Cours MPRI 2-12-2

Lecture 5/5: NFS

(lecturer for part 2/3): E. Thomé

Dec. 17th, 2012

Plan

Teaser: factoring with cubic integers

General principle

Another rosy example (skipped)

Doing it seriously

Complexity analysis

The initial idea

Factoring $F_7 = 2^{128} + 1$ was one of the early achievements of CFRAC in the 1970's.

Is there another way ?

Pollard noticed:

$$2F_7 = 2^{129} + 2 = m^3 + 2, \text{ with } m = 2^{43}.$$

Factoring $2F_7$

We have $2F_7 = 2^{129} + 2 = m^3 + 2$, with $m = 2^{43}$.

Define the number field $K = \mathbb{Q}(\alpha = \sqrt[3]{-2})$.

- K is one of the textbook examples of number fields.
- The algebraic integers in K are $\mathbb{Z}[\alpha]$. These possess unique factorization. (lucky !)

Assume we have many (a, b)'s such that:

- The integer $a - bm$ is smooth (w.r.t some bound B).
  $\Rightarrow$ write $a - bm$ as a product of primes (and possibly -1).
- The algebraic integer $a - b\alpha$ too.
  $\Rightarrow$ write $a - b\alpha$ as a product of algebraic integers (and possibly units).

Collect sufficiently many, and combine to make all valuations even!

Obstructions ?

Even in the simple example of $2F_7$, we have possible complications.

$$\begin{aligned}\mathrm{Norm}(2\alpha^2 - 3\alpha + 1) &= 51 = 3 \times 17, \\ (2\alpha^2 - 3\alpha + 1) &= (\alpha - 1) \times (2\alpha^2 + \alpha - 1) \times \text{unit}, \\ &= (\alpha - 1) \times (2\alpha^2 + \alpha - 1) \times (-\alpha^2 + \alpha - 1).\end{aligned}$$

The units which appear have to be taken into account.

- Not too frightening for $\mathbb{Q}(\sqrt[3]{2})$, but problematic for bigger fields.
- Units are only one of the obstructions encountered.

Plan

NFS as a factoring algorithm

NFS is among the algorithms which search for solutions to:

$$X^2 \equiv Y^2 \mod N,$$

as a means to factor N.

- For $N = pq$, such a congruence reveals a non-trivial factor $\gcd(X - Y, N)$ with probability $1/2$.
- Several congruences of squares are needed.
- NFS will never factor p^2q as $p \times pq$. Always $p^2 \times q$. (but anyway, detecting prime powers is trivial).

Strategy (1)

Goal: let squares modulo N appear as images of squares in something else via ring morphisms from two different strutures.

NFS:
- these ring morphisms come from number fields
- usually, we take one of these number fields to be $\mathbb{Q}$.

NFS as a framework also embraces NFS-DL and FFS. (although we care less about squares in that case).

Strategy (2)

$$\varphi(\text{a square somewhere}) = \text{a square in } \mathbb{Z}/N\mathbb{Z}.$$

Fabricating a square in this "somewhere":

- Focus on smooth objects which can be written in factored form.
- Restrict to those which factor over a factor base (set of prescribed size).
- Gather sufficiently many.
- Combine in order to build a square (all exponents even).
- Recover the square root of this square.

NFS takes long routes to achieve this.

Which "somewhere" do we choose?

Consider:

- a number field $K = \mathbb{Q}(\alpha)$ defined by $f(\alpha) = 0$, for f irreducible over $\mathbb{Q}$ and $\deg f = d$;
- extra constraint: $\exists m \in \mathbb{Z},\ f(m) \equiv 0 \mod N$.

This provides a ring morphism: $\begin{cases} \mathbb{Z}[\alpha] & \rightarrow & \mathbb{Z}/N\mathbb{Z}, \\ \alpha & \mapsto & m \mod N. \end{cases}$

The pair (f, m) is well suited to factoring N.

Broader NFS terminology refers to (f, g), with $g = x - m$.

The GNFS setup

For factoring "general" N, GNFS uses:

- a number field $K = \mathbb{Q}(\alpha)$ defined by $f(\alpha) = 0$, for f irreducible over $\mathbb{Q}$ and $\deg f = d$;
- Another irreducible polynomial g such that f and g have a common root m mod N (example: $g = x - m$).

g defines the rational side, f defines the algebraic side.

Restating with the resultant

The following restatement can be useful.

$$f \text{ and } g \text{ share a root modulo } N \Leftrightarrow \mathrm{Res}_x(f, g) = 0 \mod N.$$

Choosing f and g is referred to as the polynomial selection step.

Structures

- f defines $K = \mathbb{Q}(\alpha)$ (and the ring $\mathbb{Z}[\alpha] \subset K$).
- g defines $\mathbb{Q}$, but in a fancy way (and the ring $\mathbb{Z}[m] \subset \mathbb{Q}$).

