9 Local fields and Hensel's lemmas

9.1 Extending valuations

Recall from Lecture 3 that each prime $\mathfrak{p}$ of a Dedekind domain A determines a discrete valuation (a surjective homomorphism) $v_\mathfrak{p}\colon \mathcal{I}_A \to \mathbb{Z}$ that assigns to a nonzero fractional ideal I the exponent $n_\mathfrak{p}$ appearing in the unique factorization of $I = \prod \mathfrak{p}^{n_\mathfrak{p}}$ into prime ideals; equivalently, $v_\mathfrak{p}(I)$ is the unique integer n for which $IA_\mathfrak{p} = \mathfrak{p}^n A_\mathfrak{p}$. This induces a discrete valuation $v_\mathfrak{p}(x) := v_\mathfrak{p}(xA)$ on the fraction field K, and a corresponding absolute value $|x|_\mathfrak{p} := c^{v_\mathfrak{p}(x)}$ (with $0 < c < 1$). In the $AKLB$ setup, where L/K is a finite separable extension and B is the integral closure of A in L, the primes $\mathfrak{q}|\mathfrak{p}$ of B similarly give rise to discrete valuations $v_\mathfrak{q}$ on L, each of which restricts to a valuation on K; we would like to understand how these relate to $v_\mathfrak{p}$.

Definition 9.1. Let L/K be a finite separable extension, and let v and w be discrete valuations on K and L respectively. If $w|_K = ev$ for some $e \in \mathbb{Z}_{>0}$ then we say that w *extends* v *with index* e.

Theorem 9.2. *Assume $AKLB$ and let $\mathfrak{p}$ be a prime of A. For each prime $\mathfrak{q}|\mathfrak{p}$, the discrete valuation $v_\mathfrak{q}$ extends $v_\mathfrak{p}$ with index $e_\mathfrak{q}$, and every discrete valuation on L that extends $v_\mathfrak{p}$ arises in this way. In other words, the map $\mathfrak{q} \mapsto v_\mathfrak{q}$ gives a bijection from the set of primes $\mathfrak{q}|\mathfrak{p}$ to the set of discrete valuations of L that extend $v_\mathfrak{p}$.*

Proof. For each prime $\mathfrak{q}|\mathfrak{p}$ we have $v_\mathfrak{q}(\mathfrak{p}B) = e_\mathfrak{q}$ (by definition of the ramification index $e_\mathfrak{q}$), while $v_\mathfrak{q}(\mathfrak{r}B) = 0$ for all primes $\mathfrak{r} \neq \mathfrak{p}$ of A (since $\mathfrak{q}$ lies above only one prime of A). If $I = \prod_\mathfrak{r} \mathfrak{r}^{n_\mathfrak{r}}$ is any nonzero fractional ideal of A then

$$v_\mathfrak{q}(IB) = v_\mathfrak{q}\left(\prod_\mathfrak{r} \mathfrak{r}^{n_\mathfrak{r}} B\right) = v_\mathfrak{q}(\mathfrak{p}^{n_\mathfrak{p}}B) = v_\mathfrak{q}(\mathfrak{p}B)n_\mathfrak{p} = e_\mathfrak{q} n_\mathfrak{p} = e_\mathfrak{q} v_\mathfrak{p}(I),$$

so $v_\mathfrak{q}(x) = v_\mathfrak{q}(xB) = e_\mathfrak{q} v_\mathfrak{p}(xA) = e_\mathfrak{q} v_\mathfrak{p}(x)$ for all $x \in K^\times$; thus $v_\mathfrak{q}$ extends $v_\mathfrak{p}$ with index $e_\mathfrak{q}$.

If $\mathfrak{q}$ and $\mathfrak{q}'$ are two distinct primes above $\mathfrak{p}$ then neither contains the other and for any $x \in \mathfrak{q} - \mathfrak{q}'$ we have $v_\mathfrak{q}(x) > 0 \geq v_{\mathfrak{q}'}(x)$, thus $v_\mathfrak{q} \neq v_{\mathfrak{q}'}$ and the map $\mathfrak{q} \mapsto v_\mathfrak{q}$ is injective..

