Hensel's Lemma

Alejandro De Las Peñas Castaño

1/June/ 2017

Contents

Introduction

Newton's method for approximating the zeros of a smooth function $f : \mathbb{R}rightarrow\mathbb{R}$ says that, given an initial guess $x_0 \in \mathbb{R}$, the sequence of real numbers

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converges to some nearby root of the equation $f(x) = 0$. The idea is that each x_n is the root of a polynomial approximation of f. In Newton's case, these approximations are the partial sums of the Taylor series:

$$f_1(x) = f(0), \quad f_2(x) = f(0) + f'(0)x, \quad f_2(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 +, \ldots.$$

The error $\epsilon_N(x) = f(x) - f_N(x)$ is a function that vanishes to order N around 0. That is, it is an element of the Nth power of the maximal ideal

$$\mathfrak{m} = \{g : \mathbb{R}rightarrow\mathbb{R} \mid g(0) = 0\}$$

of the ring of smooth real-valued functions $C^\infty(\mathbb{R})$. The sequence of ideals $\mathfrak{m} \supset \mathfrak{m}^2 \supset \mathfrak{m}^3 \supset \cdots$ contains the sequence of errors of each successive approximation. Newton's method works precisely because

$$\epsilon_N(x) = \sum_{k \geq N} \frac{f^{(k)}(x)}{k!}x^k \ \in \mathfrak{m}^N.$$

Thus we are tempted to declare that the sequence of approximations $\{f_N\}$ *converges* to f because each approximation becomes better and better, or in other words: for larger N, the ideal $\mathfrak{m}^N$ gets 'smaller and smaller' because any function $f_\epsilon \in \mathfrak{m}^N$ produces very small perturbations of an arbitrary function f around a sufficiently small neighbourhood of 0, i.e. $f + f_\epsilon \approx f$ locally around 0. In metric spaces it is clear what convergences means because we can actually measure how small these errors are.

We should note that in $C^\infty(\mathbb{R})$, these polynomials don't always converge, or they can converge to a different function then the one we started with. For example the Taylor series of $f(x) = e^{-1/x^2}$ is identically zero. However, if we restrict our attention to polynomials, then Newton's Method works perfectly from this perspective.

So how can we apply Newton's method to other rings A? For example how do we apply Newton's method to Diophantine equations, that is roots of polynomials in $\mathbb{Z}[x]$ or $\mathbb{Z}[x_1, \ldots, x_n]$? Clearly not every polynomial in $\mathbb{Z}[x]$ has roots in $\mathbb{Z}$, but maybe we can still approximate these roots by defining a suitable topology on $\mathbb{Z}$ that makes Newton's method always converge.

It's very probable that questions like these motivated Kurt Hensel (1861-1941) to introduce the p-adic integers. He wanted to approximate the roots of polynomials with integer coefficients by reducing $\mod p^n$ for successively larger powers of some suitable prime number p. We will see (cf. Section 4)

that Newton's Method is very closely related with Hensel's Lemma and how it is used to approximate the roots of a polynomial.

Hensel's methods generalized Newton's so that they could be applied to commutative algebra and algebraic geometry where Hensel's Lemma proves to be a central theorem for studying complete rings. The crossover between real analysis and algebra, created by Hensel's Lemma, is what allowed Irvin S. Cohen (1917-1955) to mine Hensel's Lemma and classify complete local rings [1, Part II, pg 70-85]. Cohen also classified complete regular local rings [1, Theorem 17, pg 92].

In Section 1, we will study *filtrations* and how they allow us to define topologies on a ring A so that we may speak of convergence. Once this has been settled, in Section 2, we define the *completion* of a ring. Next we prove Hensel's Lemma in Section 3. In the last section, we study the first applications Hensel made with his lemma and later we exhibit a special case of The Cohen Structure Theorem which can be directly proven with Hensel's Lemma.

1 Filtrations

Notation: Unless otherwise specified, M will be an $A-$module and $\mathfrak{M} = \mathfrak{m}_0 \supset \mathfrak{m}_1 \supset \mathfrak{m}_2 \supset \cdots$ a descending chain of submodules of M.

Definition 1. Let $I \leq A$. An I-filtration of M is a decreasing sequence of submodules

$$\mathfrak{M}_I = \{\mathfrak{m}_n\} \ : \ M = \mathfrak{m}_0 \supset \mathfrak{m}_1 \supset \mathfrak{m}_2 \supset \cdots$$

which satisfies $I\mathfrak{m}_k \subseteq \mathfrak{m}_{k+1}$ for all $k \in \mathbb{N}$. The filtration $\{\mathfrak{m}_n\}$ provides the abelian group $(M, +)$ with the induced topology:

$$\tau(\mathfrak{M}_I) := \{U \subseteq M \mid f + \mathfrak{m}_n \subseteq U \text{ for some } f \in M \text{ and } n \in \mathbb{N}\},$$

called the $\mathfrak{M}_I$-topology of $M = (M, \mathfrak{M}_I)$. This topology makes M into a topological group.

Most commonly, the filtration $\mathfrak{M}_I$ will be the sequence of powers of an ideal I of some ring A; we call

$$\mathcal{I} \ : \quad I \supseteq I^2 \supseteq I^3 \supseteq \cdots$$

the $I - adic\ filtration$ of A and $\tau(\mathcal{I})$, its associated topology, is called the $I-adic\ topology$ of A. The basis for this topology is the set of cosets $f + I^n$ for all the $f \in A$ and $n \in \mathbb{N}$. More precisely, if $A \overset{\nu_n}{\twoheadrightarrow} A/I^n$ is the natural projection, then the basis for the $I-$adic topology is the set of all the preimages $\nu_n^{-1}[f] \subseteq A$.

First we verify that $(M, \tau(\mathfrak{M}_I))$ is indeed a topological group. Consider the operation maps:

$$M \times M \xrightarrow{+} M \quad \text{and} \quad M \xrightarrow{-} M$$

Let $U = f + \mathfrak{m}_N$ be a basic open set, and $(g, h) \in (+)^{-1}[U] \subseteq M \times M$; note that this implies $g + h - f \in \mathfrak{m}_N$. An open neighbourhood of (g, h), inside $(+)^{-1}[U]$, is the basic open set $(g + \mathfrak{m}_N) \times (h + \mathfrak{m}_N)$ because $g + h + \mathfrak{m}_N = f + \mathfrak{m}_N$. To verify that the inversion function $f \mapsto -f$ is continuous, simply observe that the inverse image of U, namely $-U = -f + \mathfrak{m}_N$ is a basic open set.

Since M is a topological group, we can describe its basis via the translation maps $0 \mapsto f$ for all $f \in M$. Since translations are homeomorphisms, every basic open set $f + \mathfrak{m}_N$ is homeomorphic to $0 + \mathfrak{m}_N = \mathfrak{m}_N$ via the translation $0 \mapsto -f$. Thus when studying the $\mathfrak{M}_I$-topology we need only consider the open neighbourhoods of 0, i.e. $\aleph(0) = \{U \in \tau_I \mid 0 \in U\}$.

By definition of an I-filtration, we only require that $I\mathfrak{m}_n \subseteq \mathfrak{m}_{n+1}$, but if equality is achieved for all sufficiently large n, then we say that the I-filtration $\mathfrak{M}_I$ is *stable*. For example the I-adic filtration of a ring A is stable since $I(I^n) = I^{n+1}$ for all $n \in \mathbb{N}$. In general the I-filtration $\mathfrak{M}_I = \{I^n M\}$ is stable; we also call this the I-adic filtration of M.

A very important property of stable $I-$filtrations is that the induced topology on M depends only on I and not on the specific descending chain $\mathfrak{m}_0 \supset \mathfrak{m}_1 \supset \mathfrak{m}_2 \supset \cdots$. More precisely:

Proposition 1. Let $\mathfrak{M}_I = \{\mathfrak{m}_n\}$ and $\mathfrak{M}'_I = \{\mathfrak{m}'_n\}$ be two stable I-filtrations of M. Then they have *bounded difference*, that is: there exists an integer k such that

$$\mathfrak{m}'_{n+k} \subseteq \mathfrak{m}_n \quad \text{and} \quad \mathfrak{m}_{n+k} \subseteq \mathfrak{m}'_n \quad \forall n \in \mathbb{N}.$$

Therefore the $\mathfrak{M}_I$-topology and the $\mathfrak{M}'_I$-topology are equivalent.

Proof. First we reduce the problem. Since 'being of bounded difference' is transitive among I-filtrations (simply take the sum of both integer bounds) we may assume that one of the stable I-filtrations, say $\mathfrak{M}'$, is $\{I^n M\}$.

Since $\mathfrak{M}$ is an I-filtration, then $I\mathfrak{m}_n \subseteq \mathfrak{m}_{n+1}$ so that inductively we have $I^n M = I^n \mathfrak{m}_0 \subseteq \mathfrak{m}_n$ for all $n \in \mathbb{N}$. Now, since $\mathfrak{M}_I$ is stable there is an integer k such that $I\mathfrak{m}_n = \mathfrak{m}_{n+1}$ for all $n \geq k$, so that by induction we have $\mathfrak{m}_{n+k} = I^n \mathfrak{m}_k$ for all n. Putting this together we conclude:

$$\mathfrak{m}_{n+k} = I^n \mathfrak{m}_k \subseteq I^n \mathfrak{m}_0 = I^n M \quad \text{and} \quad I^{n+k} M \subseteq I^n M \subseteq \mathfrak{m}_n$$

So that k is the bound for the difference between $\{\mathfrak{m}_n\}$ and $\{I^n M\}$.

Both topologies are equivalent because any open set U of the $\mathfrak{M}_I$-topology contains some coset of $\mathfrak{m}_n$, but by bounded difference it contains some $\mathfrak{m}'_{n+k}$ so that U is open in the $\mathfrak{M}'_I$-topology. □

The proof of this proposition allows us to observe two facts. First, the property of being stable implies that the integer k from which the I-filtration $\mathfrak{M}_I$ stabilizes is the same bound for the difference between itself and the prototypical I-filtration $\{I^n M\}$. Secondly, and more importantly, any stable I-filtration defines the same topology as the I-adic filtration, so that we may define the *I-adic topology* of M to be the topology induced by *any* stable I-filtration of M. Thus whenever M is said to have the I-adic topology, then we mean that M

is equipped with some stable I-filtration who induces this topology; in general we will take this filtration to be the I-adic filtration. Because of this, we will suppress the term 'adic' and simply call them I-topology/filtration.

