MATH 361: NUMBER THEORY — SIXTH LECTURE

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

$$f \in \mathbb{Z}[X_1, \ldots, X_d] \overset{\text{call}}{=} \mathbb{Z}[X].$$

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

(A) *The equation $f(X) = 0$ has solutions in $\mathbb{Z}^d$.*

(B) *For all $m \in \mathbb{Z}^+$, the congruence $f(X) = 0 \bmod m$ has solutions.*

(C) *For all $p \in \mathcal{P}$ and $n \in \mathbb{Z}^+$, the congruence $f(X) = 0 \bmod p^n$ has solutions.*

(D) *For each $p \in \mathcal{P}$ there exists some $n \in \mathbb{Z}^+$ such that the congruence $f(X) = 0 \bmod p^n$ has solutions.*

Thus we have the three implications

$$\text{(A)} \implies \text{(B)} \implies \text{(C)} \implies \text{(D)},$$

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 $f(x) = 0 \bmod p^n$ for some suitable n.
- To say that $f'(x)$ is not so small is to say that $f'(x) \neq 0 \bmod p^{k+1}$ 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 $y = x \bmod p^m$ for some suitable m.
- We generate y from x by essentially the Newton–Raphson method.

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

$$\left\{\begin{array}{l} f(x) = 0 \bmod p^n \\ f'(x) = 0 \bmod p^k \\ f'(x) \neq 0 \bmod p^{k+1} \end{array}\right\}.$$

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

$$\left\{\begin{array}{r} y = x \bmod p^{n-k} \\ f(y) = 0 \bmod p^{n+1} \\ f'(y) = 0 \bmod p^k \\ f'(y) \neq 0 \bmod p^{k+1} \end{array}\right\}.$$

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

$$f(x) = 0 \bmod p^n, \qquad f'(x) \neq 0 \bmod p$$

then we get y such that

$$y = x \bmod p^n, \qquad f(y) = 0 \bmod p^{n+1}, \qquad f'(y) \neq 0 \bmod p.$$

The second most common case, $k = 1$ and $n \geq 3$, arises naturally for $p = 2$ when f is quadratic.

A second remark before the proof is that for any integer-coefficient polynomial $\varphi[X] \in \mathbb{Z}[X]$ and any integer a, we have (exercise, in which it suffices by linearity to take $\varphi(X) = X^m$, so that $\varphi^{(n)}(a)/n! = \binom{m}{n} a^{m-n}$)

$$\varphi(a + H) = \sum_{n=0}^{\deg \varphi} \frac{\varphi^{(n)}(a)}{n!} H^n \quad \text{in } \mathbb{Z}[H].$$

For quick reference we call this result *Taylor's theorem for polynomials*.

Proof. Provisionally define

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

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

By Taylor's Theorem for polynomials with $\varphi = f$ and $a = x$ and $H = zp^{n-k}$,

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

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

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

But we are given that $f(x) = bp^n$ for some b, and that $f'(x) = ap^k$ for some $a \neq 0 \bmod p$, so the previous display gives

$$f(y) = (az + b)p^n \bmod p^{n+1}, \quad a \neq 0 \bmod p.$$

Thus, setting $z = -a^{-1}b \bmod p$ gives $f(y) = 0 \bmod p^{n+1}$, 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, not modulo a higher power of p.

Again by Taylor's Theorem for polynomials, this time with $\varphi = f'$ and $a = x$ and $H = zp^{n-k}$,

$$f'(y) = f'(x) \bmod p^{n-k},$$

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

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

Thus $f'(y) = f'(x) = 0 \bmod p^k$ and $f'(y) = f'(x) \neq 0 \bmod p^{k+1}$, and the third and fourth desired conditions are established. Incidentally the proof has shown that that the value $a^{-1} \bmod p$ in the congruence $z = -a^{-1}b \bmod p$ that determines $y = x + zp^{n-k}$ from x can be reused in setting the next x to y, iterating n, and repeating the procedure to get the next y, as many times as desired. □

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

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

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,

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

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 in the real number system setting, 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 $\mathbb{R}$ is a field while $\mathbb{Z}$ is only a ring. However, the number theoretic context actually has a certain advantage in this regard. In the Newton–Raphson method, we divide by $f'(x_1)$ to get x_2, then by $f'(x_2)$ to get x_3, and so on. In the number-theoretic context, closer inspection of the proof just given shows that to find x_{n+1} (y in the lemma) given x_n (x in the lemma), the only inverse that we really need is $a^{-1} \bmod p$ where $f'(x_1) = ap^k$. The presence of x_1 rather than x_n in the previous equality means that using Hensel's Lemma to generate a sequence $\{x_n\}$ requires only one inversion modulo p.