Let w be a discrete valuation on L that extends $v_\mathfrak{p}$, let $W = \{x \in L : w(x) \geq 0\}$ be the associated DVR, and let $\mathfrak{m} = \{x \in L : w(x) > 0\}$ be its maximal ideal. Since $w|_K = ev_\mathfrak{p}$, the discrete valuation w is nonnegative on A and positive on $\mathfrak{p}$, so $A \subseteq W$ and $\mathfrak{p} = \mathfrak{m} \cap A$. The DVR W is integrally closed in its fraction field L, so $B \subseteq W$. Let $\mathfrak{q} = \mathfrak{m} \cap B$. Then $\mathfrak{q}$ is prime (since $\mathfrak{m}$ is), and $\mathfrak{p} = \mathfrak{m} \cap A = \mathfrak{q} \cap A$, so $\mathfrak{q}$ lies over $\mathfrak{p}$. The ring W contains $B_\mathfrak{q}$ and is properly contained in L, which is the fraction field of $B_\mathfrak{q}$. But there are no intermediate rings between a DVR and its fraction field, so $W = B_\mathfrak{q}$ and $w = v_\mathfrak{q}$ (and $e = e_\mathfrak{q}$). □

9.2 Local fields

Definition 9.3. A *local field* is a field K with a nontrivial absolute value $|\ |$ that is locally compact under the topology induced by $|\ |$.

Recall that a topological space is *locally compact* if every point has a compact neighborhood.[1] The topology induced by $|\ |$ is given by the metric $d(x, y) := |x - y|$. A metric space is locally compact if and only if every point lies in a compact closed ball.

[1] Weaker definitions of locally compact are sometimes used, but they all imply this one and when X is Hausdorff (as it always will be for us) these alternative definitions are equivalent to ours.

Lecture by Andrew Sutherland

Example 9.4. Under the standard archimedean absolute value both $\mathbb{R}$ and $\mathbb{C}$ are local fields but $\mathbb{Q}$ is not. Indeed no closed ball in $\mathbb{Q}$ is compact, it is always missing limit points (in a metric space a compact set must contains all its limit points). Finite fields are not local fields because they have no nontrivial absolute values.

Our next goal is to give an alternative characterization of local fields which shows that they are precisely the fields we get by completing a global field. As in the previous lecture, we use $B_{<r}(x) := \{y : |y-x| < r\}$ to denote the open ball of radius $r \in \mathbb{R}_{>0}$ about x, and use $B_{\leq r}(x) := \{y : |y-x| \leq r\}$ to denote a closed ball. In any metric space, open balls are open sets and closed balls are closed sets; in a nonarchimedean metric space, open balls are both open and closed, as are closed balls.

Lemma 9.5. *Let K be a field with a nontrivial absolute value $|\;|$. Then K is a local field if and only if every (equivalently, any) closed ball in K is compact.*

Proof. Suppose K is a local field. Then $0 \in K$ lies in a compact closed ball $B_{\leq s}(0)$. Pick $c \in K^\times$ with $|c| > 1$ (this is possible because $|\;|$ is nontrivial). The map $x \mapsto cx$ is continuous and $|\;|$ is multiplicative, so $B_{\leq |c|^n s}(0)$ is compact for every $n \in \mathbb{Z}_{>0}$ (recall that the continuous image of a compact set is compact). We thus have compact balls about 0 of arbitrarily large radii, implying that every closed ball $B_{\leq r}(0)$ is a closed subset of a compact set, hence compact. For every $z \in K$ the translation map $x \mapsto x + z$ is continuous, so every closed ball $B_{\leq r}(z)$ is compact. This proves the forward implication, and the reverse implication is immediate. For the parenthetical, note that the argument above still works if we replace $B_{\leq s}(0)$ by any closed ball. □

Corollary 9.6. *Let K be a local field with absolute value $|\;|$. Then K is complete.*