We can follow this result further. If we have an I-filtration $\mathfrak{M}_I$ of M and $N \leq M$ any submodule, then we can restrict $\mathfrak{M}_I$ to N to produce the restriction filtration:

$$\mathfrak{M}_I|_N := \{N \cap \mathfrak{m}_n\}$$

which is also an I-filtration. If the original filtration $\mathfrak{M}_I$ was stable, there are conditions for which the restriction filtration is also stable; these are given by the celebrated Artin-Rees Lemma:

Proposition 2. (Artin Rees Lemma) Let M be a finitely generated A-module over a noetherian ring and $\mathfrak{M}$ a stable I-filtration. Then for all submodules $N \leq M$, the restriction filtration $\mathfrak{M}|_N$ is a stable I-filtration of N.

An immediate consequence of this result is that the I-topology of N is the subspace topology of the I-topology on M. To prove the Artin-Rees lemma we must characterize the property of being stable with the help of graded modules.

Lemma 1. *Let M be a finitely generated A-module over a noetherian ring and $\mathfrak{M} = \{\mathfrak{m}_n\}$ an I-filtration. If we denote by*

$$A^* := \bigoplus_{n=0}^{\infty} I^n \quad \textit{and} \quad M^* := \bigoplus_{n=0}^{\infty} \mathfrak{m}_n$$

for the graded ring A^* and the graded A^*-module M^*, then:

$$\mathfrak{M} \textit{ is a stable } I\textit{-filtration} \iff M^* \textit{ is a finitely generated } A^*\textit{-module.}$$

Proof. First we observe that the subgroup

$$N_n := \mathfrak{m}_0 \oplus \mathfrak{m}_1 \oplus \cdots \oplus \mathfrak{m}_n$$

of M^* generates the following A^*-submodule of M^*:

$$M_n^* := \langle N_n \rangle_{A^*} = \mathfrak{m}_0 \oplus \mathfrak{m}_1 \oplus \cdots \oplus \mathfrak{m}_n \oplus I\mathfrak{m}_n \oplus I^2\mathfrak{m}_n \oplus \cdots$$

which is a finitely generated A^*-module. Since $I\mathfrak{m}_n \subseteq \mathfrak{m}_{n+1}$ by definition, these submodules form an ascending chain that stabilizes simultaneously with the noetherian condition. This chain stops if and only if $M^* = M_N^*$ for a sufficiently large $N \in \mathbb{N}$ which is equivalent to $\mathfrak{m}_{N+n} = I^n\mathfrak{m}_N$ (that is the tail of $\mathfrak{M}_N^*$ coincides with the tail of $\mathfrak{M}^*$) which is another way to write the definition of the stability of an I-filtration. Thus the lemma follows. □

Now we turn to the proof of the Artin-Rees Lemma:

Proof. (of Proposition 2) Since $\{N \cap \mathfrak{m}_n\}$ is an I-filtration of N it defines N^*, a graded A^*-submodule of M^*. On the other hand $\{\mathfrak{m}_n\}$ is a stable filtration, so that Lemma (1) guarantees that M^* is a finitely generated A^*-module. We also know that A^* is a noetherian graded ring because A is noetherian. We conclude that M^* is noetherian so that N^* is finitely generated; by Lemma (1) again we conclude that $\{N \cap \mathfrak{m}_n\}$ is a stable I-filtration of N. □

We end this section by putting together these results:

Proposition 3. Let M be a finitely generated module over a noetherian ring A and let N be a submodule. Then the I-topology of N is equivalent to the subspace topology induced by the I-topology of M, in general, both I-filtrations $\{I^n N\}$ of N and $\{(I^n M \cap N)\}$ of M have bounded difference.

2 Completions

Now that we have an adequate topology we can start generalizing Newton's method and the Inverse Function Theorem by rigorously defining what it means for a sequence $\{f_n\}$ to converge to an element f. We use the general definition of convergence in a topological space, particularly the topological group M with the I-topology induced by some stable I-filtration $\mathfrak{M}_I = \{\mathfrak{m}_n\}$. We will regard the decreasing chain $\mathfrak{m}_0 \supseteq \mathfrak{m}_1 \supseteq \cdots$ as decreasing in 'size' in the I-topology. With this in mind we can now define what it means for a sequence to arbitrarily approximate something.

Definition 2. Let M, as a topological group, be endowed with the I-topology for some stable I-filtration $\mathfrak{M} = \{\mathfrak{m}_n\}$. A sequence $\mathcal{F} = \{f_n\}$ of elements $f_n \in M$ is called a *Cauchy sequence* if for every open neighbourhood U of zero (i.e. $U = \mathfrak{m}_n$), there is a sufficiently large integer N such that $f_s - f_t \in U$ for all $s, t > N$. In symbols: for all $U = U_n = \mathfrak{m}_n$ we have

$$f_s \equiv f_t \mod \mathfrak{m}_n \text{ for all sufficiently large } s, t \in \mathbb{N}.$$

Furthermore, if there is an element $f \in M$ such that $f - f_n \in U$ for all sufficiently large n we say that the Cauchy sequence *converges* to f.

The set of Cauchy sequences can be embedded into $M^{\mathbb{N}}$ as a submodule. Furthermore, we define the classic equivalence relation:

$$\{f_n\} \sim \{g_n\} \quad \Longleftrightarrow \quad \{f_n - g_n\} \longrightarrow 0.$$

We call the set of Cauchy sequences in the I-topology, modulo $\sim$, the *I-completion of M* and denote it by $\widehat{M} = \widehat{M}_I$. This is again a submodule of $M^{\mathbb{N}}$ with the natural operations:

$$\{f_n\} + \{g_n\} = \{f_n + g_n\} \quad \text{and} \quad \{f_n\} \cdot \{g_n\} = \{f_n g_n\}.$$

We can also define the completion of a module in a purely algebraic way: we will prove that $\widehat{M}$ can be obtained via an inverse limit.

Lets consider the natural projection maps

$$M \xrightarrow{\nu_n} \frac{M}{\mathfrak{m}_n} \qquad f \longmapsto f + \mathfrak{m}_n$$

and take $\mathcal{F} = \{f_n\}$ to be any Cauchy sequence in M. Projecting it onto $M/\mathfrak{m}_k$ produces the sequence

$$\widehat{M} \xrightarrow{\widehat{\nu}_n} \prod_k \frac{M}{\mathfrak{m}_n} \qquad \{f_k\} \longmapsto \{\nu_n(f_k)\}$$

that stabilizes into a constant. Indeed, by definition we have that for all $U = \mathfrak{m}_n$ there is a sufficiently large N such that $f_s \equiv f_t \mod \mathfrak{m}_n$ for all $s, t > N$, thus the sequence $\nu_n(\{f_k\}) \in M/\mathfrak{m}_n$ is constant after the Nth term, say $\delta_n(\mathcal{F}) \in M/\mathfrak{m}_n$. We therefore have a family of homomorphisms:

$$\left\{ \widehat{M} \xrightarrow{\delta_n} \frac{M}{\mathfrak{m}_n} \right\}_{n \in \mathbb{N}}$$

that induces the following maps:

$$\widehat{M} \xrightarrow{\delta} \prod \frac{M}{\mathfrak{m}_n} \xrightarrow{p_n} \frac{M}{\mathfrak{m}_n} \qquad \mathcal{F} = \{f_n\} \longmapsto \{\delta_n(\mathcal{F})\} \longmapsto \delta_n(\mathcal{F}).$$

Since $\delta_n(\{f_k\})$ is the constant to which $\widehat{\nu}_n(\{f_k\})$ stabilizes we have that $\delta_n(\{f_k\}) = \nu_n(f_N) = f_N + \mathfrak{m}_n$ for some sufficiently large $N = N(n) \in \mathbb{N}$. Similarly, $\delta_{n+1}(\{f_k\}) = \nu_{n+1}(f_N) = f_N + \mathfrak{m}_{n+1}$ where we may take the same N without loss of generality. With this notation and given the natural projection maps:

$$\frac{M}{\mathfrak{m}_{n+1}} \xrightarrow{\theta_{n+1}} \frac{M}{\mathfrak{m}_n} \qquad f + \mathfrak{m}_{n+1} \longmapsto f + \mathfrak{m}_n$$

we can easily prove that:

$$\theta_{n+1}(\delta_{n+1}(\{f_k\})) = \theta_{n+1}(f_N + \mathfrak{m}_{n+1}) = f_N + \mathfrak{m}_n = \delta_n(\{f_k\})$$

or equivalently, we have the following commutative diagram:

$$\begin{array}{ccc} \widehat{M} & \xrightarrow{\delta_{n+1}} & \dfrac{M}{\mathfrak{m}_{n+1}} \\ & \searrow^{\delta_n} & \downarrow \theta_{n+1} \\ & & \dfrac{M}{\mathfrak{m}_n} \end{array} \tag{1}$$

By the third Isomorphism Theorem, the family $\{M/\mathfrak{m}_n\}$ of quotient groups, together with the family of group homomorphisms:

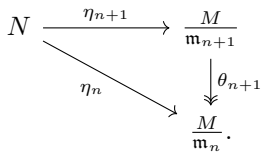
$$\left\{ \frac{M}{\mathfrak{m}_{n+1}} \xrightarrow{\theta_{n+1}} \frac{M}{\mathfrak{m}_n} \right\}_{n \in \mathbb{N}}$$

form an inverse system so that the canonical nature of the definition of the δ_n's that appear in (1) strongly suggests that $\widehat{M}$ is the inverse limit of the quotient groups $M/\mathfrak{m}_n$:

$$\widehat{M} \cong \varprojlim \frac{M}{\mathfrak{m}_n}. \tag{2}$$

To formally prove (2) we only need to prove that $\widehat{M}$ has the universal property of the inverse limit. To this end, let N be another A-module together with

a family of homomorphisms $\{\eta_n\}$ compatible with the $\{\theta_n\}$, that is they satisfy the following commutative diagram:



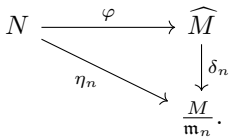
For an arbitrary $g \in N$, let $f_n(g)$ be a preimage of $\eta_n(g)$ in M, that is $f_n(g) + \mathfrak{m}_n = \eta_n(g)$ and similarly for $f_{n+1}(g)$, then the third isomorphism law and the above commutative diagram tell us that:

$$\begin{aligned} \theta_{n+1}(\eta_{n+1}(g)) =& \theta_{n+1}(f_{n+1}(g) + \mathfrak{m}_{n+1}) = f_{n+1}(g) + \mathfrak{m}_n, \\ \eta_n(g) =& f_n(g) + \mathfrak{m}_n \\ \therefore \quad f_n(g) \equiv& f_{n+1}(g) \mod \mathfrak{m}_n \end{aligned} \tag{3}$$

We can extend the above inductively to conclude that $f_s \equiv f_t \mod \mathfrak{m}^n$ for all sufficiently large s, t. It follows immediately that the sequence $\{f_n(g)\}$ is Cauchy.

