**6.S897 Algebra and Computation** March 4, 2015

Lecture 9

*Lecturer: Madhu Sudan* *Scribe: Sung Min Park*

# 1 Overview

Today, we will continue our quest for factorization of bivariate polynomials. In particular, we will see the use of resultants that we learned about last time, and see Hensel lifting in more detail.

# 2 Factoring $\mathbb{F}_q[x, y]$

Recall from last time our overall algorithm for factoring:

1. Factor $f = g \cdot h \,(\text{mod } y)$ where $g, h \in \mathbb{F}_q[x]$. Here, $g$ and $h$ should be relatively prime[1] and of positive degree. WLOG $g$ is irreducible.

2. A procedure called Hensel lifting is used to lift the above factors to $f = \tilde{g} \cdot \tilde{h} \,(\text{mod } y^{2^t})$. Noe that now $\tilde{g}, \tilde{y}$ may have terms involving $y$. We iterate Hensel lifting $t$ times to find $\tilde{g}, \tilde{h}$ s.t. $f = \tilde{g} \cdot \tilde{h} \,(\text{mod } y^{2^t})$. The hope is that if $t$ is large enough $\tilde{g}$ will contain enough information to recover the original factor $g$. Hensel lifting maintains the invariant that $\tilde{g} = g \,(\text{mod } y)$ and $\tilde{h} = h \,(\text{mod } y)$.

3. The "Jump" step: We focus on $\tilde{g}$ from this point on. Find the lowest $x$ degree polynomial $A \in \mathbb{F}_q[x, y]$ s.t. $\exists \tilde{A} \in \mathbb{F}_q[x, y]$ s.t. $A = \tilde{g} \cdot \tilde{A} \,(\text{mod } y^{2^{t-1}})$. The hope is that $A$ will be a factor of $f$.

The last step in the above algorithm is quite mysterious. Why should $A$ be factor of $f$? To get some intuition, suppose $f = A \cdot B$ (here $f$ is square-free and $A$ is irreducible).

When we mod $f$ by $(y)$, in general the factors will split further[2]:

$$\begin{aligned} f(x, y) &= A \cdot B \\ f \,(\text{mod } I) &= g \cdot \underbrace{h_1 \cdot b}_{h} \end{aligned}$$

After iterations of Hensel lifting, $A = g \cdot h_1 \,(\text{mod } y))$ is lifted to $A = g' \cdot h_1' \,(\text{mod } y^{2t})$. Then, if we consider $\tilde{g} = g'$ and $\tilde{h} = h_1' \cdot B \,(\text{mod } y^{2^t})$, $A$ is in fact a solution of the above "Jump" problem.

One concern here is the uniqueness of $\tilde{g}$ and $\tilde{h}$. In fact, given solutions $\tilde{g}$ and $\tilde{h}$, $\tilde{g}(1 + u)$ and $\tilde{h}(1 - u)$ for any $u \in (y^{2^{t-1}})$ are also equal to $f$ modulo $y^{2^t}$. This is the reason why we use $y^{2^{t-1}}$; this kills other candidates.

## 2.1 The "Jump" problem

Now that we have seen some plausibility of why $A$ might be a factor, we give some more details regarding the solutions of the "Jump" problem . It's easy to observe that solutions to the "Jump" problem form a linear subspace:

[1] To see why we need $g, h$ to be relatively prime, consider $f = (x + y)(x - y)$. Modulo $y$, this becomes $x^2$. If we let $g = x$ and $h = x$, then we lose information and don't know which of the original factor to recover

[2] Example: there are bivariate complex polynomials that are irreducible, but we know that complex polynomial in single variable will split completely

**Claim 1** *If $(A, \tilde{A})$ and $(B, \tilde{B})$ are solutions, then so are $(p_1 \cdot A + p_2 \cdot B, p_1 \cdot \tilde{A} + p_2 \cdot \tilde{B})$ for any $p_1, p_2 \in \mathbb{F}_q[x, y]$.*

This means that we can find $A$ and $\tilde{A}$ just by solving a linear system.