Ring morphisms (because m is a root of both modulo N):

$$\varphi_f : \begin{cases} \mathbb{Z}[\alpha] & \rightarrow & \mathbb{Z}/N\mathbb{Z}, \\ T(\alpha) & \mapsto & T(m) \bmod N, \end{cases} \qquad \varphi_g : \begin{cases} \mathbb{Z}[m] & \rightarrow & \mathbb{Z}/N\mathbb{Z}, \\ t & \mapsto & t \bmod N. \end{cases}$$

These morphism are arrows inside a commutative diagram.

Note: having $\deg g > 1$ is also allowed (but making up examples is harder).

The diagram

$$
\begin{tikzcd}
 & \mathbb{Z}[x] \arrow[dl, "x \mapsto m"'] \arrow[dr, "x \mapsto \alpha"] & \\
\mathbb{Z}[m] \arrow[dr, "\varphi_g : t \mapsto t \mod N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\varphi_f : \alpha \mapsto m \mod N"] \\
 & \mathbb{Z}/N\mathbb{Z} &
\end{tikzcd}
$$

This diagram commutes.

Relations in NFS

$$
\begin{tikzcd}
 & \mathbb{Z}[x] \arrow[dl, "\psi^{(1)} : x \mapsto m"'] \arrow[dr, "\psi^{(2)} : x \mapsto \alpha"] & \\
\mathbb{Z}[m] \arrow[dr, "\varphi_g : t \mapsto t \mod N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\varphi_f : \alpha \mapsto m \mod N"] \\
 & \mathbb{Z}/N\mathbb{Z} &
\end{tikzcd}
$$

Take for example $a - bx$ in $\mathbb{Z}[x]$. Suppose for a moment that:

- the integer $a - bm$ is smooth: product of factor base primes;
- the algebraic integer $a - b\alpha$ is also a product.
- factors occuring on both sides belong to a small set (factor base).

NFS collects many such "good pairs" (a, b).

Collecting relations

Suppose factor bases are:
- $\{p_1, \ldots, p_{99}\}$ (rational),
- $\{\pi_1, \ldots, \pi_{99}\}$ (algebraic).

Good pairs could lead to:

$$a_1 - b_1 m = p_2 \times p_4^3 \times p_{12} \times p_{22},$$
$$a_2 - b_2 m = p_1 \times p_3 \times p_5^2 \times p_{47},$$
$$a_3 - b_3 m = p_2 \times p_7 \times p_{12},$$
$$a_4 - b_4 m = p_1^6 \times p_4 \times p_7 \times p_{22},$$

Collecting relations

Suppose factor bases are:
- $\{p_1, \ldots, p_{99}\}$ (rational),
- $\{\pi_1, \ldots, \pi_{99}\}$ (algebraic).

Good pairs could lead to:	and at the same time:
$a_1 - b_1 m = p_2 \times p_4^3 \times p_{12} \times p_{22}$,	$a_1 - b_1\alpha = \pi_1 \times \pi_3^2 \times \pi_6^2 \times \pi_{35}$,
$a_2 - b_2 m = p_1 \times p_3 \times p_5^2 \times p_{47}$,	$a_2 - b_2\alpha = \pi_2 \times \pi_8^2 \times \pi_{29}$,
$a_3 - b_3 m = p_2 \times p_7 \times p_{12}$,	$a_3 - b_3\alpha = \pi_1^3 \times \pi_3 \times \pi_{23} \times \pi_{35}$,
$a_4 - b_4 m = p_1^6 \times p_4 \times p_7 \times p_{22}$,	$a_4 - b_4\alpha = \pi_2^4 \times \pi_3 \times \pi_{23}$,

Mission

Our plan is to have something which is a square on both sides.
NFS intends to achieve this by combining relations.

Combining relations

$$
\begin{array}{l|l}
a_1 - b_1 m = p_2 \times p_4^3 \times p_{12} \times p_{22}, & a_1 - b_1\alpha = \pi_1 \times \pi_3^2 \times \pi_6^2 \times \pi_{35}, \\
a_2 - b_2 m = p_1 \times p_3 \times p_5^2 \times p_{47}, & a_2 - b_2\alpha = \pi_2 \times \pi_8^2 \times \pi_{29}, \\
a_3 - b_3 m = p_2 \times p_7 \times p_{12}, & a_3 - b_3\alpha = \pi_1^3 \times \pi_3 \times \pi_{23} \times \pi_{35}, \\
a_4 - b_4 m = p_1^6 \times p_4 \times p_7 \times p_{22}, & a_4 - b_4\alpha = \pi_2^4 \times \pi_3 \times \pi_{23},
\end{array}
$$

- Find a combination which makes all exponents even.
- Evaluating $(a_1 - b_1 x)(a_3 - b_3 x)(a_4 - b_4 x)$ at both m and α leads to a square on both sides.
- Apply φ_g and φ_f: we get a congruence of squares in $\mathbb{Z}/N\mathbb{Z}$.