Now we define the map $N \xrightarrow{\varphi} \widehat{M}$ that simply sends g to the sequence $\varphi(g) = \{f_n(g)\}$. The election of $f_n(g) \in M$ does not affect φ because if $\{f'_n(g)\}$ is any other choice then $\{f_n(g)\} \sim \{f'_n\}$ because $f_n - f'_n \in \mathfrak{m}^n$ for all n.

Lastly we observe that, by definition of φ, we have the following diagram:



If the above diagram were commutative then it would follow that $\widehat{M}$ satisfies the universal property of the inverse limit and we would be able to define the completion of M as the inverse limit of the quotient groups $M/\mathfrak{m}_n$.

The above diagram is indeed commutative: if $g \in N$ then by (3) the Cauchy sequence $\varphi(g) = \{f_n(g)\}$ stabilizes to $f_n(g) + \mathfrak{m}_n \in M/\mathfrak{m}_n$ so that $\delta_n(\varphi(g)) = f_n(g) \equiv \eta_n(g) \mod \mathfrak{m}_n$. We conclude that the completion of M is the inverse limit of the quotient groups $M/\mathfrak{m}_n$. This way, we may define the completion of a module in a purely algebraic manner:

Definition 3. Let M be equipped with the I-topology induced by some stable I-filtration $\mathfrak{M} = \{\mathfrak{m}_n\}$. The $I-$*completion* $\widehat{M}$ of M is defined to be the inverse limit:

$$\widehat{M} := \varprojlim \frac{M}{\mathfrak{m}_n}.$$

We also say that M is $I-$*complete* if the natural homomorphism $M \xrightarrow{\Sigma} \widehat{M}$ is an isomorphism.

The natural homomorphism $Mrightarrow\widehat{M}$ can be described with Cauchy sequences too: we simply map every element $f \in M$ to the constant Cauchy sequence $\{f\} \in \widehat{M}$. Observe that the kernel of this homomorphism is the intersection of all the open neighbourhoods about zero. Indeed, if a constant Cauchy sequence $\{f\}$ is equivalent to $\{0\}$, then for all open neighbourhoods U of zero we have $f - 0 = f \in U$.

Thus a necessary condition so that M can be I-complete is that the kernel:

$$\ker(\Sigma) = \bigcap_{0 \in U} U = \bigcap_{n=1}^{\infty} \mathfrak{m}_n = 0.$$

Since a topological group M is Hausdorff iff $\{0\}$ is closed, and clearly $\ker(\Sigma) = \overline{\{0\}}$, we have the following proposition:

Proposition 4. The I-topology of M is Hausdorff if and only if $\cap \mathfrak{m}_n = 0$ and if M is I-complete, then M is Hausdorff with the I-topology.

Remark. The converse is not necessarily true. It does hold true for the discrete topology though. Indeed, in the discrete topology, the Cauchy sequences are exactly the constant sequences. This means that if if the I-topology of M is discrete, then M is trivially complete.

Here is another important reminder: given any two stable I-filtrations $\mathfrak{M}$ and $\mathfrak{M}'$, the induced topologies on M are equivalent so that the set of Cauchy sequences are identical thus both completions of M are isomorphic. Furthermore, M is complete with respect to $\mathfrak{M}$ if and only if M is complete with respect to $\mathfrak{M}'$.

Now consider the I-topology and the I^N-topology given by the following (stable) filtrations:

$$M \supseteq IM \supseteq I^2M \supseteq I^3M \supseteq \cdots \quad \text{and} \quad M \supseteq I^NM \supseteq I^{2N}M \supseteq I^{3N}M \supseteq \cdots.$$

The equivalence relation of a Cauchy sequences is identical in either topology. Because the latter filtration is a subsequence of the I-filtration, then the convergence $\{f_n - g_n\}rightarrow0$ in the I-topology, implies convergence in the I^N-topology. Now suppose that $\{f_n - g_n\}rightarrow0$ in the I^N-topology, and let $m \in \mathbb{N}$. Take an integer k such that $m < kN$. Thus for all m, we have $f_n - g_n \in I^{kN} \subseteq I^m$ for all sufficiently large n.

We have proven that the equivalence relation of Cauchy sequences is independent of the power of the ideal I that defines the I-completion. In particular, if a ring A is complete with respect with an ideal I, then it is also complete with respect to any power of I.

The inverse limit definition is immediately useful because of the following proposition:

Proposition 5. Let

$$0 \longrightarrow \{A_n\} \longrightarrow \{B_n\} \longrightarrow \{C_n\} \longrightarrow 0$$

be an *exact sequence of inverse systems*, that is, for all $n \in \mathbb{N}$ we have the following commutative diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A_n & \longrightarrow & B_n & \longrightarrow & C_n & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha_n} & & \downarrow{\scriptstyle \beta_n} & & \downarrow{\scriptstyle \gamma_n} & & \\ 0 & \longrightarrow & A_{n-1} & \longrightarrow & B_{n-1} & \longrightarrow & C_{n-1} & \longrightarrow & 0. \end{array}$$

If $\{A_n\}$ is a *surjective* system (i.e. each inverse system homomorphism α_n is surjective), then

$$0 \longrightarrow \varprojlim A_n \longrightarrow \varprojlim B_n \longrightarrow \varprojlim C_n \longrightarrow 0.$$

is a short exact sequence. In other words, the inverse limit functor $\{A_n\} \mapsto \varprojlim A_n$ is exact.

Proof. First we embed $\varprojlim A_n$ into $A = \prod A_n$ as the kernel of the map:

$$A \xrightarrow{\alpha_A} A \qquad \{x_n\} \longmapsto \{x_n - \alpha_{n+1}(x_{n+1})\}.$$

This is indeed the desired embedding because of the equivalent characterization of the inverse limit as a subring of A modulo the equivalence relation given by the above function. We define β_B and γ_C in the same manner where $\{\beta_n\}$ and $\{\gamma_n\}$ are the inverse system homomorphisms of $\{B_n\}$ and $\{C_n\}$ respectively.

By hypothesis, for each n we have homomorphisms $A_nrightarrowB_nrightarrowC_n$ so that the induced homomorphisms

$$A = \prod_{n\in\mathbb{N}} A_n \longrightarrow B = \prod_{n\in\mathbb{N}} B_n \longrightarrow C = \prod_{n\in\mathbb{N}} C_n$$

produce the commutative diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha_A} & & \downarrow{\scriptstyle \beta_B} & & \downarrow{\scriptstyle \gamma_C} & & \\ 0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0. \end{array}$$

By the *Snake Lemma*, we have the following exact sequence:

$$0 \longrightarrow \ker(\alpha_A) \longrightarrow \ker(\beta_B) \longrightarrow \ker(\gamma_C) \longrightarrow \operatorname{coker}(\alpha_A) \longrightarrow \operatorname{coker}(\beta_B) \longrightarrow \operatorname{coker}(\gamma_C) \longrightarrow 0.$$

In particular, if we concentrate on the first four terms we conclude:

$$0 \longrightarrow \varprojlim A_n \longrightarrow \varprojlim B_n \longrightarrow \varprojlim C_n \tag{4}$$

or in other words, the inverse limit functor is left exact. Observe that we still haven't use the hypothesis that $\{A_n\}$ is a surjective system. Thus (4) holds in general.

To finish the proof we need only prove that if $\{A_n\}$ is a surjective system, then α_A is surjective so that $\operatorname{coker}(\alpha_A) \cong 0$ and thus (4) can be completed into a short exact sequence. To this end, let $\{x_n\} \in A$.

Let $y_1 \in A_1$ be arbitrary and choose $y_2 \in A_2$ such that $\alpha_2(y_2) = y_1 - x_1 \in A_1$. This is possible because each α_n is surjective by hypothesis. In this case

$$y_1 \mapsto y_1 - \alpha_2(y_2) = y_1 - y_1 + x_1 = x_1$$

Now choose $y_3 \in A_3$ such that $\alpha_3(y_3) = y_2 - x_2 \in A_2$, thus:

$$y_2 \mapsto y_2 - \alpha_3(y_3) = x_2.$$

It is clear that we can inductively produce a sequence $\{y_n\}$ such that $\{y_n\} \mapsto \{x_n\}$. We conclude that α_A is surjective and we are done. □

We apply this proposition to the inverse system $\{M/\mathfrak{m}_n\}$ whose homomorphisms are the canonical projections of $M/\mathfrak{m}_{n+1}$ onto $M/\mathfrak{m}_n$. In particular, $\{M/\mathfrak{m}_n\}$ is a surjective system and the strong part of the previous proposition applies:

Corollary 1. *Let* $0 \longrightarrow M' \longrightarrow M \xrightarrow{\pi} M'' \longrightarrow 0$ *be short exact sequence of finitely generated modules over a noetherian ring* A *and endow each module with its own* I*-topology for some ideal* $I \leq A$*. Then the sequence of* I*-completions*

$$0 \longrightarrow \widehat{M'} \longrightarrow \widehat{M} \longrightarrow \widehat{M''} \longrightarrow 0$$

is exact.

