PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 125, Number 11, November 1997, Pages 3185–3189
S 0002-9939(97)04112-9

# A NOTE ON HENSEL'S LEMMA IN SEVERAL VARIABLES

BENJI FISHER

(Communicated by Wolmer V. Vasconcelos)

ABSTRACT. The standard hypotheses for Hensel's Lemma in several variables are slightly stronger than necessary, in the case that the Jacobian determinant is not a unit. This paper shows how to weaken the hypotheses for Hensel's Lemma and some related theorems.

## 0. INTRODUCTION

The most familiar version of Hensel's Lemma states that if $f$ is a polynomial with coefficients in $\mathbb{Z}_p$, $a \in \mathbb{Z}_p$ is an approximate root of $f$ (*i.e.*, $f(a) \equiv 0 \pmod{p}$), and $f'(a)$ is a unit (*i.e.*, $f'(a) \not\equiv 0 \pmod{p}$) then there is a unique root $a'$ near $a$: $f(a') = 0$ and $a' \equiv a \pmod{p}$. The first application of this theorem is to the polynomial $f(X) = X^p - X$: by Fermat's Little Theorem, every $p$-adic integer $a$ is an approximate root of $f$; clearly, $f'(a) \equiv -1 \pmod{p}$ for all $a \in \mathbb{Z}_p$. Thus there are $p$ roots of this polynomial in $\mathbb{Z}_p$; they are known as the Teichmüller representatives of the residue classes (mod $p$).

One way to generalize Hensel's Lemma is to relax the requirement that $f'(a)$ be a unit. Let $h = v_p\big(f'(a)\big)$. Letting $r = v_p\big(f(a)\big)$, one can show that if $r > 2h$ then there is a unique root $a'$ such that $a' \equiv a \pmod{p^{r-h}}$. The first example of this version of Hensel's Lemma is extracting square roots 2-adically: if $a^2 \equiv c \pmod{2^r}$, with $c$ odd and $r > 2$, then there is a square root of $c$ congruent to $a \pmod{2^{r-1}}$. In particular, an odd number has a square root in $\mathbb{Z}_2$ if and only if it is congruent to 1 (mod 8).

A further generalization is to consider $n$ polynomial equations in $n$ variables. The analogous statement, with the derivative replaced by the Jacobian determinant, is true. Let bold-face letters denote $n$-tuples: $\mathbf{a} = (a_1, \ldots, a_n)$ and so on. Let $\mathbf{f}$ be an $n$-tuple of polynomials in $n$ variables and $J_{\mathbf{f}} = \partial(f_1, \ldots, f_n)/\partial(X_1, \ldots, X_n)$ the Jacobian determinant. Let $\mathbf{a} \in {\mathbb{Z}_p}^n$ and $h = v_p\big(J_{\mathbf{f}}(\mathbf{a})\big)$. If $r = \min_i\{v_p\big(f_i(\mathbf{a})\big)\} > 2h$ then there is a unique $\mathbf{a}' \in {\mathbb{Z}_p}^n$ such that $f_i(\mathbf{a}') = 0$ and $a'_i \equiv a_i \pmod{p^{r-h}}$ for all $i$.

The point of this note is that the hypothesis $r > 2h$ above is stronger than necessary. The first improvement is to note how $h = v_p\big(J_{\mathbf{f}}(\mathbf{a})\big)$ is used in the proof: by Cramer's Rule, the Jacobian matrix times its adjoint (or adjoint-transpose) has the form $p^h \times (\text{unit}) \times I_n$; in other words, the Jacobian matrix $M_{\mathbf{f}}(\mathbf{a})$ divides $p^h I_n$. Any value of $h$ for which this is true will work just as well as $v_p\big(J_{\mathbf{f}}(\mathbf{a})\big)$. For example,

Received by the editors May 20, 1996.
1991 *Mathematics Subject Classification.* Primary 13J15; Secondary 13J05, 13B40.
*Key words and phrases.* Hensel's lemma, power series, Henselian rings.

if $M_{\mathbf{f}}(\mathbf{a})$ is diagonal, say $M_{\mathbf{f}}(\mathbf{a}) = \operatorname{diag}(p, \dots, p, 1)$ then one can take $h = 1$ instead of $h = v_p\big(J_{\mathbf{f}}(\mathbf{a})\big) = n - 1$, a big improvement. Sometimes it is convenient to allow larger values of $h$.

