Lectures on
Expansion Techniques In
Algebraic Geometry

By

S.S. Abhyankar

Tata Institute Of Fundamental Research

Bombay

1977

Lectures on

Expansion Techniques In Algebraic Geometry

By

S.S. Abhyankar

Notes by

Balwant Singh

Tata Institute Of Fundamental Research

Bombay

1977

Printed in India

By

K.P. Puthran at Tata Press Limited, Bombay 400025

and published by K.G. Ramanathan for the

Tata Institute of Fundamental Research,

Bombay 400005

Preface

These notes are based upon my lectures at the Tata Institute from November 1975 to March 1976 and further oral communication between me and the note taker.

The notes are divided into two parts. In §8 or Part One we prove the Fundamental Theorem on the structure of the coordinate ring of a meromorphic curve and its value group. We then give some applications of the Fundamental Theorem, the principal one among them being the Epimorphism Theorem. The proof of the Main Lemmas (§ 7) presented here is a simplified version of the original proof of Abhyankar and Moh. The process of simplification started with my lectures at Poona University in 1975 and culminated into the present version during my lectures at the Tata Institute. The simplification resulted mainly from the keen and stimulating interest in my lectures shown by the audience at these two places, especially at the Tata Institute.

In Part Two we record some progress on the Jacobian problem, which is as yet unsolved. The results presented here were obtained by me during 1970-71. Partial notes on these were prepared by M. van der Put and W. Heinzer at Purdue University in 1971. However, since the notes were not complete, they were never formally circulated.

I wish to thank the Tata Institute for inviting me and providing me with an opportunity to give these lectures. My special thanks go to Balwant Singh who took over the task of recording the lectures and preparing these notes entirely on his own even to the extent of relieving me of the tedium of having to read and check the manuscript.

S. S. Abhyankar

Notation

The following notation is used in the sequel.

The set of integers (resp. non-negative integers, positive integers, real numbers) is denoted by $\mathbb{Z}$ (resp. $\mathbb{Z}^+$, $\mathbb{N}$, $\mathbb{R}$). We write card (S) for the cardinality of a set S and we write $\inf(S)$ (resp. $\sup(S)$) for the infimum (resp. supremum) of a subset S of $\mathbb{R}$. If T is a subset of a set S then $S - T$ denotes the complement of T in S. If k is a field and n is a positive integer, we denote by $\mu_n(k)$ the group of nth roots of unity in k. For $w \in \mu_n(k)$ we write $\mathrm{ord}(w)$ for the order of w i.e., $\mathrm{ord}(w)$ is the least positive integer r such that $w^r = 1$.

Suppose. in a given context, k is a fixed field. We then denote by the symbol $\varnothing$ a generic (i.e. unspecified) non-zero element of k. Thus if k' is a ring containing k and $a \in k'$ then $a = \varnothing$ means that $a \in k$ and $a \neq 0$. Similarly, $b = \varnothing c$ means that $b = ac$ for some $a \in k$, $a \neq 0$. Note that $a = \varnothing$, $b = \varnothing$ does not mean that $a = b$.

Contents

Part I

Meromorphic Curves

Chapter 1

G-Adic Expansion and Approximate Roots

1 Strict Linear Combinations

(1.1) NOTATION. Let e be a non-negative integer and let $r = (r_0, r_1,$ 1
$\dots, r_e)$ be an $(e+1)$-tuple of integers such that $r_0 \neq 0$. We define

$$d_i(r) = g.c.d.(r_0, \dots, r_{i-1}), \quad 1 \leq i \leq e+1.$$

Since $r_0 \neq 0$, we have $d_i(r) > 0$ for every i. Moreover, it is clear that $d_{i+1}(r)$ divides $d_i(r)$ for $1 \leq i \leq e$. We put $n_i(r) = d_i(r)/d_{i+1}(r)$ for $1 \leq i \leq e$.

(1.2) LEMMA. *Let j, c be integers such that $1 \leq j \leq e$ and $0 \leq c < n_j(r)$. If $n_j(r)$ divides $cr_j/d_{j+1}(r)$ then $c = 0$.*

Proof. Since g.c.d. $(d_j(r), r_j) = g.c.d.\ (r_0, \dots r_j) = d_{j+1}(r)$, we have g.c.d. $(n_j(r), r_j/d_{j+1}(r)) = 1$. Therefore if $n_j(r)$ divides $cr_j/d_{j+1}(r)$ then $n_j(r)$ divides c. Therefore, since $0 \leq c < n_j(r)$, we get $c = 0$. □

(1.3) LEMMA. *Let j, c be integers $1 \leq j \leq e$ and let $c = \sum_{i=0}^{j} c_i r_i$ with $c_i \in \mathbb{Z}$ for $0 \leq i \leq j$. Assume that $0 < c_j < n_j(r)$. Let*

$$j' = \inf \left\{ i \,\middle|\, 1 \leq i \leq e+1, d_i(r) \text{ divides } c \right\}.$$

Then $j' = j + 1$. In particular, $d_1(r)$ does not divides c and $c \neq 0$.

Proof. Since $d_{j+1}(r)$ divides r_i for $0 \leq i \leq j$, it is clear that $d_{j+1}(r)$ divides c. Therefore $j' \leq j + 1$. Next, since $0 < c_j < n_j(r)$, we see by lemma (1.2) that $n_j(r)$ does not divide $c_j r_j / d_{j+1}(r)$. Therefore $d_j(r)$ does not divide $c_j r_j$. Since $d_j(r)$ divides $\sum_{i=0}^{j-1} c_i r_i$, we conclude that $d_j(r)$ does not divide c. This proves that $j' \geq j + 1$. □

2 **(1.4) DEFINITION.** Let Γ be a subsemigroup of $\mathbb{Z}$. By a Γ- *strict linear combination a* of r we mean an expression of the form

$$a = \sum_{i=0}^{e} a_i r_i$$

with $a_0 \in \Gamma$ and $a_i \in \mathbb{Z}$, $0 \leq a_i < n_i(r)$ for $1 \leq i \leq e$. If $\Gamma = \mathbb{Z}^+$ then we call a Γ-strict linear combination of r simply a *strict linear combination* of r.

(1.5) PROPOSITION. Let Γ be a subsemigroup of $\mathbb{Z}$ and let

$$a = \sum_{i=0}^{e} a_i r_i, \quad b = \sum_{i=0}^{e} b_i r_i$$

be Γ- strict linear combinations of r. If $a = b$ then $a_i = b_i$ for every i, $o \leq i \leq e$.

Proof. If the assertion is false then there exists an integer $j, 0 \leq j \leq e$, such that $a_j \neq b_j$ and $a_i = b_i$ for $j + 1 \leq i \leq e$. We may assume without loss of generality that $a_j > b_j$. Writing $c = a - b$ and $c_i = a_i - b_i$ for every i, we get

$$c = \sum_{i=0}^{j} c_i r_i, \quad c_j > 0.$$

Since $c = 0$ and $r_0 \neq 0$, we have $j \geq 1$. Therefore we have $0 \leq a_j < n_j(r)$ and $0 \leq b_j < n_j(r)$, which shows that $0 < c_j < n_j(r)$, since $c_j > 0$. Therefore $c \neq 0$ by Lemma (1.3). This is a contradiction. □

(1.6) COROLLARY. If an integer a can be expressed as a Γ- strict linear combination of r then such an expression of a is unique.

(1.7) DEFINITION. Let Γ, G be subsemigroups of $\mathbb{Z}$. We say G is *strictly generated* (resp. Γ- *strictly generated*) by r if G coincides with the set of all strict (resp. Γ- strict) linear combinations of r.

(1.8) PROPOSITION. Assume that $e \geq 1$ and $r_i \leq 0$ for $i = 0, 1$. If 3
$-d_2(r)$ can be expressed as a strict linear combination of r then r_0 divides r_1 or r_1 divides r_0.

Proof. Let $d_i = d_i(r)$, $1 \leq i \leq e + 1$. Suppose $-d_2$ is a strict linear combination of r. Then

$$-d_2 = \sum_{i=0}^{e} c_i r_i$$

with $c_0 \in \mathbb{Z}^+$, $c_i \in \mathbb{Z}$, $0 \leq c_i < n_i(r)$ for $1 \leq i \leq e$. Since $-d_2 \neq 0$, there exists i, $0 \leq i \leq e$, such that $c_i \neq 0$. Let

$$j = \sum \left\{ i \middle| 0 \leq i \leq e, \quad c_i \neq 0 \right\}.$$

$\square$

Then we have

$$-d_2 = \sum_{i=0}^{j} c_i r_i, \quad c_j \neq 0.$$

Note that, since $r_0 \neq 0$, we have $r_0 < 0$ by assumption. Now, if $j = 0$ then $-d_2 = c_0 r_0$, so that r_0 divides d_2. Therefore in this case r_0 divides r_1. Assume now that $j \geq 1$. Then $0 < c_j < n_j(r)$. Since d_2 divides $-d_2$, it follows from Lemma (1.3) that $j \leq 1$. Therefore $j = 1$ and we have

$$-d_2 = c_0 r_0 + c_1 r_1 \tag{1.8.1}$$

with $c_0 \in \mathbb{Z}^+$, $c_1 \in \mathbb{Z}$ and $0 < c_1 < n_1(r)$. The last inequalities mean, in particular, that $d_1/d_2 = n_1(r) > 1$, so that

$$-r_0 = d_1 > d_2 = \text{g.c.d. } (r_0, r_1).$$

This shows that $r_1 \neq 0$, so that by assumption $r_1 < 0$. Therefore, since d_2 divides r_1, we get

$$-d_2 \geq r_1 \geq c_1 r_1 \geq c_0 r_0 + c_1 r_1 = -d_2.$$

4 This gives $-d_2 = c_1 r_1$, so that r_1 divides d_2. Therefore r_1 divides r_0.

(1.9) PROPOSITION. Let p be a positive integer and let $(u_1, \ldots, u_p)$ be a p-tuple of positive integers such that u_i divides u_{i+1} for $1 \leq i \leq p-1$. Let $a_1, \ldots, a_p, b_1, \ldots, b_p$ be non-negative integers such that

$$a_i < u_{i+1}/u_i \text{ and } b_i < u_{i+1}/u_i \text{ for } 1 \leq i \leq p-1. \tag{1.9.1}$$

If

$$\sum_{i=1}^{p} a_i u_i = \sum_{i=1}^{p} b_i u_i \tag{1.9.2}$$

then $a_i = b_i$ for every i, $1 \leq i \leq p$.

Proof. Let $e = p - 1$ and let $r = (r_0, \ldots, r_e)$, where $r_1 = u_{e+1-i}$ for $0 \leq i \leq e$. Then $d_r(r) = u_{e+2-i}$ for $1 \leq i \leq e+1$. Therefore $n_i(r) = u_{e+2-i}/u_{e+1-i}$ for $1 \leq i \leq e$. Let $a'_i = a_{e+1-i}, b'_i = b_{e+1-i}$ for $0 \leq i \leq e$. Then the equality (1.9.2) takes the form

$$\sum_{i=0}^{e} a'_i r_i = \sum_{i=0}^{e} b'_1 r_i$$

and conditions (1.9.1) take the form

$$a'_i < n_i(r) \textit{ and } b'_i < n_i(r)$$

for $1 \leq i \leq e$. Moreover, we have $a'_0 \in \mathbb{Z}^+$ and $b'_0 \in \mathbb{Z}^+$. Now the assertion follows from Proposition (1.5) by taking $\Gamma = \mathbb{Z}^+$. □

2 *G*-Adic Expansion of a Polynomial

(2.1)

5 Let R be a ring (commutative, with unity) and let $R[Y]$ be the poly nomial ring in one variable Y over R. For $F \in R[Y]$, we write $\deg F$ for its Y-degree. We use the convention that $\deg 0 = -\infty$.

(2.2)

Let p be a positive integer and let $G = (G_1, \ldots, G_p)$ be a p-tuple of elements of $R[Y]$ such that the following three conditions are satisfied:

(i) G_i is monic in Y and $\deg G_i > 0$ for every i, $1 \leq i \leq p$.

(ii) $\deg G_i$ divides $\deg G_{i+1}$ for every i, $1 \leq i \leq p-1$.

(iii) $\deg G_1 = 1$.

We put $u_i(G) = \deg(G_i)$ for $1 \leq i \leq p$, and $u_{p+1}(G) = \infty$. We then define $n_i(G) = u_{i+1}(G)/u_i(G)$ for $1 \leq i \leq p$. Note that $n_p(G) = \infty$ and $n_i(G)$ is a positive integer for $1 \leq i \leq p-1$. Let

$$A(G) = \left\{ a = (a_1, \ldots, a_p) \in \mathbb{Z}^p \middle| 0 \leq a_i < n_i(G) \text{ for } 1 \leq i \leq p \right\}.$$

For $a \in A(G)$, we put $G^a = G_1^{a_1} \cdots G_p^{a_p}$.

(2.3) DEFINITION. An element $F \in R[Y]$ is called a strict polynomial in G if F has an expression of the form

$$F = \sum_{a \in A(G)} F_a G^a$$

with $F_a \in R$ for every a and $G_a = 0$ for almost all a. We write $R[G^A]$ for the set of strict polynomials in G. Note that $R[G^A]$ is the R-submodule of $R[Y]$ generated by the set $G^A = \left\{ G^a \middle| a \in A(G) \right\}$.

(2.4) LEMMA. *Let $a, b \in A(G)$. If $a \neq b$ then* $\deg G^a \neq \deg G^b$.

Proof. This is immediate from Proposition (1.9). For, by taking $u_i = u_i(G)$, $1 \leq i \leq p$, we have

$$\deg G^a = \sum_{i=1}^{p} a_i u_i, \deg G^b = \sum_{i=1}^{p} b_i u_i.$$

□

(2.5) COROLLARY. Let 6

$$F = \sum_{a \in A(G)} F_a G^a$$

be a strict polynomial in G. Then

$$\deg F = \sup_{a \in a(G)} \deg(F_a G^a).$$

In particular, if $G = 0$ then $F_a = 0$ for all $a \in A(G)$.

(2.6) COROLLARY. $R[G^A]$ is a free R-module with G^A as a free basis.

(2.7) DEFINITION. Let $F \in R[G^A]$. The expression

$$F = \sum_{a \in A(G)} F_a G^a,$$

which is unique by Corollary (2.6), is called the G-adic expansion of F.

(2.8) DEFINITION. For $F \in R[G^A]$, we define

$$\operatorname{Supp}_G(F) = \left\{ a \in A(G) \middle| F_a \neq 0 \right\}.$$

(2.9) COROLLARY. Let F be a non-zero element of $R[G^A]$. Then

$$\deg F = \sup_{a \in \operatorname{Supp}_G(F)} \deg G^a.$$

More precisely, there exists a unique element $a \in \operatorname{Supp}_G(F)$ such that

$$\deg F = \deg G^a > \deg G^b$$

for every *bin* $\operatorname{Supp}_G(F)$, $b \neq a$.

Proof. Immediate from Lemma (2.4). ◻

7 **(2.10) LEMMA.** *Let e be an integer, $1 \leq e \leq p$, and let $a_1, \ldots, a_e$ be non-negative integers such that $a_i < n_i(G)$ for $1 \leq i \leq e$. Then* $\sum_{i=1}^{e} a_i u_i(G) < u_{e+1}(G)$.

Proof. We use induction on e. If $e = 1$ then $a_1 < n_1(G)$ implies that $a_1 u_1(G) < n_1(G) u_1(G) = u_2(G)$. Now, suppose $e \geq 2$. By induction hypothesis, we have $\sum_{i=1}^{e-1} a_i u_i(G) < u_e(G)$. Therefore

$$\begin{aligned}\sum_{i=1}^{e} a_i u_i(G) &< u_e(G) + a_e u_e(G) \\ &= (1 + a_e) u_e(G) \\ &= n_e(G) u_e(G) \quad (\text{since } a_e < n_e(G)) \\ &= u_{e+1}(G).\end{aligned}$$

□

(2.11) LEMMA. *Let e be an integer, $1 \leq e \leq p$. Let $a = (a_1, \ldots, a_p)$ be an element of $A(G)$ such that $a_e \neq 0$ and $a_i = 0$ for $e + 1 \leq i \leq p$. Then $u_e(G) \leq \deg G^a < u_{e+1}(G)$.*

Proof. we have $\deg G^a = \sum_{i=1}^{p} a_i u_i(G) = \sum_{i=1}^{e} a_i u_i(G)$. Therefore, since $a_e > 0$ and $a_i \geq 0$ for all i, we get $i_e(G) \leq \deg G^a$. The inequality $\deg G^a < u_{e+1}(G)$ follows from Lemma (2.10). □

(2.12) LEMMA. *Let F be an element of $R[G^A]$ such that $F \notin R$. Let*

$$e = \sup \left\{ i \middle| 1 \leq i \leq p, \exists\, a \in \mathrm{Supp}_G(F) \text{ with } a_i \neq 0 \right\}.$$

Then $u_e(G) \leq \deg F < u_{e+1}(G)$.

Proof. By Corollary (2.9) there exists $a \in \mathrm{Supp}_G(F)$ such that

$$\deg F \deg G^a \geq G^b \tag{2.12.1}$$

for every $b \in \mathrm{Supp}_G(F)$. Since $F \notin R$, we have $a \neq 0$. For $b \in \mathrm{Supp}_G(F)$, $b \neq 0$, let

$$e_b = \sup \left\{ i \middle| 1 \leq i \leq p, b_i \neq 0 \right\}.$$

□

Then Lemma (2.11) $u_{e_b}(G) \leq \deg G^b < u_{e_b} + (G)$. Therefore it 8
follows from (2.12.1) that we have

$$(2.12.2) \qquad u_{e_a}(G) \leq \deg F < u_{e_a+1}(G)$$

and that $u_{e_b}(G) u_{e_a+1}(G)$ for every $b \in \mathrm{Supp}_G(F)$, $b \neq 0$. This last inequality shows that $e_b \leq e_a$, so that we get

$$e \sup \left\{ e_b \middle| b \in \mathrm{Supp}_G(F), b \neq 0 \right\} = e_a.$$

Now the lemma follows from (2.12.2).

(2.13) THEOREM. $R[G^A] = R[Y]$.

Proof. We have to show that every element F of $R[Y]$ belongs to $R[G^A]$. We do this by induction on $\deg F$. The assertion being clear for $\deg F \leq 0$, let us assume that $\deg F \geq 1$. Since $u_1(G) = 1$ and $u_{p+1}(G) = \infty$, there exists a unique integer e, $1 \leq e \leq p$, such that

$$u_e(G) \leq \deg F < u_{e+1}(G).$$

□

Then there exists a unique positive integer b_e such that

$$(2.13.2) \qquad b_e u_e(G) \leq \deg F < (b_e + 1) u_e(G).$$

If follows from (8) that we have

$$(2.13.3) \qquad b_e < n_e(G)$$

Since G_e and hence $G_e^{b_e}$ is monic, there exist $Q, P \in R[Y]$ such that

$$(2.13.4) \qquad F = Q G_e^{b_e} + P$$

9 and

$$(2.13.5) \qquad \deg P < \deg G_e^{b_e} = b_e u_e(G) \leq \deg F.$$

By induction hypothesis, $P \in R[G^A]$. Therefore it is enough to prove that $QG_e^{b_e} \in R[G^A]$. From (2.13.4) and (2.13.5) we see that $\deg F = \deg(QG_e^{b_e})$, which shows that we have

$$\deg Q = \deg F - b_e u_e(G) < \deg F. \tag{2.13.6}$$

Therefore, by induction hypothesis, $Q \in R[G^A]$. Writing

$$Q = \sum_{a \in A(G)} Q_a G^a, \quad Q_a \in R,$$

we get

$$QG_e^{b_e} = \sum_{a \in A(G)} Q_a G^a G_e^{b_e}.$$

It is therefore enough to show that

$$a + (0, \ldots, b_e, \ldots, 0) \in A(G)$$

for every $a \in \mathrm{Supp}_G(Q)$. Since $b_e < u_e(G)$ by (2.13.3), it is enough to prove that $a_e = 0$ for every $a \in \mathrm{Supp}_G(Q)$. This last assertion is clear if $Q \in R$. Assume therefore that $Q \notin R$. Then, since

$$\begin{aligned} \deg Q &= \deg F - b_e u_e(G) && \text{(by (2.13.6))} \\ &< u_e(G) && \text{(by (2.13.2)),} \end{aligned}$$

we see by Lemma (2.12) that $a_e = 0$ for every $a \in \mathrm{Supp}_G(Q)$. This completes the proof of the theorem.

(2.14) COROLLARY. Every element of $R[Y]$ has a unique G-adic expansion.

Proof. Clear from Theorem (2.13) and Corollary (2.6) □ 10

3 Tschirnhausen Operator

We preserve the notation of (2.1)

(3.1)

Let $g \in R[Y]$ be a monic polynomial of positive degree. Let $G_1 = Y$, $G_2 = g$. Then the conditions (i) - (iii) of (2.2) are satisfied by $G = (G_1, G_2)$ with $p = 2$, and we note that we have $n_1(G) = \deg g$, $n_2(G) = \infty$ and

$$A(G) = \left\{ a = (a_1, a_2) \in \mathbb{Z}^+ \times \mathbb{Z}^+ \middle| < \deg g \right\}.$$

By Corollary (2.14) every element of $R[Y]$ has a unique $G = (Y, g)-$ adic expansion. Let $f \in R[Y]$ and let

(3.1.1) $$f = \sum_{a \in A(G)} f_a Y^{a_1} g^{a_2}$$

be its G-adic expansion. For $i \in \mathbb{Z}^+$, let

$$C_f^{(i)}(g) = \sum_{\substack{a \in A(G) \\ a_2 = i}} f_a Y^{a_1}$$

Then we can rewrite (3.1.1) in the form

(3.1.2) $$f = \sum_{i=0}^{\infty} C_f^{(i)}(g) g^i$$

with $C_f^{(i)}(g) \in R[Y]$, $\deg C_f^{(i)}(g) < \deg g$ and $C_f^{(i)}(g) = 0$ for almost all i. The expression (3.1.2) is called the g-adic expansion of f. It follows from Corollary (2.14) that every element f of $R[Y]$ has a unique g-adic expansion. In particular, if $f = \sum_{i=0}^{\infty} C_i g^i$ with $C_i \in R[Y]$, $\deg C_i < \deg g$ and $C_i = 0$ for almost all i, then $C_i = C_f^{(i)}(g)$ for every i and $f = \sum_{i=0}^{\infty} C_i g^i$ is the g-adic expansion of f.

11 **(3.2) LEMMA.** *Let $f \in R[Y]$. Suppose $f = \sum_{i=0}^{e} C_i g^i$, where e is a nonnegative integer, $C_i \in R[Y]$ with $\deg C_i < \deg g$ for $0 \le i \le e$, and $C_e \neq 0$. Then $\deg f = e \deg g + \deg C_e$. In particular, we have*

$$e \deg g \le \deg f < (e+1) \deg g.$$

Proof. For every i, $0 \le i \le e-1$, we have

$$\begin{aligned}
\deg(C_i g^i) &= i \deg g + \deg C_i \\
&\le (e-1)\deg g + \deg C_i \\
&< e \deg g \qquad (\text{since } \deg C_i < \deg g) \\
&\le e \deg g + \deg C_e \qquad (\text{since } C_e \ne 0) \\
&= \deg(C_e g^e).
\end{aligned}$$

This shows that $\deg f = e \deg g + \deg C_e$. The asserted inequalities now follow from the fact that $0 \le \deg C_e < \deg g$. □

(3.3) COROLLARY. Let f be an element of $R[Y]$ such that f is monic and $\deg f = d \deg g$ for some non-negative integer d. Then

$$f = g^d + \sum_{i=0}^{d-1} C_i^{(i)}(g) g^i.$$

Proof. Since $\deg f = d \deg g$, Lemma (3.2) shows that

$$f = \sum_{i=0}^{d} C_f^{(i)}(g) g^i$$

with $\deg C_f^{(d)}(g) = 0$. This means that $C = C_f^{(d)}(g) \in R$. By Lemma (3.2) again, we have

$$\deg(f - C_g^d) = \deg\left(\sum_{i=0}^{d-1} C_f^{(i)}(g) g^i\right) < d \deg g. \tag{3.3.1}$$

Since $\deg(Cg^d) = d \deg g = \deg f$ and since both f and g are monic, 12
it follows from (3.3.1) that $C = 1$. □

(3.4) DEFINITION. Let d be a positive integer. Let $g \in R[Y]$ be a monic polynomial of positive degree and let $f \in R[Y]$ be a monic polynomial of degree $d \deg g$. Then we have

$$f = g^d + \sum_{i=0}^{d-1} C_f^{(i)}(g) g^i \tag{3.4.1}$$

by Corollary (3.3). We call $C_f^{(d-1)}(g)$ the *Tschirnhausen coefficient* in the g-adic expansion of f and denote it simply by $C_f(g)$. *If d is a unit in R* then the *Tschirnhausen transform* of g with respect to f, denoted $\tau_f(g)$, is defined to be

$$\tau_f(g) = g + d^{-1}C_f(g).$$

We call τ_f the *Tschirnhausen operator* with respect to f. Note that $\deg C_f(g) < \deg g$ and $\tau_f(g) \in R[Y]$ is monic with $\deg \tau_f(g) = \deg g$.

In (3.5) to (3.7) below, we preserve the notation of (3.4). We assume, moreover, that d is a unit in R.

(3.5) LEMMA. *If $C_f(g) \neq 0$ then*

$$\deg C_f(g) = \deg(f - g^d) - (d-1)\deg g.$$

Proof. By 3.4.1 we have

$$f - g^d = \sum_{i=0}^{d-1} C_f^{(i)}(g)g^i.$$

Since $\deg C_f^{(i)}(g) < \deg g$ for every i, the above expression is the g-adic expansion of $f - g^d$. Therefore, since $C_f^{(d-1)}(g) = C_f(g) \neq 0$, we see by Lemma (3.2) that

$$\deg(f - g^d) = (d-1)\deg g + \deg C_f(g).$$

$\square$

(3.6) PROPOSITION.

13 (i) If $C_f(g) = 0$ then $C_f(\tau_f(g)) = 0$.

(ii) If $C_f(g) \neq 0$ then $\deg C_f(\tau_f(g)) < \deg C_f(g)$.

Proof.

(i) is clear, since $\tau_f(g) = g$ if $C_f(g) = 0$.

(ii) Let $h = \tau_f(g) = g + d^{-1}C_f(g)$. Then we have

$$h^d = g^d + C_f(g)g^{d-1} + k, \tag{3.6.1}$$

where

$$k = \sum_{i=2}^{d} \binom{d}{i} d^{-i} C_f(g)^i g^{d-i}.$$

□

Let $c = \deg C_f(g)$. Then $0 \leq c < \deg g$. Therefore we have

$$\deg k \leq 2c + (d-2)\deg g < c + (d-1)\deg g.$$

Now, from (3.6.1) we get

$$\begin{aligned} f - h^d &= f - g^d - C_f(g)g^{d-1} - k \\ &= \sum_{i=0}^{d-2} C_f^{(i)}(g)g^i - k \qquad \text{(by (3.4.1))}. \end{aligned}$$

Since

$$\deg\left(\sum_{i=0}^{d-2} C_f^{(i)}(g)g^i\right) < (d-1)\deg g \leq c + (d-1)\deg g$$

by Lemma (3.2) and since $\deg k < c + (d-1)\deg g$, we get

$$\deg(f - h^d) < (d-1)\deg h + c.$$

Therefore if $C_f(h) \neq 0$ then $\deg C_f(h) < c$ by Lemma (3.5). If $C_f(h) = 0$ then $\deg C_f(h) = -\infty < c$. **14**

(3.7) COROLLARY. $C_f((\tau_f)^j(g)) = 0$ for all $j \geq \deg g$.

Proof. This is clear from Proposition (3.6), since $\deg C_f(g) < \deg(g)$. □

4 Approximate Roots

(4.1)

Let R be a ring (commutative, with unity) and let $R[y]$ be the polynomial ring in one variable Y over R.

(4.2) PROPOSITION. Let n, d be positive integers such that d divides n. Let $f \in R[Y]$ be a monic polynomial of degree n. Let $g \in R[Y]$ be a monic polynomial. Then the following two conditions are equivalent:

(i) $\deg(f - g^d) < n - (n/d)$.

(ii) $\deg g = n/d$ and $C_f(g) = 0$.

Proof. (i) $\Rightarrow$ (ii). Since g is monic, it is clear from (i) that $\deg g = n/d$. Therefore we get $\deg(f - g^d) < (d-1)\deg g$. this shows (by Lemma (3.2)) that the g-adic expansion of $f - g^d$ has the form

$$f - g^d = \sum_{i=0}^{d-2} C^{(i)}_{f-g^d}(g)g^i.$$

□

It follows that

$$f = g^d + \sum_{i=0}^{d-2} C^{(i)}_{f-g^d}(g)g^i$$

is the g-adic expansion of f and $C_f(g) = C_f^{(d-1)}(g) = 0$.

(ii) $\Rightarrow$ (i). Since $\deg g = n/d$, we have $\deg f = d \deg g$. Therefore, since $C_f(g) = 0$, we get

$$f = g^d + \sum_{i=0}^{d-2} C^{(i)}_f(g)g^i$$

15 by Corollary (3.3). Therefore

$$\deg(f - g^d) = \deg\left(\sum_{i=0}^{d-2} C^{(i)}_f(g)g^i\right)$$

$$
\begin{aligned}
&< (d-1)\deg g \qquad\qquad \text{(by Lemma (3.2))}\\
&= n - (n/d).
\end{aligned}
$$

(4.3) DEFINITION. Let $f \in R[Y]$ be a monic polynomial of positive degree n. Let d be a positive integer such that d divides n. An element g of $R[Y]$ is called an *approximate dth root of f (with respect to Y)* if g is monic and satisfies the equivalent conditions (i) and (ii) of Proposition (4.2).

(4.4) THEOREM. *Let $f \in R[Y]$ be a monic polynomial of positive degree n. Let d be a positive integer such that d divides n. Assume that d is a unit in R. Then there exists a unique approximate dth root of f with respect to Y.*

Proof. Let $g = (\tau_f)^{n/d}(Y^{n/d})$. Then g is monic of degree n/d and $C_f(g) = 0$ by Corollary (3.7). This proves the existence of an approximate dth root of f with respect to Y. □

Now, suppose g_1, g_2 are approximate dth roots of f with respect to Y. Then

$$\deg(f - g_1^d) < n - (n/d) \text{ and } \deg(f - g_2^d) < n - (n/d).$$

Therefore

$$\deg(g_1^d - g_2^d) < n - (n/d). \tag{4.4.1}$$

Now, we have

$$g_1^d - g_2^d = (g_1 - g_2) \sum_{i+j=d-1} g_1^i g_2^j. \tag{4.4.2}$$

Since both g_1 and g_2 are monic of $\deg n/d$, $g_1^i g_2^j$ is monic of degree $(d-1)(n/d)$ for $i + j = d - 1$. Therefore $d^{-1} \sum_{i+j=d-1} g_1^i g_2^i$ is monic with 16

$$\deg\left(d^{-1} \sum_{i+j=d-1} g_1^i g_2^j\right) = (d-1)(n/d) = n - (n/d). \tag{4.4.3}$$

It follows from 4.4.1, 4.4.2 and 4.4.3 that $g_1 - g_2 = 0$.

(4.5) NOTATION. We denote the approximate dth root of f with respect to Y by $App_Y^d(f)$.

(4.6) COROLLARY. Let $f \in R[Y]$ be a monic polynomial of positive degree n. Let d be a positive integer such that d divides n. Assume that d is a unit in R. Let $g \in R[Y]$ be any monic polynomial of degree n/d. Then

$$(\tau_f)^j(g) = App_Y^d(f)$$

for all $j \geq n/d$.

Proof. Immediate from Corollary (3.7). □

Let S be a ring (commutative, with unity) and let $\sigma : R \to S$ be a (unitary) ring homomorphism. Denote again by σ on R. Let $f \in R[Y]$ be a monic polynomial of positive degree n. Then $\sigma(f) \in S[Y]$ is also a monic polynomial of degree n. Let d be a positive integer such that d divides n. Assume that d is a unit in R. Then d is also a unit in S, and we have

(4.7) PROPOSITION. $App_Y^d(\sigma(f)) = \sigma(App_Y^d(f))$.

Proof. Put $g = App_Y^d(f)$. Then $\sigma(g)$ is monic of degree n/d. Moreover, we have $\sigma(f) - (\sigma(g))^d = \sigma(f - g^d)$. Therefore

$$\deg(\sigma(f) - (\sigma(g))^d) < n - (n/d).$$

This shows that $\sigma(g) = App_Y^d(\sigma(f))$. □

Chapter 2

Characteristic Sequences of a Meromorphic Curve

5 Newton-puiseux Expansion

(5.1) NOTATION. Let k be a field. If n is a positive integer we denote 17
by $\mu_n(k)$ (or simply by μ_n if no confusion is likely) the group of nth roots of unity in k. We use the letters X, Y, t to denote indeterminates. As usual, $k[[t]]$ denotes the ring of formal power series in t over k. We denote by $k((t))$ the quotient field of $k[[t]]$. Recall that every element a of $k((t))$ has a unique expression of the form $a = \sum_{j \in \mathbb{Z}} a_j t^j$ with $a_j \in k$ for every j and $a_j = 0$ for $j \ll 0$. We denote by $\mathrm{ord}_t a$ the t-order of a. Recall that if $a \neq 0$ then writing $a = \sum a_j t^i$ with $a_j \in k$, we have

$$\mathrm{ord}_t\, a = \inf \left\{ j \in \mathbb{Z} \middle| a_j \neq 0 \right\}.$$

If $a = 0$ then $\mathrm{ord}_t\, a = \infty$. If $a = \sum a_j t^j \in k((t))$ (with $a_j \in k$) we define

$$\mathrm{Supp}_t\, a = \left\{ j \in \mathbb{Z} \middle| a_j \neq 0 \right\}.$$

If R is a ring and $f \in R[Y]$, we write $\deg_Y f$ (or simply $\deg f$ if no confusion is likely) for the Y-degree of f. We use the convention: $\deg 0 = -\infty$.

(5.2) HENSEL'S LEMMA.

Let $f = f(X,Y)$ be an element of $k[[X]][Y]$ such that f is monic in Y. Suppose $f(0,Y) = \overline{g}\overline{h}$, where $\overline{g}, \overline{h}$ are elements of $k[Y]$, both monic in Y, and $g.c.d.\ (\overline{g},\ \overline{h}) = 1$. Then there exist elements $g = g(X,Y)$, $h = h(X,Y)$ of $k[[X]][Y]$, both monic in Y, such that $g(0,Y) = \overline{g}$, $h(0,Y) = \overline{h}$ and $f = gh$.

Proof. Let $n = \deg_Y f$. we can write $f = \sum_{q=0}^{\infty} f_q X^q$ with $f_q \in k[Y]$ for every q. Then f_0 is monic in Y of degree n and $\deg f_q < n$ for $q \geq 1$. Let
18 $r = \deg \overline{g}$, $s = \deg \overline{h}$. Then $r + s = n$. Now, in order to prove the lemma. it is enough to find, for every $i \in \mathbb{Z}^+$, elements g_i, h_i of $k[Y]$ such that

1. $g_0 = \overline{g}$ and $h_0 = \overline{h}$.
2. $\deg g_i < r$ and $\deg h_i < s$ for all $i \geq 1$.
3. $f_q = \sum_{i=0}^{q} g_i h_{q-i}$ for all $q \geq 0$.

□

For, then $g = \sum_{i=0}^{\infty} g_i X^i, h = \sum_{i=0}^{\infty} h_i X^i$ would meet the requirements of the lemma.

We define g_i, h_i by induction on i, these being already defined for $i = 0$ by condition (i). Let q be a positive integer and suppose g_i, h_i are already defined for $i < q$. Let

$$e_q = f_q - \sum_{i=1}^{q-1} g_i h_{q-i}.$$

Then $\deg e_q < n$. Since $g.c.d.\ (g_0, h_0) = 1$, there exist $G_q, H_q \in k[Y]$ such that $e_q = H_q g_0 + G_q h_0$. Let $G_q = g_0 Q + g_q$ with Q, $g_q \in k[Y]$ and $\deg g_q < \deg g_0 = r$. Then $e_q = h_q g_0 + g_q h_0$, where $h_q = H_q + Q h_0$. Since $\deg e_q < n = r + s$, we get $\deg h_q < s$. Now

$$f_q = \sum_{i=0}^{q} g_i h_{q-i}$$

and the lemma is proved.

(5.3) COROLLARY. Let k be an algebraically closed field. Let u be an element of $k((X))$ such that $\operatorname{ord}_X u = 0$. Let n be an integer such that char k does not divide n. Then there exists $v \in k((X))$ such that $u = v^n$.

Proof. Since $\operatorname{ord}_X u = 0$ if and only if $\operatorname{ord}_X u^{-1} = 0$ and since $u = v^n$ if **19**
and only if $u^{-1} = v^{-n}$, we may assume that n is positive. Since $\operatorname{ord}_X u = 0$, we have $u = u(X) \in k[[X]]$ and $u(0) \neq 0$. Let $f(X,Y) = Y^n - u$. Then $f(X,Y) \in k[[X]][Y]$ and $f(0,Y) = Y^n - u(0)$. Since k is algebraically closed, there exist $v_i \in k$, $1 \leq i \leq n$, such that $Y^n - u(0) = \prod_{i=1}^{n}(Y - v_i)$. Since $u(0) \neq 0$ and char k does not divide n, we have $v_i \neq v_j$ for $i \neq j$. Therefore if we let $\overline{g} = Y - v_1$ and $\overline{h} = \prod_{i=2}^{n}(Y - v_i)$ then g.c.d. $(\overline{g}, \overline{h}) = 1$ and $f(0,Y) = \overline{g}\overline{h}$. Therefore by Hensel's Lemma (5.2) there exists an element $g(X,Y)$ in $k[[X]][Y]$ such that $g(X,Y)$ is monic in Y, $g(0,Y) = \overline{g}$ and $g(X,Y)$ divides $f(X,Y)$ in $f(X,Y)$ in $k[[X]][Y]$. From the equality $g(0,Y) = \overline{g} = Y - v_1$ and the fact that $g(X,Y)$ is monic in Y, we get $g(X,Y) = Y - v$ for some $v \in k((X))$. Now $g(x,v) = 0$. Therefore $f(X,v) = 0$. This means that $v^n = u$. □

(5.4) COROLLARY. Let k be an algebraically closed field. Let a be a nonzero element of $k((X))$ and let $n = \operatorname{ord}_X a$. Assume that char k does not divide n. Then there exists $z \in k((X))$ such that:

(i) $a = z^n$.

(ii) $\operatorname{ord}_X z = 1$.

(iii) $k[[z]] = k[[X]]$ and $k((z)) = k((X))$.

Proof. (iii) is immediate from (ii), and (ii) is immediate from (i). Therefore it is enough to prove (i). Write $a = X^n u$ with $u \in k((X))$. Then $\operatorname{ord}_X u = 0$. Therefore by Corollary (5.3) there exists $v \in k((X)))$ such that $u = v^n$. Let $z = Xv$. Then $a = z^n$. □

(5.5) NEWTON'S LEMMA

Let k be an algebraically closed field. Let $f(X,Y)$ be a non-zero element
20 of $k((X))[Y]$. Assume that char k does not divide $\deg_Y f(X,Y)$. Then there exists a positive integer m and an element $y(t) \in k((t))$ such that $f(t^m, y(t)) = 0$.

Proof. Without loss of generality, we may assume that $f(X,Y)$ is irreducible. Let $N = \deg_Y f(X,Y)$. We shall prove the result by induction on n. If $n = 1$ then the assertion is clear with $m = 1$. Assume therefore that $n \geq 2$. Write $f(X,Y) = \sum_{i=0}^{n} f_i Y^{n-i}$ with $f_i = f_i(X) \in k((X))$ for $0 \leq i \leq n$, $f_0 \neq 0$. Now, for the moment, grant the following □

(5.4.1) CLAIM. In order to prove the lemma, we may, without loss of generality, make the following three assumptions:

(i) $f_0 = 1$.

(ii) $f_1 = 0$.

(iii) $f_1 \in k[[X]]$ for every i and $f_i(0) \neq 0$ for some i, $2 \leq i \leq n$.

Then (5.4.1) implies that $f(X,Y) \in k[[X]][Y]$ and we have

$$f(0,Y) = Y^n + f_2(0)Y^{n-2} + \cdots + f_n(0)$$

with $f_i(0) \neq 0$ for some i, $2 \leq i \leq n$. Since char k does not divide n, it follows from the above expression for $f(0,Y)$ that $f(0,Y)$ is not the nth power of an element of $k[Y]$. Therefore, since k is algebraically closed, there exist $\overline{g}, \overline{h} \in k[Y]$, both of them monic in Y of degree less than n, such that $f(0,Y) = \overline{g}\overline{h}$ and g.c.d. $(\overline{g}, \overline{h}) = 1$. It follows by Hensel's Lemma (5.2) that there exist $g(X,Y)$, $h(X,Y) \in k[[X]][Y]$, both of them monic in Y, such that $f(X,Y) = g(X,Y)h(X,Y)$ and $g(0,Y) = \overline{g}$, $h(0,Y) = \overline{h}$. Let $r = \deg_Y g(x,Y) = \deg \overline{g}$, $s = \deg_Y H(X,Y) = \deg \overline{h}$. Then $r < n$, $s < n$ and $r+s = n$. Since char k does not divide n, char k does not divide at least one of r and s, say r. Then, by induction hypothesis, there exists
21 a positive integer m and an element $y(t) \in k((t))$ such that $g(t^m, y(t)) = 0$.

Therefore $f(t^m, y(t)) = 0$, and the lemma is proved modulo the Claim (5.4.1).

Proof of (5.4.1).

(i) Since $f_0 \neq 0$, we may replace $f(X, Y)$ by $f_0^{-1} f(X, Y)$.

(ii) Assume (i), i.e. $f_0 = 1$. Let $Z = Y + n^{-1} f_1$. Then $f(X, Y) = f(X, Z - n^{-1} f_1) = g(X, Z)$, say. It is clear that $g(X, Z)$ has the form

$$g(X, Z) = Z^n + g_2 Z^{n-2} + \cdots + g_n$$

with $g_i \in k((X))$, $2 \leq i \leq n$. If m is a positive integer and $y(t)$ is an element of $k((t))$ such that $g(t^m, y(t)) = 0$ then we have $f(t^m, z(t)) = 0$, where $z(t) = y(t) - n^{-1} f_1(t^m)$.

(iii) Assume that f already satisfies (i) and (ii). Since $f(X, Y)$ is irreducible and $n \geq 2$, there exists $i, 2 \leq i \leq n$, such that $f_i \neq 0$. Let $u_i = \operatorname{ord}_X f_i$ and let

$$u \inf \left\{ u_i/i \middle| 2 \leq i \leq n \right\}.$$

Let r be an integer, $2 \leq r \leq n$, such that $u = u_r/r$. Let W be an indeterminate and let $Z = W^{-u_r} Y$. Let $g(W, Z) = W^{-nu_r} f(W^r, Y) = Z^n + \sum_{i=2}^{n} g_i Z^{n-i}$, where $g_i = g_i(W) = f_i(W^r) W^{-iu_r}$. Now $\operatorname{ord}_W g_i = ru_i - iu_r \geq r\,iu - iu_r = 0$ with equality for $i = r$. This means that $g_i \in k[[W]]$ for all $i, 2 \leq i \leq n$, and $g_r(0) \neq 0$. Now, if m is a positive integer and $y(t)$ is an element of $k((t))$ such that $g(t^m, y(t)) = 0$ then we have

$$0 = g(t^m, y(t)) = t^{-mnu_r} f(t^{mu_r} y(t)),$$

so that $f(t^{mr}, t^{mu_r} y(t)) = 0$.

