Homework: **(due November 9 by class time)**
Page 20, the one Homework problem there
Page 22, Problem (1) and (2)

Berlekamp's Method

This algorithm determines the factorization of a polynomial $f(x)$ modulo a prime p.

$$f(x) \equiv u(x)v(x) \pmod{p}$$

Hensel Lifting

Hensel Lifting will produce, for any positive integer k, monic polynomials $u_k(x)$ and $v_k(x)$ in $\mathbb{Z}[x]$ satisfying

$$u_k(x) \equiv u(x) \pmod{p}, \quad v_k(x) \equiv v(x) \pmod{p},$$

and

$$f(x) \equiv u_k(x)v_k(x) \pmod{p^k}.$$

```
> f:=x^7+x^6+2*x^3+x^2+x+1:
  u:=x^4+x+1: v:=x^3+x^2+1:
> for k from 1 to 20 do
    w:=expand((f-u*v)/p^k):
    cf:=cfrac(v/u,`quotients`):
    m:=nops(cf):
    conv:=simplify(nthconver(cf,m-2)):
    a:=numer(conv) mod p: b:=denom(conv) mod p:
    newa:=Rem(a*w,v,x,'q') mod p:
    newb:=expand(b*w+q*u) mod p:
    u:=sort(expand(u-p^k*newb) mod p^(k+1)):
    v:=sort(expand(v-p^k*newa) mod p^(k+1)):
  od:
> expand(f-u*v) mod 2^21;
```

$$0$$

```
> expand(f-u*v) mod 2^22;
```

$$2097152\,x^6 + 2097152\,x^3 + 2097152\,x^2$$

```
> mods(u,2^21);
```

$$x^4 + 2\,x^3 + 2\,x^2 + x + 1$$

```
    m:=nops(cf):
    conv:=simplify(nthconver(cf,m-2)):
    a:=numer(conv) mod p: b:=denom(conv) mod p:
    newa:=Rem(a*w,v,x,'q') mod p:
    newb:=expand(b*w+q*u) mod p:
    u:=sort(expand(u-p^k*newb) mod p^(k+1)):
    v:=sort(expand(v-p^k*newa) mod p^(k+1)):
  od:
> expand(f-u*v) mod 2^21;
```

$$0$$

```
> expand(f-u*v) mod 2^22;
```

$$2097152\,x^6 + 2097152\,x^3 + 2097152\,x^2$$

```
> mods(u,2^21);
```

$$x^4 + 2\,x^3 + 2\,x^2 + x + 1$$

```
> mods(v,2^21);
```

$$x^3 - x^2 + 1$$

```
> factor(f);
```

$$\left(x^3 - x^2 + 1\right)\left(x^4 + 2\,x^3 + 2\,x^2 + x + 1\right)$$

```
> f:=x^12+x^6+2*x^3+x^2+x+1: p:=2:
  u:=x^2+x+1: v:=x^10+x^9+x^7+x^6+1:
> for k from 1 to 20 do
    w:=expand((f-u*v)/p^k):
    cf:=cfrac(v/u,`quotients`):
    m:=nops(cf):
    conv:=simplify(nthconver(cf,m-2)):
    a:=numer(conv) mod p: b:=denom(conv) mod p:
    newa:=Rem(a*w,v,x,'q') mod p:
    newb:=expand(b*w+q*u) mod p:
    u:=sort(expand(u-p^k*newb) mod p^(k+1)):
    v:=sort(expand(v-p^k*newa) mod p^(k+1)):
  od:
> expand(f-u*v) mod 2^21;
```

$$0$$

```
> expand(f-u*v) mod 2^22;
```

$$2097152\,x^{11}+2097152\,x^{9}+2097152\,x^{8}+2097152\,x^{2}$$

```
> mods(u,2^21);
```

$$x^{2}-281583\,x+231781$$

```
    a:=numer(conv) mod p: b:=denom(conv) mod p:
    newa:=Rem(a*w,v,x,'q') mod p:
    newb:=expand(b*w+q*u) mod p:
    u:=sort(expand(u-p^k*newb) mod p^(k+1)):
    v:=sort(expand(v-p^k*newa) mod p^(k+1)):
  od:
> expand(f-u*v) mod 2^21;
```

