Nonlinear polynomials for NFS factorisation

Nicholas Coxon

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- f_1 and f_2 produce many smooth values in the sieve stage.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- f_1 and f_2 produce many smooth values in the sieve stage.

Very roughly speaking, smoothness probabilities are correlated with

- Coefficient size,
- Number of real roots,
- Roots modulo small primes.

See **[Brent, Montgomery & Murphy $\approx$1997]** for more details.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- f_1 and f_2 produce many smooth values in the sieve stage.

Very roughly speaking, smoothness probabilities are correlated with

$$\left.\begin{array}{l}\cdot\ \text{Coefficient size,}\\ \cdot\ \text{Number of real roots,}\end{array}\right\}\ \text{Size properties}$$

- Roots modulo small primes.

See **[Brent, Montgomery & Murphy $\approx$1997]** for more details.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- f_1 and f_2 produce many smooth values in the sieve stage.

Very roughly speaking, smoothness probabilities are correlated with

- Coefficient size,
- ~~Number of real roots~~,

Size properties

- Roots modulo small primes.

See **[Brent, Montgomery & Murphy ≈1997]** for more details.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- f_1 and f_2 produce many smooth values in the sieve stage.

Quantifying size properties:

If $f = \sum_{i=0}^{d} a_i x^i y^{d-i}$ has degree d, define its *s-skewed* 2*-norm* to be

$$\|f\|_{2,s} = \left(s^{-d} \cdot \sum_{i=0}^{d} \left| a_i s^i \right| \right)^{1/2} \quad \text{for } s > 0.$$

We want $|a_d|$ to be small and $|a_{d-1}|, |a_{d-2}|, \ldots, |a_0|$ to grow at most geometrically with ratio s. The *skew* of f is the s that minimises $\|f\|_{2,s}$.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- $\|f_1\|_{2,s}$ and $\|f_2\|_{2,s}$ are small for some large $s > 0$.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- $\|f_1\|_{2,s}$ and $\|f_2\|_{2,s}$ are small for some large $s > 0$.

Quantifying root properties:

For homogeneous $f \in \mathbb{Z}[x, y]$, define

$$\alpha(f, B) = \sum_{p \leq B} \left(1 - \sigma(f_i, p)\frac{p}{p+1}\right) \frac{\log p}{p-1},$$

where $\sigma(f, p) := \#\{(r_1 : r_2) \in \mathbb{P}^1(\mathbb{F}_p) \mid f(r_1, r_2) \equiv 0 \pmod{p}\}$.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- $\|f_1\|_{2,s}$ and $\|f_2\|_{2,s}$ are small for some large $s > 0$.

Quantifying root properties:

For homogeneous $f \in \mathbb{Z}[x, y]$, define

$$\alpha(f, B) = \sum_{p \leq B} \left(1 - \sigma(f_i, p)\frac{p}{p+1}\right) \frac{\log p}{p-1}.$$

[Brent & Murphy 1997]: $f(a, b)$ behaves like $f(a, b)\cdot e^{\alpha(f,B)}$ w.r.t. B-smoothness.

The problem

Given an integer N that we want to factor with the number field sieve, find two **homogeneous** polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 + \deg f_2 = \delta$, where $\delta = \delta(N)$ ($\in \{6, 7\}$ in practice),
- f_1 and f_2 are distinct and irreducible,
- $\exists\, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$ such that $f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}$,
- $\|f_1\|_{2,s}$ and $\|f_2\|_{2,s}$ are small for some large $s > 0$.
- $\alpha(f_1, B)$ and $\alpha(f_2, B)$ are small (-ve), where B is the smoothness bound.

Room for improvement

[Crandall and Pomerance 2001]:

- In the sieve stage, smooth values $f_1(a, b) \cdot f_2(a, b)$ are found.
- As these values are a product of two integers, they are more likely to be smooth than a random integer of the same size that is not necessarily a product of two integers.
- This effect is maximised when f_1 and f_2 produce values that are of the same magnitude.

