6.966 Algebra and Computation	October 13, 1998
Lecture 10	
Lecturer: Madhu Sudan	*Scribe: Prahladh Harsha*

1 Factoring Univariate Polynomials over Integers

In this lecture, we shall discuss factorisation of univariate polynomials over integers. Factoring polynomials over the rationals can be reduced to this case by a clearing of denominators. This factorisation can be further extended to factoring multivariate polynomials as discussed in the previous 3 lectures.

Given a polynomial $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{Z}[x], (a_n \neq 0)$, factoring f involves finding irreducible polynomials $f_1, f_2, \ldots f_k \in \mathbb{Z}[x]$ and $c \in \mathbb{Z}$ such that $f(x) = cf_1(x)f_2(x)\ldots f_k(x)$. As integer factorisation is hard, we will relax the uniqueness restrictions to just the following: each of the f_i's is an irreducible polynomial over $\mathbb{Z}$ of degree greater than 1 and $c \in \mathbb{Z}$. In this lecture, we shall present a deterministic algorithm which on input a polynomial $f \in \mathbb{Z}[x]$, determines whether f is irreducible or not and if not outputs a non-trivial factorisation of f and which runs in polynomial time in terms of the length of the coefficients of f.

One possible direction we could proceed along to factorise polynomials would be to consider the factorisation of $f \pmod p$ for several choices of primes p and find out whether these factorisations reveal anything about the factorisation of f over $\mathbb{Z}$. However, not much information is obtained in this line as there do exist ireducible polynomials $f \in \mathbb{Z}[x]$ such that $f \pmod p$ factors for every prime p.

Another line of approach could be to consider the factorisation of f modulo a suitable prime p and then employ Hensel's Lemma to lift this factorisation to a factorisation modulo p^k. We shall adopt this method which runs parallel to the bivariate polynomial factorisation discussed in Lecture 7.

2 Factoring Algorithm

The factoring algorithm is outlined below. We shall expand on the individual steps later in this lecture.
Input: A polynomial $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{Z}[x]$ with $|a_i| \approx l$-bits for each i.
Output: A non-trivial factorisation of f if f is reducible over $\mathbb{Z}[x]$.

1. Preprocessing of f: Firstly perform Square Free Factorisation and reduce the problem to the case of factoring a square free polynomial f. Then obtain a suitable "small" prime p such that $a_n \neq 0 \pmod p$. (Such a prime p polynomial in terms of l and n exists and can be found out easily.)
2. (a) Factor $f(x) \pmod p$ as $f(x) = g_0(x)h_0(x) \pmod p$ where g_0 is an irreducible monic polynomial relatively prime to h_0. (Note such a g_0 exists as f is square free)

   (b) Iteratively lift this factorisation using Hensel's Lemma to obtain
   $$f(x) = g_k(x)h_k(x) \pmod{p^k}$$
3. Find $\tilde{g}(x), l_k(x)$ such that
   $$\begin{aligned} \tilde{g}(x) &= g_k(x)l_k(x) \pmod{p^k} \\ \deg(\tilde{g}), \deg(l_k) &\leq \deg(f) \end{aligned} \tag{1}$$
   and $\tilde{g}$ has "small" coefficients.
4. Compute $h = \gcd(\tilde{g}, f)$. If h is non-trivial output $h, \frac{f}{h}$ as a non-trivial factorisation of f, else output that f is irreducible.

Before dwelling into the details of this algorithm, let us first make an observation regarding the size of the coefficients of any factor of f.

Claim 1 *For any factor g of f in $\mathbb{Z}[x]$ where $f = \sum_{i=0}^{n} a_i x^i$ and $|a_i| \approx l-bits$, there is a polynomial p in l and n such that the coefficients of g are not more than $p(l,n)$ bits long.*

Proof We shall first show that the roots (complex ones including) of f are not very large compared to the size of the coefficients of f. In fact, for all α such that $f(\alpha) = 0$, we have $|\alpha| \leq n \max_i |a_i|$.

