Polynomial Selection

Thorsten Kleinjung

École Polytechnique Fédérale de Lausanne

Contents

Brief summary of polynomial selection (no root sieve)

Motivation (lattice sieving, monic algebraic polynomial)

General case (reduction to monic algebraic polynomial)

Some results

Brief summary of polynomial selection

Given $N \in \mathbb{Z}$

Find co-prime polynomials $f, g \in \mathbb{Z}[x]$ with common zero modulo N

Degrees and coefficients as small as possible

Brief summary of polynomial selection

Given $N \in \mathbb{Z}$

Find co-prime polynomials $f, g \in \mathbb{Z}[x]$ with common zero modulo N

Degrees and coefficients as small as possible

Restriction to $\deg(f) = d$, $\deg(g) = 1$

Easy: coefficients of size $N^{\frac{1}{d+1}}$:

Choose $m = [N^{\frac{1}{d+1}}] + 1$, set $g = x - m$, $f = \sum_{i=0}^{d} a_i x^i$ where

$N = \sum_{i=0}^{d} a_i m^i$ is the base-m-expansion of N.

Skewness:

Change sieving area from $-A \leq a \leq A, 0 < b \leq A$ to $-A\sqrt{s} \leq a \leq A\sqrt{s}, 0 < b \leq \frac{A}{\sqrt{s}}$ for some s (skewness)

$\Rightarrow$ want to minimise $\max(|a_i| \cdot s^{i-\frac{d}{2}})$ $\qquad (f = \sum_{i=0}^{d} a_i x^i)$

Skewness:

Change sieving area from $-A \leq a \leq A, 0 < b \leq A$ into $-A\sqrt{s} \leq a \leq A\sqrt{s}, 0 < b \leq \frac{A}{\sqrt{s}}$ for some s (skewness)

$\Rightarrow$ want to minimise $\max(|a_i| \cdot s^{i-\frac{d}{2}})$ $\qquad (f = \sum_{i=0}^{d} a_i x^i)$

Choose a_d smaller than $N^{\frac{1}{d+1}}$, choose m near $\left(\frac{N}{a_d}\right)^{\frac{1}{d}}$

$\Rightarrow |a_{d-1}|$ roughly of size a_d, small enough

Remaining coefficients of size $\left(\frac{N}{a_d}\right)^{\frac{1}{d}}$

ok for a_0, a_1 (perhaps also for a_2)

Coefficients $a_{d-2}, \ldots, a_3(, a_2)$ too big $\qquad$ biggest problem a_{d-2}

Motivation

Lattice sieving for 768 bit numbers:

e.g.: factor base bounds $1.1 \cdot 10^9$ (for f), $2 \cdot 10^8$ (for g)

$\Rightarrow$ ca. 67 million factor base elements

gnfs-lasieveI16e needs 20 byte per factor base element:

- prime ideal $(p, x - r)$: 4 byte for p and 4 byte for r
- two vectors in special q lattice: $2 \cdot 4$ byte
- current location in special q lattice: 4 byte

could reduce this:

- use 1 byte for storing differences of $p \Rightarrow 17$ byte
- handle larger p in a different way $\Rightarrow 15$ or 16 byte

How can we reduce this further?

If skewness were equal to size of sieving area:

form of sieving area: $-A \leq a \leq A, b = 1$ (one line)

If skewness were equal to size of sieving area:

form of sieving area: $-A \leq a \leq A, b = 1$ (one line)

Storage requirements for lattice siever (12 byte per factor base element):

- prime ideal $(p, x - r)$: 4 byte for p and 4 byte for r
- current location in special q lattice: 4 byte

We can

- recalculate r from last location in special q lattice $\Rightarrow$ 8 byte
- store 1 byte differences of primes $\Rightarrow$ 5 byte

Reduced storage for factor base from 1GB (or 1.3GB) to 350MB

How can we find such polynomials?