Proof. To be able to apply Proposition 5 to the exact sequence of inverse systems

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \frac{M'}{(I^{n+1}M)\cap M'} & \longrightarrow & \frac{M}{I^{n+1}M} & \longrightarrow & \frac{M''}{\pi[I^{n+1}M]} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \frac{M'}{(I^{n}M)\cap M'} & \longrightarrow & \frac{M}{I^{n}M} & \longrightarrow & \frac{M''}{\pi[I^{n}M]} & \longrightarrow & 0.
\end{array}$$

we only need to prove that the I-topology of M' is equivalent to the subspace topology induced by M. But this is immediate for the Artin-Rees Lemma. □

This corollary allows us to 'embed' the descending sequence of any stable I-filtration into a descending sequence of submodules of the I-completion. More precisely:

Let $\{\mathfrak{m}_n\}$ be a stable I-filtration of M and restrict it to some fixed $\mathfrak{m}_N$; with the notation of the previous corollary this is $M' = \mathfrak{m}_N$. This means that $M'' = M/\mathfrak{m}_N$ and that

$$0 \longrightarrow \mathfrak{m}_N \longrightarrow M \longrightarrow \frac{M}{\mathfrak{m}_N} \longrightarrow 0$$

is a short exact sequence. By Corollary 1,

$$0 \longrightarrow \widehat{\mathfrak{m}}_N \longrightarrow \widehat{M} \longrightarrow \frac{\widehat{M}}{\mathfrak{m}_N} \longrightarrow 0$$

is a short exact sequence. Thus

$$\frac{\widehat{M}}{\widehat{\mathfrak{m}}_N} \cong \frac{\widehat{M}}{\mathfrak{m}_N}. \tag{5}$$

On the other hand, the induced I-topology on $M/\mathfrak{m}_N$ is the discrete topology because the inverse image of a singleton $\{[f]\} \subseteq M/\mathfrak{m}_N$ under the natural projection is the basic open set $f + \mathfrak{m}_N$. Therefore $M/\mathfrak{m}_N$ is complete, so that the right hand side of (5) reduces to $M/\mathfrak{m}_N$.

We have just proven:

Corollary 2. *Let M be a finitely generated module over a noetherian ring A with a stable I-filtration $\mathfrak{M} = \{\mathfrak{m}_n\}$ together with the induced I-topology. Then the completion of $\mathfrak{m}_n$ is a submodule of $\widehat{M}$, i.e. $\widehat{\mathfrak{m}}_n \leq \widehat{M}$, and:*

$$\frac{\widehat{M}}{\widehat{\mathfrak{m}}_n} \cong \frac{M}{\mathfrak{m}_n}.$$

Taking inverse limits on each side we get

$$\widehat{\widehat{M}} = \widehat{M}$$

or in other words: the I-completion of module M is always a I-complete.

We now turn our focus to the ring A of the A-module M. Lets consider the I-filtration $\mathcal{I} = \{I^n\}$ in A and set $\widehat{A}$ to be the I-completion of A. Since there exists the natural A-module homomorphism $A \xrightarrow{\Sigma} \widehat{A}$, then $\widehat{A}$ can be regarded as an A-algebra. Thus we can form the tensor product:

$$\widehat{A} \otimes_A M. \tag{6}$$

We already know that

$$\widehat{A} \otimes_{\widehat{A}} \widehat{M} \cong \widehat{M}$$

because $\widehat{M}$ is clearly a $\widehat{A}$-module. So how does (6) compare with $\widehat{M}$? The existence of the completion homomorphism Σ and the fact that $\widehat{A}$ is an A-algebra, imply the existence of the natural composite map

$$\widehat{A} \otimes_A M \xrightarrow{1 \otimes \Sigma} \widehat{A} \otimes_A \widehat{M} \longrightarrow \widehat{M}$$

we call τ. Our intuition would suggest that τ should be the I-completion of M, and for noetherian rings this is true:

Proposition 6. If M is a finitely generated A-module over a noetherian ring, then τ is an isomorphism, that is:

$$\widehat{A} \otimes_A M \cong \widehat{M}.$$

Proof. First we observe that the case $M = A$ follows trivially. For the general case we first note that completion commutes with finite direct sums because if $M = M_1 \oplus M_2$, then the short exact sequence $0rightarrowM_1rightarrowMrightarrowM_2rightarrow0$ is converted, via completion, into the short exact sequence:

$$0 \longrightarrow \widehat{M_1} \longrightarrow \widehat{M} \longrightarrow \widehat{M_2} \longrightarrow 0.$$

Since the original short exact sequence is split, then the above sequence is also split because inverse limits preserve direct sums, so that $\widehat{M} \cong \widehat{M_1} \oplus \widehat{M_2}$. By induction we conclude that completion commutes with finite direct sums.

Now let $FrightarrowGrightarrowMrightarrow0$ be a free presentation of M where F and G are free A-modules of finite dimension. Hence, if we apply the case $M = A$ to $F \cong A^N$ and $G \cong A^M$ we obtain:

$$\widehat{A} \otimes_A F \cong \widehat{A} \otimes_A A^N \cong (\widehat{A} \otimes_A A)^N \cong (\widehat{A})^N \cong \widehat{A^N} \cong \widehat{F}$$

With this in mind, lets apply τ to the short exact sequence $0rightarrowFrightarrowGrightarrowMrightarr$ so that we get the following commutative diagram:

$$\begin{array}{ccccccc}
\widehat{A} \otimes_A F & \longrightarrow & \widehat{A} \otimes_A G & \longrightarrow & \widehat{A} \otimes_A M & \longrightarrow & 0 \\
\| & & \| & & \downarrow \tau & & \\
\widehat{F} & \longrightarrow & \widehat{G} & \longrightarrow & \widehat{M} & \longrightarrow & 0.
\end{array}$$

where both left-most vertical arrows are isomorphisms. The top row is exact because tensoring is right exact. The bottom row is exact at because of Corollary 1. This means that τ is surjective because its composite with the previous homomorphism $\widehat{A} \otimes_A G \longrightarrow \widehat{A} \otimes_A M$ is equal to the composite $\widehat{A} \otimes_A G \longrightarrow \widehat{M}$ which is surjective. One thing is left:

To prove that τ is injective, set $x \in \ker(\tau)$ and consider the following path starting at x and ending at $x = 0$, derived from the diagram chasing method:

$$\begin{array}{ccccccccc}
 & & z' & \hookrightarrow & y & \twoheadrightarrow & x & \longrightarrow & 0 \\
 & & \| & & \| & & \downarrow \tau & & \\
0 & \longrightarrow & z & \hookrightarrow & y' & \longrightarrow & 0 & \longrightarrow & 0.
\end{array}$$

□

The preceding proposition, together with Corollary 1, tells us that the functor

$$M \longmapsto \widehat{A} \otimes_A A$$

on the category of finitely generated A-modules is an exact functor or in other words: $\widehat{A}$ is a flat A-module.

3 Hensel's Lemma

Before stating and proving Hensel's Lemma, we will apply our previous results to local noetherian rings $(A, \mathfrak{m}, k)$. We want to find out what can be said about the $\mathfrak{m}$-completion of a local noetherian ring.

Since A is noetherian, $\mathfrak{m}$ is finitely generated so that Proposition 6 applies and

$$\widehat{A} \otimes_A \mathfrak{m} \cong \widehat{\mathfrak{m}}$$

is an isomorphism whose image in $\hat{A}$ is $\mathfrak{m}\hat{A}$. Thus $\hat{\mathfrak{m}}$ is the extension of $\mathfrak{m}$ in $\hat{A}$ under the natural completion map $Arightarrow\hat{A}$. Thus Corollary 2 implies that:

$$k = \frac{A}{\mathfrak{m}} \cong \frac{\widehat{A}}{\widehat{\mathfrak{m}}}$$

so that $\hat{\mathfrak{m}}$ is a maximal ideal of $\hat{A}$. It will turn out the $\widehat{\mathfrak{m}}$ is $\hat{A}$'s only maximal ideal.

To this end we will prove that $\hat{\mathfrak{m}}$ is contained in the Jacobson Radical of $\hat{A}$ which is equivalent to proving that $\hat{A}$ is a local ring with maximal ideal $\hat{\mathfrak{m}}$. We use the characterization [2, proposition 1.9, pg 6]:

$$x \in \mathrm{Jac}(\widehat{A}) \quad \iff \quad 1 - xy \text{ is a unit for all } y \in \widehat{A}.$$

Let $x \in \mathfrak{m}$ and $y \in \hat{A}$. By long division we have:

$$\frac{1}{1-xy} = 1 + xy + x^2y^2 + x^3y^3 + \cdots$$

thus the sequence $\{1+\cdots+x^ny^n\}$ is Cauchy in the $\mathfrak{m}$-topology because $x^n \in \mathfrak{m}^n$. Thus it converges in $\hat{A}$ to the inverse of $1 - xy$ so that it is a unit. Now we conclude $\mathfrak{m} \subseteq \mathrm{Jac}(\widehat{A})$. We have proven that $(\hat{A}, \hat{\mathfrak{m}}, k)$ is a local ring.

Now consider the natural map $Arightarrow\hat{A}$ and its kernel

$$H = \cap_{n=1}^{\infty} \mathfrak{m}^n$$

The subspace topology of H has only one open neighbourhood of 0, namely H. On the other hand, the Artin-Rees Lemma states that this subspace topology is the $\mathfrak{m}$-topology of H as an A-module. This means that the $\mathfrak{m}$-topology of H has only one open neighbourhood around 0, itself. However, $\mathfrak{m}H$ is also an open neighbourhood around 0 so that $\mathfrak{m}H = \mathfrak{m}$. Since A is local, $\mathfrak{m}$ is equal to the Jacobson radical of A so that Nakayama's Lemma guarantees that $H = 0$.

Putting this together we have the following proposition:

Proposition 7. If $(A, \mathfrak{m}, k)$ is a noetherian local ring, then the $\mathfrak{m}$-topology is Hausdorff and its $\mathfrak{m}$-completion $\hat{A}$ is also a (noetherian[1]) local ring with maximal ideal $\hat{\mathfrak{m}}$ and with the same residue field k. The $\hat{\mathfrak{m}}$-topology on $\hat{A}$ is also Hausdorff.

