MPRI – Cours 2.12.2

F. Morain

Lecture III: Integer factorization

2013/06/27

The slides are available on http://www.lix.polytechnique.fr/Labo/Francois.Morain/MPRI/2012

I. Basics

Kraitchik (1920): find x s.t. $x^2 \equiv 1 \bmod N$, $x \neq \pm 1$.

Ex. For $N = 143$, there are 4 solutions ± 1, ± 12 and $\gcd(12-1, 143) = 11$.

A general scheme

Step 0: build a prime basis $\mathcal{B} = \{p_1, p_2, \ldots, p_k\}$.

Step 1: find a lot of relations $(R_i)_{i \in I}$: $R_i = \prod_{j=1}^{k} p_j^{a_{i,j}} \equiv 1 \bmod N$

Step 2: find $I' \subset I$ s.t.

$$\prod_{i \in I'} R_i = x^2$$

over $\mathbb{Z}$, which is equivalent to

$$\forall j, \sum_{i \in I'} a_{i,j} \equiv 0 \bmod 2,$$

which is a classical linear algebra problem.

Step 3: x is a squareroot of 1 and with probability $\geq 1/2$, $\gcd(x-1, N)$ is non-trivial.

II. Naive methods

A very naive one:

0. Build $\mathcal{B} = \{p_1 = 2, 3, \ldots, p_k\}$.

1. Generate k random relations

$$p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \bmod N$$

and hope to factor the residue to get:

$$p_1^{f_1} p_2^{f_2} \cdots p_k^{f_k} \bmod N$$

from which

$$p_1^{e_1-f_1} p_2^{e_2-f_2} \cdots p_k^{e_k-f_k} \equiv 1 \bmod N.$$

Store the $(e_i - f_i) \bmod 2$ in the matrix $\mathcal{M}$.

2. Find dependancies relations of $\mathcal{M}$ and deduce solutions of $x^2 \equiv 1 \bmod N$.

Hypothesis:

$$x(e) = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \bmod N$$

is a random integer in $[1..N[$.

A numerical example

Let $N = 143$, $\mathcal{B} = \{2, 3, 5\}$. We compute:

$$(R_1) : 2^3 \times 3 \times 5^4 \equiv 2^7 \bmod N,$$
$$(R_2) : 2^3 \times 3^3 \times 5^4 \equiv 2^3 \bmod N,$$
$$(R_3) : 3^3 \times 5^4 \equiv 1 \bmod N.$$

Combining (R_1) and (R_2), we get:

$$(2^{-2} \times 3^2 \times 5^4)^2 \equiv 1 \bmod N$$

or $12^2 \equiv 1 \bmod N$ and $\gcd(12 - 1, N) = 11$.

De Bruijn's function

Define
$$\psi(x, y) = \#\{z \leq x, z \text{ is } y-\text{smooth}\}.$$

Thm. (Candfield, Erdős, Pomerance) $\forall\, \varepsilon > 0$, uniformly in $y \geq (\log x)^{1+\varepsilon}$, as $x \to \infty$

$$\psi(x, y) = \frac{x}{u^{u(1+o(1))}}$$

with $u = \log x / \log y$.

Prop. Let
$$L(x) = \exp\left(\sqrt{\log x \log\log x}\right).$$

For all real $\alpha > 0$, $\beta > 0$, as $x \to \infty$

$$\frac{\psi(x^\alpha, L(x)^\beta)}{x^\alpha} = L(x)^{-\frac{\alpha}{2\beta}+o(1)}.$$

Analysis

Prop. The cost of the naive algorithm is $O(L^{2+o(1)})$.
Proof.
Proba($x(e)$ is p_k-smooth)$= \frac{\psi(N, p_k)}{N} \Rightarrow$ we need $k\frac{N}{\psi(N,p_k)}$ relations.
Using trial division, testing p_k-smoothness costs k divisions.
Linear algebra costs $O(k^r)$ with $2 \leq r \leq 3$ (see later).
Total cost is:
$$O\left(k^2 \frac{N}{\psi(N, p_k)}\right) + O(k^r).$$

Put $k = L(N)^b$, from which $p_k \approx k \log k = O(L(N)^{b+o(1)})$. Cost is now:

$$O(L^{2b} L^{1/(2b)}) + O(L^{rb}) = O(L^{\max(2b+1/(2b), rb)}).$$

$2b + 1/(2b)$ is minimal for $b = 1/2$ and has value 2, which is larger than rb for all r. $\square$

III. Quadratic sieve

Pb. The above methods are not practical, since factoring the relations is too costly. Can we build residues of size N^α for $\alpha < 1$?

