J. Math. Soc. Japan
Vol. 34, No. 4, 1982

On two fundamental theorems for the concept of approximate roots

By T. T. MOH*)

(Received April 15, 1981)

Introduction.

In our previous work [1, 2, 3 in collaboration with S. S. Abhyankar] we introduced the concept of approximate root which has been equivalently defined as follows (cf. [7]).

DEFINITION. Let $\theta=y^{-1}$ and $R[y]\subset R((\theta))$. Let $f(y)$ be a monic polynomial. Let d be a unit in R and a factor of $\deg f(y)$. Then the d-th root of $f(y)$ exists in $R((\theta))$ and let it be $\theta^{-(n/d)}+a_1\theta^{-(n/d)+1}+\cdots+a_{n/d}+a_{-1}\theta+\cdots$. The d-th approximate root, $g_d(y)$, of $f(y)$ is defined to be

$$y^{(n/d)}+a_1y^{(n/d)-1}+\cdots+a_{n/d}.$$

The central theorem proved in [1, 2] (cf. [2. §7]) is the following:

THEOREM. *Let $f(y, x)$ be a monic irreducible polynomial in y with coefficients in $K((x))$, where K is an algebraically closed field of char p. Let d_r be a characteristic g.c.d. with $p\nmid d_r$. Let $g_{d_r}(y)$ be the d_r-th approximate root of $f(y)$. Then*

$$\mathrm{ord}_T\, g_{d_r}(y(T), T^n)=\lambda_r/d_r$$

where $x=T^n$ and $y(T)$ is a solution of $f(y)$ in $K((T))$.

The embedding theorem of affine lines [3, 6] follows from the above theorem. Later on S. S. Abhyankar has given a simplified version of [1, 2] in [4] and we published a generalized version of the above theorem in [7]. In our generalized version, we drop the irreducible restriction on the polynomial $f(y)$ and replace the field $K((x))$ by any field with a real discrete valuation.

H. Hironaka has used the concept of approximate roots in his work on the resolution of singularities (cf. [5]). We have used a stronger version of the above theorem in our work ([8]) on the Jacobian conjecture, in fact, a part of Theorem 2 has been announced in [8]. In this article besides dropping the irreducible restriction on the polynomial $f(y)$ we prove the above theorem for any non-archimedean valued field. Moreover, we establish a strong property of the

*) The author was partially supported by N.S.F. at Purdue University.

g_d-expression of $f(y)$ (cf. Theorems 1 and 2 in § 3).

It is in fact shorter and simpler to prove the general theorems. Only a fraction of work as comparing with [1, 2, 4] is required. Conceptually, the proof of the present article is easier than [7]. The cumbersome method of Newton polygon is gotten rid of. The present method only involves some counting principles. Due to the non-availability of a suitable reference, we include all necessary materials in the present article to make it self-contained. For the applications, the reader is referred to [3, 4, 7, 8]. We want to thank Laura Zeman for typing this manuscript.

§ 1. Preliminaries.

Let K be any non-archimedean valued field, i.e. K is a field together with a map $\underline{v}: K \twoheadrightarrow G \subset \boldsymbol{R}_+$ the non-negative real numbers such that $\underline{v}(\tau_1 \cdot \tau_2) = \underline{v}(\tau_1) \cdot \underline{v}(\tau_2)$, $\underline{v}(\tau_1 + \tau_2) \leqq \max(\underline{v}(\tau_1), \underline{v}(\tau_2))$ for all τ_1, τ_2 in K and $\underline{v}(\tau) = 0 \Leftrightarrow \tau = 0$.

By a disc D we mean a closed disc, i.e. $D = \{\tau : \underline{v}(\tau - \tau^*) \leqq r_1\}$ for some $\tau^* \in K$ and some $r_1 \in G$. The number r_1 is the radius of D, in symbol $r(D)$. Note that it follows from the non-archimedean property of the value $\underline{v}$ every element in D can serve as a center of D.

Let $f(y)$ be a polynomial in $K[y]$ which splits completely over K. We introduce the concept of *the tree of discs* $\mathcal{D}(f)$ of $f(y)$ as follows. Let

$$f(y) = a \prod_{i=1}^{n} (y - \tau_i)$$

and D_1 be the minimal disc which contains all τ_i's. We define

$$D_1 \in \mathcal{D}(f).$$

Successively let us define $D_{1j\cdots l} \in \mathcal{D}(f)$. Let us consider the partition of τ_i's in $D_{1j\cdots l}$ by the following equivalence relation

$$\tau_i \sim \tau_j \Longleftrightarrow \underline{v}(\tau_i - \tau_j) < r(D_{1j\cdots l}).$$

Let the partition be

$$\{\tau_1, \cdots, \tau_n\} \cap D_{1j\cdots l} = \bigcup_m E_{1j\cdots lm}.$$

For each $E_{1j\cdots lm}$ we pick the minimal disc $D_{1j\cdots lm}$ which contains all τ_i's in $E_{1j\cdots lm}$. We define

$$D_{1j\cdots lm} \in \mathcal{D}(f).$$

It is easy to see that there are only finite members in $\mathcal{D}(f)$.

Let $\beta(y)$ be any polynomial which splits completely over K with

$$\beta(y) = b \prod (y - \tau_j^*).$$

Let D be any disc in K. We introduce the multiplicity of $\beta(y)$ in D, in symbol $l(\beta, D)$, as follows

$$l(\beta, D)=\text{the number of } \tau_j^*\text{'s in } D\,.$$

We introduce *the quasi-multiplicity of* $\beta(y)$ *in* D, in symbol $l^*(\beta, D)$, for $D_{1j\cdots l}\supsetneqq D\supset D_{1j\cdots lm}$ where $D_{1j\cdots l}$ and $D_{1j\cdots lm}$ are successive members of $\mathcal{D}(f)$ as follows

$$l^*(\beta, D)=\text{the number of } \tau_j^*\text{'s in } D_{1j\cdots l} \text{ with } \underline{v}(\tau_j^*-D)<r(D_{1j\cdots l})\,.$$

We define *the tree of radii of* $f(y)$ as $\{r(D): D\in\mathcal{D}(f)\}$ and *the tree of multiplicities of* $f(y)$ as $\{l(f, D): D\in\mathcal{D}(f)\}$.