[1]The proof that $\hat{A}$ is again noetherian detours away from this work and it is not used in the proof of Hensel's Lemma. I mention it for the sake of completeness. A simple proof can be found in [3, Theorem 8.12 pg 61, pg]. Another proof with graded algebras is given in [2, Theorem 10.26, pg 113].

Now that we have studied local noetherian rings, we are ready to state Hensel's Lemma in full generality. Then we sketch out a path towards the proof through a series of lemmas. Finally, after we've proven Hensel's Lemma, we apply it to the various examples the illustrate the usefulness of the lemma.

Instead of beginning with Hensel's original result, we shall prove a more general result which exemplifies the important concepts behind Hensel's Lemma.

Theorem 3. *(Hensel's Lemma) Let $(A, \mathfrak{m}, k)$ be a noetherian local ring which is $\mathfrak{m}$-complete, also let $F \in A[x]$ be a polynomial of degree d. Also, lets consider the natural projection map:*

$$A[x] \xrightarrow{\pi} k[x] \qquad f \longmapsto \bar{f}$$

where $\bar{f}$ is obtained by reducing its coefficients mod $\mathfrak{m}$. Now suppose there are polynomials $g, h \in A[x]$, g monic, of degrees r and $d-r$ respectively such that:

$$\bar{f} = \bar{g}\bar{h} \in k[x]$$

is a coprime factorization, that is $\langle \bar{g}, \bar{h} \rangle = \langle 1 \rangle$. This coprime factorization can be 'lifted' to a non-trivial coprime factorization of F in $A[x]$ of the same degrees, that is there exist coprime polynomials $G, H \in A[x]$, G monic, of degrees r and $d-r$ respectively such that $F = GH$ with $\bar{G} = \bar{g}$ and $\bar{H} = \bar{h}$.

The proof will follow three steps:

1. We will establish a generalized division algorithm that will allow us to calculate any polynomial as a linear combination of $\bar{g}$ and $\bar{h}$. This tool will prove useful for approximating the factorization of F in the third step.

2. The next step is to guarantee that coprime polynomials in a ring $(B/I)[x]$ (for a certain class of ideals $I \leq B$) can indeed be lifted to coprime polynomials in $B[x]$. The idea is to begin lifting $g, h \in (A/\mathfrak{m})[x]$ to $(A/\mathfrak{m}^2)[x]$ and then to $(A/\mathfrak{m}^3)[x]$ and so on; thus 'approximating' the pair of coprime polynomials $G, H \in A[x]$ that will produce the desired factorization of F. We will call this method 'Coprime Lifting'.

3. Finally, with both previous steps we construct a sequence of non-trivial coprime factorizations of F in each successive $(A/\mathfrak{m}^n)[x]$ so that the 'limit' in $A[x]$ is again a non-trivial factorization of F; here the completeness of A will be essential so that the limit of the factors are again polynomials in $A[x]$.

Now that we have established the path of the proof, we begin with step 1:

Lemma 2. *(Generalized Division Algorithm) Let $h \in S[x]$ be any polynomial in an arbitrary polynomial ring and suppose there exist $g_1, g_2 \in S[x]$ such that $\langle g_1, g_2 \rangle = \langle 1 \rangle$ and also that g_1 is monic of degree $\deg(g_1) = r$. Then there exists a unique linear combination:*

$$h = h_1 g_1 + h_2 g_2$$

such that $\deg(h_2) < r$.

Proof. First observe that, since $\langle g_1, g_2\rangle = \langle 1\rangle$, then there is a linear combination $1 - \lambda g_2 = \mu g_1$ so that $g_2 + \langle g_1\rangle$ is a unit in $S[x]/\langle g_1\rangle$. Furthermore, since g_1 is monic, then $S[x]/\langle g_1\rangle$ is a finitely generated free S-module with basis $\{1 + \langle g_1\rangle, x + \langle g_1\rangle, \ldots, x^{r-1} + \langle g_1\rangle\}$. This means that $h + \langle g_1\rangle$ can be written as:

$$h + \langle g_1\rangle = \sum_{k=0}^{r-1}(s_k x^k + \langle g_1\rangle) = \left(\sum_{k=0}^{r-1} s_k x^k\right) + \langle g_1\rangle,$$

where $s_k \in S$. If we multiply each side by $1 + \langle g_1\rangle = g_2 \cdot (g_2)^{-1} + \langle g_1\rangle$ we get:

$$h + \langle g_1\rangle = g_2 \left(\sum_{k=0}^{r-1}(g_2^{-1})(x) s_k x^k\right) + \langle g_1\rangle.$$

The sum inside the parenthesis can be further reduced to a linear combination of the basis of $S[x]/\langle g_1\rangle$, so that it equals $h_2(x) = \sum s'_k x^k$ for some polynomial with $\deg(h_2) < r$. We therefore have:

$$h + \langle g_1\rangle = g_2 h_2 + \langle g_1\rangle \quad \Longrightarrow \quad h - g_2 h_2 \in \langle g_1\rangle,$$

and the linear combination $h = h_1 g_1 + h_2 g_2$ follows.

We have proven the existence of such linear combinations, so now we prove uniqueness. Suppose

$$h_1 g_1 + h_2 g_2 = h = h'_1 g_1 + h'_2 g_2. \tag{7}$$

with $\deg(h_2), \deg(h'_2) < \deg(g_1) = r$.

This equality implies that $g_1(h'_1 - h_1) = g_2(h_2 - h'_2)$, or in other words, $g_2(h_2 - h'_2) + \langle g_1\rangle = 0$, but since g_2 is a unit mod $\langle g_1\rangle$ we conclude that

$$h_2 - h'_2 + \langle g_1\rangle = 0 \quad \Longrightarrow \quad \mu g_1 = h_2 - h'_2 \text{ for some } \mu \in S[x].$$

Recall that g_1, being monic, is not a zero divisor in $B[x]$ so that, if $\mu \neq 0$, then $\mu g_1 = h_2 - h'_2 \neq 0$ and thus $\deg(\mu g_1) \geq \deg(g_1) = r$. However $\deg(h_2 - h'_2) \leq \max\{\deg(h_2), \deg(h'_2)\} < r$ since $\deg(h_2), \deg(h'_2) < r$ by hypothesis; but this is a contradiction so that $\mu = 0$ and we conclude that $h_2 = h'_2$. To finish the proof, we substitute this last equality in (7) to obtain: $g_1(h'_1 - h_1) = 0$ and because g_1 is monic, we may cancel it out to conclude that $h_i = h'_i$ for $i = 0, 1$. □

Remark. Lemma 2 is called the Generalized Division Algorithm because it reduces to the usual division algorithm if we set $g_2 = 1$.

Lemma 3. *(Coprime Lifting) Let $S = R/I$ for some ideal $I \leq R$ contained in Jac(R) and set $g_1, g_2 \in S[x]$ as before, i.e. $\langle g_1, g_2\rangle = \langle 1\rangle$ with g_1 monic of degree r. If G_1 and G_2 are any preimages of g_1 and g_2 in $R[x]$ via the natural projection $\pi : R[x] \twoheadrightarrow S[x]$ such that G_1 is monic, then G_1 and G_2 are also coprime. In symbols we have*

$$\langle \pi(G_1), \pi(G_2)\rangle = \langle g_1, g_2\rangle = \langle 1\rangle \text{ and } G_1 \text{ is monic} \quad \Longrightarrow \quad \langle G_1, G_2\rangle = \langle 1\rangle = R[x].$$

Proof. We will prove that no maximal ideal $\mathfrak{m} \leq R[x]$ contains both G_1 and G_2. So let $\mathfrak{m} \leq R[x]$ be a maximal ideal such that $G_1 \in \mathfrak{m}$. Now, the ring $R[x]/\langle G_1 \rangle$ is an integral extension of R via the composite map $RrightarrowR[x]rightarrowR[x]/\langle G_1 \rangle$; this is because G_1 is monic and (as before) $R[x]/\langle G_1 \rangle$ is a finitely generated R-module, so that it is generated by finitely many integral elements [4, Corollary 4.5, pg 122].

In an integral extension, maximal ideals contract to maximal ideals [2, Corollary 5.8, pg 61]. Therefore the maximal ideal $\mathfrak{m}/\langle G_1 \rangle$ contracts to a maximal ideal $\mathfrak{n} = R \cap \mathfrak{m}$ of R. Since I is contained in the Jacobson radical of R, in particular it is also contained in $\mathfrak{n}$. With this in mind, we extended $\mathfrak{n}$ to $R[x]$ via the natural inclusion $RrightarrowR[x]$ so that

$$IR[x] = I[x] \subseteq \mathfrak{n}R[x] \subseteq \mathfrak{m}.$$

Thus we may safely extend $\mathfrak{m}$ to $S[x] \cong R[x]/I[x]$ as $\mathfrak{m}/I[x]$ which again is maximal. Furthermore, we know that $g_1 = G_1 + I[x] \in \mathfrak{m}/I[x]$.

Now, if $G_2 \in \mathfrak{m}$ we would also have $g_2 = G_2 + I[x] \in \mathfrak{m}/I[x]$, but this is a contradiction because $\langle g_1, g_2 \rangle = S[x]$ which would imply that $\mathfrak{m}/I[x] = R[x]/I[x]$ so that $\mathfrak{m} = R$! Therefore $G_2 \notin \mathfrak{m}$ and we're done. □

Lemma 4. *(Approximations of Factorizations) Let $(A, \mathfrak{m}, k)$ be a noetherian local ring and $F \in A[x]$ any polynomial of degree d. Suppose there are polynomials $g, h \in A[x]$ of degrees r and $d - r$ respectively, with g monic such that: $\bar{f} = \bar{g}\bar{h} \in k[x]$ where $\langle \bar{g}, \bar{h} \rangle = \langle 1 \rangle$. Then there exists two sequences of polynomials $\{g_n\}$ and $\{h_n\}$ in $A[x]$, each bounded by degree:* $\deg(g_n) \leq r$ *and* $\deg(h_n) \leq d - r$*, such that:*

$$F - g_n h_n \in \mathfrak{m}^n A[x] \quad \textit{and} \quad \widehat{g_n} = \widehat{g}, \ \widehat{h_n} = \widehat{h}$$

for all $n \in \mathbb{N}$, that is, the sequence of approximations $\{g_n h_n\}$ approaches F in the $\mathfrak{m}[x]$-topology, but still induces the same factorization in $k[x]$.

