Algorithms in Algebra and Number Theory

Fabian Gundlach

gundlach@math.harvard.edu fabiangundlach.org/21-fall/288y

OH: Tentatively TuTh 2-3pm room 233

References:
- A Course in Computational Algebraic Number Theory,
  Cohen (1993) Pari
- Algorithmic Algebraic Number Theory,
  Pohst-Zassenhaus (1989) [later topics...]
- The Art of Computer Programming,
  Vol. 1 (Fundamental Algorithms)
  + Vol. 2 (Seminumerical Algorithms),
  Knuth (1973 + 1981) [earlier topics]

Implementation exercises:
projecteuler.net

HW ungraded

Final paper

some inefficient algorithms – how not to...

1) ... add integers? [Joke...]

Stupid schoolbook addition:

```
  16345825   } n digits
+  2659076   }
= 18994891
+    10010
= 18904801
+   100100
= 18004901
+  1000000
= 19004901
```

Worst case: $\Omega(n)$ lines, $\Omega(n^2)$ digits

```
  99...9
+      1
```

~~How not to ...~~

1) ... multiply polynomials $f, g \in \mathbb{F}_p[X]$
(= given the coeffs. of f, g, determine the coeffs of fg):

Schoolbook mult:

$$f = \sum_{i=0}^{n} a_i X^i, \quad g = \sum_{j=0}^{m} b_j X^j$$

$$\rightsquigarrow fg = \sum a_i b_j X^{i+j} = \sum_{k=0}^{n+m} c_k X^k$$

with $c_k = \sum_{\substack{i \le n,\\ j \le m:\\ i+j=k}} a_i b_j$.

The total number of summands in $c_0, \dots, c_{n+m}$ is $n \cdot m$.

But we can actually compute $c_0, \dots, c_{n+m}$ in "roughly linear" time $\mathcal{O}_{p,\varepsilon}((n+m)^{1+\varepsilon})$.

3) ... divide polynomials $f, g \in \mathbb{F}_p[X]$

(given the coeffs. of f, g, determine the coeffs. of $q, r \in \mathbb{F}_p[X]$

with $f = gq + r$, $\deg(r) < \deg(g)$):

"$r = f \bmod g$"

Schoolbook division:

let $n = \deg(f)$, $m = \deg(g)$

$h_n := f$

For $i = n, \ldots, m$, let

$$c_i := \frac{X^i\text{-coeff. of } h_i}{\underbrace{\text{leading}}_{X^m\text{-}} \text{ coeff. of } g},$$

$$h_{i-1} := h_i - c_i X^{i-m} \cdot g.$$

$\deg(h_n) = n \Rightarrow \deg(h_{n-1}) \le n-1 \Rightarrow \deg(h_{n-2}) \le n-2$

$\Rightarrow \ldots \Rightarrow \deg(h_{m-1}) \le m-1$

$$f = \underbrace{\left(\sum_{i=m}^{n} c_i X^{i-m}\right)}_{\text{quotient } q} \cdot g + \underbrace{h_{m+1}}_{\text{remainder}}$$

~~In the worst case, we~~

$\rightsquigarrow$ running time $\Theta_p((n-m)m)$.

But can be done in $O_p((n+m)^{1+\varepsilon})$.

⑥

4 B)... find the gcd of polynomials $f, g \in \mathbb{F}_p[X]$:

Euclidean algorithm:

$a_0 := f$

$a_1 := g$

$a_2 := a_0 \bmod a_1$

$a_3 := a_1 \bmod a_2$

$\vdots$

$a_{i+2} := a_i \bmod a_{i+1}$

$\vdots$

$a_k = \dots \neq 0$

$a_{k+1} = 0$

$\gcd(f,g) = a_k$

Thm There are pol. f, g of degrees $n, n-1$ such that

$\deg(a_i) = n-i$, $k = n$.

(~~So~~ $\sum \deg(a_i) = \Theta(n^2)$.)

Pf Work backwards:

$a_n := 1$, $a_{n-1} := X$,

$a_i := a_{i+2} + X \cdot a_{i+1}$ for $i = n-2, \dots, 0$.

$f := a_0$, $g := a_1$. □

But $\gcd(f,g)$ can be computed in $O_p((n+m)^{\varepsilon})$.

5) .. find $\gcd(f,g)$ for $f,g \in \mathbb{Q}[X]$:

W.l.o.g. $f,g \in \mathbb{Z}[X]$, $\deg(f) \geq \deg(g)$.

Euclidean algorithm:

$a_0 := f$

$a_1 := g$

$a_{i+1} := \left(lc(a_{i+1})^{\deg(a_{i-1}) - \deg(a_{i+1}) + 1} \cdot a_{i-1} \bmod a_{i+1}\right) \in \mathbb{Z}[X]$

$\vdots$

until

$a_{k+1} = 0$.

$\Rightarrow \gcd(f,g) = a_k$.

Thm For any (large) n, there are relatively prime pol. $f,g \in \mathbb{Z}(x)$ of degrees $n, n-1$ such that each coeff. of f,g has $\mathcal{O}(n)$ digits, $k=n$, and $a_n \in \mathbb{Z}$ has $\Omega\left((1+\sqrt{2})^n\right)$ digits.

Pf let $b_n = 1$, $b_{n-1} = X$,

$b_i = b_{i+1} + x \cdot b_{i+1}$ for $i = n-2, \dots, 0$.

$\Rightarrow \deg(b_{n-i}) = i$,

$lc(b_{n-i}) = 1$

$b_{i+1} = (b_{i-1} \bmod b_i)$

$(\max|\text{coeff}| \text{ of } b_i) \leq (\max|\text{coeff}| \text{ of } b_{i+1}) + (\max|\text{coeff}| \text{ of } b_{i+2})$

$\Rightarrow (\max|\text{coeff}| \text{ of } b_{n-i}) \leq \underbrace{F_i}_{i\text{-th Fibonacci nr.}} = \mathcal{O}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n\right)$

$\Rightarrow$ Each coeff. of $b_0, b_1 \in \mathbb{Z}(x)$ has $\mathcal{O}(n)$ digits.

Let $\underset{\substack{\shortparallel \\ a_0}}{f} = b_0$, $\underset{\substack{\shortparallel \\ a_1}}{g} = 2\cdot b_1$.

~~By induction, $a_i =$~~

~~Pf $a_i =$ [illegible] $2^{r_i}\cdot b_i$~~

Claim ~~By induction~~, $a_i = 2^{r_i}\cdot b_i$, where

$r_0 = 0$, $r_1 = 1$, $r_{i+2} = 2r_{i+1} + r_i$:

Pf by ind:

$$\{ a_{i+2} = \Big(\underbrace{lc(a_{i+1})}_{2^{r_{i+1}}}{}^{\overbrace{deg(a_i)-deg(a_{i+1})+1}^{}}\underbrace{}_{2} \cdot \underbrace{a_i}_{2^{r_i}\cdot b_i} \bmod \underbrace{a_{i+1}}_{2^{r_{i+1}}\cdot b_{i+1}}\Big)$$

$$= \left(2^{2r_{i+1}+r_i}\cdot b_i \bmod b_{i+1}\right)$$

$$= 2^{2r_{i+1}+r_i}\cdot b_{i+2}\}\ \}$$ □

We have

$r_i = \Theta((1+\sqrt{2})^i)$, so in particular

$a_n = 2^{r_n}$ has $\Theta((1+\sqrt{2})^n)$ digits. □

~~$x^2-2x-1=0$~~

~~$\frac{2\pm\sqrt{4+4}}{2} = 1\pm\sqrt{2}$~~

```
a_0 = 3 * X^10 + 5 * X^9 + 4 * X^8 + 5 * X^7 + 3 * X^6 + 3 * X^5 + 2 * X^4 + 2 * X^3 + 4 * X^2
+ 1 * X^1 + 5 * X^0

a_1 = 2 * X^9 + 1 * X^8 + 2 * X^7 + 2 * X^6 + 1 * X^5 + 2 * X^4 + 2 * X^3 + 2 * X^2 + 2 * X^1
+ 2 * X^0

a_2 = - 3 * X^8 - 6 * X^7 - 8 * X^6 - 7 * X^5 - 18 * X^4 - 18 * X^3 - 10 * X^2 - 22 * X^1 + 6
* X^0

a_3 = 24 * X^7 + 48 * X^6 - 36 * X^5 + 72 * X^4 + 120 * X^3 - 24 * X^2 + 252 * X^1 - 36 * X^0

a_4 = - 7200 * X^6 + 1152 * X^5 - 1728 * X^4 - 12096 * X^3 + 12384 * X^2 - 15264 * X^1 + 3456
* X^0

a_5 = - 1734856704 * X^5 + 997318656 * X^4 + 3845947392 * X^3 + 740524032 * X^2 + 7963619328 *
X^1 - 576294912 * X^0

a_6 = - 5840866074848919552000 * X^4 - 6558576496531734528000 * X^3 -
6603834622407081984000 * X^2 - 8001058891452383232000 * X^1 + 1338805978908327936000 * X^0

a_7 = 852945579880322241623233930750099591282745671680000000 * X^3 -
726961313894884840151011020364099725068284723200000000 * X^2 +
1204729930557464691762278379088570099784283586560000000 * X^1 +
337129438656679591598394665444340006494247321600000000 * X^0

a_8 =
7599594390054607432864709535942269852898406230064084364169061461682007532582839448414444538766
95863107772638822400000000000000000000 * X^2 +
1597984990559247565932903117430723401512089792652232663589046921980206485675730428013523181085
966664947478403481600000000000000000000 * X^1 +
1176911662883991329365765977135044627459270056704063379858621559618882588055956358178152519652
011378502811202355200000000000000000000 * X^0

a_9 =
2199225620576316007900622518688059723119305735435859826921836703987836014614104059681517753856
3381829903280441551180561269477705029821797035905418594215319029358766239513162464181051827585
6972144568026728845938149032562218719284479753663092673386683829137864617379107962880000000000
00000000000000000000000000000000000000 * X^1 +
1688614121495653978622562278233711274905897121117831871180008868805039630162777797292030539955
2393671611387989490394361183128380390047536071596103161250969147643560966003869086724263125016
3575836890896161081045250301578261940016006399274792020294525499580074073158613729280000000000
00000000000000000000000000000000000000 * X^0

a_10 =
1924858381375085554380884649278103377738530626880459029135908323517221387155732187052679568588
7604936439383600317487311931539136653013190708312248437208541914022842424353640180477203860662
9372067863039275501901013956187034256634147646812634864756586557568123083873624554651135464465
5374684073389602664610009707536732744625482978290307108652521848394025498070602400707690207644
1418892707917523417113618315801119512533274619448551153310997766354495952279233330066811543097
3188661003783107389578260436837168512780040760453682437452488700877070468178542434865223343441
1834672151608736518313706971328938069858572561202286370780995870900404456495652442870906880000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000 * X^0
```

6) ... find the roots in $\mathbb{F}_p$ of a pol. $f(X) \in \mathbb{F}_p[X]$ degree n:

For each $x \in \mathbb{F}_p$, check whether $f(x) = 0$.

Time $O(p\, n)$ even if you could do arithmetic in $\mathbb{F}_p$ in $O(1)$.

Can be done $O((\log p)^{\cdots} \cdot n \cdot (\log n)^{\cdots})$

with a nondeterministic alg. in expected time

7) ... find the number of primes $p \leq n$:

Use the sieve of Erathosthenes

Running time $O\left(\sum_{p \leq n} \frac{n}{p}\right) = O(n \log\log n)$

Some problems that have no "obvious" algorithm at all:

8) Factor pol. in $\mathbb{Q}[X]$.

9) Find the ring of integers, class group, unit grou of a number field $K = \mathbb{Q}[X]/f(X)$.

10) Find the Galois closure of a field ext. $L|K$ and the Galois group.

11) Find the dimension, number of irred. comp, ... of a variety $\{ P \in K^n \mid f_1(P) = \dots = f_m(P) = 0 \}$.

Some problems are undecidable:

12) Does a given pol. $f \in \mathbb{Z}[X_1, \dots, X_n]$ have a root $(x_1, \dots, x_n) \in \mathbb{Z}^n$?

Our favorite computational model:

b-bit random access machine (RAM)

2^b working registers $r_0, \ldots, r_{2^b-1}$ with values in $\{0, \ldots, 2^b-1\}$

2^b input $in_0, \ldots, in_{2^b-1}$ ——"——

2^b output $out_0, \ldots, out_{2^b-1}$ ——"——

A program consists of s steps of the following form:

- "$r_i := x$" $(0 \le i < 2^b)$, $0 \le x < 2^b$ " assign register i the value x, then go to the next step.
- "$r_i := r_j \overset{+}{\underset{\times}{-}} r_k$" $(0 \le i, j, k < 2^b)$ undef. if result $\notin \{0, \ldots, 2^b-1\}$
- "$r_i := \left\lfloor \frac{r_j}{r_k} \right\rfloor$" —— or $r_k = 0$
- "$r_i := r_j \bmod r_k$" ———
- "go to step t" $(1 \le t \le s)$
- "if $r_i = r_j$, go to step t, otherwise to step u"
  "if $r_i < r_j$ ——"—— "
- "$r_i := r_{r_j}$"
- "$r_{r_i} := r_j$"
- "$r_i := in_{r_j}$"
- "$out_{r_i} := r_j$"
- "halt"

Initially, all registers ~~are undef~~ could … except the appropriate input registers.

After the program halts, the output should be in the output registers.

Running time = total number of steps taken

Memory usage = smallest $m \geq 0$ s.t. only registers r_i, in_i, out_i with $0 \leq i < m$ were used (read from or written to).

Upshot: "It's the intuitive running time". ("On current computers, often $n=64$".)

We'll write pseudocode …

<u>Ex</u> There is a program which adds two binary integers with $\leq n$ digits in time $O(n)$, memory $O(n)$ on an $O(\log n)$-bit RAM.

Input: Encode $\sum_{i=0}^{n-1} a_i 2^i$, $\sum_{i=0}^{n-1} b_i 2^i$ as $n, a_0, \dots, a_{n-1}, b_0, \dots, b_{n-1}$

Output: Encode $\sum_{i=0}^{m-1} c_i 2^i$ as $m, c_0, \dots, c_{m-1}$.

<u>Ex</u> Instead of binary encoding, use base t. $\rightsquigarrow$ can add two integers with $\leq n$ base t digits in time $O(n)$, memory $O(n)$ on an $O(\max(\log n, \log t))$-RAM

Some other interesting models:

- Turing machines
- Multitape Turing machines
- Multiple$^{(k)}$ communicating RAM working in parallel

  (Running time = max. number of steps taken by any of the RAM

(Of course, if we can do sth. in time $O(T)$ on k parallel RAM, we can do it in time $O(kT)$ on one RAM. But the converse doesn't always hold!)

- Quantum RAM

⋮

1. Fast multiplication

Let R be a ring with unit (not necessarily commutative).

We'll assume the ring op. in R can be done in $\mathcal{O}(1)$ by augmenting the b-bit RAM:

Registers can take values in $\overbrace{\{0,\dots,2^b-1\}}^{\in \mathbb{Z}} \cup R$.

The op. "$r_i := r_j \overset{\pm}{\times} r_k$" apply if ~~both~~ both $r_j, r_k \in \{0,\dots,2^b-1\}$ or both $r_j, r_k \in R$.

There's an op. "$r_i :=$ image of $\underset{\overset{\cap}{\mathbb{Z}}}{r_j}$ under the hom. $\mathbb{Z} \to R$".

Question (for given n,) how quickly can we multiply two pol. $f, g \in R[X]$

~~$f(X) = \sum_{i=0}^{n-1} a_i X^i$, $g(X) = \sum_{i=0}^{n-1} b_i X^i$ [illegible]~~

of degree $\leq n$ on an $\mathcal{O}(\log n)$-bit RAM?

(Given coeff. of f, g, find coeff. of $f \cdot g$.)

Idea (Toom-Cook?) $\deg(fg) \leq 2n$.

$\Rightarrow$ ~~over a field~~ If R is a field, one can reconstruct fg from its value $(fg)(p_i) = f(p_i)g(p_i)$ at $2n+1$ (distinct) points $p_0, \dots, p_{2n} \in R$.

But how to compute $f(p_i), g(p_i)$ for $i = 0, \dots, 2n$? (evaluation)

And how to compute fg from these $2n+1$ values? (interpolation)

Idea 2 This is easier for powers $p_i = \zeta_k^i$ of a root of unity ζ_k (with $k > 2n$).

1.1. Fourier transform

Let $n \geq 1$ and assume that $\zeta_n \in R$ is ~~a primitive~~ ~~n-th root of unity~~ a root of the n-th cyclotomic polynomial $\Phi_n \in \mathbb{Z}[X]$ (def. recursively by $X^n - 1 = \prod_{d|n} \Phi_d(X)$).

the monic pol.

("a prim. n-th root of unity")

Lemma 1.1.1 a) $\zeta_n^n = 1$

b) For any $d|n$, ζ_n^d is a root of $\Phi_{n/d}$.

c) For any $a \in \mathbb{Z}$,

$$\sum_{i \in \mathbb{Z}/n\mathbb{Z}} \zeta_n^{ai} = \begin{cases} 0, & a \not\equiv 0 \bmod n \\ n, & a \equiv 0 \bmod n. \end{cases}$$

Def The Fourier transform of $a=(a_i)_{i\in\mathbb{Z}/n\mathbb{Z}} \in \prod_i R$ (w.r.t. ζ_n)

is $\mathcal{F}_{\zeta_n}(a) = (b_j)_{j\in\mathbb{Z}/n\mathbb{Z}} \in \prod_i R$, where

$$b_j = \sum_{i\in\mathbb{Z}/n\mathbb{Z}} a_i \zeta_n^{ij} .$$

Lemma 1.1.2

$$\mathcal{F}_{\zeta_n}(\mathcal{F}_{\zeta_n}(a)) = (n\cdot a_{-i})_{i\in\mathbb{Z}/n\mathbb{Z}} .$$

Pf $\mathcal{F}_{\zeta_n}(a)=b$ with $b_j = \sum_i a_i \zeta_n^{ij}$

$\mathcal{F}_{\zeta_n}(b)=c$ with $c_k = \sum_j b_j \zeta_n^{jk} = \sum_{i,j} a_i \underbrace{\zeta_n^{(i+k)j}}_{\sum_j = n \text{ if } i+k=0,\ 0 \text{ otherwise}} = n a_{-i} .$ $\square$

Cor 1.1.3 If n is invertible in R, the problem of evaluating at roots of unity is equiv. to the problem of interpolating from roots of unity.

~~**Complx** The triv. alg. to compute $\mathcal{F}_{\zeta_n}(a)$ (given a, ζ_n) needs time~~

Question "given" $a=(a_i)_i$, and ζ_n, how quickly can we compute $\mathcal{F}_{\zeta_n}(a)$?

~~$\Rightarrow c_k = n \cdot a_{-k \bmod n}$ for where $i \equiv -k \bmod n$.~~ $\square$

~~Lemma~~ Thm 1.1.3 (Cooley-Tukey '1965, Gauß)

Let $n = pq$ for integers $p, q \geq 1$. Let $\zeta_p = \zeta_n^q$, $\zeta_q = \zeta_n^p$.

~~If ζ_n is a root of Φ_n, then $\zeta_p := \zeta_n^q$ is a root of Φ_p and $\zeta_q := \zeta_n^p$ is a root of Φ_q.~~

We can reduce the problem of computing a length n FT to computing p length q FT:

Let $a \in R^n$.

For $l = 0, \dots, p-1$, let $b^{(l)} :=$ ~~$\mathcal{F}_{\zeta_q}(a_{ql}, a_{ql+1}, \dots, a_{q(l+1)-1})$~~

$$\mathcal{F}_{\zeta_q}(a_l, a_{p+l}, \dots, a_{(q-1)p+l})$$

Then, $\mathcal{F}_{\zeta_n}(a) = (c)_{j \in \mathbb{Z}}$ where $c_j = \sum_{l=0}^{p-1} b_j^{(l)} \zeta_n^{lj}$ for $j \in \mathbb{Z}/n\mathbb{Z}$.

Pf

$$b_j^{(l)} = \sum_{i=0}^{q-1} a_{ip+l} \zeta_q^{ij}$$

$$= \sum_{i=0}^{q-1} a_{ip+l} \zeta_n^{ipj}$$

~~$\Rightarrow \sum_l b_j^{(l)} = \sum_{i=0}^{n-1} a_i \zeta_q^{ij}$~~

$$\Rightarrow \sum_{l=0}^{p-1} b_j^{(l)} \zeta_n^{lj} = \sum_{l=0}^{p-1} \sum_{i=0}^{q-1} a_{ip+l} \zeta_n^{(ip+l)j}$$

$$\underset{\uparrow}{=} \sum_{k \in \mathbb{Z}/n\mathbb{Z}} a_k \zeta_n^{kj}$$

any $k \in \mathbb{Z}/n\mathbb{Z}$ can be written uniquely as $k = ip + l$

$\square$

Cor 1.1.4 (Radix r Cooley-Tukey ^{FFT} alg.)

Let $r \geq 2$, $e \geq 0$, $n = r^e$.

Then, you can compute $\mathcal{F}_{\zeta_n}(a)$ with $\mathcal{O}(r^{e+1}(e+1))$
$= \mathcal{O}(n \cdot r \cdot (\log_r n+1))$

add./mult. op. in R.

(Think of r as fixed, usually $r=2$. $\rightsquigarrow$ time $\mathcal{O}(n \log n)$ for large n)

Alg ~~Apply the C-T Thm recursively, with $p=r$, $q=r^{e-1}$~~

Idea: Apply the C-T Thm. (recursively) with $p=r$, $q=r^{e-1}$.

To compute $\mathcal{F}_{\zeta_{r^e}}(a)$:

1) For $l = 0, \dots, r-1$:
   Recursively compute $b^{(l)} = \mathcal{F}_{\zeta_{r^{e-1}}}(a_l, a_{r+l}, \dots, a_{(r^{e}-1)r+l})$. (time $\leq C \cdot r^{e} \cdot e$ by induction)

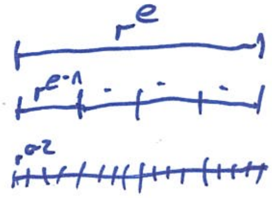
2) Compute $1, \zeta_n, \dots, \zeta_n^{n-1}$. (time $\mathcal{O}(n) = \mathcal{O}(r^e)$)

3) For $j = 0, \dots, r^e - 1$: Compute $c_j = \sum_{l=0}^{r-1} b_j^{(l)} \zeta_n^{lj}$. (time $\mathcal{O}(r)$)

(time $\mathcal{O}(r^{e+1})$)

(time $\leq C \cdot r^{e+1} \cdot$

4) Return $\mathcal{F}_{\zeta_{r^e}}(a) = (c_j)_j$.



~~to~~tal time

$C r^{e+1} \cdot e + \mathcal{O}(r^{e+1})$

$\leq C r^{e+1}(e+1)$

if $C \geq$ the constant in $\mathcal{O}(\dots)$. $\square$

Rmk ~~These~~

~~The algorithm only multipl~~

All multiplications in R performed in the alg. are mult. by powers of ζ_n.

def The <u>product</u> of — — — is $a \cdot b := c\,c \ldots$,
where $c_k := a_k b_k$.

The <u>convolution</u> of $(a_i)_i, (b_i)_i \in \prod_{i \in \mathbb{Z}/n\mathbb{Z}} R$ is $\{a * b\} := (c_i)_i \in \prod_{i \in \mathbb{Z}/n\mathbb{Z}} R$,

where $c_k := \sum_{\substack{i,j \in \mathbb{Z}/n\mathbb{Z}: \\ i+j=k}} a_i b_j$.

<u>**Lemma 1.1.6**</u> Assume that ζ_n lies in the center of R (commutes with every $x \in R$).

a) $\mathcal{F}_{\zeta_n}(a * b) = \mathcal{F}_{\zeta_n}(a) \cdot \mathcal{F}_{\zeta_n}(b)$

b) ~~$\mathcal{F}_{\zeta_n}(a) * \mathcal{F}_{\zeta_n}(b) =$~~

$n \cdot \mathcal{F}_{\zeta_n}(a \cdot b) = \{\mathcal{F}_{\zeta_n}(a) * \mathcal{F}_{\zeta_n}(b)\}$.

<u>Pf</u> a) Let $c = a * b$.

$$\underset{\text{LHS}}{\sum_k c_k \zeta_n^{kl}} = \sum_{i,j} a_i b_j \zeta_n^{(i+j)l} = \underset{\text{RHS}}{\Big(\sum_i a_i \zeta_n^{il}\Big)\Big(\sum_j b_j \zeta_n^{jl}\Big)}.$$

b)

$$\text{RHS} = \sum_{\substack{r,s: \\ r+s=l}} \Big(\sum_i a_i \zeta_n^{ir}\Big)\Big(\sum_j b_j \zeta_n^{js}\Big)$$

$$= \sum_{i,j} a_i b_j \sum_{\substack{r,s: \\ r+s=l}} \zeta_n^{ir+js}$$

$$= \sum_{i,j} a_i b_j \underbrace{\sum_r \underbrace{\zeta_n^{ir+j(l-r)}}_{\zeta_n^{jl+(i-j)r}}}_{= \begin{cases} n \cdot \zeta_n^{il} & \text{if } i=j \pmod n \\ 0 & \text{else} \end{cases}}$$

$$= n \cdot \sum_i a_i b_i \zeta_n^{il} = \text{RHS}$$

1.2. Multiplying polynomials

Thm 1.2.1 Let $r \geq 2$, and n large. Let $k = \lceil \log_r n \rceil$ (so $r^k \geq n$), $t = r^{k+1}$. If r is invertible in R and R contains a root $\zeta = \zeta_t$ of $\Phi_t(X)$, then (given f, g) we can multiply any two pol. $f, g \in R[X]$ of degrees $< n$ in time $O_r(n \log n)$ on an $O(\log n)$-bit RAM.

Alg Let $f(X) = \sum_{i=0}^{t-1} a_i X^i$, $g(X) = \sum_{i=0}^{t-1} b_i X^i$.

Write $a = (a_i)_i \in \prod_{i \in \mathbb{Z}/t\mathbb{Z}} R$, $b = (b_i)_i$

1) Use radix r Cooley-Tukey to compute the FT

$$\hat{a} := \mathcal{F}_\zeta(a), \quad \hat{b} := \mathcal{F}_\zeta(b).$$

2) Compute $\hat{a} \cdot \hat{b}$:

For each $j \in \mathbb{Z}/t\mathbb{Z}$, compute $\hat{a}_j \cdot \hat{b}_j$.

3) Use C-T to compute

$$c := \mathcal{F}_\zeta(\hat{a} \cdot \hat{b}).$$

4) Return $\underbrace{\frac{1}{t}}_{\frac{1}{r^{k+1}}} \cdot \sum_{i=0}^{t-1} c_i X^i.$

Pf correctness

$$c = \mathcal{F}(\hat{a}\cdot\hat{b}) = \mathcal{F}(\mathcal{F}(a)\cdot\mathcal{F}(b))$$

$$\underset{\substack{\uparrow\\ \boxed{\text{Lemma 1.1.6 a)}}}}{=} \mathcal{F}(\mathcal{F}(a*b))$$

$$\underset{\substack{\uparrow\\ \boxed{\text{Lemma 1.1.2}}}}{=} t\cdot(a*b)$$

For $0\le k<t$,

$$\Rightarrow \frac{1}{t}\cdot c_k = (a*b)_k = \sum_{\substack{i,j\in\mathbb{Z}/t\mathbb{Z}:\\ i+j=k}} a_i b_j = \sum_{\substack{0\le i,j<t:\\ i+j\equiv k \bmod t}} a_i b_j$$

$$\underset{\substack{\uparrow\\ \boxed{\substack{a_i,\, b_i=0\\ \text{unless } 0\le i<n\le r^k<\frac{1}{2}\cdot r^{k+1}=\frac{1}{2}\cdot t}}}}{=} \sum_{\substack{0\le i,j<t:\\ i+j=k}} a_i b_j .$$

$$\Rightarrow \frac{1}{t}\cdot\sum_{k=0}^{t-1} c_k x^k = \sum_{i,j} a_i b_j x^{i+j} = f(x)\cdot g(x).$$

Running time

Step 1) $O_r(n\log n)$

2) $O(n)$

3) $O_r(n\log n)$

4) $O(n)$.

$\square$

How to get rid of the assumption that Φ_t has a root in R?

Idea 1 Work in the ring $S = R[Y]/\Phi_t(Y)$.

$\rightsquigarrow$ $\zeta_t := [Y] \in S$ is a root of Φ_t.

Problem: Adding two el. of S takes time $\Theta(\deg(\Phi_t)) = \Theta(n)$.

In C-T, we do $\Theta(n \log n)$ such additions.

$\rightsquigarrow$ total time $\Theta(n^2 \log n)$, worse than schoolbook multiplication!