Proof. Suppose not. Then there is a Cauchy sequence (x_n) in K that converges to a limit $x \in \widehat{K} - K$. Pick $N \in \mathbb{Z}_{>0}$ so that $|x_n - x| < 1/2$ for all $n \geq N$ (here we are using the extension of $|\;|$ to $\widehat{K}$), and consider the closed ball $S := B_{\leq 1}(x_N)$ in K, which is compact by Lemma 9.5. The Cauchy sequence $(x_n)_{n \geq N}$ in S has a convergent subsequence whose limit lies in $S \subseteq K$, since S is compact, but this limit must be equal to $x \notin K$, a contradiction. □

Proposition 9.7. *Let K be a field with absolute value $|\;|_v$ induced by a discrete valuation v, let A be its valuation ring, and let π be a uniformizer. Then K is a local field if and only if K is complete and the residue field $A/\pi A$ is finite.*

Proof. If K is a local field then K is complete, by Corollary 9.6, and the valuation ring

$$A = \{x \in K : v(x) \geq 0\} = \{x \in K : |x|_v \leq 1\} = B_{\leq 1}(0)$$

is a closed ball, hence compact, by Lemma 9.5. The cosets $x + \pi A$ of the subgroup $\pi A \subseteq A$ are open balls $B_{<1}(x)$, since $y \in x + \pi A$ if and only if $|x - y|_v \leq |\pi|_v < 1$. The set $\{x + \pi A : x \in A\}$ of cosets of πA is thus an open cover of A by disjoint sets which must be finite, since A is compact. Thus $A/\pi A$ is finite.

Now suppose that K is complete and $A/\pi A$ is finite. Then $A = \hat{A}$ is complete and Proposition 8.11 gives an isomorphism of topological rings

$$A = \hat{A} \simeq \varprojlim_n \frac{A}{\pi^n A}.$$

Each quotient $A/\pi^n A$ is finite, since $A/\pi A$ is, and therefore compact; it follows that the inverse limit, and therefore A, is compact, by Proposition 8.10. Thus K contains a compact closed ball $B_{\leq 1}(0) = A$ and is locally compact by Lemma 9.5, hence a local field. □

Recall that a global field L is a finite extension of $\mathbb{Q}$ or $\mathbb{F}_q(t)$ that we can always assume to be separable (if L has positive characteristic p we take $\mathbb{F}_q$ to be the algebraic closure of $\mathbb{F}_p$ in L and choose t so that it is a separating transcendental element). In particular, we are always in an $AKLB$ setting, where $A = \mathcal{O}_K$ is either $\mathbb{Z}$ or $\mathbb{F}_q[t]$ (both Dedekind domains) and $B = \mathcal{O}_L$ is the integral closure of A in L. The residue fields of A thus have the form $\mathbb{Z}/p\mathbb{Z}$ or $\mathbb{F}_q[t]/(f)$ for some irreducible $f \in \mathbb{F}_q[t]$ and are all finite, and it follows that the residue fields of B are all finite, since they are all finite extensions of a residue field of A.

Corollary 9.8. *Let L be a global field with a nontrivial absolute value $|\ |_v$. Then the completion L_v of L with respect to $|\ |_v$ is a local field.*

Proof. Let L/K be a finite separable extension with $K = \mathbb{Q}$ or $K = \mathbb{F}_q(t)$, let $A = \mathcal{O}_K$, and let $B = \mathcal{O}_L$. If $|\ |_v$ is archimedean, then $K = \mathbb{Q}$ and the completion of L with respect to $|\ |_v$ must be a finite extension of $\mathbb{R}$, the completion of $\mathbb{Q}$ with respect to its archimedean absolute value (which is unique up to equivalence, by Ostrowski's theorem). Therefore L_v is isomorphic to either $\mathbb{R}$ or $\mathbb{C}$ (as a topological field), both of which are local fields.

