NEW ZEALAND JOURNAL OF MATHEMATICS
Volume 46 (2016), 83-88

ZERO ORDER ESTIMATES FOR MAHLER FUNCTIONS

Michael Coons

(Received 2 June, 2016)

Abstract. We give an upper bound for the zero order of the difference between a Mahler function and an algebraic function. This complements estimates of Nesterenko, Nishioka, and Töpfer, among others, who considered polynomials evaluated at Mahler functions.

1. Introduction

When Kurt Mahler was quite young and sick in bed, he set out to prove that the number $\sum_{n\geqslant 0} 2^{-2^n}$ is transcendental. He did so by exploiting the functional equation

$$F(z^2) = F(z) - z,$$

which the series $\sum_{n\geqslant 0} z^{2^n}$ satisfies. In doing so, Mahler discovered an important method in the theory of Diophantine approximations with applications to transcendence and algebraic independence.

Mahler's results were generalised and extended by several authors. The most general of generalisations, which seems to capture all previous versions considered, was given by Töpfer [**15, 16**], who considered formal power series $f_1(z), \ldots, f_d(z) \in \mathbb{C}[[z]]$ that satisfy functional equations

$$A_0(z, f_1(z), \ldots, f_d(z)) \cdot f_i(T(z)) = A_i(z, f_1(z), \ldots, f_d(z)) \quad (1 \leqslant i \leqslant d),$$

where $T(z) \in \mathbb{C}(z)$ and $A_i(z, y_1, \ldots, y_d) \in \mathbb{C}[z, y_1, \ldots, y_d]$ for $i = 0, \ldots, d$. For this version, Töpfer produced both a zero order estimate [**16**] for $Q(z, f_1(z), \ldots, f_d(z))$ with $Q(z, y_1, \ldots, y_d) \in \mathbb{C}[z, y_1, \ldots, y_d]$ as well as algebraic independence results for some special cases of his generalisation [**15**].

Of all of the generalisations, two stand out and are arguably the most important; they are also the simplest. The first was given by Mahler himself [**11**], who considered[1] functions $f(z) \in \mathbb{C}[[z]]$ satisfying

$$f(z^k) = R(z, f(z)),$$

for an integer $k \geqslant 2$ and a rational function $R(z, y) \in \mathbb{C}(z, y)$. The second is harder to attribute, but goes back at least to the 1960s or 1970s. In this case, one considers a function $f(z)$ for which there are integers $k \geqslant 2$ and $d \geqslant 1$ such that

$$a_0(z)f(z) + a_1(z)f(z^k) + \cdots + a_d(z)f(z^{k^d}) = 0, \qquad (1)$$

2010 *Mathematics Subject Classification* Primary 11J82; Secondary 11J25.
Key words and phrases: Mahler functions, rational approximation, algebraic approximation.
The research of M. Coons was supported by ARC grant DE140100223.

[1]Actually, Mahler only considered the base field $\overline{\mathbb{Q}}$ and not all of $\mathbb{C}$.

for some polynomials $a_0(z), \ldots, a_d(z) \in \mathbb{C}[z]$. These two generalisations coincide when $d = 1$, and for this value of d the strongest results have been shown. While there are few natural examples of the other classes, functions satisfying (1) are readily available and certain cases are of particular importance in theoretical computer science; the generating functions of automatic and regular sequences satisfy (1). See the works of Allouche and Shallit [**1, 2, 3**], Christol, Kamae, Mendès France, and Rauzy [**8**], Dekking, Mendès France, and van der Poorten [**9**], Loxton [**10**], and Becker [**5**] for further details and specific examples.

In this paper, we are concerned with the algebraic approximation of functions satisfying (1). We call a function satisfying (1) a *k-Mahler function*, or just a *Mahler function* when k is clear. The minimal such d for which (1) holds for $f(z)$ is called the *degree* of the Mahler function $f(z)$, denoted d_f, and we define the *height* of the Mahler function $f(z)$ by $A_f := \max\{\deg a_i(z) : i = 0, \ldots, d_f\}$.

Our main result is a zero order estimate for the difference of a Mahler function with an algebraic function. To this end, let $\nu : \mathbb{C}((z)) \to \mathbb{Z} \cup \{\infty\}$ be the valuation defined by $\nu(0) := \infty$ and

$$\nu\left(\sum c_n z^n\right) := \min\{i : c_i \neq 0\}$$

when $\sum_n c_n z^n$ is nonzero. Also, for $G(z)$ an algebraic function with minimal polynomial $P(z,y) \in \mathbb{C}[z,y]$, we call $\deg_y P(z,y)$ the *degree* of $G(z)$ and we call $\exp(\deg_z P(z,y))$ the height of $G(z)$.

