# CSE291-14: The Number Field Sieve

https://cseweb.ucsd.edu/classes/wi22/cse291-14

Emmanuel Thomé

January 20, 2022

# Part 3c

## NFS in the not-so-easy case

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# Recap from last time

We learned a lot from the algebraic number theory background.

How do we get back on our feet, and think about a factoring algorithm?

- The roadmap of the too-easy algorithm seemed very simple.
- We learned about multiple roadblocks that we have to circumvent to make this work:
  - Beyond the entirely-trivial cases (how do we factor $F_7$?)
  - and also in greater generality (how do we factor general numbers?)
- And then, assuming all this can be overcome, can we really make this a sieving algorithm?

# Plan

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha\rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$.
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$.
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$.
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$. **This is tricky!**
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$.
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# Plan

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# NFS

Not all fields are as cool as $\mathbb{Q}(\sqrt{-11})$ (see lecture 4).

- The ring of integers is not always obvious.
  Sometimes, it is even extremely hard to compute $\mathcal{O}_K$!
- In general, we do not have unique factorization of elements.
- We're not certain that we'll always like to restrict ourselves to a monic definition polynomial. (Spoiler alert: indeed, we won't!)
- The units can be much more complicated than $\pm 1$.

We expect some difficulties!

# Pollard's $F_7$ example

In the cubic integers example, Pollard only had the units issue to deal with.

- The field $\mathbb{Q}(\alpha) = \mathbb{Q}[x]/(x^3 + 2)$ does have a unit of infinite order.
- Fortunately, this generator is easy to find: $1 + \alpha$.
  This is easy to see: $\mathrm{Res}(1 + x, x^3 + 2) = (-1)^3 + 2 = 1$.

So there's no really annoying difficulty here.

We can simply add a column with the valuation in $(1 + \alpha)$.

What is a real pain, however, is how to factor algebraic numbers into elements. We'll leave that aside.

# Pollard's $F_7$ example

In the case of $\mathbb{Q}(\sqrt[3]{-2})$, we would need the following preparation work.

- Choose a smoothness bound $B$.
- List all primes below $B$.
- List all primes in $\mathbb{Q}(\alpha)$ whose norm is below $B$.
- List the known units ($-1$ and $1+\alpha$)

Then we would need to find pairs $(a, b)$ such that we have simultaneous smoothness.

- Can we do that with sieving? Yes.
- Will this end up giving us a factorization? Yes.

# Sieving for smooth $(a, b)$

We are interested in many possible polynomials $\phi = a - bx$.

Note: it is useless to consider the case $\gcd(a, b) > 1$, since it brings no useful new information compared to the coprime case.

Pollard used sieving in a simple way:

- For each $b$ from 1 to 2000, sieve the range $-4800 \leq a < 4800$ in order to detect the smooth values of $a - bm$.
  See file pollard.sage on Canvas.
- For each apparently smooth $a - bm$, compute and try to factor $\mathrm{Norm}(a - b\alpha)$.
- In cases where $\mathrm{Norm}(a - b\alpha)$ is smooth, factor it, and record this information.

# $F_7$

$F_7$ was first factored with CFRAC in 1970.

Pollard's method: 1988.

- Is it significantly faster? Not really.
- Is it a general factoring method? Not at all.
- But it does bring something new.

First, we'll see how it can work with a number fields where not all ideals are principal.

# Plan

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# Sieving: later

We're definitely going to describe a sieving algorithm.

But: for the moment (next few slides), our description will be trial-division-based.

Remember that conceptually, sieving can be introduced after the fact by swapping two loops.

# Two sides

Whenever we want to create a relation, there are clearly two sides to consider. Similarities are very strong.

- On the rational side, we compute $a - bm$.
  - For each prime number $p$, see if $p \mid a - bm$. If yes, record the valuation.
  - If $a - bm$ is fully factored, we're happy.
  - Pay attention to $\pm 1$.
- On the algebraic side, we "compute" $a - b\alpha$.
  - For each prime ideal $\mathfrak{p}$, see if $\mathfrak{p} \mid \langle a - b\alpha \rangle$. If yes, record the valuation.
  - If $\langle a - b\alpha \rangle$ is fully factored this way, we're happy.
  - Pay attention to units.

