JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 77B, Nos. 1 & 2, January–June 1973

Extrapolation Techniques Related to Transcendence Proofs*

Kenneth B. Stolarsky**

(November 16, 1972)

An extrapolation method used by A. Baker to study linear forms in the logarithms of algebraic numbers is further refined, and used to study a general extrapolation problem involving a function holomorphic in a large disk.

Key words: Baker's extrapolation method; extrapolation; Hermite interpolation formula; transcendental numbers; transcendence proofs.

1. Introduction

Say that $f \in \mathfrak{F}$, a set of very smooth functions which are "discrete" (e.g., assume integral values at the integers) and "flat" (e.g., $f^{(i)}(0)=0$, $0 \leq i \leq I$ and $|f'(x)|$ grows slowly). Then it is possible that $f(j)=0$ for $0 \leq j \leq \phi(I)$, j an integer, where $\phi(x) \to \infty$ as $x \to \infty$. For example, say $f \in \mathfrak{F}$ implies $f(j)$ is integral for integral j, $f(0)=0$, and $|f'(x)|<1$. Then $f(j)=0$ for every integral j. Many proofs that specific numbers are transcendental involve showing that a certain discrete flat entire function of exponential type has a large number of consecutive integral zeros ([2], pp. 102–106; [1],[1] pp. 211–215; note the equation

$$\Psi(l)=0, \qquad 1 \leq l \leq (L+1)^n$$

on p. 215). This paper shows a general such result.

References in this paragraph are to [1]; however our notation is different. As in [1] our proof uses an iterated extrapolation procedure based on a Hermite type formula like (15), p. 212. As in Lemma 4, p. 211, the general step in the procedure is to show that if $f^{(i)}(j)$ is small for $0 \leq i \leq I$, $0 \leq j \leq J$ then it is also small for $0 \leq i \leq I'$, $0 \leq j \leq J'$ where $I'<I$ but $J<J'$. In Baker's paper $I'=I/2$ so I decreases *exponentially* with u, the number of iterations. In our paper I decreases only *linearly* with u, so the extrapolation can be carried out much further. For a similar technique see [3].

Our theorem is stated in section 2; (ii) and (iii) correspond to flatness and discreteness respectively. The proof is given in section 3. In section 4 our extrapolation procedure is explained independently of the theorem, and shown to be optimal in some sense.

2. The Theorem

Let $\Psi = \langle A, I, J, \sigma, \tau; g, r \rangle$ be a set of positive real numbers with $A, I, J, \sigma, \tau \geq 1$ and $1+I^{-1}\sigma \leq r$. Let i and j denote integers.

AMS Subject Classification: Primary 10 F 35, Secondary 30 A 80.
*An invited paper. Partially supported by NSF Contract GP 12363.
**Present address: Mathematics Department, University of Illinois, Urbana, Illinois 61801.
[1] Figures in brackets indicate the literature references at the end of this paper.

DEFINITION: Let $\mathfrak{F}=\mathfrak{F}(\Psi)$ be the set of all functions $\varphi(z)$ holomorphic in $|z|\leq r$ such that

$$\text{(i)}\ |\varphi^{(i)}(z)|\leq \exp\{A\tau(|z|+1)\} \qquad \text{for } 0\leq i\leq I,$$
$$\text{(ii)}\ |\varphi^{(i)}(j)|\leq e^{-\sigma} \text{ for } 0\leq i\leq I, 0\leq j\leq J-1,$$

and

$$\text{(iii) if } |\varphi^{(i)}(j)|\leq e^{-\tau} \text{ then } |\varphi^{(i)}(j)|\leq e^{-\sigma} \text{ provided } 0\leq i\leq I \text{ and } 0\leq j\leq g.$$

THEOREM: *If* $\varphi\epsilon\mathfrak{F}$ *then* $|\varphi(\mathrm{j})|\leq \mathrm{e}^{-\sigma}$ *for*

$$0\leq \mathrm{j}\leq \mathrm{J}\, min\left\{\frac{1}{5}exp\ (\mathrm{I}/2000\mathrm{A}\tau),\ \frac{1}{5}(\sigma\tau/10\mathrm{JI}^2)^{1/3},\ \mathrm{g/J},\ \mathrm{r/J}\right\}. \tag{2.1}$$