CFRAC: (Morrison and Brillhart) use the continued fraction expansion of $\sqrt{N}$, leads to residues of size $N^{1/2}$; first real-life algorithm, factored F_7 in 1970.

Schroeppel's linear sieve: relations $F(a, b) = (\lfloor\sqrt{N}\rfloor + a)(\lfloor\sqrt{N}\rfloor + b) - N$ for small a and b satisfy

$$F(a, b) \equiv (\lfloor\sqrt{N}\rfloor + a)(\lfloor\sqrt{N}\rfloor + b) \bmod N$$

and $N = \lfloor\sqrt{N}\rfloor^2 + R$, $R = O(\sqrt{N})$. All numbers have size $O(\sqrt{N})$. Moreover, if $p \mid F(a, b)$, then $p \mid F(a + p, b)$, etc.

Pomerance's quadratic sieve:

Use $a = b$

$$\left(a + \lfloor\sqrt{N}\rfloor\right)^2 \equiv \left(a + \lfloor\sqrt{N}\rfloor\right)^2 - N \approx 2a\sqrt{N}.$$

$$p \mid F(a) \iff \left(a + \lfloor\sqrt{N}\rfloor\right)^2 \equiv N \bmod p$$

implies N is a square modulo p and $p \mid F(a) \Leftrightarrow a \equiv a_-$ or $a \equiv a_+ \bmod p$.

Prop. The cost is $O(L(N)^{r/\sqrt{4(r-1)}})$.

Proof. Precomputing all roots of $F(a) \bmod p$ costs L^b.
The cost of sieving over $|a| \leq L^c$ is

$$\sum_{p \leq L^b} \frac{2L^c}{p} = L^{c+o(1)}.$$

The number of L^b-smooth values of $F(a)$ in the interval is $L^{c-1/(4b)} \Rightarrow$ take $c = b + 1/(4b)$ and optimize $L^{\max(b, b+1/(4b), rb)}$ which yields $b = 1/\sqrt{4(r-1)}$. □

Large primes

Idea: suppose we end up with

$$x(e)^2 = (\prod p) C$$

for some $p_k < C(e) < p_k^2$. Then we know that $C(e)$ is prime. We can keep the relation and hope for another

$$x(e')^2 = (\prod p) C$$

so that $(x(e)x(e')/C)^2$ is factored over $\mathcal{B}$. Works due to the birthday paradox. Use hashing to store C's.

More than one prime: filtering (highly technical to implement).

Real sieving in QS

Never factor residues, but test

$$\mathcal{R}(a) = \log|F(a)| - \sum_{\substack{p^e \mid F(a) \\ p \in \mathcal{B}}} \log p^e < \log p_k$$

and replace $\log|F(a)|$ by $\log|2a\sqrt{N}|$. In practice, fits in a `char`; use integer approximations; ignore small primes.

Large primes: relax $\mathcal{R}(a) < 2\log p_k$, say.

MPQS: (Montgomery, 1985) use families of quadratic polynomials
$\Rightarrow$ massive computations become possible: email (A. K. Lenstra & M. S. Manasse, 1990), INTERNET (RSA-129).

Rem. a lot more tricks exist (SIQS, etc.).

IV. Linear algebra

Fundamental property: combination matrices are sparse, since $\Omega(N) \leq \log_2 N$.

N	size	#coeffs $\neq 0$ per relation
RSA-100	$50,000 \times 50,000$	
RSA-110	$80,000 \times 80,000$	
RSA-120	$252,222 \times 245,810$ $(89,304 \times 89,088)$	
RSA-129	$569,466 \times 524,338$ $(188,614 \times 188,160)$	47
RSA-130	$3,504,823 \times 3,516,502$	39
RSA-140	$4,671,181 \times 4,704,451$	32
RSA-155	$6,699,191 \times 6,711,336$	62
$6,353-$	$19,591,108 \times 19,590,832$	229
RSA-768	$192,796,550 \times 192,795,550$	144

Gauß

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Computations:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ \color{blue}{1} & \color{blue}{0} & \color{blue}{0} & \color{blue}{1} \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \color{blue}{0} & \color{blue}{1} & \color{blue}{0} & \color{blue}{0} \\ x & x & x & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ 1 & 0 & 1 & \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

Finally:

$$L_1 + L_3 = 0, L_1 + L_2 + L_4 = 0$$

Advanced linear algebra

Rem. The companion matrix can be merged into A.

