General Number Field Sieve

Mohammadreza Motabar

École Polytechnique Fédérale de Lausanne (EPFL)

Shool of Computer and Communication Science (IC)

September 2023

Supervisor: Dr. Tako Boris Fouotsa

Outlines

Introduction

RSA is a very popular public key cryptosystem. This algorithm is known to be secure, but this fact relies on the difficulty of factoring large numbers.

GNFS is the **fastest** known method for factoring large given integer N, where large is generally mean over **110** digits.

RSA is a very popular public key cryptosystem. This algorithm is known to be secure, but this fact relies on the difficulty of factoring large numbers.

GNFS is the **fastest** known method for factoring large given integer N, where large is generally mean over **110** digits.

What Is The Idea?

The "difference of squares" method relies upon that if integers x and y are such that $x \not\equiv \pm y \pmod{N}$ and $x^2 \equiv y^2 \pmod{N}$.

Then $gcd(x-y, N)$ and $gcd(x+y, N)$ are non-trivial factors of N.

What Is The Idea?

The "difference of squares" method relies upon that if integers x and y are such that $x \not\equiv \pm y \pmod{N}$ and $x^2 \equiv y^2 \pmod{N}$.

Then $gcd(x-y, N)$ and $gcd(x+y, N)$ are non-trivial factors of N.

Historical Background

1. Dixon's Algorithm
2. Quadratic Sieve (QS)
3. Special Number Field Sieve (SNFS)
4. On April 10, 1996, GNFS was used to factorize RSA130.

Quadratic Sieve

Smooth Numbers

- Factor Base (F) $\rightarrow$ A set of prime numbers.
- k is smooth $\rightarrow$ All primes dividing k are in F.
- k is B-smooth $\rightarrow$ All primes dividing k are less than B.

Smooth Numbers

- Factor Base (F) $\rightarrow$ A set of prime numbers.
- k is smooth $\rightarrow$ All primes dividing k are in F.
- k is B-smooth $\rightarrow$ All primes dividing k are less than B.

Smooth Numbers

- Factor Base (F) $\rightarrow$ A set of prime numbers.
- k is smooth $\rightarrow$ All primes dividing k are in F.
- k is B-smooth $\rightarrow$ All primes dividing k are less than B.

Smoothness Bound & Smooth Integers

- Choose a smoothness bound B.
- Select a range.
- Sieve to locate $\pi(B)+1$ numbers a_i such that $b_i = (a_i^2 \pmod{N})$ is B-smooth.

Let $N = 227179$ and $B = 25$. ($\sqrt{N} \approx 476$)

x	$x^2 \pmod{N}$	Factorization	Smooth?
470	-6279	$-3 \times 7 \times 13 \times 23$	Yes
473	-3450	$-2 \times 3 \times 5^2 \times 23$	Yes
476	-603	$-3^2 \times 67$	No
477	350	$2 \times 5^2 \times 7$	Yes
482	5145	$3 \times 5 \times 7^3$	Yes
493	15870	$2 \times 3 \times 5 \times 23^2$	Yes

$$\begin{aligned}(477 \times 482 \times 493)^2 &= 477^2 \times 482^2 \times 493^2 \\ &\equiv (2 \times 5^2 \times 7)(3 \times 5 \times 7^3)(2 \times 3 \times 5 \times 23^2) \\ &\equiv 2^2 \times 3^2 \times 5^4 \times 7^4 \times 23^2 \\ &\equiv (2 \times 3 \times 5^2 \times 7^2 \times 23)^2 \pmod{227179}\end{aligned}$$

And also:

$$\begin{aligned}477 \times 482 \times 493 &\equiv 212460 \pmod{227179} \\ 2 \times 3 \times 5^2 \times 7^2 \times 23 &\equiv 169050 \pmod{227179}\end{aligned}$$

Finally:

$$\left.\begin{aligned}gcd(227179, 212460 + 169050) &= 157 \\ gcd(227179, 212460 - 169050) &= 1447\end{aligned}\right\} 227179 = 157 \times 1447$$