$$0$$

```
> expand(f-u*v) mod 2^22;
```

$$2097152\,x^{11} + 2097152\,x^{9} + 2097152\,x^{8} + 2097152\,x^{2}$$

```
> mods(u,2^21);
```

$$x^{2} - 281583\,x + 231781$$

```
> mods(v,2^21);
```

$$x^{10} + 281583\,x^{9} - 368708\,x^{8} - 419783\,x^{7} + 683787\,x^{6} - 568120\,x^{5} - 848798\,x^{4} + 379542\,x^{3} + 1032032\,x^{2} + 1021812\,x - 229267$$

```
> factor(f);
```

$$x^{12} + x^{6} + 2\,x^{3} + x^{2} + x + 1$$

```
>
```

An Inequality of Landau

Definitions and Notations. For

$$f(x) = \sum_{j=0}^{n} a_j x^j = a_n \prod_{j=1}^{n} (x - \alpha_j),$$

with $a_n \neq 0$, we set

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \quad \text{and} \quad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\},$$

the latter being the Mahler measure of the polynomial $f(x)$.

An Inequality of Landau

Definitions and Notations. For

$$f(x) = \sum_{j=0}^{n} a_j x^j = a_n \prod_{j=1}^{n} (x - \alpha_j),$$

with $a_n \neq 0$, we set

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \quad \text{and} \quad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\},$$

the latter being the Mahler measure of the polynomial $f(x)$. We also define the reciprocal of $f(x)$ as

$$\widetilde{f}(x) = x^{\deg f} f(1/x).$$

An Inequality of Landau

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \qquad\qquad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}$$

$$\widetilde{f}(x) = x^{\deg f} f(1/x)$$

An Inequality of Landau

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \qquad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}$$

$$\widetilde{f}(x) = x^{\deg f} f(1/x)$$

Useful Related Items:

- If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.
- If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \geq 1$.
- The reciprocal of f is f in reverse; $\tilde{f}(x) = \sum_{j=0}^{n} a_{n-j} x^j$.
- The coefficient of x^n in $f(x)\tilde{f}(x)$ is $\|f\|^2$.

An Inequality of Landau

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \qquad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}$$

$$\widetilde{f}(x) = x^{\deg f} f(1/x)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

Comment: So the size of the coefficients of a factor of a polynomial $f(x) \in \mathbb{Z}[x]$ cannot be too large in comparison to the degree and coefficients of $f(x)$.

An Inequality of Landau

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \qquad M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}$$

$$\widetilde{f}(x) = x^{\deg f} f(1/x)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

Comment: So the size of the coefficients of a factor of a polynomial $f(x) \in \mathbb{Z}[x]$ cannot be too large in comparison to the degree and coefficients of $f(x)$. (Note that, for every B, there is an n such that $x^n - 1$ has a factor with a coefficient larger than B.)

Proof. **We begin by proving that for $f(x) \in \mathbb{R}[x]$,**

$$(*) \qquad M(f) \leq \|f\| \leq 2^{\deg f} M(f).$$

Proof. We begin by proving that for $f(x) \in \mathbb{R}[x]$,

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f).$$

Let

$$w(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j x - 1).$$

Proof. We begin by proving that for $f(x) \in \mathbb{R}[x]$,

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f).$$

Let

$$w(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j x - 1).$$

Then

$$\widetilde{w}(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (1 - \alpha_j x) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j - x).$$

Proof. We begin by proving that for $f(x) \in \mathbb{R}[x]$,

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f).$$

Let

$$w(x) = a_n \prod_{\substack{1\le j\le n \\ |\alpha_j|>1}} (x - \alpha_j) \prod_{\substack{1\le j\le n \\ |\alpha_j|\le 1}} (\alpha_j x - 1).$$

Then

$$\widetilde{w}(x) = a_n \prod_{\substack{1\le j\le n \\ |\alpha_j|>1}} (1 - \alpha_j x) \prod_{\substack{1\le j\le n \\ |\alpha_j|\le 1}} (\alpha_j - x).$$

Therefore,

$$w(x)\tilde{w}(x) = a_n^2 \prod_{j=1}^{n} (x - \alpha_j) \prod_{j=1}^{n} (1 - \alpha_j x) = f(x)\tilde{f}(x).$$

Proof. We begin by proving that for $f(x) \in \mathbb{R}[x]$,

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f).$$

Let

$$w(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j x - 1).$$

Then

$$\widetilde{w}(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (1 - \alpha_j x) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j - x).$$

Therefore,

$$w(x)\tilde{w}(x) = a_n^2 \prod_{j=1}^{n} (x - \alpha_j) \prod_{j=1}^{n} (1 - \alpha_j x) = f(x)\tilde{f}(x),$$

so that $\|w\| = \|f\|$.

