Proceedings of the Royal Society of Edinburgh, **148A**, 1297–1311, *2018*
DOI:10.1017/S0308210518000148

On linear independence measures of the values of Mahler functions

Keijo Väänänen
Department of Mathematical Science, University of Oulu,
P.O. Box 3000, 90014 Oulu, Finland (keijo.vaananen@oulu.fi)

Wen Wu*
School of Mathematics, South China University of Technology,
Guangzhou 510641, People's Republic of China
(hust.wuwen@gmail.com)

(MS received 21 June 2016; accepted 14 November 2016)

We estimate the linear independence measures for the values of a class of Mahler functions of degrees 1 and 2. For this purpose, we study the determinants of suitable Hermite–Padé approximation polynomials. Based on the non-vanishing property of these determinants, we apply the functional equations to get an infinite sequence of approximations that is used to produce the linear independence measures.

Keywords: linear form; Mahler function; Hermite–Padé approximation

2010 *Mathematics subject classification:* Primary 11J13
Secondary 11J72; 39B32

1. Introduction and results

Rational approximations of the values of Mahler functions have been an active research area in the past few years. Since Bugeaud's remarkable paper [4], where he proved that the irrationality exponent of Thue–Morse numbers is 2, several papers have appeared on the irrationality exponents of the values of degree 1 Mahler functions (see [2, 5, 8, 11, 13, 16]). A good overview of these results is given in [5]. In particular, we know that the irrationality exponents of the numbers in theorems 1.1, 1.3 and 1.4 are equal to 2.

As a natural generalization of the above results, our aim here is to obtain linear independence measures for the values of a class of Mahler functions $F(z), G(z) \in \mathbb{Q}[\![z]\!]$ converging on some open disc $D_r := \{z\colon |z| < r \leqslant 1\}$ and satisfying a system of Mahler-type functional equations:

$$\left.\begin{aligned} F(z^d) &= p_{11}(z)F(z) + p_{12}(z)G(z) + p_{10}(z),\\ G(z^d) &= p_{21}(z)F(z) + p_{22}(z)G(z) + p_{20}(z),\end{aligned}\right\} \tag{1.1}$$

where $p_{ij}(z) \in \mathbb{Q}(z)$ and $p_{11}(z)p_{22}(z) - p_{12}(z)p_{21}(z) \neq 0$. Note that Mahler functions of degree 1 or 2 satisfy functional equations of the above type: if $F(z)$ and $G(z)$

*Corresponding author.

are Mahler functions of degree 1, then $p_{12}(z) = p_{21}(z) = 0$, and if $F(z)$ is of degree 2, then we choose $G(z) = F(z^d)$, $p_{12}(z) = 1$ and $p_{11}(z) = p_{10}(z) = 0$. Our general result (theorem 4.1) needs some technical notation (to be presented later), and therefore to introduce our results we now demonstrate applications to some well-known functions.

The linear independence measures studied here are lower bounds for linear forms (in 1 and certain numbers γ_1 and γ_2) of the form

$$|h_0 + h_1\gamma_1 + h_2\gamma_2| > CH^{-\mu}, \tag{1.2}$$

valid for any integers h_0, h_1, h_2, not all zero, where the exponent μ is given explicitly, $H = \max\{|h_1|, |h_2|, H_0\}$ and positive constants C and H_0 are independent of h_i. In our results, γ_1 and γ_2 are the values of the functions under consideration at rational points $a/b \in D_r\backslash\{0\}$, where

$$\log|a| = \lambda \log b \quad (0 \leqslant \lambda < \log(rb)/\log b).$$

We note that generally [12, theorem 4.4.1] implies the existence of a μ ($\mu \geqslant 2$) in our cases below, and here our aim is to obtain an explicit upper bound for the linear independence exponent

$$\mu(\gamma_1, \gamma_2) := \inf\{\mu\colon (1.2) \text{ holds for some } C > 0,\ H_0 > 0\}.$$

It is known that $\mu(\gamma_1, \gamma_2) = 2$ for Lebesgue almost all $(\gamma_1, \gamma_2) \in \mathbb{R}^2$ and $\mu(\gamma, \gamma^2) = 2$ for Lebesgue almost all $\gamma \in \mathbb{R}$, but if a pair $(\gamma_1, \gamma_2) \in \mathbb{R}^2$ is given, it is usually difficult to determine $\mu(\gamma_1, \gamma_2)$ or even an upper bound for it. By the Schmidt subspace theorem, $\mu(\gamma_1, \gamma_2) = 2$ if γ_1 and γ_2 are real algebraic numbers such that 1, γ_1 and γ_2 are linearly independent over the rationals, and $\mu(\gamma, \gamma^2) = 2$ for all real algebraic numbers γ of degree greater than or equal to 3. We also know pairs of transcendental numbers having linear independence exponent 2. For instance, Popken proved in 1929 that $\mu(e, e^2) = 2$, and there are similar results for the values of more general Siegel E-functions. Furthermore, the theory of linear forms in logarithms implies that $\mu(\gamma_1, \gamma_2)$ has an effectively computable upper bound if γ_i are values of the logarithmic function at algebraic points. All these results are presented, for example, in [10]; note also that often analogous results hold more generally for similarly defined $\mu(\gamma_1, \ldots, \gamma_m)$ with $m \geqslant 3$.

In the following applications of theorem 4.1, we choose five pairs of Mahler functions. The first three pairs, which are degree 1 Mahler functions, satisfy different shapes of Mahler-type functional equations. In our fourth example, the functions are related to the Rudin–Shapiro sequence and satisfy more general functional equations of type (1.1). The Mahler functions in the last pair are degree 2 Mahler functions.

