J. Austral. Math. Soc. (Series A) 39 (1985), 227–240

EFFECTIVITY IN INDEPENDENCE MEASURES FOR VALUES OF E-FUNCTIONS

W. DALE BROWNAWELL

(Received 12 July 1983; revised 20 September 1984)

Communicated by P. E. Blanksby

Dedicated to Professor K. Mahler on the occasion of his 80th birthday

Abstract

We establish a measure of algebraic independence for values of E-functions which is more nearly effectively computable than the previous one. When the system of equations meets either of two criteria, then the measure becomes entirely effectively computable.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 10 F 37.

1. Statement of main result

In 1929 C. L. Siegel [11] proved the algebraic independence of values of certain functions at an algebraic point. These functions satisfy a system of first order linear differential equations over $\mathbb{C}(z)$ and in addition are KE-functions, so the coefficients of their Maclaurin expansion are of the form $a_\nu/\nu!$ with the a_ν from a fixed algebraic number field K (see Section 3 below). This new class of functions included the exponential function (thus generalizing the celebrated theorem of Lindemann) and certain hypergeometric and Bessel functions. In 1949 Siegel [12] formalized this approach for functions satisfying a more general normality condition, whose verification however proved very elusive for further classes.

In 1959 A. B. Shidlovsky [9] removed that imperfection in an ingenious way by relating the rank of a certain matrix representing the derivatives of a function to

This research was supported in part by the National Science Foundation.

the order of vanishing of the function at the origin (see Section 2 below). Excellent accounts of the method can be found in Feldman and Shidlovsky [3], Mahler [6] and Shidlovsky [10]. Shidlovsky was thereby able to establish the following basic result.

THEOREM (*Siegel-Shidlovsky*). *Let the KE-functions* $f_1(z),\dots,f_m(z)$ *be algebraically independent over* $\mathbf{C}(z)$ *and constitute a solution of the system of linear differential equations*

$$y_k' = q_{k0} + \sum_{i=1}^{m} q_{ki} y_i, \tag{1}$$

$k = 1,\dots,m$, q_{ki} *in* $\mathbf{C}(z)$. *Let* $\alpha \neq 0$ *be a non-zero algebraic number which is not a pole of any* q_{ki}. *Then the numbers* $f_1(\alpha),\dots,f_m(\alpha)$ *are algebraically independent over* $\mathbf{Q}$.

Without loss of generality we may assume that $\alpha \in K$. So when $\alpha \in K$ and the hypotheses of the above theorem are satisfied, we will say that we are in the *Siegel-Shidlovsky setting*. We take $\kappa = [K:\mathbf{Q}]$.

In 1962 S. Lang [4] showed that from this method one can deduce the following result, which, in its joint dependence on d and H, is comparable to the measure of algebraic independence established for the exponential case by K. Mahler [5].

THEOREM (*Lang*). *In the Siegel-Shidlovsky setting, there exist an effective constant* c_1 *and a function* $\Omega(d) \geqslant 0$ *such that for every non-zero polynomial* $P(X_1,\dots,X_m)$ *in* $\mathbf{Z}[X_1,\dots,X_m]$ *of degree at most* d *and having coefficients of maximum modulus* $H(P) \leqslant H$,

$$|P(f_1(\alpha),\dots,f_m(\alpha))| > \Omega(d) H^{-c_1 d^m}. \tag{2}$$

An easy application of the Dirichlet box principle shows that $c_1 \geqslant 1$. In a remarkable pair of papers, Yu. V. Nesterenko [7, 8] was able to make $\Omega(d)$ in (2) explicit in its dependence on d.

THEOREM (*Nesterenko*). *In the Siegel-Shidlovsky setting, in* (2) *one can take* $c_1 = 4^m \kappa^m (m\kappa^2 + \kappa + 1)$ *and* $\Omega(d) = \exp(-\exp(\tau_0 d^{2m} \ln(d+1)))$, *where* $\tau_0 > 0$ *is a constant independent of d and H.*

Unfortunately τ_0 is not, in general, an effective constant. The ineffectivity of τ_0 arises first of all from our inability to determine in general the minimal order of vanishing of all polynomials belonging to certain ideals whose existence is given by the Picard-Vessiot theory of solutions of linear differential equations. A second source of ineffectivity is the reduction of the system (1) to a system with coefficients in $K(z)$ (see Lemma 10 of Shidlovsky [9] or Section 87 of Mahler [6]) by selecting a K-basis for the finite dimensional vector space $\Sigma\, Kq_{ki}$. However one cannot in general perform this reduction effectively. The purpose of this paper is to show nevertheless that a variation of the proof of Nesterenko's result allows one to completely isolate the ineffectivity of the measure in a form which, after the reduction just mentioned, contains Nesterenko's theorem.

