Divulgaciones Matemáticas Vol. 22, No. 1 (2021), pp. 90–95

A proof of a version of Hensel's lemma

Una prueba de una versión del lema de Hensel

Dinamérico P. Pombo Jr. (dpombojr@gmail.com)
Instituto de Matemática e Estatística
Universidade Federal Fluminense
Rua Professor Marcos Waldemar de Freitas Reis, s/nº
Bloco G, Campus do Gragoatá
24210-201 Niteri, RJ Brasil

Abstract

By using a few basic facts, a proof of a known version of Hensel's lemma in the context of local rings is presented.

Key words and phrases: local rings, discrete valuation rings, Hensel's lemma.

Resumen

Usando algunos pocos hechos básicos, se presenta una demostración de una versión del lema de Hensel en el contexto de los anillos locales.

Palabras y frases clave: anillos locales, anillos de valoración discretos, lema de Hensel.

1 Introduction

A classical and fundamental result, known as Hensel's lemma, is discussed in [1], [3], [5], [6] and [7], for instance. A quite general form of Hensel's lemma may be found in Chapter III of [2], although special cases of it may also be very important, as the one valid in the framework of local rings and presented in Chapter II of [6]. The main purpose of this note is to offer an elementary proof of the last-mentioned form of Hensel's lemma, as well as to derive a few consequences of it.

2 A proof of a version of Hensel's lemma

Definition 2.1 (cf. [2, p. 80]). A commutative ring R with and identity element $1 \neq 0$ is said to be a *local ring* if it contains a unique maximal ideal I_1 , namely, the set of non-invertible elements of R. If K is the quotient ring R/I_1 , which is a field,

$$\lambda \in R \longmapsto \bar{\lambda} \in K$$

will denote the canonical surjection. For $f(X) = a_0 + a_1\, X + \cdots + a_n\, X^n \in R[X]$, we will write $\bar{f}(X) = \bar{a}_0 + \bar{a}_1\, X + \cdots + \bar{a}_n\, X^n \in K[X]$.

Received 29/06/2021. Revised 30/06/2021. Accepted 26/07/2021.
MSC (2010): Primary 12J25, 13F30; Secondary 13H99, 13J10, 13B25.
Corresponding author: Dinamérico P. Pombo Jr.

Example 2.1 (cf. [6]). *Let R be a discrete valuation ring and I_1 the maximal ideal of R, which may be written as $I_1 = \pi R$. We have that*

$$I_1 = \pi R \supset I_2 = \pi^2 R \supset \cdots \supset I_n = \pi^n R \supset I_{n+1} = \pi^{n+1} R \supset \ldots.$$

is a decreasing sequence of ideals of R such that $I_n I_1 \subset I_{n+1}$ for each integer $n \geq 1$ and $\bigcap_{n\geq 1} I_n = \{0\}$.

Example 2.2 (cf. [3]). *Let $\mathbb{K}$ be a field endowed with a non-trivial discrete valuation $|\cdot|$, $R = \{\lambda \in R;\, |\lambda| \leq 1\}$ the ring of integers of $(\mathbb{K}, |\cdot|)$ and $I_1 = \{\lambda \in R;\, |\lambda| < 1\}$ the maximal ideal of R. Let $\mu \in I_1$ be such that $|\mu| = \sup\{|\lambda|;\, \lambda \in I_1\}$. Then*

$$I_1 = \mu R \supset I_2 = \mu^2 R \supset \cdots \supset I_n = \mu^n R \supset I_{n+1} = \mu^{n+1} R \supset \ldots.$$

is a decreasing sequence of ideals of R such that $I_n I_1 \subset I_{n+1}$ for each integer $n \geq 1$ and $\bigcap_{n\geq 1} I_n = \{0\}$.

It may be seen that every discrete valuation ring may be regarded as the ring of integers of a field endowed with a non-trivial discrete valuation.

