

Werk

Titel: Journal für die reine und angewandte Mathematik

Verlag: de Gruyter

Jahr: 1983

Kollektion: Mathematica

Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen

Werk Id: PPN243919689_0340

PURL: http://resolver.sub.uni-goettingen.de/purl?PPN243919689_0340

LOG Id: LOG_0011

LOG Titel: On the Jacobian conjecture and the configurations of roots.

LOG Typ: article

Übergeordnetes Werk

Werk Id: PPN243919689

PURL: <http://resolver.sub.uni-goettingen.de/purl?PPN243919689>

Terms and Conditions

The Goettingen State and University Library provides access to digitized documents strictly for noncommercial educational, research and private purposes and makes no warranty with regard to their use for other purposes. Some of our collections are protected by copyright. Publication and/or broadcast in any form (including electronic) requires prior written permission from the Goettingen State- and University Library.

Each copy of any part of this document must contain there Terms and Conditions. With the usage of the library's online system to access or download a digitized document you accept the Terms and Conditions.

Reproductions of material on the web site may not be made for or donated to other repositories, nor may be further reproduced without written permission from the Goettingen State- and University Library.

For reproduction requests and permissions, please contact us. If citing materials, please give proper attribution of the source.

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

On the Jacobian conjecture and the configurations of roots

By *T. T. Moh**) at West Lafayette

Introduction

The well-known Jacobian conjecture states “let k be an algebraically closed field of characteristic zero and $f(x, y)$, $g(x, y)$ two polynomials over k . If

$$f_x g_y - f_y g_x = 1$$

then $k[x, y] = k[f, g]$.” In other words a polynomial map from $\mathbb{A}^2$ to $\mathbb{A}^2$ is one-one and onto iff the Jacobian of the map is a nonzero constant. Clearly this conjecture is false if the characteristic p of the field k is positive. For instance $f = x + x^p$, $g = y$ furnish a counter example. Furthermore this conjecture is false in the analytic case. The pair $f = e^x$, $g = ye^{-x}$ furnishes a counter-example. For the history of this problem the reader is referred to [B-C-W.1] and [W.2]. Throughout this article the field k is assumed to be algebraically closed and of characteristic zero.

The treatment in the present article is rather classical. Let K be a valued field and $g(y)$ a polynomial which factors completely over K i.e.

$$g(y) = a_0 \prod (y - \tau_i)$$

where $\tau_i \in K$. Then we may associate the following configuration data of real numbers $\{V_{ij}\}$ with $g(y)$:

$$V_{ij} = |\tau_i - \tau_j|$$

where $||$ denotes the value of the field K . In the case of complex numbers the above data and the leading coefficients a_0 uniquely determine the polynomial $g(y)$ up to the translations, the rotations and the reflection. In the case of a non-archimedean valued field K the above data give some weak but valuable information about the polynomial $g(y)$.

Let the valued field K be $k((t))$ the meromorphic field of one variable and the polynomial $g(y)$ be irreducible over K . Then it can be shown that the configuration data of $g(y)$ over its splitting field is equivalent to the classical data of the characteristic pairs of an algebroid curve.

*) This work was partially supported by N.S.F. at Purdue University.

In general, given a polynomial $g(x, y) \in k[x, y]$, we may factor $g(x, y)$ completely over the Puiseux field $k\langle t \rangle$ where $t = x^{-1}$ and $k\langle t \rangle = \bigcup_{n=1}^{\infty} k((t^{\frac{1}{n}}))$. Note the field $k\langle t \rangle$ is algebraically closed (cf. [W.1]). Similarly we define $k[\pi]\langle t \rangle = \bigcup_{n=1}^{\infty} k[\pi]((t^{\frac{1}{n}}))$ with π a symbol. Let us assume that

$$\deg_{x,y} g(x, y) = \deg_y g(x, y).$$

Then it is easy to see that the configuration data determine the number of points of the plane curve $g(x, y) = 0$ at ∞ and the desingularization processes at those points.

Let K be a non-archimedean field and $\{\tau_1, \dots, \tau_n\}$ n elements in K . It is more convenient to consider the following tree data (cf. §1) of $\{\tau_1, \dots, \tau_n\}$ in place of the configuration data. Let D_1 be the minimal disc which contains all τ_i . Let $\{D_{1,i_2}\}$ be the collection of the minimal subdiscs of D_1 with

- (1) $\bigcup D_{1,i_2} \supset \{\tau_i : \tau_i \in D_1\} \quad (= \{\tau_1, \dots, \tau_n\}),$
- (2) $D_{1,i_2} \cap D_{1,j_2} = \emptyset \quad \text{if } i_2 \neq j_2,$
- (3) the distance between D_{1,i_2} and D_{1,j_2} is the radius of D_1 if $i_2 \neq j_2$.

Due to the non-archimedean property of the field K the collection $\{D_{1,i_2}\}$ uniquely exists. We may repeat the above process for D_{1,i_2} . Finally we stop when all discs are with radii zero. In this way we obtain a tree of discs in K . Then the tree data consist of the radii and the numbers of τ_i 's in those discs. It is easy to see this tree of numbers is the characteristic pairs for an algebroid curve $g(x, y) = 0$ if we set $K = k\langle x \rangle$ and $\{\tau_1, \dots, \tau_n\}$ the complete set of roots of $g(x, y) = 0$ over K .

The major portion of this article consists of the study of the tree data of two polynomials $f(x, y)$ and $g(x, y)$ satisfying the Jacobian condition

$$f_x g_y - f_y g_x = 1.$$

It can be illustrated that the Jacobian condition implies some similarities between the two tree data. Namely consider any disc $D \subset K = k\langle t \rangle$ where $t = x^{-1}$ with

$$\min_{\tau \in D} \{\text{ord } f(t^{-1}, \tau)\} \leq 0.$$

Then the number of roots of $f(t^{-1}, y) = 0$ and $g(t^{-1}, y) = 0$ in D are proportional to the degrees of $f(x, y)$ and $g(x, y)$ if both numbers of roots are positive. This qualitative observation is not useful for our purpose. To obtain the useful quantitative properties of those similarities we have to adapt the concept of the approximate roots for the present task (cf. [M.1], [A-M.1]).

In § 2 we show that the Jacobian condition implies the existence of a finite sequence of polynomials monic in y $T_i^\psi(f(x, y), g(x, y)) \in k[f(x, y), g(x, y)]$ with y -degree $(-\mu_i)$ (cf. § 2). The polynomials T_i^ψ are specialized approximate roots of the defining equation of $f(x, y)$ and $g(x, y)$ over $k[x]$. This sequence $\{-\mu_i\}$ determines and is determined by the characteristic sequence of the rational curve parametrized by $(f(0, y), g(0, y))$ at $y = \infty$. Our analysis of the distribution of roots of $f(t^{-1}, y) = 0$ and $g(t^{-1}, y) = 0$ over $k\langle\langle t \rangle\rangle$ solely depends on the existence of the said polynomials T_i^ψ .

In § 3 we relate the sequence of polynomials T_i^ψ with the Jacobian condition for $f(x, y)$ and $g(x, y)$. It is easy to see that the task is to study the partial derivatives of $T_i^\psi(f, g)$ with respect to f . Proposition 3.3 shows that if the similarity of the distribution of roots holds for $g(x, y)$, $T_1^\psi(f(x, y), g(x, y)), \dots, T_r^\psi(f(x, y), g(x, y))$, then it holds for the extra polynomials $\frac{\partial}{\partial f} T_{r+1}^\psi(f(x, y), g(x, y))$. It furnishes an important step in our study. *From the next section on the Jacobian condition is assumed.*

In § 4 we study the basic splitting phenomena and nonsplitting phenomena. Inductively we locate a certain disc D in $k\langle\langle t \rangle\rangle$. The question is to determine if some roots

τ_i, τ_j of $g(t^{-1}, y) \prod_{i=1}^r T_i^\psi(f(t^{-1}, y), g(t^{-1}, y))$ in D have a distance which equals the radius of D . If all distances are less than the radius of D , then the disc D is not one of the discs which determine the tree data of the said polynomial. In this case we have a non-splitting phenomenon. To locate the tree data we have to further shrink the radius of the disc. Proposition 4.4 establishes some sufficient numerical criteria for the non-splitting phenomena. On the other hand if there are some roots τ_i, τ_j having a distance which equals the radius of D , then the disc D is clearly one of the discs which determine the tree data of the said polynomial. In this case we have a splitting phenomenon. Proposition 4.6 establishes some sufficient numerical criteria for the splitting phenomenon and some information about the size of the disc D and the number of the subdiscs of D in the next step of the decomposition in the construction of the tree of discs. Later on we implicitly establish that the two numerical criteria of Propositions 4.4 and 4.6 cover all possibilities.

According to the number of roots in it, the disc D can be classified as either a major disc or a minor disc. A major disc is one which contains more than a predetermined number of roots (cf. Definition 4.1). Otherwise it is a minor one. In § 5 we study the towers of major discs. Propositions 5.1, 5.2 and 5.3 establish the existence and the critical numerical properties of them. Namely a new sequence $\{v_s, v_{s-1}, \dots, v_2\}$ of restricted integers which are the relative numbers of roots in a tower of major discs $D_s \supset D_{s-1} \cdots \supset D_2 \supset D_1$ will determine the radii of those discs completely.

In the remaining part of § 5 we show various applications of Proposition 5.3. Proposition 5.4 establishes that in the case of $\delta_i > -1$ (i.e. there is only one point at ∞ for $f(x, y)$) there is a reduction of $\deg f(x, y)$ and $\deg g(x, y)$. Proposition 5.5 establishes the similar result for the case of two characteristic pairs.

In § 6 we study those minor discs. In Proposition 6.1 we show that in a minor disc the proportional property of the numbers of roots will be carried to its extreme limit. Then we compile the results in Propositions 5.4 and 6.1 to be the main theorem of this article. The other propositions in § 6 are dealing with the phenomena of reducing the degrees of the polynomials involved while increasing the Jacobian. This trade-off is sometimes very useful as illustrated in Appendix II.

In this article we make an effort to compare our results and our methods with the published one ([A.1]). The explicit references are included in various places.

In Appendix I we collect all needed propositions of the polynomial solutions of the differential equations involved. There is a simple computer program essentially based on the discussion at the end to determine the possible counterexamples to the Jacobian conjecture for all polynomials of degrees less than or equal to 100. In fact there are only very few possibilities. However the number of coefficients of those possible counterexamples range from 3370 to 7328. In Appendix II we give some theoretic arguments to reduce the number of coefficients to at most 12. Thus it becomes a computable problem to show that the Jacobian conjecture is valid for all polynomials of degrees less than or equal to 100. The direct computation is presented in Appendix II.

The computer program is a routine one and will be furnished to anyone upon request. It uses 147.148 sec of a CDC 6500 and costs only \$2.41.

After many years of research we feel little information about the distribution of roots can be added to the present article. As we pointed out before, these distribution of roots correspond to the desingularization at ∞ . The geometric approach of desingularization will probably add only little information. Our Appendix II shows that the Jacobian conjecture is quite plausible. Even if it is false a counterexample may require a large number of coefficients and is thus almost uncheckable. If the Jacobian conjecture is valid, the logical sequence to the present article is to show that a part of the data collected thus far is not permissible algebraically. The arguments in Appendix II implicitly imply that the tree data deduced thus far are permissible for two individual polynomials $f(x, y)$ and $g(x, y)$. This is a fact. The only hope, in our opinion, is to put all these tree data in the defining equation of $f(x, y)$ and $g(x, y)$ over $k[x]$ to establish their non-permissibility. However we have had limited success in this complicated task.

The materials of this article are compiled from our research during years 1971—1978. We wish to thank our friends for their encouragement. We wish to express our appreciation to Professor D. Drasin for reading carefully part of our manuscript.

§ 1. The roots configurations and the tree data

Let K be a non-archimedean valued field with value $|\cdot|$. Let $f(y)$ be a polynomial in $K[y]$ which factors completely with roots $\tau_1, \dots, \tau_m$. Then the data

$$\{V_{ij} = |\tau_i - \tau_j|: i, j = 1, \dots, m\}$$

determine the configuration of the roots of $f(y)$. Equivalently let τ_i be any one of the roots, consider the discs centered at τ_i with various radii. As the radius of the disc shrinks the other roots will gradually be left outside. When the number of roots inside the disc is just about to drop we record the number of roots inside the disc and its radius.

For a fixed τ_i , let $S_i = \{|\tau_i - \tau_j| : j = 1, \dots, m\}$. We see that the set S_i is the set of these critical radii. Let the elements in S_i be ordered as $W_{i1} > W_{i2} > \dots > W_{ih} > W_{ih+1} = 0$. Let

$$d_{ij} = \text{card} \{s : |\tau_i - \tau_s| \leq W_{ij}\},$$

$$D_{ij} = \text{the disc centered at } \tau_i \text{ with radius } W_{ij}.$$

Then the numbers $\{d_{ij}, W_{ij}\}$ will be called the tree data of the polynomial $f(y)$ along its root τ_i .

We may define the tree data of the polynomial $f(y)$ as follows. Let D_1 be the smallest disc which contains all roots of $f(y)$ and let

$$d_1^* = \text{the number of roots in } D_1 = \deg f(y),$$

$$W_1^* = \text{the radius of } D_1.$$

Due to the non-archimedean property of the distance function $|\cdot|$, the disc D_1 uniquely exists.

Let us consider the partition of all roots of $f(y)$ induced by the following equivalence relation

$$\tau_i \sim \tau_j \Leftrightarrow |\tau_i - \tau_j| < W_1^*.$$

Using the strong triangle inequality it is a simple exercise to verify the above statement. Let the set R_1 of all roots of $f(y)$ be thus decomposed into disjoint union of subsets R_{1j}

$$R_1 = \bigcup R_{1j}.$$

For each R_{1j} let D_{1j} be the smallest disc which contains all elements in it and let

$$d_{1j}^* = \text{the number of roots in } D_{1j},$$

$$W_{1j}^* = \text{the radius of } D_{1j}.$$

Note that the distance between two roots τ_i and τ_j which belong to two distinct D_{1j} 's is W_1^* .

We may repeat the above discussion for each D_{1j} and cover the roots in it by a disjoint union of subdiscs D_{1jr} with the roots from different subdiscs having distances precisely W_{1j}^* . Thus we formulate the following definition.

Definition 1. 1. Let $f(y) \in K[y]$ with roots $\tau_1, \dots, \tau_m$. The tree data of $f(y)$ consist of $\{d_{1,i_2,\dots,i_s}^*, W_{1,i_2,\dots,i_s}^* \text{ for some } s \leq m\}$ where $d_{1,i_2,\dots,i_s}^*$ are positive integers and $W_{1,i_2,\dots,i_s}^*$ are numbers satisfying the following conditions:

- 1) $d_1^* = m$.
- 2) $\sum_j d_{1,i_2,\dots,i_s,j}^* = d_{1,i_2,\dots,i_s}^*$.
- 3) The smallest disc D_1 which contains all roots is of radius W_1^* .

- 4) For any subindices $1, i_2, \dots, i_s$ there is a disc $D_{1, \dots, i_s}$ with radius $W_{1, \dots, i_s}^*$ which contains $d_{1, \dots, i_s}^*$ roots of $f(y)$.
- 5) Any proper subdisc of $D_{1, \dots, i_s}$ contains less than $d_{1, \dots, i_s}^*$ roots of $f(Y)$.
- 6) $D_{1, \dots, i_s, j} \subsetneq D_{1, \dots, i_s}$.
- 7) Any two roots of $f(y)$ in two $D_{1, \dots, i_s, j}$'s with distinct last indices have distance $W_{1, \dots, i_s}^*$.
- 8) The minimal discs $D_{i_1, \dots, i_s}$ are with radii zero.

In fact, it is elementary to establish the following proposition. Since it will not be used in this article the proof is omitted.

Proposition 1. 1. *Given a polynomial $f(y)$ which factors completely over a non-archimedean valued field K . Then the roots configuration $V_{ij} = \{|\tau_i - \tau_j| : \tau_i, \tau_j \text{ roots of } f(y)\}$ and the tree data $\{d_{1, \dots, i_s}^*, W_{1, \dots, i_s}^*\}$ determine each other.*

Definition 1. 2. Given a polynomial $f(y) \in K[y]$ with roots $\tau_1, \dots, \tau_m$. Let the tree data of $f(y)$ be $\{d_{1, i_2, \dots, i_s}^*, W_{1, i_2, \dots, i_s}^*\}$ and the associated discs be $\{D_{1, i_2, \dots, i_s}\}$. Let τ_i be any root of $f(y)$ and $D_1 \supset D_{1, j_2} \supset \dots \supset D_{1, j_2, \dots, j_{h+1}} \ni \tau_i$ be the tower of associated discs containing τ_i . Then the tree data along τ_i are

$$\{d_1^*, d_{1, j_2}^*, \dots, d_{1, j_2, \dots, j_{h+1}}^*; W_1^*, W_{1, j_2}^*, \dots, W_{1, j_2, \dots, j_{h+1}}^*\}.$$

Lemma 1. 1. *Given a polynomial $f(y) \in K[y]$ with roots $\tau_1, \dots, \tau_n$. Let the tree data of $f(y)$ be $\{d_{1, i_2, \dots, i_s}^*, W_{1, i_2, \dots, i_s}^*\}$ and the associated discs be $\{D_{1, i_2, \dots, i_s}\}$. Let D be any disc in K . Then $D \in \{D_{1, i_2, \dots, i_s}\}$ iff*

- (1) $\tau_j \in D$ for some j ,
- (2) there exist τ_i, τ_j in D with $|\tau_i - \tau_j| = \text{radius of } D$.

Proof. Suppose that $D \in \{D_{1, i_2, \dots, i_s}\}$. If D is one of the minimal discs, then D is of radius zero. The above conditions are satisfied with $\tau_i = \tau_j = \text{any root in } D$. On the other hand if D is not among the minimal discs in $\{D_{1, i_1, \dots, i_s}\}$, say, $D = D_{1, i_1, \dots, i_s}$, then it follows from the conditions (2) (4) (5) of Definition 1. 1 that

$$D = D_{1, i_1, \dots, i_s} \supsetneq D_{1, i_1, \dots, i_s, j}$$

for two different indices j . The conditions (1) and (2) of this lemma follow from the condition (7) of Definition 1. 1.

Let us consider the converse. Let D be a disc with the conditions (1) & (2) of this lemma fulfilled. Since

$$\begin{aligned} \text{radius of } D &= |\tau_i - \tau_j| \leq \max \{|\tau_r - \tau_s|; r, s = 1, \dots, m\} \\ &= \text{radius of } D_1 \end{aligned}$$

and

$$\tau_j \in D \cap D_1.$$

Then we have

$$D_1 \supset D.$$

Let us use the notations of Definition 1.2 with

$$D_1 \supset D_{1,j_2} \supset \cdots \supset D_{1,j_2,\dots,j_{h+1}} \ni \tau_i.$$

Let

$$D_{1,j_2,\dots,j_s} \supseteq D \supseteq D_{1,j_2,\dots,j_s,j_{s+1}}.$$

We claim

$$D = D_{1,j_2,\dots,j_{s+1}}.$$

Otherwise suppose for a moment that we have

$$\text{radius of } D > W_{1,j_2,\dots,j_{s+1}}^* = \text{radius of } D_{1,j_2,\dots,j_{s+1}}.$$

Then it follows from

$$\tau_i \in D_{1,\dots,j_s} \setminus D_{1,\dots,j_{s+1}}$$

and the condition (7) of Definition 1.1 that

$$\text{radius of } D = |\tau_i - \tau_j| = W_{1,i_2,\dots,i_s}^* = \text{radius of } D_{1,\dots,i_s}.$$

Then we must have that $D = D_{1,\dots,i_s}$. A contradiction. Q.E.D.

One famous example of the concept of the tree data can be deduced from the classical characteristic pairs of algebroid curves (cf. [Z.1]).

Remark. Our primary concern in this article is the case that $K = k((t))$ or $K = k\langle t \rangle$ = the Puiseux field of $k((t))$. We shall use $|\sigma_i - \sigma_j| = 2^{-\text{ord}(\sigma_i - \sigma_j)}$ and denote $-\log_2 W_{ij}$ by M_{ij} . We henceforth reformulate the tree data along τ_i as the set $\{M_{ij}, d_{ij}\}$ with $\{\text{ord}(\tau_i - \tau_j) : j = 1, \dots, m\} = \{M_{i1} < M_{i2} < \cdots < M_{ih} < M_{ih+1} = \infty\}$ and

$$d_{ij} = \text{card} \{s : \text{ord}(\tau_i - \tau_s) \geq M_{ij}\}.$$

The tree data of a polynomial will be reformulated accordingly with order $M_{1,i_2,\dots,i_s}$ replacing distance $W_{1,i_2,\dots,i_s}^*$.

The following elementary algebraic tool will be useful.

Definition 1.3. Let K be $k\langle t \rangle$ and π a symbol. Let $\sigma = \sum_{j < \delta} a_j t^j + \pi t^\delta \in k[\pi]\langle t \rangle$ with $a_j \in k$. Then σ is said to be a π -approximate root (abbr. π -root) of a polynomial $f(y)$ if in the expression in $k[\pi]\langle t \rangle$

$$f(\sigma) = f_\sigma(\pi) t^\lambda + \text{higher terms in } t$$

we have $1 \leq \deg_\pi f_\sigma(\pi) < \infty$. The multiplicity of σ as an π -root of $f(y)$ is defined to be $\deg_\pi f_\sigma(\pi)$ as polynomial in π .

A π -root $\sigma = \sum_{j < \delta} a_j t^j + \pi t^\delta$ is a general point in the disc with radius $2^{-\delta}$ centered at $\sum_{j < \delta} a_j t^j$. Any disc will uniquely determine a π -root as a general point in it. The multiplicity of σ is the number of roots of $f(y)$ in the said disc. More precisely we have

Proposition 1. 2. *The element $\sigma = \sum_{j < \delta} a_j t^j + \pi t^\delta$ is a π -root of $f(y) \in K[y]$ iff there is a root τ_j such that $\text{ord}(\sigma - \tau_i) = \delta$. Moreover let $\{\tau_1, \dots, \tau_s\}$ be the collection of roots of $f(y)$ with $\text{ord}(\sigma - \tau_i) = \delta$ and let the coefficient of t^δ in $\tau_1, \dots, \tau_s$ be $\alpha_1, \dots, \alpha_s$ respectively. Then, with $f_\sigma(\pi)$ as defined in Definition 1. 3, we have*

$$f_\sigma(\pi) = c \prod_{j=1}^s (\pi - \alpha_j)$$

where c is a nonzero constant.

Proof. Let us extend $\{\tau_1, \dots, \tau_s\}$ to a complete set of roots $\{\tau_1, \dots, \tau_m\}$ of $f(y)$. It is clear that $f_\sigma(\pi) = \pi$ (the leading coefficient of $(\sigma - \tau_i)$). If $\text{ord}(\sigma - \tau_i) < \delta$, then the leading coefficient of $(\sigma - \tau_i)$ is a nonzero constant. Moreover $\text{ord}(\sigma - \tau_i) \geq \delta$ iff $1 \leq i \leq s$ and the leading coefficient of $(\sigma - \tau_i)$ is $(\pi - \alpha_i)$ in this case. Our statement follows trivially. Q.E.D.

Corollary. *Let $\sigma = \sum a_j t^j + \pi t^\delta$ be a π -root of $f(y)$ with $f(\sigma) = f_\sigma(\pi) t^\lambda + \text{higher terms in } t$. Let τ_i be a root of $f(y)$ with $\text{ord}(\sigma - \tau_i) = \delta$. Then the only permissible degree for $f_\sigma(\pi)$ is d_{ij} for some j . Moreover $f_\sigma(\pi)$ has only one root iff $\delta \neq M_{ij}$ for all j .*

Proof. Proposition 1. 2 and Lemma 1. 1.

As in the theory of the classical characteristic pairs for plane algebroid curves we shall introduce the following auxiliary data:

$$\begin{aligned} q_{i1} &= M_{i1} \\ q_{ij} &= M_{ij} - M_{ij-1} \quad j = 2, \dots, h. \\ \lambda_{ij} &= \sum_{s=1}^j q_{is} d_{is} \end{aligned}$$

The following proposition and its proof indicate the meaning of the derived numbers q_{ij} and λ_{ij} .