The second improvement is to note that the hypothesis that $\mathbf{f}(\mathbf{a})$ be divisible by $p^{2h+1}$ can be replaced by the hypothesis that $\mathbf{f}(\mathbf{a})$ be divisible by $p^{h+1}M_{\mathbf{f}}(\mathbf{a})$, in the sense of matrix multiplication: that is, $\mathbf{f}(\mathbf{a}) = p^h M_{\mathbf{f}}(\mathbf{a}) \cdot \mathbf{b}$ for some $\mathbf{b} \in (p\mathbb{Z}_p)^n$. For example, again suppose that $M_{\mathbf{f}}(\mathbf{a})$ is diagonal, say $M_{\mathbf{f}}(\mathbf{a}) = \operatorname{diag}(p^h, 1, \dots, 1)$ (so that the first improvement had no effect). Instead of requiring that $f_i(\mathbf{a}) \equiv 0 \pmod{p^{2h+1}}$ for all $i$, one may require this for $i = 1$ and $f_i(\mathbf{a}) \equiv 0 \pmod{p^{h+1}}$ for all $i > 1$.

For a number-theoretic application of this improved form of Hensel's Lemma, see my paper with R. Dabrowski, [Da-F, Theorem 1.8]. I would be interested to see a more geometric application. In the situation described above, the proof is given in [Da-F, Lemma 1.20]. In the rest of this paper, I will verify that the same improvements can be made in more general settings.

I would like to thank Romuald Dabrowski for his insistence, while we were preparing [Da-F], that we give definitive statements of our results. It was this insistence that forced me to formulate this improved version of Hensel's Lemma.

## 1. Definitions and notations

Following Bourbaki [B2, §III.4.5], I will say that a commutative ring $A$ and an ideal $\mathfrak{m} \subseteq A$ **satisfy Hensel's conditions** if $A$ is complete and separated with respect to a linear topology (*i.e.*, a topology defined by ideals) and $\mathfrak{m}$ is closed, consisting of topologically nilpotent elements. (In the Introduction, $A = \mathbb{Z}_p$ and $\mathfrak{m} = p\mathbb{Z}_p$.) To avoid confusion with the $n^{\text{th}}$ power of $\mathfrak{m}$, the $n$-fold Cartesian product of $\mathfrak{m}$ with itself will be denoted $\mathfrak{m}^{(n)}$. (In [B2], this is denoted $\mathfrak{m}^{\times n}$.)

Let $A\{\mathbf{X}\}$ denote the ring of restricted (formal) power series in $n$ variables, where $\mathbf{X} = (X_1, \dots, X_n)$ [B2, §III.4.2]. If $\mathbf{f} = (f_1, \dots, f_m)$ is an $m$-tuple in $A\{\mathbf{X}\}$, let $M_{\mathbf{f}}$ denote the Jacobian matrix; if $m = n$, let $J_{\mathbf{f}} = \det M_{\mathbf{f}}$ be the Jacobian determinant. Let $M_{\mathbf{f}}^{(r)}$ denote the matrix consisting of the first $r$ columns of $M_{\mathbf{f}}$; let $M_{\mathbf{f}}^{(-r)}$ denote the matrix consisting of the last $r$ columns of $M_{\mathbf{f}}$. (In [B2], there is no notation for the first $r$ columns; the last $r$ columns are denoted $M_{\mathbf{f}}^{(r)}$.) In particular, $M_{\mathbf{f}} = \big[M_{\mathbf{f}}^{(r)}\ M_{\mathbf{f}}^{(r-n)}\big]$. Similarly, if $\mathbf{a} \in A^m$, let $\mathbf{a}^{(r)}$ and $\mathbf{a}^{(-r)}$ denote the first and last $r$ entries of $\mathbf{a}$. Let $\mathbf{1}_n$ denote the $n$-tuple $\mathbf{1}_n = \mathbf{X} = (X_1, \dots, X_n)$, so that $M_{\mathbf{1}_n} = I_n$ (the identity matrix). I will always think of $\mathbf{f}$ and $\mathbf{a}$ as column vectors.

