6.S897 Algebra and Computation	March 12, 2012
Lecture 10	
Lecturer: Madhu Sudan	*Scribe: Rachel Lin*

Today, we will continue our approach to factoring bivariate polynomials. We will first focus on the tool of Hensel's Lifting; and then describe how we perform the factoring.

1 Hensel's Lifting

Suppose $f(x,y) = g(x)h(x) \pmod{y}$. 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, $I \subseteq R$ is an ideal, and there exist f, g, h, a, b in R such that*

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

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

Then, for every positive interger s that is a power of 2, there exist $\tilde{g}, \tilde{h}$ and $\tilde{a}, \tilde{b}$ in R, such that

(C1) $f = \tilde{g}\tilde{h} \pmod{I^s}$

(C2) $\tilde{g}\tilde{a} + \tilde{b}\tilde{h} = 1 \pmod{I^s}$

(C3) $g = \tilde{g} \pmod{I}$ *and* $h = \tilde{h} \pmod{I}$

Furthermore, the solution satisfying the above three conditions is "unique" in the following sense:

Uniqueness: *We say that an ideal J is* special *if for all all interger k, and a, b such that $ab \in J^k$, there is an integer l such that $a \in J^l$ and $b \in J^{k-l}$.*

Assume that I is special. Then, for every two solutions g_1, h_1 and g_2, h_2 satisfying contidions C1 to C3, there exists $u \in I^t$, such that:

$$g_2 = g_1(1+u) \pmod{I^s}, \qquad h_2 = h_1(1-u) \pmod{I^s}$$

As we will see later, towards factoring a bivariate polynomial $f \in \mathbb{F}_q[x,y]$, we will apply the Hensel's lifting to $R = F_q[x,y]$, $I = (y)$ and a factorization of f mod I. When $I = (y)$, $f \pmod{y}$ is simply a univariate polynomial on x over $\mathbb{F}_q$, which we know how to factor from previous lectures. Next we proceed to prove Lemma 1.

Proof We prove this lemma by induction.

Base Case $s = 2$: We first show that there exists $\tilde{g}, \tilde{h}$ and $\tilde{a}, \tilde{b}$ satisfying C1 to C3, and then establish the uniqueness of the solution. By condition H1 and H2, we have

$$f = gh + q, \text{ for some } q \in I$$

$$ag + bh = 1 + r \text{ for some } r \in I$$

Write $\tilde{g} = g + g_1$, $\tilde{h} = h + h_1$, for some $g_1, h_1 \in I$ to be set. Then

$$\tilde{g}\tilde{h} = gh + g_1 h + h_1 g + h_1 g_1$$

Since $h_1 g_1 \in I^2$, in order to satisfy condition C1, we want $g_1 h + h_1 g + h_1 = q \pmod{I^2}$. To satisfy this, set $g_1 = bq$, $h_1 = aq$, and get that $g_1 h + h_1 g = q(bh + ag) = q(1+r)$, which equals to q modulo I^2 as required. By construction $\tilde{g} = g \pmod{I}$ and $\tilde{h} = h \pmod{I}$, satisfying condition C3. To show that $\tilde{g}$ and $\tilde{h}$ are also relatively prime, observe that $a\tilde{g} + b\tilde{h} = ag + bh + r' = 1 + r + r'$, for some $r' \in I$. Let $r'' = r + r' \in I$. Now we can take $\tilde{a} = a(1 - r'')$ and $\tilde{b} = b(1 - r'')$, and get that:

$$\tilde{a}\tilde{g} + \tilde{b}\tilde{h} = (1 - r'')(a\tilde{g} + b\tilde{h}) = (1 - r'')(1 + r'') = 1 - r''^2 = 1 \pmod{I^2}$$

Now it remains to show that $\tilde{g}, \tilde{h}$ is the unique solution satisfying C1 to C3. That is, if g^*, h^* is a different solution satisfying C1 to C3, then there is $u \in I$ such that $g^* = \tilde{g}(1+u)$ and $h^* = \tilde{h}(1-u)$. By contidion C3, we have $g^* = \tilde{g} + g_2$ and $h^* = \tilde{h} + h_2$ for some $g_2, h_2 \in I$ (because, modulo I, we know that $g^* = g = \tilde{g}$ and $h^* = h = \tilde{h}$). Therefore, we have:

$$g^* h^* = \tilde{g}\tilde{h} + g_2\tilde{h} + h_2\tilde{g} + g_2 h_2$$

By condition C1, we know that $g^* h^* = f = \tilde{g}\tilde{h} \pmod{I^2}$. Thus, the above equation modulo I^2 gives,

$$g_2\tilde{h} + h_2\tilde{g} = 0 \pmod{I^2}$$

By condition C2, we have $\tilde{a}\tilde{g} + \tilde{b}\tilde{h} = 1 \pmod{I^2}$. Therefore,

$$\begin{aligned}
&\tilde{b}(g_2\tilde{h} + h_2\tilde{g}) = 0 \pmod{I^2} \\
&g_2\tilde{b}\tilde{h} + \tilde{b}h_2\tilde{g} = 0 \pmod{I^2} \\
&g_2(1 - \tilde{a}\tilde{g}) + \tilde{b}h_2\tilde{g} = 0 \pmod{I^2} \\
&g_2 = (\tilde{a}g_2 - \tilde{b}h_2)\tilde{g} \pmod{I^2}
\end{aligned}$$

Let $u = \tilde{a}g_2 - \tilde{b}h_2$. Since g_2 and h_2 are all elements in I, so is u. Furthermore we have $g^* = \tilde{g} + g_2 = \tilde{g}(1+u)$. Similarly, by symmetry, we obtain that

$$h_2 = (\tilde{b}h_2 - \tilde{a}g_2)\tilde{h} \pmod{I^2}$$

Therefore, $h^* = \tilde{h}(1-u)$. This concludes the proof for the base case.