Proposition 1. 3. *Let $\sigma = \sum a_j t^j + \pi t^\delta$ be a π -root for $f(y) = \prod (y - \tau_s)$. Suppose $\text{ord}(\sigma - \tau_i) = \delta$ for some i and $M_{ij} < \delta \leq M_{ij+1}$. Then*

$$\begin{aligned} \text{ord } f(\sigma) &= \lambda_{ij} + d_{ij+1}(\delta - M_{ij}) \quad \text{if } j \geq 1, \\ &= m\delta \quad \text{if } j = 0. \end{aligned}$$

Proof. Let us compute $\text{ord}(\sigma - \tau_s)$ for various s . If $\text{ord}(\tau_i - \tau_s) < \delta$ then $\text{ord}(\sigma - \tau_s) = \text{ord}(\tau_i - \tau_s)$. On the other hand if $\text{ord}(\tau_i - \tau_s) \geq \delta$ then we have $\text{ord}(\sigma - \tau_s) = \delta$. In the case $j \geq 1$ we have

$$\begin{aligned} \text{ord } f(\sigma) &= \sum \text{ord}(\sigma - \tau_s) \\ &= (d_{i1} - d_{i2})M_{i1} + \cdots + (d_{ij} - d_{ij+1})M_{ij} + d_{ij+1}\delta \\ &= d_{i1}q_{i1} + d_{i2}q_{i2} + \cdots + d_{ij}q_{ij} + d_{ij+1}(\delta - M_{ij}) \\ &= \lambda_{ij} + d_{ij+1}(\delta - M_{ij}). \end{aligned}$$

In the case $j=0$ we have $\text{ord}(\sigma - \tau_s) = \delta$ for all s . We clearly have, with $j=0$, the following

$$\text{ord } f(\sigma) = m\delta. \quad \text{Q.E.D.}$$

Two natural properties of a system of π -roots are the complete and the coherent conditions (cf. Definition 5 & 8 of [M.1]) which are defined as follows.

Definition 1.4. A system of π -roots $(\sigma_1, \dots, \sigma_s)$ of $f(y)$ is said to be complete if with

$$\sigma_i = \sum a_{ij}t^j + \pi t^{\delta_i} \quad \text{for } i=1, \dots, s$$

we have

- (1) $\text{ord}(\sigma_i - \sigma_j) < \min(\delta_i, \delta_j),$
- (2) $\sum \deg_{\pi} f_{\sigma_i}(\pi) = \deg f(y) = m.$

Moreover the complete system is said to be coherent if

- (3) $\text{ord } f(\sigma_i) = \lambda \quad \text{for } i=1, \dots, s$

and the common value λ of (3) will be called the accuracy of the coherent complete system.

In the above definition the condition (1) means that the $2^{-\delta_i}$ -discs associated with σ_i 's are all disjoint. The condition (2) means that the numbers of roots of $f(y)$ in those disjoint $2^{-\delta_i}$ -discs add up to the total degree of $f(y)$. The condition (3) means, in case $f(y)$ is monic, the product of the distances between the *general* point σ_i and the roots of $f(y)$ is $2^{-\lambda}$ for all σ_i 's.

Recall the following concept of the approximate roots (cf. [M.1]).

Definition. Let $f(y)$ be a monic polynomial of degree $m \cdot d$ over a ring R with d a unit in R . Let $\theta = y^{-1}$ and $R[y] \subset R((\theta))$. Then the d -th root of $f(y)$ exists in $R((\theta))$ and let it be $\theta^{-m} + a_1 \theta^{-m+1} + \cdots + a_{-1} \theta + \cdots$. Then the d -th approximate root $f_d(y)$ of $f(y)$ is defined to be $y^m + a_1 y^{m-1} + \cdots + a_m = \theta^{-m} + a_1 \theta^{-m+1} + \cdots + a_m$.

We shall quote a simpler version of the main result of [M.1] for later references.

Theorem 1. 1. *Let $f(y)$ be monic in $K[y]$ with $K = k\langle\langle t \rangle\rangle$. Let $(\sigma_1, \dots, \sigma_s)$ be a coherent complete system of π -roots of $f(y)$ with multiplicities $(l_1, \dots, l_s)$. Let $d | \text{g.c.d.}(l_1, \dots, l_s)$ and $f_d(y)$ the d -th approximate root of $f(y)$. Then $(\sigma_1, \dots, \sigma_s)$ is a coherent complete system of π -roots of $f_d(y)$ with multiplicities $\left(\frac{l_1}{d}, \dots, \frac{l_s}{d}\right)$.*

The following further refinement of Theorem 1. 1 and the generalization of the approximate roots are interesting. We shall show some of their applications in Appendix II. A proof of Theorem 1. 2 can be found in [M.4].

Definition 1. 5. Let $\{\sigma_1, \dots, \sigma_s\}$ be a coherent complete system of π -roots for a polynomial $f(y)$ in $k\langle\langle t \rangle\rangle[y]$. Let the multiplicities be $\{l_1, \dots, l_s\}$ and the accuracy be λ . Let d be an integer with $d | \text{g.c.d.}(l_1, \dots, l_s)$. A monic polynomial $h(y)$ is said to be a d -th quasi-approximate root of $f(y)$ with accuracy λ if $\{\sigma_1, \dots, \sigma_s\}$ is a coherent complete system of π -roots of $h(y)$ with multiplicities $\left\{\frac{l_1}{d}, \dots, \frac{l_s}{d}\right\}$ and accuracy $\frac{\lambda}{d}$.

Theorem 1. 2. *Let $f(y)$ be monic in $k\langle\langle t \rangle\rangle[y]$ and $\{\sigma_1, \dots, \sigma_s\}$ be a coherent complete system of π -roots with multiplicities $\{l_1, \dots, l_s\}$ and accuracy λ . Let $d | \text{g.c.d.}(l_1, \dots, l_s)$ and $h(y)$ a d -th quasi-approximate root of $f(y)$ with accuracy λ . Let*

$$f(y) = (h(y))^d + \sum_{j=1}^d h_j(y) (h(y))^{d-j}$$

where

$$\deg h_j(y) < \deg h(y) = \frac{\deg f(y)}{d}.$$

Then we have for $i = 1, \dots, s$ and $j = 1, \dots, d$

$$\text{ord } h_j(\sigma_i) \geq \frac{\lambda}{d} j.$$

§ 2. The characteristic data for a pair of polynomials

The purpose of this section is to introduce the characteristic data for a pair of polynomials $f(x, y)$ and $g(x, y)$. It will be illustrated in the sections 4, 5 & 6 that the Jacobian condition

$$J_{x,y}(f(x, y), g(x, y)) = 1$$

implies certain relations between the root configurations of $f(x, y) = 0$ and $g(x, y) = 0$ as equations over $k\langle\langle x^{-1} \rangle\rangle$. Furthermore we shall show that the characteristic data introduced here will control the tree data of the two polynomial equations within certain natural limitations.

Let $f(x, y)$ and $g(x, y)$ be two polynomials monic in y with

$$\deg_y f(x, y) = m, \quad \deg_y g(x, y) = n.$$

Then we may view $(f(x, y), g(x, y))$ as a parametrization over $k(x)$ of a plane rational curve with y as the parameter. Clearly the said curve has only one place at $y = \infty$. Analytically, if we let

$$g(x, y) = \eta^{-n},$$

$$\eta = g(x, y)^{-\frac{1}{n}} = y^{-1} + \alpha_2(x)y^{-2} + \dots + \epsilon \in k[x]((y^{-1})) = k[x]((\eta))$$

then

$$f(x, y) = \eta^{-m} + \sum_{j > -m} f_j(x)\eta^j \in k[x]((\eta)).$$

Let Ω_i be the automorphism of $k[x]((\eta))$ defined by

$$\Omega_i(x) = x, \quad \Omega_i(\eta) = \omega^i \eta$$

where ω is a fixed n -th primitive root of unity. Suppose

$$y \in k(x, f(x, y), g(x, y)).$$

Since y is integral over $k[x, g(x, y)]$ which is integrally closed, we deduce that $f(x, y)$ satisfies a monic polynomial of degree n over $k[x, g(x, y)]$. The equation is irreducible over $k[x]((g^{-1})) = k[x]((\eta^n))$. Then the defining equation of $f(x, y)$ over

$$k[x]((\eta^n)) = k[x]((g^{-1}))$$

is the following

$$(1) \quad \prod_{i=1}^n (f - \Omega_i f(x, y)) = 0$$

where f is a symbol as usual. In fact the above equation is the analytic factorization of the defining equation of the said rational curve at $g = \infty$ or $g^{-1} = 0$.

Due to the existence of the automorphisms Ω_i , the tree data associated with the configuration of the roots of the equation (1) over $k[x]((\eta))$ are completely symmetric. Hence the tree data along any root will determine the tree data completely. Let the tree data along one root (hence all roots) be (reindexed as) $\{M_j, d_j\}$. We shall call $\{M_j, d_j\}$ the *characteristic data* of the pair of polynomials $(f(x, y), g(x, y))$ with respect to x . Recall (cf. [A-M.1] p. 68)

$$d_1 = n, \quad d_{j+1} = \text{g.c.d.}(n, M_1, \dots, M_j),$$

$$M_j = \min \{i, f_i(x) \neq 0, d_j \nmid i\}, \quad M_{h+1} = \infty$$

and the following auxiliary notations

$$n_0 = 1, \quad n_j = \frac{d_j}{d_{j+1}},$$

$$q_1 = M_1, \quad q_j = M_j - M_{j-1},$$

$$\lambda_j = \sum_{1 \leq i \leq j} q_i d_i, \quad \mu_j = \frac{\lambda_j}{d_j}, \quad \theta_j = \mu_j - M_j.$$

To make some sense out of the long list of notations introduced above, we shall insert the following lemma first observed by S.S. Abhyankar ([A.1], p. 149).

Lemma 2. 1. Let $f(x, y), g(x, y) \in k[x, y]$ be two polynomials monic in y with y -degree m, n respectively. Let

$$g(x, y) = \eta^{-n}, \quad f(x, y) = \eta^{-m} + \sum_{i > -m} f_i(x) \eta^i.$$

Then $J_{x,y}(f(x, y), g(x, y)) = c$ a nonzero constant iff

$$\text{ord}_\eta(\sum f_i(x) \eta^i - \sum f(0) \eta^i) = n - 1$$

and

$$\deg_x f_{n-1}(x) = 1.$$

Proof. It follows from chain rule that

$$\begin{aligned} J_{x,y}(f(x, y), g(x, y)) &= J_{x,\eta}(f(x, y), g(x, y)) J_{x,y}(x, \eta) \\ &= -n(\sum f'_i(x) \eta^i) \eta^{-n-1} \cdot (-\eta^2 + \dots) \\ &= n(\sum f'_i(x) \eta^i) (\eta^{-n+1} + \dots). \end{aligned}$$

The right hand side is in $k[x, y]$ and a unit in $k[x][[\eta]]$, hence a nonzero constant, iff

$$f'_i(x) = 0 \quad \forall i < n - 1$$

and

$$f'_{n-1}(x) = \text{nonzero constant}.$$

Q.E.D.

Note that since $f_{n-1}(x) \neq 0$ and $(n, n-1) = 1$, then $M_i \leq n-1$ for all $i \leq h$. Thus we know all M_i 's terminate at $n-1$ if the Jacobian condition is satisfied. Note that this implies that $f(x, y)$ has n conjugates over $k[x]$ ((g^{-1})) and $y \in k(x, g(x, y), f(x, y))$. We shall call the reader's attention to our ignorance about the numbers M_i . For instance, after long computation we (as verified by A. Sathaye) established that for the following polynomials over k

$$\begin{aligned} X &= T^6 + \dots = \theta^{-6}, \\ Y &= T^8 + \dots = \theta^{-8} + \dots + \theta^{M_2} + \dots \end{aligned}$$

the value M_2 could never be 13. It shows how little we know about M_i in the polynomial case.

The characteristic data $\{d_i; M_i\}$ are analytic in nature. We shall replace them, and thus algebraicize them, by the approximate roots. This method has been well demonstrated in [A-M.1]. In the present article we are interested in some specializations of the approximate roots. Let

$$F(f, g) = \prod_{i=1}^n (f - \Omega_i(\eta^{-m} + \sum_{j > -m} f_j(x) \eta^j)) = 0$$

be the defining equation of $f(x, y)$ and $g(x, y)$ over $k[x]$. Let ψ be the specialization $k[x] \rightarrow k$ defined by

$$\psi(x) = 0,$$

we shall extend ψ naturally to $k[x][f, g]$ with f, g symbols. Then clearly we have

$$F^\psi(f, g) = \prod_{i=1}^n [f - \Omega_i(\eta^{-m} + \sum f_j(0)\eta^j)].$$

We shall establish the following proposition.

Proposition 2. 1. *Let $T_r(f, g)$ be the d_r -th approximate root of $F(f, g)$ with f viewed as the variable. Then $T_r^\psi(f, g)$ is the d_r -th approximate root of $F^\psi(f, g)$ with f viewed as the variable.*

Proof. Recall the definition of the approximate roots in §1. Let the d_r -th root of $F(f, g)$ in $k[x, g]((f^{-1}))$ be

$$F(f, g)^{\frac{1}{d_r}} = f^{\frac{n}{d_r}} + \sum_{i>0} \alpha_i(x, g) f^{\frac{n}{d_r}-i}.$$

Then we have

$$F^\psi(f, g)^{\frac{1}{d_r}} = f^{\frac{n}{d_r}} + \sum_{i>0} \alpha_i(0, g) f^{\frac{n}{d_r}-i}.$$

Naturally we conclude the following

$$\begin{aligned} T_r^\psi(f, g) &= f^{\frac{n}{d_r}} + \sum_{\frac{n}{d_r} \geq i > 0} \alpha_i(0, g) f^{\frac{n}{d_r}-i} \\ &= \text{the } d_r\text{-th approximate root of } F^\psi(f, g). \quad \text{Q.E.D.} \end{aligned}$$

The advantage of the specialization $\psi: x \rightarrow 0$ is that the polynomials $T_r^\psi(f, g)$ are in $k[f, g]$ while the polynomials $T_r(f, g)$ are with coefficients in $k[x]$. Thus the differentiation of the polynomial $T_r(f(x, y), g(x, y))$ with respect to x which will be used later on will be rather complicated. The most useful information provided by the polynomials $T_r(f, g)$ are the y -degrees of the polynomials $T_r(f(x, y), g(x, y))$. This valuable information is largely retained by the simpler polynomials $T_r^\psi(f, g)$. To be more precise let e be an integer defined by (cf. Lemma 2. 1).

$$e = \text{ord}_\eta(\sum f_j(x)\eta^j - \sum f_j(0)\eta^j).$$

We are primarily interested in $T_r^\psi(f, g)$ for $M_r \leq e$. Let us prove the following proposition (cf. [A.1], p. 155).

Proposition 2. 2. *Let e be an integer defined by*

$$e = \text{ord}_\eta(\sum f_j(x)\eta^j - \sum f_j(0)\eta^j).$$

Let $M_r \leq e$. Then we have

$$(1) \quad \deg_y T_r^\psi(f(x, y), g(x, y)) = \deg_y T_r(f(x, y), g(x, y)) = -\mu_r.$$

If $M_r < e$, then we have

$$(2) \quad T_r(f(x, y), g(x, y)) \quad \text{and} \quad T_r^\psi(f(x, y), g(x, y)) \quad \text{are monic in } y.$$

If $M_r = e$, then we have for some nonzero constant c the following

$$(3) \quad T_r^\psi(f(x, y), g(x, y)) = cf_e(x)y^{-\mu_r} + \text{lower terms in } y.$$

Proof. Let

$$(4) \quad \sigma = \eta^{-m} + \sum_{-m < j < M_r} f_j(x) \eta^j + \pi \eta^{M_r}.$$

Due to the restriction of $M_r \leq e$, i.e. $f_j(x) \in k$, we have that σ is also a π -root of the equation

$$F^\psi(f, g) = 0.$$

Let us set

$$\bar{\sigma} = \eta^{-m} + \sum_{-m < j < M_r} f_j(x) \eta^j.$$

Then the greatest common divisor of all exponents of the terms with nonzero coefficients in $\bar{\sigma}$ and the integer n is d_r according to the definition of d_r . The automorphisms Ω_j will produce $\left(\frac{n}{d_r}\right)$ distinct conjugates of $\bar{\sigma}$. Let the corresponding π -roots for $F(f, g)$ in the M_r -discs be $\{\sigma_1, \dots, \sigma_{\frac{n}{d_r}}\}$. Note that σ_i is simply the summation of a conjugate of $\bar{\sigma}$ and $\pi \eta^{M_r}$. Then this system of π -roots is coherent and complete for $F(f, g)$ with multiplicities $\{d_r, \dots, d_r\}$ and accuracy λ . It follows from Proposition 1.3 that

$$\lambda = \lambda_r = \mu_r d_r.$$

As we pointed out before, the inequality

$$M_r \leq e$$

implies $f_j(x) \in k$ in the expression (4). Then the above reasoning works for $F^\psi(f, g)$. Thus the same π -root system $\{\sigma_1, \dots, \sigma_{\frac{n}{d_r}}\}$ is coherent and complete for $F^\psi(f, g)$ with identical multiplicities $\{d_r, \dots, d_r\}$ and $\lambda = \lambda_r = \mu_r d_r$.

Let us consider the polynomials $T_r(f, g)$ and $T_r^\psi(f, g)$ with f viewed as the variable. It follows from Theorem 1.1 that the π -roots system $\{\sigma_1, \dots, \sigma_{\frac{n}{d_r}}\}$ is coherent and complete for $T_r^\psi(f, g)$ (resp. $T_r(f, g)$) with multiplicities $\{1, \dots, 1\}$. Its accuracy can be easily determined to be $\frac{\lambda_r}{d_r} = \mu_r$ as follows. Let $\{\tau_1, \dots, \tau_n\}$ be the complete set of roots of $F(f, g)$ (resp. $F^\psi(f, g)$) and $\{\tau_1^*, \dots, \tau_{\frac{n}{d_r}}^*\}$ be the complete set of roots of $T_r(f, g)$ (resp. $T_r^\psi(f, g)$). Since the M_r -discs associated with σ_i 's are all disjoint (cf. Definition 1.4), then it follows from the strong triangle inequality that for τ_s, τ_s^* belonging to the M_r -disc associated with σ_s with $s \neq i$ we always have

$$\text{ord}(\sigma_i - \tau_s) = \text{ord}(\sigma_i - \sigma_s), \quad \text{ord}(\sigma_i - \tau_s^*) = \text{ord}(\sigma_i - \sigma_s)$$

while in each M_r -disc there are d_r τ_j 's and only one τ_j^* . A simple computation verifies that the accuracy is indeed $\frac{\lambda_r}{d_r} = \mu_r$.

Let us reorder if necessary the complete set of roots $\{\tau_1^*, \dots, \tau_{\frac{n}{d_r}}^*\}$ so that

$$\text{ord}(\sigma_i - \tau_i^*) = M_r.$$

We claim the following

$$\sigma_i - \tau_i^* = \pi \eta^{M_r} + \text{higher terms in } \eta.$$

Otherwise the leading coefficient is of the form $(\pi - \alpha)$ with some nonzero α which is the coefficient of η^{M_r} in τ_i^* . Note that the automorphisms Ω_j will leave g fixed and change η to $\omega^j \eta$. Thus the automorphisms Ω_j will produce too many roots for $T_r(f, g)$ (resp. $T_r^\psi(f, g)$). A contradiction. Let $\sigma_1 = \sigma$ be the one defined by equation (4). Then it follows from the fact $f_{M_r}(x) \neq 0$ that

$$\text{ord}(f(x, y) - \tau_1^*) = M_r = \text{ord}(\sigma_1 - \tau_1^*).$$

Moreover it follows from the disjointness of the M_r -discs associated with σ_i 's that for $i \neq 1$

$$\text{ord}(f(x, y) - \tau_i^*) = \text{ord}(\sigma_1 - \tau_i^*).$$

Thus we conclude

$$\begin{aligned} \deg_y T_r(f(x, y), g(x, y)) &= -\text{ord}_\eta T_r(f(x, y), g(x, y)) \\ &= -\sum \text{ord}(f(x, y) - \tau_i^*) \\ &= -\sum \text{ord}(\sigma_1 - \tau_i^*) = -\mu_r \end{aligned}$$

and

$$\deg_y T_r^\psi(f(x, y), g(x, y)) = -\mu_r.$$

We have established the conclusion (1).

To prove the conclusions (2) and (3) we observe that the leading coefficient $(f(x, y) - \tau_1^*)$ is $f_{M_r}(x)$ while the leading coefficients of $(f(x, y) - \tau_i^*)$ for $i \neq 1$ are all nonzero constants. Thus the product of all leading coefficients is either a nonzero constant if $f_{M_r}(x)$ is a constant or of the form $c f_{M_r}(x)$. We shall remind the reader that the term monic means the coefficient of the highest term is a unit in our terminology.

Q.E.D.

In the tradition of Pythagoras we compute numbers. In fact we show that there is a sequence of polynomials $T_r^\psi(f(x, y), g(x, y))$ with y -degree $(-\mu_r)$. One can convince oneself that the sequences $\{n, \mu_1, \dots, \mu_r\}$ and $\{n, M_1, \dots, M_r\}$ determine each other. Especially one has

$$\text{g.c.d. } \{n, \mu_1, \dots, \mu_r\} = \text{g.c.d. } \{n, M_1, \dots, M_r\}.$$

Moreover the existence of such a sequence of polynomials $T_r^\psi(f(x, y), g(x, y))$ implies that the characteristic data $\{d_i, M_i\}$ can be recovered by (cf. [A-M.1], [S.1], [AN.1]) the following formula:

$$\begin{aligned} d_1 &= n, \quad M_1 = \mu_1 \\ d_{j+1} &= \text{g.c.d.}(n, \mu_1, \dots, \mu_j), \quad M_j = \frac{1}{d_j - d_{j+1}} \left[\mu_j d_j - \sum_{i=1}^{j-1} (d_i - d_{i+1}) M_i \right]. \end{aligned}$$

The mere y -degrees of the polynomials $T_r^\psi(f(x, y), g(x, y))$ will form a foundation of our investigation of the Jacobian problem.

§ 3. The partial derivatives of T_i^ψ

Let $H(f, g)$ be a polynomial in $k[f, g]$. We may compute the Jacobian of $H(f(x, y), g(x, y))$ and $g(x, y)$ by the chain rule as follows

$$\begin{aligned} J_{x,y}(H(f(x, y), g(x, y)), g(x, y)) &= J_{f,g}(H(f, g), g) J_{x,y}(f(x, y), g(x, y)) \\ &= \left(\frac{\partial}{\partial f} H(f, g) \right) J_{x,y}(f(x, y), g(x, y)) \\ &= H(f(x, y), g(x, y))_f \cdot J_{x,y}(f(x, y), g(x, y)). \end{aligned}$$

We are primarily interested in the case that the Jacobian condition is satisfied

$$J_{x,y}(f(x, y), g(x, y)) = 1$$

and for $H(f, g) = T_r^\psi(f, g)$ as introduced in the last section. In those cases the Jacobians become the partial derivatives. We shall deduce the necessary information about the forms and the y -degrees of the partial derivatives in this section.

Let us list the three basic arithmetic properties of the polynomials $T_r(f, g)$ and the sequence $\{n, \mu_1, \dots, \mu_r\}$ as follows (cf. [A-M.1], p. 51 and Proposition 2. 2):

- (1) $\text{g.c.d.}(n, \mu_1, \dots, \mu_j) = d_{j+1} < d_j$,
- (2) $T_r(f, g) = f^{\frac{n}{d_r}} + \text{lower terms in } f$,
- (3) $T_r(f(x, y), g(x, y)) = c_r y^{-\mu_r} + \text{lower terms in } y \text{ for all } r \text{ with } M_r < e$.

For the definition of the number e the reader is referred to Proposition 2. 2. Recall from § 2

$$\begin{aligned} n_0 &= \frac{n}{d_1} = 1, \quad n_j = \frac{d_j}{d_{j+1}}, \\ \theta_j &= \mu_j - M_j. \end{aligned}$$

Then clearly we have

$$(4) \quad \deg_f T_r(f, g) = \frac{n}{d_r} = \prod_{j=1}^{r-1} n_j.$$

We shall use the three basic arithmetic properties listed above and the technique of the decimal expansions to establish a canonical expansion for $T_{r+1}(f, g)$ in terms of $T_j(f, g)$ for $j \leq r$.

We may use the integers $1, n_1, n_1 \cdot n_2, \dots, \prod_{j=1}^{r-1} n_j$ to express any nonnegative integer l uniquely as

$$l = \sum_{i=1}^r \alpha_i \prod_{j=1}^{i-1} n_j$$

with the following restrictions

$$(5) \quad 0 \leq \alpha_i < n_i \quad \text{for } i < r,$$

$$(6) \quad 0 \leq \alpha_r < \infty.$$

The above expansion of the nonnegative integer l is not different in principle from the usual decimal expansions. We shall use the following common vector notation

$$\alpha = (\alpha_1, \dots, \alpha_r), \quad T_\alpha = \prod_{i=1}^r T_i^{\alpha_i}.$$

Then it follows from equation (4) that

$$(7) \quad \deg T_\alpha = \sum \alpha_i \prod_{j=1}^{i-1} n_j.$$

We shall establish the following lemma (cf. [A-M.1], p. 40).