Let us fix a representative set $\{t_\delta: \delta\in G\}$ of the value group of G in K where $\underline{v}(t_\delta)=\delta$. We define *the leading coefficient* of a nonzero element $\tau\in K$ as follows. Let (V, M) be the valuation ring of $\underline{v}$, i.e.

$$V=\{\tau\in K: \underline{v}(\tau)\leqq 1\}\,,$$

$$M=\{\tau\in K: \underline{v}(\tau)<1\}\,.$$

Let Ω be the canonical map from V to V/M. Given any nonzero element $\tau\in K$ with $\underline{v}(\tau)=1/\delta$ we define *the leading coefficient of* τ, in symbol $\mathcal{L}(\tau)$, as $\Omega(t_\delta\tau)$.

We introduce *the general elements* or *the π-elements* of K as follows. Let π be a symbol. We extend the value $\underline{v}$ to $K(\pi)$ by assigning the value 1 to π and extend Ω accordingly. A *general element* or a *π-element* σ is an element in $K[\pi]$ of the following form

$$\sigma=\tau+\pi t_\delta$$

where $\tau\in K$ and δ may or may not be the value of τ. The disc D in K centered at τ with radius δ is the associate disc of σ and σ is a general element for D. We have the following useful algebraic lemmas.

LEMMA 1.1. *Let $f(y)$ be a polynomial which factors completely over K with*

$$f(y)=a\prod_{i=1}^{n}(y-\tau_i)\,.$$

Let $\sigma=\tau+\pi t_\delta$ be a general element and $\mathcal{L}(f(\sigma))$ the leading coefficient of $f(\sigma)$. Then we have

(1) $\mathcal{L}(f(\sigma))=b\prod_{i=1}^{n}\mathcal{L}(\sigma-\tau_i)\,,\quad 0\neq b\in V/M\,,$

(2) $\deg_\pi \mathcal{L}(\sigma-\tau_i)=1$ *iff* $\tau_i\in D$ *the disc associated with* σ ,

(3) $\deg_\pi \mathcal{L}(\sigma-\tau_i)=0$ *iff* $\tau_i\notin D$,

(4) $\mathcal{L}(\sigma-\tau_i)=\mathcal{L}(\sigma-\tau_j)$ *and* $\deg_\pi \mathcal{L}(\sigma-\tau_i)=1$
iff $\underline{v}(\tau_i-\tau_j)<\sigma$ *and* $\tau_i\in D$.

PROOF. Routine.

LEMMA 1.2. *Let σ and σ^* be two general elements for disc D and D^*, respectively. Suppose that $D\subset D^*$. Then we have for any $\tau'\in K$ the following*

(1) $\underline{v}(\sigma^*-\tau')\geqq\underline{v}(\sigma-\tau')\geqq\underline{v}(\sigma^*-\tau')\quad(r(D)/r(D^*))$,

(2) $\tau'\notin D^*\Longrightarrow\underline{v}(\sigma^*-\tau')=\underline{v}(\sigma-\tau')$,

(3) $\underline{v}(\sigma-\tau')=r(D^*)\Longrightarrow\underline{v}(\sigma^*-\tau')=\underline{v}(\sigma-\tau')$,

(4) $\tau'\in D\Longrightarrow\underline{v}(\sigma-\tau')=\underline{v}(\sigma^*-\tau')\quad(r(D)/r(D^*))$.

PROOF. Routine.

LEMMA 1.3. *Let σ be a general element for a disc D. Let the following chain be maximal in $\mathfrak{D}(f)$, i.e. no more members of $\mathfrak{D}(f)$ can be inserted in between*

(1) $D_u\supsetneqq D_{u+1}\supsetneqq\cdots\supsetneqq D_w\supsetneqq D$

with σ_i's as general element for D_i's. Then for any polynomial $\beta(y)$ which splits completely over K we have

(2) $$\underline{v}(\beta(\sigma_u))\prod_{i=u+1}^{w}(r(D_i)/r(D_{i-1}))^{l^*(\beta,D_i)}(r(D)/r(D_w))^{l^*(\beta,D)}$$

$$\leqq\underline{v}(\beta(\sigma))\leqq\underline{v}(\beta(\sigma_u))\prod_{i=u+1}^{w}(r(D_i)/r(D_{i-1}))^{l(\beta,D_i)}(r(D)/r(D_w))^{l(\beta,D)}.$$

PROOF. Clearly it suffices to prove for the case $u=w$ and $\beta(y)$ is linear and monic. Our lemma follows from Lemma 1.2. Q.E.D.

§ 2. The approximate roots and the quasi-approximate roots of a polynomial.

Let $f(y)$ be a monic polynomial of degree n. Let d be a factor of n and $d\neq 0$ in K. Recall the following definition of the d-th approximate root of $f(y)$ from [7].

DEFINITION 2.1. Let $\theta=y^{-1}$ and $K[y]\subset K((\theta))$. Then the d-th root of $f(y)$ exists in $K((\theta))$ and let it be $\theta^{-(n/d)}+a_1\theta^{-(n/d)+1}+\cdots+a_{n/d}+a_{-1}\theta+\cdots$. Then the d-th approximate root, $g_d(y)$, of $f(y)$ is defined to be

$$y^{(n/d)}+a_1y^{(n/d)-1}+\cdots+a_{n/d}.$$

To clarify the significance of the d-th approximate root of $f(y)$, we give the following definitions.

