Factorization of RSA Numbers

Logan Blinco

University of York

lb1642@york.ac.uk

June 9, 2022

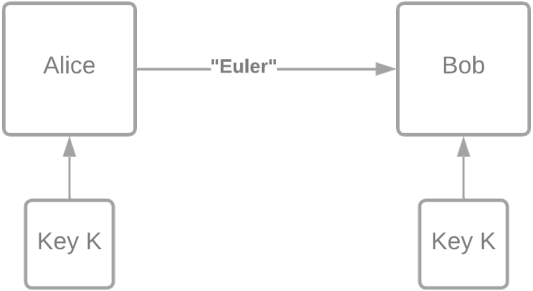


Figure: Motivation for RSA

Character	Binary	Hex
E	1000101	45
u	1110101	75
l	1101100	6C
e	1100101	65
r	1110010	72

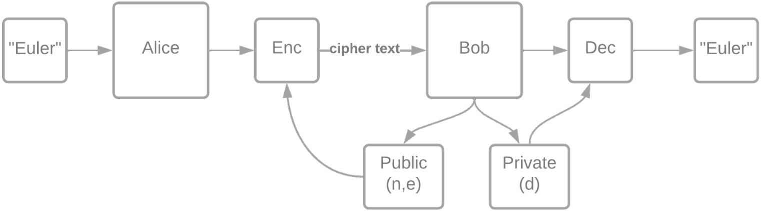
Figure: ASCII table for ”Euler”

RSA Scheme

- p, q big primes
- $n = pq$
- Calculate $\phi(n) = (p-1)(q-1)$.
- Select an encipher exponent. $1 < e < \phi(n)$ and $\gcd(e, \phi(n)) = 1$.
- Calculate a decipher exponnet d

$$d \cdot e = 1 \pmod{\phi(n)}$$

$e = 65537$ is frequently used.


"Euler"
Alice
Enc
cipher text
Bob
Dec
"Euler"
Public
(n,e)
Private
(d)

RSA Encryption

$$c = m^e \pmod{n}$$

Theorem (Euler's theorem)

Suppose $\gcd(m, n) = 1$ *then:*

$$m^{\phi(n)} \equiv 1 \pmod{n}.$$

RSA Decryption

$$\begin{aligned}
c^d \pmod{n} &= (m^e)^d \pmod{n} \\
&\equiv m^{d \cdot e} \pmod{n} \\
&\equiv m^{k\phi(n)+1} \pmod{n} \\
&\equiv \left(m^{\phi(n)}\right)^k \cdot m \pmod{n} \\
&\equiv 1 \cdot m \pmod{n} \\
&\equiv m \pmod{n}.
\end{aligned}$$

Since

$$d \cdot e \equiv 1 \pmod{\phi(n)}$$

If we can factor the RSA modulus, then we can calculate d. Thus read the message.

Fermats Factorisation Algorithm (17th Century)

if $N = a^2 - b^2$, then we have factored N.

1. $a_0 = \lceil \sqrt{n} \rceil$
2. $b_i = a_i^2 - N$. Perfect square?
3. If not, $a_{i+1} = a_i + 1$. Goto 2.

Lets Factor $n = 2993$.

$$\begin{aligned} a_0 &= \left\lceil \sqrt{2993} \right\rceil = 55 \\ b_0 &= 55^2 - 2993 = 32 \text{✗} \end{aligned}$$

Lets Factor $n = 2993$.

$$a_1 = 55 + 1 = 56$$

$$b_1 = 56^2 - 2993 = 143 \text{✗}$$

Lets Factor $n = 2993$.

$$\begin{aligned} a_2 &= 56 + 1 = 57 \\ b_2 &= 57^2 - 2993 = 256 = 16^2 \end{aligned}$$

Our factor is then $57 - 16 = 41$ so $2993 = 41 \cdot 73$.

- Can take up to $n - \sqrt{n}$ iterations
- $b \approx \log_2 n \implies 2^b - 2^{\frac{b}{2}}$ iterations
- very slow

Quadratic Sieve

- Lets loosen the condition
- Fermat requires $a^2 - b^2 = 1 \cdot n$
- Loosen the condition to

$$a^2 - b^2 = k \cdot n \implies a^2 \equiv b^2 \pmod{n}$$

Theorem (General Factoring Congruence)

If we have two integers x, y *such that*

$$x^2 \equiv y^2 \pmod{n}$$

then we have at least a 50% *chance that* $\gcd(x-y, n)$ *or* $\gcd(x+y, n)$ *is a non-trivial factor of* n*.*

- Finding pairs $x^2 \equiv y^2 \pmod{n}$ is hard
- Much easier if we construct them

Definition (Factor Base)

$$FB = \{p \,|\, p \text{ is prime, and } p \leq B\} \cup \{-1\}$$

Example (Factor Base of $B = 13$)

$$FB = \{-1, 2, 3, 5, 7, 11, 13\}$$

B-smoothness

A number is B-smooth if we can write it as a product of factors from the factor base.

$$x = \prod_{p_i \in FB} p_i^{\alpha_i}$$

Quadratic Sieve Equivalence relation

$$\begin{aligned} Q(r) &= r^2 - n \\ r^2 &\equiv Q(r) \pmod{n} \end{aligned}$$

Start sieving at $r_0 = \lfloor \sqrt{n} \rfloor$.