Current best methods generate polynomial with $\deg f_1 \geq 5$ and $\deg f_2 = 1$. Thus, they produce values that are *not* of the same magnitude.

Better smoothness probabilities could be obtained by using two nonlinear polynomials with $\deg f_1 \approx \deg f_2$.

The resultant bound

[Montgomery?]: Suppose that $f_1, f_2 \in \mathbb{Z}[x, y]$ are non-constant coprime polynomials with a common root modulo N. Then

$$N \leq \|f_1\|_{2,s}^{\deg f_2} \cdot \|f_2\|_{2,s}^{\deg f_1} \quad \text{for all } s > 0.$$

- Obtained by bounding $\left|\mathrm{Res}(f_1, f_2)\right|$ above and below.
- Small degrees used in NFS imply there must be large coefficients.
- Current best methods give f_1 and f_2 with $\|f_1\|_{2,s}^{\deg f_2} \|f_2\|_{2,s}^{\deg f_1} = O(N)$.
- **[Prest & Zimmermann 2010]** give heuristic evidence that for each N there exist pairs of NFS polynomials such that

$$\deg f_1 = \deg f_2 = d \quad \text{and} \quad \|f_i\|_{2,s} = O\left(N^{1/(2d)}\right) \quad \text{for } i = 1, 2.$$

This talk

Given an integer N that we want to factor with the number field sieve, find two homogeneous polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ such that

- $\deg f_1 = \deg f_2 = d$, where $d = \delta(N)/2$;
- f_1 and f_2 are distinct and irreducible;
- f_1 and f_2 have a common root modulo N; and
- $\|f_1\|_{2,s} \cdot \|f_2\|_{2,s} = O(N^{1/d})$ for some large $s > 0$.
- $\alpha(f_1, B)$ and $\alpha(f_2, B)$ are small.

Part i: montgomery-type algorithms

Lattices

A *lattice* is a subgroup $L \subset \mathbb{R}^n$ of the form

$$L = \boldsymbol{b}_1\mathbb{Z} + \ldots + \boldsymbol{b}_k\mathbb{Z},$$

where $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k \in \mathbb{R}^n$ are linearly independent.

Key invariants:

- k — the *dimension* of L
- $\det L := \left(\det(\boldsymbol{b}_i \cdot \boldsymbol{b}_j)_{1 \leq i,j \leq k}\right)^{1/2}$ — the *determinant* of L

[Lenstra, Lenstra & Lovász 1982]: Given $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k \in \mathbb{Z}^n$, there exists an algorithm (now called *LLL-reduction*) that can be used to compute $\boldsymbol{a}_1, \boldsymbol{a}_2 \in L$ such that

$$\|\boldsymbol{a}_1\|_2 \leq 2^{(k-1)/4} \det(L)^{1/k} \quad \text{and} \quad \|\boldsymbol{a}_2\|_2 \leq 2^{k/4} \det(L)^{1/(k-1)}$$

in time polynomial in k, n and $\max_{1 \leq i \leq k} \log \|\boldsymbol{b}_i\|_2$

Geometric progressions

[Montgomery 1993] introduced a method for constructing NFS polynomials with small coefficients which relies on construction of modular geometric progressions.

Definition. A vector $[c_0, c_1, \ldots, c_{\ell-1}] \in \mathbb{Z}^\ell$ is called a geometric progression (GP) of length ℓ and ratio r modulo N if

$$c_i \equiv c_0 r^i \pmod{N} \quad \text{and} \quad \gcd(c_i, N) = 1 \quad \text{for } i = 0, \ldots, \ell - 1.$$

Length d+1 GPs are special:

If $[c_0, c_1, \ldots, c_d]$ is a length $d + 1$ GP with ratio m_1/m_2 modulo N, then a vector $(a_0, a_1, \ldots, a_d) \in \mathbb{Z}^{d+1}$ satisfies

$$\sum_{j=0}^{d} a_j c_j \equiv 0 \pmod{N}$$

iff the polynomial $f = \sum_{i=0}^{d} a_i x^i y^{d-i}$ satisfies $f(m_1, m_2) \equiv 0 \pmod{N}$.

