TIGHTENING THE BASIC VERSION OF HENSEL'S LEMMA

KEITH CONRAD

In its simplest form, Hensel's Lemma says that a polynomial $f(X) \in \mathbf{Z}_p[X]$ with a simple root mod p has a lifting to a p-adic root:

$$f(a_0) \equiv 0 \bmod p, f'(a_0) \not\equiv 0 \bmod p \Rightarrow \text{ there is } a \in \mathbf{Z}_p \text{ such that } a \equiv a_0 \bmod p, f(a) = 0.$$

The hypotheses can also be written using absolute values: $|f(a_0)|_p < 1$, $|f'(a_0)|_p = 1$. And the conclusion is $f(a) = 0$ with $|a - a_0|_p < 1$.

As a simple application, consider $f(X) = X^p - X$. For each integer j between 0 and $p-1$, $f(j) \equiv 0 \bmod p$ and $f'(j) = pj^{p-1} - 1 \equiv -1 \not\equiv 0 \bmod p$. So there is some $\omega_j \in \mathbf{Z}_p$ such that $\omega_j^p = \omega_j$ and $\omega_j \equiv j \bmod p\mathbf{Z}_p$. This accounts for p different roots (they are incongruent mod $p\mathbf{Z}_p$, hence unequal), so we've found all the roots: $X^p - X$ splits completely in $\mathbf{Z}_p[X]$. Its nonzero roots are $(p-1)$-th roots of unity.

Let's consider the following more general setting for Hensel's Lemma: K is a field complete with respect to a nonarchimedean absolute value $|\cdot|$, $\mathfrak{o}$ is its valuation ring.

The version of Hensel's Lemma in [3, Chap. II, §2, Prop. 2] goes as follows: if $a_0 \in \mathfrak{o}$ and

$$|f(a_0)| < |f'(a_0)|^2, \tag{1}$$

then the sequence

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$$

converges in $\mathfrak{o}$, and its limit a satisfies

$$f(a) = 0, \quad |a - a_0| \leq \left|\frac{f(a_0)}{f'(a_0)^2}\right|.$$

The need for having a Hensel's Lemma like this, where we start with a possible nonsimple root at the level of the residue field (*i.e.*, with $|f(a_0)| < 1$ *and* $|f'(a_0)| < 1$), is seen when trying to determine which elements of $\mathbf{Q}_2$ are squares: the derivative of $X^2 - b$ vanishes mod 2, so Hensel's Lemma stated only for simple roots in the residue field is useless. A similar difficulty arises when deciding which elements of $\mathbf{Q}_p$ are p-th powers for any p.

The estimate (1) only makes sense if $f'(a_0) \neq 0$. In the course of the proof in [3], it is seen that $|f'(a)| = |f'(a_0)|$, so the theorem will produce only simple roots of f in $\mathfrak{o}$ (which may not be simple in $\mathfrak{o}/\mathfrak{m}$, *i.e.*, perhaps $|f'(a_0)| < 1$).

We now give a version of Hensel's Lemma which strengthens the above form in three respects:

1) it improves the upper bound on $|a - a_0|$ to an exact formula,

2) it proves uniqueness of the root in a disc around a_0 which is slightly larger than $|a-a_0|$,

3) it provides a converse of sorts, showing that the basic inequality which gets things going is guaranteed to hold in a neighborhood of any simple root.

Theorem 1. *Let $f \in \mathfrak{o}[X]$ and $a_0 \in \mathfrak{o}$ with $|f(a_0)| < |f'(a_0)|^2$. Then the recursion*

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$$

converges in $\mathfrak{o}$ to a root a of f and

1) $|a - a_0| = |f(a_0)/f'(a_0)| < |f'(a_0)| \leq 1$,

2) *a is the unique root of f whose distance from a_0 is $< |f'(a_0)|$.*

Moreover, if f has a simple root r in $\mathfrak{o}$, then for all x such that $|r-x| < |f'(r)|$, we have $|f'(x)| = |f'(r)|$ and $|f(x)| < |f'(x)|^2$.

Two consequences of the theorem are worth noting, concerning the uniqueness aspect:

1) while a is the unique root of f in the open ball around a_0 of radius $|f'(a_0)|$, it is in fact *closer* to a_0 than that distance.

2) for a simple root $r \in \mathfrak{o}$ of f, there are no other roots s of f in $\mathfrak{o}$ satisfying $|s-r| < |f'(r)|$.