  Note: this does not mean that we factor $a - b\alpha$.

  Note2: we need to think a bit about the interpretation of the relation that we obtain.

# Factoring on the algebraic side

In order to be able to factor things on the algebraic side:

- we need to determine all "small" prime ideals that will define our factor base.
  "small": their norm must be below some bound $B$.
- we need be able to check if an ideal divides another.

We're also aware of the gap between factoring an element (which is not well-defined), and factoring an ideal into prime ideals. Units are part of this gap.

# Plan

Prime ideals and factorization of $\langle a - b\alpha \rangle$

- Hard things vs doable things
- Describing prime ideals
- Factoring into ideals

# Bad news, first

Real-life example (from DLP-240):

$$\begin{aligned} f = {} & 286512172700675411986966846394359924874576536408786368056\, x^3 \\ & + 24908820300715766136475115982439735516581888603817255539890\, x^2 \\ & - 18763697560013016564403953928327121035580409459944854652737\, x \\ & - 236610408827000256250190838220824122997878994595785432202599 \end{aligned}$$

$\operatorname{disc} f =$ A 236-digit integer (not an RSA modulus!).

**Computing $\mathcal{O}_K$ is very hard**

It is very hard to be absolutely sure that we have computed $\mathcal{O}_K$.

**Computing $\mathcal{O}_K^*$ is infeasible**

The computation of a system of generators for $\mathcal{O}_K^*$ is completely out of reach.

# Good news

While the global objects (such as $\mathcal{O}_K$ and $\mathcal{O}_K^*$) are hard to compute, everything that is local (attached to a prime $p$) is much more tractable (polynomial in $\log p$ and $\deg f$).

- For any prime $p$, we can describe the prime ideals of $\mathcal{O}_K$ that are above $p$, even if we do not know $\mathcal{O}_K$.
- For any prime ideal $\mathfrak{p}$, finding the $\mathfrak{p}$-valuation of an ideal such as $\langle a - b\alpha \rangle$ is doable, even if we do not know $\mathcal{O}_K$.
- For most primes $p$, these tasks are actually very easy.

The other bit of good news is that we can work around the fact that computing $\mathcal{O}_K^*$ is out of reach.

# Plan

Prime ideals and factorization of $\langle a - b\alpha \rangle$

- Hard things vs doable things
- Describing prime ideals
- Factoring into ideals

# What are the prime ideals above $p$?

Preliminary question: does $p$ divide $f_n$ or $\operatorname{disc}(f)$?
If yes, you'll have to ask an expert (they won't charge much).

If not, then $\mathbb{Z}[\alpha]$ (or $\mathbb{Z}[\hat{\alpha}]$ if $f$ not monic) can be used in lieu of $\mathcal{O}_K$. We can really do as if they were the same.

- If $f$ factors modulo $p$ into irreducible factors of degrees $d_1 + \cdots + d_k = n$, then there are $k$ prime ideals above $p$, of residue class degrees $d_1$ to $d_k$.
- Repeated factors cannot appear (because $p \nmid \operatorname{disc} f$).

**Example**

$f = x^3 + 2$, $p = 31$: $f$ splits completely $\bmod p$.
There are three prime ideals of degree 1 above $p$.

$f = x^3 + 2$, $p = 41$: $f$ splits $\bmod p$ into $(\deg = 1) \times (\deg = 2)$.
There are two prime ideals, of degrees 1 and 2, above $p$.

# What are the prime ideals above $p$?

## Identifying most prime ideals

In the easy case ($p \nmid f_n \operatorname{disc} f$), a prime ideal above $p$ is uniquely determined by

- The prime number $p$
- One of the irreducible factors of $f$ mod $p$.

The most typical case is when the residue class degree is 1.
Such a prime ideal can be identified as $(p, x - r)$, or $(p, \alpha - r)$, or $(p, r)$ depending on notations.