In particular, if $A$ is a discrete ring then it is automatically complete and separated; $\mathfrak{m}$ is automatically closed; and a restricted power series is simply a polynomial. In any case, polynomials are special cases of restricted power series.

## 2. Hensel's Lemma in several variables

Let $A$ be a ring and $\mathfrak{m} \subseteq A$ an ideal. I will assume that $A$ and $\mathfrak{m}$ satisfy the simplest version of Hensel's Lemma in several variables and derive a more general version that incorporates the points discussed in the Introduction. The argument closely follows one in Greenberg [Gr]. I will then show that if $A$ is a Henselian local ring and $\mathfrak{m}$ is its maximal ideal then this theorem applies. (This is surely well-known, but I do not know a reference.)

Assume that $A$ is complete and separated and that $\mathfrak{m}$ is closed, with respect to a linear topology. Let us say that $(A, \mathfrak{m})$ satisfies "condition (H)" if the following version of Hensel's Lemma holds:

**Condition (H).** *Let $\mathbf{f} = (f_1, \dots, f_n)$ be an $n$-tuple of restricted power series $f_i \in A\{\mathbf{X}\}$ and let $J = J_{\mathbf{f}}$ be the Jacobian determinant. Let $\mathbf{a} \in A^n$; assume that $\mathbf{f}(\mathbf{a}) \in \mathfrak{m}^{(n)}$ and that $J(\mathbf{a}) \in A^\times$. Then there is a unique $\mathbf{a}' \in A^n$ such that $\mathbf{f}(\mathbf{a}') = \mathbf{0}$ and $\mathbf{a}' \equiv \mathbf{a}$ (mod $\mathfrak{m}^{(n)}$).*

*Remark.* If we further assume that every element of $\mathfrak{m}$ is topologically nilpotent then $(A, \mathfrak{m})$ satisfy Hensel's conditions, and so condition (H) is satisfied by [B2, §III.4.5, Corollaire 3] (*cf.* §3, below).

**Theorem 1.** *Assume that $(A, \mathfrak{m})$ satisfies condition* (H). *Let $\mathbf{f} = (f_1, \dots, f_n)$ be an $n$-tuple with $f_i \in A\{\mathbf{X}\}$. Let $\mathbf{a} \in A^n$ and let $e \in A$ be such that $M_{\mathbf{f}}(\mathbf{a}) \cdot M' = eI_n$ for some matrix $M'$ (with entries in $A$). Assume that $\mathbf{f}(\mathbf{a}) \in eM_{\mathbf{f}}(\mathbf{a})\mathfrak{m}^{(n)}$: i.e., $f(\mathbf{a}) = eM_{\mathbf{f}}(\mathbf{a}) \cdot \mathbf{b}$ for some $\mathbf{b} \in \mathfrak{m}^{(n)}$. Then there is some $\mathbf{a}' \in A^n$ such that $\mathbf{f}(\mathbf{a}') = \mathbf{0}$ and $\mathbf{a}' \equiv \mathbf{a}$ (mod $e\mathfrak{m}$). If $e$ is not a zero-divisor then $\mathbf{a}'$ is unique.*

*Proof.* For a single formal power series $f \in A[[\mathbf{X}]]$, Taylor's theorem [B1, §IV.5.8, Proposition 9] gives

$$f(\mathbf{X} + \mathbf{Y}) = f(\mathbf{X}) + M_f(\mathbf{X}) \cdot \mathbf{Y} + \sum_{1 \le i \le j \le n} G_{ij}(\mathbf{X}, \mathbf{Y})\, Y_i Y_j$$