But it remains of question whether the solution will be unique. The sketch of our uniqueness argument is as follows. Suppose $(A, \tilde{A})$ is a solution with minimal degree in $x$. If there is another solution $(B, \tilde{B})$, then $(\mathrm{Res}_x(A, B), \tilde{R})$ is also a solution since $\mathrm{Res}_x(A, B)$ is in the ideal generated by $A$ and $B$. Because $\mathrm{Res}_x(A, B) \in \mathbb{F}_q[y]$, [3] this means the following form of identity has to hold:

$$R(y) = \tilde{g}(x, y) \cdot \tilde{R}(x, y) \pmod{y^{2^{t-1}}} \text{ where } \tilde{g} \text{ has positive degree in } x$$

But this is impossible[4]. So $A$ must in fact be unique.

## 3 Factoring $\mathbb{Z}[x]$

Before we continue our discussion of the factoring algorithm for $\mathbb{F}_q[x, y]$, we note that the algorithm can be readily adapted for factoring integer polynomials $\mathbb{Z}[x]$, modulo some differences in the last step. The following is a rough outline of the algorithm for factoring $\mathbb{Z}[x]$:

1. Factor $f = gh \pmod{p}$. Prime $p$ plays the role of ideal $(y)$ from before. We pick $p$ s.t. $f$ has no repeated factors modulo $p$.
2. $f = \tilde{g} \cdot \tilde{h} \pmod{p^{2^t}}$
3. Find polynomial $A$ with lowest degree and smallest coefficients s.t. $A = \tilde{g} \cdot \tilde{A} \pmod{y^{2^{t-1}}}$.

There are two concerns here. One is that the coefficients of factors of $A$ may be very large; we can show that we don't have to worry about this. Secondly, the algebraic properties of the solutions to the "Jump" problem in the last step are different. Whereas before we had a nice linear subspace to work with, since the sum of two degree $n$ polynomials is still degree $n$, we now have the issue of coefficients need to be bounded. The "Jump" problem in this new setting reduces to finding a short basis in a certain lattice. This can be solved by the Lenstra-Lenstra-Lovasz (LLL) algorithm, which we will see later.

## 4 Hensel's Lifting

We now go back in see how to do Hensel lifting from step 2 of the factoring algorithm. We state the properties of the lift in the following lemma, and will see how it's actually done in the proof.

**Lemma 2** *Let $R$ be a UFD (ex. $R = \mathbb{F}_q[x, y], \mathbb{Z}[x]$), and $I \subseteq R$ an ideal.*

$$\begin{array}{rl} \textit{Given } f \textit{ and relatively prime } g, h \textit{ s.t.} & f = g \cdot h \pmod{I} \\ & 1 = a \cdot + b \cdot h \pmod{I} \\ & \quad\quad \Downarrow \\ \textit{Then we can find } \tilde{g}, \tilde{h} \textit{ s.t.} & f = \tilde{g}\tilde{h} \pmod{I^2} \\ & 1 = \tilde{a}\tilde{g} + \tilde{b}\tilde{h} \pmod{I} \\ & \tilde{g} = g \pmod{I} \\ & \tilde{h} = h \pmod{I} \end{array}$$

*Moreover, if we repeat this operation to get $\tilde{\tilde{g}}\tilde{\tilde{h}}$ modulo $I^4$, then $\tilde{g}$ and $\tilde{\tilde{h}}$ will be unique modulo $I^2$.*

[3] See the previous lecture for more on properties of the resultant.

[4] This mostly has to do with the fact that $\tilde{g}(x, y)$ has positive degree in $x$. More formal proof will be shown later.