1.1. The Thue–Morse number and its square

Our first application is to study the product

$$T(z) = \prod_{j=0}^{\infty}(1 - z^{2^j}),$$

the generating function of the Thue–Morse sequence on $\{-1, 1\}$, satisfying

$$T(z) = (1 - z)T(z^2). \tag{1.3}$$

THEOREM 1.1. *We have*

$$\mu\left(T\left(\frac{1}{b}\right), T^2\left(\frac{1}{b}\right)\right) \leqslant \frac{91}{32} \approx 2.843\ldots.$$

More generally, if $0 \leqslant \lambda < \frac{7}{29}$, *then*

$$\mu\left(T\left(\frac{a}{b}\right), T^2\left(\frac{a}{b}\right)\right) \leqslant \frac{91}{32 - 104\lambda}.$$

It is well known that a bound for $\mu(\gamma, \gamma^2)$ implies a bound for the approximation of γ by quadratic algebraic numbers α. Namely, if $p(x) = h_0 + h_1x + h_2x^2 \in \mathbb{Z}[x]$ is the minimal polynomial of α and $|\gamma - \alpha| < 1$, then

$$|p(\gamma)| = |p(\gamma) - p(\alpha)| \leqslant ch|\gamma - \alpha|,$$

where $h = \max\{|h_i|\}$ is the height of α and $c = 2(1 + |\gamma|)$ is a positive constant independent of α. If we take $\gamma = T(1/b)$ here, then theorem 1.1 implies the following corollary on the approximation of $T(1/b)$ by quadratic numbers.

COROLLARY 1.2. *If* $\varepsilon > 0$ *is given, then there exist positive constants* $C_1 = C_1(b, \varepsilon)$ *and* $H_1 = H_1(b, \varepsilon)$ *such that, for all algebraic numbers* α *of degree less than or equal to* 2 *and height less than or equal to* h,

$$\left|T\left(\frac{1}{b}\right) - \alpha\right| > C_1 H^{-\omega-\varepsilon},$$

where $\omega = \frac{123}{32} \approx 3.843\ldots$ *and* $H = \max\{h, H_1\}$.

1.2. Stern's sequence and its twisted version

Next, let $A(z)$ and $B(z)$ be the generating functions of Stern's diatomic sequence and its twisted version, respectively. These functions satisfy the functional equations

$$A(z) = (1 + z + z^2)A(z^2), \qquad B(z) = 2 - (1 + z + z^2)B(z^2), \tag{1.4}$$

of type (1.1) (see, for example, [6]).

THEOREM 1.3. *We have*

$$\mu\left(A\left(\frac{1}{b}\right), B\left(\frac{1}{b}\right)\right) \leqslant \frac{26}{9} \approx 2.888\ldots.$$

More generally,

$$\mu\left(A\left(\frac{a}{b}\right), B\left(\frac{a}{b}\right)\right) \leqslant \begin{cases} \dfrac{130}{45 - 149\lambda} & \text{if } \lambda < \dfrac{145}{1289}, \\[2ex] \dfrac{69}{25 - 89\lambda} & \text{if } \dfrac{145}{1289} \leqslant \lambda < \dfrac{5}{29}. \end{cases}$$

1.3. Lambert series $G_3(z)$ and $F_3(z)$

The functions

$$G_3(z) = \sum_{j=0}^{\infty} \frac{z^{3^j}}{1-z^{3^j}}, \qquad F_3(z) = \sum_{j=0}^{\infty} \frac{z^{3^j}}{1+z^{3^j}} = -G_3(-z)$$

satisfy

$$(1-z)G_3(z) - (1-z)G_3(z^3) - z = 0, \qquad (1+z)F_3(z) - (1+z)F_3(z^3) - z = 0. \tag{1.5}$$

The following result studies the values of these typical examples of Mahler functions.

THEOREM 1.4. *We have*

$$\mu\left(G_3\left(\frac{1}{b}\right), F_3\left(\frac{1}{b}\right)\right) \leqslant \frac{129}{37} \approx 3.486\ldots.$$

More generally,

$$\mu\left(G_3\left(\frac{a}{b}\right), F_3\left(\frac{a}{b}\right)\right) \leqslant \begin{cases} \dfrac{129}{37-119\lambda} & \text{if } \lambda < \dfrac{25}{443}, \\ \dfrac{83}{24-80\lambda} & \text{if } \dfrac{25}{443} \leqslant \lambda < \dfrac{43}{337}, \\ \dfrac{57}{17-59\lambda} & \text{if } \dfrac{43}{337} \leqslant \lambda < \dfrac{7}{29}. \end{cases}$$

1.4. The Rudin–Shapiro sequence

Let $(r_n)_{n\geqslant 0}$ be the Rudin–Shapiro sequence defined by $r_0 = 1, r_{2n} = r_n, r_{2n+1} = (-1)^n r_n$. Its generating function, $R(z) = \sum_{n\geqslant 0} r_n z^n$, satisfies

$$R(z) = R(z^2) + zR(-z^2). \tag{1.6}$$

We shall investigate the values of $R(z)$ and $R(-z)$ at some rational points.

THEOREM 1.5. *We have*

$$\mu\left(R\left(\frac{1}{b}\right), R\left(-\frac{1}{b}\right)\right) \leqslant \frac{13}{4} = 3.25.$$

More generally,

$$\mu\left(R\left(\frac{a}{b}\right), R\left(-\frac{a}{b}\right)\right) \leqslant \begin{cases} \dfrac{39}{12-40\lambda} & \text{if } \lambda < \dfrac{21}{187}, \\ \dfrac{47}{15-53\lambda} & \text{if } \dfrac{21}{187} \leqslant \lambda < \dfrac{3}{13}. \end{cases}$$

1.5. A degree 2 Mahler function

As an example of degree 2 Mahler functions we take the function $S(z)$ satisfying $S(0) = 1$ and

$$zS(z) - (1 + z + z^2)S(z^4) + S(z^{16}) = 0. \tag{1.7}$$

This function was introduced by Dilcher and Stolarsky [9], and it has been studied recently in, for example, [1, 3, 7]. In particular, the algebraic independence of $S(\alpha)$, $S'(\alpha)$, $S(\alpha^4)$ and $S'(\alpha^4)$ is proved in [3] for all algebraic α, $0 < |\alpha| < 1$. Note also that, in [7], an upper bound 5 is obtained for the irrationality exponent of $S(1/b)$.

THEOREM 1.6. *We have*

$$\mu\left(S\left(\frac{1}{b}\right), S\left(\frac{1}{b^4}\right)\right) \leqslant \frac{167}{25} = 6.68.$$

More generally, if $0 \leqslant \lambda < \frac{1}{5}$, *then*

$$\mu\left(S\left(\frac{a}{b}\right), S\left(\left(\frac{a}{b}\right)^4\right)\right) \leqslant \frac{167}{25 - 93\lambda}.$$

This work is a continuation of [14], in which we the studied simultaneous approximations of similar numbers. These results and Khintchine's transference theorem can be used to obtain linear independence measures for the numbers in theorems 1.3, 1.4 and 1.6, but the results are weaker than those obtained in this paper.