(5.6) NOTATION. Let m be a positive integer. We write $k((t^m))$ for the set of those $a \in k((t))$ for which $\operatorname{Supp}_t a \subset m\mathbb{Z}$. Note that $k((t^m))$ is a subfield of $k((t))$. **22**

(5.7) LEMMA. *Let m be a positive integer. Then $k((t))/k((t^m))$ is a finite algebraic extension of degree m.*

Proof. The set $\{1, t, \ldots t^{m-1}\}$ is clearly a $k((t^m))$-vector space basis of $k((t))$. $\square$

(5.8) DEFINITION. Let m be a positive integer and let $y = y(t)$ be an element of $k((t))$. By Lemma (5.7), y is algebraic over $k((t^m))$. Let $f(t^m, Y)$ in $k((t^m))[Y]$ be the minimal monic polynomial of y over $k((t^m))$. Put $f = f(X, Y)$. Then $f \in k((X))[Y]$. By abuse of language, we shall call f the *minimal monic polynomial* of y over $k((t^m))$.

(5.9) LEMMA. *Let m be a positive integer and let $y = y(t)$ be an element of $k((t))$. Let $f = f(X, Y) \in k((X))[Y]$ be the minimal monic polynomial of y over $k((t^m))$. Then we have:*

(i) *f is monic in Y and f is irreducible in $k((X))[Y]$.*

(ii) $f(t^m, y) = 0$.

(iii) *If $g = g(X, Y)$ is any element of $k((X))[Y]$ such that $g(t^m, y) = 0$ then f divides g in $k((X))[Y]$.*

(iv) $\deg_Y f = [k((t^m))(y) : k((t^m))]$.

(v) *$\deg_Y f$ divides m.*

Proof. (i), (ii), (iii) and (iv) are clear from Definition (5.8). To prove (v), we note that since $y \in k((t))$, we have

$$\begin{aligned} m &= [k((t)) : k((t^m))] \\ &= [k((t)) : k((t^m))(y)][k((t^m))(y) : k((t^m))] \\ &= [k((t)) : k((t^m))(y)] \deg_Y f. \end{aligned}$$

$\square$

23 **(5.10) LEMMA.** *Let m be a positive integer and let $y = y(t)$ be an element of $k((t))$. Let $f(X, Y) \in k((X))[Y]$ be the minimal monic polynomial of y over $k((t^m))$. Assume that char k does not divide m and that*

$$g.c.d.\ (\{m\} \cup \mathrm{Supp}_t\, y) = 1.$$

Then we have:

(i) $f(t^m, Y) = \prod_{w \in \mu_m(\overline{k})} (Y - y(wt))$, where $\overline{k}$ is the algebraic closure of k. Moreover, the m roots $y(wt)$, $w \in \mu_m(\overline{k})$, of $f(t^m, Y) = 0$ are distinct.

(ii) $[k((t^m))(y) : k((t^m))] = \deg_Y f(X, Y) = m$.

Proof. By Lemma (5.9) (v) we have $\deg_Y f(X, Y) \le m$. Therefore it is enough to prove the following two statements:

(1) $f(t^m, y(wt)) = 0$ for every $w \in \mu_m(\overline{k})$.

(2) If $w_1, w_2 \in \mu_m(\overline{k})$, $w_1 \ne w_2$, then $y(w_1 t) \ne y(w_2 t)$.

For, given (1) and (2), $f(t^m, Y)$ will have at least m distinct roots $y(wt)$, $w \in \mu_m(\overline{k})$., Since $\deg_Y f(X, Y) \le m$ and $f(X, Y)$ is monic in Y, both (i) and (ii) would be proved. □

Proof of (1). Since $w^m = 1$, substituting wt for t in the equality $f(t^m, y(t)) = 0$, we get $f(t^m, y(wt)) = 0$.

Proof of (2). Write $y = \sum y_j t^j$ with $y_j \in k$. Then $y(wt) = \sum y_j w^j t^j$. Therefore if $y(w_1 t) = y(w_2 t)$ then we have $w_1^j = w_2^j$ for every $j \in \mathrm{Supp}_t\, y$. Writing $w = w_1 w_2^{-1}$, we get get $w^j = 1$ for every $j \in$ Supp $j \in \mathrm{Supp}_t\, y$. **24**
Since also $w^m = 1$ and

$$\text{g.c.d.} \ (\{m\} \cup \mathrm{Supp}_t\, y) = 1,$$

we get $w = 1$. This means that $w_1 = w_2$.

(5.11) REMARK. A more general version of the above lemma appears in Proposition (5.16).

(5.12) LEMMA. *Let $p = char\ k$. Let $f = f(X, Y)$ be an irreducible element of $k((X))[Y]$ such that $f \notin k((X))[Y^p]$. Let m be a positive integer and let $y = y(t)$ be an element of $k((t))$ such that $f(t^m, y) = 0$. If p divides m then $y \in k((t^p))$.*

Proof. Write $y = \sum y_j t^j$ with $y_j \in k$. Suppose $y \notin k((t^p))$. Then, since $y^p = \sum y_j^p t^{jp} \in k((t^p))$, the minimal monic polynomial of y over $k((t^p))$ is $g(X,Y) = Y^p - z(X)$, where $z(X) = \sum y_j^p X^j$. Note that $g(t^p, Y) = Y^p - z(t^p) = (Y-y)^p$. Let $m = pr$ and let $h(X,Y) = f(X^r, Y)$. Then $h(t^p, y) = f(t^m, y) = 0$. Therefore $g(X,Y)$ divides $h(X,Y)$ in $k((X))[Y]$, so that $g(t^p, Y) = (Y-y)^p$ divides $h(t^p, Y) = f(t^m, Y)$ in $k((t^p))[Y]$. This implies that in the algebraic closure of $k((t^m))y$ occurs as a root of the polynomial $f(t^m, Y)$ in Y with multiplicity at least p. But this is a contradiction, since $f(t^m, Y)$, being irreducible in $k((t^m))[Y]$ and being not an element of $k((t^m))[Y^p]$, is a separable polynomial over $k((t^m))$. This contradiction proves that $y \in k((t^p))$. $\square$

(5.13) LEMMA. *Let k be an algebraically closed field. Let $f = f(X,Y)$ be an irreducible element of $k((X))[Y]$ such that f is monic in Y and char k does not divide $\deg_Y f$. Then there exists an element $y(t)$ of $k((t))$ and a positive integer m such that char k does not divide m and $f(t^m, y(t)) = 0$.*

25 *Proof.* By Newton's Lemma (5.5) thee exists a positive integer m and an element $y(t)$ of $k((t))$ such that $f(t^m, y(t)) = 0$. Let us choose m to be the least positive integer for which there exists an element $y(t)$ of $k((t))$ with $f(t^m, y(t)) = 0$. We then claim that char k does not divide m. For, let $p =$ char k and suppose p divides m. Then by Lemma (5.12) $y(t) \in k((t^p))$. Therefore there exists $z(t) \in k((t))$ such that $y(t) = z(t^p)$. Now, we get $0 = f(t^m, y(t)) = f((t^p)^{m/p}, z(t^p))$, which shows that $f(t^{m/p}, z(t)) = 0$. This contradicts the minimality of m. $\square$

(5.14) NEWTON'S THEOREM

Let k be an algebraically closed field. Let $f = f(X,Y)$ be an irreducible element of $k((X))[Y]$ such that f is monic in Y. Let $n = \deg_Y f$, and assume that char k does not divide n. Then there exists an element $y(t)$ of $k((t))$ such that $f(t^n, y(t)) = 0$. Moreover, for any such $y(t)$ we have:

(i) $f(t^n, Y) = \prod_{w \in \mu_k(k)} (Y - y(wt))$.

(ii) The n roots $y(wt)$, $w \in \mu_n(k)$, of $f(t^n, Y) = 0$ are distinct.

(iii) g.c.d. $(\{n\} \cup \mathrm{Supp}_t\, y(wt)) = 1$ for every $w \in \mu_n(k)$.

Proof. By Lemma (5.13) there exists a positive integer m such that

(5.13.2) CLAIM. char k *does not divide* m *and* $f(t^m, y(t)) = 0$ *for some* $y(t) \in k((t))$.

□

Let us assume that m is the smallest positive integer satisfying (5.13.2). Let

$$d = \text{g.c.d. } (\{m\} \cup \mathrm{Supp}_t\, y(t)).$$

We claim that $d = 1$. for, since d divides every $j \in \mathrm{Supp}_t\, y(t)$, there exists $z(t) \in k((t))$ such that $y(t) = z(t^d)$. Now, we have

$$0 = f(t^m, y(t)) = f((t^d)^{m/d}, z(t^d)),$$

which shows that $f(t^{m/d}, z(t)) = 0$. Therefore by the minimality of m 26
we get $d = 1$. Since $f(X, Y)$ is monic in Y and irreducible in $k((X))[Y]$ and since $f(t^m, y(t)) = 0$, f is the minimal monic polynomial of $y(t)$ over $k((t^m))$. Therefore, since $d = 1$, by Lemma (5.10) we get $n = \deg_Y f(X, Y) = m$. Now, (i) and (ii) follow directly from Lemma (5.10). Since, $\mathrm{Supp}_t\, y(wt) = \mathrm{Supp}_t\, y(t)$ for every $w \in \mu_n(k)$, (ii) follows from the fact $d = 1$ proved above.

(5.15) REMARK. With the notation of Theorem (5.14). let $y(t) = \sum y_j t^j$ with $y_j \in k$. If we write $X^{1/n}$ for t then $y(X^{1/n}) = \sum y_j X^{j/n}$ and $f(X, y(X^{1/n})) = 0$. Note that $y(X^{1/n})$ is a power series in X with *fractional* exponents, in fact with exponents in $(1/n)\mathbb{Z}$. The equality $f(X, y(X^{1/n})) = 0$ can thus be interpreted to mean that given an equation $f(X, Y) = 0$ (where $f(X, Y)$ is an irreducible element of $k((X))[Y]$), we can expand Y as a fractional power series in X with exponents in $(1/n)\mathbb{Z}$. We call $y(X^{1/n})$ a *Newton-Puiseux expansion* of Y in fractional powers of X. Note that there are n distinct Newton-Puiseux expansions of Y, given by the n distinct roots $y(wt)$, $w \in \mu_n(k)$.

(5.16) PROPOSITION. Let m be a positive integer such that char k does not divide m, and let $y = y(t)$ be an element of $k((t))$. Let $f(X, Y) \in k((X))[Y]$ be the minimal monic polynomial of y over $k((t^m))$. Let

$$d = \text{g.c.d.}\ \ (\{m\} \cup \operatorname{Supp}_t y).$$

Then

$$(f(t^m, Y))^d = \prod_{w \in \mu_m(\overline{k})} (Y - y(wt)).$$

27 where $\overline{k}$ is the algebraic close of k. In particular, we have

$$[k((t^m))(y) : k((t^m))] = \deg_Y f(X, Y) = m/d.$$

Proof. Since d divides j for every $j \in \operatorname{Supp}_t y(t)$, there exists $z(t) \in k((t))$ such that $y(t) = z(t^d)$. Let $\tau = t^d$. Then $y(t) = z(\tau)$ and clearly we have

$$\text{g.c.d.}\ \ (\{m/d\} \cup \operatorname{Supp}_\tau z(\tau)) = 1.$$

□

Therefore by Lemma (5.10) we have

$$f(\tau^{m/d}, Y) = \prod_{w \in \mu_{m/d}} (Y - z(w\tau)). \tag{5.16.1}$$

where $\mu_{m/d} = \mu_{m/d}(\overline{k})$. Let v be a primitive mth root of unity in $\overline{k}$. Then v^d is a primitive (m/d)th root of unity of $\overline{k}$. Therefore

$$\mu_{m/d} = \left\{ v^{di} \middle| 1 \leq i \leq m/d \right\}$$

and from 5.16.1 we get

$$\begin{aligned} f(t^m, Y) &= \prod_{i=1}^{m/d} (Y - z(v^{di}\tau)) \\ &= \prod_{i=1}^{m/d} (Y - z((v^i t)^d)) \\ &= \prod_{i=1}^{m/d} (Y - y(v^i t)). \end{aligned} \tag{5.16.2}$$

Let $n = m/d$. Since d divides j for every $j \in \operatorname{Supp}_t y(t)$, m divides nj for every $j \in \operatorname{Supp}_t y(t)$. It follows that $y(v^{rn+i}t) = y(v^i t)$ for all integers i, r. Therefore we get

$$\begin{aligned}
\prod_{w \in \mu_m(\bar{k})} (Y - y(wt)) &= \prod_{j=1}^{m} (Y - y(v^j t)) \\
&= \prod_{r=0}^{d-1} \prod_{i=1}^{n} (Y - y(v^{rn+i}t)) = \left(\prod_{i=1}^{n} (Y - y(v^i t)) \right)^d \\
&= (f(t^m, Y))^d \qquad \text{(by 5.16.2).}
\end{aligned}$$

6 Characteristic Sequences

Throughout this section, we shall preserve the notation introduced in **28**
(6.1) below

(6.1)

Let k be an algebraically closed field and let X, Y, t be indeterminates. Let $f = f(X, Y)$ be an irreducible element of $k((X))[Y]$ such that f is monic in Y. We call such an f a *meromorphic curve* over k. Let $n = \deg_Y f$, and assume that char k does not divide n. Then by Newton's Theorem (5.14) there exists an element $y(t) \in k((t))$ such that $f(t^n, y(t)) = 0$ and

$$f(t^n, Y) = \prod_{w \in \mu_n(k)} (Y - y(wt)).$$

Therefore if $z(t)$ is any element of $k((t))$ such that $f(t^n, z(t)) = 0$ then $z(t) = y(wt)$ for some $w \in \mu_n(k)$. In particular, we have $\operatorname{Supp}_t z(t) = \operatorname{Supp}_t y(t)$. Thus the set $\operatorname{Supp}_t y(t)$ depends only on f and not on a root $y(t)$ of $f(t^n, Y) = 0$. Therefore we can make

(6.2) DEFINITION. The *support* of f denoted $\operatorname{Supp}(f)$ is defined by

$$\operatorname{Supp}(f) = \operatorname{Supp}_t y(t)$$

where $y(t)$ is any element of $k((t))$ such that $f(t^n, y(t)) = 0$.

29 **(6.3) CONVENTION.** We extend the notion of divisibility in $\mathbb{Z}$ to the set $\mathbb{Z} \cup \{\infty, -\infty\}$ by postulating that:

(i) ∞ and $-\infty$ divide every element of $\mathbb{Z} \cup \{\infty, -\infty\}$.

(ii) No integer divides ∞ or $-\infty$.

Note that "a divides b" is still a reflexive and transitive relation on $\mathbb{Z} \cup \{\infty, -\infty\}$. If I is a subset of $\mathbb{Z}$ we denote, as usual, by g.c.d. (I) the unique non-negative generator of the ideal of $\mathbb{Z}$ generated by I. If I is a subset of $\mathbb{Z} \cup \{\infty, -\infty\}$ such that $I \not\subset \mathbb{Z}$ then we *define* g.c.d. $(I) = -\infty$. For a subset I of $\mathbb{Z}$ we denote by $\inf(I)$ the infimum of I. As usual, we set $\inf(\phi) = \infty$.

(6.4) DEFINITION. Let J be a subset of $\mathbb{Z}$ bounded below and let ν be a non-zero integer. We define $m_i(\nu, J)$ and $d_{i+1}(\nu, J)$ for every $i \in \mathbb{Z}^+$ by induction on i as follows: $m_0(\nu, J) = \nu, d_1(\nu, J) = |\nu|$, $m_1(\nu, J) = \inf(J)$ and, $i \geq 2$,

$$d_i(\nu, J) = \text{g.c.d. } (d_{i-1}(\nu, J), m_{i-1}(\nu, J)),$$
$$m_i(\nu, J) = \inf \left\{ j \in J \middle| j \not\equiv 0(\text{mod } d_i(\nu, J)) \right\}.$$

Note that we have $d_i(-\nu, J) = d_i(\nu, j)$ for every $i \geq 1$.

(6.5) LEMMA. *With the notation of (6.4), let $J_1 = J$ and, for $i \geq 2$, let*

$$J_i = \left\{ j \in J_1 \middle| j \not\equiv 0(\mathit{mod}\ d_i(\nu, J)) \right\}.$$

Let $d =$ g.c.d. $(\{\nu\} \cup J)$. Then we have:

(i) $d_{i+1}(\nu, J) =$ g.c.d. $(m_0(\nu, J), \ldots, m_i(\nu, J))$ for all $i \geq 0$.

(ii) $d_{i+1}(\nu, J)$ divides $d_i(\nu, J)$ for every $i \geq 1$.

(iii) $J_i \supset J_{i+1}$ and $m_i(\nu, J) \notin J_{i+1}$ for every $i \geq 1$. In particular, if $J_i \neq \phi$ then $J_i \underset{\neq}{\supset} J_{i+1}$ and $m_i(\nu, J) < m_{i+1}(\nu, J)$.

30 (iv) If $i \geq 2$ and $J_i \neq \phi$ then $d_i(\nu, J) > d_{i+1}(\nu, J) \geq d$. If $i \geq 1$ and $j_i = \phi$ then $d_{i+1}(\nu, J) = -\infty$.

Moreover, there exists a unique non-negative integer h such that we have:

(v) $d_1(\nu, J) \geq d_2(\nu, j) > d_3(\nu, J) > \cdots > d_{h+1}(\nu, J) = d$.

(vi) $d_i(\nu, J) = -\infty$ for $i \geq h + 2$.

(vii) $m_i(\nu, J) \in \mathbb{Z}$ for $0 \leq i \leq h$ and $m_i(\nu, J) = \infty$ for $i \geq h + 1$.

(viii) $m_1(\nu, J) < \cdots < m_h(\nu, J) < m_{h+1}(\nu, J) = \infty$.

(ix) $d_i(\nu, J) =$ g.c.d. $(\{\nu\} \cup \{j \in J | j < m_i(\nu, J)\})$ for $1 \leq i \leq h + 1$.

Proof.

(i) Clear from the definition by induction on i.

(ii) Follows from (i).

(iii) Let $i \geq 1$. It follows from (ii) that $J_i \supset J_{i+1}$. Moreover, since $d_{i+1}(\nu, J)$ divides $m_i(\nu, J)$, we have $m_i(\nu, J) \notin J_{i+1}$. If $J_i \neq \phi$ then $m_i(\nu, J) = \inf(J_i)$ belongs to J_i, so that we get $J_i \underset{\neq}{\supset} J_{i+1}$ and $m_i(\nu, J) < m_{i+1}(\nu, J)$.

(iv) Let $i \geq 2$. If $J_i \neq \phi$ then $m_i(\nu, J) \in J_i$, so that $d_i(\nu, J)$ does not divide $m_i(\nu, J)$. This shows that $d_i(\nu, J) > d_{i+1}(\nu, J)$. Moreover, since $J_i \neq \phi$, by (iii) we have $J_p \neq \phi$ for $1 \leq p \leq i$. Therefore $m_p(\nu, J \in J)$ for $1 \leq p \leq i$, so that $d =$ g.c.d. $(\{\nu\} \cup J)$ divides g.c.d. $(m_0(\nu, J), \ldots, m_i(\nu, J)) = d_{i+1}(\nu, J)$. This shows that $d_{i+1} \geq d$. Now, suppose $i \geq 1$ and $J_i = \phi$. Then $m_i(\nu, J) = \inf(J_i) = \infty$. Therefore $d_{i+1}(\nu, J) = -\infty$. This proves (iv).

We now claim that there exists $i \geq 1$ such that $J_i = \phi$. For, if $J_i \neq \phi$ for every i then, by (iv), $\{d_i(\nu, J) | i \geq 2\}$ is a strictly decreasing infinite sequence of integers bounded below by d. This is not possible. Therefore there exists i such that $J_i = \phi$. Let

$$h + 1 = \inf\left\{i \geq 1 \middle| J_i = \phi\right\}.$$

31 Then, since $J_i \supset J_{i+1}$ for every $i \geq 1$, we have $J_i \neq \phi$ for $1 \leq i \leq h$ and $J_i = \phi$ for $i \geq h+1$. This proves (vi), (vii) and (viii) in view of (iii) and (iv).

(v) Since $J_p \neq \phi$ for $1 \leq p \leq h$, we have $m_p(\nu, J) \in J$ for $1 \leq p \leq h$. Therefore d divides $d_{h+1}(\nu, J)$. On the other hand, since $J_{h+1} = \phi$, $d_{h+1}(\nu, J)$ divides j for every $j \in J$. Since $d_{h+1}(\nu, J)$ also divides ν, we see that $d_{h+1}(\nu, J)$ divides d. Therefore we get $d_{h+1}(\nu, J) = d$. Now, (v) follows from (i) and (iv).

(ix) Fix an i, $1 \leq i \leq h+1$. Let

$$J' = \left\{ j \in J \,\middle|\, j < m_i(\nu, J) \right\}$$

and let $d' =$ g.c.d. $(\{\nu\} \cup J')$. If $i = 1$ then $J' = \phi$ and we have $d' = |\nu| = d_i(\nu, J)$. Assume therefore that $2 \leq i \leq h+1$. Since $m_i(\nu, J) = \inf(J_i)$, we have $J' \cap J_i = \phi$. This means that $d_i(\nu, J)$ divides j for every $j \in J'$. Therefore $d_i(\nu, J)$ divides d'. On the other hand, by (viii) $m_p(\nu, J) \in J'$ for $1 \leq p \leq i-1$. Therefore, since $\nu = m_0(\nu, J), d'$ divides

$$\text{g.c.d. } (m_0(\nu, J), \ldots, m_{i-1}(\nu, J)),$$

which is equal to $d_i(\nu, J)$ by (i). Thus we get $d' = d_i(\nu, J)$.

$\square$

(6.6) DEFINITION. Let J be a subset of $\mathbb{Z}$ bounded below and let ν be a non-zero integer. The *m-sequence* of J *with respect to* ν, denoted $m(\nu, J)$, is defined to be

$$m(\nu, J) = (m_0(\nu, J), \ldots, m_h(\nu, J), m_{h+1}(\nu, J)),$$

where $m_i(\nu, J)$ is defined as in Definition (6.4) and where h is the unique non-negative integer of Lemma (6.5). If ν and J are not clear from
32 the context then we shall write $h(\nu, J)$ for h. Note then that $h(-\nu, J) = h(\nu, J)$. Note also that by Lemma (6.5) we have $m_i(\nu, J) \in \mathbb{Z}$ for $0 \leq i \leq h$ and $m_{h+1}(\nu, J) = \infty$.

(6.7) LEMMA. *Let J be a subset of $\mathbb{Z}$ bounded below and let ν be a non-zero integer. Let e be an integer such that $1 \le e \le h(\nu, J) + 1$. Let*

$$J' = \left\{ j/d_e \middle| j \in J, j < m_e(\nu, J) \right\},$$

where $d_e = d_e(\nu, J)$. Let $\nu' = \nu/d_e$. Then $J' \subset \mathbb{Z}$, J' is bounded below, ν' is a non-zero integer and we have

$$\begin{aligned} h(\nu', J') &= e - 1, \\ m_i(\nu', J') &= m_i(\nu, J)/d_e, \\ d_{i+1}(\nu', J') &= d_{i+1}(\nu, J)/d_e \end{aligned}$$

for $0 \le i \le h(\nu', J')$.

Proof. A straightforward verification.

In the remainder of this section we let ν be an integer such that $|\nu| = n$. □

(6.8) DEFINITION. The *m-sequence* $m(\nu, f)$ of f *with respect to* ν is defined by

$$m(\nu, f) = m(\nu, \mathrm{Supp}(f)).$$

Note that, since $|\nu| = \deg_Y f, h(\nu, \mathrm{Supp}(f))$ depends only on f an does not depend upon ν. We shall write $h(f)$ for $h(\nu, \mathrm{Supp}(f))$ and $m_i(\nu, f)$ for $m_i(\nu, \mathrm{Supp}(f))$ for $0 \le i \le h(f) + 1$. Note that $m_i(\nu, f) = \mathrm{ord}_t\, y(wt)$ for every $w \in \mu_n(k)$.

(6.9) DEFINITION. The *d-sequence* $d(f)$ of f is defined to be

$$d(f) = (d_1(f), \dots, d_{h+1}(f), d_{h+2}(f)),$$

where $h = h(f)$ and $d_i(f) = d_i(\nu, \mathrm{Supp}(f))$ as defined in Definition (6.4), **33**
$1 \le i \le h + 2$. We note that, since $|\nu| = \deg_Y f$, $d(f)$ depends only on f and does not depend upon ν.

(6.10) DEFINITION. The *q-sequence* $q(\nu, f)$ *of* f *with respect to* ν is defined to be

$$q(\nu, f) = (q_0(\nu, f), \dots, q_n(\nu, f), q_{h+1}(\nu, f)),$$

where $h = h(f)$, $q_i(v, f) = m + i(v, f)$ for $i = 0, 1$, and $q_j(v, f) = m_j(v, f) - m_{j-1}(v, f)$ for $2 \leq j \leq h + 1$.

(6.11) DEFINITION. The *ssequence* $s(v, f)$ *of* f *with respect to* v is defined to be

$$s(v, f) = (s_0(v, f), \dots, s_h(v, f), s_{h+1}(v, f)),$$

where $h = h(f)$, $s_0(v, f) = q_0(v, f)$ and

$$s_i(v, f) = \sum_{p=1}^{i} q_p(v, f) d_p(f)$$

for $1 \leq i \leq h + 1$.

(6.12) DEFINITION. The *r-sequence* $r(v, f)$ *of* f *with respect to* v is defined to be

$$r(v, f) = (r_0(v, f), \dots, r_h(v, f), r_{h+1}(v, f)),$$

where $h = h(f), r_0(v, f) = s_0(v, f)$ and $r_i(v, f) = s_i(v, f)/d_i(f)$ for $1 \leq i \leq h + 1$.

Some properties of the various sequences defined above are listed in the following proposition. These will be used in the sequel, mostly without explicit reference.

34 **(6.13) PROPOSITION.** Let v be an integer such that $|v| = n$. Let $h = h(f)$ and for every $i, 0 \leq i \leq h + 1$, let $m_i = m_i(v, f)$, $q_i = q_i(v, f)$, $s_i = s_i(v, f)$, $r_i = r_i(v, f)$ and $d_{i+1} = d_{i+1}(f)$. Then:

(i) d_{i+1} divides d_i for $1 \leq i \leq h + 1$.

(ii) $d_1 \geq d_2 > d_3 > \cdots > d_h > d_{h+1} = 1$.

(iii) $d_1 = n$ and $d_{h+2} = -\infty$.

(iv) $r_0 = s_0 = q_0 = m_0 = v$ and $r_1 = q_1 = m_1$.

(v) $r_{h+1} = s_{h+1} = q_{h+1} = m_{h+1} = \infty$.

(vi) m_i, q_i, s_i, r_i are integers for $0 \le i \le h$.

(vii) $m_1 < m_2 < \cdots < m_h < m_{h+1} = \infty$.

(viii) q_i is a positive integer for $2 \le i \le h$.

(ix) $d_i =$ g.c.d. $(\{n\} \cup \{j \in \mathrm{Supp}(f) | j < m_i\})$ for $1 \le i \le h+1$.

(x) For $0 \le i \le h+1$, we have

(1) $d_{i+1} =$ g.c.d. $(m_0, \dots, m_i)$,
(2) $d_{i+1} =$ g.c.d. $(q_0, \dots, q_i)$,
(3) $d_{i+1} =$ g.c.d. $(r_0, \dots, r_i)$,
(4) $d_{i+1} =$ g.c.d. $(s_0, s_1/d_1 \dots, s_i/d_i)$.

In particular, each of the four sequences $m(\nu, f)$, $q(\nu, f)$, $s(\nu, f)$ and $r(\nu, f)$ determines $d(f)$, the sequence $s(\nu, f)$ determining $d(f)$ by the recursive formula (4).

(xi) each one of the four sequences $m(\nu, f)$, $q(\nu, f)$, $s(\nu, f)$ and $r(\nu, f)$ determines the other three.

Proof.

(i) Follows from Lemma (6.5).

(ii) Follows from Lemma (6.5) and Theorem (5.14).

(iii) Clear from the definition and Lemma (6.5).

(iv) Clear from the definition. **35**

(v) Clear from the definition.

(vi) By Lemma (6.5) m_i is an integer for $0 \le i \le h$. Therefore it follows the definition that q_i, s_i are integers for $0 \le i \le h$ and that r_0 is an integer. Now by (i) d_p/d_i is an integer for $1 \le p \le i \le h$. Therefore for $1 \le i \le h$

$$r_i = s_i/d_i = \sum_{p=1}^{i} q_p(d_p/d_i)$$

is an integer.

(vii) Follows from Lemma (6.5).

(viii) Follows from (vi) and (vii).

(ix) Follows from Lemma (6.5), since $n = |\nu|$.

(x) (i) follows from Lemma (6.5). (2) follows easily from (1), since $q_0 = m_0$, $q_1 = m_1$ and $q_i = m_i - m_{i-1}$ for $1 \leq i \leq h+1$. To prove (3), we note that we have

$$r_i \sum_{p=1}^{i-1} q_p(d_p/d_i) + q_i \tag{6.13.1}$$

for $1 \leq i \leq h+1$. Therefore, since d_p/d_i is an integer and since d_i divides q_p for $1 \leq p \leq i-1$, we get

$$\begin{aligned} \text{g.c.d. } (d_i, r_i) &= \text{g.c.d. } (d_i, q_i) \\ &= \text{g.c.d. } (q_0, \ldots, q_{i-1}, q_i) && \text{(by (2))} \\ &= d_{i+1} && \text{(by (2))} \end{aligned}$$

for $1 \leq i \leq h+1$. Therefore, since $d_1 = |q_0| = |r_0|$, we get (3) for $0 \leq i \leq h+1$ by induction on i, (4) is immediate from (3).

36 (xi) Since each of four sequences determines $d(f)$ by (x), it is enough to show that each one of them *together with* $d(f)$ determines the other three. It is clear from the definition that $m(\nu, f)$ determines $q(\nu, f)$, $q(\nu, f)$ and $d(f)$ determine $s(\nu, f)$, and $s(\nu, f)$ and $d(f)$ determine $r(\nu, f)$. Moreover, $q(\nu, f)$ clearly determines $m(\nu, f)$ by the formulas

$$\begin{aligned} m_0 &= q_0, \\ m_i &= \sum_{p=1}^{i} q_p, \quad 1 \leq i \leq h+1. \end{aligned}$$

□

Therefore, to complete the cycle, it is enough to show that $r(\nu, f)$ and $d(f)$ determine $q(\nu, f)$. But this is clear from the recursive formulas

$$q_0 = r_0,$$
$$q_i = r_i - \sum_{p=1}^{i-1} q_p(d_p/d_i), 1 \le i \le h+1,$$

which we get from 6.13.1.

(6.14) LEMMA. *Let ν be an integer such that $|\nu| = n$. Let $h = h(f)$ and let $m_i = m_i(\nu, f)$, $d_{i+1} = d_{i+1}(f)$ for $0 \le i \le h+1$. Let $y(t)$ be an element of $k((t))$ such that $f(t^n, y(t)) = 0$. Let e be an integer such that $1 \le e \le h+1$. Let w be an nth root of unity in k and let $p = \mathrm{ord}(w)$. Then we have:*

(i) $\mathrm{ord}_t(y(t) - y(wt)) \ge m_e$ *if and only if p divides d_e.*

(ii) $\mathrm{ord}_t(y(t) - y(wt)) \le m_e$ *if and only if p does not divide d_{e+1}.*

(iii) $\mathrm{ord}_t(y(t) - y(wt)) = m_e$ *if and only if p divides d_e and p does not divide d_{e+1}.*

Proof. It is clearly enough to prove (i) and (ii). Since $\mathrm{ord}_t(y(t) = m_1 = \mathrm{ord}_t\, y(wt)$ and since p divides $n = d_1$, (i) is obvious for $e = 1$. Since $m_{h+1} = \infty$ and since p does not divide $-\infty = d_{h+2}$, (ii) is obvious for **37**
$e = h+1$. Therefore it is enough to prove (i) for $e \ge 2$ and (ii) for $e \le h$. Now, for the moment, grant the following two statements:

(i′) *If* $2 \le e \le h+1$ *and p divides d_e then* $\mathrm{ord}_t(y(t) - y(wt)) \ge m_e$.

(ii′) *If* $1 \le e \le h$ *and p does not divide d_{e+1} then* $\mathrm{ord}_t(y(t) - y(wt)) \le m_e$.

Then if $2 \le e \le h+1$ and $\mathrm{ord}_t(y(t) - y(wt)) \ge m_e$ we get $\mathrm{ord}_t(y(t) - y(wt)) > m_{e-1}$, since $m_e > m_{e-1}$. This shows by (ii′) that p divides d_e. If $1 \le e \le h$ and $\mathrm{ord}_t(y(t) - y(wt)) \le m_e$ then we get $\mathrm{ord}_t(y(t) - y(wt)) < m_{e+1}$ since $m_e < m_{e+1}$. This shows by (i′) that p does not divide d_{e+1}. Thus, in order to complete the proof of the lemma, it is enough to prove (i′) and (ii′).

(i′) Let $J = \mathrm{Supp}(f) = \mathrm{Supp}_t\, y(t)$. Write $y(t) = \sum_{j \in J} y_j t^j$ with $y_j \in k$, $h_j \neq 0$ for every $j \in J$. Then $y(wt) = \sum_{j \in J} w^j y_j t^j$. Therefore we have

$$\begin{aligned} \mathrm{ord}_t(y(t) - y(wt)) &= \inf \left\{ j \in J \middle| w^j \neq 1 \right\} \\ &= \inf \left\{ j \in J \middle| j \not\equiv 0 (\mathrm{mod}\ p) \right\} \\ &= \inf \left\{ j \in J \middle| j \not\equiv 0 (\mathrm{mod}\ d_e) \right\} \\ & m_e, \end{aligned}$$

where the inequality follows from the fact that p divides d_e.

(ii′) Let

$$c = \inf \left\{ i \middle| 1 \leq i \leq h, p \text{ does not divide } d_{i+1} \right\}.$$

Then, since p divides $n = d_1$, we see that p divides d_c and p does not divide d_{c+1}. Moreover, $c \leq e$. Now, $d_{c+1} = \text{g.c.d.}\ (d_c, m_c)$. Since p divides d_c and p does not divide d_{c+1}, we see that p does not divide m_c. Therefore $w^{m_c} \neq 1$, which shows that

$$\mathrm{ord}_t(y(t) - y(wt)) \leq m_c \leq m_e.$$

□

38 **(6.15) PROPOSITION.** Let ν be an integer such that $|\nu| = n$. Let $h = h(f)$ and let $m_i = m_i(\nu, f)$, $d_{i+1} = d_{i+1}(f)$ for $0 \leq i \leq h+1$. Let $y(t)$ be an element of $k((t))$ such that $f(t^n, y(t)) = 0$. Let

$$E = \left\{ \mathrm{ord}_t(y(w_1 t) - y(w_2 t)) \middle| w_1, w_2 \in \mu_n(k), w_1 \neq w_2 \right\},$$

$$M_1 = \{m_1, \ldots, m_h\}$$

and $$M_2 = \{m_2, \ldots, m_h\}.$$

Then $M_2 \subset E \subset M_1$. Moreover, we have

$$E = \begin{cases} M_1, & \text{if } d_1 > d_2, \\ M_2, & \text{if } d_1 = d_2. \end{cases}$$

Proof. If $h = 0$ then $d_1 = 1$ and $E = M_1 = M_2 = \phi$. We may therefore assume that $h \geq 1$. Since $\mathrm{ord}_t(y(w_1 t) - y(w_2 t)) = \mathrm{ord}_t(y(t) - y(w_2 w_1^{-1} t))$, it is clear that $E = \left\{ \mathrm{ord}_t(y(t) - y(wt)) \middle| w \in \mu_n(k), w \neq 1 \right\}$. Let $w \in \mu_n(k)$, $w \neq 1$, and let $p = \mathrm{ord}(w)$. Then p divides $n = d_1$ and p does not divide $1 = d_{h+1}$. Therefore there exists e, $1 \leq e \leq h$, such that p divides d_e and p does not divide d_{e+1}. Therefore by Lemma (6.14) we get

$$\mathrm{ord}_t(y(t) - y(wt)) = m_e \in M_1.$$

This proves that $E \subset M_1$. Now, let i be an integer such that $2 \leq i \leq h$. Since d_i divides $d_1 = n$, there exists $w \in \mu_n(k)$ such that $\mathrm{ord}(w) = d_1$. Since $i \geq 2$, d_i does not divide d_{i+1} by Proposition (6.13). Therefore by Lemma (6.14) we have

$$m_i = \mathrm{ord}_t(y(t) - y(wt)) \in E.$$

This proves that $M_2 \subset E$. Now, suppose $d_1 > d_2$. Then, if w is a **39**
primitive nth root of unity in k, $\mathrm{ord}(w) = d_1$ does not divide d_2, so that by Lemma (6.14) we get

$$m_1 = \mathrm{ord}_t(y(t) - y(wt)) \in E,$$

which proves that $E = M_1$. Finally, suppose $d_1 = d_2$. Then, since $d_2 =$ g.c.d. $(d_1, m_1), d_1$ divides m_1. Therefore $w^{m_1} = 1$ for every $w \in \mu_n(k)$. Since $\mathrm{ord}_t\, y(t) = m_1 = \mathrm{ord}_t\, y(wt)$, it follows that

$$\mathrm{ord}_t(y(t) - y(wt)) > m_1$$

for every $w \in \mu_n(k)$. This means that $m_1 \notin E$, which proves that $E = M_2$. □

(6.16) PROPOSITION. Let ν be an integer such that $|\nu| = n$. Let e be an integer such that $1 \leq e \leq h(f) + 1$. Let $d_e = d_e(f)$ and let $n' = n/d_{e'} \nu' = \nu/d_e$. Let f' be an irreducible element of $k((X))[Y]$ such that f' is monic in Y and $\deg_Y f' = n'$. Assume that

$$\mathrm{Supp}(f') = \left\{ j/d_e \middle| j \in \mathrm{Supp}(f), j < m_e(\nu, f) \right\}.$$

Then $h(f') = e - 1$, and for $o \leq i \leq h(f')$ we have:

(i) $m_i(\nu', f') = m_i(\nu, f)d_e$.

(ii) $d_{i+1}(f') = d_{i+1}(f)/d_e$.

(iii) $q_i(\nu', f') = q + i(\nu, f)/d_e$.

(iv) $s_i(\nu', f') = s_i(\nu, f)/d_e^2$ (if $i \neq 0$).

(v) $r_i(\nu', f') = r_i(\nu, f)/d_e$.

Proof. (i) and (ii) follow from Lemma (6.7). (iii), (iv) and (v) follow
40 immediately from (i) and (ii). □

(6.17) PROPOSITION. Let ν be an integer such that $|\nu| = n$. Let f' be an irreducible element of $k((X))[Y]$ such that f' is monic in Y and $\deg_Y f' = n$. Suppose there exists $z(t) \in k((t))$ such that $f'(t^n, z(t)) = 0$ and $\mathrm{ord}_t(z(t) - y(t)) > m_h(\nu, f)$, where $h = h(f)$. Then we have:

(i) $h(f') = h(f)$.

(ii) $m(\nu, f') = m(\nu, f)$.

(iii) $q(\nu, f') = q(\nu, f)$.

(iv) $s(\nu, f') = s(\nu, f)$.

(v) $r(\nu, f') = r(\nu, f)$.

(vi) $d(f') = d(f)$.

Proof. Let $J = \mathrm{Supp}(f)$, $J' = \mathrm{Supp}(f')$. Then the hypothesis implies that we have

(6.17.1) $$\left\{ j \in J \,\middle|\, j \leq m_h(\nu, f) \right\} = \left\{ j \in J' \,\middle|\, j \leq m_h(\nu, f) \right\}.$$

We shall prove the lemma under the weaker assumption (6.17.1). Note that it is enough to prove (ii). For, the rest then follows from (ii) and the definition. We first prove by induction on i that we have

(6.17.2) $$i \leq h(f') \text{ and } m_i(\nu, f) = m_i(\nu, f')$$

for $0 \leq i \leq h = h(f)$. For $i = 0$, this is clear. Suppose now that p is an integer, $1 \leq p \leq h$, such that (6.17.2) holds for $0 \leq i \leq p-1$. Then by Proposition (6.13) (x) we have

$$d_p(f') = d_p(f) = d_p, \text{ say. } \textit{Let}$$

$$J_p = \begin{cases} \left\{ j \in J \,\middle|\, j \neq 0 \pmod{d_p} \right\}, & \text{if } p \geq 2, \\ J, & \text{if } p = 1, \end{cases}$$

$$J'_p = \begin{cases} \left\{ j \in J' \,\middle|\, j \neq 0 \pmod{d_p} \right\}, & \text{if } p \geq 2, \\ J', & \text{if } p = 1. \end{cases}$$

Then we have $m_p(\nu, f) = \inf(J_p)$, $m_p(\nu, f') = \inf(J'_p)$. Since $m_p(\nu, f)$ **41**
$\leq m_h(\nu, f)$, we have $m_p(\nu, f) \in J'_p$ by (6.17.1). This shows that $m_p(\nu, f') \leq m_p(\nu, f) \leq m_h(\nu, f)$. Therefore by (6.17.1) $m_p(\nu, f') \in J_{p'}$ so that $m_p(\nu, f') \leq m_p(\nu, f)$. This proves that $m_p(\nu, f') = m_p(\nu, f) < \infty$, which shows also that $p \leq h(f')$. Thus (6.17.2) is proved for $0 \leq i \leq h$. In particular, we get $h \leq h(f')$ and $d_{h+1}(f') = d_{h+1}(f) = 1$ by Proposition (6.13). This means that

$$J'_{h+1} = \left\{ j \in J \,\middle|\, j \notin 0 \,(\mathrm{mod}\; d_{h+1}(f')) \right\}$$

is empty, so that $h \geq h(f')$. Thus we have $h(f') = h = h(f)$ and by (6.17.2) we get $m(\nu, f) = m(\nu, f')$. $\square$

Chapter 3

The Fundamental Theorem

7 The Main Lemmas

Throughout this section, we preserve the notation introduced in (7.1) 42
and (7.2) below

(7.1) NOTATION. Let k be an algebraically closed field and let X, Y, t be indeterminates. Let n be a positive integer such that char k does not divide n. Let $f = f(X, Y)$ be an irreducible element of $k((X))[Y]$ such that f is monic in Y and $\deg_Y f = n$. Let ν be an integer such that $|\nu| = n$. Let $h = h(f)$ and for every i, $0 \leq i \leq h + 1$, let

$$\begin{aligned} m_i &= m_i(\nu, f) \\ q_i &= q_i(\nu, f) \\ s_i &= s_i(\nu, f) \\ r_i &= r_i(\nu, f) \\ d_{i+1} &= d_{i+1}(f). \end{aligned}$$

Also, let

$$n_i = d_i / d_{i+1}$$

for $1 \leq i \leq h$. (Note that by Proposition (6.13)) n_i is a positive integer for every i and $n_i \geq 2$ for $2 \leq i \leq h$. Finally, we fix a root $y(t)$ of $f(t^n, Y) = 0$, i.e., we fix an element $y(t)$ of $k((t))$ such that $f(t^n, y(t)) = 0$. Recall then that by Newton's Theorem (5.14) we have

$$f(t^n, Y) = \prod_{w \in \mu_n} (Y - y(wt)),$$

43 where for a positive integer m we write μ_m for $\mu_m(k)$. Let

$$y(t) = \sum_j y_j t^j$$

with $y_j \in k$ for every j.