Suppose the contrary,(ie., $|\alpha| > n \max_i |a_i|$), then

$$\begin{aligned}
|a_n \alpha^n| &\geq |\alpha|^n \\
|\sum_{i=0}^{n-1} a_i \alpha^i| &\leq n \max_i |a_i||\alpha|^{n-1} \\
\text{Hence } |f(\alpha)| &= |a_n \alpha^n + \sum_{i=0}^{n-1} a_i \alpha^i| \\
&\geq |\alpha|^n - n \max_i |a_i||\alpha|^{n-1} \\
&= |\alpha|^{n-1}(|\alpha| - n \max_i |a_i|) \\
&> 0
\end{aligned}$$

This contradicts the fact that α is a root of f. If $S = \{\alpha_1, \ldots, \alpha_n\}$ is the set of roots of f then any factor g of f is of the form $g(x) = \Pi_{\alpha_i \in S'}(x - \alpha_i)$ for some $S' \subseteq S$. Hence the coefficients of g are polynomials in $\alpha_1, \ldots, \alpha_n$ of degree atmost n. This implies that there exists a polynomial p in l and n such that the coefficients of g are smaller than $2^{p(l,n)}$. ■

The "lifting" of the factorisation performed in step 2(b) is exactly similar to that performed in the bivariate polynomial factorisation (ref: Lecture 7). We recall Hensel's Lemma

Lemma 2 (Hensel's lifting) *Let R be a UFD and $I \subset R$ be an ideal. (Here $I = (p^k)$) For any $f \in R$ and for any factorisation$f = gh(\bmod I)$ of f and there exist a, b such that $ag + bh = 1(\bmod I)$, then there exist g^*, h^*, a^*, b^* such that*

$$\begin{gathered}
f = g^* h^*(\bmod I^2) \\
g^* = g(\bmod I) \\
h^* = h(\bmod I) \\
a^* g^* + b^* h^* = 1(\bmod I^2)
\end{gathered}$$

and the solution g^*, h^* is unique in the sense: If any other g', h' also satisy these properties, then

$$\begin{gathered}
g' = g^*(1 + u)(\bmod I^2) \\
h' = h^*(1 - u)(\bmod I^2)
\end{gathered}$$

for some $u \in I$.

Coming to steps 3 and 4 of the algorithm, as before we have the following questions:

- How do we know that $\tilde{g}$ and l_k satisfying (1) exist?
- How do we find such a pair $\tilde{g}$ and l_k?
- Why does the existence of $\tilde{g}$ and l_k yield anything in step 4?

The answers to the first two questions mentioned above are precisely the same as those discusse in the bivariate polynomial factorisation. The set of $(\tilde{g}, l_k)$ is a linear space and thus obtaining $\tilde{g}$ and l_k is essentially solving a linear system (with unknowns being the coeffcients of $\tilde{g}$ and l_k)

Now, we need to see why the gcd of f and $\tilde{g}$ gives a non-trivial factor of f, if f were reducible.

Claim 3 *Let $\tilde{g}, l_k$ be any pair of polynomials satisfying (1). If the coefficients of both f and $\tilde{g}$ are not more than l bits long, and $p^k > (2n)!2^{2nl}$ then* $\gcd(f, \tilde{g})$ *is non-trivial.*

Proof We clearly have $\gcd(f, \tilde{g}) \neq f$ as $\deg(\tilde{g}) < \deg(f)$. Suppose for contradiction that $\gcd(f, \tilde{g}) = 1$ over $\mathbb{Q}[x]$. Then there exists polynomials $u, v \in \mathbb{Z}[x]$ and $N \in \mathbb{Z} - \{0\}$ such that $N < Res(f, \tilde{g})$ and

$$u(x)f(x) + v(x)\tilde{g}(x) = N$$

Recall that the resultant of 2 polynomials of degree atmost n and coefficients not larger than l bits is atmost $(2n)!2^{2nl}$. (see Lecture 4). Now, using the fact that $p^k > (2n)!2^{2nl}$, we have $N \neq 0(\mod p)$. Thus reducing the above equation modulo p^k, we have

$$\begin{aligned} N &= u(x)f(x) + v(x)\tilde{g}(x)(\mod p^k) \\ &= u(x)g_k(x)h_k(x) + v(x)g_k(x)l_k(x)(\mod p^k) \\ &= g_k(x)(u(x)h_k(x) + v(x)l_k(x))(\mod p^k) \end{aligned}$$

But this is a contradiction. As $g_k(x)$ is a polynomial of degree greater than 1. On the RHS, we have either the constant 0 or polynomial in x depending on whether $u(x)h_k(x) + v(x)l_k(x)$ is identically equal to 0 or not. However on the LHS we have a non-zero constant N. Thus, we should have $\gcd(f, \tilde{g}) \neq 1$.■

Thus step 4 computes a non-trivial factor of f if the coefficients of $\tilde{g}$ are not very large. We know from Claim 1 that the coefficients of any factor g of f are not very large. Thus while solving the linear system in step 3(b) of the algorithm, we should obtain a $\tilde{g}$ with small sized coefficients. It is not necessary that we obtain the $\tilde{g}$ with the smallest coefficients, but one close enough would do. Let us look at the linear system once again.