Rem. additings rows use XOR's on `unsigned long` (in C). Still in $O(k^3)$ but with a very small constant.

Structured Gaussian elimination: use a sparse encoding of M and perform elimination so as to slow the fill-in down as much as possible.

Going further: Lanczos and Wiedemann benefit from sparse encoding, and cost $O(k^{2+\varepsilon})$. Many subtleties.

Biggest open problem: how to distribute this phase in a clean an efficient way? Currently the bottleneck of this kind of algorithms.

Wiedemann's algorithm in a nutshell

M is $n \times n$

f^M: minimal polynomial of M, and/or that of $\{M^i\}_{i=0}^{\infty}$.

$f^{M,b}$: minimal polynomial of $\{M^i b\}_{i=0}^{\infty}$; of course $f^{M,b} \mid f^M$.

If u is any row vector, the minimal polynomial of $\{uM^i b\}_{i=0}^{\infty}$ is $f_u^{M,b}$ and $f_u^{M,b} \mid f^{M,b}$.

Rem. for random b, $f^{M,b} = f^M$ with high probability.

Idea: compute $f_u^{M,b}$ using $\{uM^i b\}_{i=0}^{2n-1}$ and use Berlekamp-Massey in time $O(n^2)$ (or faster $O(\mathsf{M}(n) \log n)$).

Cost: $2n$ applications Mb (black-box operation). If M has $n^{1+\varepsilon}$ non-zero coeffs, then this is $O(n^{2+\varepsilon})$.

Proofs: Kaltofen, etc.

Application to factoring

Trick: find minimal polynomial of z for random z. Then (probably)

$$P_M(X) = X + p_2 X^2 + \cdots + p_r X^r$$

and $P_M(M)(z) = 0 = M(z + p_2 Mz + \cdots + p_r M^{r-1} z)$, so that we probably have an element of the kernel.

Distributed version: cut M into slices, evaluate Mz this way.

V. Some hints on NFS

- Z. I. Borevitch and I. R. Chafarevitch.
  Théorie des nombres. Gauthiers-Villars, Paris, 1967.
- I. N. Stewart and D. O. Tall.
  Algebraic number theory.
  Chapman and Hall, London, New-York, 2nd edition, 1987.
- M. Pohst and H. Zassenhaus.
  Algorithmic algebraic number theory.
  Cambridge Univ. Press, 1989.
- H. Cohen.
  A course in algorithmic algebraic number theory, volume 138 of *Graduate Texts in Mathematics*.
  Springer–Verlag, 1996. Third printing.
- H. Cohen.
  Advanced topics in computational number theory, volume 193 of *Graduate Texts in Mathematics*.
  Springer-Verlag, 2000.

A) Factorization in (euclidean) quadratic fields

We consider the case where $\mathbb{Q}(\sqrt{d})$ is euclidean.

Thm. Let d be such that $\mathbb{Q}(\sqrt{d})$ is euclidean and p be a rational prime.

(a) If $\left(\frac{d}{p}\right) = -1$, p is irreducible in $\mathcal{O}_K$ and p is unramified.

(b) If $\left(\frac{d}{p}\right) = 1$, $p = u\pi_p\pi_p'$ with $u \in \mathcal{U}$, $\pi_p = x - y\sqrt{d}$ and $\pi_p' = x + y\sqrt{d}$ are two irreducible non associate factors in $\mathcal{O}_K$; p splits.

(c) If $\left(\frac{d}{p}\right) = 0$, $p = u(x+y\sqrt{d})^2$ where $x+y\sqrt{d}$ is irreducible in $\mathcal{O}_K$ and $u \in \mathcal{U}$; p is ramified.

Rem. For small p's, any trivial algorithm will work.

A numerical example: $\mathbb{Q}(\sqrt{6})$

Fundamental unit: $\varepsilon = 5 + 2\sqrt{6}$.

$$2 = -(2+\sqrt{6})(2-\sqrt{6}) = (5-2\sqrt{6})(2+\sqrt{6})^2 = \varepsilon^{-1}(2+\sqrt{6})^2.$$