(7.2) NOTATION. We shall use the symbol $\varnothing$ to denote a generic (i.e. unspecified) non-zero element of k. Thus if k' is an overfield of k and $a \in k'$ then $a = \varnothing$ means that $a \in k$ and $a \neq 0$. Similarly, $b = \varnothing c$ means that $b = ac$ for some $a \in k$, $a \neq 0$. Note that $a = \varnothing$ and $b = \varnothing$ does not mean that $a = b$.

(7.3) DEFINITION. Let k' be an overfield of k and let z be a non-zero element of $k'((t))$. If $m = \operatorname{ord}_t z$, we can write z in the form

$$z = at^m + t^{m+1}z'$$

with $a \in k$, $a \neq 0$ and $z' \in k'((t))$. We define the initial form (resp. initial co-efficient) of z, denoted info (z) (resp. inco (z)), by info $(z) = at^m$ (resp. inco $(z) = a$). We also define info $(0) = 0$, inco $(0) = 0$.

(7.4) DEFINITION. Let i be an integer with $1 \leq i \leq h+1$. We define

$$A(i) = \left\{ w \in \mu_n \middle| \operatorname{ord}_t(y(t) - y(wt)) < m_i \right\},$$
$$R(i) = \left\{ w \in \mu_n \middle| \operatorname{ord}_t(y(t) - y(wt)) \geq m_i \right\},$$
$$S(i) = \left\{ w \in \mu_n \middle| \operatorname{ord}_t(y(t) - y(wt)) = m_i \right\}.$$

(7.5) LEMMA. *Let i be an integer, $1 \leq i \leq h+1$. Then:*

(i) $R(i) = \mu_{d_i}$. *In particular, card* $(R(i)) = d_i$.

(ii) *Let $i \leq h$. Then* $S(i) = R(i) - R(i+1) = \mu_{d_i} - \mu_{d_{i+1}}$. *In particular,*
44 *card* $(S(i)) = d_i - d_{i+1}$.

(iii) $S(h+1) = \{1\}$.

Proof.

(i) By Lemma (6.14) we have

$$R(i) = \left\{ w \in \mu_n \middle| \operatorname{ord}(w) \text{ divides } d_i \right\}$$
$$= \left\{ w \in \mu_n \middle| w^{d_i} = 1 \right\} = \mu_{d_i}.$$

(ii) Since, for every $w \in \mu_n$, $\operatorname{ord}_t(y(t) - y(wt))$ belongs to the set $\{m_1, \ldots, m_{h+1}\}$ by Proposition (6.15), we see that $S(i) = R(i) - R(i+1)$, for $1 \leq i \leq h$. Therefore (ii) follows from (i).

(iii) This is clear, since $m_{h+1} = \infty$ and the roots $y(wt)$, $w \in \mu_n$, are distinct.

□

(7.6) LEMMA. *Let e be an integer,* $1 \leq e \leq h$*, and let* $m = m_e$*. Let z be an element of an overfield of k. Then we have*

$$\prod_{w \in R(e)} (z - w^m y_m) = (z^{n_e} \cdot y_m^{n_e})^{d_{e+1}}.$$

Proof. Let u be a primitive d_eth root of unity in k and let $v = u^m$. Then, since $d_{e+1} = \text{g.c.d. } (d_e, m)$, we see that v is a primitive n_e^{th} root of unity. Therefore, since

$$R(e) = \mu_{d_e} = \left\{ u^i \middle| 1 \leq i \leq d_e \right\}$$

by Lemma (7.5), we get

$$\begin{aligned}
\prod_{w \in R(e)} (z - w^m y_m) &= \prod_{i=1}^{d_e} (z - v^i y_m) \\
&= \prod_{i=0}^{d_{e+1}-1} \prod_{j=1}^{n_e} (z - v^{j+in_e} y_m) \\
&= \left(\prod_{j=1}^{n_e} (z - v^i y_m) \right)^{d_{e+1}} \qquad (\text{since } v^{n_e} = 1) \\
&= (z^{n_e} - y_m^{n_e})^{d_{e+1}},
\end{aligned}$$

since v is a primitive n_e^{th} root of unity. □

(7.7) LEMMA. *Let i be an integer,* $1 \le i \le h+1$. *Then we have*

$$\operatorname{ord}_t\left(\prod_{w\in Q(i)}(y(t)-y(wt))\right) = \begin{cases} s_{i-1}-m_{i-1}d_i, & \textit{if } i \ge 2,\\ 0, & \textit{if } i = 1.\end{cases}$$

Proof. Since, for every $w \in \mu_n$, $\operatorname{ord}_t(y(t)-y(wt))$ belongs to the set $\{m_1,\dots,m_{h+1}\}$ by Proposition (6.15), we get

$$\prod_{w\in Q(i)}(y(t)-y(wt)) = \prod_{j=1}^{i-1}\prod_{w\in S(j)}(y(t)-y(wt)). \tag{7.7.1}$$

From this the assertion is clear for $i = 1$. Assume now that $i \ge 2$. Since card $(S(j)) = d_j - d_{j+1}$ by Lemma (7.5), we have

$$\operatorname{ord}_t\left(\prod_{w\in S(j)}(y(t)-y(wt))\right) = (d_j-d_{j+1})m_j$$

for $1 \le j \le h$. Therefore from (7.7.1) we get

$$\begin{aligned}\operatorname{ord}_t\left(\prod_{w\in Q(i)}(y(t)-y(wt))\right) &= \sum_{j=1}^{i-1}\left(d_j-d_{j+1}\right)m_j\\ &= \sum_{j=1}^{i-1} q_j d_j - m_{i-1}d_i\\ &= s_{i-1}-m_{i-1}d_i.\end{aligned}$$

□

(7.8) COROLLARY.

$$\operatorname{ord}_t\left(\prod_{\substack{w\in\mu_n\\ w\ne 1}}(y(t)-y(wt))\right) = \sum_{j=1}^{h} q_j(d_j-1) = s_h - m_h.$$

Proof. The equality $\sum_{j=1}^{h} q_j(d_j-1) = s_h-m_h$ is clear. Now, if $h = 0$ then $n = d_1 = d_{h+1} = 1$, so that the assertion is clear in this case, since in the middle we have an empty sum and on the left hand side the order of an empty product. Assume now that $h \geq 1$. Taking $i = h + 1$ in Lemma (7.7), we get $Q(i) = \mu_n - \{1\}$ and

$$s_{i-1} - m_{i-1}d_i = s_h - m_h d_{h+1} = s_h - m_h = \sum_{j=1}^{h} q_j(d_j - 1).$$

□

(7.9) COROLLARY. Let $f_Y(X, Y)$ denote the Y-derivative of $f(X, Y)$. Then we have

$$\mathrm{ord}_t(f_Y(t^n, y(t))) = \sum_{j=1}^{h} q_j(d_j - 1) = s_h - m_h.$$

Proof. Since

$$f(t^n, Y) = \prod_{w \in \mu_n} (Y - y(wt))$$

we get

$$f_Y(t^n, y(t)) = \prod_{\substack{w \in \mu_n \\ w \neq 1}} (y(t) - y(wt))$$

and the assertion follows from Corollary (7.8). □

(7.10) COROLLARY. Let $u(t)$ be an element of $k((t))$ such that $\mathrm{ord}_t(u(t) - y(t)) > m_h$. Then

$$\mathrm{ord}_t(f(t^n, u(t))) = s_h - m_h + \mathrm{ord}_t(u(t) - y(t)).$$

Proof. Let $w \in \mu_n$, $w \neq 1$. Then $\mathrm{ord}_t(y(t) - y(wt)) \leq m_h$ by Lemma (6.14). Therefore, since

$$u(t) - y(wt) = (u(t) - y(t)) + (y(t) - y(wt))$$

and since $\mathrm{ord}_t(u(t) - y(t)) > m_h$, we get

$$\mathrm{ord}_t(u(t) - y(wt)) = \mathrm{ord}_t(y(t) - y(wt))$$

for every $w \in \mu_n$, $w \neq 1$. Therefore

$$\begin{aligned}
\operatorname{ord}_t(f(t^n, u(t))) &= \operatorname{ord}_t\left(\prod_{w\in\mu_n}(u(t) - y(wt))\right) \\
&= \operatorname{ord}_t(u(t) - y(t)) + \operatorname{ord}_t\left(\prod_{w\neq 1}(y(t) - y(wt))\right) \\
&= \operatorname{ord}_t(u(t) - y(t)) + s_h - m_h
\end{aligned}$$

by Corollary (7.8). □

(7.11) LEMMA. *Let i be an integer, $1 \leq i \leq h+1$. Let $\overline{y}(t) = \sum\limits_{j<m_i} y_j t^j$. Let $G_i = G_i(X, Y) \in k((X))[Y]$ be the minimal monic polynomial of $\overline{y}(t)$ over $k((t^n))$. (See Definition (5.8).) Then we have:*

(i) $\deg_Y G_i = n/d_i$.

(ii) *G_i is also the minimal monic polynomial of $\overline{y}(wt)$ over $k((t^n))$ for every $w \in \mu_n$.*

Proof.

(i) We have

$$\operatorname{Supp}_t \overline{y}(t) = \left\{ j \in \operatorname{Supp}_t y(t) \,\middle|\, j < m_i \right\}.$$

Therefore by Proposition (6.13) (ix) we have

$$d_i = \text{g.c.d.}\ \ (\{n\} \cup \operatorname{Supp}_t \overline{y}(t)).$$

Now, the assertion follows from Proposition (5.16).

(ii) Substituting wt for t in the equality $G_i(t^n, \overline{y}(t)) = 0$ we get
48 $G_i(t^n, \overline{y}(wt)) = 0$. This proves (ii).

□

(7.12) DEFINITION. Let i be an integer, $1 \leq i \leq h+1$. The element $G_i = G_i(X, Y)$ of Lemma (7.11) is called the *pseudo f_ith root* of f. By Lemma (7.11) we note that G_i depends only on f and i and does not depend upon the root $y(t)$ of $f(t^n, Y)$ and that G_i is an irreducible element of $k((X))[Y]$, monic in Y, and $\deg_Y G_i = n/d_i$.

(7.13) LEMMA. *Let i be an integer, $1 \leq i \leq h$, and let $G_i(x, Y)$ be the pseudo d_ith root of f. Let k' be an overfield of k and let y^* be an element of $k'((t))$ such that*

$$\text{info } (y^* - \sum_{j<m_i} y_j t^j) = zt^{m_i}$$

with $z \in k'$, $z \neq 0$. Then info $(G_i(t^n, y^*)) = \varnothing zt^{r_i}$.

Proof. Let $\overline{y}(t) = \sum_{j<m_i} y_j t^j$. Then by Proposition (6.13) (ix) we have

$$d_i = \text{g.c.d. } (\{n\} \cup \text{Supp}_t \overline{y}(t)).$$

Therefore by Proposition (5.16) we get

(7.13.1) $$\prod_{w \in \mu_n} (Y - \overline{y}(wt)) = G_i(t^n, Y)^{d_i}.$$

Now, $\text{ord}_t(y(wt) - \overline{y}(wt)) = m_i$ for every $w \in \mu_n$. Therefore, since

$$y^* - \overline{y}(wt) = (y^* - y(t)) + (\overline{y}(t) - y(t)) + (y(t) - y(wt)) + (y(wt) - \overline{y}(wt))$$

and since $\text{ord}_t(y^* - \overline{y}(t)) = m_i$ by assumption, we have

(7.13.2) $$info(y^* - \overline{y}(wt)) = info(y(t) - y(wt)) \text{ for } w \in Q(i).$$

Next, if $w \in R(i)$ then $w^j = 1$ for all j in $\text{Supp}_t\, y(t)$ such that $j < m_i$.

Therefore $\overline{y}(t) = \overline{y}(wt)$ for all $w \in R(i)$ and we get 49

(7.13.3) $$info(y^* - \overline{y}(wt)) = info(y^* - \overline{y}(t)) = zt^{m_i} \; for\ w \in R(i).$$

From 7.13.1 we get

(7.13.4)

$$\begin{aligned}
info(G_i(t^n, Y^*)^{d_i}) &= \left(\prod_{w \in \mu_n} (y^* - \overline{y}(wt)) \right) \\
&= info \left(\prod_{w \in Q(i)} (y^* - \overline{y}(wt)) \right) \prod_{w \in R(i)} info(y^* - \overline{y}(wt)) \\
&= info \left(\prod_{w \in Q(i)} (y(t) - y(wt)) \right) z^{d_i} t^{m_i d_i}
\end{aligned}$$

by 7.13.2 and 7.13.3, since card $(R(i)) = d_i$ by Lemma (7.5). Since $y(t)$ and $y(wt)$ belong to $k((t))$, we have

$$inco\left(\prod_{w\in Q(i)}(y(t)-y(wt))\right)\in k.$$

Therefore by Lemma (7.7) we have

$$info\left(\prod_{w\in Q(i)}(y(t)-y(wt))\right)=\begin{cases}\varnothing t^{s_{i-1}-m_{i-1}d_i}, & \text{if } i\geq 2,\\ \varnothing, & \text{if } i=1.\end{cases}$$

Therefore from 7.13.4 we get

$$info(G_i(t^n,y^*)^{d_i})=\varnothing z^{d_i}t^s,$$

where

$$s=\begin{cases}s_{i-1}-m_{i-1}d_i+m_id_i, & \text{if } i\geq 2,\\ m_id_i, & \text{if } i=1.\end{cases}$$

We see that in either case we have $s=s_i=r_id_i$. Thus we have

$$(info(G_i(t^n,y^*)))^{d_i}=info(G_i(t^n,y^*)^{d_i})=\varnothing z^{d_i}t^{r_id_i}.$$

50 It follows that we have

$$info(G_i(t^n,y^*))=\varnothing zt^{r_i}.$$

□

(7.14) DEFINITION. Let e be an integer, $1\leq e\leq h$, and let Z be an indeterminate. By an (e,Z)-*deformation* of $y(t)$ we mean an element y^* of $k'(Z)((t))$, where k' is an overfield of k, such that

$$info(y^*-\sum_{j<m_e}y_jt^j)=Zt^{m_e}.$$

(7.15) COROLLARY. Let $i.e.$, be integers such that $1\leq i\leq e\leq h$. Let $G_i(X,Y)$ be the pseudo d_ith root of f. Let y^* be an (e,Z)- deformation of $y(t)$. then we have

$$\text{info }(G_i(t^n,y^*))=\begin{cases}\varnothing t^{r_i}, & \text{if } i<e,\\ \varnothing Zt^{r_i}, & \text{if } i=e.\end{cases}$$

Proof. Let k' be an overfield of k such that $y^* \in k'(Z)((t))$. Let $\overline{y}(t) = \sum_{j<m_i} y_j t^j$ and $y'(t) = \sum_{m_i \leq j \leq m_e} y_j t^j$. Then, since y^* is an (e, Z)-deformation of $y(t)$, we have

$$y^* = \overline{y}(t) + y'(t) + Zt^{m_e} + u(t)$$

for some $u(t) \in k'(Z)((t))$ with $\operatorname{ord}_t u(t) > m_e \geq m_i$. It follows that if $i < e$ then

$$\text{info}\,(y^* - \overline{y}(t)) = y_{m_i} t^{m_i} = \varnothing t^{m_i},$$

whereas if $i = e$ then

$$\text{info}\,(y^* - \overline{y}(t)) = Zt^{m_e} = Zt^{m_i}.$$

Now, the corollary follows from Lemma (7.13). □

(7.16) LEMMA. *Let e be an integer, $1 \leq e \leq h$, and let y^* be an (e, Z)-deformation of $y(t)$. Then we have* 51

$$\textit{info}\,(f(t^n, y^*)) = \varnothing (Z^{n_e} - y_{m_e}^{n_e})^{d_{e+1}} t^{s_e}.$$

Proof. The assumption on y^* means that we can write y^* in the form

$$y^* = y(t) + (Z - y_{m_e}) t^{m_e} + u(t)$$

with $u(t) \in k'(Z)((t))$ and $\operatorname{ord}_t u(t) > m_e$, where k' is some overfield of k. Therefore for every $w \in \mu_n$ we have

$$y^* - y(wt) = (Z - y_{m_e}) t^{m_e} + (y(t) - y(wt)) + u(t). \tag{7.16.1}$$

It follows that if $w \in Q(e)$ then

$$\text{info}\,(y^* - y(wt)) = \text{info}\,(y(t) - y(wt)). \tag{7.16.2}$$

Since $y(t)$ and $y(wt)$ belong to $k((t))$, we have

$$\text{inco}\left(\prod_{w \in Q(e)} (y(t) - y(wt))\right) \in k.$$

Therefore it follows from (7.16.2) and Lemma (7.7) that we have

$$\text{info}\left(\prod_{w\in Q(e)}(y^* - y(wt))\right) = \begin{cases}\varnothing t^{s_{e-1}-m_{e-1}d_e}, & \text{if } e \geq 2,\\ \varnothing, & \text{if } e = 1.\end{cases} \tag{7.16.3}$$

Next, let $w \in R(e)$. Then it follows from 7.16.1 that $\text{ord}_t(y^* - y(wt)) \geq m_e$ and that the coefficient of t^{m_e} in $y^* - y(wt)$ is

$$(Z - y_{m_e}) + (y_{m_e} - w^{m_e}y_{m_e}) = Z - w^{m_e}y_{m_e}$$

which is non-zero, since Z is an indeterminate. This shows that

$$\text{info}\,(y^* - y(wt)) = (Z - w^{m_e}y_{m_e})t^{m_e}$$

52 for every $w \in R(e)$. Therefore by Lemma (7.6) we get

$$\begin{aligned}\text{info}\left(\prod_{w\in R(e)}(y^* - y(wt))\right) &= \prod_{w\in R(e)}(Z - w^{m_e}y_{m_e})t^{m_e}\\ &= (Z^{n_e} - y_{m_e}^{n_e})^{d_{e+1}}t^{m_e d_e},\end{aligned} \tag{7.16.4}$$

since card $(R(e)) = d_e$ by Lemma (7.5). Since

$$\begin{aligned}f(t^n, y^*) &= \prod_{w\in\mu_n}(y^* - y(wt))\\ &= \prod_{w\in Q(e)}(y^* - y(wt)) \prod_{w\in R(e)}(y^* - y(wt))\end{aligned}$$

and since

$$s_e = \begin{cases}s_{e-1} - m_{e-1}d_e + m_e d_e, & \text{if } e \geq 2,\\ m_e d_e, & \text{if } e = 1,\end{cases}$$

the lemma follows from (7.16.3) and (7.16.4). □

(7.17) MAIN LEMMA 1.

Let e be an integer, $1 \leq e \leq h$. Let $C = C(X, Y)$ be a non-zero element of $k((X))[Y]$ such that $\deg_Y C < n/d_e$. Let y^* be an (e, Z)-deformation of $y(t)$. Then inco $(C(t^n, y^*)) = \varnothing$.

Proof. Suppose $e = 1$. Then $n/d_e = n/d_1 = 1$, so that $\deg_Y C = 0$. This means that $C(X, Y)$ is a non-zero element of $k((X))$. Therefore $C(t^n, y^*)$ is a non-zero element of $k((t))$ and the assertion is clear in this case. □

Assume now that $e \geq 2$. Let $G_i = G_i(X, Y)$ be the pseudo d_ith root of f, $1 \leq i \leq e-1$, and let $G = (G_1, \ldots, G_{e-1})$. Since, by Lemma (7.11), G_i is monic in Y with $\deg_Y G_i = n/d_i$, $1 \leq i \leq e-1$, we see that the three conditions

(i)-(iii) of (2.2) are satisfied by G with $R = k((X))$ and $p = e-1$. **53**
With the notation of (2.2), we note that $n_{e-1}(G) = \infty$ and

(7.17.1) $$n_i(G) = (n/d_{i+1})/(n/d_i) = d_i/d_{i+1}$$

for $1 \leq i \leq e-2$. By Corollary (2.14), let

(7.17.2) $$C = \sum_{a \in A(G)} C_a(X) G^a, \quad C_a(X) \in k((X)),$$

be the G-adic expansion of C. Since $\deg_Y C < n/d_e$ be hypothesis, we have, by Corollary (2.9), $\deg_Y G^a < n/d_e$ for every $a \in \mathrm{Supp}_G(C)$. Since $\deg_Y(G^a) = \sum_{i=1}^{e-1} a_i \deg_Y G_i$, we get, in particular, $a_{e-1} \deg_Y G_{e-1} < n/d_e$ for every $a \in \mathrm{Supp}_G(C)$. Since $\deg_Y G_{e-1} = n/d_{e-1}$, we get $a_{e-1} n/d_{e-1} n/d_{e-1} < n/d_e < n/d_e$, which gives

(7.17.3) $$a_{e-1} < d_{e-1}/d_e$$

for every $a \in \mathrm{Supp}_G(C)$. Now, substituting $X = t^n$, $Y = y^*$ in (7.17.2), we get

(7.17.4) $$C(t^n, y^*) = \sum_{a \in \mathrm{Supp}_G(C)} C_a(t^n) G(t^n, y^*)^a.$$

For $a \in \mathrm{Supp}_G(C)$, let $a_0 = (n/r_0) \operatorname{ord}_X C_a(X)$. Then we have

(7.17.5) $$\operatorname{ord}_t C_a(t^n) = n \operatorname{ord}_X C_a(X) = a_0 r_0.$$

Moreover, for $1 \leq i \leq e-1$, we have

(7.17.6) $$\text{info } (G_i(t^n, y^*)) = \varnothing t^{r_i}$$

by Corollary (7.15). From (7.17.5) and (7.17.6) we get

$$\operatorname{ord}_t(C_a(t^n)G(t^n, y^*)^a) = \sum_{i=0}^{e-1} a_i r_i. \tag{7.17.7}$$

54 Now, let $r = (r_0, \ldots, r_{e-1})$. Then, with the notation of (1.1), we have $n_i(r) = d_i(r)/d_{i+1}(r) = d_i/d_{i+1}$ for $1 \leq i \leq e-1$. Let $a \in \mathrm{Supp}_G(C)$. Then for $1 \leq i \leq e-2$, we have

$$0 \leq a_i < n_i(G) = d_i/d_{i+1}$$

by (7.17.1). Moreover, $a_{e-1} < d_{e-1}/d_e$ by (7.17.3). Thus (7.17.7) expresses $\operatorname{ord}_t(C_a(t^n)G(t^n, y^*)^a)$ as a strict linear combination of r. Therefore it follows from Proposition (1.5) that

$$\operatorname{ord}_t(C_a(t^n)G(t^n, y^*)^a) \neq \operatorname{ord}_t(C_b(t^n)G(t^n, y^*)^b)$$

if a, $b \in \mathrm{Supp}_G(C)$ and $a \neq b$. Therefore, in view of (7.17.4), we see that there exists $a \in \mathrm{Supp}_G(C)$ such that

$$\text{info}\,(C(t^n, y^*)) - \text{info}\,(C_a(t^n)G(t^n, y^*)^a)$$

and, in particular,

$$\text{inco}\,(C(t^n, y^*)) = \text{inco}\,(C_a(t^n)G(t^n, y^*)^a). \tag{7.17.8}$$

Now, inco $(C_a(t^n)) = \varnothing$, since $C_a(t^n) \in k((t))$ and $C_a(X) \neq 0$. Also,

$$\begin{aligned} \text{inco}\,(G(t^n, y^*)^a) &= \prod_{i=1}^{e-1} \text{inco}\,(G_i(t^n, y^*)^{a_i}) \\ &= \varnothing \qquad\qquad \text{(by (7.17.6))}. \end{aligned}$$

Therefore by (7.17.8) inco $(C(t^n, y^*)) = \varnothing$, and the lemma is proved.

(7.18) MAIN LEMMA 2.

Let $R = k((X))$. Let e be an integer, $2 \leq e \leq h$. Let $g = g(x, Y)$ be an element of $R[Y]$ such that g is monic in Y and $\deg_Y g = n/d_e$.

55 Let y^* be an $(e.Z)$ deformation of $y(t)$ such that info $(g(t^n, y^*)) = \varnothing Z t^{r_e}$. Then info $(\tau_f(g)(t^n, y^*) = \varnothing Z t^{r_e}$, where τ_f is the Tschirnhausen operator with respect to $f \in R[Y]$. (See § 3 for definition of τ_f.)

Proof. Let $d = d_e$ and let

$$f = g^d + \sum_{i=0}^{d-1} C_i g^i \tag{7.18.1}$$

be the g-adic expansion of f, where $C_i = C_f^{(i)}(g)$, $0 \leq i \leq d-1$. (See (3.4.1).) Then, by definition, $\tau_f(g) = g + d^{-1}C_{d-1}$. Therefore, in order to prove the lemma, it is enough to prove that we have

$$\operatorname{ord}_t C_{d-1}(t^n, y^*) > r_e. \tag{7.18.2}$$

□

Now, from (7.18.1) we get

$$f(t^n, y^*) = \sum_{i=0}^{d} C_i(t^n, y^*) g(t^n, y^*)^i,$$

where $C_d = 1$. Let

$$u \in f \left\{ \operatorname{ord}_t (C_i(t^n, y^*) g(t^n, y^*)^i) \middle| 0 \leq i \leq d \right\}. \tag{7.18.3}$$

Since $C_d = 1$ and $\operatorname{ord}_t g(t^n, y^*)^d = dr_e$, we see that $u < \infty$. Let

$$I = \left\{ i \middle| o \leq i \leq d, \operatorname{ord}_t(C_i(t^n, y^*) g(t^n, y^*)^i) = u \right\}. \tag{7.18.4}$$

Then $C_i(t^n, y^*) \neq 0$ for every $i \in I$. Let $a_i =$ inco $(C_i(t^n, y^*))$, $i \in I$. Then, since $\deg_Y c_i < \deg g = n/d_e$, it follows from Main Lemma (7.17) that $a_i \in k$ and $a_i \neq 0$ for every $i \in I$. Also, by hypothesis we have info $(g(t^n, y^*)^i) = b_i Z^i t^{ir_e}$ for some $b_i \in k$, $b_i \neq 0$. Therefore we get

$$inco(C_i(t^n, y^*) g(t^n, y^*)^i) = a_i b_i Z^i$$

for every $i \in I$. It follows that the coefficient of t^u in $f(t^n, y^*)$ is **56**

$$\sum_{i \in I} a_i b_i Z^i$$

, which is non-zero, since $I \neq \phi$ and Z is an indeterminate. Therefore we have

$$info(f(t^n, y^*)) = \left(\sum_{i \in I} a_i b_i Z^i\right) t^u.$$

On the other hand, by Lemma (7.16) we have

$$info(f(t^n, y^*)) = \varnothing \left(Z^{n_e} - y_{m_e}^{n_e}\right)^{d_{e+1}} t^{s_e}.$$

Therefore we get $u = s_e = d_e r_e$ and

$$\sum_{i \in I} a_i b_i Z^i = \varnothing \left(Z^{n_e} - y_{m_e}^{n_e}\right)^{d_{e+1}}.$$

This last equality shows that we have

$$\sum_{i \in I} a_i b_i Z^i \in k[Z^{n_e}]. \tag{7.18.5}$$

Now, we have $n_e = d_e/d_{e+1} \geq 2$ by Proposition (6.13) (ii), since $e \geq 2$. Therefore n_e does not divide $d_e - 1 = d - 1$, and it follows from (7.18.5) that $d - 1 \notin I$. This means that

$$\begin{aligned} u &< \operatorname{ord}_t \left(C_{d-1}(t^n, y^*) g(t^n, y^*)^{d-1}\right) \\ &= \operatorname{ord}_t C_{d-1}(t^n, y^*) + (d-1) r_e. \end{aligned}$$

since $u = d_e r_e$, we get

$$r_e < \operatorname{ord}_t C_{d-1}(t^n, y^*).$$

which proves (7.18.2).

(7.19) THEOREM. *Let e be an integer, $2 \le e \le h$, and let $g_e(X,Y) = App_Y^{d_e}(f)$. (See (4.5).) Let y^* be an (e,Z)-deformation of $y(t)$. Then*

$$\textit{info}(g_e(t^n, y^*)) = \varnothing Z t^{r_e}.$$

Proof. Let $G_e(X,Y)$ be the pseudo d_e^{th} root of f. Then we have **57**

$$\text{info } (G_e(t^n, y^*)) = \varnothing Z t^{r_e} \tag{7.19.1}$$

by Corollary (7.15). Now, G_e is monic in Y with $\deg_Y G_e = n/d_e$ (Lemma (7.11)). Therefore by Corollary (4.6) we have $g_e(X,Y) = (\tau_f)^j (G_e)$, where $j = n/d_e$. Now, the theorem follows from (7.19.1) by n/d_e applications of Main Lemma (7.18). □

(7.20) COROLLARY. Let e be an integer, $2 \le e \le h$, and let $g_e(X,Y) = App_Y^{d_e}(f)$. Let k' be an overfield of k. Let $a \in k'$ and let u be an element of $k'((t))$ such that $\text{ord}_t\, u > m_e$. Let

$$\overline{y} = \sum_{j<m_e} y_j t^i + a t^{m_e} + u.$$

Then there exist $c \in k$, $c \neq 0$, and an element v of $k'((t))$ such that $\text{ord}_t\, v > r_e$ and

$$g_e(t^n, \overline{y}) = c a t^{r_e} + v.$$

Proof. Let Z be an indeterminate and let

$$y^* = \sum_{j<m_e} y_j t^j + Z t^{m_e} + u.$$

Then y^* is an (e,Z)-deformation of $y(t)$. Note that $y^* \in k'((t))[Z] \subset k'(Z)((t))$. Therefore $g_e(t^n, y^*) \in k'((t))[Z]$ and we can write

$$g_e(t^n, y^*) = \sum_{i=0}^{p} b_i(t) Z^i, \tag{7.20.1}$$

where p is a non-negative integer and $b_i(t) \in k'((t))$ for $0 \le i \le p$. Now, we have info $(g_e(t^n, y^*)) = \varnothing Z t^{r_e}$ by Theorem (7.19). This means that **58** we have

$$\text{info}\,(b_1(t)) = ct^{r_e} \tag{7.20.2}$$

for some $c \in k$, $c \neq 0$, and

$$\text{ord}_t\, b_i(t) > r_e \quad \text{for } i \neq 1. \tag{7.20.3}$$

Let $\varphi : k'((t))[Z] \to k'((t))$ be the $k'((t))$-algebra homomorphism defined by $\varphi(Z) = a$. Then $\varphi(y^*) = \overline{y}$. Therefore we have

$$\begin{aligned} g_e(t^n, \overline{y}) &= \varphi(g_e(t^n, y^*)) \\ &= \sum_{i=0}^{p} b_i(t) a^i \qquad \text{(by 7.20.1).} \end{aligned} \tag{7.20.4}$$

Let

$$v = b_0(t) + (b_1(t) - ct^{r_e})a + \sum_{i=2}^{p} b_i(t) a^i.$$

Then by (7.20.2) and (7.20.3) we have $\text{ord}_t\, v > r_e$, and from (7.20.4) we get $g_e(t^n, \overline{y}) = cat^{r_e} + v$. □

8 The Fundamental Theorem

Throughout this section we preserve the notation of (7.1). In addition, we also fix the following notation:

(8.1) NOTATION. For an integer e, $1 \leq e \leq h+1$, we get

$$g_e = g_e(X, Y) = \begin{cases} Y, & \text{if } e = 1, \\ App_Y^{d_e}(f), & \text{if } e \geq 2. \end{cases}$$

We note that $g_{h+1} = f$.

(8.2) Fundamental Theorem (Part One).

59 Let e be an integer, $1 \leq e \leq h+1$. Then we have $\operatorname{ord}_t g_e(t^n, y(t)) = r_e$.

Proof. Since $g_{h+1} = f$ and $r_{h+1} = \infty$, the assertion is clear for $e = h+1$. Next, we have $g_1(t^n, y(t)) = y(t)$, and $\operatorname{ord}_t y(t) = m_1 = r_1$, which proves the assertion for $e = 1$. Assume now that $2 \leq e \leq h$. Then the assertion is immediate from Corollary (7.20) by taking $a = y_{m_e}$ and $u = \sum\limits_{j>m_e} y_j t^i$ and noting that $y_{m_e} \neq 0$. $\square$

(8.3) Fundamental Theorem (Part Two)

Let R be a subring of $k((X))$ such that n is a unit in R and $f \in R[Y]$. Then:

(i) $g_i \in R[Y]$ for every i, $1 \leq i \leq h+1$.

Further, let $\overline{R[Y]} = R[Y]/fR[Y]$ and let $\overline{g_i}$ be the image of g_i under the canonical map $R[Y] \to \overline{R[Y]}$. Then:

(ii) $\overline{R[Y]}$ is a free R-module with the set $\left\{\overline{g}^b \middle| b \in B\right\}$ as a free basis, where $\overline{g} = (\overline{g_1}, \ldots, \overline{g_h})$ and

$$B = \left\{b = (b_1, \ldots, b_h) \in \mathbb{Z}^h \middle| 0 \leq b_i < d_i/d_{i+1} \text{ for } 1 \leq i \leq h\right\}.$$

(For $h = 0$ interpret this notation as $B = \{\phi\}$ and $\left\{\overline{g}^h \middle| b \in B\right\} = \{1\}$.)

Proof.

1. For $i = 1$, $g_1 = Y \in R[Y]$. For $i \geq 2$ the assertion follows from the uniqueness of $App_Y^{d_i}(f)$.

2. We first note that since $\deg_Y f = n > 0$ and since f is monic in Y, the restriction of the canonical map $\eta : R[Y] \to \overline{R[Y]}$ to R is injective. We identify R with its image in $\overline{R[Y]}$. Then, writing $\overline{F} = \eta(F)$ for $F \in R[Y]$, we have

   (8.3.1) $\qquad \overline{F} = F$ for every $F \in R$.

□

Now, let $G_i = g_i$ for $1 \leq i \leq h+1$. Then the $(h+1)$-tuple $G = (G_1, \ldots, G_{h+1})$ satisfies conditions (i)-(iii) of (2.2) with $p = h+1$.
60 Therefore by Corollary (2.14) every element F of $R[Y]$ has a unique expression of the form

(8.3.2) $$F = \sum_{a \in A(G)} F_a G^a, \quad F_a \in R,$$

where

$$A(G) = \left\{ a = (a_1, \ldots, a_{h+1}) \in \mathbb{Z}^{h+1} \middle| 0 \leq a_i < n_i(G) \text{ for } 1 \leq i \leq h+1 \right\}.$$

Recall that with the notation of (2.2) we have $n_{h+1}(G) = \infty$ and

(8.3.3) $$n_i(G) = (n/d_{i+1})/(n/d_i) = d_i/d_{i+1}$$

for $1 \leq i \leq h$. Now, let $\overline{F}$ be any element of $\overline{R[Y]}$ and let $F \in R[Y]$ be a lift of $\overline{F}$. Then from (8.3.1) and (8.3.2) we get

(8.3.4) $$\overline{F} = \sum_{a \in A(G)} F_a \overline{G}^a.$$

Now $\overline{G}_{h+1} = 0$. Therefore, if $a \in A(G)$ is such that $a_{h+1} \neq 0$ then $\overline{G}^a = 0$. Therefore, in view of (8.3.3), the expression (8.3.4) reduces to the form

$$\overline{F} = \sum_{b \in B} F'_b \overline{g}^b,$$

where $F'_b = F_{(b_1, \ldots, b_h, 0)}$ for $b \in B$. This proves that $\overline{R[Y]}$ is generated as an R-module by the set $\left\{ \overline{g}^b \middle| b \in B \right\}$. Now, to prove that this set is a free basis, suppose

(8.3.5) $$0 = \sum_{b \in B} F'_b \overline{g}^b$$

with $F'_b \in R$ for every b and $F'_b = 0$ for almost all b. For $a \in A(G)$, define

(8.3.6) $$F_a = \begin{cases} F'_{(a_1, \ldots, a_h)}, & \text{if } a_{h+1} = 0, \\ 0, & \text{if } a_{h+1} \neq 0. \end{cases}$$

61 Let

$$F = \sum_{a \in A(G)} F_a G^a.$$

It is enough to prove that $F = 0$. For, this would imply by the uniqueness of the expression (8.3.2) that $F_a = 0$ for every $a \in A(G)$, which would prove, in view of (8.3.6), that $F'_b = 0$ for every $b \in B$. Now, suppose $F \neq 0$. Then, since f divides F in $F[Y]$ by (8.3.5), we have $F \notin R$ and $\deg F \geq \deg f = \deg G_{h+1}$. But this is a contradiction by (8.3.6) and Lemma (2.12). Therefore $F = 0$, and the proof of the theorem is complete.

(8.4) LEMMA. *Let $k((X))$ be identified with the subfield $k((t^n))$ of $k((t))$ by putting $X = t^n$. Let R be a subring of $k((X))$ such that $f \in R[Y]$. Let $R[y(t)]$ be the R-subalgebra of $k((t))$ generated by $y(t)$. Then:*

(i) $R[y(t)] = \left\{ F(t^n, y(t)) \middle| F(X, Y) \in R[Y] \right\}$.

(ii) *There exists an R-algebra isomorphism*

$$\overline{u} : R[Y]/fR[Y] \to R[y(t)]$$

which fits in a commutative diagram

$$\begin{tikzcd} R[Y] \arrow[rr, "\eta"] \arrow[dr, "u"'] & & R[Y]/fR[Y] \arrow[dl, "\overline{u}"] \\ & R[y(t)] & \end{tikzcd}$$

where η is the canonical homomorphism and u is defined by $u(F(x, Y)) = F(t^n, y(t))$ for $F(X, Y) \in R[Y]$.

Proof.

(i) This is clear.

(ii) It is clear that u is an R-algebra homomorphism. Since $u(f) = f(t^n, y(t) = 0$, u factors via η to give $\overline{u}$. Since u is surjective by (i), so is $\overline{u}$. To show that $\overline{u}$ is injective, it is enough to show that $\ker u =$ 62

$fR[Y]$. Let $F(X,Y) \in \ker u$. Then $F(t^n, y(t)) = 0$. Therefore, since f is the minimal monic polynomial of $y(t)$ over $k((t^n))$, f divides $F(X,Y)$ in $k((X))[Y]$. Since f is monic, it that follows f divides $F(x,Y)$ in $R[Y]$.

□

(8.5) Fundamental Theorem (Part Three)

Let $k((X))$ be identified with the subfield $k((t^n))$ of $k((t))$ by putting $X = t^n$. Let R be a subring of $k((X))$ such that n is a unit in R and $f \in R[Y]$. Let

$$R[y(t)] = \left\{ F(t^n, y(t)) \middle| F(X,Y) \in R[Y] \right\}.$$

Let $\overline{g}_i = g_i(t^n, y(t))$, $1 \leq i \leq h$. Then:

(i) $R[y(t)]$ is a free R-module with the set $\{\overline{g}^b | b \in B\}$ as a free basis, where $\overline{g} = (\overline{g}_1, \ldots, \overline{g}_h)$ and

$$B = \left\{ b = (b_1, \ldots, b_h) \in \mathbb{Z}^h \middle| 0 \leq b_i < d_i/d_{i+1} \text{ for } 1 \leq i \leq h \right\}.$$

(ii) Let $F \in R[y(t)]$ and let

$$F = \sum_{b \in B} F_b \overline{g}^b, \quad F_b \in R.$$

If $b, b' \in B$, $b \neq b'$ and $F_b \neq 0$, $F_{b'} \neq 0$ then

$$\operatorname{ord}_t(F_b \overline{g}^b) \neq \operatorname{ord}_t(F_{b'} \overline{g}^{b'}).$$

In particular, if $F \neq 0$ then there exists a unique $b \in B$ such that $\operatorname{ord}_t F = \operatorname{ord}_t(F_b \overline{g}^b)$.

(iii) With the notation of (ii), let $b \in B$ be the unique element such that $\operatorname{ord}_t(F) = \operatorname{ord}_t(F_b \overline{g}^b)$. Then

$$\operatorname{ord}_t F = \operatorname{ord}_t F_b + \sum_{i=1}^{h} b_i r_i.$$

Proof.

63 (i) We first note that by Theorem (8.3) we have $g_i \in R[Y]$ for $1 \leq i \leq h$. Let us now identify $R[Y(t)]$ and $\overline{R[Y]} = R[Y]/fR[Y]$ as R-algebras via the isomorphism $\overline{u}$ of Lemma (8.4). With this identification, $\overline{g}_i$ is the image of g_i under the canonical map $R[Y] \to \overline{R[Y]}$. Therefore (i) follows directly from Theorem (8.3).

(ii) Let $\Gamma_+(R) = \left\{(n/r_0)\operatorname{ord}_X G \,\middle|\, G \in R, G \neq 0\right\}$. Then, since $n = |r_0|$, it is clear that $\Gamma_+(R)$ is a subsemigroup of $\mathbb{R}$ is a subsemigroup of $\mathbb{Z}$. For $b = (b_1, \ldots, b_h) \in B$ such that $F_b \neq 0$, let us define $b_0 = (n/r_0)\operatorname{ord}_X F_b$. Then $b_0 \in \Gamma_+(R)$. Since $r_i = \operatorname{ord}_t \overline{g}_i$ by Theorem (8.2), we get

$$\text{(8.5.1)} \qquad \operatorname{ord}_t(F_b\overline{g}^b) = \operatorname{ord}_t F_b + \sum_{i=1}^{h} b_i r_i = \sum_{i=0}^{h} b_i r_i,$$

since $b_0 r_0 = n \operatorname{ord}_X F_b = \operatorname{ord}_t F_b$. Similarly, if $b' \in B$ and $F_{b'} \neq 0$ then

$$\text{(8.5.2)} \qquad \operatorname{ord}_t(F_{b'}\overline{g}^{b'}) = \sum_{i=0}^{h} b'_i r_i, \quad b'_0 \in \Gamma_+(R).$$

Now, since b, $b' \in B$, we have $0 \leq b_i < d_i/d_{i+1}$, $0 \leq b'_i < d_i/d_{i+1}$ for $1 \leq i \leq h$. Thus (8.5.1) and (8.5.2) are $\Gamma_+(R)$-strict linear combinations of $r = (r_0, \ldots, r_h)$. Therefore (ii) follows from Proposition (1.5).

(iii) This was proved in (8.5.1) above.

□

(8.6) DEFINITION. Let R be a subring of $k((X))$ such that $f \in R[Y]$. Let $w \in \mu_n(k)$. The set

$$\left\{\operatorname{ord}_t F(t^n, y(wt)) \,\middle|\, F(X,Y) \in R[Y], F(t^n, y(wt)) \neq 0\right\}.$$

which is clearly independent of $w \in \mu_n(k)$ and is a subsemigroup of $\mathbb{Z}$, is called the *value semigroup* of f *with respect to* R and is denoted $\Gamma_R(f)$. 64

(8.7) Fundamental Theorem (Part Four).

