MULTIVARIATE HENSEL’S LEMMA FOR COMPLETE RINGS

YIFEI ZHAO

CONTENTS

1. Completion ........................................... 1
2. The inverse function theorem ......................... 6
3. Proof of Hensel’s lemma and the implicit function theorem . 7
References .............................................. 8

In this set of notes, we prove that a complete ring satisfies the multivariate Hensel’s lemma (Theorem 1.11). The result can be seen as a formal version of the implicit function theorem. We make this connection precise in §3. The section on completion follows the Stacks project [4], since most textbooks on this subject assumes Noetherian property. Our exposition of Hensel’s lemma largely follows Eisenbud’s book [2, Chapter 7], although he only proves the single-variable version, remitting multivariate Hensel’s lemma to exercises.

1. Completion

1.1. Inverse limit definition. Let $R$ be a (commutative, unital) ring, and $a$ be an ideal of $R$. The $a$-adic completion of $R$ is defined as the $R$-algebra.

$$
\hat{R} = \lim_{\leftarrow} R/a^n.
$$

Let $M$ be an $R$-module. Then the $a$-adic completion of $M$ is the $\hat{R}$-module

$$
\hat{M} = \lim_{\leftarrow} M/a^n M.
$$

The collection of maps $M \to M/a^n M$ induces a map $M \to \hat{M}$. It fits into the following exact sequence:

$$
0 \to \bigcap_{n \geq 1} a^n M \to M \to \hat{M}.
$$

Lemma 1.1. Let $M_n, N_n, P_n \ (n \geq 1)$ be inverse systems of $R$-modules fitting into an exact sequence (of inverse systems):

$$
0 \to M_n \to N_n \to P_n \to 0.
$$

Then the sequence

$$
0 \to \lim M_n \to \lim N_n \to \lim P_n
$$

is exact. Furthermore, if $M_n$ is a surjective system, i.e. the maps $M_n \to M_{n-1}$ are all surjective, then the sequence

$$
0 \to \lim M_n \to \lim N_n \to \lim P_n \to 0
$$

is exact.

Proof. An element of $\lim M_n$ can be seen as a sequence $(x_n)$ where each $x_n \in M_n$, and the image of $x_n$ in $M_{n-1}$ is precisely $x_{n-1}$. Hence $\lim M_n$ is the kernel of the morphism

$$
\varphi_M : \prod_{n \geq 1} M_n \to \prod_{n \geq 1} M_n, \quad \varphi_M(x_1, x_2, \cdots) = (x_1 - \bar{x}_2, x_2 - \bar{x}_3, \cdots)
$$

where $\bar{x}_{n+1}$ denotes the image of $x_{n+1}$ in $M_n$. The snake lemma shows that

$$
0 \to \text{Ker}(\varphi_M) \to \text{Ker}(\varphi_N) \to \text{Ker}(\varphi_P) \to \text{Coker}(\varphi_M)
$$
is an exact sequence. Now, if $M_n$ is a surjective system, then $\varphi_M$ is surjective, so $\text{Coker}(\varphi_M) = 0$. \qed

1.2. Cauchy sequence definition. We can alternatively define $a$-adic completion via a topology on $R$. We set a base of neighborhoods of each $x$ by $x + a^k$. This is the $a$-adic topology on $R$; it makes $R$ into a topological ring.

**Definition 1.2.** The topological $a$-adic completion of $R$ is the set of Cauchy sequences $(x_1, x_2, \cdots)$ in $R$, i.e. sequences with elements in $R$ and

\[ \forall k \in \mathbb{N}, \exists N \in \mathbb{N}, \text{ such that } m, n \geq N \implies x_m - x_n \in a^k \]

modulo the equivalence relation: $(x_n) \sim (y_n)$ if

\[ \forall k \in \mathbb{N}, \exists N \in \mathbb{N}, \text{ such that } n \geq N \implies x_n - y_n \in a^k \]

We temporarily denote this quotient set by $\hat{R}$. It is naturally equipped with an $R$-algebra structure.

