ACTA ARITHMETICA
LVIII.3 (1991)

Effective measures for algebraic independence of the values of Mahler type functions

by

PAUL-GEORG BECKER (Köln)

In 1929/30 Mahler introduced a new method for investigating the arithmetic and algebraic properties of functions satisfying certain type of functional equations. During the last ten years several papers concerning quantitative results connected with Mahler's method were published. Galochkin [4], Miller [5], Molchanov [6], and the author [2] derived transcendence measures for the numbers studied by Mahler, whereas Nesterenko [10], Molchanov and Yanchenko [7], and the author [1, 3] gave measures for the algebraic independence of these numbers. The sharpest of these measures for algebraic independence was proved by Nesterenko who used commutative algebra to deduce his result. In this note we shall prove the following theorem.

THEOREM 1. *Let $A(z)$ be an $m \times m$ matrix and $B(z)$ be an m-dimensional vector whose entries are rational functions of z with algebraic coefficients. Let $F(z) = (f_1(z), \ldots, f_m(z))^t$ be a vector of formal power series with algebraic coefficients which converge in some neighborhood U of the point $z = 0$ and satisfy*

$$F(z^d) = A(z)F(z) + B(z),$$

where $d \geqslant 2$ is an integer, and which are algebraically independent over $\mathbf{C}(z)$. Suppose that α is an algebraic number, $\alpha \in U$, $0 < |\alpha| < 1$, and none of the numbers $\alpha, \alpha^d, \alpha^{d^2}, \ldots$ is a pole of $A(z)$ or $B(z)$.

Then, for any H and $s \geqslant 1$ and for any polynomial $R \in \mathbf{Z}[x_1, \ldots, x_m] \setminus \{0\}$ whose degree does not exceed s and whose coefficients are not greater than H in absolute value, the following inequality holds:

$$|R(f_1(\alpha), \ldots, f_m(\alpha))| > \exp(-\gamma s^m (\ln H + s^{2m+2})),$$

where γ is a positive constant depending only on α and on the functions $f_1, \ldots, f_m$.

Theorem 1 is an improvement of Nesterenko's result, Theorem 2 in [10], who showed for a smaller class of functions f_i that

$$|R(f_1(\alpha), \ldots, f_m(\alpha))| > \exp(-Cs^m \ln H)$$

under the additional hypothesis $H > \Psi(s)$, where the dependence of Ψ on s is not effectively computable.

If $F(z)$ satisfies the functional equation

$$F(z) = A(z)F(z^d) + B(z)$$

where the entries of $A(z)$ and $B(z)$ are polynomials in z with algebraic coefficients, we can improve the bound given in Theorem 1.

THEOREM 2. *Let $A(z)$ be an $m \times m$ matrix and $B(z)$ be an m-dimensional vector whose entries are polynomials in z with algebraic coefficients. Let $F(z) = (f_1(z), \ldots, f_m(z))^t$ be a vector of formal power series with algebraic coefficients which converge in some neighborhood U of the point $z = 0$ and satisfy*

$$F(z) = A(z)F(z^d) + B(z),$$

where $d \geq 2$ is an integer, and which are algebraically independent over $C(z)$. Suppose that α is an algebraic number, $\alpha \in U$, $0 < |\alpha| < 1$, and none of the numbers $\alpha, \alpha^d, \alpha^{d^2}, \ldots$ is a zero of $\det A(z)$.

Then for any H and $s \geq 1$ and for any polynomial $R \in Z[x_1, \ldots, x_m] \setminus \{0\}$ whose degree does not exceed s and whose coefficients are not greater than H in absolute value, the following inequality holds:

$$|R(f_1(\alpha), \ldots, f_m(\alpha))| > \exp\left(-\gamma s^m (\ln H + s^{m+2} \ln(s+1))\right),$$

where γ is a positive constant depending only on α and on the functions $f_1, \ldots, f_m$.

The algebraic independence measures given in Theorem 1 and Theorem 2 allow us to estimate the transcendence type (see [12], p. 100) of a special field of transcendence degree m. In the case of a transcendence degree $m > 2$ this is apparently the first algebraic independence measure which is sharp enough to give an upper bound for the transcendence type.

We will prove Theorem 1 by the method used by Nesterenko in [10]. The essential fact that allows us to improve his result is an estimate for the zero order of $P(z, f_1(z), \ldots, f_m(z))$ where P is a polynomial. This estimate was proved by Kumiko Nishioka [11].

Theorem 2 can be proved in the same way as Theorem 1. Here one only has to apply the same idea which was used in [2] to improve the result of Miller [5]. Hence we omit the proof of Theorem 2.

1. Preliminaries. We say that a prime ideal $\mathfrak{p} \subset Z[x_0, \ldots, x_m] =: Z[x]$ has *height* q if there exists in $\mathfrak{p}$ a strictly increasing chain of q prime ideals $(0) = \mathfrak{p}_0 \subset \ldots \subset \mathfrak{p}_{q-1} \subset \mathfrak{p}_q = \mathfrak{p}$ and there does not exist any longer chain of this type. We let $h(\mathfrak{p})$ denote the height of $\mathfrak{p}$ and we define the height of an arbitrary ideal $\mathfrak{a}$ as follows: $h(\mathfrak{a}) := \min\{h(\mathfrak{p}) | \ \mathfrak{p} \supset \mathfrak{a}\}$. Suppose that r is an integer, $1 \leq r \leq m$, and u_{ij}, $1 \leq i \leq r$, $0 \leq j \leq m$, are variables which are algebraically independent over the field Q of rational numbers.