Finding Square

For each number construct the vector $(x_{-1}, x_2, x_3, x_5, x_7, x_{11}, x_{13}, x_{17}, x_{19}, x_{23})$ where x_p is exponent of p parity (and $x_{-1} = 1$ if $x^2 - N < 0$ and is 0 if $x^2 - N > 0$).

x	$x^2 \pmod{N}$	Factorization	Vector
470	-6279	$-3 \times 7 \times 13 \times 23$	(1, 0, 1, 0, 1, 0, 1, 0, 0, 1)
473	-3450	$-2 \times 3 \times 5^2 \times 23$	(1, 1, 1, 0, 0, 0, 0, 0, 0, 1)
477	350	$2 \times 5^2 \times 7$	(0, 1, 0, 0, 1, 0, 0, 0, 0, 0)
482	5145	$3 \times 5 \times 7^3$	(0, 0, 1, 1, 1, 0, 0, 0, 0, 0)
493	15870	$2 \times 3 \times 5 \times 23^2$	(0, 1, 1, 1, 0, 0, 0, 0, 0, 0)

Try to solve $M^T x = 0$ and then split N.

The GNFS Algorithm

Generalizing The QS

Suppose a ring R and a ring homomorphism $\phi : R \to \mathbb{Z}/N\mathbb{Z}$ exist. If $\beta \in R$ with $\phi(\beta^2) = y^2 \pmod{N}$ and $x := \phi(\beta) \pmod{N}$ then:

$$x^2 \equiv \phi(\beta)^2 \equiv \phi(\beta^2) \equiv y^2 \pmod{N}$$

Fields And Roots of Irreducible Polynomials

Suppose a monic, irreducible polynomial $f(x)$ of degree d with rational coefficients and a root $\theta \in \mathbb{C}$ of $f(x)$, is known.

Then for the associated ring $\mathbb{Q}(\theta)$, the following hold:

- $\mathbb{Q}(\theta) \cong \mathbb{Q}[x]/\langle f(x) \rangle$ and it is a field.
- The set $\{1, \theta, \theta^2, \cdots, \theta^{d-1}\}$ forms a basis for $\mathbb{Q}(\theta)$ as a vector space over $\mathbb{Q}$.

Rings of Algebraic Integers

- $\alpha \in \mathbb{C}$ is called an algebraic integer if it is the root of a monic polynomial with **integer** coefficients.
- The set of all algebraic integers in $\mathbb{Q}(\theta)$, denoted $\mathfrak{O}$, forms a subring of the field $\mathbb{Q}(\theta)$.
- The set of all $\mathbb{Z}$-linear combinations of the elements $\{1, \theta, \theta^2, \cdots, \theta^{d-1}\}$, denoted $\mathbb{Z}[\theta]$.

Producing a Difference of Squares

If $m \in \mathbb{Z}/N\mathbb{Z}$ for which $f(m) \equiv 0 \pmod{N}$, the mapping $\phi : \mathbb{Z}[\theta] \to \mathbb{Z}/N\mathbb{Z}$ with $\phi(1) = 1$ and $\phi(\theta) = m$ is a surjective ring homomorphism.

$$\prod_{(a,b)\in U} (a + b\theta) = \beta^2 \quad , \quad \prod_{(a,b)\in U} (a + bm) = y^2$$

with $\beta \in \mathbb{Z}[\theta]$, $y \in \mathbb{Z}$ and $\phi(\beta) = x \in \mathbb{Z}/N\mathbb{Z}$ then we have:

$$x^2 \equiv \phi(\beta)^2 \equiv \phi(\beta^2) \equiv \phi(\prod_{(a,b)\in U} (a + b\theta))$$

$$\equiv \prod_{(a,b)\in U} \phi(a + b\theta) \equiv \prod_{(a,b)\in U} (a + bm) \equiv y^2 \pmod{N}$$

Smoothness And The Algebraic Factor Base

- It seams we can use irreducible elements of the ring $\mathbb{Z}[\theta]$ in the factor base.
- But $\mathbb{Z}[\theta]$ may not be a UFD.
- We can go around this problem by considering ideals of $\mathbb{Z}[\theta]$ of a special form

The high-level idea then is to choose a set I of prime ideals of $\mathfrak{O}$, which will use as **algebraic factor base**.