Thm 1.2.2 (Schönhage–Strassen)

Let r be a prime number. For large n, ~~if you can~~ ~~multiply~~ given two pol. $f, g \in R(X]$ of degree $< n$, you can compute $r^{k+2} \cdot fg$ in time $\mathcal{O}(n \log n \log\log n)$ on an $\mathcal{O}(\log n)$-bit RAM, where $k = \lceil \frac{1}{2} \log_r n \rceil$.

Cor 1.2.3 You can compute $f\cdot g$ in time $O(n\log n\log\log n)$.

[Clear if r is invertible in R (and its inverse known).]

Pf Apply the Thm with $r=2,3$.

~~Since 2^{k_2}~~

$\rightsquigarrow$ can compute $2^{r_2+2}\cdot fg$, $3^{r_3+2}\cdot fg$

for ~~some $r_2, r_3 = O(\log n)$~~. $r_2=\lceil \frac{1}{2}\log_2 n\rceil$, $r_3=\lceil \frac{1}{2}\log_3 n\rceil$.

Since 2^{k_2+2}, 3^{k_3+3} are relatively prime,

~~we can write~~ there exist $u, v\in\mathbb{Z}$ such that

$$1=u\cdot 2^{k_2+2}+v\cdot 3^{k_3+2}$$

(and $0\le u<3^{k_3+2}=3^{\frac{1}{2}\log_3 n+O(1)}=O(\sqrt{n})$).

You can find u,v by trying all $0\le u<3^{k_3+2}$ in time $O(\sqrt{n})$. (Or use the extended Euclidean algorithm.)

Then, $f\cdot g= u\cdot(2^{r_2+2}\cdot fg)+v\cdot(3^{r_3+2}\cdot fg)$.

$\square$

RETURN

Alg for Thm 1.2.2

If $k \leq 3$, use the schoolbook algorithm.
Otherwise:

Let $m = r^k$, $t = r^{k+2}$

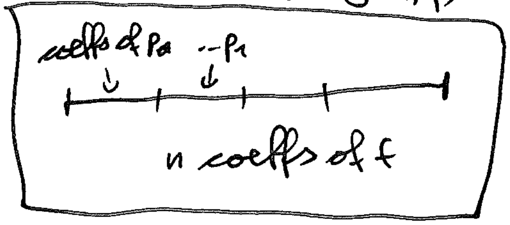
$$m \overset{}{=} \underset{\theta_r(\sqrt{n})}{\overset{\|}{r^k}}, \qquad t = \underset{\theta_r(\sqrt{n})}{\overset{\|}{r^{k+2}}}$$

1) Write $f(X) = \sum_{i=0}^{t-1} p_i(X) \cdot X^{i \cdot m}$

with $\deg(p_i) < m$ (possible because $m \cdot t = r^{2k+2} > r^{2k} \geq n$).

Similarly, $g(X) = \sum_{i=0}^{t-1} q_i(X) \cdot X^{i \cdot m}$

with $\deg(q_i) < m$.



Let $S = \mathbf{R}[Y]/\Phi_t(Y)$ and let $\zeta := \zeta_t := [Y] \in S$.
We have $\Phi_t(Y) = \frac{Y^{r^{k+2}}-1}{Y^{r^{k+1}}-1} = 1 + Y^{r^{k+1}} + \ldots + Y^{(r-1)r^{k+1}}$.
Let $a = (a_i)_i \in \prod_{i \in \mathbb{Z}/t\mathbb{Z}} S$ with $a_i = \underbrace{[p_i(Y)]}_{p_i(Y) \bmod \Phi_t(Y)} \in S$,

$b = (b_i)_i$ $\qquad b_i = [q_i(Y)] \in S$.

(Note that $\deg(p_i), \deg(q_i) < m = r^k < (r-1)r^{k+1} = \deg(\Phi_t)$, so p_i, q_i are already reduced mod Φ_t.)

2) Use radix r Cooley-Tukey to compute the FT

$$\hat{a} = \mathcal{F}_\zeta(a) \in \prod_j S, \qquad \hat{b} = \mathcal{F}_\zeta(b) \in \prod_j S.$$

In the C-T alg., we have to add elements of S and multiply el. of S by powers of $\zeta = [Y] \in S$. We do this by working in the ring

$$S' = \mathbf{R}[Y]/(Y^t - 1)$$

and only reducing modulo $\Phi_t(Y)$ (which divides $Y^t - 1$) in the end.

Addition in S': $$\underbrace{\sum_{d=0}^{t-1} u_d Y^d}_{\text{red. mod } Y^t-1} + \underbrace{\sum_{d=0}^{t-1} v_d Y^d}_{\dots} = \underbrace{\sum_{d=0}^{t-1} (u_d + v_d) Y^d}_{\dots}.$$

Mult. by powers of Y: $$\underbrace{\left(\sum_{d=0}^{t-1} u_d Y^d\right)}_{\dots} \cdot Y^l \equiv \sum_{d=0}^{t-1} u_d Y^{d+l}$$

$$\equiv \underbrace{\sum_{d=0}^{t-1} u_d Y^{(d+l) \bmod t}}_{\text{reduced mod } Y^t-1} \bmod Y^t - 1.$$

3) ~~Compute $(\hat{a}\cdot\hat{b})_i$~~

For all $j \in \mathbb{Z}/t\mathbb{Z}$, compute $\hat{a}_j \cdot \hat{b}_j \in S$ as follows:

Let $\hat{a}_j = [A_j] \in S$, $\hat{b}_j = [B_j] \in S$

with $A_j, B_j \in R[Y]$ $\deg(A_j), \deg(B_j) < \deg(\Phi_t) = (r-1)\cdot r^{k+1} < r^{k+2}$.

$\overset{k\geq 4}{\leq} r^{2k-2} < n$.

a) Recursively apply the mult. alg. to compute $A_j(Y)\cdot B_j(Y) \in R[Y]$.

b) Reduce $A_j(Y)\cdot B_j(Y) \bmod \Phi_t(Y) = 1 + Y^{r^{k+1}} + \ldots + Y^{(r-1)r^{k+1}}$ using the schoolbook algorithm.

4) Use Cooley-Tukey (like before) to compute the FT

$$c = \mathcal{F}_\zeta(\hat{a}\cdot\hat{b}) \in \prod_{i\in\mathbb{Z}/t\mathbb{Z}} S.$$

5) Let $c_i = [C_i] \in S$ with $C_i \in R[Y]$, $\deg(C_i) < \deg(\Phi_t)$.

Return $\sum_{i=0}^{t-1} C_i(X)\cdot X^{im}$ $\quad (= t\cdot f(X)\cdot g(X))$.

1.3. Multiplying integers

Thm 1.3.1 We can multiply two {binary} integers x, y with $\leq n$ digits in time $O(n)$ on an $O(\log n)$-bit RA
~~We can mult two base 2^k integers x, y with n digits in time O~~

Alg (sketch) W.l.o.g. $x, y \geq 0$.

W.l.o.g. $n = 2^k \cdot k$ with $k \geq 1$. $(\Rightarrow k = \Theta(\log n))$.

Write x, y in base 2^k:

$$x = \sum_{i=0}^{2^k-1} a_i 2^{ki}, \quad y = \sum b_i 2^{ki},$$

$0 \leq a_i, b_i < 2^k$.

Let $R = \mathbb{C}$, $t = 2^{k+1}$, ζ_t any prim. t-th root of unity.

By Thm 1.2.1, we can compute

$$c_\ell := \sum_{\substack{i,j \\ i+j=\ell}} a_i b_j \quad \text{for } k = 0, \dots, 2^{k+1}-1$$

in time $O(2^k \cdot k) = O(n)$, assuming the operations in $\mathbb{C}$ can be done in time $O(1)$. It turns out that it suffices to do the computations in $\mathbb{C}$ with precision $O(n)$ (rounding intermediate results to $O(n)$ digits) and round the result c_k to the nearest integer. (These can be done in $O(1)$ on an $O(\log n)$-bit RAM)

cf. Schönhage-Strassen 'Schnelle Mult. großer Zahlen'

Now $x \cdot y = \sum_{\ell=0}^{2^{k+1}-1} c_\ell \cdot 2^{k\ell}$, where $0 \leq c_\ell \leq (\ell+1) \cdot 2^k \cdot 2^k \leq 2^{3k+1}$ $k \geq 1$ has at most 4 digits in base 2^k.

You can add these 2^{k+1} integers with $\mathcal{O}(1)$ nonzero digits in time $\mathcal{O}(2^{k+1})$.

Rmk 1.3.2 Harvey and van Hoeven recently showed that you can multiply two binary int. x, y with $< n$ digits in time $O(n \log n)$ on a multitape Turing machine. This is conjectured to be optimal.

[Their algorithm also uses FFT and several ingenious tricks!]

REFERENCE: Fast multiplication and its applications
Daniel J. Bernstein

2. Quotients

Let K be a field and assume $\pm, \times, \cdot^{-1}$ in K and the image of an integer under the hom. $\mathbb{Z} \to K$ can be computed in $\mathcal{O}(1)$.

2.1. ~~[illegible] power series~~ Mult. inverse

~~Def $K[[X]] := \{f = \sum_{n=0}^{\infty} a_n X^n \mid a_0, a_1, \dots \in K\}$ ring of power series~~

~~Prob $K[[X]]^\times = \{f = \sum a_n X^n \mid a_0 \neq 0\}$.~~

[Assume that we can multiply two pol. $f, g \in K[X]$ of degree $< n$ in time $\mathcal{O}(\mu(n))$, where $\mu(n) \geq n$, $\mu(n+m) \geq \mu(n) + \mu(m)$. (We've shown that $\mu(n) = n \log n \log\log n$ works for large n.)

Thm 2.1 Let $f \in (K[X]/X^n)^\times$ with $f \equiv a_0 + a_1 X + \dots + a_{n-1} X^{n-1} \bmod X^n$. ($\Rightarrow a_0 \neq 0$)

We can compute $f^{-1} \pmod{X^n}$ $(= b_0 + \dots + b_{n-1} X^{n-1})$ in time $\mathcal{O}(\mu(n))$ on an $\mathcal{O}(\log n)$-bit RAM.

Alg W.l.o.g. $n = 2^k$, $k \geq 1$.

Recursively compute $g := (f^{-1} \bmod X^{2^{k-1}})$.

Return $h := (2 - fg) g \bmod X^{2^k}$.

Pf by induction. $g \equiv f \bmod X^{2^{k-1}}$.

$\Rightarrow fg \equiv 1 \bmod X^{2^{k-1}}$. $\Rightarrow fg$

$\Rightarrow fh \equiv (2 - fg) \cdot fg \bmod X^{2^{k-1}}$

$\Rightarrow 1 - fh \equiv (1 - fg)^2 \equiv 0 \bmod X^{2^k}$.

Total time: $\ll \mu(2^k) + \underbrace{\mu(2^{k-1})}_{\leq \frac{1}{2}\mu(2^k)} + \dots + \underbrace{\mu(1)}_{\leq \frac{1}{2^k}\mu(2^k)} \leq 2\mu(2^k) \ll \mu(n)$.

□

Rmk This is Newton's approximation alg. for the function

$$\varphi(t) = \frac{1}{t} - f.$$

$$\rightsquigarrow t - \frac{\varphi(t)}{\varphi'(t)} = t - \frac{\frac{1}{t} - f}{-\frac{1}{t^2}} = t + (t - ft^2) = (2 - ft)t^2$$

Rmk The same algorithm can be used to invert an element k of $(\mathbb{Z}/p^n\mathbb{Z})^\times$ ~~integer~~ $\left(x = \sum_{n=0}^{\infty} a_n \cdot p^n, \; a_0, a_1, \ldots \in \{0, \ldots, p-1\}, a_0 \neq 0\right)$.

(Just replace X by p everywhere!)

Rmk ~~It can also be used~~
Similarly, Newton's method can be used to find the inverse of a real number $k \in \mathbb{R}$, given its leading $O(n)$ digits in time $O(n^2)$.

(up to a relative error of $O(2^{-n})$)

2.2. Quotient and remainder

Thm 2.2.1 given pol. $f, g \in K[X]$ of degree $< n$ (with $g \neq 0$), we can compute the quotient $\underset{\lfloor \frac{f}{g} \rfloor}{\underset{\shortparallel}{q}} \in K(X)$ and remainder $\underset{f \bmod g}{\underset{\shortparallel}{r}} \in K(X)$

(such that $f = gq + r$, $\deg(r) < \deg(g)$) in time $O(\mu(n))$ on an $O(\log n)$-bit RAM.

Pf Let $f(X) = X^{u} \cdot \tilde{f}(\frac{1}{X})$, $g(X) = X^{v} \cdot \tilde{g}(\frac{1}{X})$,
$\tilde{f}, \tilde{g} \in K[Y]$, $\tilde{f}(0), \tilde{g}(0) \neq 0$.

(If $f(X) = a_u X^u + \dots + a_0$, then $\tilde{f}(Y) = a_u + a_{u-1}Y + \dots + a_0 Y^u$.)
$(a_u \neq 0)$

~~Compute~~ w.l.o.g. $u \geq v$. (Otherwise, $q = 0$, $r = f$.)

~~Let $h \in K(Y)$ s.t. $\tilde{g} h(Y) \equiv 1 \bmod X^{u-v+1}$.~~

~~Let $i := \tilde{f}(Y) \cdot h(Y) \bmod Y^{u-v+1}$.~~
~~$\Rightarrow \tilde{g}(Y) i(Y) \equiv \tilde{f}(Y) \bmod Y^{u-v+1}$~~
~~Then,~~

~~$q(X) = X^{u-v} \cdot i(\frac{1}{X}) \in K[X]$~~

~~and $g(X) q(X) = X^u \tilde{g}(\frac{1}{X}) i(\frac{1}{X})$~~

Let $\tilde{q}(Y) = (\tilde{f}(Y) \cdot \tilde{g}(Y)^{-1} \bmod Y^{u-v+1})$. (This can be computed in $O(\mu(n))$ because products and inverses can.)

Then, $q(X) := X^{u-v} \cdot \tilde{q}(\frac{1}{Y})$ is the quotient pol:

- It's a polynomial because $\deg(\tilde{q}) \leq u-v$.
- Since $Y^u(f(\frac{1}{Y}) - g(\frac{1}{Y}) q(\frac{1}{Y})) = \tilde{f}(Y) - \tilde{g}(Y)\tilde{q}(Y)$ is divisible by Y^{u-v+1} in $K[Y]$, we have $\deg(f - gq) \leq v-1$.

$r := f - gq$ can also be computed in $O(\mu(n))$. □

A similar argument over $\mathbb{R}$ shows:

<u>Thm 2.2.2</u> ~~we can~~ For (binary) integers x, y with $< n$ bits, $(y \neq 0)$ we can compute $q = \left\lfloor \frac{x}{y} \right\rfloor$ and $r = (x \bmod y)$ in $O(n)$...

Pf It suffices to compute $\frac{x}{y} \in \mathbb{R}$ to absolute precision 1, so relative precision $\sim 2^{-n}$.

This leaves just ≤ 3 integers q to try.

$\square$

3. Greatest common divisor

Recall the Euclidean algorithm:

$a_0 = f$

$a_1 = g$

$a_{i+2} = a_i \bmod a_{i+1} = a_i - \left\lfloor \frac{a_i}{a_{i+1}} \right\rfloor \cdot a_{i+1}$ until $a_{k+1} = 0$.

$\Rightarrow \gcd(f,g) = a_k$.

Let $q_i = \left\lfloor \frac{a_i}{a_{i+1}} \right\rfloor$.

$$\Rightarrow \begin{pmatrix} a_{i+1} \\ a_{i+2} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix}}_{M_i} \begin{pmatrix} a_i \\ a_{i+1} \end{pmatrix} \quad \Rightarrow \begin{pmatrix} a_i \\ a_{i+1} \end{pmatrix} = M_{i-1} \cdots M_0 \begin{pmatrix} f \\ g \end{pmatrix}$$

until $\begin{pmatrix} \gcd(f,g) \\ 0 \end{pmatrix} = M_{k-1} \cdots M_0 \begin{pmatrix} f \\ g \end{pmatrix}$

Rmk $\deg(q_i) = \deg(a_i) - \deg(a_{i+1})$

$$\sum_i \deg(q_i) = \deg(f) - \deg(\gcd(f,g)) \leq \deg(f),$$

so at least the total number of coefficients in the pol. q_i is linear (unlike the total number of coeff. in the pol. a_i).

~~Rough idea: To compute a matrix $M \in GL_2(K[X])$ with $\det(M) = \pm 1$~~

~~and such~~

Rmk If $M \in GL_2(K[X])$ is a matrix with $\det(M) = \pm 1$ and such that $M \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} h \\ 0 \end{pmatrix}$, then $\gcd(f,g) = h$.

Pf Let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. $\Rightarrow$ $h = af + bg$, $0 = cf + dg$ $\Rightarrow \gcd(f,g) \mid h$.

On the other hand, $dh = adf + bdg = (\det(M) + bc)f + bdg$
$= \pm f + b(cf + dg) = \pm f$,

so $h \mid f$.

Similarly, $h \mid g$. □

~~Idea Let $\deg(f), \deg(g) < n$.~~

~~Idea 1 [illegible] To compute a matrix $M \in GL_2(K[x])$ with $\det(M) = \pm 1$ and such that $M\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} h \\ r \end{pmatrix}$ for some r with $\deg(r) <$ [illegible] $n - k$, we only need to know at most the top $2k$ coefficients of f and g.~~

Idea ~~[illegible]~~, Recursively find ~~better and~~ approximations to M: matrices M' s.t. $\det(M') = \pm 1$ and

$M'\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} t \\ r \end{pmatrix}$ for some pol. t, r with [illegible] smaller and smaller $\deg(t)$ (starting with $M' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, where $r = g$, and finishing with $M' = M$, where $r = 0$.)

Lemma 3.1 Let $f, g \in K[x]$, $\deg(f), \deg(g) \leq n$ and let $k \geq 1$ with $s := n - 2k \geq 0$. Let $M \in GL_2(K[x])$ s.t. $M = \begin{pmatrix} \deg \leq k & \deg \leq k \\ \deg \leq k & \deg \leq k \end{pmatrix}$

and $M \begin{pmatrix} \lfloor f/x^s \rfloor \\ \lfloor g/x^s \rfloor \end{pmatrix} = \begin{pmatrix} * \\ * \text{ of deg} < \underset{(n-s-k)=k}{} \end{pmatrix}$.

Then,

$$M \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} * \\ * \text{ of deg.} < n-k \end{pmatrix}.$$

(Moral: To find M s.t. the lower entry has degree $< n-k$, we only need the top $2k$ coefficients of f, g.)

Pf

$$M \begin{pmatrix} f \\ g \end{pmatrix} = M \begin{pmatrix} x^s \cdot \lfloor f/x^s \rfloor + (f \bmod x^s) \\ \cdots \end{pmatrix}$$

$$= \underbrace{x^s \cdot \underbrace{M \begin{pmatrix} \lfloor f/x^s \rfloor \\ \lfloor g/x^s \rfloor \end{pmatrix}}_{\begin{pmatrix} * \\ \deg < n-s-k \end{pmatrix}}}_{\begin{pmatrix} * \\ \deg < n-k \end{pmatrix}} + \underbrace{\underbrace{M}_{\deg \leq k} \underbrace{\begin{pmatrix} f \bmod x^s \\ g \bmod x^s \end{pmatrix}}_{\deg < s}}_{\deg < k+s = n-k}$$

$\square$

Thm 3.2 (Schönhage, using ideas of Lehmer, Knuth) Let $n \geq k \geq 1$. For polynomials $f, g \in K[X]$ of degree $\leq n$, you can find a matrix $M \in \underset{\uparrow}{GL_2^{\pm 1}}(K[X])$ whose entries have

$\{M \in GL_2 : \det(M) = \pm 1\}$

degree $\leq k$ and such that $M\binom{f}{g} = \binom{a}{b}$ for some pol. $a, b \in K[X]$ with $\deg(b) \leq n-k-1$, in time $O(n + \mu(k)\log k)$ (for large k) on an $O(\log n)$ bit RAM.

Cor 3.3 (Fast extended Euclidean algorithm) You can find the gcd^a of polynomials $f, g \in K[X]$ of degree $\leq n$, and pol. c, d s.t. $a = cf + dg$ in time $O(\mu(n)\log n)$ (for large n).

Pf of Cor Apply the Thm with $k = n$.

Then, $\underbrace{M\binom{f}{g}}_{\substack{\text{computable}\\ \text{in } O(\mu(n))}} = \binom{a}{0}$, so $\gcd(f, g) = a$. ∎

If $M = \begin{pmatrix} c & d \\ * & * \end{pmatrix}$, then $a = cf + dg$. □

Cor 3.4 If $\gcd(f,g) = 1$, you can find the inverse of g mod f in time $O(\mu(n)\log n)$...

Pf of Cor $dg \equiv \underset{\substack{\uparrow \\ \text{constant pol.} \neq 0}}{a} \mod f$. □

Alg for Thm 3.2 (Strassen)

W.l.o.g. $k \leq n$. (Otherwise, replace k by n.)

If $n=k=0$, it's easy: Take $M=\begin{pmatrix} 0 & 1 \\ 1 & -f/g \end{pmatrix}$.

(so f, g are constant polynomials)

If $n > 2k$, we can, according to Lemma 3.1, replace n by $2k$, f by $\lfloor f/x^{n-2k} \rfloor$, g by $\lfloor g/x^{n-2k} \rfloor$. ~~n by 2k~~

Assume $1 \leq k \leq n \leq 2k$. Let $k' = \lfloor \frac{k}{2} \rfloor$.

1) Recursively apply the alg. to find M_1 s.t.

$$\underbrace{M_1 \begin{pmatrix} f \\ g \end{pmatrix}}_{\text{computable in } \mathcal{O}(k)} = \begin{pmatrix} c \\ d \end{pmatrix} \text{ with } \deg(d) \leq n-k'-1.$$

Time $\leq C \cdot \mu(\frac{k}{2}) \log_2(\frac{k}{2})$
$\leq \frac{C}{2} \mu(k)(\log_2(k)-1)$

2) If $d=0$, we're done. Otherwise:

~~$M_2 M_1 \begin{pmatrix} f \\ g \end{pmatrix}$~~

Let $M_2 = \begin{pmatrix} 0 & 1 \\ 1 & -\lfloor c/d \rfloor \end{pmatrix}$. $\Rightarrow M_2 M_1 \begin{pmatrix} f \\ g \end{pmatrix} = M_2 \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} d \\ c \bmod d \end{pmatrix}$

both have $\deg \leq n-k'-1$

Time $\mathcal{O}(\mu(k))$

3) Recursively apply the alg. to find M_3 s.t.

$$M_3 M_2 M_1 \begin{pmatrix} f \\ g \end{pmatrix} = M_3 \begin{pmatrix} d \\ c \bmod d \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \text{ with } \deg(b) \leq (n-k'-1)-k'-1$$
$$\leq n-k-1.$$

Time $\leq \frac{C}{2} \cdot \mu(k)(\log_2(k)-1)$

Total: $C\mu(k)(\log_2 k-1) + \mathcal{O}(\mu(k))$
$\leq C\mu(k)\log_2 k$ for suff. large C. $\square$

Rmk The algorithm also computes the quotient polynomials q_i that arise in the Euclidean alg. s.t.

$$M = \begin{pmatrix} 0 & 1 \\ 1 & q_r \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & q_1 \end{pmatrix}.$$

$$\frac{f}{g} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 \ddots q_r}}$$

Thm 3.5 (Knuth–Schönhage)

For binary integers x, y with $\le n$ bits, we can compute $a = \gcd(x,y)$, and c, d s.t. $a = cx + dy$ in time $O(n \log n)$ on an $O(\log n)$-bit RAM.

Pf1 Similar to polynomials, replacing $\deg(f)$ by $\log|f|$.

It's more complicated:

Polynomials satisfy the nice ineq. $\deg(f+g) \le \max(\deg(f), \deg(g))$, (nonarch. triangle ineq.)

whereas we only have $|x+y| \le |x|+|y|$. (arch. triangle ineq.)

$\log|x+y|$ can be slightly larger than $\max(\log|x|, \log|y|)$

$\leadsto$ Lemma 3.1 fails "slightly".

But you can carefully deal with it! "□"

Pf2 (Stehlé–Zimmerman: A Binary Recursive GCD algorithm)

Idea: Instead of the usual division with remainder

$(x = qy + r$ with $q, r \in \mathbb{Z}$, $0 \le r < |y|)$

use generalised binary division

$(x = qy + r$ with $r \in \mathbb{Z}$, $v_2(r) > v_2(y)$, $q \in \mathbb{Q}$, $|q| < 1$, denominator of q is a power of 2)

nr. of times r is divisible by 2 (∞ if $r = 0$)

where $v_2(x) < v_2(y)$

The Euclidean algorithm still terminates (after $O(n)$ steps) with this division.

There is something analogous to Lemma 3.1 that can be used to speed up the algorithm very similar to Thm 3.2. "□"

4. Fast exponentiation

Thm 4.1 Let G be a semigroup and assume we can multiply two el. of G in $O(1)$. Then, we can compute x^k for $x \in G$ and $k \geq 1$ in time $O(\log k)$ (for large n) on an $O(\log k)$-bit RAM.

Pf If $2 \mid n$, then $x^k = (x^{k/2})^2$.

If $2 \nmid n$, then $x^k = (\underbrace{x^{(k-1)/2}}_{\substack{\text{computable} \\ \text{in} \leq C \log_2 \frac{k}{2} \\ = C(\log_2 k - 1)}})^2 \cdot x$

□

(See also pf 2 on following page.)

Thm 4.2 We can compute x^k ~~for any binary~~ integers $x \in \mathbb{Z}$ with $\leq n$ bits and $k \geq 1$ in time $O(nk)$ on an $O(\log(nk))$-bit RAM.

Pf As for Thm 4.1, using that x^k has $O(nk)$ bits, so after computing $x^{\lfloor k/2 \rfloor}$ recursively in time $\leq C\, n \cdot \frac{k}{2}$, we can compute x^k in time $O(nk)$. ⇝ total time: geom. series. □

Rmk The obvious method ($x^k = x^{k-1} \cdot x$) would (usually) have running time $\Theta(nk^2)$ because the nr. of digits in step i is $\sim ni$ and there are k steps.

Pf 2 of Thm 4.1 Write $k = \sum_{i=0}^{m} a_i 2^i$, $a_i \in \{0,1\}$

$$\Rightarrow x^k = \prod_{\substack{i: \\ a_i=1}} x^{2^i} .$$

Compute $b_i := x^{2^i}$ for $i = 0, \dots, m$ using the recurrence $b_{i+1} = b_i^2$.

Then, compute the product $\prod_{\substack{i: \\ a_i=1}} b_i$. □

Rmk Similar for $(\mathbb{Q},\cdot)$, $(M_n(\mathbb{Q}),\cdot)$, if you allow nonreduced fractions $\frac{p}{q}$ (reducing the final result would take time $O(nk\log(nk))$).

$(K[X],\cdot), (K(X),\cdot), \ldots$

Warning You can often do faster!

E.g. for the semigroup $(\mathbb{Q},+)$: you can compute $k\cdot x$ in $O(n+\log k)$.

Cor 4.3 a) You can compute the n-th Fibonacci number F_n in time $O(n)$

$\uparrow$ $\Theta(n)$ digits

b) You can compute $F_n \bmod p$ in time $O_p(\log n)$.

Pf
$$\begin{pmatrix} F_{n+1} \\ F_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n}_{\substack{O(n)\text{-bit matrix} \\ \text{computable} \\ \text{in } O(n).}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

□

5. Multiplying more than two things

Thm 5.1 We can compute the prod. $x_1 \cdots x_k$ for any bin. int. $x_1, \dots, x_k$ with $\le n$ bits in time $O(nk \log k)$ on an $O(\log(nk))$-bit RAM.

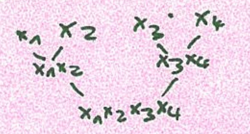
Pf w.l.o.g. $k = 2^a$.

$$x_1 \cdots x_k = \underbrace{\underbrace{(x_1 \cdots x_{2^{a-1}})}_{\substack{O(n \cdot 2^{a-1}) \text{ bits} \\ \text{time } \cancel{O} \le C\, n \cdot 2^{a-1}(a-1)}} \underbrace{(x_{2^{a-1}+1} \cdots x_{2^a})}_{\cdots}}_{O(n \cdot 2^a) \text{ bits}}$$

$$\text{time } \le C n \cdot 2^a (a-1) + O(n \cdot 2^a)$$

$$\le C n 2^a a \quad \text{for large } C.$$

"Product tree" / "Segment tree"



□

Rmk Obvious alg.: time $O(nk^2)$

Rmk Similar for $(\mathbb{Q}, \cdot)$, $(M_n(\mathbb{Q}), \cdot)$, $(\mathbb{Q}, +)$, ...

Rmk There are better alg. for $(\mathbb{Z}, +)$, $(\mathbb{Z}, \max)$, $(F_2, \cdot)$, ...

free group over two elements

Don't reduce intermediate results! (Computing the gcd would take nonlinear time.