**Proof** Proof is by induction. The work below is for the base case.

Let $\tilde{g} = g + g_1$ and $\tilde{h} = h + h_1$, where $g_1, h_1 \in I$ so that

$$f = (g + g_1)(h + h_1) \,(\text{mod } I^2)$$

Since $f = g \cdot h \,(\text{mod } I)$, let $f - gh = p \in I$. Then, we can rewrite the above as

$$p + g_1 + h_1 g + g_1 h_1 = 0 \,(\text{mod } I^2)$$

The last term disappears since $g_1, h_1 \in I$, their product is in $I^2$

$$g_1 + h_1 g = -p \,(\text{mod } I^2)$$

Is there a solution? Yes, in fact for any $p$ since the gcd of $g$ and $h$ is 1. Choosing $g_1 = -bp$ and $h_1 = -ap$, we can verify that $f = \tilde{g}\tilde{h} \,(\text{mod } I^2)$. Also, by definition $\tilde{g} = g \,(\text{mod } I)$ and $\tilde{h} = h \,(\text{mod } I)$.

Now we show that $\tilde{g}$ and $\tilde{h}$ are relatively prime modulo $I^2$. Let $1 = ag + bh + q$ where $q \in I$ by the gcd assumption. Then, $a\tilde{g} + b\tilde{h} = ag + bh + r = 1 + q + r$ for some $r \in I$. Let $s = q + r \in I$. Now take $\tilde{a} = a(1 - s)$ and $\tilde{b} = b(1 - s)$, and get:

$$\tilde{a}\tilde{g} + \tilde{b}\tilde{h} = (1 - s)(a\tilde{g} + b\tilde{h}) = (1 - s)(1 + s) = 1 - s^2 = 1 \,(\text{mod } I^2)$$

We have found lifts $\tilde{g}, \tilde{h}$ with the desired properties.

Finally, we must show that these are unique. Suppose there is another solution $g^* = \tilde{g} + g_2$ and $h^* = \tilde{h}$. [5] So we have:

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

We know that $g^* h^* = f = \tilde{g}\tilde{h} \,(\text{mod } I^2)$. Thus, modulo $I^2$ we have:

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

Define $u = \tilde{a}g_2 - \tilde{b}h_2$. Since $g_2, h_2 \in I$, $u \in I$. So we have $g^* = \tilde{g} + g_2 = \tilde{g}(1 + u)$. By symmetry, we find that

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

Therefore, $h^* = \tilde{h}(1 - u)$.

For the inductive case, the existence part is exactly the same. The uniqueness part requires a bit more work.[6] ■

The rest of these notes are written by Madhu Sudan.

[5] Note that notation was been slightly altered from the presentation in lecture

[6] See the notes from 2012 for rest of the proof.

# 5 Madhu's Addendum/Erratum

First I'd like to apologize for an error in the lecture.

Lemma 2 as stated is incorrect in full generality. It is probably the case that by considering the kind of ideal $I$ that need to be considered for our applications we might be able to patch it, but it turns out a simpler lemma is cleaner to state, actually provable (:-)), and sufficient. Lets start with the lemma statement.

**Definition 3** *Let $R$ be a* ring (ex. $R = \mathbb{F}_q[x,y], \mathbb{Z}[x]$), and $I \subseteq R$ an ideal. Given $f, g, h \in R$ such that $f = gh \,(\mathrm{mod}\ I)$ we say that *$(\tilde{g}, \tilde{h})$ lift $(g,h)$* if $f = \tilde{g}\tilde{h} \,(\mathrm{mod}\ I^2)$, $g = \tilde{g} \,(\mathrm{mod}\ I)$, $h = \tilde{h} \,(\mathrm{mod}\ I)$, and there exist $\tilde{a}, \tilde{b} \in R$ such that $\tilde{a}\tilde{g} + \tilde{b}\tilde{h} = 1 \,(\mathrm{mod}\ I^2)$.

