

MATH 361: NUMBER THEORY — SIXTH LECTURE

Let d be a positive integer. Consider a polynomial in d variables with integer coefficients,

$$f \in \mathbf{Z}[X_1, \dots, X_d] \stackrel{\text{call}}{=} \mathbf{Z}[\vec{X}].$$

Consider also a succession of conditions, each stronger than the next:

- (A) *The equation $f(\vec{X}) = 0$ has solutions in \mathbf{Z}^d .*
- (B) *For all $m \in \mathbf{Z}^+$, the congruence $f(\vec{X}) = 0 \pmod{m}$ has solutions.*
- (C) *For all $p \in \mathcal{P}$ and $n \in \mathbf{Z}^+$, the congruence $f(\vec{X}) = 0 \pmod{p^n}$ has solutions.*
- (D) *For each $p \in \mathcal{P}$ there exists some $n \in \mathbf{Z}^+$ such that the congruence $f(\vec{X}) = 0 \pmod{p^n}$ has solutions.*

Thus we have the three implications

$$(A) \implies (B) \implies (C) \implies (D),$$

and we naturally wonder about their converses. The converse implication $(C) \implies (B)$ follows from the Sun Ze Theorem. This lecture discusses the converse implication $(D) \implies (C)$. The main result is called *Hensel's Lemma*.

1. HENSEL'S LEMMA

Recall the *Newton–Raphson method* of finding roots by sliding along tangents: *Given a suitably smooth function $f(x)$, and given an initial guess x_1 , iterate*

$$x_{n+1} = x_n - f(x_n)/f'(x_n).$$

If x_1 is close enough to a root x of f such that $f'(x) \neq 0$, then the iteration converges to x .

Hensel's Lemma is closely analogous to the Newton–Raphson method. Fix a prime p , and work now with one variable rather than the d variables above. (With d variables we may always freeze all but one of them.) The idea is that

Small means congruent to zero modulo a high power of p .

Thus:

- To say that $f(x)$ is small is to say that $f(x) = 0 \pmod{p^n}$ for some suitable n .
- To say that $f'(x)$ is not so small is to say that $f'(x) \neq 0 \pmod{p^{k+1}}$ for some suitable k .
- Given such x , n , and k , we would like to find some y close to x so that $f(y)$ is smaller than $f(x)$ but $f'(y)$ is no smaller than $f'(x)$. To say that y is close to x is to say that $y = x \pmod{p^m}$ for some suitable m .
- The idea is to generate y from x by essentially the Newton–Raphson method.

Theorem 1.1 (Hensel's Lemma). *Let $f \in \mathbf{Z}[X]$ be a polynomial with integer coefficients. Suppose that we have $k, n \in \mathbf{Z}$ with $0 \leq 2k < n$ and $x \in \mathbf{Z}$ such that*

$$\left\{ \begin{array}{l} f(x) = 0 \pmod{p^n} \\ f'(x) = 0 \pmod{p^k} \\ f'(x) \not\equiv 0 \pmod{p^{k+1}} \end{array} \right\}.$$

Then there exists $y \in \mathbf{Z}$ such that

$$\left\{ \begin{array}{l} y = x \pmod{p^{n-k}} \\ f(y) = 0 \pmod{p^{n+1}} \\ f'(y) = 0 \pmod{p^k} \\ f'(y) \not\equiv 0 \pmod{p^{k+1}} \end{array} \right\}.$$

Before the proof, it deserves mention that the easiest and most common case is the case $k = 0$. In this case, if we have $x \in \mathbf{Z}$ and $n \in \mathbf{Z}^+$ such that

$$f(x) = 0 \pmod{p^n}, \quad f'(x) \not\equiv 0 \pmod{p}$$

then we get y such that

$$y = x \pmod{p^n}, \quad f(y) = 0 \pmod{p^{n+1}}, \quad f'(y) \not\equiv 0 \pmod{p}.$$

Proof. Provisionally define

$$y = x + zp^{n-k}, \quad z \text{ to be determined.}$$

Then $y = x \pmod{p^{n-k}}$ independently of z , and the first of the four desired conditions is established.

By Taylor's Theorem for polynomials,

$$f(y) = f(x) + f'(x)zp^{n-k} \pmod{p^{2n-2k}},$$

and so, because $2n - 2k \geq 2n - (n - 1) = n + 1$, it follows that

$$f(y) = f(x) + f'(x)zp^{n-k} \pmod{p^{n+1}}.$$

But $f(x) = 0 \pmod{p^n}$ and $f'(x) \not\equiv 0 \pmod{p^{k+1}}$, so that the previous display gives

$$\begin{aligned} f(y) &= bp^n + ap^k zp^{n-k} \pmod{p^{n+1}} \\ &= (az + b)p^n \pmod{p^{n+1}} \quad \text{where } a \not\equiv 0 \pmod{p}. \end{aligned}$$

Thus setting $z = -a^{-1}b \pmod{p}$ gives $f(y) = 0 \pmod{p^{n+1}}$, and the second of the four desired conditions is established. Note that finding z required only solving a congruence modulo p , independently of k and n .

