p-ADIC NUMBERS, $\mathbb{Q}_p$, AND HENSEL'S LEMMA

YIDUAN ZHENG

Abstract. This paper starts by introducing the concepts of p-adic numbers and p-adic absolute value. It then discusses three different interpretations of the field of p-adic numbers and how they correspond to each other. Finally, it will present two versions of Hensel's Lemma based on knowledge from preceding sections. It will also explain briefly the difference between the versions and provide examples of applications for both.

Contents

1. Introduction

The main goal of this paper is to introduce Hensel's Lemma. Formulated by Kurt Hensel, it predicts the existence of roots to a polynomial in the ring of p-adic integers given an initial approximated solution modulo prime p. For example, a typical question would be "given a polynomial $f(x) = x^2 - u$ and an integer approximation u, does there exist v close to u such that $f(v) = 0$?"

Hensel's Lemma also can be used to answer questions about properties of p-adic numbers, such as when a p-adic number u is a square or a pth power in the field of p-adic numbers. An example will be given after the proofs of the basic version. Also note the proofs of the general version of Hensel's Lemma assume at least some familiarity with real analysis.

To introduce Hensel's Lemma, we will first need to build the notion of the field of p-adic numbers $\mathbb{Q}_p$ and the p-adic absolute value. An explanation of the three interpretations of $\mathbb{Q}_p$ sheds light on the structure of the field and the p-adic integers, which in turn elucidates the mechanisms of Hensel's Lemma.

Date: August 29, 2020.

2. p-adic Absolute Value

Definition 2.1. (p-adic valuation) Let p be a prime number. Define p-adic valuation $\mathrm{ord}_p : \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$,

$$\mathrm{ord}_p(a) = \begin{cases} m, & \text{if } a = p^m \cdot \frac{u}{v} \\ 0, & \text{if } a = 0 \end{cases}$$

where $u, v, m \in \mathbb{Z}$, $u, v \not\equiv 0 \bmod p$.

In other words, p-adic valuation indicates the power of p that divides a. We consider $a, b \in \mathbb{Q}$ to be close in the p-adic sense if $\mathrm{ord}_p(a-b)$ is large in value. In fact, as we will see in the next section, a sequence of rational numbers (x_n) converges to $a \in \mathbb{Q}$ p-adically if $\mathrm{ord}_p(x_n - a) \to \infty$ as $n \to \infty$.

Remark 2.2. (Properties of p-adic valuation) For any $a, b \in \mathbb{Q}$, $a, b \neq 0$, we have
(1) $\mathrm{ord}_p(ab) = \mathrm{ord}_p(a) + \mathrm{ord}_p(b)$,
(2) $\mathrm{ord}_p(a+b) \geq \min(\mathrm{ord}_p(a), \mathrm{ord}_p(b))$,
(3) if $\mathrm{ord}_p(a) \neq \mathrm{ord}_p(b)$, then $\mathrm{ord}_p(a+b) = \min(\mathrm{ord}_p(a), \mathrm{ord}_p(b))$.

Note that if one of a, b does equal 0, properties (1) and (3) do not hold.

Proof. Let $a = p^m \frac{u}{v}$, $b = p^n \frac{w}{z}$. By Definition 2.1, $\mathrm{ord}_p(a) = m$ and $\mathrm{ord}_p(b) = n$.
(1) Then $ab = p^{m+n} \frac{uw}{vz}$, where $uw, vz \not\equiv 0 \bmod p$. Hence

$$\mathrm{ord}_p(ab) = m + n = \mathrm{ord}_p(a) + \mathrm{ord}_p(b).$$

(2) Assume without loss of generality that $m \geq n$, then

$$a + b = p^n \left(\frac{p^{m-n}uz + wv}{vz}\right).$$

By Definition 2.1, the denominator is not divisible by p, but we cannot say the same for the numerator because there is a chance that $m = n$. Therefore,

$$\mathrm{ord}_p(a+b) \geq n = \min(\mathrm{ord}_p(a), \mathrm{ord}_p(b)).$$

(3) If $m \neq n$, then the numerator itself is not divisible by p. The proof here remains largely the same as (2) except now $\mathrm{ord}_p(a+b) = n$. □

Definition 2.3. (p-adic absolute value) For any rational number a, $a \neq 0$, its p-adic absolute value $|a|_p$ is defined by

$$|a|_p = p^{-ord_p(a)}.$$

Hence $|a|_p$ represents the size of a in the p-adic sense. We will present a simple example below.

Example 2.4. Observe that

$$|p|_p = \frac{1}{p} \text{ and } |\frac{1}{p}|_p = p,$$

so as a p-adic number, p is smaller compared to $\frac{1}{p}$.

