ACTA ARITHMETICA
L (1988)

Maße für algebraische Unabhängigkeit nach einer Methode von Mahler

von

PAUL-GEORG BECKER-LANDECK (Köln)

1. Einleitung. Mahler zeigte um 1930 in mehreren Arbeiten [4], [5], [6] die Transzendenz und algebraische Unabhängigkeit der Werte gewisser transzendenter Funktionen an algebraischen Stellen. Bei diesen Funktionen handelt es sich um in einer Umgebung des Nullpunktes von C^n konvergente Potenzreihen, die Funktionalgleichungen eines bestimmten Typs genügen. Weitreichende Verallgemeinerungen fanden diese Ergebnisse im Laufe des vergangenen Jahrzehnts; wie von Loxton und van der Poorten in ihrem Übersichtsartikel [3] ausführlich beschrieben, gelang es dabei, die Klasse der für die Betrachtung zulässigen Funktionen beträchtlich zu erweitern.

Es war nun naheliegend, nach Verschärfungen dieser rein qualitativen Ergebnisse über Transzendenz und algebraische Unabhängigkeit in quantitativer Hinsicht zu fragen. Erste Resultate in dieser Richtung stammen von Galochkin [1] und Miller [8], die für einige der von Mahler untersuchten Zahlen Transzendenzmaße angeben konnten. Andererseits blieben Maße für die algebraische Unabhängigkeit derartiger Zahlen unbekannt, worauf etwa Waldschmidt ([10], Seite II, 10) ausdrücklich hinwies.

In der vorliegenden Arbeit wird nun ein solches Unabhängigkeitsmaß hergeleitet. Dabei wird im wesentlichen eine quantitative Version des Ergebnisses von Loxton und van der Poorten aus [2] gezeigt.

Da die Formulierung des Hauptsatzes einiger Vorbereitungen und Definitionen bedarf, soll unser Resultat zunächst nur an folgendem Beispiel illustriert werden. Seien für $k \in N$ die Funktionen G_k durch

$$G_k(z) = \sum_{h=0}^{\infty} \left(\frac{z^{2^h}}{1+z^{2^h}}\right)^k$$

erklärt und sei α eine algebraische Zahl mit $0 < |\alpha| < 1$. Dann gilt für jedes Polynom $P \in Z[X_1, \ldots, X_K]$, $P \neq 0$, mit Grad in jeder Variablen X_i höchstens d und Höhe höchstens H die Abschätzung

$$|P(G_1(\alpha), \ldots, G_K(\alpha))| > C_1 H^{-C_2}.$$

Dabei sind C_1 und C_2 von d abhängige positive Zahlen und die d-Abhängigkeit von C_2 kann effektiv angegeben werden.

Der Beweis unseres Satzes orientiert sich am Beweis des Theorems in [2]. Zunächst wird eine Transformation Ω eingeführt, die die Funktionalgleichungen der verschiedenen zu untersuchenden Funktionen miteinander verbindet. Mittels dieser Transformation erklärt man den Index einer holomorphen Funktion als Verallgemeinerung der üblichen Nullstellenordnung. Eine neue Version des Siegelschen Lemmas erlaubt dann die Konstruktion einer Hilfsfunktion mit „hohem" Index. Ausgehend von diesem „hohen" Index erhält man eine analytische obere Abschätzung, die in Verbindung mit einer Liouville-Abschätzung nach unten die im Hauptsatz behauptete Ungleichung liefert.

Die vorliegende Arbeit gibt einen Einblick in den zweiten Teil meiner bei Prof. Dr. P. Bundschuh angefertigten Dissertation [0]. Ihm möchte ich für seine Anregungen und seine Unterstützung herzlich danken.

2. Bezeichnungen und Formulierung des Hauptsatzes. Für eine algebraische Zahl α wird das Maximum der Konjugierten von α als Haus von α bezeichnet, symbolisch $\overline{|\alpha|}$. Ist P ein Polynom mit algebraischen Koeffizienten, so ist die Höhe $H(P)$ von P erklärt als das Maximum der Beträge der Koeffizienten von P und die Länge $\Lambda(P)$ von P ist die Summe der Häuser aller Koeffizienten.

Für Teilmengen L von C und $s_1, s_2 \in N$ bezeichne $M(L, s_1, s_2)$ die Menge aller $(s_1 \times s_2)$-Matrizen mit Elementen aus L.

Sind $\boldsymbol{v}, \boldsymbol{u} \in \boldsymbol{R}^n$, $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{Z}^n$, so seien folgende Vereinbarungen getroffen:

$$\boldsymbol{v} \geqslant \boldsymbol{0} \quad \text{bedeute} \quad v_1 \geqslant 0, \ldots, v_n \geqslant 0 \qquad (\text{analog } \boldsymbol{v} > \boldsymbol{0},\ \boldsymbol{v} \geqslant \boldsymbol{u},\ \boldsymbol{v} > \boldsymbol{u});$$

$$\sum_{i=a}^{b} := \sum_{i_1=a_1}^{b_1} \ldots \sum_{i_n=a_n}^{b_n}, \qquad \text{falls } \boldsymbol{b} \geqslant \boldsymbol{a}.$$

Sind $\boldsymbol{w} \in \boldsymbol{C}^n$ und $\boldsymbol{\mu} \in \boldsymbol{N}_0^n$, $N_0 := N \cup \{0\}$, so bezeichnet $\boldsymbol{w}^{\boldsymbol{\mu}}$ das Potenzprodukt $w_1^{\mu_1} \cdot \ldots \cdot w_n^{\mu_n}$.

Sind $f_1(z), \ldots, f_l(z)$ Funktionen, so ist unter $f(z)$ der Vektor $(f_1(z), \ldots, f_l(z))$ zu verstehen.

Die folgenden Definitionen und Erläuterungen bereiten die Formulierung des Hauptsatzes vor.

Definition 1. Bei $T \in M(N_0, n, n)$ und $z \in C^n$ erklärt man $w = Tz$ durch $w_i := \prod_j z_j^{t_{ij}}$ für $1 \leqslant i \leqslant n$.

Im Falle $n = 1$ entspricht T der Transformation $z \to z^r$ mit $r \in N_0$. Für $T \in M(N_0, n, n)$ setzt man als Spektralradius $r(T)$ das Maximum der Beträge der Eigenwerte von T fest. Die weiteren Untersuchungen werden auf eine Teilmenge $\mathscr{T}$ von $M(N_0, n, n)$ beschränkt.

