POLSKA AKADEMIA NAUK, INSTYTUT MATEMATYCZNY

DISSERTATIONES MATHEMATICAE

(ROZPRAWY MATEMATYCZNE)

KOMITET REDAKCYJNY

BOGDAN BOJARSKI redaktor
WIESŁAW ŻELAZKO zastępca redaktora
ANDRZEJ BIAŁYNICKI-BIRULA, ZBIGNIEW CIESIELSKI,
JERZY ŁOŚ, ZBIGNIEW SEMADENI

CCCIII

N. Ch. WASS

Algebraic independence of the values at algebraic points of a class of functions considered by Mahler

WARSZAWA 1990

5.7133

Published by the Institute of Mathematics, Polish Academy of Sciences
Typeset in TEX at the Institute
Printed and bound by M. & K. Herman, Spokojna 1, Raszyn

PRINTED IN POLAND

© Copyright by Instytut Matematyczny PAN, Warszawa 1990

ISBN 83-85116-02-8 ISSN 0012-3862

CONTENTS

Abstract

This thesis is concerned with the problem of determining a measure of algebraic independence for a particular m-tuple $\theta_1, \ldots, \theta_m$ of complex numbers. Specifically, let K be a number field and let $f_1(z), \ldots, f_m(z)$ be elements of $K[[z]]$ algebraically independent over $K(z)$ satisfying equations of the form

$$(*) \qquad f_j(z^b) = \sum_{i=1}^{m} f_i(z) a_{ij}(z) + b_j(z) \qquad (j = 1, \ldots, m)$$

for $b \geq 2$, $a_{ij}(z)$, $b_j(z)$ in $K(z)$. Suppose finally that $\alpha \in K$ is such that $0 < |\alpha| < 1$, the $f_j(z)$ converge at $z = \alpha$ and the $a_{ij}(z)$, $b_j(z)$ are analytic at $z = \alpha, \alpha^b, \alpha^{b^2}, \ldots$ Then the $\theta_i = f_i(\alpha)$ are algebraically independent numbers. This was essentially proved by Yu. V. Nesterenko for particular system $(*)$. He gave an ineffective measure of algebraic independence. The purpose of this thesis is to determine an effective measure of algebraic independence for the general case. In certain cases the estimate obtained implies that $(\theta_1, \ldots, \theta_m)$ has finite transcendence type in the sense of S. Lang.

Acknowledgements

This thesis would not have been possible without the advice and support of my thesis advisor Professor David W. Masser. In particular, he introduced the work of Yu. V. Nesterenko to me and supplied the idea behind the proof of Theorem 3.1. I would also like to thank the Department of Mathematics at the University of Michigan for its financial support during the preparation and publication of this thesis.

To my family

I. Introduction

§ 1.1. Algebraic independence

This thesis is concerned with transcendental numbers and more generally with algebraic independence of complex numbers. A complex number α is said to be *algebraic* if there exist rational integers $a_0, a_1, \ldots a_n$ not all zero such that

$$a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_n = 0\,.$$

If the number α is not algebraic it is said to be *transcendental.* Some examples of transcendental numbers are the following:

(1) $\sum_{n=0}^{\infty} 2^{-n!}$ (Liouville, 1826).

(2) e^{α} for α a nonzero algebraic number (in particular $e = e^1$ is transcendental and π is transcendental since $e^{\pi i} = -1$ (Lindemann, 1882).

(3) $\sum_{n=0}^{\infty} \alpha^{2^n}$, α algebraic, $0 < |\alpha| < 1$ (Mahler, 1926).

Before going any further we recall the following standard notation. The ring of rational integers is denoted by $\mathbf{Z}$ and $\mathbf{Q}$ denotes its quotient field. The algebraic closure of $\mathbf{Q}$ is denoted by $\overline{\mathbf{Q}}$, and is considered as a subfield of the field $\mathbf{C}$ of complex numbers. If $\mathcal{F}$ is a ring, for example, $\mathcal{F} = \mathbf{Z}$ or $\mathcal{F} = \mathbf{C}$, then $\mathcal{F}[x_1, \ldots, x_m]$ denotes the ring of polynomials in the indeterminates $x_1, \ldots, x_m$ and $\mathcal{F}[[x_1 \ldots, x_m]]$ denotes the power series ring in the indeterminates $x_1, \ldots, x_m$. If $\mathcal{F}$ is an integral domain then $\mathcal{F}[x_1, \ldots, x_m]$ is an integral domain and its quotient field is written as $\mathcal{F}(x_1, \ldots, x_m)$. Again if $\mathcal{F}$ is an integral domain then $\mathcal{F}[[x_1, \ldots, x_m]]$ is an integral domain and its quotient field is written as $\mathcal{F}((x_1, \ldots, x_m))$. Notice that if $\mathcal{F}$ is a field then $\mathcal{F}((z)) = \bigcup_{n=0}^{\infty} z^{-n}\mathcal{F}[[z]]$.

If $P \in \mathbf{Z}[x_1, \ldots, x_m]$, the maximum of the absolute values of the coefficients of P is called the *height* of P and written as $H(P)$. The degree of P, defined only for $P \neq 0$, is written as $\deg P$.

The definition of transcendence may now be phrased as follows. If α is a complex number then α is transcendental if

$$P(\alpha) = 0 \quad \text{implies that} \quad P = 0$$

for a polynomial $P \in \mathbf{Z}[x]$.

1985 *Mathematics Subject Classification* 11J85

The previous definition may be generalized. Let $\alpha_1, \ldots, \alpha_m$ be m complex numbers. Then $\alpha_1, \ldots, \alpha_m$ are algebraically independent if

$$P(\alpha_1, \ldots, \alpha_m) = 0 \quad \text{implies that} \quad P = 0$$

for a polynomial $P \in \mathbb{Z}[x_1, \ldots, x_m]$. Some examples of algebraically independent numbers will now be given.

(1) $e^{\alpha_1}, \ldots, e^{\alpha_m}$ where $\alpha_1, \ldots, \alpha_m$ are in $\overline{\mathbb{Q}}$ and linearly independent over $\mathbb{Q}$ (Lindemann, 1882).

(2) α^β, α^{β^2} where $\alpha \in \overline{\mathbb{Q}}$, $\alpha \neq 0, 1$ and β is a cubic irrational (Gel'fond, 1949).

(3) $f(\alpha), f'(\alpha), \ldots, f^{(m-1)}(\alpha)$ where $\alpha \in \overline{\mathbb{Q}}$, $0 < |\alpha| < 1$, $m \geq 1$, and $f(z) = \sum_{n=0}^{\infty} z^{k^n}$ with $k \geq 2$ an integer (Mahler, 1930).

It will be necessary to generalize the above definition of algebraic independence even further. Let k, K be fields with $k \subseteq K$. Then elements $\alpha_1, \ldots, \alpha_m$ of K are said to be *algebraically independent over* k if

$$P(\alpha_1, \ldots, \alpha_m) = 0 \quad \Rightarrow \quad P = 0$$

for any P in $k[x_1, \ldots, x_m]$. Note that for $K = \mathbb{C}$, $k = \mathbb{Q}$ this reduces to the above definition of algebraic independence of complex numbers since we can always clear denominators. In this thesis another important special case is when $k = \mathbb{C}(z)$ and $K = \mathbb{C}((z))$. As usual $\mathbb{C}((z))$ is the field of formal Laurent series in z with coefficients in $\mathbb{C}$.

Some examples of elements of $\mathbb{C}((z))$ algebraically independent over $\mathbb{C}(z)$ are the following

(1) $e^{\alpha_1 z}, \ldots, e^{\alpha_m z}$ with $\alpha_1, \ldots, \alpha_m$ complex numbers linearly independent over $\mathbb{Q}$.

(2) $f(z), \ldots, f^{(m-1)}(z)$ where $m \geq 1$ is an integer and

$$f(z) = \sum_{n=0}^{\infty} z^{k^n}$$

with $k \geq 2$ an integer (Mahler, 1930).

If $\alpha_1, \ldots, \alpha_m$ are algebraically independent over $\mathbb{Q}$ we know that

$$P(\alpha_1, \ldots, \alpha_m) \neq 0$$

for any $0 \neq P \in \mathbb{Z}[x_1, \ldots, x_m]$. A *measure of algebraic independence* for $\alpha_1, \ldots, \alpha_m$ is an explicit function $f(D, H)$ such that

$$|P(\alpha_1, \ldots, \alpha_m)| \geq f(D, H) > 0$$

for any $0 \neq P \in \mathbb{Z}[x_1, \ldots, x_m]$ satisfying $\deg P \leq D$, $H(P) \leq H$ with $D, H \geq 1$.

Some examples of algebraic independence measures will now be given.

I. *E-functions satisfying differential equations.* We refer to [La2] for the definition of an E-function: it is an entire function whose Taylor expansion at $z = 0$ has coefficients in a number field k with certain growth properties. Suppose that $f_1(z), \ldots, f_m(z)$ are E-functions algebraically independent over $\mathbf{C}(z)$ satisfying

$$\frac{d}{dz} f_j(z) = \sum_{i=1}^{n} a_{ij}(z) f_i(z) + b_i(z) \qquad (1 \le i \le m)$$

with $a_{ij}(z)$, $b_i(z)$ in $K(z)$. Let $\alpha \in K$ be such that $\alpha \neq 0$ and α is distinct from the poles of the $a_{ij}(z)$, $b_i(z)$. Shidlovskiĭ established that subject to the previous conditions the $\theta_i = f_i(\alpha)$ are algebraically independent. S. Lang (1962) refined Shidlovskiĭ's work to establish a measure of algebraic independence for θ_i of the type

$$\ln |P(\theta_1, \ldots, \theta_m)| \ge -\phi(D) - cD^m \ln H$$

for $0 \neq P \in \mathbf{Z}[x_1, \ldots, x_m]$ with $\deg P \le D$, $H(P) \le H$ and $D \ge 1$, $H \ge 1$.

Here the constant c is independent of D and H, but $\phi(D)$ is allowed to depend on D. It is interesting to note that while c can be explicitly computed in terms of $\theta_1, \ldots, \theta_m$, the function $\phi(D)$ at present cannot be. Nesterenko recently showed that one may take

$$\phi(D) < \exp(\tau D^{2m} \ln(D+1))$$

with τ independent of D and H. However, τ remains ineffective in the sense mentioned above (but see [Br] for the latest developments).

II. *Functions satisfying functional equations.* This is the main topic of the present work. Let K be a fixed algebraic number field. Let $f_1(z), \ldots, f_m(z)$ be elements of $K[[z]]$ algebraically independent over $K(z)$ satisfying

$$(*) \qquad f_j(z^b) = \sum_{i=1}^{m} a_{ij}(z) f_i(z) + b_j(z) \qquad (1 \le j \le m).$$

Let $\alpha \in K$ be such that $0 < |\alpha| < 1$, the $f_i(z)$ all converge at $z = \alpha$ and the powers $\alpha, \alpha^b, \alpha^{b^2}, \ldots$ are all distinct from the poles of the $a_{ij}(z)$, $b_i(z)$. Put $\theta_i = f_i(\alpha)$ $(1 \le i \le m)$.

For $m = 1$ the work of Galochkin [G] implies the transcendence measure

$$\ln |P(\theta)| > -\phi(D) - cD \ln H$$

for $0 \neq P \in \mathbf{Z}[x]$ where $\theta = f_1(\alpha)$, $\deg P \le D$, $H(P) \le H$. Here c is effective. W. Miller ([M1], [M2]) shows that for $m = 1$

$$\ln |P(\theta)| > -cD^2(D^2 + \ln H)$$

and determines explicitly the constant c that appears there.

In his paper [Ne4], Nesterenko shows that for arbitrary $m \geq 1$ subject to the hypotheses of $(*)$ the $\theta_i = f_i(\alpha)$ satisfy

$$\ln|P(\theta_1, \ldots, \theta_m)| \geq -\phi(D) - cD^m \ln H .$$

Actually, he considers only diagonal matrices in $(*)$ but an examination of the proof shows that this hypothesis is unnecessary. In fact, by using zero estimates and building on the work of Nesterenko it will be shown that subject to the hypotheses of $(*)$:

THEOREM 5.2.2.

$$\ln|P(\theta_1, \ldots, \theta_m)| \geq -\exp(cD^m) - cD^m \ln H$$

for some effective constant c.

THEOREM 5.2.3. *If the matrix* $[a_{ij}(z)]_{1 \leq i,j \leq m}$ *in* $(*)$ *is upper triangular then*

$$\ln|P(\theta_1, \ldots, \theta_m)| \geq -cD^m(D^{2^{m+1}} + \ln H)$$

for some effective constant c.

In both these theorems it is understood that $0 \neq P \in \mathbf{Z}[x_1, \ldots, x_m]$ satisfies $\deg P \leq D$, $H(P) \leq H$, $D \geq 1$ and that $\theta_i = f_i(\alpha)$ $(1 \leq i \leq m)$. (Actually $D \geq 0$ suffices.)

It is interesting to note that one may find, by means of Theorem 5.2.3, m-tuples $\theta_1, \ldots, \theta_m$ of complex numbers for arbitrary $m \geq 1$ of finite transcendence type. Here we adopt the definition of Serge Lang ([La1], p. 51). We define $\theta_1, \ldots, \theta_m$ to be of *finite transcendence type* if

$$\ln|P(\theta_1, \ldots, \theta_m)| \geq -c(D + \ln H)^\tau$$

for constants $c, \tau > 0$ and all nonzero $P \in \mathbf{Z}[x_1, \ldots, x_m]$ with $H(P) \leq H$, $\deg P \leq D$.

For $m = 1$ there are many known examples. For instance e, π are separately of finite transcendence type and by W. Miller's result if $m = 1$ in $(*)$ then $\theta_1 = f_1(\alpha)$ is of transcendence type ≤ 4. For $m = 2$ it was established by Chudnovsky in some work of his on elliptic functions ([Ch]) that π, $\Gamma(1/4)$ for example are of finite transcendence type. To construct m-tuples $\theta_1, \ldots, \theta_m$ of finite type one may take

$$g(z) = \sum_{n=0}^{\infty} z^{2^n} ,$$

$$f_i(z) = g^{(i-1)}(z) = \frac{d^{i-1}}{dz^{i-1}} g(z) \qquad (1 \leq i \leq m)$$

and α any algebraic number satisfying $0 < |\alpha| < 1$. Theorem 5.2.3 applies and the $\theta_i = f_i(\alpha)$ $(1 \leq i \leq m)$ are of finite transcendence type $\leq 2^{m+1} + m$.

Finally, note that functional equations of the type $(*)$ are related to the concepts of regular sequence and finite automata. For a discussion of this, see the survey article in [L]. Further references are given there.

§ 1.2. Notation and some estimates

In this section some of the notation used in the later chapters will be introduced. For the most part this notation is standard and is used frequently throughout transcendence theory.

Recall that the algebraic closure of $\mathbb{Q}$ is denoted by $\overline{\mathbb{Q}}$ and is considered as a subfield of the field $\mathbb{C}$ of complex numbers. If $\alpha \in \overline{\mathbb{Q}}$ then α is a root of some polynomial $F(x) \in \mathbb{Z}[x]$ of positive degree. If the degree of $F(x)$ is minimal and $F(x)$ is primitive we call $F(x)$ the *minimum polynomial* of α. In this case $F(x)$ is necessarily irreducible and is a divisor of any polynomial T for which $T(\alpha) = 0$. If $F(x)$ has leading coefficient ± 1 we call α an *algebraic integer*. If α satisfies any monic polynomial then α is an algebraic integer. The algebraic integers form a ring. If a is the leading coefficient of $F(x)$ then $|a|\alpha$ is an algebraic integer. The smallest positive integer d such that $d\alpha$ is an algebraic integer is called the *denominator* of α and written $\operatorname{den}\alpha$. It is easy to show that if $F(x) = a_g x^g + \ldots + a_0$ is the minimum polynomial of α then

(1) $\operatorname{den}\alpha$ divides into a_g,
(2) a_g divides into $(\operatorname{den}\alpha)^g$.

To show (2) for instance $\beta = d\alpha$ is an algebraic integer where $d = \operatorname{den}\alpha$ and β satisfies the monic polynomial

$$\sum_{i=0}^{g} a_g^{-1} d^i a_{g-i} x^{g-i}$$

so $a_g^{-1} d^i a_{g-i} \in \mathbb{Z}$ $(0 \le i \le g)$. Let $1 = \sum_{i=0}^{g} c_i a_{g-i}$ with $c_i \in \mathbb{Z}$. Then

$$a_g^{-1} d^g = \sum_{i=0}^{g} c_i \left(a_g^{-1} d^i a_{g-i}\right) d^{g-i} \in \mathbb{Z}.$$

The size of an algebraic number will be defined next. If $\alpha \in \overline{\mathbb{Q}}$ we define the *size* of α to be

$$\lceil\alpha\rceil = \max_{1 \le i \le g} |\alpha^{(i)}|$$

where $\alpha^{(i)}$ $(1 \le i \le g)$ are the *conjugates* of α, that is, they are the roots of the minimum polynomial of α. These conjugates are precisely those complex numbers of the form $\sigma_i\alpha$ for $i = 1, \ldots, g$ where $\{\sigma_1, \ldots, \sigma_g\}$ is the set of $\mathbb{Q}$-monomorphisms of $\mathbb{Q}(\alpha)$ into $\overline{\mathbb{Q}}$.

Notice that if $0 \neq \alpha \in \overline{\mathbb{Q}}$ then

$$\max\left(\operatorname{den}(\alpha^{-1}), \overline{|\alpha^{-1}|}\right) \leq (\max(\operatorname{den}\alpha, \overline{|\alpha|}))^{2g}$$

where $g = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. To see this notice that if

$$F(x) = a_g x^g + \ldots + a_0 = a_g(x - \alpha^{(1)}) \ldots (x - \alpha^{(g)})$$

is the minimum polynomial of α then

$$G(x) = a_0 x^g + \ldots + a_g = a_0(x - (\alpha^{(1)})^{-1}) \ldots (x - (\alpha^{(g)})^{-1})$$

is the minimum polynomial of α^{-1}. Since

$$|(\alpha^{(i)})^{-1}| = \left|\frac{a_g}{a_0}\right| \prod_{j \neq i} |\alpha^{(j)}| \leq |a_g| \overline{|\alpha|}^{\,g-1}$$

$$\leq (\operatorname{den}\alpha)^g \overline{|\alpha|}^{\,g-1} \leq (\max(\operatorname{den}\alpha, \overline{|\alpha|}))^{2g}$$

then $\overline{|\alpha^{-1}|} \leq (\max(\operatorname{den}\alpha, \overline{|\alpha|}))^{2g}$ follows. The other inequality to be proved is left to the reader.

The inequalities

$$\overline{|\alpha \cdot \beta|} \leq \overline{|\alpha|}\,\overline{|\beta|}\,, \quad \overline{|\alpha + \beta|} \leq \overline{|\alpha|} + \overline{|\beta|}$$

for α, β in $\overline{\mathbb{Q}}$ are obvious.

If $\alpha_1, \ldots, \alpha_t \in \overline{\mathbb{Q}}$, we define *the denominator* of $\alpha_1, \ldots, \alpha_t$ to be the smallest positive integer d with the property that $d\alpha_i$ is an algebraic integer $(1 \leq i \leq t)$. Notice that any other integer with this property must be a multiple of d. Such an integer is called *a denominator* of $\alpha_1, \ldots, \alpha_t$. If $f \in \overline{\mathbb{Q}}[x_1, \ldots, x_t]$ define $\operatorname{den} f$ to be the denominator of the coefficients of f. Define $\overline{|f|}$ to be the maximum of the sizes of the coefficients of f. With these definitions, it is not difficult to establish the following inequality.

If $\alpha_1, \ldots, \alpha_t \in \overline{\mathbb{Q}}$, $0 \neq f \in \overline{\mathbb{Q}}[x_1, \ldots, x_t]$ with

$$\deg_{x_\nu} f \leq k_\nu \quad (1 \leq \nu \leq t) \qquad \text{and} \qquad \beta = f(\alpha_1, \ldots, \alpha_t) \in \overline{\mathbb{Q}}$$

then

(i) $\overline{|\beta|} \leq 2^t \overline{|f|} \prod_{\nu=1}^t \max(2, \overline{|\alpha_\nu|})^{k_\nu}$,

(ii) $(\operatorname{den} f) \prod_{\nu=1}^t (\operatorname{den}\alpha_\nu)^{k_\nu}$ is a denominator of β.