GPs $\to$ Polynomials

Suppose we have $1 \le k \le d-1$ linearly independent length $d+1$ GPs

$$\boldsymbol{c}_1 = [c_{1,0}, \ldots, c_{1,d}], \boldsymbol{c}_2 = [c_{2,0}, \ldots, c_{2,d}], \ldots, \boldsymbol{c}_k = [c_{k,0}, \ldots, c_{k,d}]$$

that have the same ratio m_1/m_2 modulo N.

Then any vector $(a_0, \ldots, a_d) \in \mathbb{Z}^{d+1}$ satisfying

$$\sum_{j=0}^{d} a_j c_{i,j} = 0 \quad \text{for } i = 1, \ldots, k$$

gives rise to a polynomial $f = \sum_{i=0}^{d} a_i x^i y^{d-i}$ with $f(m_1, m_2) \equiv 0 \pmod{N}$.

Moreover, if $s^{-d/2}(a_0, a_1 s \ldots, a_d s^d)$ is a short vector, then $\|f\|_{2,s}$ is small.

GPs $\to$ Polynomials

The set of all such vectors,

$$L := \left\{ s^{-d/2} \left(a_0, a_1 s, \ldots, a_d s^d \right) \mid \left(a_0, a_1, \ldots, a_d \right) \in \mathbb{Z}^{d+1} \right.$$

$$\left. \text{and} \quad \sum_{j=0}^{d} a_j c_{i,j} = 0 \text{ for } i = 1, \ldots, k \right\},$$

is a $(d - k + 1)$–dimensional lattice with determinant

$$\det L \leq N^{1-k} \cdot \prod_{i=1}^{k} s^{-d/2} \left\| \left(c_{i,0} s^d, c_{i,1} s^{d-1}, \ldots, c_{i,d} \right) \right\|_2 .$$

If the product on the right is sufficiently small, then we can use LLL-reduction to find two polynomials with common root (m_1, m_2) and norms of size $O\left(N^{1/(2d)}\right)$.

In particular, if $k = d - 1$, then we require the product to be $O\left(N^{(d-1)^2/d}\right)$.

Polynomials $\to$ GPs

Montgomery showed that the converse holds for $k = d - 1$:

If there exists two degree d polynomials $f_1, f_2 \in \mathbb{Z}[x, y]$ with common root (m_1, m_2) modulo N and norms of size $O\big(N^{1/(2d)}\big)$ (+ some other conditions), then there exists $d - 1$ linearly independent length $d + 1$ geometric progressions $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{d-1}$ with ratio m_1/m_2 modulo N and

$$\prod_{i=1}^{d-1} s^{-d/2} \left\|(c_{i,0}s^d, c_{i,1}s^{d-1}, \ldots, c_{i,d})\right\|_2 = O\left(N^{(d-1)^2/d}\right).$$

$k = 1$: constructions

[Montgomery] + [Williams] + [Prest & Zimmermann] + [Koo, Jo & Kwon] + [C] construct a single GP as follows:

$$\left[am_2^{d-1}, am_2^{d-2}m_1, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2} \right],$$

where $a, v \in \mathbb{Z}$, $am_1^d \equiv vN \pmod{m_2}$ and $m_1 \approx (vN/a)^{1/d}$.

[Prest & Zimmermann]: By imposing conditions on the size of the parameters, we can obtain degree d polynomials f_1 and f_2 such that

$$\|f_i\|_{2,s} = O\left(N^{(1/d)(d^2-2d+2)/(d^2-d+2)}\right) \quad \text{for } i = 1, 2,$$

where $s = O\big(N^{2/(d(d^2-d+2))}\big)$.

Need to use sub-optimal s in order to avoid LLL returning polynomials of degree $< d$ (which are all multiples of $m_2x - m_1y$).

[Koo, Jo & Kwon]: Very easy to generate many parameters that give this bound.