Proof. Starting with the given polynomials $g, h \in A[x]$ that induce the factorization $\bar{f} = \bar{g}\bar{h}$ in $k[x]$ we will construct inductively a sequence of polynomials $\{g_n\}$

$$g_1 = g, \quad g_2 = g + \varepsilon_1, \quad g_3 = g + \varepsilon_1 + \varepsilon_2, \quad g_4 = g + \varepsilon_1 + \varepsilon_2 + \varepsilon_3, \quad \dots$$

and similarly for $\{h_n\}$:

$$h_1 = h, \quad h_2 = h + \delta_1, \quad h_3 = h + \delta_1 + \delta_2, \quad h_4 = g + \delta_1 + \delta_2 + \delta_3, \quad \dots$$

Such that each product $g_n h_n$ is a better approximation of F, more precisely, we want:

$$F - g_n h_n \in \mathfrak{m}^n A[x].$$

We also want $\varepsilon_n, \delta_n \in \mathfrak{m}^n A[x]$ (i.e. they are small polynomials in the $\mathfrak{m}$-adic topology) and want them of bounded degree: $\deg(\varepsilon_n) < d - r$ and $\deg(\delta_n) < r$ so that $\deg(g_n) = \deg(g) = r$ and $\deg(h_n) = \deg(h) = d - r$.

We begin with $n = 1$ so that $g_1 = g$ and $h_1 = h$. By hypothesis:

$$\bar{F} - \bar{g_1}\bar{h_1} = \bar{F} - \bar{g}\bar{h} = 0 \in k[x],$$

and thus $F - g_1 h_1 \in \mathfrak{m}A[x]$.

Now suppose that we have constructed the nth term of the sequences:

$$g_n = g + \varepsilon_1 + \cdots \varepsilon_{n-1} \quad \text{and} \quad h_n = h + \delta_1 + \cdots + \delta_{n-1}$$

with the desired properties. That is:

$$F - g_n h_n \in \mathfrak{m}^n A[x] \tag{8}$$

$$\varepsilon_k, \delta_k \in \mathfrak{m}^k A[x] \tag{9}$$

$$\deg(\varepsilon_k) < d - r \text{ and } \deg(\delta_k) < r \tag{10}$$

Observe that we have left out the necessary property: $g_k \equiv \bar{g} \mod \mathfrak{m}$ because this is implied by always holds by construction of the sequences $\{g_n\}$ and $\{h_n\}$.

Since each $\varepsilon_n \in \mathfrak{m}^n A[x]$, then they all reduce to 0 modulo $\mathfrak{m}$, so that g_n reduces to g, ie. $\bar{g}_n = \bar{g}$; similarly for h_n. Since $\bar{g}$ and $\bar{h}$ are coprime in $k[x]$ we can use the Coprime Lifting Lemma to lift this factorization to $(A/\mathfrak{m}^{n+1})[x]$ (by setting $R = A/\mathfrak{m}^{n+1}$ and $I = \mathfrak{m}/\mathfrak{m}^{n+1}$ in Lemma 3).

Therefore the preimages of g_n and h_n in $(A/\mathfrak{m}^{n+1})[x]$ are coprime. We can now use the Generalized Division Algorithm in $(A/\mathfrak{m}^{n+1})[x]$ to express the error $\overline{F - g_n h_n}$ in $\mathfrak{m}^{n+1}[x]$. There exist $\varepsilon_n, \delta_n \in A[x]$ with $\deg(\varepsilon_n) < \deg(g_n) = \deg(g) = r$ that project onto the unique linear combination:

$$F - g_n h_n \equiv \varepsilon_n h_n + \delta_n g_n \mod \mathfrak{m}^{n+1}, \tag{11}$$

guaranteed by Lemma 2.

Furthermore, $\deg(\delta_n) < \deg(h_n) = \deg(h) = d - r$ because otherwise both sides of the above equation would have different degrees (as polynomials in $A[x]$); thus the new polynomials ε_n and δ_n satisfy the bound for their degrees (property (10)).

By the inductive hypothesis we also have $F - g_n h_n \in \mathfrak{m}^n A[x]$ (property (8)). Then, if we project the equation (11) over $\mathfrak{m}^n/\mathfrak{m}^{n+1}$, then the left hand side reduces to zero and we obtain:

$$0 \equiv \varepsilon_n g_n + \delta_n h_n \mod \mathfrak{m}^n.$$

Since coprimality is preserved by projection onto ideals, the projection of the preimages of g_n and h_n onto $(A/\mathfrak{m}^n)[x]$ is also coprime. Thus, by Lemma 2, we may write $0 \in (A/\mathfrak{m}^n)[x]$ as the trivial linear combination (each coefficient is zero). By uniqueness the above equation and the trivial linear combination must be equal modulo $\mathfrak{m}^n$. Thus we conclude that

$$\varepsilon_n \equiv 0 \equiv \delta_n \mod \mathfrak{m}^n \quad \Longrightarrow \quad \varepsilon_n, \delta_n \in \mathfrak{m}^n A[x],$$

so that ε_n and δ_n also satisfy property (9)

With ε_n and δ_n we can define the next polynomials in our sequences:

$$g_{n+1} = g + \varepsilon_1 + \cdots + \varepsilon_n = g_n + \varepsilon_n \quad \text{and} \quad h_{n+1} = h + \delta_1 + \cdots + \delta_n = h_n + \delta_n.$$

To finish the proof by induction, we need only prove property (8). Indeed:

$$F - g_{n+1}h_{n+1} = F - (g_n + \varepsilon_n)(h_n + \delta_n) = (F - g_n h_n - g_n\delta_n - h_n\varepsilon_n) - \varepsilon_n\delta_n.$$

By equation (11), the sum inside the parenthesis is an element of $\mathfrak{m}^{n+1}A[x]$. Furthermore, since $\varepsilon_n, \delta_n \in \mathfrak{m}^n A[x]$ then their product will certainly be an element of $\mathfrak{m}^{n+1}$. We may therefore conclude that $F - g_{n+1}h_{n+1} \in \mathfrak{m}^{n+1}$ and with this the induction is over. □

We are now ready to prove Hensel's Lemma:

Proof. (Hensel's Lemma) First we must observe that while A is $\mathfrak{m}$-complete, it is not necessarily true that $A[x]$ is $\mathfrak{m}[x]$-complete. However not all is last.

Let

$$\{q_n(x)\} = \left\{ \sum_{j=0}^{d} a_{n,j}x^j \right\}$$

be a Cauchy sequence in $A[x]$ where every element has degree bounded by d. Observe that this sequence induces $d+1$ sequences in A, one for each coefficient:

$$\{a_{n,j}\}_{n\in\mathbb{N}} \subseteq A.$$

Since $\{q_n\}$ is Cauchy in the $\mathfrak{m}[x]$-topology, for all $N \in \mathbb{N}$ then $q_s - q_t \in \mathfrak{m}^N[x]$ for all sufficiently large s, t. More precisely, if we fix $N \in \mathbb{N}$ and write:

$$q_s(x) - q_t(x) = \sum_{j=0}^{d} (a_{s,j} - a_{t,j})x^j.$$

we immediately see that $q_s - q_t \in \mathfrak{m}^N[x]$ for sufficiently large s, t is equivalent to $(a_{s,j} - a_{t,j}) \in \mathfrak{m}^N$ for all $j = 0, 1, \ldots, d$ and for all sufficiently large s, t. Thus each sequence $\{a_{n,j}\}_{n\in\mathbb{N}}$ is Cauchy in A.

By hypothesis, each sequence converges:

$$\lim_{n\to\infty} a_{n,j} = a_j$$

for some $a_0, a_1, \ldots, a_d \in A$ which then implies

$$\{a_{n,d}x^d + \cdots + a_{n,0}\} \longrightarrow a_d x^d + \cdots + a_0. \tag{12}$$

We've proven that any Cauchy sequence in $A[x]$ of bounded degree converges provided A is $\mathfrak{m}$-complete. This is why the upper bound for the degree in Lemma 4 is necessary. We continue with the proof:

Let $\{g_n\}$ and $\{h_n\} \subseteq A[x]$ be the sequences of polynomials given by the previous lemma. Since $F - g_n h_n \in \mathfrak{m}^n A[x]$ for all n, then clearly:

$$0 = \lim_{n\to\infty}(F - g_n h_n) = F - \left(\lim_{n\to\infty} g_n\right)\left(\lim_{n\to\infty} h_n\right) \tag{13}$$

in the $\mathfrak{m}[x]$-topology. On the other hand we know that:

$$g_n = g + \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_{n-1}$$

where each summand $\varepsilon_j \in \mathfrak{m}^j A[x]$. Thus by construction, g_n is a Cauchy sequence in $A[x]$ because for sufficiently large $s > t$:

$$g_s - g_t = \varepsilon_t + \cdots + \varepsilon_{s-1} \in \mathfrak{m}^t[x] + \cdots + m^{s-1}[x] \subseteq \mathfrak{m}^t[x] \subseteq \mathfrak{m}^n[x]$$

for all $n \in \mathbb{N}$. By the previous lemma we also have that this sequence is of bounded degree, so that g_n converges to some $G \in A[x]$ of degree $\leq r$. Also, since each g_n was monic of degree r, then the limit is also monic of degree r; this follows from equation (12). On the other hand $\{\widehat{g}_n\} = \{\widehat{g}\} \subseteq k[x]$ by construction, so that the limit also projects onto $\widehat{g}$, that is $\widehat{G} = \widehat{g}$ in $k[x]$.

Similarly, $\{h_n\}$ converges to some $H \in A[x]$ of degree $d-r$ such that $\widehat{H} = \widehat{h}$. This means that $G, H \in k[x]$ satisfy the Lifting Lemma and thus are coprime in $A[x]$. Finally we substitute the limits in (13):

$$F = GH.$$

We conclude that $\lim g_n = G$ and $\lim h_n = H$ satisfy Hensel's Lemma. □

4 Applications

Theorem 3 is not Hensel's original statement of his Lemma. He began working on the p-adic integers, i.e. the $p\mathbb{Z}$-completion of $\mathbb{Z}$, in the 1890's (for example [5]). And unified is work on the p-adic integers and his lemma in 1908 with his book *Theorie Der Algebraischen Zhalen* [6, Chapter 4]. Throughout his life he published many versions of his lemma which makes it hard to pin point the original statement, but one of the original motives for his work on the p-adic integers was to solve polynomial equations mod p^n for some prime power.