Also note $|0|_p = 0$ because $\mathrm{ord}_p(0) = \infty$.

Remark 2.5. (Properties of p-adic absolute value) For any $a, b \in \mathbb{Q}$, we have

(1) $|ab|_p = |a|_p \cdot |b|_p$,

(2) $|a+b|_p \leq \max(|a|_p, |b|_p) \leq |a|_p + |b|_p$,

where the first part of the inequality is the Archimedean Property in $\mathbb{Q}_p$, and the latter part forms the Triangle Inequality.

Proof. The two statements come from Remark 2.2.

(1) By Definition 2.3 and an exponent product rule,

$$|ab|_p = p^{-ord_p(ab)} = p^{-ord_p(a)} \cdot p^{-ord_p(b)} = |a|_p \cdot |b|_p.$$

(2) The latter part of the inequality follows naturally, so we only need to prove the former part:

$$|a+b|_p = p^{-ord_p(a+b)} \leq p^{-\min(\mathrm{ord}_p(a), \mathrm{ord}_p(b))} = \max(|a|_p, |b|_p).$$

□

Furthermore, the p-adic absolute value function brings a sense of distance. For instance, one can also interpret Example 2.4 as p being closer to the origin p-adically than $1/p$. Similarly, for any two rational numbers, we can calculate the p-adic distance between them, and hence the p-adic absolute value forms a metric.

Definition 2.6. (p-adic metric) Given a prime number p, the p-adic metric is

$$d_p(a, b) = |a - b|_p$$

and d_p satisfies the following three properties:

- $d_p(a, b) \geq 0$, and $d_p(a, b) = 0$ if and only if $a = b$,
- $d_p(a, b) = d_p(b, a)$,
- $d_p(a, c) \leq d_p(a, b) + d_p(b, c)$.

Note that $\mathbb{Q}$ is a metric space with respect to the p-adic metric, and a sequence (x_n) converges p-adically to a rational number a if and only if $d_p(x_n, a) \to 0$.

In the rest of the paper, we will more often use the absolute value notation instead of the metric one, but readers should bear this concept of distance in mind.

3. p-adic Number Fields $\mathbb{Q}_p$

There are in total three ways one can define the field of p-adic numbers. The first definition can be understood better in the light of the relationship between $\mathbb{R}$ and $\mathbb{Q}$, so we need to examine the latter first.

First notice that in $\mathbb{R}$, a sequence of rational numbers (x_n) may converge to a number $a \notin \mathbb{Q}$. For example, the sequence

$$3, 3.1, 3.14, 3.141, 3.1415...$$

converges to π. This sequence would diverge if we take the domain to be $\mathbb{Q}$. Hence, $\mathbb{R}$ is simply an extension or a completion of $\mathbb{Q}$ in which all the Cauchy sequences converge with respect to the Archimedean absolute value. The analogy here is that for a fixed p, we can regard $\mathbb{Q}_p$ as a completion of of $\mathbb{Q}$ in which all the p-adic Cauchy sequences converge.

Definition 3.1. (p-adic Cauchy sequence) Given a sequence (x_n) with $x_n \in \mathbb{Q}$ for all n, we call (x_n) a p-adic Cauchy sequence if for any $\epsilon \in \mathbb{Q}$, $\epsilon > 0$, there exists a natural number N such that $m, n \geq N$ implies $d_p(x_m, x_n) = |x_m - x_n|_p < \epsilon$.

Therefore, we can create an injective map from $\mathbb{Q}$ to $\mathbb{Q}_p$ by associating an element $a \in \mathbb{Q}$ with an element $b \in \mathbb{Q}_p$ given by the same p-adic Cauchy sequence. Hence $\mathbb{Q}$ is dense in $\mathbb{Q}_p$.

A second definition of $\mathbb{Q}_p$ involves $\mathbb{Z}_p$, or the p-adic integers.

Definition 3.2. (p-adic integers) Define

$$\mathbb{Z}_p = \{z \in \mathbb{Q}_p : |z|_p \leq 1\} = \{z \in \mathbb{Q}_p : \mathrm{ord}_p(z) \geq 0\}.$$

Note that $\mathbb{Z}_p$ is a commutative subring of $\mathbb{Q}_p$. In fact, it is the unit disc in $\mathbb{Q}_p$.

Here is a related definition that we will see in later sections.

Definition 3.3. Define

$$\mathbb{Z}_p^\times = \{z \in \mathbb{Q}_p : |z|_p = 1\} = \{z \in \mathbb{Q}_p : \mathrm{ord}_p(z) = 0\}.$$

Note that $\mathbb{Z}_p^\times$ is a subset of $\mathbb{Z}_p$.

