PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 124, Number 8, August 1996

WHEN IS A p-ADIC POWER SERIES AN ENDOMORPHISM OF A FORMAL GROUP?

HUA-CHIEH LI

(Communicated by William W. Adams)

ABSTRACT. If $f(x)$ is a noninvertible endomorphism of a formal group, then we have that $f(x)$ commutes with an invertible series and $\overline{\mathcal{O}}[[x]]$ is Galois over $\overline{\mathcal{O}}[[f^n(x)]]$ for all $n \in \mathbf{N}$. We shall prove that the converse of this statement is also true.

1. INTRODUCTION

Let K be an algebraic extension of $\mathbf{Q}_p$, and let $\mathcal{O}$ be its integer ring with maximal ideal $\mathcal{M}$. If $\overline{K}$ is an algebraic closure of K, we denote by $\overline{\mathcal{O}}$ and $\overline{\mathcal{M}}$ the integral closure of $\mathcal{O}$ in $\overline{K}$ and the maximal ideal of $\overline{\mathcal{O}}$, respectively. The set of all power series over $\mathcal{O}$ without constant term is a monoid (noncommutative, associative, with unit) under composition. We denote the composition of two functions by $f \circ g(x) = f(g(x))$. We denote by $f^n(x)$ the n-fold composition of $f(x)$ with itself. If the inverse of $f(x)$ exists, then we denote it by $f^{-1}(x)$. Note that $f^n(x)$ does not mean $f(x)$ raised to the n-th power, a function which we will denote by $(f(x))^n$. A series $h(x) \in \mathcal{O}[[x]]$ without constant term is called invertible if there exists a series $g(x) \in \mathcal{O}[[x]]$ such that $h \circ g(x) = x$. A necessary and sufficient condition for $h(x)$ to be invertible is that $h'(0) = \mathcal{O}^*$. If $f(x) \in \mathcal{O}[[x]]$ without constant term and $0 \neq f'(0) \in \mathcal{M}$, then we call $f(x)$ a noninvertible stable series.

Let $f(x) \in \mathcal{O}[[x]]$, a noninvertible power series, such that the extension of the ring $\overline{\mathcal{O}}[[x]] \supset \overline{\mathcal{O}}[[f(x)]]$ is Galois with Galois group Γ. The set Γ is in one-to-one correspondence with the set of roots of $f(x)$, all of which lie in $\overline{\mathcal{M}}$; say that for $\gamma \in \Gamma$, γ corresponds to the root ρ_γ of $f(x)$. These satisfy, for each $\gamma : x^\gamma = g_\gamma(x) \in \overline{\mathcal{O}}[[x]]$, $g_\gamma(0) = \rho_\gamma$ and $f(g_\gamma(x)) = f(x)$. It is easy to check that in this case all roots of $f(x)$ are simple. If $f(x)$ is a noninvertible endomorphism of a formal group $\mathcal{F}(x, y)$, then since the endomorphism ring of $\mathcal{F}$ always contains $\mathbf{Z}_p$, f commutes with an invertible series. For every n, the extension of the ring $\overline{\mathcal{O}}[[x]] \supset \overline{\mathcal{O}}[[f^n(x)]]$ is Galois with Galois group Γ_n. The group Γ_n is isomorphic to the group of roots of $f^n(x)$. For each $\gamma \in \Gamma_n$, $x^\gamma = \mathcal{F}(\rho_\gamma, x)$, where ρ_γ is the root of $f^n(x)$ which corresponds to γ. In this paper we shall find that this is also a sufficient condition for a noninvertible series to be an endomorphism of a formal group.

Received by the editors June 25, 1994 and, in revised form, February 9, 1995.
1991 *Mathematics Subject Classification.* Primary 11S99; Secondary 11S31, 14L05.

2. Notational conventions and basic tools

Recall that K is a field which is complete with respect to a valuation, v. We normalize the valuation v such that $v(\pi) = 1$, where π is a generator of $\mathcal{M}$. There is a unique extension of v to $\overline{K}$, and this will likewise be denoted v.