**Lemma 1.3.** There is an $R$-algebra isomorphism

\[ \hat{R} \cong \hat{R} \]

**Proof.** Given a Cauchy sequence $(x_1, x_2, \cdots) \in \hat{R}$ and any $k \in \mathbb{N}$, pick sufficiently large $N$ so that

\[ x_N = x_{N+1} = \cdots \mod a^k. \]

Send $(x_n)$ to the element in $R/a^k$ defined by $x_N$. This process defines a compatible system of maps $\tilde{R} \to R/a^k$, which induces a map $\hat{R} \to \tilde{R}$.

Conversely, an element in $\tilde{R}$ is given by a sequence $(\tilde{x}_1, \tilde{x}_2, \cdots)$ where each $\tilde{x}_k \in R/a^k$ and its image in $R/a^{k-1}$ is exactly $\tilde{x}_{k-1}$. Pick a lift $x_k \in R$ of each $\tilde{x}_k$. Then

\[ x_k = x_{k-1} \mod a^{k-1} \]

for each $k$. So $(x_1, x_2, \cdots)$ is a Cauchy sequence. This process induces a well-defined map $\hat{R} \to \tilde{R}$. We leave it to the reader to check that these two maps are mutual inverses. \qed

Similarly, $\hat{M}$ can also be defined via the $a$-adic topology which associates to each $x \in M$ a base of neighborhoods $x + a^kM$.

1.3. Complete rings and modules. $R$ (respectively $M$) is complete with respect to $a$ if the natural map

\[ R \to \hat{R} \quad \text{(respectively } M \to \hat{M}) \]

is an isomorphism. It is not true in general that the completion $\hat{R}$ (or $\hat{M}$) is complete (with respect to $a\hat{R}$).

**Lemma 1.4.** $\hat{M}$ is complete with respect to $a\hat{R}$ if and only if the sequence

\[ 0 \to a^n\hat{M} \to \hat{M} \to M/a^nM \to 0 \]

is exact for each $n$.

The map $\hat{M} \to M/a^nM$ is always surjective. So the content of exactness is that the kernel of this map is given by $a^n\hat{M}$.

**Proof.** Let $K_n$ be the kernel of $\hat{M} \to M/a^nM$. Note that $K_n$ contains $a^n\hat{M}$, so we have exact sequences

\[ 0 \to K_n/a^n\hat{M} \to \hat{M}/a^n\hat{M} \to M/a^nM \to 0 \]

It suffices to show that the $K_n/a^n\hat{M}$ form a surjective system, since Lemma 1.1 will then give an exact sequence

\[ 0 \to \lim_{\leftarrow} K_n/a^n\hat{M} \to \hat{M}/\hat{M} \to \hat{M} \to 0, \]

so $\hat{M} \cong \hat{M}$ if and only if each $K_n/a^n\hat{M} = 0$.

Indeed, since $\hat{M} \to M/a^{n+1}M$ is surjective, after multiplying by $a^n$, the map

\[ a^n\hat{M} \to a^nM/a^{n+1}M \]
is still surjective. Take \( x \in K_n \). Then \( x \) maps to \( a^nM \) by definition. So there is some \( y \in a^n \hat{M} \) such that \( x - y \) maps to \( a^{n+1}M \). This means that \( x - y \in K_{n+1} \). In other words, \( K_n = K_{n+1} + a^n \hat{M} \). It follows that the map

\[
K_{n+1}/a^{n+1} \hat{M} \to K_n/a^n \hat{M}
\]

is surjective.

In particular, if \( \hat{M} \) is complete, we have an isomorphism

\[
\hat{M}/a^n \hat{M} \cong M/a^n M
\]

for each \( n \). Define the associated graded module of \( M \) by

\[
\text{Gr}_a(M) = \bigoplus_{n \geq 0} a^n M/a^{n+1} M, \quad \text{with } \deg(a^n M/a^{n+1} M) = n
\]

By an application of the five-lemma, we see that \( \text{Gr}_a(\hat{M}) \cong \text{Gr}_a(M) \) if \( \hat{M} \) is complete.