Smoothness And The Algebraic Factor Base

- It seams we can use irreducible elements of the ring $\mathbb{Z}[\theta]$ in the factor base.
- But $\mathbb{Z}[\theta]$ may not be a UFD.
- We can go around this problem by considering ideals of $\mathbb{Z}[\theta]$ of a special form

The high-level idea then is to choose a set I of prime ideals of $\mathfrak{O}$, which will use as **algebraic factor base**.

Smoothness And The Algebraic Factor Base

- It seams we can use irreducible elements of the ring $\mathbb{Z}[\theta]$ in the factor base.
- But $\mathbb{Z}[\theta]$ may not be a UFD.
- We can go around this problem by considering ideals of $\mathbb{Z}[\theta]$ of a special form

The high-level idea then is to choose a set I of prime ideals of $\mathfrak{O}$, which will use as **algebraic factor base**.

Proposition:

There are exactly d ring monomorphisms from the field $\mathbb{Q}(\theta)$ into the field $\mathbb{C}$. These embeddings are given by $\sigma_i(\mathbb{Q}) = \mathbb{Q}$ and $\sigma_i(\theta) = \theta_i$ for $1 \leq i \leq d$, assuming $f(x)$ split over $\mathbb{C}$ as:

$$f(x) = (x - \theta_1)(x - \theta_2) \cdots (x - \theta_d)$$

Definition:

Given an element $\alpha \in \mathbb{Q}(\theta)$, the norm of the element α, denoted by $N(\alpha)$, is defined as

$$N(\alpha) = \sigma_1(\alpha)\sigma_2(\alpha) \cdots \sigma_d(\alpha)$$

$$\begin{aligned} N(a + b\theta) &= \sigma_1(a + b\theta)\sigma_2(a + b\theta) \cdots \sigma_d(a + b\theta) \\ &= (a + b\theta_1)(a + b\theta_2) \cdots (a + b\theta_d) \\ &= (-b)^d(ab^{-1} - \theta_1)(ab^{-1} - \theta_2) \cdots (ab^{-1} - \theta_d) = (-b)^d f(ab^{-1}) \end{aligned}$$

Definition:

Given a ring R, and an ideal $\mathfrak{J}$ of R, the $N(\mathfrak{J})$ is defined to be $[R : \mathfrak{J}]$, the number of cosets of $\mathfrak{J}$ in R.

Definition:

A first-degree prime ideal $\mathfrak{p}$ of a Dedekind domain $\mathfrak{O}$ is a prime ideal of $\mathfrak{O}$ such that $N(\mathfrak{p}) = p$ for some prime integer p.

Proposition:

The set of pairs (r, p) where p is a prime integer and $r \in \mathbb{Z}/p\mathbb{Z}$ with $f(r) \equiv 0 \pmod{p}$ is in bijective correspondence with the set of all first-degree prime ideals of $\mathbb{Z}[\theta]$.

- It can be proved that the only prime ideals of $\mathbb{Z}[\theta]$ occurring in the ideal factorization of a principal ideal of the form $\langle a + b\theta \rangle$ for **coprime** integers a and b are the **first-degree prime ideals** of $\mathbb{Z}[\theta]$.
- And even more important first-degree prime ideal (r, p) occuring in the ideal factorization of $\langle a + b\theta \rangle$ if and only if $a \equiv -br \pmod{p}$.
- The algebraic factor base consists of pairs of all (r,p) corresponding to a first-degree prime ideal with p less than some integer B'.
- And we say $\langle a + b\theta \rangle$ is smooth if the norm of the ideal is smooth.