The theorem is trivial unless

$$I>2000\,A\tau \tag{2.2}$$

and

$$\sigma>10\tau, \tag{2.3}$$

so we may assume (2.2) and (2.3) in our proof. The function

$$\varphi(z)=\exp\,(-\sigma-1+J^{-1}z)$$

shows that $\mathfrak{F}(\Psi)$ is nonempty when

$$\Psi=\{A, I, J, \sigma, \tau; J, \infty\}.$$

Two special cases of interest are

$$\Psi=\{1, \tau^2, 1, e^{\tau}, \tau; e^{\tau}, \infty\} \tag{2.4}$$

and

$$\Psi=\{1, \tau^{1+\epsilon}, 1, \tau^{1+3\epsilon}, \tau; \tau^{\epsilon}, \infty\} \tag{2.5}$$

where $\epsilon>0$. For (2.4) the inequality (ii) extends to very large values of j when $i=0$, while for (2.5) a nontrivial conclusion is based on a very weak discreteness hypothesis (iii). It is clear, of course, that we need hypothesis (iii); if it were dropped then

$$\varphi(z)=\exp\,\{-e^{\tau}+\tau(\tau+1)^{-1}z\}$$

would satisfy the hypothesis but not the conclusion of the theorem in the case (2.4).

By setting $J=\tau=1$ and $\sigma=g=r=\infty$ in the theorem, we obtain the following

COROLLARY: *Let* A, I ≥ 1 *and let* i *and* j *denote integers. Say* $\varphi(\mathrm{z})$ *is an entire function such that for* $0\leq \mathrm{i}\leq \mathrm{I}$,

$$\text{(i)}\ |\varphi^{(\mathrm{i})}(\mathrm{z})|\leq exp\,\{\mathrm{A}(|\mathrm{z}|+1)\}$$

$$\text{(ii)}\ \varphi^{(\mathrm{i})}(0)=0$$

and

$$\text{(iii) } if\ |\varphi^{(i)}(j)| < \frac{1}{4}\ then\ \varphi^{(i)}(j) = 0.$$

Then $\varphi(j) = 0$ for $0 \leqslant j \leqslant \frac{1}{5} exp\ (I/2000A)$.

A peculiarity of our theorem is that "beyond some point" knowing that I is extremely large does not strengthen the conclusion. In applications it might be advisable to replace a given I by a smaller one (however, see the comment regarding λ at the beginning of sec. 4).

3. The Proof

Let $[x]$ denote the greatest integer in x, $I_1 = [I]$, $J_1 = [J]$, and let i, j, s, t, u denote integers. The first two lemmas below are stated for the sake of completeness.

LEMMA 1: *Say* $|\zeta - j| = \frac{1}{2}$, $0 \leqslant j \leqslant J-1$, *and* $0 \leqslant z \leqslant S$ *where* j, J, S *are integers and* ζ *is a complex number. Then*

$$|\Pi| \leqslant 8 \cdot 2^{S+4J}$$

where

$$\Pi = \prod_{\nu=0}^{J-1} \left(\frac{z-\nu}{\zeta-\nu}\right). \tag{3.1}$$

LEMMA 2: (The Hermite interpolation formula). *Let the simple closed contour* C *contain* z *and* $0, 1, \ldots, J-1$. *If* $\varphi(z)$ *is analytic on and inside* C, *and* z *lies outside closed disks of radius* R *about* $0, 1, \ldots, J-1$, *then*

$$\varphi^{(t)}(z) = \sum_{j=0}^{J-1} \sum_{i=0}^{I'-1} \frac{\varphi^{(i+t)}(j)}{i!} \cdot \frac{1}{2\pi i} \int_{|\zeta-j|=R} (\zeta - j)^i (z-\zeta)^{-1} \tag{3.2}$$

$$\cdot \prod_{\nu=0}^{J-1} \left(\frac{z-\nu}{\zeta-\nu}\right)^{I'} d\zeta + \frac{1}{2\pi i} \int_C \varphi^{(t)}(\zeta)(\zeta - z)^{-1} \prod_{\nu=0}^{J-1} \left(\frac{z-\nu}{\zeta-\nu}\right)^{I'} d\zeta$$

$$= T_1 + T_2.$$

LEMMA 3: *Let* $\varphi \in \mathfrak{F}(\Psi)$. *Assume (2.2), (2.3), and part (i) of the definition. Assume also that*

$$|\varphi^{(i)}(j)| \leqslant e^{-\sigma}\ for\ 0 \leqslant i \leqslant I_u,\ 0 \leqslant j \leqslant J_u - 1 \tag{3.3}$$

where

1. $0 \leqslant I_u \leqslant I_1$, *and*

2. $1 \leqslant J_1 \leqslant J_u \leqslant \sigma\tau / 10 I_1^2$.