The results on irrationality exponents mentioned above are based on the non-vanishing property of certain Hankel determinants. Analogously, all results here are based on the non-vanishing property of the determinants of suitable Hermite–Padé approximation polynomials, which are given explicitly in [15, appendix]. The non-vanishing property is verified here by computing the determinants, but it would be of great interest to find a more general criterion for this. Once some non-zero determinants have been obtained, the functional equations can be used to produce a sufficiently dense infinite sequence of approximations with non-zero determinants. It is well known that such approximations can be used to produce linear independence measures along the lines of Siegel's method. Section 3 contains this consideration, and it is then applied to prove a general result in § 4. The proofs of theorems 1.1 and 1.3–1.6 are given in § 5.

2. Important determinants

We first note that system (1.1) can be given in the form

$$P(z)F(z^d) = P_{11}(z)F(z) + P_{12}(z)G(z) + P_{10}(z), \tag{2.1}$$

$$P(z)G(z^d) = P_{21}(z)F(z) + P_{22}(z)G(z) + P_{20}(z), \tag{2.2}$$

where $P(z)$, the least common denominator of $p_{ij}(z)$, and $P_{ij}(z) = P(z)p_{ij}(z)$ belong to $\mathbb{Z}[z]$ and satisfy $P_{11}(z)P_{22}(z) - P_{12}(z)P_{21}(z) \neq 0$.

For an integer $k \geqslant 1$, let $A_k(z), B_k(z), C_k(z) \in \mathbb{Z}[z]$ denote $(d_1, d_2, d_3) = (d_1(k), d_2(k), d_3(k))$ Hermite–Padé approximation polynomials of $F(z)$, $G(z)$ and 1, so

$$A_k(z)F(z) + B_k(z)G(z) + C_k(z) = R_k(z), \tag{2.3}$$

where $\deg A_k(z) \leqslant d_1$, $\deg B_k(z) \leqslant d_2$, $\deg C_k(z) \leqslant d_3$ and the order of zero of the remainder term $R_k(z)$ at $z = 0$ satisfies $\operatorname{ord} R_k(z) =: o(k) \geqslant d_1 + d_2 + d_3 + 2$. By comparing the coefficients of z^j for all $0 \leqslant j < o(k)$, (2.3) yields a system of $o(k)$ homogeneous equations for $o(k) + 1$ unknown coefficients of $A_k(z)$, $B_k(z)$ and

$C_k(z)$, which implies that such polynomials exist and at least one of $A_k(z)$, $B_k(z)$ is non-zero. Substituting z^d for z in (2.3) and applying (2.1) and (2.2), we obtain

$$\begin{aligned}(P_{11}(z)A_k(z^d) + P_{21}(z)B_k(z^d))F(z)\\ + (P_{12}(z)A_k(z^d) + P_{22}(z)B_k(z^d))G(z)\\ + P_{10}(z)A_k(z^d) + P_{20}(z)B_k(z^d) + P(z)C_k(z^d) = P(z)R_k(z^d).\end{aligned}$$

Repeating this procedure m times yields

$$A_{k,m}(z)F(z) + B_{k,m}(z)G(z) + C_{k,m}(z) = R_{k,m}(z), \quad m = 0, 1, \dots, \tag{2.4}$$

where $A_{k,0}(z) = A_k(z)$, $B_{k,0}(z) = B_k(z)$, $C_{k,0}(z) = C_k(z)$, $R_{k,0}(z) = R_k(z)$ and, for $m = 1, 2, \dots,$

$$\left.\begin{aligned}A_{k,m}(z) &= P_{11}(z)A_{k,m-1}(z^d) + P_{21}(z)B_{k,m-1}(z^d),\\ B_{k,m}(z) &= P_{12}(z)A_{k,m-1}(z^d) + P_{22}(z)B_{k,m-1}(z^d),\\ C_{k,m}(z) &= P_{10}(z)A_{k,m-1}(z^d) + P_{20}(z)B_{k,m-1}(z^d) + P(z)C_{k,m-1}(z^d),\\ R_{k,m}(z) &= P(z)R_{k,m-1}(z^d).\end{aligned}\right\} \tag{2.5}$$

We are interested in the determinants

$$\Delta(\boldsymbol{k}, m, z) := \det\begin{pmatrix}A_{k_1,m}(z) & B_{k_1,m}(z) & C_{k_1,m}(z)\\ A_{k_2,m}(z) & B_{k_2,m}(z) & C_{k_2,m}(z)\\ A_{k_3,m}(z) & B_{k_3,m}(z) & C_{k_3,m}(z)\end{pmatrix},$$

where $1 \leqslant k_1 < k_2 < k_3$. By the above recursions (2.5),

$$\Delta(\boldsymbol{k}, m, z) = \Phi(z)\Delta(\boldsymbol{k}, m-1, z^d), \quad \Phi(z) := (P_{11}(z)P_{22}(z) - P_{12}(z)P_{21}(z))P(z),$$

and so

$$\Delta(\boldsymbol{k}, m, z) = \Delta(\boldsymbol{k}, 0, z^{d^m})\prod_{j=0}^{m-1}\Phi(z^{d^j}). \tag{2.6}$$

In particular, for degree 1 functions we have $\Phi(z) = P_{11}(z)P_{22}(z)P(z)$, and for the degree 2 function $F(z)$ with $G(z) = F(z^d)$ we have $\Phi(z) = -P_{21}(z)P^2(z)$, since $P_{11}(z) = 0$ and $P_{12}(z) = P(z)$.