Definition 2. Eine Matrix $T \in M(N_0, n, n)$ soll zur Menge $\mathscr{T}$ gehören, wenn sie regulär ist, einen positiven Eigenvektor zum Eigenwert $r(T)$ besitzt und kein Eigenwert eine Einheitswurzel ist.

Sei im folgenden $T \in \mathscr{T}$ vorausgesetzt. Man betrachtet eine teilweise Spektralzerlegung von T mit

$$T = \sum_{j=1}^{s} \lambda_j E_j + F.$$

$\lambda_1, \ldots, \lambda_s$ sind dabei die Eigenwerte von T mit Betrag $r(T)$, $E_1, \ldots, E_s$ sind die kanonischen Projektionen auf die zugehörigen Eigenräume und F ist eine Matrix mit $r(F) < r(T)$. Mit $U := \bigoplus_{j=1}^{s} \operatorname{im} E_j$ und $V := \bigcap_{j=1}^{s} \ker E_j$ gilt $C^n = U \oplus V$. Die Abbildung $\Pi\colon C^n \to U$, die als Projektion auf U entlang V definiert wird, bezeichnet man dann als Projektion auf den dominanten Eigenraum von T.

Man definiert nun $U(T) \subset C^n$ so, daß für $z \in U(T)$ die Folge $(T^k z)_{k \in N_0}$ in geeigneter Weise gegen $\boldsymbol{0}$ konvergiert.

Definition 3. Erklärt man für $z \in (C^*)^n$ die Abbildung

$$L(z) := (-\log|z_1|, \ldots, -\log|z_n|),$$

so bezeichnet $U(T)$ die Menge aller z mit $\Pi(L(z)) > \boldsymbol{0}$.

Um später das Nichtverschwinden der Hilfsfunktion an geeigneten Stellen garantieren zu können, muß eine weitere Forderung an die betrachteten Punkte gestellt werden.

Definition 4. Genau dann soll $\boldsymbol{\alpha} \in C^n$ die *Eigenschaft* $(\mathscr{B})$ besitzen, wenn für jede in einer Umgebung von $\boldsymbol{0}$ holomorphe, nicht identisch verschwindende Funktion f ein $k_0 = k_0(f, \boldsymbol{\alpha})$ existiert mit $f(T^k \boldsymbol{\alpha}) \neq 0$ für alle $k \geqslant k_0$.

Die Eigenschaft $(\mathscr{B})$ stellt eine Verschärfung der von Loxton und van der Poorten [2] verlangten Eigenschaft $(\mathscr{A})$ dar; wie von Masser [7] gezeigt wurde, ist sie stets dann erfüllt, wenn kein $\boldsymbol{\mu} \in \boldsymbol{Z}^n \setminus \{\boldsymbol{0}\}$ existiert mit $(T^k \boldsymbol{\alpha})^{\boldsymbol{\mu}} = 1$ für alle k einer festen arithmetischen Folge. Im Falle $n = 1$ jedoch besitzt jedes $\alpha \in C^*$ trivialerweise die Eigenschaft $(\mathscr{B})$ und damit auch die Eigenschaft $(\mathscr{A})$.

Satz. *Seien $a_1, \ldots, a_p$ von Null verschiedene algebraische Zahlen und $b_1(z), \ldots, b_p(z)$ rationale Funktionen in $z_1, \ldots, z_n$ mit algebraischen Koeffizienten. Seien für eine Matrix $T \in \mathscr{T}$ die algebraisch unabhängigen Funktionen $f_1(z), \ldots, f_p(z)$ in einer Umgebung von $\boldsymbol{0}$ holomorphe Lösungen des Funktionalgleichungssystems*

$$(1) \qquad f_i(z) = a_i f_i(Tz) + b_i(z) \qquad (1 \leqslant i \leqslant p),$$

und seien sämtliche $f_i(\boldsymbol{0})$ algebraisch. Es möge $\boldsymbol{\alpha} \in U(T)$ die Eigenschaft $(\mathscr{B})$ besitzen und $f_i(\boldsymbol{\alpha})$, $b_i(T^k \boldsymbol{\alpha})$ seien für $1 \leqslant i \leqslant p$, $k \in N$ definiert.

Dann gilt für jedes Polynom $Q \in Z[X_1, \ldots, X_p]$, $Q \neq 0$, *mit* $\partial_{X_i} Q \leqslant d$ $(1 \leqslant i \leqslant p)$ *und* $H(Q) \leqslant H$ *die Abschätzung*

$$|Q(f_1(\alpha), \ldots, f_p(\alpha))| > C_2(d) H^{-C_1(d)}.$$

Dabei ist mit positiven, von Q unabhängigen reellen Zahlen $\varkappa_1, \varkappa_2$ *gesetzt*

$$\log C_1(d) := \begin{cases} \varkappa_1^{d^p}, & \text{falls alle } a_i \text{ komplexe Einheitswurzeln sind,} \\ \varkappa_2^{d^p} & \text{sonst;} \end{cases}$$

$C_2(d)$ *ist eine von H unabhängige, reelle positive Zahl.*

Während der Satz ein Maß für die algebraische Unabhängigkeit von Funktionswerten mehrerer Funktionen an einer festen algebraischen Stelle liefert, erhält man ein analoges Resultat für die Werte einer festen Funktion an mehreren algebraischen Stellen. Dies kann in der von Loxton und van der Poorten in [2] beschriebenen Weise aus dem Satz abgeleitet werden. In der zitierten Arbeit finden sich weitere Funktionen, auf die unser Resultat anwendbar ist.

3. Vorbemerkungen zum Beweis. (a) Um technische Kompliziertheit zu vermeiden, beschränkt sich der Beweis auf den Fall $n = 1$, d.h. es werden nur Funktionen einer komplexen Variablen betrachtet.

(b) Durch eventuellen Übergang von $f_i(z)$ zu $f_i(z) - f_i(0)$ ist es möglich für alle Funktionen o.B.d.A. $f_i(0) = 0$, $1 \leqslant i \leqslant p$, anzunehmen.