$(p, x - r)$ is the prime ideal above $p$ that contains all algebraic integers that are $\mathcal{O}_K$-multiples of $(\alpha - r)$.

This is an implicit description, but it is sufficient for NFS.
Caveat: when $f_n \neq 1$, $(p, x - r) \neq \langle p, \alpha - r \rangle$.

# Identifying most prime ideals

```
ideals=[]
f=K.defining_polynomial()
Disc=f.discriminant()
for p in prime_range(10000):
    if gcd(p,Disc) != 1:
        continue
    fp=f.change_ring(GF(p)).factor()
    for g,m in fp:
        assert m == 1
        if p^(g.degree()) < 10000:
            ideals.append((p,g))
```

Cado-NFS has a program called `makefb` which does just this.

# What are the ideals that we miss?

There are prime ideals above the prime divisors of $f_n \operatorname{disc} f$.
Cado-NFS calls them "bad ideals".

- Whenever we look at what happens above a given $p$, everything is doable with a bit of code.
- We are only interested in prime ideals of small norm, and finding the prime numbers $p$ in this range that divide $f_n \operatorname{disc} f$ is easy because they're small.

Note: in some cases, the simple mechanism can be extended.

There are a few "bad ideals" in $\mathcal{O}_K$.
With some effort, we can find and describe them.

# Plan

Prime ideals and factorization of $\langle a - b\alpha \rangle$

- Hard things vs doable things
- Describing prime ideals
- Factoring into ideals

# Divisibility by easy ideals

Question: is some ideal above $p$ a divisor of the ideal $\langle a - b\alpha \rangle$?

Preliminary question: does $p$ divide $f_n$ or $\mathrm{disc}(f)$?
If yes, you'll have to ask an expert (they won't charge much).

If not, we are in the easy case, and it is quite simple.

# Divisibility by easy ideals

Assume that
- $p \nmid f_n \operatorname{disc} f$ (easy case).
- $p$ is coprime to $\gcd(a, b)$.
- $\mathfrak{p}$ is identified by $(p, g(x))$.
- We want to check if $\mathfrak{p} \mid \langle a - b\alpha\rangle$.

$$\begin{aligned}\mathfrak{p} \mid \langle a - b\alpha\rangle &\Leftrightarrow g(a/b) \equiv 0 \mod p\\ &\Leftrightarrow \operatorname{Res}(a - bx, g(x)) \equiv 0 \mod p\end{aligned}$$

Side-effect: at most one matching $\mathfrak{p}$ above a given $p$, and $\nu_{\mathfrak{p}}(\langle a - b\alpha\rangle) = \nu_p(\operatorname{Res}(a - bx, f(x)))$.

## Only ideals of degree 1 matter

This can happen only if $\deg g = 1$.
As long as we are factoring $\langle a - b\alpha\rangle$, only ideals of the form $(p, x - r)$ can appear.

# Data format

To represent the factorization of $\langle a - b\alpha\rangle$, we typically store this information:

- The integers $a$ and $b$.
- All the prime factors of $\mathrm{Res}(a - bx, f(x))$.

This is concise, and sufficient to precisely identify all prime ideals in the factorization (when we need to do so).

- For most primes, this boils down to computing $a/b \bmod p$.
- For "bad primes", this is doable as well.

All this identification work can be done basically as fast as `printf`.

# Plan

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# Example from Cado-NFS

To do an $F_7$ factorization with Cado-NFS:

```
git clone https://gitlab.inria.fr/cado-nfs/cado-nfs
cd cado-nfs
make -j4
[download f7.params]
[download f7.poly]
./cado-nfs.py --wdir /tmp/F7 f7.params slaves.hostnames=localhost
```

We find in one of the `/tmp/F7/F7.upload/F7.*.gz` files:

```
-1044,509:2,2,d,13,10f,119,fa7,3a03:2,b,1f,161,e2f
```

# Example from Cado-NFS

A relation: `-1044,509:2,2,d,13,10f,119,fa7,3a03:2,b,1f,161,e2f`

- `-1044,509`: These are $a = -1044$ and $b = 509$ (in decimal).
- `2,2,d,13,10f,119,fa7,3a03`: The prime factors of $a - b \times 2^{43}$.
- `2,b,1f,161,e2f`: The prime factors of $\mathrm{Res}(a - bx, x^3 + 2)$.