[redacted] [redacted]

Cor 5.2 Given integers $x_1, \ldots, x_k$ with $\leq n$ bits, you can compute the ~~reduced~~ numerator and denom. p, q of a continued fraction

$$\frac{p}{q} = x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots \; \frac{1}{x_k}}}$$

in [redacted] $O(nk\log(k))$ on an $O(\log(nk))$-bit RAM.

Pf By induction,

[redacted]

$$\begin{pmatrix} p \\ q \end{pmatrix} = \underbrace{\begin{pmatrix} x_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} x_k & 1 \\ 1 & 0 \end{pmatrix}}_{\substack{\text{compute this prod.} \\ \text{in } O(nk\log k)}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

$\square$

Cor 5.3 Given integers $a_0, \dots, a_n$ and x with $\leq m$ bits, you can compute $\sum_{i=0}^{n} a_i x^i$ (in binary) in time $\mathcal{O}(mn \log n)$.

Pf By induction,

$$\begin{pmatrix} \sum_{i=0}^{n} a_i x^i \\ x^{n+1} \end{pmatrix} = \begin{pmatrix} 1 & a_n \\ 0 & x \end{pmatrix} \cdots \begin{pmatrix} 1 & a_0 \\ 0 & x \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \qquad \square$$

What if $x_1, \dots, x_k$ have very different numbers of bits?

Thm 5.4 Let $x_1, \dots, x_k$ be integers with $n_1, \dots, n_k$ bits.

We can compute $x_1 \cdots x_k$ in time

$$O\left(\sum_{i=1}^{k} n_i \left(\log \frac{n_1 + \dots + n_k}{n_i} + 1 \right) \right)$$ on an $O(\log(\sum n_i))$-bit RAM.

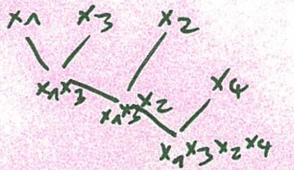
Pf ~~Start with the list of numbers.~~

~~[illegible]~~ W.l.o.g. $x_1, \dots, x_n \neq 0$. $\Rightarrow n_i = \log|x_i| + O(1)$.

Start with the list $x_1, \dots, x_n$.

In each step, replace the two integers t from the list with the smallest $\log|t|$ by their product, until there's just one number.

$\leadsto$ Product tree



Running time $\leq \sum_i \log|x_i| \cdot$ (distance of x_i from root)

$=$ [$(n_1 + \dots + n_k)$ times the average length of Huffman code over $x_1, \dots, x_n$ if the probability of x_i is $p_i = \frac{n_i}{n_1 + \dots + n_k}$]

~~Shannon entropy~~

$$= (n_1 + \dots + n_k) \cdot \left[\underbrace{\text{Shannon entropy}}_{\sum p_i \log \frac{1}{p_i}} + O(1) \right]$$

$\square$

A Theorem about Shannon codes
(Shannon: A Mathematical Theory of Communication; or look up "Shannon-Fano coding" on Wikipedia)

Thm 5.5 Let x be an int, with n bits and let $y_1, \dots, y_k$ be integers with $m_1, \dots, m_k$ bits.

We can compute $x \bmod y_1, \dots, x \bmod y_k$ in time

$$O\left(n + \sum m_i \left(\log \frac{m_1 + \dots + m_k}{m_i} + 1\right)\right) \quad \dots$$

Pf Consider the product tree for $y_1, \dots, y_k$ constructed in the proof of Thm 5.4.

For each node (labeled t), compute $x \bmod t$, starting from the root. Note that if the parent node is labeled s, then $(x \bmod t) = ((\underbrace{x \bmod s}_{<|s|}) \bmod t)$.

$t | s$, so

☐

6. Matrix operations

Let K be a field and assume $+, -, \times, ^{-1}$, $\mathbb{Z} \to K$ can be done in $O(1)$.

6.1. Multiplication

Q How quickly can we multiply two $n \times n$-matrices?

Rmk Triv. alg.: $O(n^3)$

Thm 6.1.1 (Strassen, Gaussian Elimination is not optimal)
You can multiply $A, B \in M_{n \times n}(K)$ in time $O(n^{\log_2 7})$ (on an $O(\log n)$-bit RAM).

Idea W.l.o.g. $n = 2^k$.

Write $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$

with $\frac{n}{2} \times \frac{n}{2}$-matrices A_{ij}, B_{ij}.

$$\Rightarrow AB = \underbrace{\begin{bmatrix} \sum_{j=1}^{2} A_{1j} B_{j1} & \sum A_{1j} B_{j2} \\ \sum A_{2j} B_{j1} & \sum A_{2j} B_{j2} \end{bmatrix}}_{\text{total: 8 mult. of } \frac{n}{2}\times\frac{n}{2}\text{-matrices}}.$$

$\rightsquigarrow$ time $O(n^{\log_2 8}) = O(n^3)$.

Actually, 7 mult. are enough! [Similar to Karatsuba!]

(and some number of additions/subtractions)

$\Rightarrow$ total time $\asymp 2^{2k} + 7 \cdot 2^{2(k-1)} + 7^2 \cdot 2^{2(k-2)} + \ldots + 7^k \asymp 7^k = n^{\log_2 7}$.

$\square$

Rmks The exponent ω s.t time $O(n^{\omega})$ suffices has been improved many times.

Strassen: $\omega = \log_2 7 \approx 2.807$

current record: $\omega \lesssim 2.373$

It's very unclear if $O_{\varepsilon}(n^{2+\varepsilon})$ is possible for all $\varepsilon>0$.

6.2. Determinant, rank, inverse

(Strassen, Gaussian Elimination is not Optimal)

Thm 6.2.1 Assume we can multiply $n\times n$-matrices in $\mathcal{O}(n^\omega)$, with $\omega > 2$. ~~Then, we can compute [illegible] n×n matrix~~

Then, given an $n\times n$-matrix A, we can compute an invertible $n\times n$-matrix B (and its determinant) (its inverse B^{-1}) such that BA is in reduced row echelon form in time $\mathcal{O}_\omega(n^\omega)$.

↑ first nonzero entry in each row is 1. Entries above and below this 1 are 0. Each row has at least as many leading zeros as the previous row.

Rmk Gaussian elimination does this in $\mathcal{O}(n^2 m)$ for an $n\times m$-matrix A (and $n\times n$-matrix B).

Cor 6.2.2 We can compute $\det(A)$, $\operatorname{rk}(A)$, A^{-1}, bases for $\ker(A)$, $\operatorname{im}(A)$, in $\mathcal{O}_\omega(n^\omega)$.

Pf of Cor

$\det(A) = \det(B)^{-1}\cdot\det(AB) = \det(B)^{-1}\cdot$ prod. of diagonal entries of AB

$\operatorname{rk}(A) =$ number of nonzero rows in A

↑ $\boxed{AB \text{ is upper triangular}}$

If $\operatorname{rk}(A) = n$, then $AB = I_n$, so $A^{-1} = B$.

$\ker(A) = \ker(BA) = \ldots$

$\operatorname{im}(A) = B^{-1}\cdot\operatorname{im}(BA) = \ldots$

□

Alg for Thm W.l.o.g $n=2^k$.

1) Find $B_1=\begin{bmatrix} * & 0 \\ 0 & I\end{bmatrix}$ s.t. $B_1 A = \begin{bmatrix} \text{RREF} & * \\ * & * \end{bmatrix}$

by recursively applying the alg. to the top left $\frac{n}{2}\times\frac{n}{2}$-matrix

2) Find $B_2 = \begin{bmatrix} I & 0 \\ * & I \end{bmatrix}$ s.t. in B_2B_1A below any leading 1 in the top left, there are just 0s in the bottom left.

3) Find $B_3 = \begin{bmatrix} I & 0 \\ 0 & * \end{bmatrix}$ s.t. $B_3B_2B_1A = \begin{bmatrix} \text{RREF} & * \\ \text{RREF} & * \end{bmatrix}$

by recursively applying the alg. to the bottom left $\frac{n}{2}\times\frac{n}{2}$-matrix.

4) Find $B_4 = \begin{bmatrix} I & * \\ 0 & I \end{bmatrix}$ s.t. in $B_4\cdots B_1 A$ above any leading 1s in the bottom left, there are just 0s in the top left.

~~5) Find a permutation matrix B_5 s.t. in $B_5\cdots B_1A$, the number of leading 0s among the first $\frac{n}{2}$ columns is non decreasing. Then, the left half of $B_5\cdots B_1 A$ is in RRE~~

5) Apply steps 1-4 to the right half of the matrix, ignoring all rows that have nonzero entries in the left half.

6) Apply a permutation matrix to ensure that the number of leading 0s in each row is non-decreasing as you move downwards.

The resuling matrix is in RREF.

Total: 4 recursive calls with $\frac{n}{2}\times\frac{n}{2}$-matrices and a bdd. nr. of mult. of $n\times n$-matrices

$$\Rightarrow \text{Time} \lesssim 2^{k\omega} + 4\cdot 2^{(k-1)\omega} + \dots + 4^k \underset{\omega}{\lesssim} 2^{k\omega} = n^\omega.$$

□

6.3. Characteristic polynomial

Thm 6.3.1 (Hessenberg, cf. section 2.2.4 in Cohen)

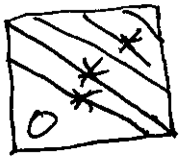
~~For an n×~~ We can compute the char. pol. $\chi_A(X) = \det(XI_n - A)$ of an $n \times n$-matrix A in $\mathcal{O}(n^3)$.

Rmk This can also be done in $\mathcal{O}_\omega(n^\omega \log n)$ (Keller-Gehrig, Fast algorithms for the char. pol.) and randomisedly in $\mathcal{O}_\omega(n^\omega)$ (Pernet-Storjohann, Faster alg. for the char. pol.) if $\#K \geq 2n^2$.

Pf First, find a similar matrix $B = (b_{ij})_{i,j}$ in

Hessenberg form: $b_{ij} = 0 \quad \forall\, 1 \leq i, j \leq n$ such that $i \geq j+2$

(Lemma 6.3.2)

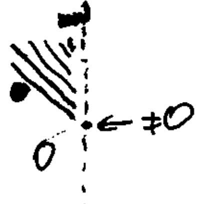


Then, $\chi_A(X) = \chi_B(X)$ ~~$\det(XI_n - A) = \det(XI_n - B)$~~ can be computed in $\mathcal{O}(n^3)$ using Laplace expansion. (Lemma 6.3.3)

$\square$

Lemma 6.3.2 You can compute a matrix B in Hessenberg form which is similar to $\underset{(a_{ij})_{i,j}}{A}$ in $\mathcal{O}(n^3)$.

Alg Start with $B=A$. We'll "fix" the columns starting from the left.



```
For $j=1,\dots,n-2$:
    If $b_{i,j} \neq 0$ for some $i \geq j+2$:
        Let $i_0$ be the smallest $i \geq j+1$ s.t. $b_{ij} \neq 0$.
        Exchange rows $j+1$ and $i_0$
            and columns $j+1 (>j)$ and $i_0 (>j)$.
        (Then, $b_{j+1,j} \neq 0$.)
        For each $i \geq j+2$:
            Subtract $v := \frac{b_{ij}}{b_{j+1,j}}$ times row $j+1$ from row $i$ and
            add $v$ times column $i^{(>j)}$ to column $j+1 (>j)$.
            (Then, $b_{i,j} = 0$.)
```

$\square$

Lemma 6.3.3

~~[illegible] can compute $\det(xI_n - B)$ in $O(n^3)$.~~

~~Pf~~ Let $B = (b_{ij})$ be an $n\times n$-matrix in Hessenberg form and let B_m be ~~[illegible]~~ its ~~[illegible]~~ top left $m\times m$ minor for $0\leq m\leq n$. Write $p_m(x) = \chi_{B_m}(x)$.

~~Then, $\chi_B(x) = (x-b_{nn})\chi_{B_{m-1}}$~~

Then, $p_m(x) = (x-b_{mm})p_{m-1}(x) - \sum_{i=1}^{m-1} b_{im}\cdot(b_{i+1,i}\cdots b_{m,m-1})\cdot p_{i-1}(x)$.

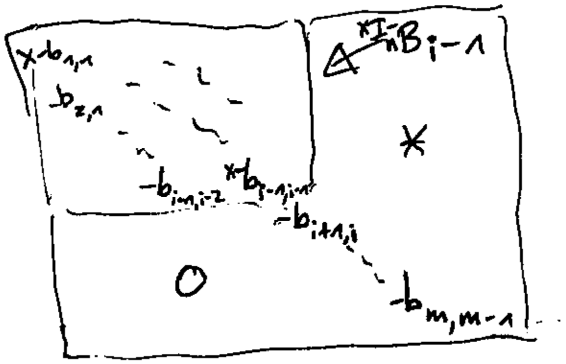
Rmk Since $\deg(p_m)=m$, we can compute $p_0,\dots,p_n$ in $O(n^3)$.

Pf Use Laplace expansion on column m. This involves picking a row.

~~[illegible]~~

~~[illegible]~~ Row m ~~[illegible]~~ gives $(x-b_{mm})p_{m-1}(x)$.

Removing row $1\leq i<m$ and column m leaves the matrix



which produces the summand

$(-1)^{i+m}\cdot(-b_{im})(-b_{i+1,i})\cdots(-b_{m,m-1})\cdot p_{i-1}(x)$

$= -b_{im}\, b_{i+1,i}\cdots b_{m,m-1}\cdot p_{i-1}(x)$.

□

6.4. Frobenius normal form

Let M be an $n\times n$-matrix over a field K.

Make K^n a (left) $K[X]$-module by defining

$$f.v := f(M)v \quad \text{for } f\in K[X].$$

(so $3.v=3v$, $X.v=Mv$, $X^2.v=M^2v$, ...)

$K[X]$ is a principal ideal domain, so the structure th for f.gen modules over PIDs shows that

$$K^n \cong \bigoplus_{i=1}^{r} K[X]/(f_i) \quad \text{for polynomials } 0\neq f_1,\dots,f_r\in K[X]$$

satisfying $f_i \mid f_{i+1}$ for $i=1,\dots,r-1$.

The pol. are unique up to units. Unique if we assume w.l.o.g. that $f_1,\dots,f_r$ are monic. These pol. are called the <u>invariant factors</u> of M.

<u>Def</u> The <u>companion matrix</u> C_f of a monic pol. $f(X)=X^n+a_{n-1}X^{n-1}+\dots+a_0$ is the matrix representing mult. by X in the vector spac $K[X]/(f(X))$ w.r.t. the basis $1, X, \dots, X^{n-1}$:

$$C_f = \begin{bmatrix} 0 & & & -a_0 \\ 1 & & 0 & \vdots \\ & \ddots & & \vdots \\ 0 & & 1 & -a_{n-1} \end{bmatrix}$$

<u>Rmk</u> The char. and min. pol. of C_f are both $f(X)$.

Pf The matrices $I, C_f, C_f^2, \dots, C_f^{n-1}$ are linearly independent

because $\underset{e_1}{\underset{\shortparallel}{Ie_1}}, \underset{e_2}{\underset{\shortparallel}{C_f e_1}}, \dots, \underset{e_n}{\underset{\shortparallel}{C_f^{n-1} e_1}}$ are.

$\Rightarrow \deg(\text{min. pol.}) = n.$

But $f(C_f) = 0$ because mult. by $f(x)$ in $K[x]/(f)$ is the zero map.

$\Rightarrow$ min. pol. $= f(x)$.

$\text{min. pol.} \mid \underset{\overset{\uparrow}{\deg = n}}{\text{char. pol.}}$

$\Rightarrow$ char. pol. $= f(x)$. □

We have shown:

Thm 6.4.1

Any $n \times n$-matrix M is similar to exactly one matrix of the form

$$\begin{bmatrix} C_{f_1} & & 0 \\ & \ddots & \\ 0 & & C_{f_r} \end{bmatrix}$$ for monic pol. $f_1 | \cdots | f_r$.

This is called the Frobenius/rational normal form of M.

The char. pol. of M is $f_1(X) \cdots f_r(X)$.

The min. pol. of M is $f_r(X)$.

Rmk Two matrices are similar iff they have the same Fr.n.f.

Cor 6.4.2 If two matrices are similar over a field $L \supseteq K$, they are similar over K.

Thm 6.4.3 (Storjohann, An $O(n^3)$ Algorithm for the Frobenius Normal Form)

- - -

7. The CRT trick

7.1 Determinants

Let M be an $n \times n$-matrix with integer entries.

Q Compute $\det(M)$.

Rmk Gaussian elimination doesn't work well because the intermediate results can be rational numbers with many digits (nr. of digits could grow exponentially in n).

Idea Compute $(\det(M) \bmod p)^{\in \mathbb{F}_p}$ for sufficiently many primes p to be able to reconstruct $\det(M)$ using the Chinese remainder theorem.

Lemma 7.1.1 For large N,

$$\underbrace{\log \prod_{p \le N \text{ prime}} p}_{= \sum_{p \le N} \log p} \asymp N \text{ and } \#\{p \le N\} \times \frac{N}{\log N}.$$

Pf This is an immediate consequence of the prime number theorem. □

Lemma 7.1.2 Any matrix $\underset{\underset{(m_{ij})_{i,j}}{\shortparallel}}{M} \in M_{n \times n}(\mathbb{R})$ satisfies

$$|\det(M)| \le \prod_{i=1}^{n} \sqrt{\sum_{j=1}^{n} m_{ij}^2} =: B(M).$$

Pf $|\det(M)|$ is the volume of the parallelepiped spanned by the rows of M. □

Hence:

Thm 7.1.3 (cf. section 2.2.3 in Cohen) For any $M \in M_{n\times n}(\mathbb{Z})$, we can compute

$\det(M)$ in time $\mathcal{O}\left((n+\log B(M))\cdot(n^{\omega}+(\log\log B(M))^2)\right)$

on an $\mathcal{O}(\log n + \log\log B(M))$-bit RAM.

Pf First, compute $B'(M) := \underset{\geq B(M)}{\prod_i \left\lceil \sqrt{\sum_j m_{ij}^2}\right\rceil} \leq 2^n B(M)$.

Then, find some $N \leq (n+\log B)$ s.t. ~~$\prod_{p\leq N} p > 2B'(M)$.~~ $\sum_{p\leq N} \lfloor \log_2 p\rfloor > \lceil \log_2(2B'(M))\rceil$

For each $p \leq N$, compute $\det(M \bmod p) \in \mathbb{F}_p$.

$\Downarrow$

p has $\mathcal{O}(\log n + \log\log B(M))$ digits

and there are $\mathcal{O}(\log B(M))$ such primes p

Finally, find the integer $x \in [-B'(M), B'(M)]$ such that $x \equiv \det(M \bmod p) \bmod p \ \forall p \leq N$ similar to Problem 4 on Pset 3. $\square$

For $N = 1, 2, 4, 8, \ldots$, compute all primes $p \leq N$ using the sieve of Erathosthenes in time $\mathcal{O}(N \log\log N)$, until you find an N that works.

<u>Rmk</u> You can also compute the determinant without reducing modulo primes using the Bareiss algorithm (Alg. 2.2.6 in Cohen)

7.2. Rank

Omk The rank of an $n\times n$-matrix M is the largest $0 \le r \le n$ s.t. some $r\times r$-minor of M (made from r not necessarily consecutive rows and columns) has nonzero determinant.

Cor 7.2.1 $\mathrm{rk}(M) \geq \mathrm{rk}(M \bmod p) \quad \forall$ primes p

with equality if p doesn't divide the (nonzero) det of a particular $r\times r$-minor of M, where $r = \mathrm{rk}(M)$.

Cor 7.2.2 $\mathrm{rk}(M) = \max_{p \le N} \mathrm{rk}(M \bmod p)$ if $\prod_{p\le N} p > T(M) := \prod_i \max\left(1, \left(\sum_j m_{ij}^2\right)^{1/2}\right)$ ($\ge B$(any minor

which can be computed in time

$O((n + \log T(M)) \cdot n^{\omega})$ ...

Omk If $\prod_{p \le N} p > T(M)$ and $N' \ge N$, then the probability that a random prime $p \le N'$ doesn't satisfy

$\mathrm{rk}(M) = \mathrm{rk}(M \bmod p)$

is at most $\dfrac{\#\{p \le N\}}{\#\{p \le N'\}}$.

This gives rise to a Monte Carlo alg. with running time $O\Big(n^{\omega} + \underbrace{(n+\log T(M)) \cdot \log\log(n + \log T(M))}_{\substack{\text{time to find} \\ \text{primes } p \ll n + \log T(M)}}\Big)$

7.3. Resultants

Rmk Let $f, g \in \mathbb{Z}[X]$ be monic pol. ~~[illegible] relatively prime in $\mathbb{Q}[X]$~~

~~[illegible]~~ If $f, g \bmod p$ are relatively prime in $\mathbb{F}_p[X]$, then f, g are rel. prime in $\mathbb{Q}[X]$.

The converse doesn't hold:

~~E.g. X^2+1 is irred. in $\mathbb{Q}[X]$, but $X^2+1 = (X+1)^2 \bmod 2$.~~

~~Q Could there be so many bad~~

~~[illegible]~~

E.g. X^2+1, $X+1$ are rel. prime in $\mathbb{Q}[X]$, but $X^2+1 = (X+1)^2 \bmod 2$

Q If f, g are rel. prime over $\mathbb{Q}$, for which p are they not rel. prime ~~[illegible]~~ mod p?

Def For any $d \geq 0$, let $K[X]_{<d} := \{ f \in K[X] : \deg(f) < d \}$.

Lemma 7.3.1 Let $f, g \in K[X]$ be pol. of degrees n, m.

Then, ~~f, g are rel. prime~~ $\gcd(f,g) = 1$ if and only if the map

$$\underset{K[X]_{<m}}{\cancel{\{a \in K[X] : \deg(a) < m\}}} \times \underset{K[X]_{<n}}{\cancel{\{b \in K[X] : \deg(b) < n\}}} \to \underset{K[X]_{<n+m}}{\cancel{\{c \in K[X] : \deg(c) < n+m\}}}$$

$$(a, b) \mapsto fa + gb$$

is an isomorphism.

Pf Note that $\dim(LHS) = m + n = \dim(RHS)$.
$K[X]_{<m} \cong K[X]/f$, $K[X]_{<n} \cong K[X]/g$, $K[X]_{<n+m} \cong K[X]/fg$.

"$\Leftarrow$" ~~If $\gcd(f,g) = c$ [illegible] is not constant, then~~ The image only contains multiples of $\gcd(f,g)$. $\Rightarrow$ The map isn't surjective.

"$\Rightarrow$" The map is an isom. according to ~~[illegible]~~ Bézout's identity. $\square$

Def The *resultant* $\operatorname{Res}(f,g)$ of pol. $f,g \in K[X]$ of deg. n,m is the determinant of the map in Lemma 7.3.1 w.r.t the basis $((1,0),(X,0),\dots,(X^{m-1},0),(0,1),\dots,(0,X^{n-1}))$ of the LHS and the basis $(1,X,\dots,X^{n+m-1})$ of the RHS.

Cor 7.3.2 $\gcd(f,g)=1 \Leftrightarrow \operatorname{Res}(f,g) \neq 0$.

Cor 7.3.3 Let $f,g \in \mathbb{Q}[X]$ be pol. and let p be a prime not dividing the denominator of any coeff. of f or g. Then, f,g are rel. prime mod p iff $p \nmid \operatorname{Res}(f,g)$.

Remk Let $f(x) = X^n + a_{n-1}X^{n-1} + \dots + a_0$ and $g(x) = X^m + b_{m-1}X^{m-1} + \dots + b_0$.

Then $$\operatorname{Res}(f,g) = \det \underbrace{\left.\begin{pmatrix} \overbrace{a_0 \quad\quad 0}^{m} & \overbrace{b_0 \quad\quad 0}^{n} \\ \begin{matrix} \vdots & \ddots & \\ & & a_0 \\ a_n & & \vdots \\ & \ddots & \\ 0 & & a_n \end{matrix} & \begin{matrix} \vdots & \ddots & b_0 \\ & & \\ b_m & & \vdots \\ & \ddots & \\ 0 & & b_m \end{matrix} \end{pmatrix}\right\}}_{} n+m$$

Sylvester matrix

Lemma 7.3.3

a) $\mathrm{Res}(g,f) = (-1)^{nm}\,\mathrm{Res}(f,g)$

c) $\mathrm{Res}(f,g) = a_n^m\, b_m^n \cdot \prod_{\substack{1\le i\le n\\ 1\le j\le m}} (\alpha_i - \beta_j) = a_n^m \prod_{1\le j\le n} g(\alpha_i)$

if $\alpha_1,\dots,\alpha_n \in \overline{K}$ are the roots of f (with mult.)
and $\beta_1,\dots,\beta_m \in \overline{K}$ — " — g — " —

b) $\mathrm{Res}(rf, sg) = r^m s^n\, \mathrm{Res}(f,g) \quad \forall r,s \in K^\times$

Pf a), b) clear

c) w.l.o.g. f and g are monic: $a_n = b_m = 1$.

$\Rightarrow f(X) = \prod_i (X-\alpha_i)$, $\quad g(X) = \prod_j (X-\beta_j)$

$\Rightarrow$ coeff. a_k of f is hom. pol. in $\alpha_1,\dots,\alpha_n$ of deg $n-k$.

-- b_l of g --- $\beta_1,\dots,\beta_m$ --- $m-l$.

$\Rightarrow$ $\mathrm{Res}(f,g)$ is hom. pol. in $\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_m$ of deg. nm.

Expand the determinant

If $(\alpha_i)_i \in \overline{K}^n$, $(\beta_j)_j \in \overline{K}^m$ satisfy $\alpha_i = \beta_j$ for some i,j, then $X-\alpha_i \mid f,g$, so $\gcd(f,g) \neq 1$, so $\mathrm{Res}(f,g) =$

$$\Longrightarrow \underbrace{\prod_{i,j} (\alpha_i - \beta_j)}_{\deg nm} \text{ divides } \underbrace{\mathrm{Res}(f,g)}_{\deg.\, nm} \text{ as a pol. in } \alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n$$

$$\Longrightarrow \mathrm{Res}(f,g) = C_{n,m} \cdot \prod_{i,j} (\alpha_i - \beta_j) \text{ for some constant } C_{n,}$$

To show $C_{n,m} = 1$, it suffices to check the equality for one pair (f,g) of pol. f,g of deg. n,m.

For example, look at $f(x) = x^n$, $g(x) = x^m + 1$

$$\cancel{\mathrm{Res}(f,g) = \begin{matrix} n\{ \\ m\{ \end{matrix} \underbrace{\begin{bmatrix} & 1 & & \\ & 1 & \ddots & \\ 1 & & \ddots & 1 \\ & \ddots & & \\ & 1 & & 1 \end{bmatrix}}_{m \quad n} \}n = 1}$$

~~$\alpha_1 = \dots = \alpha_n = 0$, β~~

$$\mathrm{Res}(f,g) = \det \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & 1 & \\ 1 & & \wr & & \ddots \\ & \ddots & & & \\ & & 1 & & 1 \end{bmatrix} = 1$$

$\alpha_1 = \dots = \alpha_n = 0$

$$\Rightarrow \prod_{i,j} (\alpha_i - \beta_j) = \Big(\underbrace{\prod_j (-\beta_j)}_{\substack{\text{const. coeff.} \\ \text{of } g}} \Big)^n = 1.$$

$\square$

Rmk Resultants can be computed using the (fast) Euclidean algorithm. (HW)

over fields K
with $O(1)$
arithmetic

The CRT trick then allows us to compute resultants of polynomials in $\mathbb{Z}[X]$, in part. to determine whether two pol. in $\mathbb{Q}[X]$ are relatively prime.

Def The <u>discriminant</u> of $f(x) = a_n X^n + \dots + a_0 \in K[X]$

$$\operatorname{disc}(f) = \frac{(-1)^{n(n-1)/2} \operatorname{Res}(f, f')}{a_n} .$$

Ex $\operatorname{disc}(ax^2 + bx + c) = b^2 - 4ac$

Lemma 7.3.4

$$\operatorname{disc}(f) = a_n^{2n-2} \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2$$

if $\alpha_1, \dots, \alpha_n \in \bar{K}$ are the roots of f (with mult.)

Pf $\deg(f) = n$, $\deg(f') = n-1$ $\qquad f(x) = a_n \cdot \prod_j (x - \alpha_j)$

$$\underset{\substack{\uparrow \\ \text{Lemma} \\ 7.3.3}}{\Rightarrow} \operatorname{Res}(f, f') = (-1)^{n(n-1)/2} a_n^{n-2} \cdot \prod_{i=1}^{n} f'(\alpha_i)$$

$$= (-1)^{n(n-1)/2} a_n^{2n-2} \cdot \prod_i \prod_{j \ne i} (\alpha_i - \alpha_j)$$

$$= a_n^{2n-2} \cdot \prod_{i<j} (\alpha_i - \alpha_j).$$

□

~~7.4. Greatest common divisor~~

7.4. Bounds on polynomial factors

Thm 7.4.1 ~~(cf. Cohen, section 3.5.1, for a better thm)~~

Let $f(X) = a_n X^n + \dots + a_0 \in \mathbb{C}[X]$,

$g(X) = b_m X^m + \dots + b_0 \in \mathbb{C}[X]$ be pol.

with $g \mid f$. Then,

$$\left|\frac{b_i}{b_m}\right| \le \binom{m}{i} \cdot \left(\sum_{j=0}^{n} \left|\frac{a_j}{a_n}\right|^2 \right)^{1/2} \quad \text{for } i = 0, \dots, m.$$

Rmk There are better bounds; see for example Thm 3.5.1 in Cohen.

