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 $\delta = \delta(N) (\in \{6, 7\} \text{ in practice})$,
- $f_1$ 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

- 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 \textbf{homogeneous} polynomials \( f_1, f_2 \in \mathbb{Z}[x, y] \) such that

\[
\begin{align*}
\cdot \quad & \text{deg } f_1 + \text{deg } f_2 = \delta, \text{ where } \delta = \delta(N) \in \{6, 7\} \text{ in practice,} \\
\cdot \quad & f_1 \text{ and } f_2 \text{ are distinct and irreducible,} \\
\cdot \quad & \exists m_1, m_2 \in \mathbb{Z} \setminus \{0\} \text{ such that } f_1(m_1, m_2) \equiv f_2(m_1, m_2) \equiv 0 \pmod{N}, \\
\cdot \quad & f_1 \text{ and } f_2 \text{ produce many smooth values in the sieve stage.}
\end{align*}
\]

Very roughly speaking, smoothness probabilities are correlated with

\[
\begin{align*}
\cdot \quad & \text{Coefficient size,} \\
\cdot \quad & \text{Number of real roots,} \\
\cdot \quad & \text{Roots modulo small primes.}
\end{align*}
\]

\[
\begin{align*}
\{\text{Size properties}\}
\end{align*}
\]

See \cite{Brent, Montgomery & Murphy} 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 $\delta = \delta(N) (\in \{6, 7\} \text{ in practice})$,
- $f_1$ 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

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

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

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

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 $\delta = \delta(N) (\in \{6, 7\} \text{ in practice})$,
- $f_1$ 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

- $\text{deg } f_1 + \text{deg } f_2 = \delta$, where $\delta = \delta(N) (\in \{6, 7\} \text{ in practice})$,
- $f_1$ 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 \( |\operatorname{Res}(f_1, f_2)| \) 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

$$L = b_1 \mathbb{Z} + \ldots + b_k \mathbb{Z},$$

where $b_1, \ldots, b_k \in \mathbb{R}^n$ are linearly independent.

**Key invariants:**

- $k$ — the *dimension* of $L$
- $\det L := (\det(b_i \cdot b_j)_{1 \leq i, j \leq k})^{1/2}$ — the *determinant* of $L$

**[Lenstra, Lenstra & Lovász 1982]:** Given $b_1, \ldots, b_k \in \mathbb{Z}^n$, there exists an algorithm (now called *LLL-reduction*) that can be used to compute $a_1, a_2 \in L$ such that

$$\|a_1\|_2 \leq 2^{(k-1)/4} \det(L)^{1/k} \quad \text{and} \quad \|a_2\|_2 \leq 2^{k/4} \det(L)^{1/(k-1)}$$

in time polynomial in $k, n$ and $\max_{1 \leq i \leq k} \log \|b_i\|_2$. 
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 \(\ell\) 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 \(f(m_1, m_2) \equiv 0 \pmod{N}\).
Suppose we have $1 \leq k \leq d - 1$ linearly independent length $d + 1$ GPs

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

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 → Polynomials**

The set of all such vectors,

\[
L := \left\{ s^{-d/2} \left( a_0, a_1 s, \ldots, a_d s^d \right) \mid (a_0, a_1, \ldots, a_d) \in \mathbb{Z}^{d+1} \right. \\
\left. \text{and } \sum_{j=0}^{d} a_j c_{i,j} = 0 \text{ for } i = 1, \ldots, k \right\},
\]

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 $\rightarrow$ 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 $O(N^{1/(2d)})$ (+ some other conditions), then there exists $d - 1$ linearly independent length $d + 1$ geometric progressions $c_1, c_2, \ldots, c_{d-1}$ 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

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

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

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]:

<table>
<thead>
<tr>
<th>$d$</th>
<th>$|f_i|_{2,s}$</th>
<th>$s$</th>
<th>Optimal?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$O(N^{1/4})$</td>
<td>$O(N^{1/4})$</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>$O(N^{5/24})$</td>
<td>$O(N^{1/12})$</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>$O(N^{5/28})$</td>
<td>$O(N^{1/28})$</td>
<td>No</td>
</tr>
</tbody>
</table>

