The number field sieve in the medium prime case

Frederik Vercauteren
ESAT/COSIC - K.U. Leuven

Magma Conference Berlin

Joint work with Antoine Joux, Reynald Lercier, Nigel Smart

Number field sieve

- Index calculus algorithm used for factoring, then DLOGs
- Let $q = p^n$ with p prime, then complexity expressed by

$$L_q(\alpha, c) = \exp((c + o(1))(\log q)^\alpha (\log \log q)^{1-\alpha})$$

- Large p: number fi eld sieve with running time

$$L_q(1/3, (64/9)^{1/3})$$

  as long as $\log p > n^{2+\varepsilon}$
- Small p: function fi eld sieve with running time

$$L_q(1/3, (32/9)^{1/3})$$

  as long as $p \leq n^{o(\sqrt{n})}$
- In the gap, i.e. $\log p < n^{2+\varepsilon}$ and $p > n^{o(\sqrt{n})}$, have to resort to Adleman - DeMarrais with complexity $L_q(1/2)$

Classical number field sieve for $\mathbb{F}_p$: setup

- To compute discrete logarithms in $\mathbb{F}_p^*$
- Two number fields $K_1 = \mathbb{Q}$ and $K_2 = \mathbb{Q}[X]/(f(X))$ with:
  - The degree of f is $d \simeq 3^{1/3} \left(\frac{\log p}{\log \log p}\right)^{1/3}$
  - Exists $m \in \mathbb{Z}$ with $f(m) \equiv 0 \bmod p$, i.e. ring homomorphism

$$\phi_2 : \mathcal{O}_2 \to \mathbb{F}_p$$

  - E.g. f can be obtained by base $m \simeq p^{1/d}$ expansion of p
- Choose two factor bases $\mathcal{F}_1$ and $\mathcal{F}_2$
  - $\mathcal{F}_1$: integer primes $p < B$ for some bound B
  - $\mathcal{F}_2$: degree 1 prime ideals of norm $< B$

Classical number field sieve for $\mathbb{F}_p$: sieving

- Sieve over pairs of integers (a, b) with
  - $\gcd(a, b) = 1$ and $|a|, |b| < S$ for some bound S
  - $a - bm$ is B-smooth
  - $\mathrm{No}(a - \theta_2 b)$ is B-smooth with $f(\theta_2) = 0$ and

$$\mathrm{No}(a - \theta_2 b) = b^d f(\frac{a}{b})$$

- Since $\mathrm{No}(a - \theta_2 b)$ is B-smooth the ideal $\langle a - \theta_2 b \rangle$ factors over $\mathcal{F}_2$ since only degree 1 (or index divisors) appear

$$\langle a - \theta_2 b \rangle = \prod_i \mathfrak{p}_i^{e_i}$$

Classical number field sieve for $\mathbb{F}_p$: relations

- (a, b) with $a - bm$ and $a - \theta_2 b$ B-smooth gives relation
- Need to get rid of ideals and work with elements only ...
- Simplicity: assume class number $h(K) = 1$ and computable unit group, then

$$a - \theta_2 b = \prod_{i=0}^{r} u_i^{\lambda_i} \prod_i \gamma_i^{e_i}$$

with $u_1, \dots, u_r$ fundamental units and $\mathfrak{p}_i = \langle \gamma_i \rangle$

- Finally, by using ϕ_2 from $\mathcal{O}_2$ to $\mathbb{F}_p^*$ obtain

$$a - bm \equiv \prod_j p_j^{e_j} \equiv \prod_{i=0}^{r} \phi_2(u_i)^{\lambda_i} \prod_i \phi_2(\gamma_i)^{e_i} \bmod p$$

Classical number field sieve for $\mathbb{F}_p$: relations

- Take logs of both sides, obtain relation between DLOGs

$$\sum_j e_j \log_g p_j \equiv \sum_{i=0}^{r} \lambda_i \log_g \phi_2(u_i) + \sum_i e_i \log_g \phi_2(\gamma_i) \bmod (p-1)$$

- Need to collect $\#\mathcal{F}_1 + \#\mathcal{F}_2 + d + \varepsilon$ relations
- Solve sparse linear system using Lanczos or Wiedemann
- Individual DLOGs: descent procedure (see more later)

Schirokauer's extension for $n > 1$

- Number field K_1 is chosen such that $\mathcal{O}_1/p\mathcal{O}_1 \cong \mathbb{F}_q$, so K_1 has degree at least n
- Number field K_2 is **extension** of K_1, i.e. $K_2 = K_1[X]/(f(X))$
- Collect pairs $(a, b) \in \mathcal{O}_1 \times \mathcal{O}_1$ with similar properties as before:
  - $a - bm$ is B-smooth where $m \in \mathcal{O}_1$ such that $f(m) \in p\mathcal{O}_1$
  - $a - \theta_2 b$ is B-smooth with $f(\theta_2) = 0$
- Leads to $L_q(1/3)$-algorithm for fixed n and $p \to \infty$
- Main disadvantage: not really practical (only $n = 2$ has been attempted by Weber)
- Choice of polynomial f depends on input DLOG problem