DEFINITION 2.2. A system of discs $\mathcal{U}=\{U_1,\cdots,U_s\}$ is said to be complete for $f(y)=\prod_{i=1}^{n}(y-\tau_i)$ iff the following conditions are satisfied:

(1) $U_i\cap U_j=\emptyset$ for $i\neq j$,

(2) $l(f,U_j)\geqq 1$ for $j=1,\cdots,s$,

(3) $\sum_{j=1}^{s} l(f, U_j)=n$.

A complete system of discs $\mathcal{U}=\{U_1, \cdots, U_s\}$ is said to be coherent iff there exists a system of π-elements $\{\sigma_1, \cdots, \sigma_s\}$ satisfying

(1) σ_j is a general element for U_j,

(2) $\underline{v}(f(\sigma_j))=\lambda$ for $j=1, \cdots, s$.

The common number λ is called the accuracy of $\mathcal{U}=\{U_1, \cdots, U_s\}$.

REMARK 1. It follows from Lemma 1.2 that the coherent condition is independent of the choice of the general elements $\{\sigma_1, \cdots, \sigma_s\}$.

REMARK 2. If the field K is algebraically closed, i.e. the value group $G\setminus\{0\}$ is divisible, then any $\lambda\in G$ can be used as accuracy to determine a unique coherent complete system of discs $\mathcal{U}$.

DEFINITION 2.3. Let $\mathcal{U}=\{U_1, \cdots, U_s\}$ be a coherent complete system of discs for $f(y)=\prod_{i=1}^{n}(y-\tau_i)$. Let d be a factor of n. Suppose that

$$d\,|\,l(f, U_j) \qquad \text{for} \quad j=1, \cdots, s.$$

A monic polynomial $g(y)$ of degree (n/d) is said to be a quasi-approximate root of $f(y)$ with respect to $\mathcal{U}$ iff

$$dl(g, U_j)=l(f, U_j) \qquad \text{for} \quad j=1, \cdots, s.$$

If $d\neq 0$ in V/M, we prove that the d-th approximate root is a d-th quasi-approximate root of $f(y)$, thus guaranteeing the existence of the quasi-approximate roots. We shall prove the following preliminary lemmas.

LEMMA 2.1. *Let $\beta(y)$ be a polynomial which factors completely over K. Let $D\in\mathcal{D}(f)=\{D_{1j\cdots l}\}$ where $f(y)=\prod_{i=1}^{n}(y-\tau_i)$. Let r^* be the number with*

$$l(\beta, D)=r^*l(f, D).$$

Then we can construct a chain starting with D

$$D=D_u^*\supset D_{u+1}^*\supset \cdots \supset D_i^*\supset \cdots$$

with

(1) $D_i^*\in\mathcal{D}(f)$,

(2) *the chain is maximal i.e. no more disc in $\mathcal{D}(f)$ can be added,*

(3) $l(\beta, D_i^*)\leqq l^*(\beta, D_i^*)\leqq r^*l(f, D_i^*)$.

PROOF. Let $\{E_j\}$ be the members of $\mathcal{D}(f)$ which are directly included in $D=D_u^*$, i.e. no member of $\mathcal{D}(f)$ is properly between D and E_j. It follows from the definitions that

$$l(f, D_u^*) = \sum l(f, E_j),$$

$$l(\beta, D_u^*) \geqq \sum l^*(\beta, E_j) \geqq \sum l(\beta, E_j).$$

Then we may choose one of the E_j's as D_{u+1}^*. Clearly we can continue this procedure. Q. E. D.

LEMMA 2.2. *Let $\beta(y)$ be a polynomial which factors completely over K. Let $\mathcal{U} = \{U_1, \cdots, U_s\}$ be a coherent complete system of discs for $f(y) = \prod_{i=1}^{n}(y - \tau_i)$. Suppose that there are two chains* (1) *and* (2) *in $D(f) \cup \mathcal{U}$*

(1) $D_u \supset D_{u+1} \supset \cdots \supset D_v \supset D_{v+1} \supset \cdots \supset D_w = U_i$,

(2) $D_u = D_u^* \supset \cdots \supset D_{w^*}^* = U_j$,

with both chains maximal with respect to the end discs. Let σ_i^* be a general point for D_i^* and r, q be numbers with $0 \leqq r \leqq q$. If we have

(3) $l(\beta, D_i^*) \leqq l^*(\beta, D_i^*) \leqq rl(f, D_i^*) \quad \forall u < i \leqq w^*$,

(4) $\underline{v}(\beta(\sigma_{w^*}^*)) \leqq \underline{v}(f(\sigma_{w^*}^*))^q$,

then we have the following

(5) $\underline{v}(\beta(\sigma_u^*)) \leqq \underline{v}(f(\sigma_u^*))^q$

with the equality only if all inequalities in (3) *and* (4) *are equalities and $r = q$. Moreover, let D be a disc satisfying*

(6) $D_v \supsetneqq D \supset D_{v+1}$

with σ a general point for it. Suppose that we have

(7) $l(\beta, D_u) \leqq rl(f, D_u)$,

(8) $l(\beta, D_j) \geqq rl(f, D_j) \quad \forall u < j \leqq v$,

(9) $l(\beta, D) \geqq rl(f, D)\ (= rl(f, D_{v+1}))$.

Then we have

(10) $\underline{v}(\beta(\sigma)) \leqq \underline{v}(f(\sigma))^q$

with the equality only if all inequalities in (3), (4), (7), (8), (9) *are equalities and $r = q$.*

PROOF. Let σ_j be a general point for D_j for $j = u+1, \cdots, w$. The coherent condition on $\mathcal{U}$ states

(11) $\underline{v}(f(\sigma_{w^*}^*)) = \underline{v}(f(\sigma_w))$.

It follows straightforward from Lemma 1.3 that

(12) $\underline{v}(f(\sigma_{w^*}^*)) = \underline{v}(f(\sigma_u^*)) \prod_{i=u+1}^{w^*} (r(D_i^*)/r(D_{i-1}^*))^{l(f, D_i^*)}$,

(13) $\underline{v}(f(\sigma_w))=\underline{v}(f(\sigma_u^*))\prod_{j=u+1}^{w}(r(D_j)/r(D_{j-1}))^{l(f,D_j)}$.