for some $G_{ij} \in A[[\mathbf{X}, \mathbf{Y}]]$. Arguing as in [B2, §III.4.5], if $f$ is restricted then so are the entries $\partial f / \partial X_i$ of $M_f$ and the $G_{ij}$. Therefore, taking $f = f_i$, one can replace $\mathbf{X}$ with $\mathbf{a}$ and $\mathbf{Y}$ with $e\mathbf{X}$, obtaining

$$\mathbf{f}(\mathbf{a} + e\mathbf{X}) = \mathbf{f}(\mathbf{a}) + M_{\mathbf{f}}(\mathbf{a}) \cdot e\mathbf{X} + e^2\mathbf{R}(\mathbf{X}) = eM_{\mathbf{f}}(\mathbf{a}) \cdot [\mathbf{b} + \mathbf{X} + M'\mathbf{R}(\mathbf{X})], \tag{1}$$

where the remainder terms satisfy $R_i \in (\mathbf{X})^2 A\{\mathbf{X}\}$. Let $\mathbf{h}(\mathbf{X}) = \mathbf{b} + \mathbf{X} + M'R(\mathbf{X})$ and apply condition (H) to $\mathbf{h}$: since $M_{\mathbf{h}}(\mathbf{0}) = I_n$ and $\mathbf{h}(\mathbf{0}) = \mathbf{b} \in \mathfrak{m}^{(n)}$, there is a unique $\mathbf{x} \in \mathfrak{m}^{(n)}$ such that $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Let $\mathbf{a}' = \mathbf{a} + e\mathbf{x}$, so that $\mathbf{a}' \equiv \mathbf{a}$ (mod $e\mathfrak{m}^{(n)}$) and $\mathbf{f}(\mathbf{a}') = \mathbf{0}$.

Now assume that $e$ is not a zero-divisor. Since $M_{\mathbf{f}}(\mathbf{a}) \cdot M' = eI_n$, $J_{\mathbf{f}}(\mathbf{a}) = \det M_{\mathbf{f}}(\mathbf{a})$ is also not a zero-divisor. If $\mathbf{a}' = \mathbf{a} + e\mathbf{x}$ is a root of $\mathbf{f}$, with $\mathbf{x} \in \mathfrak{m}^{(n)}$, then multiplying both sides of Equation (1) (with $\mathbf{X}$ replaced by $\mathbf{x}$) by the adjoint of $M_{\mathbf{f}}(\mathbf{a})$ gives $\mathbf{0} = eJ(\mathbf{a})\mathbf{h}(\mathbf{x})$, which implies $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. This proves uniqueness. □

**Proposition 2.** *Let $A$ be a (discrete) local ring with maximal ideal $\mathfrak{m}$. Then $A$ is Henselian if and only if $(A, \mathfrak{m})$ satisfy condition* (H).

*Proof.* Assume that $A$ is a Henselian local ring. The first step is to reduce to the case $n = 1$, using the structure theory of étale $A$-algebras. Following the notation of Raynaud [R], let $C = A[\mathbf{X}]$, $I = (\mathbf{f}) \subseteq C$, and $B = (C/I)_J = A[\mathbf{X}, T]/(\mathbf{f}, TJ - 1)$. If $R$ is any $A$-algebra then $\mathrm{Hom}_A(B, R)$ is the set of $\mathbf{a} \in R^n$ such that $\mathbf{f}(\mathbf{a}) = \mathbf{0}$ and $J(\mathbf{a}) \in R^\times$. Thus we must show that the canonical map $\mathrm{Hom}_A(B, A) \longrightarrow \mathrm{Hom}_A(B, k)$ is an isomorphism, where $k = A/\mathfrak{m}$.

According to the Jacobian criterion [R, V, Théorème 5] and the fact that being étale is a local condition [R, II, Proposition 6], $B$ is an étale $A$-algebra. The question is local on $\mathrm{Spec}\, B$: choosing a homomorphism $\phi : B \longrightarrow k$ determines a prime ideal $\mathfrak{q} = \ker \phi$, and lifting $\phi$ to a map $B \longrightarrow A$ is the same as lifting it to a map

$B_g \longrightarrow A$, for any $g \in B - \mathfrak{q}$. By the local structure theorem [R, V, Théorème 1], one can choose $g$ so that $B_g$ is isomorphic to a standard étale $A$-algebra:

$$B_g \cong B' = \big(A[X]/(f)\big)_h$$

for a single monic polynomial $f \in A[X]$ (in one variable), where $f'$ is invertible in $B'$. This completes the reduction step.