THEOREM 1. *In the Siegel-Shidlovsky setting for a system* (1) *whose coefficients* q_{ki} *lie in* $K(z)$, *there are*

(i) *constants* c_1, $c_2 > 0$ *depending effectively on* m, α, *the coefficients* q_{ki}, κ *and the constants C in the KE-function criterion and*

(ii) *an in general ineffective constant* $\tau > 0$ *depending only on* $f_1(z), \ldots, f_m(z)$ *and the system* (1), *such that in* (2) *one has*

$$|P(f_1(\alpha), \ldots, f_m(\alpha))| > \min\{\Omega_1(d) H^{-c_1 d^m}, \tau\},$$

where $\Omega_1(d) = \exp(-\exp(c_2 d^{2m} \ln(d+1)))$. *The constant* c_1 *may be chosen as above.*

The author thanks K. Väänänen for his comments on a preliminary sketch of this result. The author also thanks the referee for his suggestion to state the proposition below separately from Theorem 1 and for his helpful advice on the amount of detail appropriate to our exposition. In addition, the author is indebted to the referee and to D. Bertrand for alerting him to the very recent paper of N. T. Tai [153], which extends Nesterenko's results in another direction. We hope to comment on this interesting work elsewhere.

2. Effectivity of τ

In order to allow easy comparison with the fundamental results of Nesterenko [8] and thus permit a condensed presentation, we retain his outline and much of his notation. We shall however distinguish explicitly between effective and ineffective constants. In fact all our constants depend effectively on m, α, κ, the q_{ki}'s, the constant C from the E-function criterion below, all of which we regard as given, and the constant τ_1, whose existence is guaranteed by the fundamental result of Nesterenko below. We denote the constants depending only on the former, effective constants by c's and those depending on τ_1 by τ's. We begin by stating Nesterenko's strengthening (Theorem 3 of Nesterenko [8]) of Shidlovsky's fundamental result, which makes the dependence on h explicit.

https://doi.org/10.1017/S1446788700022497 Published online by Cambridge University Press

THEOREM (*Shidlovsky-Nesterenko*). *Let the functions $f_1(z), \dots, f_m(z)$ constitute a solution of the system* (1) *of linear differential equations. We assume that these functions are analytic at $z = 0$ and algebraically independent over $\mathbf{C}(z)$. Let P be non-trivial in $\mathbf{C}[z, X_1, \dots, X_m]$ with $\deg_z P \leqslant n$, $\deg_X P \leqslant h$. Set $R(z) = P(z, f_1(z), \dots, f_m(z))$ and let s_0 be the dimension of the vector space over $\mathbf{C}(z)$ spanned by the derivatives $R^{(i)}(z)$, $i = 0, 1, 2, \dots$. Then the order $\mathcal{O}(P)$ of zero at $z = 0$ of the analytic function $R(z)$ is bounded as follows*:

$$\mathcal{O}(P) \leqslant s_0 n + \tau_1 h^{\gamma},$$

where the constant τ_1 depends only on $f_1, \dots, f_m$ and the system (1) *and $\gamma = (m+1)^{m+1} + m + 1$.*

It is possible to sharpen this to $\gamma = (m+1)! + m + 1$ using the improvement in Brownawell [2] of Corollary 2 of Nesterenko [8] and even further to $\gamma = 2m$ by D. Bertrand and F. Beukers [1]. However this does not affect our discussion of the effectivity of the measure. Indeed an examination of the proof shows that improvements in estimations of γ have only the disappointingly slight effect of improving the bounds on the constants c_1, c_2, τ appearing in Theorem 1.

Definitions (6) and (9) in the proof of Theorem 2, together with our choice of δ express τ explicitly in terms of τ_1. The last displayed line of the proof of Theorem 3 of Nesterenko's [8] shows that

$$\tau_1 \leqslant c'\tau' + c'',$$

where c', c'' are explicitly given constants and the constant τ' (called C in Theorem 1 of Nesterenko [7] and Theorem 2 of Nesterenko [8]) arises in the following way.

Let $t(z)$ in $K[z]$ be a common denominator for the coefficients q_{ki} of the system (1). Then the operator on $\mathbf{C}(z)[X_0, \dots, X_m]$ defined as

$$D_1 := \frac{\partial}{\partial z} + \sum_{j=0}^{m} \left(\sum_{i=0}^{m} q_{ji} X_i \right) \frac{\partial}{\partial X_j}$$

reflects differentiation for solutions of the homogenized system of differential equations

$$\text{(1h)} \qquad y_0' = 0, \qquad y_j' = \sum_{i=0}^{m} q_{ji} y_i.$$

So for any polynomial $Q(z, X_0, \dots, X_m)$ in $\mathbf{C}(z)[X_0, \dots, X_m]$,

$$\frac{d}{dz} Q(z, f_0(z), \dots, f_m(z)) = (D_1 Q)(z, f_0(z), \dots, f_m(z)).$$

The related operator $D := t(z) D_1$ acts on $\mathbf{C}[z, X_0, \dots, X_m]$. Let Φ be a fundamental solution of the homogenized system of equations (1h). We may as well choose Φ so that the first column is the transpose of $(1, f_1(z), \dots, f_m(z))$ and the remaining columns have first entry zero, and thus correspond to solutions of the *truncated* homogeneous system obtained from (1) by dropping all the inhomogeneous terms. Then every non-trivial solution of (1h) can be written uniquely in the form $\Phi \cdot \mathfrak{x}$ with $\mathfrak{x}$ in $\mathbb{C}^{m+1} \setminus \{0\}$. Those solutions which corrspond to solutions of (1) will have $\mathfrak{x} = (x_0, \ldots, x_m)$ with $x_0 \neq 0$. We denote the set of elements they determine in $\mathbb{P}^m$ by U. Since the action of Picard-Vessiot group G is given by right multiplication $\Phi \to \Phi \cdot \sigma$, when we realize G as an algebraic matrix group over $\mathbb{C}$, we may alternatively think of G as acting on $\mathbb{P}^m$ through multiplication on the left. From the definitions we see that U is closed under the action of G. Let G_0 denote the connected component of the identity in G and let F denote its fixed field in $\mathbb{C}(z, \Phi)$. Since F is a finite extension field of $\mathbb{C}(z)$, D can be thought of as acting on $F[X_0, \ldots, X_m]$ as well.