Let's factor $\xi = 1010 + 490\sqrt{6}$. We first have

$$\mathrm{N}(\xi) = 1010^2 - 6\cdot 490^2 = -420500 = -2^2\cdot 5^3\cdot 29^2,$$

$5 = -(1+\sqrt{6})(1-\sqrt{6})$, $29 = -(5+3\sqrt{6})(5-3\sqrt{6})$; therefore

$$\xi = u(2+\sqrt{6})^{\alpha}(1+\sqrt{6})^{\gamma_1}(1-\sqrt{6})^{\delta_1}(5+3\sqrt{6})^{\gamma_2}(5-3\sqrt{6})^{\delta_2}.$$

$$\begin{array}{ll}
\alpha = 2, & \xi_1 = \frac{\xi}{(2+\sqrt{6})^2} = -415 + 215\sqrt{6}\\
\gamma_1 = 1, & \xi_2 = \frac{\xi_1}{1+\sqrt{6}} = 341 - 126\sqrt{6}\\
\delta_1 = 2, & \xi_3 = \frac{\xi_2}{(1-\sqrt{6})^2} = 35 - 8\sqrt{6}\\
\gamma_2 = 2, & \xi_4 = \frac{\xi_3}{(5+3\sqrt{6})^2} = 5 - 2\sqrt{6}\\
\gamma_3 = 0, & u = \xi_4 = \varepsilon^{-1}
\end{array}$$

$$\xi = \varepsilon^{-1}(2+\sqrt{6})^2(1+\sqrt{6})(1-\sqrt{6})^2(5+3\sqrt{6})^2.$$

B) NFS: basic idea

Pollard's idea: let $f(X) \in \mathbb{Z}[X]$ and m s.t.

$$f(m) \equiv 0 \bmod N.$$

Let θ be a root of f in $\mathbb{C}$ and $K = \mathbb{Q}[X]/(f(X)) = \mathbb{Q}(\theta)$.
To simplify things: $\mathcal{O}_\mathbf{K}$ is supposed to be $\mathbb{Z}[\theta]$ and euclidean. Let

$$\begin{array}{cccc}\phi: & \mathbb{Z}[\theta] & \rightarrow & \mathbb{Z}/N\mathbb{Z}\\ & \theta & \mapsto & m \bmod N.\end{array}$$

ϕ is a ring homomorphism.
Look for algebraic integers of the form $a - b\theta$ s.t.

$$a - b\theta = \prod_{\pi\in\mathcal{B}_K} \pi^{v_\pi(a-b\theta)}$$

where $v_\pi(a-b\theta) \in \mathbb{Z}$ and

$$a - bm = \prod_{p\in\mathcal{B}} p^{w_p(a-bm)}$$

with $\mathcal{B}$ a prime basis and $w_p(a-bm) \in \mathbb{Z}$.

NFS: basic idea (cont'd)

We then look for $\mathcal{A}$ s.t.

$$\prod_{(a,b)\in\mathcal{A}} (a - b\theta)$$

is a square in $\mathcal{O}_K$ and at the same time

$$\prod_{(a,b)\in\mathcal{A}} (a - bm)$$

is a square in $\mathbb{Z}$. Then

$$\prod_{(a,b)\in\mathcal{A}} (a - bm) = Z^2, \quad \prod_{(a,b)\in\mathcal{A}} (a - b\theta) = (A - B\theta)^2.$$

Applying ϕ, we get:

$$\phi((A - B\theta)^2) \equiv (A - Bm)^2 \equiv Z^2 \bmod N$$

and $\gcd(A - Bm \pm Z, N)$ might factor N.

Numerical example

Let's factor $N = 5^8 - 6 = 390619 = m^2 - 6$ (surprise!) with $m = 5^4 = 625$, hence we will work in $\mathbb{Q}(\theta) = \mathbb{Q}[X]/(f(X))$ with $f(X) = X^2 - 6$ and $\theta = \sqrt{6}$.

Rational basis: $\mathcal{B} = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$.

Algebraic basis: $\mathcal{B}_K$ given as

p	c_p	π, π'
2	0	$2 + \theta = \pi_2$
3	0	$3 + \theta = \pi_3$
5	± 1	$1 + \theta = \pi_5, 1 - \theta = \pi_5'$
19	± 5	$5 + \theta = \pi_{19}, 5 - \theta = \pi_{19}'$
23	± 11	$1 + 2\theta = \pi_{23}, 1 - 2\theta = \pi_{23}'$
29	± 8	$5 + 3\theta = \pi_{29}, 5 - 3\theta = \pi_{29}'$

with c_p s.t. $f(c_p) \equiv 0 \bmod p$. All these obtained via factoring of $\mathrm{N}(a - b\theta)$ for small a's and b's.

Free relations: $2 = \varepsilon^{-1}(2 + \theta)^2$, or $5 = -\pi_5\pi_5'$.