We deduce at once from (11), (12), (13)

(14) $\prod_{i=u+1}^{w^*}(r(D_i^*)/r(D_{i-1}^*))^{l(f,D_i^*)}=\prod_{j=u+1}^{w}(r(D_j)/r(D_{j-1}))^{l(f,D_j)}$

Lemma 1.3 implies

(15) $\underline{v}(\beta(\sigma_{w^*}^*))\geqq\underline{v}(\beta(\sigma_u^*))\prod_{i=u+1}^{w^*}(r(D_i^*)/r(D_{i-1}^*))^{l^*(\beta,D_i^*)}$

$\geqq\underline{v}(\beta(\sigma_u^*))\prod_{i=u+1}^{w^*}(r(D_i^*)/r(D_{i-1}^*))^{l(\beta,D_i^*)}$.

It then follows from (13), (11), (4), (15), (3), (14)

(16) $\underline{v}(f(\sigma_u^*))^q\prod_{j=u+1}^{w}(r(D_j)/r(D_{j-1}))^{ql(f,D_j)}$

$=\underline{v}(f(\sigma_w))^q$

$=\underline{v}(f(\sigma_{w^*}^*))^q$

$\geqq\underline{v}(\beta(\sigma_{w^*}^*))$

$\geqq\underline{v}(\beta(\sigma_u^*))\prod_{i=u+1}^{w^*}(r(D_i^*)/r(D_{i-1}^*))^{l(\beta,D_i^*)}$

$\geqq\underline{v}(\beta(\sigma_u^*))\prod_{i=u+1}^{w^*}(r(D_i^*)/r(D_{i-1}^*))^{rl(f,D_i^*)}$

$=\underline{v}(\beta(\sigma_u^*))\prod_{j=u+1}^{w}(r(D_j)/r(D_{j-1}))^{rl(f,D_j)}$.

The inequality (5) follows by observing $0\leqq r\leqq q$. Clearly the inequality (5) is an equality only if the inequalities in (3), (4) are equalities and $r=q$.

For the second part of our lemma, it follows from (8), (9) and Lemma 1.3 that

(17) $\underline{v}(\beta(\sigma))\leqq\underline{v}(\beta(\sigma_u^*))\prod_{j=u+1}^{v}(r(D_j)/f(D_{j-1}))^{rl(f,D_j)}\cdot(r(D)/r(D_r))^{rl(f,D_{v+1})}$,

(18) $\underline{v}(f(\sigma))^q=\underline{v}(f(\sigma_u^*))^q\prod_{j=1}^{v}(r(D_j)/r(D_{j-1}))^{ql(f,D_j)}\cdot(r(D)/r(D_v))^{ql(f,D_{v+1})}$.

It follows from our assumption $0\leqq r\leqq q$ and (16) that we have the following

(19) $\underline{v}(f(\sigma_u^*))^q\prod_{j=u+1}^{v+1}(r(D_j)/r(D_{j-1}))^{ql(f,D_j)}$

$\geqq\underline{v}(\beta(\sigma_u^*))\prod_{j=u+1}^{v+1}(r(D_j)/r(D_{j-1}))^{rl(f,D_j)}$.

The inequality (10) follows from (17), (18), (19) by observing

(20) $r(D_{v+1})/r(D_v)\leqq r(D)/r(D_v)$.

Clearly the inequality (10) is an equality only if all inequalities in (3), (4), (7), (8), (9) are equalities and $r=q$. Q. E. D.

LEMMA 2.3. *Let $\beta(y)$ be a polynomial which factors completely over K. Let $\mathcal{U}=\{U_1, \cdots, U_s\}$ be a coherent complete system of discs for $f(y)=\prod_{i=1}^{n}(y-\tau_i)$ with accuracy λ. Let σ_j's be general elements for U_j's and σ a general element for a disc D which properly contains some member of $\mathcal{U}$. If for some fixed positive integer j we have*

(1) $\deg \beta(y)\leqq(1/d)n$,

(2) $\underline{v}(\beta(\sigma_i))\leqq\lambda^{j/d}=\underline{v}(f(\sigma_i))^{j/d}$ *for* $i=1, \cdots, s$,

then we always have

(3) $\underline{v}(\beta(\sigma))\leqq\underline{v}(f(\sigma))^{j/d}$ *with the equality only*

if the inequality in (1) *is an equality. Furthermore a strict inequality for* (1) *and* $l(\beta, U_i)\geqq(1/d)l(f, U_i)$ *imply the inequality in* (2) *in a strict inequality.*

PROOF. We divide the proof of (3) into several cases.

Case 1. Let us assume that

(4) $l(\beta, D)<(1/d)l(f, D)$.

Let D_u^* be the largest member of $\mathcal{D}(f)\cup\mathcal{U}$ which is contained in D. Then we have

(5) $l(\beta, D_u^*)\leqq l(\beta, D)<(1/d)l(f, D)=(1/d)l(f, D_u^*)$.

If $D_u^*=U_i\in\mathcal{U}$, then we have

(10) $\underline{v}(\beta(\sigma_u^*))\leqq\underline{v}(f(\sigma_u^*))^{j/d}$.

Otherwise, let $D_u^*\in\mathcal{D}(f)$. It then follows from Lemma 2.1 that there are two maximal chains of the following form

(6) $D_u\supset\cdots$,

(7) $D_u=D_u^*\supset\cdots\supset D_{w^*}^*=U_j \quad (\supset D_{w^*+1}^*)$

with the following conditions satisfied (with $U_j=D_{w^*+1}^*$ if $U_j\in\mathcal{D}(f)$):

(8) $l(\beta, D_i^*)\leqq l^*(\beta, D_i^*)<(1/d)l(f, D_i^*) \quad i=u+1, \cdots, w^*-1$,

(9) $l(\beta, D_{w^*}^*)\leqq l^*(\beta, D_{w^*}^*)=l^*(\beta, D_{w^*+1}^*)<(1/d)l(f, D_{w^*+1}^*)=(1/d)l(f, D_{w^*}^*)$.