Let R be a subring of $k((X))$ such that n is a unit in R and $f \in R[Y]$. Let

$$\Gamma_+(R) = \left\{ (n/r_0) \operatorname{ord}_X F \middle| F \in R, F \neq 0 \right\}.$$

Then we have:

(i) $\Gamma_|(R)$ is a subsemigroup of $\mathbb{Z}$.

(ii) $\Gamma_+(R) r_0 \subset \Gamma_R(f)$ and $r_i \in \Gamma_R(f)$ for every i, $1 \leq i \leq h$.

(iii) $\Gamma_R(f)$ is $\Gamma(R)$-strictly generated by $r = (r_0, \ldots, r_h)$.

In particular, suppose we are in one of the following two cases:

(1) The ALGEBROID CASE: $R = k'[[X]]$ for some subfield k' of k, $f \in R[Y]$ and $r_0 = n$.

(2) The PURE MEROMORPHIC CASE: $R = k'[X^{-1}]$ for some subfield k' of k, $f \in R[Y]$ and $r_0 = -n$.

Then we have:

(i') $\Gamma_+(R) = \mathbb{Z}^+$.

(ii') $r_i \in \Gamma_R(f)$ for every $i, 0 \leq i \leq h$.

(iii') $\Gamma_R(f)$ is strictly generated by $r = (r_0, \ldots, r_h)$.

(For the definition of $\Gamma_+(R)$-strict generation, see (1.7))

Proof.

(i) This is clear, since $n = |r_0|$.

(ii) Let $\gamma \in \Gamma_+(R)$. Then there exists $F = F(X) \in R$ such that $F \neq 0$ and $\gamma = (n/r_0) \operatorname{ord}_X F$. This gives $\gamma r_0 = n \operatorname{ord}_X F = \operatorname{ord}_t f(t^n)$, which shows that $\gamma r_0 \in \Gamma_R(f)$. Next, since $g_i \in R[Y]$ by Theorem (8.3) and since $\operatorname{ord}_t g_i(t^n, y(t)) = r_i$ by Theorem (8.2), we get $r_i \in \Gamma_R(f)$ for $1 \leq i \leq h$.

(iii) Let $\gamma \in \Gamma_R(f)$ and let $F(X, Y) \in R[Y]$ be such that $\gamma = \operatorname{ord}_t F(t^n, y(t))$. Put $F = F(t^n, y(t))$. Then $F \neq 0$. Therefore by Theorem (8.5) (iii) we have

$$\gamma = \operatorname{ord}_t F = \operatorname{ord}_t F_b(t^n) + \sum_{i=1}^{h} b_i r_i,$$

where $F_b = F_b(X) \in R$, $F_b \neq 0$, and $b_i \in \mathbb{Z}$, $0 \leq b_i < d_i/d_{i+1}$ for **65**
$1 \leq i \leq h$. Let $b_0 = (n/r_0) \operatorname{ord}_X F_b$. Then $b_0 \in \Gamma_+(R)$ and we have $\operatorname{ord}_t F_b(t^n) = b_0 r_0$. Therefore $\gamma = \sum\limits_{i=0}^{h} b_i r_i$, which shows that γ is a $\Gamma_+(R)$-strict linear combination of r. Conversely, if $\gamma = \sum\limits_{i=0}^{h} \gamma_i r_i$ is a $\Gamma_+(R)$-strict linear combination of r then it follows (ii) that $\gamma \in \Gamma_R(f)$. This proves (iii).

(i') is clear, and (ii'), (iii') follow from (i'), (ii) and (iii). □

(8.8) COROLLARY. With the notation of Theorem (8.7), suppose R contains an element of X-order 1 or -1. (This condition is satisfied, for example, if $X \in R$ or $X^{-1} \in R$). Them g.c.d. $(\Gamma_R(f)) = 1$, i.e., the subgroup of $\mathbb{Z}$ generated by $\Gamma_R(f)$ coincides with $\mathbb{Z}$.

Proof. By assumption, we have $n/r_0 \in \Gamma_+(R)$ or $-n/r_0 \in \Gamma_+(R)$. Therefore by Theorem (8.7) (ii), $n \in \Gamma_R(f)$ or $-n \in \Gamma_R(f)$. Since $n = |r_o|$, we get $r_o \in \Gamma_R(f)$ or $-r_o \in \Gamma_R(f)$. Also $r_i \in \Gamma_R(f)$ for $1 \leq i \leq h$ by Theorem (8.7)(ii). Now, since

$$g.c.d.\ (-r_0, r_1, \ldots r_h) = g.c.d.\ (r_0, r_1, \ldots, r_h) = d_{h+1} = 1,$$

the corollary follows. □

Chapter 4

Applications of The Fundamental Theorem

9 Epimorphism Theorem

Let k be a field and let X, Y, Z, τ be indeterminates. **66**

(9.1) DEFINITION. Let C be a finitely generated k-subalgebra of $k[Z]$ such that the quotient field of C is $k(Z)$. We call C (the coordinate ring of) an *affine polynomial curve* over k and we call $k(Z)$ the *function field* of C. If, moreover, C is generated as a k-algebra by two elements then we call C an affine polynomial *plane* curve. A k-algebra epimorphism (i.e., surjective homomorphism) $\alpha : k[X,Y] \to C$ is called an *embedding* of C in the affine plane over k.

Note that if C has an embedding in the affine plane then C is a plane curve. Moreover, the mapping $\alpha \mapsto (\alpha(X), \alpha(Y))$ gives a bijective correspondence between the embeddings of C in the affine plane and ordered pair (x,y) of elements of C such that $C = k[x,y]$.

(9.2) DEFINITION. An embedding $\alpha : k[X,Y] \to C$ is said to be *permissible* if $\alpha(X) \neq 0$ and char k does not divide $\deg_Z \alpha(X)$.

(9.3) Equation of an Embedding

Let $\alpha : k[X,Y] \to C$ be a permissible embedding of an affine plane polynomial curve C. Let $\overline{x} = \alpha(X)$, $\overline{y} = \alpha(Y)$. Then $\alpha(F) = F(\overline{x}, \overline{y})$ for every $F = F(X, Y) \in k[X, Y]$. Let $n = \deg_Z \overline{x}$. Let $\overline{k}$ be the algebraic closure of k and let $\theta : \overline{k}[Z] \to \overline{k}((\tau))$ be the $\overline{k}$-algebra monomorphism defined by $\theta(Z) = \tau^{-1}$. Then it is clear that we have

$$\text{(9.3.1)} \qquad \operatorname{ord}_\tau \theta(F(\overline{x}, \overline{y})) = -\deg_Z F(\overline{x}, \overline{y})$$

67 for every $F(X, Y) \in \overline{k}[X, Y]$. In particular, we have $\operatorname{ord}_\tau \theta(\overline{x}) = -n$. Since char k does not divide n, there exists, by Corollary (5.4), an element $t \in \overline{k}((\tau))$ such that $\operatorname{ord}_\tau t = 1$ and $\theta(\overline{x}) = t^{-n}$. Note then that we have $\overline{k}((t)) = \overline{k}((\tau))$ and $\operatorname{ord}_t a = \operatorname{ord}_\tau a$ for every $a \in \overline{k}((t))$. Write $x = x(t) = \theta(\overline{x}) = t^{-n}$ and $y = y(t) = \theta(\overline{y})$. We call $y(t)$ a *Newton-Puiseux expansion* of $\overline{y}$ in fractional powers of $\overline{x}^{-1}$. Let $f = f(x, Y) \in \overline{k}((X))[Y]$ be the minimal monic polynomial of y over $\overline{k}((t^n))$ (Definition (5.8)). Recall that f is the unique irreducible element of $\overline{k}((X))[Y]$ such that f is monic in Y and $f(t^n, y) = 0$. We call f the *meromorphic equation* of the embedding α.

(9.4) LEMMA. *With the notation of (9.3), we have:*

(i) $\deg_Y f = n$.

(ii) $f \in k[X^{-1}, Y]$.

Proof. (i) We have $\deg_Y f = [\overline{k}((t^n))(y) : \overline{k}((t^n))]$. Therefore, since $y \in \overline{k}((t))$, we get

$$\deg_Y f \leq [\overline{k}((t)) : \overline{k}((t^n))] = n.$$

On the other hand, since α is surjective, we have $Z \in k(\overline{x}, \overline{y})$. Therefore $\tau^{-1} \in k(x, y) \subset \overline{k}((t^n))(y)$, so that $\tau \in \overline{k}((t^n))(y)$. Therefore

$$\deg_Y f \geq [\overline{k}((t^n))(\tau) : \overline{k}((t^n))] = n$$

by Lemma (5.10), since $1 \in \operatorname{Supp}_t(\tau)$. This proves (i).

(ii) Since $\deg_Z \overline{x} = n > 0$, $\overline{x}$ is transcendental over k and $k(Z)$ is algebraic over $k(\overline{x})$ with $[k(Z) : k(\overline{x})] = n$. Let $g(\overline{x}, Y) \in k(\overline{x})[Y]$ be the minimal monic polynomial of $\overline{y}$ over $k(\overline{x})$. Since α is surjective, we have $k(\overline{x})(\overline{y}) = k(Z)$. Therefore $\deg_Y g(\overline{x}, Y) = n$. We claim that
68 $g(\overline{x}, Y \in k[\overline{x}][Y])$. In order to prove the claim, we have only to show that

$\overline{y}$ is integral over $k[\overline{x}]$. Now, writing $\overline{x} = \sum_{i=1}^{n} a_i Z^i$, $a_i \in k$ for $0 \leq i \leq n$, $a_n \neq 0$, we have

$$Z^n + \sum_{i=1}^{n-1} a_i a_n^{-1} Z^i + (a_0 - \overline{x}) a_n^{-1} = 0,$$

which shows that Z is integral over $k[\overline{x}]$. Since $\overline{y} \in k[Z]$, $\overline{y}$ is also integral over $k[\overline{x}]$. Thus $g(\overline{x}, Y) \in k[\overline{x}][Y]$. Put $h(X,Y) = g(X^{-1}, Y)$. Then $h(X,Y) \in k[X^{-1}][Y] \subset \overline{k}((X))[Y]$ and $h(X,Y)$ is monic in Y with $\deg_Y h(X,Y) = n$. Now, $h(t^n, y) = g(t^{-n}, y) = g(\theta(\overline{x}), \theta(\overline{y})) = \theta(g(\overline{x}, \overline{y})) = 0$. This shows that $f(X,Y) = h(X,Y)$ and (ii) is proved. $\square$

(9.5) REMARK. Put $\varphi = \varphi(X,Y) = f(X^{-1}, Y)$. Then by Lemma (9.4) $\varphi \in k[X,Y]$. We claim that φ generates $\ker \alpha$. To see this we note that $\ker \alpha$ is a principal prime ideal of $k[X,Y]$ and, since f is irreducible in $k[X^{-1}, Y]$, φ is irreducible in $k[X,Y]$. Therefore it is enough to show that $\varphi \in \ker \alpha$. Now, $\theta(\varphi(\overline{x}, \overline{y})) = \varphi(t^{-n}, y) = f(t^n, y) = 0$. Since θ is a monomorphism, our claim is proved. Noting that φ is the unique generator of $\ker \alpha$ which is monic in Y, we call φ the *algebraic equation* of the embedding α. If ψ is any generator of $\ker \alpha$ then, clearly, we have $\psi = \varnothing \varphi$ for some $\varnothing$.

(9.6) REMARK. With the notation of (9.3), suppose S is a subring of k such that $\overline{x}$ and $\overline{y}$ belong to $S[Z]$. Consider the pair $(X - \overline{x}, Y - \overline{y})$ of elements of $S[Z][X,Y]$ and let $g = g(X,Y) \in S[X,Y]$ be the Z-resultant of $X - \overline{x}$ and $Y - \overline{y}$. Then clearly $\varnothing g$ is monic in Y and, since $\deg_Z \overline{x} = n$, we have $\deg_Y g = n$. Moreover, we have $g(\overline{x}, \overline{y}) = 0$, so that $0 = \theta(g(\overline{x}, \overline{y})) = g(t^{-n}, y)$. therefore it follows from Lemma (9.4) (i) that $\varnothing f(X,Y) = g(X^{-1}, Y) \in S[X^{-1}, Y]$. This gives an alternative proof of part (ii) of Lemma (9.4).

(9.7) Characteristic Sequences of an Embedding

Continuing with the notation of (9.3), let $R = k[X^{-1}]$. Then $f \in R[Y]$ by **69**
Lemma (9.4). Let $h = h(f)$ and let $m_i = m_i(-n, f)$, $q_i = q_i(-n, f)$,

$s_i = s_i(-n, f)$, $r_i = r_i(-n, f)$, $d_{i+1} = d_{i+1}(f)$ for $0 \leq i \leq h+1$. The sequence $(m_0, \ldots, m_{h+1})$, (resp. $(q_0, \ldots, q_{h+1})$, resp. $(s_0, \ldots, s_{h+1})$, resp. $(r_0, \ldots, r_{h+1})$, resp. $(d_1, \ldots, d_{h+2})$) is called the *characteristic* m (resp. q, resp. s, resp. r, resp. d)- *sequence* of the permissible embedding α. Note that we have

$$r_0 = -n = -\deg_Z \alpha(X). \tag{9.7.1}$$

Moreover, by (9.3.1) we have

$$r_1 = \operatorname{ord}_t y = \operatorname{ord}_\tau y = -\deg_Z \alpha(Y). \tag{9.7.2}$$

Let

$$\mathbb{Z}^- = \{a \in \mathbb{Z} \big| a \leq 0\}.$$

Recall that $\Gamma_R(f)$ is the subsemigroup of $\mathbb{Z}$ defined by

$$\Gamma_R(f) = \left\{\operatorname{ord}_t F(t^n, y) \Big| F(X, Y) \in R[Y], F(t^n, y) \neq 0\right\}.$$

(9.8) LEMMA. *With the notation of (9.7), we have:*

(i) $\Gamma_R(f) \subset \mathbb{Z}^-$.

(ii) If $C = k[Z]$ *then* $\Gamma_R(f) = \mathbb{Z}^-$.

(iii) $\Gamma_R(f)$ *is strictly generated by* $r = (r_0, r_1, \ldots, r_h)$.

(iv) $r_0 < 0$, $r_1 = \infty$ *or* $r_1 \leq 0$, *and* $r_i < 0$ *for* $2 \leq i \leq h$.

(v) If $C = k[Z]$ *and* $h \geq 2$ *then* $r_h = -1$.

Proof. (i) Let $F(X, Y) \in R[Y]$ be any element such that $F(t^n, y) \neq 0$. Put
70 $G(X, Y) = F(X^{-1}, Y)$. Then $G(X, Y) \in k[X, Y]$ and, with the notation of (9.3), we have

$$\begin{aligned} \operatorname{ord}_t F(t^n, y) &= \operatorname{ord}_t G(t^{-n}, y) \\ &= \operatorname{ord}_\tau G(t^{-n}, y) \\ &= \operatorname{ord}_\tau G(\theta(\overline{x}), \theta(\overline{y})) \\ &= \operatorname{ord}_\tau \theta(G(\overline{x}, \overline{y})) \\ &= -\deg_Z G(\overline{x}, \overline{y}) \end{aligned} \tag{9.8.1}$$

by (9.3.1). Therefore $\mathrm{ord}_t F(t^n, y) \leq 0$. this proves (i)

(ii) In view of (i), it is enough to prove that $-1 \in \Gamma_R(f)$. Since α is surjective, thee exists $G(X, Y) \in k[x, Y]$ such that $G(\overline{x}, \overline{y}) = Z$. Put $F(X, Y) = G(X^{-1}, Y)$. Then $F(X, Y) \in R[Y]$ and $G(X, Y) = F(X^{-1}, Y)$. Therefore by the computation (9.8.1) we get $\mathrm{ord}_t F(t^n, y) = -\deg_Z Z = -1$. This shows that $-1 \in \Gamma_R(f)$.

(iii) This is immediate from Theorem (8.7) (iii′).

(iv) The assertion about r_0 and r_1 follows from (9.7.1) and (9.7.2). Now suppose $2 \leq i \leq h$. Then we have

$$d_i > d_{i+1} \tag{9.8.2}$$

by Proposition (6.13) (ii). Therefore $1 < d_i/d_{i+1}$, so that r_i is a strict linear combination of $r = (r_0, r_1, \ldots, r_h)$. Therefore $r_i \in \Gamma_R(f)$ by (iii), which shows by (i) that $r_i \leq 0$. Since d_i does not divide r_i by (9.8.2), we have $r_i \neq 0$. Therefore $r_i < 0$.

(v) It follows from (ii) and (iii) that $r_i = -1$ for some i, $0 \leq i \leq h$. Since $h \geq 2$ and since d_h divides r_i for $i \leq h - 1$ it follows from (9.8.2) that $r_i \neq -1$ for $0 \leq i \leq h - 1$. Therefore $r_h = -1$. □ **71**

(9.9) LEMMA. *With the notation of (9.7), suppose $-d_2 \in \Gamma_R(f)$. Then r_0 divides r_1 or r_1 divides r_0.*

Proof. Since $-d_2 \in \mathbb{Z}$, we have $d_2 \neq -\infty$. This means that $h \geq 1$. Therefore $r_1 \neq \infty$ and it follows from Lemma (9.8) (iv) that $r_i \leq 0$ for $i = 0, 1$. since $-d_2 \in \Gamma_R(f)$, Lemma (9.8) (iii) shows that $-d_2$ is a strict linear combination of $r = (r_0, \ldots, r_h)$. Now, the assertion follows from Proposition (1.8). □

(9.10) DEFINITION. If $C = k[Z]$, we call C the *affine line* over k.

In Theorem (9.11) and (9.19) below we study the embeddings of the affine line in the affine plane.

(9.11) Epimorphism Theorem (First Formulation)

Let k be any field and let $\alpha : k[X, Y] \to k[Z]$ be a k-algebra epimorphism such that $\alpha(X) \neq 0$, $\alpha(Y) \neq 0$. Let $n = \deg_Z \alpha(X)$, $m = \deg_Z \alpha(Y)$.

Suppose char k does not divide g.c.d. (m, n). Then n divides m or m divides n.

Proof. By the symmetry of the assertion, we may assume that char k does not divide n. Then α is a permissible embedding. We now use the notation of (9.3) and (9.7) with $C = k[Z]$. By (9.7.1) and (9.7.2) we have $r_0 = -n$ and $r_1 = -m \neq \infty$. Therefore $h \geq 1$. By Lemma (9.8) (ii) we have $\Gamma_R(f) = \mathbb{Z}^-$. Therefore $-d_2 \in \Gamma_R(f)$, so that r_0 divides r_1 or r_1 divides r_0 by Lemma (9.9). This means that n divides m or m divides n, and the theorem is proved. □

The following example shows that in Theorem (9.11) we cannot relax the condition "char k does divide g.c.d. (m, n)".

(9.12) EXAMPLE. Let p = char k. Let e, s be positive integers and let

$$x = Z^{p^e}$$

$$y = Z + \sum_{i=0}^{s} a_i Z^{ip}$$

72 with $a_i \in k$ for $0 \leq i \leq s$ and $a_s \neq 0$. Let $\alpha : k[X, Y] \to k[Z]$ be the k-algebra homomorphism defined by $\alpha(X) = x$, $\alpha(Y) = y$. We claim that α is surjective. To prove our claim, it is enough to show that $Z \in k[x, y]$. In fact, we show by descending induction on j that $Z^{p^j} \in k[x, y]$ for $0 \leq j \leq e$, this assertion being clear for $j = e$. Suppose now that $j \geq 0$ and $Z^{p^{j+1}} \in k[x, y]$. We have

$$y^{p^j} = Z^{p^j} + \sum_{i=0}^{s} a_i^{p^j} (Z^{p^{j+1}})^i.$$

This shows that $Z^{p^j} \in k[x, y]$, and our claim is proved. Now, let $n = \deg_Z x = p^e$, $m = \deg_Z y = sp$. It is clear that we can choose e, s to be such that neither n divides m nor m divides n. Specifically, take $e \geq 2$ and $s = qp^c$ where q, c are integers such that $q \geq 2$, $q \not\equiv 0 \pmod{p}$ and $0 \leq c \leq e - 2$.

(9.13) QUESTION. Let $\alpha : k[X, Y] \to k[Z]$ be a k-algebra epimorphism such that $\alpha(X) \neq 0$, $\alpha(Y) \neq 0$. Let $n = \deg_Z \alpha(X)$, $m = \deg_Z \alpha(Y)$. Let $p = \operatorname{char} k$, and let $n = n'p^e$, $m = m'p^d$, whee n', m', e, d are integers such that $n' \not\equiv 0 \pmod{p}$, $m' \not\equiv 0 \pmod{p}$, $e \geq 0$, $d \geq 0$. Is it then true that n' divides m' or m' divides n'?

(9.14) DEFINITION. Let $A = k[X, Y]$ and let σ be a k-algebra automorphism of A. We say σ is *primitive* if there exists $P(Z) \in k[Z]$ such that

$$\begin{array}{lll} \text{either} & \sigma(X) = X, & \sigma(Y) = Y + P(X); \\ \text{or} & \sigma(X) = X + P(Y), & \sigma(Y) = Y. \end{array}$$

We say σ is *linear* if there exist $a_i, b_i, c_i \in k$, $i = 1, 2$, such that

$$\sigma(X) = a_1X + b_1Y + c_1, \quad \sigma(Y) = a_2X + b_2Y + c_2.$$

We say σ is *elementary* if σ is primitive or linear. We say σ is *tame* 73
if σ is a finite product of elementary automorphisms.

(9.15) REMARK. It is easily checked that the set of all tame automorphisms of A is a subgroup of the group of all k-algebra automorphisms of A. In fact, it is true that all k-algebra automorphisms of A are tame. In the next section we shall deduce this fact from the Epimorphism Theorem in case $\operatorname{char} k = 0$ (Theorem (10.1))

(9.16) DEFINITION. Let $\alpha, \beta : k[X, Y] \to k[z]$ be k-algebra epimorphisms. We say α is *equivalent* (resp. *tamely equivalent*) to β if there exists a k-algebra automorphism (resp. tame automorphism) σ of $k[X, Y]$ such that the diagram

$$\begin{tikzcd} k[x, Y] \arrow[dr, "\alpha"] \arrow[dd, "\sigma"'] & \\ & k[Z] \\ k[X, Y] \arrow[ur, "\beta"'] & \end{tikzcd}$$

is commutative, i.e., $\alpha = \beta\sigma$.

(9.17) REMARK. It is clear that both equivalence and tame equivalence are equivalence relations and that tame equivalence implies equivalence.

(9.18) DEFINITION. Let $\alpha : k[X, Y] \to k[Z]$ be a k-algebra epimorphism. We say α is *wild* if $\alpha(X) \neq 0$, $\alpha(Y) \neq 0$ and char k divides both $\deg_Z \alpha(X)$ and $\deg_Z \alpha(Y)$.

(9.19) EPIMORPHISM THEOREM (SECOND FORMULATION).

Let $\alpha, \beta : k[X, Y] \to k[Z]$ be k-algebra epimorphisms. Assume that
74 neither α nor β is wild. Then α and β are tamely equivalent. In particular, α and β are equivalent.

Proof. Let $\gamma : k[X, Y] \to k[Z]$ be the k-algebra epimorphism defined by $\gamma(X) = Z$, $\gamma(Y) = 0$. Then, since tame equivalence is an equivalence relation, it is enough to prove the following assertion: $\square$

(9.19.1)

If α is not wild then α and γ are tamely equivalent.

Given α, we define the *transpose* α^t of α to be the k-algebra epimorphism $\alpha^t : k[X, Y] \to k[Z]$ given by $\alpha^t(X) = \alpha(Y)$, $\alpha^t(Y) = \alpha(X)$. Clearly, α and α^t are tamely equivalent and α is wild if and only if α^t is wild. Put $D(\alpha) = \deg_Z \alpha(X) + \deg_Z \alpha(Y)$. Then $D(\alpha) = D(\alpha^t)$. We now prove (9.19.1) by induction on $D(\alpha)$. First, suppose $D(\alpha) \leq 1$. Replacing α by α^t, if necessary, we may assume that $\deg_Z \alpha(X) \geq \deg_Z \alpha(Y)$. Then, since α is surjective, the assumption $D(\alpha) \leq 1$ implies that $\deg_Z \alpha(Y) \leq 0$ and $\deg_Z \alpha(X) = 1$. This means that there exist $a, b, c, \in k$, $a \neq 0$, such that $\alpha(X) = aZ + b$ and $\alpha(Y) = c$. Let σ be the k-algebra automorphism of $k[X, Y]$ defined by $\sigma(X) = a(X) + b$, $\sigma(Y) = Y + c$. Then σ is tame and clearly we have $\alpha = \gamma\sigma$.

Now, suppose $D(\alpha) \geq 2$. Again, replacing α by α^t, if necessary, we may assume that $\deg_Z \alpha(X) \geq \deg_Z \alpha(Y)$. This means, in particular, that $\alpha(X) \notin k$. If $\alpha(Y) \in k$ then $\deg_Z \alpha(X) \geq 2$. This is not possible, since

α is surjective. Therefore $\alpha(X) \notin k$ and $\alpha(Y) \notin k$. Let $n = \deg_Z \alpha(X)$, $m = \deg_X \alpha(Y)$. Since α is not wild and $n \geq m \geq 1$, it follows from Theorem (9.11) that m divides n. Let $n = rm$, where r is a positive integer. Write

$$\alpha(X) = \sum_{i=0}^{rm} a_i Z^i, \quad \alpha(Y) = \sum_{j=0}^{m} b_j Z^j$$

with $a_i, b_j \in k$ for $0 \leq i \leq rm$, $0 \leq j \leq m$ and $b_m \neq 0$. Let σ be the 75
$k-$algebra automorphism of $k[X, Y]$ defined by $\sigma(X) = X - a_{rm} b_m^{-r} Y^r$ and $\sigma(Y) = Y$. Then σ is primitive, therefore tame. Let $\alpha' = \alpha\sigma$. Then $\alpha' : k[X, Y] \to k[Z]$ is a k-algebra epimorphism and α and α' are tamely equivalent. Now, we have

$$\begin{aligned} \alpha'(X) &= \alpha(\sigma(X)) \\ &= \alpha(X - a_{rm} b_m^{-r} Y^r) \\ &= \sum_{i=0}^{rm} a_i Z^i - a_{rm} b_m^{-r} \left(\sum_{j=0}^{m} b_j X^j \right)^r . \end{aligned}$$

This shows that $deg_X \alpha'(X) < rm = n$. Moreover, $\alpha'(Y) = \alpha(\sigma(Y)) = \alpha(Y)$. Therefore $\deg_Z \alpha'(Y) = m$, and we get $D(\alpha') < D(\alpha)$. Now, since α is not wild, char k does not divide g.c.d. $(n, m) = m = \deg_Z \alpha'(Y)$. This shows that α' is not wild, so that α' and γ are tamely equivalent by induction hypothesis. Therefore α and γ are tamely equivalent, and (9.19.1) is proved.

(9.20) COROLLARY. If char $k = 0$ then any two k-algebra epimorphisms $k[X, Y] \to k[Z]$ are tamely equivalent.

Proof. Immediate from Theorem (9.19), sine if char $k = 0$ then there are no wild k-algebra epimorphisms. □

(9.21) COROLLARY. Let char $k = 0$. Let φ be an element of $k[X, Y]$ such that $k[X, Y]/(\varphi)$ is isomorphic (as a k-algebra) to $k[Z]$. Then there exists an element ψ of $k[X, Y]$ such that $k[\psi, \varphi] = k[X, Y]$.

Proof. Let $\alpha : k[X,Y] \to k[Z]$ be the k-algebra epimorphism defined by $\alpha = vu$, where $u : k[X,Y] \to k[X,Y]/(\varphi)$ is the natural surjection and $v : k[X,Y]/(\varphi) \to k[Z]$ is a k-algebra isomorphism. Then $\ker\alpha = (\varphi)$. Let $\beta : k[X,Y] \to k[z]$ be the k-algebra epimorphism defined by
76 $\beta(X) = Z, \beta(Y) = 0$. then $\ker\beta = (Y)$. By Corollary (9.20) there exists a k-algebra automorphism σ of $k[X,Y]$ such that $\beta = \alpha\sigma$. This gives $(\varphi) = \ker\alpha = \sigma(\ker\beta) = (\sigma(Y))$. Therefore $\sigma(Y) = \varnothing\varphi$. Let $\Psi = \sigma(X)$. Then $k[x,Y] = k[\sigma(X),\sigma(Y)] = k[\psi,\varnothing\varphi] = k[\psi,\varphi]$. $\square$

(9.22) LEMMA. *Let the assumptions be those of Corollary (9.21). Assume, moreover, that* $\deg_Y \varphi > 0$. *Then:*

(i) $\varnothing\varphi$ *is monic in* Y *for some* $\varnothing$.

(ii) $\varphi(X^{-1},Y)$ *is irreducible in* $\overline{k}((X))[Y]$, *where* $\overline{k}$ *is the algebraic closure of* k.

Proof. Let $\alpha : k[X,Y] \to k[Z]$ be the k-algebra epimorphism defined at the beginning of the proof of Corollary (9.21). Then $\ker\alpha = (\varphi)$. Since $\deg_Y \varphi > 0$, we have $X - a \not\equiv 0 \pmod{\varphi}$ for every $a \in k$. This shows that $\deg_Z \alpha(X) > 0$. Therefore α is a permissible embedding. Let $f = f(X,Y) \in \overline{k}((X))[Y]$ be the meromorphic equation of α. It follows from Remark (9.5) that $\ker\alpha = (f(X^{-1},Y))$. Therefore $f(X^{-1},Y) = \varnothing\varphi$ for some $\varnothing$, and the lemma is proved. $\square$

(9.23) COROLLARY. Let the assumptions be those of Corollary (9.21). Assume, moreover, that $\deg_Y \varphi > 0$. Then there exists an element ψ of $k[X,Y]$ such that $\deg_Y \psi < \deg_Y \varphi$ and $k[\psi,\varphi] = k[X,Y]$.

Proof. By Corollary (9.21) there exists $\psi \in k[X,Y]$ such that $k[\psi,\varphi] = k[X,Y]$. It is now enough to show that if $\deg_Y \psi \ge \deg_Y \varphi$ then there exists $\psi' \in k[X,Y]$ such that $\deg_Y \psi' < \deg_Y \psi$ and $k[\psi',\varphi] = k[X,Y]$. Let $n = \deg_Y \varphi$, $m = \deg_Y \psi$ and suppose $m \ge n$. In view of Lemma (9.22), replacing φ by $\varnothing\varphi$, we may assume that φ is monic in Y. Similarly, since

$$k[X,Y]/(\psi) = k[\psi,\varphi]/(\psi) \approx k[\varphi] \approx k[Z].$$

77 we may replace ψ by $\varnothing\psi$ and assume that ψ is monic in Y. Now,

$k[\psi,\varphi] = k[X,Y]$ implies that $k'[\psi,\varphi] = k'[Y]$, where $k' = k(X)$. Therefore if S, T are indeterminates then the k'−algebra homomorphism $\gamma : k'[S,T] \to k'[Y]$ defined by $\gamma(S) = \psi$, $\gamma(T) = \varphi$, is surjective. Therefore by Theorem (9.11) n divides m or m divides n. Since $m \geq n$, we get $m = pn$ for some positive integer p. Let $\psi' = \psi - \varphi^p$. Then $k[\psi',\varphi] = k[\psi,\varphi] = k[X,Y]$. Moreover, since both ψ and φ are monic in Y, we have $\deg_Y \psi' < m$. □

(9.24) THEOREM. *Let char $k = 0$. Let $\varphi = \varphi(X,Y)$ be an element of $k[X,Y]$ such that $n = \deg_Y \varphi > 0$, φ is monic in Y and $k[X,Y]/(\varphi)$ is isomorphic (as a k-algebra) to $k[Z]$. Let $f = f(X,Y) = \varphi(X^{-1},Y)$. Then f is irreducible in $\overline{k}((X))[Y]$. Let $h = h(f)$ and let $\psi = App_Y^d(\varphi)$, where $d = d_h(f)$. If $h \geq 2$ then $k[\psi,\varphi] = k[X,Y]$. (As usual, $\overline{k}$ denotes the algebraic closure of k.)*

Proof. Let $\alpha : k[X,Y] \to k[Z]$ be the k-algebra epimorphism defined by $\alpha = vu$, where $u : k[X,Y] \to k[x,Y]/(\varphi)$ is the natural surjection and $v : k[X,Y]/(\varphi) \to k[z]$ is a k-algebra isomorphism. Then $\ker\alpha = (\varphi)$ and, since $n > 0$, α is a permissible embedding. Since φ is monic in Y, it follows from Remark (9.5) that f is the meromorphic equation of α. We now use the notation of (9.3) and (9.7). Let $g = g(X,Y) = App_Y^d(f)$. Then by Proposition (4.7) $g(X,Y) = \psi(X^{-1},Y)$. Since $h \geq 2$, we have $\operatorname{ord}_t \psi(t^{-n},y) = \operatorname{ord}_t g(t^n,y) = r_h$ by Theorem (8.2). Since $\psi(t^{-n},y) = \theta(\psi(\overline{x},\overline{y}))$, it follows from (9.3.1) that $\deg_Z \psi(\overline{x},\overline{y}) = -r_h$. By Lemma (9.8) ($v$) we have $r_h = -1$. Therefore we have

$$\deg_Z \alpha(\psi) = 1. \tag{9.24.1}$$

□

Now, by Corollary (9.23) there exists an element ψ' of $k[X,Y]$ such that $\deg_Y \psi' < n$ and $k[\psi',\varphi] = k[X,Y]$. It follows that $k[Z] = k[\alpha(\psi')]$. **78**
Therefore we have

$$\deg_Z \alpha(\psi) = 1. \tag{9.24.2}$$

It follows from (9.24.1) and (9.24.2) that we have $\alpha(\psi') = a\alpha(\psi) + b$ for some a, $b \in k$, $a \neq 0$. This means that

$$\psi' = a\psi + b + \lambda\varphi$$

for some $\lambda \in k[X, Y]$. Since $\deg_Y \psi' < n$ and $\deg_Y \psi = n/d < n$, we get $\lambda = 0$ and $\psi' = a\psi + b$. This shows that $k[\psi', \varphi] = k[\psi, \varphi]$, and the theorem is proved.

With the notation and assumptions of Theorem (9.24) we have the following four corollaries:

(9.25) COROLLARY. If $h \geq 2$ then $r_n(-n, f) = -1$.

Proof. This was noted in the proof of the theorem above. □

(9.26) COROLLARY. $\deg_Y \varphi$ divides $\deg_X \varphi$ or $\deg_X \varphi$ divides $\deg_Y \varphi$.

Proof. Let $\alpha : k[X, Y] \to k[Z]$ be the permissible embedding defined in the proof of Theorem (9.24). Then, since $f(X, Y) = \varphi(X^{-1}, Y)$ is the meromorphic equation of α (Remark (9.5)), it follows from Lemma (9.4) that $\deg_Z \alpha(X) = \deg_Y \varphi = n$, Let $m = \deg_X \varphi$. If $m = 0$ then n divides m. If $m > 0$ then by the argument above, we get $\deg_Z \alpha(Y) = m$. Now, it follows from Theorem (9.11) that n divides m or m divides n. □

(9.27) COROLLARY. $d_2(f) = d_1(f)$ or $d_2(f) = -q_1(-n, f)$.

79 *Proof.* As seen in the proof of Corollary (9.26), we have $n = \deg_Z \alpha(X)$. Therefore $d_1(f) = \deg_Z \alpha(X)$. Moreover, by (9.7.2) we have $\deg_Z \alpha(Y) = -q_1(-n, f)$. Now, the corollary follows from Theorem (9.11). □

(9.28) COROLLARY. $k[X, Y]/(\psi)$ is isomorphic (as a k-algebra) to $k[Z]$.

Proof. This is clear, since $k[X, Y] = k[\psi, \varphi]$. □

(9.29) REMARK. The results proved in (9.21) - (9.28) above hold also for char $k > 0$ (and, infact, the same proof goes through), provided we make the assumption that $\deg_Y \varphi$ (or, by symmetry, $\deg_X \varphi$) is not divisible by char k.

10 Automorphism Theorem

As in §9, k ia an arbitrary field and X, Y, Z are indeterminates.

(10.1) Automorphism Theorem.

Every k-algebra automorphism of $k[X, Y]$ is tame.

(For the definition of a tame automorphism, see (9.14). In the proof below we deduce the Automorphism Theorem from the Epimorphism Theorem in case char $k = 0$. For a proof in the general case the reader is referred to [5].)

Proof of (10.1) in char $k = 0$. Let φ be a k-algebra automorphism of $k[X, Y]$. Let $\gamma : k[X, Y] \to k[Z]$ be the k-algebra epimorphism defined by $\gamma(X) = Z$, $\gamma(Y) = 0$, and let $\alpha = \gamma\varphi$. Then $\alpha : k[X, Y] \to k[Z]$ is also an epimorphism. Therefore by Corollary (9.20) there exists a tame k-algebra automorphism σ of $k[X, Y]$ such that $\alpha = \gamma\sigma$. Thus we get $\gamma\varphi = \gamma\sigma$. Put $\psi = \varphi\sigma^{-1}$. Then $\varphi = \psi\sigma$, and it is enough to prove that ψ is tame. Now, $\gamma\psi = \gamma$. Therefore $\psi(\ker\gamma) = \ker\gamma$. Now, $\ker\gamma = (Y)$. Therefore we have **80**

(10.1.1) $$\psi(Y) = aY$$

for some $a \in k$, $a \neq 0$. Now,

$$k[Y][X] = k[\psi(Y), \psi(X)] = k[aY, \psi(X)] = k[Y][\psi(X)].$$

Therefore there exist $P(Y) \in k[Y]$ and $b \in k$, $b \neq 0$, such that

(10.1.2) $$\psi(X) = bX + P(Y).$$

It is clear from (10.1.1) and (10.1.2) that ψ is tame.

(10.2) THEOREM. *Let* f, g *be elements of* $k[X, Y]$ *such that* $k[f, g] = k[X, Y]$. *Then* $\deg f$ *divides* $\deg g$ *or* $\deg g$ *divides* $\deg f$.

(Here deg denotes total degree with respect to X, Y. In the proof below we deduce Theorem (10.2) from the Epimorphism Theorem in case char $k = 0$. For a proof in the general case the reader is referred to [5].)

Proof of (10.2) in case char $k = 0$. Let $n = \deg f$, $m = \deg g$. Let f^+ be the homogeneous component of f of degree n, i.e., f^+ is a homogeneous polynomial in X, Y of degree n such that $f = f^+ + f'$ with $f' \in k[X,Y]$ and $\deg f' < n$. It is then clear that $\deg_Y f < n$ if an only if X divides f^+. Similarly, $\deg_Y g < m$ if and only if X divides g^+, where g^+ is the homogeneous component of g of degree m. Since $\left\{X + aY \middle| a \in k\right\}$ is an infinite set of mutually coprime elements of $k[X,Y]$, there exists $a \in k$, $a \neq 0$, such that $X' = X + aY$ divides neither f^+ nor g^+. Therefore, replacing X by X' we may assume that $n = \deg_Y f$, $m = \deg_Y g$. Let $k' = k(X)$ and let S, T be indeterminates. Let $\alpha : k'[S,T] \to k'[Y]$ be the k'-algebra
81 homomorphism defined by $\alpha(S) = f$, $\alpha(T) = g$. Then the assumption $k[f,g] = k[X,Y]$ implies that α is an epimorphism. Therefore it follows from Theorem (9.11) that n divides m or m divides n.

11 Affine Curves with One Place at Infinity

(11.1)

Throughout this section, by a *valuation* we shall mean a *real discrete* valuation with value group $\mathbb{Z}$. Thus if K is a field then a valuation v of K is a map $v : K \to \mathbb{Z} \cup \{\infty\}$ satisfying the following three conditions:

(i) $v(a) = \infty$ if an only if $a = 0$

(ii) $v|K^* : K^* \to \mathbb{Z}$ is a surjective homomorphism of groups, where K^* is the group of units of K

(iii) $v(a + b) \geq \min(v(a), v(b))$ for all $a, b \in K$.

We denote by R_v the ring of v and by m_v the maximal ideal of R_v. Recall that $R_v = \{a \in k | v(a) \geq 0\}$ and $m_v = \{a \in K | v(a) > 0\}$. The ring R_v is a discrete valuation ring with quotient field K. If k is a subfield of K such that $v(a) = 0$ for every non-zero element a of k then we say, as usual, that v is a valuation of K/k. Note that in this case the residue field R_v/m_v of v is an overfield of k. We say v is *residually rational over* k if $k = R_v/m_v$. Let L/K be a field extension. Let v be a valuation of K and let w be a valuation of L. We say *w extends (or lies over) v* if $R_w \cap K = R_v$.

(11.2) DEFINITION. Let k be a field and let A be a k-algebra. We say A is an *affine curve* over k (more precisely, the *coordinate ring* of an *integral affine curve* over k) if the following three conditions are satisfied:

(i) A is finitely generated as a k-algebra.

(ii) A is an integral domain.

(iii) A has Krull dimension one, i.e. if K is the quotient field of A then **82**
tr. $\deg_k K = 1$

(11.3) DEFINITION. Let A be an affine curve over k. We say A is a *plane* affine curve (resp. *the affine line*) if A is generated as a k-algebra by two elements (resp. one element). Note that the affine line is the polynomial ring in one variable over k.

(11.4) DEFINITION. Let A be an affine curve over k. We say A has only *one place at infinity* if the following two conditions are satisfied:

(i) There exists exactly one valuation v of K/k, where K is the quotient field of A, such that $A \not\subset R_v$.

(ii) The unique valuation v of condition (i) is residually rational over k.

We call v the *place* (or *valuation) of A* at infinity.

(11.5) EXAMPLE. An affine polynomial curve over k (Definition (9.1)) has only one place at infinity. For, if A is such a curve then $A \subset k[Z]$ and the quotient field of A is $k(Z)$, where Z is an indeterminate. If v is the Z^{-1}-adic valuation of $k(Z)/k$ then it is clear that v is residually rational over k and is the unique place of A at infinity. In particular, the affine line has only one place at infinity.

(11.6) LEMMA. *Let v be a valuation of K/k. Let x be a non-zero element of K. If x is algebraic over k then $v(x) = 0$.*

Proof. Suppose $v(x) \neq 0$. Since x is algebraic over k if an only if x^{-1} is algebraic over k, we may assume that $v(x) > 0$. If $x$is algebraic over k then, since $x \neq 0$, there exist $n \geq 1$ and $a_i \in k$, $0 \leq i \leq n-1$, such that $a_0 \neq 0$ and

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x = a_0.$$

83 Since $v(x) > 0$, we have $v(x^n + a_{n-1} + \ldots + a_1 x) > 0$. But $v(a_0) = 0$. This contradiction proves that $v(x) = 0$. $\square$

(11.7) LEMMA. *Let v be a valuation of K/k such that v is residually rational over k. Then k is algebraically closed in K.*

Proof. Let $x \in K$ be algebraic over k. We want to show that $x \in k$. We may assume that $x \neq 0$. Then x^{-1} is also algebraic over k. Since $x \in R_v$ or $x^{-1} \in R_v$, we may assume, without loss of generality, that $x \in R_v$. Then since v is residually rational over k, there exists $a \in k$ such that $v(x - a) > 0$. Now, since $x - a$ is algebraic over k, it follows from Lemma (11.6) that $x - a = 0$, which shows that $x \in k$. $\square$

(11.8) LEMMA. *Let A be an affine curve over k with only one place v at infinity. Let K be the quotient field of A. Let $x \in A$, $x \notin k$. Then:*

(i) *x is transcendental over k and v is the unique valuation of K/k extending the x^{-1}-adic valuation of $k(x)/k$.*

(ii) $v(x) = -[K : k(x)]$. *In particular,* $v(x) < 0$.