**Proposition 1.5.** Suppose \( a \) is finitely generated. Then \( \hat{M} \) is complete with respect to \( a\hat{R} \); in particular, \( \hat{R} \) is complete with respect to \( a\hat{R} \).

**Proof.** By Lemma 1.4, we just need to show that the kernel \( K_n \) of the surjection \( \hat{M} \to M/a^n M \) is contained in \( a^n \hat{M} \). Indeed, let \( x \in K_n \). It is represented by a sequence

\[
x = (0, \ldots, 0 = \bar{x}_n, \bar{x}_{n+1}, \bar{x}_{n+2}, \ldots)
\]

where each \( \bar{x}_m \in a^n M/a^n M \) (\( m \geq n \)). Choose a lift \( x_m \in a^n M \) of each \( \bar{x}_m \). Then

\[
\delta_m = x_{m+1} - x_m \in a^n M
\]

and \( x \) can be written as the sum of the convergent power series

\[
x = x_n + \delta_n + \delta_{n+1} + \cdots
\]

Suppose \( a \) is generated by elements \( f_1, \ldots, f_r \). We may write \( x_n \) and each \( \delta_m \) as sums:

\[
x_n = \sum_{|J|=n} f^J x_{m,J}, \quad \delta_m = \sum_{|J|=n} f^J \delta_{m,J}
\]

where \( x_{m,J} \in M, \delta_{m,J} \in a^{m-n} M \) and \( f^J = f_1^{j_1} \cdots f_r^{j_r} \). Therefore,

\[
x = \sum_{|J|=n} f^J (x_{m,J} + \delta_{m,J} + \delta_{n+1,J} + \cdots) = \sum_{|J|=n} f^J x_{J}
\]

where each \( x_J = x_{n,J} + \delta_{n,J} + \cdots \) lies in \( \hat{M} \). So \( x \in a^n \hat{M} \), as required.

Examples of complete rings include

(i) The \( p \)-adic integers: \( \mathbb{Z}_p \). By definition,

\[
\mathbb{Z}_p = \lim \limits_{\leftarrow} \mathbb{Z}/p^n \mathbb{Z}
\]

It is complete since the ideal \( (p) \) is principal (so finitely generated).

(ii) Formal power series rings: \( A[[x_1, \ldots, x_n]] \) where \( A \) is any ring. By definition,

\[
A[[x_1, \ldots, x_n]] = \lim \limits_{\leftarrow} A[x_1, \ldots, x_n]/(x_1, \ldots, x_n)^k
\]

It is complete with respect to the ideal \( a = (x_1, \ldots, x_n) \). This ring is Noetherian if and only if \( A \) is \(^1\).

\(^1\)For the “only if” direction, mod out \( a \). The “if” direction is the formal Hilbert basis theorem; see [1] or [2].
1.4. More on completion. A map of $R$-modules $f : M \to N$ induces a map $\hat{f} : \hat{M} \to \hat{N}$ of $\mathfrak{a}$-adic completions. This construction is functorial. We study this functor in more details in this subsection.

**Proposition 1.6.** Suppose $\hat{f} : \hat{M} \to \hat{N}$ is the map induced by $f : M \to N$. Then

(i) $\hat{f}$ is surjective if the map $M/\mathfrak{a}M \to N/\mathfrak{a}N$ is surjective. In particular, completion preserves surjection.

(ii) $\hat{f}$ is injective if the map $\text{Gr}_\mathfrak{a}(M) \to \text{Gr}_\mathfrak{a}(N)$ is injective.

Before proving Proposition 1.6, we first need a version of Nakayama lemma that does not assume finite generation.

**Lemma 1.7 ("Nilpotent Nakayama").** Let $\mathfrak{a}$ be a nilpotent ideal of $R$, i.e. $\mathfrak{a}^n = 0$ for some $n$. Suppose $N \subset M$ are $R$-modules with $M = \mathfrak{a}M + N$. Then $M = N$.