Combining relations

$$\begin{array}{l|l}
a_1 - b_1 m = p_2 \times p_4^3 \times p_{12} \times p_{22}, & a_1 - b_1\alpha = \pi_1 \times \pi_3^2 \times \pi_6^2 \times \pi_{35}, \\
a_2 - b_2 m = p_1 \times p_3 \times p_5^2 \times p_{47}, & a_2 - b_2\alpha = \pi_2 \times \pi_8^2 \times \pi_{29}, \\
a_3 - b_3 m = p_2 \times p_7 \times p_{12}, & a_3 - b_3\alpha = \pi_1^3 \times \pi_3 \times \pi_{23} \times \pi_{35}, \\
a_4 - b_4 m = p_1^6 \times p_4 \times p_7 \times p_{22}, & a_4 - b_4\alpha = \pi_2^4 \times \pi_3 \times \pi_{23},
\end{array}$$

- Find a combination which makes all exponents even.
- Evaluating $(a_1 - b_1 x)(a_3 - b_3 x)(a_4 - b_4 x)$ at both m and α leads to a square on both sides.
- Apply φ_g and φ_f: we get a congruence of squares in $\mathbb{Z}/N\mathbb{Z}$.

Caveat

This is too rosy. $\mathbb{Z}[\alpha]$ not a UFD. Complications ahead.

Plan

Exemple de factorisation par NFS

On s'intéresse à $N = 16259 = 16384 - 125 = 16384 - 128 + 3$.
On pose:

$$f(x) = x^2 - x + 3,\ \ m = 128,\ \ g(x) = x - m.$$

On a ainsi: $f(m) = N$ et $g(m) = 0$.
Soit α une racine de f dans $\mathbb{C}$ ($\alpha = \frac{1}{2}(1 + \sqrt{-11})$).

$$\begin{tikzcd}
 & \mathbb{Z}[x] \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & \\
\mathbb{Z}[m] \arrow[dr, "\varphi_g\,:\, t \to t \mod N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\varphi_f\,:\, \alpha \to m \mod N"] \\
 & \mathbb{Z}/N\mathbb{Z} &
\end{tikzcd}$$

Nombres premiers dans $\mathbb{Z}[\alpha]$

Coup de chance, $\mathbb{Z}[\alpha]$ est un anneau euclidien.
Certains nombres premiers dans $\mathbb{Z}$ se factorisent dans $\mathbb{Z}[\alpha]$.
Les nombres premiers de $\mathbb{Z}[\alpha]$ sont:

$$\begin{array}{c|c}
2, & 17, \\
3 = \alpha \times (1-\alpha), & 19, \\
5 = (1+\alpha) \times (2-\alpha), & 23 = (4+\alpha) \times (5-\alpha), \\
7, & 29, \\
11 = -(1-2\alpha)^2, & 31 = (4-3\alpha) \times (1+3\alpha), \\
13, & 37 = (2+3\alpha) \times (5-3\alpha),
\end{array} \quad \ldots$$

Note: comme $\alpha^2 - \alpha + 3 = 0$, on a $\bar{\alpha} = 1 - \alpha$.

Factoriser des deux côtés

$$
\begin{tikzcd}
 & \mathbb{Z}[x] \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & \\
\mathbb{Z}[m] \arrow[dr, "\varphi_g : t \to t \mod N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\varphi_f : \alpha \to m \mod N"] \\
 & \mathbb{Z}/N\mathbb{Z} &
\end{tikzcd}
$$

- On part de $a - bx \in \mathbb{Z}[x]$.
- On espère avoir $a - bm$ et $a - b\alpha$ simultanément friables.
- Par résolution d'un système linéaire, on fabrique un carré de chaque côté.

Relations

On veut des nombres premiers inférieurs à $B = 40$.

$$
\begin{aligned}
1 - 1m &= -127 = -127,\\
1 - 2m &= -255 = -3 \times 5 \times 17,\\
1 - 3m &= -383 = -383,\\
1 - 4m &= -511 = -7 \times 73,\\
1 - 5m &= -639 = -3^2 \times 71,\\
2 - 1m &= -126 = -2 \times 3^2 \times 7,\\
2 - 3m &= -382 = -2 \times 191,\\
2 - 5m &= -638 = -2 \times 11 \times 29,\\
3 - 1m &= -125 = -5^3,\\
3 - 2m &= -253 = -11 \times 23,\\
3 - 4m &= -509 = -509,\\
3 - 5m &= -637 = -7^2 \times 13,\\
4 - 1m &= -124 = -2^2 \times 31,\\
4 - 3m &= -380 = -2^2 \times 5 \times 19,
\end{aligned}
$$