We apply Lemma 2.2 by setting $r=1/d\leqq j/d=q$ and noting that our conditions (2), (8), (9) are equivalent to the conditions (3), (4) in Lemma 2.2. We thus conclude

(10) $\underline{v}(\beta(\sigma_u^*))\leqq\underline{v}(f(\sigma_u^*))^{j/d}$.

On the other hand, it follows from Lemma 1.3 that

(11) $\underline{v}(\beta(\sigma_u^*))\geqq\underline{v}(\beta(\sigma))(r(D_u^*)/r(D))^{l(\beta,D)}$,

(12) $\underline{v}(f(\sigma_u^*))^{j/d}=\underline{v}(f(\sigma))^{j/d}(r(D_u^*)/r(D))^{(j/d)l(f,D)}$.

We conclude from (4), (10), (11), (12)

(13) $\underline{v}(\beta(\sigma))<\underline{v}(f(\sigma))^{j/d}$.

Case 2. Let us consider the possibility that

(14) $l(\beta, D)\geqq(1/d)l(f, D)$,

(15) $D\supset D_1$, the maximal disc in $\mathscr{D}(f)$.

Then we have the following

$$l(\beta, D)\leqq\deg\beta(y)\leqq(1/d)n=(1/d)l(f, D).$$

In light of (14) we conclude

(17) $l(\beta, D)=\deg\beta(y)=(1/d)n=(1/d)l(f, D)$.

The arguments in the case (1) can be repeated verbatim with the strict inequalities in (4), (5), (8), (9), (10) replaced by the inequalities to conclude

(18) $\underline{v}(\beta(\sigma))\leqq\underline{v}(f(\sigma))^{j/d}$.

Note that in the case the inequality in (1) must be an equality.

Case 3. We shall consider the final possibility that

(19) $l(\beta, D)\geqq(1/d)l(f, D)$,

(20) $D\subsetneqq D_v\in\mathscr{D}(f)$.

Certainly we may extend (20) to a maximal chain

(21) $D_u\supset D_{u+1}\supset\cdots\supset D_v\supset D_{v+1}\supset\cdots\supset D_w=U_i$

satisfying the following conditions (22), (23) and (24):

(22) $l(\beta, D_u)\leqq(1/d)l(f, D_u)$,

(23) $l(\beta, D_j)\geqq(1/d)l(f, D_j)\quad\forall u<j\leqq v$,

(24) $D_v\supsetneqq D\supset D_{v+1}$.

Note that if (1) is a strict inequality then we may demand (22) to be a strict inequality. As in the case (1) we construct a second maximal chain

(25) $D_u=D_u^*\supset\cdots\supset D_{w^*}^*=U_j\quad(\supset D_{w^*+1}^*)$

with the following conditions satisfied:

(26) $l(\beta, D_i^*)\leqq l^*(\beta, D_i^*)\leqq(1/d)l(f, D_i^*)$,

(27) $l(\beta, D_{w^*}^*)\leqq l^*(\beta, D_{w^*}^*)\leqq(1/d)l(f, D_{w^*}^*)$.

Lemma 2.2 states

(28) $\underline{v}(\beta(\sigma)) \leqq \underline{v}(f(\sigma))^{j/d}$.

Note that the equality can happen in (28) only if (1) is an equality.

For the last part of our lemma, we observe that $U_i \supsetneqq D_1$ the maximal disc in $\mathcal{D}(f)$ by (1) and the new inequality. Our proof follows case (3) verbatim with D replaced by U_i.

§ 3. Two fundamental theorems.

Our first theorem establishes the existence of the quasi-approximate roots by showing the approximate roots are the quasi-approximate roots. Our second theorem establishes a strong property of the g-adic expansion of $f(y)$ in terms of a quasi-approximate root g.

THEOREM 1. *Let $\mathcal{U}=\{U_1, \cdots, U_s\}$ be a coherent complete system of discs for $f(y)=\prod_{i=1}^{n}(y-\tau_i)$. Let d be a positive integer which is not zero in V/M. Suppose that for $i=1, \cdots, s$ we have*

(1) $d \mid l(f, U_i)$

and the d-th approximate root $g_d(y)$ of $f(y)$ factors completely over K. Then the polynomial $g_d(y)$ is a d-th quasi-approximate root of $f(y)$.

First, we prove the following technical lemma.

LEMMA 3.1. *We shall use the assumptions of Theorem 1. Let σ be a general point for a disc $D \supsetneqq U_i$ for some i. If in the following equation*

$$f(\sigma)=g_d(\sigma)^d+h(\sigma)$$

we have

(1) $\underline{v}(h(\sigma)) \leqq \underline{v}(f(\sigma))$,

(2) $l(h, D)<(d-1/d)l(f, D)$,

then we have

(3) $\underline{v}(h(\sigma))<\underline{v}(f(\sigma))$.

PROOF. It follows from (1) that we have

(4) $\underline{v}(g_d(\sigma))^d \leqq \underline{v}(f(\sigma))$.

It follows from Lemma 1.1 that

(5) $\deg_\pi \mathcal{L}(f(\sigma))=l(f, D)$,

(6) $\deg_\pi \mathcal{L}(h(\sigma))=l(h, D)$.

In view of (2) we conclude that (4) must be an equality. Let us assume the lemma is false, i.e. (1) is an equality. Then we shall deduce a contradiction.

Since both (1) and (4) are assumed to be equalities then we have the following equation of the leading coefficients with a, b nonzeroes in V/M:

(7) $\mathcal{L}(f(\sigma))=a\mathcal{L}(g_d(\sigma))^d+b\mathcal{L}(h(\sigma))$.

