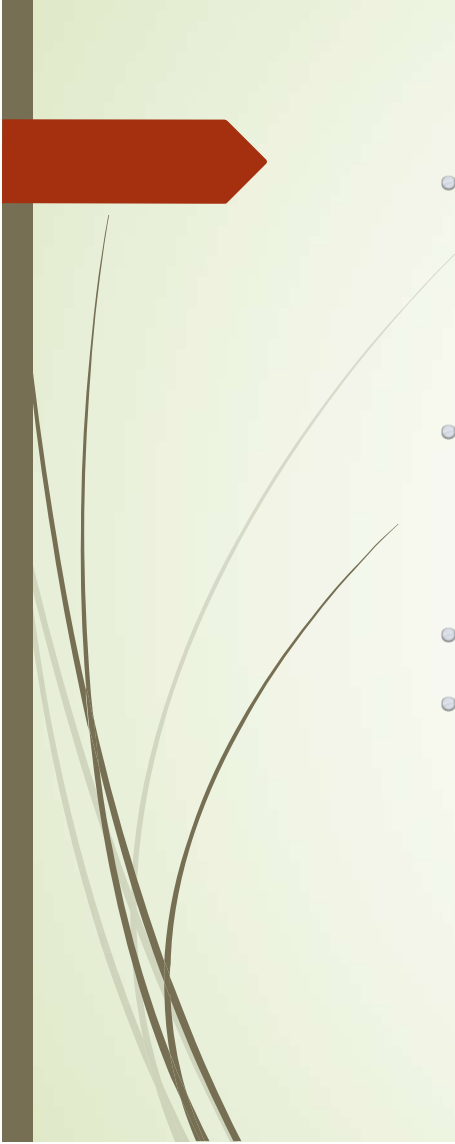
- **Método de Pollard:** Tras fijar la cota B, supongamos que  $n = p \cdot q$  con p y q primos y tales que p-1 es B-suave en potencias. El objetivo es hallar p y así factorizar n calculando  $q = n/p$ .
- Escogemos un entero  $1 < a < n$  al azar. Supondremos que p no divide a a, pues esto ocurre con una probabilidad muy alta, y además si p dividiera a a bastaría con calcular  $\text{mcd}(a, n)$  con el algoritmo de Euclides para hallar p.
- Por el Pequeño Teorema de Fermat, sabemos que:  $a^{p-1} \equiv 1 \pmod{p}$
- Calculamos (ya hemos explicado como se hace esto):
- $m = \text{mcm}(2, 3, \dots, B)$ . Como p-1 es B-suave en potencias, si:
- $p - 1 = p_1^{a_1} \cdot p_2^{a_2} \cdot \dots \cdot p_r^{a_r} \Rightarrow p_i^{a_i} \leq B$ , para todo  $i = 1, 2, \dots, r \Rightarrow$   
 $p_i^{a_i} \mid m$ , para todo  $i = 1, 2, \dots, r \Rightarrow$
- $p - 1 \mid m$


- 
- Hemos obtenido una herramienta muy poderosa: a pesar de no conocer  $p$ , y por lo tanto no conocer el valor de  $p-1$ , tenemos un entero  $m$  que es múltiplo de  $p-1$ , es decir, sabemos que existe  $k$  entero con:  $m = (p-1) \cdot k$ .
  - Por lo tanto, como  $a^{p-1} \equiv 1 \pmod{p} \Rightarrow a^m \equiv (a^{p-1})^k \equiv 1 \pmod{p} \Rightarrow$
  - $p$  divide a  $a^m - 1$ . Por otro lado sabemos que  $p$  divide a  $n = p \cdot q$ . Si calculamos con el algoritmo de Euclides el máximo común divisor  $d = \text{mcd}(a^m - 1, n)$  hay sólo dos posibilidades:
  - $$d = \begin{cases} p \\ n = p \cdot q \end{cases}$$
  - En el primer caso,  $d = p$  y por lo tanto ya hemos hallado el factor primo  $p$  de  $n$ , con lo cual calculando  $q = n/p$  tenemos la factorización de  $n$ .


- 
- En el segundo caso como  $d = n = p \cdot q$  esto no nos da la factorización de  $n$ . Cuando ocurre esto es porque  $a^m - 1$  también es divisible por  $q$ , y variando el valor de la base  $a$  y repitiendo el algoritmo para varias bases esto dejará de ocurrir, excepto si el exponente  $m$  es también múltiplo de  $q-1$ , cosa que sólo puede ocurrir si también  $q-1$  es  $B$ -suave en potencias.
  - Es decir: suponiendo que  $p-1$  es  $B$ -suave en potencias, si caemos en el segundo caso, en el cual no conseguimos factorizar  $n$ , esto se debe a que  $q-1$  también es  $B$ -suave en potencias. Cuando ocurre esto, se puede hacer una variación del algoritmo que acabará funcionando: la idea es cambiar la cota  $B$  por cotas menores, hasta llegar a una nueva cota  $B'$  tal que sólo uno de los factores primos de  $n$ , por ejemplo  $p$ , tenga la propiedad de que  $p-1$  es  $B'$ -suave en potencias, y el otro no. Por lo tanto, calculando y utilizando el correspondiente  $m$  para esta cota  $B'$  el algoritmo funcionará.