Let $\bar{d}(k) := \max\{d_1(k), d_2(k), d_3(k)\}$. By our assumption $k_1 < k_2 < k_3$, it is natural to assume that $\bar{d}(k_1) \leqslant \bar{d}(k_2) \leqslant \bar{d}(k_3)$ and $o(k_1) \leqslant o(k_2) \leqslant o(k_3)$. Since

$$\Delta(\boldsymbol{k}, 0, z) = \det\begin{pmatrix}A_{k_1}(z) & B_{k_1}(z) & R_{k_1}(z)\\ A_{k_2}(z) & B_{k_2}(z) & R_{k_2}(z)\\ A_{k_3}(z) & B_{k_3}(z) & R_{k_3}(z)\end{pmatrix},$$

it follows that $o(k_1) \leqslant \operatorname{ord}\Delta(\boldsymbol{k}, 0, z) \leqslant \deg\Delta(\boldsymbol{k}, 0, z) \leqslant \bar{d}(k_1) + \bar{d}(k_2) + \bar{d}(k_3)$ if $\Delta(\boldsymbol{k}, 0, z) \neq 0$. Thus, in this case

$$\Delta(\boldsymbol{k}, 0, z) =: z^{o(k_1)}D(\boldsymbol{k}, z) \tag{2.7}$$

with some polynomial $D(\boldsymbol{k}, z) \neq 0, \deg D(\boldsymbol{k}, z) \leqslant \bar{d}(k_1) + \bar{d}(k_2) + \bar{d}(k_3) - o(k_1)$. Furthermore, if $o(k_1) > \bar{d}(k_1) + \bar{d}(k_2) + \bar{d}(k_3)$, then $\Delta(\boldsymbol{k}, 0, z) = 0$.

We note that the condition $D(\boldsymbol{k}, z) \neq 0$ gives a strong restriction to $o(k_1)$. For example, if $d_j(k_1) = k$, $d_j(k_2) = k+1$, $d_j(k_3) = k+2$ $(j = 1, 2, 3)$, then $\deg \Delta(\boldsymbol{k}, 0, z) \leqslant 3k+3$ and $o(k_1) \geqslant 3k+2$. Thus, the condition $D(\boldsymbol{k}, z) \neq 0$ is possible only if $3k+2 \leqslant o(k_1) \leqslant 3k+3$.

The above condition means that one determinant $\Delta(\boldsymbol{k}, 0, z) \neq 0$ gives an infinite sequence of determinants $\Delta(\boldsymbol{k}, m, z) \neq 0$, $m = 0, 1, \dots$. When considering the values of the functions at rational points $z = a/b$ we need to know that $\Delta(\boldsymbol{k}, m, a/b) \neq 0$ at least for all sufficiently large m. This condition can be verified in many concrete cases by using (2.6) and (2.7), since $\deg D(\boldsymbol{k}, z)$ is small.

3. Fundamental lemma

In this section γ_1 and γ_2 denote real numbers and $b \geqslant 2$ is an integer. Let $\boldsymbol{k} = \boldsymbol{k}(\ell) = (k_{\ell,1}, k_{\ell,2}, k_{\ell,3})$ $(\ell = 1, \dots, L)$ be vectors with positive integer components $k_{\ell,i}$ satisfying $k_{\ell,1} < k_{\ell,2} < k_{\ell,3}$ and $k_{\ell,3} \leqslant k_{\ell+1,1}$ $(\ell = 1, \dots, L-1)$, $k_{L,3} \leqslant dk_{1,1}$. Assume that for each $k = k_{\ell,i}$ there exists an integer $m_0(k)$ such that for all $m \geqslant m_0(k)$ we have the linear forms

$$a_{k,m}\gamma_1 + b_{k,m}\gamma_2 + c_{k,m} = r_{k,m}$$

with the following properties.

(i) The coefficients $a_{k,m}, b_{k,m}, c_{k,m} \in \mathbb{Z}$ and satisfy

$$\max\{|a_{k,m}|, |b_{k,m}|\} \leqslant c_1(k) b^{E(k)d^m}, \tag{3.1}$$

where $E(k)$ and $c_1(k)$ (and, later, $c_2(k), \dots$) are positive constants independent of m.

(ii) We have

$$|r_{k,m}| \leqslant c_2(k) b^{-V(k)d^m}, \tag{3.2}$$

where $V(k) > 0$ is independent of m.

(iii) The determinant

$$\det \begin{pmatrix} a_{k_{\ell,1},m} & b_{k_{\ell,1},m} & c_{k_{\ell,1},m} \\ a_{k_{\ell,2},m} & b_{k_{\ell,2},m} & c_{k_{\ell,2},m} \\ a_{k_{\ell,3},m} & b_{k_{\ell,3},m} & c_{k_{\ell,3},m} \end{pmatrix} \neq 0$$

for all $\ell = 1, \dots, L$; $m \geqslant m_0(\boldsymbol{k}(\ell)) = \max_{1 \leqslant i \leqslant 3}\{m_0(k_{\ell,i})\}$.

For the following fundamental lemma, we finally define, for all $\ell = 1, \dots, L$, the notation

$$\theta(\ell) = \max_{1 \leqslant i < j \leqslant 3}\{E(k_{\ell,i}) + E(k_{\ell,j})\},$$
$$\nu(\ell) = \min_{\substack{1 \leqslant i,j \leqslant 3 \\ i \neq j}}\{V(k_{\ell,i}) - E(k_{\ell,j})\},$$

and set $K := (\boldsymbol{k}(1), \dots, \boldsymbol{k}(L))$.

https://doi.org/10.1017/S0308210518000148 Published online by Cambridge University Press

Lemma 3.1. *Suppose that* $0 < \nu(1) < \cdots < \nu(L) < d\nu(1)$. *Then there exist positive constants* $C = C(K)$ *and* $H_0 = H_0(K)$ *such that for any integers* h_0, h_1, h_2 *not all zero,*

$$|h_0 + h_1\gamma_1 + h_2\gamma_2| > CH^{-\mu},$$

where $H = \max\{|h_1|, |h_2|, H_0\}$ *and*

$$\mu = \max_{1 \leqslant \ell \leqslant L} \mu(\ell), \quad \mu(\ell) := \frac{\theta(\ell+1)}{\nu(\ell)}, \quad \theta(L+1) := d\theta(1).$$

Proof. Let

$$\Lambda = h_0 + h_1\gamma_1 + h_2\gamma_2.$$

By condition (iii) above, for all $\ell = 1, \dots, L$ there exist $1 \leqslant i < j \leqslant 3$ such that

$$D(\boldsymbol{k}(\ell), \boldsymbol{h}) := \det\begin{pmatrix} h_1 & h_2 & h_0 \\ a_{k_{\ell,i},m} & b_{k_{\ell,i},m} & c_{k_{\ell,i},m} \\ a_{k_{\ell,j},m} & b_{k_{\ell,j},m} & c_{k_{\ell,j},m} \end{pmatrix}$$
$$= \det\begin{pmatrix} h_1 & h_2 & \Lambda \\ a_{k_{\ell,i},m} & b_{k_{\ell,i},m} & r_{k_{\ell,i},m} \\ a_{k_{\ell,j},m} & b_{k_{\ell,j},m} & r_{k_{\ell,j},m} \end{pmatrix} \neq 0.$$