Then

$$\left|\varphi^{(t)}(j)\right| \leqslant e^{-2\tau} \tag{3.4}$$

provided

$$0 \leq t \leq [I_u - 650A\tau] \text{ and } 0 \leq j \leq [2.5J_u] + 1.$$

PROOF: We first show (3.4) for

$$0 \leq t \leq [\tfrac{1}{2}(I_u + s)] + 1,\ J_u \leq j \leq S_u = S_u(s) \tag{3.5}$$

where

$$0 \leq s \leq I_u - 1250A\tau \text{ and } S_u = [\tfrac{1}{2}(I_u - s)J_u/100A\tau]. \tag{3.6}$$

Then (3.4) follows by letting $s = s_0 = [I_u - 1250A\tau]$. Note that

$$S_u(s) \geq S_u(s_0) \geq 2.5J_u > 2J_u.$$

For the rest of the proof set $I' = [\tfrac{1}{2}(I_u - s)]$, $J = J_u$, $S = S_u$ and $\rho = \rho_u = 2S_u$. For $J \leq z \leq S$ and $0 \leq t \leq I'$ Lemma 2 is applicable for C a circle of radius ρ because

$$\rho = 2S \leq I_1 J/100A\tau \leq I_1\,\sigma\tau/100A\tau \cdot 10I_1^2 \leq r$$

by the definition of section 2. Thus (3.2) holds and clearly $i + t \leq I_u$ in (3.2). Hence by Lemma 1, (3.3), and $S > 2J$ (set $R = \tfrac{1}{2}$),

$$|T_1| \leq I'Je^{-\sigma} \cdot \tfrac{1}{2} \cdot 2 \cdot 8^{I'}2^{I'(S+4J)} \leq e^{-\sigma}2^{10I'J + I'S}$$

$$\leq e^{-\sigma}2^{6I'S} < e^{-(1/2)\sigma}$$

since

$$6I'S \leq 6I_1 \cdot (I_1/200A\tau) \cdot \sigma\tau/10I_1^2 < \tfrac{1}{2}\sigma \cdot$$

Next,

$$|T_2| \leq \frac{\rho}{\rho - S}\ e^{A\tau(\rho+1)} \left(\frac{S}{\rho - J}\right)^{I'J}$$

$$\leq 2 \exp\left(A\tau + 2A\tau S - I'J \ln \tfrac{3}{2}\right).$$

Since $I' \geq 600A\tau > 1$, we have

$$2A\tau S \leq 2A\tau \cdot 2I'J/100A\tau = I'J/25.$$

Therefore $(J \geq 1)$

$$|T_2| \leq \exp(2A\tau - .36I') \leq \exp(2A\tau - 216A\tau) < e^{-200\tau}.$$

Finally, by (2.3),

$$|\varphi^{(t)}(z)| \leq |T_1| + |T_2| \leq e^{-5\tau} + e^{-200\tau} \leq e^{-2\tau}.$$

In particular this holds when $z = j$, so Lemma 3 is proved.

For the proof of the theorem, set

$$I_{u+1}=[I_u-650A\tau] \text{ and } J_{u+1}=[2.5J_u]+1$$

for $u \geqslant 1$. Since $e^{-2\tau}<e^{-\tau}$, whenever the conclusion of Lemma 3 is valid we expect by (iii) that

$$|\varphi^{(i)}(j)| \leqslant e^{-\sigma} \text{ for } 0 \leqslant i \leqslant I_{u+1}, 0 \leqslant j \leqslant J_{u+1}-1. \tag{3.7}$$

Hypothesis (ii) asserts that (3.7) is valid for $u=0$. Thus (3.7) is valid for $0 \leqslant u \leqslant U$ by induction provided conditions 1. and 2. of Lemma 3 hold for $u \leqslant U$, and $J_{U+1} \leqslant \max(g, r)$ so (i) and (iii) are applicable. We choose $U=[\min(U_1, U_2)]$ where

$$U_1=I/1400A\tau \text{ and } U_2=\frac{1}{2}\ln(\sigma\tau/10JI^2).$$

The inequality $I_{u+1} \geqslant I_u-700A\tau$ implies $I_U \geqslant I_1-\frac{1}{2}I \geqslant 0$ so condition 1. follows. For 2., note that

$$2.5 \leqslant J_{u+1}/J_u \leqslant 3.5$$

so

$$(2.5)^U \leqslant J_{U+1}/J_1 \leqslant (3.5)^U \leqslant (\sigma\tau/10JI^2)^{(1/2)\ln(3.5)}.$$

Thus

$$J_{U+1} \leqslant J_1(\sigma\tau/10JI_1^2)$$

and condition 2. follows. If $U=[U_1]$ then

$$\begin{aligned} J_{U+1} &\geqslant \frac{1}{2} J \exp\{((1400A\tau)^{-1}I-1)\ln 2.5\} \\ &\geqslant \frac{1}{5} J \exp(I/2000A\tau). \end{aligned}$$

If $U=[U_2]$ then

$$\begin{aligned} J_{U+1} &\geqslant \frac{1}{2} J \exp\left\{\left(\frac{1}{2}\ln(\sigma\tau/10JI^2)-1\right)\ln 2.5\right\} \\ &\geqslant \frac{1}{5} J(\sigma\tau/10JI^2)^{1/3}. \end{aligned}$$

This proves the theorem.