Indeed,

$$\beta = \sum_{i_1=0}^{k_1} \cdots \sum_{i_t=0}^{k_t} a_{i_1 \ldots i_t} \alpha_1^{i_1} \ldots \alpha_t^{i_t} \quad \text{where} \quad f = \sum_{i_1=0}^{k_1} \cdots \sum_{i_t=0}^{k_t} a_{i_1 \ldots i_t} x_1^{i_1} \ldots x_t^{i_t}$$

and a denominator of β is

$$(\operatorname{den} f)(\operatorname{den}\alpha_1)^{k_1} \ldots (\operatorname{den}\alpha_t)^{k_t}\,,$$

proving (ii). Also

$$\lceil\beta\rceil \le \lceil f\rceil \sum_{i_1=0}^{k_1} \cdots \sum_{i_t=0}^{k_t} A_1^{i_1} \cdots A_t^{i_t}$$

where $A_\nu = \max(2, \lceil\alpha_\nu\rceil)$ $(1 \le \nu \le k)$, which proves (i). These estimates will be used in Chapter V.

Finally, we introduce some notation. For a polynomial $P \in \mathbb{C}[x_1, \ldots, x_m]$ the *height* $H(P)$ is the maximum of the absolute values of the coefficients of P. For $P, Q \in \mathbb{C}[[x_1, \ldots, x_m]]$ we write

$$P \ll Q$$

to mean

$$P = \sum a_{i_1 \ldots i_m} x_1^{i_1} \ldots x_m^{i_m}, \qquad Q = \sum b_{i_1 \ldots i_m} x_1^{i_1} \ldots x_m^{i_m}$$

with $|a_{i_1 \ldots i_m}| \le b_{i_1 \ldots i_m}$ for all $i_1, \ldots, i_m \ge 0$. In particular this notation is to be used in the special cases when P, Q are both in $\mathbb{C}[x_1, \ldots, x_m]$ or both in $\mathbb{C}[[z]]$.

With the above representations of P, Q we write

$$P \lll Q$$

to mean that $\lceil a_{i_1 \ldots i_m}\rceil \le b_{i_1 \ldots i_m}$ for all $i_1, \ldots, i_m \ge 0$.

If Γ is a ring and m, n are positive integers we denote by $M_{mn}(\Gamma)$ the set of all $m \times n$ matrices with coefficients in Γ.

Finally, if K is an algebraic number field we denote by $\mathcal{O}_K$ the ring of integers of K.

II. Formal series

§ 2.1. A class to which the solutions belong

In this section we introduce a class of formal Laurent series $\mathcal{G}$ and show that solutions $\underline{E}(z)$ of

$$\underline{E}(z) = \underline{E}(z^b)\underline{\underline{A}}(z) + \underline{B}(z)$$

with

$$\underline{\underline{A}}(z) \in M_{mm}(K(z)) \qquad \text{nonsingular},$$

$$\underline{B}(z) \in M_{1m}(K(z)),$$

$$\underline{E}(z) \in M_{1m}(K((z)))$$

are necessarily in $M_{1m}(\mathcal{G})$. Here K is a fixed algebraic number field and $b \geq 2$ an integer. This will be useful for simplifying the hypotheses needed for the Theorems of Chapter IV.

Let K be a fixed number field and $\mathcal{R} = K[[z]]$ be the ring of formal power series with coefficients in K, and let $\mathcal{F} = K((z)) = \bigcup_{n=0}^{\infty} z^{-n} K[[z]]$ be the quotient field of $\mathcal{R}$.

We define the subset $\mathcal{G}(c, C)$ of $\mathcal{F}$ to be the set of all $f(z) = \sum c_\mu z^\mu$ for which

(i) $\overline{|c_\mu|} \leq cC^\mu$ $(\mu = 0, 1, 2, \ldots)$,
(ii) $cC^\mu c_\mu \in \mathcal{O}_K$ $(\mu = 0, 1, 2, \ldots)$.

Here $\mathcal{O}_K$ denotes the ring of algebraic integers of K and c, C are positive integers. Now define

$$\mathcal{G} = \bigcup_{c=1}^{\infty} \bigcup_{C=1}^{\infty} \mathcal{G}(c, C) \subseteq \mathcal{F}.$$

LEMMA 2.1.1. *$\mathcal{G}$ is a subfield of $\mathcal{F}$.*

Proof. It is easy to see that $\mathcal{G}$ is a ring; the proof will be left to the reader. To prove that it is a field it must be shown that if $f \neq 0$ is in $\mathcal{G}$ then f^{-1} is in $\mathcal{G}$. For this we can evidently suppose f is in $\mathcal{R}$ and $f(0) = 1$. Now $f = 1 + g$ for

$$g \lll G = cC + cC^2 z^2 + \ldots = cCz/(1 - Cz)$$

so

$$f^{-1} = 1 - g + g^2 - \ldots \lll 1 + G + G^2 + \ldots = 1/(1 - G).$$

Since

$$\frac{1}{1-G} = \frac{c}{(c+1)(1-(c+1)Cz)} + \frac{1}{c+1}$$

property (i) follows immediately for f^{-1}.

As for property (ii), we merely note that $f(cCz) = 1 + h$ for h in $z\mathcal{O}_K[[z]]$. Therefore $1/f(cCz)$ is in $\mathcal{O}_K[[z]]$, and (ii) follows easily.

COROLLARY. *$\mathcal{G}$ contains $K(z)$.*

THEOREM 2.1.1. *Let $\underline{\underline{A}}(z) \in M_{mm}(\mathcal{G})$ and $\underline{B}(z) \in M_{1m}(\mathcal{G})$ be such that $\underline{E}(z)$ in $M_{1m}(\mathcal{F})$ satisfies the functional equation*

$$\underline{E}(z) = \underline{E}(z^b)\underline{\underline{A}}(z) + \underline{B}(z) \tag{2.1.1}$$

for some integer $b > 1$. Then $\underline{E}(z)$ is in $M_{1m}(\mathcal{G})$.

Proof. Suppose this has been proved in the special case when $\underline{E}(z) \in M_{1m}(K[[z]])$, $\underline{\underline{A}}(z) \in M_{mm}(K[[z]])$ and $\underline{B}(z) \in M_{1m}(K[[z]])$. The general case can be deduced from this.

Pick $n \geq 0$ such that

$$\underline{\underline{A}}_1(z) = z^{n(b-1)}\underline{\underline{A}}(z) \in M_{mm}(K[[z]]) .$$

Write

$$z^{-n}\underline{E}(z) = \sum_{\mu} e_\mu z^\mu \qquad \text{and} \qquad \underline{E}_1(z) = \sum_{\mu \geq 0} e_\mu z^\mu$$

so that $\underline{F}(z) = z^{-n}\underline{E}(z) - \underline{E}_1(z) \in M_{1m}(\mathcal{G})$ by the Corollary of Lemma 1. Then (2.1.1) can be written as

$$\underline{E}_1(z) = \underline{E}_1(z^b)\underline{\underline{A}}_1(z) + \underline{B}_1(z) \tag{2.1.2}$$

where

$$\underline{B}_1(z) = z^{-n}\underline{B}(z) + \underline{F}(z^b)\underline{\underline{A}}_1(z) - \underline{F}(z) . \tag{2.1.3}$$

From (2.1.2) we see that $\underline{B}_1(z)$ is in $M_{1m}(K[[z]])$ and from (2.1.3) it is in $M_{1m}(\mathcal{G})$. It follows from the special case above that $\underline{E}_1(z)$ is in $M_{1m}(\mathcal{G})$. Therefore $\underline{E}(z) = z^n(\underline{E}_1(z) + \underline{F}(z))$ is also in $M_{1m}(\mathcal{G})$.

We may therefore assume that $\underline{E}(z) \in M_{1m}(K[[z]])$, $\underline{\underline{A}}(z) \in M_{mm}(K[[z]])$ and $\underline{B}(z) \in M_{1m}(K[[z]])$. We may write

$$\underline{E}(z) = \sum_{\mu=0}^{\infty} \underline{e}_\mu z^\mu , \qquad \underline{\underline{A}}(z) = \sum_{\mu=0}^{\infty} a_\mu z^\mu , \qquad \underline{B}(z) = \sum_{\mu=0}^{\infty} \underline{b}_\mu z^\mu .$$

For a matrix A with entries a we write $\lceil A \rceil$ for the maximum of the $\lceil a \rceil$. Since the entries of $\underline{\underline{A}}$ and $\underline{B}$ are in $\mathcal{G}$ we can suppose that

$$\lceil \underline{\underline{a}}_\mu \rceil \leq cC^\mu , \quad \lceil b_\mu \rceil \leq cC^\mu \qquad (\mu \geq 0) \tag{2.1.4}$$

$$cC^\mu \underline{\underline{a}}_\mu \in M_{mm}(\mathcal{O}_K) , \quad cC^\mu \underline{b}_\mu \in M_{1m}(\mathcal{O}_K) \qquad (\mu \geq 0) \tag{2.1.5}$$

for positive integers c, C. We have to prove that there are positive integers e, E such that

$$\lceil \underline{e}_\mu \rceil \leq eE^\mu \qquad (\mu \geq 0) \tag{2.1.6}$$

$$eE^\mu \underline{e}_\mu \in M_{1m}(\mathcal{O}_K) \qquad (\mu \geq 0) \tag{2.1.7}$$

To prove (2.1.6) we choose E such that

$$E \geq 2C^b , \tag{2.1.8}$$

$$E \geq 3mc , \tag{2.1.9}$$

and then choose e so that (2.1.6) holds for $0 \leq \mu < b$. We verify by induction on μ that (2.1.6) holds for all $\mu \geq b$.

Suppose now that $\mu \geq b$ and that (2.1.6) holds with $0, 1, \ldots, \mu - 1$ in place of μ. The equations (2.1.1) may be written as

$$\underline{e}_\mu = \sum_{\nu=0}^{M} \underline{e}_\nu \underline{\underline{a}}_{\mu-\nu b} + \underline{b}_\mu , \qquad M = [\mu/b] . \tag{2.1.10}$$

Since $M < \mu$ we can use (2.1.6) in (2.1.10) to get

$$\overline{|\underline{e}_\mu|} \leq \sum_{\nu=0}^{M} meE^\nu cC^{\mu-\nu} + cC^\mu = C^\mu mec \sum_{\nu=0}^{M} (E/C^b) + cC^\mu . \tag{2.1.11}$$

By (2.1.8) we have

$$\sum_{\nu=0}^{M} (E/C^b)^\nu < 2(E/C^b)^M \leq 2(E/C^b)^{\mu/b} = 2c^{-\mu}E^{\mu/b}$$

and

$$C^\mu \leq (E/2)^{\mu/b} < E^{\mu/b} .$$

Putting these into (2.1.11) we find that

$$\overline{|\underline{e}_\mu|} \leq 3mecE^{\mu/b} .$$

But $\mu \geq b$ and so $\mu/b \leq \mu - 1$, and using (2.1.9) the above reduces to the desired inequality (2.1.6).

The inequality (2.1.7) can be proved in a similar way using nonarchimedean valuations. This time choose E so that

$$C^b \text{ divides into } E , \tag{2.1.12}$$

$$c \text{ divides into } E , \tag{2.1.13}$$

and then choose e so that (2.1.7) holds for $0 \leq \mu < b$. Therefore for any nonarchimedean valuation $|\ |$ on K and with the obvious extension to matrices we have from (2.1.5)

$$|\underline{\underline{a}}_\mu| \leq |c^{-1}||C^{-1}|^\mu , \quad |\underline{b}_\mu| \leq |c^{-1}||C^{-1}|^\mu \qquad (\mu \geq 0)$$

and

$$|\underline{e}_\mu| \leq |e^{-1}||E^{-1}|^\mu \qquad (0 \leq \mu < b) .$$

Using the same inductive scheme we get from (2.1.10)

$$|\underline{e}_\mu| \leq \max\{\max_{0\leq\nu\leq M} |e^{-1}||E^{-1}|^\nu |c^{-1}||C^{-1}|^{\mu-\nu b}, |c^{-1}||C^{-1}|^\mu\}$$

$$= |e^{-1}c^{-1}||C^{-1}|^\mu \max_{0\leq\nu\leq M} \left(|E^{-1}C^b|^\nu\right) .$$

By (2.1.12) we have $|E^{-1}C^b| \geq 1$, and so

$$|\underline{e}_\mu| \leq |e^{-1}c^{-1}||C^{-1}|^\mu|E^{-1}C^b|^M \leq |e^{-1}c^{-1}||C^{-1}|^\mu|E^{-1}C^b|^{\mu/b}$$

$$= |e^{-1}c^{-1}||E^{-1}|^{\mu/b} \leq |e^{-1}c^{-1}||E^{-1}|^{\mu-1} = |e^{-1}||E^{-1}|^\mu|c^{-1}E|\,.$$

By (2.1.13) we see that $|c^{-1}E| \leq 1$ and so $|eE^\mu\underline{e}_\mu| \leq 1$. Since this holds for all nonarchimedean valuations, the desired inclusion (2.1.7) follows.

Remark 1. Consider $A(z)$ in $M_{mm}(\mathcal{G})$, $\underline{B}(z)$ in $M_{nm}(\mathcal{G})$, Λ in $M_{nn}(K)$ nonsingular and $\underline{E}(z)$ in $M_{nm}(\mathcal{F})$ such that

$$\Lambda\underline{E}(z) = \underline{E}(z^b)\underline{A}(z) + \underline{B}(z)\,.$$

By following a similar proof it may easily be shown that $\underline{E}(z)$ is in $M_{mn}(\mathcal{G})$.

Remark 2. The idea of nonarchimedean valuations is not essential but simplifies the proof.

III. Zero estimates

§ 3.1. The general case

We are concerned in this chapter with an estimate of the following type. Suppose that L is a field and $f_1(z), \ldots, f_m(z)$ are in $L((z))$ and algebraically independent over $L(z)$. Let $0 \neq P \in L[z, x_1, \ldots, x_m]$. Define

$$\deg_{\underline{x}} P = \deg_{x_1,\ldots,x_m} P$$

to be the degree in $x_1, \ldots, x_m$ of P. Similarly we define $\deg_z P$ to be the degree in z of P. Now suppose that $\deg_z P \leq D$, $\deg_{\underline{x}} P \leq N$ with $D, N \geq 0$. We know that

$$0 \neq \varphi(z) = P(z, f_1(z), \ldots, f_m(z))\,.$$

Let $\varphi(z) = \sum a_\mu z^\mu$ where $a_t \neq 0$ and $a_l = 0$ for $l < t$. We define $\operatorname{ord}_{z=0}\varphi(z) = t$ and also write

$$\operatorname{ord} P = \operatorname{ord} P(z, \underline{f}(z)) = \operatorname{ord}_{z=0}\varphi(z)\,.$$

Here $\underline{f}(z) = (f_1(z), \ldots, f_m(z))$ so that $P(z, \underline{f}(z))$ is an abbreviation for $P(z, f_1(z), \ldots, f_m(z))$. A *zero estimate* is an estimate of the type

$$\operatorname{ord} P \leq W(D, N)$$

for W depending only on D and N.

The first zero estimate of this chapter (Theorem 3.1) is based on an idea of D. W. Masser. It gives a weak upper bound since it is exponential in the mth power of $\deg P$. However, it does show that Nesterenko's version

of Mahler's method is completely effective in principle. We may combine Theorem 3.1 with Theorem 4.1 of Chapter IV to obtain Theorem 4.3 for which the constant c can be effectively computed.

The other zero estimate of this chapter is Theorem 3.3. This was obtained by using resultants in a similar manner to the paper of Galochkin [G], although he treats a slightly different type of functional equation with $m = 1$ and with z replaced by an s-tuple $(z_1, \ldots, z_s)$. In contrast to Theorem 3.1, the estimate of Theorem 3.3 is a polynomial in its dependence on $\deg_z P$ and $\deg_{\underline{x}} P$ although it applies only to the situation where the matrix $A(z)$ is upper triangular. As with Theorem 3.1 we may combine Theorem 3.2 with Theorem 4.1 of Chapter IV to obtain Theorem 4.4 for which the constant c can be effectively computed. Of course Theorem 4.1 is an effective version of a theorem of Nesterenko [Ne4].

The two theorems, Theorem 3.1 and Theorem 3.3, will now be described. Both are concerned with the following situation. Let $b \geq 2$ be an integer, L be a field and $f_1(z), \ldots, f_m(z)$ be elements of $L[[z]]$ algebraically independent over $L(z)$ that satisfy

$$\underline{f}(z^b) = \underline{f}(z)\underline{\underline{A}}(z) + \underline{B}(z) \tag{3.1.1}$$

where

$$\underline{f}(z) = (f_1(z), \ldots, f_m(z)),$$

$$\underline{\underline{A}}(z) = [a_{ij}(z)]_{1 \leq i,j \leq m} \in M_{mm}(L(z)),$$

$$\underline{B}(z) = [b_j(z)]_{1 \leq j \leq m} \in M_{1m}(L(z)).$$

Let $0 \neq T(z) \in L[z]$ be such that

$$T(z)a_{ij}(z) \in L[z], \quad T(z)b_i(z) \in L[z] \tag{3.1.2}$$

with all these $m^2 + m + 1$ polynomials being either zero or having degree at most $q \geq 0$.

THEOREM 3.1. *Suppose that the conditions* (3.1.1) *and* (3.1.2) *hold. In addition suppose that* $q \geq 1$. *Let* $0 \neq P \in L[z, x_1, \ldots, x_m]$ *satisfy*

$$\deg_z P \leq d_0, \quad \deg_{x_i} P \leq d_i \qquad (1 \leq i \leq m).$$

Then $\operatorname{ord} P \leq qM^2 b^M (d_0 + 1)$ *where* $M = (d_1 + 1) \ldots (d_m + 1)$.

THEOREM 3.3. *Suppose that the conditions* (3.1.1) *and* (3.1.2) *hold, and in addition suppose that the matrix* $\underline{\underline{A}}(z)$ *is upper triangular. Let* $0 \neq P \in L[z, x_1, \ldots, x_m]$ *and suppose that*

$$\deg_z P \leq D, \qquad \deg_{\underline{x}} P \leq N \qquad (N \geq 1).$$

Then $\operatorname{ord} P \leq (b+2)^m N^{2^m - 1}(D + mqN)$.

We will now proceed to give a proof of Theorem 3.1. As mentioned before, after some preliminary results in § 3.2 on resultants, we will give a proof of Theorem 3.3 in § 3.3.

LEMMA 3.1.1. *Let $b \geq 2$ be an integer, L be a field and $\phi_j(z)$ ($j = 1, \ldots, M$) be elements of $L[[z]]$ satisfying*

$$\phi_j(z^b) = \sum_{i=1}^{M} \phi_i(z)\varsigma_{ij}(z) \qquad (j = 1, \ldots, M).$$

Let $0 \neq T(z) \in L[z]$ be such that $T(z)\varsigma_{ij}(z) \in L[z]$ for all i, j and suppose that these $M^2 + 1$ polynomials are either zero or have degree at most $\varsigma \geq 0$. Let

$$\psi(z) = \sum_{s=1}^{M} \theta_s(z)\phi_s(z)$$

where the $\theta_s(z) \in L[z]$ ($s = 1, \ldots, M$) are polynomials which are either zero or have degree at most $r \geq 0$. Then

$$\operatorname{ord}_{z=0} \psi(z) \leq \frac{M}{b-1}\left(b^M r + \varsigma \frac{b^M - 1}{b-1}\right) \leq M b^M (r + \varsigma).$$

Proof.