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

\[ c_{91} = \begin{array}{c}
4567176039894108704358752160655628192034927306 \\
969828397739074346628988327155475222843793393.
\end{array} \]

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

\[
\begin{align*}
    f_1 &= 21545x^3 + 3349054x^2 - 10356871479051937193x + 1263295294354066431546642250 \\
    f_2 &= 1356640x^3 + 210882368x^2 - 652118673869097609994x - 11972068980454909092333428939
\end{align*}
\]

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:

\[
\begin{bmatrix}
am_2^{d-1}, am_2^{d-2} m_1, am_2^{d-3} m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, m_1\left(\frac{am_1^d - vN}{m_2}\right)
\end{bmatrix}
\]

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

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

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

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

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

\[
c_1 = \begin{cases} am_2^{d-1}, am_2^{d-2}m_1, am_2^{d-3}m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, m_1\left(\frac{am_1^d - vN}{m_2}\right) \end{cases}
\]

\[
c_2 = \]

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

<table>
<thead>
<tr>
<th>( d )</th>
<th>( | f_i |_{2,s} )</th>
<th>( s )</th>
<th>Optimal?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>( O\left(N^{1/6}\right) )</td>
<td>( O\left(N^{1/9}\right) )</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>( O\left(N^{3/20}\right) )</td>
<td>( O\left(N^{1/20}\right) )</td>
<td>No</td>
</tr>
</tbody>
</table>
\[ k = 2: \text{construction} \]

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

\[
\begin{align*}
\mathbf{c}_1 &= \left\{ am_2^{d-1}, am_2^{d-2} m_1, am_2^{d-3} m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d - vN}{m_2}, \frac{m_1(am_1^d - vN)}{m_2^2} \right\} \\
\mathbf{c}_2 &= \ldots
\end{align*}
\]

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

<table>
<thead>
<tr>
<th>( d )</th>
<th>( | f_i |_{2,s} )</th>
<th>( s )</th>
<th>Optimal?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>( O(N^{1/6}) )</td>
<td>( O(N^{1/9}) )</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>( O(N^{3/20}) )</td>
<td>( O(N^{1/20}) )</td>
<td>No</td>
</tr>
</tbody>
</table>

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

\[
\left| m_1 - \frac{vN}{a} \right|^{1/d} = \begin{cases} 
O(m_2^{3/2}) & \text{for } d = 3; \\
O(m_2^{5/4}) & \text{for } d = 4;
\end{cases}
\]

\[
m_2 = \begin{cases}
O(N^{2/9}) & \text{for } d = 3, \\
O(N^{1/5}) & \text{for } d = 4.
\end{cases}
\]
$k = 2$: construction

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

\[
\begin{align*}
    \mathbf{c}_1 &= \begin{bmatrix}
        am_2^{d-1}, am_2^{d-2} m_1, am_2^{d-3} m_1^2, \ldots, am_1^{d-1}, \frac{am_1^d}{m_2}, \frac{am_1^{d+1} - (vm_1 + um_2)N}{m_2^2}
    \end{bmatrix} \\
    \mathbf{c}_2 &= \begin{bmatrix}
    \end{bmatrix}
\end{align*}
\]

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

<table>
<thead>
<tr>
<th>$d$</th>
<th>$|f_i|_{2,s}$</th>
<th>$s$</th>
<th>Optimal?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>$O\left(N^{1/6}\right)$</td>
<td>$O\left(N^{1/9}\right)$</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>$O\left(N^{3/20}\right)$</td>
<td>$O\left(N^{1/20}\right)$</td>
<td>No</td>
</tr>
</tbody>
</table>

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

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

\[
m_2 = \begin{cases} 
    O\left(N^{2/9}\right) & \text{for } d = 3, \\
    O\left(N^{1/5}\right) & \text{for } d = 4.
\end{cases}
\]
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 $p_{p,r} = (p, x - ry)$ for $r \in \mathbb{F}_p$ and $p_{p,\infty} = (p, y)$.