We claim

(8) $\mathcal{L}(f(\sigma))=af^*(\pi)^d$.

Note that in view of (2), (5), (6), (7), (8) the polynomial $b\mathcal{L}(h(\sigma))$, which is of degree $<(d-1/d)\deg_\pi \mathcal{L}(f(\sigma))$, would be expressible as the difference of the d-th powers of two polynomials of degree $(1/d)\deg_\pi \mathcal{L}(f(\sigma))$. This is clearly impossible. Here we use the fact that d is nonzero in V/M.

We are left to prove (8). Let $\{E_j\}$ be the set of maximal members of $\mathcal{D}(f)\cup\mathcal{U}$ which are properly contained in D. It is easy to see that every $E_j\supset$ some U_i and the following

(9) $l(f, E_j)=\sum_{E_j\supset U_i} l(f, U_i)$.

Thus we have

$$d\,|\,l(f, E_j)\,.$$

Then (8) follows from Lemma 1.1. Q.E.D.

PROOF OF THEOREM 1. We have the following equations

(2) $f(y)=g_d(y)^d+h(y)$,

(3) $\deg h(y)<(d-1/d)n$.

It suffices to show that for any σ with associate disc $D\supsetneqq U_i$ we have the following

(4) $\underline{v}(h(\sigma))<\underline{v}(f(\sigma))$.

Note that it then follows from (2) that

(5) $\mathcal{L}(f(\sigma))=\mathcal{L}(g(\sigma)^d)$.

In the case that value group G is discrete we may take D to be the minimal disc $\supsetneqq U_i$. Otherwise we let D run through all discs $\supsetneqq U_i$. Then Lemma 1.1 establishes that the numbers of roots of $f(y)$ and $g(y)^d$ in U_i are equal.

We separate the proof of the inequality (4) into several cases.

Case 1. $D\supseteqq D_1$ the maximal disc in $\mathcal{D}(f)$. Let σ^* be a general element for D^* the minimal disc which contains all roots of $f(y)g_d(y)$. Then we have

(6) $\underline{v}(f(\sigma^*))=\underline{v}(g_d(\sigma^*)^d)=nr(D^*)$.

It follows from (2) and the strong triangle inequality

(7) $\underline{v}(h(\sigma^*))\leqq\underline{v}(f(\sigma^*))$.

We trivially have

(8) $l(h, D^*)\leqq\deg h(y)<(d-1/d)n=(d-1/d)l(f, D^*)$.

Thus it follows from Lemma 3.1 that

(9) $\underline{v}(h(\sigma^*))<\underline{v}(f(\sigma^*))$,

(10) $\mathcal{L}(f(\sigma^*))=\mathcal{L}(g(\sigma^*)^d)$.

It follows from Lemma 1.1 and (10) that

$$\begin{aligned}D^*\supsetneqq D_1 &\Longleftrightarrow \mathcal{L}(f(\sigma^*)) \text{ is a power of a linear polynomial}\\ &\Longleftrightarrow \mathcal{L}(g(\sigma^*)) \text{ is a power of the same linear polynomial}\\ &\Longrightarrow \text{the distance of any two roots of } f(y)g_d(y)<r(D^*).\end{aligned}$$

We must have $D^*=D_1$. It follows from Lemma 1.3 and (9) that

(11) $$\begin{aligned}\underline{v}(f(\sigma^*))&=\underline{v}(f(\sigma))(r(D_1)/r(D))^{l(f,D^*)}\\ &>\underline{v}(h(\sigma^*))\\ &\geqq\underline{v}(h(\sigma))(r(D_1)/r(D))^{l(h,D^*)}.\end{aligned}$$

We thus conclude

(12) $\underline{v}(h(\sigma))<\underline{v}(f(\sigma))$.

Case 2. $D\subsetneqq D_v\in\mathcal{D}(f)$. Inductively we assume that for any $D\subsetneqq D_j\in\mathcal{D}(f)$ with a general element σ_j the following

(13) $\underline{v}(h(\sigma_j))<\underline{v}(f(\sigma_j))$.

Let D_u be the smallest disc in $\mathcal{D}(f)$ which contains D properly and satisfies

(14) $l(h, D_u)<(d-1/d)l(f, D_u)$.

Note that the maximal disc D_1 in $\mathcal{D}(f)$ satisfies (14) thus guarantees the existence of D_u. It follows from Lemma 2.1 and (14) that we may construct two maximal chains as follows:

(15) $D_u\supset D_{u+1}\supset\cdots\supset D_v\supset D_{v+1}\supset\cdots\supset D_w=U_i$,

(16) $D_u=D_u^*\supset\cdots\supset D_{w^*}^*=U_j$

with

(17) $l(h, D_i^*)\leqq l^*(\beta, D_i^*)<(d-1/d)l(f, D_i^*)\quad\forall u<i\leqq w^*$,

(18) $D_v\supsetneqq D\supset C_{v+1}$,

(19) $l(h, D_j)\geqq(d-1/d)l(f, D_j)\quad\forall u<j\leqq v$.

Let σ_i^* be a general element for D_i^*. We claim

(20) $\underline{v}(h(\sigma_i^*))\leqq\underline{v}(f(\sigma_i^*))\quad\forall u<i\leqq w^*$.

Note that if $i=u$ then (20) follows from (13). Inductively we assume the inequality (20) is true for some i we shall prove it for $i+1$. Note that it follows from Lemma 3.1, (20) and (17) that

(21) $\underline{v}(h(\sigma_i^*))<\underline{v}(f(\sigma_i^*))=\underline{v}(g_d(\sigma_i^*)^d)$,

(22) $\mathcal{L}(f(\sigma_i^*))=\mathcal{L}(g(\sigma_i^*)^d)$.