Again by Taylor's Theorem for polynomials,

$$f'(y) = f'(x) + f''(x)zp^{n-k} \pmod{p^{2n-2k}},$$

and so, because $2n - 2k > n - k \geq 2k + 1 - k = k + 1$, it follows that

$$f'(y) = f'(x) \pmod{p^{k+1}}.$$

Thus $f'(y) = f'(x) = 0 \pmod{p^k}$ and $f'(y) = f'(x) \not\equiv 0 \pmod{p^{k+1}}$, and the third and fourth desired conditions are established. \square

With Hensel's Lemma proved, we return to the analogy between it and the Newton–Raphson method. The proof of Hensel's Lemma took x and found a corresponding y such that

$$f(x) + (y - x)f'(x) = 0 \quad \text{in } \mathbf{Z}/p^{n+1}\mathbf{Z}.$$

Meanwhile, the Newton–Raphson formula for the next iterate $y = x_{n+1}$ in terms of the current iterate $x = x_n$ is

$$y = x - f(x)/f'(x),$$

or, almost identically to the formula from proving Hensel's Lemma,

$$f(x) + (y - x)f'(x) = 0 \quad \text{in } \mathbf{R}.$$

The algebra of the two methods is very similar, but it is not quite identical. On the one hand, we can in some sense better quantify the difference $f(y) - f(x) - f'(x)(y - x)$ in the number-theoretic context than over the real number system because we know that it vanishes up to a certain power of p . On the other hand, we can divide by $f'(x)$ in the real number system but not in the integers, because \mathbf{R} is a field while \mathbf{Z} is only a ring.

As mentioned earlier, usually we start with $n = 1$ and $k = 0$ in Hensel's Lemma, i.e., usually we start with some $x \in \mathbf{Z}$ such that $f(x) \equiv 0 \pmod{p}$ and $f'(x) \not\equiv 0 \pmod{p}$. The repeatedly applying Hensel's Lemma gives a sequence $\{x_1, x_2, x_3, \dots\}$ in \mathbf{Z} such that

$$\left\{ \begin{array}{l} x_1 = x \\ f(x_n) \equiv 0 \pmod{p^n} \quad \text{for all } n \in \mathbf{Z}^+ \\ x_{n+1} \equiv x_n \pmod{p^n} \quad \text{for all } n \in \mathbf{Z}^+ \end{array} \right\}$$

For example, if we let $f(X) = X^2 + 1$ and take $p = 5$ and $x = 2$ then the sequence is

$$\{2, 7, 57, 182, 2057, 14557, 45807, 280182, 280182 \text{ (yes, again), } 6139557, \dots\}$$

To our eyes the sequence may not appear to be converging, but it *is* converging in the sense that

$$\begin{aligned} x_n &\equiv x_m \pmod{5} \quad \text{and} \quad x_n^2 \equiv y_n^2 \equiv -1 \pmod{5} \quad \text{for all } n, m \geq 1, \\ x_n &\equiv x_m \pmod{5^2} \quad \text{and} \quad x_n^2 \equiv y_n^2 \equiv -1 \pmod{5^2} \quad \text{for all } n, m \geq 2, \\ x_n &\equiv x_m \pmod{5^3} \quad \text{and} \quad x_n^2 \equiv y_n^2 \equiv -1 \pmod{5^3} \quad \text{for all } n, m \geq 3, \end{aligned}$$

and so on. The sequence is **5-adically Cauchy**. However, the integers \mathbf{Z} are not complete with respect to 5-adic convergence. The obvious remedy is to complete them. Thus

Definition 1.2. *The p -adic integers \mathbf{Z}_p are the completion of the integers with respect to p -adic convergence.*

The p -adic integers form a ring that is similar to \mathbf{Z} in some regards and very different from \mathbf{Z} in others. The sequence $\{2, 7, 57, \dots\}$ from above converges to a square root of -1 in \mathbf{Z}_5 . The only prime of \mathbf{Z}_p is p . All \mathbf{Z}_p -sided triangles are isosceles. And so on.

The p -adic integers also have an algebraic construction as a *limit*,

$$\mathbf{Z}_p = \lim_n \mathbf{Z}/p^n\mathbf{Z} = \lim (\dots \longrightarrow \mathbf{Z}/p^3\mathbf{Z} \longrightarrow \mathbf{Z}/p^2\mathbf{Z} \longrightarrow \mathbf{Z}/p\mathbf{Z}).$$

Many texts on the p -adic numbers are extant. See for example the book by Koblitz. The first chapter of **Number Theory** by Borevich and Shafarevich proves the following result.

Theorem 1.3 (Hasse–Minkowski Principle). *Consider a quadratic form with rational coefficients,*

$$f(X_1, \dots, X_d) = \sum_{i < j} a_{ij} X_i X_j.$$

Then f has a nonzero root in \mathbf{Z}^d if and only if f has a nonzero root in \mathbf{Q}_p^d for each prime p and f has a nonzero root in \mathbf{R}^d .

The field \mathbf{Q}_p in the theorem is similar to the ring \mathbf{Z}_p except that it has been augmented by denominators. The virtue of the principle is that each \mathbf{Q}_p and \mathbf{R} is a complete field where it suffices to find an approximate solution and then iterate—using Hensel’s Lemma in \mathbf{Q}_p and the Newton–Raphson method in \mathbf{R} .

The word *quadratic* in the theorem is crucial. Selmer showed that the equation

$$3X^3 + 4Y^3 + 5Z^3 = 0$$

has nonzero solutions in each \mathbf{Q}_p and in \mathbf{R} , but not in \mathbf{Q} .