(c) Ebenso kann o.B.d.A. vorausgesetzt werden, daß die von $a_1, \ldots, a_p$ erzeugte multiplikative Gruppe isomorph zu $Z^{\tilde{p}}$ ist bei geeignetem $\tilde{p} \in N_0$. Dies zeigt man, indem man gegebenenfalls mit geeignetem $l \in N$ die Transformation T durch die Transformation T^l ersetzt.

(d) Die im folgenden auftretenden Konstanten $C_3, C_4, \ldots$ und $\gamma_1, \gamma_2, \ldots$ bezeichnen reelle Zahlen größer als Eins. Mit $D_1, D_2, \ldots$ und $\delta_1, \delta_2, \ldots$ seien natürliche Zahlen bezeichnet. Sie alle sind lediglich von den algebraischen Zahlen $a_1, \ldots, a_p$, α, den rationalen Funktionen $b_1(z), \ldots, b_p(z)$ sowie den Elementen der Matrix T abhängig. Insbesondere sind sie unabhängig von den Hilfsparametern ϱ und k und den durch das Polynom Q des Satzes festgelegten Größen. Der einzige Unterschied zwischen den C_i bzw. D_i und γ_j bzw. δ_j besteht darin, daß die Bedeutung der C_i bzw. D_i fest ist, während mit der Numerierung der γ_j bzw. δ_j bei jedem Beweis eines Lemmas erneut mit γ_1 bzw. δ_1 begonnen wird.

(e) K bezeichnet denjenigen algebraischen Zahlkörper, den die in der Formulierung des Satzes auftretenden algebraischen Größen erzeugen; unter I_K versteht man den Ring der ganzen Elemente von K.

4. Einige vorbereitende Lemmata. Sei unter den Voraussetzungen des Satzes

$$Q(X) = \sum_{\mu=0}^{m} \omega_\mu X^\mu \in Z[X_1, \ldots, X_p].$$

Man führt nun $m := (m_1+1) \cdot \ldots \cdot (m_p+1)$ neue Variablen w_μ, $0 \leqslant \mu \leqslant m$, ein und setzt $\hat{m} := m_1 + \ldots + m_p$ sowie

$$F(z, w) := \sum_{\mu=0}^{m} w_\mu f(z)^\mu. \tag{2}$$

Durch einmalige Anwendung der Funktionalgleichungen (1) erhält man

$$F(z, w) = \sum_{\mu=0}^{m} a^\mu \Big(\sum_{\nu=\mu}^{m} \binom{\nu}{\mu} b(z)^{\nu-\mu} w_\nu \Big) f(Tz)^\mu, \quad \text{wo} \quad \binom{\nu}{\mu} := \prod_{i=1}^{p} \binom{\nu_i}{\mu_i}. \tag{3}$$

Erklärt man mit Variablen $u_1, \ldots, u_p$ und $y_1, \ldots, y_p$ für $k \in N$ die Transformation

$$\Omega_k(y, u, w) := \Big(y^{k\mu} \sum_{\nu=\mu}^{m} \binom{\nu}{\mu} u^{\nu-\mu} w_\nu \Big)_{0 \leqslant \mu \leqslant m}$$

und definiert für $1 \leqslant i \leqslant p$ und $k \in N$

$$b_i^{(k)}(z) := \sum_{j=0}^{k-1} a_i^j b_i(T^j z),$$

so folgt durch iterierte Anwendung der Funktionalgleichungen (1) analog zu (3)

$$F(z, w) = F\big(T^k z, \Omega_k(a, b^{(k)}(z), w)\big)$$

für $k \in N$. Mit ω, dem Vektor der Koeffizienten von Q, sei $\Omega_k(u) := \Omega_k(a, u, \omega)$ und $\omega^{(k)} := \Omega_k(b^{(k)}(\alpha))$ gesetzt; es ergibt sich damit für $k \in N$

$$Q(f(\alpha)) = F(\alpha, \omega) = F(T^k \alpha, \omega^{(k)}).$$

Um im folgenden F an den Stellen $(T^k \alpha, \omega^{(k)})$ zu untersuchen, erweist sich die folgende Definiton als nützlich.

Definition 5. Sei $\mathscr{P} := K[w]$ der Ring der Polynome in w mit Koeffizienten in K. Es bezeichne $\mathscr{P}(\omega)$ die Menge aller Polynome $p \in \mathscr{P}$, für die $p(\Omega_k(u))$ für alle $k \in N$ identisch verschwindet. $\mathscr{P}_\varrho$ und $\mathscr{P}_\varrho(\omega)$ bezeichnen für jedes $\varrho \in R_+$ die Teilmengen der Polynome aus $\mathscr{P}$ und $\mathscr{P}(\omega)$, deren Grad in jedem w_μ höchstens ϱ ist.

Das folgende Lemma gestattet es, die Menge $\mathscr{P}_\varrho(\omega)$ dadurch zu charakterisieren, daß die Koeffizientenvektoren der Polynome aus $\mathscr{P}_\varrho(\omega)$ gewisse lineare Bedingungen erfüllen müssen.

Lemma 1. *Für jedes* $\varrho \in N$ *gibt es eine Matrix* M_ϱ *mit folgenden Eigenschaften:*

(a) *Es ist* $M_\varrho \in M(Z, r_1, r_2)$ *mit* $r_2 := (\varrho+1)^m$ *und*

$$\prod_{i=1}^{p} (mm_i \varrho + 1) \leqslant r_1 \leqslant \prod_{i=1}^{p} (mm_i \varrho + 1)^2.$$

(b) *Für alle* $p(w) = \sum_{s=0}^{\varrho} Y_s w^s \in \mathscr{P}_\varrho$ *mit* $\varrho = (\varrho, \ldots, \varrho) \in N_0^m$ *gilt* $p \in \mathscr{P}_\varrho(\omega)$ *genau dann, wenn* $M_\varrho \cdot Y = 0$.

(c) $\log \|M_\varrho\|_\infty \leqslant C_3 m\varrho(\hat{m} + \log H)$.

(d) $\operatorname{rg} M_\varrho > 0$.

(e) $\operatorname{rg} M_{2\varrho} \leqslant 2^m \operatorname{rg} M_\varrho$.