Cor 7.4.2 If $f \in \mathbb{Z}[X]$ is divisible by $g \in \mathbb{Z}(x)$ in the ring $\mathbb{Z}(x)$ ~~are primitive (gcd of coeff. = 1)~~

then $$|b_i| \le \underbrace{\binom{m}{i}}_{\le 2^m} \cdot \underbrace{\left(\sum_{j=0}^{n} |a_j|^2 \right)^{1/2}}_{\le \sqrt{n+1} \cdot (\max a_j)} \quad \text{for } i = 0, \dots, m.$$

Pf of Cor ~~$g \mid f$ in $\mathbb{C}[X] \xRightarrow{\text{Gauss's lemma}} g \mid f$ in $\mathbb{Z}[X]$~~ ~~g, f primitive~~

$$g \mid f \text{ in } \mathbb{Z}(X) \Rightarrow b_m \mid a_n \Rightarrow |b_m| \le |a_n|.$$ ☐

The Thm follows from:

Lemma 7.4.3 (Landau's inequality) Let $r_1, \dots, r_n \in \mathbb{C}$ be the roots of $f(X) = a_n X^n + \dots + a_0 \in \mathbb{C}[X]$. Then,

$$\prod_{\substack{1 \le i \le n: \\ |r_i| \ge 1}} |r_i| \le \left(\sum_j \left|\frac{a_j}{a_n}\right|^2 \right)^{1/2}.$$

7.5. Gcd of integer polynomials

Thm 7.5.1 Let $0 \neq f, g \in \mathbb{Z}[X]$ be polynomials of degree $\leq n$ whose coefficients c satisfy $|c| \leq B$. We can compute $\gcd(f,g) \in \mathbb{Z}[$ in average time $\tilde{O}(n(n + \log B))$ on a randomized $O(\log(n + \log B))$-bit RAM.

Here, $\tilde{O}(X)$ means $O(X(\log X)^k)$ for some fixed $k \geq 0$.

Rmk There's a subtle difference between gcd in $\mathbb{Q}[X]$ and in $\mathbb{Z}[$
The gcd in $\mathbb{Q}[X]$ is only defined up to mult. by elements of $\mathbb{Q}^\times$
but the gcd in $\mathbb{Z}[X]$ is defined up to mult. by el. of $\mathbb{Z}^\times = \{$
For example, $\gcd_{\mathbb{Z}[X]}(2X, 6X^3) = 2X$.
But the correct multiple is easy to determine, so it suffices to find $\gcd(f,g)$ up to mult. by a scalar.

Rmk Let [illegible] $\hat{h}$ [illegible] $= \gcd_{\mathbb{Q}[X]}(f,g) \in \mathbb{Z}[X]$ [illegible] be primitive (relatively prime coefficients). Then, $\hat{h} \mid f, g$ by Gauss's ~~lemma. Let $t = \gcd(lc(f), lc(g))$~~
lemma, so in part. $lc(\hat{h}) \mid lc(f), lc(g)$.
Let $t = \gcd(lc(f), lc(g))$. We'll explain how to compute the gcd $h(X) = \frac{t}{lc(\hat{h})} \cdot \hat{h}(X) \in \mathbb{Z}[X]$ of f, g (over $\mathbb{Q}[X]$) that has leading coefficient $lc(h) = t$.

~~Let $k = \mathcal{O}($ large enough so that~~

~~$p_1 \cdots p_k > \underbrace{2^n \cdot \sqrt{n+1} \cdot B}_{\text{upper bd. on coeff. of } \tilde{h}(x)}$.~~

Pf of Thm 7.5.1

Find $K = \mathcal{O}(n + \log B)$ large enough so that

$$\prod_{\substack{p \le K \\ p \nmid t}} p > \underbrace{2^n \cdot \sqrt{n+1} \cdot B}_{\text{upper bd. on coeff. of } \tilde{h}(x)}. \qquad \text{(Note that } \prod_{p \mid t} p \le |t| \le B.\text{)}$$

Find $L = \mathcal{O}(\log(nB))$ large enough so that

$$\prod_{\substack{K < p \le L \\ p \nmid t}} p > \underbrace{(2n)! \cdot B^{2n}}_{\text{upper bd. for } |s_d(f,g)|}.$$

Find $M = \mathcal{O}(n \log(nB))$ large enough so that

$$\#\{K < p \le M, p \nmid t\} > 2 \cdot \#\{K < p \le L, p \nmid t\}.$$

~~$\# A := 2\#\{p \le K, p \nmid t\} = \mathcal{O}\left(\frac{n+\log B}{\log(n+\log B)}\right)$ Pick different primes $p_1, \dots, p_A \le M$ uniformly at random. For $i = 1, \dots, A$ Compute $h_i := \gcd(f \bmod p_i, g \bmod p_i)$ where w.l.o.g. $lc(h_i) = t \bmod p_i$.~~

Let $A = \#\{p \le K, p \nmid t\} = \mathcal{O}\left(\frac{n + \log B}{\log(n + \log B)}\right)$.

Pick a random prime $p_0 \le M$, $p_0 \nmid t$ and compute $d' := \deg(\gcd(f \bmod p_0, g \bmod p_0))$.

(With prob. $\ge \frac{1}{2}$, we have $d' = d$. Always $d' \ge d$.)

Compute $\gcd(f \bmod p, g \bmod p)$ for random $p \le M$, $p \nmid t$ until you found $p_1, \dots, p_A \le M$ such that

$$\deg(\gcd(f \bmod p_i, g \bmod p_i)) = d' \text{ for } i = 1, \dots, A.$$

The expected nr. of primes to try is $\mathcal{O}(A)$.

Let $h_i = \gcd(\dots)$ where w.l.o.g. $lc(h_i) \equiv t \bmod p_i$.

Note that $p_1 \cdots p_A \ge \prod_{\substack{p \le k \\ p \nmid t}} p > 2^n \cdot \sqrt{n+1} \cdot B$, so there is at most one pol. $\tilde{h}'$ with coeff. $\le 2^n \cdot \sqrt{n+1} \cdot B$ s.t.

$$\tilde{h}' \equiv h_i \bmod p_i \text{ for } i = 1, \dots, A.$$

If there is one and it divides f and g, then it must be the gcd of f and g in $\mathbb{Q}[X]$.

Otherwise (with prob. $\le \frac{1}{2}$), start over!

$\square$

Rmk There's another "efficient" alg. that avoids reduction modulo primes: The subresultant algorithm (cf. section 3.3 in Cohen). It's basically the Euclidean alg., but avoids exponential growth of coefficients by dividing by an appropriate (easy to compute) integer (dividing all coeffs) at each step!

Rmk You can use a similar alg. as in Thm 7.5.1 for example to compute the gcd of polynomials $f, g \in \mathbb{F}_q[T][X]$. The running time is better: $\tilde{\mathcal{O}}(\underbrace{nD}_{\text{sise of input}})$, where all coeff. of f, g are pol. in T of deg. $\le D$.

The reason is again that the triangle ineq. in $\mathbb{F}_q(T)$ is stronger than in $\mathbb{R}$. (Instead of Cor. 7.4.2, you have the obvious fact that the degree of any coeff. of $\gcd(f,g)$ is also at most D)

8. ~~Factoring polynomials~~ Finding roots over finite fields

~~8.1. [illegible] fast [illegible] fields~~

Assume we can do arithmetic in $\mathbb{F}_q$ in time $\mathcal{O}(1)$.

and select ~~an~~ element of $\mathbb{F}_q$ uniformly at random

distinct

Thm 8.1 We can determine the number of roots of a pol. $f \in \mathbb{F}_q[X]$ of degree n in $\mathbb{F}_q$ in time $\tilde{\mathcal{O}}(n \log q)$.

Rmk Realistically we can't do arithmetic in $\mathbb{F}_q$ in $\mathcal{O}(1)$, but only in $\tilde{\mathcal{O}}(\log q)$, so there would be an additional factor of $\log q$.

Pf

$$\prod_{t \in \mathbb{F}_q} (X - t) = X^q - X$$

$$\Rightarrow \prod_{\substack{t \in \mathbb{F}_q: \\ f(t)=0}} (X-t) = \gcd(f, X^q - X) = \gcd(f, \underbrace{X^q - X \bmod f}_{\substack{\text{compute } X^q \bmod f \\ \text{using fast exponentiation} \\ \tilde{\mathcal{O}}(n \log q)}})$$

$$\Rightarrow \#\{t \in \mathbb{F}_q : f(t) = 0\} = \deg(\gcd(\text{--}))$$

compute gcd using fast Eucl. alg. $\tilde{\mathcal{O}}(n)$

□

Joke $f \in \mathbb{Z}[X]$. $\Rightarrow \#\{t \in \mathbb{Z} : f(t) = 0\} = \deg(\gcd(f, g))$,

where $g(X) = \sin(\pi X)$.

Lemma 8.2 Let $f \in \mathbb{F}_q[x]$ be a pol. of degree n ~~[illegible]~~ with n distinct roots in $\mathbb{F}_q$ (i.e. dividing $X^q - X$).
We can find a random splitting $f = gh$ into pol. $g, h \in \mathbb{F}_q[x]$ in time $\mathcal{O}(n \log q)$ [on an $\mathcal{O}(n)$-bit RAM], where the probability that $\deg(g) = k$ is given by a binomial distribution:

$$\mathbb{P}(\deg(g) = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0, \dots, n,$$

where $p = \frac{\lceil \frac{1}{2} q \rceil}{q}$ $\left(\approx \frac{1}{2}\right)$.

[So generally $\deg(g) \approx \frac{n}{2}$.]

[We'll use:]

Lemma 8.3 The following pol. has $\lceil \frac{1}{2} q \rceil$ (distinct) roots in $\mathbb{F}_q$:

$$u_q(X) = \begin{cases} X^{\frac{q+1}{2}} - X, & q \text{ odd} \\ \sum_{i=0}^{r-1} X^{2^i}, & q = 2^r \end{cases}$$

Pf If q is odd, the roots of $u_q(x)$ are the squares in $\mathbb{F}_q$.

(Also, $\underbrace{u_q(x)}_{\frac{q+1}{2} \text{ roots}} \underbrace{\left(X^{\frac{q-1}{2}} - 1\right)}_{\frac{q-1}{2} \text{ roots}} = X(X^{q-1} - 1) = \underbrace{X^q - X}_{q \text{ roots}}$).

If $q = 2^r$, then

$$\underbrace{u_q(x)}_{2^{r-1} = \frac{q}{2} \text{ roots}} \underbrace{(u_q(x) + 1)}_{2^{r-1} = \frac{q}{2} \text{ roots}} = u_q(x)^2 + u_q(x)$$

$$= u_q(X^2) + u_q(x)$$

$x \mapsto x^2$ is a hom. in $\mathbb{F}_q$

$$= X^{2^r} - X = \underbrace{X^q - X}_{q \text{ roots}}$$

□

[We'll use the following special case:]

Lemma 8.3 We have $X^q - X = u_q(X)\, v_q(X)$, where

$$u_q(X) = \begin{cases} X^{\frac{q+1}{2}} - X, & q \text{ odd} \\ \sum_{i=0}^{r} X^{2^i}, & q = 2^r \end{cases}$$

$$v_q(X) = \begin{cases} X^{\frac{q-1}{2}} - 1, & q \text{ odd} \\ u_q(X) + 1, & q = 2^r . \end{cases}$$

Rmk $\deg(u_q) = \lceil \frac{q}{2} \rceil$, so u_q has $\lceil \frac{q}{2} \rceil$ distinct roots in $\mathbb{F}_q$
and v_q has $\lfloor \frac{q}{2} \rfloor$ —"— .

Rmk If q is odd, the roots of u_q are exactly the squares in $\mathbb{F}_q$.

Pf of Lemma 8[redacted].2 Let $r_1, \dots, r_n \in \mathbb{F}_q$ be the roots of f.

Consider the [redacted] (linear) Vandermonde map

$$\begin{aligned} \mathbb{F}_q^n &\longrightarrow \mathbb{F}_q^n \\ a=(a_0,\dots,a_{n-1}) &\longmapsto \big(\underbrace{a_0+a_1r_i+\dots+a_{n-1}r_i^{n-1}}_{\varphi_a(r_i)}\big)_{i=1,\dots,n} \end{aligned}$$

It is an isomorphism because $r_1, \dots, r_n$ are distinct.

Pick $(a_0, \dots, a_{n-1}) \in \mathbb{F}_q^n$ uniformly at random.

$\Rightarrow (s_i)_{i=1,\dots,n} = (\varphi_a(r_i))_{i=1,\dots,n}$ is a uniformly random el. of $\mathbb{F}_q^n$.

Compute

$$g(x) := \gcd\big(f(x), u_q(\varphi_a(x))\big) = \prod_{\substack{1 \le i \le n: \\ u_q(\underset{s_i}{\text{[redacted]}})=0}} (X - r_i)$$

and $h(x) := \frac{f(x)}{g(x)}$. (Note that we can compute $u_q(\varphi_a(x)) = \begin{cases} \varphi_a(x)^{\text{[redacted]}} - \varphi_a(x) \\ \sum_i \varphi_a(x)^{2^i} \end{cases}$ modulo $f(x)$ in $\mathcal{O}(n \log q)$!)

The probability that $\deg(g) = k$ is the probability that [redacted] exactly k coordinates s_i of a random elements of $\mathbb{F}_q^n$ are roots of $u_q(x)$, which is $\binom{n}{k} p^k (1-p)^{n-k}$. $\square$

Thm 8.1.4 We can find all roots of a pol. $f(x) \in \mathbb{F}_q(x)$ of degree n in average time $\tilde{O}(n \log q)$ using randomization.

Rmk It's unknown whether there's a deterministic alg. that does this in polynomial time (in n, $\log q$).

Pf ~~[illegible]~~

W.l.o.g. $f(X) \mid X^q - X$ (replace f by $\gcd(f, X^q - X)$).

Use Lemma 8.1.2 to find a splitting $f = gh$ and recursively apply the alg. to g and h.

We have $\mathbb{E}(\deg(g)) = np$ and

$$\mathbb{P}\left((\deg(g) - \mathbb{E}(\deg(g)))^2 \geq \Delta \right) \leq \frac{\mathrm{Var}(\deg(g))}{\Delta} = \frac{np(1-p)}{\Delta},$$

where $p = \frac{\lceil \frac{1}{2} q \rceil}{q} \in [\frac{1}{2}, \frac{2}{3}]$.

$$\Rightarrow \mathbb{P}\left(\deg(g) \in [\tfrac{1}{4}n, \tfrac{3}{4}n] \right) \geq \frac{1}{2}$$

for sufficiently large n.

This shows that the average running time is $\tilde{O}(n \log q)$ (with one more factor of $\log n$ than in Lemma 8.1.2.) $\square$

9. Squarefree factorisation

Let K be a perfect field and assume we can do arithmetic in K in $\mathcal{O}(1)$. (for $K = \mathbb{F}_q$, realistically $\tilde{\mathcal{O}}(\log q)$)

If $\mathrm{char}(K) = p > 0$, assume we can compute the p-th root of $x \in K$ in $\mathcal{O}(1)$. (for $K = \mathbb{F}_q$ realistically $\mathcal{O}(\log_2 q \cdot \log \frac{q}{p})$ using the formula

$$x^{1/p} = x^{q/p}$$ and fast exponentiation)

Rmk If $p \neq 0$, then $(\sum a_i x^i)^p = \sum a_i^p x^{ip}$, so we can determine whether a pol. $f(x) \in K[x]$ of deg n is a p-th power, and if so determine its p-th root, in $\mathcal{O}(n)$.

Thm 9.1 Let $f(x) \in K[x]$ be a monic pol. of degree n.

We can compute all polynomials

$$s_k(x) = \prod_{\substack{t \in K[x] \text{ monic irred.} \\ \underbrace{v_t(f)}_{\boxed{\text{nr. of times } t(x) \text{ divides } f(x)}} = k}} t(x) \qquad \text{for } k = 1, \dots, n$$

(so that $f(x) = \prod_{k=1}^{n} s_k(x)^k$ with squarefree $s_k(x)$)

in time $\tilde{\mathcal{O}}(n)$.

~~INSERT (1)~~ 9.2

~~Alg for lemma~~

~~[All gcds are assumed to be monic!]~~

~~Compute $g = \gcd(f, f')$, $b_0 = \frac{f}{g}$, $c_0 = \frac{f'}{g}$.~~

~~For $k = 1, \dots, \min(p-1, n)$, compute~~

~~$a_k = \gcd(b_{k-1}, c_{k-1})$, $b_k = \frac{b_{k-1}}{a_k}$, $c_k = \frac{c_{k-1}}{a_k} - b_k'$.~~

~~Then, $r_k = a_k$ for all $k \leq \min(p-1, n)$.~~

(*)

[This follows from:]

Lemma 9.2

Let char(k)=p (≥0). We can compute [redacted] all pol.

$$a_k(X) = \prod_{\substack{t:\,[redacted] \\ \underbrace{v_t(f)\equiv k \bmod p}_{\iff v_t(f)=k \text{ if } \mathrm{char}(k)=p=0 \text{ or } p>n}}} t(X) \qquad \text{for } 1\le k\le N := \begin{cases}\min(p-1,n), & p\neq 0\\ n, & p=0.\end{cases}$$

in time $\tilde{O}(n)$.

Pf of Thm 9.1 (using Lemma 9.2)

Clear if [redacted] $p=0$, or $p>n$, so assume [redacted] $2\le p\le n$.

[redacted]

[redacted] The polynomial

$$h(X)=\frac{f(X)}{\prod_{1\le k\le p-1\,[redacted]} a_k(X)^k}$$ is a p-th power.

Recursively apply the alg. (from the Thm.) to the pol. $\sqrt[p]{h(X)}$ of degree $\le \frac{n}{p}$.

$$\rightsquigarrow \sigma_l(X) = \prod_{\substack{t: \\ \underbrace{v_t(\sqrt[p]{h})=l}_{\iff v_t(h)=lp}}} t(X) \qquad \text{for } l=1,\dots,\left\lfloor \frac{n}{p}\right\rfloor.$$

$$\Rightarrow s_{k+lp}=\gcd(a_k,\sigma_l) \text{ for } 1\le k\le p-1,\ 1\le l\le \left\lfloor\frac{n}{p}\right\rfloor.$$

$$s_k=\frac{a_k}{\prod_{l\ge 1} s_{k+lp}} \qquad \text{for } 1\le k\le p-1$$

$$s_{lp}=\frac{\sigma_l}{\prod s_{k+lp}} \qquad \text{for } 1\le l\le \left\lfloor\frac{n}{p}\right\rfloor. \quad \square$$

Alg for Lemma 9.2

(All gcds are assumed to be monic)

Compute $g = \gcd(f, f')$, $b_0 = \frac{f}{g}$, $c_0 = \frac{f'}{g} - b_0'$

For $k = 1, \dots, N$ [illegible] :

Compute $a_k = \gcd(b_{k-1}, c_{k-1})$, $b_k = \frac{b_{k-1}}{a_k}$, $c_k = \frac{c_{k-1}}{a_k} - b_k'$

Claim (correctness) We have

$$a_k = \prod_{\substack{t:\\ v_t(f) \equiv k \bmod p}} t(x) \qquad \text{for } k = 1, \ldots, N$$

$$b_k = \prod_{\substack{t:\\ v_t(f) \not\equiv 0, \ldots, k \bmod p}} t(x) \qquad \text{for } k = 0, \ldots, N$$

$$c_k = \sum_{\substack{t:\\ v_t(f) \not\equiv 0, \ldots, k \bmod p}} (v_t(f) - (k+1)) \cdot \frac{t'(x)}{t(x)} \cdot b_k(x). \qquad \text{for } k = 0, \ldots, N$$

Pf (by ind. over k)

k=0:

$$f(x) = \prod_t t(x)^{v_t(f)}$$

$$\Rightarrow f'(x) = \sum_{t:} v_t(f) \cdot \frac{t'(x)}{t(x)} \cdot f(x)$$

If $v_t(f)=0$, ~~then~~ $v_t(g)=0$.

Otherwise: If $v_t(f) \not\equiv 0 \bmod p$, then $v_t(f') = v_t(f) - 1$. $\Rightarrow v_t(g) = v_t(f) - 1$. (because $t' \not\equiv 0$)

(so $v_t(f) \not\equiv 0$ in k)

If $v_t(f) \equiv 0 \bmod p$, then $v_t(f') \geq v_t(f)$. $\Rightarrow v_t(g) = v_t(f)$.

(so $v_t(f) = 0$ in k)

$$\Rightarrow g(x) = \prod_{\substack{t:\\ v_t(f) \not\equiv 0}} t(x)^{v_t(f)-1} \cdot \prod_{\substack{t:\\ v_t(f) \equiv 0}} t(x)^{v_t(f)}$$

$$\Rightarrow b_0(x) = \frac{f(x)}{g(x)} = \prod_{\substack{t:\\ v_t(f) \not\equiv 0}} t(x)$$

$$c_0(x) = \frac{f'(x)}{g(x)} \cdot b_0'(x) = \sum_{\substack{t:\\ v_t(f) \not\equiv 0}} (v_t(f)-1) \frac{t'(x)}{t(x)} \cdot b_0(x)$$

$k-1 \to k$:

- Let $t \mid b_{k-1}$.

Then, $v_t(c_{k-1}) = \begin{cases} 1, & v_t(f) \equiv k \mod p \\ 0, & v_t(f) \not\equiv k \mod p \end{cases}$

$\Rightarrow a_k$ is as claimed.

$\Rightarrow b_k$ — —

$$b_k'(x) = \sum_{\substack{t: \\ v_t(f) \not\equiv 0,\dots,k}} \frac{t'(x)}{t(x)} \cdot b_k(x).$$

$\Rightarrow c_k$ is as claimed. ☐

~~Running time:~~

Claim: The alg. has running time $\tilde{O}(n)$.

Running time $= \tilde{O}\left(\sum \deg(a_k) + \sum \deg(b_k) + \sum \deg(c_k)\right)$.

$\deg(a_k) \le \deg(b_{k-1})$

$\deg(c_k) \le \deg(b_k)$

$$\sum_k \deg(b_k) \le \sum_t v_t(f) \cdot \deg(t)$$
$$= \deg\left(\prod t(x)^{v_t(f)}\right)$$
$$= \deg(f) = n.$$

☐

10. Factoring over finite fields

~~10.1: Distinct degree factors~~

You've seen one method (Berlekamp-Zassenhaus) on problem set

(Expected running time $\tilde{O}(n^{\omega} + n \log q)$)

There are faster algorithms that work more like the root-finding alg. in section 8:

10.1. Distinct-degree factorisation

Lemma 10.1.1

$$X^{q^k} - X = \prod_{\substack{t \in \mathbb{F}_q(X) \\ \text{monic irred} \\ \deg(t) \mid k}} t(X).$$

Pf If t is irred. of degree $d \mid k$, then its splitting field is $\mathbb{F}_{q^d} \subseteq \mathbb{F}_q$

$\Rightarrow$ Each ~~root~~ r of t satisfies $r^{q^k} = r$

$\Rightarrow$ RHS | LHS.

On the other hand, each root r of $X^{q^k} - X$ lies in $\mathbb{F}_{q^k}$.

Now, $\mathbb{F}_q \subseteq \mathbb{F}_q(r) \subseteq \mathbb{F}_{q^k}$, so $\mathbb{F}_q(r) = \mathbb{F}_{q^d}$ for some $d \mid k$. The min. pol. of r has degree d.

$\Rightarrow$ LHS | RHS. $\square$

Cor 10.1.2 Let $f \in \mathbb{F}_q[X]$ be a pol. of degree n and assume we are given the n polynomials $X^{q^k} \bmod f$ for $k=1,\dots,n$. Then, we can compute the degree k parts

$$g_k(X) = \prod_{\substack{t \mid f \text{ monic irred} \\ \deg(t)=k}} t(X) \quad \text{of } f(X) \text{ for } k=1,\dots,n$$

in time $\tilde{O}(n^2)$.

Rmk If f is squarefree, then $f(X) = g_1(X) \cdots g_n(X)$.

Alg Let $h_0 = f$. w.l.o.g. f is squarefree (after using Thm 9.1 and replacing $f(X)$ by $s_1(X) \cdots s_n(X)$.)

For $k = 1, \dots, n$:

compute $g_k = \gcd(h_{k-1}, X^{q^k} - X)$

and $h_k = \dfrac{h_{k-1}}{g_k}$. □

~~If $g_k = 1$, we "compute" $h_k = h_{k-1}$ in $\tilde{O}(1)$.~~

Q How to compute $\alpha_k = X^{q^k}$ (mod f) for $k = 1, \dots, n$?

~~Both: Fast exponentiation takes time $\tilde{O}(n \log q^k) = \tilde{O}(nk \log q)$.~~

Rmk $\alpha_k(X) = \alpha_{k-1}(X)^q$, so using fast exponentiation, we can compute α_k from α_{k-1} in $\tilde{O}(n \log q)$.

$\Rightarrow$ Total time $\tilde{O}(n^2 \log q)$.

We can do faster!

Rmk $\alpha_{k+l}(X) \equiv X^{q+l} \equiv (X^{q^k})^{q^l} \equiv \alpha_l(\alpha_k(X)) \bmod f(X)$.

Warning In general, if $\alpha(X) \equiv \beta(X) \bmod f(X)$, then $\alpha(\gamma(X)) \equiv \beta(\gamma(X)) \bmod f(\gamma$

not $\bmod f(X)$

Pf of Rmk $\alpha_l(X) \equiv X^{q^l} \bmod f(X)$

$\Rightarrow \alpha_l(\alpha_k(X)) \equiv \alpha_k(X)^{q^l} \bmod f(\alpha_k(X))$.

Since $f(\alpha_k(X)) \equiv f(X^{q^k}) \equiv f(X)^{q^k} \equiv 0 \bmod f(X)$,

$\alpha_k(X) \equiv X^{q^k} \bmod f(X)$

$Y \mapsto Y$[illegible] is a hom. on $\mathbb{F}_q(X)$ and fixes the coeff. of f

this implies

$$\alpha_l(\alpha_k(X)) \equiv \alpha_k(X)^{q^l} \equiv (X^{q^k})^{q^l} \equiv X^{q^{k+l}} \equiv \alpha_{k+l}(X) \bmod f(X$$

$\square$

~~$\alpha_{k+l}(X) = X^{q^{k+l}} = X^{q^k \cdot q^l} = (X^{q^k})^{q^l} = \alpha_l(\alpha_k(X))$ mod f.~~

~~Pf $\alpha_l(X) - X^{q^l} = f(X) g_l(X)$ for some pol. $g(X)$~~

~~$\alpha_l(X) \equiv X^{q^l}$ mod f~~

~~$\Rightarrow \alpha_l(\alpha_k(X)) = \alpha_l(X^{q^k}) = (X^{q^k})^{q^l} + \underbrace{f(X^{q^k})}_{f(X)^{q^k}} g_l(X^{q^k}) \equiv X^{q^{k+l}}$ mod f(X~~

~~Warning: in general, $\beta(X)$ mod f(X) doesn't imply $\alpha(\gamma(X)) \equiv \beta(\gamma(X))$ mod f(X)!~~

Modular composition problem

Given polynomials $\alpha(X), \beta(X), f(X)$ of degree $< n$,

compute $\alpha(\beta(X))$ mod f.

(Note that it's in general not enough to know $\alpha(X)$ mod f(X)!)

Rmk Evaluating α at $\beta(X)$ using "Cor 5.3" takes time $\tilde{O}(n^2)$.
↑ degree of $\beta(X)$

It can be done faster, but I won't explain a better alg. for (a single) modular composition. Instead, we'll use a "cheat":

Evaluating a pol. of degree n at n points is not much harder than evaluating it at a single point(!):

Lemma 10.1.3 Assume we can do arithmetic in R in $O(1)$.

Let $f \in R[X]$ be a pol. of degree $\leq n$ and let $c_1, \ldots, c_n \in R$.

We can compute $f(c_1), \ldots, f(c_n)$ in $\tilde{O}(n)$.

Pf $f(c_i) = f(X)$ mod $X - c_i$.

Using the modulo tree ("Thm 5.5"), we can compute f mod $X - c_1, \ldots, f$ mod $X - c_n$ in $\tilde{O}(n)$. □

Cor 10.1.4 Let $f \in \mathbb{F}_q[X]$ be a pol. of degree n. We can compute $\alpha_k(X) = X^{q^k} \bmod f$ for $k=1,\dots,n$ in $\tilde{\mathcal{O}}(n^2 + n\log q)$.