For an arbitrary unmixed homogeneous ideal of $Z[x]$ with $r = m+1 - h(I) \geq 1$, we can define a nonzero principal ideal $\bar{I}(r)$ of the ring $Z[u_{10}, \ldots, u_{rm}]$ (see [8], Proposition 2). Let F be the generator of this ideal. We let $H(I)$ denote the maximum absolute value of the coefficients of the polynomial F, and let $N(I) = \deg_{u_1} F$, where $u_j := (u_{j0}, \ldots, u_{jm})$, $1 \leq j \leq r$. Suppose that $\omega = (\omega_0, \ldots, \omega_m) \in C^{m+1}$ is a nonzero vector, and $S^{(i)} = (s_{jk}^{(i)})_{0 \leq j,k \leq m}$, $1 \leq i \leq r$, are skew symmetric matrices whose entries are not connected by any algebraic relation over $Z[x, u_1, \ldots, u_r]$ except for the skew symmetry $s_{jk}^{(i)} + s_{kj}^{(i)} = 0$. For any polynomial $E \in Z[u_1, \ldots, u_r]$ we let $\varkappa(E)$ denote the polynomial in the variables $s_{jk}^{(i)}$, $j < k$, $1 \leq i \leq r$, which is obtained by substituting the vectors $S^{(i)}\omega$, $1 \leq i \leq r$, in place of the variables u_i in E. If $C[S^{(1)}, \ldots, S^{(r)}]$ is the ring of polynomials in the variables $s_{jk}^{(i)}$ with complex coefficients, and F is the generator of the ideal $\bar{I}(r)$, then we define

$$|I(\omega)| := |\omega|_0^{-rN(I)} H(\varkappa(F)),$$

where $|\omega|_0 := \max_{0 \leq i \leq m} |\omega_i|$ and $H(\varkappa(F))$ is the maximum of the absolute values of the coefficients of the polynomial $\varkappa(F)$.

We will derive Theorem 1 from the following theorem. The main task of this paper will then be the proof of Theorem 3.

THEOREM 3. *Suppose that $D \geq 1$, $H \geq 1$ and the conditions of Theorem 1 are fulfilled. Suppose that $I \subset Z[x]$ is an unmixed homogeneous ideal such that $I \cap Z = (0)$, $r = m+1-h(I) \geq 1$, and*

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

where the constant λ is greater than some bound which depends on the functions $f_1, \ldots, f_m$ and the number α.

Then there exists a constant γ depending on $f_1, \ldots, f_m$, α and λ such that for all H with $\ln H \geq \gamma D^{2m+2}$ the following inequality holds

$$\ln|I(1, f_1(\alpha), \ldots, f_m(\alpha))| \geq -\lambda^r (D \ln H(I) + N(I) \ln H) D^{r-1}. \tag{1}$$

Proof of Theorem 1. We now show how Theorem 1 follows from this theorem. Suppose that the conditions of Theorem 1 are fulfilled and let $P \in Z[x]$ be a homogeneous polynomial with $P(1, x_1, \ldots, x_m) = R(x_1, \ldots, x_m)$, $\deg P = \deg R$, and $H(P) = H(R)$. By Lemma 1 we have for the ideal $I \subset Z[x]$ which is generated by P the following inequalities:

$$N(I) = \deg R \leq s,$$

$$\ln H(I) \leq \ln H(R) + m^2 \deg R \leq \ln H + C_1 s,$$

$$|I(1, f(\alpha))| \leq |R(f(\alpha))| C_2 s.$$

The constants $C_1, C_2, \ldots$ depend only on $f_1, \ldots, f_m$ and α. Let $r = m$, $D = s$ and $H_1 = H \exp(C_3 s^{2m+2})$. With H_1 instead of H the requirements of Theorem 3 are now fulfilled and we obtain

$$\ln|R(f(\alpha))| \geqslant -C_4 s^m(\ln H + s^{2m+2}),$$

the inequality asserted in Theorem 1.

In the statements of the following lemmas we denote by $\omega \in C^{m+1}$ a nonzero vector.

LEMMA 1. *Let $I = (P)$ be the principal ideal of $Z[x]$ which is generated by the homogeneous polynomial P. Then*

$$N(I) = \deg P, \quad \ln H(I) \leqslant \ln H(P) + m^2 \deg P,$$
$$|I(\omega)| \leqslant |P(\omega)|\,|\omega|_0^{-\deg P}(m+1)^{2m\deg P}.$$

For the proof, see [9], the proof of Proposition 1.

LEMMA 2. *Suppose that I is an unmixed homogeneous ideal in $Z[x]$, $h(I) \leqslant m$; $I = I_1 \cap \ldots \cap I_s \cap \ldots \cap I_t$ is its irreducible primary decomposition, in which for $l \leqslant s$ we have $I_l \cap Z = (0)$, $I_{s+1} \cap \ldots \cap I_t \cap Z = (b)$, $b \neq 0$; for $l \leqslant s$ let $\mathfrak{p}_l = \sqrt{I_l}$, 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) $\ln|b| + \sum_{l=1}^{s} k_l \ln H(\mathfrak{p}_l) \leqslant \ln H(I) + m^2 N(I)$,

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

When $s = t$ the term $\ln|b|$ does not occur in 2) and 3).