4. The General Extrapolation Procedure

Proposition 4.2 (with $\lambda(x)=x$) idealizes our extrapolation procedure, and shows that the first term inside the min(,,,) of (2.1) cannot be improved in its general from (at least by these methods). Under certain conditions the second term inside the min(,,,) can be replaced by

$$(\sigma\lambda(\tau)/10I\lambda(I)J)^{1/3}$$

but we shall omit the details. If we set $t=0$ in (4.4) we get the old extrapolation procedure.

Let $\lambda(x)$ be an increasing positive function such that $\lambda'(1) > 0$ and $\lambda''(x) \leq 0$ (for example $\lambda(x) = x^{1/3}$). Let μ be the inverse function so that $\mu(\lambda(x)) = \lambda(\mu(x)) = x$.

PROPOSITION 4.1: *Let* $x \geq A > 0$ *and define*

$$\theta(x, A) = min \sum_{n=1}^{N-1} \mu(A_{n+1}/A_n) \tag{4.1}$$

where the minimum is over all integers $N \geq 2$ *and all real* N*-tuples* $(A_1, \ldots, A_N)$ *such that* $A = A_1 \leq A_2 \leq \ldots \leq A_{N-1} \leq A_N = x$. *Then*

$$C_{20}(\mu) \ln (x/A) \leq \theta(x, A) \leq C_{21}(\mu) \ln (x/A) \tag{4.2}$$

where C_{20}, C_{21} *depend only on* μ.

PROOF: Let $A_n = x^{n/N} A^{(N-n)/N}$. Then $A_{n+1}/A_n = (x/A)^{1/N}$,

so

$$\theta(x, A) \leq \min_N \{(N-1)\mu(\{x/A\}^{1/N})\}.$$

The upper bound follows by taking $N = [\ln (x/A)]$ for x large. On the other hand, by the convexity of μ, the inequality of the arithmetic-geometric means, and some simple calculus,

$$\begin{aligned} \theta(x, A) &\geq \min_N \left\{ (N-1)\mu\left(\frac{1}{N-1}\sum_{n=1}^{N-1} A_{n+1}/A_n\right)\right\} \\ &\geq \min_N \{(N-1)\mu(\{x/A\}^{1/(N-1)})\} \\ &\geq \min_N \{C_{22}(\mu)(N-1)(x/A)^{1/(N-1)}\} \geq C_{20}(\mu) \ln (x/A) \end{aligned}$$

and the result follows.

PROPOSITION 4.2: *Let* S *be the smallest subset of* $[0, \infty] \times [0, \infty]$ *such that*

$$[0, J] \times [0, I] \subseteq S \text{ for some } I, J > 0. \tag{4.3}$$

(4.4) *If* $[0, A] \times [t, B] \subseteq S$, *then*

$$[0, A\lambda(B-t)] \times [t, \tfrac{1}{2}(B+t)] \subseteq S.$$

Then

$$Je^{C_{23}(\lambda)I} \leq max \{x | (x, y) \in S\} \leq Je^{C_{24}(\lambda)I}$$

where $C_{23}(\lambda)$, $C_{24}(\lambda)$ *depend only on* λ.

PROOF: Let $x = J\lambda(I-t)$, $y = \frac{1}{2}(I+t)$. Then it is clear that $S = U_{i=0}^{\infty} S_i$ where $S_0 = [0, J] \times [0, I]$, S_1 consists of all points in the first quadrant under the curve

$$y = I - \mu(x/J),$$

S_2 of all such points under all curves

$$y = I - \tfrac{1}{2}\{\mu(A_2/J) + \mu(x/A_2)\} \qquad J \leqslant A_2 \leqslant x$$

and in general S_N of all such points under

$$y = I - \tfrac{1}{2} \sum_{n=1}^{N-1} \mu(A_{n+1}/A_n)$$

where $A_1 = J$, $A_N = x$ and $A_1 \leqslant A_2 \leqslant \ldots \leqslant A_N$. In the notation of the previous lemma S contains all points under the curve

$$y = I - \tfrac{1}{2}\,\theta(x, J).$$

This crosses the x-axis at $\theta^{-1}(2I)$, so the result follows from Proposition 4.1.

5. References

[1] Baker, A., Linear forms in the logarithms of algebraic numbers, Mathematika **13**, 204–216 (1966).
[2] Gelfond, A. O., Transcendental and Algebraic Numbers (1960).
[3] Ramachandra, K., A note on Baker's method, J. Austral. Math. Soc. **10**, 197–203 (1969).

(Paper 77B1 & 2–377)