Lecture 11: Factoring Bivariate Polynomials

Algorithmic Number Theory (Fall 2014)
Rutgers University
Swastik Kopparty
Scribe: Andrew Lohr

Roughly, factoring bivariate polynomials is like factoring univariate polynomials over $\mathbb{Q}$.

Given a field $\mathbb{F}$, $F[T,X] \subseteq F(T)[X] \approx \mathbb{Q}[x]$

1 General Idea

1) Find an approximate root g of $F(T,X)$, that is, $X = g(T)$. This will be a power series in T that, if allowed to be an infinite power series, we can hope for a true root. Instead, we will just truncate it to find an approximate root.

2) Find a minimal polynomial, $G(T,X)$ of $g(T)$ found in previous step.

2 Algorithm

Given a polynomial $F \in F[T,X]$ to be factored,

1) Make $F(T,X)$ monic in X by doing a linear change of variables.

2) Make $F(T,X)$ squarefree in $\mathbb{F}(T)[X]$ by using the derivative trick.

3) Find $t_0 \in \mathbb{F}$ such that $F(t_0, X)$ is a squarefree univariate polynomial in $\mathbb{F}[X]$. We know that such a t_0 exists by the discriminant argument, and how to find it follows from the proof done last class. Try $2d^2$ different choices of $t \in \mathbb{F}$. If the size of $\mathbb{F}$ is too small, extend the base field to one that has size at least $2d^2$, where d is the total degree of $F(T,X)$. Being able to factor over this extension gets us factoring over $\mathbb{F}$.

3.5) Shift the origin of T so that $t_0 = 0$.

4) For some extension $\mathbb{K}$ of $\mathbb{F}$, find a root $\alpha \in \mathbb{K}$ of $F(0,X)$. As an example, suppose that $F(0,X) = (X^2-2)(x^3-7)$. Then, we could extend $\mathbb{F}$ to $\mathbb{K} = \mathbb{F}[y]/\langle y^2-2\rangle$. Then $F(0,X)$ has a root in $\mathbb{K}$, namely y, and we can continue working in $\mathbb{K}$ from here on.

5) Find $\alpha_0 + \alpha_1 T + \alpha_2 T^2 + \cdots + \alpha_{k-1}T^{k-1} = g_k(T)$ such that $F(T, g_k(T)) \equiv 0 \mod T^k$. Note that this is where we use the squarefreeness of $F(T, g_k(T))$.

6) Find $G(T,X)$ such that $deg_X(G(T,X)) < deg_X(F(T,X))$, $deg_T(G(T,X)) \leq deg_T(F(T,X))$, and $G(T, g_k(T)) \equiv 0 \mod T^k$. We can find this because it can be written as a system of linear equations because $G(T,X) = \Sigma a_{ij} T^i X^j$ and then we just need solve for a_{ij}. This process is like finding the minimal polynomial of an approximate root that we saw in earlier lectures. If there exists such a G, the G of minimal X-degree is a factor of F. Note that we will end up taking k to be approximately $2d^2$.

3 Analysis

3.1 Making $F(T, X)$ monic in X

Suppose that the total degree of $F(T, X)$ is d. We want to find a, b, c, e such that

$$F(aT + bX, cT + eX) = Q(a, b, c, e)X^d + H(T, X)$$

Where Q is some nonzero polynomial, and $deg_X(H) < d$. Note that only b and e will be affecting Q since any of the terms containing an a or a c will also contain T, and therefore not be full X-degree. We then write $F = F_d + F_{<d}$, separating out the terms that have total degree d from those which have smaller total degree. Then, we just try letting $a = c = 0$, we have

$$F_d(bX, eX) = F_d(b, e)X^d$$

We know that $F_d(b, e)$ is a nonzero polynomial because there had to of been terms in F that attained total degree d. So, there is some choice of b and e that makes $F_d(b, e) \neq 0$. So, $Q(b, e) = F_d(b, e) \neq 0$. Scaling the entire polynomial by $Q(b, e)^{-1}$ completes this step

3.2 Make $F(T, X)$ squarefree

There is not much to say for this step, as we've seen many similar things before. If $\frac{\delta F}{\delta x}(T, X) = 0$, then use the trick as in the univariate factoring case. Otherwise, take $GCD(F(T, X), \frac{\delta F}{\delta x}(T, X)) \in \mathbb{F}(T)[X]$. If it is degree 0 in X, then F is squarefree, otherwise, we just found a factor of F and are done.

3.3 Finding t_0 to make $F(t_0, X)$ squarefree

Not much will be said on step 3, as we covered how to do it last class. How step 3.5 works is obvious.

3.4 Finding a root of $F(0, X)$

We know from step 1 that $F(0, X)$ is a non-zero degree d polynomial (that is squarefree from step 2). Since this is a univariate polynomial, we can factor it as

$$F(0, X) = \prod F_i(X)$$

where each F_i is irreducible.

Then, take any $F_i(X)$ and consider the field $\mathbb{K} = \mathbb{F}[Y]/\langle F_i(Y)\rangle$. Now, we know that $F(0, X)$ has $\alpha = Y$ as a root over $\mathbb{K}$. We then consider $\mathbb{K}$ to be the field we are working over from here on. By squarefreeness, we know that $\frac{\delta F}{\delta x}(0, \alpha) \neq 0$. This completes step 4.