This says that (blue and red are hex above, decimal below):

$$
\begin{aligned}
-1044 - 509 \cdot 2^{43} &= \pm 2^2 \times 13 \times 19 \times \cdots \\
\langle -1044 - 509\alpha \rangle &= \text{a ``bad ideal'' of norm 2} \\
&\quad \times (11, \alpha - 4) \\
&\quad \times (31, \alpha - 27) \\
&\quad \times (353, \alpha - 292) \\
&\quad \times (3631, \alpha - 1389).
\end{aligned}
$$

Violet numbers such as 1389 are implicit: $a/b \bmod 3631 = 1389$.

# Things to pay attention to

- Repeated factors are rare and when a prime divides multiple times, it is printed multiple times (see the 2 in the example).
- The unit on the rational side does not appear in the relation. It's easy enough to find out the sign!
- There is some information about "bad ideals". We might provide it to our expert so that they can identify these ideals properly.
- On the algebraic side, we only have a factorization into ideals.

# Important caveat for non-monic $f$

Reminder:

$$\mathrm{Norm}\langle a - b\alpha\rangle = \mathrm{Norm}(a - b\alpha) = \frac{1}{f_n}\mathrm{Res}(a - bx, f(x)).$$

- We claim that we are writing down the factorization of $\langle a - b\alpha\rangle$.
- But the prime factors that we list are those of $\mathrm{Res}(a - bx, f(x))$.
- There's got to be something missing.

**The ideal $J$ is here to square things up**

When $f_n \neq 1$, we are actually writing down the factorization of $J \times \langle a - b\alpha\rangle$, with $J = \langle 1, \alpha\rangle^{-1} = \{x,\ x \in \mathcal{O}_K \text{ and } x\alpha \in \mathcal{O}_K\}$.

- $J = \langle 1, \alpha\rangle^{-1}$ is an integral ideal of norm $f_n$.
  $J$ has no reason to be prime (e.g., if $f_n$ isn't, $J$ isn't either).
- This is hardly ever mentioned in the literature.

# Example with non-monic $f$

The number $2^{199} + 3^{109}$ is a nice 60-digit number to play with.

```
./cado-nfs.py --wdir /tmp/c60 $(bc<<<2^199+3^109)
```

# Summary of the information we have

On the algebraic side, we have:

- in a straightforward manner, the ideals and valuations in the factorization of $\langle a - b\alpha \rangle \times J$, when $p \nmid f_n \operatorname{disc}(f)$ (all $p$ but finitely many).
- with some extra work, the full factorization of $\langle a - b\alpha \rangle$ can be obtained, but we'll have to ask our expert for that.

# What remains to be done

If we follow our basic workplan, we can see how linear algebra will produce a subset of the $(a - bx)$ such that

- $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$ (we will add a column with the sign for that).
- $\prod_i \langle a_i - b_i \alpha \rangle$ has even valuations at
  - all easy prime ideals if we only look at these.
  - all prime ideals with some extra effort.

Therefore $\langle \prod_i (a_i - b_i \alpha) \rangle$ is the square of an ideal, but we do not know if $\prod_i (a_i - b_i \alpha)$ is the square of an element!

We will see how to work around this difficulty when we address the square root computation.

# Plan

A roadmap for NFS

Stumbling blocks

Prime ideals and factorization of $\langle a - b\alpha \rangle$

Making sense of a relation

The main steps of NFS and the NFS diagram

# What we have done so far

We have a few ideas of how an NFS algorithm could look like.

- So far, we mentioned ad hoc numbers, but our demo gives away the fact that it also works in greater generality.
- Factoring into prime ideals is doable.
- We mentioned some possibilities down the road, but I claim that these can be circumvented.

Now: list (and name) all the different steps of the General Number Field Sieve (GNFS).

We're going to repeat blocks of our sketch slide "How would we factor $N$?"