$$
\begin{aligned}
4 - 5m &= -636 = -2^2 \times 3 \times 53,\\
5 - 1m &= -123 = -3 \times 41,\\
5 - 2m &= -251 = -251,\\
5 - 3m &= -379 = -379,\\
5 - 4m &= -507 = -3 \times 13^2,\\
6 - 1m &= -122 = -2 \times 61,\\
6 - 5m &= -634 = -2 \times 317,\\
7 - 1m &= -121 = -11^2,\\
7 - 2m &= -249 = -3 \times 83,\\
7 - 3m &= -377 = -13 \times 29,\\
7 - 4m &= -505 = -5 \times 101,\\
7 - 5m &= -633 = -3 \times 211,\\
8 - 1m &= -120 = -2^3 \times 3 \times 5,\\
8 - 3m &= -376 = -2^3 \times 47,
\end{aligned}
$$

On ne garde que ce qui est bon

$1 - 2m = -255 = -3 \times 5 \times 17,$
$2 - 1m = -126 = -2 \times 3^2 \times 7,$
$2 - 5m = -638 = -2 \times 11 \times 29,$
$3 - 1m = -125 = -5^3,$
$3 - 2m = -253 = -11 \times 23,$
$3 - 5m = -637 = -7^2 \times 13,$
$4 - 1m = -124 = -2^2 \times 31,$
$4 - 3m = -380 = -2^2 \times 5 \times 19,$
$5 - 4m = -507 = -3 \times 13^2,$
$7 - 1m = -121 = -11^2,$
$7 - 3m = -377 = -13 \times 29,$
$8 - 1m = -120 = -2^3 \times 3 \times 5,$
$9 - 1m = -119 = -7 \times 17,$
$9 - 2m = -247 = -13 \times 19,$

$10 - 3m = -374 = -2 \times 11 \times 17,$
$11 - 1m = -117 = -3^2 \times 13,$
$11 - 2m = -245 = -5 \times 7^2,$
$11 - 5m = -629 = -17 \times 37,$
$12 - 1m = -116 = -2^2 \times 29,$
$13 - 1m = -115 = -5 \times 23,$
$13 - 2m = -243 = -3^5,$
$13 - 5m = -627 = -3 \times 11 \times 19,$
$14 - 1m = -114 = -2 \times 3 \times 19,$
$14 - 3m = -370 = -2 \times 5 \times 37,$
$16 - 1m = -112 = -2^4 \times 7,$
$16 - 3m = -368 = -2^4 \times 23,$
$16 - 5m = -624 = -2^4 \times 3 \times 13,$
$17 - 1m = -111 = -3 \times 37,$

Côté algébrique

On fait pareil.

Pour factoriser $a - b\alpha$, on commence par calculer la norme:

$$N(a - b\alpha) = (a - b\alpha)(a - b\bar{\alpha}) = b^{\deg f} f(a/b).$$

En fonction de la factorisation de la norme, on détermine les facteurs présents.

$$\begin{aligned}
1 - \alpha &= (1 - \alpha), \\
1 - 2\alpha &= (1 - 2\alpha), \\
1 - 3\alpha &= (2 - \alpha)^2, \\
1 - 4\alpha &= (1 - \alpha)^2 \times (1 + \alpha), \\
1 - 5\alpha &= (1 - 5\alpha), \\
2 - \alpha &= (2 - \alpha), \\
2 - 3\alpha &= -(1 + \alpha)^2, \\
2 - 5\alpha &= (1 - \alpha) \times (5 - \alpha), \\
3 - \alpha &= -(\alpha)^2,
\end{aligned}$$

$$\begin{aligned}
3 - 2\alpha &= -(\alpha) \times (1 + \alpha), \\
3 - 4\alpha &= -(\alpha)^2 \times (2 - \alpha), \\
3 - 5\alpha &= -(\alpha) \times (4 + \alpha), \\
4 - \alpha &= (1 - \alpha) \times (1 + \alpha), \\
4 - 3\alpha &= (4 - 3\alpha), \\
4 - 5\alpha &= (4 - 5\alpha), \\
5 - \alpha &= (5 - \alpha), \\
5 - 2\alpha &= -(1 - \alpha)^3, \\
5 - 3\alpha &= (5 - 3\alpha),
\end{aligned}$$