- The coefficient of x^n in $f(x)\tilde{f}(x)$ is $\|f\|^2$.

$$(*) \qquad M(f) \leq \|f\| \leq 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| \leq 1}} (\alpha_j x - 1) \qquad \|w\| = \|f\|$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j x - 1) \qquad \|w\| = \|f\|$$

The definition of $w(x)$ implies $|w(0)| = M(f)$.

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1\le j\le n \\ |\alpha_j|>1}} (x-\alpha_j) \prod_{\substack{1\le j\le n \\ |\alpha_j|\le 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

The definition of $w(x)$ implies $|w(0)| = M(f)$. Writing $w(x) = \sum_{j=0}^{n} c_j x^j$, we obtain

$$M(f) = |c_0| \le (c_0^2 + c_1^2 + \cdots + c_n^2)^{1/2} = \|w\| = \|f\|,$$

establishing the first inequality in $(*)$.

$$(*) \qquad M(f) \leq \|f\| \leq 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| \leq 1}} (\alpha_j x - 1) \qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\leq M(f)/|a_n|$.

$$\boxed{M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |\alpha_j|\}}$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1\le j\le n\\ |\alpha_j|>1}} (x-\alpha_j) \prod_{\substack{1\le j\le n\\ |\alpha_j|\le 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\le M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\le \binom{n}{k} \times M(f)/|a_n|$.

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1\le j\le n \\ |\alpha_j|>1}} (x-\alpha_j) \prod_{\substack{1\le j\le n \\ |\alpha_j|\le 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\le M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\le \binom{n}{k} \times M(f)/|a_n|$. Hence,

$$|a_{n-k}| \le \binom{n}{k} M(f) = \binom{n}{n-k} M(f).$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1\le j\le n \\ |\alpha_j|>1}} (x-\alpha_j) \prod_{\substack{1\le j\le n \\ |\alpha_j|\le 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\le M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\le \binom{n}{k} \times M(f)/|a_n|$. Hence,

$$|a_{n-k}| \le \binom{n}{k} M(f) = \binom{n}{n-k} M(f).$$

The second inequality in $(*)$ now follows from

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2}$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1 \le j \le n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \le j \le n \\ |\alpha_j| \le 1}} (\alpha_j x - 1) \qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\le M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\le \binom{n}{k} \times M(f)/|a_n|$. Hence,

$$|a_{n-k}| \le \binom{n}{k} M(f) = \binom{n}{n-k} M(f).$$

The second inequality in $(*)$ now follows from

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \le \sum_{j=0}^{n} |a_j|$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1\le j\le n\\ |\alpha_j|>1}} (x-\alpha_j) \prod_{\substack{1\le j\le n\\ |\alpha_j|\le 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\le M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\le \binom{n}{k} \times M(f)/|a_n|$. Hence,

$$|a_{n-k}| \le \binom{n}{k} M(f) = \binom{n}{n-k} M(f).$$

The second inequality in $(*)$ now follows from

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2\right)^{1/2} \le \sum_{j=0}^{n} |a_j| \le \sum_{j=0}^{n} \binom{n}{j} M(f)$$

$$(*) \qquad M(f) \leq \|f\| \leq 2^{\deg f} M(f)$$

$$w(x) = a_n \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| > 1}} (x - \alpha_j) \prod_{\substack{1 \leq j \leq n \\ |\alpha_j| \leq 1}} (\alpha_j x - 1) \qquad\qquad \|w\| = \|f\|$$