Theorem 1. *If $F(z)$ is an irrational k-Mahler function of degree d_F and height A_F, and $G(z)$ is an algebraic function of degree at most n and height at most H_G, then*

$$\nu\big(F(z) - G(z)\big) \leqslant (d_F + 1) \cdot A_F \cdot n^{d_F+1} + \frac{k^{d_F+1} - 1}{k - 1} \cdot \log H_G \cdot n^{d_F}.$$

Previous results on zero estimates of Mahler functions focussed on upper bounds for $\nu(Q(z, F(z)))$ for polynomials $Q(z,y) \in \mathbb{C}[z,y]$ and used quite deep methods, relying on Nesterenko's elimination-theoretic method [**12, 13**]; see also Becker [**4**], Nishioka [**14**], and Töpfer [**16**]. While the estimate provided by Theorem 1 is essentially of the same order as the best bounds for $\nu(Q(z, F(z)))$, our proof is much simpler—it avoids the use of Nesterenko's and Nishioka's methods—and is by all means, elementary.

2. Algebraic Approximation of Mahler Functions

In recent work with Jason Bell [**6**], we proved the following result.

Lemma 1 (Bell and Coons)**.** *Let $F(z)$ be an irrational k-Mahler function of degree d_F and height A_F, and let $P(z)/Q(z)$ be any rational function with $Q(0) \neq 0$. Then*

$$\nu\left(F(z) - \frac{P(z)}{Q(z)}\right) \leqslant A_F + \frac{k^{d_F+1} - 1}{k - 1} \cdot \max\{\deg P(z), \deg Q(z)\}.$$

Theorem 1 is the generalisation of this result to approximation by algebraic functions. To prove this generalisation, we use a resultant argument.

Lemma 2. *Let $f(z)$ and $g(z)$ be two algebraic functions of degrees at least 2 satisfying polynomials of degrees Δ_f and Δ_g with coefficients of degree at most δ_f*

and δ_g, respectively. Then the algebraic function $f(z)+g(z)$ satisfies a polynomial of degree

$$\Delta_{f+g} \leqslant \Delta_f \Delta_g$$

with coefficients of degree

$$\delta_{f+g} \leqslant \delta_f \Delta_g + \delta_g \Delta_f.$$

Proof. This result follows by using the Sylvester matrix to calculate a certain resultant. For R a ring and $P, Q \in R[y]$ with

$$P(y) = \sum_{i=0}^{\deg_y P} p_i y^i \quad \text{and} \quad Q(y) = \sum_{i=0}^{\deg_y Q} q_i y^i,$$

the resultant of P and Q with respect to the variable y is denoted by $\operatorname{res}_y(P,Q)$ and may be calculated as the determinant of the $(\deg_y Q + \deg_y P) \times (\deg_y Q + \deg_y P)$ Sylvester matrix; that is

$$\operatorname{res}_y(P,Q) := \det \begin{pmatrix} p_0 & p_1 & p_2 & \cdots & p_{\deg_y P} & & & & \\ & p_0 & p_1 & p_2 & \cdots & p_{\deg_y P} & & & \\ & & \ddots & \ddots & \ddots & & \ddots & & \\ & & & p_0 & p_1 & p_2 & \cdots & p_{\deg_y P} \\ q_0 & q_1 & q_2 & \cdots & q_{\deg_y Q} & & & & \\ & q_0 & q_1 & q_2 & \cdots & q_{\deg_y Q} & & & \\ & & \ddots & \ddots & \ddots & & \ddots & & \\ & & & q_0 & q_1 & q_2 & \cdots & q_{\deg_y Q} \end{pmatrix},$$

where there are $\deg_y Q$ rows of the coefficients of P and $\deg_y P$ rows of the coefficients of Q. Now suppose $R = \mathbb{C}[z,x]$, so that the entries of the above Sylvester matrix are polynomials in the variables z and x, and set $D(x,z) := \operatorname{res}_y(P,Q)$. Since polynomial degrees are additive, using the Leibniz formula for the determinant, we have immediately that

$$\deg_z D(x,z) \leqslant \deg_y Q \deg_z P + \deg_y P \deg_z Q \tag{2}$$

and

$$\deg_x D(x,z) \leqslant \deg_y Q \deg_x P + \deg_y P \deg_x Q. \tag{3}$$

The lemma now follows immediately by combining (2) and (3) with the fact that given algebraic functions $f(z), g(z) \in \mathbb{C}[[z]]$ and polynomials $P_f(z,y), P_g(z,y) \in \mathbb{C}[z,y]$ with $P_f(z,f) = P_g(z,g) = 0$, the algebraic function $f(z)+g(z)$ is a root of the polynomial $\operatorname{res}_y(P_f(z,y), P_g(z,x-y))$ viewed as a polynomial in x. □

Because of Lemma 1, we may focus on algebraic functions of degree at least 2.