We now assume that $|\ |_v$ is nonarchimedean. We claim that in this case $|\ |_v$ is induced by a discrete valuation. Let $C := \{x \in L : |x|_v \le 1\}$. be the valuation ring corresponding to $|\ |_v$. Then C is a local ring equal to the union of its maximal ideal $\mathfrak{m} := \{x \in L : |x|_v < 1\}$ and its unit group $C^\times = \{x \in L : |x|_v = 1\}$. The restriction of $|\ |_v$ to K is a nonarchimedean absolute value, and from the classification of absolute values on $\mathbb{Q}$ and $\mathbb{F}_q(t)$ proved on Problem Set 1, we can assume it is induced by a discrete valuation on A; in particular, $|x|_v \le 1$ for all $x \in A$, and therefore $A \subseteq C$. Like all valuation rings (discrete or not), C is integrally closed in its fraction field L, and C contains A, so C contains B. The ideal $\mathfrak{q} = \mathfrak{m} \cap B$ is therefore maximal, and the DVR $B_\mathfrak{q}$ lies in C and is therefore equal to C, since there are no intermediate rings between a DVR and its fraction field (we cannot have $C = L$ because $|\ |_v$ is nontrivial). It follows that the absolute value induced by $v_\mathfrak{q}$ is equivalent to $|\ |_v$, since they have the same valuation rings, and by choosing $0 < c < 1$ appropriately, we can assume $|\cdot|_v = c^{v_\mathfrak{q}(\cdot)}$ is induced by $v_\mathfrak{q}$. The residue field $B_\mathfrak{q}/\mathfrak{q}B_\mathfrak{q} \simeq B/\mathfrak{q}$ is finite, since it is a finite extension of the finite field $A/\mathfrak{p}$.

If we now consider the completion L_v with valuation ring B_v, we can take any uniformizer π for $\mathfrak{q} \subseteq B \subseteq B_v$ as a uniformizer for B_v, and we have

$$\frac{B}{\mathfrak{q}} \simeq \frac{B_\mathfrak{q}}{\mathfrak{q}B_\mathfrak{q}} = \frac{B_\mathfrak{q}}{\pi B_\mathfrak{q}} \simeq \frac{B_v}{\pi B_v},$$

so $B_v/\pi B_v$ is finite. Thus L_v is a complete field with an absolute value induced by a discrete valuation with finite residue field, and therefore a local field, by Proposition 9.7. □

In order to classify all local fields we require the following result from topology.

Proposition 9.9. *A locally compact topological vector space over a nondiscrete locally compact field has finite dimension.*

Proof. See [3, Prop. 4-13.iv]. □

Theorem 9.10. *Let L be a local field. If L is archimedean then it is isomorphic to $\mathbb{R}$ or $\mathbb{C}$; otherwise, L is isomorphic to a finite extension of $\mathbb{Q}_p$ or $\mathbb{F}_p((t))$ for some prime p.*

Proof. Let L be a local field with nontrivial absolute value $|\ |$; then L is complete, by Corollary 9.6. If $\operatorname{char}(L) = 0$, then the prime field of L is $\mathbb{Q}$, and L contains the completion of $\mathbb{Q}$ with respect to $|\ |$, by the universal property of completions. By Ostrowski's theorem (see Problem Set 1), L contains a subfield K isomorphic to $\mathbb{Q}_p$ for some prime p, or to $\mathbb{R}$.

If $\operatorname{char}(k) = p > 0$ then the prime field of L is $\mathbb{F}_p$, and L must contain a transcendental element t, since no algebraic extension of $\mathbb{F}_p$ has a nontrivial absolute value: in an algebraic extension L of $\mathbb{F}_p$, every nonzero $\alpha \in L^\times$ has some finite order n, and this implies $|\alpha| = 1$ because $\alpha^n = 1$ implies $|\alpha^n| = |\alpha|^n = 1$ and therefore $|\alpha| = 1$, because the only nth root of 1 in $\mathbb{R}_{\geq 0}$ is 1. Thus L contains the completion of $\mathbb{F}_p(t)$ with respect to $|\ |$, and every completion of $\mathbb{F}_p(t)$ is isomorphic to $\mathbb{F}_q((t))$ for some q a power of p, each of which is a finite extension of $\mathbb{F}_p((t))$. So in this case L contains a subfield K isomorphic to $\mathbb{F}_p((t))$.