**Proof.** The hypothesis means that $M/N = \mathfrak{a}(M/N)$. Hence $M/N = \mathfrak{a}^n(M/N) = 0$. $\square$

**Proof of Proposition 1.6.**

(i) There is a commutative diagram

\[
\begin{array}{ccc}
M/\mathfrak{a}^nM & \longrightarrow & M/\mathfrak{a}M \\
\downarrow & & \downarrow \\
N/\mathfrak{a}^nN & \longrightarrow & N/\mathfrak{a}N \\
\end{array}
\]

where all but the left vertical arrow are surjective a priori. So the diagonal arrow $M/\mathfrak{a}^nM \to N/\mathfrak{a}N$ is surjective. On the other hand, $N/\mathfrak{a}N$ is the quotient of $N/\mathfrak{a}^nN$ by the nilpotent ideal $\mathfrak{a}/\mathfrak{a}^n$ in the ring $R/\mathfrak{a}^n$. Hence, $M/\mathfrak{a}^nM \to N/\mathfrak{a}^nN$ is also surjective by Lemma 1.7.

Now, write down exact sequences

\[0 \to K_n \to M/\mathfrak{a}^nM \to N/\mathfrak{a}^nN \to 0.\]

It again suffices, by Lemma 1.1, to show that the $K_n$ form a surjective system. An application of the snake lemma boils this down to showing that

\[\mathfrak{a}^nM/\mathfrak{a}^{n+1}M \to \mathfrak{a}^nN/\mathfrak{a}^{n+1}N\]

is surjective. Indeed, each $y \in \mathfrak{a}^nN$ is of the form

\[y = \sum_{i=1}^{r} a_i y_i, \quad \text{for} \quad a_i \in \mathfrak{a}^n, \quad y_i \in N.\]

But $M \to N/\mathfrak{a}N$ is surjective, so each

\[y_i = f(x_i) + \sum_{j=1}^{s_i} a_{ij} y_{ij}, \quad \text{for} \quad x_i \in M, \quad a_{ij} \in \mathfrak{a}, \quad y_{ij} \in N.\]

This shows that

\[y = \sum_{i,j} a_{ij} f(x_i) + a_i a_{ij} y_{ij}.\]

The first term on the right hand side is in the image of $\mathfrak{a}^n M$, while the second term lies in $\mathfrak{a}^{n+1} N$.

(ii) Suppose $x \in \hat{M}$ is mapped to zero under $\hat{f}$. Write $x = (\bar{x}_1, \bar{x}_2, \cdots)$, where each $\bar{x}_n \in M/\mathfrak{a}^nM$ is represented by some $x_n \in M$. The hypothesis is that $f(x_n) \in \mathfrak{a}^n N$ for all $n$. If $x_1 \notin \mathfrak{a} M$, then the element

\[(\bar{x}_1, 0, 0, \cdots) \in \text{Gr}_\mathfrak{a}(M)\]

is nonzero, and maps to zero in $\text{Gr}_\mathfrak{a}(N)$. This is impossible, so $x_1 \in \mathfrak{a} M$, and $\bar{x}_1 = 0$.

We now assume by induction that $\bar{x}_1, \cdots, \bar{x}_{n-1} = 0$. Since $x_n = x_{n+1} \mod \mathfrak{a}^n M$, we must have $x_{n+1} \in \mathfrak{a}^n M$. If $x_{n+1} \notin \mathfrak{a}^{n+1} M$. Then

\[(0, \cdots, 0, \bar{x}_{n+1}, 0, 0, \cdots) \in \text{Gr}_\mathfrak{a}(M)\]

is nonzero, and maps to zero in $\text{Gr}_\mathfrak{a}(N)$. This is again impossible, so $x_{n+1} \in \mathfrak{a}^{n+1} M$, and $\bar{x}_{n+1} = 0$. $\square$
The rest of this subsection is devoted to general results on completions that will not be used in the sequel.