Polynomials with large skewness

Example: 768-bit integer N, size of sieving area $\approx 2^{64} \approx$ skewness,

$f = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$, $g = lx - m$

$N = a_4m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$

Polynomials with large skewness

Example: 768-bit integer N, size of sieving area $\approx 2^{64} \approx$ skewness,

$f = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$, $g = lx - m$

$N = a_4m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$

coefficient	a_4	a_3	a_2	a_1	a_0	l	m
bit size	0	64	128	192	256	128	192

$\Rightarrow$ values of polynomials: ca. 256 bit and 192 bit

seems too be slightly worse than current degree 6 polynomials

Polynomials with large skewness

Example: 768-bit integer N, size of sieving area $\approx 2^{64} \approx$ skewness,

$f = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$, $g = lx - m$

$N = a_4m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$

coefficient	a_4	a_3	a_2	a_1	a_0	l	m
bit size	0	64	128	192	256	128	192

$\Rightarrow$ values of polynomials: ca. 256 bit and 192 bit

seems too be slightly worse than current degree 6 polynomials

Check: $64 + 128 + 192 + 256 + 128 + 192 - 64 - 64 = 768 + 64$

$\Rightarrow$ expect to find 2^{64} such polynomial pairs

How can we find such polynomial pairs (with cost $\ll 2^{64}$)?

$$f = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0,\ g = lx - m$$

$$N = m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$$

$f = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0,\ g = lx - m$

$N = m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$

translation $\Rightarrow$ can assume $a_3 \in \{0, 1, 2, 3\}$

$f = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0,\ g = lx - m$

$N = m^4 + a_3lm^3 + a_2l^2m^2 + a_1l^3m + a_0l^4$

translation $\Rightarrow$ can assume $a_3 \in \{0, 1, 2, 3\}$

Restrict to $a_3 = 0$, assume $l \ll \frac{m}{2^{64}}$:

$f = x^4 + a_2x^2 + a_1x + a_0,\ g = lx - m$:

$N = m^4 + a_2l^2m^2 + a_1l^3m + a_0l^4 = m^4 + l^2R \qquad a_2 \approx \frac{R}{m^2}$

New problem: to find l, m such that $l^2 | N - m^4$ and $\frac{|N-m^4|}{l^2m^2}$ is small

General problem: N, d and bound B given, find l, m such that $l^2 | N - m^d$ and $\frac{|N-m^d|}{l^2 m^{d-2}} < B$

General problem: N, d and bound B given, find l, m such that $l^2 | N - m^d$ and $\frac{|N-m^d|}{l^2 m^{d-2}} < B$

Set $m_0 = \sqrt[d]{N}$, $m = m_0 + i$, $i \in [-M, M]$

$\Rightarrow |N - m^d| \overset{\approx}{<} dMm_0^{d-1}$

want i, l such that $l^2 | N - (m_0 + i)^d$ and $\frac{dMm_0}{l^2} < B$

General problem: N, d and bound B given, find l, m such that $l^2 | N - m^d$ and $\frac{|N-m^d|}{l^2 m^{d-2}} < B$

Set $m_0 = \sqrt[d]{N}$, $m = m_0 + i$, $i \in [-M, M]$

$\Rightarrow |N - m^d| \overset{\approx}{<} dMm_0^{d-1}$

want i, l such that $l^2 | N - (m_0 + i)^d$ and $\frac{dMm_0}{l^2} < B$

Set $l = p_1 p_2, \quad p_i \in \mathcal{P}$ primes, $\mathcal{P} = [P, 2P]$

1. generate pairs (p, i) such that $p^2 | N - (m_0 + i)^d$
2. sort pairs w. r. t. second entry
3. for each collision, i. e., pairs (p_1, i), (p_2, i) with $p_1 \neq p_2$: output $l = p_1 p_2$, $m = m_0 + i$

result: $|a_{d-2}| \approx \frac{|N-m^d|}{l^2 m^{d-2}} \overset{\approx}{<} \frac{dM}{P^4} m_0$

Analysis

$m_0 = \sqrt[d]{N}$, $m = m_0 + i$, $i \in [-M, M]$

$l = p_1 p_2, \quad p_i \in \mathcal{P}$ primes, $\mathcal{P} = [P, 2P]$

number of pairs $\approx \frac{M}{P \log P}$, number of collisions $\approx \frac{M}{4P^2 (\log P)^2}$