For each $k \in \{1, 2, \ldots, n\}$, the product of any k of the α_j has absolute value $\leq M(f)/|a_n|$. It follows that $|a_{n-k}|/|a_n|$, which is the sum of the products of the roots taken k at a time, is $\leq \binom{n}{k} \times M(f)/|a_n|$. Hence,

$$|a_{n-k}| \leq \binom{n}{k} M(f) = \binom{n}{n-k} M(f).$$

The second inequality in $(*)$ now follows from

$$\|f\| = \left(\sum_{j=0}^{n} a_j^2 \right)^{1/2} \leq \sum_{j=0}^{n} |a_j| \leq \sum_{j=0}^{n} \binom{n}{j} M(f) = 2^n M(f).$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

- If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.
- If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \ge 1$.

Landau's inequality follows from

$$\|g\| \le 2^{\deg g} M(g)$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

- If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.
- If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \ge 1$.

Landau's inequality follows from

$$\|g\| \le 2^{\deg g} M(g) \le 2^{\deg g} M(g)M(h)$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

- If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.
- If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \ge 1$.

Landau's inequality follows from

$$\begin{aligned} \|g\| &\le 2^{\deg g} M(g) \le 2^{\deg g} M(g) M(h) \\ &= 2^{\deg g} M(gh) \end{aligned}$$

$$(*) \qquad M(f) \le \|f\| \le 2^{\deg f} M(f)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \le 2^{\deg g} \|f\|.$$

- If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.
- If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \ge 1$.

Landau's inequality follows from

$$\|g\| \le 2^{\deg g} M(g) \le 2^{\deg g} M(g)M(h)$$

$$= 2^{\deg g} M(gh) = 2^{\deg g} M(f)$$

$$(*) \qquad M(f) \leq \|f\| \leq 2^{\deg f} M(f)$$

Theorem. *If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then*

$$\|g\| \leq 2^{\deg g} \|f\|.$$

- **If $g(x)$ and $h(x)$ are in $\mathbb{C}[x]$, then $M(gh) = M(g)M(h)$.**
- **If $g(x)$ is in $\mathbb{Z}[x]$, then $M(g) \geq 1$.**

Landau's inequality follows from

$$\begin{aligned}\|g\| &\leq 2^{\deg g} M(g) \leq 2^{\deg g} M(g) M(h) \\ &= 2^{\deg g} M(gh) = 2^{\deg g} M(f) \leq 2^{\deg g} \|f\|. \quad \blacksquare\end{aligned}$$

An Approach of Zassenhaus

We explain a method for factoring a given $f(x) \in \mathbb{Z}[x]$ with the added assumptions that $f(x)$ is monic and squarefree.

- **Set**
$$B = 2^{\lfloor (\deg f)/2 \rfloor} \|f\|.$$
(If $f(x)$ has a nontrivial factor $g(x)$ in $\mathbb{Z}[x]$, it has such a factor of degree $\le \lfloor (\deg f)/2 \rfloor$ so that by Landau's inequality, we can use B as a bound on $\|g\|$.)

Theorem. ***If $f(x)$, $g(x)$, and $h(x)$ in $\mathbb{Z}[x]$ are such that $f(x) = g(x)h(x)$, then***
$$\|g\| \le 2^{\deg g} \|f\|.$$

An Approach of Zassenhaus

We explain a method for factoring a given $f(x) \in \mathbb{Z}[x]$ with the added assumptions that $f(x)$ is monic and squarefree.

- **Set**
$$B = 2^{\lfloor (\deg f)/2 \rfloor} \|f\|.$$
(If $f(x)$ has a nontrivial factor $g(x)$ in $\mathbb{Z}[x]$, it has such a factor of degree $\le \lfloor (\deg f)/2 \rfloor$ so that by Landau's inequality, we can use B as a bound on $\|g\|$.)
- **Find a prime p for which $f(x)$ is squarefree modulo p.**
- **Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)**
- **Factor $f(x)$ modulo p by Berlekamp's algorithm and use Hensel lifting to obtain the complete factorization of $f(x)$ modulo p^r. Given our conditions on $f(x)$, we can take all irreducible factors to be monic and do so.**

- Set $B = 2^{\lfloor (\deg f)/2 \rfloor} \|f\|$.
- Find a prime p for which $f(x)$ is squarefree modulo p.
- Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)
- Factor $f(x)$ modulo p by Berlekamp's algorithm and use Hensel lifting to obtain the complete factorization of $f(x)$ modulo p^r. Given our conditions on $f(x)$, we can take all irreducible factors to be monic and do so.
- Loop through all possible products of irreducible factors of $f(x)$ modulo p^r to consider all possible factors $u(x)$ of $f(x)$ modulo p^r.
- For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.