Hensel realized that a root to a polynomial $f(x)$ with integer coefficients could be approximated by the roots of $f(x) \mod p^n$ for higher and higher powers of a suitable prime p. To this end, he would lift simple roots of $f \mod p^n$ to simple roots of $f \mod p^{n+1}$, thus creating a sequence of numbers that approached the desired root.

To rigorously define 'approach' he introduced the p-adic measure in $\mathbb{Q}$ and began using the powerful tool of Analysis in Number Theory. In 1913, Hensel published a text book on Number Theory where he carefully lays out his theory of the p-adic integers [7]. The p-adic measure induces the p-topology on $\mathbb{Q}$ as a $\mathbb{Z}$-module, so that we have simply generalized his ideals to Commutative Algebra and Topology.

Once his p-adic measure was established, the sequence of lifted simple roots did indeed converge to the desired root. With our definition of the p-topology, this is deduced from Lemma 3. Hensel's statement of this fact can be distilled as follows:

If $f \in \mathbb{Z}[x]$ and $a \in \mathbb{Z}$ is a simple root of $f \mod p^n$, that is

$$f(a) \equiv 0 \mod p^n \quad \text{and} \quad f'(a) \not\equiv 0 \mod p,$$

then there exists an integer b such that $f(b) \equiv 0 \mod p^{n+1}$ and $a \equiv b \mod p^n$.

A short an elemental proof of this statement can be given as follows: Let

$$f(a + p^n x) = f(a) + f'(a)p^n x + \frac{f'(a)}{2!}p^{2n}x^2 + \cdots$$

be the Taylor series of f at $a + p^n x$. If we reduce this formula $\mod p^{n+1}$, then all terms after the second one vanish so that:

$$f(a + p^n x) \equiv f(a) + f'(a)p^n x \mod p^{n+1}.$$

Let $b = a + p^n x$. Since clearly $b \equiv a \mod p^n$, we need only solve for x to calculate the lifted root b. Set $f(a + p^n x) \equiv 0 \mod p^{n+1}$. Recall that, since $f'(a) \not\equiv 0 \mod p$, then $f'(a)$ is a unit in $\mathbb{Z}/p\mathbb{Z}$ and thus it is a unit in $\mathbb{Z}/p^n\mathbb{Z}$. This way we can easily solve for $p^n x$:

$$0 \equiv f(a+p^n x) \equiv f(a)+f'(a)p^n x \mod p^{n+1} \quad \Longrightarrow \quad p^n x \equiv -\frac{f(a)}{f'(a)} \mod p^{n+1}.$$

This means

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

is a simple root of $f \mod p^{n+1}$ obtained by lifting a to a better approximation in the p-adic topology.

The above formula is identical to Newton's Method for approximating roots (polynomials in this case) and the proof is identical; just switch $\mod p^n$ to 'vanishing to order n'. The idea behind these methods is simply being able to lift simple roots. We now state this result in our context of complete local rings:

Corollary 4. *Let $(A, \mathfrak{m}, k)$ be an $\mathfrak{m}$-complete local noetherian ring. Let $f \in A[x]$ be a polynomial such that $\bar{f} \in k[x]$ has a simple root $\zeta \in k$, or equivalently $\bar{f}'(\zeta) \neq 0$, such that $\bar{f}(\zeta) = 0$. Then there exists a simple root $\xi \in A$ of f such that $\bar{\xi} = \zeta$. In other words, it is possible to lift the simple root $\zeta \in k$ of $\bar{f}$ to a simple root $\xi \in A$ of f.*

Proof. The proof is as follows, a simple root of $\bar{f} \in k[x]$ will induce a factorization of the type $\bar{f} = (x-\zeta)q(x)$ for some $q \in k[x]$ such that $\langle x - \zeta, q(x)\rangle = k[x]$. We will then be able to use Hensel's Lemma to lift the factorization in $k[x]$ to a coprime factorization $f(x) = (x - \xi)Q(x)$; this gives ξ as a simple root of f that projects onto the simple root ζ of $\bar{f}$.

Suppose $\zeta \in k$ is a simple root of $\bar{f} \in k[x]$. By the division algorithm we may right $\bar{f}(x) = (x-\zeta)q(x)$ where $q(\zeta) \neq 0$, or equivalently: the polynomial $(x-\zeta)$ does not divide $q(x)$. Since $x - \zeta$ is irreducible and $k[x]$ is a Principal Ideal Domain, the ideal $I = \langle x - \zeta \rangle$ is maximal and since $(x - \zeta) \nmid q(x)$ then I does not contain q. Thus they both generate the entire ring, i.e. they're coprime.

We may now use Hensel's Lemma to lift the factorization of $\bar{f}$ to f: there exist coprime $P, Q \in A[x]$ of the same degrees as $x - \zeta$ and q respectively, P monic, such that $f = PQ$, $\bar{P} = (x-\zeta)$ and $\bar{Q} = q$. This means that $P(x) = x-\xi$ for some $\xi \in A$ such that $\bar{\xi} = \zeta$. Since P and Q are coprime, $P \nmid Q$ so that $Q(\xi) \neq 0$. We may now conclude that $f = PQ = (x - \xi)Q(x)$ so that ξ is a simple root of f such that $\bar{\xi} = \zeta$. □

A direct consequence of this lemma is the Implicit Function Theorem. We will prove the Implicit Function Theorem for polynomials in two variables. The proof of the general Implicit Function Theorem is a simple generalization that requires a deeper study of power series rings in multiple variables so we omit the proof (see for example [3, Chapter 3, Section 8] or [4, Theorem 7.16 and Corollary 7.17]).

Corollary 5. *Let $f \in k[x, y]$ for some field k. If $\partial f/\partial x \neq 0$ at $(0, a)$, then there is a unique power series*

$$y(x) = a + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \in k[[x]]$$

such that $f(x, y(x)) = 0$ identically.

Proof. Let $A = k[[x]]$ and $\mathfrak{m} = \langle x \rangle$ its maximal ideal. We view $f \in k[x, y]$ as a polynomial in y with coefficients in $k[[x]]$, that is $f \in A[y]$. Similarly to the previous corollary, the hypothesis

$$\frac{\partial f}{\partial x}(0, a) \neq 0$$

implies that $a \in k$ is a simple root of the polynomial $F(y) = f(0, y) \in k[y] = (A/\mathfrak{m})[y]$; observe that making $x = 0$ is the same as reducing $\bmod \mathfrak{m}$ in A so that $\bar{f} = f(0, y) = F(y)$. Thus by Corollary 4 there exists a power series $g(x) \in A$ that is a simple root of the polynomial $f \in A[y]$, that is:

$$f(x, g(x)) = 0.$$

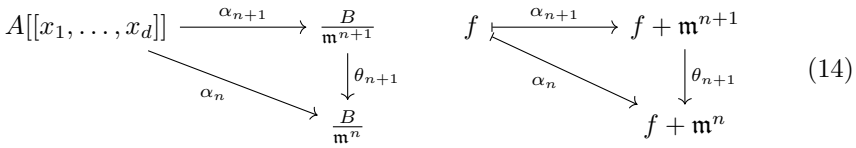
Finally, observe that $g(x)$ reduces to $a \bmod \mathfrak{m}$ (that is evaluating at $x = 0$), thus its constant term is a as desired. □

Now we turn our attention to formal power series. If A is a ring, then the ring of formal power series in d variables with coefficients in A is denoted as:

$$A[[x_1, \ldots, x_d]].$$

This ring is closely related to complete rings. Let $B = A[x_1, \ldots, x_d]$ and $\mathfrak{m} = \langle x_1, \ldots, x_n \rangle$ and give B the $\mathfrak{m}$-topology. The $\mathfrak{m}$-completion of B is exactly $A[[x_1, \ldots, x_d]]$.

To prove this, first observe that the ring homomorphisms $\{\alpha_n\}$ defined as:



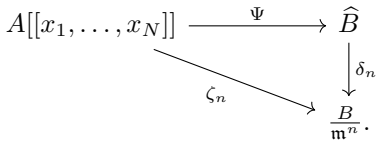
(14)

Form an inverse system that satisfies the universal property of the inverse limit. The proof of this fact is very similar to the discussion preceding Definition 3. Thus:

$$A[[x_1,\ldots,x_d]] \cong \varprojlim \frac{A}{\mathfrak{m}^n} \cong \widehat{B} \tag{15}$$

We also verify this with an explicit isomorphism Ψ: again consider the family of maps $\{\alpha_n\}$ which can be thought of as forgetting the terms of order greater than n of a power series f. Together with the natural projection maps θ_n (cf. Section 2), we have the commutative diagram (14).

This makes $A[[x_1,\ldots,x_N]]$ and the family $\{\alpha_n\}$ compatible with the inverse system $\{B/\mathfrak{m}^n\}$ so that the universal property of the inverse limit implies the existence of a unique homomorphism $A[[x_1,\ldots,x_N]] \longrightarrow \widehat{B}$ that commutes with the inverse system:



where δ_n are the inverse system homomorphisms of $\varprojlim B/\mathfrak{m}_n$ (cf. Section 2). Thus Ψ acts as follows:

$$f \mapsto (f+\mathfrak{m}, f+\mathfrak{m}^2, f+\mathfrak{m}^3, \ldots) \in \prod \frac{B}{\mathfrak{m}^n}.$$

Clearly $f+\mathfrak{m}^{n+1}$ projects onto $f+\mathfrak{m}^n$ so that the image is indeed an element of the inverse limit if we consider it via the characterization of the inverse limit as a subring of the product.

The inverse map is given by:

$$(f_1+\mathfrak{m}, f_2+\mathfrak{m}^2, f_3+\mathfrak{m}^3, \ldots) \mapsto f_1+(f_2-f_1)+(f_3-f_2)+\cdots \in A[[x_1,\ldots,x_N]].$$

the entries of the vector $(f_1+\mathfrak{m}, f_2+\mathfrak{m}^2, f_3+\mathfrak{m}^3, \ldots)$ satisfy $f_{n+1} \equiv f_n \mod \mathfrak{m}^n$ so that the nth summand is of order at least n. The choice of f_n for $f_n+\mathfrak{m}^n$ does not influence the function because each summand $f_{n+1}-f_n$ is constant in $B/\mathfrak{m}^n$ and thus does not affect the image. We conclude that Ψ is an isomorphism and equation (15) is true.