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

$$\alpha(f, B) = \sum_{\substack{p_{p,r} \leq B \leq \infty \mathbb{F}_p \setminus \{p\} \setminus \{p_{p,\infty}\} \setminus \{p_{p,r}\} \setminus \{p, x - ry\} \setminus \{p, y\}}} (1 - \sigma(f, p_{p,r}p)) \frac{\log p}{p^2 - 1}.$$
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
\begin{itemize}
  \item $f$ is a non-constant and irreducible;
  \item $f$ is homogeneous of degree $\leq d$;
  \item $f \in (N, m_2x - m_1y)$; and
  \item $\|f\|_{2,s} \leq (CN)^{1/2d}$.
\end{itemize}

Lemma. If $f_1, f_2 \in \mathcal{M}$ satisfy
\[
\sum_{\substack{p, r \not\in (N) \\ p \leq B}} \sigma(f_1, p, r)\sigma(f_2, p, r) \log p > \log C
\]
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 $|\text{Res}(f_1, f_2)|$ used in the resultant bound.
Lemma

Let \( M = M(N, m_2, m_1; d, s, C) \) 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 \( f_1, f_2 \in M \) satisfy

\[
\sum_{\substack{p, r \in \mathbb{P}(N) \\
p \leq B \land \ p \nmid \ (f_1, p, r)}} \sigma(f_1, p, r) \sigma(f_2, p, r) \log p > \log C
\]

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

\[\Rightarrow\] If \( p_{r_1}, \ldots, p_{r_n} \nmid (N) \) are distinct and \( \prod_{i=1}^n p_i > C \), then the vectors

\[
f \cdot (1 - \sigma(f, p_{r_1}), 1 - \sigma(f, p_{r_2}), \ldots, 1 - \sigma(f, p_{r_n})) \quad \text{for} \ f \in M/ \sim,
\]

have a nonzero minimum “distance”.
A combinatorial bound

Given distinct \( p_1, \ldots, p_n \notin (N) \), positive real weights \( \beta_1, \ldots, \beta_n \) and a real number \( \ell \geq 1 \), there are at most \( 2\ell \) polynomials \( f \in \mathcal{M} \) such that

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

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

Given distinct \( p_1, \ldots, p_n \) \( \in \mathbb{R}^N \), positive real weights \( \beta_1, \ldots, \beta_n \) and a real number \( \ell \geq 1 \), there are at most \( 2\ell \) polynomials \( f \in \mathcal{M} \) such that

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

Example. \#\{ \( f \in (N, m_2x - m_1y) \mid \deg f \leq 3, \|f\|_{2,s} \leq (CN)^{1/6}, \overline{\alpha}(f, B) \leq -2 \} \}

<table>
<thead>
<tr>
<th>( C^{1/6} )</th>
<th>( B = 100 )</th>
<th>( B = 1000 )</th>
<th>( B = 10000 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>860</td>
<td>83463</td>
<td>7299206</td>
</tr>
<tr>
<td>2</td>
<td>1484</td>
<td>130046</td>
<td>10499454</td>
</tr>
<tr>
<td>3</td>
<td>2581</td>
<td>193086</td>
<td>14121084</td>
</tr>
<tr>
<td>4</td>
<td>5434</td>
<td>294311</td>
<td>18696869</td>
</tr>
<tr>
<td>5</td>
<td>38188</td>
<td>496011</td>
<td>24973925</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>1127183</td>
<td>34414014</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>50578542</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>85275302</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>215937570</td>
</tr>
</tbody>
</table>

Ignores roots at \( \infty \).
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 $F_q[x]$.

[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 \( p_1, \ldots, p_n \not\supset (N) \) and integer weights \( z_1, \ldots, z_n > 0 \).  

**Output:** All \( f \in \mathcal{M} \) such that \( \sum_{i=1}^{n} \sigma(f, p_i)z_i \log p_i \) is “sufficiently large”.

1. Construct a homogeneous polynomial \( h \in (N, m_2x - m_1y)^{z_0} \cap (\bigcap_{i=1}^{n} p_i^{z_i}) \) 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 \( \ell \) polynomials in \( (N, m_2x - m_1y)^{z_0} \cap (\bigcap_{i=1}^{n} p_i^{z_i}) \).  
   b. Scale it appropriately, then LLL-reduce.