If K is archimedean then $K = \mathbb{R}$ is a local field, and if K is nonarchimedean then Proposition 9.7 implies that K is a local field, since $\mathbb{Q}_p$ and $\mathbb{F}_\mathfrak{p}((t))$ are both complete fields with discrete valuations that have finite residue fields. Thus K is a local field and therefore locally compact; it is nondiscrete because its absolute value is nontrivial. Proposition 9.9 implies that L has finite degree over K. If K is archimedean then $K = \mathbb{R}$, and L must be $\mathbb{R}$ or $\mathbb{C}$; otherwise, L is a finite extension of $\mathbb{Q}_p$ or $\mathbb{F}_p((t))$ as claimed. □

9.3 Hensel's lemmas

Definition 9.11. Let R be a (commutative) ring, and let $f(x) = \sum_{i=0}^d f_i x^i \in R[x]$ be a polynomial. The (*formal*) *derivative* f' of f is the polynomial $f'(x) := \sum_{i=1}^d i f_i x^{i-1} \in R[x]$.

Note that the canonical ring homomorphism $\mathbb{Z} \to R$ defined by $1 \mapsto 1$ allows us to view the integers i as elements of R (the map $\mathbb{Z} \to R$ will be injective only when R has characteristic zero, but it is well defined in any case). It is easy to verify that the formal derivative satisfies the following identities:

$$\begin{aligned}(f+g)' &= f' + g',\\ (fg)' &= f'g + fg',\\ (f\circ g)' &= (f'\circ g)g'.\end{aligned}$$

When the characteristic of R is positive, we may have $\deg f' < \deg f - 1$. For example, if R has characteristic $p > 0$ and $g(x) = f(x^p)$ for some $f \in R[x]$, then $g' = f'(x^p)px^{p-1} = 0$ (conversely, one can show that $g' = 0$ implies $g(x) = f(x^p)$ for some $f \in R[x]$).

Lemma 9.12. *Let R be a ring, let $f = \sum_{i=0}^d f_i x^i \in R[x]$ be a polynomial, and let $a \in R$. Then $f(x) = f(a) + f'(a)(x-a) + g(x)(x-a)^2$ for a unique $g \in R[x]$.*

Proof. Without loss of generality, we assume $d \geq 2$ (let $f_i = 0$ for $i > \deg f$). We have

$$\begin{aligned}
f(x) &= f(a + (x-a)) \\
&= \sum_{i=0}^{d} f_i (a + (x-a))^i \\
&= \sum_{i=0}^{d} f_i \sum_{j=0}^{i} \binom{i}{j} a^j (x-a)^{i-j} \\
&= f(a) + \sum_{i=1}^{d} f_i \sum_{j=0}^{i-1} \binom{i}{j} a^j (x-a)^{i-j} \\
&= f(a) + f'(a)(x-a) + \sum_{i=2}^{d} f_i \sum_{j=0}^{i-2} \binom{i}{j} a^j (x-a)^{i-j} \\
&= f(a) + f'(a)(x-a) + \left(\sum_{i=2}^{d} f_i \sum_{j=0}^{i-2} \binom{i}{j} a^j (x-a)^{i-2-j} \right)(x-a)^2,
\end{aligned}$$

so we can take $g(x) = \sum_{i=2}^{d} f_i \sum_{j=0}^{i-2} \binom{i}{j} a^j (x-a)^{i-2-j} \in R[x]$. □

Remark 9.13. The lemma can be viewed as giving the first two terms of a formal Taylor expansion of $f(x)$ about a. Note that the binomial coefficients $\binom{i}{j}$ are integers, hence well defined elements of R under the canonical homomorphism $\mathbb{Z} \to R$, even if $j!$ is divisible by the characteristic of R. In the usual Taylor expansion

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

used in characteristic zero, if f is a polynomial then $f^{(i)}(a)$ is necessarily a multiple of $i!$, so $f^{(i)}(a)/i!$ is actually a well defined element of R in any characteristic.

Corollary 9.14. *Let R be a ring, $f \in R[x]$, and $a \in R$. Then $f(a) = f'(a) = 0$ if and only if a is (at least) a double root of f, that is, $f(x) = (x-a)^2 g(x)$ for some $g \in R[x]$.*

Definition 9.15. Let $f \in R[x]$ be a polynomial over a ring R and let $a \in R$. If $f(a) = 0$ and $f'(a) \neq 0$ then a is a *simple root* of f.