(iii) *A is integral over $k[x]$.*

Proof.

(i) Since v is residually rational over k and since $x \notin k$, x is transcendental over k by Lemma (11.7). Let v'be any valuation of K/k extending the x^{-1}-adic valuation of $k(x)/k$. Then x^{-1} is a non-unit in the ring $R_{v'}$ of v'. This means that $x \notin R_{v'}$. Therefore $A \not\subset R_{v'}$, and the hypothesis on A implies that $v = v'$.

84 (ii) Since v is the only valuation of K/k extending the x^{-1}-adic valuation of $k(x)/k$ and since the residue field of v is k, $[K : k(x)]$ equals the ramification index of v over the x^{-1}-adic valuation of $k(x)/k$, i.e., $[K : k(x)] = v(x^{-1}) = -v(x)$.

(iii) Let $y \in A$. To show that y is integral over $k[x]$, it is enough to show that y is integral over each valuation ring of $k(x)/k$ containing $k[x]$. Let then R_w be such a valuation ring with valuation w, and let $w_1, \ldots, w_r$ be all the extensions of w to K. Then, if $\overline{R}_w$ is the integral closure of R_w in K, we have $\overline{R}_w = \bigcap_{i=1}^{r} R_{w_i}$. Therefore it is enough to prove that $y \in R_{w_i}$ for every i, $1 \leq i \leq r$. Since $A \subset R_{v'}$ for every valuation v' of K/k other than v, we have only to show that $w_i \neq v$ for every i, $1 \leq i \leq r$. But this is clear, since $x \in R_{w_i}$ for every i, $1 \leq i \leq r$, and $x \notin R_v$ by (ii).

□

(11.9) COROLLARY. Let A be an affine curve over k with only one place v at infinity. Then $v(A - \{0\}) = \left\{ v(a) \middle| a \in A, a \neq 0 \right\}$ is a subsemigroup of the semigroup of non-positive integers. Moreover, the only units of A are the non-zero elements of k.

Proof. The first assertion is immediate from Lemma (11.8) (ii). To prove the second assertion, let $x \notin k$. Then x is transcendental over k, hence a non unit in $k[x]$. Since A is integral over $k[x]$, x is a non-unit in A. □

(11.10) REMARK. In view of Corollary (11.9), we may omit explicit mention of k in Definition (11.4). That is, we may say A to have only one place at infinity if *there exists* a subfield k of A such that A is an affine curve over k with only one place at infinity in the sense of Definition (11.4). The subfield k is then uniquely determined by A. viz, it is the set of all units of A together with zero. We call k the *ground field* of A. 85

(11.11) DEFINITION. Let R be a ring and let $R[Y]$ be the polynomial ring in one variable Y over R. An element f of $R[Y]$ is said to be *almost* *monic* in Y if $f \neq 0$ and the leading coefficient of f is a unit in R, i.e. $f \neq 0$ and there exists a unit a in R such that $\deg(f - aY^n) < n$, where $n = \deg_Y f$.

(11.12) PROPOSITION. Let k' be a field and let k be its algebraic closure. Let $\varphi = \varphi(X, Y)$ be an element of $k'[X, Y] \subset k((X^{-1}))[Y]$ such that $\deg_Y \varphi > 0$. Let $A = k'[X, Y]/(\varphi)$, where $(\varphi) = \varphi k'[x, Y]$. Assume that A is an affine curve over k' with only one place v at infinity. Then:

(i) φ is almost monic in Y.

(ii) $\deg_Y \varphi = -v(X + (\varphi))$.

(iii) φ is irreducible in $k((X^{-1}))[Y]$.

Proof. Let $x = X + (\varphi)$. Since $\deg_Y \varphi > 0$, we have $x \notin k'$. Therefore by Lemma (11.8) x is transcendental over k' and A is integral over $k'[x]$. In particular $y = Y + (\varphi)$ is integral over $k'[x]$, and (i) is proved. Now, if K is the quotient field of A then we have $\deg_Y \varphi = [K : k'(x)]$. By Lemma (11.8) we have $[K : k'(x)] = -v(x)$. This proves (ii). In order to prove (iii), we may, in view of (i), replace φ by $a\varphi$ for a suitable non-zero element a of k' to assume that φ is monic in Y. Then $\varphi(x, Y) \in k'[x][Y]$ is the minimal monic polynomial of y over $k'(x)$. Let L be an overfield of $k((x^{-1}))$ such that we have a $k'(x)$-monomorphism $u : K \to L$ and L is generated over $k((x^{-1}))$ by $u(y)$. (Here we regard $k((x^{-1}))$ as an overfield of $k'(x)$ via the natural inclusions $k' \hookrightarrow k(x) \hookrightarrow k((x^{-1}))$.) Let $\psi(x, Y) \in k((x^{-1}))[Y]$ be the minimal monic polynomial of $u(y)$ over $k((x^{-1}))$. In
86 order to prove (iii), it is enough to show that $\varphi(x, Y) = \psi(x, Y)$. Now, since $\varphi(x, u(y)) = u(\varphi(x, y)) = 0$, $\psi(x, Y)$ divides $\varphi(x, Y)$ in $k((x^{-1}))[Y]$. Therefore it is now enough to show that $\deg_Y \varphi(x, Y) \leq \deg_Y \psi(x, Y)$. Let $n = \deg_Y \varphi(x, Y)$, $m = \deg_Y \psi(x, Y)$. Then $n = v(x^{-1})$ by (ii), and $m = [L : k((x^{-1}))]$. Let w be a valuation of L extending the x^{-1}-adic valuation of $k((x^{-1}))/k$. We claim that there exists a (unique) valuation v' of K such that w is an extension of v'. For, let $w' : K \to \mathbb{Z} \cup \{\infty\}$ denote the restriction of w to K. Then, writing K^* for the group of units of K, $w'(K^*)$ is a subgroup of $\mathbb{Z}$. Since $w(x^{-1}) > 0$ and $x^{-1} \in K$, we have $w'(K^*) \neq 0$. If r is the positive generator of $w'(K^*)$, we put $v' = r^{-1}w'$.

Then $v' : K \to \mathbb{Z} \cup \{\infty\}$ is surjective and our claim is proved. Now, since $v'(x^{-1}) > 0$, v' is an extension of the x^{-1}-adic valuation of $k'(x)/k'$. Therefore $v' = v$ by Lemma (11.8). Now, we get $n = v(x^{-1}) = v'(x^{-1}) = r^{-1}w(x^{-1}) \leq w(x^{-1}) \leq [L : k((x^{-1}))] = m$, and (iii) is proved.

This completes the proof of the proposition. □

(11.13) NOTATION. Let k be an algebraically closed field and let $\varphi = \varphi(X,Y)$ be an element of $k[X,Y]$ such that φ is monic in Y and char k does not divide $\deg_Y \varphi$. In particular, this means that $\deg_Y \varphi > 0$. Let $n = \deg_Y \varphi$. Assume that φ is irreducible in $k((X^{-1}))[Y]$. Put $f = f(X,Y) = \varphi(X^{-1}, Y)$. Then f is a irreducible element of $k((X))[Y]$ and f is monic in Y with $\deg_Y f = n$. Therefore by Newton's Theorem (5.14) there exists $y(t) \in k((t))$ such that $f(t^n, y(t)) = 0$. Let k' be a subfield of k such that $\varphi \in k'[x, Y]$. Let $R = k'[X^{-1}]$. Then $f \in R[Y]$. Let $\overline{R[Y]} = R[Y]/fR[Y]$ and let $A = k'[X,Y]/\varphi k'[X,Y]$. It is then clear that the k'-algebra isomorphism $\theta' : k'[X,Y] \to R[Y]$ defined by $\theta'(X) = X^{-1}$, $\theta'(Y) = Y$, induces a k'-algebra isomorphism $\overline{\theta'} : A \to \overline{R[Y]}$. Recall also that if $k'[t^{-n}, y(t)]$ denotes the k'-subalgebra of $k((t))$ generated by 87
t^{-n} and $y(t)$ then by Lemma (8.4) there exists k'-algebra isomorphism $\overline{u} : \overline{R[Y]} \to k'[t^{-n}, y(t)]$ given by $\overline{u}(\overline{F(X,Y)}) = F(t^n, y(t))$, where $\overline{F(X,Y)}$ denotes the image of an element $F(X,Y)$ of $R[Y]$ under the canonical homomorphism $R[Y] \to \overline{R[Y]}$. Putting $\theta = \overline{u}\overline{\theta'}$, we get a k'-algebra isomorphism

$$\theta : A = k'[X,Y]/\varphi k'[X,Y] \to k'[t^{-n}, y(t)]$$

given by $\theta(F(x,Y)) = F(t^{-n}, y(t))$ for $F(X,Y) \in k'[X,Y]$, where x (resp. y) is the canonical image of X (resp. Y) in A. In the sequel we shall

(11.13.1) Identify A with $k'[t^{-n}, y(t)]$ via θ.

Note that under this identification we have $x = t^{-n}$ and $y = y(t)$. Let $K = k'(t^n, y(t))$ be the quotient field of A. Since K is a subfield of $k((t))$, we have a map

$$\mathrm{ord}_t : K \to \mathbb{Z} \cup \{\infty\}.$$

Let $h = h(f)$ and let $r_i = r_i(-n, f)$, $d_{i+1} = d_{i+1}(f)$ for $0 \leq i \leq h+1$. Let $\Gamma_R(f)$ be the value semigroup of f with respect to R. Recall that

$$\Gamma_r(f) = \left\{ \operatorname{ord}_t F(t^n, y(t)) \middle| F(X, Y) \in R[Y], F(t^n, y(t)) \neq 0 \right\}.$$

(11.14) LEMMA. *With the notation of (11.13), we have:*

(i) $\operatorname{ord}_t(A - \{0\}) = \Gamma_R(f)$.

(ii) ord_t *is a valuation of* K/k'.

(iii) *A is an affine curve over k' with only one place ord_t at infinity.*

(iv) $\operatorname{ord}_t(A - \{0\})$ *is strictly generated by* $r = (r_0, \ldots, r_h)$

88 (v) $r_0 < 0$, $r_1 = \infty$ *or* $r_1 \leq 0$, *and* $r_i < 0$ *for* $2 \leq i \leq h$.

Proof.

(i) In view of the identification of A with $k'[t^{-n}, y(t)]$ via θ, we have

$$\begin{aligned}
\Gamma_R(f) &= \left\{ \operatorname{ord}_t F(t^n, y(t)) \middle| F(X, Y) \in R[Y], F(t^n, y(t)) \neq 0 \right\} \\
&= \left\{ \operatorname{ord}_t F(t^n, y(t)) \middle| F(X, Y) \in k'[X, Y], F(t^{-n}, y(t)) \neq 0 \right\} \\
&= \left\{ \operatorname{ord}_t F(x, y) \middle| F(X, Y) \in k'[X, Y], F(x, y) \neq 0 \right\} \\
&= \left\{ \operatorname{ord}_t a \middle| a \in A, \neq 0 \right\} \\
&= \operatorname{ord}_t(A - \{0\}).
\end{aligned}$$

(ii) We have only to show that $\operatorname{ord}_t : K \to \mathbb{Z} \cup \{\infty\}$ is surjective or, equivalently, that $\operatorname{ord}_t(K^*) = \mathbb{Z}$, where $K^* = K - \{0\}$. Now $\operatorname{ord}_t(K^*)$ is clearly the subgroup of $\mathbb{Z}$ generated by the semigroup $\operatorname{ord}_t(A - \{0\})$, hence by $\Gamma_R(f)$ in view of (i). Since $X^{-1} \in R$, the assertion now follows from Corollary (8.8).

(iii) Since φ is monic in Y, A is integral over $k'[x]$. We have $\operatorname{ord}_t(x) = \operatorname{ord}_t(t^{-n}) = -n$. Therefore

$$\operatorname{ord}_t(x^{-1}) = n = \deg_Y \varphi = [K : k'(x)].$$

This shows that ord_t is the only valuation of K/k' extending the x^{-1}-adic valuation of $k'(x)/k$ and that ord_t is residually rational over k'. Now, let w be any valuation of K/k' such that $A \not\subset R_w$. Then, since A is integral over $k'[x]$, we have $k'[x] \not\subset R_w$. This means that $w(x) < 0$, so that $w(x^{-1}) > 0$. Therefore w extends the x^{-1}-adic valuation of $k'(x)$, and we get $w = \mathrm{ord}_t$.

(iv) This is immediate from Theorem (8.7) (iii$'$), since we have $\mathrm{ord}_t(A - \{0\}) = \Gamma_R(f)$ by (i) and we are in the pure meromorphic case.

(v) We have $r_0 = -n < 0$. Next, $r_1 = \mathrm{ord}_t(y)$. If $y \in k'$ then $\mathrm{ord}_t(y) = 0$ **89**
or ∞. If $y \notin k'$ then, since $y \in A$, we get $\mathrm{ord}_t(y) < 0$ by (iii) and lemma (11.8) (ii). Now, let $g_i = g_i(X, Y) = App_Y^{d_i}(f)$, $2 \le i \le h$. Then $g_i \in k'[X^{-1}][Y]$ for every i by Theorem (8.3)(i). Put $\psi_i = \psi_i(X, Y) = g_i(X^{-1}, Y)$, $2 \le i \le h$. Then $\psi_i \in k'[X, Y]$ for every i. Now, for $2 \le i \le h$, we have

$$\begin{aligned} r_i &= \mathrm{ord}_t\, g_i(t^n, y(t)) && \text{(by Theorem (8.2))} \\ &= \mathrm{ord}_t\, \psi_i(t^{-n}, y(t)) \\ &= \mathrm{ord}_t\, \psi_i(x, y) && \text{(by (11.13.1)).} \end{aligned}$$

Therefore by (iii) and Lemma (11.8) (ii) it is enough to prove that $\psi_i(x, y) \notin k'$ for every i, $2 \le i \le h$. Now, we have $\deg_Y \psi_i = n/d_i$. This shows that $1 \le \deg_Y \psi_i < n = \deg_Y \varphi$ for every i, $2 \le i \le h$. Therefore, for every $a \in k'$, φ does not divide $\psi_i - a$ in $k'[X, Y]$. This means that $\psi_i(x, y) \notin k'$. □

(11.15) THEOREM. *Let k be an algebraically closed field and let φ be an element of $k[X, Y]$ such that $\deg_Y \varphi > 0$. Consider the following four conditions.*

(i) *For every subfield k' of k such that $\varphi \in k'[X, Y]$, $k'[X, Y]/\varphi k'[X, Y]$ is an affine curve k' with only one place at infinity.*

(ii) *$k[X, Y]/\varphi k[x, Y]$ is an affine curve over k with only one place at infinity.*

(iii) *There exists a subfield k' of k such that $\varphi \in k'[X, Y]$ and $k'[X, Y]/\varphi k'[X, Y]$ is an affine curve over k' with only one place at infinity.*

(iv) *φ is almost monic in Y and φ is irreducible in $k((X^{-1}))[Y]$.*

We have (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). Moreover, if char k does not divide $\deg_Y \varphi$ *then (iv) $\Rightarrow$ (i).*

Proof. (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). Trivial.

(iii) $\Rightarrow$ (iv). Immediate from Proposition (11.12).

90 (iv) $\rightarrow$ (i). Assume that char k does not divide $\deg_Y \varphi$. Let k' be a subfield of k such that $\varphi \in k'[X, Y]$. Then, replacing φ by $a\ \varphi$ for a suitable $a \in k'$, we may assume that φ is monic in Y. Now, (i) follows from Lemma (11.14) (iii). $\square$

(11.16) COROLLARY. Let k' be a field and let k be its algebraic closure. Let $\varphi = \varphi(X, Y)$ be a non-zero element of $k'[X, Y]$ such that char k does not divide $\deg_Y \varphi$ and $k'[X, Y]/\varphi k'[X, Y]$ is an affine curve over k' with only one place at infinity. Then for every $\lambda \in k$, $k'(\lambda)[X, Y]$ is an affine curve over $k'(\lambda)$ with only one place at infinity

Proof. Since char k does not divide $\deg_Y \varphi$, we have $\deg_Y \varphi > 0$. Therefore by Theorem (11.15) φ is almost monic in Y, i.e. there exists $a \in k'$, $a \neq 0$, such that $a\ \varphi$ is monic in Y. Since $k = \{a\lambda \big| \lambda \in k\}$, we may replace φ by $a\ \varphi$ and assume that φ is monic in Y. By Theorem (11.15) φ is irreducible in $k((X^{-1}))[Y]$. Since $\deg_Y(\varphi + \lambda) = \deg_Y \varphi$ is not divisible by char k for every $\lambda \in k$, it is enough, by Theorem (11.15), to prove that $\varphi + \lambda$ is irreducible in $k((x^{-1}))[Y]$ for every $\lambda \in k$. Let $n = \deg_Y \varphi$. Put $f = f(X, Y) = \varphi(X^{-1}, Y)$. Then f is an irreducible element of $k((X))[Y]$ and f is monic in Y with $\deg_Y f = n$. Clearly, it is enough to prove that $f + \lambda$ is irreducible in $k((X))[Y]$ for every $\lambda \in k$. By Newton's Theorem (5.14) there exists an element $y(t)$ of $k((t))$ such that $f(t^n, y(t)) = 0$. Let $h = h(f)$, $s_h = s_h(-n, f)$ and $r_i = r_i(-n, f)$ for $0 \leq i \leq h$. Then by Lemma (11.14)(v) we have $r_h \leq 0$. First, suppose that $r_h = 0$. Then by Lemma (11.14)(v) we have $h = 1$. Therefore we get $1 = d_{h+1}(f) = d_2(f) = \text{g.c.d. } (r_0, r_1) = \text{g.c.d. } (-n, 0) = n$. Thus in this case we have $\deg_Y(f + \lambda) = 1$, which clearly implies that

$f + \lambda$ is irreducible in $k((X))[Y]$. Now, suppose that $r_h < 0$. Then $s_h < 0$. Let $f_\lambda = f + \lambda$. Then $f_\lambda(t^n, y(t)) = \lambda \in k$. Therefore $\mathrm{ord}_t\, f_\lambda(t^n, y(t)) \geq 0 > s_h$. Now, it follows from the Irreducibility Cri- **91**
terion (Theorem (12.4)) proved in the next section that f_λ is irreducible in $k((X))[Y]$. $\square$

(11.17) REMARK. Let us justify the use of a result from § 12 in proving Corollary (11.16) by declaring that the result of Corollary (11.16) will not be used anywhere in the sequel.

(11.18) QUESTION. Is Corollary (11.16) true without the assumption that char k does not divide $\deg_Y \varphi$?

(11.19) PROPOSITION. Let k be a field and let n be a positive integer such that char k does not divide n. Let

$$\varphi = \varphi(X, Y) = a_0(X)Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X)$$

with $a_i(X) \in k[X]$ for $0 \leq i \leq n$, $a_0(X) \neq 0$. Let $m = \deg_X \varphi$. Assume that $k[X, Y]/\varphi k[X, Y]$ is an affine curve over k with only one place at infinity. Then $a_0(X) \in k$ and we have $n \deg_X a_i(X) \leq im$ for every $i, 0 \leq i \leq n$. Moreover, if $m \geq 1$ then we have $\deg_X a_n(X) = m$ and $n \deg_X a_i(X) \leq i \deg_X a_n(X)$ for every i, $0 \leq i \leq n$.

Proof. By Proposition (11.12) φ is almost monic in Y. This means that $a_0(X) \in k$. Therefore, replacing φ by $a_0(X)^{-1}\varphi$, we may assume that $a_0(X) = 1$. Now, if $m = 0$ then the assertion is clear. Assume therefore that $m \geq 1$. Then by Proposition (11.12) φ is almost monic in X. This shows that $\deg_X a_n(X) = m$. $\square$

Now, by Proposition (11.12) φ is irreducible in $\overline{k}((X^{-1}))[Y]$, where $\overline{k}$ is the algebraic closure of k. Therefore by Newton's Theorem (5.14) there exists $y(t) \in \overline{k}((t))$ such that

$$\varphi(t^{-n}, Y) = \prod_{w \in \mu_n(\overline{k})} (Y - y(wt)).$$

Let $q = \mathrm{ord}_t\, y(wt)$ for all $w \in \mu_n(\overline{k})$. Then, since $a_i(t^{-n})$ equals $(-1)^i$ **92** times the i^{th} elementary symmetric function of $\{y(wt) | w \in \mu_n(\overline{k})\}$, we have $\operatorname{ord}_t a_i(t^{-n}) \geq iq$ for $1 \leq i \leq n$. Moreover, since

$$a_n(t^{-n}) = (-1)^n \prod_{w \in \mu_n(\overline{k})} y(wt),$$

we have $\operatorname{ord}_t a_n(t^{-n}) = nq$, which gives $q = \operatorname{ord}_X a_n(X^{-1}) = -\deg_X a_n(X)$. Therefore for every i, $1 \leq i \leq n$, we get

$$\begin{aligned} n \deg_X a_i(X) &= -n \operatorname{ord}_X a_i(X^{-1}) \\ &= -\operatorname{ord}_t a_i(t^{-n}) \\ &\leq -iq \\ &= i \deg_X a_n(X) \\ &= im. \end{aligned}$$

(11.20) COROLLARY. Let k be a field of characteristic zero and let f, g be elements of $k[X, Y]$ such that $k[f, g] = k[X, Y]$. Let $m = \deg_X f$, $n = \deg_Y f$ and let

$$f = a_0(X)Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X)$$

with $a_i(x) \in k[X]$ for $0 \leq i \leq n$. Then we have $n \deg_X a_i(X) \leq im$ for $0 \leq i \leq n$. Moreover, if $m \geq 1$ (resp. $n \geq 1$) then f is almost monic in X (resp. Y).

Proof. The inequality $n \deg_X a_i(X) \leq im$ is obvious for $n = 0$. We may therefore assume that $n > 0$. Then, since $k[X, Y]/fk[X, Y]$ is isomorphic
93 to $k[g]$, which is an affine curve over k with only one place at infinity (Example (11.5)), the corollary follows from Propositions (11.19) and (11.12). $\square$

(11.21) DEFINITION. Let k be a field and let f be a non-zero element of $k[X, Y]$. Write $f = \sum a_{ij} X^i Y^j$ with $a_{ij} \in k$. The *degree form* of f, denoted f^+, is defined by

$$f^+ = \sum_{i+j=n} a_{ij} X^i Y^j$$

where $n = \deg f$. (Note that $\deg f$ and f^+ depend only on the k-vector subspace $kX \oplus kY$ of $k[X, Y]$ and do not depend upon a k-basis X, Y of $kX \oplus kY$.)

(11.22) DEFINITION. Let $f \in k[X, Y]$, $f \notin k$. We say f has *only one point at infinity* if f^+ is a power of a linear polynomial in $\overline{k}[x, Y]$, where $\overline{k}$ is the algebraic closure of k. (Note that this definition depends only on the k-vector subspace $kX \oplus kY$ of $k[X, Y]$ and is independent of the choice of a k-basis X, Y of $kX \oplus kY$.)

(11.23) PROPOSITION. Let k be a field of characteristic zero and let f be an element of $k[X, Y]$ such that $k[X, Y]/fk[X, Y]$ is an affine curve over k with only one place at infinity. Then f has only one point at infinity.

Proof. We may assume that k is algebraically closed. For, by interchanging X and Y, if necessary, we may assume that $\deg_Y f > 0$ and then apply Theorem (11.15).

Now, suppose f^+ is not a power of a linear polynomial in $k[X, Y]$. Then, replacing X, Y by a suitable k-basis of $kX \oplus kY$, we may assume that f^+ is of the form

$$f^+ = X^r \prod_{i=1}^{q} (X + a_i Y),$$

where r, q are positive integers and $a_i \in k$, $a_i \neq 0$, for $1 \leq i \leq q$. Let **94**
$m = \deg_X f$ and $n = \deg_Y f$. Then $m = r + q$ and $m > n \geq q \geq 1$. By Proposition (11.12) f is almost monic in Y. Therefore $n > q$ and we can write f in the form

$$f = f_1 + f_2 + f_3,$$

where $f_1 = f^+$, $f_2 = bY^n$ for some $b \in k$, $b \neq 0$, and

$$f_3 = \sum_{\substack{i+j<m \\ j<n}} c_{ij} X^i Y^i$$

with $c_{ij} \in k$. Let $A = k[X, Y]/fk[X, Y]$ and let v be the valuation of A at infinity. Let $\overline{F}$ denote the image of an element F of $k[X, Y]$ under the canonical map $k[X,Y] \to A$. Then by Proposition (11.12) we have $v(\overline{X}) = -n$, $v(\overline{Y}) = -m$. Since $-m < -n$, we have $v(\overline{X} + a_i\overline{Y}) = -m$ for every i, $1 \leq i \leq q$, and we get

$$v(\overline{f}_1) = -rn - qm < -rn - qn = -mn.$$

Therefore, since $v(\overline{f}_2) = -mn$, we get

$$v(\overline{f}_1 + \overline{f}_2) < -mn. \tag{11.23.1}$$

Now, let $(i, j) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ be such that $c_{ij} \neq 0$. Then by Proposition (11.19) we have $ni \leq (n - j)m$. This gives $-in - jm \geq -mn$. Therefore we get

$$v(\overline{f}_3) \geq \inf\{-in - jm | c_{ij} \neq 0\} \geq -mn. \tag{11.23.2}$$

Since $\overline{f}_1 + \overline{f}_2 = -\overline{f}_3$, (11.23.1) and (11.23.2) together give a contradiction. □

95 **(11.24) COROLLARY.** Let k be a field of characteristic zero and let f, g be elements of $k[X,Y]$ such that $k[f,g] = k[X,Y]$. Then f has only one point at infinity.

Proof. Since $k[X,Y]/fk[X,Y] \approx k[g]$ is an affine curve over k with only one place at infinity, the corollary follows from Proposition (11.23). □

(11.25) REMARK. Proposition (11.23) and Corollary (11.24) are, in fact, true even without the assumption that char $k = 0$.

(11.26) REMARK. Let us call $k[X,Y]$ the *affine plane* over k. Let A be an affine curve over k. By an *embedding* α of A in the affine plane we mean a k-algebra epimorphism (i.e. surjective homomorphism) $\alpha : k[X,Y] \to A$. (See Definition (9.1).) We say two such embedding α, *beta* are *equivalent* if there exists a k-algebra automorphism σ of $k[X,Y]$ such that $\alpha = \beta\sigma$. With this terminology, the Epimorphism Theorem (9.19) says that if char $k = 0$ (or, more generally, if we restrict our attention to non-wild embeddings) then all embeddings of the *affine line* in the affine plane are equivalent to each other. This statement is not true for more general affine curves. However, if A is an affine curve with only one place at infinity then to each embedding of A in the affine plane we can associate certain characteristic sequences and, using the Fundamental Theorem of § 8, we can classify the equivalence classes of the embeddings in terms of these characteristic sequences. It can be deduced from this classification that if char $k = 0$ (or, more generally, if we restrict our attention to certain "non-wild" embeddings) then the number of these equivalence classes is finite. For precise statements and proofs of these assertions, the reader is referred to [3]. However, in Theorems (11.26.1) and (11.26.2) below we state (without proof) a simplified version of these results. **96**

Suppose char $k = 0$ and A is an affine curve k with only one place v at infinity. Let α be an embedding of A (in the affine plane) such that $\alpha(X) \notin k$. Let $x = \alpha(X)$, $y = \alpha(Y)$. Then by Lemma (11.8) x is transcendental over k and $A = k[x, y]$ is integral over $k[x]$. Therefore the minimal monic in polynomial $\varphi(x, Y) \in k(x)[Y]$ of y over $k(x)$ belongs to $k[x, Y]$. Let $\varphi = \varphi(X, Y)$. Then φ is monic in Y and $\deg_Y \varphi = n$, where $n = -v(x)$ (Lemma (11.8)). Moreover, it is clear that $\ker \alpha = \varphi k[X, Y]$. Therefore it follows from Proposition (11.12) that φ is irreducible in $\overline{k}((X^{-1}))[Y]$, where $\overline{k}$ is the algebraic closure of k. Let $f = \varphi(X^{-1}, Y)$. Put $h(\alpha) = h(f)$, $d_2(\alpha) = d_2(f)$, $q_i(-n, f)$ for $0 \leq i \leq h(\alpha) + 1$, and $q(\alpha) = (q_0(\alpha), q_1(\alpha), \ldots, q_{h+1}(\alpha)) = q(-n, f)$, where $h = h(\alpha)$.

For an embedding α of A we define its *transpose* α^t to be the embedding of A given by $\alpha^t(X) = \alpha(Y)$, $\alpha^t(Y) = \alpha(X)$. Note that α and α^t are equivalent embeddings. If $\alpha(X) \in k$ then $\alpha^t(X) \notin k$, and in this case we define: $h(\alpha) = h(\alpha^t)$, $d_2(\alpha) = d_2(\alpha^t)$, $q_0(\alpha) = q_1(\alpha^t)$, $q_1(\alpha) = q_0(\alpha^t)$, $q_i(\alpha) = q_i(\alpha^t)$ for $2 \leq i \leq h + 1$ and

$$q(\alpha^t) = (q_0(\alpha^t), q_1(\alpha^t), \ldots, q_{h+1}(\alpha^t)),$$

where $h = h(\alpha^t)$.

Let α be an embedding of A. Then $v(\alpha(X)) = q_0(\alpha(Y)) = q_1(\alpha)$. We call the pair $(-v(\alpha(X)), -v(\alpha(Y)))$ the *bidegree* of α and denote it by bideg (α). Let bideg $(\alpha) = (m, n)$. We say α is *principal* if $m \neq -\infty$, $n \neq -\infty$ and m divides n or n divides m. Otherwise, we say α is *non-principal*. Note that $d_2(\alpha) =$ g.c.d. (m, n). We now state

(11.26.1) THEOREM

97 Let k be a field of characteristic zero and let A be an affine curve over k with only one place at infinity. Then any embedding of A (in the affine plane) is equivalent to a non-principal embedding. If α, β are non-principal embeddings of A then the following four conditions are equivalent:

(1) α and β are equivalent.

(2) $q(\alpha) = q(\beta)$ or $q(\alpha) = q(\beta^t)$.

(3) bideg (α) = bideg (β) or bideg (α) = bideg (β^t).

(4) $d_2(\alpha) = d_2(\beta)$.

(11.26.2) THEOREM

Let A be as in Theorem (11.26.1). Then the number of equivalence classes of embeddings of A in the affine plane is finite.

Chapter 5

Irreducibility, Newton's Polygon

12 Irreducibility Criterion

(12.1)

Let k be an algebraically closed field. Let $f = f(X, Y)$ be an irreducible 98
element of $k((X))[Y]$ such that f is monic in Y and char k does not divide $\deg_Y f$. Let $n = \deg_Y f$. By Newton's Theorem (5.14) there exists an element $y(t)$ of $k((t))$ such that

$$f(t^n, Y) = \prod_{w \in \mu_n} (Y - y(wt)).$$

where $\mu_n = \mu_n(k)$. Let ν be an integer such that $|\nu| = n$. Let $h = h(f)$ and let $m_i = m_i(\nu, f)$, $q_i = q_i(\nu, f)$, $s_i = s_i(\nu, f)$, $r_i = r_i(\nu, f)$, $d_{i+1} = d_{i+1}(f)$ for $0 \le i \le h + 1$.

(12.2)

Let L be an overfield of $k((t))$ and let v be a valuation of L extending the valuation ord_t of $k((t))/k$. (As in § 11, by a valuation we mean a real discrete valuation with value group $\mathbb{Z}$, as defined in (11.1).) Let $e = v(t)$. Then we have $v(a) = e \operatorname{ord}_t a$ for every $a \in k((t))$.

With the notation of (12.1) and (12.2), we have

(12.3) LEMMA. *Let z be an element of L such that $v(z - y(wt)) \leq em_h$ for every $w \in \mu_n$. Then $v(f(t^n, z)) \leq es_h$.*

Proof. Let $m = \sup\left\{v(z - y(wt))\middle| w \in \mu_n\right\}$. Then $m \leq em_h$. We may assume, without loss of generality, that $m = v(z-y(t))$. Then $v(z-y(t)) \geq v(z - y(wt))$ for every $w \in \mu_n$. Therefore, since

$$y(t) - y(wt) = (y(t) - z) + (z - y(wt)),$$

99 we get

$$v(y(t) - y(wt)) \geq v(z - y(wt)) \tag{12.3.1}$$

for every $w \in \mu_n$. Now, we have

$$\begin{aligned}
v(f(t^n, z)) &= v\left(\prod_{w \in \mu_n} (z - y(wt))\right) \\
&= v(z - y(t)) + v\left(\prod_{w \neq 1} (z - y(wt))\right) \\
&\leq em_h + v\left(\prod_{wm \neq 1} (y(t) - y(wt))\right) && \text{(by (12.3.1))} \\
&= em_h + e\,\mathrm{ord}_t\left(\prod_{w \neq 1} (y(t) - y(wt))\right) \\
&= em_h + e(s_h - m_h) && \text{(by (7.8))} \\
&= es_h.
\end{aligned}$$

□

(12.4) Theorem (Irreducibility Criterion).

Let k be an algebraically closed field and let n be a positive integer such that char k does not divide n. Let $f = f(X, Y)$, $\varphi = \varphi(X, Y)$ be elements of $k((X))[Y]$ such that f and φ are monic in Y and $\deg_Y f = \deg_Y \varphi = n$. Assume that f is irreducible in $k((X))[Y]$, and let $y(t)$ be an element of

$K((t))$ such that $f(t^n, y(t)) = 0$. Let ν be an integer such that $|\nu| = n$. Suppose that

$$\operatorname{ord}_t \varphi(t^n, y(t)) > s_h(\nu, f),$$

where $h = h(f)$. Then:

(i) φ is irreducible in $k((X))[Y]$.

(ii) There exists $z(t) \in k((t))$ such that $\varphi(t^n, z(t)) = 0$ and $\operatorname{ord}_t(z(t) - y(t)) > m_h(\nu, f)$. **100**

Proof. We shall use the notation of (12.1).

(i) Let L be a finite algebraic normal extension of $k((t))$ such that L contains the splitting field of $\varphi(t^n, Y)$ over $k((t^n))$. Then there exist $z_1, \ldots, z_n \in L$ such that we have

$$\varphi(t^n, Y) = \prod_{i=1}^{n} (Y - z_i). \tag{12.4.1}$$

Let v be a valuation of L extending the valuation ord_t of $k((t))$. (See (12.2).) Let $e = v(t)$. Then we have $v(a) = e \operatorname{ord}_t a$ for every $a \in k((t))$. Now, we have

$$\operatorname{ord}_t \varphi(t^n, y(wt)) = \operatorname{ord}_t \varphi(t^n, y(t)) > s_h$$

for every $w \in \mu_n$. Therefore $v(\varphi(t^n, y(wt))) > es_h$ for every $w \in \mu_n$ and it follows that

$$\begin{aligned}
nes_h &< v\left(\prod_{w \in \mu_n} \varphi(t^n, y(wt))\right) \\
&= v\left(\prod_{i=1}^{n} \prod_{w \in \mu_n} (z_i - y(wt))\right) \qquad \text{(by (12.4.1))} \\
&= v\left(\prod_{i=1}^{n} f(t^n, z_i)\right) \\
&= \sum_{i=1}^{n} v(f(t^n, z_i)).
\end{aligned}$$

Therefore there exists i_0, $1 \leq i_0 \leq n$, such that, writing $z = z_{i_0}$, we have $v(f(t^n, z)) > es_h$. It therefore follows from Lemma (12.3) that there exists $w' \in \mu_n$ such that we have

(12.4.2) $$v(z - y(w't)) > em_h.$$

Put $y' = y(w't)$. For $w \in \mu_n$, let σ_w be the $k((t^n))$-automorphism
101 of $k((t))$ defined by $\sigma_w(t) = wt$. Let τ_w be an extension of σ_w to an automorphism of L. Since $k((t))$ it complete with respect to the valuation ord_t, v is the only valuation of L extending ord_t. Therefore, since $\mathrm{ord}_t = \mathrm{ord}_t \circ \sigma_w$, we have $v = v \circ \tau_w$ for every $w \in \mu_n$. In particular, from (12.4.2) we get

(12.4.3) $$v(\tau_w(z) - \tau_w(y')) = v(z - y') > em_h$$

for every $w \in \mu_n$. Moreover, if $w_1, w_2 \in \mu_n$, $w_1 \neq w_2$, then by Proposition (6.15) we have

(12.4.4) $$v(\tau_{\omega_1}(y') - \tau_{\omega_2}(y')) = e\,\mathrm{ord}_t(y(w_1w't) - y(w_2w't)) \leq em_h.$$

Therefore, since

$$\begin{aligned}\tau_{w_1}(z) - \tau_{w_2}(z) = (\tau_{w_1}(z) &- \tau_{w_1}(y')) \\ &+ (\tau_{w_1}(y') - \tau_{w_2}(y')) + (\tau_{w_2}(y') - \tau_{w_2}(z)),\end{aligned}$$

it follows from (12.4.3) and (12.4.4) that $v(\tau_{w_1}(z) - \tau_{w_2}) \leq em_h$ if $w_1 \neq w_2$. In particular, $\tau_{w_1}(z) \neq \tau_{w_2}(z)$ if $w_1 \neq w_2$. Therefore the set $S = \left\{\tau_w(z) \middle| w \in \mu_n\right\}$ consists of n distinct elements. Since all the n elements of S are conjugates of z over $k((t^n))$, the minimal polynomial of z over $k((t^n))$ has degree at least n. On the other hand, $\varphi(t^n, Y) \in k((t^n))[Y]$, $\deg_Y \varphi(t^n, Y) = n$ and $\varphi(t^n, z) = 0$. Therefore $\varphi(t^n, Y)$ is irreducible in $k((t^n))[Y]$. This means that $\varphi(X, Y)$ is irreducible in $k((X))[Y]$. This proves (i).

(ii) Since φ is irreducible by (i), all the roots of $\varphi(t^n, Y)$ belong to $k((t))$ by Newton's Theorem (5.14). Therefore $\tau_w(z) \in k((t))$ for every $w \in \mu_n$. Now, taking $z(t) = \tau_w(z)$ with $w = w'^{-1}$, (ii) follows from (12.4.3).

□

13 Irreducibility of the Approximate Roots

(13.1)

Let k be an algebraically closed field and let $f = f(X,Y)$ be an irre- **102**
ducible element of $k((X))[Y]$. Assume that f is monic in Y and that char k does not divide $n = \deg_Y f$. Let ν be an integer such that $|\nu| = n$. With this notation, we have the following theorem:

(13.2) THEOREM. *Let $y(t)$ be an element of $k((t))$ such that $f(t^n, y(t)) = 0$. Let e be an integer such that $1 \le e \le h(f)+1$ and let*

$$g_e = g_e(X,Y) = App_Y^{d_e}(f).$$

where $d_e = d_e(f)$. Then:

(i) *g_e is irreducible in $k((X))[Y]$.*

(ii) *If $e \ge 2$ then there exists an element $z(t)$ of $k((t))$ such that $g_e(t^{n/d_e}, z(t)) = 0$ and $\mathrm{ord}_t(z(t^{d_e}) - y(t)) = m_e(\nu, f)$.*

Proof.

(i) If $e = 1$ then $\deg_Y g_e = n/d_1 = 1$, so that the assertion is clear in this case. If $e = h(f)+1$ then $g_e = f$, so that the assertion is clear also in this case. We assume now that $2 \le e \le h(f)$. Write $y(t) = \sum y_j t^j$ with $y_j \in k$ for every j, and let $\overline{y}(t) = \sum\limits_{j<m_e} y_j t^j$, where $m_e = m_e(\nu, f)$. Let $G_e = G_e(X,Y)$ be the pseudo d_e^{th} root of f. Recall that G_e is the minimal monic polynomial of $\overline{y}(t)$ over $k((t^n))$. Now, by Proposition (6.13) (ix) d_e divides j for every $j \in \mathrm{Supp}_t\, \overline{y}(t)$. Therefore there exists $y'(t) \in k((t))$ such that $\overline{y}(t) = y'(t^{d_e})$. Put $n' = n/d_e, t' = t^{d_e}$. Then we have $G_e(t'^{n'}, y'(t')) = G_e(t^n, \overline{y}(t)) = 0$. Let $\nu' = \nu/d_e$. Now, in order to prove (i), it is enough to show that

$$\text{(13.2.1)} \qquad ord_{t'}\, g_e(t'^{n'}, y'(t')) > s_{h'}(\nu', G_e),$$

where $h' = h(G_e)$. For, given (13.2.1), we can apply Theorem **103**

(12.4) with f (resp. φ) replaced by G_e (resp. g_e) and conclude that g_e is irreducible. Now, (13.2.1) is clearly equivalent to

$$\text{(13.2.2)} \qquad \mathrm{ord}_t\, g_e(t^n, \overline{y}(t)) > s_{h'}(\nu', G_e)d_e.$$

By Proposition (6.16) we have $h' = e - 1$ and

$$s_{h'}(\nu', G_e)d_e = s_{e-1}(\nu, f)/d_e < s_e(\nu, f)/d_e = r_e(\nu, f).$$

Therefore, in order to prove (13.2.2), it is enough to prove that

$$\text{(13.2.3)} \qquad \mathrm{ord}_t\, g_e(t^n, \overline{y}(t)) > r_e(\nu, f).$$

Now, (13.2.3) follows from Corollary (7.20) by taking $a = 0$ and $u = 0$. This completes the proof of (i).