(1) *Define*

$$g^{(k)}(z) = \begin{cases} 1 & \text{for } k = 0, \\ T(z)T(z^b)\ldots T(z^{b^{k-1}}) & \text{for } k \geq 1. \end{cases}$$

Then for each $k \geq 0$

$$\phi_j(z^{b^k}) = \sum_{i=0}^{M} \phi_i(z)\varsigma_{ij}^{(k)}(z) \qquad (j = 1, \ldots, M)$$

where each $g^{(k)}(z)\varsigma_{ij}^{(k)}(z)$ is in $L[z]$ and is either zero or of degree at most $\varsigma(b^k - 1)/(b-1)$. Further

$$\deg g^{(k)}(z) \leq \varsigma \frac{b^k - 1}{b-1}.$$

Proof. We prove (1) by induction on k, $k \geq 0$. For $k = 0$ put $\varsigma_{ij}^{(0)}(z) = \delta_{ij}$. Suppose that $k \geq 1$ and the $\varsigma_{ij}^{(k-1)}(z)$ are defined with the correct properties. Put

$$\varsigma_{ij}^{(k)}(z) = \sum_{l=1}^{M} \varsigma_{il}(z)\varsigma_{lj}^{(k-1)}(z^b) \in L[z].$$

BU
W

Notice that

$$\sum_{i=1}^{M} \phi_i(z)\varsigma_{ij}^{(k)}(z) = \sum_{l=1}^{M} \varsigma_{ij}^{(k-1)}(z^b) \sum_{i=1}^{M} \phi_i(z)\varsigma_{il}(z)$$

$$= \sum_{l=1}^{M} \phi_l(z^b)\varsigma_{lj}^{(k-1)}(z^b) = \phi_j(z^{b^k}) .$$

Also

$$g^{(k)}(z)\varsigma_{ij}^{(k)}(z) = \sum_{l=1}^{M} (T(z)\varsigma_{il}(z))(g^{(k-1)}(z^b)\varsigma_{lj}^{(k-1)}(z^b)) \in L[z]$$

and if $g^{(k)}(z)\varsigma_{ij}^{(k)}(z) \neq 0$ then

$$\deg\,(g^{(k)}(z)\varsigma_{ij}^{(k)}(z)) \leq \varsigma + b\varsigma\frac{b^{k-1}-1}{b-1} = \varsigma\frac{b^k-1}{b-1} .$$

Since

$$\deg g^{(k)}(z) \leq \varsigma\frac{b^k-1}{b-1} ,$$

the inductive proof is complete.

(2) *For any* $\mu \geq 0$

$$\psi(z^{b^\mu}) = \sum_{r=1}^{M} p_{\mu r}(z)\phi_r(z)$$

where

$$q_{\mu r}(z) = g^{(\mu)}(z)p_{\mu r} \in L[z]$$

is either zero or has degree at most

$$b^\mu r + \varsigma\frac{b^\mu-1}{b-1} .$$

Proof.

$$\psi(z^{b^\mu}) = \sum_{s=1}^{M} \theta_s(z^{b^\mu})\phi_s(z^{b^\mu})$$

$$= \sum_{s=1}^{M} \theta_s(z^{b^\mu}) \sum_{r=1}^{M} \phi_r(z)\varsigma_{rs}^{(\mu)}(z) = \sum_{r=1}^{M} p_{\mu r}(z)\phi_r(z)$$

where

$$p_{\mu\tau}(z) = \sum_{s=1}^{M} \theta_s(z^{b^\mu}) \varsigma_{\tau s}^{(\mu)}(z) .$$

The desired conclusion follows immediately from this result and (1).

(3) *There exist polynomials* $h_\mu(z) \in L[z]$ *not all zero with degrees at most* MD *where*

$$D = b^M r + \varsigma \frac{b^M - 1}{b - 1}$$

such that

$$\sum_{\mu=0}^{M} p_{\mu\tau}(z) h_\mu(z) = 0 \qquad (\tau = 1, \ldots, M) .$$

Consequently

$$\sum_{\mu=0}^{M} h_\mu(z) \psi(z^{b^\mu}) = \sum_{\mu=0}^{M} \sum_{\tau=1}^{M} h_\mu(z) p_{\mu\tau}(z) \phi_\tau(z)$$

$$= \sum_{\tau=1}^{M} \sum_{\mu=0}^{M} p_{\mu\tau}(z) h_\mu(z) \phi_\tau(z) = 0 .$$

Proof. Put $q_{\mu\tau}^*(z) = g^{(M)}(z) p_{\mu\tau}(z)$ $(\mu = 0, 1, \ldots, M,\ \tau = 1, \ldots, M)$. Notice that either $q_{\mu\tau}^*(z) = 0$ or

$$\deg q_{\mu\tau}^*(z) \le \deg (g^{(M)}(z)/g^{(\mu)}(z)) + \varsigma(b^\mu - 1)/(b - 1)$$

$$\le \varsigma(b^M - b^\mu)/(b - 1) + \varsigma(b^\mu - 1)/(b - 1) = D .$$

We now proceed to solve

$$\sum_{\mu=0}^{M} q_{\mu\tau}^*(z) h_\mu(z) = 0 \qquad (\tau = 1, \ldots, M) .$$

This may be done with $h_\mu(z) \in L[z]$ not all zero and either $h_\mu(z) = 0$ or $\deg h_\mu(z) \le MD$. This is because if we require $h_\mu(z) = 0$ or $\deg h_\mu(z) \le Q$ for all μ we obtain

$$q_{\mu\tau}^*(z) = \sum_{\kappa=0}^{D} \alpha_{\mu\tau\kappa} z^\kappa , \qquad h_\mu(z) = \sum_{\nu=0}^{Q} \beta_{\mu\nu} z^\nu .$$

Put $\alpha_{\mu\tau\kappa} = 0$ if $\kappa > D$ or $\kappa < 0$. The equations to be solved may be written

as

$$\sum_{\mu=0}^{M}\sum_{\nu=0}^{Q}\beta_{\mu\nu}\alpha_{\mu\tau(\theta-\nu)}=0 \qquad (\tau=1,\dots,M\,,\quad \theta=0,\dots,D+Q)\,.$$

There are $M(D+Q+1)$ equations and $(M+1)(Q+1)$ unknowns $\beta_{\mu\nu}$. For a nontrivial solution to exist we require

$$(M+1)(Q+1)>M(D+Q+1)$$

and so we may take $Q=MD$. Consequently

$$\sum_{\mu=0}^{M}p_{\mu\tau}(z)h_\mu(z)=0 \qquad (\tau=1,\dots,M)$$

and the assertions of (3) are now proved.

(4)

$$\operatorname{ord}_{z=0}\psi(z)\le\frac{M}{b-1}\left(b^M r+\varsigma\frac{b^M-1}{b-1}\right).$$

Proof. Choose $m\ge 0$ minimal with $h_m(z)\ne 0$. Then $h_m(z)\psi(z^{b^m})=-\sum_{\mu\ge m+1}h_\mu(z)\psi(z^{b^\mu})$. Let $T=\operatorname{ord}_{z=0}\psi(z)<\infty$. Then

$$\operatorname{ord}_{z=0}(\text{left-hand side})\le MD+b^mT\,,$$

$$\operatorname{ord}_{z=0}(\text{right-hand side})\ge b^{m+1}T\,,$$

so that

$$T\le MD/(b^m(b-1))\le\frac{M}{b-1}\left(b^M r+\varsigma\frac{b^M-1}{b-1}\right).$$

The lemma is now proved.

THEOREM 3.1. *Let $b\ge 2$ be an integer, L be a field and $f_1(z),\dots,f_m(z)$ be elements of $L[[z]]$ algebraically independent over $L(z)$ that satisfy*

$$\underline{f}(z^b)=\underline{f}(z)\underline{\underline{A}}(z)+\underline{B}(z)$$

where

$$\underline{f}(z)=(f_1(z),\dots,f_m(z))\,,$$

$$\underline{\underline{A}}(z)=[a_{ij}(z)]_{1\le j\le m}\in M_{mm}(L(z))\,,$$

$$\underline{B}(z)=[b_j(z)]_{1\le j\le m}\in M_{1m}(L(z))\,.$$

Let $0\ne T(z)\in L[z]$ be such that

$$T(z)a_{ij}(z)\in L[z]\,,\qquad T(z)b_i(z)\in L[z]$$

with all of these m^2+m+1 polynomials being either zero or having degree at most $q \geq 1$. Let $0 \neq P \in L[z, x_1, \ldots, x_m]$ be such that

$$\deg_z P \leq d_0\,, \quad \deg_{x_i} P \leq d_i \qquad (i = 1, \ldots, m)\,.$$

Let $\varphi(z) = P(z, f_1(z), \ldots, f_m(z))$. Then

$$\operatorname{ord} P = \operatorname{ord}_{z=0} \varphi(z) \leq qM^2 b^M (d_0 + 1)$$

where $M = (d_1+1)\ldots(d_m+1)$.

Proof. It is enough to show that the conclusion of the theorem holds if we suppose that $f_1(z), \ldots, f_m(z)$ are no longer algebraically independent over $L(z)$ but we assume in its place that

$$\varphi(z) = P(z, f_1(z), \ldots, f_m(z)) \neq 0\,.$$

To establish the theorem in this form we may even suppose that $\underline{B}(z) = 0$. This is because if we define $f_0(z) = 1$ then

$$[f_0(z^b), \ldots, f_m(z^b)] = [f_0(z), \ldots, f_m(z)] \begin{bmatrix} 1 & b_1(z) & \ldots & b_m(z) \\ 0 & a_{11}(z) & \ldots & a_{1m}(z) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{m1}(z) & \ldots & a_{mm}(z) \end{bmatrix}$$

and if we define $Q \in L[z, x_0, \ldots, x_m]$ by

$$Q(z, x_0, \ldots, x_m) = P(z, x_1, \ldots, x_m)$$

then

$$Q(z, f_0(z), \ldots, f_m(z)) = P(z, f_1(z), \ldots, f_m(z)) = \varphi(z) \neq 0$$

and therefore one may apply the above quoted result with $(d_0, 0, d_1, \ldots, d_m)$ in place of $(d_0, d_1, \ldots, d_m)$ to obtain the desired conclusion.

The above remarks show that it is enough to prove the following assertion. Let $f_1(z), \ldots, f_m(z)$ be elements of $L[[z]]$ that satisfy

$$f_j(z^b) = \sum_{i=1}^{m} a_{ij}(x) f_i(z) \qquad (i = 1, \ldots, m)\,.$$

Let $0 \neq T(z) \in L[z]$ be such that

$$T(z) a_{ij}(z) \in L[z]$$

with all of these m^2+1 polynomials being either zero or having degree at most $q \geq 1$. Let $P \in L[z, x_1, \ldots, x_m]$ be such that

$$\deg_z P \leq d_0\,, \quad \deg_{x_i} P \leq d_i \qquad (i = 1, \ldots, m)$$

and

$$0 \neq \varphi(z) = P(z, f_1(z), \ldots, f_m(z)) .$$

Then

$$\operatorname{ord}_{z=0} \varphi(z) \leq qM^2 b^M (d_0 + 1)$$

where $M = (d_1 + 1) \ldots (d_m + 1)$.

To prove this last assertion put $G_{i_1 \ldots i_m}(z) = f_1(z)^{i_1} \ldots f_m(z)^{i_m}$ for $0 \leq i_\mu \leq d_\mu$ $(\mu = 1, \ldots, m)$. Let

$$P(z, x_1, \ldots, x_m) = \sum_{i_1=0}^{d_1} \cdots \sum_{i_m=0}^{d_m} q_{i_1 \ldots i_m}(z) x_1^{i_1} \ldots x_m^{i_m}$$

so that

$$\begin{aligned} 0 \neq \varphi(z) &= P(z, f_1(z), \ldots, f_m(z)) \\ &= \sum_{i_1=0}^{d_1} \cdots \sum_{i_m=0}^{d_m} q_{i_1 \ldots i_m}(z) G_{i_1 \ldots i_m}(z) . \end{aligned}$$

Now

$$\begin{aligned} G_{i_1 \ldots i_m}(z^b) &= \Big(\sum_{j_1=1}^{m} a_{j_1 1}(z) f_{j_1}(z) \Big)^{i_1} \cdots \Big(\sum_{j_m=1}^{m} a_{j_m m}(z) f_{j_m}(z) \Big)^{i_m} \\ &= \sum_{j_1=0}^{d_1} \cdots \sum_{j_m=0}^{d_m} \varsigma_{j_1 \ldots j_m i_1 \ldots i_m}(z) G_{j_1 \ldots j_m}(z) \end{aligned}$$

where the $\varsigma_{j_1 \ldots j_m i_1 \ldots i_m}(z)$ are rational functions in $L(z)$. Put

$$0 \neq f(z) = T(z)^{d_1 + \ldots + d_m} \in L[z] .$$

Then each $f(z)\varsigma_{j_1 \ldots j_m i_1 \ldots i_m}(z)$ is in $L[z]$ and is either zero or of degree at most $\varsigma = q(d_1 + \ldots + d_m)$. Also $f(z)$ is of degree at most ς. Now apply Lemma 3.1.1 with

$$\begin{aligned} M &= (d_1 + 1) \ldots (d_m + 1) , \\ r &= d_0 , \\ \varsigma &= q(d_1 + \ldots + d_m) . \end{aligned}$$

Then

$$\operatorname{ord}_{z=0} \varphi(z) \leq M b^M (r + \varsigma) ,$$

from which the desired inequality follows without difficulty.

§ 3.2. Resultants

This section deals with the basic properties of resultants. I have included proofs of all the properties needed with the exception of Lemma 3.2.1 of which a reference is given. These properties will be used without further comment in § 3.3.

DEFINITION. Let R be a ring. Define for f, g

$$f(x) = \sum_{i=0}^{m} a_i x^{m-i}, \qquad g(x) = \sum_{i=0}^{n} b_i x^{n-i},$$

the *resultant*

$$\text{(3.2.1)} \qquad \operatorname{Res}(f,g) = \det \begin{bmatrix} a_0 & \dots & a_{m-1} & a_m & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & a_0 & \dots & a_{m-1} & a_m & \\ b_0 & \dots & b_n & & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & b_0 & & \dots & b_n \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

The resultant is an element of R.

LEMMA 3.2.1. *Suppose that the ring R is an integral domain. Let K be the quotient field of R and $\overline{K}$ the algebraic closure of K. Then for*

$$f(x) = a\prod_{i=1}^{m}(x - t_i), \quad g(x) = b\prod_{j=1}^{n}(x - s_j), \qquad t_i, s_j \in \overline{K},$$

we have

$$\operatorname{Res}(f,g) = a^n b^m \prod_{i,j}(t_i - s_j).$$

Proof. For a proof of the above lemma see [Wae], Volume I.

LEMMA 3.2.2. *Let R be a ring and x an indeterminate over R. Let $f, g \in R[x]$ with $\deg_x f = m \geq 1$, $\deg_x g = n \geq 1$. Then*

$$\operatorname{Res}(f,g) = A(x)f(x) + B(x)g(x)$$

with $A(x)$, $B(x)$ in $R[x]$.

Proof. Apply column operations to the definition (3.2.1). If C_i denotes

column i then replace C_{m+n} by $C_1 x^{m+n-1} + \ldots + C_{m+n}$ to obtain

$$\operatorname{Res}(f,g) = \det \begin{bmatrix} a_0 & \ldots & a_{m-1} & a_m & x^{n-1}f(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & a_0 & \ldots & a_{m-1} & f(x) \\ b_0 & \ldots & b_n & & x^{m-1}g(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & b_0 & \ldots & g(x) \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

$$= A(x)f(x) + B(x)g(x)$$

with $A(x)$, $B(x)$ in $R[x]$.

LEMMA 3.2.3. *Let $R = R_0[y_1, \ldots, y_t]$ where R_0 is an integral domain and let $1 \le i_1 < \ldots < i_s \le t$. Let $f, g \in R[x]$ with $\deg_x f \ge 1$, $\deg_x g \ge 1$. Suppose that* $\operatorname{Res}(f,g) \ne 0$. *Then*

$$\deg_{y_{i_1} \ldots y_{i_s}} \operatorname{Res}(f,g) \le (\deg_x g)(\deg_{y_{i_1} \ldots y_{i_s}} f) + (\deg_x f)(\deg_{y_{i_1} \ldots y_{i_s}} g)\,.$$

Proof. This follows immediately from the definition (3.2.1).

GAUSS' LEMMA. *Let R be a UFD (unique factorization domain). A polynomial $0 \ne f \in R[x_1, \ldots, x_n]$ is called primitive if 1 is GCD (greatest common divisor) of its coefficients. If $f, g \in R[x_1, \ldots, x_n]$ are primitive then $fg \in R[x_1, \ldots, x_n]$ is also primitive.*

COROLLARY 1. *Let K be the quotient field of R and let f, $g \in R[x_1, \ldots, x_n]$ with f primitive and $f \mid g$ in $K[x_1, \ldots, x_n]$. Then $f \mid g$ in $R[x_1, \ldots, x_n]$.*

COROLLARY 2. *Let K be the quotient field of R and let $f \in R[x_1, \ldots, x_n]$. Then the following conditions are equivalent:*

(1) *f is irreducible in $R[x_1, \ldots, x_n]$ and $\deg_{x_1 \ldots x_n} f \ge 1$.*

(2) *f is primitive and irreducible in $K[x_1, \ldots, x_n]$.*

Proof. Apply Corollary 1.

COROLLARY 3. *Let K be the quotient field of R and let $0 \ne a \in R$ be a GCD (greatest common divisor) of f, g in $R[x_1, \ldots, x_n]$. Then 1 is a GCD of f, g in $K[x_1, \ldots, x_n]$.*

Proof. If not let $r|f$, $r|g$ in $K[x_1, \ldots, x_n]$ with r primitive and of positive degree. By Corollary 1 we obtain a contradiction.

LEMMA 3.2.4. *Let R be a UFD and let $f, g \in R[x]$ with $\deg_x f = m \ge 1$, $\deg_x = n \ge 1$. Then the following conditions are equivalent:*

(1) $\operatorname{Res}(f,g) \ne 0$.

(2) *f, g have no common factor in $R[x]$ of positive degree.*

Proof. Let K be the quotient field of R and $\overline{K}$ the algebraic closure of K. Then for $f = a\prod_{i=1}^{m}(x - t_i)$, $g = b\prod_{j=1}^{n}(x - s_j)$ we have

$$\operatorname{Res}(f,g) = a^n b^m \prod_{i,j}(t_i - s_j)\,.$$

(1) $\Rightarrow$ (2). If f, g have a common factor in $R[x]$, say $k(x)$ with $\deg_x k \geq 1$, let $\xi \in \overline{K}$ with $k(\xi) = 0$. Since $k|f$ and $k|g$ in $R[x]$ and so also in $\overline{K}[x]$, $f(\xi) = g(\xi) = 0$ must hold. From $\operatorname{Res}(f,g) = A(x)f(x) + B(x)g(x)$ and $\operatorname{Res}(f,g) \in R$ it follows that $\operatorname{Res}(f,g) = 0$, which contradicts (1).

(2) $\Rightarrow$ (1). It assumed that $0 \neq \alpha \in R$ is a GCD of f, g in $R[x]$. By Corollary 3 of Gauss' lemma we see that 1 is a GCD of f, g in $K[x]$ and so $(f,g) = (1)$ in $K[x]$ since $K[x]$ is a PID (principal ideal domain). Consequently $(f,g) = (1)$ in $\overline{K}[x]$ and so f, g have no common zero in $\overline{K}$. This means that $t_i \neq s_j$ for all i, j so that $\operatorname{Res}(f,g) \neq 0$ in $\overline{K}$ and so also in R.

LEMMA 3.2.5. *Let L be a field, $k \geq 2$ be an integer and let $P, Q \in L[x_1, \ldots, x_k]$ be such that*

$$\deg_{x_1,\ldots,x_k} P \leq M\,, \qquad \deg_{x_1,\ldots,x_k} Q \leq N\,,$$

$$\deg_{x_k} P = m \geq 1\,, \qquad \deg_{x_k} Q = n \geq 1\,.$$

Suppose that $R = \operatorname{Res}_{x_k}(P,Q) \neq 0$. Then

$$\deg_{x_1,\ldots,x_{k-1}} R \leq MN\,.$$

Proof. Write

$$P = \sum_{i=1}^{m} P_i(x_1,\ldots,x_{k-1})x_k^{m-i}\,, \qquad Q = \sum_{i=0}^{n} Q_i(x_1,\ldots,x_{k-1})x_k^{n-i}$$

where the P_i, Q_i are in $L[x_1,\ldots,x_{k-1}]$. Introduce an additional indeterminate y and consider

$$R(y) = \operatorname{Res}_{x_k}(P(x_1,\ldots,x_{k-1},yx_k), Q(x_1,\ldots,x_{k-1},yx_k))$$

$$= \det \begin{bmatrix} y^m P_0 & \ldots & y^1 P_{m-1} & y^0 P_m & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & y^m P_0 & \ldots & y^1 P_{m-1} & y^0 P_m & \\ y^n Q_0 & \ldots & y^0 Q_n & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & y^n Q_0 & & \ldots & y^0 Q_n \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

Now multiply the rows of the above matrix by

$$y^{n-1}, y^{n-2}, \ldots, y, 1, y^{m-1}, \ldots, 1$$

and divide the columns by

$$y^{m+n-1}, y^{m+n-2}, \ldots, 1.$$

We obtain $y^{-mn}R(y) = R(1)$ after a little simplification. But

$$\deg_{x_1,\ldots,x_{k-1},y} P(x_1, \ldots, x_{k-1}, yx_k) \leq M,$$

$$\deg_{x_1,\ldots,x_{k-1},y} Q(x_1, \ldots, x_{k-1}, yx_k) \leq N,$$

and so by Corollary 2 of Theorem 3.2.2

$$\deg_{x_1,\ldots,x_{k-1},y} R(y) \leq nM + mN.$$

Consequently

$$\begin{aligned}\deg_{x_1,\ldots,x_{k-1}} R(1) &= \deg_{x_1,\ldots,x_{k-1},y} (y^{-mn}R(y)) \leq -mn + nM + mN \\ &= MN - (M-m)(N-n) \leq MN.\end{aligned}$$

§ 3.3. The upper triangular case

As remarked in § 3.1, this section will be devoted to a proof of Theorem 3.3.