If R is a ring and I is an ideal, by a *lift* of an element of R/I, we mean a preimage under the quotient map $R \to R/I$. We now state the (apparently) weakest form of what is known as *Hensel's Lemma.*

Lemma 9.16 (Hensel's Lemma I). *Let A be a complete DVR with maximal ideal $\mathfrak{p}$ and residue field $k := A/\mathfrak{p}$. Suppose $f \in A[x]$ is a monic polynomial whose reduction to $k[x]$ has a simple root $r \in k$. Then r can be lifted to a root of f in A.*

Proof. We work in the quotient field K of A. Let a_0 be any lift of r to A; the element a_0 is not necessarily a root of f, but it is a root modulo $\mathfrak{p}$. We will show that a_0 is the first term of a Cauchy sequence (a_n), where each a_n is a root of f modulo $\mathfrak{p}^{2^n}$. In terms of the absolute value $|\cdot| := c^{v_{\mathfrak{p}}(\cdot)}$ (for some $0 < c < 1$) on K, we have $|f(a_n)| \leq c^{2^n}$ rapidly converging to 0. The assumption that r is a simple root means that $|f'(a_0)| = 1$, and we have $\epsilon := |f(a_0)/f'(a_0)^2| \le c < 1$.

Our proof only requires $\epsilon < 1$, so it actually works in many cases where r is not a simple root (see Lemma 9.17 below); we also don't need f to be monic. For each $n \ge 0$ we define

$$a_{n+1} := a_n - f(a_n)/f'(a_n).$$

We will prove by induction on n that

- **(a)** $|a_n| \le 1$ (so $a_n \in A$);
- **(b)** $|a_n - a_0| \le \epsilon < 1$ (so $a_n \equiv a_1 \bmod \mathfrak{p}$, equivalently, a_n is a lift of r);
- **(c)** $|f'(a_n)| = |f'(a_0)| \ne 0$ (so a_{n+1} is well defined);
- **(d)** $|f(a_n)| \le \epsilon^{2^n}|f'(a_0)|^2$ (so $|f(a_n)|$ and therefore $f(a_n)$ converges rapidly to 0).

The base case $n = 0$ is clear. We now assume (a), (b), (c), (d) for n and prove them for $n + 1$:

- (a) $|a_{n+1} - a_n| = |f(a_n)/f'(a_n)| \le \epsilon^{2^n}|f'(a_0)^2|/|f'(a_0)| = \epsilon^{2^n}|f'(a_0)| \le \epsilon^{2^n}$ (by (c),(d)), therefore $|a_{n+1}| = |a_{n+1} - a_n + a_n| \le \max(|a_{n+1} - a_n|, |a_n|) \le 1$ (by (a)).
- (b) $|a_{n+1} - a_0| \le \max(|a_{n+1} - a_n|, |a_n - a_0|) \le \max(\epsilon^{2^n}, \epsilon) = \epsilon$ (as above and using (b)).
- (c) Applying Lemma 9.12 to $f'(x)$ at a_n and substituting a_{n+1} for x yields

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

  where $f''(a_n) \in A$ and $\alpha = g(a_{n+1}) \in A$ for some $g \in A[x]$ (so $|f''(a_n)|, |\alpha| \le 1$). We have $|f(a_n)/f'(a_n)| = |f(a_n)|/|f'(a_0)| \le \epsilon^{2^n}|f'(a_0)|$, by (d), so the absolute values of the last two terms on the RHS are strictly smaller than first $|f'(a_n)| = |f'(a_0)|$, thus $|f'(a_{n+1})| = |f'(a_n)| = |f'(a_0)| \ne 0$.
- (d) Applying Lemma 9.12 to $f(x)$ and substituting a_{n+1} for x yields