Definition 3.4. (Inverse limits) Given a sequence of sets X_n and maps $f_n : X_{n+1} \to X_n$, the *inverse limit* is defined by

$$\varprojlim_n X_n = \{(x_n) \in \prod_{n=1} X_n : f_n(x_{n+1}) = x_n\}.$$

If we set X_n to be $\mathbb{Z}/p^n\mathbb{Z}$ and f_n to be the natural projection from $\mathbb{Z}/p^{n+1}\mathbb{Z}$ to $\mathbb{Z}/p^n\mathbb{Z}$, then we obtain $\varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$. An element (x_n) of $\varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$ is a sequence composed of elements such that $x_1 \in \mathbb{Z}/p$, $x_2 \in \mathbb{Z}/p^2$, and so on, which are compatible with each other (i.e. x_n can be reduced to x_{n-1}). The elements of this sequence correspond to the coefficients of the p-adic expansion of a number in $\mathbb{Z}_p$ (see Theorem 3.5 below), so there exists a bijection between $\mathbb{Z}_p$ and $\varprojlim_n \mathbb{Z}/p^n\mathbb{Z}$. As a detailed explanation would involve more advanced algebra, it will not be discussed here. Readers should consult Kato's book [1, p.67-68] for a formal proof of this bijection.

Since $\mathbb{Q}_p$ is the quotient field of $\mathbb{Z}_p$, $\mathbb{Q}_p$ can be seen as the quotient of inverse limits. One can study p-adic integers mod p^n for various n and get $\mathbb{Q}_p$.

The third and final way of defining $\mathbb{Q}_p$ concerns p-adic expansion.

Theorem 3.5. *$\mathbb{Q}_p$ is the set of all numbers that can be written in the form of a p-adic expansion, namely,*

$$\mathbb{Q}_p = \{\sum_{n=m}^{\infty} a_n p^n : m \in \mathbb{Z}, a_n \in \{0, 1, ..., p-1\}\}.$$

To associate a p-adic integer with its p-adic expansion, choose a p-adic integer z and an integer m such that $\mathrm{ord}_p(z) \geq m$. Then $p^{-m}z$ is also a p-adic integer, and $p^{-m}z \equiv a_m \bmod p$, for some $a_m \in \{0, 1, ..., p-1\}$. Therefore, $z - p^m a_m \equiv 0 \bmod p^{m+1}$, which means $\mathrm{ord}_p(z - p^m a_m) \geq m + 1$. Repeating the step above, we will get $\mathrm{ord}_p(z - p^m a_m - p^{m+1}a_{m+1}) \geq m + 2$ for some a_{m+1}. Further repetition gives the unique p-adic expansion. In this way, the three definitions are equivalent.

4. Hensel's Lemma: Basic Version

As we will see in this section, Hensel's Lemma implies that congruence in the p-adic sense means approximation. Given $a \equiv b \bmod p^n$, we have

$$a - b \equiv 0 \bmod p^n,$$

which is equivalent to

$$|a-b|_p \leq \frac{1}{p^n}$$

by Definition 2.1 and 2.3. Again, the larger n is, the closer a and b are.

Theorem 4.1. *(Hensel's Lemma) Let $f(x) \in \mathbb{Z}_p[x]$ and $a \in \mathbb{Z}/p$ for a fixed p. Suppose $f(a) \equiv 0 \mod p$ but $f'(a) \not\equiv 0 \mod p$. Then there exists a unique $z \in \mathbb{Z}_p$ such that*

- *$f(z) = 0$,*
- *$z \equiv a \mod p$, or equivalently, $|z-a|_p \leq 1/p$.*

To prove the theorem, we will need the lemma below.

Lemma 4.2. *Let $f(x) \in F[x]$, where F is a field. Then*

$$f(x+y) = f(x) + f'(x)y + g(x,y)y^2$$

for some polynomial $g(x,y) \in F[x,y]$.

Proof. This polynomial identity follows from the Binomial Theorem and isolating the first two terms. First, we know any polynomial can be written in the form

$$f(x) = a_0 + a_1x + ... + a_nx^n = \sum_{i=0}^{n} a_ix^i$$

for $n < \infty$. Plugging in $x+y$ into the polynomial, we get

$$\begin{aligned} f(x+y) &= \sum_{i=0}^{n} a_i(x+y)^i \\ &= a_0 + a_1x + a_1y + \sum_{i=2}^{n} a_i(x^i + ix^{i-1}y + g_i(x,y)y^2) \end{aligned}$$

where $g_i(x,y)$ equals the sum of all terms of the expansion except for the first two then divided by y^2. By Binomial Theorem, each term after the first two in the expansion has a y^k component, where $k \geq 2$, so $g_i(x,y)$ is a polynomial for each $i \geq 2$. Therefore,

$$\begin{aligned} f(x+y) &= \sum_{i=0}^{n} a_ix^i + \sum_{i=1}^{n} ia_ix^{i-1}y + \sum_{i=2}^{n} g_i(x,y)y^2 \\ &= f(x) + f'(x)y + g(x,y)y^2 \end{aligned}$$

where $g(x,y) \in R[x,y]$. □

Now we can prove the main theorem.