THEOREM 3.3. *Let $b \geq 2$ be an integer, L be a field and $f_1(z), \ldots, f_m(z)$ be elements of $L[[z]]$ algebraically independent over $L(z)$ that satisfy*

$$\underline{f}(z^b) = \underline{f}(z)\underline{\underline{A}}(z) + \underline{B}(z) \tag{3.3.1}$$

where

$$\underline{\underline{A}}(z) = [a_{ij}(z)]_{1\leq i,j\leq m} \in M_{mm}(L(z)) \quad \textit{is upper triangular,}$$

$$\underline{B}(z) = [b_j(z)]_{1\leq j\leq m} \in M_{1m}(L(z)).$$

Let $0 \neq T(z) \in L[z]$ be such that

$$T(z)a_{ij}(z) \in L[z], \quad T(z)b_i(z) \in L[z],$$

with all of these $m^2 + m + 1$ polynomials being either zero or having degree at most $q \geq 0$. Then for any $0 \neq P \in L[z, x_1, \ldots, x_m]$ with $\deg_z P \leq D$, $\deg_{x_1,\ldots,x_m} P \leq N$ $(N \geq 1)$ we have

$$\operatorname{ord} P \leq (b+2)^m N^{2^m-1}(D + mqN).$$

Proof. We will use the notation

$$F_m(N, D) = (b+2)^m N^{2^m - 1}(D + mqN)$$

for $N \geq 1$, $D \geq 0$ and m a nonnegative integer. Here $q \geq 0$, $b \geq 2$ are regarded as fixed. For brevity we will also use the notation

$$\deg_{\underline{x}} P = \deg_{x_1, \ldots, x_m} P$$

for $0 \neq P \in L[z, x_1, \ldots, x_m]$.

Notice that equations (3.3.1) may be written equivalently as

$$f_j(z^b) = \sum_{i=1}^{j} f_i(z) a_{ij}(z) + b_j(z) \qquad (j = 1, \ldots, m)\,. \tag{3.3.2}$$

The proof of Theorem 3.3 is by induction on m, $m \geq 1$. We suppose that $m \geq 1$ and that in the case $m \geq 2$ the conclusion of the Theorem holds with $m-1$ in place of m. We use the convention that $L[z, x_1, \ldots, x_{m-1}] = L[z]$ if $m = 1$. Let $D \geq 0$, $N \geq 1$ and suppose that $0 \neq P \in L[z, x_1, \ldots, x_m]$ satisfies

$$\deg_z \leq D\,, \qquad \deg_{\underline{x}} P \leq N\,.$$

It will be proved that $\operatorname{ord} P \leq F_m(N, D)$ which will complete the inductive proof of the Theorem.

Now the inductive hypothesis implies that if $m \geq 1$, $D_1 \geq 0$, $N_1 \geq 1$ and $0 \neq S \in L[z, x_1, \ldots, x_{m-1}]$ satisfies

$$\deg_z S \leq D_1\,, \quad \deg_{\underline{x}} S \leq N_1 \quad (\text{in the case } m \geq 2)$$

then

$$\operatorname{ord} S \leq F_{m-1}(N_1, D_1)\,. \tag{3.3.3}$$

Indeed, for $m = 1$ we have $0 \neq S \in L[z]$ so that

$$\operatorname{ord} S \leq D_1 = F_0(N_1, D_1)$$

and for $m \geq 2$ we use the inductive hypothesis. Notice that in the case $m \geq 2$ the equations (3.3.2) still hold but with $m-1$ in place of m.

As a first step we will show the following. Given a polynomial $0 \neq R$ in $L[z, x_1, \ldots, x_m]$ there exists $0 \neq R^*$ in $L[z, x_1, \ldots, x_m]$ with

$$\begin{aligned} &\deg_{\underline{x}} R^* \leq \deg_{\underline{x}} R\,, \\ &\deg_z R^* \leq b \deg_z R + q \deg_{\underline{x}} R\,, \\ &\operatorname{ord} R^* \geq b \operatorname{ord} R\,. \end{aligned} \tag{3.3.4}$$

Indeed, we put $l = \deg_{\underline{x}} R$ and define

$$R^*(z, x_1, \ldots, x_m) = R_{(1)}(z^b, S_0(z), S_1(z, \underline{x}), \ldots, S_m(z, \underline{x}))$$

where

$$R_{(1)}(z, X_0, \ldots, X_m) = X_0^l R(z, X_1/X_0, \ldots, X_m/X_0)$$

and

$$S_0(z) = T(z)\,,$$

$$S_j(z, \underline{x}) = \sum_{i=1}^{j} x_i T(z) a_{ij}(z) + T(z) b_j(z) \quad (j = 1, \ldots, m)\,.$$

Notice that

$$R^*(z, f_1(z), \ldots, f_m(z)) = T(z)^l R(z^b, f_1(z^b), \ldots, f_m(z^b))\,.$$

All the properties (3.3.4) are now immediate.

We recall that the polynomial $0 \neq P$ in $L[z, x_1, \ldots, x_m]$ satisfies $\deg_z P \leq D$, $\deg_{\underline{x}} P \leq N$ with $D \geq 0$, $N \geq 1$ and that we wish to prove the inequality

$$(*) \qquad \operatorname{ord} P \leq F_m(N, D)$$

to complete the inductive proof of the theorem.

If $\deg_{x_m} P = 0$ then $0 \neq P \in L[z, x_1, \ldots, x_{m-1}]$ and the required inequality follows from (3.3.3). It may therefore be supposed that $\deg_{x_m} P \geq 1$.

Let $P = aP_1 \ldots P_g$ where $P_1, \ldots, P_g \in L[z, x_1, \ldots, x_m]$ are irreducible of positive degree in x_m and a is in $L[z, x_1, \ldots, x_{m-1}]$. Now

$$\sum \deg_{\underline{x}} P_i = \deg_{\underline{x}} P - \deg_{\underline{x}} a \leq \deg_{\underline{x}} P\,,$$

$$\sum \operatorname{ord} P_i = \operatorname{ord} P - \operatorname{ord} a \geq \operatorname{ord} P - F_{m-1}(N, D) \qquad \text{(using (3.3.3))}\,.$$

We may suppose that

$$\operatorname{ord} P \geq (b+2) F_{m-1}(N, D)\,,$$

since otherwise

$$\operatorname{ord} P < (b+2) F_{m-1}(N, D) \leq F_m(N, D)\,.$$

Then

$$\sum \deg_{\underline{x}} P_i \leq \deg_{\underline{x}} P\,, \qquad \sum \operatorname{ord} P_i \geq \frac{b+1}{b+2} \operatorname{ord} P$$

and so there exists an index i, $1 \le i \le g$ with

$$\operatorname{ord} P_i \ge \frac{b+1}{b+2} \frac{\deg_{\underline{x}} P_i}{\deg_{\underline{x}} P} \operatorname{ord} P .$$

We relabel the indices so that $P_i = P_1$ and define

$$N_1 = \deg_{\underline{x}} P_1 \ge 1 , \quad D_1 = \deg_z P_1$$

so that in particular $1 \le N_1 \le N$, $0 \le D_1 \le D$. Further

$$\operatorname{ord} P \le \frac{b+2}{b+1} \frac{N}{N_1} \operatorname{ord} P_1 .$$

We now distinguish three cases:

Case I. $R = P_1$ satisfies $R \mid R^*$. In this case

$$\operatorname{ord} R \le \frac{q}{b-1} N_1 + \frac{b}{b-1} D_1$$

since $\deg_{\underline{x}} R^* \le \deg_{\underline{x}} R$ so that $R^* = R\varphi$ with $\varphi \in L[z]$ and therefore if $T = \operatorname{ord} R$ then

$$bT \le \operatorname{ord} R^* \le T + \deg_z R^* \le T + bD_1 + qN_1 ,$$

which gives the above estimate. Consequently

$$\operatorname{ord} P \le \frac{b+2}{b+1} \frac{N}{N_1} \frac{1}{b-1} (bD_1 + qN_1) \le (b+2)(D+qN) ,$$

which implies the required estimate.

Case II. $R = P_1$ satisfies $R^* \in L[z, x_1, \ldots, x_{m-1}]$. In this case

$$\begin{aligned} \operatorname{ord} R \le \operatorname{ord} R^* &\le F_{m-1}(N_1, bD_1 + qN_1) \\ &= (b+2)^{m-1} N_1^{2^{m-1}-1} (bD_1 + mqN_1) \end{aligned}$$

and so

$$\begin{aligned} \operatorname{ord} P &\le \frac{b+2}{b+1} \frac{N}{N_1} (b+2)^{m-1} N_1^{2^{m-1}-1} (bD_1 + mqN_1) \\ &\le (b+2)^m N^{2^{m-1}-1} (D + mqN) , \end{aligned}$$

which again implies the required estimate.

Case III. $R = P_1$ satisfies $R \nmid R^*$ and $R^* \notin L[z, x_1, \ldots, x_{m-1}]$. In this case, $\deg_{x_m} R \ge 1$, $\deg_{x_m} R^* \ge 1$, and R, R^* are relatively prime in $L[z, x_1, \ldots, x_m]$ so that their resultant $Q = \operatorname{Res}_{x_m}(R, R^*)$ is nonzero as an element of $L[z, x_1, \ldots, x_m]$. We know that $Q = AR + BR^*$ with A, B in

$L[z, x_1, \ldots, x_m]$ so that

$$\operatorname{ord} Q \geq \min(\operatorname{ord} R, \operatorname{ord} R^*) \geq \operatorname{ord} R\,.$$

Also

$$\begin{aligned}\deg_z Q &\leq (\deg_z R)(\deg_{x_m} R^*) + (\deg_z R^*)(\deg_{x_m} R)\\ &\leq D_1 N_1 + (bD_1 + N_1)N_1 = (b+1)D_1 N_1 + N_1^2\,,\end{aligned}$$

and in the case $m \geq 2$,

$$\deg_{x_1,\ldots,x_{m-1}} Q \leq (\deg_{\underline{x}} R)(\deg_{\underline{x}} R^*) \leq N_1^2\,.$$

Therefore

$$\begin{aligned}\operatorname{ord} Q &\leq F_{m-1}(N_1^2, (b+1)D_1 N_1 + qN_1^2)\\ &\leq (b+2)^{m-1} N_1^{2^m-1}((b+1)D_1 + mqN_1)\,.\end{aligned}$$

Consequently,

$$\begin{aligned}\operatorname{ord} P &\leq \frac{b+2}{b+1}\,\frac{N}{N_1}(\operatorname{ord} R) \leq \frac{b+2}{b+1}\,\frac{N}{N_1}(\operatorname{ord} R)\\ &\leq \frac{b+2}{b+1}\,\frac{N}{N_1}(\operatorname{ord} Q) \leq F_m(N, D)\,,\end{aligned}$$

which is the required inequality.

This completes the inductive proof of $(*)$ and so Theorem 3.3 is proved.

IV. Preliminaries

§ 4.1. Ideals

The work of Yu. V. Nesterenko is ideally suited for estimating the transcendence measure of certain sets of numbers. In his paper [Ne1] he introduces the basic definitions and applies these in different contexts ([Ne1]–[Ne4]). It will be seen from the applications he makes, some of which have been mentioned in Chapter I, that his theory is indeed brilliant and will very likely play an important role in future algebraic independence investigations.

Let R be a PID (principal ideal domain) whose quotient field has characteristic zero. For the most part Nesterenko applies his theory to $R = \mathbb{Z}$ (the ring of rational integers) and to $R = \mathbb{C}[z]$. Let I be a homogeneous ideal of the ring $R[X] = R[X_0, \ldots, X_m]$ where $m \geq 1$ is an integer. We denote the height of the ideal I by $h(I)$; $h(\mathfrak{p}) = \sup\{d : \ 0 = \mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_d = \mathfrak{p}$ where the $\mathfrak{p}_\mu$ are prime ideals of $R[X]\}$ for $\mathfrak{p}$ a prime ideal and $h(I) =$

$\inf\{h(\mathfrak{p}) : \mathfrak{p} \subseteq I, \mathfrak{p}$ is a homogeneous prime ideal of $R[X]\}$. We take $h(I) = m+1$ if I is an irrelevant ideal, that is,

$$(X_0, \ldots, X_m) \subseteq \sqrt{I}$$

where $\sqrt{I}$ denotes the radical of I.

An ideal I is said to be *unmixed* if all of its associated prime ideals have the same height $h(I)$.

Let r be an integer satisfying $1 \leq r \leq m$ and let U_{ij} $(1 \leq i \leq r, 0 \leq j \leq m)$ be variables algebraically independent over the quotient field of $R[X]$. We use the notation

$$\overline{U}_i = (U_{i0}, \ldots, U_{im}) \qquad (1 \leq i \leq r).$$

Put $L_i(X) = \sum_{j=0}^{m} U_{ij}X_j$ $(1 \leq i \leq r)$ and introduce the ideals $(I, L_1, \ldots, L_r)$ and $\chi = (X_0, \ldots, X_m)$ of $R[X, \overline{U}_1, \ldots, \overline{U}_r] = R[X_0, \ldots, X_m, U_{10}, \ldots, U_{rm}]$. Now define $\bar{I} = \bar{I}(r)$ to be the ideal of $R[U] = R[U_{10}, \ldots, U_{rm}]$ consisting of those polynomials G for which there exists an M such that

$$G\chi^M \subseteq (I, L_1, \ldots, L_r).$$

We begin by stating some properties of the ideal $\bar{I}$.

LEMMA 4.1.1. *Let R be a PID and let I be an unmixed homogeneous ideal of $R[X] = R[X_0, \ldots, X_m]$ having height $h(I)$ and $I \cap R = (0)$. Let r be an integer with $1 \leq r \leq m$.*

(1) *If $h(I) \leq m - r$ then $\bar{I}(r) = (0)$.*

(2) *If $h(I) = m - r + 1$ then $\bar{I}(r)$ is a nonzero principal ideal in the ring $R[\overline{U}_1, \ldots, \overline{U}_r]$.*

LEMMA 4.1.2. *Let R be a PID whose quotient field has characteristic zero. Let I be a homogeneous ideal of $R[X] = R[X_0, \ldots, X_m]$ having height $h(I)$ and $I \cap R = (0)$. We suppose that $r = m+1-h(I)$ satisfies $1 \leq r \leq m$.*

(1) *If I is a prime ideal then $\bar{I}(r)$ is a principal prime ideal.*

(2) *If I is a $\mathfrak{p}$-primary ideal then $\bar{I}(r)$ is a principal $\overline{\mathfrak{p}}(r)$-primary ideal. The exponent of $\bar{I}$ in $\overline{\mathfrak{p}}(r)$ is equal to the exponent of I in $\mathfrak{p}$.*

(3) *If I is unmixed, $I = I_1 \cap \ldots \cap I_s \cap \ldots \cap I_t$ where $I_l \cap R = (0)$ for $1 \leq l \leq s$ and $I_l \cap R \neq (0)$ for $l > s$ then*

$$\bar{I} = \bar{I}_1 \cap \ldots \cap \bar{I}_t$$

and this decomposition is irreducible in $\bar{I}_l$ $(1 \leq l \leq s)$. Let $\mathfrak{p}_l = \sqrt{I_l}$ and let K_l be the exponent of I_l $(1 \leq l \leq s)$. Then k_l is the exponent of $\bar{I}_l$ $(1 \leq l \leq s)$. Write $I_{s+1} \cap \ldots \cap I_t \cap R = (a) \subseteq R$ with $a \neq 0$. Then if $\overline{\mathfrak{p}}_l = (F_l)$ $(1 \leq l \leq s)$ one has $\bar{I} = (F)$ where $F = aF_1^{k_1} \ldots F_s^{k_s}$.

Lemma 4.1.1 follows from Lemma 5 and Proposition 2 of [Ne1]. Lemma 4.1.2 follows from Lemma 5 and Propositions 1–3 and the corollary of Proposition 3 of [Ne1].

Once these results are established we make some further definitions. In what follows we will specialize by taking $R = \mathbf{Z}$ to be the ring of rational integers. We know that the polynomial F in $\mathbf{Z}[\overline{U}_1, \ldots, \overline{U}_r]$ such that $\bar{I}(r) = (F)$ is determined up to sign and so one may define $H(I)$ to be the maximum of the absolute values of the coefficients of F. We also know that F is homogeneous in each system of variables $\overline{U}_1 \ldots, \overline{U}_r$ and since the definition of $\bar{I}(r)$ is symmetric in the $\overline{U}_k$ it follows that the homogeneous degree of F in the variables $\overline{U}_k$ is the same for every k. We may therefore define

$$N(I) = \deg_{\overline{U}_1} F = \ldots = \deg_{\overline{U}_r} F\,.$$

Now introduce additional indeterminates $S_{jk}^{(i)}$ $(j, k = 0, \ldots, m;\ i = 1, \ldots, r)$ that have no algebraic relation over $\mathbf{Z}[X, \overline{U}_1, \ldots, \overline{U}_r]$ except for the skew symetry

$$S_{jk}^{(i)} + S_{kj}^{(i)} = 0 \qquad (j, k = 0, 1, \ldots, m\,; \quad i = 1, \ldots, r)\,.$$

We write $S^{(i)} = [S_{jk}^{(i)}]_{0 \le j,k \le m}$ so that $S^{(i)}$ is a skew-symmetric matrix and put

$$\mathbf{C}[S] = \mathbf{C}[S^{(1)}, \ldots, S^{(r)}]\,.$$

We next define a ring homomorphism

$$K : \mathbf{Z}[\overline{U}_1, \ldots, \overline{U}_r] \to \mathbf{C}[S^{(1)}, \ldots, S^{(r)}]\,.$$

It will depend on a fixed nonzero vector $\overline{\omega} = (\omega_0, \ldots, \omega_m) \in \mathbf{C}^{m+1}$ and is defined by

$$K(\overline{U}_i) = S^{(i)}\overline{\omega}\,, \quad \text{that is}\,, \quad K(U_{ij}) = \sum_{k=0}^{m} S_{jk}^{(i)}\omega_k\,.$$

If I is a homogeneous ideal of $\mathbf{Z}[X_0, \ldots, X_m]$ with $h(I) = m + 1 - r$ and $\bar{I}(r) = (F)$ define $|I(\overline{\omega})|$ by

$$|I(\overline{\omega})| = |\overline{\omega}|^{-rN(I)} H(K(F))$$

where $|\overline{\omega}| = \max_{0 \le i \le m} |\omega_i|$ and $H(K(F))$ is the maximum of the absolute values of the coefficients of the polynomial $K(F)$ (the height of $K(F)$). Notice that $K(F)$ is homogeneous of degree $N(I)$ in each set of variables $S_{jk}^{(i)}$ for $1 \le i \le r$ so that $K(\alpha F) = \alpha^{rN(I)} K(F)$ and also

$$|I(\alpha\overline{\omega})| = |I(\overline{\omega})|$$

for $0 \ne \alpha \in \mathbf{C}$.

The following results due essentially to Nesterenko will now be stated.