For any D-invariant homogeneous radical ideal I of $F[X_0, \ldots, X_m]$, let $N(I)$ denote the set of all x in $\mathbb{P}^m$ for which $\Phi \cdot x$ is a zero of I. On the other hand, for any G_0-invariant subvariety V of $\mathbb{P}^m$, let $1(V)$ denote the ideal of all polynomials of $F[X_0, \ldots, X_m]$ vanishing on all of $\Phi \cdot V$. Theorem 2 of Nesterenko [7] shows that the maps N and l establish a bijection between the set of all D-invariant homogeneous radical ideals of $F[X_0, \ldots, X_m]$ and the G_0-invariant subvarieties of $\mathbb{P}^m$.

Lemma 8 of Nesterenko [7] shows that there is a unique minimal radical D-invariant non-zero ideal $\mathfrak{I}_0$ of $\mathbb{C}[z, X_0, \ldots, X_m]$ which is homogeneous in the X's. Then in Lemma 16 in the proof of Theorem 1 of Nesterenko [7], we find that

$$\tau' := \operatorname{ord} \mathfrak{I}_0 = \min \operatorname{ord} Q(1, f_1(z), \ldots, f_m(z)),$$

where Q ostensibly runs through all non-zero polynomials in $\mathfrak{I}_0$. However in the application for Theorem 2 of Nesterenko [8], we see that Q need only run through some prime component $\mathfrak{P}$ of $\mathfrak{I}_0$. Further the proof of Lemma 8 of Nesterenko [7] shows that the prime components $\mathfrak{P}$ of $\mathfrak{I}_0$ are of two types:

(i) $\mathfrak{P}$ for which $\mathfrak{P} \cap \mathbb{C}[z] = (0)$,

(ii) $\mathfrak{P}$ for which $\mathfrak{P} \cap \mathbb{C}[z] \neq (0)$.

In the second case, since $\mathfrak{P}$ is D-invariant, $\mathfrak{P} \cap \mathbb{C}[z] = (c(z))$, where $c(z)$ divides $t(z)$ in $\mathbb{C}[z]$ (see the proof of Lemma 8 cited above). Moreover since $\mathfrak{P}$ is prime, $\deg c(z) = 1$, and $\operatorname{ord} \mathfrak{P} \leqslant 1$.

The first case can be treated in the special situation where any of the following (equivalent) conditions holds:

LEMMA 1. *In the above situation, the following are equivalent.*

(a) *Every solution of* (1) *has components y_i which are algebraically independent over* $\mathbb{C}(z)$.

(b) *The G_0-orbit of any x in U is Zariski dense in $\mathbb{P}^m$.*

(c) *All non-zero D-invariant homogeneous radical ideals in $F[X_0, \ldots, X_m]$ contain X_0.*

https://doi.org/10.1017/S1446788700022497 Published online by Cambridge University Press

PROOF. (a) $\Rightarrow$ (b). Take the closure C of an orbit $G_0 x$ with x in U. By the correspondence of Nesterenko cited above, $C = N(\mathfrak{I})$ for some D-invariant homogeneous radical ideal $\mathfrak{I}$ of $F[X_0, \dots, X_m]$. Then in particular $\mathfrak{I}$ vanishes at $\Phi \cdot x$. However by assumption (a), then $\mathfrak{I} = (0)$. From the other half of Nesterenko's correspondence, we find that $C = \mathbf{P}^m$, as desired.

(b) $\Rightarrow$ (c). Let $\mathfrak{I}$ be a D-invariant homogeneous radical ideal in $F[X_0, \dots, X_m]$. Then $N(\mathfrak{I})$ is a G_0-invariant algebraic variety. So if $N(\mathfrak{I}) \cap U \neq \varnothing$, then by assumption $N(\mathfrak{I}) = \mathbf{P}^m$, and thus $\mathfrak{I} = (0)$. If on the other hand $N(\mathfrak{I}) \cap U = \varnothing$, then from the correspondence we find that X_0 vanishes on all elements of $\Phi \cdot N(\mathfrak{I})$, so X_0 lies in $\mathfrak{I}$.

(c) $\Rightarrow$ (a). For a solution $(g_1(z), \dots, g_m(z))$, take the prime D-invariant homogeneous ideal $\mathfrak{I}$ vanishing at $(1, g_1(z), \dots, g_m(z))$. Since X_0 does not vanish there, X_0 does not lie in $\mathfrak{I}$. Thus by assumption $\mathfrak{I} = (0)$. Since then no non-zero homogeneous polynomial over F vanishes at $(1, g_1(z), \dots, g_m(z))$, the functions $g_1(z), \dots, g_m(z)$ are algebraically independent over F.

Thus in case (i) above, if any of the equivalent conditions of Lemma 1 holds, then X_0 lies in $\mathfrak{P}$ and ord $\mathfrak{P} = 0$. We note in passing that the homogeneous analogue of Lemma 1 (occurring when (1) is already homogeneous, i.e. $q_{10} = \cdots = q_{m0} = 0$) has the three following equivalent conditions.