Proof. (Hensel's Lemma) To show the existence of such a root z, we will first use induction to construct a p-adic Cauchy sequence. Then we will show that the limit of this sequence corresponds to the root we are looking for.

More specifically, we need to find a sequence $(a_n) \in \mathbb{Z}_p$ such that

- $f(a_n) \equiv 0 \mod p^n$,
- $a_n \equiv a \mod p$

for all n. The base case is relatively straightforward, since setting $a_1 = a$ yields the desired result. For the inductive step, we assume the nth case is true and we want to find $a_{n+1} \in \mathbb{Z}_p$ such that

- $f(a_{n+1}) \equiv 0 \bmod p^{n+1}$,
- $a_{n+1} \equiv a \bmod p$.

Since the nth case holds true, there exists a root a_n divisible by p^n. For the sequence to be Cauchy, the next term a_{n+1} has to satisfy the condition $a_{n+1} \equiv a_n \bmod p^n$ or $|a_{n+1} - a_n| \leq 1/p^n$, which is equivalent to Definition 3.1. Note that this condition also fulfills the second requirement, since by Remark 2.5,

$$|a_{n+1} - a|_p \leq \max(|a_{n+1} - a_n|_p, |a_n - a|_p) = \frac{1}{p}$$

or $a_{n+1} \equiv a \bmod p$. Write $a_{n+1} = a_n + p^n t_n$ for some $t_n \in \mathbb{Z}_p$. Now we want to show $f(a_n + p^n t_n) \equiv 0 \bmod p^{n+1}$. By Lemma 4.2, we get

$$\begin{aligned} f(a_n + p^n t_n) &= f(a_n) + f'(a_n)p^n t_n + g(a_n, p^n t_n)p^{2n} t_{2n} \\ &\equiv f(a_n) + f'(a_n)p^n t_n \bmod p^{n+1}, \end{aligned}$$

or equivalently,

$$f'(a_n)t_n \equiv -\frac{f(a_n)}{p^n} \bmod p.$$

Note that $-f(a_n)/p^n \in \mathbb{Z}_p$ since $f(a_n) \equiv 0 \bmod p^n$. Also, such $t_n \neq 0$ exists because $f'(a_n) \equiv f'(a) \not\equiv 0 \bmod p$. Thus it is possible to find a_{n+1} such that $f(a_{n+1}) \equiv 0 \bmod p^{n+1}$, and we can complete the rest of the Cauchy sequence in this way.

Next, let the limit of the Cauchy sequence be z. We want to show $f(z) = 0$ and $z \equiv a \bmod p$. Since $a_{n+1} \equiv a_n \bmod p^n$, $a_m \equiv a_n \bmod p^n$, for all $m > n$. As $m \to \infty$, we get $z \equiv a_n \bmod p^n$. As this holds true for all n, we set $n = 1$ and $z \equiv a \bmod p$.

We also note that $z \equiv a_n \bmod p^n$ means that $f(z) \equiv f(a_n) \equiv 0 \bmod p^n$, or $|f(z)|_p \leq 1/p^n$. Thus $|f(z)|_p = 0$ as $n \to \infty$, and $f(z) = 0$.

Next, we need to show the uniqueness of such a root z. Suppose there exists α such that $f(\alpha) = 0$ and $\alpha \equiv z \bmod p$. We will show that the two roots have the same p-adic expansion, that is, $\alpha \equiv z \bmod p^n$ for all n. The $n = 1$ case holds true because both roots are congruent to $a \bmod p$. For $n > 1$, we know that $\alpha \equiv z \bmod p^n$, so let $\alpha = z + p^n s_n$, $s_n \in \mathbb{Z}_p$. Applying Lemma 4.2 again, we get

$$f(\alpha) = f(z + p^n s_n) \equiv f(z) + f'(z)p^n s_n \bmod p^{n+1}.$$

We know $f(\alpha) = f(z) = 0$, so cancelling them out, we get

$$-f'(z)s_n \equiv 0 \bmod p.$$

Since $f'(z) \not\equiv 0 \bmod p$, $s_n \equiv 0 \bmod p$, and we conclude $\alpha \equiv z \bmod p^n$ for all n. □

Example 4.3. Let $f(x) = x^2 - 11$. We know $f(1) \equiv 0 \bmod 5$ and $f'(1) \equiv 2 \not\equiv 0 \bmod 5$. By Hensel's Lemma, 11 has a unique square root in $\mathbb{Z}_5$ which is congruent to 1 mod 5. The exact square root goes on indefinitely, and it can be found by calculating its 5-adic expansion.

$$\begin{aligned} &11 \equiv 1^2 \bmod 5 \\ &11 \equiv (1+5)^2 \bmod 25 \\ &11 \equiv (1 + 5 + 2 \cdot 5^2)^2 \bmod 125 \\ &... \end{aligned}$$

So the exact solution is $z = 1 + 5 + 2 \cdot 5^2 + ...$.