Beweis. Für jedes $p \in \mathscr{P}_\varrho$ gilt

$$p(\Omega_k(u)) = \sum_{j=1}^{N} d_j^k p_j(u)$$

mit N paarweise verschiedenen $d_j \in C^*$, $N \leqslant \prod_{i=1}^{p} (mm_i \varrho + 1)$, und von k unabhängigen Polynomen $p_j(u)$ in $u_1, \ldots, u_p$, mit $\partial_{u_i} p_j \leqslant mm_i \varrho$, $1 \leqslant i \leqslant p$. (Vgl. Loxton und van der Poorten [2], Beweis zu Lemma 4.) Aus der Bedingung des identischen Verschwindens von $p(\Omega_k(u))$ für alle $k \in N$ folgt

$$\begin{bmatrix} d_1 & \ldots & d_N \\ \ldots & \ldots & \ldots \\ d_1^N & \ldots & d_N^N \end{bmatrix} \begin{bmatrix} p_1(u) \\ \vdots \\ p_N(u) \end{bmatrix} = 0.$$

Aus der Regularität dieser Matrix erhält man $p_j(u) \equiv 0$, $1 \leqslant j \leqslant N$, als notwendige und hinreichende Bedingung für $p \in \mathscr{P}_\varrho(\omega)$. Ist $p(w)$ wie in (b) dargestellt, so überlegt man leicht, daß die Koeffizienten der Polynome $p_j(u)$ Linearformen in den Größen Y_s, $0 \leqslant s \leqslant (\varrho, \ldots, \varrho) \in N^m$, sind, und die Bedingungen an die $p_j(u)$ daher äquivalent dazu sind, daß die Y_s einem homogenen linearen Gleichungssystem mit ganzrationalen Koeffizienten genügen. Aus der Definition von Ω_k ergibt sich die in (c) behauptete Abschätzung für die Koeffizienten der Matrix M_ϱ, die dieses Gleichungssystem beschreibt. Damit sind die Teile (a) bis (c) des Lemmas gezeigt. Teil (d) folgt sofort, da M_ϱ wegen $1 \notin \mathscr{P}_\varrho(\omega)$ stets von der Nullmatrix verschieden ist. Teil (e) ist wegen $\operatorname{rg} M_\varrho = \dim_K \mathscr{P}_\varrho - \dim_K \mathscr{P}_\varrho(\omega)$ äquivalent zu Lemma 5 bei Loxton und van der Poorten [2].

Definition 6. Sei $\mathscr{A}$ der Ring der Potenzreihen

$$E(z, w) = \sum_{i=0}^{\infty} p_i(w) z^i$$

in z mit Koeffizienten $p_i(w) \in \mathscr{P}$, die in einer Umgebung des Ursprungs $z = 0$ konvergieren. Dann definiert man den Index der Funktion $E(z, w) \in \mathscr{A}$ im Punkt ω durch $\operatorname{ind} E(z, w) := \min\{i \mid p_i(w) \notin \mathscr{P}(\omega)\}$ und $\operatorname{ind} E(z, w) := \infty$, falls dieses Minimum nicht existiert.

Lemma 2. *Bei* $E_1, E_2 \in \mathscr{A}$ *gelten die Aussagen*

(a) $\operatorname{ind}(E_1 \cdot E_2) = \operatorname{ind} E_1 + \operatorname{ind} E_2$,

(b) $\operatorname{ind}(E_1 + E_2) \geqslant \min(\operatorname{ind} E_1, \operatorname{ind} E_2)$ *mit Gleichheit, falls*

$$\operatorname{ind} E_1 \neq \operatorname{ind} E_2.$$

(c) *Für* $F(z, w)$, *wie in* (2) *erklärt, ist* $\operatorname{ind} F(z, w) = 0$.

Beweis. Aus den in Vorbemerkung (c) beschriebenen Annahmen läßt sich zeigen, daß $\mathscr{P}(\omega)$ ein Primideal in $\mathscr{P}$ ist. Die Eigenschaften (a) und (b) lassen sich daraus dann mittels der üblichen Schlüsse herleiten. Aus Vorbemerkung (b) weiß man, daß $f_i(0) = 0$, $1 \leqslant i \leqslant p$, vorausgesetzt werden darf; Teil (c) ergibt sich daraus durch elementares Nachrechnen.

Zur Konstruktion der Hilfsfunktion ist es notwendig, die Beträge und Nenner der Potenzreihenkoeffizienten von $F(z, w)^j$ geeignet nach oben abzuschätzen.

Lemma 3. *Für jedes* $j \in N_0$ *besitzt* $F(z, w)^j$ *eine Darstellung der Form*

$$F(z, w)^j = \sum_{i=0}^{\infty} B_{ij}(w) z^i$$

mit Polynomen B_{ij} *aus* $\mathscr{P}_j$. *Diese haben die Eigenschaften*

(a) $\Lambda(B_{ij}) \leqslant (C_4 i)^{C_5 j \hat{m}} C_6^i$ *für* $i \in N$,

(b) $\Lambda(B_{0j}) \leqslant C_7$,

(c) $D_1^i B_{ij} \in I_K[w]$ *für* $i \in N_0$.

Beweis. Sei $b(z) \in K(z)$ mit $b(0) = 0$ eine in einer Umgebung von 0 holomorphe Funktion mit der Potenzreihenentwicklung

$$b(z) = \sum_{\lambda=0}^{\infty} b_\lambda z^\lambda.$$

Dann gilt

$$\boxed{b_\lambda} \leqslant \gamma_1^{\lambda+1} \quad \text{und} \quad \delta_1^{\lambda+1} b_\lambda \in I_K. \tag{4}$$

Gilt für eine bei 0 holomorphe Funktion $f(z) = \sum f_i z^i$ mit $a \in K$ die Funktionalgleichung $f(z) = af(Tz) + b(z)$, so folgt für jedes $i \in N$

$$f_i = \sum_{k,\lambda} a^k b_\lambda,$$

wobei über alle Paare (k, λ) mit $r^k \lambda = i$ summiert wird; man beachte $T = (r)$ im Fall $n = 1$. Aus (4) ergibt sich

$$\boxed{f_i} \leqslant \gamma_2^i \quad \text{und} \quad \delta_2^i f_i \in I_K. \tag{5}$$