Since $D(\boldsymbol{k}(\ell), \boldsymbol{h})$ is an integer, we obtain, by (3.1) and (3.2),

$$\begin{aligned} 1 \leqslant {} & 2|\Lambda| c_1(k_{\ell,i}) c_1(k_{\ell,j}) b^{(E(k_{\ell,i})+E(k_{\ell,j}))d^m} \\ & + 2hc_1(k_{\ell,j}) c_2(k_{\ell,i}) b^{-(V(k_{\ell,i})-E(k_{\ell,j}))d^m} \\ & + 2hc_1(k_{\ell,i}) c_2(k_{\ell,j}) b^{-(V(k_{\ell,j})-E(k_{\ell,i}))d^m} \end{aligned} \tag{3.3}$$

with $h = \max\{|h_1|, |h_2|\}$. The definitions of $\theta(\ell)$ and $\nu(\ell)$ then give

$$1 \leqslant C_1(K)|\Lambda| b^{\theta(\ell)d^m} + C_2(K) h b^{-\nu(\ell)d^m} \tag{3.4}$$

for all $m \geqslant M_0 := \max\{m_0(\boldsymbol{k}(1)), \dots, m_0(\boldsymbol{k}(L))\}$, and $C_1(K)$ and $C_2(K)$ (and also $C_3(K)$ later) are positive constants depending on K. Note that $C_1(K)$ and $C_2(K)$ here are the same for all ℓ.

We now choose H_0 such that

$$2C_2(K)H_0 \geqslant b^{\nu(1)d^{M_0}},$$

and fix the pair (ℓ, m) from the sequence $(1, M_0), \dots, (L, M_0), (1, M_0+1), \dots, (L, M_0+1), (1, M_0+2), \dots$ to be the first one satisfying

$$2C_2(K)H < b^{\nu(\ell)d^m},$$

where $H = \max\{h, H_0\}$. Then $(\ell, m) \neq (1, M_0)$, and the pair just before it is $(\ell-1, m)$ if $\ell > 1$ or $(L, m-1)$ if $\ell = 1$. The above choice means that

$$2C_2(K)H \geqslant \begin{cases} b^{\nu(\ell-1)d^m}, & \ell > 1, \\ b^{\nu(L)d^{m-1}}, & \ell = 1. \end{cases}$$

https://doi.org/10.1017/S0308210518000148 Published online by Cambridge University Press

In the first case, by (3.4),

$$\tfrac{1}{2} < C_1(K)|\Lambda|b^{\theta(\ell)d^m} = C_1(K)|\Lambda|(b^{\nu(\ell-1)d^m})^{\theta(\ell)/\nu(\ell-1)} \leqslant C_3(K)|\Lambda|H^{\mu}.$$

In the $\ell = 1$ case, we similarly have

$$\tfrac{1}{2} < C_1(K)|\Lambda|b^{\theta(1)d^m} = C_1(K)|\Lambda|(b^{\nu(L)d^{m-1}})^{d\theta(1)/\nu(L)} \leqslant C_3(K)|\Lambda|H^{\mu}.$$

□

4. General theorem

We now assume that $F(z), G(z) \in \mathbb{Q}[\![z]\!]$ converge in some disc D_r and satisfy (2.1) and (2.2). Our aim is to apply lemma 3.1 to consider the function values $F(a/b)$ and $G(a/b)$ at non-zero rational points $a/b \in D_r$, where $\log|a| = \lambda \log b$, $0 \leqslant \lambda < \log(rb)/\log b$. We also assume that

$$\left(P_{11}\left(\left(\frac{a}{b}\right)^{d^j}\right)P_{22}\left(\left(\frac{a}{b}\right)^{d^j}\right) - P_{12}\left(\left(\frac{a}{b}\right)^{d^j}\right)P_{21}\left(\left(\frac{a}{b}\right)^{d^j}\right)\right)P\left(\left(\frac{a}{b}\right)^{d^j}\right) \neq 0, \quad j = 0, 1, \ldots. \tag{4.1}$$

The approximation forms we use are obtained from (2.4) at $z = a/b$. The recursions (2.5) imply, for all $m \geqslant 1$, Journal style is to use centred dots only for “placeholders” and scalar products. I have assumed that the centred dot denoted simple multiplication of scalars and deleted it here– OK?

$$\deg A_{k,m}(z), \deg B_{k,m}(z), \deg C_{k,m}(z) \leqslant \left(\bar{e}(k) + \frac{\tau}{d-1}\right)d^m - \frac{\tau}{d-1}, \tag{4.2}$$

where $\bar{e}(k)$ and τ are non-negative integers satisfying $\bar{e}(k) \leqslant \bar{d}(k) := \max\{d_1(k), d_2(k), d_3(k)\}$ and $\tau \leqslant \nu$, the maxima of the degrees of $P_{ij}(z)$ and $P(z)$. Thus, the multiplication of (2.4) at $z = a/b$ by

$$Q_{k,m} := b^{(\bar{e}(k)+\tau/(d-1))d^m - \tau/(d-1)}$$

leads to the linear forms

$$a_{k,m}F\left(\frac{a}{b}\right) + b_{k,m}G\left(\frac{a}{b}\right) + c_{k,m} = r_{k,m}, \quad m = 0, 1, \ldots,$$

where all $a_{k,m}$, $b_{k,m}$ and $c_{k,m}$ are integers. To be able to apply lemma 3.1 with $\gamma_1 = F(a/b)$ and $\gamma_2 = G(a/b)$, we need to estimate the coefficients $a_{k,m}$ and $b_{k,m}$ and the remainders $r_{k,m}$. For this we apply the recursions (2.5).