(ii) If $e = h(f) + 1$ then $d_e = 1$, $g_e = f$ and $m_e = \infty$. Therefore in this case the assertion is clear by taking $z(t) = y(t)$. Now, suppose $2 \leq e \leq h(f)$. Then, in view of (13.2.1), it follows from Theorem (12.4) that there exists $z'(t') \in k((t'))$ such that $g_e(t'^{n'}, z'(t')) = 0$ and

$$\text{(13.2.4)} \qquad \mathrm{ord}_{t'}(z'(t') - y(t')) > m_{h'}(\nu', G_e).$$

Therefore by Proposition (6.17) we get

$$\text{(13.2.5)} \quad h(g_e) = h', m(\nu', g_e) = m(\nu', G_e), S(\nu', g_e) = s(\nu', G_e).$$

In particular, from (13.2.4) we get

$$\text{(13.2.6)} \qquad \mathrm{ord}_t(y'(t') - z'(t')) > m_{h'}(\nu', g_e).$$

104 Now, by Corollary (7.10) applied to (13.2.6) by replacing f (resp. $y(t)$, resp. $u(t)$) by g_e (resp. $z'(t')$, resp. $y'(t')$), we get

$$\mathrm{ord}_{t'}(g_e(t'^{n'}, y'(t'))) = s_{h'}(\nu', g_e) - m_{h'}(\nu', g_e) + \mathrm{ord}_{t'}(y'(t') - z'(t')).$$

From this, by (13.2.5) we get

$$\mathrm{ord}_{t'}(g_e(t'^{n'}, y'(t'))) = s_{h'}(\nu' G_e) - m_{h'}(\nu' G_e) + \mathrm{ord}_{t'}(y'(t') - z'(t')).$$

Now, since $t' = t^{d_e}$, there exists $z(t) \in k((t))$ such that $z'(t') = z(t^{d_e})$, and we get

$$\operatorname{ord}_t(g_e(t^n, \overline{y}(t))) = d_e s_{h'}(\nu', G_e) - d_e m_{h'}(\nu', G_e) + \operatorname{ord}_t(\overline{y}(t) - z(t^{d_e})).$$

Therefore by (13.2.3) we get

$$\begin{aligned} r_e(\nu, f) &< d_e s_{h'}(\nu', G_e) - d_e m_{h'}(\nu', G_e) + \operatorname{ord}_t(\overline{y}(t) - z(t^{d_e})) \\ &= s_{e-1}(\nu, f)/d_e - m_{e-1}(\nu, f) + \operatorname{ord}_t(\overline{y}(t) - z()t^{d_e}) \end{aligned}$$

by Proposition (6.16). This gives

$$\begin{aligned} \operatorname{ord}_t(\overline{y}(t) - z(t^{d_e})) &> m_{e-1}(\nu, f) + f_e(\nu, f) - s_{e-1}(\nu, f)/d_e \\ &= m_{e-1}(\nu, f) + (s_e(\nu, f) - s_{e-1}(\nu, f))/d_e \\ &= m_{e-1}(\nu, f) + q_e(\nu, f) \\ &= m_e(\nu, f). \end{aligned}$$

Therefore, since $\operatorname{ord}_t(\overline{y}(t) - y(t)) = m_e(\nu, f)$, we get

$$\begin{aligned} \operatorname{ord}_t(z(t^{d_e}) - y(t)) &= \operatorname{ord}_t((z(t^{d_e}) - \overline{y}(t)) + (\overline{y}(t) - y(t))) \\ &= m_e(\nu, f). \end{aligned}$$

Also, from $g_e(t'^{n'}, z'(t')) = 0$ we get $g_e(t^{n/d_e}, z(t)) = 0$. This completes the proof of (ii). □ **105**

(13.3) COROLLARY. Let f and ν be as in (13.1). Let e be an integer, $2 \le e \le h(f) + 1$. Let $g_e = App_Y^{d_e}(f)$, where $d_e = d_e(f)$. Let $\nu' = \nu/d_e$. Then $h(g_e) = e - 1$ and for $0 \le i \le e - 1$ we have

$$\begin{aligned} m_i(\nu', g_e) &= m_i(\nu, f)/d_e, \\ q_i(\nu', g_e) &= q_i(\nu, f)/d_e, \\ s_i(\nu', g_e) &= s_i(\nu, f)/d_e^2 \quad (\text{if } i \ne 0). \\ r_i(\nu', g_e) &= r_i(\nu, f)/d_e, \\ d_{i+1}(g_e) &= d_{i+1}(f)/d_e. \end{aligned}$$

Proof. This is immediate from Theorem (13.2) (ii). □

(13.4) COROLLARY. Let char $k = 0$. Let $\varphi = \varphi(X, Y)$ be an element of $k[X, Y]$ such that $n = \deg_Y \varphi > 0$, φ is monic in Y and $k[X, Y]/(\varphi)$ is isomorphic (as a k-algebra) to $k[Z]$, where Z is an indeterminate. Let $f = f(X, Y) = \varphi(X^{-1}, Y)$. Then f is irreducible in $k((X))[Y]$. Let $h = h(f)$ and for $1 \leq e \leq h + 1$ let $\psi_e = App_Y^{d_e}(\varphi)$, where $d_e = d_e(f)$. Then $k[X, Y]/(\psi_e)$ is isomorphic (as a k-algebra) to $k[Z]$ for every e, $1 \leq e \leq h + 1$.

Proof. The irreducibility of f follows from Theorem (9.24). Now, since
106 $d_1(f) = n$, ψ_1 is monic in Y of Y-degree one. Therefore the assertion is clear for $e = 1$. For $2 \leq e \leq h + 1$ we prove the assertion by decreasing induction on e. If $e = h + 1$ then $d_e = 1$, so that $\psi_e = \varphi$ and the assertion follows from the hypothesis. Now, let $2 \leq e \leq h(f)$ and suppose $k[X, Y]/(\psi_{e+1})$ is isomorphic to $k[Z]$. Let $g_{e+1} = App_Y^{d_{e+1}}(f)$. Then by Proposition (4.7) we have $g_{e+1}(X, Y) = \psi_{e+1}(X^{-1}, Y)$. Let $h' = h(g_{e+1})$. Then by Corollary (13.3) we have $h' = e$ and $d_{h'}(g_{e+1}) = d_e/d_{e+1}$. If follows that $\psi_e = App_Y^{d_{h'}}(\psi_{e+1})$, where $d_{h'} = d_{h'}(g_{e+1})$. Now it follows from Corollary (9.28) that $k[X, Y]/(\psi_e)$ is isomorphic to $k[Z]$. □

(13.5) COROLLARY. With the notation and assumptions of Corollary (13.4) , let $h = h(f)$ and let $m_i = m_i(-n, f)$, $q_i = q_i(-n, f)$, $s_i(-n, f)$, $r_i = r_i(-n, f)$ and $d_{i+1} = d_{i+1}(f)$ for $0 \leq i \leq h$. Then we have:

(i) $r_i = -d_{i+1}$ for $2 \leq i \leq h$.

(ii) $s_i = -d_i d_{i+1}$ for $2 \leq i \leq h$.

(iii) $q_i = d_{i-1} - d_{i+1}$ for $3 \leq i \leq h$.

(iv) $m_i = d_1 - d_i - d_{i+1}$ for $2 \leq i \leq h$.

(v) If $h \geq 2$ then $m_i < n - 2$ for every i, $1 \leq i \leq h$.

Proof.

(i) Fix an e, $2 \leq e \leq h$, and let $\psi = App_Y^{d_{e+1}}(\varphi)$. Then by Corollary (13.4) $k[X, Y]/(\psi)$ is isomorphic to $k[Z]$. Let $g = g(X, Y) =$

$\psi(X^{-1}, Y)$. Then $g = App_Y^{d_{e+1}}(f)$. Let $h' = h(g)$. Then by Corollary (13.3) we have $h' = e$ and $d_{h'}(g) = d_e/d_{e+1}$. Noting that $\deg_Y \psi = n/d_{e+1}$ and $h' = e \geq 2$, it follows from Corollary (9.25) that we have $r_{h'}(-n/d_{e+1}, g) = -1$. By Corollary (13.3) we have $r_{h'}(-n/d_{e+1}, g) = r_e(-n, f)/d_{e+1} = r_e/d_{e+1}$. Thus we have $-1 = r_e/d_{e+1}$, and (i) is proved.

(ii) This is immediate from (i), since $s_i = d_i r_i$. **107**

(iii) By (ii) we have

$$\begin{aligned} -d_i d_{i+1} &= s_i \\ &= s_{i-1} + q_i d_i \\ &= -d_{i-1} d_i + q_i d_i, \end{aligned}$$

since $i \geq 3$. This gives $q_i = d_{i-1} - d_{i+1}$.

(iv) For $i \geq 3$ we have

$$\begin{aligned} m_i &= m_{i-1} + q_i \\ &= m_{i-1} + d_{i-1} - d_{i+1} \end{aligned}$$

by (iii). Therefore, by induction on i, it is enough to prove that $m_2 = d_1 - d_2 - d_3$. Now, by (ii) we have $-d_2 d_3 = s_2 = q_1 d_1 + q_2 d_2$. Therefore we get

$$m_2 = q_1 + q_2 = -q_1((d_1/d_2) - 1) - d_3.$$

Now, by Corollary (9.27) we have $d_2 = d_1$ or $d_2 = -q_1$. We consider the two cases separately.

Case(1). $d_2 = d_1$. Then $m_2 = -d_3 = d_1 - d_2 - d_3$.

Case (2). $d_2 = -q_1$. Then

$$m_2 = d_2((d_1/d_2) - 1) - d_3 = d_1 - d_2 - d_3.$$

(v) Suppose $h \geq 2$. It is enough to prove that $m_h < n - 2$. By (iv) we have $m_h = d_1 - d_h - d_{h+1} < d_1 - 2 = n - 2$, since $d_{h+1} = 1$ and $d_h \geq 2$.

□

(13.6) REMARK. Corollaries (13.4) and (13.5) hold also for char $k > 0$ (and, in fact, the same proof goes through) provided we assume that n is not divisible by char k.

108 **(13.7) PROPOSITION.** Let f and ν be as in (13.1). Let e be an integer, $1 \le e \le h(f)$. Let $y(t)$ be an element of $k((t))$ such that $f(t^n, y(t)) = 0$. Let k' be an overfield of k and let $y^*(t)$ be an element of $k'((t))$ such that $\mathrm{ord}_t(y^*(t) - y(t)) \ge m_e(\nu, f)$ and $m_e(\nu, f) \in \mathrm{Supp}_t\, y^*(t)$. Let $g_e = g_e(X, Y)$ be defined as follows: If $e \ge 2$ then $g_e = App_Y^{d_e}(f)$, whereas if $e = 1$ then $g_1 = App_Y^{d_1}(f)$ or $g_1 = Y$, where $d_e = d_e(f)$. Let g'_e denote the Y-derivative of g_e. Then we have

$$\mathrm{ord}_t\, g'_e(t^n, y^*(t)) = r_e(\nu, f) - m_e(\nu, f).$$

Proof. With either definition of g_1 we have $g'_1 = 1$. Therefore, since $r_1(\nu, f) = m_1(\nu, f)$, the assertion is clear in case $e = 1$. Assume now that $e \ge 2$. By Theorem (13.2) g_e is irreducible in $k((X))[Y]$. Put $d = d_e$, $g = g_e, h' = h(g)$, $\nu' = \nu/d$, $s'_{h'} = s_{h'}(\nu', g)$, $m'_{h'} = m_{h'}(\nu', g)$. Then by Corollary (7.9) applied to g we have

$$\mathrm{ord}_t\, g'(t^{n/d}, z(t)) = s'_{h'} - m'_{h'}, \tag{13.7.1}$$

where $g' = g'_e$ and $z(t) \in k((t))$ is any zero of $g(t^{n/d}, Y)$. Put $m_i = m_i(\nu, f)$, $q_i = q_i(\nu, f)$, $s_i = s_i(\nu, f)$ and $r_i = r_i(\nu, f)$ for $0 \le i \le h(f)$. then by Corollary (13.3) we have $h' = e - 1$, $s'_{h'} = s_{e-1}/d^2$, $m'_{h'} = m_{e-1}/d$. Therefore

$$\begin{aligned} d(s'_{h'} - m'_{h'}) &= s_{e-1}/d - m_{e-1} \\ &= s_e/d - q_e - m_{e-1} \\ &= r_e - m_e. \end{aligned}$$

Therefore it follows from (13.7.1) that we have

$$\mathrm{ord}_t\, g'(t^n, z(t^d)) = r_e - m_e \tag{13.7.2}$$

109 for any zero $z(t)$ of $g(t^{n/d}, Y)$. By Theorem (13.2) we may choose $z(t)$

such that $\operatorname{ord}_t(y(t) - z(t^d)) = m_e$. Then, since $\operatorname{ord}_t(y^*(t) - y(t)) \geq m_e$ and $m_e \in \operatorname{Supp}_t y^*(t)$ by assumption and since $m_e \notin \operatorname{Supp}_t z(t^d)$, we get

$$\operatorname{ord}_t(y^*(t) - z(t^d)) = m_e. \tag{13.7.3}$$

Now, we have

$$g(t^{n/d}, Y) = \prod_{w \in \mu_{n/d}} (Y - z(wt)),$$

where $\mu_{n/d} = \mu_{n/d}(k)$. Therefore

$$g(t^n, Y) = \prod_{w \in \mu_{n/d}} (Y - z(wt^d)).$$

differentiating with respect to Y and then substituting $y = y^*(t)$, we get

$$g'(t^n, y^*(t)) = \sum_{v \in \mu_{n/d}} \prod_{w \neq v} (y^*(t) - z(wt^d)$$

$$P_1 + \sum_{\substack{v \in \mu_{n/d} \\ v \neq 1}} P_v,$$

where $P_v = \prod_{w \neq v} (y^*(t) - z(wt^d))$. Thus, in order to complete the proof of the proposition, it is now enough to prove the following two statements:

(i) $\operatorname{ord}_t P_1 = r_e - m_e$.

(ii) $\operatorname{ord}_t P_v > r_e - m_e$ for every $v \in \mu_{n/d} - \{1\}$.

Since we have

$$y^*(t) - z(wt^d) = (y^*(t) - z(t^d)) + (z(t^d) - z(wt^d))$$

and since for $w \neq 1$ **110**

$$\operatorname{ord}_t(z(t^d) - z(wt^d)) \leq dm'_{h'} \qquad \text{(Proposition (6.15))}$$

$$m_{e-1} \qquad \text{(Corollary (13.3))}$$

$$=< m_e,$$

it follows from (13.7.3) that we have

$$\operatorname{ord}_t(y^*(t) - z(wt^d)) = \operatorname{ord}(z(t^d) - z(wt^d)) < m_e \tag{13.7.4}$$

for $w \neq 1$. Therefore

$$\begin{aligned} \operatorname{ord}_t P_1 &= \operatorname{ord}_t \prod_{w\neq 1}(z(t^d) - z(wt^d)) \\ &= \operatorname{ord}_t g'(t^n, z(t^d)) \\ &= r_e - m_e \end{aligned}$$

by (13.7.2). This proves (i). Now, let $v \in \mu_{n/d}, v \neq 1$. We have

$$P_v = P_1(y^*(t) - z(t^d))(y^*(t) - z(vt^d))^{-1}.$$

Therefore by (i) we have

$$\operatorname{ord}_t P_v = r_e - m_e + \operatorname{ord}_t(y^*(t) - z(t^d)) - \operatorname{ord}_t(y^*(t) - z(vt^d)).$$

Therefore (ii) will be proved if we show that

$$\operatorname{ord}_t(y^*(t) - z(t^d)) > \operatorname{ord}_t(y^*(t) - z(vt^d)).$$

Since $v \neq 1$, this last inequality is clear from (13.7.3) and (13.7.4). □

14 Newton's Algebraic Polygon

(14.1)

111 We revert to the notation of (7.1), (7.2) and (7.3). In addition, we fix the following notation: for an integer m, we put

$$p(m) = \inf \left\{ i \middle| 1 \leq i \leq h+1, m < m_i \right\}.$$

Let $d^*(m) = d_{p(m)}$ and let

$$s^*(m) = \begin{cases} s_{p-1} + (m - m_{p-1})d_p, & \text{if } p = p(m) \geq 2. \\ md_1, & \text{if } p(m) = 1. \end{cases}$$

Note that $p(m_i) = i + 1$, $d^*(m_i) = d_{i+1}$ and $s^*(m_i) = s_i$ for $1 \leq i \leq h$. If Z is an indeterminate, define

$$P(m, Z) = \begin{cases} Z - y_m, & \text{if } m \notin \{m_1, \ldots, m_h\}, \\ Z^{n_e} - y_{m_e}^{n_e}, & \text{if } m \in \{m_1, \ldots, m_h\}, \end{cases}$$

where $e = p(m) - 1$.

with the above notation, we have

(14.2) THEOREM. *Let m be an integer. Let Z be an indeterminate and let $k' = k(Z)$. Let y^* be an element of $k'((t))$ such that*

$$\textit{info}\ (y^* - y(t)) = (Z - y_m)t^m.$$

Then

$$\textit{info}\ (f(t^n, y^*)) = \varnothing P(m, Z)^{d^*(m)} t^{s^*(m)}.$$

Proof. Suppose $m \in \{m_1, \ldots, m_h\}$. say $m = m_e$. Then $p(m) = e + 1$. Let $\overline{y}(t) = \sum_{j<m_e} y_j t^j$. Then it easily follows from the assumption on y^* that we **112**
have

$$\text{info}\ (y^* - \overline{y}(t)) = Zt^{m_e}.$$

Therefore y^* is an (e, Z)-deformation of $y(t)$ and it follows from Lemma (7.16) that we have

$$\text{info}\ (f(t^n, y^*)) = \varnothing \left(Z^{n_e} - y_{m_e}^{n_e}\right)^{d_{e+1}} t^{s_e}.$$

Since $d^*(m_e) = d_{e+1}$ and $s^*(m_e) = s_e$, the assertion is proved in case $m \in \{m_1, \ldots, m_h\}$. □

Now, suppose $m \notin \{m_1, \ldots, m_h\}$. Let $p = p(m)$. Let $Q(p)$, $R(p)$ be the sets defined in Definition (7.4). If $w \in R(p)$ then $\operatorname{ord}(y(t) - y(wt)) \geq m_p > m$. Therefore, since

(14.2.1) $$y^* - y(wt) = (y^* - y(t)) + (y(t) - y(wt)),$$

we get info $(y^* - y(wt)) = \text{info } (y^* - y(t)) = (Z - y_m)t^m$ for $w \in R(p)$. This shows that we have

$$\begin{aligned} \text{info} \left(\prod_{w \in R(p)} (y^* - y(wt)) \right) &= \prod_{w \in R(p)} (Z - y_m)t^m \\ &= (Z - y_m)^{d^*(m)} t^{md^*(m)}, \end{aligned} \tag{14.2.2}$$

since by Lemma (7.5) card $(R(p)) = d_p = d^*(m)$. Now, suppose $w \in Q(p)$ and $p \geq 2$. Then by Proposition (6.15) we get $\text{ord}_t(y(t)) \leq m_{p-1}$. Since $m \notin \{m_1, \ldots, m_h\}$, we have $m_{p-1} < m$. Therefore from (14.2.1) we get

$$\text{info } (y^* - y(wt)) = \text{info } (y(t) - y(wt)) \text{ for } w \in Q(p). \tag{14.2.3}$$

Since $Q(1) = \phi$, (14.2.3) holds also for $p = 1$. Now, clearly, inco

113 $(y(t) - y(wt)) = \varnothing$ for every $w \in Q(p)$. Therefore we get

$$\begin{aligned} \text{info} \left(\prod_{w \in Q(p)} (y^* - y(wt)) \right) &= \text{info} \left(\prod_{w \in Q(p)} (y(t) - y(wt)) \right) \\ &= \varnothing t^s, \end{aligned} \tag{14.2.4}$$

where by Lemma (7.7) we have

$$s = \begin{cases} s_{p-1} - m_{p-1}d_p, & \text{if } p \geq 2, \\ 0, & \text{if } p = 1. \end{cases}$$

From (14.2.2) and (14.2.4) we get

$$\begin{aligned} \text{info } (f(t^n, y^*)) &= \text{info} \left(\prod_{w \in \mu_n(k)} (y^* - y(wt)) \right) \\ &= \varnothing (Z - y_m)^{d^*(m)} t^{md^*(m)+s} \\ &= \varnothing P(m, Z)^{d^*(m)} t^{s^*(m)}. \end{aligned}$$

(14.3) REMARK. The above theorem is an algebraic version of the method of Newton's polygon for constructing a root in $k((t))$ of the equation $f(t^n, Y) = 0$. The successive coefficients y_j of a root $y(t) = \sum y_j t^j$

are found by induction on j. Thus, suppose we know y_j for j less than a certain integer m. Let Z be an indeterminate and let $y^* = \sum_{j<m} y_j t^j + Zt^m$. Find inco $(f(t^n, y^*))$. This will be a certain polynomial $F(Z) \in k[Z]$, viz. $F(Z) = \varnothing P(m, Z)^{d^*(m)}$. Take y_m to be any root of the equation $F(Z) = 0$. Note that if $m \notin \{m_1, \ldots, m_h\}$ then $F(Z) = 0$ will have a unique root, whereas if $m = m_e$ for some e, $1 \leq e \leq h$, then $F(Z) = 0$ will have n_e distinct roots. Let us remark that, since $f(t^n, 0) = (-1)^n \prod y(wt)$, we have $m_1 = \text{ord}_X\, f(X, 0)$. Therefore we may *start* the inductive construction of y_j by taking $y_j = 0$ for all $j < \text{ord}_X\, f(x, 0)$.

Part II

The Jacobian Problem

Chapter 6

The Jacobian Problem

15 Statement of the Problem

(15.1)

Let k be a field and let $A = k[x_1, x_2]$ be the polynomial ring in two vari- 117
ables x_1, x_2 over k. Let K be the quotient field of A. A pair (u_1, u_2) of elements of A is an *automorphic pair* (for A) if $A = k[u_1, u_2]$. Note that (u_1, u_2) is an automorphic pair if and only if the k-algebra homomorphism $\sigma : A \to A$ defined by $\sigma(x_i) = u_i$, $i = 1, 2$, is an automorphism. A pair (u_1, u_2) of elements of K is a *transcendence base* (of K over k) if K is algebraic over $k(u_i, u_2)$. Clearly, every automorphic pair is a transcendence base.

Let $u = (u_1, u_2)$ be a transcendence base. Then u_1, u_2 are algebraically independent over k. Therefore there exist k-derivations $D_{u,1}$, $D_{u,2}$ of $k(u_1, u_2)$ defined by $D_{u,i}(u_j) = \delta_{ij}$ (Kronecker delta). Suppose now that K is separable over $k(u_1, u_2)$. Then for each $i = 1, 2, D_{u,i}$ extends to a unique k-derivation of K. We shall denote this extension also by $D_{u,i}$, $i = 1, 2$. In particular, for each automorphic pair $u = (u_1, u_2)$ we have k-derivations $D_{u,i}$ of K, $i = 1, 2$. We shall often write simply D_i for $D_{x,i}$, $i = 1, 2$, where $x = (x_1, x_2)$. Note that if u is an automorphic pair then $d_{u,i}(A) \subset A$, $i = 1, 2$.

(15.2) DEFINITION. Let $u = (u_1, u_2)$ be an automorphic pair and let $f, g \in A$. The *Jacobian of* (f, g) *with respect to* u, denoted $J_u(f, g)$, is defined by

$$J_u(f,g) = \det\begin{pmatrix} D_{u,1}(f) & D_{u,2}(f) \\ D_{u,1}(g) & D_{u,2}(g) \end{pmatrix} = D_{u,1}(f)D_{u,2}(g) - D_{u,2}(f)D_{u,1}(g).$$

We shall write simply $J(f,g)$ for $J_x(f,g)$.

118 **(15.3) LEMMA.** *Let $u = (u_1, u_2)$, $v = (v_1, v_2)$ be automorphic pairs for A and let $f, g \in A$. Then we have*

$$J_u(f,g) = J_v(f,g)J_u(v_1, v_2).$$

Proof. This is immediate from the chain rule for derivations, namely

$$D_{u,i}(a) = D_{u,1}(a)D_{u,i}(v_1) + D_{v,2}(a)D_{u,i}(v_2) \quad \text{for } a \in A, i = 1, 2.$$

□

(15.4) COROLLARY. Let $u = (u_1, u_2)$, $v = (v_1, v_2)$ be automorphic pairs for A. Then $J_u(v_1, v_2)$ is a unit of A.

Proof. By Lemma (15.3) we have

$$1 = J_u(u_1, u_2) = J_v(u_1, u_2)J_u(v_1, v_2)$$

and the corollary is proved. □

(15.5)

Noting that the units of A are the non-zero elements of k, it follows from Corollary (15.4) that if (f,g) is an automorphic pair for A then $J(f,g)$ is a non-zero element of k. Then Jacobian problem asks whether the converse is true in case char $k = 0$:

The Jacobian Problem. Suppose char $k = 0$. Let f, g be elements of A such that $J(f,g)$ is a non-zero element of k. Is (f,g) then an automorphic pair for A?

(15.6) REMARK. Suppose char $k = p > 0$. Let $f = x_1 + x_1^p$, $g = x_2$. Then $J(f,g) = 1$. Then $J(f,g) = 1$. However, (f,g) is not an automorphic pair. For, $k[x_1, x_2]/(g) = k[x_1] \neq k[x_1 + x_1^p]$, which shows that $k[f,g] \neq k[x_1, x_2]$. This explains the assumption char $k = 0$ made in the Jacobian problem.

16 Notation

(16.1)

Let $A = k[x_1, x_2]$ as in § 15. *We assume henceforth that* char $k = 0$. Let **119**
$w = (w_1, w_2)$ be a pair of integers. By the *w-gradation* on A we mean the gradation on A obtained by giving weight w_i to x_i, $i = 1, 2$. Recall that this means that we write A in the form

$$A = \bigoplus_{n \in \mathbb{Z}} A_w^{(n)},$$

where $A_w^{(n)}$ is the k-subspace of A generated by monomials $x_1^{i_1} x_2^{i_2}$ with $i_1 w_1 + i_2 w_2 = n$. The elements of $A_w^{(n)}$ are called *w-homogeneous* elements of *w-degree n*. Note that by this definition 0 is w-homogeneous of w-degree n for every n. Every element f of A can be written uniquely in the form $f = \sum_n f_w^{(n)}$, where $f_n^{(n)}$ is w-homogeneous of w-degree n and $f_w^{(n)} = 0$ for almost all n. We call $f_w^{(n)}$ the n^{th} *w-homogeneous component* of f. Suppose $f \neq 0$. Then there exists $m \in \mathbb{Z}$ such that $f_w^{(m)} \neq 0$ and $f_w^{(n)} = 0$ for all $n > m$. We call this m the *w-degree* of f and denote it by $d_w(f)$. Thus

$$d_w(f) = \sup \left\{ n \in \mathbb{Z} \,\middle|\, f_w^{(n)} \neq 0 \right\}.$$

If $f = 0$, we define $d_w(f) = -\infty$. If $f \neq 0$ then the *w-degree form* of f, denoted f_w^+, is defined by $f_w^+ = f_w^{(m)}$, where $m = d_w(f)$. If $f = 0$, we define $f_w^+ = 0$. Note that f is w-homogeneous if and only if $f = f_w^+$.

Suppose now that $w = (1, 1)$. then the w-gradation on A is called the *usual gradation* on A. In this case we often omit the symbol w in the notation introduced above. Thus we write $d(f)$, $f^{(n)}$, $f^+, \ldots$ etc. for $d_w(f)$, $f_w^{(n)}$, $f_w^+, \ldots$. when $w = (1, 1)$.

(16.2)

The w-gradation on A defined in (16.1) above is with respect to the au- **120**
tomorphic pair $x = (x_1, x_2)$. If $u = (u_1, u_2)$ is any automorphic pair then we can also define a gradation on A by giving weight w_i to u_i, $i = 1, 2$.

However, in the sequel we will mostly need to consider only the $(1,1)$-gradation on A with respect to an arbitrary automorphic pair u. In order to distinguish this from the usual gradation, we fix the following notation:

$\deg f$ *denotes* $d_{(1,1)}(f)$ with respect to x,
$\deg_u f$ *denotes* $d_{(1,1)}(f)$ with respect to u,

If $u = (u_1, u_2)$ is an automorphic pair and $f \in A$, we write $\deg_{u_1} f$ (resp. $\deg_{u_2} f$) for the u_1- degree (resp. u_2-degree) of f regarded as a polynomial in u_1 (resp. u_2) with coefficients in $k[u_2]$ (resp. $k[u_1]$).

(16.3)

One final piece of notation: We denote by k^* the set of non-zero elements of k and, as noted in (7.2), we use the symbol $\varnothing$ to denote a generic (i.e., unspecified element of k^*.)

17 w-Relation

We preserve the notation of §15 and §16. In particular, we have char $k = 0$. Let $w = (w_1, w_2)$ be a pair of integers.

(17.1) LEMMA. *Let F, G be non-zero w-homogeneous elements of A. The following two conditions are equivalent:*

(1) $F^r = \varnothing G^s$ *for some* $r, s \in \mathbb{Z}^+$; $r + s > 0$.

(2) *There exist p, $q \in \mathbb{Z}^+$ and a w-homogeneous element H of A such that* $F = \varnothing H^p$, $G = \varnothing H^q$.

121 *Proof.* (1) $\Rightarrow$ (2). Write $F = \varnothing H_1^{p_1} \dots H_n^{p_n}$, $G = \varnothing H_1^{q_1} \dots H_n^{q_n}$, where H_i is an irreducible w-homogeneous element of A, p_i, $q_i \in \mathbb{Z}^+$ for $1 \le i \le n$ and g.c.d. $(H_i, H_j) = 1$ for $i \ne j$. Then (1) implies that $rp_i = sq_i$ for every i, $1 \le i \le n$. Now, if $r = 0$ or $s = 0$, say $r = 0$, then $s > 0$ and $q_i = 0$ for every i, so that $G = \varnothing$. In this case (2) follows by taking $H = F$, $p = 1$, $q = 0$. We may therefore assume that $r > 0$ and $s > 0$.

Then for any i, $p_i = 0$ if and only if $q_i = 0$. Therefore we may assume that $p_i > 0$ and $q_i > 0$ for every i, $1 \le i \le n$. Then for every i we have $p_i/q_i = s/r = p/q$, say, where p, q are positive integers such that g.c.d. $(p, q) = 1$. For every i, $1 \le i \le n$, there exists a positive integer t_i such that $p_i = pt_i$, $q_i = qt_i$. Let $H = H_1^{t_1} \dots H_n^{t_n}$. Then $F = \varnothing H^p$, $G = \varnothing H^q$.

(2) $\Rightarrow$ (1). If $p = 0 = q$ then $F = \varnothing$, $G = \varnothing$, so that $F = \varnothing G$, which implies (1) in this case. Assume therefore that $p + q > 0$. Now, (2) implies that $F_q = \varnothing G^p$, which implies (1). □

(17.2) DEFINITION. Let f, $g \in A$. We say f and g are *w-related* if $f \neq 0$, $g \neq 0$, and $F = f_w^+$ and $g = g_W^+$ satisfy the equivalent conditions (1) and (2) of Lemma (17.1). We say f and g are *related* if f and g are (1,1)-related.

(17.3) LEMMA. *Let $f, g_1, \dots, g_e$ be elements of A.*

(i) *If $f_w^+ = \varnothing$ and $g_1 \neq 0$ then f and g_1 are w-related.*

(ii) *If f and g_i are w-related for every i, $1 \le i \le e$, then f and $g_1 \dots g_e$ are w-related.*

Proof.

(i) We have $f_w^+ = \varnothing = \varnothing(g_{1w}^+)^\circ$.

(ii) By induction on e, it is enough to consider the case $e = 2$. There exist r_i, $s_i \in \mathbb{Z}^+$, $r_i + s_i > 0$, such that $F^{r_i} = \varnothing G_i^{s_i}$, where $F = f_w^+$, $G_i = g_{iw}^+$, $i = 1, 2$. This gives

(17.3.1) $$F^{r_1 s_2 + r_2 s_1} = \varnothing (G_1 G_2)^{s_1 s_2}.$$

If $s_i = 0$ for $i = 1$ or 2, say $s_1 = 0$, then $r_1 > 0$ and $F^{r_1} = \varnothing G_1^\circ = \varnothing$ **122**
shows that $F = \varnothing$. Therefore in this case f is related to $g_1 g_2$ by (i). We may therefore assume that $s_1 > 0$, $s_2 > 0$. Then $s_1 s_2 > 0$, and it follows from (17.3.1) that f and $g_1 g_2$ are w-related. □

(17.4) PROPOSITION. Let F, G be non-zero w - homogeneous elements of A of w-degrees m, n, respectively. Consider the following five conditions:

(1) F and G are w-related.

(2) F and G are algebraically dependent over k.

(3) $J(F, G) = 0$.

(4) $F^n = \varnothing G^m$.

(5) $F^{|n|} = \varnothing G^{|m|}$.

Among these five conditions we have the following implications:

(1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5).

Assume, moreover, that at least one of the following two conditions is satisfied:

(i) $w_1 w_2 > 0$ and $F \notin k$ or $G \notin k$.

(ii) $m \neq 0$ or $n \neq 0$.

Then the above five conditions (1) - (5) are equivalent to each other. Further, let d = g.c.d. (m, n). Then $d > 0$ and the conditions (1) - (5) are also equivalent to each of the following two conditions:

(6) $F^{n/d} = \varnothing G^{m/d}$.

(7) We have $mn \geq 0$ and there exists a w-homogeneous element H of A such that $F = \varnothing H^{|m|/d}, G = \varnothing H^{|n|/d}$.

123 In order to prove the proposition, we need the following three lemmas:

(17.14.1) LEMMA.

Let L be a field and let $L(t)$ be the field of rational functions in one variable t over L. Let D_t be the L-derivation of $L(t)$ defined by $D_t(t) = 1$. If h is an element of $L(t)$ such that $D_t(h) = 0$ then $h \in L$.

(17.14.2) LEMMA.

Let f, g be non-zero elements of A of w-degrees m, n, respectively. If $f^n = \varnothing g^m$ then $mn \geq 0$.

(17.14.3) LEMMA.

Let F, G be non-zero w homogeneous elements of A of w-degrees m, n respectively. Then we have:

(i)
$$mF = w_1 x_1 D_1(F) + w_2 x_2 D_2(F),$$
$$nG = w_1 x_1 D_1(G) + w_2 x_2 D_2(G).$$

(ii)
$$w_1 x_1 J(F, G) = mFD_2(G) - nGD_2(F),$$
$$w_2 x_2 J(F, G) = nGD_1(F) - mFd_1(G).$$

Proof of Lemma (17.14.1). The assertion is clear if $h \in L[t]$. In general, we can write $h = f/g$ with f, $g \in L[t]$ and g.c.d. $(f, g) = 1$. Then we have
$$0 = D_t(h) = (gD_t(f) - fD_t(g))/g^2,$$
which shows that $gD_t(f) = fD_t(g)$. Thus g divides $fD_t(g)$ in $L[t]$. Therefore, since g.c.d. $(f, g) = 1$, g divides $D_t(g)$. Since $\deg_t D_t(g) < \deg_t g$, we get $D_t(g) = 0$, so that $g \in L$. Therefore $h \in L[t]$, and the assertion follows.

Proof of Lemma (17.14.2). Suppose $mn < 0$. then one of m, n is positive and the other is negative. We may suppose that $m < 0$ and $n > 0$. Then $f^n g^{-m} = \varnothing$ implies that f (also g) is a unit of A. Therefore $f \in k^*$. But this means that $m = 0$, which is a contradiction. **124**

Proof of Lemma (17.14.3). (i) We have only to observe that
$$w_1 x_1 D_1\left(x_1^{i_1} x_2^{i_2}\right) + \omega_2 x_2 D_2\left(x_1^{i_1} x_2^{i_2}\right) = (i_1 w_1 + i_2 w_2)\, x_1^{i_1} x_2^{i_2}.$$

(ii) We have
$$\begin{aligned} w_1 x_1 J(F, G) &= \det\begin{pmatrix} w_1 x_1 D_1(F) & D_2(F) \\ w_1 x_1 D_1(G) & D_2(G) \end{pmatrix} \\ &= \det\begin{pmatrix} w_1 x_1 D_1(F) + w_2 x_2 D_2(F) & D_2(F) \\ w_1 x_1 D_1(G) + w_2 x_2 D_2(G) & D_2(G) \end{pmatrix} \\ &= \det\begin{pmatrix} mF & D_2(F) \\ nG & D_2(G) \end{pmatrix} \qquad \text{(by (i))} \end{aligned}$$

$$= m\, FD_2(G) - nGD_2(F).$$

This proves the first equality of (ii). The second is proved similarly.

Proof of Proposition (17.4). (1) ⇒ (2). We have $F^r = \varnothing G^s$ for some non-negative integers r, s, $r+s > 0$. Therefore F and G are algebraically dependent over k.

(2) ⇒ (3). Let X_1, X_2 be indeterminates and let $\varphi = \varphi(X_1, X_2) \in k[X_1, X_2]$ be such that $\varphi \neq 0$ and $\varphi(F, G) = 0$. Then $\varphi \notin k$. Therefore $\deg_{X_1} \varphi + \deg_{X_2} \varphi > 0$. We may choose φ to be such that $\deg_{X_1} \varphi + \deg_{X_2} \varphi$ is the least possible. Let $\varphi_i = D_{X,i}(\varphi)$, $i = 1, 2$, where $X = (X_1, X_2)$. Then we have $\deg_{X_1} \varphi_i + \deg_{X_2} \varphi_1 < \deg_{X_2} \varphi + \deg_{X_2} \varphi$, $i = 1, 2$. Moreover
125 $\varphi_1 \neq 0$ or $\varphi_2 \neq 0$. Ir follows that we have $\varphi_1(F, G) \neq 0$ or $\varphi_2(F, G) \neq 0$. Now, we have

$$\begin{aligned} 0 = D_1(\varphi(G, G)) = \varphi_1(F, G)D_1(F) + \varphi_2(F, G)D_1(G), \\ 0 = D_2(\varphi(F, G)) = \varphi_1(F, G)D_2(F) + \varphi_2(F, G)D_2(G). \end{aligned}$$

Since $\varphi_1(F, G) \neq 0$ or $\varphi_2(F, G) \neq 0$, we get

$$0 = \det \begin{pmatrix} D_1(F) & D_1(G) \\ D_2(F) & D_2(G) \end{pmatrix} = J(F, G),$$

which proves (3).

(3) ⇒ (4). We have $J(F, G) = 0$ and we want to show that $F^n/G^m \in k$. Since $k = k(x_1) \cap k(x_2)$, it is enough, by symmetry, to show that $F^n/G^m \in k(x_1)$. By lemma (17.14.3), we have

$$0 = w_1 x_1 J(F, G) = mFD_2(G) - nGD_2(F).$$

This gives

$$D_2(F^n/G^m) = F^{n-1}G^{m-1}(nGD_2(F) - mFD_2(G))/G^{2m} = 0.$$

Therefore $F^n/G^m \in k(x_1)$ by Lemma (17.14.1).

(4) ⇒ (5). Since $F^n = \varnothing G^m$ if and only if $F^{-n} = \varnothing G^{-m}$, it is enough to show that we have $m \geq 0$, $n \geq 0$ or $m \leq 0$, $n \leq 0$. But this is immediate, since $mn \geq 0$ by Lemma (17.14.2).

Assume now that one of the conditions (i) and (ii) is satisfied. It is then enough to prove that $d < 0$, (5) $\Rightarrow$ (1) and (1) $\Rightarrow$ (7) $\Rightarrow$ (6) $\Rightarrow$ (2).

We first note that if condition (i) is satisfied then either $w_1 > 0$, $w_2 > 0$ or $w_1 < 0$, $w_2 < 0$. In either case, since $F \notin k$ or $G \notin k$, we get $m \neq 0$ or $n \neq 0$. Therefore we may assume that condition (ii) is satisfied. It is then clear that $d > 0$. **126**

(5) $\Rightarrow$ (1). Trivial, since $m \neq 0$ or $n \neq 0$.

(1) $\Rightarrow$ (7). There exist $p, q \in \mathbb{Z}^+$ and a w-homogeneous element E of A such that $F = \varnothing E^p$, $G = \varnothing E^q$. Let $e = d_w(E)$. Then $m = pe$, $n = qe$. It follows that $mn \geq 0$. Also, since condition (ii) is satisfied, we have $p > 0$ or $q > 0$, say $q > 0$. Let d' =g.c.d. (p, q) and write $p = p'd'$, $q = q'd'$, so that g.c.d. $(p', q') = 1$. Let $H = E^{d'}$. Then $F = \varnothing H^{p'}$, $G = \varnothing H^{q'}$. It is now enough to show that $p' = |m|/d$, $q' = |n|/d$. Since $q > 0$ and since

$$\text{g.c.d. } (p', q') = 1 = \text{g.c.d. } (|m|/d, |n|/d),$$

it is enough to prove that $p'|n| = q'|m|$. Now, since $F = \varnothing E^p$, $G = \varnothing E^q$, we have $pn = d_w(G^p) = d_w(E^{pq}) = d_w(F^q) = qm$. This shows that $p'|n| = q'|m|$.

(7) $\Rightarrow$ (6). Immediate, since $mn \geq 0$.

(6) $\Rightarrow$ (2). Immediate, since $m \neq 0$ or $n \neq 0$.

(17.5) COROLLARY. Let $f, g_1, \ldots, g_e$ be non-zero elements of A. Let $m = d_w(f)$, $n_i = d_w(g_i)$, $1 \leq i \leq e$, and let $d =$ g.c.d. $(m, n_1, \ldots, n_e)$. Assume that $m > 0$ and that f and g_i are w related for every i, $1 \leq i \leq e$. Then there exists a w-homogeneous element $H \in A$ of w-degree d such that $f_w^+ = \varnothing H^{m/d}$.

Proof. We prove the assertion by induction on e. Since f and g_1 are w-related, there exists, by Proposition (17.4), a w-homogeneous element H_1 of A such that $f_w^+ = \varnothing H_1^{m/d_1}$, where $d_1 =$ g.c.d. (m, n_1). It follows that $d_w(H_1) = d_1$, so that the assertion is proved for $e = 1$. Now, let $e > 1$ and let $d' =$ g.c.d. $(m, n_1, \ldots, n_{e-1})$, $d''=$ g.c.d. (m, n_e). By induction hypothesis and by the case $e = 1$, there exist w-homogeneous elements H_2, H_3 of A with $d_w(H_2) = d'$, $d_w(H_3) = d''$, such that $f_w^+ = \varnothing H_2^{m/d'} =$ **127**
$\varnothing H_3^{m/d''}$. This shows that H_2 and H_3 are w-related. Therefore by the

case $e = 1$, there exists a w-homogeneous element $H \in A$ of w-degree d such that $H_2 = \varnothing H^{d'/d}$. (Note that $d =$ g.c.d. (d', d'').) Thus we get $f_w^+ = H^{m/d}$. □

18 Structure of the w-Degree Form

We preserve the notation §15 and §16. In particular, we have char $k = 0$. Let $w = (w_1, w_2)$ be a pair of integers.

(18.1) DEFINITION. For non-zero elements f, g of A we define

$$\delta_w(f,g) = d_w(fg) - d_w(x_1x_2) - d_w(J(f,g)).$$

(18.2) LEMMA. *Let f, g be non-zero elements of A. Then we have:*

(i) $\delta_w(f,g) \geq 0$.

(ii) $J(f_w^+, g_w^+) = \begin{cases} J(f,g)_w^+, & \text{if } \delta_w(f,g) = 0 \\ 0, & \text{if } \delta_w(f,g) > 0. \end{cases}$

Proof.

(i) Clearly, we have

$$d_w(D_i(f)) \leq d_w(f) - w_i,$$
$$d_w(D_i(g)) \leq d_w(g) - w_i$$

for $i = 1, 2$. Therefore

$$d_w(D_1(f)D_2(g) - D_2(f)D_1(g)) \leq d_w(fg) - w_1 - w_2 = d_w(fg) - d_w(x_1x_2).$$

which proves (i).

128 (ii) Let $f' = f - f_w^+$, $g' = g - g_w^+$. Then $d_w(f') < d_w(f)$, $d_w(g') < d_w(g)$. An easy computation shows that

$$J(f,g) = J(f_w^+, g_w^+) + h.$$

where $h \in A$ with $d_w(h) < d_w(fg) - d_w(x_1x_2)$. Now, (ii) follows, since $J(f_w^+, g_w^+)$ is (either zero or) w-homogeneous of w-degree $d_w(fg) - d_w(x_1x_2)$.