$k = 1$: constructions

[Montgomery] + [Williams] + [Prest & Zimmermann] + [Koo, Jo & Kwon] + [C] construct a single GP as follows:

$$\left[am_2^{d-1}, am_2^{d-2}m_1, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}\right],$$

where $a, v \in \mathbb{Z}$, $am_1^d \equiv vN \pmod{m_2}$ and $m_1 \approx (vN/a)^{1/d}$.

[Prest & Zimmermann]:

d	$\|f_i\|_{2,s}$	s	Optimal?
2	$O\left(N^{1/4}\right)$	$O\left(N^{1/4}\right)$	Yes
3	$O\left(N^{5/24}\right)$	$O\left(N^{1/12}\right)$	No
4	$O\left(N^{5/28}\right)$	$O\left(N^{1/28}\right)$	No

Need to use sub-optimal s in order to avoid LLL returning polynomials of degree $< d$ (which are all multiples of $m_2x - m_1y$).

[Koo, Jo & Kwon]: Very easy to generate many parameters that give this bound.

$k = 1$: example

Let N be the 91-digit composite number

$$\mathrm{c}91 = 4567176039894108704358752160655628192034927306\backslash$$
$$9698283977390743466289883271554752228437933 93.$$

The following pair was found by using parameters that satisfy the size requirements that give the bound on the previous slide:

$$f_1 = 21545x^3 + 3349054x^2 - 10356871479051937193x + 1263295294354066431546642250$$

$$f_2 = 1356640x^3 + 210882368x^2 - 65211867386909760999 4x - 1197206898045490909233342893 9$$

The product $\|f_1\|_{2,s} \cdot \|f_2\|_{2,s}$ is approximately $N^{0.368}$ for $s \approx N^{1/12}$.

$k = 2$: construction

[Koo, Jo & Kwon]+**[C]** construct two GPs as follows:

$$\left[\ \overbrace{am_2^{d-1}, \mathrlap{\underbrace{\phantom{am_2^{d-2}m_1, am_2^{d-3}m_1^2, \dots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, \frac{m_1\left(am_1^d - vN\right)}{m_2^2}}}_{\boldsymbol{c}_2=}} am_2^{d-2}m_1, am_2^{d-3}m_1^2, \dots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}}^{\boldsymbol{c}_1=}, \frac{m_1\left(am_1^d - vN\right)}{m_2^2}\ \right]$$

where $a, v \in \mathbb{Z}$, $am_1^d \equiv vN \pmod{m_2^2}$ and $m_1 \approx (vN/a)^{1/d}$.

By imposing conditions on the size of the parameters, we can obtain degree d polynomials f_1 and f_2 such that

$$\|f_i\|_{2,s} = O\left(N^{(1/d)(d^2-4d+6)/(d^2-3d+6)}\right) \quad \text{for } i = 1, 2,$$

where $s = O\left(N^{2/(d(d^2-3d+6))}\right)$.

$k=2$: construction

[Koo, Jo & Kwon]+[C] construct two GPs as follows:

$$\left[\rlap{$\displaystyle\overbrace{\phantom{am_2^{d-1}, am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}}}^{\boldsymbol{c}_1=}$} am_2^{d-1}, \underbrace{am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, \frac{m_1\left(am_1^d - vN\right)}{m_2^2}}_{\boldsymbol{c}_2=} \right]$$

where $a, v \in \mathbb{Z}$, $am_1^d \equiv vN \pmod{m_2^2}$ and $m_1 \approx (vN/a)^{1/d}$.

d	$\lVert f_i \rVert_{2,s}$	s	Optimal?
3	$O\left(N^{1/6}\right)$	$O\left(N^{1/9}\right)$	Yes
4	$O\left(N^{3/20}\right)$	$O\left(N^{1/20}\right)$	No

$k = 2$: construction

[Koo, Jo & Kwon]+**[C]** construct two GPs as follows:

$$\left[\overbrace{am_2^{d-1}, \underbrace{am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}}_{}}^{\boldsymbol{c}_1=}\underbrace{, \frac{m_1\left(am_1^d - vN\right)}{m_2^2}}_{\boldsymbol{c}_2=}\right]$$

where $a, v \in \mathbb{Z}$, $am_1^d \equiv vN \pmod{m_2^2}$ and $m_1 \approx (vN/a)^{1/d}$.