Let $\tilde{P}(z)$ denote the polynomial, where the coefficient of z^j is the maximum of the absolute values of the corresponding coefficients in $P_{ij}(z)$, $1 \leqslant i, j \leqslant 2$. Then, for all $m = 1, 2, \ldots$,

$$|A_{k,m}(z)| \leqslant \tilde{P}(|z|)(|A_{k,m-1}(z^d)| + \delta|B_{k,m-1}(z^d)|),$$
$$|B_{k,m}(z)| \leqslant \tilde{P}(|z|)(\delta|A_{k,m-1}(z^d)| + |B_{k,m-1}(z^d)|),$$

https://doi.org/10.1017/S0308210518000148 Published online by Cambridge University Press

where $\delta = 0$ for degree 1 functions $F(z)$ and $G(z)$, and $\delta = 1$ otherwise. Applying these inequalities, we obtain

$$\max\{|A_{k,m}(z)|, |B_{k,m}(z)|\} \leqslant (1+\delta)^m \max\{|A_k(z^{d^m})|, |B_k(z^{d^m})|\} \prod_{j=0}^{m-1} \tilde{P}(|z|^{d^j}).$$

Therefore, for all $m \geqslant m_1(k)$,

$$\max\{|a_{k,m}|, |b_{k,m}|\} \leqslant c_3(k) b^{(\bar{e}(k)+\tau/(d-1))d^m}$$

if the condition

$$(1+\delta)|\tilde{P}(0)| \leqslant 1 \tag{4.3}$$

holds. Generally, for any given $\delta_1 > 0$, there exists an $m_2(k, \delta_1) > m_1(k)$ such that

$$(1+\delta)^m \prod_{j=0}^{m-1} \tilde{P}(|z|^{d^j}) < \Big((1+\delta) \max_{z \in D_r} \tilde{P}(|z|)\Big)^m < (b^{\delta_1})^{d^m}$$

for all $m \geqslant m_2(k, \delta_1)$. So, for any given $\delta_1 > 0$,

$$\max\{|a_{k,m}|, |b_{k,m}|\} \leqslant c_3(k) b^{(\bar{e}(k)+\tau/(d-1)+\delta_1)d^m} \tag{4.4}$$

for all $m \geqslant m_2(k, \delta_1)$, and under condition (4.3) we may choose $\delta_1 = 0$ here.

Since

$$R_{k,m}(z) = R_k(z^{d^m}) \prod_{j=0}^{m-1} P(z^{d^j}),$$

we also have

$$|r_{k,m}| \leqslant c_4(k) \max\{1, |P(0)|^m\} b^{-((1-\lambda)o(k)-\bar{e}(k)-\tau/(d-1))d^m}$$

for all $m \geqslant m_3(k)$. Thus, for any given $\delta_2 > 0$,

$$|r_{k,m}| \leqslant c_4(k) b^{-((1-\lambda)o(k)-\bar{e}(k)-\tau/(d-1)-\delta_2)d^m} \tag{4.5}$$

for all $m \geqslant m_4(k, \delta_2)$, and we may use the value $\delta_2 = 0$ here if the condition

$$|P(0)| \leqslant 1 \tag{4.6}$$

holds.

Thus, we obtain the estimates (3.1) and (3.2) for all $m \geqslant m_5(k, \delta_1, \delta_2)$, where

$$E(k) = \bar{e}(k) + \frac{\tau}{d-1} + \delta_1, \qquad V(k) = (1-\lambda)o(k) - \bar{e}(k) - \frac{\tau}{d-1} - \delta_2. \tag{4.7}$$

By using these values with lemma 3.1 we get the following theorem (we need only to note that the condition $D(\boldsymbol{k}, z) \neq 0$ implies $D(\boldsymbol{k}, (a/b)^{d^m}) \neq 0$ for all $m \geqslant m_6(\boldsymbol{k}, a/b)$).

Theorem 4.1. *Assume that condition (4.1) holds and that $D(\boldsymbol{k}, z) \neq 0$ for all $\ell = 1, \dots, L$. Let $\theta(\ell)$ and $\nu(\ell)$ be defined as in lemma 3.1, with $E(k)$ and $V(k)$ given in (4.7). If $0 < \nu(1) < \cdots < \nu(L) < d\nu(1)$, then there exist positive constants*

$\lambda_0 = \lambda_0(K, F, G)$, $C = C(K, a/b, F, G)$ and $H_0 = H_0(K, a/b, F, G)$ such that, for all $0 \leqslant \lambda < \lambda_0$ and any integers h_0, h_1, h_2 not all zero,

$$\left| h_0 + h_1 F\left(\frac{a}{b}\right) + h_2 G\left(\frac{a}{b}\right) \right| > CH^{-\mu},$$

with H and μ as in lemma 3.1.

5. Proof of theorems 1.1 and 1.3–1.6

We are ready to prove theorems 1.1 and 1.3–1.6. We start by giving the following formulae, which follow from (4.7):

$$\left.\begin{aligned} \theta(\ell) &= \max_{1\leqslant i<j\leqslant 3}\{\bar{e}(k_{\ell,i}) + \bar{e}(k_{\ell,j})\} + \frac{2\tau}{d-1} + 2\delta_1, \\ \nu(\ell) &= \min_{\substack{1\leqslant i,j\leqslant 3 \\ i\neq j}}\{(1-\lambda)o(k_{\ell,i}) - \bar{e}(k_{\ell,i}) - \bar{e}(k_{\ell,j})\} - \frac{2\tau}{d-1} - \delta_1 - \delta_2. \end{aligned}\right\} \tag{5.1}$$

Thus, we should choose τ and δ_i as small as possible while applying lemma 3.1.

Proof of theorem 1.1. To prove theorem 1.1, we apply theorem 4.1 with $F(z) = T(z)$ and $G(z) = T^2(z)$. Now, by (1.3),

$$(1-z)^2 F(z^2) = (1-z)F(z), \qquad (1-z)^2 G(z^2) = G(z).$$

Therefore, $r = 1$, $\delta = 0$, $P(z) = (1-z)^2$, $\tilde{P}(z) = 1 + z$ and $\tilde{P}(0) = P(0) = 1$ give $\delta_1 = \delta_2 = 0$. We shall use $(k, k+1, k-1)$ approximations and we may take $\bar{e}(k) = k+1$ and $\tau = 0$. Our $\boldsymbol{k}(\ell)$ are $(k_{\ell,1}, k_{\ell,1}+1, k_{\ell,1}+2)$ and the choices for $k = k_{\ell,1}$ are 29, 31, 34, 43 and 49. For all these values, $o(k) = 3k+2$. Since $\deg \Delta(\boldsymbol{k}(\ell), z) \leqslant 3k+3$, we have $D(\boldsymbol{k}(\ell), z) = s_{\ell,0} + s_{\ell,1}z$, where

$$s_{\ell,0} = \det\begin{pmatrix} A_{k+1}(0) & B_{k+1}(0) \\ A_{k+2}(0) & B_{k+2}(0) \end{pmatrix} c \neq 0$$

and c is the coefficient of z^{3k+2} in $R_k(z)$ (see [15, appendix]). In fact $s_{\ell,0}$ is non-zero in all of our cases, including the proofs of theorems 1.3–1.6. By using (5.1) we get

$$\theta(\ell) = 2k+5, \quad \nu(\ell) = k - 2 - \lambda(3k+2)$$

for all $\lambda < \frac{2}{3}$. So we have table 1.

