HENSEL'S LEMMA AND SOME APPLICATIONS

D. Katz

The purpose of this note is to record Hensel's Lemma and a few of its immediate applications for my algebra class. In the early part of the 20th century Hensel used the notion of completion in a number theoretic context in order to bring to bear techniques of analysis on some purely algebraic problems in number theory. One of his most crucial discoveries (in modern terms) was that when a ring is complete in a suitable ideal-adic topology, solutions to polynomial equations with appropriate nondegeneracy conditions exist in the ring provided solutions already exist modulo the ideal. We will concern ourselves with the case that the ring in question is a local ring.

Throughout A will denote a local ring with maximal ideal m. We will assume that A is complete in the m-adic topology and we also write $k := A/m$ for the residue field of A. Thus, the paragraph above suggests that we will be able to lift solutions to polynomial equations from k to A.

We begin with some preliminary remarks. Suppose that S is a commutative ring and $J \subseteq S$ is an ideal. Given a polynomial $f(X) \in S[X]$, we can reduce its coefficients modulo J and thereby obtain a polynomial $\overline{f}(X)$ in $(S/J)[X] = S[X]/J[X]$. A solution to the equation $\overline{f}(X) = 0$ means there exists $\alpha \in S/J$ such that $\overline{f}(\alpha) = 0$ in S/J. Of course, $\alpha = \overline{a}$, for some $a \in A$, not necessarily unique. To say that we can lift the solution α of the equation $\overline{f}(X) = 0$ to a solution in S just means that $f(a) = 0$ in S for some a satisfying $\overline{a} = \alpha$. We now record some elementary facts that will play a crucial role in the development of Hensel's Lemma.

(a) Suppose that $g(X)$, $h(X)$ are polynomials in $S[X]$, with $g(X)$ monic, whose images in $(S/J)[X]$ generate the unit ideal. If J is contained in the Jacobson radical of A, then $g(X)$ and $h(X)$ generate the unit ideal in $S[X]$. To see this, suppose that $M \subseteq S[X]$ is a maximal ideal containing $g(X)$. Since $S[X]/(g(X)))$ is an integral extension of S, M contracts to a maximal ideal in S, which means that M contains J and hence $J[X]$. By hypothesis, M

cannot contain $h(X)$. Thus, no maximal ideal in $S[X]$ contains both $g(X)$ and $f(X)$, which is what we needed to show.

(b) Suppose $g(X)$ and $h(X)$ generate the unit ideal in $S[X]$. Let $f(X) \in S[X]$. Then there exist $a(X), b(X) \in S[X]$ such that $f(X) = a(X)g(X) + b(X)h(X)$ with $deg(b(X))$ less than $deg(g(X))$. To see this, since $g(X)$ and $h(X)$ generate the unit ideal, there exist $u(X), v(X) \in S[X]$ such that $f(X) = u(X)g(X) + v(X)h(X)$. If $deg(v(X)) < deg(g(X))$, this gives what we want. Otherwise, since $g(X)$ is monic, we may divide $v(X)$ by $g(X)$ and write $v(X) = t(X)g(X) + b(X)$, with $deg(b(X)) < deg(g(X))$. Substituting for $v(X)$, we get $f(X) = (u(X) + t(X))g(X) + g(X)h(X)$, as desired.

(c) Let $g(X)$ and $h(X)$ generate the unit ideal. Suppose that $a(X)g(X)+b(X)h(X) = 0$, for $a(X), b(X) \in S[X]$ with $deg(b(X)) < deg(g(X))$. Then $a(X) = b(X) = 0$. To see this, suppose that $1 = u(X)g(X) + v(X)h(X)$. Then from $a(X)v(X)g(X) + b(X)v(X)h(X) = 0$, we obtain $a(X)v(X)g(X) + (1 - u(X)g(X))b(X) = 0$, so $b(X) = (u(X)b(X) - a(X)v(X))g(X)$. Since $deg(b(X)) < deg(g(X))$ and $g(X)$ is monic, we get $b(X) = 0$. Thus, $a(X)g(X) = 0$, from which it follows that $a(X) = 0$.