d	$\lVert f_i \rVert_{2,s}$	s	Optimal?
3	$O\left(N^{1/6}\right)$	$O\left(N^{1/9}\right)$	Yes
4	$O\left(N^{3/20}\right)$	$O\left(N^{1/20}\right)$	No

It is much harder to generate parameters that give this bound: we are required to find a parameters such that $am_1^d \equiv vN \pmod{m_2^2}$ and

$$\left| m_1 - \left|\frac{vN}{a}\right|^{1/d} \right| = \begin{cases} O\left(m_2^{3/2}\right) & \text{for } d = 3, \\ O\left(m_2^{5/4}\right) & \text{for } d = 4; \end{cases} \quad m_2 = \begin{cases} O\left(N^{2/9}\right) & \text{for } d = 3, \\ O\left(N^{1/5}\right) & \text{for } d = 4. \end{cases}$$

$k = 2$: construction

[Koo, Jo & Kwon]+[C] construct two GPs as follows:

$$\left[\, \overbrace{am_2^{d-1},\ \rlap{$\displaystyle\underbrace{\phantom{am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, \frac{am_1^{d+1} - (vm_1 + um_2)\,N}{m_2^2}}}_{\boldsymbol{c}_2=}$} am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}}^{\boldsymbol{c}_1=}, \frac{am_1^{d+1} - (vm_1 + um_2)\,N}{m_2^2} \,\right]$$

where $a, v \in \mathbb{Z}$, $am_1^{d+1} \equiv (vm_1 + um_2)\,N \pmod{m_2^2}$ and $m_1 \approx (vN/a)^{1/d}$.

d	$\lVert f_i \rVert_{2,s}$	s	Optimal?
3	$O\left(N^{1/6}\right)$	$O\left(N^{1/9}\right)$	Yes
4	$O\left(N^{3/20}\right)$	$O\left(N^{1/20}\right)$	No

It is much harder to generate parameters that give this bound: we are required to find a parameters such that $am_1^d \equiv vN \pmod{m_2^2}$ and

$$\left| m_1 - \left|\frac{vN}{a}\right|^{1/d} \right| = \begin{cases} O\left(m_2^{3/2}\right) & \text{for } d = 3, \\ O\left(m_2^{5/4}\right) & \text{for } d = 4; \end{cases} \quad m_2 = \begin{cases} O\left(N^{2/9}\right) & \text{for } d = 3, \\ O\left(N^{1/5}\right) & \text{for } d = 4. \end{cases}$$

Part ii: impractical polynomial generation

Current best methods involve extensive searches, are guided by experience, helped by luck, and profit from patience.

Kleinjung et al. 2010

Notation

For any ideal proper $\mathfrak{a} \subset \mathbb{Z}[x, y]$ and nonzero $f \in \mathbb{Z}[x, y]$, define

$$\sigma(f, \mathfrak{a}) = \begin{cases} 1 & \text{if } f \in \mathfrak{a}, \\ 0 & \text{if } f \notin \mathfrak{a}. \end{cases}$$

For prime p, define $\mathfrak{p}_{p,r} = (p, x - ry)$ for $r \in \mathbb{F}_p$ and $\mathfrak{p}_{p,\infty} = (p, y)$.

Note. For homogeneous $f \in \mathbb{Z}[x, y]$, we have

$$\alpha(f, B) = \sum_{\substack{\mathfrak{p}_{p,r} \\ p \leq B}} \left(1 - \sigma(f, \mathfrak{p}_{p,r})p\right) \frac{\log p}{p^2 - 1}.$$

Lemma

Let $\mathcal{M} = \mathcal{M}(N, m_2, m_1; d, s, C)$ be the set of all $f \in \mathbb{Z}[x, y]$ such that

- f is a non-constant and irreducible;
- f is homogeneous of degree $\leq d$;
- $f \in (N, m_2 x - m_1 y)$; and
- $\|f\|_{2,s} \leq (CN)^{1/2d}$.