For the proof, see [9], Proposition 2.

DEFINITION. For any two nonzero vectors $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_m)$, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_m) \in C^{m+1}$, we define

$$\|\boldsymbol{\alpha} - \boldsymbol{\beta}\| := |\boldsymbol{\alpha}|_0^{-1}|\boldsymbol{\beta}|_0^{-1} \max_{0 \leqslant i,j \leqslant m} |\alpha_i\beta_j - \beta_i\alpha_j|.$$

LEMMA 3. *Suppose that $Q \in Z[x]$, $Q \not\equiv 0$, is a homogeneous polynomial; $\mathfrak{p} \subset Z[x]$ is a homogeneous prime ideal, $\mathfrak{p} \cap Z = (0)$, $\mathfrak{p} \neq (0)$, $r = m+1-h(\mathfrak{p}) \geqslant 1$;*

$$|\mathfrak{p}(\omega)| \leqslant e^{-X}, \quad X > 0, \tag{2}$$

$$|Q(\omega)|\,|\omega|_0^{-\deg Q} \leqslant H(Q)^{-1}(\deg Q + 1)^{-(2m+2)}, \tag{3}$$

and, finally, the following equality holds for some $\sigma \geqslant 1$:

$$\min(X, \tfrac{1}{2}\ln(1/\varrho)) = -\sigma \ln(|Q(\omega)|\,|\omega|_0^{-\deg Q}),$$

where $\varrho = \min\|\omega - \boldsymbol{\beta}\|$, and the minimum is taken over all zeros $\boldsymbol{\beta} \in C^{m+1}$, $\boldsymbol{\beta} \neq 0$, of the ideal $\mathfrak{p}$.

Then for $r \geqslant 2$ there exists an unmixed homogeneous ideal $J \subset Z[x]$, $J \cap Z = (0)$, $h(J) = m-r+2$, with

1) $N(J) \leqslant N(\mathfrak{p})\deg Q$,

2) $\ln H(J) \leqslant \deg Q \ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H(Q) + m(r+1)N(\mathfrak{p})\deg Q$,

3) $\ln|J(\omega)| \leqslant -X/2\sigma + \deg Q \ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H(Q) + 8m^2 N(\mathfrak{p})\deg Q$.

In the case $r = 1$, the right hand side of the last inequality is nonnegative.

For the proof, see [10], Lemma 5.

LEMMA 4. *Suppose that $I \subset Z[x]$ is an unmixed homogeneous ideal, $I \cap Z = (0)$, and $r = m+1-h(I) \geqslant 1$. For every nonzero vector $\boldsymbol{\omega} \in C^{m+1}$, there exists a zero $\boldsymbol{\beta} \in C^{m+1}$, $\boldsymbol{\beta} \neq \mathbf{0}$, of the ideal I such that*

$$N(I)\ln\|\omega - \boldsymbol{\beta}\| \leqslant (\ln|I(\omega)|)/r + 3m^3 N(I).$$

For the proof, see [10], Lemma 6.

LEMMA 5. *Let $\omega_0 = 1$ and suppose that $\zeta \in C$ is algebraic over $Q(\omega_1, \ldots, \omega_m)$ and integral over $Z[\omega_1, \ldots, \omega_m]$. Furthermore, suppose that γ_1, γ_2, X and Y are positive numbers, P is a polynomial in $Z[x_0, \ldots, x_m, y]$ which is homogeneous in the variables $x_0, \ldots, x_m$ and satisfies*

$$P = \sum_{j=0}^{\nu-1} P_j(x_0, \ldots, x_m)y^j, \quad \deg P_j + \ln H(P_j) \leqslant Y,$$
$$-\gamma_1 X \leqslant \ln|P(\omega_0, \ldots, \omega_m, \zeta)| \leqslant -\gamma_2 X.$$

If $Y > \gamma_3$ and $X > \gamma_4 Y$, where γ_3 and γ_4 are constants depending on $\omega_1, \ldots, \omega_m$, ζ, ν, γ_1 and γ_2, then there exists a homogeneous polynomial $Q \in Z[x]$ satisfying the conditions

$$\deg Q \leqslant \gamma_5 \deg_x P, \quad \ln H(Q) \leqslant \gamma_6 Y, \quad -\gamma_7 X \leqslant \ln|Q(\boldsymbol{\omega})| \leqslant -\gamma_8 X$$

where $\gamma_5, \gamma_6, \gamma_7$ and γ_8 are positive constants depending on $\omega_1, \ldots, \omega_m, \zeta, \nu, \gamma_1$ and γ_2.