We are now ready to state and prove Hensel's Lemma.

Hensel's Lemma. *Let (A, m, k) be a local ring that is complete and Hausdorff in its m-adic topology. Let $f(X) \in A[X]$ be a monic polynomial of degree d. Suppose that $g(X), h(X)$ are monic polynomials in $A[X]$ of degrees r and $d-r$ such that $\overline{f}(X) = \overline{g}(X)\overline{h}(X)$ in $k[X]$, with $\overline{g}(X)$ and $\overline{h}(X)$ relatively prime. Then there exist monic polynomials $\hat{g}(X), \hat{h}(X) \in A[X]$ of degrees r and $d - r$ such that $f(X) = \hat{g}(X)\hat{h}(X)$ and $\hat{g}(X)$, $\hat{h}(X)$ reduce to $\overline{g}(X)$, $\overline{h}(X)$ modulo m.*

Proof. We begin with an observation about the $m[X]$-adic topology on $A[X]$. It is not the case that $A[X]$ is complete in this topology. However, if $\{g_n(X)\}$ is a Cauchy sequence of polynomials of *bounded degree*, then this sequence converges. Indeed, suppose d bounds the degrees of the $g_n(X)$ and write each $g_n(X) := a_{n,d}X^d + \cdots + a_{n,0}$. Then for each $0 \leq j \leq d$, the sequence $\{a_{n,j}\}$ is a Cauchy sequence in A, and thus converges to an element $a_j \in A$. It follows immediately from this that the sequence $\{g_n(X)\}$ converges to $g(X) := a_dX^d + \cdots + a_0$.

We now start with the factorization $f(X) \equiv g(X)h(X)$ modulo $m[X]$. By induction on $n \geq 1$, we will find polynomials $a_n(X), b_n(X) \in A[X]$ such that : $a_n(X), b_n(X) \in m^n[x]$,

$deg(a_n(X)) < d - r$, $deg(b_n(X)) < r$, and such that for

$$g_{n+1}(X) := g(X) + b_1(X) + \cdots + b_n(X) \text{ and } h_{n+1}(X) := h(X) + a_1(X) + \cdots + a_n(X),$$

$f(X) - g_{n+1}(X)h_{n+1}(X) \in m^{n+1}[X]$. Suppose that this can be accomplished. Then by the remarks in the preceding paragraph, the Cauchy sequence $\{g_{n+1}(X)\}$ must converge to some $\hat{g}(X)$ having degree r and the Cauchy sequence $\{h_{n+1}(X)\}$ converges to $\hat{h}(X)$ having degree $d - r$. On the one hand, by definition, $0 = \lim_n(f(X) - g_{n+1}(X)h_{n+1}(X))$. On the other hand

$$\lim_n(f(X) - g_{n+1}(X)h_{n+1}(X)) = f(X) - \lim_n(g_{n+1}(X)) \cdot \lim_n(h_{n+1}(X)) = f - \hat{g}(X)\hat{h}(X).$$

Thus, $f(X) = \hat{g}(X)\hat{h}(X)$, as desired.

It remains to construct the $a_n(X)$ and $b_n(X)$. We do so by induction on n. Since the construction for the base case $n = 1$ is essentially the same as the case for general n, we assume $a_1(x), b_1(x), \ldots, a_n(x), b_n(x)$ with the required properties have already been constructed, and find the required $a_{n+1}(x)$ and $b_{n+1}(x)$.