- 
- A partir del B inicial, y suponiendo que este B no sirve para factorizar n porque se cae en el caso  $d = \text{mcd}(a^m - 1, n) = n$ , los siguientes valores de la cota se calculan según el método dicotómico:
  - Probamos con  $B/2$ : Si para esta cota  $\text{mcd}(a^m - 1, n)$  es igual a un factor primo de n, el algoritmo acaba. Si  $\text{mcd}(a^m - 1, n)$  es igual a 1, es que la cota es muy pequeña y hay que probar un valor mayor: en el siguiente paso se prueba con  $\frac{3}{4} \cdot B$ . Si en cambio  $\text{mcd}(a^m - 1, n) = n$  es porque esta nueva cota sigue siendo muy alta, y como siguiente valor cogemos  $B/4$ .
  - Iterando este proceso llegaremos al valor de  $B' < B$  que permite factorizar n. Además, el método dicotómico garantiza que como máximo probaremos un total de  $\log_2 B$  valores para la cota.

- 
- Por lo tanto, el algoritmo acaba funcionando, y de modo eficiente, siempre que se cumpla la hipótesis inicial: tiene que existir un valor de la cota  $B$  que no sea demasiado grande, y por lo tanto permita calcular el  $m$  como se indicó, tal que para al menos uno de los factores primos de  $n$ , por ejemplo  $p$ , se cumpla que  $p - 1$  es  $B$ -suave en potencias.
  - Respecto al cálculo de  $d = \text{mcd}(a^m - 1, n)$ , se hace con el algoritmo de Euclides, pero hay un problema: el número  $a^m - 1$  es demasiado grande, y sería muy lento, o incluso imposible, calcularlo. Para resolver este problema hacemos uso del siguiente resultado elemental:
  - Si  $x \equiv y \pmod{n} \Rightarrow \text{mcd}(x, n) = \text{mcd}(y, n)$ .
  - Por lo tanto, para calcular  $d$  no hace falta conocer el valor de  $a^m - 1$ , basta con conocer su residuo módulo  $n$ : para hallarlo, calculamos la potencia  $a^m$  módulo  $n$  utilizando la exponenciación modular binaria (es decir, que utiliza la descomposición de  $m$  en base 2) y así calculamos de manera eficiente la clase módulo  $n$  de  $a^m$  y de  $a^m - 1$ , sin tener que hacer ningún cálculo que involucre a números mayores que  $n^2$ .

- 
- Observación: En el Mathematica la función  $\text{PowerMod}(a, m, n)$  se ocupa de calcular la clase módulo  $n$  de  $a^m$ .
  - De este modo, una vez calculado el residuo módulo  $n$  de  $a^m - 1$ , calculamos con el algoritmo de Euclides  $d = \text{mcd}(a^m - 1, n)$  y, como explicamos antes, bajo las apropiadas hipótesis y probando varias bases si fuera preciso, acabaremos obteniendo el factor primo  $p$  y la factorización de  $n$ .
  - Ejemplo:  $n = 5917$ . Probamos con la cota  $B = 5$ .
  - Calculamos  $m = \text{mcm}(2, 3, 4, 5) = 60$ . Cogemos  $a = 2$ .
  - Aplicando  $\text{PowerMod}(2, 60, 5917)$  calculamos:  $2^{60} \equiv 3417 \pmod{5917} \Rightarrow$
  - $2^{60} - 1 \equiv 3416 \pmod{5917}$ .
  - Calculamos  $\text{mcd}(2^{60} - 1, 5917) = \text{mcd}(3416, 5917) = 61 \Rightarrow 61 \mid 5917$ .
  - Calculamos:  $5917/61 = 97$ , y concluimos que  $n = 5917 = 61 \cdot 97$ .

- 
- Observación: Con la cota  $B = 5$  hemos podido factorizar  $n$  “a la primera” (es decir, la única cota que hubo que considerar fue esta). Esto se debe a que el factor 61 de  $n$  cumple:
  - $61 - 1 = 60 = 2^2 \cdot 3 \cdot 5$  y  $2^2$ , 3 y 5 son todos menores o iguales que 5, con lo cual  $61 - 1$  es 5-suave en potencias.
  - Además para el otro factor primo de  $n$  se tiene:
  - $97 - 1 = 96 = 2^5 \cdot 3$  y como  $2^5 = 32 > 5 \Rightarrow 97 - 1$  no es 5-suave en potencias.
  - En general, si para una cota  $B$  (y para el  $m$  correspondiente) se obtiene  $\text{mcd}(a^m - 1, n) = 1 \Rightarrow$  se tiene que volver a intentar con valores de  $B$  más grandes hasta conseguir (si hay suerte!!) que este mcd deje de ser 1.
  - También, como ya vimos, si para una cota  $B$  obtenemos  $\text{mcd}(a^m - 1, n) = n$ , se tiene que seguir probando con valores más pequeños de la cota hasta conseguir factorizar  $n$ .