It then follows from Lemma 1.1 that

(23) $l(f\,;D_{i+1}^*)=l^*(f,D_{i+1}^*)=l^*(g_d{}^d,D_{i+1}^*)$.

We copy the proof for case (1) as follows. Let σ^* be a general element for D^* the smallest disc which contains all root τ^* of $f(y)g_d(y)$ satisfying

(24) $\underline{v}(\tau^*,\tau)<r(D_i^*)$, $D^*\supset D_{i+1}^*$

where τ is any element in D_{i+1}^*. Then it follows from Lemma 1.3, (21) and (23) that

(25) $$\begin{aligned}\underline{v}(f(\sigma^*))&=\underline{v}(f(\sigma_i^*))(r(D^*)/r(D_i^*))^{l^*(f,D^*_{i+1})}\\&=\underline{v}(g(\sigma_i^*))(r(D^*)/r(D_i^*))^{l^*(g_d{}^d,D^*_{i+1})}\\&=\underline{v}(g_d(\sigma_i^*)^d)\,.\end{aligned}$$

It follows from (2) and the strong triangle inequality

(26) $\underline{v}(h(\sigma^*))\leqq\underline{v}(f(\sigma^*))$.

It trivially follows from (17)

(27) $l(h,D^*)\leqq l^*(h,D_{i+1}^*)<(d-1/d)l(f,D_{i+1}^*)=(d-1/d)l(f,D^*)$.

The rest of the proof of case (1) can be copied almost verbatim to show

(28) $D^*=D_{i+1}^*$,

(29) $\underline{v}(h(\sigma_{i+1}^*))=\underline{v}(h(\sigma^*))\leqq\underline{v}(f(\sigma^*))=\underline{v}(f(\sigma_{i+1}^*))$.

We establish our claim (20).

Note that our conditions (17), (20), (18), (14), (19) fulfill the requirement (3), (4), (6), (7), (8) of Lemma 2.2 for $r=(d-1/d)$, $q=1$. Thus if we have

(30) $l(h,D)\geqq(d-1/d)l(f,D)$

then the requirement (9) of Lemma 2.2 is also fulfilled and our theorem follows from the conclusion (10) of Lemma 2.2. We shall assume the opposite of (30), i. e.

(31) $l(h,D)<(d-1/d)l(f,D)$.

Let the following chain be the maximal one in $\mathcal{D}(fg_d{}^d)$

(32) $D_v=E_1\supset\cdots\supset E_i\supset\cdots\supset E_q=D$

with $\bar{\sigma}_i$ a general element for E_i. It suffices to prove inductively the following:

(33) $\underline{v}(h(\bar{\sigma}_i))<\underline{v}(f(\bar{\sigma}_i))$ $i=1,\cdots,q$.

Note that if we have

(34) $l(h,E_{i+1})\geqq(d-1/d)l(f,E_{i+1})$

then the proof used to handle (30) can be copied verbatim to settle this case. We hence assume for some i

(35) $l(h, E_{i+1})<(d-1/d)l(f, E_{i+1})$,

(36) $\underline{v}(h(\bar{\sigma}_i))<\underline{v}(f(\bar{\sigma}_i))$ $(=\underline{v}(g_d(\bar{\sigma}_i)^d)$.

It then trivially follows (cf. Lemmas 1.1 and 1.2)

(37) $\mathcal{L}(f(\bar{\sigma}_i))=\mathcal{L}(g_d(\bar{\sigma}_i)^d)$,

(38) $l(f, E_{i+1})=l(g_d{}^d, E_{i+1})$,

(39) $\underline{v}(f(\bar{\sigma}_{i+1}))=\underline{v}(g_d(\bar{\sigma}_{i+1})^d)\geqq\underline{v}(h(\bar{\sigma}_{i+1}))$.

It then follows from Lemma 3.1 that

(40) $\underline{v}(h(\bar{\sigma}_{i+1}))<\underline{v}(f(\bar{\sigma}_{i+1}))$. Q. E. D.

THEOREM 2. *We use the assumptions of Theorem* 1. *Let σ_i's be general elements for U_i's and $g(y)$ be a d-th quasi-approximate root of $f(y)$. Let $\alpha_j(y)$ be the polynomials of degrees less than (n/d) and defined by the following equation:*

(1) $f(y)=g(y)^d+\sum_{j=1}^{d}\alpha_j(y)g(y)^{d-j}$.

Let σ be a general element for D which properly contains some member of $\mathcal{U}$. Then we have for all i, j

(2) $\underline{v}(\alpha_j(\sigma))<\underline{v}(f(\sigma))^{j/d}$,

(3) $\underline{v}(g(\sigma)^d)=\underline{v}(f(\sigma))$,

(4) $\underline{v}(\alpha_j(\sigma_i))\leqq\underline{v}(f(\sigma_i))^{j/d}$,

(5) $l(\alpha_j, U_i)\geqq(1/d)l(f, U_i)\Longrightarrow\underline{v}(\alpha_j(\sigma_i))<\underline{v}(f(\sigma_i))^{j/d}$,

(6) $\underline{v}(g(\sigma_i)^d)=\underline{v}(f(\sigma_i))^{j/d}$.

PROOF. By a finite field extension of K we may assume that $\alpha_j(y)$'s factor completely over K. Note that the value $\underline{v}$ may be extended accordingly. It follows from Lemma 2.3 that (4) implies (2) and (5). Moreover, (2) trivially implies (3). It thus suffices to prove (4) and (6).

Let $\bar{D}$ be any disc which contains some U_i. Note that $\bar{D}\cap U_j\neq\emptyset$ implies $\bar{D}\supset U_j$ or $\bar{D}\subset U_j$. It then follows $\bar{D}\supset U_j$. We thus conclude

(7) $l(g, \bar{D})=\sum_{U_j\subset\bar{D}}l(g, U_j)=(1/d)l(f, \bar{D})$.

Let the following chain be maximal in $\mathcal{D}(f)$

(8) $D_1\supset D_2\supset\cdots\supset D_w\supset U_i$.