# How would we factor $N$?

- Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$. Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.
  - The analysis will help us see that in greater detail.
  - For some numbers, some very nice values exist.
- Fix a smoothness bound $B$.
- Find many pairs $(a, b)$ such that:
  - $a - bm$ factors into primes below $B$.
  - $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.
- Using linear algebra, find a subset of the $(a - bx)$ such that:
  - $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
  - $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$. **This is tricky!**
- Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

# Finding $f$ and $m$

Find $f \in \mathbb{Z}[x]$ and $m \in \mathbb{Z}$ such that $f(m) \equiv 0 \mod N$.
Neither $m$, nor $\deg f$, nor the coefficients of $f$ should be too large.

- The analysis will help us see that in greater detail.
- For some numbers, some very nice values exist.

This is called **Polynomial Selection**: next lecture.

Here's a simple method called base-$m$ to do it for arbitrary $N$:

- Choose the degree $d$ of $f$ s.t. $N > 2^{d^2}$.
- Set $m = \lceil N^{1/(d+1)} \rceil$.
- Write $N$ in base $m$: $N = \sum_{i=0}^{d} f_i m^i$ where $0 \leq f_i < m$.
- Set $f = \sum_{i=0}^{d} f_i x^i$. (not monic!).
- Notation-wise, we sometimes write "the rational polynomial" as $g = x - m$.

# Parameters

Remark that $d$ is a free parameter in the previous slide.

So is, for example, the bound $B$.

As well as many, many other parameters!

**This is called parameter selection**

Parameter selection is among the black arts in NFS!

- Asymptotic analysis gives asymptotic guidelines.
- In practice, it's a complicated matter which requires a lot of global understanding of how NFS works.

We'll tentatively cover a bit of the practical side of this by the end of the quarter.

# Finding pairs $a, b$

Find many pairs $(a, b)$ such that:

- $a - bm$ factors into primes below $B$.
- $\langle a - b\alpha \rangle$ factors into prime ideals of norm below $B$.

This is called **Relation Collection**: beginning of February.

One of the ways to do relation collection is sieving.

- It is actually possible to sieve for rational primes $p \in \mathbb{Z}$ but also for prime ideals $\mathfrak{p} \subset \mathcal{O}_K$.
- There are many, many, many parameters.
- Most of the old knowledge of sieving from the QS era is relevant.
- This is the most expensive part, computationally speaking.

# Combining pairs

Using linear algebra, find a subset of the $(a - bx)$ such that:

- $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
- $\prod_i (a_i - b_i \alpha)$ is a square in $\mathbb{Z}[\alpha]$. **This is tricky!**

# Combining pairs

Using linear algebra, find a subset of the $(a - bx)$ such that:

- $\prod_i (a_i - b_i m)$ is a square in $\mathbb{Z}$.
- $\prod_i (a_i - b_i \alpha)$ is (almost) a square in $\mathbb{Z}[\alpha]$.

This comprises two steps: We will see both mid-February.

- The **Filtering** step is a pre-processing step.
- Then we have **Linear Algebra** proper.

Linear algebra is the second most expensive step, and requires expensive hardware, too.

# Factoring $N$, at last

Arrange so that $\prod_i (a_i - b_i\alpha)$ really is a square in $\mathbb{Z}[\alpha]$.
Write down both square roots in $\mathbb{Z}$ and $\mathbb{Z}[\alpha]$, map them to $\mathbb{Z}/N\mathbb{Z}$, and hopefully get a factor.

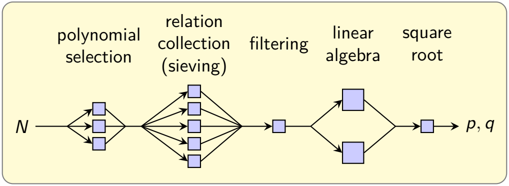
Again, two steps here. End of February.

- A pre-processing step called **the characters step**.
- Then **the square root step**.

This step will entail some more algebraic number theory, as well asymptotically fast algorithms.

