6.885 Algebra and Computation	October 12, 2005
Lecture 9	
Lecturer: Madhu Sudan	*Scribe: Brendan Juba*

Today, we will continue our approach to factoring bivariate polynomials. We will begin by recalling Hensel's Lifting Lemma, and we will discuss a few new things about the Lemma. We will then give an outline of how we will perform the factoring, and develop another concept: the resultant. Once we have these concepts, the algorithm will be easy, and moreover, these concepts will help us in other domains.

1 Hensel's Lifting

Suppose $f(x, y) = g(x, y)h(x, y) \pmod{y^k}$. We wanted a factorization for higher powers of y: $f(x, y) = \tilde{g}(x, y)\tilde{h}(x, y) \pmod{y^{2k}}$. Hensel's Lifting says we can obtain this if g and h are "relatively prime," and that the $\tilde{g}$ and $\tilde{h}$ we obtain are essentially unique. Formally:

Lemma 1 (Hensel's Lifting) *If R is a ring, $J \subseteq R$ is an ideal, and there exist f, g, h, a, b such that*

(H1) $f = gh \pmod{J}$

(H2) $ag + bh = 1 \pmod{J}$

then there exist $\tilde{g}, \tilde{h}, \tilde{a}, \tilde{b}$ such that

(C1) $g = \tilde{g} \pmod{J}$ *and* $h = \tilde{h} \pmod{J}$

(C2) $f = \tilde{g}\tilde{h} \pmod{J^2}$

(C3) $\tilde{g}\tilde{a} + \tilde{b}\tilde{h} = 1 \pmod{J^2}$

Uniqueness: if $g^\star$, $h^\star$ also satisfy (C1) and (C2), then $\exists u \in J$ such that $g^\star = g(1+u)$ and $h^\star = h(1-u)$.

These conditions allow us to continue to y^{4k} and so on.

We are planning to apply this with $R = \mathbb{F}_q[x, y]$ and $J = (y - y_0)^k$ for some $y \in \mathbb{F}_q$. We'll start with a factorization in $\mathbb{F}_q[x]$, and we'll increase the powers of y. We ultimately hope to get a factorization that is not modulo any ideals.

Remark We were given explicit formulas for $\tilde{g}$ and $\tilde{h}$. So, given g, h, a, b, f, we can compute $\tilde{g}, \tilde{h}, \tilde{a}, \tilde{b}$ in $O(1)$ operations in R, $\pmod{J}$, $\pmod{J^2}$, and so on.

For example, $\tilde{g} = g + g_1$ where $g_1 = a(f - gh)$.

Although we need to be capable of efficiently carrying out ring operations, and the operations of reducing modulo an ideal, $\tilde{g}$ and $\tilde{h}$ don't require any serious computations.

One more clarification: uniqueness will take on a nicer form for us.

Corollary 2 *Suppose $R = \mathbb{F}_q[x, y]$ and $J = (y - y_0)^k$, and assume f, g, h are monic in x. Then, if g and h are relatively prime and $f = gh \pmod{J}$, there exist monic polynomials $\tilde{g}$ and $\tilde{h}$ such that $f = \tilde{g}\tilde{h} \bmod (y - y_0)^{2k}$ and $\tilde{g}, \tilde{h}$ are unique $\bmod (y - y_0)^{2k}$.*

Naturally, we can't expect any better than this modulo $(y - y_0)^{2k}$.

Proof Existence and the fact that the polynomials are monic follow immediately.

Suppose $\tilde{g}, \tilde{h}, g^\star, h^\star$ satisfy the conditions. Then, by Hensel's Lifting, $g^\star = \tilde{g}(1 + u)$, $h^\star = \tilde{h}(1 - u)$. If $u \neq 0 \bmod (y - y_0)^{2k}$, then $g^\star$ is not monic. ■

2 Outline of Factoring

We now turn to the question of how we can use Hensel's Lifting for bivariate factoring—specifically, how we initiate it, and what we do after applying it.

Suppose we are given $f(x, y)$ and we wish to factor it into $f_1(x, y)f_2(x, y)\cdots f_k(x, y)$. Suppose we pick some y_0 and substitute it in for y to obtain $f(x, y_0) = p^{(0)}(x)$, a polynomial of one variable, which we therefore know how to factor into $p_1^{(0)}(x)\cdots p_k^{(0)}(x)$. Suppose now that we are very lucky, and by picking $y_1 \neq y_0$, we obtain $f(x, y_1) = p^{(1)}(x) = p_1^{(1)}(x)\cdots p_k^{(1)}(x)$. Then we could "interpolate" pairs $(p_1^{(0)}(x), p_1^{(1)}(x)), (p_2^{(0)}(x), p_2^{(1)}(x)), \ldots (p_k^{(0)}(x), p_k^{(1)}(x))$ to obtain polynomials in y. Of course, we haven't formally defined what we mean by this, and moreover, it's not clear how we should pair up the factors of $p^{(0)}(x)$ and $p^{(1)}(x)$. Even more troublesome, depending on our choice of y_0, we could get many more than k factors!