Da die Funktionen $f_j(z)$ aus dem Satz Funktionalgleichungen des obigen Typs erfüllen, gilt auch für ihre Potenzreihenkoeffizienten $f_i^{(j)}$ eine zu (5) analoge Aussage. Daraus erhält man für die durch

$$f_1(z)^{\mu_1} \cdot \ldots \cdot f_p(z)^{\mu_p} = \sum_{i=0}^{\infty} f_{i\mu} z^i \quad (\mu \in N_0^p)$$

erklärten Koeffizienten $f_{i\mu}$, $i \in N$, mit $|\mu| := \sum_{j=1}^{p} \mu_j$

(6) $$\overline{|f_{i\mu}|} \leqslant i^{\gamma_3|\mu|}\gamma_4^i \quad \text{und} \quad \delta_3^i f_{i\mu} \in I_K.$$

Für die in der Formulierung des Lemmas erklärten $B_{i1}(w)$ gilt

$$B_{i1}(w) = \sum_{\mu=0}^{m} w_\mu f_{i\mu},$$

und aus (6) folgen für $j = 1$ die Behauptungen (a) und (c). Teil (b) ist wegen $f_j(0) = 0$, $1 \leqslant j \leqslant p$, trivial. Man setzt nun

$$B_{ij}(w) := \sum B_{i_1 1}(w) \ldots B_{i_j j}(w),$$

wobei über alle j-Tupel $(i_1, \ldots, i_j) \in N_0^j$ mit $i_1 + \ldots + i_j = i$ summiert wird. Daraus ergibt sich die Behauptung für alle $j \in N_0$.

Lemma 4. *Seien $D, s_1, s_2, n \in N$ und L ein algebraischer Zahlkörper. Es gelte für die Matrizen $A_1 \in M(L, s_1, n)$ und $A_2 \in M(L, s_2, n)$:*

(a) $r_1 := \operatorname{rg} A_1 < \operatorname{rg}\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} =: r_2$.

(b) $D \cdot A_1 \in M(I_L, s_1, n)$.

(c) *$S \in R^+$ ist eine obere Schranke für die Häuser der Elemente von A_1. Dann besitzt das System*

$$A_1 X = 0 \quad \text{und} \quad A_2 X \neq 0$$

eine Lösung $x \in L^n$ mit ganzen Koordinaten, die mit einer nur von L abhängigen Konstanten C_8 der folgenden Abschätzung genügen

$$\max_{1 \leqslant \nu \leqslant n} \overline{|x_\nu|} \leqslant (C_8 r_2 DS)^{r_1/(r_2 - r_1)}.$$

Beweis. $D \cdot S$ ist eine obere Schranke für die Häuser der Elemente von $D \cdot A_1$. Man setzt nun

$$A := \begin{bmatrix} DA_1 \\ A_2 \end{bmatrix} \in M(L, s_1 + s_2, n)$$

und konstruiert aus A eine neue Matrix B, indem man, beginnend mit der ersten Zeile von A, die Zeilen streicht, die L-linear abhängig von den über ihnen stehenden Zeilen sind. Für B gilt

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \in M(L, r_2, n), \qquad \operatorname{rg} B = r_2,$$

mit einer Matrix $B_1 \in M(I_L, r_1, n)$, die bei Anwendung des gleichen Verfahrens aus $D \cdot A_1$ entsteht. Wählt man in analoger Weise aus den Spalten von B nun r_2 linear unabhängige aus, so gilt für die entstehende Matrix C:

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \in M(L, r_2, r_2), \qquad \operatorname{rg} C = r_2,$$

und $C_1 \in M(I_L, r_1, r_2)$. Mittels des Siegelschen Lemmas (vgl. Schneider [9], Hilfssatz 31) läßt sich nun das homogene lineare Gleichungssystem $C_1 u = 0$ nichttrivial durch ein $u \in I_L^{r_2}$ mit

$$\max_{1 \leqslant i \leqslant r_2} \overline{|u_i|} \leqslant (C_8 r_2 DS)^{r_1/(r_2 - r_1)}$$

lösen. Aus der Invertierbarkeit von C folgt $Cu \neq 0$ und damit $C_2 u \neq 0$. Man erhält daher eine Lösung $x \in I_L^n$ des Ausgangssystems, indem man $x_i = 0$ setzt, wenn beim Übergang von B nach C die i-te Spalte gestrichen wurde und $x_i = u_j$ setzt, wenn die i-te Spalte von B in die j-te Spalte von C übergegangen ist.

5. Konstruktion einer Hilfsfunktion. Nach den Vorbereitungen durch die vorangestellten Lemmata ist es nun möglich, eine geeignete Hilfsfunktion zu konstruieren.

Lemma 5. *Zu jedem $\varrho \in N$ mit $\varrho \geqslant C_9^m$ gibt es Polynome*

$$p_0(z, w), \ldots, p_\varrho(z, w) \in I_K[z, w],$$

so daß gilt:

(a) $\operatorname{ind} p_0(z, w) \neq \infty$.

(b) $\operatorname{ind} E_\varrho(z, w) \geqslant C_{10}^{-m} \varrho^{3/2} =: G$, *wobei gesetzt ist*

$$E_\varrho(z, w) := \sum_{j=0}^{\varrho} p_j(z, w) F(z, w)^j.$$

(c) *Für $0 \leqslant j \leqslant \varrho$ und $0 \leqslant \mu \leqslant m$ ist $\partial_z p_j \leqslant \varrho$ und $\partial_{w_\mu} p_j \leqslant \varrho$ sowie*

$$\log A(p_j) \leqslant C_{11} \varrho^{1/2}(m \log H + \varrho^{1/2}).$$

Beweis. Mit Unbekannten x_{jls} für $0 \leqslant j \leqslant p$, $0 \leqslant l \leqslant \varrho$, $0 \leqslant s \leqslant (\varrho, \ldots, \varrho) \in N^m$ setzt man

(7) $$p_j(z, w) = \sum_{l=0}^{\varrho} \sum_{s=0}^{\varrho} x_{jls} z^l w^s.$$

Die in Lemma 3 erklärten B_{ij} stellt man mit $b_{jit} \in K$ in der folgenden Form dar

(8) $$B_{ij}(w) = \sum_{t=0}^{j} b_{jit} w^t \qquad (j = (j, \ldots, j) \in N_0^m).$$

Substitution von (7) und (8) in der Definition von $E_\varrho(z, w)$ liefert dann

$$E_\varrho(z, w) = \sum_{i=0}^{\infty} A_i(w) z^i$$

mit

$$A_i(w) := \sum_{j=0}^{\varrho} \sum_{l=0}^{\min(i,\varrho)} \Big(\sum_{\tau=0}^{\varrho+j} \Big(\sum_{s=0}^{\min(\tau,\varrho)} x_{jls} b_{j,i-l,\tau-s}\Big) w^{\tau}\Big)$$

wobei $\min(\tau, \varrho) := \big(\min(\tau_\nu, \varrho)\big)_{1 \leqslant \nu \leqslant m}$.