Pf First, compute $\alpha_1(X) = X^q$ in $\tilde{\mathcal{O}}(n \log q)$ using fast exponentiation. Afterwards:

Claim: We can compute $\alpha_1, \dots, \alpha_{2^r}$ in $\tilde{\mathcal{O}}(n^2 r)$ for $r \in \lceil \log_2 n \rceil$.

Pf Assume we've computed $\alpha_1, \dots, \alpha_{2^{r-1}}$.

Then, $\alpha_{2^{r-1}+i}(X) = \underbrace{\alpha_{2^r}(\alpha_i(X))^{\bmod f}}_{\substack{\text{value of the pol. } \alpha_{2^r}(X) \\ \text{at } \alpha_i(X) \text{ in the ring} \\ \mathbb{F}_q[X]/(f).}}$ for $i = 1, \dots, 2^{r-1}$.

Arithmetic in $\mathbb{F}_q[X]/(f)$ takes time $\tilde{\mathcal{O}}(n)$.

$\Rightarrow$ Since $2^{r-1} \le n$, by Lemma 10.1.3, we can compute $\alpha_{2^{r-1}+i}$ for $i = 1, \dots, 2^{r-1}$ in $\tilde{\mathcal{O}}(n^2)$ after computing α_j for $j = 1, \dots, 2^{r-1}$ in $\tilde{\mathcal{O}}(n^2(r-1))$. □

□ (Cor)

Cor 10.1.5 Let $f \in \mathbb{F}_q[X]$ of degree n and $g \in \mathbb{F}_q[X]$ of deg. $< n$. We can compute $g(x)^{q^k} \bmod f(x)$ for $k = 1, \dots, n$ in $\tilde{\mathcal{O}}(n^2 + n \log q)$.

Pf $g(x)^{q^k} \equiv g(x^{q^k}) \equiv g(\alpha_k) \bmod f$.

$\Rightarrow$ It suffices to evaluate g at $\alpha_1, \dots, \alpha_n \in \mathbb{F}_q[X]/(f)$. □

Summary We can compute the degree k parts of f for $k = 1, \dots, n$ in $\tilde{\mathcal{O}}(n^2 + n \log q)$.

Rmk ~~This can actually be done in $\tilde{\mathcal{O}}(n^{\frac{3}{2}+\varepsilon}(\log q)^{1+\varepsilon} + n^{1+\varepsilon}(\log q)^{2+\varepsilon})$ bit operations (not the more expensive operations in $\mathbb{F}_q$!) (see Kedlaya, Umans: Fast pol. factorization and modular composition)~~

10.2. Equal-degree factorisation

Lemma 10.2.1 Let $f \in \mathbb{F}_q[X]$ be ~~a pol. of degree n = km which~~ the product of m irred. pol. of degree d (so $n = \deg(f) = km$, $f \mid \frac{X^{q^d} - X}{\prod_{e \neq d}(X^{q^{de}} - X)}$).

Assume we are given the pol. $\alpha_i = (X^{q^i} \bmod f)$ for $i = 0, \dots, d-1$. Then, we can find a random splitting $f =$ into pol. $g, h \in \mathbb{F}_q(X)$ in time $\tilde{O}(n^2 + n \log q)$, where ~~the prob~~ ~~that deg(g)~~

$$\mathbb{P}(\deg(g) = kd) = \binom{m}{k} p^k (1-p)^{m-k} \text{ for } l = 0, \dots, m,$$

where $p = \frac{\lceil \frac{1}{2} q \rceil}{q}$.

Pf Let $f = f_1 \cdots f_m$ be the factorisation of f.

$$\underset{CRT}{\Rightarrow} \mathbb{F}_q[X]/(f) \cong \prod_{i=1}^{m} \mathbb{F}_q[X]/(f_i) \cong \prod_{i=1}^{m} \mathbb{F}_{q^d}.$$

Pick $a_0, \dots, a_{n-1} \in \mathbb{F}_q$ uniformly at random.

$\Rightarrow \varphi_a := a_0 + \dots + a_{n-1} X^{n-1} \bmod f$ is a uniformly random element of $\mathbb{F}_q(X)/(f) \cong \prod_{i=1}^{m} \mathbb{F}_{q^d}$.

Consider the trace map (linear) Tr sending X to $X + X^q + X^{q^2} + \dots + X^{q^{d-1}}$
$= \alpha_0 + \alpha_1 + \dots + \alpha_{d-1}$.

On $\mathbb{F}_{q^d}$, it's the (field) trace map $\mathrm{Tr}_{\mathbb{F}_{q^d}|\mathbb{F}_q} : \mathbb{F}_{q^d} \to \mathbb{F}_q$.

$\leadsto$ We get a map $\prod \mathbb{F}_{q^d} \xrightarrow{} \prod \mathbb{F}_q$.

linear surjective

Each element of $\prod \mathbb{F}_q$ has the same number of preimages.

$\Rightarrow \mathrm{Tr}(\varphi_a)$ is a uniformly random element of $\prod \mathbb{F}_q$.

$$\overset{\mathrm{Tr}(\varphi_a)}{\parallel}$$

$\varphi_a(x) + \varphi_a(x)^q + \dots + \varphi_a(x)^{q^{d-1}}$

[redacted]

can be computed in $\tilde{O}(n^2)$.

Let $v_q(X) = \begin{cases} X^{(q+1)/2} - X \\ \sum X^{2^i} \end{cases}$ as in Lemma 8.3.

Now, $\gcd(f, v_q(\mathrm{Tr}(\varphi_a)))$ is divisible by f_i if and only if the image of $v_q(\mathrm{Tr}(\varphi_a))$ in the i-th factor $\mathbb{F}_q$ is 0.

Since $v_q(x)$ has $\lceil \frac{1}{2} q \rceil$ roots in $\mathbb{F}_q$, this happens with prob. p.

The events for different i are all independent. $\square$

<u>Cor 10.2.2</u> [redacted] We can factor any f as in Lemma 10.2.1 in expected time $\tilde{O}(n^2 + n \log q)$.

<u>Pf</u> like Thm 8.4. $\square$

[redacted]

Combining all factorization steps (squarefree, distinct-degree, equal-degr

Thm 10.2.3 (von zur Gathen, Shoup: Computing Frobenius maps and factoring polynomials ~~Let $f \in \mathbb{F}_q[X]$ be a pol. of degree n~~

We can factor a pol. $f \in \mathbb{F}_q[X]$ of degree n in time $\tilde{O}(n^2 + n \log q)$.

Rmk This is a factor of $(n + \log q)$ worse than the triv. lower bound $\Theta(n)$.

There are faster algorithms (improving n, but not $\log q$)

Kaltofen-Shoup: Subquadratic-time factoring of polynomials over finite fields (baby-step/giant step alg.)

Kedlaya-Umans: Fast polynomial factorization and modular composition

(better modular comp. + baby step/giant step)

essentially: $n + \log q \rightsquigarrow n^{1/2} + \log q$

[Don't know how to improve the $\log q$ factor even when just counting linear factors!]

11. Factoring over nonarchimedean local fields

Let K be a nonarch. local field with

normalised valuation v : map $v: K \to \mathbb{Z} \cup \{\infty\}$ s.t. $v(x)=\infty \Leftrightarrow x=0$
$v(xy)=v(x)+v(y)$
$v(x+y) \geq \min(v(x),v(y))$

uniformiser π : el. $\pi \in K$ s.t. $v(\pi)=1$

ring of integers $\mathcal{O} = \{x \in K : v(x) \geq 0\}$

prime ideal $\mathfrak{q} = \{x \in K : v(x) \geq 1\} = (\pi)$

(finite) residue field $k = \mathcal{O}/\mathfrak{q} = \mathbb{F}_q$

~~Let $r_1, \dots, r_q \in K$ be representatives of the~~

Ex $K = \mathbb{Q}_p = \{\frac{x}{y} : x, y \in \mathbb{Z}_p, y \neq 0\}$, $v(x)$ = nr. of times x is divisible by p,
$\pi = p$, $\mathcal{O} = \mathbb{Z}_p$, $\mathfrak{q} = (p)$, $k = \mathbb{F}_p$, $q = p$.

~~Assume we can do arithmetic in $k = \mathbb{F}_q$ in $O(1)$.~~

~~Let $a_1, \dots, a_q \in K$ be representatives~~

~~Let $r_1, \dots, r_q \in K$ be~~

In computations, we won't work with elements of $\mathcal{O}$ (or K), but with mod $\mathfrak{q}^k$-approximations in $\mathcal{O}/\mathfrak{q}^k$.

Assume we can do arithmetic in $\mathcal{O}/\mathfrak{q}^k$ in $O(k)$.
[In part., we can do arithmetic in $k = \mathcal{O}/\mathfrak{q}$ in $O(1)$.]

Ex For $K = \mathbb{Q}_p$, this involves arithmetic on base p integers with $O(k)$ digits.

Pf. of Hensel's lemma

Assume $f \equiv gh \mod \mathfrak{p}^k$ with f, g, h monic.

Let $\tilde{g} = g + \pi^k r$, $\hat{h} = h + \pi^k s$.

$$\Rightarrow \tilde{g}\hat{h} \equiv \underbrace{gh}_{f \bmod \mathfrak{p}^k} + \pi^k(rh + sg) \mod \mathfrak{p}^{2k}.$$

If g, h are relatively prime modulo $\mathfrak{p}$, then they are rel. prime modulo $\mathfrak{p}^k$, so the residue class $\frac{f-gh}{\pi^k}$ mod p^k can be written uniquely as $rh + sg \mod \mathfrak{p}^k$ with polynomials r, s, where

$\deg(r) < \deg(g)$, $\deg(s) < \deg(h)$.

~~There are unique pol. $\tilde{g}, \hat{h} \mod \mathfrak{p}^{2k}$ s.t. $f \equiv \tilde{g}\hat{h} \mod \mathfrak{p}^{2k}$, $\deg(\tilde{g}) = \deg(g)$, $\deg(\hat{h}) = \deg(h)$.~~

Then, $\tilde{g}\hat{h} \equiv f \mod \mathfrak{p}^{2k}$. □

Proceed by induction.

Hensel's lemma Let $f, g, h \in \mathcal{O}(x)$ be monic polynomials such that $f \equiv gh \mod \mathfrak{p}$, where g, h are relatively prime mod $\mathfrak{p}$. Then, there are unique pol. $\tilde{g}, \hat{h} \in \mathcal{O}[x]$ s.t. $f = \tilde{g}\hat{h}$, $\tilde{g} \equiv g \mod \mathfrak{p}$, $\hat{h} \equiv h \mod \mathfrak{p}$.

Thm 11.1 We can compute $\hat{g}, \hat{h} \bmod \varphi^k$ in time $\tilde{O}(nk)$.

Pf HW. $\square$

More generally:

Thm 11.2 Let $f, g_1, \dots, g_r \in \mathcal{O}(x)$ be monic pol. such that $f \equiv g_1 \cdots g_r \bmod \varphi$, where $g_1, \dots, g_r$ are pairwise relatively prime mod φ. Then, we can compute the (unique) pol. $\hat{g}_1, \dots, \hat{g}_r \bmod \varphi^k$ such that $f = \tilde{g}_1 \cdots \tilde{g}_r$, $\hat{g}_i \equiv g_i \bmod \varphi$ in time $\tilde{O}(n)$.

Pf HW. $\square$

Rmk In general, knowing a (monic) polynomial $f \in \mathcal{O}[X]$ of degree n modulo $\mathfrak{p}^k$ isn't enough to determine the structure of the factorisation of f in $K(X)$ no matter how large k is.

For example, a monic degree 2 pol. $f(X) \equiv X^2 \bmod \mathfrak{p}^k$ could be

- a square: $f(X) = X^2$
- a product of two lin. pol.: $f(X) = (X - \pi^i)(X + \pi^i) = X^2 - \pi^{2i}$ for $2i \geq k$
- irreducible: $f(X) = X^2 - \pi^{2i+1}$ for $2i+1 \geq k$.

(Similarly, $X^2 + \epsilon \in \mathbb{R}[X]$ could be a square, prod. of lin., or irred. for arbitrarily small ϵ.)

But if f is squarefree, then knowing $f \bmod \mathfrak{p}^k$ suffices for sufficiently large k (depending of f).

12. Factoring pol. over $\mathbb{Z}$ (attempt 1)

Let $f \in \mathbb{Z}[X]$ ~~reduce f mod p for a~~ be squarefree and monic.
Factor f mod p for a suitable prime p.
How does that factorization relate to that of f?

Let $f = f_1 \cdots f_r$.

$\Rightarrow f \equiv f_1 \cdots f_r \mod p$.

But $f_1, \dots, f_r$ could factor further mod p.

Rmk If $p \nmid \operatorname{disc}(f)$, then f mod p is still squarefree, and vice-versa.

Lemma 12.1 Let K be a nr. field, $\mathfrak{q}$ a prime ideal of K, $f \in \mathcal{O}_K[X]$ be a pol. whose leading coeff. and discriminant are not divisible by $\mathfrak{q}$.
Consider the number field $L = K[X]/(f)$.
The polynomial f splits mod $\mathfrak{q}$ in the same way as the prime ideal splits in L:

$f \equiv g_1 \cdots g_t \mod \mathfrak{q}$ with $g_1, \dots, g_t$ irreducible mod $\mathfrak{q}$ $\in (\mathcal{O}_K/\mathfrak{q})[X]$

$\mathfrak{q} = \mathfrak{Q}_1 \cdots \mathfrak{Q}_t$ with prime ideals $\mathfrak{Q}_1, \dots, \mathfrak{Q}_t$ of L

with $\mathcal{O}_L/\mathfrak{Q}_i \cong (\mathcal{O}_K/\mathfrak{q})[X]/(g_i)$.

Pf See e.g. Prop I.8.3 in Neukirch's Algebraic Number Theory. ∎

Def Let $L|K$ be a Galois ext. of number fields, $\mathfrak{p}$ a prime of K and $\mathfrak{q}$ a prime of L dividing $\mathfrak{p}$.

The decomposition group is $D(\mathfrak{q}|\mathfrak{p}) = \{\sigma \in G : \sigma(\mathfrak{q}) = \mathfrak{q}\}$.

The inertia group is $I(\mathfrak{q}|\mathfrak{p}) = \{\sigma \in D(\mathfrak{q}|\mathfrak{p}) : \sigma(x) \equiv x \bmod \mathfrak{q} \ \forall x \in \mathcal{O}_L\}$.

Thm 12.2 a) G acts transitively on the primes $\mathfrak{q}$ of L dividing $\mathfrak{p}$.

b) $D(\tau\mathfrak{q}|\mathfrak{p}) = \tau D(\mathfrak{q}|\mathfrak{p})\tau^{-1}$

c) $I \ \ldots \ = \tau I \ \ldots \ \tau^{-1}$

d) $\mathfrak{q}$ divides $\mathfrak{p}$ exactly $|I(\mathfrak{q}|\mathfrak{p})|$ times.

e) $I(\mathfrak{q}|\mathfrak{p})$ is a normal subgroup of $D(\mathfrak{q}|\mathfrak{p})$ with $D(\mathfrak{q}|\mathfrak{p})/I(\mathfrak{q}|\mathfrak{p}) \cong \mathrm{Gal}(\mathcal{O}_L/\mathfrak{q} \mid \mathcal{O}_K/\mathfrak{p})$.

Cor 12.3 If $e = |I(\mathfrak{q}|\mathfrak{p})|$ and $ef = |D(\mathfrak{q}|\mathfrak{p})|$ and $efr = |G| = [L:K]$ then $\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1^e \cdots \mathfrak{q}_r^e$ with $[\mathcal{O}_L/\mathfrak{q}_i : \mathcal{O}_K/\mathfrak{p}] = f$.

Unfortunate Cor 12.4 Let $f \in \mathbb{Z}[X]$ be an irreducible monic pol. such that $L = \mathbb{Q}[X]/(f)$ is a Galois ext. of $\mathbb{Q}$ with Galois group G. Unless G is cyclic, f splits modulo every prime p.

Pf If $p \mid \mathrm{disc}(f)$, then $f \bmod p$ is not squarefree.

If $p \nmid \mathrm{disc}(f)$, then $f \bmod p$ is squarefree and splits like p in L. $\Rightarrow I(\mathfrak{q}|p) = 1$ (unram.) If $p\mathcal{O}_L = \mathfrak{q}$, then $e = 1$, $f = |$

$\Rightarrow D(\mathfrak{q}|\mathfrak{p}) \cong \mathrm{Gal}(\mathcal{O}$

Let $p_1, \dots, p_k$ be distinct prime numbers.

Extreme Ex Then $L = \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_k})$ is a Galois ext. of $\mathbb{Q}$ with Galois group $G = (\mathbb{Z}/2\mathbb{Z})^k$. The largest cyclic subgroups of G have size 2.

$L = \mathbb{Q}(\underbrace{\sqrt{p_1} + \dots + \sqrt{p_k}}_{\alpha :=})$. Let $f \in \mathbb{Z}[X]$ be the min. pol. of $\alpha \in \mathcal{O}_L$

For any $p \nmid \mathrm{disc}(f)$, the pol. $f \bmod p$ splits either into 2^k linear factors (if $|D| = 1$) or into 2^{k-1} quadratic factors (if $|D| = 2$).

Rmk " ~~For~~ a random monic pol. $f \in \mathbb{Z}[X]$ of degree n, with probabilit

a) f is irreducible.

b) The Galois closure of $\mathbb{Q}[X]/(f)$ over $\mathbb{Q}$ has Galois group S_n

c) For a random prime p, $f \bmod p$ is irreducible with probability $\frac{1}{n}$."

(The proof of c) uses the Chebotarev density theorem.)

13. Lattice reduction

Def A lattice $\Lambda \subset \mathbb{R}^n$ is a set of the form

$$\Lambda = \mathbb{Z}v_1 + \dots + \mathbb{Z}v_n = \{a_1v_1 + \dots + a_nv_n \mid a_1, \dots, a_n \in \mathbb{Z}\}$$

with linearly independent vectors $v_1, \dots, v_n$.

Such $v_1, \dots, v_n$ are called a basis of Λ.

Rmk We can encode a basis $(v_1, \dots, v_n)$ of Λ as a matrix

$$\begin{pmatrix} - v_1 - \\ \vdots \\ - v_n - \end{pmatrix} \in GL_n(\mathbb{R}).$$

A change of basis corresponds to left multiplication by an element of $GL_n(\mathbb{Z})$.

Hence, we obtain a bijection

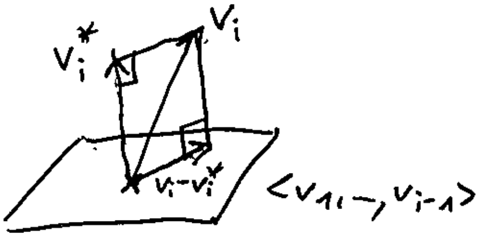
$$\{\Lambda \subset \mathbb{R}^n \text{ lattice}\} \Longleftrightarrow GL_n(\mathbb{Z}) \backslash GL_n(\mathbb{R}).$$

Goal For a given lattice Λ with basis $(v_1, \dots, v_n)$, find a basis $(w_1, \dots, w_n)$ consisting of "nearly as short as possible" vectors $w_1, \dots, w_n$.

Def Let $v_1, \dots, v_n$ be a basis of $\mathbb{R}^n$.
For $i = 1, \dots, n$, let v_i^* be the component of v_i orthogonal to the subspace $\langle v_1, \dots, v_{i-1} \rangle$, i.e.:

$$v_i^* \perp v_1, \dots, v_{i-1}$$

and $v_i - v_i^* \in \langle v_1, \dots, v_{i-1} \rangle$.



Write $v_i = v_i^* + \sum_{j=1}^{i-1} \mu_{ij} v_j^*$
with $\mu_{ij} \in \mathbb{R}$ (for $j < i$).

The vectors $v_1^*, \dots, v_n^*$ are the Gram-Schmidt basis for $v_1, \dots, v_n$.

The numbers μ_{ij} $(j<i)$ are the Gram-Schmidt coefficien

Lemma 13.1
a) $\langle v_1, \dots, v_i \rangle = \langle v_1^*, \dots, v_i^* \rangle$ for $i = 1, \dots, n$.
In part., $v_1^*, \dots, v_n^*$ form a basis of $\mathbb{R}^n$.

b) $v_i^* \perp v_j^*$ for all $i \neq j$.

c) $\mu_{ij} = \dfrac{v_i \cdot v_j^*}{|v_j^*|^2}$.

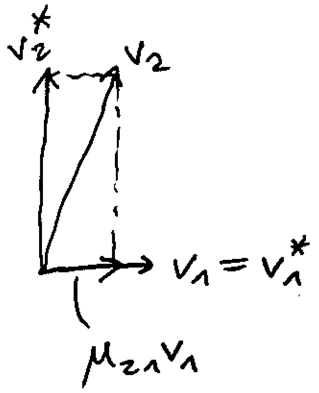
d) $|v_i|^2 = |v_i^*|^2 + \sum_{j=1}^{i-1} \mu_{ij}^2 \, |v_j^*|^2$.

e) $\begin{pmatrix} - v_1 - \\ \vdots \\ - v_n - \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ \mu_{21} & \ddots & \\ \vdots & \ddots & \\ \mu_{n1} & \cdots \mu_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} - v_1^* - \\ \vdots \\ - v_n^* - \end{pmatrix}$

Pf a) induction over i
b) clear from a)
c) projection formula
d) Pythagoras.
e) clear

$\square$

Ex $(n=2)$



Thm 13.3 Let $v_1,\dots,v_n$ be a basis of $\mathbb{R}^n$. There are integers $a_{ij} \in \mathbb{Z}$ $(j<i)$ such that the G-S coeff. for the basis $w_1,\dots,w_n$ given by $w_i = v_i - \sum_{j=1}^{i-1} a_{ij} v_j$ satisfy $|\mu_{ij}| \leq \frac{1}{2}$ for all $j<i$.

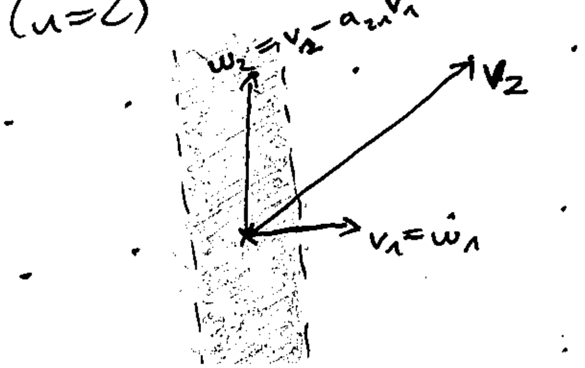
They can be computed using $\mathcal{O}(n^3)$ operations in $\mathbb{R}$.

Rmks a)
$$\begin{pmatrix} - w_1 - \\ \vdots \\ - w_n - \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ a_{ij} & & 1 \end{pmatrix} \begin{pmatrix} - v_1 - \\ \vdots \\ - v_n - \end{pmatrix}$$

b) $w_i^* = v_i^*$ for $i = 1,\dots,n$.

Pf of Thm
```
For $i = 1,\dots,n$:
    For $j = i-1,\dots,1$:
        Subtract an appropriate integer multiple of
        row $j$ from row $i$ to make $|\mu_{ij}| \leq \frac{1}{2}$.
```
□

Ex $(n=2)$

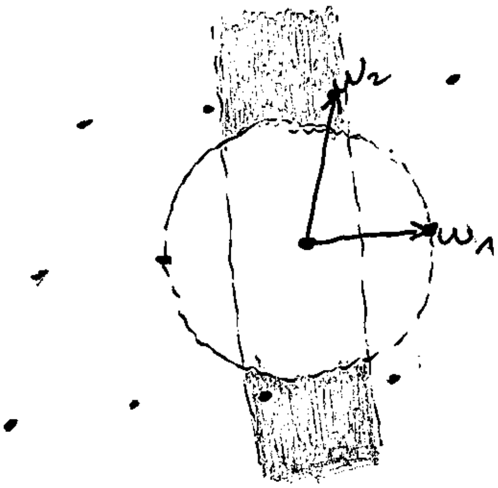


Thm 13.4 Let $n=2$. The following algorithm computes a basis w_1, w_2 of $\Lambda = \mathbb{Z}v_1 + \mathbb{Z}v_2$ such that w_1 is a shortest nonzero vector in Λ:

Alg
```
1) Replace $v_1, v_2$ by the basis computed in Thm 13.3
   such that $|\mu_{21}| \le \frac{1}{2}$.
2) If $|v_1| \le |v_2|$:
       Return $w_1 = v_1$, $w_2 = v_2$.
   If $|v_1| > |v_2|$:
       Swap $v_1, v_2$ and return to step 1.
```

Pf correctness: Assume the alg. returned w_1, w_2.

Clearly, w_1, w_2 still form a basis of Λ.
We have $|\mu_{21}| \le \frac{1}{2}$ and $|w_1| \le |w_2|$.



That there is no shorter nonzero vector in Λ than u is "clear from the picture".

Formally:

$$|b_1 w_1 + b_2 w_2|^2 = b_1^2 |w_1|^2 + b_2^2 |w_2|^2 + 2 b_1 b_2 \underbrace{(w_1 \cdot w_2)}_{\mu_{21}|w_1|^2}$$

$$\ge \underbrace{(b_1^2 + b_2^2 - b_1 b_2)}_{\ge 1} |w_1|^2 \ge |w_1|^2$$

for all $(0,0) \ne (b_1, b_2) \in \mathbb{Z}^2$.

algorithm terminates: $|v_1|$ gets smaller in every iteration. But Λ has only finitely many vectors of length less than the original $|v_1|$. $\square$

Thm 13.5 Assume $v_1, v_2 \in \mathbb{Z}^2$ and the coordinates c of v_1, v_2 satisfy $|c| \leq B$. Then, the algorithm from Thm 13.4 takes $\mathcal{O}(\log B)$ steps (for large B).
($\Rightarrow$ polynomial running time in size of the input!)

Pf Rephrase the alg. as follows:

w.l.o.g. $|v_1| \geq |v_2|$.

$u_1 := v_1, \quad u_2 := v_2.$

$u_{i+2} := u_i - k_i u_{i+1}$ with $k_i = \mathrm{round}\left(\frac{u_i \cdot u_{i+1}}{|u_{i+1}|^2}\right) \in \mathbb{Z}$

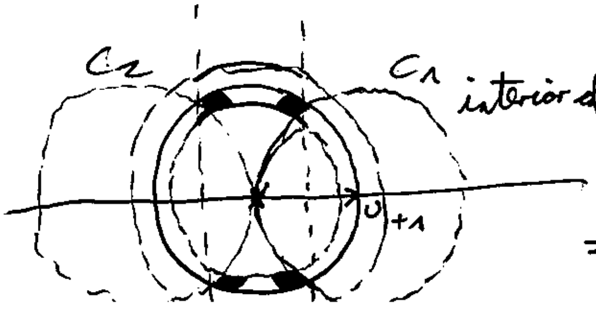
until $|u_{j+1}| > |u_j|$.

Then, we return $w_1 = u_j$, $w_2 = u_{j+1}$.

Clearly, $|u_1| \geq |u_2| > \ldots > |u_j|$. Let $\delta = \frac{11}{10}$.

Claim: ~~$|u_i| > \delta |u_{i+1}|$~~
For all $1 \leq i \leq j-3$, we have $|u_i| > \delta |u_{i+2}|$.

Pf Assume $|u_i| \leq \delta |u_{i+2}|$.
$\Rightarrow |u_i| \leq \delta |u_{i+1}|$ and $|u_{i+1}| \leq \delta |u_{i+2}|$.



u_{i+2} lies in the vertical strip and in th interior of the inner annulus.
u_i lies in the outer annulus.
$u_{i+2} \in u_i + \mathbb{Z} u_{i+1}$.
$\Rightarrow u_{i+2}$ lies in the shaded region.

In particular, v_{i+2} doesn't lie in the interior of the balls C_1 or C_2.

$\Rightarrow$ The projection of v_{i+1} onto v_{i+2} has length $\leq \frac{1}{2}|v_{i+2}$

$\Rightarrow v_{i+3} = v_{i+1}$.

$\Rightarrow |v_{i+3}| = |v_{i+1}| > |v_{i+2}|$

$\Rightarrow j = i+2$ ↯ $\square$

The claim implies that the total number of steps is $O(\log_\delta B)$ because $|v_1|^2 \leq O(B)$ and each $|v_i|^2$ is an integer. $\square$

Def A basis $v_1,\dots,v_n$ of $\mathbb{R}^n$ is LLL-reduced if its G-S basis and coeff.

satisfy $|\mu_{ij}| \le \frac{1}{2}\quad \forall\, j<i$

Lenstra, Lenstra, Lovász
Arjen, Hendrik, László

and $\|v_{i+1}^*\|^2 \ge \frac{1}{2}\|v_i^*\|^2$.

Reference chapter 16 of "Modern Computer Algebra".

Lemma 13.5 Any $0 \neq r \in \Lambda$ satisfies

$$\|r\|^2 \ge \frac{1}{2^{n-1}} \cdot \|v_1\|^2.$$

[v_1 is "almost" as short as possible.]