Lemma. If $f_1, f_2 \in \mathcal{M}$ satisfy

$$\sum_{\substack{\mathfrak{p}_{p,r} \not\supseteq (N) \\ p \leq B}} \sigma(f_1, \mathfrak{p}_{p,r}) \sigma(f_2, \mathfrak{p}_{p,r}) \log p > \log C$$

for some $B > 0$, then $f_1 = \pm f_2$.

Proved by using a result of Jouanolou (1990) + some trickery to sharpen the lower bound on $|\mathrm{Res}(f_1, f_2)|$ used in the resultant bound.

Lemma

Let $\mathcal{M} = \mathcal{M}(N, m_2, m_1; d, s, C)$ be the set of all $f \in \mathbb{Z}[x, y]$ such that

- f is a non-constant and irreducible;
- f is homogeneous of degree $\leq d$;
- $f \in (N, m_2x - m_1y)$; and
- $\|f\|_{2,s} \leq (CN)^{1/2d}$.

Lemma. If $f_1, f_2 \in \mathcal{M}$ satisfy

$$\sum_{\substack{\mathfrak{p}_{p,r} \not\supseteq (N) \\ p \leq B}} \sigma(f_1, \mathfrak{p}_{p,r})\sigma(f_2, \mathfrak{p}_{p,r}) \log p > \log C$$

for some $B > 0$, then $f_1 = \pm f_2$.

$\Rightarrow$ If $\mathfrak{p}_{p_1,r_1}, \ldots, \mathfrak{p}_{p_n,r_n} \not\supseteq (N)$ are distinct and $\prod_{i=1}^{n} p_i > C$, then the vectors

$$f \cdot (1 - \sigma(f, \mathfrak{p}_{p_1,r_1}), 1 - \sigma(f, \mathfrak{p}_{p_2,r_2}), \ldots, 1 - \sigma(f, \mathfrak{p}_{p_n,r_n})) \quad \text{for } f \in \mathcal{M}/\sim,$$

have a nonzero minimum “distance”.

A combinatorial bound

Given distinct $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \not\supseteq (N)$, positive real weights $\beta_1, \ldots, \beta_n$ and a real number $\ell \geq 1$, there are at most 2ℓ polynomials $f \in \mathcal{M}$ such that

$$\sum_{i=1}^{n} \sigma(f, \mathfrak{p}_i)\beta_i \geq \sqrt{\left(\left(1 - \frac{1}{\ell}\right) \log C + \frac{1}{\ell} \sum_{i=1}^{n} \log p_i\right) \sum_{i=1}^{n} \frac{\beta_i^2}{\log p_i}}.$$

Obtained by applying a generic coding bound of **[Guruswami 2000]**.

A combinatorial bound

Given distinct $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \not\supseteq (N)$, positive real weights $\beta_1, \ldots, \beta_n$ and a real number $\ell \geq 1$, there are at most 2ℓ polynomials $f \in \mathcal{M}$ such that

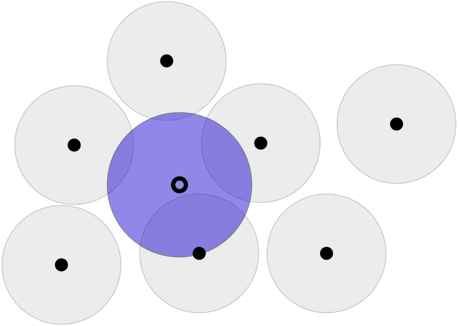
$$\sum_{i=1}^{n} \sigma(f, \mathfrak{p}_i)\beta_i \geq \sqrt{\left(\left(1 - \frac{1}{\ell}\right) \log C + \frac{1}{\ell} \sum_{i=1}^{n} \log p_i\right) \sum_{i=1}^{n} \frac{\beta_i^2}{\log p_i}}.$$

Example. $\#\Big\{f \in (N, m_2x - m_1y) \mid \deg f \leq 3, \|f\|_{2,s} \leq (CN)^{1/6}, \underbrace{\bar{\alpha}(f, B)}_{\text{Ignores roots at } \infty} \leq -2\Big\}$