Results for $|a| \leq 60$, $1 \leq b \leq 30$

rel	a	b	$N(a - b\theta)$	$a - b\theta$	$a - bm$
L_1	2	0	2^2	$\varepsilon^{-1} \cdot \pi_2^2$	2
L_2	3	0	3^2	$\varepsilon^{-1} \cdot \pi_3^2$	3
L_3	5	0	5^2	$-\pi_5 \cdot \pi_5'$	5
L_4	19	0	19^2	$\pi_{19} \cdot \pi_{19}'$	19
L_5	23	0	23^2	$-\pi_{23} \cdot \pi_{23}'$	23
L_6	29	0	29^2	$-\pi_{29} \cdot \pi_{29}'$	29
L_7	-21	1	$3 \cdot 5 \cdot 29$	$-\varepsilon^{-1} \cdot \pi_3 \cdot \pi_5 \cdot \pi_{29}$	$-2 \cdot 17 \cdot 19$
L_8	-12	1	$2 \cdot 3 \cdot 23$	$-\varepsilon^{-1} \cdot \pi_2 \cdot \pi_3 \cdot \pi_{23}$	$-7^2 \cdot 13$
L_9	-5	1	19	$-\pi_{19}$	$-2 \cdot 3^2 \cdot 5 \cdot 7$
L_{10}	-2	1	-2	$-\pi_2$	$-3 \cdot 11 \cdot 19$
L_{11}	0	1	$-2 \cdot 3$	$-\varepsilon^{-1} \cdot \pi_2 \cdot \pi_3$	-5^4
L_{12}	1	1	-5	π_5'	$-2^4 \cdot 3 \cdot 13$
L_{13}	4	1	$2 \cdot 5$	$\varepsilon^{-1} \cdot \pi_2 \cdot \pi_5$	$-3^3 \cdot 23$

rel	a	b	$N(a - b\theta)$	$a - b\theta$	$a - bm$
L_{14}	9	1	$3 \cdot 5^2$	$\varepsilon^{-1} \cdot \pi_3 \cdot \pi_5^2$	$-2^3 \cdot 7 \cdot 11$
L_{15}	16	1	$2 \cdot 5^3$	$-\pi_2 \cdot \pi_5'^3$	$-3 \cdot 7 \cdot 29$
L_{16}	-10	3	$2 \cdot 23$	$\pi_2 \cdot \pi_{23}'$	$-5 \cdot 13 \cdot 29$
L_{17}	5	3	-29	π_{29}'	$-2 \cdot 5 \cdot 11 \cdot 17$
L_{18}	13	3	$5 \cdot 23$	$\pi_5' \cdot \pi_{23}'$	$-2 \cdot 7^2 \cdot 19$
L_{19}	1	4	$-5 \cdot 19$	$-\pi_5 \cdot \pi_{19}'$	$-3 \cdot 7^2 \cdot 17$
L_{20}	25	4	23^2	$\pi_{23}'^2$	$-3^2 \cdot 5^2 \cdot 11$
L_{21}	-11	5	-29	$\varepsilon \cdot \pi_{29}'$	$-2^6 \cdot 7^2$
L_{22}	-7	9	$-19 \cdot 23$	$\pi_{19} \cdot \pi_{23}'$	$-2^9 \cdot 11$
L_{23}	-27	11	3	$-\varepsilon \cdot \pi_3$	$-2 \cdot 7 \cdot 17 \cdot 29$
L_{24}	-2	11	$-2 \cdot 19^2$	$-\pi_2 \cdot \pi_{19}'^2$	$-13 \cdot 23^2$
L_{25}	33	13	$3 \cdot 5^2$	$\varepsilon^{-1} \cdot \pi_3 \cdot \pi_5'^2$	$-2^2 \cdot 7 \cdot 17^2$

	u_0	ε	π_2	π_3	π_5	π'_5	π_{19}	π'_{19}	π_{23}	π'_{23}	π_{29}	π'_{29}	2	3	5	7	11	13	17	19	23	29
L_1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
L_2	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L_3	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
L_4	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
L_5	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0
L_6	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1
L_7	0	1	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	1	0	0
L_8	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
L_9	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0
L_{10}	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
L_{11}	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L_{12}	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
L_{13}	1	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0