(a) Every non-trivial solution has components which are homogeneously independent over $\mathbf{C}(z)$.

(b) G_0 acts transitively on $\mathbf{P}^{m-1}$.

(c) There is no proper D-invariant homogeneous radical ideal in $F[X_1, \dots, X_m]$.

The corresponding changes can also be made in the following result.

PROPOSITION. *The constant τ is effectively computable whenever we can make the zero estimate of Shidlovsky-Nesterenko effective. In particular τ is effectively computable whenever either* (a) $z = 0$ *is a non-singular point of* (1), *or* (b) *every solution of* (1) *has components y_i which are algebraically independent over* $\mathbf{C}(z)$.

PROOF. It only remains to check case (a). However then D lowers positive, finite ords of ideals in $\mathbf{C}[z, X_0, \dots, X_m]$ by 1. Thus no D-invariant ideals there can have positive but finite ord; consequently $\tau' = 0$.

3. Notation and preliminary lemmas

Since all known E-functions satisfying linear differential equations also satisfy somewhat stronger and more convenient properties than those of Siegel [11], we follow the convention of Lang [4] to incorporate those properties into the definition. For an element ρ of the algebraic number field K, denote the maximum modulus of its conjugates by $\overline{|\rho|}$. Then we call an analytic function $f(z) = \sum_{\nu=0}^{\infty} c_\nu z^\nu/\nu!$ a (K)*E-function* if there is a constant $C > 0$ such that for each $\nu = 0, 1, 2, \ldots$, (a) $c_\nu \in K$, (b) $\overline{|c_\nu|} \leqslant C^{\nu+1}$, (c) there is a sequence q_ν from $\mathbb{N}$ such that (i) for $k = 0, 1, \ldots, \nu$, $q_\nu c_k$ is an algebraic integer and (ii) $0 \leqslant q_\nu \leqslant C^{\nu+1}$.

We assume in the Siegel-Shidlovsky setting for $k = 1, 2, \ldots, m$ that $f_k(z) = \sum_{\nu=0}^{\infty} c_{k\nu} z^\nu/\nu!$ is a KE-function and that C is effective and chosen large enough so that conditions (b) and (c) hold simultaneously for all m of our functions. The proof follows the usual Siegel-Shidlovsky outline, so we first gather together the three standard supplementary lemmas.

LEMMA 2. *Let* $0 < \delta < 1$. *For arbitrary* h, n *in* $\mathbb{N}$ *there exist polynomials*

$$P_{\mathfrak{k}}(z) = n! \sum_{i=0}^{n} P_{\mathfrak{k},i} \frac{z^i}{i!},$$

$\mathfrak{k} = (k_1, \ldots, k_m)$ *with* $k_1 + \cdots + k_m \leqslant h$, *not all* $P_{\mathfrak{k}}$ *zero, such that*

(1) *the* $p_{\mathfrak{k},i}$ *are algebraic integers of* K *satisfying* $\overline{|p_{\mathfrak{k},i}|} \leqslant c_3^{M^2(n+1)\ln(h+1)}$,

(2) *the function*

$$P(z, f_1(z), \ldots, f_m(z)) := \sum_{k_1 + \cdots + k_m \leqslant h} P_{\mathfrak{k}}(z) f_1(z)^{k_1} \cdots f_m(z)^{k_m} = \sum_{\mu=0}^{\infty} a_\mu \frac{z^\mu}{\mu!}$$

has a zero at the origin of order at least $[(M - \delta)(n+1)]$ *where* $M = \binom{h+m}{m}$, *and*

$$\overline{|a_\mu|} \leqslant n^n c_4^{M^2(n+1)\ln(h+1)} (c_5 h)^\mu. \tag{3}$$

This is a straightforward Thue-Siegel application of the box principle (see Lemma 21 of Nesterenko [8]) using properties (a), (b), (c) above. For $k = 1, 2, \ldots$, set $P_k := D^{k-1}P$, so that $P_k(z, f_1(z), \ldots, f_m(z)) = (t(z)d/dz)^{k-1}R(z)$. Then since $\mathcal{O}(P) \geqslant [(M - \delta)(n+1)] > (M - \delta)n$, the Shidlovsky-Nesterenko theorem shows that when $(1 - \delta)n > \tau_1 h^\gamma$, the polynomials $P_1, \ldots, P_M$, considered as linear forms in the $M = \binom{h+m}{m}$ monomials $X_1^{k_1} \cdots X_m^{k_m}$, $k_1 + \cdots + k_m \leqslant h$, are linearly independent over $\mathbb{C}(z)$. Calling these monomials $v_1, \ldots, v_M$ and setting $P_k = \sum_{i=1}^{M} P_{ki}(z) v_i$, $k = 1, 2, \ldots$, we see from this linear independence that

$$\det(P_{ki}(z))_{1 \leqslant k, i \leqslant M} \neq 0.$$

Moreover as usual (see e.g. the proof of Theorem 1 of Nesterenko [7] for the details of the argument), we use the multilinearity of the determinant to express the determinant we have as a $\mathbb{C}[z]$-linear combination of the $P_i(z, f_1(z), \ldots, f_m(z))$. Since the latter have order of vanishing at $z = 0$ at least $\mathcal{O}(P) - (M - 1)$, dividing $\det(P_{ki}(z))$ by the maximal power of z leaves a polynomial of degree at most $\delta n + c_6 M^2$, where $c_6 = \max(\det t, \deg tq_{ki})/2$. We set $t := [\delta n + c_6 M^2]$.

https://doi.org/10.1017/S1446788700022497 Published online by Cambridge University Press

Then Lemma 7 of Shidlovsky [9] (or Section 60 of Mahler [6]) shows that the matrix

$$(P_{k,i}(\alpha))_{\substack{i=1,\dots,M\\ k=1,\dots,M+t}}$$

has rank M.