When $f(x) \in \mathcal{O}[[x]]$, but not all coefficients of $f(x)$ are in $\mathcal{M}$, then the lowest degree in which a unit coefficient appears will be called the Weierstrass degree of $f(x)$, denoted $\mathrm{wideg}(f)$. According to the Weierstrass Preparation Theorem there exist a unit power series $U(x) \in \mathcal{O}[[x]]$ and a distinguished polynomial $P(x) \in \mathcal{O}[[x]]$ such that $f(x) = P(x)U(x)$ and $\deg(P) = \mathrm{wideg}(f)$.

The Newton polygon is a natural tool to study the roots of p-adic power series (see Koblitz [1]). Another geometric object, which contains some information as the Newton polygon, is the *Newton copolygon.* Let $f(x) = \sum_{i=1}^{\infty} a_i x^i$. The Newton copolygon of $f(x), \mathcal{N}^*(f)$, is defined to be the intersection in the Cartesian plane of all half-planes defined by the inequalities $y \le ix + v(a_i)$.

The valuation function of $f(x)$, denoted $\Psi_f(x)$, is a real-valued polygonal function defined for nonnegative values whose graph is the upper boundary of the Newton copolygon. We shall see that for any $\alpha \in \overline{\mathcal{M}}$, the relation $v(f(\alpha)) \ge \Psi_f(v(\alpha))$ holds. If $v(\alpha)$ is not the x-coordinate of any vertex of $\mathcal{N}^*(f)$, then the equality $v(f(\alpha)) = \Psi_f(v(\alpha))$ holds.

Definition. Let $f(x)$ be a noninvertible stable series in $\mathcal{O}[[x]]$. We define $\Lambda(f) = \{\alpha \in \overline{\mathcal{M}} | f^n(\alpha) = 0, \text{ for some } n\}$, the set of all roots of iterates of $f(x)$.

An *f-consistent sequence* is a sequence $(\alpha_1, \alpha_2, \dots)$ of elements of $\overline{\mathcal{M}}$ with $\alpha_1 \ne 0$, $f(\alpha_1) = 0$, and for all $i > 1$, $f(\alpha_i) = \alpha_{i-1}$.

If $f(x) \in \mathcal{O}[[x]]$ is a noninvertible series with $\mathrm{wideg}(f) = d < \infty$, then the valuation function of $f(x), \Psi_f(x)$, is a strictly increasing polygonal function with finitely many segments. The leftmost segment is the line $y = dx$, and the rightmost segment continuing to infinity is the line $y = x + v(f'(0))$. All the segments of Ψ_f lie entirely above the line $y = x$ (i.e. $\Psi_f(x) > x$). Let $(\alpha_1, \alpha_2, \dots)$ be an f-consistent sequence. Since $v(\alpha_i) = v(f(\alpha_{i+1})) \ge \Psi_f(v(\alpha_{i+1})) > v(\alpha_{i+1})$, there exists m such that when $j > m$, $\mathcal{N}^*(f(x) - \alpha_j)$ is the intersection of $y \le dx$ and $y \le v(\alpha_i)$. The x-coordinate of the only vertex of $\mathcal{N}^*(f(x) - \alpha_j)$ is $v(\alpha_j)/d$. Thus $v(\alpha_{j+1}) = v(\alpha_j)/d$. This tells us that $\lim_{j\to\infty} v(\alpha_j) = 0$.

On the ring $K[[x]]$ there are rank-one valuations of a particularly simple kind. If ρ is any nonnegative real number and $f(x) = \sum_{i=0}^{\infty} a_i x^i$, then we may define $w_\rho(f) = \Psi_f(\rho) = \inf\{v(a_i) + i\rho\}_i$. Define $\mathbf{A}_\rho$ to be the subring of $K[[x]]$ such that for a series $h(x) \in \mathbf{A}_\rho$, $w_\rho(h) > -\infty$. Informally, $\mathbf{A}_\rho$ is the set of K-series whose coefficients grow in a controlled enough manner that we may substitute an element $\alpha \in \overline{\mathcal{M}}$ for the variable whenever $v(\alpha) > \rho$. We are interested in the intersection of all the rings $\mathbf{A}_\rho$ for positive ρ. Let us call this ring $\mathbf{A}$. Let $f(x)$ be a noninvertible stable series in $\mathcal{O}[[x]]$, with finite Weierstrass degree. Then there is a unique power series $L_f(x) \in \mathbf{A}$ with $L_f(x) \equiv x \pmod{x^2}$ and $L_f(f(x)) = f'(0) \cdot L_f(x)$. Furthermore, $L_f(x) = \lim_{n\to\infty}(f^n(x)/f'(0)^n)$. This convergence is with respect to w_ρ for all $\rho > 0$ (Lubin [2, Proposition 1.2 and Proposition 2.2]).