**Corollary 1.8.** If $M$ is a finite $R$-module, then the natural map $M \otimes_R \hat{R} \to \hat{M}$ is surjective.

**Proof.** There is a surjection $R^r \to M$. By Proposition 1.6, the map $\hat{R}^r \to \hat{M}$ is surjective. This map also arises as the composition

$$\hat{R}^r \twoheadrightarrow R^r \otimes_R \hat{R} \to M \otimes_R \hat{R} \to \hat{M}.$$ 

So the last map in the chain is again surjective. □

We say $M$ is separated in the $a$-adic topology if

$$\bigcap_{n \geq 1} a^n M = 0.$$ 

This is equivalent to $M$ being Hausdorff in the $a$-adic topology. Since the intersection of all $a^n M$ is precisely the kernel of $M \to \hat{M}$, the module $M$ is separated if and only if $M \to \hat{M}$ is injective.

**Corollary 1.9.** If $R$ is complete with respect to $a$, then any finite, separated $R$-module $M$ satisfies $M \cong \hat{M}$.

**Proof.** The map

$$M \twoheadrightarrow M \otimes_R \hat{R} \to \hat{M}$$

is injective by hypothesis, and surjective by Corollary 1.8. □

Here is another version of Nakayama lemma that does not assume finite generation.

**Corollary 1.10** ("Complete Nakayama"). Suppose $R$ is complete with respect to $a$, and $M$ is a separated $R$-module. If the images of $m_1, \ldots, m_n \in M$ generate $M/aM$, then $m_1, \ldots, m_n$ generate $M$.

**Proof.** Let $N$ be the submodule of $M$ generated by $m_1, \ldots, m_n$. There is a commutative diagram

$$\begin{array}{ccc}
N & \to & M \\
\downarrow & & \downarrow \\
N & \to & \hat{M}
\end{array}$$

where the left vertical arrow is an isomorphism because $N$ is separated and finite, so Corollary 1.9 applies to $N$. Since $N \to M/aM$ is surjective, the bottom arrow is surjective by Proposition 1.6(i). Hence every arrow in the diagram is an isomorphism. □

1.5. **Hensel’s lemma.** Our objective is to show that complete rings satisfy the multivariate Hensel’s lemma.

Let $f_1, \ldots, f_n \in R[x_1, \ldots, x_n]$. Use $f$ to denote the $n$-tuple $(f_1, \ldots, f_n)$, and $J_a(f)$ to denote the Jacobian matrix evaluated at $a \in R^n$:

$$J_a(f) = \left( \frac{\partial f_i}{\partial x_j}(a) \right)_{1 \leq i, j \leq n}$$

**Theorem 1.11.** Suppose $R$ is complete with respect to some ideal $a$. If there is some $a \in R^n$ satisfying

$$f(a) = 0 \mod \det(J_a(f))^2 a,$$

then there is some $b \in R^n$ such that $f(b) = 0$ and

$$b = a \mod \det(J_a(f)) a$$

Furthermore, $b$ is unique if $\det(J_a(f))$ is not a zero-divisor.

Note that no Noetherian hypothesis is required for $R$. 

1.6. Variants. There are other statements that go under the name of Hensel’s lemma. In fact, for any local ring \((R, m)\) with residue field \(k\), the following statements are equivalent:

(i) For every monic \(f \in R[x]\) and every simple root \(\bar{a} \in k\) of the reduced polynomial \(\bar{f} \in k[x]\), there exists some \(a \in R\) such that

\[
    f(a) = 0
\]

and

\[
    a = \bar{a} \mod m.
\]

(ii) For every monic \(f \in R[x]\) and any factorization \(\bar{f} = \bar{g}\bar{h}\) into coprime polynomials \(\bar{g}, \bar{h} \in k[x]\), there exist polynomials \(g, h \in R[x]\) such that

\[
    f = gh
\]

and

\[
    g = \bar{g}, \quad h = \bar{h} \mod m.
\]

(iii) For any étale ring map \(\varphi : R \to S\) and prime \(q\) of \(S\) lying over \(m\) with residue field \(k\), there exists a section \(\tau : S \to R\) of \(\varphi\) with \(q = \tau^{-1}(m)\).