Friabilité simultanée

$$\begin{array}{rl}
1+3m = 5\times 7\times 11 & 1+3\alpha = (3\alpha+1),\\
1-2m = -3\times 5\times 17 & 1-2\alpha = -(2\alpha-1),\\
2+1m = 2\times 5\times 13 & 2+1\alpha = -(-\alpha+1)^2,\\
2-1m = -2\times 3^2\times 7 & 2-1\alpha = (-\alpha+2),\\
2-5m = -2\times 11\times 29 & 2-5\alpha = (-\alpha+1)\times(-\alpha+5),\\
3+2m = 7\times 37 & 3+2\alpha = -(\alpha)^3,\\
3-1m = -5^3 & 3-1\alpha = -(\alpha)^2,\\
3-2m = -11\times 23 & 3-2\alpha = -(\alpha)\times(\alpha+1),\\
3-5m = -7^2\times 13 & 3-5\alpha = -(\alpha)\times(\alpha+4),\\
4+5m = 2^2\times 7\times 23 & 4+5\alpha = -(-\alpha+1)\times(-3\alpha+5),\\
4+1m = 2^2\times 3\times 11 & 4+1\alpha = (\alpha+4),\\
4-1m = -2^2\times 31 & 4-1\alpha = (-\alpha+1)\times(\alpha+1),\\
4-3m = -2^2\times 5\times 19 & 4-3\alpha = (-3\alpha+4),\\
5+1m = 7\times 19 & 5+1\alpha = (-\alpha+1)\times(2\alpha-1),\\
7-1m = -11^2 & 7-1\alpha = -(-\alpha+1)^2\times(-\alpha+2),\\
\end{array}$$

$$\ldots$$

Trouver un carré

Soit:

$$p(x) = (2x+3) \times (-3x+7) \times (\alpha+8) \times (-2x+9) \\ \times (-\alpha+14) \times (-\alpha+16) \times (-\alpha+17) \times (-4x+19).$$

On a

$$p(m) = 2^8 \times 3^2 \times 7^2 \times 13^2 \times 17^2 \times 19^2 \times 29^2 \times 37^2,$$
$$p(\alpha) = (\alpha)^4 \times (-\alpha+1)^8 \times (\alpha+1)^2 \times (-\alpha+2)^6 \\ \times (2\alpha-1)^2 \times (-3\alpha+5)^2.$$

Beaucoup mieux:

$$p(x) = (7-x) \times (17+4x).$$

Mais dans ce cas, on aurait $p(m) = -\square$.

This is all cheating

The example above is too easy (on purpose, of course).

- The number N comes with an "obvious" f ;
- f is chosen so that $\mathbb{Z}[\alpha]$ is the maximal order ;
- f is monic ;
- the unit group of K is $\{\pm 1\}$;
- the class group of K is trivial ;
- $\mathbb{Z}[\alpha]$ is even a euclidean ring (although not even a UFD in general !);

How does it work in real life (but still for f monic, for clarity) ?

Plan

NFS

Major obstruction: $\mathbb{Z}[\alpha]$ not a UFD.

Outline of the algorithm:

- Do the setup. Choose a factor base bound B ;
- **Relation search**

  Pick pairs a, b for coprime integers a and b ;

  - Expect $a - bm$ to be a smooth integer ;
  - Expect also the ideal $(a - b\alpha)$ to be smooth ;
- Do some combination work, recover an equality of squares.

Purpose of the next slides:
- How the identity of squares appears ;
- Analysis.

Living in number fields

The subring $\mathbb{Z}[\alpha]$ lacks some desired properties.

- The "most $\mathbb{Z}$-like" ring in K is the ring of integers $\mathcal{O}_K$.
- $\mathcal{O}_K$ is unfortunately hard to compute in general, but can be approximated.
- Even $\mathcal{O}_K$ lacks unique factorization.
    - Instead, try to factor ideals into prime ideals.
    - This also implies that $\mathcal{O}_K$ is not principal.

Living in number fields

The subring $\mathbb{Z}[\alpha]$ lacks some desired properties.

- The "most $\mathbb{Z}$-like" ring in K is the ring of integers $\mathcal{O}_K$.
- $\mathcal{O}_K$ is unfortunately hard to compute in general, but can be approximated.
- Even $\mathcal{O}_K$ lacks unique factorization.
  - Instead, try to factor ideals into prime ideals.
  - This also implies that $\mathcal{O}_K$ is not principal.

Prime ideals in $\mathcal{O}_K$ are commonly written e.g. $\mathfrak{p}, \mathfrak{q}, \mathfrak{a}, \mathfrak{b}$.

- Most ideals can be written in a simple form:

$$\mathfrak{p} = \langle p, \alpha - r \rangle.$$

- Computing the norm is a first step towards factoring, since:

$$\mathrm{Norm}(\mathfrak{a}\mathfrak{b}) = \mathrm{Norm}(\mathfrak{a})\,\mathrm{Norm}(\mathfrak{b}).$$

Fetching smooth data

Finding a, b such that $a - bm$ is smooth: easily stated.

