LECTURE 7: POLYNOMIAL CONGRUENCES TO PRIME POWER MODULI

1. Hensel Lemma for nonsingular solutions

Although there is no analogue of Lagrange's Theorem for prime power moduli, there is an algorithm for determining when a solution modulo p generates solutions to higher power moduli. The motivation comes from Newton's method for approximating roots over the real numbers.

Suppose that $x = a$ is a solution of the polynomial congruence

$$f(x) \equiv 0 \pmod{p^j},$$

and we want to use it to get a solution modulo p^{j+1}. Th idea is to search for solutions of the form $x = a + tp^j$. The Taylor expansion gives

$$f(a + tp^j) = f(a) + tp^j f'(a) + t^2 p^{2j} f''(a)/2! + \cdots + t^n p^{nj} f^{(n)}(a)/n!,$$

where n is the degree of f. Despite the presence of reciprocals of factorials, the coefficients in the above Taylor expansion are necessarily integral. Indeed, if $f(x) = x^m$ then $f^{(k)}(a)/k! = \binom{m}{k} a^{m-k} \in \mathbb{Z}$, and it follows for general f by linearity. Hence,

$$f(a + tp^j) = f(a) + tp^j f'(a) \pmod{p^{j+1}}.$$

Since $p^j | f(a)$, the congruence $f(a + tp^j) \equiv 0 \pmod{p^{j+1}}$ is equivalent to

$$tf'(a) \equiv -\frac{f(a)}{p^j} \pmod{p}.$$

This congruences have either zero, one, or p solutions. In the case when $f'(a) \not\equiv 0 \pmod{p}$, it has exactly one solution. We conclude:

Theorem 1.1 (Hensel Lemma)**.** *Let $f \in \mathbb{Z}[x]$. Suppose that*

$$f(a) \equiv 0 \pmod{p^j} \quad \text{and} \quad f'(a) \not\equiv 0 \pmod{p}.$$

Then there exists a unique $t \pmod{p}$ such that

$$f(a + tp^j) \equiv 0 \pmod{p^{j+1}}.$$

Hensel's lemma implies that every a solution x_j of $f(x) \equiv 0 \pmod{p^j}$ satisfying $f'(x_j) \not\equiv 0 \pmod{p}$ lifts to a unique solution x_{j+1} of $f(x) \equiv 0 \pmod{p^{j+1}}$ such that $x_{j+1} \equiv x_j \pmod{p^j}$. This solution could be computed using the recursive formula:

$$x_{j+1} = x_j - f(x_j) f'(x_j)^{-1} \pmod{p^{j+1}},$$

where $f'(x_j)^{-1}$ denotes the multiplicative inverse of $f'(x_j)$ modulo p.

Example 1.2. Solve the congruence $x^3 + x + 4 \equiv 0 \pmod{7^3}$.

(I) We first solve the corresponding congruence modulo 7, since any solution x modulo 7^3 must also satisfy $x^3 + x + 4 \equiv 0 \pmod 7$. By an exhaustive search (try $x = 0, \pm 1, \pm 2, \pm 3$), we find that the only solution is $x \equiv 2 \pmod 7$.
(II) Next, we try to solve the corresponding congruence modulo 7^2, since any solution x modulo 7^3 must also satisfy $x^3+x+4 \equiv 0 \pmod{7^2}$. But such solutions must also satisfy the corresponding solution modulo 7, so $x \equiv 2 \pmod 7$. Then we put $x = 2 + 7y$ and substitute. We need to solve

$$(2 + 7y)^3 + (2 + 7y) + 4 \equiv 0 \pmod{7^2}.$$

Notice that when we use the Binomial Theorem to expand the cube, any terms involving 7^2 or 7^3 can be ignored. Thus we need to solve

$$(2^3 + 3 \cdot 2^2 \cdot 7y) + (2 + 7y) + 4 = 14 + 13 \cdot 7y \equiv 0 \pmod{7^2},$$

or equivalently,

$$13y + 2 \equiv -y + 2 \equiv 0 \pmod 7.$$

Then we put $y = 2$ and find that $x = 2 + 7y = 16$ satisfies the congruence $x^3 + x + 4 \equiv 0 \pmod{7^2}$.
(III) We can now repeat the previous strategy (and in fact, we can repeat this as many times as necessary). So we substitute $x = 16 + 7^2 z$ and solve for z to obtain a solution modulo 7^3. Thus we need to solve

$$(16+7^2z)^3 + (16+7^2z) + 4 \equiv (16^3 + 3 \cdot 16^2 \cdot 7^2 z) + (16+7^2z) + 4 \equiv 0 \pmod{7^3}.$$

But $16^3 + 16 + 4$ is divisible by 7^2 (why do we know this?), and in fact is equal to $84 \cdot 7^2$. Then we need to solve

$$84 \cdot 7^2 + (3 \cdot 16^2 + 1) \cdot 7^2 z \equiv 0 \pmod{7^3},$$

which is equivalent to

$$(3 \cdot 16^2 + 1)z + 84 \equiv 0 \pmod 7,$$

or $13z \equiv 0 \pmod 7$. So we put $z = 0$, and find that $x \equiv 16 \pmod{7^3}$ solves $x^3 + x + 4 \equiv 0 \pmod{7^3}$.