LEMMA 3. *Let the algebraic number α be such that $\alpha t(\alpha) \neq 0$. Then for all $i = 1,\dots,M$, $k = 1,\dots,M + t$, we have*

$$\overline{|P_{k,i}(\alpha)|} \leqslant n^{(1+\delta)n} c_7^{M^2 n \cdot \ln(h+1)}, \qquad |P_k(\alpha)| \leqslant n^{-(M-1-2\delta)n} c_8^{M^2 n \cdot \ln(h+1)}.$$

This is the content of Lemma 22 of Nesterenko [7].

LEMMA 4. *Let $0 < \delta < 1$. For arbitrary n, h with $h \geqslant 1$ and $(1 - \delta)n > \tau_1 h^\gamma$, there are $M = \binom{h+m}{m}$ linearly independent forms L_i in the numbers $f_1(\alpha)^{k_1} \cdots f_m(\alpha)^{k_m}$, $k_1 + \cdots + k_m \leqslant h$, with coefficients a_{ij} in $K(\alpha)$, $i, j = 1, 2,\dots,M$, which are algebraic integers such that*

$$\overline{|a_{ij}|} \leqslant n^{(1+\delta)n} c_{10}^{M^2 n \ln(h+1)}, \qquad |L_i| \leqslant n^{-(M-1-2\delta)n} c_{11}^{M^2 n \ln(h+1)}.$$

We saw that the matrix

$$(P_{k,i}(\alpha))_{\substack{i=1,\dots,M\\ k=1,\dots,M+t}}$$

has rank M when $(1 - \delta)n > \tau_1 h^\gamma$. Selecting M linearly independent rows and multiplying by $a^{n+s(M+t)}$, where a is a denominator for α, gives us the linear forms L_i. The inequalities follow from the corresponding inequalities of Lemma 3.

4. Precise statement and proofs

THEOREM 2. *Under the hypotheses given in Theorem 1, there is a constant $c_3 > 0$ depending effectively on α, m, κ, the q_{ki} and the constant C in the KE-function criterion and an in general ineffective constant τ_2 depending only on $f_1(z),\dots,f_m(z)$ and the system (1) such that if*

$$\ln H > \max\left\{\tau_2 d^\gamma \ln(\tau_2 d^\gamma), \exp\left(c_3 d^{2m} \ln(d+1)\right)\right\}, \tag{4}$$

then for every non-zero $P(X_1,\dots,X_m)$ in $\mathbb{Z}[X_1,\dots,X_m]$ with degree at most d and height $H(P) \leqslant H$,

$$|P(f_1(\alpha),\dots,f_m(\alpha))| > H^{-c_1 d^m}, \tag{5}$$

where $c_1 = 4^m \kappa^m (m\kappa^2 + \kappa + 1)$.

https://doi.org/10.1017/S1446788700022497 Published online by Cambridge University Press

DEDUCTION OF THEOREM 1 FROM THEOREM 2. For d, $H \geqslant 1$ define $\Phi(d, H) := \min|P(f_1(\alpha),\ldots,f_m(\alpha))|$ where the minimum is taken over all non-zero P in $\mathbb{Z}[X_1,\ldots,X_m]$ with degree P at most d and whose coefficients have maximum modulus $H(P) \leqslant H$. Then Theorem 2 shows that $\ln\Phi(d, H) \geqslant -c_1 d^m \ln H$ when (4) is satisfied. Since $\Phi(d, H)$ is a non-increasing function of both d and H, we see that even when (4) is not satisfied,

$$\begin{aligned}\ln\Phi(d, H) &\geqslant \ln\Phi\big(d, \exp\big(\max\{\tau_2 d^\gamma \ln(\tau_2 d^\gamma), \exp(c_3 d^{2m}\ln(d+1))\}\big)\big)\\ &\geqslant -c_1 d^m \max\{\tau_2 d^\gamma \ln(\tau_2 d^\gamma), \exp(c_3 d^{2m}\ln(d+1))\}.\end{aligned}$$

We now distinguish two cases.