3.5 Finding $g_k(T)$

We will be proceeding in a manner similar to the way that the Implicit Function Theorem. First note that $F(0, \alpha) = 0$ is equivalent to $F(T, \alpha) \equiv 0(\mod T)$. So, we can let α_0 be the root found in the previous step. We will procede by induction, as an example first step, we want to find an α_1 such that

$$F(T, \alpha_0 + \alpha_1 T) \equiv 0(\mod T^2)$$

. We exapnd using Taylor's theorem for polynomials to get

$$F(T, \alpha_0) + \frac{\delta F}{\delta x}(T, \alpha_0)\alpha_1 T + \frac{1}{2!}\frac{\delta^2 F}{\delta x^2}(T, \alpha_0)(\alpha_1 T)^2 + \cdots$$

We might be concerned that $\frac{1}{2!}$ does not make sense in our field, as it may have characteristic 2. To get around this issue, we briefly introduce the Hasse derivative:

Suppose we have $H(x) \in \mathbb{F}[x]$, then, we can think of expanding $H(x+z)$ and grouping together all the terms that have z^i in them for each i. The coefficient of z^i is defined to be the i^{th} Hasse derivative.

The i^{th} Hasse derivative can take the place of $\frac{1}{i!}\frac{\delta^i F}{\delta x^i}F(T, \alpha_0)$ when applying Taylor's Theorem.

Turning our attention back to infinite polynomial obtained by Taylor's Theorem, all but the first two terms are $\equiv 0(\mod T^2)$, and so, we may drop them, and we are left with

$$F(T, \alpha_0) + \frac{\delta F}{\delta x}(T, \alpha_0)\alpha_1 T \equiv 0(\mod T^2)$$

Trouble can happen when trying to solve this if $\frac{\delta F}{\delta x}(T, \alpha_0)$ is divisible by T. However, if it was divisible by T, then we wouldn't have $\frac{\delta F}{\delta x}(0, \alpha_0) = 0$ as we found in the previous step by square-freeness of $F(0, X)$. Note also that T divides $F(T, \alpha_0)$ because, as a polynomial in T, it must have a root at 0 by the way we selected α_0.

We write

$$F(T, \alpha_0) = \beta T + \delta_1(T)T^2$$
$$\frac{\delta F}{\delta x}(T, \alpha_0) = \gamma + T\delta(T)$$

For some non-zero γ, and some polynomials δ_1, δ_2. Then, rewriting our expression, and eliminating all terms that are multiples of T^2, we get

$$\beta T + \gamma \alpha_1 T \equiv 0(\mod T^2)$$

It is then trivial to solve for $\alpha_1 = \frac{-\beta}{\gamma}$

Now, suppose that we have $g_\ell = \alpha_0 + \alpha_1 T + \cdots + \alpha_{\ell-1}T^{\ell-1}$ such that $F(T, g_\ell(T)) \equiv 0(\mod T^\ell)$. We will look for an α_ℓ such that

$$F(T, g_\ell(T) + \alpha_\ell T^\ell) \equiv 0(\mod T^{\ell+1})$$

We will just apply Taylor's Theorem again, to get that this is

$$F(T, g_\ell(T)) + \frac{\delta F}{\delta x}(T, g_\ell(T))\alpha_\ell T^\ell + T^{2\ell} poly(T) \equiv 0(\mod T^{\ell+1})$$

reducing modulo $T^{\ell+1}$, and remembering what we know about g_ℓ, we have that this expression is

$$\beta_\ell T^\ell + (\gamma_\ell + T\text{poly}(T))\alpha_\ell T^\ell \equiv \beta_\ell T^\ell + \gamma_\ell \alpha_\ell T^\ell \equiv 0(\mod T^{\ell+1}$$

Note that we have $\gamma_\ell \neq 0$ because we may write $\frac{\delta F}{\delta x}(T, X) = \frac{\delta F}{\delta x}(0, \alpha_0) + T\text{poly}(T)$. So, again, we get that $\alpha_\ell = \frac{-\beta_\ell}{\gamma_\ell}$. This completes step 5.

As an aside, tricks of this type can be used to modify the algorithm to be done very quickly in parallel.

3.6 Finding Minimal Polynomial

We now are looking for a $G(T, X) \neq 0$ such that $G(T, g_k(T)) = 0(\mod T^k)$, $deg_X(G) < d$, and $deg_T(G) \leq d$.

We make the following claims:

1. If F has a nontrivial factor, then we will find a nonzero G.
2. If we find a nontrivial, non-zero G, then we can find a nontrivial factor of F.

We start with the proof of claim 1. Suppose that $F = F_1 F_2$, then, we know:

$$F_1(T, g_k(T))F_2(T, G_k(T)) \equiv 0(\mod T^k)$$

We want that one of these two factors is zero, then we can let that one be our G. The crux of showing this is that if one of them has a non-zero constant term, the other must be zero. If we have neither with a non-zero constant term, then both are divisible by T, then F is divisible by T^2, contradicting the fact that F was made to be squarefree.

Motivating Note: Suppose we have $f(x) \in \mathbb{Z}[x], \alpha \in \mathbb{R}, g(x) \in \mathbb{Z}[x]$ with all coefficients ≤ 100 and degree 5. If $|f(\alpha)| < 2^{-100}$, and $|g(\alpha)| < 2^{-100}$, then f and g have a common factor.

The remainder of the analysis will be left until next lecture.