Let us recall that, if X is a non-empty set, a mapping

$$d\colon X \times X \longrightarrow \mathbb{R}_+$$

is an *ultrametric* on X if the following conditions hold for all $x, y, z \in X$:

(a) $d(x, y) = 0$ if and only if $x = y$;

(b) $d(x, y) = d(y, x)$;

(c) $d(x, y) \leq \max\{d(x, z), d(z, y)\}$.

By induction,

$$d(x_1, x_n) \leq \max\{d(x_1, x_2), \ldots, d(x_{n-1}, x_n)\}$$

for $n = 2, 3, \ldots$ and $x_1, \ldots, x_n \in X$. And, since $\max\{d(x, z), d(z, y)\} \leq d(x, z) + d(z, y)$, d is a metric on X.

We shall present an elementary proof of the following form of Hensel's lemma [6, p. 43]:

Proposition 2.1. *Let R be a local ring and I_1 its maximal ideal, and assume the existence of a decreasing sequence $I_1 \supset I_2 \supset \cdots \supset I_n \supset I_{n+1} \supset \ldots.$ of ideals of R such that $I_n I_1 \subset I_{n+1}$ for each integer $n \geq 1$ and $\bigcap_{n\geq 1} I_n = \{0\}$. Then there exists a translation-invariant ultrametric d on R such that $I_n = \left\{\lambda \in R;\, d(\lambda, 0) \leq \frac{1}{2^n}\right\}$ for each integer $n \geq 1$ (thus $(I_n)\_n \geq 1$ is a fundamental system of neighborhoods of 0 in R with respect to the topology defined by d) and the mappings*

$$(\lambda, \mu) \in R \times R \longmapsto \lambda + \mu \in R \quad \textit{and} \quad (\lambda, \mu) \in R \times R \longmapsto \lambda\mu \in R$$

are continuous. Moreover, if the metric space (R, d) is complete and if $f(X) \in R[X]$ is such that $\bar{f}(X)$ admits a simple root θ in K, then there exists a unique root λ of $f(X)$ in R such that $\bar{\lambda} = \theta$.

In order to prove Proposition 2.1 we shall need an auxiliary result:

Lemma 2.1. *Let $(G,+)$ be a commutative group and $H_1 \supset H_2 \supset \cdots \supset H_n \supset H_{n+1} \supset \cdots$ a decreasing sequence of subgroups of G such that $\bigcap\limits_{n\geq 1} H_n = \{0\}$. Then there exists a translation-invariant ultrametric d on G such that $H_n = \big\{x \in G;\, d(x,0) \leq \frac{1}{2^n}\big\}$ for each integer $n \geq 1$ (thus $(H_n)\_n \geq 1$ is a fundamental system of neighborhoods of 0 in G with respect to the topology defined by d) and the mapping*

$$(x,y) \in G \times G \longmapsto x + y \in G$$

is continuous.

Proof of Lemma 2.1. We shall use a classical argument. Put $H_0 = G$ and let $g\colon G \to \mathbb{R}_+$ be the mapping given by $g(0) = 0$ and $g(x) = \frac{1}{2^n}$ if $x \in H_n \backslash H_{n+1}$ $(n = 0,1,2,\dots)$. Obviously, $g(x) > 0$ if $g \in G\backslash\{0\}$, $g(-x) = g(x)$ if $x \in G$ and

$$H_n = \left\{x \in G;\, g(x) \leq \frac{1}{2^n}\right\}$$

for $n = 0,1,2,\dots$. Moreover, $g(x+y) \leq \max\{g(x), g(y)\}$ for all $x, y \in G$, which is clear if $x = 0$ or $y = 0$. Indeed, if $x, y \in G\backslash\{0\}$, $x \in H_k \backslash H_{k+1}$, $y \in H_\ell \backslash H_{\ell+1}$, with $\ell \geq k \geq 0$, then $g(x) = \frac{1}{2^k}$ and $g(y) = \frac{1}{2^\ell} \leq \frac{1}{2^k}\cdot$ But, since $H_\ell \subset H_k$, $x + y \in H_k$, and hence $g(x+y) \leq \frac{1}{2^k} = \max\{g(x), g(y)\}$.