For example, consider $f(x, y) = (x - 5)(x - 3)(x - 2) + y$, which is irreducible. If we now substitute $y = 0$, we obtain $f(x, 0) = (x - 5)(x - 3)(x - 2)$, which has three factors. We therefore see that we have a one-to-many correspondence between the factors of $f(x, y_0)$ and the factors of $p(x)$, in general:

$$\begin{array}{rcccccccc} f(x,y) & = & f_1(x,y) & \cdot & f_2(x,y) & \cdot\cdots\cdot & f_k(x,y) \\ p(x) & = & \overbrace{p_{11}(x)p_{12}(x)\cdots p_{1i_1}(x)} & \cdot & \overbrace{p_{21}(x)p_{22}(x)\cdots p_{2i_2}(x)} & \cdot\cdots\cdot & \overbrace{p_{k1}(x)p_{k2}(x)\cdot p_{ki_k}(x)} \end{array}$$

By Hensel's Lifting and a few other facts, we'll see that this is okay, that p_{11} and its many lifts will tell us f_1, provided that p_{11} is relatively prime to p_{ij} for all i and j.

Still, it is not yet clear how we will ensure that these elements are relatively prime. We know this will not happen if $f(x, y) = f_1(x, y)^5$. More generally, suppose $f(x, y) = f_1(x, y)^a f_2(x, y)\cdots f_k(x, y)$ for $a \geq 2$. It turns out that our intuitions from calculus hold here: suppose we take derivatives of f with respect to x and y. If $f_1(x, y)$ has positive degree in x, then $\frac{\partial f}{\partial x}$ and f will have a nontrivial GCD. We can argue this the same way as before, using the sum and product rules. So, it turns out that we don't need to worry about repeated factors in the factorization of f.

Now, suppose $f(x, y) = x^k + y$, and we try $f(x, 0) = p(x) = x\cdot x\cdots x$. We started with a polynomial with no repeated factors; how can we guarantee that our choice of y will not yield a polynomial with repeated factors? It's clear that not every choice of y fulfills this condition. In general, we will be looking at the GCD of a polynomial and its derivative, so we will be studying the GCDs of two polynomials more carefully.

We'd like reasoning of the form: if $f(x, y)$ and $g(x, y)$ $\left(= \frac{\partial}{\partial x}f(x, y)\right)$ have no common factors, then $\exists y_0$ such that $f(x, y_0)$ and $g(x, y_0)$ have no common factors.

This begs the question, how can we derive statements about the GCDs, when plugging in values for y changes the structure of the polynomial completely? Our answer will be the resultant.

3 Resultants

There are many ways of developing and talking about resultants, and our definition may be inconsistent with some of the others, but it will work for our purposes.

Definition 3 *Let R be a UFD with some notion of order on its elements*[1] *and let $f, g \in R[x]$. The* resultant of f and g *is then the "minimal" element in the ideal generated by (f, g): $\{af + bg | a, b \in R[x]\}$. That is, the element with the smallest leading coefficient among the lowest degree polynomials in that set (excluding the zero polynomial).*

When $R = \mathbb{Q}$, the lowest degree polynomial is easily obtained from the GCD, whereas if $R = \mathbb{Z}$ (e.g., if R is not a field), $f = 2x^3 + 3x$, and $g = 9x^{10}$, then we find that resultant(f, g) is some multiple of x (we take the smallest one).

[1] For $R = \mathbb{Z}$, this order is given by the magnitude. For $R = \mathbb{F}_q[y]$, the order is given by the degree in y

Now consider polynomials $f, g \in \mathbb{F}_q[x, y]$ that don't share a common factor. Let $\mathrm{Res}_x(f, g)$ denote the resultant of f and g when we take $R = \mathbb{F}_q[y]$. We will see that we can say that the x degree of $R(x, y) = \mathrm{Res}_x(f(x, y), g(x, y))$ is zero.

We can see this if we consider $f, g \in \widetilde{\mathbb{F}_q[y]}[x]$. Over a field, since f and g have no common factors, we know $\gcd(f, g) = 1 = R(x, y)$. Thus, over $\mathbb{F}_q[y]$, we should at least be able to get some polynomial in y (which is divided out in $\widetilde{\mathbb{F}_q[y]}$). That is, $\exists a(x, y), b(x, y)$ such that $a(x, y)f(x, y) + b(x, y)g(x, y) = R(y)$ where $R(y) \neq 0$.

Recall that we wanted to show $f(x, y_0)$ and $g(x, y_0)$ don't have a common factor; we generally show this via the extended GCD by demonstrating $af(x, y_0) + bg(x, y_0) = 1$. Now, if $R(y) \neq 0$, then in some extension field we will be able to find a y_0 such that this condition holds. We'll guarantee this by placing bounds on the degree of R.

We'd now like to find a and b to obtain $\mathrm{Res}(f, g)$. Note that our ideal $\{af + bg | a, b \in R[x]\}$ is a linear space and our conditions are linear, so we'll be able to solve for a and b easily.

Bounding the degree of $R(y)$: put $\deg_x f \leq d_1$, $\deg_y f \leq d_1$, $\deg_x g \leq d_2$, $\deg_y g \leq d_2$.