Induction Step: Assume that for the case of $s = t$, there exist $g_0, h_0 \in R[x]$ and $a_0, b_0 \in I^t$ satisfying conditions C1 to C3, and the solution to the three conditions is "unique". We show that for the case of $s = 2t$, we can construct $g_1, h_1 \in R[x]$ and $a_1, b_1 \in I^{2t}$ satisfying conditions C1 to C3, and the solution is also unique.

The existence of g_1, h_1, a_1, b_1 satisfying conditions C1 to C3 follows exactly the same proof as in the base case. Therefore, we focus on the proof of uniqueness. Let g_1, h_2 be a different solution from g_1, h_1. Then both $g_1 \pmod{I^t}$, $h_1 \pmod{I^t}$ and $g_2 \pmod{I^t}$, $h_2 \pmod{I^t}$ are solutions satisfying C1 to C3 for the case of $s = t$. Then by the induction hypothesis, we have that there is a $u_0 \in I^{t/2}$ such that,

$$(g_2 \pmod{I^t}) = (g_1 \pmod{I^t})(1 + u_0) \pmod{I^t}$$

$$(h_2 \pmod{I^t}) = (h_1 \pmod{I^t})(1 - u_0) \pmod{I^t}$$

This implies that

$$g_2 = g_1(1 + u_0) \pmod{I^t}$$

$$h_2 = h_1(1 - u_0) \pmod{I^t}$$

Notice that this is differerent from the condition in the base case where any two solutions must equal modulo I. Nevertheless, following the same argument, we can derive that there is an element $u \in I^t$ such that,

$$g_2 = g_1(1 + u_0)(1 + u) \pmod{I^{2t}}$$

$$h_2 = h_1(1-u_0)(1-u) \pmod{I^{2t}}$$

Below we show that u_0 is in fact an element in I^t, then

$$g_2 = g_1(1 + u_0 + u + u_0 u) = g_1(1 + u_0 + u) \pmod{I^{2t}}$$

$$h_2 = h_1(1 - u_0 - u + u_0 u) = h_1(1 - u_0 - u) \pmod{I^{2t}}$$

Thus $g_2 = g_1(1 + u')$ and $h_2 = h_1(1 - u')$ for $u' = u_0 + u \in I^t$ as desired.

To show that $u_0 \in I^t$, consider:

$$g_2 h_2 = g_1 h_1 (1 - u_0^2)(1 - u^2) \pmod{I^{2t}}$$
$$g_2 h_2 = g_1 h_1 (1 - u_0^2) \pmod{I^{2t}} \qquad [\text{as } u \in I^t]$$

Since g_1 and h_1 are not elements in I, for the last euqation to hold, it must be the case that $u_0 \in I^t$. Therefore, we conclude the lemma. ■

2 Outline of Factoring, revisited

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

Given a monic $f(x, y) \in \mathbb{F}_q[x, y]$, with total degree d, the factoring algorithm $SPLIT$ proceeds as follows:

1. If $g = \gcd(f, \frac{\partial f}{\partial x}) \neq 1$, then output $(g, f/g)$ and stop. Otherwise, continue the following steps.
2. Find $y_0 \in \mathbb{F}$ such that $f(x, y_0)$ has no repeated factors. This can be done by computing $\text{Res}\left(f, \frac{\partial f}{\partial x}\right)$, and pluging in $y_0 = 1, 2, \ldots$ until we find one that makes the resultant non-zero.

   We claim that this will terminiate in at most d^2 iterations, as $\text{Res}\left(f, \frac{\partial f}{\partial x}\right)$ is a polynomial in y with degree at most d^2. Futhermore, the first step ensures that f does not have repeated roots; therefore, $\text{Res}\left(f, \frac{\partial f}{\partial x}\right)$ is not a zero polynomial. Hence it has at most d^2 roots.
3. Put $f_{y_0}(x) = f(x, y) \pmod{(y - y_0)} = f(x, y_0)$ and factor it. This can be done by using the factoring algorithm for univariate polynomial over $\mathbb{F}$. Let g be an irreducable factor of $f_{y_0}(x)$, and h such that $f = gh \pmod{(y - y_0)}$.
4. Now we apply Hensel's Lifting to obtain $f = g_1 h_1 \pmod{(y - y_0)^t}$ for a $t \approx d^2$
5. Next, from g_1 we ask if we can find a nontrivial factor $\tilde{g}$ of f. This is done through the "Jump" step, which tries to find polynomials $\tilde{g}$ and $\tilde{h}$ such that $\tilde{g} = g_1 \tilde{h} \pmod{(y - y_0)^t}$, and $\tilde{g}$ has small degrees in y (smaller than d) and minimal degree in x.
6. Finally, return $\tilde{g}$ and $f/\tilde{g}$.