In fact, for each $n > 0$ not divisible by p and $u \equiv 1 \bmod p$, u is an nth power in $\mathbb{Z}_p^\times$. To prove this, let $f(x) = x^n - u$. Then $f(1) = 1 - u \equiv 0 \bmod p$, $f'(1) = n(1)^{n-1} = n \not\equiv 0 \bmod p$. Hence, by Hensel's Lemma, there exists a unique solution z such that $z^n = u$ and $z \equiv 1 \bmod p$. Letting $p = 5$, $u = 11$, and $n = 2$, we obtain the previous result.

Example 4.4. With Hensel's Lemma, we can accurately describe when an element of $\mathbb{Q}_p$ is a square. Assume $p > 2$ and $q \in \mathbb{Q}_p^\times$. In order for q to be a square, its p-adic valuation must be even, so let $\mathrm{ord}_p(q) = 2k$ for some $k \in \mathbb{Z}$. Then $q' = q/p^{2k}$, of order 0, must also be a square. Hence the question becomes when an element $u \in \mathbb{Z}_p^\times$ is a square. Since u is a p-adic square, we know there exists v such that $u \equiv v^2 \bmod p$ and $|u|_p = |v|_p^2 = 1$ ($v \in \mathbb{Z}_p^\times$).

Let $f(x) = x^2 - u$. Then we have $f(v) \equiv 0 \bmod p$ and $f'(v) = 2v \not\equiv 0 \bmod p$. We set $p \neq 2$, so by Hensel's Lemma, $f(x)$ has a root that can be reduced to $v \bmod p$. Hence $u \in \mathbb{Z}_p^\times$ is a square if and only if it can be reduced to a square mod p. In other words, it is a square if and only if its image in $\mathbb{Z}/p$ is a square by the natural map from $\mathbb{Z}_p$ to $\mathbb{Z}/p$. For a numerical example, we know from Example 4.3 that 1 is a square mod 5, so $u \in \mathbb{Z}_5^\times$ is a 5-adic square if $u \equiv 1 \bmod 5$.

Note that the conclusion may not hold for $p = 2$, since $f'(v) = 2v \equiv 0 \bmod 2$. For example, 3 is a square mod 2 but not a square mod 4, so it cannot be in $\mathbb{Z}_2$ and hence not in $\mathbb{Q}_2$.

5. Hensel's Lemma: General Version

In this section, we present a more general version of Hensel's Lemma.

Theorem 5.1. *(Hensel's Lemma) Let $f(x) \in \mathbb{Z}_p[x]$ and $a \in \mathbb{Z}_p$ for a fixed p. Suppose $|f(a)|_p < |f'(a)|_p^2$. Then there exists a unique $z \in \mathbb{Z}_p$ such that the following three conditions hold:*

- $f(z) = 0$,
- $|z - a|_p = |f(a)/f'(a)|_p < |f'(a)|_p$,
- $|f'(z)|_p = |f'(a)|_p$.

The main difference between the two versions is that the given approximated root has to be a simple root for the basic version, whereas the general version allow roots with multiplicity greater than 1 to be used as the initial approximation. To see this, note that since $f(x) \in \mathbb{Z}_p[x]$, $f'(x) \in \mathbb{Z}_p[x]$ and $f'(a) \in \mathbb{Z}_p$, which means $|f'(a)|_p \leq 1$ by Definition 3.2. If $|f'(a)|_p = 1$, then $f'(a) \equiv 0 \bmod p$, and a is at least a double root. This is clearly not allowed in Theorem 4.1. Another important distinction between the two versions is that Theorem 5.1 gives a much clearer description on the bound of the p-adic distance between the approximated and the actual root.

Now we need yet another polynomial identity to prove the theorem.

Lemma 5.2. *Let F be a field and $f(x) \in F[x]$. Then*

$$f(x) - f(y) = (x - y)g(x, y)$$

where $g(x, y) \in F[x, y]$.