Now lets consider the case where $(A,\mathfrak{m},K)$ is a local ring that contains some field k. Since each element of k is a unit in A, we have $\mathfrak{m}\cap k = 0$ so that under the projection $A \twoheadrightarrow A/\mathfrak{m} = K$, the field k extends to itself:

$$k^e = \frac{k+\mathfrak{m}}{\mathfrak{m}} \cong \frac{k}{\mathfrak{m}\cap k} = \frac{k}{0} = k.$$

By these means we can associate any field contained in local ring A with a subfield of its residue field. We define a very special type of these contained fields.

Definition 4. Let $(A, \mathfrak{m}, K)$ be a local ring. If a field $k \subseteq A$ whose associated subfield is all of K, that is k extends isomorphically to $A/\mathfrak{m}$, is called a *coefficient field* of A.

Coefficient fields are central to the study of complete local rings, evidenced by the following proposition:

Proposition 8. Let $(A, \mathfrak{m}, K)$ be an $\mathfrak{m}$-complete noetherian local ring and set $\mathfrak{m} = \langle a_1, \ldots, a_d \rangle$. If A contains a coefficient field k, then every element of A can be expressed as a power series with coefficients in k with d variables, our equivalently, there exists a surjective homomorphism $k[[x_1, \ldots, x_d]] \twoheadrightarrow A$ sending $f(x_1, \ldots, x_d) \mapsto f(a_1, \ldots, a_d)$. We can also state it as: every complete local noetherian ring which contains a coefficient field is a homomorphic image of some power series ring.

Proof. First lets assume that for every $x \in \mathfrak{m}^n$ there exists a representation $x = y_n + z_{n+1}$ where y_n is a homogeneous expression of degree n in $a_1, \ldots, a_n$ (the generators of $\mathfrak{m}$) with coefficients in k and $z_{n+1} \in \mathfrak{m}^{n+1}$. Using these representations, for all elements $x \in A$, we will be able to construct a Cauchy sequence of partial sums that converges to x thus giving us a power series representation for x in $a_1, \ldots, a_d$.

Indeed, if $x \in A = \mathfrak{m}^0$, there exist y_0, homogeneous of degree 0 that is $y_0 \in k$ and $z_1 \in \mathfrak{m}$, such that $x = y_0 + z_1$. Now apply this representation to z_1: $z_1 = y_1 + z_2$ with y_1 homogeneous of degree 1 in $a_1, \ldots, a_d$ and $z_2 \in \mathfrak{m}^2$. Repeating this argument, we can construct the following sequence:

$$\begin{aligned} x &= y_0 + z_1 =: w_1 \\ x &= y_0 + y_1 + z_2 =: w_2 \\ &\vdots \\ x &= y_0 + y_1 + \cdots + y_{n-1} + z_n =: w_n \end{aligned}$$

where each y_j is homogeneous of degree j in the generators of $\mathfrak{m}$ and $z_n \in \mathfrak{m}^n$. By construction, the sequence $\{w_n\}$ is Cauchy so that the series:

$$x' = \lim_{n \to \infty} w_n = \sum_{n=1}^{\infty} y_n$$

converges in A because the partial sums $x - z_n \in x + \mathfrak{m}^n$ approach x in the $\mathfrak{m}$-topology. Since A is $\mathfrak{m}$-complete $\sum y_n$ converges to x. Also since $\{w_n\}$ is the constant sequence $\{x\}$ we conclude that $x = y_0 + y_1 + \cdots$ as a power series in $k[[a_1, \ldots, a_d]]$ because each y_j is homogeneous of degree j in the variables $a_1, \ldots, a_d$ by construction.

The only thing left to prove is that for every $x \in \mathfrak{m}^n$, there exists a representation $x = y_n + z_{n+1}$ with the desired properties. Indeed, since $\mathfrak{m} = \langle a_1, \ldots, a_d \rangle$, then x is a finite A-linear combination of homogeneous elements of degree n, that is:

$$x = \sum \mu_e a_1^{e_1} \cdots a_d^{e_d}$$

with $n = e_1 + \cdots + e_d$.

Lets consider a general monomial of this expression i.e. $\mu_e a_1^{e_1} \cdots a_d^{e_d}$. If μ_e is an element of $\mathfrak{m}$, then the entire monomial is an element of $\mathfrak{m}^{n+1}$ because $a_1^{e_1} \cdots a_d^{e_d} \in \mathfrak{m}^n$. Thus every monomial with coefficient in $\mathfrak{m}$ is gathered into the summand z_{n+1}.

Now suppose that $\mu_e \notin \mathfrak{m}$. Since A is local, μ_e is a unit and its image, $\bar{\mu}_e$ under the projection $A \twoheadrightarrow K$ it is non-zero. Therefore there exists a non-zero element $\lambda_e \in k$ that projects onto $\bar{\mu}_e$, that is $\bar{\lambda}_e = \bar{\mu}_e$. This is possible because k projects isomorphically onto K. In other words: there exists an $m_e \in \mathfrak{m}$ such that $\mu_e = \lambda_e + m_e$. Thus:

$$\mu_e a_1^{e_1} \cdots a_d^{e_d} = \lambda_e a_1^{e_1} \cdots a_d^{e_d} + m_e a_1^{e_1} \cdots a_d^{e_d}$$

where the first summand is a homogeneous expression with coefficient in k in $a_1, \ldots, a_d$ and the second is an element of $\mathfrak{m}^{n+1}$ by the previous discussion.

We have proven that each monomial of x can be separated into a sum of elements in $\mathfrak{m}^{n+1}$, i.e. z_{n+1}, and a sum of homogeneous elements of $k[a_1, \ldots, a_d]$, i.e. y_n. Therefore $x = y_n + z_{n+1}$ and we are done.

□

This is the preliminary result used to prove the Cohen Structure Theorem which was presented by I.S. Cohen in 1942 and published in 1946 [1]. This theorem classifies complete local noetherian rings as homomorphic images of power series. We prove the important case when A contains a field of characteristic 0 because it is a direct consequence of Hensel's Lemma (or more precisely Corollary 4). When A does not contain a field, or if the field it contains has characteristic p, then the theorem requires much more machinery than Hensel's Lemma. In Cohen's article, this occupies all of Part II [1, pg 70-85].

Theorem 6. *(Cohen Structure Theorem) Let $(A, \mathfrak{m}, K)$ be a complete local noetherian ring that contains a field of characteristic 0. Then A contains a coefficient field so that it is a homomorphic image of the power series with coefficients in the residue field K with a finite amount of variables or equivalently:*

$$A \cong \frac{K[x_1, \ldots, x_d]}{I}$$

for some ideal I of the power series ring.

Proof. Suppose A contains a field k of characteristic zero, so that the set:

$$\Omega := \{L \subseteq A \mid L \text{ is a field and } k \subseteq L\}$$

is none empty. Given a chain $L_1 \subseteq L_2 \subseteq \cdots$ in Ω we may form $L = L_1 \cup L_2 \cup \cdots$ which is again a subfield of A which contains k. Thus we may apply Zorn's Lemma and guarantee the existence of a maximal element $F \in \Omega$. We prove that F is a coefficient field for A.

Suppose not, i.e. F projects onto a proper subfield of K under the map $A \twoheadrightarrow A/\mathfrak{m}$. We denote by $\bar{F}$ for the image of F in K and let $\alpha \in K - \bar{F}$. We prove that α can't be algebraic nor non-algebraic, thus a contradiction.

1. (α is algebraic over $\bar{F}$) Let $\bar{p}(x) \in \bar{F}[x]$ be its minimal polynomial where $p \in F[x]$ is some preimage of this minimal polynomial. Since $\mathrm{char}(F) = \mathrm{char}(\bar{F}) = 0$, then α is a simple root so that we can lift it to a simple root $a \in A$ such that $\bar{a} = \alpha$ and $p(a) = 0$. Since $\bar{p} \in \bar{F}[x]$ is irreducible, it is a prime element so that its preimage $p \in F[x] \subseteq A[x]$ is also prime and therefore (since $F[x]$ is a Principal Ideal Domain) it generates a maximal ideal. Thus
$$F \subsetneq F(a) \cong \frac{F[x]}{\langle p \rangle} \subseteq A$$
which contradicts the maximality of F.

2. (α is not algebraic) In this case take $a \in A$ to be any preimage of $\alpha \in K$, that is $\bar{a} = \alpha$. If $a \in A$ were algebraic over F, then it would be satisfy some polynomial equation $f(a) = 0$ for $f \in F[x]$, but by projecting it onto K we would obtain a polynomial equation $\bar{f}(\alpha) = 0$! Thus for all $f \in F[x]$ we have $\bar{f}(\alpha) \neq 0$, or in other words, $f(a) \notin \mathfrak{m}$ for all f. Since A is local, this means that $f(a)$ is a unit for all f. We can now conclude that:
$$F \subsetneq F(a) = \left\{ \frac{g(a)}{f(a)} : f, g \in F[x] \right\} \subseteq A$$
is naturally embedded into A and thus contradicts the maximality of F.

We must conclude that F is a coefficient field for A, thus Proposition 8 applies and A is a homomorphic image of $F[x_1, \ldots, x_d]$. $\square$

References

[1] Cohen I.S. On the structure and ideal theory of complete local rings. *Transactions of the American Mathematical Society*, 52, 1946.

[2] Atiyah M.F. and I.G. Macdonald. *Introduction to Commutative Algebra*. Addison-Wesley, 1969.

[3] Hideyuki Matsumura. *Commutative Ring Theory*. Cambridge University Press, 1986.

[4] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, 2004.

[5] Kurt Hensel. Über eine neue begründung der theorie der algebraischen zahlen. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, (87), 1897.

[6] Kurt Hensel. *Theorie Der Algebraischen Zhalen*. 1908, Berlin.

[7] Kurt Hensel. *Zhalentheorie*. 1913, Berlin.