LEMMA 4.1.3. *Let $I = (P)$ be the principal ideal of $\mathbb{Z}[X_0, \ldots, X_m] = \mathbb{Z}[X]$ which is generated by the homogeneous polynomial P. Then*

$$N(I) = \deg P\,, \quad \ln H(I) \le \ln H(P) + m^2 \deg P\,,$$

$$|I(\overline{\omega})| \le |P(\overline{\omega})||\overline{\omega}|^{-\deg P}(m+1)^{2m \deg P}\,.$$

LEMMA 4.1.4. *Suppose that I is an unmixed homogeneous ideal in $\mathbb{Z}[X_0, \ldots, X_m]$ with $r = m + 1 - h(I)$ satisfying $1 \le r \le m$. Let $I = I_1 \cap \ldots \cap I_s \cap \ldots \cap I_t$ be its irreducible primary decomposition in which for $l \le s$ we have $I_l \cap \mathbb{Z} = (0)$, $0 \ne I_{s+1} \cap \ldots \cap I_t = (b) \subseteq \mathbb{Z}$. Put $\mathfrak{p}_l = \sqrt{I_l}$ for $l \le s$ and let k_l be the exponent of the ideal I_l. Then*

(1) $\sum_{l=1}^{s} k_l N(\mathfrak{p}_l) = N(I)$,
(2) $\ln|b| + \sum_{l=1}^{s} k_l \ln H(\mathfrak{p}_l) \le \ln H(I) + m^2 N(I)$,
(3) $\ln|b| + \sum_{l=1}^{s} k_l \ln|\mathfrak{p}_l(\overline{\omega})| \le \ln|I(\overline{\omega})| + m^3 N(I)$.

If $s = 0$ the terms involving $\mathfrak{p}_l$ do not occur in (1), (2), (3) *while if $s = t$ the terms involving $\ln|b|$ do not occur in* (2), (3).

Lemmas 4.1.3 and 4.1.4 are proved as Propositions 1 and 2 respectively in [Ne3]. They are restated as Propositions 1 and 2 in [Ne4].

We next state three more lemmas due to Nesterenko:

LEMMA 4.1.5. *Suppose that $\overline{\omega} \in \mathbb{C}^{m+1}$, $\overline{\omega} \ne 0$, $\mathfrak{p}$ is a homogeneous prime ideal of $\mathbb{Z}[X]$, $\mathfrak{p} \cap \mathbb{Z} = (0)$, $1 \le h(\mathfrak{p}) \le m$. Let Q be a homogeneous polynomial in $\mathbb{Z}[X]$ of degree ≥ 1, $Q \in \mathfrak{p}$. If $r = m + 1 - h(\mathfrak{p}) \ge 2$ then there exists an unmixed homogeneous ideal $I \subseteq \mathbb{Z}[X]$ whose zeros coincide with the zeros of the ideal $(\mathfrak{p}, Q)$ (in P^m, projective space of dimension m) and for which $h(I) = m - r + 2$ and*

(1) $N(I) \le N(\mathfrak{p}) \deg Q$,
(2) $\ln H(I) \le (\deg Q) \ln H(\mathfrak{p}) + N(\mathfrak{p}) \ln H(Q) + m(r+1) N(\mathfrak{p}) \deg Q$,
(3) $\ln|I(\overline{\omega})| \le \ln \max(|\mathfrak{p}(\overline{\omega})|, |Q(\overline{\omega})||\overline{\omega}|^{-\deg Q})$
$\qquad + \deg Q \ln H(\mathfrak{p}) + N(\mathfrak{p}) \ln H(Q) + 6m^2 N(\mathfrak{p}) \deg Q$.

If $h(\mathfrak{p}) = m$ (equivalently $r = 1$), then the right side of the inequality in (3) *is nonnegative.*

Notation. Define $\|\overline{\theta} - \overline{\psi}\|$ for two nonzero vectors $\overline{\theta} = (\theta_0, \ldots, \theta_n)$, $\overline{\psi} = (\psi_0, \ldots, \psi_m)$ in $\mathbb{C}^{m+1}$ by

$$\|\overline{\theta} - \overline{\psi}\| = |\overline{\theta}|^{-1}|\overline{\psi}|^{-1} \max_{0 \le i,j \le m} |\theta_i \psi_j - \theta_j \psi_i|\,.$$

Here $|\overline{\theta}| = \max_{0 \le i \le m} |\theta_i|$ and $|\overline{\psi}|$ is similarly defined.

LEMMA 4.1.6. *Suppose that* $Q \in \mathbb{Z}[X_0, \ldots, X_m]$, $\deg Q \geq 1$, Q *is a homogeneous polynomial. Suppose also that* $\mathfrak{p} \subseteq \mathbb{Z}[X_0, \ldots, X_m]$ *is a homogeneous prime ideal,* $\mathfrak{p} \cap \mathbb{Z} = (0)$, $r = m + 1 - h(\mathfrak{p}) \geq 1$, $\overline{\omega} = (\omega_0, \ldots, \omega_m) \in \mathbb{C}^{m+1}$, $|\overline{\omega}| \geq 1$,

$$|\mathfrak{p}(\overline{\omega})| \leq e^{-X}, \quad X > 0,$$

$$|Q(\overline{\omega})||\overline{\omega}|^{-\deg Q} \leq H(Q)^{-1}(\deg Q + 1)^{-(2m+2)},$$

and suppose finally that for some $\sigma \geq 1$ *the following equality holds:*

$$\min\left(X, \frac{\ln(1/\rho)}{2}\right) = -\sigma \ln\left(|Q(\overline{\omega})||\overline{\omega}|^{-\deg Q}\right)$$

where $\rho = \inf\left\{\|\overline{\omega} - \overline{\beta}\| : 0 \neq \overline{\beta} \in \mathbb{C}^{m+1} \text{ is zero of the ideal } \mathfrak{p}\right\}$ *satisfies* $\rho \leq 1$.

The following conclusions hold. First $Q \in \mathfrak{p}$. *Put*

$$Y = \frac{-X}{2\sigma} + \deg Q \ln H(\mathfrak{p}) + N(\mathfrak{p}) \ln H(Q) + 8m^2 N(\mathfrak{p}) \deg Q.$$

In the case $r \geq 2$ *let* I *be the ideal constructed in Lemma 4.1.5. Then*

$$\ln |I(\overline{\omega})| \leq Y.$$

In the case $r = 1$,

$$0 \leq Y.$$

LEMMA 4.1.7. *Suppose that* $I \subset \mathbb{Z}[X_0, \ldots, X_m]$ *is an unmixed homogeneous ideal,* $I \cap \mathbb{Z} = (0)$ *and* $r = m + 1 - h(I)$ *satisfies* $1 \leq r \leq m$. *For every nonzero vector* $\overline{\omega} \in \mathbb{C}^{m+1}$ *there exists a zero* $\overline{\beta} \in \mathbb{C}^{m+1}$, $\overline{\beta} \neq 0$, *of the ideal* I *such that*

$$N(I) \ln \|\overline{\omega} - \overline{\beta}\| \leq \frac{1}{r} \ln |I(\overline{\omega})| + 3m^3 N(I).$$

§ 4.2. Some lemmas

Lemmas 4.2.2 and 4.2.5 are needed for the proof of Lemma 5.1.1 of Chapter IV which is the main tool when combined with the results of § 4.1 in establishing the proposition of § 5.2. We recall the definition of the class $\mathcal{G}$ and the class $\mathcal{G}(c, C)$ made at the beginning of Chapter II.

LEMMA 4.2.1. *Suppose that* $f_1(z), \ldots, f_m(z)$ *are elements of* $\mathcal{G}(c, C)$ *analytic at* $z = 0$ *and let* $R(z, x_1, \ldots, x_m)$ *be a polynomial of degree at most* N *in* $x_1, \ldots, x_m$ *whose coefficients are rational integers of absolute values at most* H. *Here* N *and* H *are positive integers. Then* $\phi(z) =$

$R(z, f_1(z), \ldots, f_m(z))$ lies in

$$\mathcal{G}\left(2^N c^N \binom{N+m}{N} H, 2C\right)$$

and furthermore $c^N \phi(Cz)$ lies in $\mathcal{O}_K[[z]]$.

Proof. Recall the notation $\ll$ and $\lll$ introduced at the end of § 1.2. We have the estimate $f_i(z) \lll c/(1-Cz)$ $(1 \le i \le m)$ and so

$$\phi(z) \lll H \binom{N+m}{N} (c/(1-Cz))^N \sum_{\nu=0}^{N} z^\nu$$

$$\ll c^N \binom{N+m}{N} H(1-Cz)^{-N-1} = c^N \binom{N+m}{N} H \sum_{\nu=0}^{\infty} \binom{N+\nu}{N} c^\nu z^\nu$$

$$\ll 2^N c^N \binom{N+m}{N} H \sum_{\nu=0}^{\infty} (2C)^\nu z^\nu .$$

Write $\phi(z) = \sum_{\mu=0}^{\infty} c_\mu z^\mu$ so that

$$\overline{|c_\mu|} \le 2^N c^N \binom{N+m}{N} H(2C)^\mu$$

for all $\mu \ge 0$. Notice also that each $cf_i(Cz)$ lies in $\mathcal{O}[[z]]$ $(1 \le i \le m)$ and so $c^N \phi(Cz)$ lies in $\mathcal{O}_K[[z]]$. The assertion of the lemma is now immediate.

LEMMA 4.2.2. *Let $f_1(z), \ldots, f_m(z)$ be elements of $\mathcal{G}$ analytic near $z = 0$. Then there exists a positive integer c depending only on K and $f_1, \ldots, f_m$ with the following property. For any real number $N \ge c$ we can find a nonzero polynomial $P(z, x_1, \ldots, x_m)$ of degree at most N in z and of degree at most N in $x_1, \ldots, x_m$ whose coefficients are rational integers of absolute values at most $c^{N^{m+1}}$ such that*

$$\phi(z) = P(z, f_1(z), \ldots, f_m(z))$$

satisfies

$$\operatorname{ord}_{z=0} \phi(z) \ge c^{-1} N^{m+1}$$

and such that $\phi(z)$ belongs to the class $\mathcal{G}(1, c)$.

Proof. We may suppose that N is an integer. For nonnegative integers $\nu_0, \nu_1, \ldots, \nu_m$ with $\nu_0 \le N$ and $\nu_1 + \ldots + \nu_m \le N$ we write $\underline{\nu} = (\nu_0, \nu_1, \ldots, \nu_m)$ and $F(\underline{\nu}; z) = z^{\nu_0} f_1^{\nu_1} \ldots f_m^{\nu_m}$. Then

$$F(\underline{\nu}; z) = \sum_{\mu=0}^{\infty} a_\mu(\underline{\nu}) z^\mu$$

with $a_\mu(\underline{\nu})$ in K.

Put $T = [N^{m+1}/2\delta m!]$ where $\delta = [K : Q]$ and observe that $T \geq c^{-1}N^{m+1}$ provided that c is sufficiently large. We write

$$P(z, x_1, \ldots, x_m) = \sum_{\underline{\nu}} p(\underline{\nu}) z^{\nu_0} x_1^{\nu_1} \ldots x_m^{\nu_m} .$$

For T a positive integer the condition $\operatorname{ord}_{z=0} \phi(z) \geq T$ is equivalent to the conditions

$$\sum_{\underline{\nu}} a_\mu(\underline{\nu}) p(\underline{\nu}) = 0 \qquad (0 \leq \mu < T) .$$

The number of coefficients $p(\underline{\nu})$ of P is

$$S = (N+1) \binom{N+m}{N}$$

where $S > (1/m!)N^{m+1} \geq 2\delta T$.

We apply Siegel's lemma (see later) to solve for the $p(\underline{\nu})$ over $\mathbf{Z}$ with not all the $p(\underline{\nu})$ being zero. By Lemma 4.2.1 with $R(z, x_1, \ldots, x_m) = z^{\nu_0} x_1^{\nu_1} \ldots x_m^{\nu_m}$ there are positive integers a, A such that

$$\overline{|a_\mu(\underline{\nu})|} \leq 2^N a^N \binom{N+m}{N} (2A)^\mu \qquad (\mu \geq 0)$$

and $a^N A^\mu a_\mu(\underline{\nu}) \in \mathcal{O}_K$ $(\mu \geq 0)$ so we can take

$$D \leq U = 2^N a^N \binom{N+m}{N}^2 (2A)^T$$

in Siegel's lemma to give

$$0 < \max_{\underline{\nu}} |p(\underline{\nu})| \leq (3U^2)^{\delta T/(S-\delta T)} \leq 3U^2 < c^{N^{m+1}}$$

again provided that c is sufficiently large.

SIEGEL'S LEMMA. *Suppose that $[K : Q] = \delta$ and let a_{ij} $(1 \leq i \leq T,$ $1 \leq j \leq S)$ be elements of K with $S > \delta T$. Suppose also that for an integer $D \geq 1$ and real number $U \geq 1$ we have*

$$\sum_{j=1}^{S} \overline{|a_{ij}|} \leq U , \quad Da_{ij} \in \mathcal{O}_K \qquad (1 \leq i \leq T , \quad 1 \leq j \leq S) .$$

Then there are rational integers $X_1, \ldots, X_S$ with

$$0 < \max |X_j| \leq (3DU)^{\delta T/(S-\delta T)}$$

and

$$\sum_{j=1}^{S} a_{ij} X_j = 0 \qquad (1 \le i \le T).$$

For a proof see [Wa].

To complete the proof we show that $\phi(z)$ belongs to the class $\mathcal{G}(1, c_2)$ for a suitable positive integer c_2. If $\phi(z) = 0$ there is nothing more to prove. Suppose that $\phi(z) \neq 0$ so that $\operatorname{ord}_{z=0} \phi(z) < \infty$. It must be shown that if $\mu \ge \operatorname{ord}_{z=0} \phi(z)$ then

$$\overline{|a_\mu|} \le c_2^\mu, \qquad c_2^\mu a_\mu \in \mathcal{O}_K$$

where $\phi(z) = \sum_{\mu \ge 0} a_\mu z^\mu$. We have

$$\overline{|a_\mu|} \le H(P) \cdot 2^N a^N (N+1) \binom{N+m}{N}^2 (2A)^\mu \le c_1^{N^{m+1}+\mu} \le c_2^\mu$$

and $a^N A^\mu a_\mu \in \mathcal{O}_K$, so that $c_2^\mu a_\mu \in \mathcal{O}_K$ for c_2 suitably chosen. The lemma is now established.

LEMMA 4.2.3. *Let $q, r_1, \ldots, r_\delta$ be complex numbers with $q^\delta + r_1 q^{\delta-1} + \ldots + r_\delta = 0$ and suppose that*

$$Q_2^{-1} \le |q| \le Q_1^{-1}, \quad |r_i| \le R \quad (1 \le i \le \delta)$$

for real $Q_2 \ge Q_1 \ge 1$, $R \ge 1$. Then there exists an index i with

$$(*) \qquad (\delta R)^{-1} Q_2^{-\delta} \le |r_i| \le \delta R Q_1^{-1}$$

Proof. Suppose that the result is false. Then for each i we have either

$$(I) \qquad |r_i| > \delta R Q_1^{-1}$$

or

$$(J) \qquad |r_i| < (\delta R)^{-1} Q_2^{-\delta}.$$

Let I, J be the set of i satisfying (I), (J) respectively. If I is empty we obtain a contradiction from

$$Q_2^{-\delta} \le |q^\delta| = \Big|\sum_{i \in J} r_i q^{\delta - i}\Big| < Q_2^{-\delta}.$$

Consequently I is not empty. Let i_0 be the largest element of I and let $I_0 = I \setminus \{i_0\}$. Consider

$$(**) \qquad r_{i_0} = -q^{i_0} - \sum_{i \in I_0} r_i q^{i_0 - i} - \sum_{i \in J} r_i q^{i_0 - i}.$$

Denote the right-hand side of (**) by T. Observe that

$$\Big|\sum_{i\in I_0} r_i q^{i_0-i}\Big| \le (\#I_0)R|q| \le (\#I_0)RQ_1^{-1}\,,$$

$$\Big|\sum_{i\in J} r_i q^{i_0-i}\Big| \le (\#J)(\delta R)^{-1}Q_2^{-\delta}\cdot Q_2^{\delta-1} = (\#J)(\delta R)^{-1}Q_2^{-1}\,,$$

$$|q^{i_0}| \le RQ_1^{-1}$$

where $\#S$ denotes the cardinality of the set S. Consequently

$$\begin{aligned} T &\le (\#I)RQ_1^{-1} + (\#J)(\delta R)^{-1}Q_2^{-1} \\ &\le \delta R\max(Q_1^{-1},(\delta R)^{-1}Q_2^{-1}) = \delta RQ_1^{-1}\,. \end{aligned}$$

This is impossible since $i_0 \in I$ and so

$$|T| = |r_{i_0}| > \delta RQ_1^{-1}\,.$$

This final contradiction establishes the lemma.

The examples $q^\delta + r_\delta = 0$, $q^\delta + r_1 q^{\delta-1} = 0$ show the exponents in (*) cannot be improved.

LEMMA 4.2.4. *Given K with $[K:Q]=\delta$, there is a constant c, depending only on K, with the following property. Given $Q(\underline{x})$ in $\mathcal{O}[x_1,\ldots,x_m]$ of degree at most $N\ge 1$, we can find $R_1(\underline{x}),\ldots,R_\delta(\underline{x})$ in $\mathbf{Z}[x_1,\ldots,x_m]$ of degrees at most δN such that*

$$Q^\delta + R_1Q^{\delta-1} + \ldots + R_\delta = 0\,.$$

Further

$$\overline{|R_i|} \le cN^{\delta m}\overline{|Q|}^{\,\delta} \qquad (1\le i\le\delta)\,.$$

Proof. Let $\alpha_1,\ldots,\alpha_\delta$ be an integral basis for $\mathcal{O}_K$ so that $\alpha_i\alpha_j = \sum_{k=1}^{\delta} a_{ijk}\alpha_k$ for rational integers a_{ijk} $(1\le i,j,k\le\delta)$. We can write

$$Q = \sum_{j=1}^{\delta}\alpha_j Q_j \tag{4.2.1}$$

for $Q_j(\underline{x})$ in $\mathbf{Z}[\underline{x}]$ $(1\le j\le\delta)$, and then

$$\alpha_i Q = \sum_{k=1}^{\delta}\alpha_k Q_{ik} \qquad (1\le i\le\delta) \tag{4.2.2}$$

for

$$Q_{ik} = \sum_{j=1}^{\delta} a_{ijk} Q_j \qquad (1 \le i, k \le \delta)\,. \tag{4.2.3}$$

From the equations (4.2.2) it is easy to derive the equation

$$\det\left(Q\delta_{ik} - Q_{ik}\right) = 0 \tag{4.2.4}$$

where δ_{ik} denotes the Kronecker symbol. This equation clearly has the required form.

To estimate the $\overline{|R_i|}$ we consider conjugates in (4.2.1). Then (4.2.1) is well-known to imply

$$\overline{|Q_j|} \le c_1 \overline{|Q|}$$

and so (4.2.2) gives

$$\overline{|Q_{ik}|} \le c_2 \overline{|Q|}\,.$$

The required estimate follows immediately from

$$\overline{|R_i|} \le \delta^{2\delta} \binom{N+m}{N}^{\delta} \left((1+c_2)\overline{|Q|}\right)^{\delta} \qquad (1 \le i \le \delta)\,.$$

LEMMA 4.2.5. *Let K be an algebraic number field and denote the integers of K by $\mathcal{O}_K$. Let $\underline{\underline{A}}(z)$ be in $M_{mm}(K(z))$ and $\underline{B}(z)$ be in $M_{1m}(K(z))$. Choose $0 \ne T(z) \in \mathcal{O}_K[z]$ such that*

$$T(z)\underline{\underline{A}}(z) \in M_{mm}(\mathcal{O}_K[z])\,, \qquad T(z)\underline{B}(z) \in M_{1m}(\mathcal{O}_K[z])\,.$$

Let $c_1, q \ge 0$ be such that if a denotes any entry of $T(z)$, $T(z)\underline{\underline{A}}(z)$ or $T(z)\underline{B}(z)$ then

$$a \lll c_1(1+z)^q\,.$$

(Here the notation $\lll$ introduced in § 1.2 is used.) Let $P(z, x_1, \ldots, x_m) \in \mathcal{O}_K[z, x_1, \ldots, x_m]$ satisfy

$$\deg_z P \le N\,, \qquad \deg_{x_1,\ldots,x_m} P \le N\,.$$

Define P_l recursively by $P_0 = P$ and for $l \ge 1$

$$(*) \qquad P_l(z, x_1, \ldots, x_m) = (T(z))^N P_{l-1}(z^b, (x_1, \ldots, x_m)\underline{\underline{A}}(z) + \underline{B}(z))$$

where $b \ge 2$ is an integer. Then the following conclusions hold:

(1) *$P_l(z, x_1, \ldots, x_m)$ is in $\mathcal{O}_K[z, x_1, \ldots, x_m]$ $(l \ge 0)$,*
(2) *$\deg_{x_1,\ldots,x_m} P_l \le N$, $\deg_z P_l \le Nb^l + Nq(b^l - 1)/(b-1)$,*
(3) *$P_l \lll c_2^{Nl} \overline{|P|}\, Z^{D_l} E^N$*

where $Z = 1+z$, $E = 1+x_1+\ldots+x_m$, $c_2 = c_1(m+1)$, $D_l = Nb^l + Nq\,\frac{b^l-1}{b-1}$,

(4) $\overline{|P_l|} \le c_2^{Nl}\overline{|P|}\, 2^{D_l}(m+1)^N \le c_3^{Nb^l}\overline{|P|}$

where $c_3 = c_2(m+1)2^{1+q/(b-1)}$.