Thus, $g_{n+1}(X)$ and $h_{n+1}(X)$ reduce modulo $m[x]$ to $\overline{g}(X)$ and $\overline{h}(X)$ respectively. Since $\overline{g}(X)$ and $\overline{h}(X)$ are relatively prime in $k[X]$, they generate the unit ideal in $k[X] = A/m[X]$. Now set $S := A/m^{n+2}$ and $J := m/m^{n+2}$, its unique maximal ideal. Then the images of $g_{n+1}(X)$ and $h_{n+1}(X)$ in $S[X]$ generate the unit ideal, by observation (a) above. Thus there exist $a_{n+1}(X), b_{n+1}(X) \in A[X]$, with $deg(b_{n+1}(X)) < deg(g(X))$ and

$$f - g_{n+1}(X)h_{n+1}(X) \equiv a_{n+1}(X)g_{n+1}(X) + b_{n+1}(X)h_{n+1}(X) \text{ modulo } \ m^{n+2}[X],$$

by observations (b) and (a). Since $deg(g_{n+1}(X)) = deg(g(X))$ and $deg(h_{n+1}(X)) = deg(h(X))$, it follows that $deg(a_{n+1}(X)) < deg(h_{n+1}(X))$. By hypothesis $f(X) - g_{n+1}(X)h_{n+1}(X)$ belongs to $m^{n+1}[X]$, so

$$a_{n+1}(X)g_{n+1}(X) + b_{n+1}(X)h_{n+1}(X) \equiv 0 \text{ modulo } m^{n+1}[X].$$

Since $g_{n+1}(X)$ and $h_{n+1}(X)$ also generate the unit ideal modulo $m^{n+1}[X]$, observation (c) gives that both $a_{n+1}(X)$ and $b_{n+1}(X)$ are congruent to 0 modulo $m^{n+1}[X]$, i.e., $a_{n+1}(X)$, $b_{n+1}(X)$ are in $m^{n+1}[X]$, which is what we want. Finally, for $g_{n+2}(X) := g_{n+1}(X)+b_{n+1}(X)$ and $h_{n+2}(X) := h_{n+1}(X) + a_{n+1}(X)$, we have that $f(X) - g_{n+2}(X)h_{n+2}(X) =$

$$f(X) - g_{n+1}(X)h_{n+1}(X) - a_{n+1}(X)g_{n+1}(X) - b_{n+1}(X)h_{n+1}(X) - a_{n+1}(X)b_{n+1}(X).$$

By construction, $f(X) - g_{n+1}(X)h_{n+1}(X) - a_{n+1}(X)g_{n+1}(X) - b_{n+1}(X)h_{n+1}(X)$ belongs to $m^{n+2}[X]$, as does the product $a_{n+1}(X)b_{n+1}(X)$. Thus, $f(X) - g_{n+2}(X)h_{n+2}(X)$ belongs to $m^{n+2}[X]$, as required. This completes the proof of Hensel's Lemma.

We now record a number of applications of Hensel's Lemma. The first provides the promised result about lifting roots from k to A.

Corollary A. *Let A be a complete local ring and $f(X) \in A[X]$ a monic polynomial. Suppose that $\overline{f}(X)$ admits a simple root $\alpha \in k$. Then there exists $a \in A$ such that $f(a) = 0$ and $\overline{a} = \alpha$.*

Proof. In $k[X]$ we can write $f(X) = (X - \alpha)t(X)$, for some $t(X) \in k[X]$ not divisible by $X - \alpha$. By Hensel's Lemma, there exist monic $g(X), h(X) \in A[X]$ such that $deg(g(X)) = 1$, $f(X) = g(X)h(X)$ and $g(X), h(X)$ reduce modulo m to $X - \alpha, t(X)$. It follows immediately that $g(X) = X - a$, for some $a \in A$ and that $\overline{a} = \alpha$. Of course, $f(a) = 0$.