Remark. In the case that $\omega_1, \ldots, \omega_m$ are algebraically independent a slightly different version of Lemma 5 can be found as Lemma 10 in [9]. If $\omega_1, \ldots, \omega_m$ are algebraically dependent, one can assume that there is an $n < m$ such that $\omega_1, \ldots, \omega_n$ are algebraically independent and $\omega_{n+1}, \ldots, \omega_m$ are algebraic over $Q(\omega_1, \ldots, \omega_n)$. Hence there exists a ζ' algebraic over $Q(\omega_1, \ldots, \omega_n)$ with the property that ζ' is integral over $Z[\omega_1, \ldots, \omega_n]$ and $\omega_{n+1}, \ldots, \omega_m \in Q(\omega_1, \ldots, \omega_n, \zeta')$. Now one can apply Lemma 5 as already proved for algebraically independent ω_i, to prove Lemma 5 in the general case.

For a formal power series $f(z)$ we denote by $\operatorname{ord} f(z)$ the zero order of $f(z)$ at $z = 0$.

LEMMA 6. *Let $f_1(z), \ldots, f_m(z) \in C[[z]]$ be formal power series satisfying the functional equations*

$$f_i(z^d) = A_i(z, f_1(z), \ldots, f_m(z)) \quad (1 \leqslant i \leqslant m),$$

where $d \geqslant 2$ is an integer and $A_i(z, x_1, \ldots, x_m) \in C(z)[x_1, \ldots, x_m]$, $1 \leqslant i \leqslant m$, are polynomials with $\deg_x A_i \leqslant 1$. Suppose that $Q(z, x_1, \ldots, x_m) \in C[z, x_1, \ldots, x_m]$,

$Q \not\equiv 0$, *is a polynomial with* $\deg_z Q$, $\deg_x Q \leqslant N$. *If* $Q(z, f(z)) \not\equiv 0$, *then*

$$\operatorname{ord} Q(z, f(z)) \leqslant \gamma N^{m+1}$$

where γ *is a positive constant depending only on the functions* $f_1, \ldots, f_m$.

Remark. Lemma 6 is an immediate consequence of the Theorem in [11].

2. Proof of Theorem 3. We shall prove Theorem 3 by induction on r and for each fixed r by contradiction. We assume that under the conditions of the theorem inequality (1) does not hold, i.e.

$$\ln|I(\omega)| < -\lambda^r(D\ln H(I) + N(I)\ln H)D^{r-1}. \tag{4}$$

We shall show that this is impossible if λ is larger than C_1 and if $\ln H$ is larger than $C_2 D^{2m+2}$, where the constants C_1, C_2 and all the constants C_3, $C_4, \ldots$ appearing in the proof depend only on the functions f and the algebraic number α. The proof is divided into four steps.

Step 1. Reduction to a homogeneous prime ideal. There exists a homogeneous prime ideal $\mathfrak{p} \subset \mathbf{Z}[x]$, $\mathfrak{p} \cap \mathbf{Z} = (0)$, $h(\mathfrak{p}) = m+1-r$ with

$$N(\mathfrak{p}) \leqslant \lambda^{m-r} D^{m-r+1}, \tag{5}$$

$$\ln H(\mathfrak{p}) \leqslant \lambda^{m-r} D^{m-r}(\ln H + m^2 D), \tag{6}$$

$$\ln|\mathfrak{p}(\omega)| \leqslant -(\lambda^r/3)(D\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H)D^{r-1}. \tag{7}$$

This assertion can be proved by the following argument: Suppose, on the contrary, that there exists no such homogeneous prime ideal and consider the ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ defined according to Lemma 2 for the ideal I. Since $I \cap \mathbf{Z} = (0)$ we have $s \geqslant 1$. The inequalities (5) and (6) follow from 1) and 2) of Lemma 2 and from the assumption that I satisfies the conditions of Theorem 3. If, as we assume, (7) is false for the prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ we get by 3) of Lemma 2

$$\begin{aligned}\ln|I(\omega)| + m^3 N(I) &\geqslant \sum_{l=1}^{s} k_l \ln|\mathfrak{p}_l(\omega)| \\ &\geqslant -(\lambda^r/3)\Big(D^r \sum_{l=1}^{s} k_l \ln H(\mathfrak{p}_l) + D^{r-1}\ln H \sum_{l=1}^{s} k_l N(\mathfrak{p}_l)\Big) \\ &\geqslant -(\lambda^r/3)\big(D^r(\ln H(I) + m^2 N(I)) + N(I)D^{r-1}\ln H\big).\end{aligned}$$

Now from this inequality and from (4) it follows that for $\lambda \geqslant 1$,

$$D^{r-1}(D\ln H(I) + N(I)\ln H) + (D^r + 3m)m^2 N(I) > 3D^{r-1}(D\ln H(I) + N(I)\ln H).$$

Thus we have a contradiction if $\ln H \geqslant C_3 D$ and the existence of a prime ideal with the above mentioned properties is proved.

Step 2. Construction of an auxiliary function. We shall construct an auxiliary function with a high vanishing order at the point $z = 0$. This will be done by Siegel's Lemma and therefore, we have to study the coefficients of the power series expansions of the $f_i(z)$ at $z = 0$.

From our hypothesis that the $f_i(z)$ are algebraically independent we can deduce that $\det A(z)$ is not identically zero. Hence we can find an $m \times m$ matrix $A_1(z)$, namely $A(z)^{-1}$, and an m-dimensional vector $B_1(z)$ whose entries are rational functions with algebraic coefficients such that the functional equation $F(z) = A_1(z)F(z^d) + B_1(z)$ holds.