□

(18.3) LEMMA. *Let f be a non-zero element of A such that $d_w(f) \neq 0$. Suppose there exists $g \in A$ such that f and $J(f, g)$ are w-related. Then there exists $h \in A$ such that f and $J(f, h)$ are w-related and $\delta_w(f, h) = 0$.*

Proof. If $\delta_w(f, g) = 0$ then we may take $h = g$. Assume therefore that $\delta_w(f, g) > 0$. It is then enough to prove the following assertion:

(18.3.1)

There exists $h \in A$ such that f and $J(f, h)$ are w-related and $\delta_w(f, h) < \delta_w(f, g)$.

For, then the lemma would follow by induction on $\delta_w(f, g)$. To prove (18.3.1), we note first that, since f and $J(f, g)$ are w-related, we have $j(f, g)$ are w-related, we have $J(f, g) \neq 0$ by definition. Therefore $g \neq 0$. Moreover, by Lemma (18.2) the assumption $\delta_w(f, g) > 0$ implies that $J(f_w^+, g_w^+) = 0$. Therefore by Proposition (17.4) f and g are w-related and there exists $c \in k^*$ such that $c(f_w^+)^{|n|} = (g_w^+)^{|m|}$, where $m = d_w(f)$, $n = d_w(g)$. (Note that by assumption we have $m \neq 0$.) Define $h = g^{|m|} - cf^{|n|}$, Then

$$J(f, h) = J(f, g^{|m|} - cf^{|n|}) = |m|g^{|m|}J(f, g).$$

It follows from Lemma (17.3) that f and $J(f, h)$ are w-related. Now, put $p = |m|$. Then we have

$$\begin{aligned} d_w(J(f, h)) &= d_w(g^{p-1}) + d_w(J(f, g)) \\ &= d_w(g^{p-1}) + d_w(fg) - d_w(x_1x_2) - \delta_w(f, g) \\ &= d_w(fg^p) - d_w(x_1x_2) - \delta_w(f, g) \\ &> d_w(fh) - d_w(x_1x_2) - \delta_w(f, g), \end{aligned}$$

since $(g_w^+)^p - c(f_w^+)^{|n|} = 0$. Thus we get **129**

$$\begin{aligned} \delta_w(f, g) &> d_w(fh) - d_w(x_1x_2) - d_w(J(f, h)) \\ &= \delta_w(f, h). \end{aligned}$$

This proves (18.3.1), and the lemma is proved. □

(18.4) COROLLARY. Let f be a non-zero element of A such that $d_w(f) \neq 0$. Suppose there exists $g \in A$ such that f and $J(f,g)$ are w-related. Then there exist w-homogeneous elements H, G of A, a positive integer p and a non-negative integer r such that $f_w^+ = \varnothing H^p$ and $J(H,G) = \varnothing H^r$,

Proof. By Lemma (18.3), replacing g by h we may assume that $\delta_w(f,g) = 0$. Then by Lemma (18.2) we have $J(f,g)_w^+ = J(f_w^+, g_w^+)$. Since f and $J(f,g)$ are w-related, there exist non-negative integers p, q and a w-homogeneous element H of A such that $f_w^+ = \varnothing H^p$, $J(f,g)_w^+ = \varnothing H^q$. Since $d_w(f) \neq 0$, we have $p > 0$. Let $G = g_w^+$. Then

$$\varnothing H^q = J(f,g)_w^+ = J(\varnothing H^p, G) = \varnothing p H^{p-1} J(H,G).$$

which shows that $q \geq p-1$. Let $r = q-(p-1)$. Then we have $J(H,G) = \varnothing H^r$. □

(18.5) LEMMA. *Assume the $w_1 w_2 > 0$. Let H, G be non-zero w-homogeneous elements of A such that $J(H,G) = \varnothing H^r$ for some positive*
130 *integer r. Then H^{r-1} divides G in A.*

Proof. We want to show that $G/H^{r-1} \in A$. Let $\overline{k}$ be the algebraic closure of k. Since $A = \overline{k}[x_1, x_2] \cap k(x_1, x_2)$, it is enough to prove that $G/H^{r-1} \in \overline{k}[x_1, x_2]$. We may therefore assume that $k = \overline{k}$. □

Since $w_1 w_2 > 0$, we have $w_1 > 0$, $w_2 > 0$ or $w_1 < 0$, $w_2 < 0$. Since an element F of A is (w_1, w_2)-homogeneous if and only if it is $(-w_1, -w_2)$-homogeneous, we may assume that $w_1 > 0$, $w_2 > 0$. Let $m = d_w(H)$, $n = d_w(G)$. Since $U(H,G) \neq 0$, we have $h \notin k$, $G \notin k$. Therefore $m > 0$ and $n > 0$. From $J(H,G) = \varnothing H^r$ we get $m + n - (w_1 + w_2) = mr$ (Lemma (18.2)). This gives

(18.5.1) $$n/m = r - 1 + (w_1 + w_2)/m > r - 1.$$

Next, by Lemma (17.14.3) we have

(18.5.2) $$\begin{aligned} nGD_1(H) - mHD_1(G) &= w_2 x_2 J(H,G) = \varnothing x_2 H^r, \\ nGD_2(H) - mHD_2(G) &= -w_1 x_1 J(H,g) = \varnothing x_1 H^r. \end{aligned}$$

Let u_1, u_2 be indeterminates. Identify A with the subring $k[u_1^{w_1}, u_2^{w_2}]$ of $k[u_1, u_2]$ by putting $x_i = u_i^{w_i}$, $i = 1, 2$. Then $A = k[u_1, u_2] \cap k(x_1, x_2)$. Therefore it is enough to prove the following assertion:

(18.5.3) $$H^{r-1} \text{ divides } G \text{ in } k[u_1, u_2].$$

Put $u = (u_1, u_2)$ and let $D_{u,i}$ be the k-derivation of $k(u_1, u_2)$ defined by $D_{u,i}(u_j) = \delta_{ij}$ (Kronecker delta), $i, j = 1, 2$. Then $D_{u,i}(F) = w_i u_i^{w_i - 1} D_i(F)$ for every $F \in A$. Therefore from (18.5.2) we get

(18.5.4) $$\begin{aligned} nGD_{u,1}(H) - mHD_{u,1}(G) &= \varnothing u_1^{w_1-1} u_2^{w_2} H^r, \\ nGD_{u,2}(H) - mHD_{u,2}(G) &= \varnothing u_1^{w_1} u_2^{w_2-1} H^r. \end{aligned}$$

Since H, G are w-homogeneous in A, they are $(1, 1)$-homogeneous in $k[u_1, u_2]$ of degrees m, n respectively. Now, (18.5.3) follows from (18.5.1) and (18.5.4) in view of the following **131**

(18.5.5) SUBLEMMA

Assume that k is algebraically closed. Let H, G be non-zero homogeneous elements of A of positive degrees m, n, respectively. Let r be a positive integer such that $r - 1 \leq n/m$ and H^r divides $nGD_i(H) - mHD_i(G)$ for $i = 1, 2$. Then H^{r-1} divides G.

Proof. Being homogeneous, H is a product of homogeneous linear polynomials in A. Therefore it is enough to prove that if F is a homogeneous linear polynomial in A and p is a positive integer such that F^p divides H then $F^{(r-1)p}$ divides G. So, let $F = a_1 x_1 + a_2 x_2$ with a_1, $a_2 \in k$, and suppose F^p divides H. We want to show that $F^{(r-1)p}$ divides G. We may assume that F^{p+1} does not divide H. Moreover, by interchanging x_1 and x_2, if necessary, we may assume that $a_1 \neq 0$. We may then assume that $a_1 = 1$. Write $H = F^p H'$, $G = F^q G'$ with $q \in \mathbb{Z}^+$ and H', $G' \in A$ such that $H' \not\equiv 0 \pmod F$, $G' \not\equiv 0 \pmod F$. We want to show that $q \geq (r-1)p$. We consider two cases:

CASE (1). $np = mq$. In this case we have $q/p = n/m \geq r - 1$, by assumption. Therefore $q \geq (r-1)p$.

CASE (2). $np \neq mq$. Since $D_1(F) = 1$, we have

$$D_1(H) = pF^{p-1}H' \pmod{F^p}$$
$$D_1(G) = qF^{q-1}G' \pmod{F^q}.$$

Therefore we get

$$nGD_1(H) - mHD_1(G) \equiv (np - mq)F^{p+q-1}G'H' \pmod{F^{p+q}}.$$

132 Since $np-mq \neq 0$ and $G'H' \not\equiv 0 \pmod{F}$ and since, by assumption, H^r divides $nGD_1(H) - mHD_1(G)$, we get $pr \leq p + q - 1$. This gives $(r - 1)p < q$. □

(18.6) COROLLARY. Assume that $w_1w_2 > 0$. Let H, G be non-zero w-homogeneous elements of A such that $J(H, G) = \varnothing H^r$ for some positive integer r. Then there exists a w-homogeneous element G' of A such that $J(H, G') = \varnothing H$.

Proof. By Lemma (18.5) we have $G = G'H^{r-1}$ for some $G' \in A$. Since H, G are w-homogeneous, so is G'. Now, $\varnothing H^r = J(H, G'H^{r-1}) = H^{r-1}J(H, G')$, so that $J(H, G') = \varnothing H$. □

(18.7) COROLLARY. Assume that $w_1 > 0$, $w_2 > 0$. Let f, g be elements of A such that f and $J(f, g)$ are w-related. Then there exist w-homogeneous elements H, G of A and a positive integer p such that $f_w^+ = \varnothing H^p$ and $J(H, G) = \varnothing H^s$ with $s = 0$ or 1.

Proof. Since f and $J(f, g)$ are w-related, we have $J(f, g) \neq 0$, which shows that $f \notin k$. Therefore, since $w_1 > 0$, $w_2 > 0$, we have $d_w(f) \neq 0$. Therefore by Corollary (18.4) there exist w-homogeneous elements H, G of A and a positive integer p such that $f_w^+ = \varnothing H^p$ and $J(H, G) = \varnothing H^r$ for some non-negative integer r. If $r = 0$, we are through. If $r > 0$ then by Corollary (18.6) there exists a w homogeneous element G' of A such that $J(H, G') = \varnothing H$. Replacing G by G', the assertion is proved. □

(18.8) LEMMA. *Assume that $w_1w_2 > 0$. Let H, G be w-homogeneous elements of A such that $J(H, G) = \varnothing$. Then:*

(i) *If* $|w_1| = |w_2|$ *then* $H = a_1x_1 + a_2x_2$ *with* $a_1, a_2 \in k$, $a_1 \neq 0$ *or* $a_2 \neq 0$.

(ii) *If* $|w_1| > |w_2|$ *then* $H = \varnothing z$, *where* $z = x_2$ *or* $z = x_1 + ax_2^{w_1/w_2}$ *with* $a \in k$. *Moreover, if* $a \neq 0$ *then* $w_1/w_2 \in \mathbb{N}$. 133

(iii) *If* $|w_1| < |w_2|$ *then* $H = \varnothing z$, *where* $z = x_1$ *or* $z = x_2 + ax_1^{w_2/w_1}$ *with* $a \in k$. *Moreover, if* $a \neq 0$ *then* $w_2/w_1 \in \mathbb{N}$.

Proof. By symmetry, it is enough to prove (i) and (ii). Since $w_1w_2 > 0$, we have either $w_1 > 0$, $w_2 > 0$ or $w_1 < 0$, $w_2 < 0$. We may assume, without loss of generality, that $w_1 > 0$, $w_2 > 0$. Then, since $H \notin k$, $G \notin k$, we have

$$\begin{aligned} d_w(H) &\geq \min(w_1, w_2), \\ d_w(G) &\geq \min(w_1, w_2). \end{aligned} \tag{18.8.1}$$

□

Since $J(H, G) = \varnothing$, it follows from Lemma (18.2) that $d_w(HG) = d_w(x_1x_2) = w_1 + w_2$.

(i) If $w_1 = w_2$ then $d_w(HG) = 2w_1$. Therefore from (18.8.1) we get $d_w(H) = w_1$. This means that H is a non-zero homogeneous polynomial in x_1, x_2 of degree one.

(ii) Since $w_1 = w_2$ we have $d_w(G) \geq w_2$ by (18.8.1). Therefore $d_w(H) \leq w_1$. This means that $\deg_{x_1} H \leq 1$. If $\deg_{x_1} H = 0$ then, since H is w-homogeneous, we have $H = \varnothing x_2^n$ for some $n \in \mathbb{N}$. This implies that x_2^{n-1} divides $J(H, G) = \varnothing$. Therefore $n = 1$ and $H = \varnothing x_2$. Now, suppose $\deg_{x_1} H = 1$. Then H is w-homogeneous of w-degree w_1. Therefore we have $H = bx_1 + cx_2^{w_1/w_2}$ with $b \in k^*$, $c \in k$ and $w_1/w_2 \in \mathbb{N}$ if $c \neq 0$. Let $z = x_1 + b^{-1}cx_2^{w_1/w_2}$. Then $H = \varnothing z$.

(18.9) LEMMA. *Assume that* $w_1 + w_2 \neq 0$. *Let* H *be a non-zero* w-*homogeneous element of* A *such that* $J(H, x_1x_2) = \varnothing H$. *Then* $H = \varnothing x_1^{i_1}x_2^{i_2}$ *for some non-negative integers* i_2, i_2 *with* $i_1 + i_2 > 0$.

Proof. Let $J(H, x_1 x_2) = cH$ with $c \in k^*$. Then we have 134

$$\begin{aligned} cH &= \det\begin{pmatrix} D_1(H) & D_2(H) \\ x_2 & x_1 \end{pmatrix} \\ &= x_1 D_1(H) - x_2 D_2(H). \end{aligned}$$

Let $d = d_w(H)$. We can write

$$H = \sum_{j_1 w_1 + j_2 w_2 = d} H_{j_1 j_2} x_1^{j_1} x_2^{j_2}$$

with $H_{h_1 j_2} \in k$. Then

$$x_1 D_1(H) - x_2 D_2(H) = \sum (j_1 - j_2) H_{j_1 j_2} x_1^{j_1} x_2^{j_2}$$

Therefore we have $j_1 - j_2 = c$ for all those pairs (j_1, j_2) for those pairs (j_1, j_2) for which $H_{j_1 j_2} \neq 0$. Since also $j_1 w_1 + j_2 w_2 = d$ and since

$$\det\begin{pmatrix} 1 & -1 \\ w_1 & w_2 \end{pmatrix} \neq 0$$

(because $w_1 + w_2 \neq 0$), there exists a unique pair (i_1, i_2) such that $H_{i_1 i_2} \neq 0$. This means that $H = \varnothing x_1^{i_1} x_2^{i_2}$. Since $J(H, x_1 x_2 \neq 0)$, we have $H \notin k$. Therefore $i_1 + i_2 > 0$. □

(18.10) LEMMA. *Assume that $w_1 = w_2 \neq 0$. Let H, G be w - homogeneous elements of A such that $H \neq 0$ and $J(H, G) = \varnothing H$. Then*

$$G = (a_1 x_1 + a_2 x_2)(b_1 x_1 + b_2 x_2)$$

and

$$H = \varnothing (a_1 x_1 + a_2 x_2)^{i_1} (b_1 x_1 + b_2 x_2)^{i_2},$$

135 *where i_1, i_2 are non-negative integers with $i_1 + i_2 > 0$ and a_1, a_2, b_1, b_2 are elements k such that $a_1 x_1 + a_2 x_2$ and $b_1 x_1 + b_2 x_2$ are linearly independent over k.*

Proof. We may assume, without loss of generality, that $w_1 = w_2 = 1$. Since $J(H, G) = \varnothing H$, by Lemma (18.2) we have $d(HG) = d(H) + d(x_1 x_2)$, where $d = d_w$. This gives $d(G) = 2$. Now, assume for the moment that k is algebraically closed. Then there exist $a_1, b_1, a_2, b_2 \in k$ such that $u_1 = a_1 x_1 + a_2 x_2$ and $u_2 = b_1 x_1 + b_2 x_2$ are linearly independent over k and $G = u_1^2$ or $G = u_1 u_2$. Now, $u = (u_1, u_2)$ is an automorphic pair for A. If $G = u_1^2$ then we have

$$\begin{aligned}\varnothing H = J(H, G) &= J_u(H, G) \\ &= \varnothing \det \begin{pmatrix} D_{u,1}(H) & D_{u,2}(H) \\ 2u_1 & 0 \end{pmatrix} \\ &= \varnothing u_1 D_{u,2}(H).\end{aligned}$$

This is not possible, since $\deg_{u_2} D_{u,2}(H) < \deg_{u_2} H$. Thus we have $G = u_1 u_2$. Now, since

$$\varnothing H = J(H, G) = \varnothing J_u(H, u_1 u_2)$$

and H is (1,1)-homogeneous with respect to u, it follows from Lemma (18.9) that we have $H = \varnothing u_1^{i_1} u_2^{i_2}$ with $i_1 + i_2 > 0$. Thus we have proved that we can choose elements $a_1, b_1, a_2, b_2 \in \overline{k}$ (= algebraic closure of k) which meet the requirements of our lemma. If this choice cannot be made in k then G would be irreducible in A and it would follow from the form of H that $i_1 = i_2$ and $H = \varnothing G^{i_1}$. But this is not possible, since $J(H, G) \neq 0$. □

(18.11) LEMMA. *Assume that $w_1 w_2 > 0$. Let H, G be w-homogeneous* 136
elements of A such that $H \neq 0$ and $J(H, G) = \varnothing H$. If $|w_1| > |w_2|$ (resp. $|w_1| < |w_2|$) then $G = \varnothing z x_2$ and $H = \varnothing z^{i_1} x_2^{i_2}$ (resp. $G = \varnothing x_1 z$ and $H = \varnothing x_1^{i_1} z^{i_2}$) for some non-negative integers i_1, i_2 with $i_1 + i_2 > 0$, where $z = x_1 + a x_2^{w_1/w_2}$ (resp. $z = x_2 + a x_1^{w_2/w_1}$) for some $a \in k$. If $a \neq 0$ then $w_1/w_2 \in \mathbb{N}$ (resp. $w_2/w_1 \in \mathbb{N}$).

Proof. The proof is analogous to that Lemma (18.10). First, we note that, by symmetry, it is enough to consider the case $|w_1| > |w_2|$. Since $w_1 w_2 > 0$, we may assume, without loss of generality, that $w_1 > 0$,

$w_2 > 0$. Then $w_1 > w_2$. Since $J(H, G) = \varnothing H$, we have $d_w(HG) = d_w(H) + d_w(x_1x_2)$ (Lemma (18.2)). Therefore $d_w(G) = w_1 + w_2$. Since $w_1 > w_2$, the only monomials in x_1, x_2 of w-degree $w_1 + w_2$ are x_1x_2 and (if $w_1/w_2 \in \mathbb{N}$ then) $x_2^{(w_1/w_2)+1}$. Therefore we have $G = bx_1x_2 + cx_2^{(w_1/w_2)+1}$ with $b, c \in k$ and $w_1/w_2 \in \mathbb{N}$ if $c \neq 0$. We claim that $b \neq 0$. For, if $b = 0$ then we get

$$\varnothing H = J(H, G) = \det\begin{pmatrix} D_1(H) & D_2(H) \\ 0 & \varnothing x_2^{w_1/w_2} \end{pmatrix} = \varnothing x_2^{w_1/w_2} D_1(H),$$

which is not possible, since $\deg_{x_1} D_1(H) < \deg_{x_1} H$. Thus $b \neq 0$. Let $z = x_1 + ax_2^{w_1/w_2}$, where $a = b^{-1}c$. Then $G = bzx_2$. Let $u_1 = z$, $u_2 = x_2$. Then $u = (u_1, u_2)$ is an automorphic pair for A and u_i is w-homogeneous of w-degree w_i, $i = 1, 2$. Therefore H is w-homogeneous with respect to u. Moreover, we have $\varnothing H = J(H, G) = \varnothing J_u(H, bu_1u_2) = \varnothing J_u(H, u_1u_2)$. Therefore it follows from Lemma (18.9) that we have $H = \varnothing u_1^{i_1} u_2^{i_2}$ with $i_1 + i_2 > 0$, and the lemma is proved. □

(18.12) LEMMA. *Assume that $w_1 > 0$, $w_2 > 0$ and that w_2 divides w_1 and $w_2 \neq w_1$. Let a be a non-zero element of k and let $u = (u_1, u_2)$ be*
137 *the automorphic pair defined by $u_1 = x_1 + ax_2^{w_1/w_2}$, $u_2 = x_2$. Let f be an element of A such that $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$, where i_1 is a positive integer and i_2 is a non-negative integer. Then* $\deg_u f < \deg f$.

(See (16.2) for the definition of $\deg_u f$ and $\deg f$.)

Proof. Let $n = d_w(f)$. Since u_i is w-homogeneous of w-degree w_i, $i = 1, 2$, f_w^+ is also the w-degree form of f with respect to $u = (u_1, u_2)$ (i.e., when we regard f as a polynomial in u_1, u_2 and give weight w_i to u_i, $i = 1, 2$). Since $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$, we can write f in the form

(18.12.1) $$f = \varnothing u_1^{i_1} u_2^{i_2} + \sum_{p_1w_1 + p_2w_2 < n} b_{p_1p_2} u_1^{p_1} u_2^{p_2}$$

and also in the form

(18.12.2) $$f = \varnothing (x_1 + ax_2^{w_1/w_2})^{i_1} x_2^{i_2} + \sum_{p_1w_1 + p_2w_2 < n} c_{p_1p_2} x_1^{p_1} x_2^{p_2}$$

with $b_{p_! p_2}, c_{p_1 p_2} \in k$. Let p_1, p_2 be non-negative integers such that $p_1 w_1 + p_2 w_2 < n$. Then, noting that by assumption we have $w_1/w_2 \geq 2$, we get

$$p_1 + p_2 \leq p_1(w_1/w_2) + p_2 < n/w_2 = i_1(w_1/w_2) + i_2.$$

Therefore we have

$$\begin{aligned} &\deg_u\left(\sum_{p_1 w_1 + p_2 w_2 < n} b_{p_1 p_2} u_1^{p_1} u_2^{p_2}\right) < i_1(w_1/w_2) + i_2, \\ &\deg\left(\sum_{p_1 w_1 + p_2 w_2 < n} c_{p_1 p_2} x_1^{p_1} x_2^{p_2}\right) < i_1(w_1/w_2) + i_2. \end{aligned} \tag{18.12.3}$$

□

Since $\deg_u\left(u_1^{i_1} u_2^{i_2}\right) = i_1 + i_2 < i_1(w_1/w_2) + i_2$ (because $w_1/w_2 \geq 2$ and $i_1 > 0$) and since

$$\deg\left(\left(x_1 + a x_2^{w_1/w_2}\right)^{i_1} x_2^{i_2}\right) = i_1(w_1/w_2) + i_2$$

(because $a \neq 0$), it follows from (18.12.1), (18.12.2) and (18.12.3) that **138**
$\deg_u f < i_1(w_1/w_2) + i_2 = \deg f$.

(18.13) THEOREM. *Assume that* $w_1 > 0$, $w_2 > 0$. *Let* f, g *be elements of* A *such that* $J(f,g) = \varnothing$. *Then* $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$, *where* i_1, i_2 *are non-negative integers,* $i_1 + i_2 > 0$, *and* $u = (u_1, u_2)$ *is an automorphic pair for* A *which has one of the following three forms:*

(i) *If* $w_1 = w_2$ *then* u_i *is homogeneous linear in* $x_1, x_2, i = 1, 2$.

(ii) *If* $w_1 > w_2$ *then* $u_1 = x_1 + a x_2^{w_1/w_2}, u_2 = x_2$, *with* $a \in k$ *and* $w_1/w_2 \in \mathbb{N}$ *if* $a \neq 0$.

(iii) *If* $w_1 < w_2$ *then* $u_1 = x_1$, $u_2 = x_2 + a x_1^{w_2/w_1}$ *with* $a \in k$ *and* $w_2/w_1 \in \mathbb{N}$ *if* $a \neq 0$.

Moreover, if u *is given by (ii) (resp. (iii)) and if* $i_1 \neq 0$ *(resp.* $i_2 \neq 0$*) and* $\neq 0$ *then* $\deg_u f < \deg f$.

Proof. By symmetry, it is enough to consider the cases $w_1 = w_2$ and $w_1 > w_2$. If $w_1 > w_2$ and if $i_1 \neq 0$ and $a \neq 0$ in (ii) then the last assertion of the theorem follows immediately from Lemma (18.12). □

Now, $J(f,g) = \varnothing$ implies that f and $J(f,g)$ are w-related. Therefore by Corollary (18.7) there exist w-homogeneous elements H, G of A and a positive integer r such that $f_w^+ = \varnothing H^r$ and $J(H,G) = \varnothing H^s$ with $s = 0$ or 1. Since $J(f,g) = \varnothing$, we have $f \neq 0$. Hence $H \neq 0$.

Suppose $s = 0$. Then $J(H,G) = \varnothing$. Therefore if $w_1 = w_2$ then by Lemma (18.8) H is homogeneous linear in x_1, x_2. Let $u_1 = H$ and let u_2 be any homogeneous linear polynomial in x_1, x_2 such that u_1, u_2
139 are linearly independent over k. Taking $i_1 = r$, $i_2 = 0$, we have $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$. Now, if $w_1 > w_2$ then by Lemma (18.8) $H = \varnothing z$ where $z = x_2$ or $z = x_1 + a x_2^{w_1/w_2}$ with $a \in k$ and $w_1/w_2 \in \mathbb{N}$ if $a \neq 0$. Let $u_1 = x_1 + a x_2^{w_1/w_2}$, $u_2 = x_2$ and let

$$(i_1, i_2) = \begin{cases} (r, 0), & \text{if } z = u_1, \\ (0, r), & \text{if } z = u_2. \end{cases}$$

Then we have $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$.

Now, suppose $s = 1$. Then $J(H,G) = \varnothing H$. If $w_1 = w_2$ then by Lemma (18.10) we have $H = \varnothing u_1^{j_1} u_2^{j_2}$, where u_1, u_2 are homogeneous linear and are linearly independent over k. Taking $i_1 = r j_1$, $i_2 = r j_2$, we get $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$. If $w_1 > w_2$ then by Lemma (18.11) we have $H = \varnothing u_1^{j_1} u_2^{j_2}$, where $u_2 = x_2, u_1 = x_1 + a x_2^{w_1/w_2}$ with $a \in k$ and $w_1/w_2 \in \mathbb{N}$ if $a \neq 0$, and j_1, j_2 are non-negative integers such that $j_1 + j_2 > 0$. Taking $i_1 = r j_1$, $i_2 = r j_2$, we get $f_w^+ = \varnothing u_1^{i_1} u_2^{i_2}$.

(18.14) DEFINITION. Let f be an element of A such that $f \notin k$ and let r be a positive integer. We say f has *r points at infinity with respect to the w-gradation* if f_w^+ is a product of r mutually coprime factors in $\overline{k}[x_1, x_2]$, where $\overline{k}$ is the algebraic closure of k, i.e., if $f_w^+ = h_1^{n_1} \dots h_r^{n_r}$, where $n_1, \dots, n_r$ are positive integers and $h_1, \dots, h_r$ are irreducible elements of $\overline{k}[x_1, x_2]$ with g.c.d. $(h_i, h_j) = 1$ for $i \neq j$. We say simply that f has *r points at infinity* if f has r points at infinity with respect to the usual (i.e., (1,1)-)gradation.

(18.15) COROLLARY. Let f, g be elements of A such that $J(f, g) = \varnothing$. If $w_1 > 0$, $w_2 > 0$ then f (also g) has at most two points at infinity with respect to the w-gradation. In particular, f (also g) has at most two points at infinity.

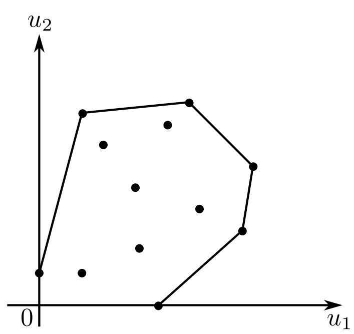
Proof. Immediate from Theorem (18.13). □

19 Various Equivalent Formulations of the Jacobian Problem

We preserve the notation of §15 and §16. In particular, we have char $k = 0$. **140**

(19.1) Newton Polygon of f

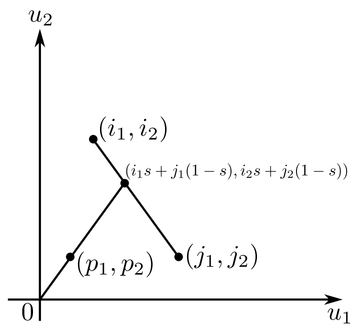
Let $u = (u_1, u_2)$ be an automorphic pair for A. Let $f \in A$. Writing $f = \sum a_{i_1 i_2} u_1^{i_1} u_2^{i_2}$ with $a_{i_1 i_2} \in k$, we put $S_u(f) = \left\{(i_1, i_2) \middle| a_{i_1 i_2} \neq 0\right\}$. We call $S_u(f)$ the *support of f with respect to u*. Let $N_u(f)$ be the smallest convex subset of the real plane $\mathbb{R}^2$ containing the set $S_u(f) \cup \{(0, 0)\}$. We call $N_u(f)$ the *Newton Polygon of f with respect to u*.



Newton Polygon of f
(Points of $S_u(f)$ are indicated by dots)

Note that $N_u(f)$ is the set of points $(p_1, p_2) \in \mathbb{R}^2$ for which there exist (i_1, i_2), (j_1, j_2) in $S_u(f)$ and $s, t \in \mathbb{R}$ with $0 \le s, t \le 1$ such that

$$(p_1, p_2) = (i_1 st + j_1(1-s)t, i_2 st + j_2(1-s)t).$$



141 We write $S(f)$ (resp. $N(f)$) for $S_x(f)$ (resp. $N_x(f)$) and call it simply the *support* (resp. *Newton Polygon*)of f.

(19.2) THEOREM. *Let f, g be elements of A such that $J(f, g) = \varnothing$. Assume that f has only one point at infinity and that* $\deg f \ge 2$. *Then there exists an automorphic pair $u = (u_1, u_2)$ for A such that* $\deg_u f < \deg f$.

Proof. Let $\overline{k}$ be the algebraic closure of k. Since f has only one point at infinity, there exists an irreducible homogeneous element F in A such that $f^+ = \varnothing F^n$ for some positive integer n and □

(19.2.1)

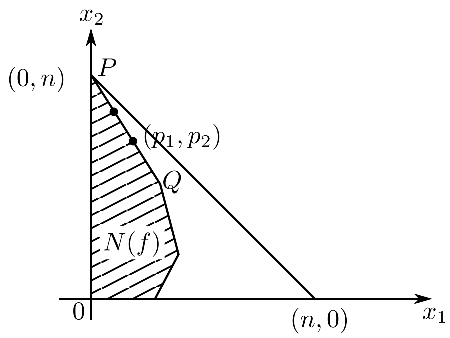
F is a power of a homogeneous linear polynomial in $\overline{k}[x_1, x_2]$.

Since char $k = 0$, the homogeneous polynomial F, being irreducible in $k[x_1, x_2]$, factors into distinct (i.e. mutually coprime) homogeneous linear polynomials in $\overline{k}[x_1, x_2]$. Therefore in view of (19.2.1) we necessarily have $\deg F = 1$, so that by a suitable homogeneous linear change of variables in A, we may assume that $F = x_2$ and $f^+ = \varnothing x + 2^n$
142 with $n = \deg f \ge 2$. Then $(0, n) \in S(f)$ and $i_1 + i_2 < n$ for all

$(i_1, i_2) \in S(f) - \{(0, n)\}$. It follows that $(0, n) \in N(f)$ and $i_1 + i_2 < n$ for all $(i_1, i_2) \in N(f) - \{(0, n)\}$. (this means that $N(f)$ lies below the line through $(0, n)$ with slope -1 and meets that line only in the point $(0, n)$. See the figure below.) Since $J(f, g) = \varnothing$ and $n \geq 2$, we have $f \notin k[x_2]$. Therefore there exists $(i_1, i_2) \in S(f)$ with $i_1 > 0$. Let

$$q = \inf \left\{ (n - i_2)/i_1 \middle| (i_1, i_2) \in S(f), i_1 > 0 \right\}$$

and let $(p_1, p_2) \in S(f)$ be such that $q = (n - p_2)/p_1$. (Note that (p_1, p_2) is one of



the points of $S(f) - \{(0, n)\}$ lying on the line PQ in the above figure and that $-q$ is the slope of the line PQ.) Let $w = (w_1, w_2)$, where $w_1 = n - p_2$, $w_2 = p_1$. Since $p_1 + p_2 < n$, we have $w_1 > w_2$. Therefore by Theorem (18.13) we have $f_w^+ = \varnothing u_1^{r_1} u_2^{r_2}$, where r_1, r_2 are non-negative integers with $r_1 + r_2 > 0$, $u_2 = x_2$ and $u_1 = x_1 + a x_2^{w_1/w_2}$ with $a \in k$ **143**
and $w_1/w_2 \in \mathbb{N}$ if $a \neq 0$. Let $(i_1, i_2) \in S(f)$. Then, since $i_2 \leq n$ and since $q = w_1/w_2$, we get $i_1 w_1 + i_2 w_2 \leq n w_2$. This, together with the fact that $p_1 w_1 + p_2 w_2 = n w_2$, shows that $d_w(f) = n w_2$ and that the two distinct points $(0, n)$ and (p_1, p_2) belong to $S(f_w^+)$. Therefore f_w^+ is not a monomial in x_1, x_2. This means that $r_1 \neq 0$ and $a \neq 0$. Therefore by Theorem (18.13) we have $\deg_u f < \deg f$, and the theorem is proved.

(19.3) REMARK. Let $u = (u_1, u_2)$ be an automorphic pair for A. Let σ be the k-algebra automorphism of A defined by $\sigma(x_i) = u_i$, $i = 1, 2$. Let us say that u is *obtained* from x by σ. We say σ is *homogeneous* *linear* if there exist $a_i, b_i \in k$ such that $u_i = a_i x_1 + b_i x_2$, $i = 1, 2$. We say σ is *very primitive* if there exist $a \in k$ and $n \in \mathbb{Z}$, $n \geq 2$, such that $u_1 = x_1 + a x_2^n$, $u_2 = x_2$ or $u_1 = x_1$, $u_2 = x_2 + a x_1^n$. We then note from the proof of Theorem (19.2) that there exists an automorphic pair u for A such that $\deg_u f < \deg f$ and u is obtained from x by a homogeneous linear automorphism followed by a very primitive automorphism.

(19.4) THEOREM. *The following four statements are equivalent:*

(i) *If* $f, g \in A$ *and* $J(f, g) = \varnothing$ *then* $k[f, g] = A$.

(ii) *If* $f, g \in A$ *and* $J(f, g) = \varnothing$ *then* f *has only one point at infinity.*

(iii) *If* $f, g \in A$ *and* $J(f, g) = \varnothing$ *then* $N(f)$ *is a triangle with vertices* $(0, n)$, $(0, 0)$, $(m, 0)$ *for some non-negative integers* m, n.

(iv) *If* $f, g \in A$ *and* $J(f, g) = \varnothing$ *then* $\deg f$ *divides* $\deg g$ *or* $\deg g$ *divides* $\deg f$.

Proof.

(I) $\Rightarrow$ (II). This follows from Corollary (11.24).

(II) $\Rightarrow$ (I). If $\deg f \geq 2$ then, since by (II) f has only one point at
144 infinity, it follows from Theorem (19.2) that there exists an automorphic pair $u = (u_1, u_2)$ for A such that $\deg_u f < \deg f$. Moreover, $J_u(f, g) = \varnothing$, so that f has only one point at infinity with respect to u. Therefore, by a repeated application of (II) and Theorem (19.2), we may assume that $\deg f = 1$. Now, by a further linear automorphism of A, we may assume that $f = x_1$. Then $\varnothing = J(f, g) = D_2(g)$, which shows that $g = \varnothing x_2 + p(x_1)$ with $p(x_1) \in k[x_1]$. Now, it is clear that $k[f, g] = A$.

(I) $\Rightarrow$ (III). Let $m = \deg_{x_1} f$, $n = \deg_{x_2} f$. Let T be the triangle with vertices $(0, n)$, $(0, 0)$, $(m, 0)$. We *claim* that $Nf = T$. This is clear if $m = 0$ or $n = 0$. Assume therefore that $m \geq 1$ and $n \geq 1$. Then by Corollary (11.20) f is almost monic in both x_1 and x_2. This means that $(m, 0) \in S(f)$ and $(0, n) \in S(f)$. Therefore $T \subset N(f)$. Now, let

$$f = a_0(x_1) x_2^n + a_1(x_1) x_2^{n-1} + \cdots + a_n(x_1)$$

with $a_i(x_1) \in k[x_1]$ for $0 \leq i \leq n$. Then by Corollary (11.20) we have $n \deg_{x_1} a_i(x_1) \leq im$ for every i, $0 \leq i \leq n$. It follows that if $(p, q) \in S(f)$

then $np \le (n-q)m$, so that $np+mq-mn \ge 0$. This shows that $(p,q) \in T$. Therefore $S(f) \subset T$ and hence $N(f) \subset T$. Thus $N(f) = T$.

(III) $\Rightarrow$ (II). We may assume that k is algebraically closed. Let $d = \deg(f)$. Suppose f has at least two points at infinity. Then by a linear homogeneous change of variables (i.e. by replacing x_1, x_2 by a suitable k-basis of $kx_1 \oplus kx_2$) we may assume that $f^+ = x_1^r G$, where r is a positive integer and G is a homogeneous element of A such that x_1 does not divide G in A and $\deg G > 0$. Since $J(f,g) = \varnothing$, $N(f)$ is a triangle with vertices $(0,n)$, (0.0), $(m,0)$, where m, n non-negative integers such that 145
$m+n > 0$. This shows that if $n \ge m$ then the monomial x_2^n appears in f^+ with a non-zero coefficient. But this is not possible, since $f^+ = x_1^r G$ with $r > 0$. Thus we have $n < m$. Therefore, since $N(f)$ is the triangle $(0,n)$, $(0,0)$, $(m,0)$, we get $f^+ = \varnothing x_1^m$. This is also not possible since x_1 does not divide G and $\deg G > 0$.

(I) $\Rightarrow$ (IV). This follows from Theorem (10.2).

(IV) $\Rightarrow$ (I). Assuming (IV), we prove (I) by induction on $\deg(fg)$. Since $J(f,g) = \varnothing$, we have $f \notin k$, $g \notin k$. Therefore $\deg f \ge 1$, $\deg g \ge 1$ and $\deg(fg) \ge 2$. If $\deg(fg) = 2$ then $\deg f = 1 = \deg g$ and the assertion is clear in this case. Now, let $m = deg f$, $n = \deg g$, and assume that $m+n \ge 3$. Without loss of generality, we may assume that $m \ge n$. Then by (IV) n divides m. Since $\deg(fg) \ge 3$ and $J(f,g) = \varnothing$, we have $J(f^+, g^+) = 0$ by Lemma (18.2). Therefore by Proposition (17.4) we have $f^+ = c(g^+)^{m/n}$ for some $c \in k^*$. Let $h = f - cg^{m/n}$. Then $\deg h < \deg f$. Moreover, clearly $J(h,g) = J(f,g) = \varnothing$. Therefore $k[h,g] = A$ by induction hypothesis. Since $k[f,g] = k[h,g]$, (I) is proved. □

(19.5) REMARK. In order to solve the Jacobian problem, we may assume that the field k is algebraically closed. For, each of statements (II), (III) and (IV) of Theorem (19.4) is unaltered if we replace k by its algebraic closure.

(19.6) REMARK. In the next section we give yet another equivalent formulation of the Jacobian problem in terms of a Newton-Puiseux expansion.

20 Jacobian Problem Via Newton-Puiseux Expansion

We preserve the notation of §15 and §16. In particular, we have char $k = 0$. We assume, in addition, that k is algebraically closed.

(20.1) Newton-Puiseux Expansion

146 Let f, g be elements of A. Assume that $n = \deg_{x_2} f > 0$ and that f is monic in x_2. By a construction analogous to the one used in §9, we can expand g in fractional powers of f^{-1} with coefficients in the algebraic closure of $k(x_1)$. Explicitly, let L be the algebraic closure of $k(x_1)$ and let τ be an indeterminate. Let $\theta : L[x_2] \to L((\tau))$ be the L-algebra monomorphism defined by $\theta(x_2) = \tau^{-1}$. It is then clear that we have $\operatorname{ord}_\tau \theta(F) = -\deg_{x_2} F$ for every $F \in L[x_2]$. In particular, we have $\operatorname{ord}_\tau \theta(f) = -n$. By Corollary (5.4) there exists $t \in L((\tau))$ such that $\operatorname{ord}_\tau(t) = 1$ and $\theta(f) = t^{-n}$. We then have $L((t)) = L((\tau))$ and $\operatorname{ord}_\tau F = \operatorname{ord}_\tau F$ for every $F \in L((t))$. Let $B = k[x_1]$. Then $B \subset L$ and we have $A = B[x_2]$. Let

$$B((t)) = \left\{ \sum a_i t^i \in L((t)) \middle| a_i \in B \ \ \forall i \right\}.$$

Then we have

(20.1.1) LEMMA

$$\theta(A) \subset B((t)).$$

Proof. We have only to show that $\theta(x_2) = \tau^{-1}$ belongs to $B((t))$. Since f is monic in x_2 with $\deg_{x_2} f = n$, we can write $f = x_2^n + f_1$ with $f_1 \in A$ and $\deg_{x_2} f_1 < n$. Therefore we get

$$t^{-n} = \theta(f) = \tau^{-n}(1 + \tau p)$$

with $p \in B[[\tau]]$. It follows that $t = \zeta\tau(1 + \tau q)$, where $\zeta \in \mu_n(k)$ (= n^{th} roots of unity in k) and

$$q \sum_{i=1}^{\infty} \binom{s}{i} \tau^{i-1} p^i \in B \in B[[\tau]].$$

where $s = -1/n$. Replacing t by $\zeta^{-1}t$, we may assume that $\zeta = 1$. Let $\tau = \sum_{i=1}^{\infty} a_i t^i$ with $a_i \in L$. Then we get **147**

$$\tau = \sum_{i=1}^{\infty} a_i \tau^i (1 + \tau_q)^i.$$

Now, we can write $(1+\tau q)^i = 1+\tau q_i$ with $q_i \in B[[\tau]]$. Let $q_i = \sum_{j=0}^{\infty} b_{ij}\tau^j$ with $b_{ij} \in B$. Then we get

$$\tau = \sum_{i=1}^{\infty} a_i \tau^i \left(1 + \sum_{j=0}^{\infty} b_{ij}\tau^{j+1}\right). \tag{20.1.1.1}$$

Comparing the coefficients of τ, we get $a_1 = 1 \in B$. Inductively, assume that $a_i \in B$ for $1 \le i \le d-1$ for some integer $d \ge 2$. Then, comparing the coefficients of τ^d in (20.1.1.1) we get $0 = a_d + c$, where

$$c = \sum_{i=1}^{d-1} a_i b_{i,d-1-i}.$$

By induction hypothesis $c \in B$. Therefore $a_d \in B$. This proves that we have

$$\tau = t(1 + \mathrm{tr.}) \tag{20.1.1.2}$$

with $r \in B[[t]]$. Therefore we get

$$\tau^{-1} = t^{-1}\left(1 + \sum_{i=1}^{\infty} (-1)^i t^i r^i\right).$$

which shows that $\tau^{-1} \in B((t))$. □

(20.1.2) COROLLARY

For any choice of $t \in L((\tau))$ such that $\theta(f) = t^{-n}$, we have $\tau = \zeta t(1 + \mathrm{tr.})$ for some $r \in B[[t]]$ and some $\zeta \in \mu_n(k)$.