Pf Write $r = b_1 v_1 + \dots + b_k v_k$ with $k \le n$, $b_1,\dots,b_k \in \mathbb{Z}$, $b_k \neq 0$

The component of r orthogonal to $\langle v_1,\dots,v_{k-1}\rangle$ is $b_k v_k^*$.

$$\Rightarrow \|r\| \ge \underbrace{|b_k|}_{\ge 1} \cdot \|v_k^*\| \ge \|v_k^*\| \ge \frac{1}{2^{k-1}} \cdot \|v_1^*\| = \frac{1}{2^{k-1}} \cdot \|v_1\|. \qquad \square$$

Thm 13.6 The following alg. computes an LLL-reduced basis of a lattice $\Lambda = \mathbb{Z} v_1 + \dots + \mathbb{Z} v_n$ (if it terminates).

Alg 13.6

```
1) Compute the G-S basis $v_1^*, \dots, v_n^*$   (which we'll keep up to date
                                                  as we change $v_1, \dots, v_n$).

Let $i \leftarrow 1$.
While $i \le n$:
    2) For $j = i-1, \dots, 1$:
        Subtract round($\underset{\substack{\shortparallel \\ \frac{v_i \cdot v_j^*}{|v_j^*|^2}}}{\mu_{ij}}$) times $v_j$ from $v_i$ to make $|\mu_{ij}| \le \frac{1}{2}$
    3) If $i \ge 2$ and $|v_i^*|^2 < \frac{1}{2} |v_{i-1}^*|^2$:
        4) Swap $v_i, v_{i-1}$. Recompute $v_i^*, v_{i-1}^*$.
           Return to $i \leftarrow i-1$.
       Otherwise:
           Proceed to $i \leftarrow i+1$.
Return $v_1, \dots, v_n$.
```

Pf Correctness is clear: At the beginning of any while loop, $v_1, \dots, v_{i-1}$ satisfy the LLL-reducedness criterion. □

Rmk The alg. always terminates, but that's less obvious. We'll show that it has polynomial running time if $v_1, \dots, v_n \in \mathbb{Z}^n$.

Lemma 13.7 Let $v_1, \dots, v_n$ be a basis of $\mathbb{R}^n$ and $2 \leq i \leq n$ with $|\mu_{i,i-1}| \leq \frac{1}{2}$ and $|v_i^*| < \frac{1}{2}|v_{i-1}^*|$.

Let $w_1, \dots, w_n$ be the same basis, but with v_i, v_{i-1} swapped.

Then:

a) $w_j^* = v_j^* \quad \forall j \neq i, i-1$. [$\Rightarrow$ We only need to update v_i^*, v_{i-1}^* in step 4.]

b) $|w_{i-1}^*|^2 < \frac{3}{4}|v_{i-1}^*|^2$ [$\Rightarrow$ Exponential decay.] [But $|v_{i-1}^*|^2 \in \mathbb{Q}$ might not be an integer!]

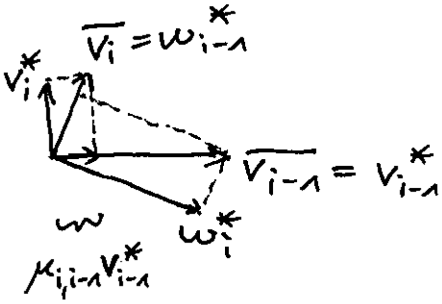
c) $|w_i^*| \leq |v_{i-1}^*|$.

d) $|w_{i-1}^*| \cdot |w_i^*| = |v_{i-1}^*| \cdot |v_i^*|$.

Pf a) $\langle w_1, \dots, w_{j-1} \rangle = \langle v_1, \dots, v_{j-1} \rangle$ and $w_j = v_j$.

b–d) Only the components of v_{i-1}, v_i orthogonal to $\langle v_1, \dots, v_{i-2} \rangle$ matter for the computation of $v_{i-1}^*, v_i^*, w_{i-1}^*, w_i^*, \mu_{i,i-1}$.

Let $\overline{v_i}, \overline{v_{i-1}}$ be these components of v_i, v_{i-1}.



b) $|w_{i-1}^*|^2 = |v_i^*|^2 + \mu_{i,i-1}^2 |v_{i-1}^*|^2$ by Pythagoras

$$< \frac{1}{2}|v_{i-1}^*|^2 + \frac{1}{4}|v_{i-1}^*|^2 = \frac{3}{4}|v_{i-1}^*|^2.$$

c) clear

d) $|v_{i-1}^*|\cdot|v_i^*|$ = area of the parallelogram spanned by $\overline{v_{i-1}}, \overline{v_i}$

$|w_{i-1}^*|\cdot|w_i^*| =$ ______ 〃 ______ ____________

□

Rmk For any $0 \le k \le n$,

$$d_k := |v_1^*|^2 \cdots |v_k^*|^2$$

$$= \begin{pmatrix}\text{k-dimensional volume of the parallelepiped} \\ \text{spanned by } v_1, \dots, v_k\end{pmatrix}^2$$

$$= \det(M_k),$$

for the $k\times k$-matrix $M_k = (v_i \cdot v_j)_{1\le i,j\le k}$.

In particular, if $v_1, \dots, v_n \in \mathbb{Z}^n$, then $d_0, \dots, d_n \in \mathbb{Z}$.

Lemma 13.8 If $v_1,\dots,v_n$ lie in $\mathbb{Z}^n$ and $|v_1|,\dots,|v_n| \le B$, then Alg. 13.6 does at most $O(n^2 \log B)$ swaps (line 4).

Pf Consider the integer $D = d_1 \cdots d_{n-1} > 0$.

In the beginning,

$$d_k = |v_1^*|^2 \cdots |v_k^*|^2 \le |v_1|^2 \cdots |v_k|^2 \le B^{2k},$$

so $D \le B^{2(1+\dots+(n-1))} = B^{n(n-1)}$

D only changes in line 4, in which it decreases at least by a factor of $\frac{4}{3}$.

(More precisely, d_{i-1} decreases, while $d_1,\dots,d_{i-2},d_i,\dots,d_{n-1}$ remain the same.) by Lemma 13.7 b

$\Rightarrow$ line 4 can only run $O(\log_{\frac{4}{3}}(B^{n(n-1)})) = O(n^2 \log B)$ times

Cor 13.9 Alg. 13.6 performs $O(n^4 \log B)$ operations in $\mathbb{Q}$.

Pf 1) $O(n^3)$

$$O(n^2 \log B) \text{ times} \left\{ \begin{array}{l} 2)\ O(n^2) \\ 3)\ O(n) \\ 4)\ O(n^2) \end{array} \right.$$

□

Lemma 13.10 For $v_1,\dots,v_n \in \mathbb{Z}^n$, we have

$$d_{k-1} v_k^* \in \mathbb{Z}^n.$$

Pf The orth. projection $v_k - v_k^*$ onto $\langle v_1,\dots,v_{k-1}\rangle$ is given by the formula

$$v_k - v_k^* = \underbrace{\begin{pmatrix} | & & | \\ v_1 & \cdots & v_{k-1} \\ | & & | \end{pmatrix}}_{\text{int. matrix}} \underbrace{M_{k-1}^{-1}}_{\frac{\text{int. mat.}}{\det(M_{k-1})}} \underbrace{\begin{pmatrix} - \; v_1 \; - \\ \vdots \\ - \; v_{k-1} \; - \end{pmatrix}}_{\text{int. matrix}} \underbrace{v_k}_{\in \mathbb{Z}^n}$$

□

Cor 13.11 (bounds on denominators) ~~For Alg. 13.6, the denominators of~~

If $v_1,\dots,v_n \in \mathbb{Z}^n$ with $|v_1|,\dots,|v_n| \le B$, then in Alg. 13.6, the vectors v_k^* at any time satisfy $t v_k^* \in \mathbb{Z}^n$ for some $t \in \mathbb{Z}$ with $\log(t) = O(n \log B)$. ("bounded denominators")

Pf d_{k-1} is nonincreasing and ~~lo~~ $d_{k-1} = O(n \log B)$ in the beginning. □

[How long can the vectors be?]

Lemma 13.12 If $|v_1|, \dots, |v_n| \le B$ in the beginning, then during Alg. 13.6:

a) $|v_1|, \dots, |v_n| \le \sqrt{n}\, B$

except possibly during step 2.

b) $|v_1|, \dots, |v_n| \le n\, (2B)^{2n}$

during step 2.

Cor 13.13 We ~~have~~ always $\log|v_1|, \dots, \log|v_n| \le O(n \log B)$ (for large B). ($\Rightarrow$ bound on numerators)

Pf a) holds in the beginning.

$\max(|v_1|, \dots, |v_n|)$ can only change during step 2 (when $|v_i|$ might chang

After step 2, $|\mu_{ij}| \le \frac{1}{2} \quad \forall j < i$.

Then, $|v_i|^2 = |v_i^*|^2 + \sum_{j<i} \mu_{ij}^2 \, |v_j^*|^2$.

By Lemma 13.7, $\max(|v_1^*|, \dots, |v_n^*|)$ is nonincreasing.

In the beginning, it's $\le B$.

$\Rightarrow |v_i|^2 \le B^2 + \sum_{j<i} \frac{1}{4} B^2 \le n B$.

b) At the begining of step 2,

$$|\mu_{ij}| = \frac{|v_i \cdot v_j^*|}{|v_j^*|^2} \le \underset{\sim\sim\sim}{\sqrt{n} \cdot B \cdot B} \cdot B^{2(j-1)} = \sqrt{n} \cdot B^{2j} \le \sqrt{n} \cdot B^{2(n-1)}$$

because $|v_i| \le \sqrt{n} \cdot B$, $|v_j^*| \le B$, $|v_j^*|^2 = \frac{d_j}{d_{j-1}} \ge \frac{1}{d_{j-1}} \ge B^{-2(j-1)}$.

Moreover, $|v_j^*| \le |v_j| \le \sqrt{n} B$.

When subtracting $\text{round}(\mu_{ij}) \cdot v_j$ from v_i,

μ_{ik} changes by $|\text{round}(\mu_{ij}) \cdot \underbrace{\mu_{jk}}_{1 \cdot 1 \le \frac{1}{2}}| \le \frac{1}{2} \mu_{ij}| + \frac{1}{2}$

$\Rightarrow$ At the beginning of the for loop body in step 2 with index j, we have

$$\max(1, |\mu_{i1}|, \dots, |\mu_{i,i-1}|) \leq 2^{i-j-1} \cdot \sqrt{n}\, B^{2(n-1)} \leq 2^{n-2} \cdot \sqrt{n} \cdot B^{2(}$$

(every time we handle an index j, the LHS at most increases by a factor of 2).

Then, $$|v_i|^2 = |v_i^*|^2 + \sum_{k<i} \mu_{ik}^2 |v_k^*|^2$$

$$\leq 2^{2(n-2)} n\, B^{4(n-1)} \cdot n B^2$$

$$\leq n^2 (2B)^{4n}.$$

$\square$

Summary

Thm 13.13 If $v_1, \dots, v_n \in \mathbb{Z}^n$ with $|v_1|, \dots, |v_n| \leq B$, then Alg. 13.6 has running time $\tilde{O}(n^5 (\log B)^2)$ (on an $O(\log(n \log B))$-b RAM).

Pf The rational numbers computed in the alg. have numerators and denominators with $O(n \log B)$ bits. This shows the claim with Cor 13.9. $\square$

14. Factoring over the integers, attempt 2

We will identify a pol. $\underset{= a_n x^n + \dots + a_0}{f} \in$ ~~[illegible]~~ $\mathbb{Z}[x]$ ~~with the ve~~ of degree $\leq n$ with the vector $(a_0, \dots, a_n) \in \mathbb{Z}^{n+1}$.

We'll write $|f| = |f|_2 = \sqrt{a_0^2 + \dots + a_n^2}$ for its Euclidean length.

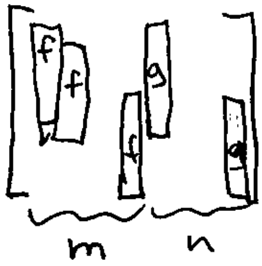
Reminders:

Lemma 14.1 If $f \in \mathbb{Z}[x]$ is divisible by ~~[illegible]~~ ~~[illegible]~~ the pol. $g \in \mathbb{Z}[x]$ of degree d in the ring $\mathbb{Z}[x]$, then $|g| \leq \sqrt{d+1} \cdot 2^d \cdot |f|$.

Pf Imediate consequence of ~~[illegible]~~ Cor 7.4.2. □

Lemma 14.2 If $f, g \in \mathbb{Z}[x]$ are pol. of degrees n, m, then $|\mathrm{Res}(f,g)| \leq |f|^m \cdot |g|^n$

Pf $\mathrm{Res}(f,g) = \det$



. □

Thm 14.3 We can factor any polynomial $f \in \mathbb{Z}(x)$ of degree n with $|f| \leq B$ in the ring $\mathbb{Q}(x]$ in time $\tilde{O}(n^{10} + n^{8} [redacted] (\log B)^2)$.

Pf Let p be a prime.

We can factor f mod p. Let [redacted] $a \in \mathbb{F}_p(X]$ be an irreducible factor [redacted] of degree t.

Goal: Find [redacted] an (the) irreducible factor $g \in \mathbb{Z}(X)$ of f which is divisible by a modulo p.

~~Let $d := \deg(g)$. We'll find d by trying $d = 1, 2, \dots, n$.~~

[redacted] We don't know $\deg(g)$, so will try $d = t, t+1, \dots, n$.

$\Rightarrow |g| \leq \sqrt{d+1} \cdot 2^d \cdot |f| =: A$.

[redacted]

Consider the set

$$\Lambda = \{ \tilde{g} \in \mathbb{Z}(x] \text{ of deg.} \leq d \mid \text{u divisible by } [redacted] \bmod p \}.$$

It is a lattice $\Lambda \subseteq \mathbb{Z}^{d+1} \subset \mathbb{R}^{d+1}$ [redacted]
(of rank $d+1$ because it contains $p \cdot \mathbb{Z}^{d+1}$).

How to find a basis?

Let $\Lambda' := \{ h \in \mathbb{F}_p(x) \text{ of deg.} \leq d \mid h \text{ divisible by } a \}$.
This $\mathbb{F}_p$-vector space is generated by

$$a(x),\ x \cdot a(x),\ \dots,\ X^{d-t} \cdot a(x).$$

[redacted] The ref of [redacted] the matrix with rows $a(X), \dots, [redacted] X^{d-t} \cdot a(X)$ [redacted] gives us a basis of Λ'.

A basis of Λ consists of the lifts of these basis vectors (to together with the vectors $p \cdot X^j$ where [redacted] the column corr. to X in the ref has no leading 1. [redacted] of the matrix with rows $X^{d-t} \cdot a(X), p, pX, \dots, pX^d$. sto

~~Note that~~

~~Of course, $g \in \Lambda$~~ If $d = \deg(g)$, then $g \in \Lambda$.

We can use Alg. 13.6 to find an LLL-reduced basis of Λ. By Lemma 13.5, the first basis vector $\tilde{g} \in \Lambda$ has "almost-minimal length", so in particular

$$|\tilde{g}| \leq 2^{d/2} \cdot |g| \leq 2^{d/2} \cdot A =: \tilde{A} \text{ if } d = \deg(g).$$

By definition, both g and $\tilde{g}$ are modulo p divisible by a.

$\Rightarrow \gcd(g \bmod p, \tilde{g} \bmod p) \neq 1$

$\Rightarrow \operatorname{Res}(g, \tilde{g}) \equiv 0 \bmod p.$ (I)

Let's choose $p > (A \cdot \tilde{A})^d$.

$\Rightarrow p > (|g| \cdot |\tilde{g}|)^d \underset{\substack{\uparrow \\ \boxed{\text{Lemma 14.2}}}}{\geq} |\operatorname{Res}(g, \tilde{g})|.$ (II)

(I), (II) $\Rightarrow \operatorname{Res}(g, \tilde{g}) = 0.$

$\Rightarrow \gcd(g, \tilde{g}) \neq 1$

$\underset{\substack{\uparrow \\ g \text{ irreducible}}}{\Rightarrow} g \mid \tilde{g}$ in $\mathbb{Q}(X)$

$\underset{\substack{\uparrow \\ \deg(\tilde{g}) \leq d \leq \deg(g)}}{\Rightarrow} \tilde{g} = \lambda \cdot g$ for some $\lambda \in \mathbb{Q}^\times$.

$\rightsquigarrow$ we have found an irred. factor g of f. Divide f by g and eliminate all mod p factors ~~of~~ dividing g. Then continue ...

If $d < \deg(g)$, then the basis vector $\tilde{g}$ obviously can't divide f.

Total running time: $\tilde{O}(n^{10} + n^{8}(\log B)^{2})$.

$\square$

Some practical remarks:

Rmk 14.4 Since the alg. for Hensel's lemma has near-linear running time but factoring in $\mathbb{F}_p$ has an extra running time factor of $\log p$, it's better to factor $f \bmod p^k$ with $p^k > (A\tilde{A})^d$ and p chosen randomly from an interval large enough to make $f \bmod p$ squarefree.

(this saves time in the factoring mod p^k step in the beginning, but doesn't change the theoretical upper bd. on the r. time

Rmk 14.5 If $(f \bmod p) = lc(f)\, a_1 \cdots a_r$ for small r (with $a_1, \dots, a_r \in \mathbb{F}_p[X]$ monic and irreducible), it can be faster to try for every subset $S \subseteq \{1, \dots, r\}$ whether there is a divisor of f divisible (mod p) exactly by a_i $(i \in S)$, but not by a_i $(i \notin S)$.

To check this, as long as $\frac{p}{2} > lc(f) \cdot A$, it suffices to just try whether the pol. $g \in \mathbb{Z}[X]$ with $g \equiv lc(f) \cdot \prod_{i \in S} a_i \bmod p$ and coeffs. $\in [-\frac{p}{2}, \frac{p}{2}]$ divides f.

But since r can be large, this can have exponential running time (cf. extreme example in section 12).

Rmk 14.6 You can (should) combine 14.4, 14.5.

Rmk 14.7 van Hoeij (Factoring polynomials and the knapsack problem) found another alg. that seems to work better in practice (but without rigorous analysis of the running time): Idea:

For simplicity, assume $lc(f) = 1$.

How can we tell whether the pol. g in Rmk 14.5 has short length ($< A$)?

For a pol. f of degree n with roots $\alpha_1, \dots, \alpha_n$, let

$$Tr^i(f) := \alpha_1^i + \dots + \alpha_n^i \qquad (i = 0, 1, \dots).$$

Note that the coeff. of f are the el. symm. pol. in $\alpha_1, \dots, \alpha_n$, which can be written as pol. in $Tr^i(f)$ $(i = 1, \dots, n)$. Conversely, we can write $Tr^i(f)$ as pol. in the coeff. of f.

Hence, $|f|$ small $\Leftrightarrow$ $\Bigg|\underbrace{\begin{pmatrix} Tr^1(f) \\ \vdots \\ Tr^n(f) \end{pmatrix}}_{Tr(f)}\Bigg|$ small.

Clearly, $Tr(fg) = Tr(f) + Tr(g)$.

Finding a short $g \equiv \prod_{i \in S} a_i \mod p$ corresponds to finding $e_1, \dots, e_n \in \{0, 1\}$ $\quad (e_i = 1 \Leftrightarrow i \in S)$ such that there is a short vector

$$v = \sum_{i \in S} Tr(a_i) + pw = \underbrace{\sum_i e_i Tr(a_i) + pw} \quad \text{with } w \in \mathbb{Z}^n.$$

If we allowed arbitrary $e_i \in \mathbb{Z}$, these would form a lattice.

Rmk Say you know a complex number $r \in \mathbb{R}$ which is approximately an algebraic number. How to find the min. pol. $f \in \mathbb{Z}[X]$?

Say $|f| \leq A$, $|f(r)| \leq B$, $\deg(f) \leq n$.

Look for a short vector of length $\in O(1)$ in the rank $n+1$ lattice

$$\Lambda = \left\{ \left(\underbrace{\frac{f}{A}}_{\in \mathbb{R}^{n+1}}, \underbrace{\frac{f(r)}{B}}_{\in \mathbb{C} \cong \mathbb{R}^2} \right) \;\middle|\; f \in \mathbb{Z}[X] \text{ of deg.} \leq n \right\} \subseteq \mathbb{R}^{n+3}.$$

Rmk You could also use this for a nonrigorous factoring alg.:
Find a complex root r of f and then find its min. pol. g.

15. Primality testing and integer factorization

Rmk If $n = p_1^{e_1} \cdots p_k^{e_k}$, then

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{e_1} \times \dots \times \mathbb{Z}/p_k^{e_k}$$

$$(\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1^{e_1})^\times \times \dots \times (\mathbb{Z}/p_k^{e_k})^\times.$$

Rmk If p is an odd prime and $e \geq 1$, then $(\mathbb{Z}/p^e)^\times$ is isomorphic to the cyclic group $C_{\varphi(p^e)}$ of order $(p-1)p^{e-1} = \varphi(p^e)$.

$\rightsquigarrow (\mathbb{Z}/n\mathbb{Z})^\times \cong C_{(p_1-1)p_1^{e_1-1}} \times \dots \times C_{(p_k-1)p_k^{e_k-1}}$ if n is odd.

~~**Lemma** We can do Given $n \geq 2$, we can determine whether n is a perfect power ($n = m^k$ for some $m \in \mathbb{Z}$, $k \geq 2$) in $\tilde{O}(\log n)$.~~

~~**Pf** For each $2 \leq k \leq \log_2(n)$, compute $\lfloor \sqrt[k]{n} \rfloor$ using Newton's method (in time $\tilde{O}(\log$~~

~~Hence, we can easily assume that n is odd and not a perfect pow.~~

~~**Rmk** Unlike for pol in $\mathbb{F}_q[T]$, we have no efficient way pol time alg. for finding the squarefree factorization of $n \in \mathbb{Z}$, or even to determine whether n is squarefree~~

Rmk Let $n \geq 2$. Then, the set

$$S := \{ a \in (\mathbb{Z}/n\mathbb{Z})^\times \mid a^{n-1} \equiv 1 \bmod n \}$$

forms a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.

In part., either

a) $S = (\mathbb{Z}/n\mathbb{Z})^\times$ or

b) $|S| \leq \frac{1}{2} \cdot |(\mathbb{Z}/n\mathbb{Z})^\times| = \frac{\varphi(n)}{2} < \frac{n}{2}$.

$\left(H \leq G \Rightarrow |H| = \frac{|G|}{[G:H]} \right)$

Def Integers $n \geq 2$ with $S = (\mathbb{Z}/n\mathbb{Z})^\times$ are called Carmichael numbers.

Rmk Any prime is a Carmichael number (little Fermat).

Lemma 15.1 An odd number $n = p_1^{e_1} \cdots p_k^{e_k}$ is a Carmichael number if and only if

$$\varphi(p_i^{e_i}) \mid n-1 \text{ for all } i.$$

Pf

"⇐"

$\varphi(p_i^{e_i}) \mid n-1$

$\Rightarrow a^{n-1} \equiv 1 \bmod p_i^{e_i} \quad \forall i$

$\Rightarrow a^{n-1} \equiv 1 \bmod n$

"⇒" Take any a s.t. a mod $p_i^{e_i}$ generates the cyclic group $(\mathbb{Z}/p_i^{e_i})^\times$ of order $\varphi(p_i^{e_i}) \not\equiv 0 \bmod n-1$. □

Ex $n = 3 \cdot 11 \cdot 17$ is a Carmichael number.

Lemma 15.2 Every Carmichael number is squarefree.

Pf. If $e_i \geq 2$, then $p_i \mid \varphi(p_i^{e_i})$, but $p \nmid n-1$. □

Thm 15.5 The following randomised Monte Carlo alg. detects whether an odd number $n \geq 3$ is Carmichael with a false pos. ~~prob~~. $\leq \frac{1}{2}$ and no false negatives and average running time $\tilde{O}((\log n)^2)$.

Alg

Pick $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ uniformly at random.

Answer Carmichael if $a^{n-1} \equiv 1 \bmod n$. □

Lemma 15.3 We can pick $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ uniformly at random in expected time $\tilde{O}(\cancel{\Sigma}\log n)$.

Alg Pick $a \in \mathbb{Z}/n\mathbb{Z}$ uniformly at random. If $\gcd(a,n) \neq 1$, start over.

Pf The ~~running~~ expected running time is $\tilde{O}\left((\log n) \cdot \frac{n}{\varphi(n)}\right)$. □

Lemma 15.4 We have $\frac{n}{\varphi(n)} \ll \log\log n$ for large n.

Pf $$\frac{n}{\varphi(n)} = \prod_{p \mid n} \frac{1}{1-\frac{1}{p}}$$

$$\Rightarrow \log \frac{n}{\varphi(n)} = \sum_{p \mid n} \log \frac{1}{1-\frac{1}{p}} = \sum_{p \mid n} \left(\frac{1}{p} + \frac{1}{2p^2} + \frac{1}{3p^3} + \dots \right)$$

$$\leq \sum_{p \mid n} \frac{1}{p} + O(1)$$

If K is the largest number s.t. $\prod_{p \leq K} p \leq n$, then

$$\sum_{p \mid n} \frac{1}{p} \leq \sum_{p \leq K} \frac{1}{p} \sim \log\log K$$

with $K \leq \log n + O(1)$. □

Let's look more at the group structure of $(\mathbb{Z}/n\mathbb{Z})^\times$:

Rmk ~~The 2-tor~~ ~~If n~~ For odd $n = p_1^{e_1} \cdots p_u^{e_u}$, the 2-torsion subgroup is

$$\begin{array}{ccc} (\mathbb{Z}/n\mathbb{Z})^\times[2] & \cong & \{\pm 1\} \times \dots \times \{\pm 1\} . \\ \cap & & \cap \qquad\qquad \cap \\ (\mathbb{Z}/n\mathbb{Z})^\times & & C_{\varphi(p_1^{e_1})} \times \dots \times C_{\varphi(p_u^{e_u})} \end{array}$$

cyclic groups of even order

Rmk Assume that n is an odd Carmichael number,

$n - 1 = 2^r \cdot s$ with $r \geq 1$ and odd s.

~~Then, the set~~

~~$T := \{ a \in (\mathbb{Z}/n\mathbb{Z})^\times \mid a^s \equiv 1 \bmod n \text{ or } a^{2^{r-1}s} \equiv -1 \bmod n,$~~
~~$a^{2^i s} \equiv 1 \bmod n \text{ for some } i \in \{1, \dots, r\} \}$~~

~~is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.~~

For $0 \leq j \leq r$, consider the subgroup

$$T_i := \{ a \in (\mathbb{Z}/n\mathbb{Z})^\times \mid a^{2^j s} \equiv 1 \bmod n \}.$$

Clearly, $T_r = (\mathbb{Z}/n\mathbb{Z})^\times$, but $-1 \notin T_0$. Let l be the largest index with $T_l \neq (\mathbb{Z}/n\mathbb{Z})^\times$. Consider the subgroup
(The l is the smallest nr. s.t. $\varphi(p_i^{e_i}) \mid 2^{l+1} s$ for all i.

$$U := \{ a \in (\mathbb{Z}/n\mathbb{Z})^\times \mid a^{2^l s} \equiv \pm 1 \bmod n \}.$$

<u>Lemma 15.6</u> Let n be an odd Carmichael number. We have $U = (\mathbb{Z}/n\mathbb{Z})^\times$ if and only if n is prime.

<u>Pf</u> "$\Leftarrow$" $a^{2^{l+1}s} \equiv (a^{2^l s})^2 \equiv 1 \implies a^{2^l s} \equiv \pm 1$

"$\Rightarrow$" ~~For some prime factor~~ For some i, every $2^{l+1}s$-th power in $(\mathbb{Z}/p_i)^\times$ is 1, but ... $\implies$ Some $2^l s$-th power is -1.

⇒ By the Chin. rem. thm., there is some $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ s.t.

$a \equiv 1 \bmod p_{i'}^{e_{i'}} \quad \forall i' \neq i$

and $a^{2^l s} \equiv -1 \bmod p_i^{e_i}$.

$\Rightarrow a^{2^l s} \not\equiv \pm 1 \bmod n.$ □

<u>Cor 15.7</u> There is a Monte Carlo alg. to determine whether n is prime with false pos. prob. $\leq \frac{1}{2}$, no false neg., avg. running time $\tilde{O}((\log n)$

<u>Alg</u> Pick $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ uniformly at random.

Compute $b = a^s$,

then b^{2^i} for $i = 1, \dots, r$.

If $b^{2^r} \not\equiv 1$, return $\boxed{\text{not prime}}$ (not even Carmichael).

If $b^{2^{i+1}} \equiv 1$ but $b^{2^i} \not\equiv \pm 1$ for some i, return $\boxed{\text{not prime}}$.

Otherwise, return $\boxed{\text{(maybe) prime}}$.