Let e be a natural number with the property that the entries of $z^e A_1(z)$ are regular at the origin, and for $1 \leqslant i \leqslant m$ let $\sum_{t=0}^{\infty} f_{it} z^t$ be the power series expansions of the functions f_i. Then we define

$$\hat{F}(z) = \Big(\sum_{t=e}^{\infty} f_{it} z^{t-e}\Big)_{1 \leqslant i \leqslant m}.$$

It is easily checked that $\hat{F}(z)$ satisfies a functional equation

$$\hat{F}(z) = \hat{A}(z)\hat{F}(z^d) + \hat{B}(z)$$

where $\hat{A}(z)$ is an $m \times m$ matrix and $\hat{B}(z)$ is an m-dimensional vector whose entries are rational functions with coefficients from K, the number field generated by α and the coefficients of the entries of $A(z)$ and $B(z)$. Furthermore, all these entries are holomorphic at $z = 0$.

For $i, j = 1, \ldots, m$ let $\sum_{t=0}^{\infty} a_{ijt} z^t$ and $\sum_{t=0}^{\infty} b_{it} z^t$ be the power series expansions of the entries of $\hat{A}(z)$ and $\hat{B}(z)$ respectively. By the functional equation we have the following recursion formula:

$$f_{it} = b_{it-e} + \sum_{j=1}^{m} \sum_{\substack{k=0 \\ d|(t-k-e)}}^{t-e} a_{ijk} f_{je+(t-k-e)/d}.$$

Since the entries of $\hat{A}(z)$ and $\hat{B}(z)$ are rational functions, we can find a constant C_4 such that for $i, j = 1, \ldots, m$, $t \geqslant 0$ the houses of the coefficients a_{ijt} and b_{it} are bounded by C_4^{t+1}. From this and from the recursion formula we see that the houses of the coefficients f_{it} are at most C_5^{t+1}. Let $\mu_1, \ldots, \mu_m$ be nonnegative integers and $f_{\mu t}$ be the power series coefficients of

$$f(z)^{\mu} := f_1(z)^{\mu_1} \ldots f_m(z)^{\mu_m} = \sum_{t=0}^{\infty} f_{\mu t} z^t.$$

Since the $f_{\mu t}$ are sums of products of the f_{it}, we have the inequality

$$\overline{|f_{\mu t}|} \leqslant C_6^{|\mu|+t} \tag{8}$$

where $|\mu| := \mu_1 + \ldots + \mu_m$.

By similar arguments one can show that there exists a natural number D_1 such that $D_1^{|\mu|+t} f_{\mu t}$ is an algebraic integer. We set $N := [\mu^{2m} D]$ where $\mu^{2m+2} = \lambda$, and we have to construct a polynomial $R \in \boldsymbol{Z}[z, x_1, \ldots, x_m]$, $R \not\equiv 0$, with $\deg_z R \leqslant N$ and $\deg_x R \leqslant N$ such that the function $G(z) := R(z, f(z))$ has a zero of a sufficiently large multiplicity at the point $z = 0$. We denote the power series coefficients of $G(z)$ by γ_t, i.e.

$$G(z) = \sum_{t=0}^{\infty} \gamma_t z^t.$$

These coefficients are linear forms in the coefficients of the polynomial R. Hence we can apply Siegel's Lemma (see [12], p. 10) to solve the system of linear equations $\gamma_t = 0$ for $0 \leqslant t \leqslant C_7 N^{m+1}$ where C_7 is a positive constant small enough to guarantee that with $\varkappa := [K : \boldsymbol{Q}]$ the inequality

$$(N+1)\binom{N+m}{m} > 2\varkappa C_7 N^{m+1}$$

holds.

Thus we can find a polynomial $R \in \boldsymbol{Z}[z, x_1, \ldots, x_m]$ with the following properties:

(a) $\deg_z R \leqslant N$, $\deg_x R \leqslant N$,

(b) $\ln H(R) \leqslant C_8 N^{m+1}$,

(c) $\operatorname{ord} R(z, f(z)) \geqslant C_7 N^{m+1}$,

(d) $\operatorname{ord} R(z, f(z)) \leqslant C_9 N^{m+1}$.

The properties (a), (b) and (c) are consequences of the construction via Siegel's Lemma and the inequality (8). Property (d) follows by the application of Lemma 6.

Step 3. Deduction of an analytic inequality. Suppose that $K = \boldsymbol{Q}(\zeta)$ where ζ is integral over $\boldsymbol{Q}$, and let a be a natural number with $a\alpha \in \boldsymbol{Z}[\zeta]$. Furthermore, let $T(z)$ be a polynomial with coefficients from $\boldsymbol{Z}[\zeta]$ such that all the entries of $T(z)A(z)$ and $T(z)B(z)$ are polynomials with coefficients from $\boldsymbol{Z}[\zeta]$.