- Set $B = 2^{\lfloor(\deg f)/2\rfloor}\|f\|$.
- Find a prime p for which $f(x)$ is squarefree modulo p.
- Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)
- Factor $f(x)$ modulo p by Berlekamp's algorithm and use Hensel lifting to obtain the complete factorization of $f(x)$ modulo p^r. Given our conditions on $f(x)$, we can take all irreducible factors to be monic and do so.
- Loop through all possible products of irreducible factors of $f(x)$ modulo p^r to consider all possible factors $u(x)$ of $f(x)$ modulo p^r.
- For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.
- If one of these $u_0(x)$ with degree in $[1, (\deg f)/2]$ divides $f(x)$, we have found a non-trivial factorization of $f(x)$. If no such $u_0(x)$ divides $f(x)$, then $f(x)$ is irreducible.

- **Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)**

- **Loop through all possible products of irreducible factors of $f(x)$ modulo p^r to consider all possible factors $u(x)$ of $f(x)$ modulo p^r.**

- **For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.**

- **If one of these $u_0(x)$ with degree in $[1, (\deg f)/2]$ divides $f(x)$, we have found a non-trivial factorization of $f(x)$. If no such $u_0(x)$ divides $f(x)$, then $f(x)$ is irreducible.**

- Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)
- For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.
- If one of these $u_0(x)$ with degree in $[1, (\deg f)/2]$ divides $f(x)$, we have found a non-trivial factorization of $f(x)$. If no such $u_0(x)$ divides $f(x)$, then $f(x)$ is irreducible.

Explanation. If $f(x)$ has a monic factor $g(x)$ of degree $\leq (\deg f)/2$, then necessarily

$$g(x) \equiv u_0(x) \pmod{p^r},$$

for some $u_0(x)$ in the algorithm.

- Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)

- For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.

- If one of these $u_0(x)$ with degree in $[1, (\deg f)/2]$ divides $f(x)$, we have found a non-trivial factorization of $f(x)$. If no such $u_0(x)$ divides $f(x)$, then $f(x)$ is irreducible.

Explanation. If $f(x)$ has a monic factor $g(x)$ of degree $\leq (\deg f)/2$, then necessarily

$$g(x) \equiv u_0(x) \pmod{p^r},$$

for some $u_0(x)$ in the algorithm. The coefficients of $g(x)$ and $u_0(x)$ are all in $(-p^r/2, p^r/2]$ so that all coefficients of $g(x) - u_0(x)$ are divisible by p^r and have absolute value $< p^r$.

- Set $r \in \mathbb{Z}^+$ minimal such that $p^r > 2B$. (Thus, each coefficient of $g(x)$ as above is in $(-p^r/2, p^r/2]$.)
- For each such $u(x)$, consider the polynomial $u_0(x) \in \mathbb{Z}[x]$ with $u_0(x) \equiv u(x) \pmod{p^r}$ and each coefficient of $u_0(x)$ in the interval $(-p^r/2, p^r/2]$. Check if $u_0(x)$ divides $f(x)$ in $\mathbb{Z}[x]$.
- If one of these $u_0(x)$ with degree in $[1, (\deg f)/2]$ divides $f(x)$, we have found a non-trivial factorization of $f(x)$. If no such $u_0(x)$ divides $f(x)$, then $f(x)$ is irreducible.

Explanation. If $f(x)$ has a monic factor $g(x)$ of degree $\leq (\deg f)/2$, then necessarily

$$g(x) \equiv u_0(x) \pmod{p^r},$$

for some $u_0(x)$ in the algorithm. The coefficients of $g(x)$ and $u_0(x)$ are all in $(-p^r/2, p^r/2]$ so that all coefficients of $g(x) - u_0(x)$ are divisible by p^r and have absolute value $< p^r$. Therefore, $g(x) = u_0(x)$.