Proof. It is clear that (4) follows from (3). The proof of (1)–(3) is by induction on l ($l \ge 0$). For $l = 0$ all these estimates are trivial. Now suppose that $l \ge 1$ and that (1)–(3) hold with $l-1$ in place of l.

Now (*) may be written as

$$P_l(z, x_1, \ldots, x_m) = \widetilde{P}_{l-1}(z, S_0, S_1, \ldots, S_m) \tag{4.2.5}$$

where

$$S_0 = T(z)\,, \qquad (S_1, \ldots, S_m) = T(z)((x_1, \ldots, x_m)\underline{\underline{A}}(z) + \underline{B}(z))$$

and

$$\widetilde{P}_{l-1}(z, X_0, X_1, \ldots, X_m) = X_0^N P_{l-1}(z, X_1/X_0, \ldots, X_m/X_0)$$

is homogenized version of P_{l-1}.

By the inductive hypothesis P_{l-1} is in $\mathcal{O}_K[z, x_1, \ldots, x_m]$ and we have $\deg_{x_1,\ldots,x_m} P_{l-1} \le N$. This implies that $\widetilde{P}_{l-1}$ is in $\mathcal{O}_K[z, X_0, X_1, \ldots, X_m]$ and this in turn implies that $P_l(z, x_1, \ldots, x_m)$ is in $\mathcal{O}_K[z, x_1, \ldots, x_m]$.

Now we have

$$S_0 \lll c_1 Z^q\,, \qquad S_i \lll c_1 Z^q E \qquad (1 \le i \le m) \tag{4.2.6}$$

and if

$$Q(z, X_0, X_1, \ldots, X_m) = c_2^{N(l-1)}\overline{|P|}\, Z^{D_{l-1}}(X_0 + X_1 + \ldots + X_m)^N$$

we have also $\widetilde{P}_{l-1} \lll Q$ by (3) for P_{l-1}. From (4.2.5), (4.2.6) and the estimate just obtained it follows that

$$\begin{aligned} P_l(z, x_1, \ldots, x_m) &\lll Q(z^b, c_1 Z^q, c_1 Z^q E, \ldots, c_1 Z^q E) \\ &= c_2^{N(l-1)}\overline{|P|}\, Z_1^{D_{l-1}} c_1^N Z^{qN} E_1^N \end{aligned}$$

with $Z_1 = 1 + z^b$, $E_1 = 1 + mE$. But $Z_1 \ll Z^b$ and $E_1 \ll (m+1)E$, so that finally

$$P_l(z, x_1, \ldots, x_m) \lll c_2^{Nl}\overline{|P|}\, Z^{bD_{l-1}+qN} E^N\,.$$

This gives the desired result (3), and (3) in turn implies the degree estimates in (2). The lemma is now proved.

V. The main results

§ 5.1. Hypothesis Hyp $(\underline{f}, \alpha)$

These are hypotheses on the functional equation that $f_1(z), \ldots, f_m(z)$ as elements of $K[[z]]$ must satisfy. We break this hypothesis up into two parts.

Hyp $(\underline{f})$: It is supposed that $[K : Q] < \infty$, and that $f_1(z), \ldots, f_m(z)$ are elements of $K[[z]]$, algebraically independent over $K(z)$, such that $\underline{f}(z) = (f_1(z), \ldots, f_m(z))$ satisfies

$$\underline{f}(z^b) = \underline{f}(z)\underline{\underline{A}}(z) + \underline{B}(z)\,.$$

Here $b \geq 1$ is an integer, $\underline{\underline{A}}(z)$ is in $M_{mm}(K(z))$, $\underline{B}(z)$ is in $M_{1m}(K(z))$.

Hyp (α): It is supposed that α is an element of K with $0 < |\alpha| < 1$ such that $\underline{f}(z)$ converges at $z = \alpha$ and that $\underline{\underline{A}}(z)$, $\underline{B}(z)$ are analytic at $z = \alpha, \alpha^b, \alpha^{b^2}, \ldots$

Consequences of Hyp $(\underline{f}, \alpha)$.

(1) From Hyp $(\underline{f})$ we find that $\underline{\underline{A}}(z)$ is nonsingular. Indeed, if $\underline{\underline{A}}(z)$ was singular then it is easy to see that there would exist $p_l(z) \in K(z)$ $(0 \leq l \leq m)$ not all zero and such that $\sum_{l=0}^{m} p_l(z) f_l(z^b) = 0$ where $f_0(z) = 1$. Clearing denominators it may be supposed that each $p_l(z)$ is in $K[z]$. Since each $p_l(z)$ is of the form

$$\sum_{j=0}^{b-1} z^j p_{l,j}(z^b) \qquad (0 \leq l \leq m)\,,$$

by considering residue classes mod b in the Laurent series expansion of $\sum p_l(z) f_l(z^b)$ it is easy to see that there exist polynomials $q_l(z)$ $(0 \leq l \leq m)$ with $q_l(z)$ not all zero and $\sum_{l=0}^{m} q_l(z) f_l(z) = 0$. This is a contradiction.

(2) Since $\underline{\underline{A}}(z)$ is nonsingular we have $\underline{f}(z) = \underline{f}(z^b)\underline{\underline{A}}(z)^{-1} - \underline{B}(z)\underline{\underline{A}}(z)^{-1}$, and so from Lemma 2.1.2 we see that $\underline{f}(z) \in M_{1m}(\mathcal{G})$.

(3) Define $Z(N)$ by

$$\begin{aligned} Z(N) = \sup\{\operatorname{ord}_{z=0} \phi(z) \mid\ & \phi(z) = P(z, \underline{f}(z)) \\ & \text{where } 0 \neq P \in \mathbb{Z}[z, x_1, \ldots, x_m] \text{ has} \\ & \deg_z P \leq N\,,\ \deg_{x_1,\ldots,x_m} P \leq N\}\,. \end{aligned}$$

Then $Z(N)$ must satisfy $Z(N) < \infty$. We also know that

$$Z(N) \geq C^{-1} N^{m+1}$$

as a consequence of linear algebra or by means of Lemma 4.2.2.

(4) Using Hyp (α) we can find $0 \neq T(z)$ in $\mathcal{O}_K[z]$ such that $T(\alpha^{b^k}) \neq 0$ $(k = 0, 1, \ldots)$ and $T(z)\underline{\underline{A}}(z) \in M_{mm}(\mathcal{O}_K[z])$, $T(z)\underline{B}(z) \in M_{1m}(\mathcal{O}_K[z])$. Here $\mathcal{O}_K$ denotes the ring of integers of the algebraic number field K.

As mentioned before our main goal is to establish an effective version of Nesterenko's Theorem 2 of [Ne4]. This is accomplished in Theorem 5.2.1 after a technical proposition. From this result we obtain Theorem 5.2.2 and Theorem 5.2.3 as immediate consequences. Our immediate goal is to prove Lemma 5.1.1 which will be needed in § 5.2 to prove the proposition.

LEMMA 5.1.1. *Assume* Hyp $(\underline{f}, \alpha)$. *Then there exists* $c > 0$ *depending only on* K, $\underline{f}$ *and* α *with the following property. For any* $N \geq c$ *there exists an integer* W *with*

$$c^{-1}N^{m+1} \leq W \leq Z(N)$$

such that for any integer l *with* $b^l \geq cW$ *we can find* $R_l(X_0, X_1, \ldots, X_m)$, *homogeneous of degree* δN *in* $Z[X_0, X_1, \ldots, X_m]$ *such that*

(i) $\ln \overline{|R_l|} \leq cNb^l$,

(ii) $-cWb^l \leq \ln |R_l(\overline{\omega})| \leq -c^{-1}Wb^l$ *for* $\overline{\omega} = (1, f_1(\alpha), \ldots, f_m(\alpha))$.

Proof. First observe that $\underline{f}(z) \in M_{1m}(\mathcal{G})$. This has already been observed as a consequence of Hyp $(\underline{f}, \alpha)$. Therefore by Lemma 4.2.2 there exists a positive integer c_1 such that for any $N \geq c_1$ we can find a nonzero polynomial $P(z, x_1, \ldots, x_m)$ in $\mathbb{Z}[z, x_1, \ldots, x_m]$ of degree at most N in z and of degree at most N in $x_1, \ldots, x_m$ such that

$$\overline{|P|} \leq c_1^{N^{m+1}} \tag{5.1.1}$$

and $\phi(z) = P(z, f_1(z), \ldots, f_m(z))$ satisfies

$$W = \operatorname{ord}_{z=0} \phi(z) \geq c_1^{-1}N^{m+1}.$$

Further

$$\phi(z) \quad \text{is in the class} \quad \mathcal{G}(1, c_1). \tag{5.1.2}$$

(Recall the definition of the class $\mathcal{G}(c, C)$ at the beginning of § 2.1.)

Now using Hyp (α) we can find $0 \neq T(z)$ in $\mathcal{O}_K[z]$ such that

$$T(\alpha^{b^k}) \neq 0 \qquad (k = 0, 1, \ldots) \tag{5.1.3}$$

and $T(z)\underline{\underline{A}}(z) \in M_{mm}(\mathcal{O}_K[z])$, $T(z)\underline{B}(z) \in M_{1m}(\mathcal{O}_K[z])$. We write $\underline{\underline{A}}(z) = [a_{ij}(z)]_{1 \leq i,j \leq m}$, $\underline{B}(z) = [b_1(z), \ldots, b_m(z)]$ and we choose $g \geq 1$ real, $q \geq 0$

an integer so that

$$
\begin{aligned}
&T(z) \lll g(1+z)^q\,,\\
(5.1.4)\qquad &T(z)a_{ij}(z) \lll g(1+z)^q\,,\\
&T(z)b_j(z) \lll g(1+z)^q \quad \text{for all } i,j\,.
\end{aligned}
$$

Now define $P_0 = P$ and for $l \geq 1$

$$P_l(z, x_1, \ldots, x_m) = T(z)^N P_{l-1}(z^b, (x_1, \ldots, x_m)\underline{\underline{A}}(z) + \underline{B}(z))\,.$$

By Lemma 4.2.5 we know that

$$
\begin{aligned}
&P_l \in \mathcal{O}_K[z, x_1, \ldots, x_m]\,,\\
(5.1.5)\qquad &\deg_z P_l \leq D_l = Nb^l + Nq\frac{b^l - 1}{b-1}\,,\\
&\deg_{x_1,\ldots,x_m} P_l \leq N
\end{aligned}
$$

and

$$(5.1.6)\qquad \overline{|P_l|} \leq c_2^{Nb^l}\overline{|P|} \leq c_3^{Nb^l + N^{m+1}}\,.$$

Now let $a \geq 1$ be such that $a\alpha$ is in $\mathcal{O}_K$ and put

$$(5.1.7)\qquad Q_l(x_1, \ldots, x_m) = a^{D_l} P_l(\alpha, x_1, \ldots, x_m) \in \mathcal{O}_K[x_1, \ldots, x_m]\,.$$

Put $q = Q_l(\underline{f}(\alpha))$. We wish to derive upper and lower bounds for $|q|$. Notice that the definitions of P_l and the functional equations for $\underline{f}$ show that

$$(5.1.8)\qquad q = a^{D_l}\psi_l(\alpha)\phi(\alpha^{b^l})$$

where

$$\psi_l(z) = \begin{cases} 1 & \text{if } l = 0 \\ T(\alpha)\ldots T(\alpha^{b^{l-1}}) & \text{if } l \geq 1 \end{cases}$$

and $\phi(z) = P(z, f_1(z), \ldots, f_m(z))$. Here we use

$$P_l(z, f_1(z), \ldots, f_m(z)) = T(z)^N P_{l-1}(z^b, f_1(z^b), \ldots, f_m(z^b))$$

for $l \geq 1$.

It will now be shown that if

$$b^l \geq c_4 W$$

where

$$c_4 = \frac{4(1+2\delta)\ln c_1}{\ln(1/|\alpha|)} \qquad \text{and} \qquad c_1 \geq e$$

then

$$(5.1.9) \qquad -\frac{5}{4}b^l W \ln(1/|\alpha|) \le \ln|\phi(\alpha^{b^l})| \le -\frac{3}{4}b^l W \ln(1/|\alpha|)\,.$$

We first observe that $\max(\operatorname{den} a_\mu, \overline{|a_\mu|}\,) \le c_1^\mu$ for all $\mu \ge 0$ where $\phi(z) = \sum_{\mu\ge 0} a_\mu z^\mu$. Put $c_1^{1+2\delta} = c_5$. Then

$$c_5^{2W}|\alpha|^{b^l} \le |\alpha|^{b^l/2}\,, \qquad |a_\mu/a_W| \le c_5^\mu \quad (\mu \ge W)\,.$$

Here the results of § 1.2 are used and so since $\mu \ge W+1$ implies that $\mu \le 2W(\mu - W)$ we see that for $\mu \ge W+1$,

$$\left|\frac{a_\mu \alpha^{b^l\mu}}{a_W \alpha^{b^l W}}\right| \le c_5^\mu |\alpha|^{b^l(\mu-W)} \le \left(c_5^{2W}|\alpha|^{b^l}\right)^{(\mu-W)} \le |\alpha|^{b^l(\mu-W)/2}\,.$$

Let

$$x = \sum_{\mu\ge W+1} \frac{a_\mu \alpha^{b^l\mu}}{a_W \alpha^{b^l W}}\,.$$

We see that

$$|x| \le \sum_{\mu\ge W+1} |\alpha|^{b^l(\mu-W)/2} \le \frac{|\alpha|^{b^l/2}}{1-|\alpha|^{b^l/2}} \le \frac{1}{2}$$

since

$$b^l \ge \frac{4\ln c_5}{\ln(1/|\alpha|)}W \ge \frac{2\ln 3}{\ln(1/|\alpha|)}$$

so that $|\alpha|^{b^l/2} \le \frac{1}{3}$.

Now $\phi(\alpha^{b^l}) = a_W \alpha^{b^l W}(1+x)$ and since $|x| \le 1/2$ we see that

$$\big|\ln|1+x|\big| \le |\ln(1+x)| \le 2|x| \le 1$$

and so

$$\begin{aligned} |\ln|\phi(\alpha^{b^l})| + b^l W \ln(1/|\alpha|)| &\le 1 + |\ln|a_W|| \\ &\le 1 + (2\delta)(\ln c_1)W \le \frac{1}{4}b^l W \ln(1/|\alpha|)\,, \end{aligned}$$

which establishes (5.1.9).

Now the following result was established in § 1.2. If $f \in \overline{\mathbb{Q}}[x_1,\ldots,x_t]$ $\alpha_1,\ldots,\alpha_t \in \overline{\mathbb{Q}}$ then $\beta = f(\alpha_1,\ldots,\alpha_t)$ satisfies

$$\max\left(\operatorname{den}\beta, \overline{|\beta|}\,\right) \le 2^t \max\left(\operatorname{den} f, \overline{|f|}\,\right) \prod_{\nu=1}^{t} A_\nu^{\deg_{x_\nu} f}$$

where the notation $A_\nu = \max(\operatorname{den}\alpha_\nu, \overline{|\alpha_\nu|}, 2)$ is used. We see that

$$\max\big(\operatorname{den}\psi_l(\alpha), \overline{|\psi_l(\alpha)|}\big) \le c_6^{Nb^l} \tag{5.1.10}$$

on using the definition of $\psi_l(z)$.

From (5.1.8), (5.1.9), (5.1.10) we find that

$$-c_7 W b^l \le \ln|q| \le -c_7^{-1} W b^l \quad \text{provided that} \quad b^l \ge c_4 W\,. \tag{5.1.11}$$

We next apply Lemma 4.2.4 to $Q(x_1,\ldots,x_m) = Q_l(x_1,\ldots,x_m)$ where $Q_l(x_1,\ldots,x_m)$ is defined in (5.1.7). There exist $R_1,\ldots,R_\delta$ in $\mathbb{Z}[x_1,\ldots,x_m]$ of degrees at most δN such that

$$(Q(\underline{x}))^\delta + R_1(\underline{x})(Q(\underline{x}))^{\delta-1} + \ldots + R_\delta(\underline{x}) = 0\,. \tag{5.1.12}$$

Further $\overline{|R_i|} \le c_8 N^{\delta m}\overline{|Q|}^{\,\delta}$, and using the estimates (5.1.6) we find that

$$\overline{|R_i|} \le c_9^{Nb^l + N^{m+1}} \le c_{10}^{Nb^l} \tag{5.1.13}$$

provided that $b^l \ge c_3 W$.

Putting $\underline{x} = \underline{f}(\alpha)$ in (5.1.12) we see that the numbers $r_i = R_i(\underline{f}(\alpha))$ satisfy

$$q^\delta + r_1 q^{\delta-1} + \ldots + r_\delta = 0$$

and $|r_i| \le R$ for

$$R \le c_{10}^{Nb^l + N^{m+1}} \le c_{11}^{Nb^l}$$

provided that $b^l \ge c_3 W$. Apply Lemma 4.2.3 to (5.1.11), (5.1.13) to conclude that for some i with $1 \le i \le \delta$, the polynomial $R(\underline{x}) = R_i(\underline{x})$ satisfies

$$e^{-\delta c_7 W b^l}(\delta R)^{-1} \le |R(\underline{f}(\alpha))| \le e^{-c_7^{-1} - W b^l}\delta R\,.$$

The required inequalities (ii) of the present lemma follow at once on taking

$$R_l(X_0, X_1, \ldots, X_m) = X_0^{\delta N} R(X_1/X_0, \ldots, X_m/X_0)\,.$$

The lemma is now established.

§ 5.2. Conclusions

We will now prove Proposition 5.2.1. followed by Theorem 5.2.1. Theorem 5.2.1 is an effective version of Nesterenko's Theorem 2 of [Ne4]. We will then apply this result together with the results of Chapter II to establish Theorems 5.2.2 and 5.2.3. These last three theorems are the main results of this thesis.

Recall the definition of Hyp $(\underline{f}, \alpha)$ and $Z(N)$ at the beginning of § 5.1. Here

$$Z(N) = \sup\{\operatorname{ord}_{z=0} \phi(z) \mid \phi(z) = P(z, f_1(z), \ldots, f_m(z))$$
$$\text{where } 0 \neq P \in \mathbb{Z}[z, x_1, \ldots, x_m] \text{ satisfies}$$
$$\deg_z P \leq N, \quad \deg_{x_1,\ldots,x_m} P \leq N\}.$$

We know by Lemma 4.2.2 or by linear algebra that

$$Z(N) \geq c^{-1} N^{m+1} \tag{5.2.1}$$

for all $N \geq c$ where $c > 0$ depends only on m and $[K : Q]$.

PROPOSITION 5.2.1. *Suppose that $\underline{f}$, α satisfy* Hyp $(\underline{f}, \alpha)$. *Then there exists an effectively computable constant μ_0 depending only on K, $\underline{f}$ and α, with the following property. Let $\mu \geq \mu_0$, $\lambda = \mu^{m+1}$ and $I \subseteq \mathbb{Z}[X] = \mathbb{Z}[X_0, X_1, \ldots, X_m]$ be an unmixed homogeneous ideal with $I \cap \mathbb{Z} = (0)$, such that $r = m + 1 - h(I)$ satisfies $1 \leq r \leq m$. Suppose also that*

(i) $N(I) \leq \lambda^{m-r} D^{m-r+1}$,

(ii) $\ln H(I) \leq \lambda^{m-r} D^{m-r} \ln H$

for some D and H with

(iii) $D \geq 1$, $\ln H \geq (Z(\mu^n D))^2$.

Then

(iv) $\ln |I(\overline{\omega})| > -\lambda^r D^{r-1}(D \ln H(I) + N(I) \ln H)$

for $\overline{\omega} = (1, f_1(\alpha), \ldots, f_m(\alpha))$.