Case 1. If $\ln\tau_2 \geqslant c_3 d^{2m}\ln(d+1)$, then clearly the first term in the right-hand side of (4) dominates. For $c_3 \geqslant 2$, $\ln\tau_2 > d^{2m}$, and so one can show that, say,

$$\tau_2 d^\gamma \ln(\tau_2 d^\gamma) < \tau_2(1 + \gamma/2m)(\ln\tau_2)^{1+\gamma/2}.$$

Consequently

$$\ln\Phi(d, H) \geqslant -c_1 d^m \tau_2 (\ln\tau_2)^{1+\gamma/2} \geqslant -c_1(1+\gamma/2m)\tau_2(\ln\tau_2)^{2+\gamma/2}.$$

Case 2. If on the other hand $\ln\tau_2 \leqslant c_3 d^{2m}\ln(d+1)$ for c_3 large enough to force the aforementioned conclusion in the first case, then

$$\tau_2 d^\gamma \ln(\tau_2 d^\gamma) \leqslant (d+1)^{2(c_3 d^{2m}+\gamma)}.$$

Consequently

$$\ln\Phi(d, H) \geqslant -c_1 d^m (d+1)^{2(c_3 d^{2m}+\gamma)} \geqslant -c_1 (d+1)^{m+2(c_3 d^{2m}+\gamma)}.$$

Taking $c_2 = m + 2(c_3+\gamma) + \log c_1$ and

$$\tau = \exp\big(-c_1(1+\gamma/2m)\tau_2(\ln\tau_2)^{2+\gamma/2}\big) \tag{6}$$

gives the following result, which easily implies Theorem 1.

COROLLARY. *Under the hypotheses of Theorem* 1, *there are effective constants* c_1, $c_2 > 0$ *and an in general ineffective constant* $\tau > 0$ *such that*

$$|P(f_1(\alpha),\ldots,f_m(\alpha))| > \min\{H^{-c_1 d^m}, \exp(-\exp(c_2 d^{2m}\ln(d+1))), \tau\}.$$

PROOF OF THEOREM 2. The proof follows the usual Siegel-Shidlovsky pattern and parallels even more exactly that of Theorem 4 of Nesterenko [8]. The only essential difference lies in our use of the fact that *the constants* c_{10}, c_{11} *of the conclusion of Lemma* 3 *are effective*. Still, for the sake of clarity, we have repeated the proof in enough detail for the reader to verify the genealogy of the constants (a misprint in Nesterenko [8] omits a factor of n in the the last occurrence of what corresponds to our c_{16} and there is called c_{28}).

We assume first for large enough τ_1 and c_3 and for $d, H \geqslant 1$ satisfying (4) that there is a non-zero polynomial R in $\mathbb{Z}[Z_1,\dots,Z_m]$ with degree at most d and height at most H for which

$$|R(f_1(\alpha),\dots,f_m(\alpha))| \leqslant H^{-c_1 d^m}.$$

We shall see that this assumption leads to a contradiction. We set $\nu = m\kappa^2 + 1$, $\lambda = m\kappa$, $h = (\lambda + 1)d$ and determine n by the condition

$$n^n \leqslant H^{(\kappa+\nu)} < (n+1)^{n+1}. \tag{7}$$

In particular we have $n \ln n \leqslant (\kappa + \nu)\ln H$. We remark that in the following proof, the constants $c_{10},\dots,c_{17}$ do not depend on c_2 or c_3. Keeping in mind that $x/\ln x$ is an increasing function of x for $x > e$, we see from the first part of the inequality (4) that $\tau_2 d^\gamma < 2\ln H/\ln\ln H$. Consequently for our choice of n and h, we have from (4) that

$$h^\gamma = (1+\lambda)^\gamma d^\gamma < (1+\lambda)^\gamma \frac{2\ln H}{\tau_2 \ln\ln H}.$$

Similarly from the monotonicity of $x/\ln x$ and our choice of n, we see that for c_3 (and hence H) large enough

$$\frac{n}{2} < \frac{(\kappa+\nu)\ln H}{\ln\ln H} < 2n. \tag{8}$$

From this and the preceding inequality we find that

$$\tau h^\gamma < (4/(\kappa+\nu))(\tau/\tau_2)(1+\lambda)^\gamma n.$$

Thus choosing

$$\tau_2 = 4\tau_1(1+\lambda)^\gamma/(\kappa+\nu)(1-\delta) \tag{9}$$

guarantees that the hypotheses of Lemma 4 are fulfilled for every $0 < \delta < 1$. Let us remark that this will be the last mention of any ineffective constant in the proof. Choosing, say $\delta = 1/4\lambda(\kappa+\nu) < 1$ puts the linear forms $L_1,\dots,L_M$ described by Lemma 4 at our disposal.

First consider the linear forms $l_{\mathfrak{r}}$ in the monomials $(f_1(\alpha))^{k_1}\cdots(f_m(\alpha))^{k_m}$ defined by

$$l_{\mathfrak{r}} = (f_1(\alpha))^{r_1}\cdots(f_m(\alpha))^{r_m} R(f_1(\alpha),\dots,f_m(\alpha))$$

for $r_1 + \cdots + r_m \leqslant \lambda d$ obtained from R simply by multiplying with the indicated power products. Obviously $k_1 + \cdots + k_m \leqslant \lambda d + d = h$. These $M_1 = \binom{\lambda d + m}{m}$ forms are linearly independent, as can be seen from considering the term of R of maximal lexicographical order.