Schreibt man die Polynome $A_i(w)$ in der Form

$$A_i(w) = \sum_{\tau=0}^{2\varrho} a_{i\tau} w^{\tau},$$

so entstehen die Koeffizientenvektoren $\boldsymbol{a}_i = (a_{i\tau})_{0 \leqslant \tau \leqslant 2\varrho}$ dabei (nach geeigneter Anordnung) als Produkte zwischen $\boldsymbol{x} = (x_{jls})$ und einer Matrix N_i, deren Elemente die $b_{ji\tau}$ mit $0 \leqslant j \leqslant \varrho$, $\boldsymbol{0} \leqslant \tau \leqslant \varrho$ sind. Man bestimmt die x_{jls} so, daß

1. mindestens eines der Polynome p_j endlichen Index hat,
2. $A_i(w) \in \mathscr{P}_{2\varrho}(\omega)$ für $0 \leqslant i \leqslant G$ gilt.

Lemma 1 gestattet nun die Rückführung der genannten Forderungen auf die Lösung eines linearen Gleichungssystems der in Lemma 4 betrachteten Art. Die erste Forderung bedeutet, daß $j, l \in \{0, \ldots, \varrho\}$ existieren sollen mit

$$\sum_{s=0}^{\varrho} x_{jls} w^s \notin \mathscr{P}_{\varrho}(\omega).$$

Dies ist aber wegen Lemma 1 (nach geeigneter Anordnung der Koordinaten von $\boldsymbol{x}$) äquivalent zu $A_2 \boldsymbol{x} \neq \boldsymbol{0}$ mit

$$A_2 := \begin{bmatrix} M_\varrho & & 0 \\ & \ddots & \\ 0 & & M_\varrho \end{bmatrix}.$$

A_2 enthält $(\varrho+1)^2$-mal die Matrix M_ϱ entlang der Hauptdiagonalen. Die zweite Forderung ist, wiederum nach Lemma 1 und geeigneter Anordnung der Koordinaten, gleichbedeutend mit $M_{2\varrho} \boldsymbol{a}_i = \boldsymbol{0}$ für $0 \leqslant i \leqslant G$. Wegen $\boldsymbol{a}_i = N_i \boldsymbol{x}$ formt man dies äquivalent um zu $A_1 \boldsymbol{x} = \boldsymbol{0}$ mit

$$A_1 := \begin{bmatrix} M_{2\varrho} N_0 \\ \vdots \\ M_{2\varrho} N_G \end{bmatrix}.$$

Um Lemma 4 auf das System $A_1 \boldsymbol{x} = \boldsymbol{0}$ und $A_2 \boldsymbol{x} \neq \boldsymbol{0}$ anzuwenden, schätzt man die Häuser und Nenner von A_1 mittels Lemma 1, Teil (c), und Lemma 3 ab. Daß die in Lemma 4 an die Ränge der Matrizen gestellte Bedingung erfüllt ist, folgt mittels (d) und (e) aus Lemma 1. Lemma 4 liefert dann eine Lösung $\boldsymbol{x}$ mit Komponenten aus I_K, die den gestellten Forderungen genügt. Mittels der Abschätzung der Häuser aus Lemma 4 folgt die behauptete Ungleichung für die Längen der Polynome $p_j(z, w)$. Nach Konstruktion hat mindestens ein $p_j(z, w)$ einen endlichen Index und mit

$$l := \min\{j \mid \operatorname{ind} p_j \neq \infty\}$$

gilt wegen Lemma 2

$$\operatorname{ind} \sum_{j=0}^{\varrho} p_j(z, w) F(z, w)^j = \operatorname{ind} \sum_{j=0}^{\varrho-l} p_{j+l}(z, w) F(z, w)^j.$$

Daher kann o.B.d.A. $\operatorname{ind} p_0 \neq \infty$ angenommen werden und sämtliche der in Lemma 5 gestellten Bedingungen sind damit erfüllt.

6. Abschätzungen. Da die in Lemma 5 konstruierte Hilfsfunktion einen „hohen" Index besitzt, gelingt die Herleitung einer geeigneten oberen Abschätzung für die Werte $E_\varrho(T^k\alpha, \omega^{(k)})$.

Lemma 6. *Für $\varrho \geqslant C_{12}^m$ und $k \in N$ mit $r^k \geqslant C_{13} \log H$ gilt*

$$\log |E_\varrho(T^k\alpha, \omega^{(k)})| \leqslant -C_{14}^{-m} \varrho^{3/2} r^k.$$

Beweis. Mit den in Abschnitt 5 erklärten Polynomen $A_i(w)$ folgt aus Teil (b) von Lemma 5

$$E_\varrho^{(k)} := E_\varrho(T^k\alpha, \omega^{(k)}) = \sum_{i=G+1}^{\infty} A_i(\omega^{(k)})(T^k\alpha)^i.$$

Aus Teil (c) von Lemma 5 folgen für $i > G$ und $\boldsymbol{0} \leqslant \boldsymbol{\mu} \leqslant \boldsymbol{m}$ die Ungleichungen

$$\log \Lambda(A_i(w)) \leqslant \gamma_1 (i + \varrho^{1/2} m \log H)$$

sowie $\partial_{w_\mu} A_i \leqslant 2\varrho$. Zusammen mit

$$\log |\omega_\mu^{(k)}| \leqslant \gamma_2 (\hat{m}k + \log H) \quad \text{und} \quad \log |T^k\alpha|^i \leqslant -\gamma_3 i r^k$$

erhält man

$$|E_\varrho^{(k)}| \leqslant \sum_{i=G}^{\infty} \exp\big(\gamma_4(-ir^k + \varrho m\hat{m}k + \varrho m \log H)\big)$$
$$\leqslant \exp\big(\gamma_4(-C_{10}^{-m} \varrho^{3/2} r^k + \varrho m\hat{m}k + \varrho m \log H)\big).$$

Daraus folgt die behauptete Abschätzung.