(iv) Any finite \(R\)-algebra is a finite product of local rings.

The local ring \((R, m)\) satisfying these properties is called a Henselian local ring. For proofs of these statements, see [4, Tag 04GG]. Note that (i) is a (very) special case of Theorem 1.11.3

2. The inverse function theorem

2.1. The inverse function theorem. We first study maps from a formal power series ring \(R[[x_1, \ldots, x_n]]\). They are uniquely determined by their actions on \(x_i\), when the target is complete.

Proposition 2.1. Let \(R\) be a ring (not necessarily complete). Let \(S\) be an \(R\)-algebra that is complete with respect to some ideal \(n\). Given \(f_1, \ldots, f_n \in n\), there is a unique \(R\)-algebra map

\[
    \varphi : R[[x_1, \ldots, x_n]] \to S
\]

sending each \(x_i\) to \(f_i\). Furthermore,

(i) \(\varphi\) sends each power series \(g(x)\) to \(g(f)\).

(ii) \(\varphi\) is surjective if the induced map

\[
    R \to S/n
\]

is surjective and \(f_1, \ldots, f_n \) generate \(n\).

(iii) \(\varphi\) is injective if the induced map

\[
    \text{Gr}(\varphi) : R[x_1, \ldots, x_n] \to \text{Gr}_n(S)
\]

is injective and \(f_1, \ldots, f_n \) generate \(n\).4

Proof. Let \(m\) be the ideal \((x_1, \ldots, x_n)\) of \(R[[x_1, \ldots, x_n]]\). For each \(k \in \mathbb{N}\), there is a unique ring map

\[
    \varphi_k : R[[x_1, \ldots, x_n]] \to S/n^k
\]

sending \(x_i\) to the class of \(f_i\), since \(\varphi_k\) factors through

\[
    R[[x_1, \ldots, x_n]]/m^k \cong R[x_1, \ldots, x_n]/(x_1, \ldots, x_n)^k.
\]

The collection \(\varphi_k\) induces the unique map \(\varphi\) since \(S\) is the inverse limit of \(S/n^k\).

(i) We ought to show that the image of \(g(x)\) in \(S\) agrees with \(g(f)\) after reducing by \(n^k\). This is clear.

(ii) The fact that \(f_1, \ldots, f_n \) generate \(n\) shows that as an \(R[x_1, \ldots, x_n]\)-submodule of \(S\),

\[
    n = aS
\]

where \(a\) is the ideal \((x_1, \ldots, x_n)\) of \(R[x_1, \ldots, x_n]\). Therefore, the \(R[x_1, \ldots, x_n]\)-module map

\[
    R[x_1, \ldots, x_n] \to S
\]

satisfies the hypothesis of Proposition 1.6(i). So the induced map on completions is surjective.

3I wonder whether the multivariate Hensel’s lemma holds for every Henselian ring?

4In (iii), the hypothesis that \(f_1, \ldots, f_n \) generate \(n\) can be removed; see [2, Theorem 7.16c].
(iii) \( \text{Gr}(\varphi) \) is exactly the map of associated graded modules of \( R[x_1, \ldots, x_n] \) and \( S \), over the ring \( R[x_1, \ldots, x_n] \) and the ideal \( \mathfrak{a} = (x_1, \ldots, x_n) \). Hence Proposition 1.6(ii) applies.

The following result is the inverse function theorem for formal power series:

Let \( R \) be a ring, and \( \mathfrak{m} \) be the ideal \( (x_1, \ldots, x_n) \) of \( R[[x_1, \ldots, x_n]] \).

**Theorem 2.2.** Suppose \( f_1, \ldots, f_n \in \mathfrak{m} \). If \( \varphi \) is the \( R \)-linear endomorphism

\[
\varphi : R[[x_1, \ldots, x_n]] \to R[[x_1, \ldots, x_n]]
\]

sending \( x_i \) to \( f_i \). Then \( \varphi \) is an isomorphism if and only if \( \det(J_0(\mathfrak{f})) \) is a unit.