If $h_1 \in \mathbf{A}$ and $h_2(0) \in \mathcal{M}$, we may not have $(h_1 \circ h_2)(\alpha) = h_1(h_2(\alpha))$ for all $\alpha \in \overline{\mathcal{M}}$. But we do have that $(h_1 \circ h_2)(0) = h_1(h_2(0))$, because $h_1(h_2(0))$ is just the constant term of $h_1 \circ h_2$. Let $f_n \in \mathbf{A}$ and $\lim_{n\to\infty} f_n = f$ with respect to w_ρ

for all ρ. We have $\lim_{n\to\infty} f_n(\alpha) = f(\alpha)$ for all $\alpha \in \overline{\mathcal{M}}$. For this reason we have the following results:

Lemma 2.1. (1) *If $h_1, h_2 \in \mathcal{O}[[x]]$ and $\alpha \in \overline{\mathcal{M}}$, then $(h_1 \circ h_2)(\alpha) = h_1(h_2(\alpha))$.*

(2) *Let $f_n \in \mathbf{A}$ and $g(x) \in \overline{K}[[x]]$ with $g(0) \in \overline{\mathcal{M}}$. If $\lim_{n\to\infty} f_n = f$ with respect to w_ρ for all $\rho > 0$, then $\lim_{n\to\infty} f_n \circ g = f \circ g$ in the sense of coefficientwise convergence.*

Proof. (1) Denote $h_{1,n}$ as the n-bud of h_1 (the polynomial consisting of all terms of h_1 of degree $\leq n$) and $h_{2,n}$ as the n-bud of h_2. Since $h_1, h_2 \in \mathcal{O}[[x]]$, we have $\lim_{n\to\infty} h_{1,n} = h_1, \lim_{n\to\infty} h_{2,n} = h_2$ and $\lim_{n\to\infty} h_{1,n} \circ h_{2,n} = h_1 \circ h_2$, with respect to w_ρ for all $\rho > 0$. Therefore $\lim_{n\to\infty} (h_{1,n} \circ h_{2,n})(\alpha) = (h_1 \circ h_2)(\alpha)$. On the other hand, since $h_{1,n}$ and $h_{2,n}$ are polynomials, we have $(h_{1,n} \circ h_{2,n})(\alpha) = h_{1,n}(h_{2,n}(\alpha))$. Therefore $\lim_{n\to\infty} (h_{1,n} \circ h_{2,n})(\alpha) = \lim_{n\to\infty} h_{1,n}(h_{2,n}(\alpha)) = h_1(h_2(\alpha))$. Our claim follows.

(2) The constant term of $\lim_{n\to\infty} f_n \circ g$ is $\lim_{n\to\infty} f_n(g(0)) = f(g(0))$, which is the constant term of $f \circ g$. The first-degree coefficient (coefficient of x) of $\lim_{n\to\infty} f_n \circ g$ is $\lim_{n\to\infty} f_n'(g(0))g'(0) = f'(g(0))g'(0)$, because $\lim_{n\to\infty} f_n' = f'$ with respect to w_ρ for all $\rho > 0$. Also, $f'(g(0))g'(0)$ is the first-degree coefficient of $f \circ g$. Therefore $\lim_{n\to\infty} f_n \circ g$ and $f \circ g$ have the same first-degree coefficients. Using this method inductively, our claim follows. □

3. Main theorem

We denote by $f'^{(n)}$ the n-th derivative of f.