The surjectivity of $\mathrm{Hom}_A(B', A) \longrightarrow \mathrm{Hom}_A(B', k)$ is just a translation of [R, VII, Proposition 3]. Injectivity follows from a standard argument: if $f(a+x) = f(a) = 0$ with $x \in \mathfrak{m}$ then Taylor's theorem gives $x[f'(a) + G(a,x)x] = 0$. Since $f'(a)$ is a unit and $x \in \mathfrak{m}$, the quantity in brackets is a unit, which implies that $x = 0$.

Conversely, if $(A, \mathfrak{m})$ satisfy condition (H) then, taking $n = 1$, the converse direction of [R, VII, Proposition 3] shows that $A$ is Henselian. □

## 3. The Inverse and Implicit Function Theorems

Theorems 3 and 4 below are improvements of Théorème 2 and Corollaire 3 of [B2, §III.4.5], to which one may refer for details of the proofs. Theorem 3 is an algebraic version of the Inverse Function Theorem; Theorem 4 is an algebraic version of the Implicit Function Theorem. Taking $r = 0$ in Theorem 4 gives Hensel's Lemma in several variables. (Instead of re-proving this theorem, one can apply Theorem 1 to produce an exact root and then use the version in [B2].)

**Theorem 3.** *Let $A$ and $\mathfrak{m}$ satisfy Hensel's conditions. Let $\mathbf{f}$ be an $n$-tuple of restricted power series $f_i \in A\{\mathbf{X}\}$ and let $\mathbf{a} \in A^n$. Let $e \in A$ and let $M'$ be an $n \times n$ matrix (with entries in $A$) such that $M_{\mathbf{f}}(\mathbf{a}) \cdot M' = eI_n$. There is an $n$-tuple $\mathbf{g}$, with $g_i \in (\mathbf{X})A\{\mathbf{X}\}$, such that*

(i) $M_{\mathbf{g}}(0) = I_n$.

(ii) *For all $\mathbf{x} \in A^n$,*

$$\mathbf{f}(\mathbf{a} + e\mathbf{x}) = \mathbf{f}(\mathbf{a}) + M_{\mathbf{f}}(\mathbf{a}) \cdot e\mathbf{g}(\mathbf{x}).$$

(iii) *Let $\mathbf{h}$ be the $n$-tuple of formal power series (not necessarily restricted) $h_i \in A[[\mathbf{X}]]$ such that $\mathbf{g} \circ \mathbf{h} = \mathbf{1}_n$. For all $\mathbf{y} \in \mathfrak{m}^{(n)}$,*

$$\mathbf{f}(\mathbf{a} + e\mathbf{h}(\mathbf{y})) = \mathbf{f}(\mathbf{a}) + M_{\mathbf{f}}(\mathbf{a}) \cdot e\mathbf{y}.$$

*Proof.* Taking $\mathbf{g} = \mathbf{1}_n + M' \cdot \mathbf{R}$, (i) and (ii) follow from Equation (1); and (iii) follows by replacing $\mathbf{x}$ with $\mathbf{h}(\mathbf{y})$. □

**Theorem 4.** *Let $A$ and $\mathfrak{m}$ satisfy Hensel's conditions. Let $\mathbf{f} = (f_{r+1}, \dots, f_n)$ be an $(n-r)$-tuple of restricted power series $f_i \in A\{\mathbf{X}\}$ and let $\mathbf{a} \in A^n$. Let $e \in A$ and let $M'$ be an $(n-r) \times (n-r)$ matrix (with entries in $A$) such that $M_{\mathbf{f}}^{(r-n)}(\mathbf{a}) \cdot M' = eI_{n-r}$. Assume that $\mathbf{f}(\mathbf{a}) = eM_{\mathbf{f}}^{(r-n)}(\mathbf{a}) \cdot \mathbf{b}$ for some $\mathbf{b} \in \mathfrak{m}^{(n-r)}$. Then there are $n - r$ power series $\phi_i \in (\mathbf{X}^{(r)})A[[\mathbf{X}^{(r)}]]$ $(r < i \le n)$ such that, for all $\mathbf{t} \in \mathfrak{m}^{(r)}$,*

$$\mathbf{f}\big(\mathbf{a}^{(r)} + e^2\mathbf{t}, \mathbf{a}^{(r-n)} + e\phi(\mathbf{t})\big) = 0.$$