Lemma 3. *Let $a_0(z), \ldots, a_d(z)$ be polynomials of degree at most A. If $G(z) \in \mathbb{C}[[z]]$ is an algebraic function of degree $\Delta_G \geqslant 2$ satisfying a minimal polynomial with coefficients of degree at most δ_g, then the function*

$$M_G(z) := \sum_{i=0}^{d} a_i(z) G(z^{k^i})$$

is an algebraic function satisfying a polynomial of degree

$$\Delta_{M_G} \leqslant \Delta_G^{d+1}$$

whose coefficients have degree

$$\delta_{M_G} \leqslant (d+1)A \cdot \Delta_G^{d+1} + \frac{k^{d+1}-1}{k-1} \cdot \delta_G \cdot \Delta_G^d.$$

Proof. Since $G(z)$ is an algebraic function, so is $\sum_{i=0}^d a_i(z)G(z^{k^i})$. One can easily gain information about the sum using the theory of resultants.

To get an upper bound on $\nu(M_G(z))$, we apply the idea of the previous paragraph by including the terms $G_i(z) := a_i(z)G(z^{k^i})$ one at a time. To do this, let

$$P_G(z,y) := g_{\Delta_G} y^{\Delta_G} + \cdots + g_1 y + g_0$$

be the minimal polynomial of $G(z)$. Here we have denoted the degree of $G(z)$ by Δ_G. Set $\delta_G := \deg_z P_G(z,y)$. Then

$$P_{G_i}(z,y) = a_i(z)^{\Delta_G} P_G(z^{k^i}, y/a_i(z))$$

is a polynomial with $P_{G_i}(z, G_i(z)) = 0$, where, of course, we only form this polynomial when $a_i(z) \neq 0$. Here, we have that $P_{G_i}(z,y)$ is still minimal with respect to the degree of y, but there is no guarantee that it is minimal with respect to the degree of z for this degree of y. However, we do have that the minimal polynomial of $G_i(z)$ divides $P_{G_i}(z,y)$ and the remainder is just a polynomial in z. In any case, the above gives that

$$\Delta_{G_i} := \deg_y P_{G_i}(z,y) = \deg_y P_G(z,y) = \Delta_G \tag{4}$$

and

$$\delta_{G_i} := \deg_z P_{G_i}(z,y) \leqslant A\Delta_G + k^i \delta_G. \tag{5}$$

The lemma now follows by combining (4) and (5) with Lemma 2. □

Lemma 4. *Let $G(z) \in \mathbb{C}[[z]]$ be an algebraic function of degree at least 2 satisfying the polynomial $P_G(z,y) = a_n(z)y^n + a_{n-1}(z)y^{n-1} + \cdots + a_1(z)y + a_0(z)$, with $a_0(z) \neq 0$. Then $\nu(G(z)) \leqslant \nu(a_0(z))$. In particular, $\nu(G(z)) \leqslant \deg_z P_G(z,y)$.*

Proof. Since $P_G(z,y)$ is a minimal polynomial, we have $a_0(z) \neq 0$. We thus have, identically,

$$\left(a_n(z)G(z)^{n-1} + a_{n-1}(z)G(z)^{n-2} + \cdots + a_1(z)\right) G(z) = -a_0(z).$$

The fact $G(z), a_n(z), \ldots, a_0(z) \in \mathbb{C}[[z]]$ then gives

$$\nu\left(a_n(z)G(z)^{n-1} + a_{n-1}(z)G(z)^{n-2} + \cdots + a_1(z)\right) + \nu(G(z)) = \nu(a_0(z)),$$

which proves the lemma, since each of the terms is a nonnegative integer. □

Proof of Theorem 1. Let $F(z)$ be a k-Mahler function satisfying (1) of degree d_F and height A_F and let $G(z)$ be an algebraic function of degree at most n and height at most H_G. Since by Lemma 1, the theorem holds for $n = 1$, we may assume without loss of generality that $n \geqslant 2$.