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

  where $\beta = h(a_{n+1})$ for some $h \in A[x]$. We have $|\beta| \le 1$, so (d) gives

$$|f(a_{n+1})| \le |f(a_n)/f'(a_n)|^2 = |f(a_n)|^2/|f'(a_0)|^2 \le \epsilon^{2^{n+1}}|f'(a_0)|^2.$$

We have $|a_{n+1} - a_n| \le \epsilon^{2^n} \to 0$ as $n \to \infty$, and for a nonarchimedean absolute value this implies (a_n) is Cauchy. Thus $a := \lim_{n\to\infty} a_n \in A$, since A is complete. We have $f(a) = \lim_{n\to\infty} f(a_n) = 0$, so a is a root of f, and $|a - a_0| = \lim_{n\to\infty} |a_n - a_0| < 1$, so a is a lift of $r \equiv a_0 \bmod \mathfrak{p}$. □

We now record the stronger form of Hensel's lemma that we actually proved above.

Lemma 9.17 (Hensel's Lemma II). *Let A be a complete DVR. Let $f \in A[x]$, and suppose $a_0 \in A$ satisfies*

$$|f(a_0)| < |f'(a_0)|^2$$

(in particular, $f'(a_1) \neq 0$), and for $n \geq 0$ define

$$a_{n+1} := a_n - f(a_n)/f'(a_n).$$

The sequence (a_n) is well-defined and converges to the unique root $a \in A$ of f for which

$$|a - a_0| \leq \epsilon := |f(a_0)|/|f'(a_0)|^2.$$

Moreover, $|f(a_n)| \leq \epsilon^{2n}|f'(a_0)|^2$ for all $n \geq 0$.

We should note the similarity between Lemma 9.17 and Newton's method for finding (or more closely approximating) a root of a polynomial given an initial approximation. Like Newton's method, the recurrence in Lemma 9.17 converges quadratically, meaning that we double the number of p-adic digits in our approximation with each iteration. But Lemma 9.17 is even better than Newton's method, for two reasons: (1) if the residue field is finite, finding an initial approximation is very easy, and (2) once we have an initial approximation with $\epsilon < 1$, convergence is guaranteed.

Remark 9.18. The hypothesis in Lemmas 9.16 and 9.17 that A is a complete DVR is not necessary, the proof generalizes to any complete local ring. But even completeness is not strictly necessary. A local ring A in which Lemma 9.16 holds without the hypothesis that A is a complete DVR is called a *henselian ring.* One can show that Lemma 9.17 necessarily also holds in any henselian ring, as do many other forms of "Hensel's Lemma", including Lemma 9.20 below. In general, any lemma that holds for a local ring if and only if it is a henselian ring may be called "Hensel's Lemma", and there are at least a dozen candidates; see [4, Tag 04GE], for example. One can define the *henselization* of a noetherian local ring R as the minimal extension of a ring that is henselian (as usual, it is minimal (and unique) in the sense of satisfying a universal property); in many cases this turns out to be the algebraic closure of R in its completion. The henselization of R is often much smaller than its completion (e.g. finite over R in cases where $\widehat{R}$ is not), and can serve as a substitute for the completion in algebraic settings.

Example 9.19. Let $A = \mathbb{Z}_5$ and $f(x) = x^2 - 6 \in \mathbb{Z}_5[x]$. Then $\bar{f}(x) = x^2 - 1 \in \mathbb{F}_5[x]$ has $r = 1$ as a simple root. By Lemma 9.16 there is a unique $a \in \mathbb{Z}_5$ such that $a^2 - 6 = 0$ and $a \equiv 1 \bmod 5$. We could also have chosen $r = -1$, which would give another distinct root of $f(x)$, which must be $-a$. Thus $\mathbb{Z}_5$ contains two distinct square roots of 6.

Now let $A = \mathbb{Z}_2$ and $f(x) = x^2 - 17$. Then $\bar{f}(x) = x^2 - 1 = (x-1)^2$ has no simple roots (note $\bar{f}' = 0$). But if we let $a_0 = 1$, then $f(a_0) = -16$ and $|f(a_0)| = 1/16$, while $f'(a_0) = 2$ and $|f'(a_0)| = 1/2$. We thus have $|f(a_0| < |f'(a_0)|^2$ and can apply Lemma 9.17 to get a square root of 17 in $\mathbb{Z}_2$.