As each square root only has probability $1/2$ to factor $N$, this step is designed to produce several independent square roots.

# The different steps of NFS




Note: there is also a version of NFS that computes discrete logarithms in $\mathbb{F}_p^*$. The main outline is similar. End of February.

# Some handwaving

- We find $f$ with a known root $m$ modulo $N$.
- Let $\mathbb{Q}(\alpha)$ be the number field defined by $f$.
- For any polynomial $P(x)$, we have:
  - the integer $P(m)$;
  - the number field element $P(\alpha)$;

  These are compatible: both map to $P(m)$ mod $N$ in $\mathbb{Z}/N\mathbb{Z}$.

$$
\begin{tikzcd}
 & & \mathbb{Z}[x] \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & & \\
\text{subring of } \mathbb{Q} & \mathbb{Z}[m] = \mathbb{Z} \arrow[dr, "\bmod N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\alpha \to m"] & \text{subring of } \mathbb{Q}(\alpha) \\
 & & \mathbb{Z}/N\mathbb{Z} & &
\end{tikzcd}
$$

# Some handwaving

- We find $f$ with a known root $m$ modulo $N$.
- Let $\mathbb{Q}(\alpha)$ be the number field defined by $f$.
- For any polynomial $a - bx$, we have:
  - the integer $a - bm$;
  - the number field element $a - b\alpha$;

  These are compatible: both map to $P(m)$ mod $N$ in $\mathbb{Z}/N\mathbb{Z}$.

$$
\begin{tikzcd}
 & \mathbb{Z}[x] \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & \\
\text{subring of } \mathbb{Q} \quad \mathbb{Z}[m] = \mathbb{Z} \arrow[dr, "\text{mod } N"'] & & \mathbb{Z}[\alpha] \quad \text{subring of } \mathbb{Q}(\alpha) \arrow[dl, "\alpha \to m"] \\
 & \mathbb{Z}/N\mathbb{Z} &
\end{tikzcd}
$$

# Some handwaving

- We find $f$ with a known root $m$ modulo $N$.
- Let $\mathbb{Q}(\alpha)$ be the number field defined by $f$.
- For any polynomial $\prod_i (a_i - b_i x)$, we have:
  - the integer $\prod_i (a_i - b_i m)$;
  - the number field element $\prod_i (a_i - b_i \alpha)$;

  These are compatible: both map to $P(m)$ mod $N$ in $\mathbb{Z}/N\mathbb{Z}$.

$$
\begin{tikzcd}
 & & \mathbb{Z}[x] \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & & \\
\text{subring of } \mathbb{Q} & \mathbb{Z}[m] = \mathbb{Z} \arrow[dr, "\text{mod } N"'] & & \mathbb{Z}[\alpha] \arrow[dl, "\alpha \to m"] & \text{subring of } \mathbb{Q}(\alpha) \\
 & & \mathbb{Z}/N\mathbb{Z} & &
\end{tikzcd}
$$

# Write something multiplicative

The NFS diagram can also be written as a multiplicative diagram, even though it is a bit awkward to write it as such.

$$\begin{tikzcd} & \text{group generated by } \mathbb{Z}[x] - \{\text{things that do not map to } \mathbb{Z}/N\mathbb{Z}^*\} \arrow[dl, "x \to m"'] \arrow[dr, "x \to \alpha"] & \\ \text{subgroup of } \mathbb{Q}^* \arrow[dr, "\text{mod } N"'] & & \text{subgroup of } \mathbb{Q}(\alpha)^* \arrow[dl, "\alpha \to m"] \\ & \mathbb{Z}/N\mathbb{Z}^* & \end{tikzcd}$$

No difference in practice between the two diagrams.

- The multiplicative one just says that we won't stumble on factors of $N$ accidentally. There is no practical difference between $\mathbb{Z}[x]$ and the structure on top.
- The multiplicative diagram does have an interest in the discrete logarithm context.

# Rundown of an NFS computation

A more detailed look at the factorization of $2^{199} + 3^{109}$.

```
./cado-nfs.py --wdir /tmp/c60 $(bc<<<2^199+3^109)
```