Swinnerton-Dyer's Example

Why is the method of Zassenhaus bad?

Why is the method of Zassenhaus good?

Swinnerton-Dyer's Example

Why is the method of Zassenhaus bad?

Let $a_1, a_2, \ldots, a_m$ be arbitrary squarefree pairwise relatively prime integers > 1. Let S_m be the set of 2^m different m-tuples $(\varepsilon_1, \ldots, \varepsilon_m)$ where each $\varepsilon_j \in \{1, -1\}$. Then the polynomial

$$f(x) = \prod_{(\varepsilon_1,\ldots,\varepsilon_m)\in S_m} \Big(x - (\varepsilon_1\sqrt{a_1} + \cdots + \varepsilon_m\sqrt{a_m})\Big)$$

has the properties:

(i) The polynomial $f(x)$ is in $\mathbb{Z}[x]$.

(ii) It is irreducible over the rationals.

(iii) It factors as a product of linear and quadratic polynomials modulo every prime p.

- **Loop through all possible products of irreducible factors of $f(x)$ modulo p^r to consider all possible factors $u(x)$ of $f(x)$ modulo p^r.**

Let $a_1, a_2, \ldots, a_m$ be arbitrary squarefree pairwise relatively prime integers > 1. Let S_m be the set of 2^m different m-tuples $(\varepsilon_1, \ldots, \varepsilon_m)$ where each $\varepsilon_j \in \{1, -1\}$. Then the polynomial

$$f(x) = \prod_{(\varepsilon_1,\ldots,\varepsilon_m)\in S_m} \big(x - (\varepsilon_1\sqrt{a_1} + \cdots + \varepsilon_m\sqrt{a_m})\big)$$

has the properties:

(i) The polynomial $f(x)$ is in $\mathbb{Z}[x]$.

(ii) It is irreducible over the rationals.

(iii) It factors as a product of linear and quadratic polynomials modulo every prime p.

Swinnerton-Dyer's Example

Why is the method of Zassenhaus bad?

Why is the method of Zassenhaus good?

Swinnerton-Dyer's Example

Why is the method of Zassenhaus good?

As we will see, there exists a polynomial time algorithm for factoring polynomials in $\mathbb{Z}[x]$.

Swinnerton-Dyer's Example

Why is the method of Zassenhaus good?

As we will see, there exists a polynomial time algorithm for factoring polynomials in $\mathbb{Z}[x]$. The method of Zassenhaus is not it.

Swinnerton-Dyer's Example

Why is the method of Zassenhaus good?

As we will see, there exists a polynomial time algorithm for factoring polynomials in $\mathbb{Z}[x]$. The method of Zassenhaus is not it. However, his method typically runs faster.

The Lattice Base Reduction Algorithm

This is a method which was developed in 1982 by Arjen Lenstra, Hendrik Lenstra and László Lovász to prove that factoring polynomials in $\mathbb{Z}[x]$ can be done in polynomial time.

The Lattice Base Reduction Algorithm

This is a method which was developed in 1982 by Arjen Lenstra, Hendrik Lenstra and László Lovász to prove that factoring polynomials in $\mathbb{Z}[x]$ can be done in polynomial time. It is sometimes called the LLL-algorithm or the L^3-algorithm.

Definitions and Notations.

The Lattice Base Reduction Algorithm

This is a method which was developed in 1982 by Arjen Lenstra, Hendrik Lenstra and László Lovász to prove that factoring polynomials in $\mathbb{Z}[x]$ can be done in polynomial time. It is sometimes called the LLL-algorithm or the L^3-algorithm.

Definitions and Notations. **Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$.**

The Lattice Base Reduction Algorithm

This is a method which was developed in 1982 by Arjen Lenstra, Hendrik Lenstra and László Lovász to prove that factoring polynomials in $\mathbb{Z}[x]$ can be done in polynomial time. It is sometimes called the LLL-algorithm or the L^3-algorithm.