As mentioned earlier, usually we start with $n = 1$ and $k = 0$ in Hensel's Lemma, i.e., usually we start with some $x \in \mathbb{Z}$ such that $f(x) = 0 \bmod p$ and $f'(x) \neq 0 \bmod p$. The repeatedly applying Hensel's Lemma gives a sequence $\{x_1, x_2, x_3, \ldots\}$ in $\mathbb{Z}$ such that

$$\left\{ \begin{aligned} &x_1 = x \\ &f(x_n) = 0 \bmod p^n \quad \text{for all } n \in \mathbb{Z}^+ \\ &x_{n+1} = x_n \bmod p^n \quad \text{for all } n \in \mathbb{Z}^+ \end{aligned} \right\}$$

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

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

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

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

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

Definition 1.2. *The ring of* **p-adic integers** $\mathbb{Z}_p$ *is the completion of the ring of integers with respect to p-adic convergence. The field of* **p-adic numbers** $\mathbb{Q}_p$ *is the field of quotients of* $\mathbb{Z}_p$.

The ring of p-adic integers is similar to the usual ring of integers in some regards but very different in others. The sequence $\{2, 7, 57, \dots\}$ from above converges to a square root of -1 in $\mathbb{Z}_5$. The only prime of $\mathbb{Z}_p$ is p. All $\mathbb{Z}_p$-sided triangles are isosceles. The exponential series does not converge everywhere, but the exponential series and the logarithm series do invert each other where they do converge.

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

$$\mathbb{Z}_p = \lim_n \mathbb{Z}/p^n\mathbb{Z} = \lim\left(\cdots \longrightarrow \mathbb{Z}/p^3\mathbb{Z} \longrightarrow \mathbb{Z}/p^2\mathbb{Z} \longrightarrow \mathbb{Z}/p\mathbb{Z}\right).$$

The limit is a mathematical structure (a group and a compact topological space, with the group multiplication and inversion continuous under the topology) that maps to all the quotients $\mathbb{Z}/p^n\mathbb{Z}$ compatibly with how they map to one another.

Many texts on the p-adic numbers exist, e.g., the book by Koblitz. Chapter 1 of **Number Theory** by Borevich and Shafarevich proves the following result.

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

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

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

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

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 $\mathbb{Q}_p$ and in $\mathbb{R}$, but not in $\mathbb{Q}$.

As an exercise with Hensel's Lemma 2-adically, let p be a 1 mod 8 prime, and let q be an odd prime. Consider a function of one variable,

$$f(X) = pX^2 + qy^2 - z^2 \quad \text{where } y = 0 \text{ and } z = 1.$$

That is, $f(X) = pX^2 - 1$. Set $x = 1$, so that

$$f(x) = p - 1 = 0 \bmod 2^3, \qquad f'(x) = 2p = 0 \bmod 2, \qquad f'(x) = 2p \neq 0 \bmod 2^2.$$

Thus Hensel's Lemma with $(n, k) = (3, 1)$ gives a 2-adic root of the three-variable polynomial $F(X, Y, Z) = pX^2 + qY^2 - Z^2$. Continuing in this vein, one can show that for distinct odd primes p and q, the equation

$$pX^2 + qY^2 = Z^2$$

has a nonzero solution in $\mathbb{Z}_2^3$ if at least one of p and q is 1 mod 4. More elementary considerations, using the surjection $\mathbb{Z}_p \longrightarrow \mathbb{Z}/p^e\mathbb{Z}$ with $p = 2$ and a well-chosen e, show that it doesn't have a nonzero solution in $\mathbb{Z}_2^3$ if p and q are 3 mod 4. Soon we will encounter the same condition on p and q,

Yes if at least one of p and q is 1 *mod* 4*, No if both are* 3 *mod* 4*,*

in a context that seems entirely different. The connection is explained in a later writeup.