It follows from (7) that the chain (8) is maximal in $\mathcal{D}(g)$. It follows from Lemma 1.3 that

(9) $\underline{v}(g(\sigma_i))=r(D_1)^{l(g,D_1)}\prod_{j=2}^{w}(r(D_j)/r(D_{j-1}))^{l(g,D_j)}\cdot(r(U_i)/r(D_w))^{l(g,U_i)}$,

(10) $\underline{v}(f(\sigma_i))=r(D_1)^{l(f,D_1)}\prod_{j=2}^{w}(r(D_j)/r(D_{j-1}))^{l(f,D_j)}\cdot(r(U_i)/r(D_w))^{l(f,U_i)}$.

Then (6) follows from (7), (9), (10). We are left to prove (4).

Let e be the largest integer $\leqq d+1$ such that (4) is satisfied for all $1\leqq j<e$ and $1\leqq i\leqq s$. Note the integer e is at least 1. Suppose $e\neq d+1$ and rewrite (1) as follows

(11) $f(y)=g(y)^d+\sum_{j=1}^{e-1}\alpha_j(y)g(y)^{d-j}+\alpha_e(y)g(y)^{d-e}+h(y)$.

Note the following trivial fact

(12) $\deg h(y)<(d-e/d)n$.

Let σ^* be a general element of a disc D^* which contains $U_i\in\mathcal{U}$. Clearly it suffices to prove

(13) $\underline{v}(h(\sigma^*))\leqq\underline{v}(f(\sigma^*))$.

Note that (11) and (13) with $\sigma^*=\sigma_i$ imply

(14) $\underline{v}(\sigma_e(\sigma_i)g(\sigma_i)^{d-e})\leqq\underline{v}(f(\sigma_i))$.

It follows from (6) and (14) that

(15) $\underline{v}(\alpha_e(\sigma_i))\leqq\underline{v}(f(\sigma_i))^{e/d}$

and this contradicts the choice of e.

Let us observe the following fact

(16) $l(h,D^*)<(d-e/d)l(f,D^*)\Longrightarrow\underline{v}(h(\sigma^*))\leqq\underline{v}(f(\sigma^*))$.

Note that it follows from Lemma 2.3 that for $1\leqq j<e$

(17) $\underline{v}(\alpha_j(\sigma^*)g(\sigma^*)^{d-j})\leqq\underline{v}(f(\sigma^*))$

and it follows from Lemma 2.3 and (6) that

(18) $\underline{v}(g(\sigma^*)^d)\leqq\underline{v}(f(\sigma^*))$.

Thus if (16) is false, then we must have

(19) $\underline{v}(h(\sigma^*))=\underline{v}(\alpha_e(\sigma^*)g(\sigma^*)^{d-e})$,

(20) $\mathcal{L}(h(\sigma^*))=\mathcal{L}(\alpha_e(\sigma^*)g(\sigma^*)^{d-e})$.

While the π-degrees of the two sides of (20) are given by

(21) $l(h,D^*)<(d-e/d)l(f,D^*)=(d-e)l(g,D^*)\leqq l(\alpha_e g^{d-e},D^*)$.

This is clearly impossible. We thus have (16).

We divide the proof of (13) into several cases.

Case 1. $D^*\supset D_1$ the maximal disc in $\mathcal{D}(f)$. Clearly we have by (12)

$$l(h, D^*) \leqq \deg h(y) < (d-e/d)n = l(f, D^*) .$$

Then (13) follows from (16).

Case 2. $D^* \subsetneqq D_1$. We construct the following maximal chain of discs

(22) $D_u \supset \cdots \supset D_v \supset D_{v+1} \supset \cdots \supset D_w = U_i$

satisfying the following requirements

(23) $D_v \supsetneqq D^* \supset D_{v+1}$,

(24) $l(h, D_u) < (d-e/d) l(f, D_u)$,

(25) $l(h, D_j) \geqq (d-e/d) l(f, D_j) \quad \forall u < j \leqq v$.

Note that D_1 satisfies (24) thus guarantees the existence of the chain (22). It follows from Lemma 2.1 that we can construct the following maximal chain (26) satisfying (27)

(26) $D_u = D_u^* \supset \cdots \supset D_{w^*}^* = U_j$,

(27) $l(h, D_i^*) \leqq l^*(h, D_i^*) < (d-e/d) l(f, D_i^*) \quad \forall u < i \leqq w^*$.

It then follows from (16) that

(28) $\underline{v}(h(\sigma_w^*)) \leqq \underline{v}(f(\sigma_{w^*}^*))$.

Note that the requirements of Lemma 2.2 are fulfilled with $r=(d-e/d)$, $q=1$. Then (13) follows from Lemma 2.2. Q. E. D.

References

[1, 2] S. S. Abhyankar and T. T. Moh, Newton-Puiseux expansion and Tschirnhausen transformation I and II, J. Reine Angew. Math., **260** (1973), 47-63 and **261** (1973), 29-54.

[3] S. S. Abhyankar and T. T. Moh, Embedding of the line in the plane, J. Reine Angew. Math., **276** (1975), 148-166.

[4] S. S. Abhyankar, On expansion techniques in algebraic geometry, Tata Institute, Bombay, 1977.

[5] H. Hironaka, Idealistic exponents of singularity, Algebraic geometry (J. J. Sylvester Sympos.), 52-125, Johns Hopkins Univ. Press, Baltimore, Maryland, 1977.

[6] M. Miyanishi, Curves on rational and irrational surfaces, Tata Institute, Bombay, 1978.

[7] T. T. Moh, On the concept of approximate roots for algebra, J. Algebra, **65** (1980), 347-360.

[8] T. T. Moh, On the Jacobian conjecture and the configuration of roots (preprint).

T. T. Moh
Department of Mathematics
Purdue University
West Lafayette
Indiana 47907
U. S. A.