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a_1', a_2', \ldots, a_n' \rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a_1' + a_2 a_2' + \cdots + a_n a_n',$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a_1', a_2', \ldots, a_n' \rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a_1' + a_2 a_2' + \cdots + a_n a_n',$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a'_1, a'_2, \ldots, a'_n \rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a'_1 + a_2 a'_2 + \cdots + a_n a'_n,$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Further, we use A^T to denote the transpose of a matrix A, so the rows and columns of A are the same as the columns and rows of A^T, respectively.

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a_1', a_2', \ldots, a_n' \rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a_1' + a_2 a_2' + \cdots + a_n a_n',$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Further, we use A^T to denote the transpose of a matrix A, so the rows and columns of A are the same as the columns and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = \big(\vec{b}_1, \ldots, \vec{b}_n\big)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$.

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n\rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n\rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a_1', a_2', \ldots, a_n'\rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a_1' + a_2 a_2' + \cdots + a_n a_n',$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Further, we use A^T to denote the transpose of a matrix A, so the rows and columns of A are the same as the columns and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = \left(\vec{b}_1, \ldots, \vec{b}_n\right)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1\mathbb{Z} + \cdots + \vec{b}_n\mathbb{Z}.$$

Definitions and Notations. Let $\mathbb{Q}^n$ denote the set of vectors $\langle a_1, a_2, \ldots, a_n \rangle$ with $a_j \in \mathbb{Q}$. For

$$\vec{b} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{Q}^n \quad \text{and} \quad \vec{b}' = \langle a'_1, a'_2, \ldots, a'_n \rangle \in \mathbb{Q}^n,$$

define the usual dot product $\vec{b} \cdot \vec{b}'$ by

$$\vec{b} \cdot \vec{b}' = a_1 a'_1 + a_2 a'_2 + \cdots + a_n a'_n,$$

and set

$$\|\vec{b}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Further, we use A^T to denote the transpose of a matrix A, so the rows and columns of A are the same as the columns and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = (\vec{b}_1, \ldots, \vec{b}_n)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1 \mathbb{Z} + \cdots + \vec{b}_n \mathbb{Z}.$$

We typically want $\vec{b}_1, \ldots, \vec{b}_n$ to be linearly independent; in this case, $\vec{b}_1, \ldots, \vec{b}_n$ is called a basis for $\mathcal{L}$.

and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = (\vec{b}_1, \ldots, \vec{b}_n)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1\mathbb{Z} + \cdots + \vec{b}_n\mathbb{Z}.$$

We typically want $\vec{b}_1, \ldots, \vec{b}_n$ to be linearly independent; in this case, $\vec{b}_1, \ldots, \vec{b}_n$ is called a basis for $\mathcal{L}$.

and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = (\vec{b}_1, \ldots, \vec{b}_n)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1\mathbb{Z} + \cdots + \vec{b}_n\mathbb{Z}.$$

We typically want $\vec{b}_1, \ldots, \vec{b}_n$ to be linearly independent; in this case, $\vec{b}_1, \ldots, \vec{b}_n$ is called a basis for $\mathcal{L}$.

Comment:

and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = (\vec{b}_1, \ldots, \vec{b}_n)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1\mathbb{Z} + \cdots + \vec{b}_n\mathbb{Z}.$$

We typically want $\vec{b}_1, \ldots, \vec{b}_n$ to be linearly independent; in this case, $\vec{b}_1, \ldots, \vec{b}_n$ is called a basis for $\mathcal{L}$.

Comment: Different A can determine the same $\mathcal{L}$.

and rows of A^T, respectively. Let $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{Q}^n$, and let $A = (\vec{b}_1, \ldots, \vec{b}_n)$ be the $n \times n$ matrix with column vectors $\vec{b}_1, \ldots, \vec{b}_n$. The lattice $\mathcal{L}$ generated by $\vec{b}_1, \ldots, \vec{b}_n$ is

$$\mathcal{L} = \mathcal{L}(A) = \vec{b}_1\mathbb{Z} + \cdots + \vec{b}_n\mathbb{Z}.$$

We typically want $\vec{b}_1, \ldots, \vec{b}_n$ to be linearly independent; in this case, $\vec{b}_1, \ldots, \vec{b}_n$ is called a basis for $\mathcal{L}$.

Comment: Different A can determine the same $\mathcal{L}$. But given $\mathcal{L}$, the value of $|\det A|$ is the same for all such A.