*Proof.* Apply Theorem 3 to $\mathbf{u} = (\mathbf{X}^{(r)} - \mathbf{a}^{(r)}, \mathbf{f})$. Note that $M_{\mathbf{u}} = \begin{bmatrix} I_r & 0 \\ M_{\mathbf{f}}^{(r)} & M_{\mathbf{f}}^{(r-n)} \end{bmatrix}$, so that $M_{\mathbf{u}}(\mathbf{a}) \begin{bmatrix} eI_r & 0 \\ -M'M_{\mathbf{f}}^{(r)}(\mathbf{a}) & M' \end{bmatrix} = eI_n$. It follows that, in Equation (1), $G_{ij} = 0$ and $R_i = 0$ for $1 \le i \le r$; and, in the proof of Theorem 3, $g_i = h_i = X_i$.

According to Theorem 3(iii),

$$\mathbf{u}\big(\mathbf{a}+e\mathbf{h}(\mathbf{y})\big)=\mathbf{u}(\mathbf{a})+M_{\mathbf{u}}(\mathbf{a})\cdot e\mathbf{y}=M_{\mathbf{u}}(\mathbf{a})\cdot e\begin{bmatrix}\mathbf{y}^{(r)}\\ \mathbf{b}+\mathbf{y}^{(n-r)}\end{bmatrix}.$$

Thus $\mathbf{f}\big(\mathbf{a}+e\mathbf{h}(\mathbf{y})\big)=\mathbf{0}$ if and only if $e\big(M_{\mathbf{f}}^{(r)}(\mathbf{a})\cdot\mathbf{y}^{(r)}+M_{\mathbf{f}}^{(r-n)}(\mathbf{a})\cdot(\mathbf{b}+\mathbf{y}^{(r-n)})\big)=\mathbf{0}$. To guarantee this, it suffices to set $\mathbf{y}^{(r)}=e\mathbf{t}$ and $\mathbf{y}^{(r-n)}=-\mathbf{b}-M'M_{\mathbf{f}}^{(r)}(\mathbf{a})\cdot\mathbf{t}$, with $\mathbf{t}\in\mathfrak{m}^{(r)}$. Therefore set $\phi(\mathbf{X}^{(r)})=\mathbf{h}^{(r-n)}(e\mathbf{X}^{(r)},-\mathbf{b}-M'M_{\mathbf{f}}^{(r)}(\mathbf{a})\cdot\mathbf{X}^{(r)})$; the theorem follows. □

## References


[B1] N. Bourbaki, *Algèbre*, Hermann, Paris, 1959.

[B2] ———, *Algèbre Commutative*, Hermann, Paris, 1962.

[Da-F] R. Dabrowski and B. Fisher, *A stationary-phase formula for exponential sums over* $\mathbb{Z}/p^m\mathbb{Z}$ *and applications to* GL(3)*-Kloosterman sums*, Acta. Arith. (to appear).

[Gr] M. J. Greenberg, *Rational points in Henselian discrete valuation rings*, Pub. Math. IHES **31** (1966), 59–64. MR **34:**7515

[R] M. Raynaud, *Anneaux Locaux Henseliens,* Lecture Notes in Math. 169, Springer-Verlag, Berlin-Heidelberg-New York, 1970. MR **43:**3252



Department of Mathematics, Columbia University, New York, New York 10027

*Current address*: The Bronx High School of Science, 75 West 205th Street, Bronx, New York 10468

*E-mail address*: benji@math.columbia.edu