- It can be proved that the only prime ideals of $\mathbb{Z}[\theta]$ occurring in the ideal factorization of a principal ideal of the form $\langle a + b\theta \rangle$ for **coprime** integers a and b are the **first-degree prime ideals** of $\mathbb{Z}[\theta]$.

- And even more important first-degree prime ideal (r, p) occuring in the ideal factorization of $\langle a + b\theta \rangle$ if and only if $a \equiv -br \pmod{p}$.

- The algebraic factor base consists of pairs of all (r,p) corresponding to a first-degree prime ideal with p less than some integer B'.

- And we say $\langle a + b\theta \rangle$ is smooth if the norm of the ideal is smooth.

- It can be proved that the only prime ideals of $\mathbb{Z}[\theta]$ occurring in the ideal factorization of a principal ideal of the form $\langle a + b\theta \rangle$ for **coprime** integers a and b are the **first-degree prime ideals** of $\mathbb{Z}[\theta]$.
- And even more important first-degree prime ideal (r, p) occuring in the ideal factorization of $\langle a + b\theta \rangle$ if and only if $a \equiv -br \pmod{p}$.
- The algebraic factor base consists of pairs of all (r,p) corresponding to a first-degree prime ideal with p less than some integer B'.
- And we say $\langle a + b\theta \rangle$ is smooth if the norm of the ideal is smooth.

Some Obstructions

- The ideal $\prod_{(a,b)\in U}(a+b\theta)\mathfrak{O}$ of $\mathfrak{O}$ may not be the square of an ideal.
- Even if it is equal to $\mathfrak{J}^2$ for some ideal $\mathfrak{J}$ of $\mathfrak{O}$ the ideal $\mathfrak{J}$ need not be principal.
- Even if $\prod_{(a,b)\in U}(a+b\theta)\mathfrak{O}=\gamma^2\mathfrak{O}$ it is not necessary that $\prod_{(a,b)\in U}(a+b\theta)=\gamma^2$.
- Even if $\prod_{(a,b)\in U}(a+b\theta)=\gamma^2$, we need not have $\gamma\in\mathbb{Z}[\theta]$.
  - If $\prod_{(a,b)\in U}(a+b\theta)=\gamma^2$ with $\gamma\in\mathbb{Q}(\theta)$, then $\gamma\in\mathfrak{O}$ and $\gamma f'(\theta)\in\mathbb{Z}[\theta]$.
  - $f'(\theta)\prod_{(a,b)\in U}(a+b\theta)$ is the square of an element of $\mathbb{Z}[\theta]$.

Some Obstructions

- The ideal $\prod_{(a,b)\in U}(a+b\theta)\mathfrak{O}$ of $\mathfrak{O}$ may not be the square of an ideal.
- Even if it is equal to $\mathfrak{J}^2$ for some ideal $\mathfrak{J}$ of $\mathfrak{O}$ the ideal $\mathfrak{J}$ need not be principal.
- Even if $\prod_{(a,b)\in U}(a+b\theta)\mathfrak{O} = \gamma^2\mathfrak{O}$ it is not necessary that $\prod_{(a,b)\in U}(a+b\theta) = \gamma^2$.
- Even if $\prod_{(a,b)\in U}(a+b\theta) = \gamma^2$, we need not have $\gamma \in \mathbb{Z}[\theta]$.
  - If $\prod_{(a,b)\in U}(a+b\theta) = \gamma^2$ with $\gamma \in \mathbb{Q}(\theta)$, then $\gamma \in \mathfrak{O}$ and $\gamma f'(\theta) \in \mathbb{Z}[\theta]$.
  - $f'(\theta)\prod_{(a,b)\in U}(a+b\theta)$ is the square of an element of $\mathbb{Z}[\theta]$.