Proof. The constants $c_1, c_2, \ldots$ will be positive and effectively computable in terms of K, $\underline{f}$ and α. For brevity we shall write

$$T(I) = D \ln H(I) + N(I) \ln H$$

for any ideal I as above. Before beginning the main proof, observe that by choosing μ_0 sufficiently large

$$\ln H > m^2 D \tag{5.2.2}$$

is seen to hold. This is a consequence of (5.2.1) and (iii).

The proof of the proposition is by induction on r. Fix r with $1 \leq r \leq m$, and assume that the result is false for some ideal I as in the proposition. If $r = 1$, we shall eventually obtain a contradiction, whereas if $r > 1$, we shall deduce a contradiction from further assuming that the proposition holds for all J with $h(J) > m + 1 - r$. It is seen that there exists an unmixed homogeneous ideal I with $h(I) = m + 1 - r$ satisfying (i)–(iii) but

$$\ln |I(\overline{\omega})| \leq -\lambda^r D^{r-1} T(I). \tag{5.2.3}$$

It will now be shown that there exists a homogeneous prime ideal $\mathfrak{p}$ of $\mathbb{Z}[X]$, with $\mathfrak{p} \cap \mathbb{Z} = (0)$ and $h(\mathfrak{p}) = m + 1 - r$, such that

$$N(\mathfrak{p}) \leq \lambda^{m-r} D^{m-r+1}, \tag{5.2.4}$$

$$\ln H(\mathfrak{p}) \leq \lambda^{m-r} D^{m-r} (\ln H + m^2 D), \tag{5.2.5}$$

$$\ln |\mathfrak{p}(\overline{\omega})| \leq -\frac{1}{3} \lambda^r D^{r-1} T(\mathfrak{p}). \tag{5.2.6}$$

For suppose this is not true. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ be the prime ideals of I satisfying $\mathfrak{p}_i \cap \mathbb{Z} = (0)$ $(1 \leq i \leq s)$. By Lemma 4.1.4 it is seen that (5.2.4) and (5.2.5) hold for the ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ and therefore

$$\ln |\mathfrak{p}_i(\overline{\omega})| > -\frac{1}{3} \lambda^r D^{r-1} T(\mathfrak{p}_i) \qquad (1 \leq i \leq s).$$

Define $k_1, \ldots, k_s$ as in Lemma 4.1.4. By that lemma

$$\sum_{l=1}^{s} k_l \ln H(\mathfrak{p}_l) \leq \ln H(I) + m^2 N(I),$$

$$\sum_{l=1}^{s} k_l N(\mathfrak{p}_l) = N(I),$$

$$\sum_{l=1}^{s} k_l \ln |\mathfrak{p}_l(\overline{\omega})| \leq \ln |I(\overline{\omega})| + m^3 N(I)$$

and so

$$\begin{aligned} \sum_{l=1}^{s} k_l T(\mathfrak{p}_l) &= \sum_{l=1}^{s} k_l (D \ln H(\mathfrak{p}_l) + N(\mathfrak{p}_l) \ln H) \\ &\leq D(\ln H(I) + m^2 N(I)) + N(I) D \ln H \\ &= T(I) + m^2 D N(I). \end{aligned}$$

Therefore

$$\ln |I(\overline{\omega})| + m^3 N(I) > -\frac{1}{3} \lambda^r D^{r-1} (T(I) + m^2 D N(I)).$$

Since $\ln |I(\overline{\omega})| \leq -\lambda^r D^{r-1} T(I)$ we deduce that

$$\begin{aligned} \frac{2}{3} \lambda^r D^{r-1} T(I) &< \left(\frac{1}{3} \lambda^r D^{r-1} m^2 D + m^3 \right) N(I) \\ &\leq \frac{2}{3} \lambda^r D^{r-1} m^2 D N(I) \qquad (\text{for } \mu_0 \geq 3m) \end{aligned}$$

and so

$$T(I) < m^2 DN(I).$$

This clearly contradicts (5.2.2) and the contradiction so obtained establishes that one of the prime ideals satisfies (5.2.4)–(5.2.6). Call this prime ideal $\mathfrak{p}$.

Next, in the notation of Lemma 4.1.6 let $\rho = \inf\{\|\overline{\omega} - \overline{\beta}\| : 0 \neq \overline{\beta} \in \mathbb{C}^{m+1}$ is a zero of the ideal $\mathfrak{p}\}$. By Lemma 4.1.7 we know that

$$N(\mathfrak{p}) \ln \rho \leq \frac{1}{r} \ln |\mathfrak{p}(\overline{\omega})| + 3m^3 N(\mathfrak{p})$$

and so (5.2.6) gives

$$\begin{aligned} N(\mathfrak{p}) \ln \rho &\leq -\frac{\lambda^r}{3r} D^{r-1} N(\mathfrak{p}) \ln H + 3m^3 N(\mathfrak{p}) \\ &\leq -\frac{\lambda^r}{4r} D^{r-1} N(\mathfrak{p}) \ln H \end{aligned}$$

provided μ is sufficiently large. In particular, $\rho \leq 1$. Notice also that

$$\frac{1}{2} \ln \frac{1}{\rho} \geq \frac{1}{8r} \lambda^r D^{r-1} \ln H.$$

Define now

$$X = \frac{1}{8r} \lambda^r D^{r-1} T(\mathfrak{p})$$

and

$$\chi = \min\left(X, \frac{\ln(1/\rho)}{2}\right)$$

so that

$$\chi \geq \frac{1}{8r} \lambda^r D^{r-1} \ln H. \tag{5.2.7}$$

Now put $N = [\mu^m D]$ and apply Lemma 5.1.1. There exists W with $c_1^{-1} N^{m+1} \leq W \leq Z(N)$ such that for any integer l with

$$b^l \geq c_1 W \tag{5.2.8}$$

we can find $R_l(X_0, \ldots, X_m)$ homogeneous of degree δN in $\mathbb{Z}[X_0, \ldots, X_m]$ such that

$$\ln \overline{|R_l|} \leq c_1 N b^l, \tag{5.2.9}$$

$$-c_1 W b^l \leq \ln |R_l(\overline{\omega})| \leq -c_1^{-1} W b^l. \tag{5.2.10}$$

Let l be the unique integer satisfying

$$2c_1 W b^{l+1} > \chi \geq 2c_1 W b^l. \tag{5.2.11}$$

By (5.2.7)

$$2c_1Wb^{l+1} > X \geq \frac{1}{8r}\lambda^r D^{r-1} \ln H$$

and since $\ln H \geq Z(N)^2 \geq W^2$, (5.2.8) holds provided μ_0 is sufficiently large.

For brevity write $Q(X) = R_l(X)$ for this polynomial. Put

$$q = |Q(\overline{\omega})||\overline{\omega}|^{-\deg Q}$$

where $|\overline{\omega}| = \max(1, |f_1(\alpha)|, \ldots, |f_m(\alpha)|)$. From (5.2.10)

$$-2c_1Wb^l \leq \ln q \leq -\frac{1}{2}c_1^{-1}Wb^l \tag{5.2.12}$$

if μ_0 is sufficiently large.

We intend to apply Lemma 4.1.6. By (5.2.6) and the definition of X we know that

$$|\mathfrak{p}(\overline{\omega})| \leq e^{-X}, \qquad X > 0.$$

By means of (5.2.11) and (5.2.8) it is easy to verify that

$$q \leq H(Q)^{-1}(\deg Q + 1)^{-(2m+2)}.$$

Define σ as in Lemma 4.1.6 by $X = -\sigma \ln q$. By (5.2.11) and (5.2.12) it follows that

$$1 \leq \sigma < 4c_1^2 b. \tag{5.2.13}$$

Indeed,

$$2c_1Wb^{l+1} > X = -\sigma \ln q \geq \frac{\sigma}{2}c_1^{-1}Wb^l,$$

$$2c_1Wb^l \leq X = -\sigma \ln q \leq 2\sigma c_1 Wb^l$$

imply that $\sigma < 4c_1^2 b$ and $\sigma \geq 1$ respectively. It has already been observed that $\rho \leq 1$. All the hypotheses of Lemma 4.1.6 are now verified. Apply Lemma 4.1.6. Its conclusion can be expressed as follows: Define

$$Y_1 = (\deg Q) \ln H(\mathfrak{p}) + N(\mathfrak{p}) \ln H(Q) + 8m^2 N(\mathfrak{p}) \deg Q,$$

$$Y = -\frac{X}{2\sigma} + Y_1.$$

Then if $r = 1$ we have $Y \geq 0$ while if $r > 1$ there exists an unmixed homogeneous ideal J of $\mathbb{Z}[X]$ with

$$h(J) = m - r + 2, \qquad N(J) \leq N(\mathfrak{p}) \deg Q, \tag{5.2.14}$$

$$\ln |J(\overline{\omega})| \leq Y. \tag{5.2.15}$$

Further, the estimate for $\ln H(J)$ given in (2) of Lemma 4.1.5 implies that

$$\ln H(J) \le Y_1 . \tag{5.2.16}$$

Before proceeding to a contradiction we first establish the inequality

$$Y_1 \le c_2 \mu^m T(\mathfrak{p}) . \tag{5.2.17}$$

Indeed, by (5.2.9),

$$\begin{aligned} Y_1 &\le (N\delta) \ln H(\mathfrak{p}) + N(\mathfrak{p}) c_1 N b^l + 8m^2 N(\mathfrak{p})(N\delta) \\ &\le c_3 (N \ln H(\mathfrak{p}) + N b^l N(\mathfrak{p})) \le c_3 \mu^m T(\mathfrak{p}) + c_3 N b^l N(\mathfrak{p}) . \end{aligned} \tag{5.2.18}$$

Also from (5.2.11), (5.2.1) and the definition of X we see that

$$b^l \le \frac{1}{2c_1} \chi W^{-1} \le c_4 \chi N^{-m-1} \le c_4 X^{-m-1}$$

Now $N = [\mu^m D] \ge \frac{1}{2}\mu^m D$ and $X = \frac{1}{8r}\lambda^r D^{r-1} T(\mathfrak{p})$ so that

$$\begin{aligned} b^l &\le c_4 \frac{1}{8r} \mu^{(m+1)r} D^{r-1} 2^{m+1} \mu^{-m(m+1)} D^{-m-1} \\ &\le c_5 \mu^{(r-m)(m+1)} D^{r-m+2} T(\mathfrak{p}) . \end{aligned}$$

Also by (5.2.4) we have

$$N \cdot N(\mathfrak{p}) \le \mu^m D \cdot \mu^{(m-r)(m+1)} D^{m-r+1}$$

so that $N b^l N(\mathfrak{p}) \le c_6 \mu^m T(\mathfrak{p})$. Using this inequality in (5.2.18) gives (5.2.17) as required.

It was shown in (5.2.13) that σ satisfies $1 \le \sigma < 4c_1^2 b$. Using this inequality, the definition of X, and (5.2.15) it follows that

$$Y = -\frac{X}{2\sigma} + Y_1 < -\frac{X}{4\sigma} \le -c_7 \lambda^r D^{r-1} T(\mathfrak{p}) \tag{5.2.19}$$

provided that μ_0 is chosen sufficiently large. If $r = 1$ this already contradicts $Y \ge 0$ and this contradiction starts off the inductive procedure.

Suppose now that $r \ge 2$. It will be verified that J satisfies the conditions (i) and (ii) of the Proposition with r replaced by $r-1$. By (5.2.14) and (5.2.4)

$$N(J) \le \delta N \cdot N(\mathfrak{p}) \le \delta \mu^m D \cdot \lambda^{m-r} D^{m-r+1} \le \lambda^{m-r+1} D^{m-r+2}$$

if $\mu \ge \delta$, which is (i) as required. Also (5.2.16) and (5.2.17) give

$$\ln H(J) \le c_2 \mu^m (D \ln H(\mathfrak{p}) + N(\mathfrak{p}) \ln H) . \tag{5.2.20}$$

From (5.2.2) and (5.2.5) we have

$$\ln H(\mathfrak{p}) < 2\lambda^{m-r} D^{m-r} \ln H .$$

Using this inequality and (5.2.4) we deduce from (5.2.20) that

$$\ln H(J) \le c_2\mu^m\lambda^{m-r}D^{m-r+1}(2\ln H + \ln H) \le \lambda^{m-r+1}D^{m-r+1}\ln H$$

if $\mu_0 \ge 3c_2$, which is (ii) as required.

We now verify that

$$\ln|J(\overline{\omega})| > -\lambda^{r-1}D^{r-1}T(J)\,. \tag{5.2.21}$$

Indeed, if $J \cap \mathbf{Z} = (0)$ then J satisfies all the conditions of the Proposition with $r-1$ in place of r so that in this case (5.2.21) follows by the inductive hypothesis. If however $J \cap \mathbf{Z} \ne (0)$ then the remarks following Lemma 4.1.4 show that

$$\ln|J(\overline{\omega})| + m^3N(J) \ge 0$$

and since $N(J) \le T(J)$ this evidently implies (5.2.21). Now (5.2.16) and (5.2.14) lead to

$$T(J) \le DY_1 + N(\mathfrak{p})\deg Q \cdot \ln H \le DY_1 + (\deg Q)T(\mathfrak{p})\,.$$

Using (5.2.17) it follows that

$$T(J) \le c_2\mu^m T(\mathfrak{p})D + \mu^m D \cdot T(\mathfrak{p}) \le c_3\mu^m D \cdot T(\mathfrak{p})$$

Substitute that last upper bound for $T(J)$ into (4.3.21) to obtain

$$\ln|J(\overline{\omega})| > -c_8\lambda^{r-1}\mu^m D^{r-1}T(\mathfrak{p})\,.$$

On comparing this inequality with (5.2.15) and (5.2.19) a contradiction is obtained provided that μ_0 is sufficiently large. This final contradiction establishes the induction step, and thereby completes the proof of the Proposition.

Recall the definition of Hyp $(\underline{f}, \alpha)$, $Z(N)$ at the beginning of § 5.1.

THEOREM 5.2.1. *Suppose that $\underline{f}$, α satisfy* Hyp $(\underline{f}, \alpha)$. *Then there exists an effective computable constant c depending only on K, $\underline{f}$ and α, with the following property. Suppose that $P(x_1, \ldots, x_m)$ is a nonzero polynomial of degree at most D with rational integer coefficients of absolute value at most H. Then*

$$\ln|(\underline{f}(\alpha))| \ge -cD^m(\ln H + Z^2(cD))\,.$$

Proof. It may be supposed that $D \ge 1$. Put

$$\widetilde{P}(X_0, X_1, \ldots, X_m) = X_0^D P(X_1/X_0, \ldots, X_m/X_0)$$

and $I = (\widetilde{P})$ in $\mathbf{Z}[X_0, X_1, \ldots, X_m]$. Then $h(I) = 1$, so $r = m$ in the Proposition. By Lemma 4.1.3 we obtain

$$N(I) \le D\,, \qquad \ln H(I) \le \ln H + m^2 D \le \ln \widetilde{H}\,, \tag{5.2.22}$$

where $\widetilde{H}$ is defined by

$$\ln \widetilde{H} = \ln H + m^2 D + Z^2$$

for $Z = Z(\mu^m D)$. Since $\ln \widetilde{H} \geq Z^2$, we may apply the Proposition to conclude that

$$\ln |I(\overline{\omega})| > -\lambda^m D^{m-1}(D \ln H(I) + N(I) \ln \widetilde{H}) .$$

Therefore by (5.2.22),

$$\ln |I(\overline{\omega})| > -2\lambda^m D^m \ln \widetilde{H} .$$

Since $Z \geq c_1(\mu^m D)^{m+1} \geq m^2 D$, we obtain

$$\ln \widetilde{H} \leq \ln H + 2Z^2$$

and so

$$\ln |I(\overline{\omega})| > -4\lambda^m D^m (\ln H + Z^2) .$$

By Lemma 4.1.3

$$\begin{aligned} \ln |P(\underline{f})(\alpha)| = \ln |\widetilde{P}(\overline{\omega})| &\geq \ln |I(\overline{\omega})| - D|\ln |\overline{\omega}|| - 2mD \ln(m+1) \\ &> -5\lambda^m (\ln H + Z^2) . \end{aligned}$$

The Theorem is now proved.

The following two theorems follow immediately from this result and Theorems 3.1, 3.3 of Chapter III.

THEOREM 5.2.2. *Assume* Hyp $(\underline{f}, \alpha)$. *Let* $0 \neq P \in \mathbb{Z}[x_1, \ldots, x_m]$ *satisfy*

$$\deg_{x_1,\ldots,x_m} P \leq D , \qquad H(P) = \text{height}(P) \leq H .$$

Then

$$\ln |P(\underline{f}(\alpha))| \geq -e^{cD^m} - cD^m \ln H$$

where c is an effective constant, computable in terms of K, $\underline{f}$ and α.

THEOREM 5.2.3. *Assume* Hyp $(\underline{f}, \alpha)$ *and in addition assume that* $\underline{\underline{A}}(z)$ *is upper triangular. Let* $0 \neq P \in \mathbb{Z}[x_1, \ldots, x_m]$ *satisfy*

$$\deg_{x_1,\ldots,x_m} P \leq D , \qquad H(P) = \text{height}(P) \leq H .$$

Then

$$\ln |P(\underline{f}(\alpha))| \geq -cD^m (D^{2^m+1} + \ln H)$$

where c is an effective constant, computable in terms of K, $\underline{f}$ and α.

COROLLARY 1. *Assume* Hyp$(\underline{f}, \alpha)$ *and in addition suppose that* $\underline{\underline{A}}(z)$ *is upper triangular. Then the numbers* $\theta_i = f_i(\alpha)$, $i = 1, \ldots, m$, *are of finite transcendence type, at most* $2^{m+1} + m$.

COROLLARY 2. *Let* $f(z) = \sum_{h=0}^{\infty} z^{k^h}$, $|z| < 1$, *where* k *is an integer* ≥ 2. *Let* $\alpha \in \overline{\mathbb{Q}}$ *satisfy* $0 < |\alpha| < 1$ *and let*

$$\theta_i = \frac{d^{i-1}}{dz^{i-1}} f(z)\Big|_{z=\alpha} \qquad (i = 1, \ldots, m)$$

for m *an integer* ≥ 1. *Then* $\theta_1, \ldots, \theta_m$ *have mutual transcendence type at most* $2^{m+1} + m$.

(We refer to Mahler for a proof of the algebraic independence over $\mathbb{C}(z)$ of the functions $\frac{d^{i-1}}{dz^{i-1}} f(z)$ $(i = 1, \ldots, m)$.)

Appendix

The purpose of this appendix is to prove Lemma C which determines in particular a lower bound for the radius of convergence for solution of

$$\Phi(z^b)A(z) = \Lambda\Phi(z) + B(z).$$

Definitions

Let K be an algebraic number field. Let $K((z))$ denote the quotient field of $K[[z]]$ where $K[[z]]$ is the ring of formal power series with coefficients in K. For $B = [b_{ij}]_{\substack{1 \leq i \leq s \\ 1 \leq j \leq t}} \in M_{st}(K)$ define

$$\lceil B \rceil = \sqrt{\sum_{i=1}^{s} \sum_{j=1}^{t} \lceil b_{ij}^2 \rceil}$$

and

$$|B| = \sqrt{\sum_{i=1}^{s} \sum_{j=1}^{t} |b_{ij}|^2}.$$

Now $K((z)) = \bigcup_{n=0}^{\infty} z^{-n} K[[z]]$. For $A(z) = \sum_{n=0}^{\infty} A_n z^n \in M_{st}(K((z)))$ we make the following further definitions:

$$\rho(A(z)) = \limsup_{n \to \infty} |A_n|^{1/n},$$

$$\overline{\rho}(A(z)) = \limsup_{n \to \infty} \lceil A_n \rceil^{1/n}.$$

For the purpose of stating the following lemma, the following further definitions will be given. Let $\mathcal{G} = \{\sigma : K \to \overline{\mathbb{Q}}, \sigma$ is a field monomorphism

fixing $\mathbb{Q}\}$ where $\overline{\mathbb{Q}}$ denotes the algebraic closure of $\mathbb{Q}$. Now $\mathcal{G}$ acts on $M_{st}(K)$ by acting on the entries of matrices and $\mathcal{G}$ acts on $M_{st}(K(z))$ by $\left(\sum_{n\geq\mu} A_n z^n\right)^\sigma = \sum_{n\geq\mu} A_n^\sigma z^n$.