Lemma 3.1. *Let $f(x) \in \mathcal{O}[[x]]$ be a noninvertible stable series, and let $\alpha \in \overline{\mathcal{M}}$ be a root of $f(x)$. Then $f'(\alpha) \neq 0$ if and only if there is a unique power series $g_\alpha \in \overline{K}[[x]]$ with $g_\alpha(0) = \alpha$ and $f(g_\alpha(x)) = f(x)$.*

Proof. Let $f(g(x)) = f(x)$ with $g(0) = \alpha$. By considering $f'(g(0))g'(0) = f'(0)$, we have $f'(g(0)) = f'(\alpha) \neq 0$ (because $f'(0) \neq 0$).

For the converse, assume $f(g(x)) = f(x)$ with $g(0) = \alpha$. Since $f'(\alpha) \neq 0$ and $f'(\alpha)g'(0) = f'(0)$, it implies $g'(0) = f'(0)/f'(\alpha)$. Hence the first-degree coefficient of $g(x)$ can be decided uniquely. By induction, suppose that $g'^{(j)}(0)$ has been defined for all $j < n$, such that $f(x) \equiv f(g(x)) \pmod{x^n}$. Consider higher order derivatives,

$$f'^{(n)} = (f' \circ g)g'^{(n)} + \sum_{i=2}^{n-1} \sum_{j_1+\cdots+j_i=n} C_{j_1\cdots j_i}(f'^{(i)} \circ g) \cdot g'^{(j_1)} \cdots g'^{(j_i)} + (f'^{(n)} \circ g)(g')^n,$$

where $C_{j_1\cdots j_i}$ is some integer. This means

$$g'^{(n)}(0) = (f'^{(n)}(0) - \cdots - f'^{(n)}(\alpha)g'(0)^n)/f'(\alpha).$$

This proves existence and uniqueness. □

Lemma 3.2. *Let $F(x,y) = L_f^{-1}(L_f(x) + L_f(y)) = x + f_1(x)y + f_2(x)y^2 + \cdots$. If $L_f'(x) \in \mathcal{O}[[x]]$, then $\forall m \in \mathbf{N}, m!\, f_m(x)$ is in $\mathcal{O}[[x]]$.*

Proof. Since the constant term of $L_f'(x)$ is 1, $L_f'(x) \in \mathcal{O}[[x]]$ implies $1/L_f'(x) \in \mathcal{O}[[x]]$. Because $L_f^{-1} \circ L_f(x) = x$, by taking the derivative, $((L_f^{-1})' \circ L_f) \circ L_f' = 1$.

Hence $(L_f^{-1})' \circ L_f = 1/L'_f \in \mathcal{O}[[x]]$. By induction, suppose that $(L_f^{-1})'^{(i)} \circ L_f \in \mathcal{O}[[x]]$, for all $i < n$. By taking the n-th derivative on both sides of $L_f^{-1} \circ L_f = x$, we can get

$$((L_f^{-1})'^{(n)} \circ L_f) \cdot (L'_f)^n + \sum_{i=1}^{n-1} \sum_{j_1+\cdots+j_i=n} C_{j_1\cdots j_i}((L_f^{-1})'^{(i)} \circ L_f) \cdot L_f'^{(j_1)} \cdots L_f'^{(j_i)} = 0.$$

Since $(L_f^{-1})'^{(i)} \circ L_f$ and $L_f'^{(j)} \in \mathcal{O}[[x]]$, it follows that $(L_f^{-1})'^{(n)} \circ L_f(x) = h_n(x)/(L'_f(x))^n$ for some $h_n(x) \in \mathcal{O}[[x]]$. Hence $(L_f^{-1})'^{(n)} \circ L_f(x) \in \mathcal{O}[[x]]$.