Basic variation $p = L_{p^n}(2/3, c)$: setup

- Finite fields $\mathbb{F}_{p^n}$ with $p = L_{p^n}(2/3, c)$ and c near $2 \cdot (1/3)^{1/3}$
- Choose polynomial f_1 of degree n
  - irreducible over $\mathbb{F}_p$
  - very small coefficients (e.g. use poly to define $\mathbb{F}_q$)
- Choose polynomial $\boxed{f_2 = f_1 + p}$
- $K_1 \simeq \mathbb{Q}[X]/(f_1(X)) \cong \mathbb{Q}[\theta_1]$ and $K_2 \cong \mathbb{Q}[X]/(f_2(X)) \cong \mathbb{Q}[\theta_2]$
- Note: $f_1 \equiv f_2 \bmod p$, so have compatible homomorphisms

$$\phi_i : \mathcal{O}_i \to \mathbb{F}_q, \quad \text{for } i = 1, 2 \text{ with } \phi_1(\theta_1) = \phi_2(\theta_2)$$

- No relative extensions necessary and f independent of input DLOG

Basic variation $p = L_{p^n}(2/3, c)$: sieving/linear algebra

- Factor bases $\mathcal{F}_1$ and $\mathcal{F}_2$ of degree 1 ideals of small norm
- Choose smoothness bound B and a sieve limit S
- Pairs (a, b) of coprime integers, $|a| \leq S$ and $|b| \leq S$

$$\mathrm{No}(a - b\theta_1) \text{ and } \mathrm{No}(a - b\theta_2) \quad B\text{-smooth}$$

- Add logarithmic maps to take into account $h(K_i) \neq 1$ and unit groups
- Obtain linear equation between "logarithms of ideals" in the smoothness bases
- Solve using SGE and Lanczos or Wiedemann

Basic variation $p = L_{p^n}(2/3, c)$: individual DLOG

- Recursive special $\mathfrak{q}$-descent procedure similar to $\mathbb{F}_p^*$
- Represent $\mathbb{F}_{p^n}$ as $\mathbb{F}_p[t]/(f_1(t))$
- Assume we want to compute $\log_t y$ with $y \in \mathbb{F}_{p^n}$
- Search for element $z = y^i t^j$ for some $i, j \in \mathbb{N}$ with
  1. lifting $z \in K_1$, norm factors into primes smaller than some bound $B_1 \in L_{p^n}(2/3, 1/3^{1/3})$,
  2. only degree one prime ideals in the factorisation of (z)
  3. E.g.: the norm of the lift of z should be squarefree
- Remark: probability of *squarefree smoothness* is about $6/\pi^2$ probability of smoothness

Basic variation $p = L_{p^n}(2/3, c)$: individual DLOG

- Factor principal ideal generated by z as

$$(z) = \prod_{p_i \in \mathcal{F}_1} \mathfrak{p}_i^{e_i} \prod_j \mathfrak{q}_j^{e_j}$$

- Ideals $\mathfrak{q}_j$ not contained in $\mathcal{F}_1$, so need to compute DLOGs
- For each $\mathfrak{q}_j$, perform special-$\mathfrak{q}_j$ descent:
  1. Sieve over pairs (a, b) such that $\mathfrak{q}_j | (a - b\theta_1)$ and

     $$\mathrm{No}(a - b\theta_1)/\mathrm{No}(\mathfrak{q}_j) \text{ and } \mathrm{No}(a - b\theta_2) \quad B_2\text{-smooth } B_2 < B_1$$

  2. Factor $(a - b\theta_1)$ and $(a - b\theta_2)$ to obtain new special q_j's
  3. Repeat until bound $B_k < B \Rightarrow$ DLOGs of all $\mathfrak{q}_j$ known
- Remark: special $\mathfrak{q}_j$ in both number fi elds K_1 and K_2

Optimisation I: Galois extensions

- p is inert in K_1, so isomorphism $\mathrm{Gal}(K_1/\mathbb{Q}) \simeq \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$
- Thus: K_1 has to be a cyclic number fi eld of degree n
- Partition factor base $\mathcal{F}_1$ in n parts $\mathcal{F}_{1,k}$ with $k = 1, \ldots, n$

$$(a - b\theta_1) = \prod_{k=1}^{n} \prod_{\mathfrak{p}_i \in \mathcal{F}_{1,1}} \psi^k(\mathfrak{p}_i)^{e_{i,k}}$$

  with $\mathrm{Gal}(K_1/\mathbb{Q}) = \langle \psi \rangle$
- Choose ψ such $\log_g \phi_1(\psi(\delta_i))) = p \log_g \phi_1(\delta_i)$ with $\mathfrak{p}_i = \langle \delta_i \rangle$
- Effectively divides factor base size by n