Suppose g_k has degree $d_0 < n(= \deg(f))$. Let

$$\tilde{g}(x) = \sum_{i=0}^{d} c_i x^i \text{ where } d = n - 1, c_i \in \mathbb{Z}$$

$$l_k(x) = \sum_{i=0}^{d-d_0} \alpha_i x^i, \alpha_i \in \mathbb{Z}$$

The linear system is then given by the equation:

$$\sum_{i=0}^{d} c_i x^i = \sum_{i=0}^{d-d_0} \alpha_i g_k(x) x^i + \sum_{i=0}^{d} \beta_i p^k x^i, \beta_i \in \mathbb{Z} \text{ where } c_i, \alpha_i, \beta_i \text{ are the unknowns.}$$

Obtaining a $\tilde{g}$ with small coefficients thus reduces to the following problem

Given $b_1 \ldots, b_m \in \mathbb{Z}^n$, find $\gamma_1, \ldots, \gamma_m \in \mathbb{Z}$, such that the vector $\sum \gamma_i b_i$ has small norm

3 Small Vectors in a Lattice

For any vectors $b_1, \ldots b_n \in \mathbb{Z}^n$, let lattice $L(b_1, \ldots b_n) = \{\sum \alpha_i b_i | \alpha_i \in \mathbb{Z}, 1 \leq i \leq n\}$. By small we mean small with respect to any suitable L_p norm[1]. Given a set of vectors $b_1, \ldots, b_m \in \mathbb{Z}^n$, it is sufficient if we look at a linearly independent subset of the vectors which also spans the lattice. We can obtain such a subset by the following reductions.

[1] L_p norm for a vector $\overline{x} = (x_1, \ldots, x_n)$ is given by $\|\overline{x}\|_p = (\sum x_i^p)^{\frac{1}{p}}$

Let each vector b_i be given by the column vector $(b_{i1}, \ldots, b_{in})^T$. Consider the matrix given by $B = (b_1, \ldots, b_m)$ (ie.,

$$B = \begin{pmatrix} b_{11} & b_{21} & \ldots & b_{m1} \\ b_{12} & b_{22} & \ldots & b_{m2} \\ \vdots & \vdots & & \vdots \\ b_{1n} & b_{2n} & \ldots & b_{mn} \end{pmatrix}$$

1. (a) Compute $g = \gcd(b_{11}, b_{21}, \ldots, b_{m1})$ and integers $a_{11}, a_{21}, \ldots, a_{m1}$ such that $\sum_{i=1}^{m} a_{i1} b_{i1} = g$
   (b) Construct a new basis $B' = (b'_1, \ldots, b'_m)$ as follows

$$b'_1 = \sum_{i=1}^{m} a_{i1} b_i$$

$$b'_k = b_k - \frac{b_{k1}}{g} b'_1, k \neq 1$$

   [Note that B' also spans $L(b_1, \ldots, b_m)$ and that $b'_{11} = g$ and $b'_{k1} = 0, k \neq 1$]

2. Repeat step 1 for $\tilde{B}$ given by

$$\tilde{B} = \begin{pmatrix} b'_{22} & \ldots & b'_{m2} \\ \vdots & & \vdots \\ b'_{2n} & \ldots & b'_{mn} \end{pmatrix}$$

   Repeat this process till B is reduced to a matrix of the form $\begin{pmatrix} \frac{A}{C} & O \end{pmatrix}$ where A is a $m' \times m'$ lower triangular matrix ($\in \mathbb{Z}_{m' \times m'}$), C is a $(n-m') \times m'$ matrix ($\in \mathbb{Z}_{(n-m') \times m'}$) and O is the $n \times (m - m')$ zero matrix.