Forming Squares

If $Q(r)$ is B-Smooth then:

$$Q(r) = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$$

Denote

$$v(r) = (\alpha_1, \alpha_2 \dots, \alpha_k)$$

$$b_i = \begin{cases} 0, & \text{if } \alpha_i \text{ is even} \\ 1, & \text{otherwise} \end{cases}$$

$$w(r) = (b_1, b_2, \ldots b_k)$$

$$r_1^2 \cdot r_2^2 \cdot \dots \cdot r_d^2 \equiv Q(r_1) \cdot Q(r_2) \cdot \ldots \cdot Q(r_d) \pmod{n} = (p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdot \dots \cdot p_d^{\lambda_d})^2$$

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \pmod{2}$$

If we have m B-smooth numbers and a factor base of size b then we have a $m \times b$ matrix.
x, y, z, w determines which ones are used in the multiplication (binary).

Finishing the process

$$(r_1 \cdot r_2 \cdot \ldots r_n)^2 = (p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdot \cdots \cdot p_d^{\lambda_d})^2 \pmod n$$

$$\gcd\left(n, (r_1 \cdot r_2 \cdot \ldots r_n) - (p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdot \cdots \cdot p_d^{\lambda_d})\right)$$
$$\gcd\left(n, (r_1 \cdot r_2 \cdot \ldots r_n) + (p_1^{\lambda_1} \cdot p_2^{\lambda_2} \cdot \cdots \cdot p_d^{\lambda_d})\right)$$

At least a 50% chance of non-trivial factor.

Sieving

- Sieving over $Q(\lfloor\sqrt{n}\rfloor + i)$ varying i.
- Must be stored in memory which is physical and finite
- split sieve interval into regions of size D

$$i \in [0, D), [D, 2D), [2D, 3D) \ldots [M - D, M]$$

- Use parallelism!

Propagating factor hits

Theorem

If $p \in FB$ is a factor of $Q(r)$ then it is also a factor for $Q(r+kp)$

Proof.

$$\begin{aligned} Q(r) &= r^2 - n \\ Q(r+kp) &= (r+kp)^2 - n = r^2 + 2rkp + (kp)^2 - n \\ Q(r+kp) &= Q(r) + 2rkp + (kp)^2 \equiv Q(r) \pmod{p} \\ Q(r) &\equiv 0 \pmod{p} \implies Q(r+kp) \equiv 0 \pmod{p} \end{aligned}$$

□

Reducing the factor base

Suppose $p \mid Q(r)$:

$$\frac{Q(r)}{p} = s \implies \frac{r^2 - n}{p} = s$$
$$\implies r^2 - n = ps \implies n \equiv r^2 \pmod{p}$$

So n is a quadratic residue modulo p. This removes about half the factor base.

Sieve Example

Lets perform a sieve on $n = 551$ with a factor base of $[-1, 2, 5, 11]$.

$$Q(\lfloor\sqrt{551}\rfloor + i) = Q(23 + i)$$

Using $i \in [0, 9]$ we get

$$[-22, 25, 74, 125, 178, 233, 290, 349, 410]$$

Sieve -1 through gives

$$[22, 25, 74, 125, 178, 233, 290, 349, 410]$$

Sieve 2 and propagate hits.

$$[11, 25, 37, 125, 89, 233, 145, 349, 205]$$

Sieve 5 and propagate hits.

$$[11, 1, 37, 1, 89, 233, 29, 349, 41]$$

etc.

Our final sieve array is

$$[1, 1, 37, 1, 89, 233, 29, 349, 41]$$

- So only $Q(23+0), Q(23+1), Q(23+3)$ are B-Smooth
- They will be added to the matrix

Breaking RSA

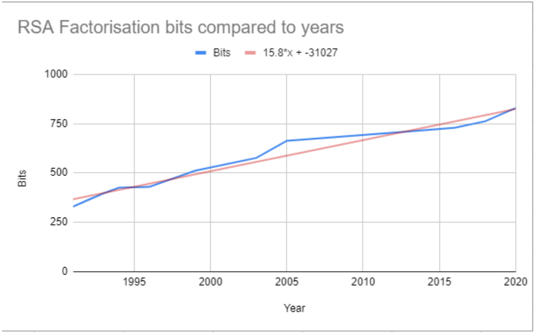
How big can we factor?