We define a sequence of polynomials by setting $R_0 := R$ and for $l \geqslant 1$

$$R_l(z, x_1, \ldots, x_m) := T(z)^N R_{l-1}(z^d, A(z)(x_1, \ldots, x_m)^t + B(z)). \tag{9}$$

Then, from the functional equation, we have

$$R_l(z, f(z)) = \Big(\prod_{j=0}^{l-1} T(z^{d^j})\Big)^N R(z^{d^l}, f(z^{d^l})). \tag{10}$$

From (9) we can deduce the following inequalities for the degrees and length of R_l:

$$\deg_x R_l \leqslant \deg_x R_{l-1}, \quad \deg_z R_l \leqslant d \deg_z R_{l-1} + C_{10} N,$$

$$\ln \Lambda(R_l) \leqslant \ln \Lambda(R_{l-1}) + C_{11} N,$$

where $\Lambda(R_l)$ denotes the sum of the houses of the coefficients of R_l. Hence we get by the result deduced in Step 2

$$\deg_x R_l \leqslant N, \qquad \deg_z R_l \leqslant C_{12} N d^l,$$

$$\ln \Lambda(R_l) \leqslant C_{13}(N^m + l)N.$$

Let $N_l := \deg_z R_l$ for $l \geqslant 0$. It is easily checked that the coefficients $r_{l\mu\nu}$ of

$$R_l(z, x_1, \ldots, x_m) = \sum_{\nu=0}^{N_l} \sum_{|\mu| \leqslant N} r_{l\mu\nu} z^\nu f(z)^\mu$$

are from $\boldsymbol{Z}[\zeta]$. Hence the same is true for $a^\nu \alpha^\nu r_{l\mu\nu}$ if $0 \leqslant \nu \leqslant N_l$ and $|\mu| \leqslant N$. For $l \geqslant 0$, $0 \leqslant \nu \leqslant N_l$, $|\mu| \leqslant N$ and $0 \leqslant n < \varkappa$ we define $\tau_{nl\mu\nu}$ by

$$a^\nu \alpha^\nu r_{l\mu\nu} = \sum_{n=0}^{\varkappa-1} \tau_{nl\mu\nu} \zeta^n.$$

With

$$P_l(y, x) := \sum_{\nu=0}^{N_l} \sum_{|\mu| \leqslant N} \sum_{n=0}^{\varkappa-1} \tau_{nl\mu\nu} a^{N_l - \nu} y^n x_0^{N-|\mu|} x_1^{\mu_1} \ldots x_m^{\mu_m}$$

we have $a^{N_l} R_l(\alpha, f(\alpha)) = P_l(\zeta, 1, f(\alpha))$. $P_l \in \boldsymbol{Z}[y, x]$ is homogeneous in the variables $x_0, \ldots, x_m$ and we get the following bounds for its degrees and length:

$$\deg_y P_l \leqslant C_{14}, \quad \deg_x P_l \leqslant N,$$

$$\ln \Lambda(P_l) \leqslant C_{15}(N d^l + N^{m+1} + lN) \leqslant C_{16} N (d^l + N^m).$$

By (10) we have a relation between $P_l(\zeta, 1, f(\alpha))$ and $G(\alpha^{d^l})$, and therefore we have to study the behaviour of $G(z)$ for small z. From the definition of the γ_t, the upper bound for $H(R)$, and (8) we conclude that $\ln \overline{|\gamma_t|} \leqslant C_{17}(N^{m+1} + t)$ and that $D_1^t \gamma_t$ is an algebraic integer for $t \geqslant 0$. But we have $\gamma_t = 0$ for $t \leqslant C_7 N^{m+1}$ and therefore $\ln \overline{|\gamma_t|} \leqslant C_{18} t$ for all $t \geqslant 0$.

We let $\tau := \operatorname{ord} G(z)$ and we get

$$\Big|\sum_{t=\tau+1}^{\infty} \gamma_t z^t\Big| \leqslant C_{19}(|z| e^{C_{18}})^{\tau+1}, \tag{11}$$

if we assume that $|z| < e^{-C_{18}}$. By the fundamental inequality we have $\ln|\gamma_\tau| \geqslant -C_{20}\tau$. If we combine this with (11), we see that

$$\tfrac{1}{2}|\gamma_\tau z^\tau| < |G(z)| < \tfrac{3}{2}|\gamma_\tau z^\tau|,$$

if only $\ln|z| \leqslant -C_{21}\tau$. Now we set $z := \alpha^{d^l}$ and the last condition is fulfilled as soon as $d^l \geqslant C_{22} N^{m+1}$. Hence we get under this assumption

$$-C_{23} N^{m+1} d^l \leqslant \ln|G(\alpha^{d^l})| \leqslant -C_{24} N^{m+1} d^l.$$

From these inequalities and from (10) we get

$$-C_{25}N^{m+1}d^l \leqslant \ln|P_l(\zeta, 1, f(\alpha))| \leqslant -C_{26}N^{m+1}d^l,$$

if $N \geqslant C_{27}$ and $d^l \geqslant C_{28}N^{m+1}$. By Lemma 5 applied to the polynomials P_l it follows that there exist homogeneous polynomials $Q_l \in \mathbf{Z}[x]$ such that

(12) $$\deg_x Q_l \leqslant C_{29}N, \quad \ln H(Q_l) \leqslant C_{30}Nd^l,$$

(13) $$-C_{31}N^{m+1}d^l \leqslant \ln|Q_l(\omega)| \leqslant -C_{32}N^{m+1}d^l$$

where $\omega := (1, f_1(\alpha), \ldots, f_m(\alpha))$.