Proof. Note that $x - y$ is always a factor of $x^n - y^n$ for $n \geq 1$. If we write the polynomial as $f(x) = \sum_{i=0}^{n} a_i x^i$ for $n < \infty$, then

$$f(x) - f(y) = \sum_{i=0}^{n} a_i(x^i - y^i) = (x-y)\sum_{i=0}^{n} a_i \sum_{j=0}^{i-1} x^{i-1-j}y^j = (x-y)g(x,y)$$

and $g(x, y) \in F[x, y]$. □

Here we will cover two different proofs of Theorem 5.1, one by Newton's Method and the other by Contraction Mapping Theorem. Both focus on constructing a Cauchy sequence using certain iterations and then proving the limit of the sequence is the root proposed by Hensel's Lemma. We start with the version that uses Newton's Method.

Proof. (Hensel's Lemma) This proof especially resembles the proof written for the basic version, so we will omit some detailed calculations, which the readers can fill in as an exercise.

To find the desired p-adic Cauchy sequence, we define (a_n) with $a_1 = a$ and

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}. \tag{5.3}$$

We will show this sequence is Cauchy later on. For now, we want to show, by induction, that all members of this sequence satisfy three properties:

(1) $|a_n|_p \leq 1$ or $a_n \in \mathbb{Z}_p$
(2) $|f'(a_n)|_p = |f'(a)|_p$
(3) $|f(a_n)|_p \leq |f'(a)|_p^2 t^{2^{n-1}}$

where $t = |f(a)/f'(a)^2|_p$. The base case is straightforward, as a is in $\mathbb{Z}_p$ by given and the last property turns out be an equality after cancelling out terms. Now we proceed to the inductive step. Given n, assume the three properties hold. We want to show they still hold for the $n + 1$ case.

For property (1) to be true, both parts on the right of (5.3) have to be in $\mathbb{Z}_p$, so we need to show $f(a_n)/f'(a_n) \in \mathbb{Z}_p$ or $|f(a_n)/f'(a_n)|_p \leq 1$. Using (2) and (3) from case n, we have

$$|\frac{f(a_n)}{f'(a_n)}|_p \leq |f'(a)|_p t^{2^{n-1}} < 1,$$

since $t < 1$ by brief calculation.

To show property (2), we apply Lemma 5.2 to the derivative of $f(x)$ along with (5.3) and (3) from n. We have

$$|f'(a_{n+1}) - f'(a_n)|_p \leq |a_{n+1} - a_n|_p = |\frac{f(a_n)}{f'(a)}|_p < |f'(a)|_p.$$

By Remark 2.5, $|f'(a_{n+1})|_p = |f'(a)|_p$.

Finally, we use the polynomial identity in Lemma 4.2 to prove property (3). Letting $x = a_n$ and $y = -f(a_n)/f'(a_n)$, we get, by property (3) of case n,

$$|f(a_{n+1})|_p = |g(a_n, \frac{-f(a_n)}{f'(a_n)})(\frac{-f(a_n)}{f'(a_n)})^2|_p \leq |\frac{f(a_n)}{f'(a)}|_p^2 \leq \frac{|f(a)|_p^4 t^{2^{n-1}}}{|f'(a)|_p} \leq |f'(a)|_p^2 t^{2^{n-1}}.$$

With all three properties proven, now we want to show (a_n) is a Cauchy sequence in $\mathbb{Q}_p$. By property (3), for any two adjacent terms in the sequence, we have

$$|a_{n+1} - a_n|_p = |\frac{f(a_n)}{f'(a_n)}|_p \leq |f'(a)|_p t^{2^{n-1}}, \tag{5.4}$$

so the upper bound of the distance gets smaller as $n \to \infty$.

Let z be the limit of this sequence. Since property (1) holds true for all n, z is also a p-adic integer. Letting $n \to \infty$ in property (3), we get $|f(z)|_p \leq 0$, which means $f(z) = 0$.

If we let n approach ∞ for property (2), we get $|f'(z)|_p = |f'(a)|_p$, which proves condition (3) in the theorem.

Thus it remains to show the second condition, $|z-a|_p = |f(a)/f'(a)|_p < |f'(a)|_p$. The inequality follows from the given, so we only need to prove the equality in the front, which we divide into two cases. If $f(a) = 0$, then $z = a_n = a$. Both sides of the equation equal ∞, and the equality follows naturally. If $f(a) \neq 0$, however, then we want to show $|a_n - a|_p = |f(a)/f'(a)|_p$ for all n, using some simple induction.

Note that by (5.4),

$$|a_{n+1} - a_n|_p < |f'(a)|_p t = |\frac{f(a)}{f'(a)}|_p.$$

So by properties of p-adic absolute value, if $|a_n - a|_p = |f(a)/f'(a)|_p$, then $|a_{n+1} - a|_p = |f(a)/f'(a)|_p$. Induction is complete.