Finding a, b such that $(a - b\alpha)\mathcal{O}_K$ is a smooth ideal:

- When $I = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_k^{e_k}$, we have $\operatorname{Norm} I = \prod_i (\operatorname{Norm} \mathfrak{p}_i)^{e_i}$.
- Look at $\operatorname{Norm}((a - b\alpha)\mathcal{O}_K) = \operatorname{Norm}_{K/\mathbb{Q}}(a - b\alpha) = b^{\deg f} f(a/b)\ (\in \mathbb{Z})$.
- If this norm is smooth, then $(a - b\alpha)\mathcal{O}_K$ is a smooth ideal.
- Note: because $a - b\alpha$ has degree 1 in α, ideals $\mathfrak{p}$ are “simple”.

Each pair a, b meeting these conditions yields a relation.

For each relation, we focus on valuations at primes / prime ideals.

Searching for relations

To search for relations, NFS uses sieving.

Old technique: line sieving.

- Decide on a search space for (a, b) values.
- For each prime p, mark (coprime) (a, b)'s s.t. $p \mid a - bm$.
- Only a subset of the search space survives.
- For each prime ideal $\mathfrak{p}$, mark $\mathfrak{p} \mid a - b\alpha$.

$$\mathfrak{p} = \langle p, \alpha - r \rangle \mid a - b\alpha, \\ \Updownarrow \\ a - br \equiv 0 \mod p.$$

- Pairs which survive both sieves yield relations.

All large prime variants allowed.

Lattice sieving

Newer technique: divide the computation into smaller ranges of interest based on a divisibility condition, e.g. $\mathfrak{q} \mid (a - b\alpha)$.

- The set of pairs (a, b) meeting the condition is a $\mathbb{Z}$-lattice.
- Pick a short basis, and take small combinations of the vectors (e.g. $i\vec{u} + j\vec{v}$, for small i, j).
- In (i, j) coordinates, sieve as before.

Lattice sieving is superior because:

- It is more cache-friendly,
- It can be optimized well,
- It allows stable yields.

All large prime variants still allowed.

Linear system

Build a matrix where each row corresponds to an a, b pair.

- First set of columns: valuations (mod2) of $a - bm$ at primes $p < B$.
- Second set of columns: valuations (mod2) of $(a - b\alpha)\mathcal{O}_K$ at unramified prime ideals $\mathfrak{p}$ of norm $< B$ (and residue class degree 1).
  - For simplicity, we completely forget about ramified ideals, and more generally, all "special ideals" (whose norm is not coprime to disc K).
  - Remaining problem: knowing $\nu_p(\mathrm{Norm}_{K/\mathbb{Q}}(a - b\alpha)) = v$, determine $\nu_{\mathfrak{p}}((a - b\alpha))$ for prime ideals above $\mathfrak{p}$.

Prop. For a, b coprime, exactly one ideal $\mathfrak{p}$ above p has $\nu_{\mathfrak{p}}((a - b\alpha)) = v$. This $\mathfrak{p}$ is the unique ideal $(p, \alpha - r)$ for which $a - br \equiv 0 \mod p$.

Linear system

Consider for example the pair $a = 61$, $b = 9$, for the NFS setup given by $f = x^3 - 39$ and $m = 1006$. We have:

- $a - bm = 61 - 9 \times 1006 = -1 \times 17 \times 23^2$;
- $\mathrm{Norm}_{K/\mathbb{Q}}(a - b\alpha) = 61^3 - 39 \times 9^3 = 2 \times 5^2 \times 11 \times 19^2$.

This yields the valuation vector:

Nullspace of the relation matrix

The left nullspace yields polynomials $R(x)$ such that:

- $R(m) = \pm\Box$ (because for all p, $\nu_p(R(m))$ is even) ;
- $(R(\alpha))$ is a product of special ideals times the square of an ideal J (for all non-special $\mathfrak{p}$, $\nu_{\mathfrak{p}}((R(\alpha)))$ is even).

This, however, is not enough:

- We haven't kept track of the sign of $R(m)$;
- $(R(\alpha))$ is not exactly the square of an ideal ;
- Even if it were, while $(R(\alpha))$ is a principal ideal by construction, its square root has no reason for being principal ;
- Even assuming we have $(R(\alpha)) = (\gamma^2)$, this defines γ only up to a unit. The equation to solve is $R(\alpha) = \gamma^2\epsilon$, and units are intractable.
- We have no guarantee that $\gamma \in \mathbb{Z}[\alpha]$.

We know how to handle all this.

Plan

Teaser: factoring with cubic integers

General principle

Another rosy example (skipped)

Doing it seriously

Complexity analysis

Simplifications for analysis

Some important improvements have no effect on the overall complexity.

- Polynomial selection.
- Large primes, cofactorization.
- Linear algebra optimizations.

OTOH, sieving does serve to eliminate the per-pair factoring.

Key figures for complexity analysis

There's one main theorem known as:

- Canfield-Erdős-Pomerance,
- Construction kit lemma,
- whatever credit people give... (Odlyzko / Balasubramanian)

It's also valid in various contexts.