Für die algebraischen Zahlen $p_j(T^k\alpha, \omega^{(k)})$ lassen sich folgende Ungleichungen herleiten.

Lemma 7. *Bei $r^k \geqslant C_{15} \log H$ gilt*

(a) $\log |p_j(T^k\alpha, \omega^{(k)})| \leqslant C_{16}\, m\hat{m}\varrho r^k$ *für* $1 \leqslant j \leqslant \varrho$,

(b) $\log |p_0(T^k\alpha, \omega^{(k)})| \geqslant -C_{17}\, m\hat{m}\varrho r^k$, *falls* $p_0(T^k\alpha, \omega^{(k)}) \neq 0$.

Beweis. Teil (a) folgt mit Schlüssen ähnlich zum vorangegangenen Beweis von Lemma 6. Zum Beweis von Teil (b) stellt man die in den Funktionalgleichungen der $f_i(z)$ auftretenden rationalen Funktionen $b_i(z)$ in der folgenden Form dar

$$b_i(z) = \frac{p_i(z)}{q_i(z)} \quad (1 \leqslant i \leqslant p)$$

mit teilerfremden Polynomen $p_i, q_i \in I_K[z]$. Weiter sei D_3 ein gemeinsamer Nenner der a_i für $1 \leqslant i \leqslant p$. Mit

$$Q_i^{(k)}(z) := \prod_{j=0}^{k-1} q_i(T^j z) \qquad (1 \leqslant i \leqslant p,\ k \in N)$$

gilt dann

$$D_3^{\hat{m}k} Q^{(k)}(z)^m \Omega_k\big(b^{(k)}(z)\big)$$

ist in jeder seiner m Komponenten ein Polynom mit Koeffizienten aus I_K. Der Grad eines jeden dieser Polynome in z ist durch $\gamma_1 \hat{m} r^k$ beschränkt und der Logarithmus der Länge ist höchstens $\gamma_2(\hat{m}k + \log H)$. Für jedes $l \in N_0^m$ mit $l \leqslant \varrho$ gilt daher

$$\big(D_3^{\hat{m}k} Q^{(k)}(z)^m\big)^{m\varrho} \Omega_k\big(b^{(k)}(z)\big)^l \in I_K[z].$$

Es besteht aber

$$P_0(z) := D_3^{\hat{m}km\varrho} Q^{(k)}(z)^{mm\varrho} p_0\big(T^k z, \Omega_k(b^{(k)}(z))\big)$$

aus einer Summe solcher zusätzlich noch mit Faktoren $(T^k z)^s$, $0 \leqslant s \leqslant \varrho$, versehener Terme. Daher ist $P_0(z) \in I_K[z]$, und sein Grad läßt sich nach oben durch $\gamma_3 \hat{m} m \varrho r^k$ abschätzen. Für die Länge folgt aus früheren Überlegungen in Verbindung mit Teil (c) von Lemma 5:

$$\Lambda(P_0) \leqslant \exp\big(\gamma_4 m\varrho(\hat{m}k + \log H)\big).$$

Die Anwendung einer Liouville-Abschätzung (vgl. Galochkin [1], Lemma 5) liefert dann das Behauptete.

7. Beweis des Satzes. Ähnlich zu Loxton und van der Poorten [2], Lemma 8 beweist man:

LEMMA 8. *Seien $d_1, \ldots, d_N \in C^*$ paarweise verschieden und $g_1(z), \ldots, g_N(z)$ in einer Umgebung von 0 holomorphe Funktionen, die nicht alle identisch verschwinden. Dann gibt es eine von **d** und **g** abhängige Zahl $\hat{C}_0 > 0$, so daß für alle $k \in N$, $k \geqslant \hat{C}_0$ mindestens eine der Zahlen*

$$\sum_{j=1}^{N} d_j^{k+i} g_j(T^{k+i}\alpha) \qquad (1 \leqslant i \leqslant N)$$

von Null verschieden ist.

LEMMA 9. *Mit einer nur von ϱ und m abhängigen Konstanten $\hat{C}_1 = \hat{C}_1(\varrho, m) > 0$ und*

$$N(\varrho, m) := \prod_{l=1}^{p} (mm_l\varrho + 1)^2$$

ist für alle $k \in N$, $k \geqslant \hat{C}_1$, mindestens eine der Zahlen $p_0(T^{k+i}\alpha, \omega^{(k+i)})$, $1 \leqslant i \leqslant N(\varrho, m)$, von Null verschieden.

Beweis. Stellt man $p_0(z, w)$ wie folgt dar

$$p_0(z, w) = \sum_{i=0}^{\varrho} p_i(w) z^i,$$

so weiß man aus den Überlegungen zum Beweis von Lemma 1, daß es $N_1 \leqslant N(\varrho, m)^{1/2}$ paarweise verschiedene Zahlen $t_1, \ldots, t_{N_1} \in C^*$ und Polynome $q_j^{(i)}(u) \in K[u]$, $0 \leqslant i \leqslant \varrho$, $0 \leqslant j \leqslant N_1$, gibt, so daß für alle $k \in N$ gilt

$$p_i\big(\Omega_k(b^{(k)}(z))\big) = \sum_{j=1}^{N_1} t_j^k q_j^{(i)}\big(b^{(k)}(z)\big). \tag{9}$$

Setzt man $h_l^{(k)}(z) := f_l(\alpha) - a_l^k f_l(z)$ für $1 \leqslant l \leqslant p$, $k \in N$, so kann $p_0(T^k\alpha, \omega^{(k)})$ wegen $b_l^{(k)}(z) = f_l(z) - a_l^k f_l(T^k z)$ umgeformt werden zu

$$p_0(T^k\alpha, \omega^{(k)}) = \sum_{i=0}^{\varrho} p_i\big(\Omega_k(h^{(k)}(T^k\alpha))\big)(T^k\alpha)^i.$$