Optimisation II: choice of polynomials

Two possible optimisations:

- Poss I: Maximise automorphism group of K_1 and K_2 simultaneously
- Example: $p \equiv 2, 5 \bmod 9$, can take

$$f_1 = x^6 + x^3 + 1 \qquad f_2 = x^6 + (p+1)x^3 + 1$$

- K_1 is Galois and K_2 has non-trivial automorphism order 2
- Poss II: balance size of coeffi cients of f_1 and f_2
- Remark: better to adapt sieving region . . .

Optimisation III: individual logarithms

- Instead of factoring $\langle z \rangle$, first write z as
$$\frac{\sum a_i t^i}{\sum b_i t^i}$$
with a_i and b_i are of the order of $\sqrt{p}$.
- Use LLL to find short vector in lattice L
$$L = \begin{pmatrix} \mathbf{z} & \mathbf{tz} & \mathbf{t^2z} & \cdots & \mathbf{t^{n-1}z} & \mathbf{p} & \mathbf{pt} & \mathbf{pt^2} & \cdots & \mathbf{pt^{n-1}} \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$
- Expect LLL finds short vector of norm $\sqrt{p}$

Definition 120-digit challenge

- Adaptation of Joux & Lercier's implementation for $\mathbb{F}_p^*$
- Finite field $\mathbb{F}_{p^3}$ with $p = \lfloor 10^{39}\pi \rfloor + 2622$

$$p = 3141592653589793238462643383279502886819$$

- Group order $p^3 - 1$ has 110-bit factor l
- Definition of number fields K_1 and K_2 by

$$f_1(X) = X^3 + X^2 - 2X - 1 \quad \text{and} \quad f_2(X) = f_1(X) + p\,,$$

where we have $\mathbb{F}_{p^3} \simeq \mathbb{F}_p[t]/(f_1(t))$

Number fields K_1 and K_2

- $\mathbb{Q}[\theta_1]$ is a cubic cyclic number fi eld with Galois group

$$\mathrm{Aut}(\mathbb{Q}[\theta_1]) = \{\theta_1 \mapsto \theta_1, \theta_1 \mapsto \theta_1^2 - 2, \theta_1 \mapsto -\theta_1^2 - \theta_1 + 1\}$$

- K_1 has class number 1 and System of fundamental units

$$u_1 = \theta_1 + 1 \text{ and } u_2 = \theta_1^2 + \theta_1 - 1$$

- $\mathbb{Q}[\theta_2]$ has signature $(1, 1)$, so only need single Schirokauer logarithmic map λ

Factor bases and sieving

- Smoothness bases with 1 000 000 prime ideals
  - in the $\mathbb{Q}[\theta_1]$ side, we include 899 999 prime ideals, but only 300 000 are meaningful due to the Galois action,
  - in the $\mathbb{Q}[\theta_2]$ side, we include 700 000 prime ideals.
- Lattice sieving: only algebraic integers $a + b\theta_2$ divisible by prime ideal in $\mathbb{Q}[\theta_2]$
- Norms to be smoothed in $\mathbb{Q}[\theta_2]$ are 150 bit integers
- Norms in $\mathbb{Q}[\theta_1]$ are 110 bit integers
- Sieving took 12 days on a 1.15 GHz 16-processors HP AlphaServer GS1280

Linear algebra

- Compute the kernel of a $1\,163\,482 \times 793\,188$ matrix
- Coeffi cients mostly equal modulo ℓ to ± 1, $\pm p$ or $\pm p^2$
- SGE: $450\,246 \times 445\,097$ matrix with $44\,544\,016$ non null entries
- Lanczos's algorithm: about one week
- $h(K_1) = 1$, check DLOGs of generators of ideals in $\mathcal{F}_1$

$$
\begin{aligned}
(t^2 + t + 1)^{(p^3-1)/l} &= G^{29406688645015596112746712243217l}, \\
(t - 3)^{(p^3-1)/l} &= G^{36422456363509538073334012349071 9}, \\
(3\,t - 1)^{(p^3-1)/l} &= G^{46887658774739638067572350292825 7},
\end{aligned}
$$

where $G = g^{(p^3-1)/1159268202574177739715462155841484\,l}$ and $g = -2t + 1$.