<u>Pf</u> False pos. can only occur when $a \in U \subsetneq (\mathbb{Z}/n\mathbb{Z})^\times$. □

(even just for $i = l$)

→

<u>Rmk</u> There is also ~~a~~ an unconditional deterministic alg. that determines whether n is prime in time $\tilde{O}((\log n)^6)$. (AKS algorithm)

<u>Rmk</u> Assuming the generalised Riemann Hypothesis, $(\mathbb{Z}/n\mathbb{Z})^\times$ is generated by $1, \dots, \lfloor 3(\log n)^2 \rfloor$, so it suffices to check $a = 1, \dots, \lfloor 3(\log n)^2 \rfloor$ for a deterministic primality test (Miller test).

Thm 15.8 There's an alg. that returns a random number $p \le N$ in expected time $\tilde{O}(k \cdot (\log N)^3)$ with $\mathbb{P}(p \text{ not prime}) \le 2^{-k} \log N$. All primes $p \le N$ are equally likely to occur.

Alg Pick $p \le N$ uniformly at random. If Rabin-Miller says "prob. prime" k times, return p. Otherwise, start over.

Pf The number of primes $p \le N$ is $\ge \Omega\left(\frac{N}{\log N}\right)$.

$\Rightarrow$ The alg. makes $\le O(\log N)$ attempts on average.

On each attempt, the prob. of returning a composite no. is $\le \frac{1}{2^k}$.

$\square$

Rmk Many alg. that require choosing a random prime p actually "work" with composite numbers as well:

Either they succeed, or they prove that p is composite (e.g. when trying to divide by a nonzero noninvertible element of $\mathbb{Z}/n\mathbb{Z}$).

For others, you may need to prove primality.

Lemma 15.9 Let $n \geq 3$ be an odd composite integer. Given a uniform random element $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ and its (multiplicative) order $\operatorname{ord}(a)$ (or the size $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$), we can with prob. $\geq \frac{1}{4}$ find a proper divisor $1 < d < n$ of n in time $\tilde{O}((\log n)^2)$.

Pf Write $\varphi(n) = 2^r s$ and $\operatorname{ord}(a) = 2^t u$.

($\operatorname{ord}(a) | \varphi(n) \Rightarrow t \leq r$ and $u | s$.)

~~Consider the map~~

~~$f: (\mathbb{Z}/n\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/n\mathbb{Z})^\times$~~

~~$x \longmapsto x^s$~~

~~Under the CRT, it corr. to~~

~~$\prod_i (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^\times \longrightarrow \prod_i \{\pm 1\}$~~

~~$(x_i)_i \longmapsto (x_i^s)_i$~~

Claim: With prob. $\geq \frac{1}{4}$, of the numbers $d_i = \gcd(a^{2^i u} - 1, n)$ for $i = 0, \ldots, t-1$, a proper divisor.

Pf As before, let c be the smallest nr. s.t. $\varphi(p_i^{e_i}) | 2^{c+1} s \quad \forall i$.

Let $\varphi(p_i^{e_i}) \nmid 2^c s$ and let $j \neq i$.

$\Rightarrow$ We have hom.

$$f_i : (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^\times \twoheadrightarrow \{\pm 1\}, \quad x \mapsto x^{2^c s} \qquad , \qquad f_j : (\mathbb{Z}/p_j^{e_j}\mathbb{Z})^\times \to \{\pm 1\}, \quad x \mapsto x^{2^c s}$$

with surjective f_i.

$$\left.\begin{array}{l} \text{With prob. } \frac{1}{2}, \ f_i(a) = -1 \ \underset{\Rightarrow a^{2^c u} \equiv -1 \bmod p_i^{e_i}}{} \\ \text{With prob. } \geq \frac{1}{2}, \ f_j(a) = +1 \ \underset{\Rightarrow a^{2^c u} \equiv +1 \bmod p_j^{e_j}}{} \end{array}\right\} \text{independent by CRT}$$

$\Rightarrow$ With prob. $\geq \frac{1}{4}$, $\gcd(a^{2^c u} - 1, n)$ is divisible by p_i, but not by p_i.

How to determine the mult. order of an element $a \in (\mathbb{Z}/n\mathbb{Z})^\times$?

Thm 15.10 (Baby-step giant-step alg.)

Assume we can perform arithmetic in the group G in $O(1)$ and we can compare elements w.r.t. some total order on G in $O(1)$. We can compute the order $k < \infty$ of a (torsion) element $a \in G$ in time $O(\sqrt{k} \log k)$ with memory $O(\sqrt{k})$.

Idea Let $w > \sqrt{k}$.

Write $k = iw + j$ with $1 \leq j \leq w$, $0 \leq i \leq w-1$.

$$a^k = 1 \iff \underset{\substack{\| \\ (\underbrace{a^w}_{\text{giant step}})^i}}{a^{iw}} = \underbrace{a^{-j}}_{\text{baby step}}$$

Alg

```
For $e = 1, 2, \dots$:
    Let $w = 2^e$.
    Compute $a^{-j}$ for $j = 1, \dots, w$ and save the pairs $(a^{-j}, j)$
    in a binary search tree (BST).
    For $i = 0, 1, \dots, w-1$:
        Compute $(a^w)^i$. If there exists some $j$ in the BST
        with $a^{-j} = (a^w)^i$, return the smallest such $iw + j$.
```

Rmk Better to use a hashtable...

Rmk Combining this with Lemma 15.9, we can find a nontriv. factor of a composite integer n in $\tilde{O}(\sqrt{n})$.

"Yay..."

Problems 1) BSGS alg. too slow. There are better algorithms for the group $(\mathbb{Z}/n\mathbb{Z})^\times$ (e.g. the index calculus algorithm).

2) $(\mathbb{Z}/n\mathbb{Z})^\times$ too large. The class group of $\mathbb{Q}(\sqrt{-n})$ has just order $O(\sqrt{n})$. Its 2-torsion elements "correspond to"

divisors of n. (Shanks's class group method).

Rmk On a quantum RAM, we can compute the mult. order of any $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ in time polynomial in $\log n$. (Shor's algorithm)

Lemma 15.11 Let $n \geq 2$, let p be a prime dividing n and let $t \geq 1$ such that $p-1 \mid t!$. Then, the following alg. returns a divisor $d \geq 2$ of n in time $\tilde{O}(t \log n)$.

Alg (Pollard's $p-1$ alg.)

```
Pick $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ at random.
For $k = 1, 2, \dots$:
    Compute $a^{k!} \bmod n$
    $\quad\overset{\shortparallel}{(a^{(k-1)!})^k}$
    If $d := \gcd(a^{k!} - 1, n) > 1$, return $d$.
```

Pf $p-1 \mid k! \Rightarrow \operatorname{ord}(a \bmod p) \mid k! \Rightarrow a^{k!} \equiv 1 \bmod p$
$\Rightarrow p \mid d$. $\square$

Problems 1) The alg. might return the trivial divisor $d = n$. (If $n = pq$ and $(p-1 \mid t! \Leftrightarrow q-1 \mid t!)$, this could happen for many $a \in (\mathbb{Z}/n\mathbb{Z})^\times$.)

2) t could be large:
E.g. if $p-1 = 2q$ for a prime q, then we need $t \geq q$, which could be $\Omega(\sqrt{n})$ even for the smallest prime factor p of n.

Rmk We can get rid of the problems by replacing $(\mathbb{Z}/n\mathbb{Z})^\times$ by groups $E(\mathbb{Z}/n\mathbb{Z})$ for elliptic curves E. (Lenstra's elliptic curve method)

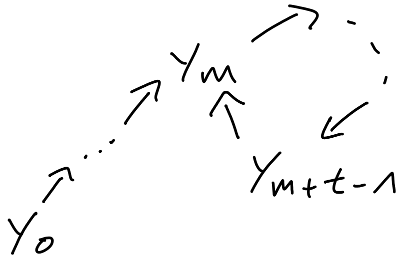
15.1. Pollard's rho algorithm (cf. Cohen)

Lemma 15.1.1 Let $f: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ be a uniformly random map. Let M, T be the preperiod and period of the sequence

$$1, f(1), f(f(1)), \dots \qquad (y_i = f^i(1)).$$

We have $\mathbb{E}(M+T) \asymp \sqrt{n}$.

Pf (sketch)



$$\mathbb{P}(M=m, T=t) = \underset{\substack{\uparrow \\ \mathbb{P}(y_k \neq y_0, \dots, y_{k-1})}}{\left(\prod_{k=1}^{m+t-1} \left(1-\frac{k}{n}\right)\right)} \cdot \underset{\substack{\uparrow \\ \mathbb{P}(f(y_{m+t-1}) = y_m)}}{\frac{1}{n}}$$

$$\sum_{k=1}^{m+t-1} \log\left(1-\frac{k}{n}\right) \approx -\sum_{k} \frac{k}{n} \approx -\frac{(m+t)^2}{2n}$$

$$\Rightarrow \mathbb{P}(M=m, T=t) \approx e^{-(m+t)^2/2n} \cdot \frac{1}{n}.$$

$$\Rightarrow \mathbb{E}(M+T) = \sum_{m,t} \mathbb{P}(M=m, T=t)\cdot(m+t)$$

$$\approx \sum_{m,t} e^{-(m+t)^2/2n}\cdot(m+t)\cdot\frac{1}{n}$$

$$\approx \int_1^\infty\int_1^\infty e^{-(m+t)^2/2n}\cdot(m+t)\cdot\frac{1}{n}\,dm\,dt$$

$$\underset{\substack{\uparrow\\ m=a\sqrt{n}\\ t=b\sqrt{n}}}{\approx} \underbrace{\int_0^\infty\int_0^\infty e^{-(a+b)^2/2}(a+b)}_{\in(0,\infty)}\cdot\sqrt{n}\,da\,db$$

$$\sim \sqrt{n}.$$

"D"

Thm 15.1.2 Let $n = p_1^{e_1}\cdots p_k^{e_k}$ (with $p_1^{e_1}<\ldots<p_k^{e_k}$), $k\geq 2$. Assume $f_1,\dots,f_k$ are (independent) uniformly random functions, $f_i: \mathbb{Z}/p_i^{e_i}\mathbb{Z}\to\mathbb{Z}/p_i^{e_i}\mathbb{Z}$. They give rise to a function $f:\mathbb{Z}/n\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$. Assuming we can evaluate f in $\mathcal{O}(1)$, the following alg. returns a divisor $1<d\leq n$ of n in expected time $\mathcal{O}(\sqrt{p_1^{e_1}}\log n)$. With probability $>\varepsilon$, we have $d<n$. (For some constant $\varepsilon>0$.)

```
Alg  Let $a = f(0 \bmod n)$, $b = f(a)$.
  For $j = 1, 2, \dots$:
    (Now, $a = f^j(0)$, $b = f^{2j}(0)$.)
    If $d := \gcd(a-b, n) > 1$, return $d$.
    Let $a \leftarrow f(a)$, $b \leftarrow f^2(b)$.
```

Pf

Let m_i, t_i be the preperiod and period of 0 for the function f_i. We have $p_i^{e_i} \mid f^j(0) - f^{2j}(0)$ if and only if $t_i \mid j$ and $j \geq m_i$.

$\Rightarrow$ The number of steps taken by the alg. is at most the smallest multiple of t_1 which $\geq m_1$, which is $\leq t_1 + m_1$, which on average is $O(\sqrt{p_1^{e_1}})$.

The prob. that the smallest j s.t. $t_i \mid j$ and $j \geq m_i$ is the same number for all i is $< 1 - \varepsilon$ for some constant $\varepsilon > 0$. $\square$

Rmk We don't know how to generate a random function f as in the Thm.

Instead, usually the following heuristic is used:

Take $f(X) = X^2 + c$ for a random (fixed) number $c \in \mathbb{Z}/n\mathbb{Z}$.

15.2. Dixon's random squares method

Reference: – Chapter 19.5 in "Modern Computer Algebra"
– Dixon: Asymptotically fast factorization of integers

Reminder.

Lemma 15.2.1 Let $n = p_1^{e_1} \cdots p_k^{e_k}$ odd $(k \geq 2)$. Let a be a uniformly random element of $(\mathbb{Z}/n\mathbb{Z})^\times[2] \cong \underbrace{\{\pm 1\} \times \ldots \times \{\pm 1\}}_{k}$. Then,

$$1 < \gcd(a-1, n) < n \text{ with probability } 1 - \frac{1}{2^{k-1}} \geq \frac{1}{2}.$$

Pf There are exactly 2 bad a:

$a = 1$: $\gcd = n$

$a = -1$: $\gcd = 1$. □

How to construct a 2-torsion element:

Find $a, b \in (\mathbb{Z}/n\mathbb{Z})^\times$ with $a^2 \equiv b^2 \bmod n$. Then, $\frac{a}{b} \in (\mathbb{Z}/n\mathbb{Z})^\times[2]$.

Rmk Birthday paradox: need to choose $\sim\sqrt{|(\mathbb{Z}/n\mathbb{Z})^\times[2]|}$ random elements $a_i \in (\mathbb{Z}/n\mathbb{Z})^\times$ until finding two with the same square.

Idea Try to find numbers $a_1, \ldots, a_r$ mod n s.t.

$$\underbrace{(a_1^2 \bmod n)}_{\in\{1,\ldots,n\}} \cdots \underbrace{(a_r^2 \bmod n)}_{\in\{1,\ldots,n\}} = b^2 \text{ for an integer } b.$$

(!!!)

This is equivalent to the condition that the LHS is divisible by every prime q an even number of times.

We'll only allow a_i such that $(a_i^2 \bmod n) \in \{1, \ldots, n\}$ has only prime factors $q < B$.

(small)

~~$\frac{\log n}{\epsilon} + (\log\log n)(2\epsilon+2)$~~

~~Alg 15.2.2~~

~~Given: odd integer $n \geq 3$ composite, not a prime power,~~

~~$\frac{\log n}{\log\log n} \approx 4\epsilon^2$~~

~~some $B > 2$.~~

~~1) compute the primes $p_1, \dots, p_r \leq B$.~~

__Lemma 15.2.2__ Let $n = p_1^{e_1} \cdots p_u^{e_u}$ odd, $k \geq 2$.

Let $q_1 < \dots < q_r$ be (small) primes numbers [not dividing n].

The following alg. returns ~~[illegible]~~

~~[illegible]~~

a uniformly random element of $(\mathbb{Z}/n\mathbb{Z})^\times[2]$.

__Alg (15.2.2)__

1) Find uniformly random elements $b_1, \dots, b_{r+1} \in (\mathbb{Z}/n\mathbb{Z})^\times$ subject to the condition that each $(b_i^2 \bmod n)$ ~~[illegible]~~ can be written as $(b_i^2 \bmod n) = q_1^{f_{i1}} \cdots q_r^{f_{ir}}$:

Just pick random $b \in (\mathbb{Z}/n\mathbb{Z})^\times$, compute $b^2 \bmod n$, find the number of times this integer is divisible by each q_j by trial division, until you've found $r+1$ good residue classes b_i. [This terminates because $b = 1$ is possible.]

2) Using Gaussian elimination, find the kernel of the map $\mathbb{F}_2^{r+1} \longrightarrow \mathbb{F}_2^r$

$$(v_i)_{i=1,\dots,r+1} \mapsto \left(\sum_i v_i f_{ij} \right)_{j=1,\dots,r}$$

3) Pick a uniformly random ~~[illegible]~~ nonzero element $v = (v_i)_i \in \mathbb{F}_2^{r+1}$ of the kernel. Let $\underset{\neq \emptyset}{S} = \{ i \mid v_i = 1 \} \subseteq \{1, \dots, r+1\}$.

$$\Big(\Rightarrow \underset{\substack{\| \\ \sum_{i \in S} f_{ij}}}{\sum v_i f_{ij}} \equiv 0 \bmod 2 \text{ for all } j. \Big)$$

4) Let $t_j = \frac{1}{2} \sum_{i \in S} f_{ij}$. ~~[illegible]~~

$$\Big(\Rightarrow \prod_{i \in S} (b_i^2 \bmod n) = \prod_{i \in S} \prod_j q_j^{f_{ij}} = \prod_j q_j^{2 t_j} = \Big(\prod_j q_j^{t_j} \Big)^2 . \Big)$$

5) Return $c := \dfrac{\prod_{i \in S} b_i}{\prod_j q_j^{t_j}} \bmod n$.

Pf The result c is ~~[illegible]~~ an el. of $(\mathbb{Z}/n\mathbb{Z})^\times[2]$ because

$$\Big(\prod_{i \in S} b_i \Big)^2 \equiv \Big(\prod_j q_j^{t_j} \Big)^2 \bmod n.$$

~~We'll show that for any fixed set $S \neq \emptyset$, all elements of $(\mathbb{Z}/n\mathbb{Z})^\times[$ are equally likely. Fix any $i_0 \in S$.~~

~~Fix the set S and any $i_0 \in S$ and the value [illegible] $d = (b_{i_0}^2 \bmod n)$. Also fix all b_j with $j \neq i_0$. $\Rightarrow b_{i_0}$ is determined by its square d up to mult. by an el. of $(\mathbb{Z}/n\mathbb{Z})$ Then [illegible] All square roots of $d \bmod n$ are equally likely to be the value of b_{i_0} [illegible]~~

~~$\Rightarrow$ All [illegible] elements of $(\mathbb{Z}/n\mathbb{Z})^\times[2]$ [illegible] [illegible] occur with the same probability.~~

$\square$

~~We'll show that for any fixed S, $i_0 \in S$, b~~

We'll show that c is a uniformly random element of $(\mathbb{Z}/n\mathbb{Z})^\times[2]$, even for

~~S, $i_0 \in S$, b_i for all $i \neq i_0$, $d_{i_0} = (b_{i_0}^2 \bmod n)$~~

~~Note that S can be determined from~~

particular fixed values $d_i = (b_i^2 \bmod n)$.

Note that the set S only depends on these values (and randomness) not on the square roots b_i of d_i.

The b_i are uniformly random square roots of the d_i.

$\Rightarrow$ $\prod_{i \in S} b_i$ is a uniformly distributed random square root of $\prod_{i \in S} d_i$ (even if we pick $i_0 \in S$ and fix all b_i with $i \neq i_0$).

[

Rmk We could have chosen v (and therefore S) deterministically, as long as the choice only depends on $d_1, \dots, d_{r+1}$, not on $b_1, \dots, b_{r+1}$.

Question ~~What fraction~~ For what fraction of elements $b \in (\mathbb{Z}/n\mathbb{Z})^\times$ can $(b^2 \bmod n)$ be written as $q_1^{f_1}\cdots q_r^{f_r}$?

I.e. how long does step 1 take?

Remark Assume $q_1 < \dots < q_r$ and $q_r^t \le n$. Then,

$$\#\{1 \le a \le n \mid a = q_1^{f_1}\cdots q_r^{f_r} \text{ for some } f_1,\dots,f_r \ge 0\}$$

$$\ge \#\{(f_1,\dots,f_r) \mid f_1,\dots,f_r \ge 0, \quad f_1+\dots+f_r \le t\}$$

$$= \binom{t+r}{t} \ge \frac{r^t}{t!} \qquad (\text{“close to } \tfrac{n}{t!}\text{”})$$

But we need to prove that many of these $1 \le a \le n$ are invertible quadratic residues mod n.

Idea (quadr. nonres. mod p_i)·(quadr. nonres) = (quad. res

Lemma 15.2.3 Assume $q_1 < \dots < q_r$ and $q_r^{2t} \le n$ and that no q_i divides n. Then,

~~$\#\{a \in (\mathbb{Z}/n\mathbb{Z})^{\times 2} \mid$~~ ~~quadr. res.~~

~~$\#\{1 \le a \le n \mid a = q_1^{f_1}\cdots q_r^{f_r}$ for some $f_1,\dots,f_r \ge 0\}$ and~~

$$\#\{b \in (\mathbb{Z}/n\mathbb{Z})^\times \mid (b^2 \bmod n) = q_1^{f_1}\cdots q_r^{f_r} \text{ for some } f_1,\dots,f_r \ge 0\}$$

$$\ge \frac{r^{2t}}{(2t)!}. \qquad (\text{“close to } \tfrac{n}{(2t)!}\text{”})$$

Write $n = p_1^{e_1} \cdots p_r^{e_r}$. $\leadsto (\mathbb{Z}/n\mathbb{Z})^\times \cong C_{\varphi(p_1^{e_1})} \times \cdots \times C_{\varphi(p_r^{e_r})}$.

<u>Pf</u> Consider the map

$$\chi : (\mathbb{Z}/n\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/n\mathbb{Z})^\times / \underset{\substack{\uparrow \\ \text{quadr. res}}}{(\mathbb{Z}/n\mathbb{Z})^{\times 2}} \cong \underbrace{C_2 \times \cdots \times C_2}_{k} =: G$$

with kernel $(\mathbb{Z}/n\mathbb{Z})^{\times 2}$.

For any $g \in G$, let

$$U_g := \chi^{-1}(g).$$

If $a_1, a_2 \in U_g$, then $\chi(a_1 a_2) = \chi(a_1)\chi(a_2) = g^2 = e$, so $a_1 a_2 \in \ker = (\mathbb{Z}/n\mathbb{Z})^{\times 2}$ is a quadratic residue, with 2^k square roots.

Let $T := \{ a = q_1^{f_1} \cdots q_r^{f_r}$ for some $f_1, \dots, f_r \geq 0$ with $f_1 + \dots + f_r \le t \}$
$(\Rightarrow 1 \le a \le \sqrt{n})$

If $a_1, a_2 \in U_g \cap T$, then

$$a_1 a_2 \in W := \left\{ 1 \le a \le n \;\middle|\; \begin{array}{l} a = q_1^{f_1} \cdots q_r^{f_r} \\ f_1 + \dots + f_r \le 2t \\ a \in (\mathbb{Z}/n\mathbb{Z})^{\times 2} \end{array} \cdots \right\}.$$

Hence, we obtain a map

$$\rho : \coprod_{g \in G} (U_g \cap T) \times (U_g \cap T) \longrightarrow W.$$
$$(a_1, a_2) \longmapsto a_1 a_2$$

[Crude estimate:]
Any $a = q_1^{f_1} \cdots q_r^{f_r} \in W$ (with $f_1 + \dots + f_r \le 2t$)
has at most $\binom{2t}{t} = \frac{(2t)!}{t!^2}$ preimages (choose which t of the $2t$ prime factors of a go into a_1).

$$\Rightarrow \sum_{g\in G} |U_g \cap T|^2 \le |W| \cdot \frac{(2t)!}{t!^2}.$$

AM-QM ineq: $$\left(\frac{\sum_{g\in G} |U_g\cap T|}{|G|}\right)^2 \le \frac{\sum_{g\in G}|U_g\cap T|^2}{|G|}.$$

$$\boxed{(\mathbb{Z}/n\mathbb{Z})^\times = \bigsqcup_{g\in G} U_g} \rightarrow \left(\frac{\sum_{g\in G} |U_g\cap T|}{|G|}\right)^2 = \left(\frac{|T|}{|G|}\right)^2$$

$$\frac{\sum_{g\in G}|U_g\cap T|^2}{|G|} \le \frac{|W|}{|G|}\cdot\frac{(2t)!}{t!^2}$$

$$\Rightarrow \#\{b\in(\mathbb{Z}/n\mathbb{Z})^\times \mid (b^2 \bmod n) = q_1^{f_1}\cdots q_r^{f_r} \cdots\}$$

$$\ge 2^k\cdot|W| \ge \frac{2^k\cdot|T|^2}{\cancel{|G|}\cdot(2t)!/t!^2} = \binom{t+r-1}{t}^2\cdot\frac{t!^2}{(2t)!}$$

$$\ge \frac{r^{2t}}{t!^2}\cdot\frac{t!^2}{(2t)!} = \frac{r^{2t}}{(2t)!}$$

□

Reminder

Alg: 0) Find all primes $q_1 < \dots < q_r \leq B$ (for a number B to be chosen later) and check that no q_i divides n (otherwise, we're done)

1) Find good $b_1, \dots, b_{r+1} \in (\mathbb{Z}/n\mathbb{Z})^\times$ s.t. $(b_i^2 \bmod n) = q_1^{f_{i1}} \cdots q_r^{f_{ir}}$ by trying $b \in \mathbb{Z}/n\mathbb{Z}$ at random until finding $r+1$ good ones.

2) Compute the kernel of the $r \times (r+1)$-matrix $(f_{ij} \bmod 2)_{j,i}$ over $\mathbb{F}_2$ using Gaussian elimination

3)-5) simple stuff...

Running time: Let $B = n^{1/2t}$ ~~for~~ t to be chosen optimally later..
$(\Rightarrow q_r^{2t} \leq n)$

$$\Rightarrow r \asymp \frac{B}{\log B}$$

Step 0 takes time $\tilde{O}(B + r \log n) \leq \tilde{O}(B \log n)$.

~~Step 1 takes time~~

A random $b \in \mathbb{Z}/n\mathbb{Z}$ is good with probability

$$\gtrsim \frac{\frac{r^{2t}}{(2t)!}}{n} \geq \frac{\left(\frac{r}{2t}\right)^{2t}}{n} \asymp \frac{\left(\frac{B}{2t \log B}\right)^{2t}}{n} = \frac{\frac{n}{(\log n)^{2t}}}{n} = \frac{1}{(\log n)^{2t}}$$ by

Lemma 15.2.3.

$\Rightarrow$ On average, we need to try $\ll (\log n)^{2t}$ random b to find a good one.

$\Rightarrow$ Need to check $\ll r (\log n)^{2t} \leq B (\log n)^{2t}$ to find $r+1$ good ones.

Checking whether some $b \in \mathbb{Z}/n\mathbb{Z}$ is good takes time

$$\tilde{O}\Big(\underbrace{(r + \underset{\substack{\uparrow \\ \sum_j f_{ij} \le \log n}}{\log n})}_{\text{nr. of trial divisions}} \log n\Big) \underset{\uparrow}{\le} \tilde{O}(r \log n) \subseteq \tilde{O}(B \log n)$$

We'll choose t so that $r \gg \log n$.

$\Rightarrow$ Step 1 takes time $\tilde{O}(B(\log n)^{2t} \cdot B \log n)$

$$= \tilde{O}(B^2 (\log n)^{2t+1}) \text{ on average.}$$

Step 2 takes time $O(r^3) \subseteq O(B^3)$.

Steps 3-5 take time $\le \tilde{O}(r^2 \log\log n) \subseteq \tilde{O}(B^2 \log\log n)$.

$\Rightarrow$ Total time $\ll \tilde{O}(B^2(\log n)^{2t+1} + B^3)$

$$= \tilde{O}(n^{1/t}(\log n)^{2t+1} + n^{3/2t})$$

$$= \tilde{O}\Big(\exp\big((\log n)\cdot\tfrac{1}{t} + (\log\log n)\cdot(2t+1)\big) + \exp\big((\log n)\cdot\tfrac{3}{2t}\big)\Big)$$

Choose t to minimise the first summand:

$$\mathbb{Z}_{\ge 1} \ni t = \sqrt{\frac{\log n}{2\log\log n}} + O(1).$$

$\leadsto$ Total time

$$\tilde{O}\Big(\exp\Big((\log n)\cdot\sqrt{\frac{2\log\log n}{\log n}} + (\log\log n)\cdot 2\sqrt{\frac{\log n}{2\log\log n}}\Big) + \exp\Big((\log n)\cdot\frac{3}{2}\cdot\sqrt{\frac{2\log\log n}{\log n}}\Big)\Big)$$

$$= \tilde{O}\Big(\exp\big(2\sqrt{2}\cdot\sqrt{(\log n)(\log\log n)}\big)\Big).$$

⇒ **Thm 15.2.4** We can find a nontrivial divisor of a composite integer n in expected time

$$\tilde{O}\left(\exp\left(C\sqrt{(\log n)(\log\log n)}\right)\right),$$

where $C = 2\sqrt{2}$.

Rmk $\exp(C(\log n)^{\delta}) \ll_{\delta,\varepsilon} n^{\varepsilon} \quad \forall\, \delta < 1,\ \varepsilon > 0$

(subexponential in $\log n$)

$\exp(C(\log n)^{\delta}) \gg (\log n)^{k} \quad \forall\, \delta > 0,\ k \geq 0$

(superpolynomial in $\log n$)

Improvements

1) Make step 1 (finding good $(b_i^2 \bmod n)$) faster (cf. HW).

2) Make step 2 (Gaussian elimination) faster:

Since$_{n\geq}$ $(b_i^2 \bmod n) = q_1^{f_{i1}} \cdots q_r^{f_{ir}}$, we have $\sum_j f_{ij} \leq \mathcal{O}(\log n)$, which is way smaller than r.

$\Rightarrow$ The matrix $(f_{ij})_{i,j}$ is sparse.

Can find the kernel more quickly using Wiedemann's alg

3) Improve the estimate in Lemma 15.2.3.

4) Heuristic improvement:

Instead of using $(b^2 \bmod n)$ for arbitrary b, use $(b^2 \bmod n)$ for $b = \lceil\sqrt{n}\rceil + c$, where $0 \leq c \ll n^{\varepsilon}$. $\Rightarrow b^2 = \lceil\sqrt{n}\rceil^2 + 2\lceil\sqrt{n}\rceil c + c^2 \approx n$

$\Rightarrow (b^2 \bmod n) = (\lceil\sqrt{n}\rceil^2 - n) + 2\lceil\sqrt{n}\rceil c + c^2$ | for suff. small c

$\ll n^{\frac{1}{2}+\varepsilon}$.