In addition to the $l_{\mathfrak{r}}$ we select $N := M - M_1$ linear forms from among $L_1,\dots,L_M$ to obtain a system of M linearly independent forms, say, $l_1,\dots,l_{M_1}$;

https://doi.org/10.1017/S1446788700022497 Published online by Cambridge University Press

$L_1, \ldots, L_N$. Let Δ denote the determinant of this system. From Lemma 4 we see that

$$\overline{|\Delta|} \leqslant M!H^{M_1}\left[n^{(1+\delta)n}c_{10}^{M^2 n \ln(h+1)}\right]^N \leqslant H^{M_1}c_{12}^{M^2 Nn \ln(h+1)}n^{(1+\delta)Nn}.$$

Furthermore, multiplying every column in Δ by the corresponding expression $(f_1(\alpha))^{k_1} \cdots (f_m(\alpha))^{k_m}$ when $k_1 + \cdots + k_m \neq 0$ and adding to the exceptional column corresponding to $(f_1(\alpha))^0 \cdots (f_m(\alpha))^0$, we obtain the expressions l_i, $i = 1, \ldots, M_1$, and L_i, $i = 1, \ldots, N$, as the entries in this column. Then expanding along it, we have

$$\Delta = \sum_{i=1}^{M_1} l_i \Delta_i + \sum_{i=1}^{N} L_i \Delta_{M+i}, \tag{10}$$

where Δ_i is the appropriate cofactor.

It is easily seen that

$$|l_i| \leqslant c_{13}^h |R(f_1(\alpha), \ldots, f_m(\alpha))| \leqslant c_{13}^h H^{-c_1 d^m},$$

$$\max_{i \leqslant M_1} |\Delta_i| \leqslant H^{M_1 - 1} c_{12}^{M^2 Nn \ln(h+1)} n^{(1+\delta)n},$$

$$\max_{i > M_1} |\Delta_i| \leqslant H^{M_1} c_{12}^{M^2 Nn \ln(h+1)} n^{(1+\delta)(N-1)n}.$$

Applying these inequalities in (10) and using the bounds for L_i from Lemma 4, we find that

$$|\Delta| \leqslant M!H^{M_1} n^{(1+\delta)Nn} c_{14}^{M^2 Nn \ln(h+1)} \max\left\{H^{-c_1 d^m - 1}, n^{-(M-\delta)n}\right\}.$$

Since Δ is a non-zero integer of the field K, its norm has modulus at least 1. Thus

$$1 \leqslant H^{\kappa M_1} n^{\kappa(1+\delta)Nn} c_{15}^{M^2 Nn \ln(h+1)} \max\left\{H^{-c_1 d^m - 1}, n^{-(M-\delta)n}\right\}. \tag{11}$$

Note that $M = \binom{h+m}{m}$, $N < M$ and $= (1 + \lambda)d$ so that

$$M^2 Nn \ln(h+1) \leqslant c_{16} n d^{3m} \ln(d+1) \leqslant \frac{c_{16}}{c_3} n d^m \ln\ln H$$

from the second part of the inequality (4) of the hypotheses. Together with (8) and our choice of n in (7), this gives

$$1 \leqslant H^{\kappa[M_1 + (1+\delta)(\kappa+\nu)N] + [c_{17}/c_3]d^m} \max\left\{H^{-c_1 d^m - 1}, n^{-(M-\delta)n}\right\}.$$

Recalling our choice (7) of n, we verify that the maximum will be contributed by the second term in case $c_1 d^m \geqslant (\kappa + \nu)M$. But

$$M < \kappa^m d^m (2m+1) \cdots (m+2)/m!.$$

Now an application of Stirling's formula to $(2m+1)(2m)!/(m+1)(m!)^2$ shows that $M < 4^m \kappa^m d^m$. Thus our choice of c_1 ensures that the maximum is indeed contributed by the second term above and thus

$$1 \leqslant H^{\kappa[M_1+(1+\delta)(\kappa+\nu)N]+[c_{17}/c_3]d^m} H^{-(M-\delta)(\kappa+\nu)}.$$

Writing in terms of exponents yields

$$0 \leqslant \kappa\left[M_1 + (1+\delta)(\kappa+\nu)N\right] + \frac{c_{17}}{c_3}d^m - (M-\delta)(\kappa+\nu) \tag{12}$$

when we use the inequality provided by the left side of (7).

By Lemma 24 of Nesterenko [8], we have the inequality $N < mM/(\lambda+1)$. Consequently from (12) it follows that

$$0 < \kappa M + \frac{(1+\delta)(\kappa+\nu)}{\lambda+1} m\kappa M + \frac{c_{17}}{c_3}d^m - (M-\delta)(\kappa+\nu).$$

Using the definition of λ and ν, one calculates the right-hand side to be

$$\begin{aligned}
&\frac{\kappa(\lambda+1) + (1+\delta)(\kappa+\nu)\lambda - (\kappa+\nu)(\lambda+1)}{\lambda+1}M + \delta(\kappa+\nu) + \frac{c_{17}}{c_3}d^m \\
&\qquad = \left(\frac{\lambda\delta}{\lambda+1}(\kappa+\nu) - \frac{1}{\lambda+1}\right)M + \delta(\kappa+\nu) + \frac{c_{17}}{c_3}d^m \\
&\qquad < \left(\frac{3\delta\lambda(\kappa+\nu)}{\lambda+1} - \frac{1}{\lambda+1}\right)M + \frac{c_{17}}{c_3}d^m \\
&\qquad < -\delta M + \frac{c_{17}}{c_3}d^m,
\end{aligned}$$

when we invoke the definition of δ. The expression is thus negative if c_3 is taken large enough. This contradiction of (12) refutes the existence of the polynomial R satisfying (4) but not (5). Hence Theorem 2 is established.