Lastly, we need to prove the uniqueness of root z. Assume α is another root. Then $f(\alpha) = 0$ and $|z-\alpha|_p < |f'(a)|_p$ by Remark 2.5, since both $|z-a|_p$, $|\alpha-a|_p < |f'(a)|_p$. Let $\alpha = z + c$ for some $c \in \mathbb{Z}_p$, $c \neq 0$. By Lemma 4.2,

$$f(\alpha) = f'(z)c + g(z,c)c^2 = 0,$$

so $f'(z) = -g(z,c)c$, and

$$|f'(z)|_p \leq |c|_p = |\alpha - z|_p < |f'(a)|_p.$$

But $|f'(z)|_p = |f'(a)|_p$, and we reach a contradiction. Hence z is unique. □

Before we move on to the second proof of Hensel's Lemma, we need to review the Contraction Mapping Theorem.

Lemma 5.5. *(Contraction Mapping Theorem) Let (X, d) be a complete metric space and $f : X \to X$ be a contraction mapping, i.e. a map such that*

$$d(f(x), f(y)) \leq cd(x, y)$$

for some $0 \leq c < 1$ and any $x, y \in X$. Then f has a unique fixed point in X.

Due to limited space, this paper will not state a detailed proof of the Contraction Mapping Theorem. Readers who desire further information should consult Conrad's paper [3].

Proof. (Hensel's Lemma) First note that this proof will leave out some steps in the deduction of various inequalities, but all of them follows naturally from the given conditions or p-adic absolute value properties. Readers are encouraged to trace these steps themselves.

In this proof, we will use this iterative function

$$\phi(x) = x - \frac{f(x)}{f'(a)}$$

to construct the Cauchy sequence. But first, we have to show this function is a contraction mapping on a ball of some radius $r \in (0,1)$ around a. Thus for r yet to be determined, let $B_a(r) = \{x \in \mathbb{Q}_p : |x-a|_p \leq r\}$. Note that this ball has to be contained in $\mathbb{Z}_p$. The radius must be such that ϕ maps $B_a(r)$ back to itself and forms a contraction, i.e.

$$|\phi(x) - \phi(y)|_p = |x - y - \frac{f(x)-f(y)}{f'(a)}|_p \leq c|x-y|_p$$

for $x, y \in B_a(r)$, $c < 1$. We proceed to determine suitable r and c.

We know that we can write $f(x)$ in terms of $x-a$ using a change of variable formula for polynomials, or

$$f(x) = \sum_{i=0}^{n} \alpha_i (x-a)^i \tag{5.6}$$

for some $\alpha_i \in \mathbb{Z}_p$. Plugging a into the derivative of $f(x)$, we get $\alpha_1 = f'(a)$. Substitution gives

$$\phi(x) - \phi(y) = -\frac{1}{f'(a)} \sum_{i=2}^{n} \alpha_i ((x-a)^i - (y-a)^i).$$

By properties of p-adic absolute value and proof of Lemma 5.2, we get

$$\begin{aligned} |(x-a)^n - (y-a)^n|_p &= |x-y|_p | \sum_{i=0}^{n-1} (x-a)^{n-1-i}(y-a)^i|_p \\ &= |x-y|_p \max(|x-a|_p, |y-a|_p). \end{aligned}$$

Since both $x-a$ and $y-a$ are p-adic integers, $\max(|x-a|_p, |y-a|_p) \leq 1$. Thus

$$|\phi(x) - \phi(y)|_p \leq \frac{|x-y|_p}{|f'(a)|_p}.$$

Note that this inequality takes $c = 1$, which does not satisfy the conditions of a contraction mapping. In order to make ϕ a contraction mapping, we have to choose $c < 1$ such that $|x-a|_p, |y-a|_p \leq c|f'(a)|_p$ in order for the terms to cancel out for inequality to be $|\phi(x) - \phi(y)|_p \leq c|x-y|_p$. Note that ϕ is a contraction mapping means if $|x-a|_p \leq c|f'(a)|_p$, then $|\phi(x) - a|_p \leq c|f'(a)|_p$, i.e.

$$|x - a - \frac{f(x)}{f'(a)}|_p \leq c|f'(a)|_p,$$

which implies

$$|\frac{f(x)}{f'(a)}|_p \leq c|f'(a)|_p.$$

Substituting $f(a)$ with the form in (5.6), we get

$$\frac{f(x)}{f'(a)} = \frac{f(a)}{f'(a)} + (x-a) + \sum_{i=2}^{n} \frac{\alpha_i}{f'(a)} (x-a)^i.$$

For $i \geq 2$, we have

$$|\frac{\alpha_i}{f'(a)}(x-a)^i|_p \leq \frac{|x-a|_p^2}{|f'(a)|_p} \leq c^2 |f'(a)|_p \leq c|f'(a)|_p.$$

So $|f(x)/f'(a)|_p \leq c|f'(a)|_p$ by Remark 2.5, and $|f(a)/f'(a)^2|_p \leq c$. Since the desired c is less than 1, we need $|f(a)/f'(a)^2|_p < 1$. Yet this condition is already given

in the theorem, so we are free to set $c = |f(a)/f'(a)^2|_p$. After rapid calculation, we get the desired radius $r = |f(a)/f'(a)|_p$.