By considering partial differentiation with respect to y of $F(x, y)$, we have

$$\frac{\partial^m}{\partial y^m} F(x, 0) = \sum_{i=1}^{m} \sum_{j_1+\cdots+j_i=m} C_{j_1\cdots j_i}((L_f^{-1})'^{(i)} \circ L_f(x)) \cdot L_f'^{(j_1)}(0) \cdots L_f'^{(j_i)}(0).$$

Since $(L_f^{-1})'^{(i)} \circ L_f(x) \in \mathcal{O}[[x]]$ and $L_f'^{(j)}(0) \in \mathcal{O}$, we have $\partial^m F/\partial y^m(x, 0) \in \mathcal{O}[[x]]$. Thus $m!\, f_m(x) \in \mathcal{O}[[x]]$. □

Let h_1, h_2 and h_3 be power series in $K[[x]]$ without constant term. Then we have the associative rule, $(h_1 \circ h_2) \circ h_3 = h_1 \circ (h_2 \circ h_3)$. Associativity may not be applied when they have constant terms. However, we still have the associative rule if $h_1, h_2 \in \mathcal{O}[[x]]$ without constant term and the constant term of $h_3(x)$ is in $\overline{\mathcal{M}}$. This can be checked by taking every order of derivatives of $(h_1 \circ h_2) \circ h_3$ and $h_1 \circ (h_2 \circ h_3)$ and by using the fact that $(h_1 \circ h_2)(\alpha) = h_1(h_2(\alpha))$ for all $\alpha \in \overline{\mathcal{M}}$. If $g(x) \in K[[x]]$ with constant term in $\overline{\mathcal{M}}$, by considering

$$L_f(f \circ g) = \lim_{n\to\infty} (f^n(f \circ g)/f'(0)^n) = \lim_{n\to\infty} f'(0)(f^{n+1}(g)/f'(0)^{n+1}) = f'(0)L_f(g)$$

(in the sense of coefficientwise convergence, see Lemma 2.1), we have $(L_f \circ f) \circ g = L_f \circ (f \circ g)$.

Lemma 3.3. *If $\overline{\mathcal{O}}[[x]] \supset \overline{\mathcal{O}}[[f(x)]]$ is a Galois extension with Galois group Γ and if $L'_f(x) \in \mathcal{O}[[x]]$, then for $\gamma \in \Gamma$, we have $g_\gamma(x) = F(\rho_\gamma, x)$.*

Proof. For any $\gamma \in \Gamma$, we have $g_\gamma(x) \in \overline{\mathcal{O}}[[x]]$ with $g_\gamma(0) = \rho_\gamma$, where $f(\rho_\gamma) = 0$ and $f(g_\gamma(x)) = f(x)$. By Lemma 3.1, $f'(\rho_\gamma) \neq 0$. From Lemma 3.2, we have $F(\rho_\gamma, x) = \rho_\gamma + f_1(\rho_\gamma)x + f_2(\rho_\gamma)x^2 + \cdots \in \overline{K}[[x]]$. Since

$$\begin{aligned}(L_f \circ f) \circ (F(x, y)) &= (L_f \circ f) \circ (L_f^{-1} \circ (L_f(x) + L_f(y))) \\ &= f'(0) \cdot (L_f(x) + L_f(y)) = L_f(f(x)) + L_f(f(y)),\end{aligned}$$

we have $L_f \circ (f \circ F(\rho_\gamma, x)) = (L_f \circ f) \circ (F(\rho_\gamma, x)) = L_f(f(x))$. Both $f \circ F(\rho_\gamma, x)$ and $f(x)$ have no constant terms; by composing with L_f^{-1}, we have $f \circ F(\rho_\gamma, x) = f(x)$. For $F(\rho_\gamma, 0) = \rho_\gamma = g_\gamma(0)$, by uniqueness (Lemma 3.1), $F(\rho_\gamma, x) = g_\gamma(x)$ follows. □

Lemma 3.4. *Let $h(x)$ be a power series in $K[[x]]$ and $m \cdot h(x) \in \mathcal{O}[[x]]$, for some $m \in \mathbf{N}$. If there exists a sequence $(\alpha_1, \alpha_2, \dots)$ in $\overline{\mathcal{M}}$ such that $\lim_{n\to\infty} v(\alpha_n) = 0$ and $v(h(\alpha_n)) \geq 0$, for all $n \in \mathbf{N}$, then $h(x) \in \mathcal{O}[[x]]$.*