Example B. Let A denote the 7-adic integers, i.e., the completion of the ring $\mathbb{Z}$ at the ideal $7\mathbb{Z}$. As mentioned in class, A is also the completion of the ring obtained by localizing $\mathbb{Z}$ at the maximal ideal $7\mathbb{Z}$ first, so A is a complete local ring and its maximal ideal is $7A$. It is also the case that $k = \mathbb{Z}_7$, the field with 7 elements. We first observe that by Corollary A, 2 has two square roots in A. For this, consider $f(X) = X^2 - 2 \in A[X]$. Then $\overline{f}(X)$ has two distinct roots in $\mathbb{Z}_7$, so it also has two distinct roots in A. From class we know that the elements of A can be written as infinite series whose terms come from increasing powers of $m = 7A$. In fact, 7-adic integers can always be written *uniquely* in the form $\sum_{n=0}^{\infty} a_n 7^n$, such that $0 \leq a_n \leq 6$. We now show how to use the proof of Hensel's Lemma to find an expression for one of the square roots of 2 – or at least the first few terms in its 7-adic expansion.

Start with $f = X^2 - 2$ which factors as $(X-3)(X+3)$ modulo $7A$. Note that $A/7A = \mathbb{Z}/7\mathbb{Z}$. So we take $g(X) = (X-3)$ and $h(X) = (X+3)$, i.e., the factors of $f(X)$ modulo 7. Now, by the proof of Hensel's Lemma we must find $a_1(X), b_1(X)$ such that $f(X) - ((X-3)(X+3))$ is congruent to $a_1(X)(X-3) + b_1(X)(X+3)$ modulo 7^2A i.e, in the ring $A/7^2A = \mathbb{Z}/7^2\mathbb{Z}$. In other words, we must solve the congruence $7 \equiv a_1(X)(X-3) + b_1(X)(X+3)$ modulo 49. A quick check shows that we can take $a_1(X) = 7$ and $b_1(X) = -7$. Thus we write $g_2(X) = (X-3) - 7$ and $h_2(X) = (X+3) + 7$, as in the proof of Hensel's Lemma. Dropping X from the notation in polynomials, we must now solve the congruence $f - g_2h_2 \equiv a_2g_2 + b_2h_2$

modulo $7^3\mathbb{Z}$. Multiplying out the polynomials in question shows that we must solve $98 \equiv -20a$ modulo 343, for some integer a. The solution to this is $a = 98$, which means we take $a_2 = 98 = 2{\cdot}7^2$ and $b_2 = -98-2{\cdot}7^2$. Thus, $g_3 = (X-3)-7-2{\cdot}7^2$ and $h_3 = (X+3)+7+2{\cdot}7^2$. Repeating once more gives the start of the 7-adic expansion of $\sqrt{2}$ as $3+1{\cdot}7+2{\cdot}7^3+6{\cdot}7^3+\cdots$.

Example C. In a similar vein, we could take $A = \mathbb{R}[[T]]$, the formal power series ring in one variable over $\mathbb{R}$. Then $m = TA$ and $k = \mathbb{R}$. The polynomial $T + 1 \in A$ certainly does not admit a polynomial square root. But it does admit a power series square root, since the polynomial $f(X) = X^2 - (T + 1)$ has distinct roots in $\mathbb{R}$, i.e., when reduced modulo TA. If one invokes the method of Hensel's Lemma to find $\sqrt{T+1}$, one ultimately observes that the resulting power series is just the Taylor series about 0 for $\sqrt{T+1}$.

Corollary D. *(Implicit function theorem) Let k be a field and let $A := k[[X_1, \ldots, X_d]]$ be the formal power series ring in d variables over k. Let $P(Y) := Y^n + a_{n-1}Y^{n-1} + \cdots + a_0$ belong to $A[Y]$. Suppose that the polynomial $Y^n + a_{n-1}(0)Y^{n-1} + \cdots + a_0(0)$ in $k[Y]$ has a simple root $\alpha \in k$. Then there exists a power series $g \in A$ such that $g(0) = \alpha$ and $P(g) = 0$. In particular, if n is a positive integer relativley prime to $char(k)$ and $f \in A$ is a power series whose constant term is an nth power in A, then f is an nth power in A, i.e., there exists a power series $h \in A$ such that $f = h^n$.*

Proof. The statements follow from Corollary A, since A is a complete local ring with respect to the m-adic topology, for $m := (X_1, \ldots, X_d)$, and its residue field is k. Note also that if $g \in A$ is a power series, then $g(0)$ is both the constant term of g and the image of g in the residue field. With these comments, the first statement is clear and the second follows since a polynomial of the type $X^n - f$ with coefficients in a field has a simple root if its derivative is relatively prime to it, and this follows if n is relatively prime to the characteristic of the field.