Individual DLOGs

- Challenge $\gamma = \sum_{i=0}^{2}(\lfloor \pi \times p^{i+1} \rfloor \bmod p)t^i$
- Using Pollard-Rho, computed DLOG modulo $(p^3 - 1)/l$,

  388953891589015189758459229369411846775349910996122146045769727138614728691028247732 8.

- To obtain a complete result, we expressed

$$\gamma = \frac{-909879803559595293 47t^2 - 1144430082485221569 10t + 1544936643733412719 98}{949127644415707714 06t^2 - 1200555698097118619 65t - 8195961996444635256 7},$$

- Numerator and denominator are both smooth in $\mathbb{Q}[\theta_1]$
- Three level tree with 80 special-$\mathfrak{q}$ ideals
- Recovered DLOG modulo l, namely 110781190155780903592153105706975
- Each special-$\mathfrak{q}$ sieving took 10 minutes or a total of 14 hours

Variation I: smaller p

- Polynomial setup same as in basic case
- Main problem: sieving space is not large enough, due to larger n
- $\Rightarrow$ cannot collect enough relations
- Solution: sieve over elements of larger degree than 1

$$\sum_{i=0}^{t} a_i \theta_1^i \quad \text{and} \quad \sum_{i=0}^{t} a_i \theta_2^i$$

- Bound on norm: $(n+t)^{n+t} {B_a}^n {B_f}^t$ with
  - B_a is an upper bound on the absolute values of the a_i
  - B_f a similar bound on the coeffi cients of f_1 (resp. f_2)

Variation II: larger p

- p is too large to simply add to f_1, so need different polynomial construction
- Only requirement is: $f_1(x) \mid f_2(x) \bmod p$
- Idea: construct $f_2(x)$ of degree $> n$ with small coeffi cients such that $f_1(x) \nmid f_2(x)$ over $\mathbb{Q}$
- Choose constant W and construct $f_1(x) = f_0(x + W)$, coeffi cient at least W^n
- Use LLL to reduce the lattice

$$L = \begin{pmatrix} \mathbf{f_1(x)} & \mathbf{xf_1(x)} & \mathbf{x^2f_1(x)} & \cdots & \mathbf{x^{D-n}f_1(x)} & \mathbf{p} & \mathbf{px} & \mathbf{px^2} & \cdots & \mathbf{px^D} \end{pmatrix}$$

- Need vector with coeffi cients smaller than W^n so

$$2^{(D+1)/4} p^{n/(D+1)} \leq W^n$$

Complexity of variations

- p can be written as $L_q(l_p, c)$ with $1/3 < l_p < 2/3$

$$L_q(1/3, (128/9)^{1/3}) \simeq L_q(1/3, 2.423\ldots)$$

- p can be written as $L_q(2/3, c)$ for a constant c

$$L_q(1/3, 2c') \quad \text{with} \quad c' = \frac{4}{3}\left(\frac{3t}{4(t+1)}\right)^{1/3}$$

sieve over elements of degree t with $3c^3 t(t+1)^2 - 32 = 0$

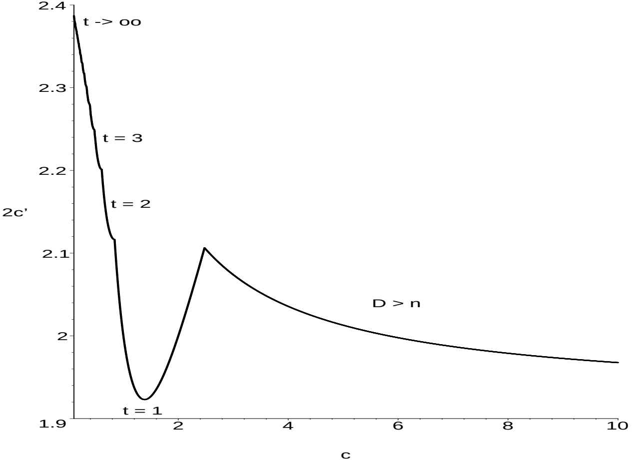
- p can be written as $L_q(2/3, c)$ for a constant c

$$L_q(1/3, 2c') \quad \text{with} \quad 9c'^3 - \frac{6}{c}c'^2 + \frac{1}{c^2}c' - 8 = 0$$

- p can be written as $L_q(l_p, c)$ with $l_p > 2/3$

$$L_q(1/3, (64/9)^{1/3}) \simeq L_q(1/3, 1.923\ldots)$$

JLSV NFS: complexity = $L_q(1/3, 2c')$



Conclusions

- New, simple and **practical** variations of NFS
- Can simply adapt existing implementations of NFS for $\mathbb{F}_p$
- More optimisations: large prime variation, multiple number fi elds, ...
- Combined with talk of Joux: obtain two families of algorithms such that DLOGs in $\mathbb{F}_{p^n}$ can be computed in $L_{p^n}(1/3)$ time