Proof. Let $\Psi_h(x)$ be the valuation function of $h(x)$. We know that $\Psi_h(x)$ is a continuous increasing function and $v(h(\alpha)) = \Psi_h(v(\alpha))$ if $v(\alpha)$ is not equal to the valuation of any root of $h(x)$. Since $m \cdot h(x) \in \mathcal{O}[[x]]$, $h(x)$ has only finitely many roots in $\overline{\mathcal{M}}$. We have $\lim_{n\to\infty} v(h(\alpha_n)) = \lim_{n\to\infty} \Psi_h(v(\alpha_n)) = \Psi_h(0) \geq 0$. Thus $\Psi_h(x) \geq 0$. □

Lemma 3.5. *Let $f(x) \in \mathcal{O}[[x]]$ be a noninvertible stable series which commutes with an invertible series in $\mathcal{O}[[x]]$. Then every root of $f'(x)$ in $\overline{\mathcal{M}}$ is a root of an iterate of $f(x)$.*

Proof. See Lubin [2, Corollary 3.2.1]. □

Theorem 3.6. *Let $f(x) \in \mathcal{O}[[x]]$ be a noninvertible stable series which commutes with an invertible series in $\mathcal{O}[[x]]$. If $\overline{\mathcal{O}}[[x]]$ is Galois over $\overline{\mathcal{O}}[[f^n(x)]]$, for all $n \in \mathbf{N}$, then $f(x)$ is an endomorphism of a formal group over $\mathcal{O}$.*

Proof. Since $f(x)$ commutes with an invertible power series, every root of $f'(x)$ is a root of $f^m(x)$ for some m. But since $f^m(x)$ has only simple roots for all $m \in \mathbf{N}$, it implies that $f'(x)$ has no root. Hence $\text{wideg}(f')$ is either 0 or infinity. But since $f'(0) \in \mathcal{M}$, it implies that $\text{wideg}(f') = \infty$. Therefore $f'(x) = f'(0) \cdot h(x)$, for some $h(x) \in \mathcal{O}[[x]]$ with $\text{wideg}(h) = 0$. Thus $f'(x)/f'(0) \in \mathcal{O}[[x]]$. By the same reasoning, we see $(f^n)'(x)/(f'(0))^n \in \mathcal{O}[[x]]$. Hence $L'_f(x) \in \mathcal{O}[[x]]$.

Let $F(x,y) = L_f^{-1} \circ (L_f(x) + L_f(y)) = x + f_1(x)y + f_2(x)y^2 + \cdots$. We claim that $f_n(x) \in \mathcal{O}[[x]]$, for all n. By Lemma 3.2, we have $\partial^n F/\partial y^n(x,0) = n!\, f_n(x) \in \mathcal{O}[[x]]$. Let Γ_m be the Galois group of $\overline{\mathcal{O}}[[x]]$ over $\overline{\mathcal{O}}[[f^m(x)]]$. By Lemma 3.3, we have $F(\rho_\gamma, y) = g_\gamma(y) \in \overline{\mathcal{O}}[[y]]$, $\forall \gamma \in \Gamma_m$. Therefore $f_n(\alpha) \in \overline{\mathcal{O}}$, for all $\alpha \in \Lambda(f)$. It follows that $f_n(x) \in \mathcal{O}[[x]]$, by Lemma 3.4. □

Acknowledgment

The work presented here is part of the author's 1994 Brown Ph.D. thesis. Without Professor Rosen's continued help and encouragement, none of this work would have been possible. Professor Lubin was the one who introduced the author to the field of p-adic dynamical systems. His guidance in this research was indispensable.

References

1. N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, Springer-Verlag, New York, 1977. MR **57:**5964
2. J. Lubin, *Nonarchimedean dynamical systems*, Compositio Math. **94** (1994), 321–346. CMP 95:06

Department of Mathematics, Brown University, Providence, Rhode Island 02912

Current address: Department of Mathematics, National Tsin Hua University, Hsin Chu, Taiwan, R.O.C.

E-mail address: li@math.nthu.edu.tw