Lemma 3.1. *Let $H(f, g)$ be a polynomial in $k[x][f, g]$. Then $H(f, g)$ can be uniquely expressed over $k[x]$ as follows*

$$(8) \quad H(f, g) = \sum_{j, \alpha} c_{j, \alpha} g^j T_\alpha$$

with $\alpha = (\alpha_1, \dots, \alpha_r)$ satisfying the restriction (5) & (6) in the preceding paragraph. Moreover we have

$$\begin{aligned} \deg_f H(f, g) &= \max \{ \deg_f T_\alpha : c_{j, \alpha} \neq 0 \text{ for some } j \} \\ &= \max \left\{ \sum \alpha_i \prod_{j=1}^{i-1} n_j : c_{j, \alpha} \neq 0 \right\}. \end{aligned}$$

Proof. Let us use the expansion technique for the following integer l

$$\deg_f H(f, g) = l = \sum \alpha_i \prod_{j=1}^{i-1} n_j.$$

We have the following expression for $H(f, g)$

$$\begin{aligned} H(f, g) &= \sum_j (c_{j, l} g^j) f^l + \text{lower terms in } f \\ &= \sum_j c_{j, \alpha} g^j T_\alpha + \text{lower terms in } f \end{aligned}$$

where $c_{j, l} = c_{j, \alpha}$ for all j . The existence of the expression (8) can be easily established by applying the preceding process to the polynomial

$$H(f, g) - \sum c_{j, \alpha} g^j T_\alpha$$

which is of f -degree less than l .

Note that in the expression (8) the terms with distinct α indices will have distinct f -degrees. This is simply due to the fact that the f -degrees of the terms of the right hand side of the equation (7) are the expansions of distinct nonnegative integers if the α 's are distinct.

For the uniqueness of the expression (8) it suffices, as usual, to give a prove for the zero polynomial. Let

$$0 = \sum_{j, \alpha} c_{j, \alpha}^* g^j T_{\alpha}.$$

Using the preceding argument about f -degrees we separate the above equation into many equations for fixed α 's as follows

$$0 = \sum_j c_{j, \alpha}^* g^j T_{\alpha}$$

or

$$0 = \sum_j c_{j, \alpha}^* g^j.$$

Certainly we have for all j and all α

$$0 = c_{j, \alpha}^*. \quad \text{Q.E.D.}$$

The first half of the following proposition follows from § 7.2 of [A-M.1]. For completeness we give a short argument.

Proposition 3. 1. *The polynomial $T_{r+1}(f, g)$ has a unique expression over $k[x]$ of the following form*

$$(9) \quad T_{r+1}(f, g) = T_r(f, g)^{n_r} + \sum_{j, \alpha} c_{j, \alpha} g^j T_{\alpha}$$

with $\alpha = (\alpha_1, \dots, \alpha_r)$ satisfying the restrictions (5) and (6). Moreover let r be an integer with $M_r < e$ (cf. Proposition 2. 2), we have the following for $c_{j, \alpha} \neq 0$

$$(1) \quad \alpha_r < n_r,$$

$$(2) \quad \deg_y g^j(x, y) T_{\alpha}(f(x, y), g(x, y)) \leq (-\mu_r) n_r.$$

Furthermore there is a unique j and a unique α for the equality to hold in (2) and this particular α is with $\alpha_r = 0$.

Proof. Applying Lemma 3. 1 to the polynomial

$$T_{r+1}(f, g) - T_r(f, g)^{n_r}$$

we establish the first part of the proposition.

Since both polynomials $T_{r+1}(f, g)$ and $T_r(f, g)^{n_r}$ have leading coefficient 1 and are of degree $\left(\frac{n}{d_{r+1}}\right)$ in f then we have

$$\deg_f (T_{r+1}(f, g) - T_r(f, g)^{n_r}) = \frac{n}{d_{r+1}} < \prod_{j=1}^r n_j.$$

Thus it follows from Lemma 3. 1 that in the expression (9) we always have

$$\deg_f T_{\alpha} = \sum \alpha_i \prod_{j=1}^{i-1} n_j < \prod_{j=1}^r n_j.$$

Hence we conclude easily for $c_{j,\alpha} \neq 0$ the conclusion (1)

$$\alpha_r < n_r.$$

Let us replace the symbols f and g in all terms of the expression (9) by $f(x, y)$ and $g(x, y)$ and compute the y -degrees of the resulting polynomials. We have (cf. Proposition 2.2)

$$(10) \quad \deg_y T_{r+1}(f(x, y), g(x, y)) = -\mu_{r+1} = (-\mu_r)n_r - q_{r+1} \\ < (-\mu_r)n_r = \deg_y T_r(f(x, y), g(x, y))^{n_r}$$

and for the other terms we have

$$(11) \quad \deg_y c_{j,\alpha} g^j(x, y) T_\alpha(f(x, y), g(x, y)) = jn + \sum (-\mu_i)\alpha_i$$

where $\alpha = (\alpha_1, \dots, \alpha_r)$. We shall prove that any two terms with distinct j and α indices will have distinct y degrees. The proof is arithmetic in nature. Let j^* and α^* be the other indices with

$$\deg_y g^j(x, y) T_\alpha(f(x, y), g(x, y)) = \deg_y g^{j^*}(x, y) T_{\alpha^*}(f(x, y), g(x, y)).$$

Then we have

$$(12) \quad j \cdot n + \sum (-\mu_i)\alpha_i = j^*n + \sum (-\mu_i)\alpha_i^*.$$

Recall

$$d_{j+1} = \text{g.c.d.}(n, \mu_1, \dots, \mu_j).$$

It is easy to see that the residue

$$\frac{-\mu_r}{d_{r+1}} \bmod \left(\frac{d_r}{d_{r+1}} \right) = \frac{-\mu_r}{d_{r+1}} \bmod (n_r)$$

is a unit in $\mathbb{Z}/(n_r)$. Let us divide the numbers $n, -\mu_1, \dots, -\mu_r$ in the equation (12) by d_{r+1} and then mod the equation by $\frac{d_r}{d_{r+1}} = n_r$. The residue equation is of the following form

$$\bar{\alpha}_r \left(\frac{-\mu_r}{d_{r+1}} \right) \equiv \bar{\alpha}_r^* \left(\frac{-\mu_r}{d_{r+1}} \right)$$

or

$$\bar{\alpha}_r \equiv \bar{\alpha}_r^* \bmod n_r.$$

Recall

$$0 \leq \alpha_r, \quad \alpha_r^* < n_r.$$

We thus conclude

$$\alpha_r = \alpha_r^*.$$

Now we may throw away α_r and α_r^* from the equation (12) and proceed likewise. It is routine to show

$$\alpha_{r-1} = \alpha_{r-1}^*, \\ \vdots \\ \alpha_1 = \alpha_1^*, \\ j = j^*.$$

Thus we establish that distinct j and α indices in the equation (1) will produce terms with distinct y -degrees. In light of the equation (10) we must have the conclusion (2) and the uniqueness of a particular term of y -degree $(-\mu_n)n_r$. We are left to show that this particular term is with $\alpha_r=0$. Let it be with indices j and $\alpha=(\alpha_1, \dots, \alpha_r)$. Then we need to solve the following arithmetic equation

$$(13) \quad jn + \sum (-\mu_i)\alpha_i = (-\mu_n)n_r.$$

Let us divide the numbers $n, -\mu_1, \dots, -\mu_r$ in the above equation by d_{r+1} . The residue equation mod $\frac{d_r}{d_{r+1}}$ is of the following form

$$\bar{\alpha}_r \left(\frac{-\mu_n}{d_{r+1}} \right) \equiv 0.$$

Since the residue of $\frac{-\mu_n}{d_{r+1}}$ is a unit and

$$0 \leq \alpha_r < n_r$$

we must have

$$\alpha_r = 0.$$

Q.E.D.

Remark. Proposition 3.1 remains valid after applying the specialization ψ with

$$\psi: x \rightarrow 0.$$

In fact the specialized proposition is the one interesting to us in the present article.

We shall give an arithmetic proof of the following proposition about the forms and the y -degrees of the partial derivatives of $T_r(f, g)$ (cf. [A.1], p. 155).

Proposition 3.2. *We have the following*

- (1) $T_1(f(x, y), g(x, y))_f = 1,$
- (2) $T_{r+1}(f(x, y), g(x, y))_f = n_r T_r(f(x, y), g(x, y))^{n_r-1} \cdot T_r(f(x, y), g(x, y))_f$
 $+ \sum C_{i,\beta} g^i(x, y) T_\beta(f(x, y), g(x, y))$

where the terms under the summation sign satisfy

$$\begin{aligned} \deg_y T_{r+1}(f, g)_f &= \deg_y T_r(f, g)^{n_r-1} T_r(f, g)_f \\ &> \deg_y g^i(x, y) T_\beta(f(x, y), g(x, y)) \end{aligned}$$

for $r \geq 1$.

$$(3) \quad \deg_y T_{r+1}(f(x, y), g(x, y))_f = (-\mu_{r+1}) + M_{r+1} = \theta_{r+1}$$

for $r \geq 1$.

Proof. The statement (1) is trivial. Let us prove the statements (2) and (3) for the case $r = 1$. The equation (9) of Proposition 3.1 is of the following form

$$T_2(f, g) = T_1(f, g)^{n_1} + \sum C_{j, \alpha} g^j T_\alpha$$

where $\alpha = (\alpha_1)$ and the unique j, α which gives the maximal y -degree is with $\alpha = (0)$. So we conclude

$$T_2(f, g)_f = n_1 T_1(f, g)^{n_1-1} + \sum \alpha_1 C_{j, \alpha} g^j T_{(\alpha_1-1)}.$$

Multiply the above equation by $T_1(f, g)$ and the right hand side is restored to the original form except for the change of coefficients and the dropping out of the unique term in $\sum C_{j, \alpha} g^j T_\alpha$ which gives the maximal y -degree and all other terms with $\alpha = 0$. Thus we conclude

$$\begin{aligned} \deg_y T_2(f, g)_f &= \deg_y T_1(f, g)^{n_1-1} = (-\mu_1)(n_1-1) \\ &= -q_1 n_1 + M_1 = -q_1 n_1 - M_2 + M_1 + M_2 \\ &= -q_1 \left(\frac{n}{d_2} \right) + q_2 + M_2 = -\mu_2 + M_2 = \theta_2 \end{aligned}$$

which proves statements (2) and (3) for $r = 1$.

In general let us assume that the statements (2) and (3) have been proved for all smaller indices and we shall proceed for the index $r+1$. From the equation (9) of Proposition 3.1 we have the following

$$T_{r+1}(f, g)_f = n_r T_r(f, g)^{n_r-1} T_r(f, g)_f + \sum C_{j, \alpha} (g^j T_\alpha)_f.$$

Note that the unique term giving the maximal y -degree in $\sum C_{j, \alpha} g^j T_\alpha$ of the equation (9) is with $\alpha = (\alpha_1, \dots, \alpha_{r-1}, 0)$. We have

$$\deg_y T_r(f, g)^{n_r-1} T_r(f, g)_f = (-\mu_r) n_r + M_r$$

by mathematical induction. And the term $(g^j T_\alpha)_f$ can be spread out as

$$(g^j T_\alpha)_f = \sum \alpha_i g^j T_{\alpha^{(i)}}(T_i)_f$$

where $\alpha = (\alpha_1, \dots, \alpha_r)$ and $\alpha^{(i)} = (\alpha_1, \dots, \alpha_i - 1, \dots, \alpha_r)$. We have by mathematical induction

$$(14) \quad \deg_y g^j T_{\alpha^{(i)}}(T_i)_f = \deg_y g^j T_\alpha + M_i$$

and there are only two possibilities; either

$$\deg_y g^j T_\alpha < (-\mu_r) n_r = \deg_y T_r(f, g)^{n_r}$$

or

$$\deg_y g^j T_\alpha = (-\mu_r) n_r = \deg_y T_r(f, g)^{n_r}.$$

In the second case the corresponding α_r is zero, we may assume the following for the equation (14)

$$M_i < M_r.$$

In either case we always have

$$\deg_y g^j T_{\alpha^{(i)}}(T_i)_f < \deg_y T_r(f, g)^{n_r-1} (T_r)_f.$$

In other words we must have

$$\begin{aligned} \deg_y T_{r+1}(f, g)_f &= \deg_y T_r(f, g)^{n_r-1} (T_r)_f \\ &= (-\mu_r)n_r + M_r = -\mu_{r+1} + M_{r+1} = \theta_{r+1} \end{aligned}$$

which establishes statement (3). To finish statement (2) we note that in the following expression (cf. Lemma 3.1)

$$\sum C_{j,\alpha}(g^j T_{\alpha})_f = \sum C_{i,\beta} g^i T_{\beta}$$

the y -degree of the left hand side is less than $\deg_y T_{r+1}(f, g)_f$ as proved above and the y -degrees of all terms on the right side are distinct (cf. the proof of Lemma 3.1). Thus we must have

$$\deg_y T_{r+1}(f, g)_f > \deg_y g^i T_{\beta}$$

for each individual term $g^i T_{\beta}$.

Q.E.D.

Remark. For r with $M_{r+1} \leq e$ we may specialize $\psi: x \rightarrow 0$ and thus replace $T_i(f, g)$ by $T_i^{\psi}(f, g)$ for all $i \leq r+1$ in proposition 3.2. All conclusions are valid. In fact, we are only interested in $T_i^{\psi}(f, g)$ for the obvious reason that $T_i^{\psi}(f, g) \in k[f, g]$.

We are in a position to connect $T_i^{\psi}(f, g)$ with the concept of π -root introduced in §1. For this purpose we give the following definition

Definition 3.1. Let $h_1(y), \dots, h_s(y)$ be polynomials in $K[y]$ where $K = k \langle\langle t \rangle\rangle$. A common π -root $\sigma = \sum \alpha_j t^j + \pi t^{\delta}$ of $h_1(y), \dots, h_s(y)$ is said to be a distribution detector if given

- (1) $\deg_y h_i(y) = p_i$,
- (2) $h_i(\sigma) = h_{i\sigma}(\pi) t^{\lambda_i^*} + \text{higher terms in } t$

we have for some λ and $c_{ij} \in k$ the following

- (3) $\lambda_i^* = \lambda p_i$,
- (4) $h_{i\sigma}(\pi)^{p_j} = c_{ij} h_{j\sigma}(\pi)^{p_i}$.

Remark. Note that condition (4) is equivalent to the statement that with $p = \text{g.c.d.}(p_1, \dots, p_s)$ there exists a polynomial $h(\pi)$ of degree p such that

$$h_{i\sigma}(\pi) = c_i h(\pi)^{r_i}$$

with

$$p_i = p r_i.$$

The existence of a distribution detector $\sigma = \sum \alpha_j t^j + \pi t^{\delta}$ implies that the roots of $h_1(y), \dots, h_s(y)$ in the δ -disc centered at $\sum \alpha_j t^j$ are all distributed proportionally according to the degrees of $h_i(y)$.

Now we are in a position to state the first connection between the approximate roots $T_i^\psi(f, g)$ and the concept of π -root introduced in § 1. Namely we have

Proposition 3. 3. *Let $g(y) = g(x, y)$, $T_i^\psi(y) = T_i^\psi(f(x, y), g(x, y))$ and*

$$T_i^\psi(y)_f = T_i^\psi(f(x, y), g(x, y))_f$$

be considered as polynomials in $k\langle\langle t \rangle\rangle[y]$ with $x = t^{-1}$. Suppose $\sigma = \sum \alpha_j t^j + \pi t^\delta$ is a distribution detector for $g(y)$, $T_i^\psi(y)$, $i = 1, \dots, r$ where r is an integer with $M_r \leq e$ and $\text{ord } g(\sigma) < 0$. Then σ is a distribution detector after we add $T_{r+1}^\psi(y)_f$ to the above list of polynomials.

Proof. Let λ be the rational negative number satisfying

$$\text{ord } g(\sigma) = n\lambda < 0.$$

Then we have for $i = 1, \dots, r$

$$\text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda < 0$$

and

$$\text{ord } T_\alpha^\psi(\sigma) = \sum (-\mu_i \alpha_i) \lambda = (\deg_y T_\alpha^\psi(y)) \lambda.$$

We shall establish for $i = 1, \dots, r+1$

$$\text{ord } T_i^\psi(\sigma)_f = \deg_y^\psi(T_i(y))_f \lambda.$$

Note that for $i = 1$ both sides of the above equation are zeroes hence equal. In general let us consider the following equation mentioned in the conclusion (2) of Proposition 3. 2 (cf. the Remark after it).

$$(14)^* \quad T_{r+1}^\psi(y)_f = n_r T_r^\psi(y)^{n_r-1} T_r^\psi(y)_f + \sum C_{i,\beta} g^i(y) T_\beta^\psi(y).$$

By mathematical induction we have

$$\text{ord } T_r^\psi(\sigma)_f = \deg_y(T_r^\psi(y))_f \lambda$$

and, note $\lambda < 0$, the following for the terms in equation (14)*

$$\begin{aligned} \text{ord}(T_r^\psi(\sigma)^{n_r-1} T_r^\psi(\sigma)_f) &= \deg_y(T_r^\psi(y)^{n_r-1} T_r^\psi(y)_f) \lambda \\ &< (\deg_y g^i(y) T_\beta(y)) \lambda = \text{ord } g^i(\sigma) T_\beta(\sigma). \end{aligned}$$

Thus

$$\begin{aligned} \text{ord}(T_{r+1}^\psi(\sigma))_f &= \text{ord}(T_r^\psi(\sigma)^{n_r-1} T_r^\psi(\sigma)_f) \\ &= \deg_y(T_r^\psi(y)^{n_r-1} T_r^\psi(y)_f) \lambda = (\deg_y T_{r+1}^\psi(y)_f) \lambda. \end{aligned}$$

We have verified the condition (3) in Definition 3. 1. To verify the condition (4) we note that the leading coefficient $T(\pi)$ of $T_{r+1}^\psi(\sigma)_f$ is the leading coefficient of $T_r^\psi(\sigma)^{n_r-1} T_r^\psi(\sigma)_f$. We know that there is a polynomial $h(\pi)$ (cf. the remark after Definition 3. 1) such that the leading coefficients of $T_r^\psi(\sigma)$ and $T_r^\psi(\sigma)_f$ are of the form $c_i h(\pi)^{r_i}$ where r_i equals (the y -degree of the corresponding polynomial in y) $\cdot$ (a common constant q). Thus $T(\pi)$, the leading coefficient of $T_{r+1}^\psi(\sigma)_f$, equals $ch(\pi)^{r^*}$ where r^* equals $(\deg_y T_{r+1}^\psi(y)_f)q$. The condition (4) of Definition 3. 1 is thus verified. Q.E.D.

§ 4. The splitting and nonsplitting phenomena

From this section on we shall assume that the following Jacobian condition is satisfied for a pair of polynomials monic in y , $f(x, y)$ and $g(x, y)$ in $k[x, y]$ if the opposite is not explicitly stated.

Jacobian condition:

$$J_{x,y}(f, g) = \det \begin{vmatrix} \frac{\partial f(x, y)}{\partial x} & \frac{\partial f(x, y)}{\partial y} \\ \frac{\partial g(x, y)}{\partial x} & \frac{\partial g(x, y)}{\partial y} \end{vmatrix} = 1.$$

As usual we set $t = x^{-1}$ and consider the polynomials $f(x, y)$ and $g(x, y)$ as polynomials in y with coefficients in $k\langle\langle t \rangle\rangle$. We shall write $f(x, y)$ as $f(y)$ and $g(x, y)$ as $g(y)$ to simplify our notations. Our problem is to investigate the tree data of $f(y)$, $g(y)$ and $T_i^\psi(y) = T_i^\psi(f(x, y), g(x, y))$.

As we pointed out in section 1, the algebraic tool for our investigation is the notion of π -root $\sigma \in k[\pi]\langle\langle t \rangle\rangle$ for a polynomial $h(y)$ with

$$\sigma = \sum_{j < \delta} a_j t^j + \pi t^\delta.$$

Then we have established in § 1 that

(1) The number of roots of $h(y) = 0$ inside the disc centered at $\sum a_j t^j$ with the logarithmic radius δ is the π -degree of the leading coefficient of $h(\sigma)$ i.e. the multiplicity of σ as a π -root of $h(y) = 0$ (cf. Proposition 1.2).

(2) The leading coefficient of $h(\sigma)$ has two distinct roots for π iff the disc centered at $\sum a_j t^j$ with radius δ is a disc associated with the tree data of $h(y) = 0$ (cf. Lemma 1.1). Hence, if the leading coefficient of $h(\sigma)$ has only one root, then we have to shrink the radius to locate the disc associated with the tree data.

(3) Let $(\pi - b)$ be a factor of the leading coefficient of $h(\sigma)$ with multiplicity l . Then there are precisely l roots $\tau_1, \dots, \tau_l$ of the equation $h(y) = 0$ (cf. Proposition 1.2) which satisfy

$$\text{ord}(\tau_i - \sum a_j t^j + b t^\delta) > \delta$$

thus

$$\text{ord}(\tau_i - \tau_j) > \delta.$$

In view of the discussion preceding Definition 1.1 there is a disc associated with the tree data and of radius less than δ which contains those l roots $\tau_1, \dots, \tau_l$. Note that the size of this disc is not determined.

To summarize the preceding discussions we have *the nonsplitting phenomena* (cf. Proposition 4. 4): If the leading coefficient of $h(\sigma)$ has only one root, then the δ -disc has to be shrunk to determine the tree data. *The splitting phenomena* (cf. Proposition 4. 6): If the leading coefficient of $h(\sigma)$ has distinct roots $b_1, \dots, b_s$ with multiplicities $l_1, \dots, l_s$ then the δ -disc can be covered by s proper subdisc containing $l_1, \dots, l_s$ roots respectively (cf. Definition 1. 1).

In this section we shall determine the conditions for the nonsplitting phenomena and the splitting phenomena for the polynomials $f(y)$, $g(y)$ and $T_i^\psi(y)$. For the splitting phenomena we shall further determine the necessary restrictions on the number of roots and the corresponding multiplicities.

Firstly we shall establish the following simple proposition.

Proposition 4. 1. *Let $x = t^{-1}$ and $\sigma = \sum a_j t^j + \pi t^\delta$ be a π -root of $g(y) = g(x, y)$. Given any polynomial $h(y) = h(f(x, y), g(x, y)) \in k[f(x, y), g(x, y)]$ with*

$$\begin{aligned} g(\sigma) &= g_\sigma(\pi) t^{\lambda_g} + \text{higher terms in } t, \\ h(\sigma) &= h_\sigma(\pi) t^{\lambda_h} + \text{higher terms in } t \end{aligned}$$

we have

$$\begin{aligned} (1) \quad J_{t,\pi}(h(\sigma), g(\sigma)) &= [\lambda_h h_\sigma(\pi) g'_\sigma(\pi) - \lambda_g g_\sigma(\pi) h'_\sigma(\pi)] t^{\lambda_g + \lambda_h - 1} + \dots \\ &= -h_f(\sigma) t^{-2+\delta}. \end{aligned}$$

Proof. The first equality follows trivially from the definition of Jacobians. For the second equality we have by the chain rule the following

$$\begin{aligned} J_{t,\pi}(h(\sigma), g(\sigma)) &= J_{f,g}(h, g) J_{x,y}(f, g) J_{t,\pi}(x, \sigma) \\ &= h_f(\sigma) \cdot 1 \cdot J_{t,\pi}(x, \sigma) = -h_f(\sigma) t^{-2+\delta}. \end{aligned} \quad \text{Q.E.D.}$$

The ordinary differential operator involving two polynomials $h_\sigma(\pi)$ and $g_\sigma(\pi)$ in the parenthesis of the equation (1) is of some interest to us. In Appendix I we shall give a detailed discussion of the polynomial solutions of some equations involving this differential operator. For the convenience of our discussions we shall introduce the following definition.