Proof. Immediate from (20.1.1.2).

In view of Lemma (20.1.1), we can restrict θ to A to get a B-algebra monomorphism $\theta : A \to B((t))$ such that $\theta(f) = t^{-n}$ and

$$\text{(20.1.3)} \qquad \operatorname{ord}_t \theta(F) = -\deg_{x_2} F$$

148 for every $F \in A$. Let

$$\text{(20.1.4)} \qquad \theta(g) = \sum_j g_j t^j$$

with $g_j = g_j(x_1) \in B$. We call (20.1.4) a *Newcon-Puiseux expansion of g in fractional powers of f^{-1}*. Note that for fixed x_1, x_2, f, g, (20.1.4) depends on the choice of an element t such that $\theta(f) = t^{-n}$. If t_1, t_2 are two such choices then we have $t_1 = \zeta t_2$ for some $\zeta \in \mu_n(k)$. Thus there are atmost n distinct Newton-Puiseux expansions of g in fractional powers of f^{-1} and any two of them are conjugate to each other under a B-automorphism of $B((t))$ given by $t \mapsto \zeta t$ for some $\zeta \in \mu_n(k)$. In particular, the condition (JC) in Definition (20.2) below depends only on $x = (x_1, x_2)$, f, g and does not depend upon t. □

(20.2) DEFINITION. With the notation of (20.1), we say the pair (f, g) *satisfies condition* (*JC*) *(with respect to X)* if the following holds:

(*JC*) $g_j \in k$ *for every* $j \le n - 2$ *and* $\deg_{x_1} g_{n-1} = 1$.

(20.3) A DERIVATION OF $L((t))$.

Continuing with the notation of (20.1), put $u_1 = x_1$, $u_2 = f$ and $u = (u_1, u_2)$. Since $\deg_{x_2} f > 0$, u is a transcendence base of $K = k(x_1, x_2)$ over k. Therefore we have k-derivations $D_{u,1}, u_{u,2}$ of K as defined in (15.1). Let d_i denote the unique extension of $D_{u,i}$ to a k-derivation of $L(x_2)$, $i = 1, 2$. Let $\delta : L((t)) \to L((t))$ be the map defined by

$$\delta\left(\sum_j a_j t^i\right) = \sum_j d_1(a_j) t^j.$$

Then δ is clearly a $k((t))$-derivation of $L((t))$. We note that $\delta(B((t))) \subset B((t))$.

149 Moreover, denoting again by θ the extension of θ to an L - monomorphism $L(x_2) \to L((t))$ of fields, we have

(20.3.1) LEMMA

$$\delta\theta = \theta d_1.$$

Proof. Since $L(x_2)$ is separable algebraic over $k(u_1, u_2)$, it is enough to show that $\delta\theta | k(u_1, u_2) = \theta d_1 | k(u_1, u_2)$. Therefore it is enough to check that $\delta\theta(u_i) = \theta d_1(u_i)$, $i = 1, 2$. Now, $\delta\theta(u_1) = \delta\theta(x_1) = \delta(x_1) = \delta(u_1) = 1$ and $\theta d_1(u_1) = \theta(1) = 1$. Next, $\delta\theta(u_2) = \delta\theta(f) = \delta(t^{-n}) = 0$ and $\theta d_1(u_2) = \theta(0) = 0$. The lemma is proved. $\square$

(20.4) THEOREM. *Let f, g be elements of A. Assume that f is monic in x_2 and that $\deg_{x_2} f > 0$. Then the following two conditions are equivalent:*

(i) $J(f, g) = \varnothing$.

(ii) (f, g) *satisfies* (JC).

Proof. We use the notation of (20.1) and (20.3). Let $D_i = D_{x.i}, i = 1, 2$, where $x = (x_1, x_2)$. By the chain rule of derivation we have

$$\begin{aligned} J(f, g) &= J_u(f, g) J_x(u_1, u_2) \\ &= J_u(f, g) J_x(x_1, f) \\ &= \det\begin{pmatrix} 0 & 1 \\ d_1(g) & d_2(g) \end{pmatrix} \det\begin{pmatrix} 1 & 0 \\ D_1(f) & D_2(f) \end{pmatrix} \\ &= -d_1(g) D_2(f). \end{aligned}$$

This gives

$$\theta(d_1(g))\theta(D_2(f)) = -\theta(J(f, g)).$$

Therefore by Lemma (20.3.1) we get $\delta(\theta(g))\theta(D_2(f)) = -\theta(J(f, g))$. Using the expression (20.1.4) for $\theta(g)$ we get

$$\left(\sum_j d_1(g_j) t^i\right) \theta(D_2(f)) = -\theta(J(f, g)). \qquad (20.4.1)$$

Now, let $n = \deg_{x_2} f$. Then $n \geq 1$. Since f is monic in x_2, we **150**
get $D_2(f) = nx_2^{n-1} + f'$ with $f' \in A$ and $\deg_{x_2} f' < n-1$. Therefore $\theta(D_2(f)) = n\tau^{1-n} + \theta(f')$ with $\operatorname{ord}_\tau \theta(f') > 1-n$. It therefore follows from Corollary (20.1.2) that $\theta(D_2(f)) = \varnothing t^{1-n} + e$, where $e \in L((t))$ and $\operatorname{ord}_t e > 1-n$. This shows that we have $\theta(D_2(f))^{-1} = \varnothing t^{n-1} + h$ with $h \in L((t))$ and $\operatorname{ord}_t h > n-1$. Therefore from (20.4.1) we get

$$\sum_j d_1(g_j)t^j = -\theta(J(f,g))(\varnothing t^{n-1} + h). \tag{20.4.2}$$

Now, suppose $J(f,g) = \varnothing$. Then we have

$$\sum_j d_1(g_j)t^j = \varnothing(\varnothing t^{n-1} + h).$$

This shows that $d_1(g_j) = 0$ for $j \leq n-2$ and $d_1(g_{n-1}) = \varnothing$, which clearly implies that (f,g) satisfies condition (JC).

Conversely, suppose that (f,g) satisfies condition (JC). Then we have $d_1(g_j) = 0$ for $j \leq n-2$ and $d_1(g_{n-1}) = \varnothing$. Therefore it follows from (20.4.2) that we have

$$\varnothing t^{n-1} + \sum_{j \leq n} d_1(g_j)t^j = -\theta(J(f,g))(\varnothing t^{n-1} + h). \tag{20.4.3}$$

This shows that $\operatorname{ord}_t \theta(J(f,g)) = 0$. Therefore by (20.1.3) we get $\deg_{x_2} J(f,g) = 0$, which means that $J(f,g) \in L$. Put $\lambda = J(f,g)$. Then $\theta(\lambda) = \lambda$. Therefore comparing the coefficients of t^{n-1} in (20.4.3) we get $\varnothing = -\lambda\varnothing$, which shows that $\lambda = \varnothing$, and the theorem is proved. $\square$

(20.5) NOTATION. Let f, g be elements of A. Assume that $n = \deg_{x_2} f > 0$ and that f is monic in x_2. Then with the notation of (20.1) we have a commutative diagram

$$\begin{tikzcd} L[x_2] \arrow[r, "\theta"] & L((t)) \\ A \arrow[u, hook] \arrow[r, "\theta"] & B((t)) \arrow[u, hook] \end{tikzcd}$$

151 where θ is a B-algebra monomorphism such that $\theta(f) = t^{-n}$ and

(20.5.1) $$\operatorname{ord}_t \theta(F) = -\deg_{x_2} F$$

for every $F \in A$. Let
$$\theta(g) = \sum_j g_j t^i$$
with $g_j = g_j(x_1) \in B$. *Assume that the pair* (f, g) *satisfies condition (JC), i.e. assume that we have*

(20.5.2) $$\begin{aligned} &g_j \in k \text{ for every } j \leq n-2, \\ &\deg_{x_1} g_{n-1} = 1. \end{aligned}$$

Then by Theorem (20.4) we have $J(f, g) = \varnothing$. Let $\tilde{\Phi} = \tilde{\Phi}(X, Y) \in L((X))[Y]$ be the minimal monic polynomial of $\theta(g)$ over $L((t^n))$. (See Definition (5.8).) Recall that $\tilde{\Phi}$ is the unique irreducible element of $L((X))[Y]$, monic in Y, such that $\tilde{\Phi}(t^n, \theta(g)) = 0$. Put $\Phi = \Phi(X, Y) = \tilde{\Phi}(X^{-1}, Y)$.

(20.5.3) LEMMA

(i) Φ is monic in Y and $\deg_Y \Phi = n$.

(ii) $\Phi \in B[X, Y]$.

(iii) $\Phi(f, g) = 0$.

(iv) $L[X, Y]/(\Phi)$ is isomorphic (as an L-algebra) to $L[f, g]$.

Proof.

(i) By definition of Φ, Φ is monic in Y. By (20.5.2) $n-1 \in \operatorname{Supp}_t \theta(g)$. Therefore
$$\text{g.c.d. } (\{n\} \cup \operatorname{Supp}_t \theta(g)) = 1.$$
Now it follows from Lemma (5.10) that $\deg_Y \tilde{\Phi} = n$. This proves **152**
(i).

(ii) Let $\Psi = \Psi(X, Y) \in B[X, Y]$ be the x_2-resultant of $(f - X, Y - g)$. Since f is monic in x_2, Ψ is monic in Y. Moreover, since $\deg_{x_2} f = n$, we have $\deg_Y \Psi = n$. Put $\tilde{\Psi}(X, Y) = \Psi(X^{-1}, Y)$.

We have $\Psi(f,g) = 0$. Therefore $0 = \theta(\Psi(f,g)) = \Psi(t^{-n}, \theta(g)) = \tilde{\Psi}(t^n, \theta(g))$. It now follows from (i) that $\tilde{\Phi} = \tilde{\Psi}$. Therefore $\Phi = \Psi \in B[X,Y]$.

(iii) Since $\Phi = \Psi$ as proved above, we have $\Phi(f,g) = \Psi(f,g) = 0$.

(iv) Let $\alpha : L[X,Y] \to L[f,g]$ be the L-algebra epimorphism defined by $\alpha(X) = f$, $\alpha(Y) = g$. then (ii) and (iii) $\Phi \in \ker\alpha$. Since Φ is irreducible in $L((X^{-1}))[Y] \supset L[X,Y]$ and is monic in Y, Φ is irreducible in $L[X,Y]$. Therefore $\ker\alpha = (\Phi)$, and (iv) is proved.

$\square$

(20.5.4) A SPECIALIZATION.

Since $\deg_{x_1} g_{n-1} = 1$ by (20.5.2), there exists $c \in k$ such that $g_{n-1}(x_1) \neq g_{n-1}(c) \neq 0$. We choose such a $c \in k$ and keep it fixed in the sequel. For an element F of $A = B[x_2]$ (resp. $B((t))$, $B[X,Y]$, $B[X^{-1},Y], \ldots$) we shall denote by $\overline{F}$ the element of $k[x_2]$ (resp. $k((t))$, $k[X,Y]$, $k[X^{-1},Y], \ldots$) obtained from F by putting $x_1 = c$. Let $\varphi = \overline{\Phi}$, $\tilde{\varphi} = \overline{\tilde{\Phi}}$.

(20.5.5) LEMMA

(i) $\varphi \in k[X,Y]$, φ is monic in Y and $\deg_Y \varphi = n$.

(ii) $\tilde{\varphi} \in k[X^{-1},Y]$ $\tilde{\varphi}$ is monic in Y and $\deg_Y \tilde{\varphi} = n$.

(iii) $\tilde{\varphi}$ is the minimal monic polynomial of $\overline{\theta(g)} = \sum_j \overline{g}_j t^j$ over $k((t^n))$.

(iv) $\operatorname{ord}_t(\theta(g) - \overline{\theta(g)}) = n - 1$.

Proof.

(i) is immediate from Lemma (20.5.3).

(ii) This follows from (i), since $\tilde{\varphi}(X,Y) = \varphi(X^{-1},Y)$.

153 (iii) Since $\tilde{\Phi}(t^n, \theta(g)) = 0$, we have $\tilde{\varphi}(t^n, \overline{\theta(g)}) = 0$. Since $\overline{g}_{n-1} \neq 0$, we have $n-1 \in \operatorname{Supp}_t \overline{\theta(g)}$. Therefore the minimal monic polynomial of $\overline{\theta(g)}$ over $k((t^n))$ has Y-degree n (Lemma (5.10)). Therefore by (ii) $\tilde{\varphi}$ is the minimal monic polynomial of $\overline{\theta(g)}$ over $k((t^n))$.

(iv) Since $g_j \in k$ for $j \leq n-2$, we have $g_j = \overline{g}_j$ for $j \leq n-2$. Moreover, we have $g_{n-1} \neq \overline{g}_{n-1}$. Therefore the assertion follows.

□

(20.5.6) Characteristic Sequences of (f, g).

(See § 6.) We define $h(f, g) = h(\tilde{\Phi})$ and we define the *characteristic sequences* of the pair (f, g) by

$$\begin{aligned} m_i(f, g) &= m_i(-n, \tilde{\Phi}), \\ q_i(f, g) &= q_i(-n, \tilde{\Phi}), \\ s_i(f, g) &= s_i(-n, \tilde{\Phi}), \\ r_i(f, g) &= r_i(-n, \tilde{\Phi}), \\ d_{i+1}(f, g) &= d_{i+1}(\tilde{\Phi}), \end{aligned}$$

for $0 \leq i \leq h(f, g) + 1$. (Note that these sequences depend not only on f, g, but also on $x = (x_1, x_2)$. However, the omission of x in the notation $m_i(f, g)$ etc. will cause no confusion.)

(20.5.7) LEMMA

We have $h(\tilde{\varphi}) = h(f, g)$ and

$$\begin{aligned} m_i(-n, \tilde{\varphi}) &= m_i(f, g), \\ q_i(-n, \tilde{\varphi}) &= q_i(f, g), \\ s_i(-n, \tilde{\varphi}) &= s_i(f, g), \\ r_i(-n, \tilde{\varphi}) &= r_i(f, g), \\ d_{i+1}(\tilde{\varphi}) &= d_{i+1}(f, g) \end{aligned}$$

for $0 \leq i \leq h(\tilde{\varphi}) + 1$.

Proof. Immediate, since g.c.d. $(n, n-1) = 1$, $n-1 \in \mathrm{Supp}_t\,\overline{\theta(g)}$, $n-1 \in$ 154
$\mathrm{Supp}_t\,\theta(g)$ and $\mathrm{ord}_t(\theta(g) - \overline{\theta(g)}) = n-1$ by Lemma (20.5.5). ☐

In the remainder of subsection (20.5) we fix the following notation:

$$\begin{aligned} h &= h(f,g), \\ m_i &= m_i(f,g), \\ q_i &= q_i(f,g), \\ s_i &= s_i(f,g), \\ r_i &= r_i(f,g), \\ d_{i+1} &= d_{i+1}(f,g) \end{aligned}$$

for $0 \leq i \leq h+1$. Also, for $1 \leq i \leq h+1$, let

$$\begin{aligned} \tilde{\psi}_i &= \begin{cases} Y, & \text{if } i = 1, \\ App_Y^{d_i}(\tilde{\psi}, & \text{if } i \geq 2 \end{cases} \\ \psi_i &= \begin{cases} Y, & \text{if } i \geq 1, \\ App_Y^{d_i}(\varphi), & \text{if } i \geq 2, \end{cases} \\ \tilde{\psi}'{}_i &= \frac{\partial \tilde{\psi} i}{\partial Y}, \\ \psi'_i &= \frac{\partial \psi_i}{\partial Y}. \end{aligned}$$

(See § 4).

(20.5.8) LEMMA

We have:

(i) $h \geq 1$.

(ii) $m_1 = -\deg_{x_2} g \leq 0$.

(iii) $m_i < n-1$ for $1 \leq i \leq h-1$ and $m_h \leq n-1$.

155 *Proof.* (i) This is clear, since $\theta(g) \neq 0$.

(ii) Follows from (20.5.1) and the fact that $g \neq 0$.

(iii) This is also clear, since $n-1 \in \mathrm{Supp}_t\, \theta(g)$ and g.c.d. $(n, n-1) = 1$.

□

(20.5.9) LEMMA

For $1 \leq i \leq h+1$, we have

(i) $\tilde{\psi}_i(X, Y) = \psi_i(X^{-1}, Y)$,

(ii) $\tilde{\psi}'_i(X, Y) = \psi'_i(X^{-1}, Y)$.

Proof. (i) Follows from Proposition (4.7).

(ii) Follows from (i)

□

(20.5.10) LEMMA

For $F(X, Y) \in k[X, Y]$, we have $\deg_{x_2} F(f, g) = -\,\mathrm{ord}_t\, F(t^{-n}, \theta(g))$.

Proof. This follows from 20.5.1, since $\theta(F(f, g)) = F(t^{-n}, \theta(g))$. □

(20.5.11) LEMMA

For $1 \leq e \leq h$, we have $\deg_{x_2} \psi_e(f, g) = -r_e$.

Proof. We have $\psi_1(X, Y) = Y$. Therefore by Lemma (20.5.10) $\deg_{x_2} \psi_1 (f, g) = -\,\mathrm{ord}_t\, \theta(g) = -m_1 = -r_1$. This proves the assertion for $e = 1$. Assume now that $e \geq 2$. Since $m_e \leq n-1$ by Lemma (20.5.8), it follows from Lemma (20.5.5) (iv) that we have

$$\theta(g) = \sum_{j<m_e} \overline{g}_j t^j + g_{m_e} t^{m_e} + \sum_{j>m_e} g_j t^j.$$

Therefore, since $g_{m_e} \neq 0$, it follows from Corollary (7.20) that $\mathrm{ord}_t\, \tilde{\psi}(t^n, \theta(g)) = r_e$. Therefore by Lemma (20.5.9) we have $\mathrm{ord}_t\, \psi_e(t^{-n}, \theta(g)) = r_e$. Now, the lemme follows from Lemma (20.5.10). □

(20.5.12) LEMMA

For $1 \leq e \leq h$, we have $\deg_{x_2} \psi'_e(f,g) = m_e - r_e$.

Proof. Since $\mathrm{ord}_t(\theta(g) - \overline{\theta(g)}) = n - 1 \geq m_e$ and $m_e \in \mathrm{Supp}_t\, \theta(g)$, it follows from Proposition (13.7) that $\mathrm{ord}_t\, \tilde{\psi}'{}_e(t^n, \theta(g)) = r_e - m_e$. Therefore by Lemmas (20.5.9) and (20.5.10) we get $\deg_{x_2} \psi'_e(f,g) = -\mathrm{ord}_t\, \psi'_e(t^{-n}, \theta(g)) = m_e - r_e$. □

156 **(20.6) DEFINITION.** An element f of A is said to be x_2 *regular* if $f \neq 0$ and $\deg f = \deg_{x_2} f$.

Note that f is x_2-regular if and only of x_1 does not divide f^+ in A.

In Lemmas (20.7) - (20.9) below, we let the notation and assumptions be those (20.5). We assume, moreover, that f ix x_2-regular.

(20.7) LEMMA. *Let e be an integer,* $2 \leq e \leq h$. *Assume that* $\psi_i(f,g)$ *is related to* f *for every* i, $1 \leq i \leq e-1$. *Let* $F = F(X,Y)$ *be a non-zero element of* $k[X,Y]$ *with* $\deg_Y F < n/d_e$. *Then* $F(F,g)$ *is related to* f.

Proof. Let $R = k[X]$. Let $p = e - 1$ and let $G = (G_1, \ldots, G_p)$, where $G_i = \psi_i$ for $1 \leq i \leq p$. Then G satisfies conditions (i)-(iii) of (2.2) and, with the notation of (2.2), we have $n_i(G) = d_i/d_{i+1}$ for $1 \leq i \leq p-1$. Let

$$A(G) = \left\{ a = (a_1, \ldots, a_p) \in (\mathbb{Z}^+)^p \,\middle|\, a_i < d_i/d_{i+1} \text{ for } 1 \leq i \leq p-1 \right\}.$$

□

Then by Corollary (2.14) we have the G-adic expansion

$$F = \sum_{a \in A(G)} F_a G^a \tag{20.7.1}$$

of F with $f_a = F_a(X) \in R$ for every $a \in A(G)$. By Corollary (2.9) we have

$$\sum_{i=1}^{p} a_i \deg_Y G_i = \deg_Y G^a \leq \deg_Y F < n/d_e = n/d_{p+1}$$

for every $a \in \mathrm{Supp}_G\, F$. In particular, we have

$$a_p n/d_p = a_p \deg_Y G_p < n/d_{p+1}.$$

This gives

$$a_p < d_p/d_{p+1} \tag{20.7.2}$$

for every $a \in \mathrm{Supp}_G(F)$. Putting $X = f$, $Y = g$ in (20.7.1), we get 157

$$F(f,g) = \sum_{a\in S} F_a(f)G(f,g)^a, \tag{20.7.3}$$

where $S = \mathrm{Supp}_G(F)$. Since $G_a(f) \in k[f]$, we can rewrite (20.7.3) in the form

$$F(f,g) = \sum_{b\in B(H)} \lambda_b H^b$$

with $\lambda_b \in k$ for every b, where $h = (h_0,\dots,h_p)$ with $h_0 = f$. $H_i = G_i(f,g)$ for $1 \le i \le p$, and where

$$B(H) = \left\{ b = (b_0,\dots,b_p) \in (\mathbb{Z}^+)^{p+1} \middle| b_i < d_i/d_{i+1} \text{ for } 1 \le i \le p \right\}.$$

Note that the condition $b_p < d_p/d_{p+1}$ for $b \in B(H)$ is justified in view of (20.7.2). Since $\deg_{x_2} H_i = -r_i$ for $1 \le i \le p$ (Lemma (20.5.11)) and $\deg_{x_2} H_0 = \deg_{x_2} f = n = -r_0$, we have, for every $b \in (B(H))$.

$$\deg_{x_2} H^b = \sum_{i=0}^{p} b_i(-r_i).$$

which is clearly a strict linear combination of $(-r_0,\dots,-r_p)$. (See § 1.) Therefore if $b, b' \in B(H)$, $b \ne b'$, then $\deg_{x_2} H^b \ne \deg_{x_2} H^{b'}$. It follows that there exists a unique $b \in B(H)$ such that $\lambda_b \ne 0$ and

$$\deg_{x_2} F(f,g) = \deg_{x_2}(\lambda_b H^b) > \deg_{x_2}(\lambda_{b'} H^{b'}) \tag{20.7.4}$$

for every $b' \in B(H)$, $b' \ne b$. Now, by assumption, H_i is related to f for every i, $0 \le i \le p$. In particular, since f is x_2-regular, so is H_i for every i, $0 \le i \le p$. Therefore we have $\deg_{x_2} H^{b'} = \deg H^{b'}$ for every $b' \in B(H)$, and it follows from (20.7.4) that we have

$$F(f,g)^+ = (\lambda_b H^b).$$

Since each H_i is related to f, so is $\lambda_b H^b$ by Lemma (17.3). Thus 158
$F(f,g)$ is related to f, and the lemma is proved.

(20.8) LEMMA. *Let e be an integer, $2 \le e \le h$. Assume that $\psi_i(f,g)$ is related to f for every i, $1 \le i \le e-1$. Then f has only one point at infinity or $\psi_e(f,g)$ is x_2-regular.*

Proof. By the chain rule for differentiation we have

$$J(f,\psi_e(f,g)) = \psi'_e(f,g)J(f,g) = \varnothing\psi'_e(f,g). \tag{20.8.1}$$

Now, if $J(f^+,\psi_e(f,g)^+) = 0$ then by Proposition (17.4) f and $\psi_e(f,g)$ are related. Therefore in this case, since f is x_2-regular, so is $\psi_e(f,g)$. Thus we may now assume that $J(f^+,\psi_e(f,g)^+) \ne 0$. Then by (20.8.1) and Lemma (18.2) we have

$$J(f^+,\psi_e(f,g)^+) = \varnothing\psi'_e(f,g)^+. \tag{20.8.2}$$

□

Since $\deg_Y \psi'_e = \deg_Y \psi_e - 1 < n/d_e$, it follows from Lemma (20.7) that $\psi'_e(f,g)$ is related to f. Therefore there exist non-negative integers p, q and a homogeneous element H of A such that $f^+ = \varnothing H^p$, $\psi'_e(f,g)^+ = \varnothing H^q$. From (20.8.2) we get $J(H^q,G) = \varnothing H^q$, where $G = \psi_e(f,g)^+$. This shows that $p-1 \le q$ and $J(H,G) = \varnothing H^r$, where $r = q-p+1$. If $r=0$ then $J(H,G) = \varnothing$ and it follows from Lemma (18.8) (i) that H is linear in x_1, x_2, which shows that f has only one point at infinity. We may therefore assume that $r > 0$. Then by Lemma (18.5) H^{r-1} divides G. Let $G = EH^{r-1}$ with $E \in A$. Then from $J(H,G) = \varnothing H^r$ we get $J(H,E) = \varnothing H$. Therefore by Lemma (18.10) we have $E = (a_1x_1+a_2x_2)(b_1x_1+b_2x_2)$ and $H = \varnothing(a_1x_1+a_2x_2)^{i_1}(b_1x_1+b_2x_2)^{i_2}$,
159 where i_1, i_2 are non-negative integers, $i_1 + i_2 > 0$, and a_1, a_2, b_1, b_2 are elements of k such that $a_1x_1 + a_2x_2$ and $b_1x_1 + b_2x_2$ are linearly independent over k. If $i_1 = 0$ or $i_2 = 0$ then H (and therefore f) has only one point at infinity. Assume therefore that $i_1 > 0$, $i_2 > 0$. Then, since f (and therefore H) is x_2-regular, we have $a_2 \ne 0$, $b_2 \ne 0$. This implies that E is x_2-regular. Therefore $G = EH^{r-1}$ is x_2-regular. This means that $\psi_e(f,g)$ is x_2-regular.

(20.9) LEMMA. *Let e be an integer, $2 \le e \le h$. Assume that $\psi_i(f,g)$ is related to f for every i, $1 \le i \le e-1$. Assume also that $m_e \ne n-2$. Then f has only one point at infinity or $\psi_e(f,g)$ is related to f.*

Proof. If f has only one point at infinity, there is nothing to prove. Therefore by Lemma (20.8) we may assume that $\psi_e(f,g)$ is x_2-regular. By Proposition (17.4) we have to show that $J(f^+,\psi_e(f,g)^+) = 0$. Suppose $J(f^+,\psi_e(f,g)^+) \neq 0$. Then, since by (20.8.1) we have

$$J(f,\psi_e(f,g)) = \varnothing \psi'_e(f,g)$$

we get

$$\deg f + \deg \psi_e(f,g) - 2 = \deg \psi'_e(f,g) \tag{20.9.1}$$

by Lemma (18.2). Since $\deg_Y \psi'_e < n/d_e, \psi'_e$, $\psi'_e(f,g)$ is related to f by Lemma (20.7). Therefore, since f is x_2-regular, so is $\psi_e(f,g)$. Also, by assumption, $\psi_e(f,g)$ is x_2-regular. Therefore we have

$$\deg \psi_e(f,g) = \deg_{x_2} \psi_e(f,g) = -r_e$$

by Lemma (20.5.11) and

$$\deg \psi'_e(f,g) = \deg_{x_2} \psi'_e(f,g) = m_e - r_e$$

by Lemma (20.5.12). Therefore, since $\deg f = n$, (20.9.1) gives $n - r_e -$ **160**
$2 = m_e - r_e$, so that $m_e = n - 2$, which is a contradiction. Therefore $J(f^+,\psi_e(f,g)^+) = 0$, and the lemma is proved. □

(20.10) THEOREM (cf. Theorem (19.4))**.** *The following three statements are equivalent:*

(I) *If* $f, g \in A$ *and* $J(f,g) = \varnothing$ *then* $k[f,g] = A$.

(V) *Let* $f, g \in A$. *Assume that* $\deg_{x_2} f > 0$ *and that* f *is* x_2*-regular and is monic in* x_2. *If the pair* (f,g) *satisfies condition (JC) then we have* $\deg_{x_2} f = 1$ *or* $m_e(f,g) < \deg_{x_2} f - 2$ *for every* e, $1 \leq e \leq h(f,g)$.

(VI) *Let* $f, g \in A$ *be as in statement (V). If the pair* (f,g) *satisfies (JC) then we have* $\deg_{x_2} f = 1$ *or* $m_e(f,g) \neq \deg_{x_2} f - 2$ *for every* e, $1 \leq e \leq h(f,g)$.

Proof. Consider the statement

(II) If $f, g \in A$ and $J(f,g) = \varnothing$ then f has only one point at infinity.

By theorem (19.4) it is enough the implications

$$\text{(I)} \Rightarrow \text{(V)} \Rightarrow \text{(VI)} \Rightarrow \text{(II)}.$$

(I) $\Rightarrow$ (V). Let f, g satisfy the hypothesis of (V). Then by Theorem (20.4) we have $J(f,g) = \varnothing$. Therefore by (I) we have $k[f,g] = A$. We now use the notation of (20.5). From the equality $k[f,g] = A$ we get $L[f,g] = L[x_2]$. This means that $L[X,Y]/(\Phi)$ is isomorphic to $L[x_2]$ (Lemma (20.5.3)) (iv)). Now, by Lemma (20.5.8) we have $h \geq 1$. If $h \geq 2$ then it follows from Corollary (13.5) (v) that $m_e(f,g) = m_e(-n, \tilde{\Phi}) < n-2$ for every e, $1 \leq e \leq h$, where $n = \deg_{x_2} f$. Suppose now that $h = 1$.
161 Let $m_1 = m_1(f,g)$. Then $h = 1$ implies that g.c.d. $(n, m_1) = 1$. Suppose $m_1 \geq n-2$. Then $n - m_1 \leq 2$. Since $m_1 \leq 0$ by Lemma (20.5.8), we get $n \leq 2$. If $n = 2$ then we must have $m_1 = 0$. This is not possible, since g.c.d. $(n, m_1) = 1$. Therefore $n = 1$, and (V) is proved.

(V) $\Rightarrow$ (VI). Trivial.

(VI) $\Rightarrow$ (II). Let f, g be elements of A such that $J(f,g) = \varnothing$. We have to show that f has only one point at infinity. To do this we may replace x_1, x_2 by any basis of the k-vector space $kx_1 \oplus kx_2$. We may therefore assume, without loss of generality, that x_1 does not divide f^+, i.e., f is x_2-regular. Then, in particular, $\deg_{x_2} f > 0$. Moreover, replacing f by $\varnothing f$ for suitable $\varnothing$, we may assume that f is monic in x_2. By Theorem (20.4), since $J(f,g) = \varnothing$, the pair (f,g) satisfies condition (JC). Let $n = \deg_{x_2} f = \deg f$. If $n = 1$ then, clearly, f has only one point at infinity. Assume therefore that $n > 1$. Then by (VI) we have $m_e(f,g) \neq n-2$ for every e, $1 \leq e \leq h$, where $h = h(f,g)$. Since $\deg f > 1$ and $J(f,g) = \varnothing$, it follows from Lemma (18.2) that $J(f^+, g^+) = 0$. Let us now use the notation of (20.5). Since $J(f^+, g^+) = 0$, it follows from Proposition (17.4) that f and $g = \psi_1(f,g)$ are related. Now, since $m_e(f,g) \neq n-2$ for every e, $1 \leq e \leq h$, it follows from Lemma (20.9) by induction on e that f has only one point at infinity or f is related to $\psi_e(f,g)$ for every, e, $1 \leq e \leq h$. If f has only one point a infinity then we have nothing more to prove. We may therefore assume that f is related to $\psi_e(f,g)$ for every e, $1 \leq e \leq h$. In particular, since f is x_2-regular, so is $\psi_e(f,g)$ for every e. Therefore, for $1 \leq e \leq h$, we have $\deg \psi_e(f,g) = \deg_{x_2} \psi_e(f,g) = -r_e$

by Lemma (20.5.11). Therefore since $\deg f = n = -r_0$ and since

$$\text{g.c.d. } (f_0, \ldots, r_h) = d_{h+1} = 1,$$

it follows from Corollary (17.5) that there exists a homogeneous element 162
H of A of degree 1 such that $f^+ = \varnothing H^n$. This means that f has only one point at infinity. ◻

21 Solution in the Galois Case

In this section we show that the answer to the Jacobian problem is in the affirmative in case $k(x_1, x_2)/k(f, g)$ is a Galois extension (Theorem (21.11)).

We preserve the notation of § 15 and § 16. In addition, we fix the following notation: Let f, g be elements of $A = k[x_1, x_2]$ such that $J(f, g) = \varnothing$. Put $B = k[f, g]$ and $L = k(f, g)$. Recall that we have $k = k(x_1, x_2)$ and that char $k = 0$.

(21.1) Definition and Notation.

As in § 11, by a *valuation* we shall mean a real discrete valuation. Let Ω be a field of characteristic zero and E. F be over fields of Ω such that E is a finite field extension of F. Let v be a valuation of E/Ω and let $V = R_v$ be the discrete valuation ring of E/Ω associated to v. Let $W = V \cap F$. We say V *lies over* (or is an *extension* of) W. We denote by $e_{V|W}$ (or simply, e_V) the *ramification index* of V over W, i.e., $e_V = v(z)$, where z is a uniformizing parameter for W. We say V is *ramified* (resp. *unramified*) in the extension E/F if $E_v > 1$ (resp. $e_V = 1$). We say W is *ramified* in E/F if there exists an extension V of W to E such that V is ramified in E/F.

In our proof of Theorem (21.11) we shall need the following well-known formula:

(21.2) Lemma (Hurwitz Formula).

Let Ω be an algebraically closed field of characteristic zero and let E, F be function fields of one variable over Ω such that E is a finite extension 163 of F. Let $n = [F : F]$ and let g_E (resp. g_F) be the genus of E/Ω (resp. F/Ω). Then we have

$$2g_E - 2 = (2g_F - 2)n + \sum_V (e_V - 1),$$

where the summation is over all discrete valuation rings V of E/Ω and $e_V = e_{V|V\cap F}$.

For a proof of this lemma see, for instance, [4].

(21.3) COROLLARY. With the notation of Lemma (21.2) suppose that $g_F = 0$ and that there exists atmost one discrete valuation ring of F/Ω ramified in E/F. Then $E = F$.

Proof. By Lemma (21.2) we have

$$2g_E - 2 = -2n + \sum_V (e_V - 1).$$

By assumption, all those V for which $e_V > 1$ lie over the same discrete valuation ring of F. Therefore we have $\sum\limits_V (e_V - 1) \leq n - 1$ and we get $2g_E - 2 \leq -n - 1$, so that $n \leq 1 - 2g_E \leq 1$. □

(21.4) LEMMA. *K/L is a finite (separable) extension.*

Proof. Since K is finitely generated over L, we have only to show that K has no non-trivial L-derivations. Let d be an L-derivation of K. Then we have

$$0 = d(f) = D_1(f)d(x_1) + D_2(f)d(x_2),$$
$$0 = d(g) = D_1(g)d(x_1) + D_2(g)d(x_2).$$

Since $J(f, g) \neq 0$, we get $d(x_1) = 0 = d(x_2)$. Therefore $d = 0$. □

164 **(21.5) COROLLARY.** f and g are algebraically independent over k and B is the polynomial ring in two variables f and g over k.

(21.6) LEMMA. *Let $\mathscr{J}$ be a prime ideal of A of height one. Then $ht(\mathscr{J} \cap B) = 1$. (Here ht denotes "height".)*

Proof. Let $\overline{k}$ be the algebraic closure of k and let $\overline{A} = \overline{k}[x_1, x_2]$, $\overline{B} = \overline{k}[f, g]$. Since $\overline{A}$ is integral over A, there exists a prime ideal $\overline{\mathscr{P}}$ of $\overline{A}$ such that $\overline{\mathscr{P}} \cap A = \mathscr{P}$. Moreover, $ht\overline{\mathscr{P}} = 1$. since $\overline{B}$ is integral over B, we have $ht(\overline{\mathscr{P}} \cap \overline{B}) = ht(\mathscr{P} \cap B)$. We may therefore assume that $k = \overline{k}$. Since K/L is algebraic (Lemma (21.4)) we have $\mathscr{P} \cap B \neq 0$. Suppose $ht(\mathscr{P} \cap B) > 1$. Then $\mathscr{P} \cap B = (f - a)B + (g - b)B$ for some $a, b \in k$. We have $\mathscr{P} = pA$ for some $p \in A$. Since p divides $f - a$ and $g - b$ in A, p divides $J(f - a, g - b) = J(f, g) = \varnothing$ in A. This is a contradiction. □

(21.7) Proposition (Birational Case)

If $L = K$ then $B = A$.

Proof. Let $\mathscr{Q}$ be any prime ideal of B of height one. Then $\mathscr{Q} = qB$ for some $q \in B$. Since $q \notin k$, q is a non-unit in A. Therefore there exists a prime ideal $\mathscr{P}$ of A of height one such that $q \in \mathscr{P}$. Then by Lemma (21.6) we have $\mathscr{P} \cap B = \mathscr{Q}$. Therefore $B_{\mathscr{Q}} \subset A_{\mathscr{P}}$. Now, both $B_{\mathscr{Q}}$ and $A_{\mathscr{P}}$ are discrete valuation rings of the same field K. Therefore we have $B_{\mathscr{Q}} = A_{\mathscr{P}}$, so that $A \subset B_{\mathscr{Q}}$. Thus

$$a \subset \bigcap_{ht\mathscr{Q}=1} B_{\mathscr{Q}} = B.$$

□

(21.8) DEFINITION. Let $\mathscr{P}$ be a prime ideal of A of height one. We say $\mathscr{P}$ is *unramified* B if the discrete valuation ring $A_{\mathscr{P}}$ is unramified in the extension K/L (Definition (21.1)).

Note that $\mathscr{P}$ is unramified over B if and only if $\mathscr{P} \cap B \not\subset \mathscr{P}^2$.

(21.9) LEMMA. *Every prime ideal of A of height one is unramified* 165
over B.

Proof. Let $\mathscr{P}$ be a prime ideal of A of height one and let $\mathscr{Q} = \mathscr{P} \cap B$. We have to show that $\mathscr{Q} \not\subset \mathscr{P}^2$. Let $\mathscr{P} = pA$, $\mathscr{Q} = qB$ with $p \in A$, $q \in B$. Since $q \notin gk$, we have

(21.9.1) $$(\partial q/\partial f)B + (\partial q/\partial g)B \not\subset qB.$$

Now, we have

$$D_i(q) = (\partial q/\partial f)D_i(f) + (\partial q/\partial g)D_i(g)$$

for $i = 1, 2$. Therefore since $J(f, g) = \varnothing$, we get

$$(\partial q/\partial f)A + (\partial q/\partial g)A \subset D_1(q)A + D_2(q)A. \tag{21.9.2}$$

Now, suppose $q \in p^2A$. then $D_i(q) \in pA$, $i = 1, 2$. Therefore by (21.9.2) we have

$$(\partial q/\partial f)B + (\partial q/\partial g)B \subset pA \cap B = qB,$$

which contradicts (21.9.1) □

(21.10) LEMMA. *Let u_1, u_2 be elements of K such that $K/k(u_1, u_2)$ is a finite extension. Then there exists $a \in K$ such that $k(u_1 + au_2)$ is algebraically closed in K.*

Proof. For a subfield F of K let $\overline{F}$ denote its algebraic closure in K. Consider the family

$$\left\{\overline{k(u_1 + au_2)}(u_2)\,\middle|\,a \in k\right\}$$

of subfields of K containing $k(u_1, u_2)$. Since there are only finitely many fields between $k(u_1, u_2)$ and K (and since k is infinite), there exist a_1,
166 $a_2 \in k$, $a_1 \neq a_2$, such that $\overline{k(v_1)}(u_2) = \overline{k(v_2)}(u_2)$, where $v_1 = u_1 + a_1u_2$, $v_2 = u_1 + a_2u_2$. Since $u_2 \in k(v_1, v_2)$, we get $\overline{k(v_1)} \subset \overline{k(v_2)}(u_2) \subset \overline{k(v_2)}(v_1)$ and $\overline{k(v_2)} \subset \overline{k(v_1)}(u_2) \subset \overline{k(v_1)}(v_2)$. Therefore we have $\overline{k(v_1)}(v_2) = k\overline{k(v_2)}(v_1)$. Since $\overline{k(v_2)} \subset K = k(x_1, x_2)$, k is algebraically closed in $\overline{k(v_2)}$. Therefore, since u_1, u_2 and hence v_1, v_2 are algebraically independent over k, $k(v_1)$ is algebraically closed in $\overline{k(v_2)}(v_1) = \overline{k(v_1)}\ (v_2)$. This means that $k(v_1) = \overline{k(v_1)}$. □

(21.11) THEOREM. *If K/L is a Galois extension then $B = A$.*

Proof. In view of Proposition (21.7), we have only to show that $L = K$. Replacing f by $f + ag$ for some $a \in k$, we may assume that $k(f)$ is algebraically closed in K(Lemma (21.10)). Then, denoting by Ω the

algebraic closure of $k(f)$, we see that Ω and K are linearly disjoint over $k(f)$. It follows that $\Omega(g)$ and K are linearly disjoint over L, so that $L = \Omega(g) \cap K$, the intersection being taken in ΩK. Therefore, putting $E = \Omega K$, $F = \Omega(g)$, it is enough to show that $E = F$. Suppose $E \neq F$. Then it follows from Corollary (21.3) that at least two (discrete) valuation rings of F/Ω are ramified in E/F. Since the (g^{-1})-adic valuation ring is the only valuation ring of F/Ω not containing $\Omega[g]$, there exists a valuation ring W' of F/Ω such that $W' \supset \Omega[g]$ and W' is ramified in E/F. Let $W = W' \cap L$. Then W is a discrete valuation ring of L containing $k(f)[g]$ and is ramified in K/L. Since K/L is Galois, the extensions of W to K are ramified over L. Now, since $W \supset k(f)[g]$, $W = B$ for some prime ideal of B of height one. Let $= qB$ with $q \in B$. then q is a non-unit in A. Therefore there exists a prime ideal $\mathscr{J}$ of A of height one such that $q \in \mathscr{J}$. By Lemma (21.6) we have $\mathscr{J} \cap B =$. Therefore $W = B = A_{\mathscr{J}} \cap L$. Thus $\mathscr{J}$ is ramified over B. This is a contradiction 167
by Lemma (21.9). □

(21.12) REMARK. The above proof shows, in fact, that there cannot exist a proper Galois extension of L contained in K.

Bibliography

[1] S.S. Abhyankar and T.T. Moh: Newton-Puiseux expansion and generalizer Tschirnhausen transformation, J. reine angew. Math., *260*, 47-83 and *261*, 29-54 (1973). **168**

[2] S.S. Abhyankar and T.T. Moh:Embeddings of the line in the plane, J. refine angew, math., *276*, 148-166 (1975).

[3] S.S. Abhyankar and B. Singh: Embeddings of certain curves in the affine plane, to appear in Amer. J. Math.

[4] H. Hasse, Zahlenthrorie, 2nd ed., Akademie-Verlag, Berlin, 1963.

[5] W. van der Kulk, On polynomial rings in two variables, Nieuw Archief voor Wiskunde, (*3*) I, 33-41 (1953).