Example 1.3. Let $f(x) = x^2 + 1$. Find the solutions of the congruence $f(x) \equiv 0 \pmod{5^4}$.

Observe that the congruence $x^2 + 1 \equiv 0 \pmod 5$ has the solutions $x \equiv \pm 2 \pmod 5$ (note that there are at most 2 solutions modulo 5, by Lagrange's theorem). Consider first the solution $x_1 = 2$ of the latter congruence. One finds that $f'(x_1) = 2x_1 \equiv -1 \pmod 5$. It follows that $5 \nmid f'(x_1)$, and since $f(x_1) = 5 \equiv 0 \pmod 5$, we may apply Hensel's iteration to find integers x_n $(n \geqslant 1)$ with $f(x_n) \equiv 0 \pmod{5^n}$. We obtain

$$\begin{aligned}
x_2 &\equiv x_1 - \frac{f(x_1)}{f'(x_1)} \equiv 2 - \frac{5}{-1} \equiv 7 \pmod{5^2}, \\
x_3 &\equiv 7 - \frac{50}{14} \equiv 7 - \frac{50}{-1} \equiv 57 \pmod{5^3} \\
x_4 &\equiv 57 - \frac{3250}{114} \equiv 57 - \frac{3250}{-1} \equiv 3307 \equiv 182 \pmod{5^4}.
\end{aligned}$$

Thus $x = 182$ provides a solution of the congruence $x^2+1 \equiv 0 \pmod{5^4}$. Proceeding similarly, one may lift the alternate solution $x = -2$ to the congruence $x^2+1 \equiv 0 \pmod 5$ to obtain the solution $x \equiv -182 \pmod{5^4}$. Note that in each instance, the lifting process provided by Hensel's lemma led to a unique residue modulo 5^4 corresponding to each starting solution modulo 5.

2. Hensel Lemma in general

Now we consider the problem of lifting solutions when $f'(a) \equiv 0 \pmod p$.

Example 2.1. Let $f(x) = x^2 - 4x + 13$. Find all of the solutions of the congruence $f(x) \equiv 0 \pmod{3^4}$.

Notice that

$$x^2 - 4x + 13 \equiv x^2 + 2x + 1 \equiv (x+1)^2 \pmod 3,$$

and hence $x \equiv -1 \pmod 3$ is the only solution of the congruence $f(x) \equiv 0 \pmod 3$. Next, since $f'(x) = 2x - 4$, we find that $3|f'(-1)$, We proceed systematically:
(i) Observe first that all solutions satisfy $x \equiv 2 \pmod 3$, and so any solution x must satisfy $x \equiv 2$, 5 or 8 modulo 9. One may verify that all three residue classes satisfy $f(x) \equiv 0 \pmod 9$.
(ii) Next we consider all residues modulo 27 satisfying $x \equiv 2$, 5 or 8 modulo 9, and find that none of these (there are 9 such residues) provide solutions of $f(x) \equiv 0 \pmod{27}$.

So there are no solutions to the congruence $x^2 - 4x + 13 \equiv 0 \pmod{3^3}$.

This example shows that solutions modulo p in general may not lift to solutions modulo some higher powers of p, but not necessarily to solutions modulo arbitrarily high powers of p. Moreover, lifts of the solutions are not unique.

Theorem 2.2. *Let $f \in \mathbb{Z}[x]$. Suppose that*

$$f(a) \equiv 0 \pmod{p^j} \quad \textit{and} \quad p^\tau \parallel f'(a).^1$$

Then if $j \geqslant 2\tau + 1$, whenever $b \equiv a \pmod{p^{j-\tau}}$, one has

$$f(b) \equiv f(a) \pmod{p^j} \quad \textit{and} \quad p^\tau \parallel f'(b).$$

Proof. Writing $b = a + hp^{j-\tau}$ and applying Taylor's expansion, we obtain

$$f(b) = f(a + hp^{j-\tau}) = f(a) + hp^{j-\tau} f'(a) + \frac{1}{2!} f''(a)(hp^{j-\tau})^2 + \dots$$

The quadratic and higher terms in the above expansion are all divisible by $p^{2(j-\tau)}$. But $j \geqslant 2\tau + 1$, whence $2(j-\tau) = j + (j - 2\tau) \geqslant j + 1$, and so

$$f(b) \equiv f(a) + hp^{j-\tau} f'(a) \pmod{p^j}.$$

Since $p^\tau \mid f'(a)$, the latter shows that $f(b) \equiv f(a) \pmod{p^j}$.

[1] Recall that $p^i \parallel A$ means that $p^i | A$ and $p^{i+1} \nmid A$.

Applying Taylor's theorem in like manner to f' one finds that

$$\begin{aligned} f'(b) = f'(a + hp^{j-\tau}) &\equiv f'(a) \pmod{p^{j-\tau}} \\ &\equiv f'(a) \pmod{p^{\tau+1}}, \end{aligned}$$

since $j - \tau \geqslant \tau + 1$. Then since $p^\tau \parallel f'(a)$, one obtains $p^\tau \parallel f'(b)$. □

A good news is that a solution $f(x) \equiv 0 \pmod{p^j}$ gives rise to a solution $f(x) \equiv 0 \pmod{p^{j+1}}$ provided that j is sufficiently large.