Definition 4. 1. Let a, b be constants in k and $p(\pi), q(\pi)$ be two polynomials in $k[\pi]$. Then the symbol $D(a, b, p(\pi), q(\pi))$ is defined as

$$D(a, b, p(\pi), q(\pi)) = ap(\pi) q'(\pi) - bq(\pi) p'(\pi).$$

In Proposition 3. 3 we have established that under the condition $\text{ord } g(\sigma) < 0$ if σ is a distribution detector for $g(y), T_1^\psi(y), \dots, T_r^\psi(y)$ then it is a distribution detector for $T_{r+1}(y)_f$. The following proposition will allow us to proceed one step further if certain conditions are satisfied.

Proposition 4. 2. Let $\sigma = \sum a_j t^j + \pi t^\delta$ be a distribution detector (cf. Definition 3. 1) for $g(y) = g(x, y)$, $T_i^\psi(y) = T_i^\psi(f(x, y), g(x, y))$ where $i = 1, \dots, r$. Let

$$\deg_y g(y) = n.$$

Suppose for some $\lambda < 0$ we have

- (1) $\text{ord } g(\sigma) = n\lambda$,
- (2) $\text{ord } T_{r+1}^\psi(\sigma) = (-\mu_{r+1})\lambda$,
- (3) $\lambda < \frac{-1 + \delta}{n - M_{r+1}}.$

Then σ is a distribution detector for $g(y)$, $T_i^\psi(y)$, $i = 1, \dots, r+1$.

Proof. We observe the following

$$\lambda < \frac{-1 + \delta}{n - M_{r+1}} \Leftrightarrow n\lambda + (-\mu_{r+1})\lambda - 1 < -2 + \delta + (-\mu_{r+1} + M_{r+1})\lambda.$$

Note it follows from the proof of Proposition 3. 3 and Definition 3. 1 and Proposition 3. 2 that

$$\text{ord } T_{r+1}^\psi(\sigma)_f = (\deg_y T_{r+1}^\psi(y)_f)\lambda = (-\mu_{r+1} + M_{r+1})\lambda.$$

Thus the left hand side of equation (1) of Proposition 4. 1 with $h = T_{r+1}^\psi$ has a smaller formal order. We must then have

$$(4) \quad D(-\mu_{r+1}\lambda, n\lambda, T_{r+1,\sigma}^\psi(\pi), g_\sigma(\pi)) = 0.$$

Let

$$\deg g_\sigma(\pi) = n^*$$

and

$$\frac{n\lambda}{n^*} = c \neq 0.$$

Then the above ordinary differential equation (4) can be reduced to the following form

with $a = -\mu_{r+1} \frac{\lambda}{c}$

$$(5) \quad cD(a, n^*, T_{r+1,\sigma}^\psi(\pi), g_\sigma(\pi)) = 0.$$

It follows from Proposition A.1 of Appendix I that

$$\deg T_{r+1,\sigma}^\psi(\pi) = \frac{(-\mu_{r+1})\lambda}{c} = (-\mu_{r+1})^*$$

and

$$g_\sigma(\pi)^{(-\mu_{r+1})^*} = C_1 (T_{r+1,\sigma}^\psi(\pi))^{n^*}$$

where $(-\mu_{r+1})^*$ is thus defined and C_1 is a nonzero constant. Our proposition follows from Definition 3. 1. Q.E.D.

Before handling the general nonsplitting phenomena, we shall state the following proposition which is illustrative (cf. [A.1], pp. 158, 159). Recall that M_h is the last M_i in our notation.

Proposition 4.3. *Suppose that $\deg g(x, y) = \deg_y g(x, y) = n > 1$, and $M_i \neq n-2$ for $i = 1, \dots, h$. Then the highest homogeneous forms of $g(x, y)$, $T_i^\psi(f(x, y), g(x, y))$ for all i with $M_i < n-2$ are powers of a common linear form (i.e. one point at infinite).*

Proof. Let δ be the number defined as

$$\delta = \min \left\{ \text{ord } \tau : \tau \text{ root of } g(y) \prod_{i=1}^h T_i^\psi(y) \right\}.$$

Let σ be the π -root of $g(y) \prod_{i=1}^h T_i^\psi(y)$ defined as

$$\sigma = \pi t^\delta.$$

Let us consider the possibility of $\delta > -1$. With this possibility we know that all roots of the equation $g(y) = 0$ are in the disc centered at zero with logarithmic radius $\delta > -1$. Then we set

$$\sigma^* = \pi t^{-1}$$

and conclude (cf. Proposition 1.3) that

$$g(\sigma^*) = \pi^n t^{-n} + \dots.$$

The above statement is equivalent to the degree n homogeneous form of $g(x, y)$ being y^n . Similarly the highest homogeneous forms of $T_i(f(x, y), g(x, y))$ are powers of y^n . Our proposition is thus trivial. Let us henceforth assume that

$$\delta \leq -1.$$

There are two cases i.e. (1) $M_h \neq n-1$, thus $M_h < n-2$ (cf. Lemma 2.1) (2) $M_h = n-1$.

Case 1. We shall apply Proposition 4.2. It follows straightforward from the definition of δ that we have

$$\begin{aligned} \deg g(\pi) &= n, \\ \text{ord } g(\sigma) &= n\delta, \\ \deg T_{i,\sigma}^\psi(\pi) &= \deg_y T_i^\psi(y) = -\mu_i \quad \text{for } i = 1, \dots, h, \\ \text{ord } T_i^\psi(\sigma) &= (-\mu_i)\delta \quad \text{for } i = 1, \dots, h. \end{aligned}$$

We shall set

$$\lambda = \delta.$$

Now it is simply assumed that

$$M_i < n-2 \quad \text{for } i = 1, \dots, h.$$

Then we must have

$$\lambda = \delta < \frac{-1}{n - M_i - 1}$$

which is equivalent to

$$\lambda = \delta < \frac{-1 + \delta}{n - M_i}.$$

Thus the conditions of Proposition 4.2 are verified. The conclusion of Proposition 4.2 states that σ is a distribution detector for $g(y)$, $T_1(y), \dots, T_h(y)$. Recall

$$\text{g.c.d.}(n, -\mu_1, \dots, -\mu_h) = 1.$$

Thus the polynomials $g_\sigma(\pi)$, $T_{1,\sigma}^\psi(\pi), \dots, T_{h,\sigma}^\psi(\pi)$ are powers of a common linear polynomial (cf. Definition 3.1).

We claim that

$$\delta = -1.$$

Otherwise suppose for a moment that

$$\delta < -1.$$

Due to the assumption

$$\deg_y g(x, y) = \deg g(x, y)$$

it is elementary to see

$$g(\sigma) = \pi^n t^{n\delta} + \dots.$$

Thus the common linear polynomial for $g_\sigma(\pi)$, $T_{1,\sigma}^\psi(\pi), \dots, T_{h,\sigma}^\psi(\pi)$ must be simply π . It contradicts the very definition of δ . Hence our claim has been proved. Let

$$g(x, y) = g_n(x, y) + \text{lower terms in } x, y.$$

Then it is easy to see

$$g(t^{-1}, \pi t^{-1}) = g_n(1, \pi) t^{-n} + \text{higher terms in } t.$$

Moreover we have proved that the polynomial $g_n(1, \pi) = g_\sigma(\pi)$ is a power of a linear polynomial in π . It follows trivially that the highest homogeneous forms of $g(x, y)$, $T_i(f(x, y), g(x, y))$ are powers of a linear form.

Case 2. $M_h = n - 1$. The following Proposition 4.4 will establish that the polynomials $g_\sigma(\pi)$, $T_{1,\sigma}^\psi(\pi), \dots, T_{h,\sigma}^\psi(\pi)$ are powers of a common linear polynomial $(\pi - a)$. Then the last paragraph of the proof for Case 1 can be used verbatim to finish Case 2. Q.E.D.

The following proposition gives some sufficient conditions for the nonsplitting phenomena.

Proposition 4. 4. Suppose that for some $r, \lambda, v, \varepsilon$ and a π -root $\sigma = \sum a_j t^j + \pi t^\delta$ of $g(y)$ we have the following:

- (1) $\deg_y g(y) = n > 1$.
- (2) $\text{ord } g(\sigma) = n\lambda < 0$,
 $\text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda \quad i = 1, \dots, r-1$.
- (3) $\lambda < \frac{-1 + \delta}{n - M_i}$.
- (4) $\deg g_\sigma(\pi) = v \left(\frac{n}{d_r} \right)$.
- (5) $\deg T_{r,\sigma}^\psi(\pi) = v \left(\frac{-\mu_r + M_r - n}{d_r} \right) + 1$.
- (6) $\text{ord } T_r^\psi(\sigma) = (-\mu_r + M_r - n)\lambda - 1 + \delta$
 $= (\deg T_{r,\sigma}^\psi(\pi) + \varepsilon) \left(\frac{d_r}{v} \right) \lambda$.

(7) The number ε defined above satisfies $\varepsilon > -1$.

Then $g_\sigma(\pi), T_{1,\sigma}^\psi(\pi), \dots, T_{r,\sigma}^\psi(\pi)$ are powers of a common linear polynomial in π .

Remark. This proposition has been applied to the second case of the preceding proposition with $r = h$, $M_r = n - 1$, $\lambda = \delta$, $v = d_h$ and $\varepsilon = -\frac{1}{\delta}$.

Proof of the proposition. It follows from the conditions (1), (2), (3), and Proposition 4. 2 that σ is a distribution detector for $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$. Thus the π -degrees of the leading coefficients of $g(\sigma), T_1^\psi(\sigma), \dots, T_{r-1}^\psi(\sigma)$ are proportional to

$$n : -\mu_1 : \dots : -\mu_{r-1}.$$

Note that $\text{g.c.d.}(n, -\mu_1, \dots, -\mu_{r-1}) = d_r$. Hence the said degrees must be $v \left(\frac{n}{d_r} \right), v \left(\frac{-\mu_1}{d_r} \right), \dots, v \left(\frac{-\mu_{r-1}}{d_r} \right)$ for the integer v satisfying condition (4). It follows from the definition of the distribution detectors that there is a polynomial $p(\pi)$ of degree V such that

$$\begin{aligned} g_\sigma(\pi) &= C_0 p(\pi)^{\frac{n}{d_r}}, \\ T_{1,\sigma}^\psi(\pi) &= C_1 p(\pi)^{\frac{-\mu_1}{d_r}}, \\ &\vdots \\ T_{r-1,\sigma}^\psi(\pi) &= C_{r-1} p(\pi)^{\frac{-\mu_{r-1}}{d_r}}, \\ (T_{r,\sigma}^\psi(\pi))_f &= C_r p(\pi)^{\frac{-\mu_r + M_r}{d_r}}. \end{aligned}$$

In view of the condition (6) after comparing the orders on both sides of equation (1) of Proposition 4.1 with $h(f, g) = T_r^\psi(f, g)$ we conclude that $T_{r,\sigma}^\psi(\pi)$ satisfies the following differential equation

$$\begin{aligned} & (\deg T_{r,\sigma}^\psi(\pi) + \varepsilon) T_{r,\sigma}^\psi(\pi) g'_\sigma(\pi) - \deg g_\sigma(\pi) g_\sigma(\pi) (T_{r,\sigma}^\psi(\pi))' \\ & = c p(\pi)^{\frac{-\mu_r + M_r}{d_r}}. \end{aligned}$$

After simplifying, the above equation becomes

$$\begin{aligned} & (\deg T_{r,\sigma}^\psi(\pi) + \varepsilon) T_{r,\sigma}^\psi(\pi) p'(\pi) - v p(\pi) (T_{r,\sigma}^\psi(\pi))' \\ & = c^* p(\pi)^{\frac{-\mu_r + M_r - n}{d_r} + 1}. \end{aligned}$$

It follows from Proposition A.2 of Appendix I that the polynomials $p(\pi)$ and $T_{r,\sigma}^\psi(\pi)$ are powers of a common linear polynomial in π . Our proposition is thus established.

Q.E.D.

Remark. With a verbatim proof, for a slightly general Jacobian condition of the following form (cf. Appendix II)

$$J_{x,y}(f(x, y), g(x, y)) = x^l$$

Proposition 4.4 is still valid with the condition (6) replaced by the following

$$\begin{aligned} (6)^* \quad \text{ord } T_r^\psi(\sigma) &= (-\mu_r + M_r - n)\lambda - 1 + \delta - l \\ &= (\deg T_{r,\sigma}^\psi(\pi) + \varepsilon) \left(\frac{d_r}{v} \right) \lambda. \end{aligned}$$

The following proposition will be proved after Proposition 4.6. The purpose of inserting Proposition 4.5 is to illustrate the splitting phenomena. One interpretation of Proposition 4.5 is that there are at most two points at infinite for $f(x, y) = 0$. Note that the conditions of Proposition 4.5 complete the conditions of Proposition 4.3 (cf. [A.1], p. 139).

Proposition 4.5. *Suppose that $\deg g(x, y) = \deg_y g(x, y) = n > 1$ and $M_r = n - 2$ for some r . Then the highest homogeneous forms of $g(x, y)$, $T_i(f(x, y), g(x, y))$ for $i < r$ are powers of a common form which has at most two distinct linear factors. Moreover the multiplicities for the two factors are different.*

The following proposition gives some sufficient conditions for the general splitting phenomena.

Proposition 4. 6. Suppose that for some r, λ, v and a π -root $\sigma = \sum a_j t^j + \pi t^\delta$ of $g(y)$ we have the following:

- (1) $\deg_y g(y) = n > 1$,
- (2) $\text{ord } g(\sigma) = n\lambda < 0$, $\text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda \quad i = 1, \dots, r$,
- (3) $\lambda = \frac{-1 + \delta}{n - m_r} < \frac{-1 + \delta}{n - m_i} \quad i = 1, \dots, r-1$,
- (4) $\deg g_\sigma(\pi) = v \left(\frac{n}{d_r} \right)$,
- (5) $\deg T_{r,\sigma}^\psi(\pi) = v \left(\frac{-\mu_r}{d_r} \right)$
 $= v \left(\frac{-\mu_r + M_r - n}{d_r} \right) + v \left(\frac{n - M_r}{d_r} \right).$

Then we have, if $r \geq 2$, the following:

(1) σ is a distribution detector for $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$. The leading coefficients of $g(\sigma), T_1^\psi(\sigma), \dots, T_{r-1}^\psi(\sigma)$ are powers of a common polynomial $p(\pi)$ of degree v ,

$$(2) \quad T_{r,\sigma}^\psi(\pi) = p(\pi)^{\frac{-\mu_r + M_r - n}{d_r}} \cdot q(\pi) \text{ where } q(\pi) \text{ is a polynomial of degree } v \left(\frac{n - M_r}{d_r} \right),$$

(3) $q(\pi)$ has all distinct roots,

(4) all roots of $p(\pi)$ are roots of $q(\pi)$,

(5) $p(\pi)$ is not a power of $q(\pi)$.

Moreover if $r = 1$ then $g_\sigma(\pi)$ and $T_{1,\sigma}^\psi(\pi)$ satisfy the following differential equation

$$D(n, -M_1, g_\sigma(\pi), T_{1,\sigma}^\psi(\pi)) = \text{nonzero constant}.$$

Proof. Let us consider the case $r \geq 2$. The conclusion (1) follows from Proposition 4. 2 and Definition 3. 1.

Let us apply Proposition 4. 1 for $h(y) = T_r^\psi(y)$. It follows from the conditions (2) and (3) that we have

$$\begin{aligned} \lambda = \frac{-1 + \delta}{n - m_r} &\Leftrightarrow (n - m_r)\lambda = -1 + \delta \\ &\Leftrightarrow n\lambda + (-\mu_r)\lambda - 1 = -2 + \delta + (-\mu_r + M_r)\lambda \\ &\Leftrightarrow \text{the order of both sides of the equation (1)} \\ &\quad \text{of proposition 4. 1 are equal.} \end{aligned}$$

Thus we conclude that the polynomials $g_\sigma(\pi)$ and $T_{r,\sigma}^\psi(\pi)$ must satisfy the following differential equation

$$(6) \quad D(n\lambda, (-\mu_r)\lambda, g_\sigma(\pi), T_{r,\sigma}^\psi(\pi)) \\ = C \text{ the leading coefficient of } T_{r,\sigma}^\psi(\pi)_f.$$

Note that the conclusion (1) and Proposition 3.3 imply

$$g_\sigma(\pi) = C_0 p(\pi)^{\frac{n}{d_r}}, \\ (T_{r,\sigma}^\psi(\pi))_f = \text{the leading coefficient of } T_{r,\sigma}^\psi(\sigma)_f \\ = C_r p(\pi)^{\frac{-\mu_r + M_r}{d_r}}.$$

The differential equation (6) can be simplified as follows

$$D\left(v\left(\frac{-\mu_r}{d_r}\right)v, p(\pi), T_{r,\sigma}^\psi(\pi)\right) = C^* p(\pi)^{\frac{-\mu_r + M_r - n}{d_r} + 1}.$$

Note that

$$\deg T_{r,\sigma}^\psi(\pi) = v\left(\frac{-\mu_r}{d_r}\right) = v\left(\frac{-\mu_r + M_r - n}{d_r}\right) + v\left(\frac{n - M_r}{d_r}\right).$$

Then the conclusions (2), (3), (4) and (5) follow at once from Propositions A.3 and A.4 of Appendix 1.

Let us consider the case $r=1$. As in the preceding argument we shall apply Proposition 4.1 for $h(y) = T_1^\psi(y)$. Note that

$$T_1^\psi = f + \text{polynomial in } g.$$

It follows

$$(T_1^\psi)_f = 1.$$

It is elementary to show that orders of both sides of the equation (1) of Proposition 4.1 are equal. Thus the leading coefficients $g_\sigma(\pi)$ and $T_{1,\sigma}(\pi)$ must satisfy the said equation.

Q.E.D.

Remark. With a verbatim proof, for a slightly general Jacobian condition of the following form (cf. Appendix II)

$$J_{x,y}(f(x, y), g(x, y)) = x^l$$

Proposition 4.6 is still valid with the condition (3) replaced by the following:

$$(3)^* \quad \lambda = \frac{-1 - l + \delta}{n - m_r} < \frac{-1 - l + \delta}{n - m_i} \quad i = 1, \dots, r-1.$$

Proof of Proposition 4.5. Let δ be the number defined as

$$\delta = \min \left\{ \text{ord } \tau : \tau \text{ root of } g(y) \prod_{i=1}^r T_i^\psi(y) \right\}.$$

Using the first paragraph of the proof of Proposition 4.3 verbatim we reduce the proof to the case that

$$\delta \leq -1.$$

We claim

$$\delta = -1.$$

Otherwise assume for a moment

$$\delta < -1.$$

Then we shall apply Proposition 4.2 with $(r-1)$ replacing r , $\lambda = \delta$ and $\sigma = \pi t^\delta$. The conditions (1) and (2) are trivially verified. To verify the condition (3) we observe

$$\lambda = \delta < \frac{-1 + \delta}{n - M_r} = \frac{-1 + \delta}{2} \Leftrightarrow \delta < -1.$$

Thus σ is a distribution detector for $g(y)$, $T_1(y), \dots, T_r(y)$. With the assumption that

$$\deg g(x, y) = \deg_y g(x, y)$$

we see at once

$$g(\sigma) = \pi^n t^{n\delta} + \dots.$$

It follows from the definition of the distribution detectors (cf. Definition 3.1) that

$$T_i^\psi(\sigma) = \pi^{-\mu_i} t^{-\mu_i \delta} + \dots.$$

It thus contradicts the definition of δ . Our claim is established.

We shall now apply Proposition 4.6 with $r = r$, $\lambda = \delta = -1$, $v = d_r$ and $\sigma = \pi t^\delta$. It is trivial to verify all conditions. Note that $r \geq 2$. Otherwise we have $M_1 = n - 2 \geq 0$. Note that we then have from the definition of M_1

$$\begin{aligned} n \nmid M_1 &\Rightarrow M_1 > 0, \\ \deg_y T_1^\psi(f(x, y), g(x, y)) &= -\mu_1 = -M_1 < 0. \end{aligned}$$

A contradiction. The polynomial $q(\pi)$ in the conclusion (2) is of degree

$$\deg q(\pi) = v \left(\frac{n - M_r}{d_r} \right) = 2.$$

Thus the polynomial $q(\pi)$ has two distinct roots. Since the polynomial $p(\pi)$ has at most two roots (cf. the conclusion (4)) with distinct multiplicities (cf. the conclusion (5)). Thus our proposition follows from the conclusion (2). Q.E.D.

§ 5. The major discs

In this section we shall study the root configurations of $g(y) \prod_{i=1}^r T_i^\psi(y)$ for $r=1, \dots, h$. Propositions 4.4 and 4.6 are our main tools. Topologically in both propositions we investigate certain δ -discs. Under the conditions of Proposition 4.4 we prove that all roots of $g(y) \prod_{i=1}^r T_i^\psi(y)$ in it have distances less than $2^{-\delta}$ or the order of their differences are bigger than δ . In the terminology of Section 1 this particular δ -disc is not one of the discs associated with the tree data of $g(y) \prod_{i=1}^r T_i^\psi(y)$ (cf. Definition 1.1). In other words we have to further shrink the size of the discs to locate the tree data. After finitely many shrinking steps we expect that the conditions of Proposition 4.6 will be satisfied. Proposition 4.6 implies that there are roots of $g(y) \prod_{i=1}^r T_i(y)$ in the corresponding δ -disc with distances equal to $2^{-\delta}$. Thus the δ -disc $D_{1,i_2,\dots,i_r}$ is associated with the tree data of $g(y) \prod_{i=1}^r T_i^\psi(y)$. Moreover if $r \geq 2$ then the roots in it will be covered by subdiscs $D_{1,i_2,\dots,i_r,j}$ satisfying the criterion of Definition 1.1. Furthermore Proposition 4.6 specifies that the number of subdiscs $D_{1,i_2,\dots,i_r,j}$ is the degree of $q(\pi)$, and the number of roots in those subdiscs are unequal. If we choose one of the subdiscs with the number of roots above a certain number (cf. Definition 5.1) and drop $T_r^\psi(y)$ from the product $g(y) \prod_{i=1}^r T_i^\psi(y)$, i.e. we shall consider $g(y) \prod_{i=1}^{r-1} T_i^\psi(y)$, then the phenomena of shrinking the size of the disc until the conditions of Proposition 4.6 being met will reappear again. This way we plan to get certain information of the root configurations of $g(y) \prod_{i=1}^r T_i^\psi(y)$, especially $g(y) T_1^\psi(y)$. The most important result of this section is Proposition 5.3 which establishes the existence of major towers ending in D_1 (cf. Definition 5.1). The complicated formula of δ_r can be used to show the reductions of "one point at infinite" (cf. Proposition 5.4) and "the two characteristic pair case" (cf. Proposition 5.5).

We shall start our process by establishing the following proposition.

Proposition 5.1. *Suppose that $g(x, y)$ is monic in y with y -degree $n > 1$. Let δ_s be defined for $s = h-1, h$ as follows*

$$\delta_s = \min \left\{ \text{ord}(\tau_i - \tau_j) : \tau_i, \tau_j \text{ roots of } g(y) \prod_{i=1}^s T_i^\psi(y) = 0 \right\}.$$

Then we have:

(1) If $M_h = n-1$ then $\delta_h = \delta_{h-1} = \frac{-1}{n - M_{h-1} - 1}$. Moreover the conditions of Proposition 4.6 are satisfied with σ the unique π -root in the δ_{h-1} -disc D_{h-1} containing all roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$, $\delta = \delta_{h-1} = \lambda$, $r = h-1$ and $v = d_{h-1}$.

(2) If $M_h < n-1$, then $\delta_h = \frac{-1}{n-M_h-1}$. Moreover the conditions of Proposition 4.6 are satisfied with σ the unique π -root in the δ_h -disc D_h containing all roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$, $\delta = \delta_h = \lambda$, $r = h$ and $v = d_h$.

Definition – Remark. The conclusion (1) implies that if $M_h = n-1$, then we should drop it from our own consideration. The “effective” characteristic pairs exclude M_h if $M_h = n-1$. The last effective characteristic pair is denoted by M_s in this article.

Proof. (1) Let us assume for a moment that

$$\delta_h < \frac{-1}{n-M_{h-1}-1}.$$

Then the conditions of Proposition 4.4 can be verified with σ the unique π -root of $g(y) \prod_{i=1}^h T_i^\psi(y)$ in the δ_h -disc and let

$$\delta = \delta_h = \lambda < 0, \quad r = h, \quad v = d_h, \quad \varepsilon = -\frac{1}{\delta_h}.$$

The conclusion of Proposition 4.4 implies that

$$\text{ord}(\tau_i - \tau_j) > \delta = \delta_h$$

for all τ_i, τ_j roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$. It contradicts the definition of δ_h . Thus we must have

$$\delta_h \geq \frac{-1}{n-M_{h-1}-1} = \delta^*.$$

Now we shall verify the conditions of Proposition 4.6 for σ the unique π -root of $g(y) \prod_{i=1}^h T_i^\psi(y)$ in the δ^* -disc and let

$$\delta = \delta^* = \lambda < 0, \quad r = h-1, \quad v = d_{h-1}.$$

This verification is elementary. Note the polynomial factor $q(\pi)$ of $T_{h-1, \sigma}^\psi(\pi)$ in Proposition 4.6 is with degree as follows

$$\deg q(\pi) = n - M_{h-1} \geq 2$$

and has distinct roots. Thus we have

$$\delta_h = \delta_{h-1} = \frac{-1}{n-M_{h-1}-1}$$

and the conclusion (1) is finished.