Step 4. Application of the results from commutative algebra and conclusion of the proof. Let $\mathfrak{p}$ be the ideal which was constructed in Step 1 and which depends on the given positive numbers D, H and λ. We define

$$\varrho := \min\{\|\omega - \beta\| \mid \beta \text{ is a zero of } \mathfrak{p}\}$$

and we deduce from Lemma 4 the inequality

$$\ln(1/\varrho) \geqslant -(rN(\mathfrak{p}))^{-1}\ln|\mathfrak{p}(\omega)| - 3m^3.$$

Together with (7) it follows that

$$\ln(1/\varrho) \geqslant C_{33}D^{r-1}\ln H - 3m^3,$$

and if we assume $D^{r-1}\ln H \geqslant C_{34}N^{m+1}d^l$, we see that $\frac{1}{2}\ln(1/\varrho) \geqslant 2C_{31}N^{m+1}d^l$.

It is easily checked that

$$\tfrac{1}{3}\lambda^r(D\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H)D^{r-1} \geqslant 2C_{31}N^{m+1}d^l,$$

as soon as $D^{r-1}\ln H \geqslant C_{35}N^{m+1}d^l$. Hence by the results of Step 3 we can find a natural number $n \geqslant n'$ with

(14) $$2C_{31}N^{m+1}d^n \leqslant \min\left(\tfrac{1}{3}\lambda^r(D\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H)D^{r-1}, \tfrac{1}{2}\ln(1/\varrho)\right) \leqslant 2C_{31}N^{m+1}d^{n+1}$$

and

(15) $$-C_{31}N^{m+1}d^n \leqslant \ln|Q_n(\omega)| \leqslant -C_{32}N^{m+1}d^n,$$

if the conditions $N \geqslant C_{36}$, $d^{n'} \geqslant C_{37}N^{m+1}$, and

(16) $$D^{r-1}\ln H \geqslant C_{38}N^{m+1}d^{n'}$$

are fulfilled. By Q we will denote a polynomial Q_n corresponding to such a number n.

We choose n' such that $d^{n'} \geqslant C_{39}N^{m+1} > d^{n'-1}$ where C_{39} is a sufficiently large constant.

Now we apply Lemma 3 to the ideal $\mathfrak{p}$ and the polynomial Q and we set

$$X := \tfrac{1}{3}\lambda^r(D\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H)D^{r-1}$$

and therefore, inequality (2) is true by the construction of $\mathfrak{p}$. Inequality (3) follows immediately from (12) and (15) if we assume that $N, n \geqslant C_{40}$, and using (15) we see that

$$-2C_{31}N^{m+1}d^n \leqslant \ln(|Q(\omega)|\,|\omega|_0^{-\deg Q}) \leqslant -C_{32}N^{m+1}d^n.$$

If we combine this with (14) and if we define σ as in Lemma 3, then we can prove that $1 \leqslant \sigma \leqslant 2C_{31}C_{32}^{-1}d$. Hence the requirements of Lemma 3 are fulfilled.

We set $\Omega := D\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H$ and we obtain from (5), (6) and (12)

(17) $$\ln H(Q)N(\mathfrak{p}) \leqslant C_{30}Nd^nN(\mathfrak{p}) \leqslant C_{41}\lambda^m D^m N^{-m}\Omega \leqslant C_{42}\mu^{2m}\Omega.$$

From (12) and the definition of N it follows that

(18) $$\deg Q \ln H(\mathfrak{p}) \leqslant C_{29}N\ln H(\mathfrak{p}) \leqslant C_{43}\mu^{2m}\Omega.$$

And if we assume that $\ln H \geqslant C_{44}D$, we get the inequalities

(19) $$8m^2\deg Q N(\mathfrak{p}) \leqslant C_{45}NN(\mathfrak{p}) \leqslant C_{46}\mu^{2m}\Omega.$$

In the case $r = 1$ we have by Lemma 3

$$X/(2\sigma) \leqslant \deg Q\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H(Q) + 8m^2\deg Q N(\mathfrak{p}).$$

The left hand side of this inequality has $C_{47}\mu^{2m+2}\Omega$ as a lower bound whereas the right hand side is at most $C_{48}\mu^{2m}\Omega$. Thus we have the desired contradiction if λ is sufficiently large and if H satisfies (16). But as (16) is a consequence of the condition $\ln H \geqslant \gamma D^{2m+2}$, Theorem 3 is proved for $r = 1$.

We now consider the case $r \geqslant 2$. By Lemma 3 there exists an unmixed homogeneous ideal $J \subset \mathbf{Z}[x]$ with $h(J) = m - r + 2$, $J \cap \mathbf{Z} = (0)$ and

$$N(J) \leqslant N(\mathfrak{p})\deg Q,$$

$$\ln H(J) \leqslant \deg Q\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H(Q) + m(r+1)N(\mathfrak{p})\deg Q,$$

$$\ln|J(\omega)| \leqslant -X/(2\sigma) + \deg Q\ln H(\mathfrak{p}) + N(\mathfrak{p})\ln H(Q) + 8m^2N(\mathfrak{p})\deg Q.$$

Hence we get by (17), (18) and (19)

$$N(J) \leqslant C_{49}\mu^{2m}\Omega,$$

$$\ln H(J) \leqslant C_{50}\mu^{2m}\Omega,$$

$$\ln|J(\omega)| \leqslant -C_{51}\mu^{(2m+2)r}D^{r-1}\Omega + C_{52}\mu^{2m}\Omega \leqslant -C_{53}\mu^{(2m+2)r}D^{r-1}\Omega.$$