**Lemma 4** *Let $R$ be a ring and $I \subseteq R$ an ideal. Given $f, g, h \in R$ such that $f = gh \,(\mathrm{mod}\ I)$ and there exist $a, b \in R$ such that $af + bh = 1 \,(\mathrm{mod}\ I)$ there exists a lift $(\tilde{g}, \tilde{h})$ of $(g,h)$. Furthermore, $(g_1, h_1)$ and $(g_2, h_2)$ both lift $(g,h)$ if and only if there exist $u \in I$ such that $g_1 = g_2(1+u)$ and $h_1 = h_2(1-u)$.*

Main changes are we no longer require $R$ to be a UFD (not even that!), but we don't go to $I^4$ nor do we prove uniqueness upto $I^2$. Instead we prove a simpler equivalence of multiple solutions which is sufficient. Note that condition is symmetric since we have $g_1(1-u) = g_2(1-u^2) = g_2 \,(\mathrm{mod}\ I^2)$.

We argue sufficiency first.

**Sufficiency.** Let $f = AB$ with $f, A, B \in \mathbb{F}[x,y]$ and $A$ irreducible. Further, let $f = gh \,(\mathrm{mod}\ y)$ with $g$ being irreducible and relatively prime to $h$. Assume $A = gh' \,(\mathrm{mod}\ y)$. Finally let $(g^{(i)}, h^{(i)}$ be the lift of $(g^{(i-1)}, h^{(i-1)})$ modulo $(y^{2^i})$ with $g^{(0)} = g$ and $h^{(0)} = h$.

Main point (that we didn't make in the lecture) is that it suffices to show that there exists $(h')^{(i)}$ such that $(g^{(i)}, (h')^{(i)})$ is the lift of $(g^{(i-1)}, (h')^{(i-1)})$ modulo $(y^{2^i})$ and $h^{(i)} = B \cdot (h')^{(i)} \,(\bmod\ y^{2^i})$. If we show this, then by chaining the conditions we have that $A = g^{(i)}(h')^{(i)} \,(\mathrm{mod}\ y^{2^i})$ and so the jump step will discover $A$ if we run it with $g^{(t)}$ for sufficiently large $t$.

So let us show the claim from the above para: Suppose $((g')^{(i)}, (h')^{(i)})$ form a lift of $(g^{(i-1)}, (h')^{(i-1)})$ modulo $I^2$, where $I = (y^{2^{i-1}})$. Then we have that $((g')^{(i)}, B \cdot (h')^{(i)})$ form a lift of $(g^{(i-1)}, B \cdot (h')^{(i-1)})$ modulo $I^2$, and so does $(g^{(i)}, h^{(i)})$. By the equivalence above it follows that $(g')^{(i)} = g^{(i)} \cdot (1+u)$ for some $u \in I$. By moving the multiplication by $1+u$ It follows that $((g)^{(i)}, (1+u) \cdot B \cdot (h')^{(i)})$ form a lift of $(g^{(i-1)}, B \cdot (h')^{(i-1)})$ modulo $I^2$ and this yields the desired claim.

**Proof** [(of Lemma 4)] We already proved existence of a lift in the proof of Lemma 2. We only do the uniqueness now. Let $g_1 = g_2 + \alpha$ and $h_1 = h_2 + \beta$ for $\alpha, \beta \in I$. Since $g_1 h_1 = g_2 h_2 \,(\mathrm{mod}\ I^2)$ we have

$$\alpha h_2 + \beta g_2 = 0 \,(\mathrm{mod}\ I^2).$$

We'd like to say $\alpha = g_2 u$ for some $u \in I$ and we would be able to say this if we could mutliply by $h_2^{-1}$ but we don't quite have an inverse for $h_2$. But we have something close — we have that $b_2 h_2 = 1 - a_2 g_2 \,(\mathrm{mod}\ I^2)$ and this is good enough. So we mutliply both sides of the displayed equation by $b_2$ to get

$$\alpha(1 - a_2 g_2) + b_2 \beta g_2 = 0 \,(\mathrm{mod}\ I^2).$$

Rearrainging above we have

$$\alpha = g_2(a_2\alpha - b_2\beta) = g_2 u,$$

if we let $u = a_2\alpha - b_2\beta$. Similar reasoning shows $\beta = -h_2 u$ and this proves the lemma.

■