(2) Let us assume for a moment that

$$\delta_h < \frac{-1}{n-M_h-1}.$$

Then the conditions of Proposition 4.2 can be easily verified with σ the unique π -root of $g(y) \prod_{i=1}^h T_i^\psi(y)$ in the δ_h -disc and let

$$\delta = \delta_h = \lambda < 0, \quad r = h, \quad v = d_h.$$

Note that

$$\text{g.c.d.}(n, -\mu_1, \dots, -\mu_h) = 1.$$

The conclusion of Proposition 4.2 implies that the leading coefficients of $g(\sigma)$, $T_1^\psi(\sigma), \dots, T_h^\psi(\sigma)$ are powers of a polynomial of degree 1. Thus

$$\text{ord}(\tau_i - \tau_j) > \delta = \delta_h$$

for all roots τ_i, τ_j of $g(y) \prod_{i=1}^h T_i^\psi(y)$. It contradicts the definition of δ_h . Thus we must have

$$\delta_h \geq \frac{-1}{n - M - 1} = \delta^*.$$

Now we shall verify the conditions of Proposition 4.6 for σ the unique π -root of $g(y) \prod_{i=1}^h T_i^\psi(y)$ in the δ^* -disc and let

$$\delta = \delta^* = \lambda < 0, \quad r = h, \quad v = d_h.$$

This verification is elementary. Note the polynomial factor $q(\pi)$ of $T_h^\psi(\pi)$ in Proposition 4.6 is with degree as follows

$$\deg q(\pi) = n - M_h \geq 2$$

and has distinct roots. Thus we have

$$\delta_h = \frac{-1}{n - M_h - 1}$$

and our proposition is finished. Q.E.D.

In fact later on we shall prove that the possibility of

$$\delta_h = \frac{-1}{n - M_1 - 1}$$

i.e. $h=2$ in the first case, $h=1$ in the second case, will imply that

$$h[f, g] = h[T_1, g] = h[x, y].$$

However we postpone the proof of the above statement to avoid the repetition of certain arguments. Let us examine the conclusion of Proposition 4.6 for $h-1 \geq 2$ in the first case and $h \geq 2$ in the second case. Let $s = h-1$ in the first case and $s = h$ in the second case, then we have the following:

(1) All roots of $p(\pi)$ are roots of $q(\pi)$ which is a polynomial of degree $n - M_s$ and with distinct roots,

(2) the polynomial $p(\pi)$ with degree d_s is not a power of $q(\pi)$.

Thus among the roots of $p(\pi)$ there must be some with multiplicities bigger than $\frac{d_s}{n - M_s}$. Let c_s be any one of them. Then we shall establish

Proposition 5.2. *Suppose that $g(x, y)$ is monic in y with y -degree $n > 1$. Let M_s be the largest M_i which is less than $n - 1$. Let δ_s be defined as in Proposition 5.1 and suppose $s \geq 2$. Let $\sigma \sum a_j t^j + \pi t^{\delta_s}$ be the unique π -root of $g(y) \prod_{i=1}^s T_i^\psi(y)$ in the δ_s -disc D_s and $\pi - c_s$ be a factor of $p(\pi)$ (cf. Proposition 4.6) with multiplicity $V_s > \frac{d_s}{m - M_s}$. Let τ and δ_{s-1} be defined as follows*

$$\begin{aligned} \tau &= \sum a_j t^j + c_s t^{\delta_s} \in D_s, \\ \delta_{s-1} &= \min \left\{ \text{ord}(\tau_i - \tau_j) : \tau_i \tau_j \text{ roots of } g(y) \prod_{i=1}^s T_i^\psi(y) \text{ with } \text{ord}(\tau_i - \tau) > \delta_s \right\} \\ &= \text{the logarithmic radius of } D_{s-1} \text{ which contains all roots } \tau_i \text{ of } g(y) \prod_{i=1}^s T_i^\psi(y) \text{ with } \\ &\quad \text{ord}(\tau_i - \tau) > \delta_s. \end{aligned}$$

Then we have the following

$$\delta_{s-1} = 1 - \frac{(n - M_{s-1}) [V_s(n - M_s) - d_s]}{(n - M_s - 1) [V_s(n - M_{s-1}) - d_s]}.$$

Moreover the conditions of Proposition 4.6 are satisfied with σ^* the unique π -root of $g(y) \prod_{i=1}^{s-1} T_i(y)$ in the δ_{s-1} -disc D_{s-1} , $\delta = \delta_{s-1} = \lambda$, $r = s - 1$ and $v = V_s \left(\frac{d_{s-1}}{d_s} \right)$.

The above proposition is a special case of Proposition 5.3 which will be stated and proved in a moment. We shall prove the following numerical lemmas. The readers are advised to skip them for the time being and return to them when are needed.

Lemma 5.1. *Let L, V_s be numbers with $L \leq M_s < n - 1$ and $V_s(n - M_s) > d_s$. Let $\delta(L)$, δ_s and $\lambda(L)$ be defined as*

$$\begin{aligned} \delta(L) &= \delta^* = 1 - \frac{(n - L) [V_s(n - M_s) - d_s]}{(n - M_s - 1) [V_s(n - L) - d_s]}, \\ \delta_s &= \delta(M_s) = 1 - \frac{n - M_s}{n - M_s - 1} = \frac{-1}{n - M_s - 1}, \\ \lambda(L) &= \lambda^* = d_s \delta_s + V_s(\delta^* - \delta_s). \end{aligned}$$

Then $\delta(L)$ is a monotonically decreasing function of L and we have

$$\left(\frac{(n - L)}{d_s} \right) \lambda^* = -1 + \delta^* < 0.$$

Proof. It is easy to see that $\delta(L)$ is a monotonically decreasing function of L . To establish

$$-1 + \delta^* < 0$$

we simply observe the following

$$-1 + \delta^* = -\frac{(n-L) [V_s(n-M_s) - d_s]}{(n-M_s-1) [V_s(n-L) - d_s]}$$

and

$$n-L \geq (n-M_s) > (n-M_s)-1 > 0,$$

$$V_s(n-L) - d_s > V_s(n-M_s) - d_s > 0.$$

We shall establish the equality

$$\begin{aligned} \frac{n-L}{d_s} \lambda^* = -1 + \delta^* &\Leftrightarrow (n-L) [(d_s - V_s) \delta_s + V_s \delta^*] = (-1 + \delta^*) d_s \\ &\Leftrightarrow (n-L) [(d_s - V_s) (1 - \delta_s) + V_s (1 - \delta^*) - d_s] = (1 - \delta^*) d_s \\ &\Leftrightarrow [(n-L) V_s - d_s] (1 - \delta^*) = (n-L) d_s - (n-L) (d_s - V_s) (1 - \delta_s) \\ &= \frac{(n-L) [d_s(n-M_s-1) - (d_s - V_s)(n-M_s)]}{n-M_s-1} \\ &= \frac{(n-L) [V_s(n-M_s) - d_s]}{n-M_s-1} \end{aligned}$$

The last one is equivalent to the definition of $\delta(L) = \delta^*$.

Q.E.D.

The following lemma is a generalization of the preceding lemma and will be used in the proof of Proposition 5.3.

Lemma 5.2. Let L and V_i for $i=r, r+1, \dots, s$ be numbers with

$$L \leq M_r < \dots < M_s < n-1$$

and

$$V_i(n-M_i) > d_i.$$

Let $\delta(L)$, δ_i and $\lambda(L)$ be defined as

$$\delta(L) = \delta^* = 1 - \frac{(n-L) \prod_{j=r}^s [V_j(n-M_j) - d_j]}{(n-M_s-1) [V_r(n-L) - d_r] \prod_{j=r+1}^s [V_j(n-M_{j-1}) - d_j]},$$

$$\delta_i = 1 - \frac{(n-M_i) \prod_{j=i+1}^s [V_j(n-M_j) - d_j]}{(n-M_s-1) \prod_{j=i+1}^s [V_j(n-M_{j-1}) - d_j]},$$

$$\lambda(L) = \lambda^* = d_r \delta_s + \sum_{i=r+1}^s V_i \left(\frac{d_r}{d_i} \right) (\delta_{i-1} - \delta_i) + V_r (\delta^* - \delta_r).$$

Then $\delta(L)$ is a monotonically decreasing function of L and we have

$$\left(\frac{n-L}{d_r}\right)\lambda^* = -1 + \delta^* < 0.$$

Proof. It is easy to see that $\delta(L)$ is a monotonically decreasing function of L . To establish

$$-1 + \delta^* < 0$$

we observe that

$$-1 + \delta^* = \frac{-(n-L) \prod_{j=r}^s [V_j(n-M_j) - d_j]}{(n-M_s-1) [V_r(n-L) - d_r] \prod_{j=r+1}^s [V_j(n-M_{j-1}) - d_j]}$$

and

$$\begin{aligned} (n-L) &\geq (n-M_s) > (n-M_s-1) > 0, \\ V_r(n-L) - d_r &\geq V_r(n-M_s) - d_r > 0, \\ V_j(n-M_{j-1}) - d_j &> V_j(n-M_j) - d_j > 0 \end{aligned}$$

which imply

$$-1 + \delta^* < 0.$$

We shall establish the equality

$$\left(\frac{n-L}{d_s}\right)\lambda^* = -1 + \delta^*.$$

Note the following identities:

$$\begin{aligned} \delta_{i-1} - \delta_i &= \frac{\prod_{j=i+1}^s [V_j(n-M_j) - d_j]}{(n-M_s-1) \prod_{j=i+1}^s [V_j(n-M_{j-1}) - d_j]} \left\{ (n-M_i) - \frac{(n-M_{i-1}) [V_i(n-M_i) - d_i]}{V_i(n-M_{i-1}) - d_i} \right\} \\ &= \frac{d_i(M_i - M_{i-1}) \prod_{j=i+1}^s [V_j(n-M_j) - d_j]}{(n-M_s-1) \prod_{j=i}^s [V_j(n-M_{j-1}) - d_j]}. \end{aligned}$$

Substitute them into the following summation

$$\begin{aligned} \frac{n-L}{d_r} \lambda^* &= \frac{n-L}{d_r} \left[d_r \delta_s + V_s \frac{d_r}{d_s} (\delta_{s-1} - \delta_s) + \dots \right. \\ &\quad \left. + V_{r+1} \frac{d_r}{d_{r+1}} (\delta_r - \delta_{r+1}) + V_r (\delta^* - \delta_r) \right]. \end{aligned}$$

Note that

$$\begin{aligned}\delta_s + \frac{V_s}{d_s}(\delta_{s-1} - \delta_s) &= \frac{-1}{n - M_s - 1} + \frac{V_s(M_s - M_{s-1})}{(n - M_s - 1)[V_s(n - M_{s-1}) - d_s]} \\ &= -\frac{V_s(n - M_s) - d_s}{(n - M_s - 1)[V_s(n - M_{s-1}) - d_s]}.\end{aligned}$$

After repeating the above computation $s - r$ times we see at once that

$$\left(\frac{n-L}{d_r}\right)\lambda^* = -1 + \delta^*.$$

Q.E.D.

Now we shall treat the generalization of Proposition 5.1 and 5.2. Inductively suppose that for some $r \geq 2$ we identify a tower of *major discs* $D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_r$, defined as

Definition 5.1. A tower of major discs $D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_r$ satisfies the following four criteria with a sequence of integers $\{V_i : i = (r+1, \dots, s+1)\}$:

(1) In the disc D_i the polynomials $g(y)$ have precisely $\frac{n}{d_{i+1}} V_{i+1}$ roots. $T_j^\psi(y)$ has precisely $\frac{-\mu_j}{d_{i+1}} V_{i+1}$ roots for $j = 1, \dots, i$,

(2) the numbers V_i satisfy

$$\frac{V_{i+1}d_i}{d_{i+1}} \geq V_i > \frac{d_i}{n - M_i} \quad \text{for } i = r+1, \dots, s, \quad V_{s+1} = d_{s+1},$$

(3) the disc D_i is with logarithmic radius δ_i where for $i = r, \dots, s$

$$\delta_i = 1 - \frac{(n - M_i) \prod_{j=i+1}^s [V_j(n - M_j) - d_j]}{(n - M_s - 1) \prod_{j=i+1}^s [V_j(n - M_{j-1}) - d_j]},$$

(4) for $i = r, \dots, s$, let $\sigma_i = \sum \alpha_j t^j + \pi t^{\delta_i}$ be the unique general point in the disc D_i then the conditions of Proposition 4.6 are satisfied for $\sigma = \sigma_i$, $\delta = \delta_i$ and $v = V_{i+1} \left(\frac{d_i}{d_{i+1}} \right)$.

Note that Proposition 5.1, which has been proved, implies that the minimal disc D_s which contains all roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$ satisfies the criteria (1), (2), (3) and (4). Proposition 5.2 simply states that among the disjoint minimal subdiscs of D_s (cf. Definition 1.1) which cover the roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$ in D_s any subdisc can be used as D_{s-1} if it contains more than the average number of roots. The existence of such a subdisc is ensured by Proposition 4.6.

Proposition 5.3. Suppose that $g(x, y)$ is monic in y with y -degree $n > 1$ and for $r \geq 2$ a tower of major discs $D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_r$ is constructed. Let $\pi - C_r$ be a factor of $p(\pi)$ as in the conclusions of Proposition 4.6 with multiplicity V_r satisfying

$$\deg p(\pi) = V_{r+1} \left(\frac{d_r}{d_{r+1}} \right) \geq V_r > \frac{d_r}{n - M_r}.$$

Let τ and δ_{r-1} be defined as follows

$$\begin{aligned} \tau &= \sum \alpha_j t^j + c_r t^{\delta_r} \in D_r, \\ \delta_{r-1} &= \min \left\{ \text{ord}(\tau_i - \tau_j): \tau_i, \tau_j \text{ roots of } g(y) \prod_{i=1}^r T_i^\psi(y) \text{ with } \text{ord}(\tau_i - \tau) > \delta_r \text{ and} \right. \\ &\quad \left. \text{ord}(\tau_j - \tau) > \delta_r \right\} \\ &= \text{the logarithmic radius of } D_{r-1} \text{ which is the minimal disc containing all roots} \\ &\quad \tau_i \text{ of } g(y) \prod_{i=1}^r T_i^\psi(y) \text{ with } \text{ord}(\tau - \tau_i) > \delta_r. \end{aligned}$$

Then we have the following:

$$\delta_{r-1} = 1 - \frac{(n - M_{r-1}) \prod_{j=r}^s [V_j(n - M_j) - d_j]}{(n - M_s - 1) \prod_{j=r}^s [V_j(n - M_{j-1}) - d_j]}.$$

Moreover the tower $D_s \supseteq \cdots \supseteq D_r \supseteq D_{r-1}$ is a tower of major discs.

Proof. The criterion (2) of Definition 5.1 is automatic. Let us handle the remaining ones. Let $\delta(L)$ denote

$$\delta(L) = 1 - \frac{(n - L) \prod_{j=r}^s [V_j(n - M_j) - d_j]}{(n - M_s - 1) [V_r(n - L) - d_r] \prod_{j=r+1}^s [V_j(n - M_{j-1}) - d_j]}.$$

Firstly we claim the following

$$\delta_{r-1} \geq \delta(M_{r-1}).$$

Let us assume the contrary for a moment. i.e.

$$\delta(M_r) = \delta_r < \delta_{r-1} < \delta(M_{r-1}).$$

Due to the monotony of $\delta(L)$ as a function of L , we have for some L^* with

$$M_r > L^* > M_{r-1}$$

the equality

$$\delta_{r-1} = \delta(L^*) = \delta^*.$$

Let σ^* be the unique π -root of $g(y) \prod_{i=1}^r T_i^\psi(y)$ in the δ_{r-1} -disc with

$$\text{ord}(\sigma^* - \tau) > \delta_r.$$

We shall verify the conditions of Proposition 4.4 with $\sigma = \sigma^*$, $r = r$, $v = V_r$ to produce a contradiction. Note that it follows from Proposition 1.3 that

$$\begin{aligned} \text{ord } g(\sigma^*) &= n\delta_s + \sum_{i=r+1}^s V_i\left(\frac{n}{d_i}\right)(\delta_{i-1} - \delta_i) + V_r\left(\frac{n}{d_r}\right)(\delta^* - \delta_r) \\ &= \left(\frac{n}{d_r}\right) \left[d_r\delta_s + \sum_{i=r+1}^s V_i\left(\frac{d_r}{d_i}\right)(\delta_{i-1} - \delta_i) + V_r(\delta^* - \delta_r) \right] \\ &= \frac{n}{d_r} \lambda^* = n\left(\frac{\lambda^*}{d_r}\right) = n\lambda. \end{aligned}$$

Then it follows from Lemma 5.2 that

$$\lambda^* < 0 \quad \text{or} \quad \lambda < 0.$$

It is straightforward to verify

$$\text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda \quad i = 1, \dots, r-1.$$

Thus the condition (2) of Proposition 4.4 is established. Note that it follows from Lemma 5.2 that

$$\lambda = \frac{\lambda^*}{d_r} = \frac{-1 + \delta^*}{n - L} < \frac{-1 + \delta^*}{n - M_{r-1}} < \dots < \frac{-1 + \delta^*}{n - M_1} < 0.$$

Thus the condition (3) of Proposition 4.4 is finished. The conditions (4) and (5) are trivial. Let us consider condition (6) of Proposition 4.4. It follows from Proposition 1.3, criterion (4) of Definition 5.1 and Proposition 4.6 that

$$\begin{aligned} \text{ord } T_r^\psi(\sigma) &= (-\mu_r)\delta_s + \sum_{i=r+1}^s V_i\left(\frac{-\mu_r}{d_i}\right)(\delta_{i-1} - \delta_i) \\ &\quad + \left[V_r\left(\frac{-\mu_r + M_r - n}{d_r}\right) + 1 \right] (\delta^* - \delta_r) \\ &= (-\mu_r + M_r - n) \frac{\lambda^*}{d_r} + \frac{n - M_r}{d_r} \left[d_r\delta_s + \sum_{i=r+1}^s V_i\left(\frac{d_r}{d_i}\right)(\delta_{i-1} - \delta_i) \right] + \delta^* - \delta_r. \end{aligned}$$

Using lemma 5.2 with $L = M_r$, we have $\delta(M_r) = \delta_r$ and

$$\frac{n - M_r}{d_r} \left[d_r\delta_s + \sum_{i=r+1}^s V_i\left(\frac{d_r}{d_i}\right)(\delta_{i-1} - \delta_i) \right] = -1 + \delta_r.$$

We thus deduce

$$\begin{aligned}\text{ord } T_r^\psi(\sigma) &= (-\mu_r + M_r - n) \frac{\lambda^*}{d_r} - 1 + \delta^* \\ &= (-\mu_r + M_r - n) \lambda - 1 + \delta^* \\ &= (\deg T_{r,\sigma}^\psi(\pi) + \varepsilon) \left(\frac{d_r}{V_r} \right) \lambda\end{aligned}$$

where

$$\varepsilon = \left(\frac{1 + \delta^*}{\lambda d_r} \right) V_r - 1 = \frac{-1 + \delta^*}{\lambda^*} V_r - 1 > -1.$$

We have verified all conditions of Proposition 4.4. The conclusion of it means

$$\text{ord}(\tau_i - \tau_j) > \delta_{r-1}$$

for all τ_i, τ_j roots of $g(y) \prod_{i=1}^s T_i^\psi(y)$ satisfying

$$\text{ord}(\tau_i - \tau) > \delta_r, \quad \text{ord}(\tau_j - \tau) > \delta_r.$$

It contradicts the very definition of δ_{r-1} . Our claim has thus been established.

Let δ^{**} denote $\delta(M_{r-1})$. Let σ^{**} be the unique π -root of $g(y) \prod_{i=1}^r T_i^\psi(y)$ in the δ^{**} -disc with

$$\text{ord}(\sigma^{**} - \tau) > \delta_r.$$

We shall verify the conditions of Proposition 4.6 (the splitting phenomena) with $\sigma = \sigma^{**}$, $\delta = \delta^{**}$, $V = V_r \left(\frac{d_{r-1}}{d_r} \right)$ and $r-1$ replacing r . Note that

$$\begin{aligned}\text{ord } g(\sigma^{**}) &= \frac{n}{d_{r-1}} \left[d_{r-1} \delta_s + \sum_{i=r+1}^s V_i \left(\frac{d_{r-1}}{d_i} \right) (\delta_{i-1} - \delta_i) + V_r \left(\frac{d_{r-1}}{d_r} \right) (\delta^{**} - \delta_r) \right] \\ &= \frac{n}{d_{r-1}} \lambda^{**} = n \left(\frac{\lambda^{**}}{d_{r-1}} \right) < 0\end{aligned}$$

and

$$\text{ord } T_i^\psi(\sigma^{**}) = (-\mu_i) \left(\frac{\lambda^{**}}{d_{r-1}} \right) \quad i = 1, \dots, r-1.$$

Conditions (1) and (2) are verified. Moreover it follows from Lemma 5.2 that

$$\frac{\lambda^{**}}{d_{s-1}} = \frac{-1 + \delta^{**}}{n - M_{s-1}} < \frac{-1 + \delta^{**}}{n - M_i} \quad i = 1, \dots, s-2.$$

We thus have condition (3). The conditions (4) and (5) are trivial. Note that the polynomials $q(\pi)$ and $p(\pi)$ in Proposition 4.6 are of degree

$$\deg q(\pi) = V_r \left(\frac{n - M_{r-1}}{d_r} \right) > 1, \quad \deg p(\pi) = V_r \left(\frac{d_{r-1}}{d_r} \right).$$

Moreover the polynomial $q(\pi)$ has distinct roots. It implies that at least two roots of $T_{r-1}^\psi(y)$ in the δ^{**} -disc have distance $2^{-\delta^{**}}$. Thus we conclude

$$\delta_{r-1} = \delta(M_{r-1}) = \delta^{**}.$$

The criteria (1), (3) and (4) follow for the extended tower. Q.E.D.

From the second criterion of Definition 5.1 we can find the smallest disc which contains all roots of $g(y) T_1^\psi(y)$. Namely (cf. [A.1], p. 14) "one point at infinite".

Proposition 5.4. *Suppose that $g(x, y)$ is monic in y with y -degree $n > 1$ and a tower of major discs $D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_1$ is constructed. Then the smallest disc which contains all roots of $g(y) T_1^\psi(y)$ is the D_i with*

$$i = \max \left\{ r : \frac{V_{r+1} d_r}{d_{r+1}} > V_r \right\}.$$

Moreover if $\delta_i > -1$ then either $k[x, y] = k[T_1^\psi, g] = k[f, g]$ or there exists an automorphism of $k[x, y]$ which reduces the degrees of $T_1^\psi(f(x, y), g(x, y))$, $g(x, y)$ and $f(x, y)$ simultaneously.

Proof. As proved in Proposition 5.3 the numbers $\frac{V_r d_{r-1}}{d_r}$ are the degrees of $p(\pi)$'s involved. Clearly $\frac{V_r d_{r-1}}{d_r} = V_{r-1}$ iff $p(\pi)$ has only one root which means the leading coefficients $g_\sigma(\pi)$ and $T_{1,\sigma}^\psi(\pi)$ have only one root. It is thus evident that D_i is the smallest disc which contains all roots of $g(y) T_1^\psi(y)$.

Suppose that we have

$$\delta_i > -1 \quad \text{and} \quad \frac{V_{j+1} d_j}{d_{j+1}} = V_j \quad \text{for} \quad j = i+1, \dots, s.$$

Then it follows from the definition of δ_i that

$$\delta_i = 1 - \frac{(n - M_i) \prod_{j=i+1}^s [V_j(n - M_j) - d_j]}{(n - M_s - 1) \prod_{j=i+1}^s [V_j(n - M_{j-1}) - d_j]} = 1 - \frac{n - M_i}{n - M_i - 1} = \frac{-1}{n - M_i - 1}.$$

There are two possibilities $i=1$ or $i>1$. In the first case we shall establish that $k[x, y] = k[T_1^\psi, g] = k[f, g]$. In the second case we shall show the existence of certain automorphisms of $k[x, y]$.