Clearly the basis $\overline{B} = \begin{pmatrix} \frac{A}{C} \end{pmatrix}$ also spans $L(b_1, \ldots, b_m)$ and it is also a linearly independent set. With this the problem of finding a small vector reduces to:

Given a linearly independent set of vectors $\{b_1 \ldots, b_m\} \in \mathbb{Z}^n$, find a small vector in $L(b_1, \ldots b_n)$

For any lattice $L = L(b_1, \ldots, b_n)$ in $\mathbb{Z}^n$, the determinant of the lattice L (denoted as $\det(L)$) is defined as:

$$\det(L) = \det \begin{pmatrix} | & | & & | \\ b_1 & b_2 & \ldots & b_n \\ | & | & & | \end{pmatrix}$$

Though this definition seems to be dependent of the choice of the basis, we shall show that $\det(L)$ (modulo $\{\pm 1\}$) is independent of the basis.

Claim 4 *For any two bases $B = (b_1, \ldots, b_n), B' = (b'_1, \ldots, b'_n)$ of a lattice L, $\det(B) = \pm \det(B')$.*

Proof As B is a basis for the lattice L and $b'_i \in L$, there exists a $n \times n$ matrix $A \in \mathbb{Z}_{n \times n}$ such that $B' = BA$. Similarly there exists a $A' \in \mathbb{Z}_{n \times n}$ such that $B = B'A'$. hence we have

$$\begin{aligned} & AA' = I \\ \Rightarrow \quad & \det(AA') = 1 \\ \Rightarrow \quad & \det(A)\det(A') = 1 \\ \Rightarrow \quad & \det(A) = \det(A') = \pm 1 \text{ since } A, A' \in \mathbb{Z}_{n \times n} \\ \Rightarrow \quad & \det(B) = \det(B'A') = \det(B')\det(A') = \pm \det(B') \blacksquare \end{aligned}$$

To obtain a small vector, we shall make a series of moves from one basis to another.(In fact the moves we adopt to do so leave the sign of the determinant unaltered) We move from one basis to another by a sequence of (i, j, q) moves. An (i, j, q) move from a basis $(b_1, \ldots, b_n)$ to $(b'_1, \ldots, b'_n)$ is defined as follows:

$$b'_i = b_i - qb_j$$
$$b'_k = b_k, k \neq i$$

We choose q such that the b'_i is the smallest such vector. For the case $n = 2$, a sequence of such moves leads us to a basis (a, b) where $\|a\| \leq \|b\| \leq \|b + qa\|$ for all $q \in \mathbb{Z}$. In fact for this case with L_2 norm, a is indeed the smallest vector in the lattice.

Claim 5 *If $a, b \in \mathbb{Z}^2$ and $\|a\|_2 \leq \|b\|_2 \leq \|b + qa\|_2$ for all $q \in \mathbb{Z}$, then for all $z \in L(a, b) - \{(0, 0)\}$, $\|a\|_2 \leq \|z\|_2$.*

Proof Let $z = ua + vb$ for some $u, v \in \mathbb{Z}$ (u, v not both 0). If either u or v is 0, then by hypothesis $\|z\|_2 \geq \|a\|_2$. If both $u \neq 0$ and $v \neq 0$

Case (i) : $u > v$

We first observe that

$$\begin{aligned}
(a + b) \cdot (a + b) &\geq b \cdot b \\
\Rightarrow 2a \cdot b &\geq -a \cdot a \\
\|z\|_2^2 &= (u^2 a \cdot a + v^2 b \cdot b + 2uva \cdot b) \\
&\geq (u^2 a \cdot a + v^2 b \cdot b - uva \cdot a) \\
&\geq (u(u - v)a \cdot a + v^2 b \cdot b) \\
&\geq (u(u - v)a \cdot a) \\
&\geq \|a\|_2^2
\end{aligned}$$

Case (i) : $u \leq v$

We now observe that

$$\begin{aligned}
(a + b) \cdot (a + b) &\geq a \cdot a \\
\Rightarrow 2a \cdot b &\geq -b \cdot b \\
\|z\|_2^2 &\geq (u^2 a \cdot a + v(v - u)b \cdot b) \\
&\geq (u^2 a \cdot a) \\
&\geq \|a\|_2^2 \qquad \blacksquare
\end{aligned}$$

For the case $n > 2$, we employ the LLL basis reduction to find a small vector. The LLL basis reduction will be discussed in the next lecture.