Set $M := \nu(F(z) - G(z))$, and write

$$F(z) - G(z) = z^M T(z),$$

where $T(z) \in \mathbb{C}[[z]]$ with $T(0) \neq 0$. Then also

$$\sum_{i=0}^{d} a_i(z)F(z^{k^i}) - \sum_{i=0}^{d} a_i(z)G(z^{k^i}) = \sum_{i=0}^{d} a_i(z)z^{k^iM}T(z^{k^i}),$$

which since $F(z)$ satisfies (1) reduces to

$$M_G(z) := \sum_{i=0}^{d} a_i(z)G(z^{k^i}) = -\sum_{i=0}^{d} a_i(z)z^{k^iM}T(z^{k^i}).$$

This immediately implies that

$$\nu(F(z) - G(z)) = M \leqslant \nu\left(M_G(z)\right) \leqslant \delta_{M_G},$$

where the last inequality follows from Lemma 4. By definition, $\delta_G = \log H_G$, hence applying Lemma 3 proves the theorem. □

3. Concluding Remark

The n-dependence in the estimate of Theorem 1 is the best that can be attained by this method, that is, a bound of n-order n^{d_F}; this is the same n-order for the best known bounds on $\nu(Q(z, F(z)))$ as well [**16**]. While at first glance, the n-dependence in Theorem 1 looks like n^{d_F+1}, when using the results one usually first takes a limit through the height H_G. With this in mind, one assumes that $\log H_G \geqslant n \geqslant 1$ so that our estimate gives

$$\nu\left(F(z) - G(z)\right) \leqslant \left((d_F + 1) \cdot A_f + \frac{k^{d_F+1} - 1}{k - 1}\right) \cdot \log H_G \cdot n^{d_F}.$$

The immediate question is whether or not this is the best bound possible; probably the answer is 'no.' Presumably a 'Roth-type' estimate holds, so that one has a bound that is linear in n. This would imply that a Mahler function $F(z)$ is an S-number in the suitable function-field analogue of Mahler's classification (see Bugeaud [**7**] for the relevant definitions), which is a question that has been circulating within the area for some time now.

References

[1] Jean-Paul Allouche and Jeffrey Shallit, *The ring of k-regular sequences*, Theoret. Comput. Sci., **98** (2) (1992), pp. 163–197.
[2] Jean-Paul Allouche and Jeffrey Shallit, *Automatic Sequences*, Cambridge University Press, Cambridge, 2003.
[3] Jean-Paul Allouche and Jeffrey Shallit, *The ring of k-regular sequences. II*, Theoret. Comput. Sci., **307** (1) (2003), pp. 3-29.
[4] Paul-Georg Becker, *Effective measures for algebraic independence of the values of Mahler type functions*, Acta Arith., **58** (3) (1991) pp. 239-250.
[5] Paul-Georg Becker, *k-regular power series and Mahler-type functional equations*, J. Number Theory, **49** (3) (1994), pp. 269-286.
[6] Jason P. Bell and Michael Coons, *Transcendence tests for Mahler functions*, Proc. Amer. Math. Soc., to appear.
[7] Yann Bugeaud, *Approximation by Algebraic Numbers*, Cambridge Tracts in Mathematics, vol. 160, Cambridge University Press, Cambridge, 2004.

[8] G. Christol, T. Kamae, M. Mendès France, and G. Rauzy, *Suites algébriques, automates et substitutions*, Bull. Soc. Math. France, **108** (4) (1980), pp. 401-419.

[9] Michel Dekking, Michel Mendès France, and Alf van der Poorten, *Folds*, Math. Intelligencer, **4** (3) (1982), pp. 130-138.

[10] J. H. Loxton, *Automata and transcendence*, New Advances in Transcendence Theory (Durham, 1986), Cambridge Univ. Press, Cambridge, 1988, pp. 215–228.

[11] K. Mahler, *Arithmetische eigenschaften der lösungen einer klasse von funktionalgleichungen*, Math. Ann., **101** (1) (1929), pp. 342-366.

[12] Ju. V. Nesterenko, *Estimate of the orders of the zeroes of functions of a certain class, and their application in the theory of transcendental numbers*, Izv. Akad. Nauk SSSR Ser. Mat., **41** (2) (1977), pp. 253-284, 477.

[13] Yu. V. Nesterenko, *Algebraic independence of algebraic powers of algebraic numbers*, Mat. Sb. (N.S.), **123(165)** (4) (1984), pp. 435-459.

[14] Kumiko Nishioka, *Algebraic independence measures of the values of Mahler functions*, J. Reine Angew. Math., **420** (1991), pp. 203-214.

[15] Thomas Töpfer, *Algebraic independence of the values of generalized Mahler functions*, Acta Arith., **70** (2) (1995), pp. 161-181.

[16] Thomas Töpfer, *Zero order estimates for functions satisfying generalized functional equations of Mahler type*, Acta Arith., **85** (1) (1998), pp. 1-12.

Michael Coons
School of Math. and Phys. Sciences
University of Newcastle
Callaghan
Australia
Michael.Coons@newcastle.edu.au