$C^{1/6}$	$B = 100$	$B = 1000$	$B = 10000$
1	860	83463	7299206
2	1484	130046	10499454
3	2581	193086	14121084
4	5434	294311	18696869
5	38188	496011	24973925
6	-	1127183	34414014
7	-	-	50578542
8	-	-	85275302
9	-	-	215937570

List decoding

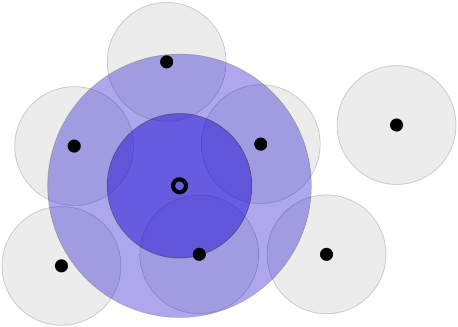
Nearest codeword/maximum likelihood: Find the codeword closest to the received word.



List decoding

Nearest codeword/maximum likelihood: Find the codeword closest to the received word.

List decoding: Find *all* codewords within a certain distance.

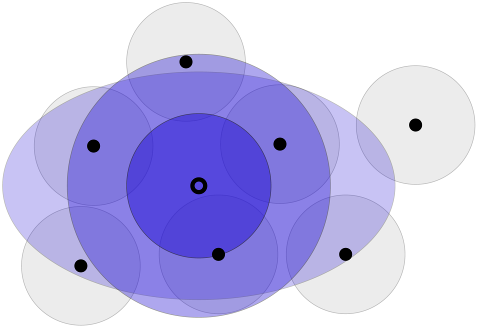


List decoding

Nearest codeword/maximum likelihood: Find the codeword closest to the received word.

List decoding: Find *all* codewords within a certain distance.

Weighted list decoding: Find all codewords within a certain *weighted* distance.



For polynomials selection, use weighted list decoding to correct the natural bias towards roots modulo large primes.

Analogues

[Cheng, Wan 2007] showed that a list decoding algorithm for Reed–Solomon codes can be used to find smooth polynomials in $\mathbb{F}_q[x]$.

[Boneh 2002] used a list decoding algorithm for CRT codes to find smooth integers.

This result generalises to number fields, giving an algorithm which finds smooth principal ideals.

Boneh used similar ideas to give an algorithm which finds smooth polynomial values.

Algorithm

Using ideas from the framework of **[Guruswami, Sahai & Sudan 2000]** + a simplification, gives the following algorithm:

INPUT: $\mathcal{M}$, distinct ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \not\supseteq (N)$ and integer weights $z_1, \ldots, z_n > 0$.

OUTPUT: All $f \in \mathcal{M}$ such that $\sum_{i=1}^{n} \sigma(f, \mathfrak{p}_i) z_i \log p_i$ is "sufficiently large".

1. Construct a homogeneous polynomial $h \in (N, m_2x - m_1y)^{z_0} \cap \left(\bigcap_{i=1}^{n} \mathfrak{p}_i^{z_i}\right)$ such that $\deg h$ and $\|h\|_{2,s}$ are small, where z_0 is chosen to exploit the fact that $\mathcal{M} \subset (N, m_2x - m_1y)$.
   a. Construct a basis for the lattice generated by the homogeneous polynomials degree ℓ polynomials in $(N, m_2x - m_1y)^{z_0} \cap \left(\bigcap_{i=1}^{n} \mathfrak{p}_i^{z_i}\right)$.
   b. Scale it appropriately, then LLL-reduce.
2. Factor h over $\mathbb{Q}$ and return all factors in $\mathcal{M}$.

Here, "sufficiently large" means $\underbrace{(CN)^{\deg h/(2d)} \cdot \|h\|_{2,s}^{d}}_{|\mathrm{Res}(f,h)| \le} < \underbrace{N^{z_0} \cdot \prod_{i=1}^{n} p_i^{\sigma(f,\mathfrak{p}_i)\, z_i}}_{\text{Divides } |\mathrm{Res}(f,\, h)|}$.