Proof. Let $c = |f(a_0)/f'(a_0)^2| < 1$. We inductively prove that

i) $|a_n| \leq 1$, *i.e.*, $a_n \in \mathfrak{o}$,

ii) $|f'(a_n)| = |f'(a_0)|$,

iii) $|f(a_n)| < |f'(a_0)|^2 c^{2^n}$,

iv) $|a_{n+1} - a_n| < |f'(a_0)| c^{2^n}$.

For $n = 0$ they are all clear. Part iv comes from the definition $a_1 = a_0 - f(a_0)/f'(a_0)$.

For the inductive step, we use the following: for any polynomial $P(X) \in \mathfrak{o}[X]$,

$$P(X+Y) = P(X) + P'(X)Y + Q(X,Y)Y^2$$

for some $Q \in \mathfrak{o}[X,Y]$ (by the binomial theorem, $(X+Y)^n = X^n + (nX^{n-1})Y + Y^2 H(X,Y)$, now use $\mathfrak{o}$-linearity). So if $x, y \in \mathfrak{o}$ then $P(x+y) = P(x) + P'(x)y + y^2 z$ where $z \in \mathfrak{o}$.

Assume parts i through iv are true for n. Then $|f(a_n)/f'(a_n)| = |f(a_n)/f'(a_0)| < |f'(a_0)| c^{2^n} \leq 1$, so $|a_{n+1}| \leq 1$.

Next, since $|f(a_n)| < |f'(a_0)|^2$,

$$|f'(a_{n+1}) - f'(a_n)| \leq |a_{n+1} - a_n| = \frac{|f(a_n)|}{|f'(a_n)|} = \frac{|f(a_n)|}{|f'(a_0)|} < |f'(a_0)|,$$

so $|f'(a_{n+1})| = |f'(a_0)|$.

To estimate $|f(a_{n+1})|$,

$$f(a_{n+1}) = f(a_n) + f'(a_n)\left(-\frac{f(a_n)}{f'(a_n)}\right) + \left(\frac{f(a_n)}{f'(a_n)}\right)^2 z = \left(\frac{f(a_n)}{f'(a_n)}\right)^2 z,$$

where $z \in \mathfrak{o}$. Thus

$$|f(a_{n+1})| \leq \left|\frac{f(a_n)}{f'(a_n)}\right|^2 = \frac{|f(a_n)|^2}{|f'(a_0)|^2} < |f(a_0)|^2 c^{2^{n+1}}.$$

Part iv is equivalent to part iii.

Thus, the sequence $\{a_n\}$ is Cauchy. Let a be its limit, so $a \in \mathfrak{o}$ by i. By iii, $f(a) = 0$. By ii, $|f'(a)| = |f'(a_0)|$. For $n \geq 1$,

$$|a_{n+1} - a_n| < |f'(a_0)| c^2 < |f'(a_0)| c = \left|\frac{f(a_0)}{f'(a_0)}\right| = |a_1 - a_0|,$$

so writing $a_{n+1} - a_0 = a_{n+1} - a_n + a_n - a_0$, we get by induction on n that $|a_n - a_0| = |f(a_0)/f'(a_0)|$, so $|a - a_0| = |f(a_0)/f'(a_0)|$.

Now we want to show the uniqueness of the root in the open ball around a_0 of radius $|f'(a_0)|$. Assume $|b - a_0| < |f'(a_0)|$ and $f(b) = 0$. Then also $|b - a| < |f'(a_0)|$.

Let $b = a + h, h \in \mathfrak{o}$. Then

$$0 = f(b) = f(a) + f'(a)h + h^2 z = f'(a)h + h^2 z$$

for $z \in \mathfrak{o}$. If $z \neq 0$ then $f'(a) = -hz$, so $|f'(a)| \leq |h| = |b - a| < |f'(a_0)|$. But $|f'(a)| = |f'(a_0)|$, so we have a contradiction and $h = 0$, *i.e.*, $b = a$.

Now assume r is a simple root of f and $|r-x| < |f'(r)|$. Then $|f'(r) - f'(x)| \le |r-x| < |f'(r)|$, so $|f'(x)| = |f'(r)|$. Also

$$f(x) = f(r) + f'(r)(x-r) + (x-r)^2 z = f'(r)(x-r) + (x-r)^2 z,$$

where $z \in \mathfrak{o}$. Each term on the right hand side has size less than $|f'(r)|^2$, so also $|f(x)| < |f'(r)|^2 = |f'(x)|^2$. □

Let's see the uniqueness aspect of Hensel's Lemma put to use to answer a concrete question: what are the roots of unity in $\mathbf{Q}_p$? Of course, if $x^n = 1$ then $|x|^n = 1$, so $|x| = 1$. This means any root of unity in $\mathbf{Q}_p$ lies in $\mathbf{Z}_p^\times$. Therefore we work in $\mathbf{Z}_p^\times$ right from the start.