For the condition $0 < \nu(1) < \cdots < \nu(5) < 2\nu(1)$ we need to assume $\lambda < \lambda_0 := \frac{7}{29} \approx 0.241\ldots$. When $\lambda < \lambda_0$, the comparison of $\mu(\ell)$ gives

$$\mu = \max_{\ell} \frac{\theta(\ell+1)}{\nu(\ell)} = \frac{\theta(4)}{\nu(3)} = \frac{91}{32 - 104\lambda}.$$

□

To prove theorems 1.3–1.6, we need to modify our choice of parameters.

https://doi.org/10.1017/S0308210518000148 Published online by Cambridge University Press

Table 1. *Selected values of k for theorem 1.1.*

ℓ	1	2	3	4	5
k	29	31	34	43	49
$\theta(\ell)$	63	67	73	91	103
$\nu(\ell)$	$27-89\lambda$	$29-95\lambda$	$32-104\lambda$	$41-131\lambda$	$47-149\lambda$

Table 2. *Selected values of k for theorem 1.3.*

ℓ	1	2	3	4	5	6
k	29	31	34	38	43	49
$\theta(\ell)$	65	69	75	83	93	105
$\nu(\ell)$	$25-89\lambda$	$27-95\lambda$	$30-104\lambda$	$34-116\lambda$	$39-131\lambda$	$45-149\lambda$

Proof of theorem 1.3. Here we apply theorem 4.1 with $F(z) = A(z)$, $G(z) = B(z)$, and the use of (1.4) gives $r = 1$, $\delta = 0$, $P(z) = 1+z+z^2$, $\tilde{P}(z) = 1$ and $\delta_1 = \delta_2 = 0$. The $(k, k+1, k-1)$ approximations give $\bar{e}(k) = k+1$ and $\tau = 1$. By choosing $\boldsymbol{k}(l)$ as above, where $k = k_{\ell,1}$ are 29, 31, 34, 38, 43 and 49, we get $o(k) = 3k+2$ and the determinants $D(\boldsymbol{k}(\ell), z) \neq 0$ (see [15, appendix]). Furthermore,

$$\theta(\ell) = 2k+7, \qquad \nu(\ell) = k-4-\lambda(3k+2)$$

for all $\lambda < \frac{2}{3}$, and this leads to table 2.

To satisfy the condition $0 < \nu(1) < \cdots < \nu(6) < 2\nu(1)$, we need to assume $\lambda < \lambda_0 := \frac{5}{29} \approx 0.172\ldots$. After the comparison of $\mu(\ell) = \theta(\ell+1)/\nu(\ell)$ we see that

$$\mu = \max_{1 \leqslant \ell \leqslant 6} \mu(\ell) = \begin{cases} \mu(6) = \dfrac{130}{45-149\lambda} & \text{if } \lambda < \dfrac{145}{1289}, \\[2ex] \mu(1) = \dfrac{69}{25-89\lambda} & \text{if } \dfrac{145}{1289} \leqslant \lambda < \dfrac{5}{29}. \end{cases}$$

□

Remark 5.1. We note that here all determinants $D(\boldsymbol{k}(\ell), z) \neq 0$, $1 \leqslant k \leqslant 50$. In all other theorems most of these determinants equal zero.

Proof of theorem 1.4. In this case we apply theorem 4.1 with $d = 3$, $F(z) = G_3(z)$ and $G(z) = F_3(z)$. Then (1.5) implies $r = 1$, $\delta = 0$, $P(z) = 1 - z^2$, $\tilde{P}(z) = 1 + z^2$ and $\delta_1 = \delta_2 = 0$. The use of (k, k, k) approximations gives $\bar{e}(k) = k$ and $\tau = 2$. If $\boldsymbol{k}(\ell)$ is the same as above and $k = k_{\ell,1}$ are 19, 26 and 39, then $o(k) = 3k+2$ and $D(\boldsymbol{k}(\ell), z) \neq 0$ (see [15, appendix]). By (5.1), if $\lambda < \frac{2}{3}$, we get

$$\theta(\ell) = 2k+5, \qquad \nu(\ell) = k-2-\lambda(3k+2).$$

Now, we have table 3.

The condition $0 < \nu(1) < \nu(2) < \nu(3) < 3\nu(1)$ holds if $\lambda < \lambda_0 := \frac{7}{29} \approx 0.241\ldots$. Similarly to the above proofs, we now get theorem 1.4. □

https://doi.org/10.1017/S0308210518000148 Published online by Cambridge University Press

Table 3. *Selected values of k for theorem 1.4.*

ℓ	1	2	3
k	19	26	39
$\theta(\ell)$	43	57	83
$\nu(\ell)$	$17-59\lambda$	$24-80\lambda$	$37-119\lambda$

Table 4. *Selected values of k for theorem 1.5.*

ℓ	1	2	3
k	17	21	26
$\theta(\ell)$	$39+2\delta_1$	$47+2\delta_1$	$57+2\delta_1$
$\nu(\ell)$	$15-\delta_1-53\lambda$	$19-\delta_1-65\lambda$	$24-\delta_1-80\lambda$

Proof of theorem 1.5. Here we may use theorem 4.1 with

$$F(z)=R(z) \quad \text{and} \quad G(z)=R(-z).$$

By (1.6), we have

$$2zF(z^2)=zF(z)+zG(z), \qquad 2zG(z^2)=F(z)-G(z).$$

Therefore, we can choose $r=1$, $\delta=1$, $P(z)=2z$ and $\tilde{P}(z)=1+z$. Since $P(0)=0$, (4.5) holds and we may take $\delta_2=0$. We use the (k,k,k) approximations and we can take $\bar{e}(k)=k$ and $\tau=1$. We also choose $\boldsymbol{k}(\ell)=(k_{\ell,1},k_{\ell,1}+1,k_{\ell,1}+2)$, where $k=k_{\ell,1}$ are 17, 21 and 26. Then we get $o(k)=3k+2$ and the determinants $D(\boldsymbol{k}(\ell),z)\neq 0$ (see [15, appendix]). Moreover,

$$\theta(\ell)=2k+5+2\delta_1, \qquad \nu(\ell)=k-2-\delta_1-\lambda(3k+2)$$

for all $\lambda<\frac{2}{3}$. This gives table 4.