All components of ϕ determined, we can now construct the Cauchy sequence (a_n) by setting $a_1 = a$ and $a_{n+1} = \phi(a_n)$. As we mentioned in Section 3, $\mathbb{Q}_p$ is complete, and by Contraction Mapping Theorem, ϕ has a unique fixed point $z \in B_a(|f(a)/f'(a)|_p)$ such that $\phi(z) = z$. Substitution gives $f(z) = 0$.

At this point, there are still three things left to prove. First, we need to show that $|z - a|_p = |f(a)/f'(a)|_p$. Note, by $\phi(x)$, we have $|a_2 - a|_p = |f(a)/f'(a)|_p$. Then for all n, we have

$$|a_{n+1} - a_n|_p = |\phi^n(a) - \phi^{n-1}(a)|_p \leq c^{n-1}|a_2 - a|_p < |f(a)/f'(a)|_p.$$

Thus by properties of p-adic absolute value, $|a_n - a|_p = |f(a)/f'(a)|_p$, and we get $|z - a|_p = |f(a)/f'(a)|_p$ as $n \to \infty$.

Next, we need to prove $|f'(z)|_p = |f'(a)|_p$. We will use induction to show $|f'(a_n)|_p = |f'(a)|_p$ for all members of the sequence (a_n), and then take $n \to \infty$. The base case is straightforward, so it remains to show $|f'(a_{n+1})|_p = |f'(a)|_p$ given $|f'(a_n)|_p = |f'(a)|_p$. Note that

$$|f'(a_{n+1}) - f'(a_n)|_p \leq |a_{n+1} - a_n|_p \leq |\frac{f(a)}{f'(a)}|_p,$$

where the first inequality arises by Lemma 5.2 and the second by the radius of the ball. By Remark 2.5, $|f'(a_{n+1}) - f'(a_n)|_p < |f'(a)|_p$, and $|f'(a_{n+1})|_p = |f'(a)|_p$.

As for the uniqueness of the root z, we simply repeat the last part from the Newton's Method proof. □

Before we proceed onto a numerical application of the general version of Hensel's Lemma, note the similarity of the iterative functions in the two proofs. The only difference is that the denominator of the fraction is fixed to be $f'(a)$ for the latter, while it varies with a_n for the former. This distinction results in different rates at which the Cauchy sequence converges. Interested readers should see Conrad's paper [2, p.13] for details.

Example 5.7. Let $f(x) = x^3 + 17$. Note $f(x) \equiv (x - 1)^3 \bmod 3$, so let $a = 1$ be the approximated root. After some brief calculation, we know $f(1) = 18$ and $f'(1) = 3$, which means $|f(1)|_p \not< |f'(1)|_p^2$, thus the solution, if it exists, cannot be approximated by 1. However, if we lift the root to $a = 4$, we find that the conditions are satisfied, as $f(4) = 81$ and $f'(4) = 12$. By Hensel's Lemma, -17 has a unique cube root in $\mathbb{Z}_3$ which is congruent to 4 mod 9. The exact cube root can be written as $z = 1 + 3 + 3^2 + \ldots$.

Acknowledgments

I would like to first thank my mentor Xingyu for meeting with me frequently throughout the program, directing me towards numerous resources and providing abundant feedback for my paper. I would also like to thank Akhil Mathew for taking the time to give the lectures on quadratic forms, which sparked my interest in p-adic numbers in the first place. I am also grateful to Peter May for organizing the REU program and Daniil Rudenko for taking the effort to produce the wonderful lightboard videos for the Apprentice program. Last but not least, I am thankful to my family for their unwavering support during this difficult time.

References

[1] Kazuya Kato, Nobushige Kurokawa, and Takeshi Saito. Number Theory 1: Fermat's Dream (Translations of Mathematical Monographs Vol 1). American Mathematical Society. 2000.
[2] Keith Conrad. Hensel's Lemma. https://kconrad.math.uconn.edu/blurbs/gradnumthy/hensel.pdf
[3] Keith Conrad. The Contraction Mapping Theorem. https://kconrad.math.uconn.edu/blurbs/analysis/contraction.pdf
[4] Robert Lewis. A Formal Proof of Hensel's Lemma over the p-adic Integers. https://robertylewis.com/padics/padics.pdf