Canfield-Erdős-Pomerance (CEP) Theorem

Let $x, y \to +\infty$ and $\epsilon > 0$ s.t. $(\log x)^{\epsilon} < \log y < (\log x)^{1-\epsilon}$.

$$\frac{1}{x}\#\{n,\ 1 \leq n \leq x,\ n \text{ is } y\text{-smooth}\} \sim \rho(u) = u^{-u(1+o(1))}$$

where $u = \frac{\log x}{\log y}$, and ρ is the Dickman-de Bruijn function.

A gross estimate for analytic number theorists, but sufficient for us.

The L function

We introduce:

$$L_x[a,\alpha] = \exp\left(\alpha(\log x)^a(\log\log x)^{1-a}\right).$$

CEP with the L function

A random integer $n \leq L_x[a,\alpha]$ is $L_x[b,\beta]$-smooth with probability:

$$\pi = L_x\left[a-b, -\frac{\alpha}{\beta}(a-b)(1+o(1))\right].$$

This formulation is very important for analyzing sieve algorithms.

Calculus with L

Basic formulae with L

$$L_x[a,\alpha] \times L_x[b,\beta] = \begin{cases} L_x[a,\alpha+o(1)] & \text{if } a > b, \\ L_x[b,\beta+o(1)] & \text{if } b > a, \\ L_x[a,\alpha+\beta] & \text{if } a = b. \end{cases}$$

$$L_x[a,\alpha] + L_x[b,\beta] = \begin{cases} L_x[a,\alpha+o(1)] & \text{if } a > b, \\ L_x[b,\beta+o(1)] & \text{if } b > a, \\ L_x[a,\max(\alpha,\beta)] & \text{if } a = b. \end{cases}$$

$$L_{L_x[b,\beta]}[a,\alpha] = L_x[ab, \alpha\beta^a b^{1-a} + o(1)].$$

$$L_x[b,\beta]^{\log_{\log x} L_x[a,\alpha]} = L_x[a+b, \alpha\beta].$$

Analysis (1)

- Let $d = \log_{\log N} L_N[\Delta, \delta]$ be the number field degree.
  The "trivial" polynomial selection yields:

$$m \approx f_i \approx N^{1/d+1} = L_N[1 - \Delta, \frac{1}{d}].$$

- Let $S = L_N[s, \sigma]$ be the bound on the (a, b) pairs.
  Then $\mathrm{Res}(a - bx, f)$ and $\mathrm{Res}(a - bx, g)$ are bounded by:

$$S^d \times \|f\| = L_N[s + \Delta, \sigma\delta] \times L_N[1 - \Delta, \frac{1}{\delta}],$$

$$S \times m = L_N[s, \sigma] \times L_N[1 - \Delta, \frac{1}{\delta}].$$

Analysis (1)

- Let $d = \log_{\log N} L_N[\Delta, \delta]$ be the number field degree.
  The "trivial" polynomial selection yields:

$$m \approx f_i \approx N^{1/d+1} = L_N[1 - \Delta, \frac{1}{d}].$$

- Let $S = L_N[s, \sigma]$ be the bound on the (a, b) pairs.
  Then $\mathrm{Res}(a - bx, f)$ and $\mathrm{Res}(a - bx, g)$ are bounded by:

$$S^d \times \|f\| = L_N[s + \Delta, \sigma\delta] \times L_N[1 - \Delta, \frac{1}{\delta}],$$

$$S \times m = L_N[s, \sigma] \times L_N[1 - \Delta, \frac{1}{\delta}].$$

Minimize the norms (fix Δ)

Set $1 - \Delta = s + \Delta$, i.e. $\Delta = \frac{1-s}{2}$, whence $1 - \Delta = s + \Delta = \frac{1+s}{2}$.

Analysis (1)

- Let $d = \log_{\log N} L_N[\Delta, \delta]$ be the number field degree. The "trivial" polynomial selection yields:

$$m \approx f_i \approx N^{1/d+1} = L_N[1-\Delta, \frac{1}{d}].$$

- Let $S = L_N[s, \sigma]$ be the bound on the (a, b) pairs. Then $\mathrm{Res}(a - bx, f)$ and $\mathrm{Res}(a - bx, g)$ are bounded by:

$$S^d \times \|f\| = L_N[\frac{1+s}{2}, \sigma\delta + \frac{1}{\delta}],$$

$$S \times m = L_N[\frac{1+s}{2}, \frac{1}{\delta}(1 + o(1))].$$

Minimize the norms (fix Δ)

Set $1 - \Delta = s + \Delta$, i.e. $\Delta = \frac{1-s}{2}$, whence $1 - \Delta = s + \Delta = \frac{1+s}{2}$.