Analysis

$$m_0 = \sqrt[d]{N},\ m = m_0 + i,\ i \in [-M, M]$$

$$l = p_1 p_2, \quad p_i \in \mathcal{P} \text{ primes}, \mathcal{P} = [P, 2P]$$

number of pairs $\approx \frac{M}{P \log P}$, number of collisions $\approx \frac{M}{4P^2 (\log P)^2}$

cost $O(\frac{M \log M}{P \log P} + \frac{P}{\log P})$

result: $|a_{d-2}| \overset{\approx}{<} \frac{dM}{P^4} m_0$

Analysis

$m_0 = \sqrt[d]{N}$, $m = m_0 + i$, $i \in [-M, M]$

$l = p_1 p_2, \quad p_i \in \mathcal{P}$ primes, $\mathcal{P} = [P, 2P]$

number of pairs $\approx \frac{M}{P \log P}$, number of collisions $\approx \frac{M}{4P^2 (\log P)^2}$

cost $O(\frac{M \log M}{P \log P} + \frac{P}{\log P})$

result: $|a_{d-2}| \stackrel{\approx}{<} \frac{dM}{P^4} m_0$

for 768 bit example choose $M = 2^{90}$, $P = 2^{39}$:
≈ 1 collision, $\frac{dM}{P^4} m_0 \approx 2^{128}$, cost 2^{46} pairs

Analysis

$m_0 = \sqrt[d]{N}$, $m = m_0 + i$, $i \in [-M, M]$

$l = p_1 p_2, \quad p_i \in \mathcal{P}$ primes, $\mathcal{P} = [P, 2P]$

number of pairs $\approx \frac{M}{P \log P}$, number of collisions $\approx \frac{M}{4P^2(\log P)^2}$

cost $O(\frac{M \log M}{P \log P} + \frac{P}{\log P})$

result: $|a_{d-2}| \overset{\approx}{<} \frac{dM}{P^4} m_0$

for 768 bit example choose $M = 2^{90}$, $P = 2^{39}$:
≈ 1 collision, $\frac{dM}{P^4} m_0 \approx 2^{128}$, cost 2^{46} pairs

choosing $M = P^2$:

cost per collision $O(P(\log P)^2)$, result $|a_{d-2}| \overset{\approx}{<} \frac{d}{P^2} m_0$

Asymptotic considerations

degree $d = \left(\frac{3\log N}{\log\log N}\right)^{\frac{1}{3}}$, sieving area $\approx L(\frac{1}{3}, \sqrt[3]{\frac{64}{9}}) \approx$ skewness

product of coefficient ranges of algebraic polynomial $= L(1, \frac{7}{8})$

$\Rightarrow$ cannot find such polynomial pairs

Remark: polynomial pairs of degree d and $d-1$ would be ok

General situation

$$N = a_d m^d + a_{d-1} l m^{d-1} + l^2 R$$

Find l, m such that and $\frac{|R|}{m^{d-2}}$ $(\approx |a_{d-2}|)$ is sufficiently small.

General situation

$$N = a_d m^d + a_{d-1} l m^{d-1} + l^2 R$$

Find l, m such that and $\frac{|R|}{m^{d-2}}$ ($\approx |a_{d-2}|$) is sufficiently small.

Reduction to $a_d = 1$, $a_{d-1} = 0$ (translation $x \mapsto x - \frac{a_{d-1}}{d a_d}$):

$$d^d a_d^{d-1} N = (d a_d m + a_{d-1} l)^d + l^2 \left(d^d a_d^{d-1} R - (d a_d m)^{d-2} \cdot \binom{d}{2} \cdot a_{d-1}^2 - \dots \right)$$

General situation

$$N = a_d m^d + a_{d-1} l m^{d-1} + l^2 R$$

Find l, m such that and $\frac{|R|}{m^{d-2}}$ $(\approx |a_{d-2}|)$ is sufficiently small.