First let's consider roots of unity of order prime to p. We've already found one in each residue class, from our factorization of $X^{p-1} - 1$. Now assume that ζ_1 and ζ_2 are roots of unity with order prime to p. So they are both roots of some $f(X) = X^m - 1$, where m is prime to p. Since $|f'(\zeta_i)|_p = 1$, the *uniqueness* in Hensel's Lemma says that the only root x of $X^m - 1$ satisfying $|x - \zeta_1|_p < 1$ is ζ_1 itself. So we can't have $|\zeta_2 - \zeta_1|_p < 1$. That is, roots of unity of order prime to p must be incongruent mod p. Since we've already found one root in each nonzero class mod p, we've found them all: the only roots of unity of order prime to p in $\mathbf{Q}_p$ are the roots of $X^{p-1} - 1$.

Now we consider roots of unity of p-power order. Of course, in $\mathbf{Q}_2$ we have a root of unity of order 2, namely -1. That's all, folks.

Theorem 2. *For p odd, there are no roots of unity of order p in $\mathbf{Q}_p$. There are no roots of unity of order 4 in $\mathbf{Q}_2$.*

Proof. We first consider odd p, showing the only root of $X^p - 1$ in $\mathbf{Q}_p$ is 1.

Suppose $r^p = 1$. From $r^p \equiv 1 \bmod p$ we get $r \equiv 1 \bmod p$. By the uniqueness in Hensel's Lemma for $f(X) = X^p - 1$, the ball

$$\{x \in \mathbf{Q}_p : |x - r|_p < 1/p^2\} = r + p^3\mathbf{Z}_p$$

around r contains no pth root of unity except for r. (Using the pth cyclotomic polynomial instead of $X^p - 1$ would make the bound on $|x - r|_p$ depend on r, which would be inconvenient.) That is, cosets mod $p^3\mathbf{Z}_p$ contain at most one pth root of unity. So we show the only solution mod p^3 to $X^p = 1$ is 1.

Any pth root of unity has the form $1 + py$. Then

$$(1+py)^p = 1 \Rightarrow \sum_{k=1}^{p} \binom{p}{k} p^k y^k = 0,$$

where all terms except for $k = 1$ are divisible by p^3 (p is odd), so $p^2 y$ is divisible by p^3, hence y is divisible by p. But then all terms except for $k = 1$ are divisible by p^4, hence $p^2 y$ is divisible by p^4, so y is divisible by p^2. Of course, we could continue *ad infinitum* to find y is arbitrarily highly divisible by p, so $y = 0$, but once we know y is divisible by p^2 we get $1 + py \equiv 1 \bmod p^3$, and by the uniqueness of Hensel's Lemma the only root of $X^p - 1$ which is $\equiv 1 \bmod p^3\mathbf{Z}_p$ is 1, so we're done.

Now we consider roots of unity of order 4 in $\mathbf{Q}_2$, *i.e.*, roots of $X^2 + 1$. This won't use Hensel's Lemma. Any putative root of $X^2 + 1$ is in $\mathbf{Z}_2^\times = 1 + 2\mathbf{Z}_2$. But $x \equiv 1 \bmod 2 \Rightarrow x^2 \equiv 1 \bmod 8$, so $x^2 + 1 \equiv 2 \bmod 8$. Therefore $x^2 + 1 \neq 0$. □

When there's no root of unity of order p^n in a field, there are none of order p^m for $m \ge n$. Since every root of unity is a (unique) product of a root of unity of p-power order and a root of unity of order prime to p, the only roots of unity in $\mathbf{Q}_p$ are the roots of $X^{p-1} - 1$ for p odd, and ± 1 for $p = 2$.

For a version of Hensel's Lemma allowing $\mathfrak{o}$ to be any complete local ring (possibly not a domain, so division is more delicate), see [2, Theorem 7.3].

For an even more elaborate form, dealing with several polynomials (or power series) in several variables, see [1, Chap. III, §4.3].

References

[1] N. Bourbaki, "Commutative Algebra," Springer-Verlag, New York, 1989.
[2] D. Eisenbud, "Commutative Algebra with a View to Algebraic Geometry," Springer-Verlag, New York, 1995.
[3] S. Lang, "Algebraic Number Theory," 3rd ed., Springer-Verlag, New York, 1994.