LEMMA A. (1) *For* $A, B \in M_{st}(K((z)))$ *and* $\alpha, \beta \in K$

$$\rho(\alpha A + \beta B) \leq \max(\rho(A), \rho(B)), \quad \overline{\rho}(\alpha A + \beta B) \leq \max(\overline{\rho}(A), \overline{\rho}(B)).$$

(2) *For* $A \in M_{st}(K((z)))$, $B \in M_{tu}(K((z)))$

$$\rho(AB) \leq \max(\rho(A), \rho(B)), \qquad \overline{\rho}(AB) \leq \max(\overline{\rho}(A), \overline{\rho}(B)).$$

(3) *For* $A \in M_{st}(K((z)))$

$$\overline{\rho}(A) = \sup\{\, \rho(A^\sigma) \mid \sigma \in \mathcal{G}\,\}.$$

Proof. (1) This follows easily from

$$|\alpha A_n + \beta B_n|^{1/n} \leq (2\max(|\alpha|, |\beta|))^{1/n} \max(|A_n|^{1/n}, |B_n|^{1/n}),$$

$$\overline{|\alpha A_n + \beta B_n|}^{1/n} \leq \left(2\max(\overline{|\alpha|}, \overline{|\beta|})\right)^{1/n} \max\left(\overline{|A_n|}^{1/n}, \overline{|B_n|}^{1/n}\right).$$

(2) Let $A(z) = \sum A_n z^n$, $B(z) = \sum B_n z^n$. Then $C(z) = A(z)B(z) = \sum c_n z^n$ where $c_n = \sum_{k+l=n} A_k B_l$. Given $\varepsilon > 0$ we have for certain constants $\tau_1, \tau_2, \tau_3, \tau_4$

$$|A_k| \leq \tau_1(\rho_1 + \varepsilon)^k \leq \tau_1(\rho + \varepsilon)^k,$$

$$|B_l| \leq \tau_2(\rho_2 + \varepsilon)^l \leq \tau_2(\rho + \varepsilon)^l \qquad \text{for all} \quad k, l.$$

Therefore for $n \geq 1$,

$$|C_n| \leq \tau_3 n(\rho + \varepsilon)^n = \tau_3(\rho + 2\varepsilon)^n \cdot n\left(\frac{\rho + \varepsilon}{\rho + 2\varepsilon}\right)^n \leq \tau_4(\rho + 2\varepsilon)^n.$$

Since $\varepsilon > 0$ is arbitrary, this proves that $\rho(AB) \leq \max(\rho(A), \rho(B))$. The proof that $\overline{\rho}(AB) \leq \max(\overline{\rho}(A), \overline{\rho}(B))$ is similar.

(3) Let $1 \leq n(1) < n(2) < \ldots$ be a subsequence such that $\overline{\rho}(A) = \lim_{k\to\infty} \overline{|A_{n(k)}|}^{1/n(k)}$. For each $k \geq 1$ choose $\sigma(k)$ such that $\sigma(k) \in \mathcal{G}$ and

$$\overline{|A_{n(k)}|} \leq st \cdot \left|A_{n(k)}^{\sigma(k)}\right|.$$

Let $1 \leq k_1 \leq k_2 \ldots$ be such that $\sigma(k_1) = \sigma(k_2) = \ldots = \sigma \in \mathcal{G}$. Then

$$\overline{\rho} = \lim_{l\to\infty} \overline{|A_{n(k_l)}|}^{1/n(k_l)} \leq \lim_{l\to\infty} |A_{n(k_l)}|^{1/n(k_l)} \leq \rho(A^\sigma)$$

and so $\overline{\rho}(A) \leq \sup\{\, \rho(A^\sigma) \mid \sigma \in \mathcal{G}\,\}$. The converse inequality $\sup\{\, \rho(A^\sigma) \mid \sigma \in \mathcal{G}\,\} \leq \overline{\rho}(A)$ is clear.

Preliminaries. Let K be an algebraic number field and $b \geq 2$ an integer. Let $\Lambda \in M_{ss}(K)$ be nonsingular and $C \in M_{st}(K)$ satisfy

$$CA_0 = \Lambda C + B_0$$

(we also write $A_0 = A(0)$, $B_0 = B(0)$) where

$$A(z) = \sum_{n=0}^{\infty} A_n z^n\,, \quad A_n \in M_{tt}(K)\,,$$

$$B(z) = \sum_{n=0}^{\infty} B_n z^n\,, \quad B_n \in M_{st}(K)\,.$$

As mentioned before, it is possible to solve

$$\Phi(z) = \sum_{n=0}^{\infty} \Phi_n z^n\,, \quad \Phi_n \in M_{st}(K)\,,$$

$$\Phi(z^b)A(z) = \Lambda\Phi(z) + B(z)\,,$$

$$\Phi_0 = C$$

in formal power series. The solution will be unique and is inductively determined by

$$\Lambda\Phi_{n+1} + B_{n+1} = \sum_{kb+l=n+1} \Phi_k A_l = \sum_{0\leq k\leq (n+1)/b} A_{n+1-kb} \quad (n \geq 0)$$

with $\Phi_0 = C$.

It is easily proved by induction that for each integer $k \geq 1$ one has in formal power series

$$\Phi(z^{b^k})A^{(k)}(z) = \Lambda^k\Phi(z) + B^{(k)}(z)\,, \qquad \Phi_0 = C\,,$$

where

$$A^{(k)}(z) = A(z^{b^{k-1}})\ldots A(z)\,,$$

$$B^{(k+1)}(z) = \Lambda^k B(z) + B^{(k)}(z^b)A(z)\,,$$

$$B^{(1)}(z) = B(z)\,, \qquad A^{(1)}(z) = A(z)$$

with $cA^{(k)}(0) = \Lambda^k C + B^{(k)}(0)$.

LEMMA B. *Let* $\Phi(z) \in M_{st}(K((z)))$, $A(z) \in M_{tt}(K((z)))$, $B(z) \in M_{st}(K((z)))$ *be such that*

$$\Phi(z^b)A(z) = \Lambda\Phi(z) + B(z)$$

with $\Lambda \in M_{ss}(K)$ *nonsingular. Then*

(1) $\rho(\Phi(Z)) \leq \max(1, \rho(A(z)), \rho(B(z)))$,

(2) $\overline{\rho}(\Phi(Z)) \leq \max(1, \overline{\rho}(A(z)), \overline{\rho}(B(z)))$.

Proof. It will first be proved that (1) implies (2). To see this choose $\sigma \in \mathcal{G}$ so that $\overline{\rho}(\Phi(z)) = \rho(\Phi^\sigma(z))$. This is possible by Lemma A. Then

$$\Phi^\sigma(z^b)A^\sigma(z) = \Lambda^\sigma \Phi^\sigma(z) + B^\sigma(z)$$

with $\Lambda^\sigma \in M_{ss}(K)$ nonsingular. Since we are assuming (1) it follows that

$$\overline{\rho}(\Phi(z)) = \rho(\Phi^\sigma(z)) \leq \max(1, \rho(A^\sigma(z)), \rho(B^\sigma(z)))$$
$$\leq \max(1, \overline{\rho}(A(z)), \overline{\rho}(B(z)))\,.$$

It remains to prove (1). It will first be shown that it suffices to prove (1) in the special case that

$$\Phi(z) \in M_{st}(K[[z]])\,, \quad A(z) \in M_{tt}(K[[z]])\,, \quad B(z) \in M_{st}(K[[z]])\,.$$

Indeed, suppose that (1) is established in this special case. Now in the general case choose a positive integer n such that $A_1(z) = A(z)z^{n(b-1)} \in M_{tt}(K[[z]])$. Then

$$\Psi(z^b)A_1(z) = \lambda\Psi(z) + z^{-n}B(z) \qquad \text{where}$$
$$\Psi(z) = z^{-n}\Phi(z)\,.$$

Write

$$\Psi(z) = \sum_\mu \Psi_\mu z^\mu\,, \quad \Psi_{\mathrm{P}}(z) = \sum_{\mu \geq 0} \Psi_\mu z^\mu\,, \quad \Psi^*(z) = \Psi(z) - \Psi_P(z)\,.$$

Then

$$\Psi_{\mathrm{P}}(z^b)A_1(z) = \Lambda\Psi_{\mathrm{P}}(z) + B_1(z) \qquad \text{where}$$
$$B_1(z) = z^{-n}B(z) + \Lambda\Psi^*(z) - \Psi^*(z^b)A(z)\,.$$

It is easy to see that $\rho(B_1(z)) \leq \rho(B(z))$ and that $B_1(z) \in M_{st}(K[[z]])$. By hypothesis it follows that

$$\rho(\Psi_{\mathrm{P}}(z)) \leq \max(1, \rho(A_1(z)), \rho(B_1(z))) \leq \max(1, \rho(A(z)), \rho(B(z)))\,.$$

Consequently $\rho(\Psi(z)) \leq \max(1, \rho(A(z)), \rho(B(z)))$, and therefore $\rho(\Phi(z)) = \rho(\Psi(z)) \leq \max(1, \rho(A(z)), \rho(B(z)))$. It remains to establish (1) in this special case.

We assume that

$$\Phi(z^b)A(z) = \Lambda\Phi(z) + B(z) \tag{1}$$

with

$$(2)\qquad \begin{aligned} &\Lambda \in M_{ss}(K) \qquad \text{nonsingular}\,,\\ &\Phi(z) \in M_{st}(K[[z]])\,, \quad B(z) \in M_{st}(K[[z]])\,, \quad A(z) \in M_{tt}(K[[z]]) \end{aligned}$$

and that $\rho \geq 1$ is such that

$$(3)\qquad \rho(A(z)) \leq \rho\,, \qquad \rho(B(z)) \leq \rho\,.$$

We wish to prove that

$$(4)\qquad \rho(\Phi(z)) \leq \rho\,.$$

It will first be shown that it is enough to prove (4) if we replace (3) by the stronger hypothesis

$$(3)'\qquad \rho \geq 1\,, \quad \rho(A(z)) \leq \rho^{(b-1)/b}\,, \quad \rho(B(z)) \leq \rho\,.$$

To see this, given $\varepsilon > 0$ choose $k \geq 1$ so that

$$\rho^{b^k/(b^k-1)} \leq \rho + \varepsilon\,.$$

Now $\rho(A^{(k)}(z)) \leq \rho$, $\rho(B^{(k)}(z)) \leq \rho$ follows by application of Lemma A and the inductive definition of the $A^{(k)}(z)$, $B^{(k)}(z)$ given in the Preliminaries. We now have

$$\Psi(z^{b^k})A^{(k)}(z) = \Lambda^k\Phi(z) + B^{(k)}(z)$$

with

$$\rho(A^{(k)}(z)) \leq (\rho+\varepsilon)^{(b^k-1)/b^k}\,, \qquad \rho(B^{(k)}(z)) \leq \rho + \varepsilon\,,$$

and so by our hypothesis we deduce that $\rho(\Phi(z)) \leq \rho + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that (4) holds.

It remains to be proved that if (1), (2), (3)′ hold then (4) also holds. This is the content of the following lemma.

LEMMA C. *Let*

$$\rho \geq 1\,, \quad \limsup_{t\to\infty} |A_t|^{1/t} \leq \rho^{(b-1)/b}\,, \quad \limsup_{t\to\infty} |B_t|^{1/t} \leq \rho\,.$$

Then

$$\limsup_{t\to\infty} |\Phi_t|^{1/t} \leq \rho$$

where as usual one solves in formal power series

$$\Phi(z^b)A(z) = \Lambda\Phi(z) + B(z)$$

satisfying (1) *and* (2).

Proof. It has already been remarked that in this case the solution is unique and is inductively determined by

$$\Lambda\Phi_{n+1} + B_{n+1} = \sum_{0\le k\le (n+1)/b} \Phi_k A_l n + 1 - kb \qquad (n \ge 0)$$

with $\Phi_0 = C$ where $C \in M_{st}(K)$ is such that

$$CA_0 = \Lambda C + B_0 .$$

Define $\delta = (3|\Lambda^{-1}|)^{-1}$. Let $\varepsilon > 0$ and determine $C_1 = C_1(\varepsilon) > 0$ such that

$$|A_t| \le C_1 \left(\rho + \frac{\varepsilon}{2}\right)^{t(b-1)/b} \quad \text{for all} \quad t \ge 0 .$$

Choose K_0 to be a positive integer so large that

(1) $\left(\frac{\rho+\varepsilon/2}{\rho+\varepsilon}\right)^{k_0(b-1)} \le \delta$,

(2) $\sum_{k>k_0} C_1(\rho+\varepsilon/2)^{-k(b-1)} \le \delta$.

Define $C = \sum_{k=0}^{k_0} |\Phi_k|$. Choose C^* to be a positive real number satisfying

(3) $C^* \ge C$, $C^* \ge CC_1$,

(4) $C \max_{0\le t\le 2k_0 b+1} |A_t| \le \delta C^*$,

(5) $|B_t| \le \delta C^*(\rho+\varepsilon)^t$ for all $t \ge 0$.

It will be shown that

$$|\Phi_n| \le C^*(\rho+\varepsilon)^n \quad \text{for all } n \ge 0 ,$$

which will complete the proof since $\varepsilon > 0$ is arbitrary.

If $0 \le n \le k_0$ then

$$|\Phi_n| \le C \le C^* \le C^*(\rho+\varepsilon)^n .$$

Now let n be a positive integer such that for all integers m, $0 \le m \le n$, one has

$$|\Phi_m| \le C^*(\rho+\varepsilon)^m .$$

It will be proved that $|\Phi_{n+1}| \le C^*(\rho+\varepsilon)^{n+1}$. We have

$$\Lambda\Phi_{n+1} + B_{n+1} = \sum_{0\le k\le (n+1)/b} \Phi_k A_{n+1-kb} = S_1 + S_2$$

where

$$S_1 = \sum_{k_0<k\le (n+1)/b} \Phi_k A_{n+1-kb} , \qquad S_2 = \sum_{0\le k_0\le k} \Phi_k A_{n+1-kb} .$$

It will now be shown that $|S_2| \le \delta C^*(\rho+\varepsilon)^{n+1}$. Indeed, if $n \ge 2k_0 b$ then

$$|S_2| \le C \max_{0\le k\le k_0} |A_{n+1-kb}| \le \delta C^* \le \delta C^*(\rho+\varepsilon)^{n+1}.$$

But if $n \ge 2k_0 b + 1$ we obtain

$$\begin{aligned}
|S_2| &\le C \max_{0\le k\le k_0} |A_{n+1-kb}| \le CC_1 \max_{0\le k\le k_0} \left(\rho+\frac{\varepsilon}{2}\right)^{(n+1-kb)(b-1)/b} \\
&= CC_1 \max_{0\le k\le k_0} (\rho+\varepsilon)^{(n+1-kb)(b-1)/b} \left(\frac{\rho+\varepsilon/2}{\rho+\varepsilon}\right)^{(n+1-kb)(b-1)/b} \\
&\le CC_1(\rho+\varepsilon)^{n+1}\left(\frac{\rho+\varepsilon/2}{\rho+\varepsilon}\right)^{k_0(b-1)} \\
&\le CC_1\delta(\rho+\varepsilon)^{n+1} \le \delta C^*(\rho+\varepsilon)^{n+1}.
\end{aligned}$$

This proves that $|S_2| \le \delta C^*(\rho+\varepsilon)^{n+1}$ in all cases.

Now by the definition of S_1,

$$|S_1| \le \sum_{k_0\le k\le (n+1)/b} C^*(\rho+\varepsilon)^{(n+1)/b} \cdot C_1 \left(\rho+\frac{\varepsilon}{2}\right)^{n+1-kb(b-1)/b}$$

using the inductive hypothesis. It follows that

$$|S_1| \le \sum_{k_0\le k\le (n+1)/b} C^* C_1(\rho+\varepsilon)^{n+1} \cdot \left(\rho+\frac{\varepsilon}{2}\right)^{-kb(b-1)/b} \le \delta C^*(\rho+\varepsilon)^{n+1}.$$

Therefore

$$|\Lambda\Phi_{n+1} + B_{n+1}| \le |S_1| + |S_2| \le 2\delta C^*(\rho+\varepsilon)^{n+1}$$

and so $|\Lambda\Phi_{n+1}| \le 3\delta C^*(\rho+\varepsilon)^{n+1}$. Consequently

$$|\Phi_{n+1}| \le 3|\Lambda^{-1}|\delta C^*(\rho+\varepsilon)^{n+1} \le C^*(\rho+\varepsilon)^{n+1}$$

and Lemma C, and so Lemma B, is now proved.

References

[B] P.-G. Becker-Landeck, *Maße für algebraische Unabhängigkeit nach einer Methode von Mahler*, Acta Arith. 50 (1988), 279–293.

[Br] W. D. Brownawell, *Effectivity in independence measures for values of E-functions*, J. Austral. Math. Soc. Ser. 39 (1985), 227–240.

[Ch] G. Chudnovsky, *Contributions to the Theory of Transcendental Numbers*, Surveys and Monographs, 19, AMS, 1984.

[G] A. I. Galochkin, *A transcendence measure for the values of functions satisfying certain functional equations*, Mat. Zametki 27 (1980), 175–183; English transl. in Math. Notes 27 (1980).

[K] K. K. Kubota, *On the algebraic independence of holomorphic solutions of certain functional equations and their values*, Math. Ann. 227 (1970), 9–50.

[La1] S. Lang, *Introduction to Transcendental Numbers*, Addison-Wesley, 1966.

[La2] —, *A transcendence measure for E-functions*, Mathematika 9 (1962), 157–161.

[L] J. H. Loxton, *Automata and Transcendence*, Chapter 13 of: New Advances in Transcendence Theory, A. Baker (ed.), Cambridge Univ. Press, 1988.

[LP1] J. H. Loxton and A. J. van der Poorten, *Algebraic independence properties of the Fredholm series*, J. Austral. Math. Soc. Ser. A 26 (1978), 31–45.

[LP2] —, —, *Arithmetic properties of certain functions in several variables III*, Bull. Austral. Math. Soc. 16 (1977), 15–47.

[LP3] —, —, *A class of hypertranscendental functions*, Aequationes Math. 16 (1977), 93–106.

[Ma1] K. Mahler, *Remarks on a paper by W. Schwarz*, J. Number Theory 1 (1969), 512–521.

[Ma2] —, *Arithmetische Eigenschaften der Lösungen einer Klasse von Funktionalgleichungen*, Math. Ann. 101 (1929), 342–366.

[Ma3] —, *Arithmetische Eigenschaften einer Klasse transzendental-transzendenter Funktionen*, Math. Z. 32 (1930), 545–585.

[Ma4] —, *Über das Verschwinden von Potenzreihen mehrerer Veränderlichen in speziellen Punktfolgen*, Math. Ann. 103 (1930), 573–587.

[M1] W. Miller, *Transcendence measures for values of analytic solutions to certain functional equations*, Ph. D. Thesis, University of Michigan 1979.

[M2] —, *Transcendence measures by a method of Mahler*, J. Austral. Math. Soc. Ser. A 32 (1982), 68–72.

[Ne1] Yu. V. Nesterenko, *Estimates for the orders of zeros of functions of a certain class and applications in the theory of transcendental numbers*, Izv. Akad. Nauk SSSR Ser. Mat. 41 (1977), 253–284; English transl. in Math. USSR-Izv. 11 (1977).

[Ne2] —, *Estimates for the characteristic function of a prime ideal*, Mat. Sb. 123 (165) (1984), 11–49; English transl. in Math. USSR-Sb. 51 (1) (1985).

[Ne3] —, *On algebraic independence of algebraic powers of algebraic numbers*, Mat. Sb. 123 (165) (1984), 435–459; English transl. in Math. USSR-Sb. 51 (1985).

[Ne4] —, *On a measure of the algebraic independence of the values of certain functions*, Mat. Sb. 128 (170) (1985), 545–568; English transl. in Math. USSR-Sb. 56 (2) (1987).

[Sh] A. V. Shidlovskiĭ, *On criteria for algebraic independence of a class of entire functions*, Izv. Akad. Nauk SSSR Ser. Mat. 23 (1959), 35–66; English transl. in Amer. Math. Soc. Transl. (2) 22 (1962).

[Sie] C. L. Siegel, *Approximation algebraischer Zahlen*, Math. Z. 10 (1921), 173–213.

[Wae] B. L. van der Waerden, *Modern Algebra*, Vols. 1 and 2, F. Ungar, New York 1950 and 1953.

[Wa] M. Waldschmidt, *Nombres Transcendants*, Springer, Berlin–New York 1974.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN 48109–1003, U.S.A.

Current address: 57 Arapito Rd, Titirangi, Auckland 7, New Zealand