$\ldots$

$L_3 \cdot L_6 \cdot L_7 \cdot L_{10} \cdot L_{15} \cdot L_{17} \cdot L_{25}$ yields

$$\phi\left((5+2\theta)^{-1}(2+\theta)(3+\theta)(1+\theta)(1-\theta)^3(5+3\theta)(5-3\theta)\right)^2$$

$$\equiv \left(2^2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 17^2 \cdot 19 \cdot 29\right)^2 \pmod N$$

and

$$2^2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 17^2 \cdot 19 \cdot 29 \equiv 148603 \pmod N,$$

gives $242016^2 \equiv 148603^2 \bmod N$ and $\gcd(242016 - 148603, N) = 1$,
$L_1 \cdot L_2 \cdot L_3 \cdot L_4 \cdot L_9 \cdot L_{10} \cdot L_{14} \cdot L_{15} \cdot L_{19} \cdot L_{23}$ leads to

$$\phi\left((5+2\theta)^{-1}(2+\theta)^2(3+\theta)^2(1+\theta)^2(1-\theta)^2(5+\theta)(5-\theta)\right)^2$$

$$\equiv \left(2^3 \cdot 3^3 \cdot 5 \cdot 7^3 \cdot 11 \cdot 17 \cdot 19 \cdot 29\right)^2 \pmod N$$

or $61179^2 \equiv 81314^2 \bmod N$, $\gcd(61179 - 81314, N) = 4027$.

Working without units

We can use factorization modulo units. We will end up with relations

$$\mathrm{N}(A + B\sqrt{6}) = \varepsilon^m = 1$$

and hope to get a square.

If we don't know ε, we can try to extract a squareroot of

$$\eta = A + B\sqrt{6}$$

using brute force: $\eta = \xi^2 = (x + y\sqrt{6})^2$, or:

$$\begin{cases} x^2 - 6y^2 & = & \pm 1 \\ x^2 + 6y^2 & = & A \end{cases}$$

which readily gives $x^2 = (A \pm 1)/2$ which is easily solved over $\mathbb{Z}$.

Over a general number field, computing units is in general difficult, and some workaround has been found.

C) a bit of complexity

SNFS: $N = r^e \pm s$ with r and s small.
Choose an extension of degree d. Put $k = \lceil e/d \rceil$, $m = r^k$ and $c = sr^{kd-e}$ s.t. $m^d \equiv c \bmod N$. Put $f(X) = X^d - c$ and use $K = \mathbb{Q}(X)/(f(X)) = \mathbb{Q}(\theta)$.

$$\mathrm{N}(a - b\theta) = b^d f(a/b).$$

For $0 \leq \alpha \leq 1$ and $\beta > 0$, we define $L_N[\alpha, \beta] = \exp((\beta + o(1))(\log n)^\alpha (\log\log n)^{1-\alpha})$, sometimes simplified to $L_N[\alpha]$.

Thm. The computing time is $L_N[1/2, \sqrt{2/d}]$.

Thm. Let d vary with N as:

$$d = K(\log N)^\varepsilon (\log\log N)^{1-\varepsilon}.$$

Optimal values are $\varepsilon = 1/3$, $K = (2/3)^{-1/3}$

$$L_N[1/3, \exp(2(2/3)^{2/3})].$$

GNFS

For a general N, we need $f(X)$ representing N and f is not sparse, nor "small".
Basic thing is to write N in base m for $m \approx N^{1/d}$ and

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

Conj. GNFS has cost $L_N[1/3, (64/9)^{1/3}]$ for optimal d as function of N.

Some problems:

- A lot of effort was put in searching for

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

  with $a_i \approx N^{1/(d+1)}$ and a_i "small" with many properties.
- Properties related to units and/or factorization solved using characters (Adleman). See LNM 1554 for details.
- As usual, linear algebra causes some trouble.

RSA-768 with NFS

- **Who?** Thorsten Kleinjung, Kazumaro Aoki, Jens Franke, Arjen K. Lenstra, Emmanuel Thomé, Joppe W. Bos, Pierrick Gaudry, Alexander Kruppa, Peter L. Montgomery, Dag Arne Osvik, Herman te Riele, Andrey Timofeev, and Paul Zimmermann.
- **Sieving:** August 2007 til April 2009 (about 1500 AMD64 years), several countries/continents. 64 334 489 730 relations (38% INRIA, 30% EPFL, 15% NTT, 8% Bonn, 3.5% CWI, 5.5% others).
- **Linear algebra** (after filtering): $192\,796\,550 \times 192\,795\,550$ (total weight 27 797 115 920) using 155 core years in 119 calendar days (block Wiedemann in parallel).

Conclusions

- A broad view of integer factorization.
- Programs are now available (cado-nfs, GMP-ECM).
- discrete log algorithms as companions to integer factorization.