Therefore the mapping

$$d\colon G \times G \longrightarrow \mathbb{R}_+ ,$$

defined by $d(x,y) = g(x-y)$, is a translation-invariant ultrametric on G such that

$$H_n = \left\{t \in G;\, d(t,0) \leq \frac{1}{2^n}\right\}$$

for each integer $n \geq 0$. Consequently,

$$x + H_n = \left\{t \in G;\, d(t,x) \leq \frac{1}{2^n}\right\}$$

if $x \in G$ and $n = 0,1,2,\dots$ are arbitrary.

Finally, if $x_0, y_0 \in G$ and $n = 0,1,2,\dots$ are arbitrary,

$$(x_0 + H_n) + (y_0 + H_n) \subset (x_0 + y_0) + H_n ,$$

proving the continuity of the mapping

$$(x,y) \in G \times G \longmapsto x + y \in G$$

at (x_0, y_0). □

Now, let us turn to the

Proof of Proposition 2.1. By Lemma 2.1 there is a translation-invariant ultrametric d on R such that

$$I_n = \left\{\lambda \in R;\, d(\lambda, 0) \leq \frac{1}{2^n}\right\}$$

for each integer $n \geq 1$, and the operation of addition in R is continuous. Moreover, if $(\lambda_0, \mu_0) \in R \times R$ and $n = 1, 2, \dots$ are arbitrary, the relations $\lambda \in \lambda_0 + I_n$, $\mu \in \mu_0 + I_n$ imply

$$\lambda\mu - \lambda_0\mu_0 = \lambda\mu - \lambda_0\mu + \lambda_0\mu - \lambda_0\mu_0 = \mu(\lambda - \lambda_0) + \lambda_0(\mu - \mu_0) \in I_n + I_n \subset I_n\,,$$

proving the continuity of the mapping

$$(\lambda, \mu) \in R \times R \longmapsto \lambda\mu \in R$$

at (λ_0, μ_0).

Now, assume that (R, d) is complete and let $f(X), \bar{f}(X), \lambda, \theta$ be as in the statement of the proposition. In order to conclude the proof we shall apply Newton's approximation method, as in p. 44 of [6]. Let us first observe that, if $h(X) \in R[X]$ and $\gamma \in R$, then $\overline{h(\gamma)} = \bar{h}(\bar{\gamma})$.

To prove the uniqueness, assume the existence of a $\mu \in R$ so that $\bar{\mu} = \theta$ and $f(\mu) = 0$. Since $\bar{\lambda} = \theta$ is a simple root of $\bar{f}(X)$, there is a $g(X) \in R[X]$ such that $f(X) = (X - \lambda)\, g(X)$ and $\bar{g}(\theta) \neq 0$; thus

$$0 = f(\mu) = (\mu - \lambda)\, g(\mu).$$

Therefore, since $\overline{g(\mu)} = \bar{g}(\theta) \neq 0$, we conclude that $g(\mu)$ is an invertible element of R; hence $\lambda = \mu$.

To prove the existence, we claim that there is a sequence $(\lambda_n)\_n \geq 1$ in R so that $\bar{\lambda}_n = \theta$, $f(\lambda_n) \in I_n$ and $\lambda_{n+1} - \lambda_n \in I_n$ for each integer $n \geq 1$. Indeed, let $\lambda_1 \in R$ be such that $\bar{\lambda}_1 = \theta$. Then $\overline{f(\lambda_1)} = \bar{f}(\theta) = 0$, that is, $f(\lambda_1) \in I_1$. Now, let $n \geq 1$ be arbitrary, and suppose the existence of a $\lambda_n \in R$ such that $\bar{\lambda}_n = \theta$ and $f(\lambda_n) \in I_n$. Then, for every $h \in I_n$, $(\lambda_n + h) - \lambda_n \in I_n$ and $\overline{(\lambda_n + h)} = \bar{\lambda}_n + \bar{h} = \theta$. We shall show the existence of an $h \in I_n$ with $f(\lambda_n + h) \in I_{n+1}$. In fact, by Taylor's formula [4, p. 387], there is a $\xi \in R$ so that