Algorithm

Using ideas from the framework of **[Guruswami, Sahai & Sudan 2000]** + a simplification, gives the following algorithm:

INPUT: $\mathcal{M}$, distinct ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \not\supseteq (N)$ and integer weights $z_1, \ldots, z_n > 0$.

OUTPUT: All $f \in \mathcal{M}$ such that $\sum_{i=1}^{n} \sigma(f, \mathfrak{p}_i) z_i \log p_i$ is "sufficiently large".

1. Construct a homogeneous polynomial $h \in (N, m_2x - m_1y)^{z_0} \cap \left(\bigcap_{i=1}^{n} \mathfrak{p}_i^{z_i}\right)$ such that $\deg h$ and $\|h\|_{2,s}$ are small, where z_0 is chosen to exploit the fact that $\mathcal{M} \subset (N, m_2x - m_1y)$.
   a. Construct a basis for the lattice generated by the homogeneous polynomials degree ℓ polynomials in $(N, m_2x - m_1y)^{z_0} \cap \left(\bigcap_{i=1}^{n} \mathfrak{p}_i^{z_i}\right)$.
   b. Scale it appropriately, then LLL-reduce.
2. Factor h over $\mathbb{Q}$ and return all factors in $\mathcal{M}$.

Here, "sufficiently large" means $\underbrace{(CN)^{\deg h/(2d)} \cdot \|h\|_{2,s}^{d}}_{|\mathrm{Res}(f,h)| \leq} < \underbrace{N^{z_0} \cdot \prod_{i=1}^{n} p_i^{\sigma(f,\mathfrak{p}_i)\, z_i}}_{\text{Divides } |\mathrm{Res}(f,\, h)|}$.

Theorem

Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \not\supseteq (N)$ be distinct, $z_1, \ldots, z_n$ be positive real weights and $\varepsilon > 0$. Then there exists an algorithm that returns all polynomials $f \in \mathcal{M}$ such that

$$\sum_{i=1}^{n} \sigma_i(f, \mathfrak{p}_i) z_i \log p_i > \sqrt{\log\left(2^{\frac{d^2}{2}} C\right)\left(\sum_{i=1}^{n} z_i^2 \log p_i + \varepsilon z_{\max}^2\right)}.$$

The algorithm runs in time $\operatorname{poly}\bigl(n, d, \log s, \log C, \sum_{i=1}^{n} \log p_i, \log N, 1/\varepsilon\bigr)$.

The problem

Example. $N = 10^{170} + 7$

$$\left\{ f \in (N, m_2 x - m_1 y) \mid \deg f \leq 3, \|f\|_{2,s} \leq (CN)^{1/6} \text{ and } \bar{\alpha}(f, B) \leq -2 \right\}$$

B	$\#\mathfrak{p}$	$C^{1/6}$	dim
10	17	1.78	809
20	77	1.99	1143
30	129	2.06	1274
40	197	2.12	1400
50	328	2.20	1579
100	1060	2.41	2153
1000	76127	3.39	12412

Have to LLL-reduced a lattice with huge dimension for each $(N, m_2 x - m_1 y)$.

Algorithmic bounds

Each output of the algorithm is a factor of h, which has degree equal to ℓ

$\Rightarrow$ The algorithm returns at most $2\ell/d$ degree d polynomials.

Example. $N = 10^{170} + 7$

$$\#\left\{f \in (N, m_2x - m_1y) \mid \deg f = 3, \|f\|_{2,s} \leq (CN)^{1/6} \text{ and } \bar{\alpha}(f, B) \leq -2\right\}$$

$C^{1/6}$	$B = 100$	$B = 1000$	$B = 10000$
1	224	1014	8267
2	383	1476	10972
3	662	2093	13952
4	1387	3075	17649
5	9756	5022	22656
6	-	11100	30117
7	-	-	42804
8	-	-	69903
9	-	-	171650

Is there a special-$\mathfrak{q}$ version?

Is there a special-$\mathfrak{q}$ version?

Yes.

Thanks!