Let the number $n - M_i - 1$ be l . Suppose that $i=1$. Then Proposition 4.6 implies that the leading coefficients $g_\sigma(\pi)$ and $T_{1,\sigma}^\psi(\pi)$, which are polynomials in π , satisfy the following differential equation (cf. Definition 4.1)

$$(1) \quad D(n, -\mu_1, g_\sigma(\pi), T_{1,\sigma}^\psi(\pi)) = \text{nonzero constant}.$$

Note that $M_1 = \mu_1 = -\deg_y T_1(y)$ a negative integer. We thus have

$$l = n - M_1 - 1 \geq n.$$

It follows from Proposition A.5 of Appendix 1 that the polynomial $g_\sigma(\pi)$ has all distinct roots. Any root of $g(y)$ with nonzero coefficient of $t^{\frac{1}{l}}$ will produce a different root for $g(y)$ under the automorphism over $k((t))$ of the following form

$$t^{\frac{1}{l}} \rightarrow \omega t^{\frac{1}{l}}$$

where ω is a l -th primitive root of unity. Thus we must have

$$n = l = n - M_1 - 1 \quad \text{or} \quad M_1 = -1.$$

Hence T_1^ψ is a monic polynomial of degree 1 in y and an automorphism of $k[x, y]$ will send T_1^ψ to y . A simple computation on the Jacobian shows at once that under the same automorphism

$$g(x, y) \rightarrow x + \text{polynomial in } y.$$

Thus we have

$$k[x, y] = k[T_1^\psi, g] = k[f, g].$$

Let us consider the possibility of $i > 1$. Let σ_i be the unique π -root of $g(y) T_1^\psi(y)$ in the δ_i -disc D_i and

$$g(\sigma) = g_\sigma(\pi) t^{n\delta_i} + \dots, \quad T_1^\psi(\sigma) = T_{1,\sigma}^\psi(\pi)^{-\mu_1\delta_i} + \dots.$$

Then Propositions 5.3 and 4.6 show that $g_\sigma(\pi)$ and $T_{1,\sigma}^\psi(\pi)$ are powers of a common polynomial $p(\pi)$. Our definition of δ_i implies that $p(\pi)$ has more than one root. Let $\pi - a$ be a factor of $p(\pi)$ with $a \neq 0$. Let the unique π -root σ_i be

$$\sigma_i = \sum \alpha_j t^j + \pi t^{\delta_i}.$$

Since a noninteger exponent j in the above expression will produce different roots for $g(y) T_1^\psi(y)$ as shown in the previous case, then all exponents must be integers. In other words for some polynomial $h(x)$ we have

$$\sigma = h(x) + \pi t^{\delta_i}.$$

The following automorphism

$$x \rightarrow x, \quad y \rightarrow y - h(x)$$

will reduce σ to the following simple form

$$\sigma \rightarrow \pi t^{\delta_i}$$

and without changes in the y -degrees of $g(y)$ and $T_i^\psi(y)$. We shall simply assume from the beginning, that the π -root σ is of the form

$$\sigma = \pi t^{\delta_i}.$$

It follows from

$$\text{ord } g(t^{-1}, \pi t^{\delta_i}) = n\delta_i = -\frac{n}{l}$$

that we have

$$\text{ord } g(t^{-l}, \pi t^{-1}) = -n \quad \text{or} \quad \deg g(x^l, \pi x) = n.$$

Similarly we have for the nonzero a defined above

$$\text{ord } g(t^{-l}, at^{-1}) > -n \quad \text{or} \quad \deg g(x^l, ax) < n.$$

This way we have established that the highest weighted form of $g(x, y)$ with x of weight l and y of weight 1 has a factor of the form

$$a^l x - y^l.$$

An automorphism of $k[x, y]$ defined by

$$x \rightarrow x + \frac{1}{a^l} y^l, \quad y \rightarrow y$$

will save the weight of every term while sending $a^l x - y^l$ to $a^l x$. Thus under the automorphism the polynomial $g(x, y)$ will have x as a factor for the highest weighted form of weight n . This proves at once the (unweighted) degree of $g(x, y)$ must drop at least 1. Similar arguments work for T_1^ψ . As for $f(x, y)$ we note that (cf. Proposition 2.2)

$$T_1^\psi = f + \text{polynomial in } g$$

and

$$\deg T_1^\psi(f(x, y), g(x, y)) = \deg_y T_1^\psi = -M_1 \notin (n).$$

Thus if both degrees of $g(x, y)$ and T_1^ψ drops, then the degree of $f(x, y)$ must drop too. Q.E.D.

The following lemma will specify the implications of Proposition 5.4 in terms of the forms of the polynomial $g(x, y)$ (cf. [A.1], p. 139).

Lemma 5.3. *Suppose that $g(x, y)$ is monic in y with $\deg g(x, y) = \deg_y g(x, y) = n > 1$. Then the smallest disc which contains all roots of $g(y) T_1^\psi(y)$ is of logarithmic radius -1 iff $M_s = n - 2$ and the highest homogeneous form of $g(x, y)$ has two roots with one root having a multiplicity $\frac{n}{d_s} v_s$ where v_s satisfies*

$$d_s > v_s > \frac{d_s}{2}.$$

Proof. Necessary. Let $D_s \supset \dots \supset D_1$ be any tower of major discs with logarithmic radii $\delta_s < \dots < \delta_1$. The proof of Proposition 5.4 shows for some i the following

$$-1 = \delta_i = 1 - \frac{n - M_i}{n - M_i - 1} = \frac{-1}{n - M_i - 1}.$$

We conclude that

$$n - M_i - 1 = 1, \quad n - 2 \geq M_s \geq M_i = n - 2.$$

Thus

$$M_s = n - 2.$$

It follows from Proposition 5.4 and 5.3

$$\frac{V_{s+1}d_s}{d_{s+1}} > V_s > \frac{d_s}{n - M_s} = \frac{d_s}{2}.$$

Recall

$$V_{s+1} = d_{s+1}.$$

We thus have

$$d_s > V_s > \frac{d_s}{2}.$$

The criteria (4) of Definition 5.1 and Proposition 4.6 specify the degree of the polynomial $q(\pi)$ (cf. Proposition 4.6) is $n - M_s = 2$. Then the polynomial $p(\pi)$ has two roots with one of the multiplicities being V_s . It is routine to deduce the stated properties of the highest homogeneous form of $g(x, y)$ (cf. Propositions 4.3 and 4.5).

Sufficient. It follows from Proposition 5.3 that there is a tower of major discs $D_s \supset D_{s-1} \supset \cdots \supset D_1$ with the corresponding V_s given by this lemma. Thus the index i as specified in Proposition 5.4 is

$$i = \max \left\{ r : \frac{V_{r+1}d_r}{d_{r+1}} > V_r \right\} = s$$

and the corresponding logarithmic radius δ_i is

$$\delta_s = \frac{-1}{n - M_s - 1} = -1. \quad \text{Q.E.D.}$$

The following proposition shows that there is a reduction for the two characteristic pairs case.

Proposition 5.5. *Suppose that $g(x, y)$ is monic in y with y -degree $n > 2$. If there are only two effective characteristic pairs i.e. $s = 2$, then either $k[x, y] = k[T_1^\psi, g] = k[f, g]$ or there exists an automorphism of $k[x, y]$ which reduces the degrees of $T_1^\psi(f(x, y), g(x, y))$, $g(x, y)$ and $f(x, y)$ simultaneously.*

Proof. It follows from Proposition 5.4 that we may assume that the smallest disc D_i which contains all roots of $g(y) T_1^\psi(y)$ is of logarithmic radius -1 i.e. for $i = 1$ or 2 the following

$$\delta_i = -1 = \frac{-1}{n - M_i - 1}.$$

We shall deduce a contradiction. It follows from $M_1 = \mu_1 = -\deg_y T_1^\psi(y)$ a negative integer that we must have $i=2$ and

$$M_2 = n - 2, \quad \delta_2 = -1.$$

Hence we have a tower of major discs $D_2 \supset D_1$ (cf. Proposition 5.3) with the corresponding numbers V_2, V_3 satisfying

$$\frac{V_3 d_2}{d_3} > V_2 > \frac{d_2}{n - M_2} \quad \text{or} \quad d_2 > V_2 > \frac{d_2}{n - M_2} = \frac{d_2}{2}.$$

Then it follows from Proposition 5.3 that the logarithmic radius δ_1 of D_1 is

$$\delta_1 = 1 - \frac{(n - M_1) [2V_2 - d_2]}{[V_2(n - M_1) - d_2]} = \frac{(n - M_1)(d_2 - V_2) - d_2}{(n - M_1)V_2 - d_2}.$$

Recall

$$d_2 = \text{g.c.d.}(n, M_1).$$

Let us write

$$n = d_2 n^*, \quad -M_1 = d_2 m^*, \quad U_2 = d_2 - V_2 < V_2$$

and write δ_1 in its reduced form as

$$\delta_1 = \frac{(n^* + m^*)U_2 - 1}{(n^* + m^*)V_2 - 1} = \frac{B}{A}.$$

Clearly we have

$$0 < \delta_1 < 1$$

i.e. the rational number δ_1 is a true fraction and $A > 1$. Let σ be the π -root in the δ_1 -disc D_1 and

$$g(\sigma) = g_\sigma(\pi) t^{n\lambda} + \dots, \quad T_1^\psi(\sigma) = T_{1,\sigma}^\psi(\pi) t^{-M_1\lambda} + \dots.$$

Then it follows from Proposition 5.3 and Proposition 4.6 that $g_\sigma(\pi)$ and $T_{1,\sigma}^\psi(\pi)$ satisfy the following differential equation

$$D(n, -M_1, g_\sigma(\pi), T_{1,\sigma}^\psi(\pi)) = \text{nonzero constant}.$$

Proposition A.5 implies that

$$(g_\sigma(\pi), T_{1,\sigma}^\psi(\pi)) = 1$$

and

$$\left(\frac{d}{d\pi} g_\sigma(\pi), \frac{d}{d\pi} T_{1,\sigma}^\psi(\pi) \right) = 1.$$

The conjugations of $k((t^{\frac{1}{A}}))$ over $k((t))$ show us that if the reduced denominator A of δ_1 is not a factor of $\deg g_\sigma(\pi) = n^* V_2$ then π is a factor of $g_\sigma(\pi)$. Moreover if A is a factor of $n^* V_2$ then $g_\sigma(\pi) \in k[\pi^A]$ and π is a factor of $\frac{d}{d\pi} g_\sigma(\pi)$. In view of the two co-prime conditions above we must have

$$A|n^* V_2 \text{ and } A \nmid m^* V_2 \text{ or } A \nmid n^* V_2 \text{ and } A|m^* V_2.$$

Moreover we always have from the very definition of A the following

$$A|(n^* + m^*) V_2 - 1.$$

Let us consider the possibility of

$$(1) \quad A|n^* V_2 \text{ and } A|m^* V_2 - 1.$$

The other possibility can be treated verbatim. Let

$$n^* V_2 = AD, \quad m^* V_2 - 1 = AE.$$

The second equality implies

$$(A, V_2) = 1.$$

Thus we must have from the first equality

$$V_2 | D.$$

Let us consider the following equation

$$1 - \delta_1 = \frac{(n^* + m^*) (V_2 - U_2)}{(n^* + m^*) V_2 - 1} = (n^* + m^*) \frac{C}{A}$$

where

$$(n^* + m^*) V_2 - 1 = A(D + E), \quad V_2 - U_2 = C(D + E).$$

Then we deduce the following

$$V_2 > V_2 - U_2 = C(D + E) \geq D \geq V_2.$$

A contradiction.

Q.E.D.

The following proposition will take Proposition 5.5 as a special case and gives a different proof of the reduction of the two characteristic pair case.

Proposition 5.6. *Suppose that $s \geq 2$ and that a tower of major discs*

$$D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_1$$

has been constructed. If the corresponding π -root σ_1 , which is the unique one in D_1 , is of the following form

$$\sigma_1 = \pi t^{\delta_1}$$

then either $k[x, y] = k[T_1^\psi, g] = k[f, g]$ or there exists an automorphism $k[x, y] \rightarrow k[x, y]$ which will reduce the degrees of $f(x, y)$, $g(x, y)$ and $T_1^\psi(f(x, y), g(x, y))$ simultaneously.

Proof. It follows from Proposition 5.4 and Lemma 5.3 that we shall only consider the case $\delta_s = -1$, $M_s = n-2$ and $d_s > v_s$. It suffices to produce a numerical contradiction for this case. Note the following

$$\begin{aligned}\delta_{s-1} &= 1 - \frac{(n - M_{s-1}) [V_s(n - M_s) - d_s]}{(n - M_s - 1) [V_s(n - M_{s-1}) - d_s]} \\ &= 1 - \frac{(n - M_{s-1}) [2V_s - d_s]}{V_s(n - M_{s-1}) - d_s} = \frac{(d_s - V_s)(n - M_{s-1}) - d_s}{V_s(n - M_{s-1}) - d_s}.\end{aligned}$$

Recall that

$$\begin{aligned}V_s(n - M_{s-1}) &> V_s(n - M_s) > d_s, \\ d_s &> V_s, \quad d_s | (n - M_{s-1}).\end{aligned}$$

Then we conclude

$$\delta_1 \geq \delta_{s-1} \geq 0.$$

We shall establish $\delta_1 < 0$ to produce a contradiction.

First we shall point out a trivial observation about the Jacobian problem. Namely the polynomials $g(x, y)$ and $T_1^\psi(f(x, y), g(x, y))$ have no common polynomial factor. Otherwise let it be $h(x, y)$. A simple computation will show at once that

$$J_{x,y}(T_1^\psi, g) \in (h(x, y)).$$

A contradiction. Especially y is not a factor of either $T_1^\psi + c_1$ or $g + c_2$, say $g + c_2$, for all possible constants c_1 and c_2 . We conclude at once that there must be a term of the form x^l in $g(x, y)$ for $l \geq 1$. Then we have

$$\text{ord } g(\sigma_1) = n\lambda \leq -l.$$

Note that the orders of the two sides of the equation (1) of Proposition 4.1 are equal for $\sigma = \sigma_1$ and $H = T_1^\psi$ (cf. the criteria of Definition 5.1 and Proposition 4.6) and

$$\text{ord } T_1^\psi(\sigma_1) = (-M_1)\lambda < 0.$$

Then it follows from $\delta_1 \geq 0$ that

$$n\lambda + (-M_1)\lambda - 1 = -2 + \delta_1 \geq -2.$$

We must have

$$-l - 1 \geq -2$$

which implies

$$l = 1, \quad n\lambda \leq -1$$

and

$$\delta_1 = n\lambda + (-M_1)\lambda - 1 + 2 \leq (-M_1)\lambda < 0.$$

A contradiction.

Q.E.D.

To deduce Proposition 5.5 from Proposition 5.6 we note that we only have to consider (cf. the first part of the proof of Proposition 5.5)

$$\delta_2 = -1, \quad 0 < \delta_1 < 1.$$

Thus we have for suitable constants a and b the following

$$\sigma_1 = at^{-1} + b + \pi t^{\delta_1}.$$

And an automorphism of the following form

$$x \rightarrow x, \quad y \rightarrow y - ax - b$$

will change σ_1 to $\pi t^{-\delta_1}$. Proposition 5.5 follows from Proposition 5.6.

§ 6. The minor discs

In the last section we have discussed the construction of the towers of major discs (cf. Definition 5.1 and Proposition 5.3). Note that the minimal disc D_s which contains all roots of $g(y) \prod_{i=1}^h T_i^\psi(y)$ is a tower of one major disc (cf. Proposition 5.1). Suppose that we have constructed a tower of major discs $D_s \supseteq D_{s-1} \supseteq \cdots \supseteq D_r$. The criterion (4) of Definition 5.1 implies the conditions of Proposition 4.6 are fulfilled. The conclusions of Proposition 4.6 thus implies the existence of a certain polynomial $p(\pi)$. Let $(\pi - C_r)$ be a factor of the polynomial $p(\pi)$ with multiplicity V_r . If the multiplicity V_r satisfies

$$\deg p(\pi) = V_{r+1} \frac{d_r}{d_{r+1}} \geq V_r > \frac{d_r}{n - M_r}$$

then we may extend the tower of major discs $D_s \supseteq \cdots \supseteq D_r \supseteq D_{r-1}$ according to the construction of D_{r-1} as described in Proposition 5.3. On the other hand the multiplicity V_r may satisfy

$$\frac{d_n}{n - M_r} \geq V_r.$$

We shall construct the corresponding disc D_{r-1}^* thus denoted a *minor disc* with the following characterizations.

Proposition 6.1. *Suppose that $g(x, y)$ is monic in y with y -degree $n > 1$ and for $r \geq 2$ a tower of major discs $D_s \supseteq \cdots \supseteq D_r$ is constructed. Let $\pi - c_r$ be a factor of $p(\pi)$ as in the conclusions of Proposition 4.6 with multiplicity V_r satisfying*

$$\frac{d_r}{n - M_r} \geq V_r \geq 1.$$

Let τ and δ_{r-1}^* be defined as

$$\begin{aligned} \tau &= \sum a_j t^j + c_r t^r \in D_r, \\ \delta_{r-1}^* &= \min \left\{ \text{ord}(\tau_i - \tau_j) : \tau_i, \tau_j \text{ roots of } g(y) \prod_{i=1}^r T_i^\psi(y) \right. \\ &\quad \left. \text{with } \text{ord}(\tau_i - \tau) > \delta_r \text{ and } \text{ord}(\tau_j - \tau) > \delta_r \right\} \\ &= \text{the logarithmic radius of } D_{r-1}^* \text{ which is the minimal disc containing all roots } \tau_i \\ &\quad \text{of} \\ &\quad g(y) \prod_{i=1}^r T_i^\psi(y) \text{ with } \text{ord}(\tau_i - \tau) > \delta_r. \end{aligned}$$

Then we have $\delta_{r-1}^* \geq 1$. Moreover let $\sigma = \sum b_j t^j + \pi t^\delta$ be any π -root of $g(y) \prod_{i=1}^{r-1} T_i^\psi(y)$ with

$$\text{ord}(\sigma - \tau) > \delta_r.$$

Then we have

- (1) the inequality $\delta < 1$ implies $\text{ord } g(\sigma) < 0$;
- (2) the inequality $\text{ord } g(\sigma) < 0$ implies that σ is a distribution detector for $g(y)$, $T_1^\psi(y), \dots, T_{r-1}^\psi(y)$.

Proof. Suppose for a moment that $\delta_{r-1}^* < 1$. We shall deduce a contradiction as follows. Let σ^* be the unique π -root of $g(y) \prod_{i=1}^r T_i^\psi(y)$ in the disc D_{r-1}^* . Then it follows from Definition 5.1 and Proposition 1.3 that

$$\begin{aligned} \text{ord } g(\sigma^*) &= n\delta_s + \sum_{i=r+1}^s V_i \left(\frac{n}{d_i} \right) (\delta_{i-1} - \delta_i) + V_r \left(\frac{n}{d_r} \right) (\delta_{r-1}^* - \delta_r) \\ &= \frac{n}{d_r} \left[d_r \delta_s + \sum_{i=r+1}^s V_i \left(\frac{d_r}{d_i} \right) (\delta_{i-1} - \delta_i) \right] + V_r \frac{n}{d_r} (\delta_{r-1}^* - \delta_r) \\ &= \frac{n}{d_r} \lambda(\delta_r) + V_r \frac{n}{d_r} (\delta_{r-1}^* - \delta_r) \\ &= \frac{n}{n - M_r} (-1 + \delta_r) + V_r \frac{n}{d_r} (\delta_{r-1}^* - \delta_r). \end{aligned}$$

For the definition of $\lambda(L)$ (or the special case of $L = \delta_r$) and the last equality the reader is referred to Lemma 5.2. Note that we have

$$\frac{d_r}{n - M_r} \geq V_r, \quad 1 > \delta_{r-1}^* > \delta_r.$$

Thus we conclude trivially the following

$$(1) \quad \text{ord } g(\sigma^*) = n \left[\frac{1}{n - M_r} (-1 + \delta_r) + \frac{V_r}{d_r} (\delta_{r-1}^* - \delta_r) \right] = n\lambda^* < 0$$

with λ^* defined by the above equation. Similarly we conclude

$$(2) \quad \text{ord } T_i^\psi(\sigma^*) = (-\mu_i)\lambda^* \quad \text{for } i = 1, \dots, r-1.$$

We have thus verified the conditions (1) and (2) of Proposition 4.4. We shall establish the condition (3) of Proposition 4.4

$$\lambda^* < \frac{-1 + \delta_{r-1}^*}{n - M_i} \quad i = 1, \dots, r-1,$$

or equivalently the following inequalities

$$\begin{aligned} \frac{1}{n - M_r} (-1 + \delta_r) + \frac{V_r}{d_r} (\delta_{r-1}^* - \delta_r) &\leq \frac{-1 + \delta_{r-1}^*}{n - M_r} < \frac{-1 + \delta_{r-1}^*}{n - M_i} \\ &\Leftrightarrow (-1 + \delta_r) + V_r \frac{n - M_r}{d_r} (\delta_{r-1}^* - \delta_r) \leq -1 + \delta_{r-1}^* \\ &\Leftrightarrow V_r (\delta_{r-1}^* - \delta_r) \leq \frac{d_r}{n - M_r} (\delta_{r-1}^* - \delta_r) \\ &\Leftrightarrow V_r \leq \frac{d_r}{n - M_r}. \end{aligned}$$

For condition (4) and (5) (cf. Definition 5.1 and Proposition 4.6) we simply let $v = V_r$. We shall finally verify conditions (6) and (7) of Proposition 4.4. Note that

$$\begin{aligned} \text{ord } T_r^\psi(\sigma^*) &= (-\mu_r)\delta_s + \sum_{i=r+1}^s V_i \left(\frac{-\mu_r}{d_i} \right) (\delta_{i-1} - \delta_i) + \left[V_r \left(\frac{-\mu_r + M_r - n}{d_r} \right) + 1 \right] (\delta_{r-1}^* - \delta_r) \\ &= (-\mu_r) \left[\frac{1}{n - M_r} (-1 + \delta_r) \right] + \left[V_r \left(\frac{-\mu_r + M_r - n}{d_r} \right) + 1 \right] (\delta_{r-1}^* - \delta_r) \\ &= (-\mu_r + M_r - n) \left[\frac{1}{n - M_r} (-1 + \delta_r) + \frac{V_r}{d_r} (\delta_{r-1}^* - \delta_r) \right] + (-1 + \delta_r) + (\delta_{r-1}^* - \delta_r) \\ &= (-\mu_r + M_r - n)\lambda^* - 1 + \delta_{r-1}^* = \left\{ \left[V_r \left(\frac{-\mu_r + M_r - n}{d_r} \right) + 1 \right] + \varepsilon \right\} \frac{d_r}{V_r} \lambda^* \end{aligned}$$

where it follows from Definition 5.1 and Proposition 4.6

$$\deg T_{r,\sigma^*}^\psi(\pi) = V_r \left(\frac{-\mu_r + M_r - n}{d_r} \right) + 1, \quad (1 + \varepsilon) \frac{d_r}{V_r} \lambda^* = -1 + \delta_{r-1}^*$$

which is equivalent to

$$\varepsilon = (-1 + \delta_{r-1}^*) \frac{V_r}{d_r \lambda^*} - 1.$$

Condition (7) follows from

$$\begin{aligned}\varepsilon > -1 &\Leftrightarrow (-1 + \delta_{r-1}^*) \frac{V_r}{d_r \lambda'^*} > 0 \\ &\Leftrightarrow -1 + \delta_{r-1}^* < 0 \Leftrightarrow \delta_{r-1}^* < 1.\end{aligned}$$

Hence conditions (6) and (7) are satisfied. The conclusions of Proposition 4.4 imply that the leading coefficient of $g(\sigma^*) \prod_{i=1}^r T_i^\psi(\sigma^*)$ has only one root. It thus contradicts the very definition of δ_{r-1}^* (cf. Proposition 1.2). We have established $\delta_{r-1}^* \geq 1$.