5. Further remarks

It is possible to deduce directly from Nesterenko's result inequalities of the same form as ours, albeit weaker in their dependence on d. For example one obtains, as the analogue of our Corollary,

$$\Phi(d, H) \geqslant \max\left\{H^{-c_1 d^m}, \exp\left(-\exp\left(d^{4m}(\ln(d+1))^2\right)\right), \tau\right\}$$

by considering which of τ or $d^{2m}\ln(d+1)$ is larger.

Finally lęt us remark that the Dirichlet box principle shows that an inequality of the form $\Phi(d, H) > H^{-cd^m}$ would be optimal. So inequality (5) cannot be improved substantially for H large enough compared to d and τ. The unfortunate doubly exponential dependence on d for smaller H comes from the necessity to use the auxiliary function constructed in the lemmas to provide the forms

$L_1, \ldots, L_N$ to complement those coming from R itself. Those forms have inordinately large coefficients since the Shidlovsky Lemma requires a function with order of zero at $z = 0$ greater that $(M-1)n + \tau_1 h$. But we only have Mn unknowns (coefficients of the auxiliary function). So although the integer coefficients of the equations are bounded by h^{cMn}, when we are forced to solve so many equations, the entries in the solution may grow to about $h^{cM^2 n}$. Finally expanding the determinant gives $h^{cM^2 Nn}$ as an upper bound for the contribution from the coefficients of the auxiliary function. This must be overcome by the known smallness n^{-Mn} of the absolute values $|L_1|, \ldots, |L_N|$. Thus $MNn \ln(h+1)$ must be dominated by $n \ln n$ in order to achieve the desired contradiction. This means that $M^2 \ln(h+1) \ll \ln n$.

On the other hand, for our contradiction of (11), we must have $Nn \ln n \ll d^m \ln H$. Since $N \gg d^m$, this gives $n \ln n \ll \ln H$. Together with $M^2 \ln(h+1) \ll \ln n$ this gives in turn

$$\ln H \gg n \ln n \gg \exp\big(cM^2 \ln(h+1)\big) M^2 \ln(h+1) \gg \exp\big(cd^{2m} \ln(d+1)\big).$$

Thus to improve this aspect of the measure, it seems that one will have to find an approach which does not depend on Shidlovsky's Lemma.

References

[1] D. Bertrand and F. Beukers, 'Equations différentielles linéaires et majorations de multiplicités', to appear.

[2] W. D. Brownawell, 'Travaux récents de Ju. V. Nesterenko', *Séminaire Delange-Pisot-Poitou*, 1977/78, no. 35, 6 p.

[3] N. I. Feldman and A. B. Shidlovsky, 'The development and present state of the theory of transcendental numbers', *Uspehi Mat. Nauk* **22** No. 3 (1967), 3–81. English transl. *Russian Math. Surveys* **22** No. 3 (1967), 1–79.

[4] S. Lang, 'A transcendence measure for E-functions', *Mathematika* **9** (1962), 157–161.

[5] K. Mahler, 'Zur Approximation der Exponentialfunktion und des Logarithmus. I' *J. Reine Angew. Math.* **166** (1932), 118–136.

[6] K. Mahler, *Lectures on transcendental numbers* (Lecture Notes in Mathematics, No. 546, Springer-Verlag, Berlin, Heidelberg, New York, 1976).

[7] Yu. V. Nesterenko, 'On the algebraic dependence of components of solutions of systems of linear differential equations', *Izv. Akad. Nauk SSSR, Ser. Mat.* **38** (1974), 495–512. English transl. *Math. USSR Izv.* **8** (1974), 501–518.

[8] Yu. V. Nesterenko, 'Bounds on the order of zeros of a class of functions and their application to the theory of transcendental numbers', *Izv. Akad. Nauk SSSR, Ser. Mat* **41** (1977), 239–270. English transl. *Math. USSR Izv.* **11** (1977), 253–284.

[9] A. B. Shidlovsky, 'A criterion for algebraic independence of the values of a class of entire functions', *Izv. Akad. Nauk SSSR Ser. Mat.* **23** (1959), 35–66. English transl. *Amer. Math. Soc. Transl.* (2) **22** (1962), 339–370.

[10] A. B. Shidlovsky, 'On arithmetic properties of values of analytic functions', *Trudy MIAN SSSR Steklova* **132** (1973), 169–202. English transl. *Proceedings of the Steklov Institute of Math.*, No. 132 (1973), 193–233.
[11] C. L. Siegel, 'Über einige Anwendungen diophantischer Approximationen', *Abh. Preuss. Akad. Wiss. Phys.-mat. Kl.* Berlin (1929), reprinted in *Carl Ludwig Siegel Gesammelte Abhandlungen*, I. Springer-Verlag, Berlin, Heidelberg, New York, 1966, 209–266.
[12] C. L. Siegel, 'Transcendental numbers', *Ann. of Math. Studies* **16** (Princeton University Press, Princeton, N. J., 1949).
[13] N. T. Tai, 'On estimates for the orders of zeros of polynomials in analytic functions and their application to estimates for the relative transcendence measure of values of E-functions', *Mat. Sb.* **120** (1983), 112–142. English transl. *Math. USSR Sb.* **48** (1984), 111–140.

Department of Mathematics
Pennsylvania State University,
University Park, Pennsylvania 16802
U.S.A.