A random number $\ll n^{\frac{1}{2}+\varepsilon}$ is more likely only divisible by small primes than a random number $\leq n$.

$\rightsquigarrow$ Heuristic running time $\mathcal{O}_\delta(\exp((1+\delta)\sqrt{(\log n)(\log\log n)}))$ $\forall \delta >$

Q Why not simply use $(b^2 \bmod n) = b^2$ for $0 \leq b \ll n^{\varepsilon}$?

A If nothing is actually reduced mod n, the alg. always returns the trivial result 1.

useless

Rmk Lenstra–Pomerance proved in 1992 that another alg. has expected running time $\tilde{\mathcal{O}}_\delta(\exp((1+\delta)\sqrt{(\log n)(\log\log n)}))$ $\forall\delta$

That's the best proven running time.

But:

Rmk The general number field siève has heuristi running time

$$O_\delta\left(\exp\left((C+\delta)(\log n)^{1/3}(\log\log n)^{2/3}\right)\right)$$

with $C=(64/9)^{1/3}$.

16. Number fields

several ways of specifying a field ext. $L|K$:

a) $L = K[X]/f(X)$, $f(X) \in K[X]$

L is a field if and only if $f(X)$ is irreducible.

L is a product of fields if and only if $f(X)$ is squarefree.

$$\underset{\dim_K(L)}{\overset{}{[L:K]}} = \deg(f) =: n.$$

b) give the multiplication table:

For a basis $\omega_1, \dots, \omega_n$ of L as a K-vector space, specify the numbers $a_{ijk} \in K$ such that

$$\omega_i \omega_j = \sum_k a_{ijk} \omega_k.$$

With respect to this basis, the mult. by ω_i map is given by the matrix $M_i = (a_{jk})_{k,j}$.

Rmk In a), a basis of $L|K$ is $1, X, \dots, X^{n-1}$, so el. of L corr. to pol. $g(X) \in K[X]$ of degree $< n$.

W.r.t. the basis $1, \dots, X^{n-1}$, the coeff. in the mult. table are given by

$$\left(X^{(i-1)+(j-1)} \bmod f\right) = \sum_k \underset{\substack{\uparrow \\ \text{only depends on } i+j}}{a_{ijk}} X^{k-1}$$

If $i-1+j-1 < n$, then $a_{ijk} = \begin{cases} 1, & k-1 = i-1+j-1 \\ 0, & \text{otherwise.} \end{cases}$

16.1. Rings of integers

References: – Cohen, Chapter 6.1

– Pohst, Chapter V

Let $K = \mathbb{Q}$, $L = \mathbb{Q}[X]/(f)$ a degree n number field with $f \in \mathbb{Z}[X]$ mon irreducible.

In other words, $L = \mathbb{Q}(\alpha)$ for a root α of f.

Q How to determine the ring of integers $\mathcal{O}_L$?

(i.e.: a basis of $\mathcal{O}_L$ as a $\mathbb{Z}$-module).

Rmk f monic, $f(\alpha)=0 \Rightarrow \alpha \in \mathcal{O}_L \Rightarrow \mathbb{Z}[\alpha] \subseteq \mathcal{O}_L$

$\mathbb{Z}[\alpha] \subseteq \mathcal{O}_L$ is an order: a subring of $\mathcal{O}_L$ of finite index.

$$\overset{\displaystyle \mathbb{Z}[\alpha]}{\parallel} \atop \mathbb{Z}+\mathbb{Z}\alpha+\dots+\mathbb{Z}\alpha^{n-1}$$

Rmk $\mathrm{disc}(\mathbb{Z}[\alpha]) = \mathrm{disc}(f)$

$\mathrm{disc}(\mathcal{O}_L) = \mathrm{disc}(L)$

Rmk For any orders $R \subseteq S \subseteq \mathcal{O}_L$, we have

$$\mathrm{disc}(R) = \mathrm{disc}(S) \cdot [S:R]^2.$$

(In particular, if $\mathrm{disc}(f)^{\in\mathbb{Z}}$ is squarefree, then $\mathbb{Z}[\alpha] = \mathcal{O}_L$.)

Def Let p be a prime number.

An order $R \subseteq \mathcal{O}_L$ is *p-maximal* if there is no $a \in \mathcal{O}_L$ such that $a \notin R$ but $p\,a \in R$.

Lemma 16.1.1 R is p-max. if and only if $p \nmid [\mathcal{O}_L : R]$.

Pf $\mathcal{O}_L/R$ is a finite abelian group of order $[\mathcal{O}_L : R]$. [redacted] it contains an element (a mod R) of ~~order~~ p if and only if $p \mid [\mathcal{O}_L : R]$. $\square$

[redacted] **Rmk** In particular, R [redacted] is p-maximal for all p with $p^2 \nmid$ [redacted] $\mathrm{disc}(R)$.

[redacted]

Cor 16.1.2 We have $R = \mathcal{O}_L$ (R is *maximal*) if and only if R is p-maximal for all p.

~~Thus, $\mathcal{O}_L/R$ is a finite abelian group of order k divisible by p. $\Rightarrow$ It contains an element of order p, say (a mod R).~~

~~$\Rightarrow$ $pa \in R$, but $a \notin R$.~~ ~~$\square$~~

Ex Let $t \in \mathbb{Z}$ be not a square.

$\Rightarrow f(X) = X^2 - t$ is irreducible, $L = \mathbb{Q}(X)/f = \mathbb{Q}(\sqrt{t})$.

W.l.o.g. $\alpha = \sqrt{t}$.

$\operatorname{disc}(f) = 4t$

$\|$

$\operatorname{disc}(\mathbb{Z}[\alpha])$

a) Let $p \neq 2$. Then, $\mathbb{Z}(\alpha)$ is p-maximal iff $p^2 \nmid t$:

If $p^2 \mid t$, then $\frac{\alpha}{p} \in \mathcal{O}_L$ (with min. pol. $X^2 - \frac{t}{p^2}$),

$p \cdot \frac{\alpha}{p} \in \mathbb{Z}(\alpha)$, $\quad \frac{\alpha}{p} \notin \mathbb{Z}[\alpha]$.

b) $\mathbb{Z}(\alpha)$ is 2-maximal iff $t \equiv 2, 3 \bmod 4$:

If $t \equiv 0 \bmod 4$, then $\frac{\alpha}{2} \in \mathcal{O}_L$ like before.

If $t \equiv 1 \bmod 4$, then $\frac{1+\alpha}{2} \in \mathcal{O}_L$ (with min. pol. $X^2 - X - \frac{t-1}{4} \in \mathbb{Z}(X)$)

If $t \equiv 2,3 \bmod 4$, then the min. pol. of $\frac{r+s\alpha}{2}$ with $r,s \in$

is $X^2 - rX - \frac{s^2t - r^2}{4}$, which only lies in $\mathbb{Z}(X)$ if r,s are both even.

$\|$

$(X - \frac{r}{2})^2 - \frac{s^2 t}{4}$

Then, $\frac{r+s\alpha}{2} \in \mathbb{Z}[\alpha]$.

Rmk: The method used in the example in principle works for any number field (and can even be used to compute the ring of integers).

$\mathbb{Z}[\alpha]$ is p-max. if and only if the min. pol. of $v := \frac{1}{p}(r_0 + r_1 \alpha + \dots + r_{n-1} \alpha^{n-1})$ with $r_0, \dots, r_{n-1} \in \mathbb{Z}$ has integer coefficients only when $r_0, \dots, r_{n-1}$ are all divisible by p.

Note: Since $\mathbb{Z}[\alpha] \subseteq \mathcal{O}_L$, whether $v \in \mathcal{O}_L$ only depends on the values $r_i \bmod p$, so it suffices to check p^n tuples $(r_0, \dots, r_{n-1}) \in \mathbb{F}_p^n$.

($\rightsquigarrow$ exponential running time in both n and $\log p$).

Better approach:

Def The radical of an ideal I of a ring R is the ideal

$$\mathrm{rad}(I) := \{ r \in R \mid r^k \in I \text{ for some } k \geq 1 \}.$$

Rmk $I \subseteq \mathrm{rad}(I) \subseteq R$.

Rmk If R is noetherian, then $\mathrm{rad}(I)^m \subseteq I$ for some $m \geq 1$.

Ex $\mathrm{rad}(\underbrace{p_1^{e_1} \cdots p_u^{e_u} \mathbb{Z}}_{\subseteq \mathbb{Z}}) = p_1 \cdots p_u \mathbb{Z}$.

Lemma 16.1.3 Let R be an order in a number field of degree n.
Then, $J_p(R) := \mathrm{rad}(pR) = \{ r \in R \mid r^v \in pR \}$ for any $v \geq n$.

Pf "$\supseteq$" clear

"$\subseteq$" As a $\mathbb{Z}$-module, $R \cong \mathbb{Z}^n$.

$\Rightarrow$ R/pR is an n-dimensional $\mathbb{F}_p$-vector space.

Let $r \in J_p(R)$. $\Rightarrow$ $r^k \in pR$ for some $k \geq 1$.

$\Rightarrow$ The mult. by r map $m_r : R/pR \longrightarrow R/pR$ is nilpotent.

$\Rightarrow$ Its n-th power is the zero map.

$\Rightarrow$ $r^n \in pR$

$\Rightarrow$ $r^v \in pR \ \forall v \geq n$. $\square$

Rmk $$\begin{aligned} R/pR &\longrightarrow R/pR \\ x &\longmapsto x^{p^s} \end{aligned}$$ is an $\mathbb{F}_p$-linear map for all $s \geq 0$.

If $p^s \geq n$, then $J_p(R)/pR$ is the kernel of this map according to the Lemma.

Hence, we can efficiently compute $J_p(R)$.

Lemma 16.2.4 Let $J_p(R) = \mathrm{rad}(pR)$ as before and let $T_p(R) := \{ x \in L \mid x J_p(R) \subseteq J_p(R) \}$.

Then,

a) [illegible] ~~$T_p(R)$ [illegible]~~
$T_p(R)$ is an order in $\mathcal{O}_L$. [illegible]

b) $R \subseteq T_p(R) \subseteq \frac{1}{p} \cdot R$

c) $R = T_p(R)$ [illegible] if and only if R is p-maximal.

Pf [illegible] b) $R \subseteq T_p(R)$ is clear because $J_p(R)$ is an ideal of R.
If $x \in T_p(R)$, then $xp \in J_p(R) \subseteq R$ because $p \in J_p(R)$.
$\Rightarrow T_p(R) \subseteq \frac{1}{p} \cdot R$

a) Since $R \subseteq T_p(R)$ is an order, it suffices to show that $T_p(R) \subseteq \mathcal{O}_L$.
Let $x \in T_p(R)$. ~~As a $\mathbb{Z}$-module, $J_p R$~~
R is a free $\mathbb{Z}$-module of [illegible] rank n.
$\Rightarrow$ So is its ideal $J_p(R)$ (a submodule of finite index).
$\Rightarrow$ The multiplication by x map $m_x : J_p(R) \to J_p(R)$ is represented by an integral $n \times n$-matrix. [illegible]
~~That pol.~~ Its char. pol. $g(X)$ is a monic integral pol. of deg. n and $g(m_x) = 0$. $\Rightarrow g(x) = 0$.
$\Rightarrow x$ is integral $\Rightarrow x \in \mathcal{O}_L$.

c) "$\Leftarrow$" clear from def.

~~"$\Rightarrow$" [illegible] Assume $x \in \mathcal{O}_L$, $px \in R$, $x \notin R$.~~
~~Let $m \geq 1$ be sufficiently large that $J_p(R)^m \subseteq pR$.~~
~~Then, $x \cdot J_p(R)^m \subseteq R$.~~

"⇒" [redacted] Consider the [redacted] p-maximal order

$R_p = \{x \in \mathcal{O}_L \mid p^k x \in R \text{ for some } k \geq 0\}$. [redacted]

$(R \subseteq R_p \subseteq \mathcal{O}_L)$.

[redacted]

Since R is a finitely generated $\mathbb{Z}$-module, there is a number $k \geq 0$ such that $p^k \cdot R_p \subseteq R$.

Also, pick $m \geq 1$ so that $\underbrace{J_p(R)}_{rad(pR)}^m \subseteq pR$.

$\Rightarrow \quad R_p \cdot J_p(R)^{km} \subseteq R_p \cdot p^k R \subseteq R$.

Assume that R is not p-maximal.

$\Rightarrow R_p$ [redacted] $\supsetneq R$, so in part. $R_p \nsubseteq R$.

$\Rightarrow$ There is a largest integer $i \geq 0$ (with $i < km$) such that $R_p \cdot J_p(R)^i \nsubseteq R$.

$\Rightarrow R_p \cdot J_p(R)^{i+1} \subseteq R$.

Let $x \in R_p \cdot J_p(R)^i$, but $x \notin R$.

$\Rightarrow x J_p(R) \subseteq R_p J_p(R)^{i+1} \subseteq R$.

For any $y \in J_p(R)$, we have

$$(xy)^{i+m+1} = \underbrace{\underbrace{x^{i+m+1}}_{\in R_p} \cdot \underbrace{y^{i+1}}_{\in J_p(R)^{i+1}}}_{\in R} \cdot \underbrace{y^m}_{\substack{\in J_p(R)^m \\ \subseteq pR}} \in pR,$$

so $xy \in rad(pR) = J_p(R)$.

Hence, $x \in T_p(R)$. But $x \notin R$, so indeed $R \neq T_p(R)$. □

Rmk This gives a procedure for computing the ring of integers:

```
Start with $R = \mathbb{Z}[\alpha]$.
For every $p$ with $p^2 \mid \mathrm{disc}(f)$:
    Keep replacing $R$ by $T_p(R)$ until it stops changing.
Return $R$.
```

can also be computed by linear algebra mod p

Note: If $R \subsetneq T_p(R)$ $(\subseteq \frac{1}{p} \cdot R)$, then $p \mid [T_p(R) : R]$.

$\Rightarrow$ $\mathrm{disc}(R)$ always decreases by a factor of at least $\frac{1}{p^2}$.

See the references for details!

16.2. Decomposition of prime numbers

[For almost all primes, we can use the following:]

Thm 16.2.1 (≈ Lemma 12.1) Let K be a number field of degree n, $\alpha \in \mathcal{O}_K$ with minimal polynomial $f(x)$ of degree n. ~~If $\mathbb{Z}[\alpha]$ is p-maximal,~~ Assume that $\mathbb{Z}[\alpha]$ is p-maximal. Let $f(x) \equiv g_1(x)^{e_1} \cdots g_t(x)^{e_t} \bmod p$ be the factorisation of f mo (with $g_i(x) \in \mathbb{Z}[X]$ monic)

Then,

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_t^{e_t}$$

with prime ideals $\mathfrak{p}_i = (p, g_i(\alpha)) = p\mathcal{O}_K + g_i(\alpha)\mathcal{O}_K$,

$$[\mathcal{O}_K/\mathfrak{p}_i : \mathbb{Z}/p\mathbb{Z}] = \deg(g_i).$$

Rmk Any ideal $\mathfrak{a} \subset \mathcal{O}_K$ is a free $\mathbb{Z}$-module of rank n (≈ a rank n lattice). It can therefore be specified by giving n basis vectors, each of which can be written as a lin. comb. of a fixed basis $\omega_1, \dots, \omega_n$ of $\mathcal{O}_K$

⇝ We can represent an ideal ~~by~~ an integer $n \times n$-matrix M, which we can put in Hermite normal form $M^{(HNF)}$ by changing the basis of $\mathfrak{a}$.

We have $Nm(I) = |\det(M)|$.

Using HNF, we can also find a basis of the $\mathbb{Z}$-module spanned by any number of elements $\beta_1, \dots, \beta_m$ of $\mathcal{O}_K$.

This allows us to add / multiply ideals.

Fractional ideals work the same but with rational coefficients

Dividing two (fractional) ideals is also not hard ("just linear algebra"). (cf. chapters 4.6-4.8 of Cohen)

<u>Alg</u> to find the decomposition $p\mathcal{O}_K = \mathfrak{p}_1^{e_1}\cdots\mathfrak{p}_t^{e_t}$ for arbitrary p:

Compute $\mathfrak{a} := J_p(\mathcal{O}_K) = \mathrm{rad}(p\mathcal{O}_K) = \mathfrak{p}_1\cdots\mathfrak{p}_t$.

It then suffices to factor the squarefree ideal $\mathfrak{a} \mid p\mathcal{O}_K$ and determine the exponents by trial division.

To factor a squarefree ideal $\mathfrak{a} \overset{\prod_{i=1}^{t}\mathfrak{p}_i}{\overset{\shortparallel}{}} \mid p\mathcal{O}_K$, we can use Berlekamp's algorithm (problem 2 on Sheet 5):

Note that $\mathcal{O}_K/\mathfrak{a} \underset{\substack{\uparrow \\ \text{CRT}}}{\cong} \prod_{i=1}^{t} \mathcal{O}_K/\mathfrak{p}_i$,

where $\mathcal{O}_K/\mathfrak{p}_i = \mathbb{F}_{p^{f_i}}$ is a fin. ext. of $\mathbb{F}_p$.

The map $\begin{array}{ccc} \mathcal{O}_K/\mathfrak{a}^{\cong\prod\mathbb{F}_{p^{f_i}}} & \to & \mathcal{O}_K/\mathfrak{a}^{\cong\prod\mathbb{F}_{p^{f_i}}} \\ x & \mapsto & x^p \end{array}$ is $\mathbb{F}_p$-linear.

Compute $V = \{x \in \mathcal{O}_K/\mathfrak{a} \mid x^p = x\}$ using linear algebra over $\mathbb{F}_p$. We have $V \cong \prod_{i=1}^{t} \mathbb{F}_p$, so in part., $\dim_{\mathbb{F}_p}(V) = t$.

If $t > 1$:

Pick a random $x \in V$ and compute $y := u_p(x) \in V$.

$\overset{\shortparallel}{}$ ($x^{(p-1)/2} - 1$ if p is odd)

~~With prob. $\geq \frac{1}{2}$~~

The projections onto the factors $\mathbb{F}_p$ (of y) are independent and each $\underbrace{\text{projection is } 0}_{\Updownarrow}$ with prob. $\approx \frac{1}{2}$.

$$\overbrace{y \mid \mathfrak{p}_i}$$

We obtain a splitting $\mathfrak{a} = \mathfrak{a}_1\mathfrak{a}_2$ and recursively factor $\mathfrak{a}_1, \mathfrak{a}_2$.

Rmk This is not the fastest alg. to decompose p!

Rmk The factorization of $f(x)$ over $\mathbb{Q}_p$ looks like the decomposition of $p\mathcal{O}_K$ in $K = \mathbb{Q}[X]/(f)$.

16.3. Ideal class group

Def The Riemann zeta function is given by

$$\zeta(s) = \sum_{n\geq 1} n^{-s} \left(= \prod_{p} \frac{1}{1-p^{-s}} \right) \quad \text{for } s \in \mathbb{C} \text{ with } \mathrm{Re}(s) > 1.$$

Def The Dedekind zeta function of a number field K is given by

$$\zeta_K(s) = \sum_{\substack{0\neq \mathfrak{a} \subset \mathcal{O}_K \\ \text{ideal}}} \mathrm{Nm}(\mathfrak{a})^{-s} = \prod_{\substack{\mathfrak{p} \subset \mathcal{O}_K \\ \text{prime} \\ \text{ideal}}} \frac{1}{1-\mathrm{Nm}(\mathfrak{p})^{-s}}$$

for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$.

Ex $\zeta_{\mathbb{Q}} = \zeta$.

Thm 16.3.1 (Class number formula)

$$\lim_{s\to 1} (s-1)\zeta_K(s) = \frac{2^{r_1}(2\pi)^{r_2}\, R_K\, |\mathcal{C}\ell_K|}{w_K\cdot\sqrt{|D_K|}}$$

if K has r_1 real embeddings,
r_2 pairs of complex embeddings,

regulator R_K, class group $\mathcal{C}\ell_K$, roots of unity $w_K \subset \mathcal{O}_K$ (torsion subgroup),

$\uparrow$
"how far apart
the units are"

discriminant D_K.

Ex $\lim_{s\to 1} (s-1)\zeta(s) = 1$

Rmk $LHS = \lim_{s\to 1} \frac{\zeta_K(s)}{\zeta(s)}$.

Rmk R_K, $|\mathcal{C}\ell_K|$ often show up together and can be hard to separate!

~~Rmk (Brauer–Sieg~~

Thm 16.3.2 (Brauer–Siegel Thm)

For fixed $n = [K:\mathbb{Q}]$ and any $\varepsilon > 0$,

$$|D_K|^{\frac{1}{2}-\varepsilon} \ll R_K|\mathcal{C}\ell_K| \ll |D_K|^{\frac{1}{2}+\varepsilon}.$$

$$\left(\frac{\log(R_K|\mathcal{C}\ell_K|)}{\log(\sqrt{|D_K|})} \longrightarrow 1.\right)$$

We will focus on <u>imaginary quadratic number fields</u> K.

(Hafner, McCurley: A rigorous subexponential alg. for computation of class group

(For general number fields K, see Buchmann: A subexp. alg. for the determination of class groups and regulators of alg. nr. fields)

Rmk $r_1 = 0$, $r_2 = 1$, $\mathcal{O}_K^\times = \mu_K$,

$R_K = 1$,

$$\omega_K = |\mu_K| = \begin{cases} 2, & \text{otherwise} \\ 4, & K = \mathbb{Q}(i) \\ 6, & K = \mathbb{Q}(\zeta_3) \end{cases}$$

$K = \mathbb{Q}(\sqrt{D_K})$, $D_K < 0$.

~~**Def** An ideal $0 \neq \mathfrak{a} \subset \mathcal{O}_K$ is reduced if~~

~~a) it is not divisible by any integer $k \geq 2$ and~~

~~b)~~

(Nonstandard)

Def A fractional ideal $\mathfrak{a}$ is <u>reduced</u> if

$1 \in \mathfrak{a}$ but there is no $x \in \mathfrak{a}$, $x \neq 0$, with $\underset{\underset{|\text{compl. emb.}(x)|}{\|}}{\mathrm{Nm}}(x) < 1$.

Rmk $1 \in \mathfrak{a} \Leftrightarrow \underset{\underset{(1)}{\|}}{\mathcal{O}_K} \subseteq \mathfrak{a} \Leftrightarrow \mathfrak{a}^{-1} \subseteq \mathcal{O}_K$,

so the inverse of a reduced fractional ideal is an (integral) ideal of $\mathcal{O}_K$.

Lemma 16.3.3 Any ideal class contains at least 1, at most 6 reduced ideals

Pf Let $\mathfrak{b}$ be any fractional ideal

and let $y \in \mathfrak{b}$ be a nonzero element of minimal norm.

Then, $\mathfrak{a} := y^{-1} \cdot \mathfrak{b}$ is reduced. Any reduced ideal in the ideal class $[\mathfrak{b}]$ is

of this form. There are at most 6 such y. $\sqsubset$

Rmk ~~For imaginary quadratic K, we can~~

We can efficiently determine the reduced ideals using Gauß's lattice reduction to find the shortest nonzero vector.

Hence, we can efficiently determine whether two ideals lie in the same ideal class.

Thm 16.3.4 (Minkowski bound)

If $\mathfrak{a}$ is reduced, then $\mathrm{Nm}(\mathfrak{a}^{-1}) \leq O(\sqrt{|D_K|})$.

Rmk This gives rise to a slow alg. to determine Cl_K:

Find all ideals $\mathfrak{b}$ with $\mathrm{Nm}(\mathfrak{b}) \leq O(\sqrt{|D_K|})$.

For each, compute the reduced ideals in the same ideal class as $\mathfrak{b}^{-1}$ to determine which $\mathfrak{b}$ are in the same class.

Thm 16.3.5 Assume the Extended Riemann Hypothesis (ERH). Then, Cl_K is generated by the (ideal classes of) prime ideals $\mathfrak{p}$ of norm $Nm(\mathfrak{p}) \leq 6\,(\log|D_K|)^2$.

Rmk Hence, if $\mathfrak{p}_1, \dots, \mathfrak{p}_r$ are the prime ideals of norm $\leq B$ with $B \geq 6\,(\log|D_K|)^2$, then we get a surjective group hom

$$\varphi : \mathbb{Z}^r \twoheadrightarrow Cl_K$$
$$(a_1, \dots, a_r) \mapsto [\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}]$$

To determine Cl_K, we need to find its kernel:

$$Cl_K \cong \mathbb{Z}^r / \ker(\varphi)$$

I.e.: Need to find elements generating the rank n lattice $\ker(\varphi)$

Idea: Find random elements ~~of $\ker(\varphi)$~~ until they generate $\ker(\varphi)$.

How to tell when we're finished?

Let $\Lambda \subset \ker(\varphi)$, the lattice generated by the elements discovered so far.

~~We either have $\Lambda \subsetneq \ker(\varphi)$,~~

If $\Lambda \subsetneq \ker(\varphi)$, then $[\ker(\varphi) : \Lambda] \geq 2$, so

$$|\mathbb{Z}^r / \Lambda| \geq 2 \cdot |\mathbb{Z}^r / \ker(\varphi)| = 2\,|Cl_K|.$$

Hence, it suffices to know $|Cl_K|$ within a factor of 2, which (assuming ERH) can be computed using the class number formula.

How to "quickly" find random elements of $\ker(\varphi)$?

Pick a random vector $a = (a_1, \dots, a_r) \in \mathbb{Z}^r$.

~~Compute~~

~~[The ideal $\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}$ is]~~

[This only lies in $\ker(\varphi)$ with (very small) prob. $\approx \frac{1}{\# Cl_K}$!]

Compute a reduced fractional ideal $\mathfrak{a}$ in the ideal class $[\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r}] = [\mathfrak{p}_1]^{a_1} \cdots [\mathfrak{p}_r]^{a_r}$ using fast exponentiation, reducing at every step to ensure that we only need to work with ideals of norm $\leq |D_K|$ at any time.

If $\mathfrak{a}^{-1} = \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_r^{b_r}$ with integers $b_1, \dots, b_r \geq 0$,

then $[\mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_r^{a_r} \, \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_r^{b_r}] = [\mathfrak{a}\mathfrak{a}^{-1}] = [1]$,

so $\boxed{a+b} = (a_1 + b_1, \dots, a_r + b_r) \in \ker(\varphi)$.

(Note that $b_1, \dots, b_r \leq O(\log |D_K|)$ because $Nm(\mathfrak{a}^{-1}) \leq O(\sqrt{|D_K|})$.)

Otherwise try again, with a new random vector $a \in \mathbb{Z}^r$.

~~First, to obtain of linearly independent~~

~~[illegible]~~

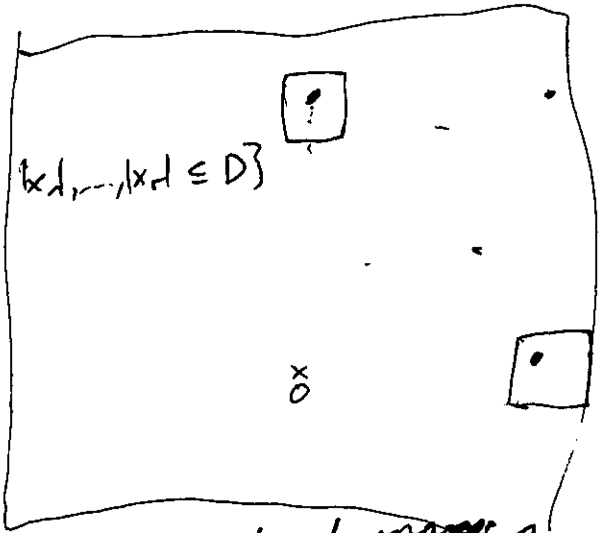
We pick the first r vectors a $\overset{\text{(uniformly)}}{\vee}$ from

$$(2r|D|,0,\dots,0)+\underbrace{B(D)}_{\{x\in\mathbb{Z}^r \mid |x_1|,\dots,|x_r|\le D\}},$$

$$(0,2r|D|,\dots,0)+B(D),$$

$$\vdots$$

$$(0,\dots,0,2r|D|)+B(D)$$



so the first r random elements of $\ker(\varphi)$ we construct ~~spa~~ a lattice $\Lambda\subset\mathbb{Z}^r$ of rank n and covolume $\le O((r|D|)^r)$.

Afterwards, we pick vectors a from $B(|D|^2)$ until $|\mathbb{Z}^r/\Lambda|$ is small enough.

Analyzing the expected running time and choosing $\underset{\uparrow}{B}$ optimally, we get:

bound on norm. a_i we consider

Thm 16.3.6 (~~[illegible]~~ McCurley) Assuming GRH, we can determine Cl_K in expected time $\tilde{O}(\exp(\underbrace{\sqrt{2}}\cdot\sqrt{\log|D_K|\log\log|D_K|}))$.

can presumably be improved...