The second part of our proposition, i.e. the inequality $\delta < 1$ implies $\text{ord } g(\sigma) < 0$, follows almost verbatim from the second paragraph of this proof with σ^* replaced by σ and δ_{r-1}^* replaced by δ . The equations (1) and (2) become

$$(3) \quad \text{ord } g(\sigma) = n \left[\frac{1}{n - M_r} (-1 + \delta_r) + \frac{V_r}{d_r} (\delta - \delta_r) \right] = n\lambda < 0$$

with λ defined by the above equation and

$$(4) \quad \text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda \quad \text{for } i = 1, \dots, r-1.$$

For the last part of our proposition if $\delta < 1$ then the first part of our proposition shows that the leading coefficients of $g(\sigma), T_1^\psi(\sigma), \dots, T_{r-1}^\psi(\sigma)$ are powers of a common linear polynomial in π of degree $V_r \frac{n}{d_r}, V_r \left(\frac{-\mu_1}{d_r} \right), \dots, V_r \left(\frac{-\mu_{r-1}}{d_r} \right)$ respectively. The second part of our proposition, i.e. the equation (3) and (4), shows

$$\text{ord } g(\sigma), \text{ord } T_1^\psi(\sigma), \dots, \text{ord } T_{r-1}^\psi(\sigma)$$

are proportional to $n, -\mu_1, \dots, -\mu_{r-1}$. Thus it follows from Definition 3.1 that σ is a distribution detector for $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$. We are left with the case $\delta \geq 1$. We shall apply Proposition 4.2 inductively. Note that it suffices to check

$$(5) \quad \text{ord } g(\sigma) = n\lambda < 0,$$

$$(6) \quad \text{ord } T_i^\psi(\sigma) = (-\mu_i)\lambda \quad i = 1, \dots, r-1,$$

$$(7) \quad \lambda < \frac{-1 + \delta}{n - M_i} \quad i = 1, \dots, r-1$$

with equation (5) as the definition of λ .

Suppose that there is a σ which is not a distribution detector. Then we may choose one with the corresponding δ minimal. Proposition 1.3 implies at once that equation (6) is satisfied. Equation (7) is trivial due to the fact $\delta \geq 1$. Thus we conclude that σ is a distribution detector. A contradiction. Q.E.D.

The above proposition does not specify the logarithmic radius δ_{r-1}^* of D_{r-1}^* . We only get an estimate $\delta_{r-1}^* \geq 1$. The significance of the minor disc D_{r-1}^* is that the roots of $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$ in it are all distributed proportionally according to the y -degrees up to the uncertainty of $2^0 = 1$. In other words if our estimate of the roots

is blurred as the order of $g, T_1^\psi, \dots, T_{r-1}^\psi$ simultaneously reach 0, then all roots in the disc D_{r-1}^* of $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$ are indeed proportional to the y -degrees of $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$. This uncertainty is intrinsic due to two different reasons. First we may change $g(y)$ to $g(y) + c$. Let σ be a π -root of $g(y)$ such that $\text{ord } g(\sigma) = 0$. Then the leading coefficient of $g(\sigma) + c$ is with a variable constant c . There is no way to control the roots of it. Second, in all our arguments we may replace the approximate roots $T_i^\psi(y)$ by the corresponding quasi-approximate roots (cf. section 1). As the approximate roots $T_1^\psi(y), \dots, T_{r-1}^\psi(y)$ simultaneously become units with y substituted by a π -root σ , the corresponding quasi-approximate roots will become units with different leading coefficients. This fact thus introduces some uncertainties in the conclusions.

We shall apply Proposition 6.1 to deduce some information about the Jacobian conjecture. As pointed out in Proposition 5.4 and Lemma 5.3 we shall only consider the case $M_s = n - 2$ and the highest homogeneous form $g(x, y)$ equals $[(y - ax)^{v_s} (y - bx)^{u_s}]^{\frac{n}{d_s}}$ for $a \neq b$ where $u_s = d_s - v_s$. Note that Proposition 6.1 implies that the highest homogeneous forms in x and y of $T_1^\psi, \dots, T_s^\psi$ are determined up to constant multiples. The major disc D_s is of logarithmic radius -1 . It contains a major disc D_{s-1} and a minor disc D_{s-1}^* (cf. Propositions 5.3 and 6.1). The logarithmic radius δ_{s-1} of D_{s-1} is the following (cf. Definition 5.1)

$$\begin{aligned} \delta_{s-1} &= 1 - \frac{(n - M_{s-1}) [v_s(n - M_s) - d_s]}{(n - M_s - 1) [v_s(n - M_{s-1}) - d_s]} \\ &= 1 - \frac{(n - M_{s-1}) (v_s - u_s)}{v_s(n - M_{s-1}) - d_s} = \frac{u_s(n - M_{s-1}) - d_s}{v_s(n - M_{s-1}) - d_s}. \end{aligned}$$

Recall that

$$d_s = \text{g.c.d.}(n, M_1, \dots, M_{s-1}).$$

We have

$$\delta_{s-1} \geq 0.$$

Lemma 6.1. *Suppose that $g(x, y)$ is monic in y with $\deg g(x, y) = \deg_y g(x, y) = n > 1$. Let $D_s \supseteq D_{s-1}$ be a tower of major discs with logarithmic radius δ_s and δ_{s-1} . If $\delta_s = -1$ then $\delta_{s-1} \geq 0$.*

Proof. Preceding discussions.

Q.E.D.

It follows from Lemma 6.1 that all roots of $g(y)$ and $T_i^\psi(y)$ for $i = 1, \dots, s$ in the major disc D_{s-1} are of the form

$$(8) \quad y = at^{-1} + c_j + \text{higher terms in } t$$

where c_j 's are constants. It follows from Proposition 6.1 that the logarithmic radius of the minor disc D_{s-1}^* is greater than or equal to 1. Thus all roots of $g(y)$ and $T_i^\psi(y)$ for $i = 1, \dots, s$ in the minor disc D_{s-1}^* are of the form

$$(9) \quad y = bt^{-1} + e + c_j^*t + \text{higher terms in } t$$

where e is common for all roots. We may choose y properly such that in equation (8) $a = 0$. Let

$$(10) \quad z = y - bx - e.$$

Then we may trade the monic property with a reduction of y -deg (or z -deg) for the polynomials involved as follows.

Proposition 6. 2. *Suppose that $g(x, y)$ is monic in y with*

$$\deg g(x, y) = \deg_y g(x, y) = n > 1, \quad \delta_s = -1.$$

Furthermore suppose that in the equation (8) $a=0$. Let z be defined as in equation (10). Let $g(x, y)$, $T_i^\psi(f(x, y), g(x, y))$ for $i=1, \dots, s-1$ be expressed as polynomials $\bar{g}$, $\bar{T}_i^\psi$ in y and z . Then we have

- (1) $\deg_y \bar{g} = v_s \frac{n}{d_s},$
- (2) $\deg_z \bar{g} = u_s \frac{n}{d_s},$
- (3) $\deg_y \bar{T}_i^\psi = v_s \left(\frac{-\mu_i}{d_s} \right) \quad \text{for } i=1, \dots, s-1,$
- (4) $\deg_z \bar{T}_i^\psi = u_s \left(\frac{-\mu_i}{d_s} \right) \quad \text{for } i=1, \dots, s-1.$

Proof. The proofs of the four equations are similar. We shall only prove the first one. Let the factorization of $g(y)$ in $k\langle\langle t \rangle\rangle$ be

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

where

$$\tau_i = \begin{cases} bt^{-1} + e + c_i^* t + \text{higher terms in } t & \text{for } i=1, \dots, u_s \frac{n}{d_s}, \\ c_i + \text{higher terms in } t & \text{for } i=u_s \frac{n}{d_s} + 1, \dots, n. \end{cases}$$

Let

$$\sigma = bt^{-1} + e + \pi t.$$

Then we have

$$g(\sigma) = g_\sigma(\sigma) t^{(u_s - v_s) \frac{n}{d_s}} + \text{higher terms in } t.$$

Let

$$g = \bar{g} = \sum a_{ij} y^i z^j.$$

Then

$$(11) \quad \bar{g}(\sigma) = \sum a_{ij} (bt^{-1} + e + \pi t)^i (\pi t)^j = \sum a_{ij} (b^i \pi^j t^{j-i} + \dots).$$

If we have in the above expression

$$i > v_s \frac{n}{d_s}$$

for some i , then we may take the largest one. In addition

$$i + j \leq n.$$

Then we have

$$j - i < (u_s - v_s) \frac{n}{d_s}.$$

Note that the coefficient π^j in equation (11) guarantees that the terms with different j indices will not cancel out. Hence we conclude

$$\text{ord } g(\sigma) < (u_s - v_s) \frac{n}{d_s}.$$

A contradiction.

Q.E.D.

The following lemma will be used in Proposition 6.3.

Lemma 6.2 (Inversion lemma). *Let x and y be two elements in $k\langle\langle t \rangle\rangle$ with*

$$x = t^{-1}, \quad y = y(t) = \sum a_j t^j = \theta^{-1}.$$

Suppose that

$$\text{ord}_t y(t) = -1 < 0.$$

Then we have $k\langle\langle t \rangle\rangle = k\langle\langle \theta \rangle\rangle$. Furthermore in the following expression

$$x = x(\theta)$$

we have for arbitrary rational q and positive integer p :

(1) $x(\theta) \bmod \theta^q$ and $y(t) \bmod t^q$ uniquely determine each other,

(2) $x(\theta) \bmod \theta^q \in k((\theta^{\frac{1}{p}})) \bmod \theta^q \Leftrightarrow y(t) \bmod t^q \in k((t^{\frac{1}{p}})) \bmod t^q$.

Proof. Since we have

$$\text{ord}_t \theta = 1 > 0$$

then it follows from the definition of Puiseux fields that $k\langle\langle t \rangle\rangle = k\langle\langle \theta \rangle\rangle$. Suppose that

$$y(t) \in k((t^{\frac{1}{m}}))$$

for some positive integer m . Then let

$$T = t^{\frac{1}{m}}, \quad X = T^{-m}, \quad Y = bT^{-m} + \dots.$$

Then the conclusions (1) and (2) follow from the classical inversion formula for plane meromorphic curves [M.3]. Q.E.D.

The following Proposition shows that in some cases the degrees of the polynomials involved can be reduced by some transformations while the Jacobian will be modified to a monomial.

Proposition 6.3. *Suppose that $g(x, y)$ is monic in y with*

$$\deg g(x, y) = \deg_y g(x, y) = n > 1,$$

$\delta_s = -1$ and the logarithmic radius of δ_{s-1}^* of the minor disc $D_{s-1}^* \geq \frac{v_s}{u_s}$. Let z be defined by equation (10), $y^{-1} = \theta$ and consider $g, T_1^\psi, \dots, T_{s-1}^\psi$ as elements $\bar{g}(z), \bar{T}_1^\psi(z), \dots, \bar{T}_{s-1}^\psi(z)$ in $k\langle\theta\rangle[z]$. Then there is a π -root σ of the following form

$$\sigma = \sum a_j \theta^j + \pi \theta^{\frac{v_s}{u_s}}$$

such that with

$$\gamma = \theta^{\frac{1}{u_s}}$$

we have

- (1) $\bar{g}(\sigma), \bar{T}_1^\psi(\sigma), \dots, \bar{T}_{s-1}^\psi(\sigma) \in k[\gamma, \pi]$,
- (2) $\bar{g}(\sigma), \bar{T}_1^\psi(\sigma), \dots, \bar{T}_{s-1}^\psi(\sigma)$ are monic in π with π -degree

$$u_s \frac{n}{d_s}, u_s \frac{-\mu_1}{d_s}, \dots, u_s \frac{-\mu_{s-1}}{d_s}$$

respectively,

$$(3) \quad J_{\gamma, \pi}(\bar{g}(\sigma), T_1^{-\psi}(\sigma)) = -\frac{u_s}{b} \gamma^{v_s - u_s - 1}.$$

Proof. It follows from Proposition 6.2 that

$$\begin{aligned} \bar{g}(y, z) &= (y^{\frac{v_s n}{d_s}} + \dots) z^{\frac{u_s n}{d_s}} + \text{lower terms in } z \\ &= (y^{\frac{v_s n}{d_s}} + \dots) \prod_{i=1}^{\frac{u_s n}{d_s}} (z - \tau_i) \end{aligned}$$

where $\tau_i \in k\langle\theta\rangle$. Let us consider all parametrizations of $g(x, y)$ in the following form

$$\begin{aligned} x &= t^{-1}, \\ y &= b t^{-1} + e + \prod_{i=1}^{< \frac{v_s}{u_s}} e_i t^i + \text{higher terms in } t. \end{aligned}$$

By the assumption on the logarithmic radius δ_{s-1}^* of the minor disc D_{s-1}^* we have all e_i 's are common. Hence it follows from Lemma 6.2 that in the following inversion

$$x = \frac{1}{b} \left[\theta^{-1} + \sum_{i=1}^{< \frac{v_s}{u_s}} c_i \theta^i + \text{higher terms in } \theta \right], \quad y = \theta^{-1}$$

all c_i 's are common. With z defined as

$$z = y - bx - e$$

we conclude that in all parametrizations of $\bar{g}(y, z)$ of the following form

$$y = \theta^{-1}, \quad z = (c_0 - e) - \sum_{i=1}^{< \frac{v_s}{u_s}} c_i \theta^i + \text{higher terms in } \theta$$

all c_i 's are common. Thus if we let σ be defined as

$$\sigma = (c_0 - e) + \sum_{i=0}^{< \frac{v_s}{u_s}} c_i \theta^i + \pi \theta^{\frac{v_s}{u_s}}$$

then we have

$$\begin{aligned} \bar{g}(\sigma) &= \left(\theta^{-\frac{v_s n}{d_s}} + \dots \right) \left[\bar{g}_\sigma(\pi) \theta^{\frac{v_s n}{d_s}} + \dots \right] \\ &= \bar{g}_\sigma(\pi) + \text{terms of higher degree in } \theta \text{ and lower degree in } \pi \end{aligned}$$

where

$$\deg \bar{g}_\sigma(\pi) = \frac{u_s n}{d_s}.$$

The conclusions (1) and (2) for $\bar{g}(\sigma)$ are evident. Similar arguments prove the conclusions (1) and (2) for $\bar{T}_i^\psi(\sigma)$ for $i = 1, \dots, s-1$. To prove conclusion (3) we note that

$$\begin{aligned} J_{\gamma, \pi}(\bar{g}(\sigma), \bar{T}_1^\psi(\sigma)) &= J_{\gamma, \pi}(x, y) J_{x, y}(g(x, y), T_1^\psi(f(x, y), g(x, y))) \\ &= J_{\gamma, \pi}(x, y) \\ &= \det \begin{bmatrix} * & -\frac{1}{b\gamma^{v_s}} \\ -u_s \gamma^{-u_s-1} & 0 \end{bmatrix} \\ &= -\frac{u_s}{b\gamma^{v_s-u_s-1}}. \end{aligned} \quad \text{Q.E.D.}$$

Proposition 6.4. Suppose that $g(x, y)$ is monic in y with

$$\deg g(x, y) = \deg_y g(x, y) = n > 1,$$

$\delta_s = -1$ and $u_s = 1$. Then the logarithmic radius δ_{s-1}^* of the minor disc $D_{s-1}^* \geq \frac{v_s}{u_s} = v_s$.

Proof. Suppose for a moment that we have

$$\delta_{s-1}^* < v_s.$$

Let σ be the unique π -root in the minor disc D_{s-1}^* with

$$\delta_{s-1}^* = \sum a_j t^j + \pi t^{\delta_{s-1}^*}.$$

Then clearly we have

$$\text{ord } g(\sigma) = -\frac{v_s n}{d_s} + \delta_{s-1}^* \frac{u_s n}{d_s} = -\frac{v_s n}{d_s} + \delta_{s-1}^* \frac{n}{d_s} < 0.$$

It follows from Proposition 6.1 that σ is a distribution detector for $g(y)$, $T_1^\psi(y)$, $\dots$, $T_{s-1}^\psi(y)$.
Let

$$\begin{aligned} g(\sigma) &= g_\sigma(\pi) t^{n\lambda} + \dots, \\ T_i^\psi(\sigma) &= T_{i,\sigma}^\psi(\pi) t^{-\mu_i \lambda} + \dots \end{aligned} \quad \text{for } i = 1, \dots, s-1$$

where

$$\lambda = \frac{1}{d_s} [-v_s + \delta_{s-1}^*] < 0$$

and

$$\deg_\pi g_\sigma(\pi) = \frac{n}{d_s}, \quad \deg T_{i,\sigma}^\psi(\pi) = \frac{-\mu_i}{d_s} \quad \text{for } i = 1, \dots, s-1.$$

Note that

$$\text{g.c.d.} \left(\frac{n}{d_s}, \frac{-\mu_1}{d_s}, \dots, \frac{-\mu_{s-1}}{d_s} \right) = 1.$$

Thus the polynomials $g_\sigma(\pi)$, $T_{1,\sigma}^\psi(\pi)$, $\dots$, $T_{s-1,\sigma}^\psi(\pi)$ are powers of a common linear polynomial. It contradicts the very definition of δ_{s-1}^* . Q.E.D.

Corollary 6.1. *Suppose that there is a pair of polynomials $(f(x, y), g(x, y))$ satisfying the Jacobian condition with*

- (1) $f(x, y)$ and $g(x, y)$ are monic in y with y -degrees m and n respectively,
- (2) the associated characteristic data $\{M_i, d_i\}$ are with $M_s = n-2$, $d_s = 3$,
- (3) $\delta_s = -1$ (i.e., more than one point at infinite).

Then there exists a transformation χ which send $f(x, y)$, $g(x, y)$, $T_i^\psi(f(x, y), g(x, y))$ for $i = 1, \dots, s-1$ to polynomials $\bar{f}(\gamma, \pi)$, $\bar{g}(\gamma, \pi)$, $\bar{T}_i^\psi$ for $i = 1, \dots, s-1$ in $k[\gamma, \pi]$ a polynomial ring in two indeterminates with

- (4) $\bar{f}(\gamma, \pi)$, $\bar{g}(\gamma, \pi)$ are monic in π with π -degrees $\frac{m}{d_s}$ and $\frac{n}{d_s}$ respectively,
- (5) the associated characteristic data are $\left\{ \frac{M_i}{d_s}, \frac{d_i}{d_s} \right\}$ for $i = 1, \dots, s-1$,
- (6) $J_{\gamma, \pi}(\bar{f}(\gamma, \pi), \bar{g}(\gamma, \pi)) = a$ nonzero constant.

Proof. Since we have (cf. Lemma 6.3)

$$3 = d_s > v_s > \frac{d_s}{2}, \quad v_s + u_s = d_s = 3.$$

Then we must have

$$v_s = 2, \quad u_s = 1.$$

Our corollary follows from Propositions 6.4 and 6.3.

Q.E.D.

The previous corollary shows that in looking for a possible counter-example for the Jacobian conjecture of the smallest possible degrees we must assume that $d_s > 3$.

We may summarize some of our results in the following theorem.

Theorem. *Let $f(x, y)$ and $g(x, y)$ be two polynomials satisfying the Jacobian condition and the following conditions:*

- (1) $\deg f(x, y) = \deg_y f(x, y) = m, \quad \deg g(x, y) = \deg_y g(x, y) = n,$
- (2) *the associated characteristic data are $\{M_i, d_i\}$ for $i = 1, \dots, h,$*
- (3) *the number M_s is the largest one $\leq n - 2$.*

Then for any $r \geq 2$ we may construct a tower of major discs $D_s \supseteq D_{s-1} \supseteq \dots \supseteq D_r$, according to Definition 5.1. The roots of $\prod_{i=1}^r g(y) T_i^\psi(y)$ in D_r can be covered by a disjoint union of subdiscs $\bigcup E_j$ satisfying the requirements of Definition 1.1. We have

(4) *if the number of roots of $g(y)$ in E_i is $> \frac{n}{n - M_r}$ then the tower of major discs $D_s \supseteq D_{s-1} \supseteq \dots \supseteq D_r$ can be extended to $D_s \supseteq D_{s-1} \supseteq \dots \supseteq E_i,$*

(5) *if the number of roots of $g(y)$ in E_i is $\leq \frac{n}{n - M_r}$, then E_i is a minor disc and any π -root σ of $g(y)$ with $\text{ord } g(\sigma) < 0$ is a distribution detector for $g(y), T_1^\psi(y), \dots, T_{r-1}^\psi(y)$ (cf. Definition 3.1).*

(6) *The number of subdiscs E_i is bounded by*

$$(n - M_r) \cdot \frac{V_{r+1}}{d_{r+1}},$$

(7) *there must be at least one E_i satisfying the requirement in (4).*

Proof. Propositions 5.4 and 6.1.

One application of the theorem and the propositions of this article is to solve the Jacobian conjecture for all polynomials of degrees less than 100. For this purpose we consider a pair of polynomials f and g with the nonrestrictive assumption that $f = T_1^\psi$. We shall search for the sequences of integers $\{n, M_1, \dots, M_s\}$ with

- (1) $\deg f(x, y) = m = -M_1 < \deg g(x, y) = n \leq 100,$
- (2) $m \nmid n, \quad M_s = n - 2,$

(3) $J_{x,y}(f(x, y), g(x, y)) = 1$ and the degrees of $f(x, y)$ and $g(x, y)$ can not be reduced simultaneously,

(4) $\{n, M_1, \dots, M_s\}$ is the part of characteristic data of (f, g) which are less than $n-2$,

$$(5) \quad d_r = \text{g.c.d.} \{n, M_1, \dots, M_{r-1}\}.$$

Note that it follows from Corollary 6.1 that $d_s \geq 4$. A simple computation shows that $s \leq 5$. Combining this with Proposition 5.5 we conclude

$$(6) \quad 3 \leq s \leq 5, \quad d_s \geq 4.$$

Furthermore we assign a sequence of integers $\{V_s, \dots, V_2\}$ for a hypothetically existing tower of major discs $D_s \supset \dots \supset D_1$ (cf. Theorem and Definition 5.1) with

$$(7) \quad V_{r+1} \left(\frac{d_r}{d_{r+1}} \right) \geq V_r > \frac{d_r}{n - M_r},$$

(8) the radii $\delta_s, \dots, \delta_1$ of $D_s, \dots, D_1$ can be computed from Definition 5.1. We shall consider the increment of the denominator A_{r-1} of δ_{r-1} with respect to $\delta_s, \dots, \delta_r$. Namely let the l.c.m. of the reduced denominators of $\delta_s, \dots, \delta_r$ be L . Then A_{r-1} is defined to be the reduced denominator of $L\delta_{r-1}$. Let us consider the following division algorithm equation

$$(9) \quad V_r \left(\frac{d_{r-1}}{d_r} \right) = \Delta_{r-1} A_{r-1} + \square_{r-1}$$

Let $(\bar{t})^{LA_{r-1}} = t$. Due to the existence of the following automorphism of $k\langle\bar{t}\rangle$ over $k\langle\bar{t}^{LA_{r-1}}\rangle$

$$\bar{t} \rightarrow \omega \bar{t}$$

where ω is a A_{r-1} -th root of unity, the value of V_{r-1} is further restricted by

$$(10) \quad V_{r-1} \leq \Delta_{r-1}$$

if the corresponding factor of $p(\pi)$ is of the form $\pi - a$ with $a \neq 0$ or for some j the following

$$(11) \quad V_{r-1} = jA_{r-1} + \square_{r-1}$$

if the corresponding factor of $p(\pi)$ is of the form π . Finally, when $r=2$ we must have the following (cf. the proof of Proposition 5.5):

$$(12) \quad A_1 \left| \left(\frac{n}{d_2} \right) V_2, \quad A_1 \left| \left(\frac{m}{d_2} \right) V_2 - 1 \right.$$

or

$$(13) \quad A_1 \left| \left(\frac{m}{d_2} \right) V_2, \quad A_1 \left| \left(\frac{n}{d_2} \right) V_2 - 1 \right.$$

Furthermore we shall note that the situation indicated by the equation (11) can not always happen as established by Proposition 5.6.

The above discussion can be easily put into a computer program. The computer program produces only the following exceptions which satisfy the necessary numerical restrictions.

n	$m = -M_1$	M_2	M_3	M_4	V_3	V_2	δ_2	δ_1
64	48	52	62	63	3	3	$\frac{1}{4}$	$\frac{9}{16}$
84	56	64 [72]	82	83	3	2 [5]	$\frac{2}{7} \left[\frac{1}{4} \right]$	$\frac{16}{21} \left[\frac{7}{12} \right]$
75	50	55	73		4	3 [2]	$\frac{1}{5}$	$\frac{1}{2} \left[\frac{1}{3} \right]$
99	66	77	97		8	8	$\frac{1}{3}$	$\frac{4}{9}$

It is easy to see that the above table is very restrictive. For instance given $\deg g = 75$ $\deg f = 50$ there are 25 possible values for M_2 and 2 possible values for V_3 etc. The above table determines only one possible value for M_2 and one for V_3 . These data are very useful in excluding these four exceptional data from the possible counterexamples to the Jacobian conjecture. Without them a brute force computation will involve the four thousands coefficients of $g(x, y)$ and $f(x, y)$. The treatment of these four exceptional cases will be presented in Appendix II.