Our next results concern the notion of 'coefficient field' for a local ring. Suppose that A is a local ring with maximal ideal m and residue field k. Let $F \subseteq A$ be a field. Note that since any non-zero element in F is a unit in A, it follows that $m \cap F = (0)$. Thus $F = F/(0) = F/(m \cap F)$ maps isomorphically onto its image $\overline{F}$ in $A/m = k$. If the image of F in k equals k, we say that F *is a coefficient field* for A. For example, if $A := k[[X_1, \ldots, X_d]]$, then k is a coefficient field for A. The Cohen structure theorem states that any complete local ring containing a field has a coefficient field. Hensel's Lemma will enable us to do this

in two crucial cases : (a) A contains a field and its residue field is finite and (b) A contains a field of characteristic zero.

Corollary E. *Let (A,m,k) be a complete local ring containing a field K. If either k is finite or K has characteristic zero, then A contains a coefficient field.*

Proof. Suppose first that k is finite. Then $k = p^n$, for some prime p and $n \geq 1$. Since K maps into k via the canonical homomorphism from A to k, K also has characteristic p. Moreover, since any subfield of A containing K maps isomorphically into a subfield of k containing the image of K, we may find a finite subfield $K_0 \subseteq A$ maximal with respect to the property of containing K. We claim K_0 is the required coefficient field. If not, writing $\overline{K_0}$ for the image of K_0 in k, we can find $\alpha \in k$ not in $\overline{K_0}$.

Now, every non-zero element of k satisfies $f(X) := X^{p^n-1} - 1$. Since there are $p^n - 1$ such elements, every non-zero element of k is a simple root of $f(X)$. In particular, α is a simple root of $f(X)$. Thus, by Corollary A, we may lift α to $a \in A$ such that $f(a) = 0$ in A. Since $\alpha \notin \overline{K_0}$, $a \notin K_0$. Thus, $K_0[a]$ strictly contains K_0. Since a is algebraic over K_0, $K_0[a]$ is a field, and this contradicts the maximality of K_0. Thus, K_0 is a coefficient field.

When K has characteristic zero, the proof is similar. Let $\mathcal{C}$ denote the set of subfields of A containing K. An easy application of Zorn's Lemma yields a maximal element $F \in \mathcal{C}$, i.e., F is a maximal subfield of A. We now argue that F is a coefficient field. Suppose not. Then the image $\overline{F}$ of F in k is properly contained in k. Take $\alpha \in k$ not in $\overline{F}$. Suppose first that α is not algebraic over $\overline{F}$. Let $a \in A$ be such that $\overline{a} = \alpha$ on k. Note that a is not algebraic over F (in A). Then for all $f(X) \in F[X]$, $\overline{f}(\alpha) \neq 0$, which means that in A, $f(a) \notin m$. Thus each such expression is a unit it A. It follows that $F(a)$, the rational function field in the indeterminate a, is contained in A, contradicting the maximality of F. Now assume that α is algebraic over $\overline{F}$. Let $f(X) \in F[X]$ be such that $\overline{f}(X)$ is the minimal polynomial for α over $\overline{F}$. Since $char(\overline{F}) = 0$, α is a simple root. By Corollary A, there exists $a \in A$ such that $\overline{a} = \alpha$ and $f(a) = 0$. Note also that $f(X)$ is irreducible over F. Thus $F[a]$ is a field in A properly containing F, a contradiction. Thus we must have $\overline{F} = k$, so that F is a coefficient field for k

Department of Mathematics, the University of Kansas