**Proof.** The endomorphism \( \varphi \) induces a map of graded rings:

\[
\text{Gr}_\mathfrak{m}(\varphi) : R[x_1, \ldots, x_n] \to R[x_1, \ldots, x_n]
\]
given by \( x \mapsto J_0(\mathfrak{f}) \cdot x \) (where \( x \) is understood as a column vector), i.e. each

\[
x_i \mapsto \hat{f}_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(0)x_j
\]

Suppose \( \varphi \) is an isomorphism. Then \( \text{Gr}_\mathfrak{m}(\varphi) \) must be surjective on the degree-1 piece of \( R[[x_1, \ldots, x_n]] \).

Hence each \( x_1 \) is a linear combination of \( \hat{f}_1, \ldots, \hat{f}_n \). This implies that we have an \( n \)-by-\( n \) matrix \( T \) with coefficients in \( R \) such that

\[
T \cdot J_0(\mathfrak{f}) = \mathbb{I}.
\]

Hence \( \det(J_0(\mathfrak{f})) \) is a unit.

Conversely, suppose \( \det(J_0(\mathfrak{f})) \) is a unit. Then \( J_0(\mathfrak{f}) \) is invertible, so \( \text{Gr}_\mathfrak{m}(\varphi) \) is an isomorphism. In particular, \( \hat{f}_1, \ldots, \hat{f}_n \) generate \( \mathfrak{m}/\mathfrak{m}^2 \) over \( R \). In order to apply Proposition 2.1 to conclude that \( \varphi \) is an isomorphism, we only need to show that \( \hat{f}_1, \ldots, \hat{f}_n \) generate the ideal \( \mathfrak{m} \) of \( R[[x_1, \ldots, x_n]] \). Note that

(i) \( \mathfrak{m} \) is contained in the Jacobson radical of \( R[[x_1, \ldots, x_n]] \). This is because \( 1 - g \) is invertible for every \( g \in \mathfrak{m} \); its inverse is given by \( 1 + g + g^2 + \cdots \).

(ii) \( \mathfrak{m} \) can be regarded as a finite \( R[[x_1, \ldots, x_n]] \)-module. Indeed, it is generated by \( x_1, \ldots, x_n \).

The fact that \( \hat{f}_1, \ldots, \hat{f}_n \) generate \( \mathfrak{m}/\mathfrak{m}^2 \) now shows that \( \hat{f}_1, \ldots, \hat{f}_n \) generate \( \mathfrak{m} \), by Nakayama lemma. \( \Box \)

### 3. Proof of Hensel’s Lemma and the Implicit Function Theorem

#### 3.1. The proof

Hensel’s lemma follows easily from Proposition 2.1 and the inverse function theorem.

**Proof of Theorem 1.11.** Let \( e = \det(J_{\mathfrak{a}}(\mathfrak{f})) \). Then we have an \( n \)-by-\( n \) matrix \( T \) with coefficients in \( R \) such that \( T \cdot J_\mathfrak{a}(\mathfrak{f}) = J_\mathfrak{a}(\mathfrak{f}) \cdot T = e \mathbb{I} \). In particular,

\[
f(a + eT \cdot x) = f(a) + J_\mathfrak{a}(\mathfrak{f}) \cdot eT \cdot x + e^2 g(T \cdot x)
\]

\[
= f(a) + e^2 (x + g(T \cdot x))
\]

for some \( g \in \mathfrak{m}^2 \). The map \( \varphi : R[[x_1, \ldots, x_n]] \to R[[x_1, \ldots, x_n]] \) determined by

\[
\varphi : x \mapsto x + g(T \cdot x)
\]

has Jacobian \( \mathbb{I} \). By the inverse function theorem 2.2, it is an isomorphism. The inverse \( \varphi^{-1} \) is also \( R \)-linear.