Da die $p_i\big(\Omega_k(h^{(k)}(z))\big)$ Polynome in den transzendenten Größen $f_l(z)$ sind, erhält man aus (9)

$$p_i\big(\Omega_k(h^{(k)}(z))\big) = \sum_{j=1}^{N_2} d_j^k \tilde{q}_j^{(i)}(f(z)) \tag{10}$$

mit $N_2 \leqslant N(\varrho, m)$ paarweise verschiedenen Zahlen $d_1, \ldots, d_{N_2} \in C^*$ und Polynomen $\tilde{q}_j^{(i)}(u) \in K[u]$. Die Zusammenfassung der in (10) erhaltenen Darstellungen liefert

$$\Gamma_k(z) := \sum_{i=0}^{\varrho} z^i p_i\big(\Omega_k(h^{(k)}(z))\big) = \sum_{j=1}^{N_2} d_j g_j(z)$$

mit Funktionen $g_j(z)$, $1 \leqslant j \leqslant N_2$, die in einer Umgebung des Ursprungs holomorph sind. Wegen $\Gamma_k(T^k\alpha) = p_0(T^k\alpha, \omega^{(k)})$ folgt dann aus Lemma 8 die Behauptung, falls nicht alle $g_j(z)$ und damit auch $\Gamma_k(z)$ identisch verschwinden. Wegen der algebraischen Unabhängigkeit der Funktionen $f_l(z)$ kann letzteres aber nur dann der Fall sein, wenn sämtliche $p_i(\Omega_k(u))$ mit $k \in N$, $0 \leqslant i \leqslant \varrho$ identisch verschwinden. Dieses bedeutet aber $p_i(w) \in \mathscr{P}_\varrho(\omega)$ für $0 \leqslant i \leqslant \varrho$ und damit $\operatorname{ind} p_0(z, w) = \infty$, was nach Konstruktion von $p_0(z, w)$ in Lemma 5 unmöglich ist.

Bemerkung. Die im Satz angegebene Verbesserung unseres Resultates bezüglich $C_1(d)$ im Falle $a_1 = \ldots = a_p = 1$ folgt, da in diesem Fall $N(\varrho, m) = 1$ gezeigt werden kann.

Seien nun die Parameter ϱ und k so gewählt, daß mit hinreichend groß gewählten Konstanten γ_1, γ_2 und $\hat{C}_2$ gilt

$$\varrho \geqslant \gamma_1^m, \qquad r^k \geqslant \gamma_2 \log H + C_2(\varrho, m);$$

außerdem genüge k der Bedingung $p_0(T^k\alpha, \omega^{(k)}) \neq 0$. Dann sind die Voraussetzungen aller Lemmata erfüllt und es folgt mit

$$\Delta := \left|\sum_{\mu=0}^{m} \omega_\mu f(\alpha)^\mu\right| = |F(T^k\alpha, \omega^{(k)})|$$

durch Kombination von Lemma 6 und Lemma 7

$$\begin{aligned}\exp(-C_{14}^{-m}\varrho^{3/2}r^k) &\geqslant |E_\varrho(T^k\alpha, \omega^{(k)})| \\ &= \left|p_0(T^k\alpha, \omega^{(k)}) + \sum_{j=1}^{\varrho} p_j(T^k\alpha, \omega^{(k)}) F(T^k\alpha, \omega^{(k)})^j\right| \\ &\geqslant \exp(-C_{17}\hat{m}m\varrho r^k) - \varrho\max(\Delta, \Delta^\varrho)\exp(C_{16}\hat{m}m\varrho r^k).\end{aligned}$$

Dies liefert

$$\Delta \geqslant \frac{1}{\varrho}\left(\exp(-(C_{16}+C_{17})\hat{m}m\varrho r^k) - \exp(-(C_{14}^{-m}\varrho^{1/2} + C_{16}\hat{m}m)\varrho r^k)\right)$$

und unter der zusätzlichen Bedingung $\varrho \geqslant \gamma_3^m$ gilt

$$\Delta \geqslant \exp(-\gamma_4 \hat{m}m\varrho r^k). \tag{11}$$

Man setzt nun $\varrho := [\gamma_5^m]$ und bestimmt davon abhängig k' gemäß

$$r^{k'} \geqslant \gamma_2 \log H + \hat{C}_2(\varrho, m) > r^{k'-1}.$$

Aus Lemma 9 folgt, daß für mindestens ein k zwischen k' und $k'+N(\varrho, m)$ die Ungleichung $p_0(T^k\alpha, \omega^{(k)}) \neq 0$ gilt. Damit erhält man aus (11) die Abschätzung

$$\Delta > C_2(d) H^{-C_1(d)}$$

mit $\log C_1(d) = \varkappa^{d^p}$.

Literatur

[0] P.-G. Becker-Landeck, *Quantitative Resultate im Zusammenhang mit der Mahlerschen Transzendenzmethode*, Dissertation, Köln, 1984.

[1] A. I. Galochkin, *Transcendence measures of values of functions satisfying certain functional equations*, Math. Notes 27 (1980), S. 83–88.

[2] J. H. Loxton and A. J. van der Poorten, *Arithmetic properties of certain functions in several variables II*, J. Austral. Math. Soc. Ser. A 24 (1977), S. 393–408.

[3] — — *Transcendence and algebraic independence by a method of Mahler*, in *Transcendence Theory: Advances and Applications*, Ed. by A. Baker and D. W. Masser, Academic Press, New York, 1977, S. 211–226.

[4] K. Mahler, *Arithmetische Eigenschaften der Lösungen einer Klasse von Funktionalgleichungen*, Math. Ann. 101 (1929), S. 342–366.

[5] — *Über das Verschwinden von Potenzreihen mehrerer Veränderlicher in speziellen Punktfolgen*, ibid. 103 (1930), S. 573–587.

[6] K. Mahler, *Arithmetische Eigenschaften einer Klasse transzendental-transzendenter Funktionen*, Math. Z. 32 (1930), S. 545–585.

[7] D. W. Masser, *A vanishing theorem for power series*, Invent. math. 67 (1982), S. 275–296.

[8] W. Miller, *Transcendence measures by a method of Mahler*, J. Austral. Math. Soc. Ser. A 32 (1982), S. 68–78.

[9] T. Schneider, *Einführung in die transzendenten Zahlen*, Springer-Verlag, Berlin 1957.

[10] M. Waldschmidt, *Indépendance algébrique de nombres transcendants*, Cours de 3ème cycle, Paris VI, 2ème Sémestre 1982/3.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT
Weyertal 86–90
D-5000 Köln 41

Eingegangen am 27.3.1986
und in revidierter Form am 4.8.1986 (1610)