$$f(\lambda_n + h) = f(\lambda_n) + hf'(\lambda_n) + h^2\xi.$$

And, by hypothesis, $h^2\xi = h(h\xi) \in I_n\, I_n \subset I_n\, I_1 \subset I_{n+1}$. But, since θ is a simple root of $\bar{f}(X)$, $\overline{f'(\lambda_n)} = (\bar{f})'(\theta) \neq 0$, that is, $f'(\lambda_n)$ is an invertible element of R. Thus, by taking $h = -f(\lambda_n)(f'(\lambda_n))^{-1} \in I_n$ and $\lambda_{n+1} = \lambda_n + h$, we arrive at $\overline{\lambda_{n+1}} = \theta$, $f(\lambda_{n+1}) \in I_{n+1}$ and $\lambda_{n+1} - \lambda_n \in I_n$, as desired.

Finally, $\big(f(\lambda_n)\big)_{n\geq 1}$ converges to 0 in R, because $d\big(f(\lambda_n), 0\big) \leq \dfrac{1}{2^n}$ for $n = 1, 2, \dots$. On the other hand, for $n, \ell = 1, 2, \dots$,

$$d(\lambda_{n+\ell}, \lambda_n) \leq \max\{d(\lambda_{n+\ell}, \lambda_{n+\ell-1}), \dots, d(\lambda_{n+1}, \lambda_n)\} \leq \max\left\{\frac{1}{2^{n+\ell-1}}, \dots, \frac{1}{2^n}\right\} = \frac{1}{2^n}\,,$$

and hence $(\lambda_n)\_n \geq 1$ is a Cauchy sequence in (R, d). By the completeness of (R, d), there is a $\lambda \in R$ for which $(\lambda_n)\_n \geq 1$ converges. Consequently, in view of the continuity of the mappings

$$(\alpha, \beta) \in R \times R \longmapsto \alpha + \beta \in R \quad \text{and} \quad (\alpha, \beta) \in R \times R \longmapsto \alpha\beta \in R,$$

$\big(f(\lambda_n)\big)\_n \geq 1$ converges to $f(\lambda)$; thus $f(\lambda) = 0$.

Now, let us consider $K = R/I_1$ endowed with the discrete ultrametric d', given by $d'(s,s) = 0$ and $d'(s,t) = 1$ if $s \neq t$ $(s,t \in K)$. Since the canonical surjection

$$\lambda \in (R,d) \longmapsto \bar{\lambda} \in (K,d')$$

is continuous ($\bar{I_1} = \{0\}$) and $(\lambda_n)\_n \geq 1$ converges to λ, $(\bar{\lambda}_n)\_n \geq 1$ converges to $\bar{\lambda}$. Therefore $\bar{\lambda} = \theta$. □

Corollary 2.1. *Let R be a discrete valuation ring which is complete under the ultrametric d given in Proposition 2.1. Let $f(X) \in R[X]$ be such that $\bar{f}(X) \in K[X]$ admits a simple root θ. Then there exists a unique root λ of $f(X)$ such that $\bar{\lambda} = \theta$.*

Proof. Follows immediately from Proposition 2.1, by recalling Example 2.1. □

Remark 2.1. *Let $(\mathbb{K}, |\cdot|)$ and $I_n \quad (n = 1, 2, \dots)$ be as in Example 2.2. Then $\widetilde{d}(\lambda,\mu) = |\lambda - \mu|$ is an ultrametric on $\mathbb{K}$, and hence its restriction to $R \times R$ is an ultrametric on R (which we shall also denote by $\widetilde{d}$). Since, for $n = 1, 2, \dots$,*

$$\left\{\lambda \in R;\, \widetilde{d}(\lambda,0) = |\lambda| \leq \frac{1}{2^n}\right\} = I_n = \left\{\lambda \in R;\, d(\lambda,0) \leq \frac{1}{2^n}\right\},$$

d being as in Proposition 2.1, it follows that $\widetilde{d}$ and d are equivalent.