In other words,

\[
f(a + eT \cdot \varphi^{-1}(x)) = f(a) + e^2 x
\]

Now, the hypothesis shows that \( f(a) = e^2 c \), where each component of \( c \) lies in the ideal \( \mathfrak{a} \). Since \( R \) is complete with respect to \( \mathfrak{a} \), there is an \( R \)-linear map \( \psi : R[[x_1, \ldots, x_n]] \to R \) determined by

\[
\psi : x \mapsto -c
\]

Applying \( \psi \) to the above equation, we obtain

\[
f(a + eT \cdot \psi \varphi^{-1}(x)) = e^2 \psi(c + x) = 0.
\]

Thus \( b = a + eT \cdot \psi \varphi^{-1}(x) \) is the desired element.
We now assume $e$ is not a zero-divisor, and prove the uniqueness of $b$. Note that

\[ T \cdot f(a + ex) = T \cdot f(a) + T \cdot J_a(f) \cdot ex + e^2 T \cdot g(x) = T \cdot f(a) + e^2 (x + T \cdot g(x)) \]

Like before, the map $\tilde{\varphi} : R[[x_1, \ldots, x_n]] \to R[[x_1, \ldots, x_n]]$ determined by

\[ \tilde{\varphi} : x \mapsto x + T \cdot g(x) \]

is an isomorphism. Now, suppose $b = a + er$ and $b' = a + er'$ satisfy $f(b) = f(b') = 0$. Then the above formula, along with cancellation by $e^2$, shows that

\[ b + T \cdot g(b) = b' + T \cdot g(b') \]

Let $\beta, \beta' : R[[x_1, \ldots, x_n]] \to R$ be determined by sending $x$ to $b$ and $b'$, respectively. We have just seen that the two compositions

\[ R[[x_1, \ldots, x_n]] \xrightarrow{\tilde{\varphi}} R[[x_1, \ldots, x_n]] \xrightarrow{\beta} R \]

agree on $x$. The uniqueness of Proposition 2.1 then shows that $\beta \tilde{\varphi} = \beta' \tilde{\varphi}$. But $\tilde{\varphi}$ is an isomorphism, so $\beta = \beta'$. Thus $b = b'$.

3.2. The implicit function theorem. Hensel’s lemma should really be understood as saying that the ring $R$ “satisfies the implicit function theorem,” i.e. the variables $x_1, \ldots, x_n$ afford a unique parametrization by elements in $R$ when the Jacobian $J_a(f)$ is invertible. The standard form of implicit function theorem is precisely the case where $R$ is a ring of power series:

Let $k$ be a field.

**Theorem 3.1.** Suppose $f_1, \ldots, f_n \in k[y_1, \ldots, y_m; x_1, \ldots, x_n]$ satisfy $f(0; a) = 0$, and

\[ \det \left( \frac{\partial f_i}{\partial x_j} (0; a) \right)_{1 \leq i, j \leq n} \neq 0 \]

Then there are unique power series $g_1, \ldots, g_n \in k[[y_1, \ldots, y_m]]$ such that $f(y; g(y)) = 0$ and $g(0) = a$.

**Proof.** Let $R = k[[y_1, \ldots, y_m]]$. It is complete with respect to the unique maximal ideal $a = (y_1, \ldots, y_m)$. Denote by $\tilde{f}_1, \ldots, \tilde{f}_n$ the images of $f_1, \ldots, f_n$ in $R[[x_1, \ldots, x_n]]$. The hypothesis $f(0; a) = 0$ means precisely that

\[ \tilde{f}(a) = 0 \mod a \]

while the non-vanishing of the determinant means that

\[ \det(J_a(\tilde{f})) \neq 0 \mod a \]

In other words, $\det(J_a(\tilde{f}))$ is a unit in the ring $R$. Hensel’s lemma 1.11 provides a unique $g \in R^n$ with the property that $\tilde{f}(g) = 0$ and $g = a \mod a$. These properties translate directly to $f(y; g(y)) = 0$ and $g(0) = a$. \qed

**References**