Quadratic Characters

Proposition:

Let U be a set of (a, b) such that $\prod_{(a,b)\in U}(a + b\theta) = \alpha^2$ for some $\alpha \in \mathbb{Q}(\theta)$. Given a first-degree prime ideal (s, q) that does not divide $\langle a + b\theta \rangle$ for any pair (a, b) and for which $f'(s) \not\equiv 0 \pmod{q}$, it follows that:

$$\prod_{(a,b)\in U} \left(\frac{a + bs}{q}\right) = 1$$

Filling in The Details

Finding a Polynomial

1. Choosing **odd** degree d for the polynomial.
2. Set $m = \lfloor N^{\frac{1}{d}} \rfloor$.
3. Consider the base-m form of N:

$$N = m^d + a_{d-1}m^{d-1} + \cdots + a_1 m + a_0$$

4. Construct the function

$$f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

Finding a Polynomial

1. Choosing **odd** degree d for the polynomial.
2. Set $m = \lfloor N^{\frac{1}{d}} \rfloor$.
3. Consider the base-m form of N:

$$N = m^d + a_{d-1}m^{d-1} + \cdots + a_1 m + a_0$$

4. Construct the function

$$f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

Finding a Polynomial

1. Choosing **odd** degree d for the polynomial.
2. Set $m = \lfloor N^{\frac{1}{d}} \rfloor$.
3. Consider the base-m form of N:

$$N = m^d + a_{d-1}m^{d-1} + \cdots + a_1 m + a_0$$

4. Construct the function

$$f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

Finding a Polynomial

1. Choosing **odd** degree d for the polynomial.
2. Set $m = \lfloor N^{\frac{1}{d}} \rfloor$.
3. Consider the base-m form of N:

$$N = m^d + a_{d-1}m^{d-1} + \cdots + a_1 m + a_0$$

4. Construct the function

$$f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$$

- $f(m) \equiv 0 \pmod{N}$.
- $f(x)$ is monic.
- $f(x)$ has integer coefficients.
- It may be **reducible**. But it likely leads to splitting N. Since if we have $f(x) = g(x)h(x)$ then $N = f(m) = g(m)h(m)$ which is a non-trivial factorization of N.

Finding First-Degree Prime Ideals of $\mathbb{Z}[\theta]$

Proposition:

When consider as a polynomial in $(\mathbb{Z}/p\mathbb{Z})[x]$, the polynomial $x^p - x$ factor as

$$x^p - x = \prod_{i=0}^{p-1}(x - i)$$

1. To strip out of $f(x)$ any quadratic or higher degree polynomial that occurs in its canonical factorization into irreducibles, let $g(x) = gcd(f(x), x^p - x)$.
2. Now let b be any random integer with $0 \leq b < p$.
3. $g(x - b) | x^p - x = x(x^{p-1} - 1) = x(x^{(p-1)/2} - 1)(x^{(p-1)/2} + 1)$.

Sieving

- Fix a value for b and then scan the values within a range $u < a < u$ for values of $a + bm$ that are smooth.
- let p be a fixed prime.

$$p|a + bm \iff a + bm \equiv 0 \pmod{p} \iff a \equiv -bm \pmod{p}$$

- We can speed up sieving by using approximate logarithm.

Solving The Equation

We are looking for vector x which satisfy $M^T x = 0$.

Since M is sparse we can use special algorithms like Lanczos's or Wiedemann's algorithms.

Computing $\phi(\beta)$ When $\beta^2 \in \mathbb{Z}[\theta]$ Is Known

- Fortunately there is a estimation for size of $\phi(\beta)$.
- We can find $\phi(\beta)$ in all $\mathbb{F}_{p_i}$.
- We can use the Chinese Remainder theorem to find $\phi(\beta)$ in $\mathbb{Z}/N\mathbb{Z}$.

Proposition:

Let $f(x)$ be a monic, irreducible polynomial of **odd** degree d with integer coefficients. Then for any $\alpha \in \mathbb{Q}[\theta]$ it follows that $N(-\alpha) = -N(\alpha)$.