Corollary 2.2. *Let $(\mathbb{K}, |\cdot|)$ and μ be as in Example 2.2, and assume that $(\mathbb{K}, \widetilde{d})$ is complete. If $f(X) \in R[X]$ and $\bar{f}(X) \in K[X]$ admits a simple root θ, then there is a unique root λ of $f(X)$ so that $|\lambda - \xi| \leq |\mu|$ (where $\xi \in R$ and $\bar{\xi} = \theta$).*

Proof. Follows immediately from Remark 2.1 and Proposition 2.1. □

Corollary 2.3 (cf. [5, p. 16])**.** *Let p be a prime number, $\mathbb{Z}_p = \{\lambda \in \mathbb{Q}_p\,; |\lambda|_p \leq 1\}$ the ring of p-adic integers and $f(X) \in \mathbb{Z}_p[X]$. If there is an $a_0 \in \mathbb{Z}_p$ such that $|f(a_0)|_p < 1$ and $|f'(a_0)|_p = 1$, then there is a unique $a \in \mathbb{Z}_p$ such that $f(a) = 0$ and $|a - a_0|_p \leq \dfrac{1}{p}\cdot$*

Proof. Since the condition "$|f(a_0)| < 1$" is equivalent to the condition "$\bar{f}(\bar{a}_0) = \overline{f(a_0)} = 0$", and the condition "$|f'(a_0)|_p = 1$" is equivalent to the condition "$(\bar{f})'(\bar{a}_0) = \overline{f'(a_0)} \neq 0$", Theorem 6, p. 391 of [4] guarantees that $\bar{a}_0$ is a simple root of $\bar{f}(X)$. Therefore the result follows from Corollary 2.2. □

Example 2.3 (cf. [3, p. 52])**.** *Let p be a prime number, $p \neq 2$, and let $b \in \mathbb{Z}_p$ with $|b|_p = 1$. If there is an $a_0 \in \mathbb{Z}_p$ such that $|a_0^2 - b|\_p < 1$, then $b = a^2$ for a unique $a \in \mathbb{Z}_p$ such that $|a - a_0|\_p \leq \dfrac{1}{p}\cdot$*

Indeed, put $f(X) = X^2 - b \in \mathbb{Z}_p[X]$. Then $|f(a_0)|\_p = |a_0^2 - b|\_p < 1$ and $|f'(a_0)|\_p = |2a_0|\_p = |2|_p\,|a_0|\_p = |a_0|\_p = 1$ (the relation $|a_0^2 - b|\_p < 1 = |b|_p = 1$ implies $\big(|a_0|\_p\big)^2 = |(a_0^2 - b) + b|_p = |b|_p = 1$). Thus the result follows from Corollary 2.9.

In the same vein one shows that if p is a prime number, $p \neq 3$, $c \in \mathbb{Z}_p$, $|c|_p = 1$, and there is an $f_0 \in \mathbb{Z}_p$ such that $|f_0^3 - c|\_p < 1$, then $c = f^3$ for a unique $f \in \mathbb{Z}_p$ such that $|f - f_0|\_p \leq \dfrac{1}{p}\cdot$

References

[1] E. Artin. *Algebraic Numbers and Algebraic Functions*, American Mathematical Society, Providence, Rhode Island, 2005.

[2] N. Bourbaki. *Commutative Algebra*, Hermann and Addison-Wesley, Paris and Reading, Massachusetts, 1972.

[3] J.W. Cassels. *Local Fields*, London Mathematical Society Student Texts 3, Cambridge University Press, Cambridge, 1986.

[4] R. Godement. *Cours d'algbre*, Troisime dition, Enseignement des Sciences, Hermann, Paris, 1966.

[5] N. Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, Second edition, Springer-Verlag, Berlin. Heidelberg. New York, 1984.

[6] J. P. Serre. *Corps Locaux*, Quatrime dition, Actualits Scientifiques et Industrielles 1296, Hermann, Paris, 1968.

[7] J. P. Serre. *A Course in Arithmetic*, Third printing, Graduate Texts in Mathematics 7, Springer-Verlag, Berlin. Heidelberg. New York, 1985.