2. Factor \( h \) over \( \mathbb{Q} \) and return all factors in \( \mathcal{M} \).

Here, “sufficiently large” means \((CN)^{\deg h/(2d)} \cdot \|h\|_{2,s}^d < N^{z_0} \cdot \prod_{i=1}^{n} p_i^{\sigma(f, p_i)z_i} \).
Algorithm

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

**INPUT:** \( \mathcal{M} \), distinct ideals \( p_1, \ldots, p_n \not\supseteq (N) \) and integer weights \( z_1, \ldots, z_n > 0 \).

**OUTPUT:** All \( f \in \mathcal{M} \) such that \( \sum_{i=1}^{n} \sigma(f, p_i) z_i \log p_i \) is “sufficiently large”.

1. Construct a homogeneous polynomial \( h \in (N, m_2x - m_1y)^{z_0} \cap (\bigcap_{i=1}^{n} p_i^{z_i}) \) 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 \( \ell \) polynomials in \( (N, m_2x - m_1y)^{z_0} \cap (\bigcap_{i=1}^{n} p_i^{z_i}) \).
   b. Scale it appropriately, then LLL-reduce.

2. Factor \( h \) over \( \mathbb{Q} \) and return all factors in \( \mathcal{M} \).

Here, “sufficiently large” means \((CN)^{\deg h/(2d)} \cdot \|h\|_{2,s}^{d} < N^{z_0} \cdot \prod_{i=1}^{n} p_i^{\sigma(f, p_i) z_i} \cdot \frac{|\text{Res}(f, h)|}{\text{Divides}|\text{Res}(f, h)|}\).
Theorem

Let $p_1, \ldots, p_n \not\in \mathcal{N}$ 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

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

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

Example. \( N = 10^{170} + 7 \)

\[
\left\{ f \in (N, m_2x - m_1y) \mid \deg f \leq 3, \|f\|_{2,5} \leq (CN)^{1/6} \text{ and } \tilde{\alpha}(f, B) \leq -2 \right\}
\]

<table>
<thead>
<tr>
<th>( B )</th>
<th>#p</th>
<th>( C^{1/6} )</th>
<th>dim</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>17</td>
<td>1.78</td>
<td>809</td>
</tr>
<tr>
<td>20</td>
<td>77</td>
<td>1.99</td>
<td>1143</td>
</tr>
<tr>
<td>30</td>
<td>129</td>
<td>2.06</td>
<td>1274</td>
</tr>
<tr>
<td>40</td>
<td>197</td>
<td>2.12</td>
<td>1400</td>
</tr>
<tr>
<td>50</td>
<td>328</td>
<td>2.20</td>
<td>1579</td>
</tr>
<tr>
<td>100</td>
<td>1060</td>
<td>2.41</td>
<td>2153</td>
</tr>
<tr>
<td>1000</td>
<td>76127</td>
<td>3.39</td>
<td>12412</td>
</tr>
</tbody>
</table>

Have to LLL-reduced a lattice with huge dimension for each \((N, m_2x - m_1y)\).
Algorithmic bounds

Each output of the algorithm is a factor of $h$, which has degree equal to $\ell$

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

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

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

<table>
<thead>
<tr>
<th>$C^{1/6}$</th>
<th>$B = 100$</th>
<th>$B = 1000$</th>
<th>$B = 10000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>224</td>
<td>1014</td>
<td>8267</td>
</tr>
<tr>
<td>2</td>
<td>383</td>
<td>1476</td>
<td>10972</td>
</tr>
<tr>
<td>3</td>
<td>662</td>
<td>2093</td>
<td>13952</td>
</tr>
<tr>
<td>4</td>
<td>1387</td>
<td>3075</td>
<td>17649</td>
</tr>
<tr>
<td>5</td>
<td>9756</td>
<td>5022</td>
<td>22656</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>11100</td>
<td>30117</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>42804</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>69903</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>171650</td>
</tr>
</tbody>
</table>
Is there a special-\( q \) version?
Is there a special-q version?

Yes.
THANKS!