Theorem 2.3 (Hensel Lemma)**.** *Let $f \in \mathbb{Z}[x]$. Suppose that*

$$f(a) \equiv 0 \pmod{p^j} \quad \text{and} \quad p^\tau \parallel f'(a).$$

Then if $j \geqslant 2\tau + 1$, there is a unique residue $t \pmod{p}$ such that

$$f(a + tp^{j-\tau}) \equiv 0 \pmod{p^{j+1}}.$$

Proof. Since $p^\tau \parallel f'(a)$, we may write $f'(a) = gp^\tau$ for a suitable integer g with $(g, p) = 1$. Let $\overline{g}$ be any integer with $g\overline{g} \equiv 1 \pmod{p}$, and write

$$a' = a - \overline{g}f(a)p^{-\tau}.$$

Then an application of Taylor's theorem on this occasion supplies the congruence

$$f(a') = f(a - \overline{g}f(a)p^{-\tau}) \equiv f(a) - p^{-\tau}f(a)\overline{g}f'(a) \pmod{p^{2(j-\tau)}},$$

since $j > \tau$ and $p^{-\tau}\overline{g}f(a) \equiv 0 \pmod{p^{j-\tau}}$. But $2(j-\tau) = j + (j - 2\tau) \geqslant j + 1$, and thus

$$f(a') \equiv f(a) - (p^{-\tau}f(a)\overline{g})(gp^\tau) = f(a)(1 - g\overline{g}) \equiv 0 \pmod{p^{j+1}}.$$

So there exists an integer t with $f(a + tp^{j-\tau}) \equiv 0 \pmod{p^{j+1}}$, and indeed one may take $t \equiv -p^{-j}f(a)(p^{-\tau}f'(a))^{-1} \pmod{p}$.

In order to establish the uniqueness of the integer t, suppose, if possible, that two such integers t_1 and t_2 exist. Then one has

$$f(a + t_1p^{j-\tau}) \equiv 0 \equiv f(a + t_2p^{j-\tau}) \pmod{p^{j+1}},$$

whence by Taylor's theorem, as above, one obtains

$$f(a) + t_1p^{j-\tau}f'(a) \equiv f(a) + t_2p^{j-\tau}f'(a) \pmod{p^{j+1}}.$$

Thus $t_1f'(a) \equiv t_2f'(a) \pmod{p^{\tau+1}}$. Since $p^\tau \parallel f'(a)$, we obtain $t_1 \equiv t_2 \pmod{p}$. This establishes the uniqueness of t modulo p, completing our proof. □

Example 2.4. Consider the polynomial $f(x) = x^2 + x + 223$. We observe that $f(4) = 3^5$ and $f'(4) = 3^2$. So $f(4) \equiv 0 \pmod{3^5}$. Searching for solutions of $f(x) \equiv 0 \pmod{3^6}$ of the form $4 + 27t$, we find that

$$f(4 + 27t) \equiv 3^5 + 3^5t \pmod{3^6},$$

and unique $t = 2$ gives such a solution $f(58) \equiv 0 \pmod{3^6}$. Moreover, for any $t = 0, 1, \ldots 8$,

$$f(58 + 81t) \equiv 0 \pmod{3^6}.$$

Some concluding observations may be of assistance:

(i) Hensel's lemma allows one to lift repeatedly. Thus, whenever

$$f(a) \equiv 0 \pmod{p^j} \text{ and } p^\tau \parallel f'(a) \text{ with } j \geqslant 2\tau + 1$$

then there exists a unique residue t modulo p such that, with $a' = a + tp^{j-\tau}$,

$$f(a') \equiv 0 \pmod{p^{j+1}} \text{ and } p^\tau \parallel f'(a') \text{ with } j + 1 \geqslant 2\tau + 1,$$

and then we are set up to repeat this process.

(ii) Notice that in Hensel's lemma, the residue t modulo p is unique, and given by

$$t \equiv -(p^{-j} f(a))(p^{-\tau} f'(a))^{-1} \pmod{p},$$

so one only needs to compute $(p^{-\tau} f'(a))^{-1}$ modulo p. Moreover,

$$p^{-\tau} f'(a') \equiv p^{-\tau} f'(a) \pmod{p},$$

so our initial inverse computation remains valid for subsequent lifting processes.

(iii) If $f(a) \equiv 0 \pmod{p^j}$ and $p^\tau \parallel f'(a)$ and $j \geqslant 2\tau + 1$, then

$$f(a + hp^{j-\tau}) \equiv f(a) \equiv 0 \pmod{p^j}.$$

So there are p^τ solutions of $f(x) \equiv 0 \pmod{p^j}$ corresponding to the single solution $x \equiv a \pmod{p^j}$, namely $a + hp^{j-\tau}$ with $0 \leqslant h \leqslant p^\tau$.