The condition $0<\nu(1)<\nu(2)<\nu(3)<2\nu(1)$ holds if $\lambda<\frac{3}{13}\approx 0.230\ldots$ and δ_1 is sufficiently small. If $\lambda<\frac{21}{187}\approx 0.112\ldots$ and δ_1 is small enough, then

$$\mu=\frac{\theta(4)}{\nu(3)}=\frac{78+4\delta_1}{24-\delta_1-80\lambda}.$$

If $\frac{21}{187}\leqslant\lambda<\frac{3}{13}$, then

$$\mu=\frac{\theta(2)}{\nu(1)}=\frac{47+2\delta_1}{15-\delta_1-53\lambda}.$$

This proves theorem 1.5, since we may choose δ_1 arbitrarily small. □

Proof of theorem 1.6. We now apply theorem 4.1 with $F(z)=S(z)$ and $G(z)=S(z^4)$. The use of (1.7) gives $d=4$, $r=1$ and

$$F(z^4)=G(z), \qquad G(z^4)=-zF(z)+(1+z+z^2)G(z).$$

Since $P(0)=1$, we may choose $\delta_2=0$ in (4.5). We shall use $(k,k-1,k)$ approximations and we may take $\bar{e}(k)=k$, $\tau=2$. Again our $\boldsymbol{k}(\ell)=(k_{\ell,1},k_{\ell,1}+1,k_{\ell,1}+2)$

Table 5. *Selected values of k for theorem 1.6.*

ℓ	1	2
k	10	26
$\theta(\ell)$	$23+\frac{2}{3}+2\delta_1$	$55+\frac{2}{3}+2\delta_1$
$\nu(\ell)$	$9-\frac{2}{3}-\delta_1-31\lambda$	$25-\frac{2}{3}-\delta_1-79\lambda$

and the choices for $k=k_{\ell,1}$ are 10 and 26. For both of these values, $o(k)=3k+1$ and the determinants $D(\boldsymbol{k}(\ell),z)\neq 0$ (see [15, appendix]). By using (5.1), if $\lambda<\frac{2}{3}$, we get

$$\theta(\ell)=2k+3+\tfrac{2}{3}+2\delta_1$$

and

$$\nu(\ell)=k-1-\tfrac{2}{3}-\delta_1-\lambda(3k+1).$$

Thus, we have table 5.

If $\lambda<\lambda_0:=\frac{1}{5}$ and $\delta_1>0$ is sufficiently small, then $0<\nu(1)<\nu(2)<4\nu(1)$. Since

$$\frac{55+\frac{2}{3}+2\delta_1}{9-\frac{2}{3}-\delta_1-31\lambda}>\frac{4(23+\frac{2}{3}+2\delta_1)}{25-\frac{2}{3}-\delta_1-79\lambda}$$

for all $0\leqslant\lambda<\lambda_0$, and $\delta_1>0$ can be arbitrarily small, theorem 1.6 follows from theorem 4.1. □

Acknowledgements

The authors are grateful to Yann Bugeaud for comments that improved the paper. The research of W.W. was supported by the Natural Science Foundation of China (Grant no. 11401188) and by the Academy of Finland's Centre of Excellence in Analysis and Dynamics Research.

References

1 B. Adamczewski. Non-converging continued fractions related to the Stern diatomic sequence. *Acta Arith.* **142** (2010), 67–78.

2 B. Adamczewski and T. Rivoal. Irrationality measures for some automatic real numbers. *Math. Proc. Camb. Phil. Soc.* **147** (2009), 659–678.

3 R. P. Brent, M. Coons and W. Zudilin. Algebraic independence of Mahler functions via radial asymptotics. *Int. Math. Res. Not.* **2016** (2016), 571–603.

4 Y. Bugeaud. On the rational approximation to the Thue–Morse–Mahler numbers. *Annales Inst. Fourier* **61** (2011), 2065–2076.

5 Y. Bugeaud, G.-N. Han, Z.-Y. Wen and J.-Y. Yao. Hankel determinants, Padé approximations, and irrationality exponents. *Int. Math. Res. Not.* **2016** (2016), 1467–1496.

6 P. Bundschuh and K. Väänänen. Algebraic independence of the generating functions of Stern's sequence and of its twist. *J. Théorie Nombres Bordeaux* **25** (2013), 43–57.

7 P. Bundschuh and K. Väänänen. Transcendence results on the generating functions of the characteristic functions of certain self-generating sets. *Acta Arith.* **162** (2014), 273–288.

8 M. Coons. On the rational approximation of the sum of the reciprocals of the Fermat numbers. *Ramanujan J.* **30** (2013), 39–65.

9 K. Dilcher and K. B. Stolarsky. Stern polynomials and double-limit continued fractions. *Acta Arith.* **140** (2009), 119–134.

10 N. I. Fel′dman and Y. V. Nesterenko. *Number theory IV: transcendental numbers* (ed. A. N. Parshin and I. R. Shafarevich). Encyclopaedia of Mathematical Sciences, vol. 44 (Springer, 1998).

11 Y.-J. Guo, Z.-X. Wen and W. Wu. On the irrationality exponent of the regular paperfolding numbers. *Linear Alg. Applic.* **446** (2014), 237–264.

12 K. Nishioka. *Mahler functions and transcendence*. Lecture Notes in Mathematics, vol. 1631 (Springer, 1996).

13 K. Väänänen. On rational approximations of certain Mahler functions with a connection to the Thue–Morse sequence. *Int. J. Number Theory* **11** (2015), 487–493.

14 K. Väänänen and W. Wu. On simultaneous approximation of the values of certain Mahler functions. Preprint, 2015. (Available at http://arxiv.org/abs/1505.00931.)

15 K. Väänänen and W. Wu. On linear independence measures of the values of Mahler functions. Preprint, 2016. (Available at http://arxiv.org/abs/1604.01630.)

16 Z.-X. Wen and W. Wu. Hankel determinants of the Cantor sequence. *Sci. Sinica Math.* **44** (2014), 1059–1072. (In Chinese.)