Proposition:

The norm of an element α in the finite field $\mathbb{F}_q$ with $q = p^d$ may be computed as

$$N_p(\alpha) = \alpha^{\frac{p^d-1}{p-1}}$$

Computing $\phi(\beta)$ When $\beta^2 \in \mathbb{Z}[\theta]$ Is Known

- Fortunately there is a estimation for size of $\phi(\beta)$.
- We can find $\phi(\beta)$ in all $\mathbb{F}_{p_i}$.
- We can use the Chinese Remainder theorem to find $\phi(\beta)$ in $\mathbb{Z}/N\mathbb{Z}$.

Proposition:

Let $f(x)$ be a monic, irreducible polynomial of **odd** degree d with integer coefficients. Then for any $\alpha \in \mathbb{Q}[\theta]$ it follows that $N(-\alpha) = -N(\alpha)$.

Proposition:

The norm of an element α in the finite field $\mathbb{F}_q$ with $q = p^d$ may be computed as

$$N_p(\alpha) = \alpha^{\frac{p^d-1}{p-1}}$$

Computing Square Root of δ in $\mathbb{F}^*_{p^d}$

Assume $p^d - 1 = 2^r s$ where s is odd, the idea is to produce a sequence of elements λ_i and ω_i in $\mathbb{F}_{p^d}$ such that $\omega_i^2 = \lambda_i \delta$, with the order o_{i+1} of λ_{i+1} strictly less than the order o_i of λ_i.

If η is quadratic non-residue in $\mathbb{F}^*_{p^d}$ then:

1. Let $\lambda_0 = \delta^s$ and $\omega_0 = \delta^{(s+1)/2}$
2. $\lambda_{i+1} = \lambda_i \eta^{s2^{r-m}}$ which m is the order of λ_i.

Complexity

The basic cost of the algorithm is $u^{2+o(1)} + y^{2+o(1)}$ as N tends to infinity. By trying to minimize this expression we got:

$$\ln y \approx \ln u \approx (\tfrac{8}{9})^{\frac{1}{3}} (\ln N)^{\frac{1}{3}} (\ln \ln N)^{\frac{2}{3}}$$

And:

$$u^{2+o(1)} + y^{2+o(1)} \approx e^{(\frac{64}{9})^{\frac{1}{3}} (\ln N)^{\frac{1}{3}} (\ln \ln N)^{\frac{2}{3}}} = L_N[\tfrac{1}{3}, (\tfrac{64}{9}))^{\frac{1}{3}}]$$

Any Question?

References

Briggs, Matthew E. "An Introduction to the General Number Field Sieve". In: *Master Thesis* (1998).

Buhler, J. P., H. W. Lenstra, and Carl Pomerance. "Factoring integers with the number field sieve". In: 1993. DOI: 10.1007/bfb0091539.

Couveignes, Jean-Marc. "Computing a square root for the number field sieve". In: 1993. DOI: 10.1007/bfb0091540.

Forbes, Tony. "Prime numbers: A computational perspective, by Richard Crandall and Carl Pomerance. Pp. 545. 2001. ISBN 0 387 94777 9 (Springer-Verlag).". In: *The Mathematical Gazette* 86 (507 2002). ISSN: 0025-5572. DOI: 10.2307/3621190.

Gerhard, Jurgen. *Modern computer algebra: Third edition*. 2011. DOI: 10.1017/CBO9781139856065.

Jarvis, Frazer. *Algebraic Number Theory*. Springer Cham, 2014. DOI: 10.1007/978-3-319-07545-7.

Knott, Ron. "Prime numbers - a computational perspective (2nd edn), by Richard Crandall and Carl Pomerance. Pp. 597. £42.50 (hbk). 2005. ISBN 0-387-25282-7 (Springer Verlag).". In: *The Mathematical Gazette* 92 (523 2008). ISSN: 0025-5572. DOI: 10.1017/s0025557200182932.

Montgomery, Peter L. "A block Lanczos algorithm for finding dependencies over GF(2)". In: vol. 921. 1995. DOI: 10.1007/3-540-49264-X_9.

Stevenhagen, Peter. "THE NUMBER FIELD SIEVE Peter Stevenhagen". In: *Wstein.Org* (2004).