Appendix I. The polynomial solutions of some ordinary differential equations

Recall Definition 4.1 of the symbol $D(a, b, p(\pi), q(\pi))$ from Section 4. We shall further assume that $p(\pi)$ and $q(\pi)$ are nonconstants in this section.

Proposition A. 1. *Suppose that at least one of the constants a, b is not zero. Then the following equation*

$$D(a, b, p(\pi), q(\pi)) = 0$$

implies for some $e, c \in k$ the following:

- (1) $a = e \deg_{\pi} p(\pi),$
- (2) $b = e \deg_{\pi} q(\pi),$
- (3) $p(\pi)^{\deg_{\pi} q(\pi)} = cq(\pi)^{\deg_{\pi} p(\pi)}.$

Proof. Let $p(\pi)$ and $q(\pi)$ be expressed as

$$p(\pi) = a_1 \pi^m + \text{lower terms in } \pi,$$

$$q(\pi) = a_2 \pi^n + \text{lower terms in } \pi.$$

Then we have

$$\begin{aligned} D(a, b, p(\pi), q(\pi)) &= ap(\pi) q'(\pi) - bq(\pi) p'(\pi) \\ &= (an - bm) a_1 a_2 \pi^{n+m-1} + \text{lower terms.} \end{aligned}$$

Clearly we must have

$$an = bm.$$

Thus the conclusions (1) and (2) follow. Clearly we have

$$D(a, b, p(\pi), q(\pi)) = 0 \Leftrightarrow D(m, n, p(\pi), q(\pi)) = 0.$$

The following identity is elementary

$$\left(\frac{q(\pi)^m}{p(\pi)^n} \right)' = \frac{p(\pi)^{n-1} q(\pi)^{m-1} D(m, n, p(\pi), q(\pi))}{p(\pi)^{2n}} = 0.$$

Thus we conclude

$$\frac{q(\pi)^m}{p(\pi)^n} = c \in k$$

and (3) follows. Q.E.D.

Proposition A. 2. Let $m = \deg p(\pi)$ and $n = \deg q(\pi)$. Suppose that $n = (l-1)m + 1$ and $p(\pi), q(\pi)$ satisfy the following equation

$$D(m, n + \varepsilon, p(\pi), q(\pi)) = cp(\pi)^l \neq 0$$

where $c \in k$ and ε is a rational number > -1 . Then $p(\pi)$ and $q(\pi)$ are powers of a common linear polynomial.

Proof. Let us prove the case $l = 1$. Then we have $n = \deg q(\pi) = 1$. We may assume $q(\pi) = \pi$. The differential equation is reduced to the following form

$$-(1 + \varepsilon) \pi p'(\pi) = (c - m) p(\pi).$$

Note that the coefficients are not zero. Consider the multiplicities of roots of $p(\pi)$. We deduce that π is the only root of $p(\pi)$. Hence the proposition has been established for the case $l = 1$.

Inductively let us assume that $l > 1$ and the proposition has been proved for smaller l . We claim that

$$p(\pi) \mid q(\pi).$$

Let π (or in general $\pi - a$) be a factor of $p(\pi)$ with

$$\begin{aligned} p(\pi) &= \pi^u p_1(\pi) \quad \text{with} \quad p_1(0) \neq 0, \\ q(\pi) &= \pi^v q_1(\pi) \quad \text{with} \quad q_1(0) \neq 0. \end{aligned}$$

To prove our claim it suffices to show that $u \leq v$. Note that the differential equation is of the following form

$$(1) \quad mp(\pi) q'(\pi) - (n + \varepsilon) q(\pi) p'(\pi) = c \pi^l (p_1(\pi))^l.$$

Let

$$\begin{aligned} p_1(\pi) &= b_0 + b_1 \pi + \dots, \\ q_1(\pi) &= c_0 + c_1 \pi + \dots. \end{aligned}$$

Then the left hand side of the differential equation is of the following form

$$[mv - (n + \varepsilon)u] b_0 c_0 \pi^{u+v-1} + \text{higher terms in } \pi.$$

The restriction

$$\varepsilon > -1$$

implies

$$n + \varepsilon \geq (l - 1)m \geq m.$$

Thus the inequality

$$(2) \quad u > v$$

implies

$$mv - (n + \varepsilon)u \neq 0.$$

After comparing π -orders on the two sides of the differential equation (1) we conclude

$$u + v - 1 = lu.$$

It contradicts the inequality (2). Thus we must have

$$u \leq v.$$

It is clear that the above argument works for an arbitrary factor $(\pi - a)$ of $p(\pi)$. Thus we conclude

$$q(\pi) = p(\pi) q^*(\pi).$$

A simple computation shows

$$D(m, n + \varepsilon, p(\pi), q(\pi)) = cp(\pi)^l \Leftrightarrow D(m, (n - m) + \varepsilon, p(\pi), q^*(\pi)) = cp(\pi)^{l-1}.$$

Thus by the induction hypothesis we conclude that the polynomials $p(\pi)$, $q^*(\pi)$ and $q(\pi)$ are all powers of a common linear polynomial. Q.E.D.

Proposition A. 3. Let $m = \deg p(\pi)$, $n = \deg q(\pi)$. Suppose that the polynomials $p(\pi)$ and $q(\pi)$ satisfy

$$D(m, n, p(\pi), q(\pi)) = cp(\pi)$$

where $c \neq 0 \in k$. Then we have:

- (1) every root of $p(\pi)$ is a root of $q(\pi)$,
- (2) $q(\pi)$ has no multiple root,
- (3) $p(\pi)$ is not a power of $q(\pi)$,
- (4) there is at least one root of $p(\pi)$ with multiplicity $> \frac{m}{n}$.

Proof. The differential equation can be rewritten as

$$nq(\pi)p'(\pi) = (mq'(\pi) - c)p(\pi).$$

Let $(\pi - a)$ be a factor of $p(\pi)$. Comparing the multiplicities of $(\pi - a)$ on both sides of the above equation, we conclude that $(\pi - a)$ is a factor of $q(\pi)$. The conclusion (1) thus follows.

It is easy to see that every multiple root of $q(\pi)$ must be a root of $p(\pi)$. Let $(\pi - a)$, say $a = 0$, be a multiple factor of $q(\pi)$ with

$$\begin{aligned} p(\pi) &= \pi^u p_1(\pi) & p_1(0) &\neq 0, \\ q(\pi) &= \pi^v q_1(\pi) & q_1(0) &\neq 0. \end{aligned}$$

Then we have

$$\pi^{u+v-1} |D(m, n, p(\pi), q(\pi)) = cp(\pi).$$

A contradiction. The conclusion (2) thus follows.

In case that $p(\pi)$ is a power of $q(\pi)$ it is elementary to verify

$$D(m, n, p(\pi), q(\pi)) = 0.$$

It contradicts our assumptions on c and $p(\pi)$.

Suppose that all roots of $p(\pi)$ are with multiplicities $\leq \frac{m}{n}$. Then it follows from conclusions (1) and (2) that all multiplicities must be equal to $\frac{m}{n}$. Thus we must have

$$p(\pi) = c_1 q(\pi)^{\frac{m}{n}}.$$

It contradicts conclusion (3).

Q.E.D.

Proposition A. 4. Let $m = \deg p(\pi)$ and $n = \deg q^*(\pi)$. Suppose that we have

$$D(m, n, p(\pi), q^*(\pi)) = cp(\pi)^{l+1}$$

where $c \neq 0 \in k$ and $l \geq 0$. Then we have

- (1) $q^*(\pi) = p(\pi)^l q(\pi)$,
- (2) $D(m, n - lm, p(\pi), q(\pi)) = cp(\pi)$.

Proof. For $l = 0$ it is Proposition A. 3. Let us assume that $l \geq 1$ and our proposition has been proved for all smaller l . Clearly we have

$$n > m.$$

Otherwise the right hand side of the differential equation has a larger π -degree. Let $\pi - a$, say $a = 0$, be a factor of $p(\pi)$ with

$$\begin{aligned} p(\pi) &= \pi^u p_1(\pi) & p_1(0) &\neq 0, \\ q^*(\pi) &= \pi^v q_1(\pi) & q_1(0) &\neq 0. \end{aligned}$$

Let

$$\begin{aligned} p_1(\pi) &= b_0 + b_1 \pi + \dots, \\ q_1(\pi) &= c_0 + c_1 \pi + \dots. \end{aligned}$$

Then we have

$$D(m, n, p(\pi), q^*(\pi)) = \pi^{u+v-1} [(mv - nu)b_0 c_0 + \dots].$$

If we have

$$v < u$$

then the leading coefficient

$$(mv - nu)b_0 c_0 \neq 0.$$

After comparing the multiplicities of π on both sides of the differential equation we deduce a contradiction. Thus we conclude

$$v \geq u.$$

It is clear that the above arguments work for any factor $(\pi - a)$ of $p(\pi)$. Thus we conclude

$$p(\pi) | q^*(\pi) = p(\pi) \bar{q}(\pi).$$

We clearly have

$$D(m, n, p(\pi), q^*(\pi)) = cp(\pi)^l \Leftrightarrow D(m, n - m, p(\pi), \bar{q}(\pi)) = cp(\pi)^{l-1}.$$

Our proposition has been established by mathematical induction.

Q.E.D.

Proposition A. 5. Let $m = \deg p(\pi)$ and $n = \deg q(\pi)$. Suppose that $p(\pi)$ and $q(\pi)$ satisfy

$$D(m, n, p(\pi), q(\pi)) = c$$

where $c \neq 0 \in k$. Then $p(\pi) \cdot q(\pi)$ has only simple roots.

Proof. Routine.

Appendix II. Some special cases

A simple computer program shows that the only possible counter-examples to the Jacobian conjecture for polynomials of degrees less than 100 are of degrees (64, 68), (84, 56), (75, 50) and (99, 66) with some special data attached. A routine computation shows that the numbers of coefficients of the pairs of polynomials involved are 3370, 5308, 4352 and 7348 respectively. The task of direct verification of the Jacobian conjecture by computation for those pairs of polynomials with the said degrees is beyond the capacity of a large computer. In this Appendix we shall illustrate some methods to reduce the numbers of coefficients, and then settle the problems of computations.

We shall apply Proposition 6.4 to the first three cases. Note that in these cases we always have

$$u_3 = d_3 - v_3 = 1.$$

Thus it follows from Proposition 6.3 that our table can be transformed to the following table.

n	$m = -M_1$	M_2	V_2	δ_2	δ_1	Jacobian
16	12	13	3	-1	$\frac{1}{4}$	X
21	14	16 [18]	2 [5]	$-\frac{1}{2} [-1]$	$\frac{7}{6} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$	X
15	10	11	3 [2]	-1	$\frac{1}{2} \begin{bmatrix} 4 \\ 3 \end{bmatrix}$	X^2

The numbers of coefficients are reduced to 244, 373 and 202 respectively.

We shall use the first case to illustrate the method of further reduction of the numbers of coefficients. We shall choose x, y properly so that the π -root in the major disc D_1 is of the following form

$$\sigma = \pi t^{\frac{1}{4}}$$

and the π -root σ^* in the minor disc D_1^* is of the following form

$$\sigma^* = t^{-1} + a_0 + a_1 t + a_2 t^2 + \pi t^3.$$

To truncate σ^* properly we define $\bar{\sigma}$ as

$$\bar{\sigma} = t^{-1} + a_0 + a_1 t + a_2 t^2 + \pi t^{\frac{11}{4}}.$$

Then it is easy to see that $\{\sigma, \bar{\sigma}\}$ is a coherent complete system of π -roots for $\bar{g}(y)$ (and $\bar{f}(y) = \bar{T}_1(y)$) with accuracy $-1 \left(\text{resp. } -\frac{3}{4} \right)$. Then it follows from Theorem 1.2 that

$$\begin{aligned}\bar{g}(y) &= h(y)^4 + \alpha_1(y) h(y)^3 + \alpha_2(y) h(y)^2 + \alpha_3(y) h(y) + \alpha_4(y), \\ \bar{f}(y) &= h(y)^3 + \beta_2(y) h(y) + \beta_3(y)\end{aligned}$$

with

$$h(y) = \text{the fourth approximate root of } \bar{f}(y)$$

$$\deg_y \alpha_i(y) \leq 3, \quad \deg_y \beta_i(y) \leq 3,$$

$$\text{ord } \alpha_i(\sigma) \geq i \left(-\frac{1}{4} \right), \quad \text{ord } \alpha_i(\bar{\sigma}) \geq i \left(-\frac{1}{4} \right),$$

$$\text{ord } \beta_i(\sigma) \geq i \left(-\frac{1}{4} \right), \quad \text{ord } \beta_i(\bar{\sigma}) \geq i \left(-\frac{1}{4} \right),$$

$$\text{ord } h(\sigma) \geq -\frac{1}{4}, \quad \text{ord } h(\bar{\sigma}) \geq -\frac{1}{4}.$$

It is easy to see that

$$\begin{aligned}(1) \quad h(x, y) &= y^3(y-x) + b_1 y^3 + b_2 y^2 + b_3 y + b_4 \\ &= yA(x, y) + b_4 \\ &= y^2B(x, y) + b_3 y + b_4\end{aligned}$$

with $A(x, y)$ and $B(x, y)$ defined by the above equations and

$$\begin{aligned}(2) \quad &\alpha_1 \text{ is a constant,} \\ (3) \quad &\alpha_2 = c_1 A + c_2, \quad \beta_2 = c_3 A + c_4, \\ (4) \quad &\alpha_3 = c_5 A + c_5 B + c_7, \quad \beta_3 = c_8 A + c_9 B + c_{10}, \\ (5) \quad &\alpha_4 = c_{11} A + c_{12} B + c_{13}(y-x).\end{aligned}$$

Thus the total number of coefficients is further reduced to 17.

Similar arguments can be applied to the second case to show the impossibility directly and to the third case to reduce the number of coefficients to 22 [15 or 13].

We shall use the following method to further simplify our computations. Let

$$\begin{aligned}f &= \eta^{-12}, \\ g &= \eta^{-16} + a_1 \eta^{-12} + a_2 \eta^{-8} + a_3 \eta^{-4} + a_4 + \dots.\end{aligned}$$

Then we may replace (f, g) by $\left(f - \frac{3}{4}a_3, g - a_1f - a_4\right)$ to reduce the above expression to

$$g = \eta^{-16} + a_2\eta^{-8} + \text{terms in } \eta \text{ with positive exponents.}$$

Note that

$$f = h^3 + \beta_2 h + \beta_3 = h^3(1 + \beta_2 h^{-2} + \beta_3 h^{-3}).$$

It is easy to conclude that

$$g = h^4 + \frac{4}{3}(\beta_2 h^2 + \beta_3 h) + \frac{2}{9}(\beta_2^2 + 2\gamma) - \frac{4}{81}\delta + a_2 h^2$$

where γ and δ are defined by

$$\beta_2 \beta_3 = \gamma h + \gamma^*, \quad \beta_2^3 = \delta h^2 + \delta^*$$

with

$$\deg_y \gamma^* < \deg h = 4, \quad \deg_y \delta^* < \deg h^2 = 8.$$

Thus the total number of coefficients are reduced to 10. This method can be used to reduce the numbers of coefficients of the third case to 12 [10 or 7].

In the fourth case there are complications. First, there are two possibilities according to the distribution of the roots of $g(y)$ in the minor disc D_2^* . Namely, there is an unique π -root σ of $g(y)$ of the following form

$$\sigma = a_{-1}t^{-1} + a_0 + a_1t + \pi t^2$$

with

$$g(\sigma) = g_\pi(\sigma)t^{-18} + \dots.$$

The polynomial $g_\pi(\pi)$ is either a power of a linear polynomial or the 9-th power of a cubic polynomial with precisely two roots. For the first possibility the data can be transformed to

n	$m = -M_1$	M_2	V_2	δ_2	δ_1	Jacobian
27	18	21	8	-1	0	X^4

The above mentioned method can be used to reduce the number of coefficients to 10. For the second possibility we can show that for suitable variables x and y there is a transformation Ω of the following form

$$x \rightarrow x^{-1}, \quad y \rightarrow a_0 + a_1x + a_2x^2 + yx^3$$

such that

$$\Omega(f(x, y)) = x^6 y^{18} + \dots,$$

$$\Omega(g(x, y)) = x^9 y^{27} + \dots,$$

$$\Omega(T_2^\psi(x, y)) = x^5 y^{15} + \dots,$$

$$\Omega(T_3^\psi(x, y)) = x^{15} y^{40} + \dots,$$

$$J_{x,y}(\Omega f(x, y), \Omega g(x, y)) = x.$$

A discussion of the solutions of $\Omega(g(x, y)) = 0$ in $k\langle\langle x^{-1} \rangle\rangle$ and in $k\langle\langle y^{-1} \rangle\rangle$ with a slightly modified version of the above method we may reduce the direct computation to involve only 11 variables.

Now we are in the position to settle these cases by direct computation. Let us handle the case $n = 15$, $m = 10$, $M_2 = 11$, $V_2 = 3$. After simplifying, we may assume

$$\begin{aligned} f &= h^2 + 2\beta = h^2(1 + 2\beta h^{-2}) = \eta^{-10}, \\ g &= \eta^{-15} + a\eta^5 + b\eta^6 + c(x^3 + d)\eta^9 + \dots \\ &= f^{\frac{3}{2}} + af^{-\frac{1}{2}} + bf^{-\frac{6}{10}} + c(x^3 + d)f^{-\frac{9}{10}} + \dots \end{aligned}$$

where a, b, c, d are constants and $b \neq 0$, $c \neq 0$. Clearly we have

$$\begin{aligned} f^{\frac{3}{2}} &= h^3(1 + 2\beta h^{-2})^{\frac{3}{2}} = h^3 + 3\beta h + \frac{3}{2} \frac{\beta^2}{h} - \frac{1}{2} \frac{\beta^3}{h^3} + \frac{3}{8} \frac{\beta^4}{h^5} + \dots, \\ af^{-\frac{1}{2}} &= a \left[\frac{1}{h} + \dots \right]. \end{aligned}$$

Substituting the above expressions into the expression for g we have

$$(1) \quad g = h^3 + 3\beta h + \frac{3}{2} \alpha + \frac{3}{2} \frac{\gamma}{h} - \frac{1}{2} \frac{\beta^3}{h^3} + \frac{a}{h} + bf^{-\frac{6}{10}} + c(x^3 + d^*)f^{-\frac{9}{10}} + \dots$$

where d^* is a constant and α, γ are defined by the following equations:

$$(2) \quad \beta^2 = \alpha h + \gamma$$

with the restriction

$$\deg_y \gamma < \deg_y h = 5.$$

Since g is a polynomial in y , then the powerseries in y^{-1} in the expression (1) shall naturally cancel out. Thus we must have the following equations (3) and (4)

$$(3) \quad \beta^3 = (3\gamma + a)h^2 + \varepsilon h + \delta$$

where ε and δ are polynomials with y -degrees less than 4,

$$(4) \quad \frac{1}{2} \varepsilon = bf^{\frac{4}{10}} + c(x^3 + d^*)y + \text{a powerseries in } y^{-1}.$$

Let us consider the possible forms of the polynomial $h(x, y)$. Since we assume that the polynomials $f(x, y)$ and $g(x, y)$ are the transforms of some polynomials of degrees 50 and 75 respectively according to Proposition 6.3, the inverse transformation implies that there is a minor disc. Thus in the major disc D_2 there are precisely three subdiscs which contains 2, 2, 6 of roots f respectively. The datum $\delta_1 = \frac{1}{2}$ implies that the poly-

nomials $h(x, y)$ and $\beta(x, y)$ must be of the following form (cf. the beginning of this Appendix)

$$(5) \quad \begin{aligned} h &= (y^2 - x^2 + a_1 y + a_2 x + a_3) y^3 + (a_4 x + a_5) y^2 + (a_6 x + a_7) y + a_8 \\ &= Ay + a_8 = By^2 + (a_6 x + a_7) y + a_8, \end{aligned}$$

$$(6) \quad \beta = a_9 A + a_{10} y + a_{11} x + a_{12}$$

with A, B thus defined. In fact the existence of the polynomials h and β of the above forms and satisfying the equations (2), (3), (4) will furnish a counterexample to the Jacobian conjecture.

From equation (3) we deduce

$$(7) \quad \deg_y \gamma \leq 2.$$

Using the equations (5) and (6) we get

$$\begin{aligned} \beta^2 &= (a_9^2 B + 2a_9 a_{10})h + [(a_9^2 a_6 + 2a_9 a_{11})x + (a_9^2 a_7 + 2a_9 a_{12})]A \\ &\quad - a_9^2 a_{10} B + (a_{10} y + a_{11} x + a_{12})^2 - 2a_8 a_9 a_{10}. \end{aligned}$$

In view of equation (2) we conclude

$$\begin{aligned} \alpha &= a_9^2 B + 2a_9 a_{10}, \\ \gamma &= [(a_9^2 a_6 + 2a_9 a_{11})x + (a_9^2 a_7 + 2a_9 a_{12})]A - a_9^2 a_{10} B \\ &\quad + (a_{10} y + a_{11} x + a_{12})^2 - 2a_8 a_9 a_{10}. \end{aligned}$$

In view of equation (7) we have the following equations

$$\begin{aligned} a_9^2 a_6 + 2a_9 a_{11} &= 0, \\ a_9^2 a_7 + 2a_9 a_{12} &= 0, \\ a_9^2 a_{10} &= 0. \end{aligned}$$

We have two possibilities $a_9 = 0$ or $a_9 \neq 0$.

Case 1. $a_9 = 0$. We then have

$$\alpha = 0, \quad \beta = a_{10} y + a_{11} x + a_{12}, \quad \gamma = \beta^2.$$

Thus equation (3) can not be fulfilled.

Case 2. $a_9 \neq 0$. We then have

$$a_{10} = 0, \quad \gamma = (a_{11} x + a_{12})^2, \quad \deg_y \beta = 4.$$

Thus equation (3) implies that the polynomial γ must be of y -degree 2. A contradiction. This establishes the impossibility for $n = 15$, $m = 10$, $M_2 = 11$ and $V_2 = 3$.

All other cases can be computed directly as above. There is no counter-example of polynomials of degrees less than or equal to 100 for the Jacobian conjecture.

References

- [A-M.1] S. S. Abhyankar and T. T. Moh, Newton-Puiseux expansion and Tschirnhausen transformation. I & II, J. reine angew. Math. **260** (1973), 47—63 and **261** (1973), 29—54.
- [A-M.2] S. S. Abhyankar and T. T. Moh, Embeddings of the line in the plane, J. reine angew. Math. **276** (1975), 148—166.
- [A.1] S. S. Abhyankar, On Expansion Techniques in Algebraic geometry, TATA Institute Bombay 1977.
- [AN.1] G. Angermüller, Die Wertehalgruppe einer ebenen irreduziblen algebroiden Kurve, Math. Z. **153** (1977), 267—282.
- [B-C-W.1] H. Bass, E. Connell and D. Wright, Polynomial automorphisms and the Jacobian conjecture, Preprint.
- [E.1] W. Engel, Ein Satz über ganze Cremona-Transformationen der Ebene, Math. Ann. **130** (1955), 11—19.
- [K.1] O. H. Keller, Ganze Cremona-Transformationen, Monatshefte Math. Phys. **47** (1973), 299—306.
- [M.1] T. T. Moh, On the concept of approximate roots for algebra, J. Alg. **65** (1980), 301—306.
- [M.2] T. T. Moh, On analytic irreducibility at ∞ , Proc. Amer. Math. Soc. **44** (1974), 22—24.
- [M.3] T. T. Moh, On characteristic pairs, Bull. Inst. Math. Acad. Sinica **1** (1973), 75—91.
- [M.4] T. T. Moh, On two fundamental theorems for the concept of approximate roots, J. Math. Soc. Japan **34** (1982), 637—652.
- [N.1] M. Nagata, On the Automorphism group of $K[X, Y]$, Lectures in Mathematics, Tokyo 1972.
- [S.1] A. Sathaye, On planar curves, Amer. J. Math. **99** (1977), 1105—1135.
- [S.2] B. Segre, Forme differenziali e loro integrali. 2, Roma 1956.
- [W.1] R. Walker, Algebraic Curves, New York 1962.
- [W.2] S. Wang, A Jacobian criterion for separability, J. Alg. **65** (1980), 453—494.
- [Z.1] O. Zariski, Algebraic Surfaces, Berlin-Heidelberg-New York 1935.

Department of Mathematics, Purdue University, West Lafayette, Indiana 47907 USA

Eingegangen 10. Januar 1979, in revidierter Form 18. August 1982