There is a another version of Hensel's Lemma we need (which is considered by some to be the "canonical" one). Recall that polynomial over a ring is *primitive* if its coefficients generate the unit ideal (over a DVR this just means that at least one coefficient is a unit).

Lemma 9.20 (Hensel's lemma III)**.** *Let A be a complete DVR with maximal ideal $\mathfrak{p}$ and residue field k, let $f \in A[x]$ be a primitive polynomial with image $\bar{f}$ in $k[x]$, and suppose $\bar{f} = \bar{g}\bar{h}$ for some coprime $\bar{g}, \bar{h} \in k[x]$. Then there exist polynomials $g, h \in A[x]$ for which $f = gh$ with $g \equiv \bar{g} \bmod \mathfrak{p}$ and $h \equiv \bar{h} \bmod \mathfrak{p}$ such that $\deg g = \deg \bar{g}$.*

Proof. See [2, Theorem II.4.6]. □

This form of Hensel's lemma has the following useful corollary.

Lemma 9.21 (Hensel-Kürschák lemma)**.** *Let A be a complete DVR with fraction field K, and let $f \in K[x]$ be an irreducible polynomial whose leading and constant coefficients lie in A. Then $f \in A[x]$.*

Proof. Let $\mathfrak{p} = (\pi)$ be the maximal ideal of A, let $k := A/\mathfrak{p}$, and write $f = \sum_{i=0}^{n} f_i x^i$ with $f_n \neq 0$. Let $m = \min\{v_{\mathfrak{p}}(f_i)\}$. Suppose for the sake of contradiction that $m < 0$, and let $g := \pi^{-m} f = \sum_{i=0}^{n} g_i x^i \in A[x]$. Then g is a primitive irreducible polynomial in $A[x]$ with $g_0, g_n \in \mathfrak{p}$, since $m < 0$ and $f_0, f_n \in A$. The reduction $\bar{g}$ of g to $k[x]$ is divisible by $\bar{u} := x^d$ for some uniquely determined d for which the quotient $\bar{v} := g/x^d \in k[x]$ is coprime. We must have $0 < d < n$ because $g_0, g_n \in \mathfrak{p}$ both reduce to zero in $k = A/\mathfrak{p}$. It follows from Lemma 9.20 that $g = uv$ for some $u, v \in A[x]$ with $0 < \deg u = \deg \bar{u} < n$. This implies that g is not irreducible, which is a contradiction. So $m \geq 0$ and $f \in A[x]$. □

Corollary 9.22. *Let A be a complete DVR with fraction field K, and let L/K be a finite extension of degree n. Then $\alpha \in L$ is integral over A if and only if $\mathrm{N}_{L/K}(\alpha) \in A$.*

Proof. Let $f = \sum_{i=0}^{d} f_i x^i \in K[x]$ be the minimal polynomial of α. If α is integral over A then $f \in A[x]$ (by Proposition 1.25) and $N_{L/K}(\alpha) = (-1)^n f(0)^e \in A$, where $e = [L : K(b)]$, by Proposition 4.44. Conversely, if $N_{L/K}(\alpha) = (-1)^n f(0)^e \in A$, then $f(0) \in A$, since $f(0) \in K$ is a root of $x^e - (-1)^n N_{L/K}(\alpha) \in A[x]$ and A is integrally closed. Thus the constant coefficient of f lies in A, as does its leading coefficient (it is monic), so $f \in A[x]$, by Lemma 9.21. □

References

[1] N. Bourbaki, *General Topology: Chapters 1-4*, Springer, 1985.

[2] J. Neukirch, *Algebraic number theory*, Springer, 1999.

[3] D. Ramakrishnan and R.J. Valenza, *Fourier analysis on number fields*, Springer, 1999.

[4] Stacks Project Authors, *Stacks Project*, `http://stacks.math.columbia.edu`.

MIT OpenCourseWare
https://ocw.mit.edu

18.785 Number Theory I
Fall 2016

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.