Reduction to $a_d = 1$, $a_{d-1} = 0$ (translation $x \mapsto x - \frac{a_{d-1}}{d a_d}$):

$$d^d a_d^{d-1} N = (d a_d m + a_{d-1} l)^d + l^2 \left(d^d a_d^{d-1} R - (d a_d m)^{d-2} \cdot \binom{d}{2} \cdot a_{d-1}^2 - \dots \right)$$

or

$$\tilde{N} = \tilde{m}^d + l^2 \tilde{R} \qquad \text{where } \tilde{N} = d^d a_d^{d-1} N, \tilde{m} = d a_d m + a_{d-1} l$$

General situation

$$N = a_d m^d + a_{d-1} l m^{d-1} + l^2 R$$

Find l, m such that and $\frac{|R|}{m^{d-2}}$ ($\approx |a_{d-2}|$) is sufficiently small.

Reduction to $a_d = 1$, $a_{d-1} = 0$ (translation $x \mapsto x - \frac{a_{d-1}}{da_d}$):

$$d^d a_d^{d-1} N = (da_d m + a_{d-1} l)^d + l^2 \left(d^d a_d^{d-1} R - (da_d m)^{d-2} \cdot \binom{d}{2} \cdot a_{d-1}^2 - \dots \right)$$

or

$$\tilde{N} = \tilde{m}^d + l^2 \tilde{R} \qquad \text{where } \tilde{N} = d^d a_d^{d-1} N, \tilde{m} = da_d m + a_{d-1} l$$

1. find l, $\tilde{m}$ as above
2. $\tilde{m} = da_d m + a_{d-1} l$: find m, $0 \le a_{d-1} < da_d$ $\quad (\gcd(l, da_d) = 1)$

Result: $|a_{d-2}| \approx \frac{|\tilde{R}|}{d^2 a_d \tilde{m}^{d-2}} \overset{\approx}{<} \frac{dM\tilde{m}_0}{d^2 a_d P^4} \approx \frac{M}{P^4} m_0$

Some tricks

Replace $l = p_1 p_2$ by $l = cp$, $c \in \mathcal{C}$, $p \in \mathcal{P}$

e. g.: $\mathcal{C} = [P_1, P_2]$, $\mathcal{P} = \{p \in [P_2, P_3] | p \text{ prime}\}$ for some $P_1 < P_2 < P_3$

1. generate pairs (c, i), $c \in \mathcal{C}$
2. generate pairs (p, j), $p \in \mathcal{P}$
3. search for collisions between c-pairs and p-pairs, and for collisions within p-pairs

many alternative approaches, e. g.:

- arbitrary $\mathcal{C}$, $\mathcal{P}$, remove multiples of primes of $\mathcal{P}$ from $\mathcal{C}$
- $\mathcal{C} = \{c \in [P_1, P_2] | p | c \Rightarrow p \equiv 1 \pmod 4\}$,
  $\mathcal{P} = \{c \in [P_1, P_2] | p | c \Rightarrow p \equiv 3 \pmod 4\}$
- ...

Special q

Choose q, $0 \leq s < q^2$ such that $q^2 | N - (m_0 + s)^d$

Search for l' with $l'^2 | N - (m_0 + s + iq^2)^d$ as above and set $l = l'q$

analysis remains the same, only l is increased by q

Advantage: Initialisation costs drop, since expensive root calculation of $N - x^d$ modulo p (resp. c) can be used for many q

Even better: can do inversion modulo p^2 for many q simultaneously $\Rightarrow$ cost drops to a few modular additions + multiplications per generated pair

Some results

number	sieving time	pol. sel. time	improvement
RSA512	≈ 0.25 a	4 d, 4 d, 4 d	0.84, 0.8, 0.84
RSA576	≈ 2.5 a	15 d	0.87
RSA640	≈ 20 a	10 d	0.77 (?)

improvement = time for new pol. pair / time for old pol. pair

RSA512: comparison with best polynomial pair found by old method

RSA576, RSA640: comparison with polynomial pairs used in factorisation