- RSA-100 (330 bits) factored in 1991 using MPQS (Lenstra)
- RSA-140 (463 bits) factored in 2020 using QS (Konsor) via using 6000 core hours (6 days total time)
- RSA-250 (829 bits) factored in 2020 using GNFS. Largest RSA number factored
- 22 bits factored by a Quantum Computer (Dash)

How big do we want to factor?

- NIST standards recommend at least 1024 bit modulus
- Modern systems are starting to use 2048 or 3072 bit modulus

Will RSA be broken soon?



- Expect 1024 bit to be factored soon breaking security for many systems
- NIST guidelines are out of date
- Far off 2048 or 3072 bit factorization

Factorization of RSA Numbers

Logan Blinco

University of York

lb1642@york.ac.uk

June 9, 2022

Appendix: Complexity

Definition (L-Notation)

For a bound variable n, we define $L_n[\alpha, c]$ for a positive c and $\alpha \in [0, 1]$ as:

$$L_n[\alpha, c] = e^{(c+o(1))\cdot(\ln n)^{\alpha}\cdot(\ln \ln n)^{1-\alpha}}$$

- $\alpha = 0 \implies \log n$ growth
- $\alpha = 1 \implies$ exponential growth
- $\alpha \in (0, 1) \implies$ sub-exponential growth. Better than exponential, worse than polynomial.

Appendix: Complexity

Quadratic Sieve

$$L_n[\frac{1}{2},1] = e^{(1+o(1))\cdot(\ln n)^{\frac{1}{2}}\cdot(\ln\ln n)^{\frac{1}{2}}}$$

Number Field Sieve

$$L_n[\frac{1}{3},(\frac{64}{9})^{\frac{1}{3}}] = e^{(\frac{64}{9})^{\frac{1}{3}}+o(1))\cdot(\ln n)^{\frac{1}{3}}\cdot(\ln\ln n)^{\frac{2}{3}}}$$

Appendix: Quadratic Sieve Example

Lets factor 2623 using the scheme. Lets choose our bound to be 13 so our factor base is

$$[2, 3, 5, 7, 11, 13]$$

We then generate a series of random numbers r and attempt to factorise them into the factor base (all positive so ignoring the sign).

r	$f(r)$	factor	vector form	$\pmod 2$
89	52	$2^2 \cdot 13$	$(2, 0, 0, 0, 0, 1)$	$(0, 0, 0, 0, 0, 1)$
93	780	$2^2 \cdot 3 \cdot 5 \cdot 13$	$(2, 1, 1, 0, 0, 1)$	$(0, 1, 1, 0, 0, 1)$
97	1540	$2^2 \cdot 5 \cdot 7 \cdot 11$	$(2, 0, 1, 1, 1, 0)$	$(0, 0, 1, 1, 1, 0)$
90	231	$3 \cdot 7 \cdot 11$	$(0, 1, 0, 1, 1, 0)$	$(0, 1, 0, 1, 1, 0)$
35	1225	$5^2 \cdot 7^2$	$(0, 0, 2, 2, 0, 0)$	$(0, 0, 0, 0, 0, 0)$
49	2401	7^4	$(0, 0, 0, 4, 0, 0)$	$(0, 0, 0, 0, 0, 0)$
42	1764	$2^2 \cdot 3^2 \cdot 7^2$	$(2, 2, 0, 2, 0, 0)$	$(0, 0, 0, 0, 0, 0)$

Appendix: Quadratic Sieve Example

$$f(89)\cdot f(93)\cdot f(97)\cdot f(90) = (89\cdot 93\cdot 97\cdot 90)^2 \equiv \\ 2^2\cdot 13\cdot 2^2\cdot 3\cdot 5\cdot 13\cdot 2^2\cdot 5\cdot 7\cdot 11\cdot 3\cdot 7\cdot 11 \pmod{2623}$$

$$(89\cdot 93\cdot 97\cdot 90)^2 \equiv (2^6\cdot 3^2\cdot 5^2\cdot 7^2\cdot 11^2\cdot 13^2) \pmod{2623}$$

$$(89\cdot 93\cdot 97\cdot 90)^2 \equiv (2^3\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13)^2 \pmod{2623}$$

So we now check the $\gcd$:

$$\gcd(2623, 89\cdot 93\cdot 97\cdot 90 - 2^3\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13) = 43$$

So we have successfully factored our value. $2623 = 43\cdot 61$.

Appendix: Quadratic Residue Removal

n	B	original Size	new Size
977	15	6	3
2993	50	15	9
79369	250	53	26
79369	1000	168	77
192421	250	53	30
192421	1000	168	87
563343097	2000	303	152

Figure: Comparing Factor base after non-quadratic residue removal

Appendix: Smoothness probability

$\Psi(x, B)$ denote number of y-smooth numbers $\leq x$

$$\Psi(x, B) \ \frac{1}{\phi(B)!} \prod_{p \leq B} \frac{\log x}{\log p}$$