Analysis (2)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1+s}{2} - b, -(\frac{1+s}{2} - b)\frac{1}{\beta}(\sigma\delta + \frac{2}{\delta}) + o(1)].$$

Analysis (2)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1+s}{2} - b, -(\frac{1+s}{2} - b)\frac{1}{\beta}(\sigma\delta + \frac{2}{\delta}) + o(1)].$$

Optimize the probability so as to fix δ

$$\sigma\delta + \frac{2}{\delta} \text{ minimal} \Rightarrow \delta = \sqrt{2/\sigma},$$

$$\Rightarrow \pi = L_N[\frac{1+s}{2} - b, -(\frac{1+s}{2} - b)\frac{1}{\beta}2\sqrt{2\sigma}].$$

Analysis (3)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1+s}{2} - b, -(\frac{1+s}{2} - b)\frac{1}{\beta} 2\sqrt{2\sigma} + o(1)].$$

Analysis (3)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1+s}{2} - b, -(\frac{1+s}{2} - b)\frac{1}{\beta}2\sqrt{2\sigma} + o(1)].$$

- Number of relations obtained: $S^2\pi$.
- Number of relations needed: $\widetilde{O}(B)$.
- Total cost of sieving: $O(S^2)$.
- Cost of linear algebra: $O(B^2)$.

Equate sieving and linear algebra

$$S \approx B \Rightarrow b = s,\ \beta = \sigma.$$

Analysis (4)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1-b}{2}, -(\frac{1-b}{2})2\sqrt{2/\beta} + o(1)].$$

- Number of relations obtained: $\widetilde{O}(B^2\pi)$.
- Number of relations needed: $\widetilde{O}(B)$.

Analysis (4)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1-b}{2}, -(\frac{1-b}{2})2\sqrt{2/\beta} + o(1)].$$

- Number of relations obtained: $\widetilde{O}(B^2\pi)$.
- Number of relations needed: $\widetilde{O}(B)$.

Just enough relations

$B^2\pi \approx B$, thus $1/\pi \approx B$. Two consequences.

$$(1-b)/2 = b \Rightarrow b = 1/3,$$

$$\Rightarrow \pi = L_N[\frac{1}{3}, -\frac{1}{3}2^{3/2}\sqrt{1/\beta} + o(1)],$$

Analysis (4)

Let $B = L_N[b, \beta]$ be the smoothness bound.

- Number of primes / prime ideals: $\widetilde{O}(B) = L_N[b, \beta + o(1)]$.
- Smoothness probability:

$$\pi = L_N[\frac{1}{3}, -\frac{1}{3}2^{3/2}\sqrt{1/\beta} + o(1)].$$

Just enough relations

$B^2\pi \approx B$, thus $1/\pi \approx B$. Two consequences ; $b = 1/3$, and:

$$\begin{aligned} \beta &= \frac{1}{3}2^{3/2}\sqrt{1/\beta}, \\ (\beta/2)^{3/2} &= \frac{1}{3}, \\ \beta &= 2\sqrt[3]{9} = \sqrt[3]{8/9}. \end{aligned}$$

Complexity of NFS

For factoring an integer N, GNFS takes time:

$$L_N[1/3, (64/9)^{1/3}] = \exp\left((1+o(1))(64/9)^{1/3}(\log N)^{1/3}(\log\log N)^{2/3}\right).$$

This is sub-exponential.

Note: some special numbers allow for a faster variant NFS, with complexity

$$L_N[1/3, (32/9)^{1/3}] = \exp\left((1+o(1))(32/9)^{1/3}(\log N)^{1/3}(\log\log N)^{2/3}\right).$$

Remarks related to analysis

- The two norms are $L_N[2/3, \frac{1}{\delta}]$ and $L_N[2/3, \frac{1}{\delta} + \sigma\delta]$.
  The algebraic norm is intrisically larger in the GNFS case.
- The 4 steps of the analysis may be done in various orders, but lead to the same thing.

The SNFS case

SNFS numbers are those for which a polynomial f exists which leads to smaller norms than the GNFS.

Example: assuming the right degree, coeffs $\ll L_N[2/3, \sqrt[3]{3}]$.

Typical example: Cunnigham numbers

Assume $N = 2^{1039} - 1$. A good choice is:
- $g = x - 2^{173}$.
- $f = 2x^6 - 1$.

Notes:

- In some cases, f is rather tiny.
- The rational norm may well become the largest one.
- Exceptional Galois groups are no longer exceptional. (e.g. above: D_6, not $\mathfrak{S}_6$).

Records with NFS

Current record for GNFS: RSA-768 (2010).

Current record for SNFS: 1024 bits (2007).

NFS variants exist for the discrete logarithm problem.

- In finite fields of small characteristic and large degree.
- In finite fields of large characteristic and small degree.
- In “balanced” finite fields.
- Also for some classes of algebraic curves or large genus.