The last inequality is true if $\mu \geqslant C_{54}$. It is easily checked that the ideal J satisfies the requirements of Theorem 3. Hence from the induction hypothesis, that the theorem is already true for $r-1$, it follows that

$$\ln|J(\omega)| \geqslant -\lambda^{r-1}(D\ln H(J)+N(J)\ln H)D^{r-2}$$
$$\geqslant -C_{55}\lambda^{r-1}D^{r-1}\mu^{2m}\Omega \geqslant -C_{56}\mu^{(2m+2)r-2}D^{r-1}\Omega.$$

But this contradicts the upper bound for $\ln|J(\omega)|$ as soon as μ is sufficiently large. Hence Theorem 3 is proved.

References

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

[2] — *Transcendence measures by Mahler's transcendence method*, Bull. Austral. Math. Soc. 33 (1986), 59–65.

[3] — *Maße für algebraische Unabhängigkeit nach einer Methode von Mahler*, Acta Arith. 50 (1988), 279–293.

[4] A. O. Galochkin, *Transcendence measures of values of functions satisfying certain functional equations*, Mat. Zametki 27 (1980), 175–183; English transl. in Math. Notes 27 (1980), 83–88.

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

[6] S. M. Molchanov, *Estimates for the measure of transcendence in the Mahler method* (in Russian), *Diophantine approximations*, Part 1, 56–65, Moskov. Gos. Univ., Mekh.-Mat. Fak., Moscow 1985.

[7] S. M. Molchanov and A. Ya. Yanchenko, *A measure of the algebraic independence of values of functions satisfying certain functional equations* (in Russian), *All-Union Conf. Theory of Transcendental Numbers and its Applications*, Izdat. Moskov. Gos. Univ., Moscow 1983, 98.

[8] 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), 235–284; English transl. in Math. USSR-Izv. 11 (1977), 239–270.

[9] — *On algebraic independence of algebraic powers of algebraic numbers*, Mat. Sb. 123 (165) (1984), 435–459; English transl. in Math. USSR-Sb. 51 (1985), 429–454.

[10] — *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 (1987), 545–567.

[11] K. Nishioka, *On an estimate for the orders of zeros of Mahler type functions*, Acta Arith. 56 (1990), 249–256.

[12] M. Waldschmidt, *Nombres transcendants*, Lecture Notes in Math. 402, Springer, Berlin 1974.

MATHEMATISCHES INSTITUT
UNIVERSITÄT ZU KÖLN
Weyertal 86-90
D-5000 Köln 41
Germany

Received on 24.4.1989
and in revised form on 4.5.1990 (1928)

ACTA ARITHMETICA
LVIII.3 (1991)

On an irreducibility theorem of I. Schur

by

MICHAEL FILASETA* (Columbia, S.C.)

Dedicated to the memory of Emil Grosswald

1. Introduction. In Grosswald's book *Bessel Polynomials* [7], he investigates various aspects of the Bessel polynomials

$$y_n(x) = \sum_{j=0}^{n} \frac{(n+j)!}{2^j(n-j)!j!}x^j.$$

In particular, he discusses several results about the irreducibility of $y_n(x)$ over the rationals (also see [8, 9]). He proves that $y_n(x)$ is irreducible if $n = p^m$, $p+1$, or $p-1$ where p is a prime and m is a positive integer. He further shows that the largest degree of an irreducible factor of $y_n(x)$ is asymptotic to n. Later, Grosswald [10] pointed out that if $p_j < n \leqslant p_{j+1}$ where p_j and p_{j+1} are consecutive primes, then $y_n(x)$ is irreducible provided that the product $n(n+1)$ has a prime factor $> \min\{n-p_j+1,\ p_{j+1}-n\}$. This fact is sufficient enough to establish that $y_n(x)$ is irreducible for every $n \leqslant 10^6$. It may in fact imply that every $y_n(x)$ is irreducible, but to prove so seems to require a much better understanding of gaps between primes than is currently known. On the other hand, with a little work (cf. [5]), one can use Grosswald's observation to show that a positive proportion of the $y_n(x)$ are irreducible. More specifically, if $k_1(t)$ denotes the number of reducible $y_n(x)$ with $n \leqslant t$, then there is a constant $c < 1$ such that $k_1(t) \leqslant ct$ for all t sufficiently large.

Mainly motivated by Grosswald's work and his encouragement, the author pursued the problem of determining when $y_n(x)$ is irreducible. He was able to show [5] that $k_1(t) = o(t)$. Later in Section 4, we will see how work of Lagarias and Odlyzko [11] can aid in establishing that $k_1(t) \ll t/l_3(t)$ where $l_m(t)$ denotes m iterations of $\log t$. Under the assumption of the Generalized Riemann Hypothesis (GRH), the same arguments lead to $k_1(t) \ll t/\log\log t$. On the other hand, we shall see that Grosswald's observation above and the Riemann Hypothesis (RH) imply the better result $k_1(t) \ll t\exp((-1/\sqrt{2}+\varepsilon)\sqrt{\log t \log\log t})$ for any $\varepsilon > 0$.

* Research was supported in part by the NSF under grant number DMS-8903123.