Claim 4 *$\exists R$, a nonzero polynomial with $\deg R \leq 2d_1 d_2$ such that $R(y) = a(x, y)f(x, y) + b(x, y)g(x, y)$.*

Proof Let the coefficients of a and b be unknowns, and $R(x, y) = a(x, y)f(x, y) + b(x, y)g(x, y)$. Note that R is temporarily a polynomial in x as well as y. We'll say that the coefficients of $x^i y^j$ are linear forms in the unknowns, and require that they be zero if $i > 0$. So we have a system of homogeneous linear constraints, where we need a solution other than $a, b = 0$ and $af + bg = 0$. First, we'll make the degrees in x small, so there's no chance of cancellation. Then, we'll "shoot for" $\deg R(y) = \Delta$. We bound $\deg_x a < \deg_x g$ (so $a \notin \{\pm g\}$) and $\deg_x b < \deg_x f$ (which similarly rules out $b \in \{\pm f\}$). We now further impose $\deg_y a < \Delta - d_1$ (so $\deg_y(af) < \Delta$) and $\deg_y b < \Delta - d_2$. Thus, now, $\deg_x R < \deg_x f + \deg_x g$ and $\deg_y R < \Delta$.

Claim 5 *$\exists a, b \neq 0$ satisfying these constraints.*

(e.g., with $\deg_x(af + bg) \leq 0$) Observe:

$$\begin{array}{llcll}
\#\text{ of unknowns}: & \underbrace{\deg_x g(\Delta - d_1)}_{\#\text{ coeffs in } a} & + & \underbrace{\deg_x f(\Delta - d_2)}_{\#\text{ coeffs in } b} & \geq \Delta(\deg_x f + \deg_x g) - 2d_1 d_2 \\
\#\text{ of constraints}: & \underbrace{(\deg_x f + \deg_x g - 1)}_{\#\text{ powers in } x} & \cdot & \underbrace{\Delta}_{\#\text{ powers in } y} & = \Delta(\deg_x f + \deg_x g) - \Delta
\end{array}$$

So we see that # unknowns > # constraints if $\Delta > 2d_1 d_2$: we can then find a and b such that there are no powers of x in R. All that remains is to show that $af + bg$ is not identically zero.

Claim 6 *We can't obtain an identically zero polynomial unless f and g share a common factor.*

Observe that, if f and g have no common factors, since we are in a UFD, then $af + bg = 0$ implies that the factorizations of af and bg are the same. Since f and g share no factors, a must have all of the factors of g and b must have all of the factors of f, which can't happen by our bounds on the degrees in x of a and b. ■

Now, by our bound on the degree of $R(y)$, for a sufficiently large extension field, there must be choices of y_0 such that $R(y_0) \neq 0$. Hence, if f and g have no repeated factors, for such a choice of y_0, $f(x, y_0)$ and $g(x, y_0)$ have no repeated factors.[2] We can take the derivative or plug in y_0 in any order we like.

[2] When $g = \frac{\partial f}{\partial x}$, $\mathrm{Res}(f, g)$ is called the "discriminant" of f. For $f = ax^2 + bx + c$, $\mathrm{disc}(f) = b^2 - 4ac$. We ignored questions about whether the discriminant was positive or negative, so our notions may not agree, but at least we know that ours is nonzero if there are no repeated roots, and is zero otherwise.

4 Outline of Factoring, revisited

We now give a more complete outline for factoring bivariate polynomials.

Given a monic $f(x, y) \in \mathbb{F}_q[x, y]$, we'll work with some extension field K of $\mathbb{F}_q$.

1. Find $y_0 \in K$ such that $f(x, y_0)$ has no repeated factors. (how do we do this? compute $\mathrm{Res}\left(f, \frac{\partial f}{\partial x}\right)$, and plug in $y_0 = 1, 2, \ldots$ until we find one that works.)
2. Put $p(x) = f(x, y_0)$, and factor it. Some factor should be linear; if not, we extend K. We know that every field can be extended until the polynomial splits completely, but we only need a low degree extension with one linear factor, say $p(x) = (x - \alpha)h(x)$. We assume $(x - \alpha)$ and $h(x)$ are relatively prime.
3. Now we start to apply Hensel's Lifting: $f(x, y) = (x - \alpha)h(x) \bmod (y - y_0)$. After some iterations, $f(x, y) = (x - A(y))H(x, y) \bmod (y - y_0)^t$.
4. Now, from $A(y)$ we ask if we can find a nontrivial factor of f. We will see that this is related to the notion of minimal polynomials. We will try to find $f_1(x, y)$ with $\deg_x f_1 < \deg_x f$ such that $f_1(x, y) = (x - A(y))H_1(x, y) \bmod (y - y_0)^t$. It will turn out that, if the f_1 we find is the same as f, then we will have a proof that f has no nontrivial factors!

We'll complete the last step by using linear algebra (which, along with the GCD, comprises our toolbox at the moment), e.g., we'll give linear constraints and show that solutions exist. Then, we'll say something along the lines of "if $\gcd(f_1, f)$ is nontrivial, report it, else f is irreducible."