Geometric and Arithmetic Postulation of the Exponential Function

J. Pila

Abstract. This paper presents new proofs of some classical transcendence theorems. We use real variable methods, and hence obtain only the real variable versions of the theorems we consider: the Hermite–Lindemann theorem, the Gelfond–Schneider theorem, and the Six Exponentials theorem. We do not appeal to the Siegel lemma to build auxiliary functions. Instead, the proof employs certain natural determinants formed by evaluating n functions at n points (alternants), and two mean value theorems for alternants. The first, due to Pólya, gives sufficient conditions for an alternant to be non-vanishing. The second, due to H. A. Schwarz, provides an upper bound.

1. Introduction

The purpose of this paper is to give new proofs of some classical results in the transcendence theory of the exponential function. We employ some determinantal mean value theorems, and some geometrical properties of the exponential function on the real line. Thus our proofs will yield only the real valued versions of the theorems we consider.

Specifically, we give proofs of (the real versions of) the six exponentials theorem, the Gelfond-Schneider theorem, and the Hermite-Lindemann theorem.

We do not use Siegel's lemma on solutions of integral linear equations. Using the data of the hypotheses, we construct certain determinants. With the aid of one of the mean value theorems, we show that these determinants are not zero. This result is, more precisely, a generalization due to Pólya [6] of Rolle's theorem to a determinant consisting of n functions evaluated at n points (i.e. an alternant). A second mean value theorem furnishes an upper estimate for this determinant, and a contradiction is reached in the usual way to conclude the transcendence proofs. This second mean value theorem, our proposition 2.1, goes back to H. A. Schwarz [11]. Versions of both results can be found in Pólya and Szegö [7].

The determinantal mean value theorems are given in section 2. None of the statements in that section relate specifically to the exponential function. The proofs of the transcendence theorems are then given in section 3.

In the Hermite-Lindemann and Gelfond-Schneider cases, our argument can be quantified to yield a transcendence measure. To do this one must give a *lower* bound for the non-vanishing alternant. The resulting transcendence measure appears to be weaker (in all aspects) than those obtainable by other methods; we nevertheless include the lower bound argument, in section 4. This involves yet another mean value theorem, one that does not hold for arbitrary sufficiently smooth functions. It depends on a curious positivity property.

Our approach was motivated by the following considerations. Let L_d denote the space of real algebraic plane curves of degree $\leq d$. The space L_d forms a real projective space of

dimension $\frac{1}{2}(d+1)(d+2)-1 = D-1$. Thus, $D-1$ points in the plane always lie on a curve in L_d, while D points in general do not. A function $f(x)$, defined on an interval I and possessing $D-1$ derivatives will be called *d-averse* if no D points on the graph Γ of $f(x)$ lie on a curve in L_d (counting multiplicity). This notion is the "Geometric Postulation" of the title. A curious fact is that the exponential function is d-averse for every d. This follows immediately from the abovementioned theorem of Pólya, our proposition 2.2; it is also readily proved using induction (on d) and the ordinary mean value theorem (see [7, vol.II, ch.V, problem 75]).

The aversity property entails the non-vanishing of a certain form of determinant (see [5]), and upper bounds furnished by Schwarz's mean value theorem can then be used to control integrality. Such considerations were applied by Bombieri and the author [1] to get upper bounds for the number of integral points on the graphs of functions under various kinds of hypotheses, and by the author to obtain some extensions in [5].

It was thus natural to attempt to link the aversion of the exponential function for algebraic curves with its aversion for algebraic points. Thus the title of the present paper.

We have not pursued the question of deducing the full complex statements of the theorems considered from the real versions; further, it seems attractive to consider other transcendence statements in these terms.

Note. Since submitting this paper we have been informed that these methods have been found independently and somewhat earlier by Laurent, but are not yet published. In [3], Laurent uses interpolation determinants to give a proof of the full six exponentials theorem. In place of our appeal to Pólya's mean value theorem, which shows that the determinant is not zero in the real case, he appeals to a zero lemma of Philippon to show that the determinant has large rank. His starting point [4] was the work of Cantor-Strauss connected with Lehmer's conjecture. Waldschmidt [15] gives a proof of six exponentials in the real case essentially identical to that given here. He also gives indications for the Hermite-Lindemann and Gelfond-Schneider theorems. They use complex variable methods to estimate the determinants, as opposed to our strictly real variable treatment using Schwarz's mean value theorem. According to a recent talk of Waldschmidt [16], they have been able to prove many other transcendence theorems by these methods. Since our treatment via mean value theorems is somewhat different, as was our motivation, we feel that the present paper may still be of some interest. We would like to thank Laurent, Waldschmidt, and the referee for drawing our attention to this work.

2. Determinantal Identities

The propositions of this section are all generalizations of the mean value theorem to determinants of matrices formed by evaluation of n functions at n points (with multiplicities). The qualitative (Rolle's theorem) and quantitative aspects are generalized separately. This distinction is meaningful in the presence of several variables.

2.1 A Mean Value Theorem

2

The following identity was established in [1]. It leads to a mean value theorem. Both are generalized in section 2.4. We begin with some notation. Let x, x_i, y_{ij} for $i, j = 1, \cdots, n$ be indeterminates, and let $V(x_1, \cdots, x_k)$ denote the Vandermonde determinant. Define

$$g_{ij}(x) = \frac{-1}{V(x_1, \cdots, x_i)} \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ & & \vdots & & \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & 0 \end{pmatrix}.$$

Note that, for an indeterminate y,

$$g_{ij}(x) = \frac{-1}{V(x_1, \cdots, x_i)} \det \begin{pmatrix} 1 & x_1 & \cdots & x_1^{i-1} & y_{1j} \\ & & \vdots & & \\ 1 & x_i & \cdots & x_i^{i-1} & y_{ij} \\ 1 & x & \cdots & x^{i-1} & y \end{pmatrix} + y ,$$

so that $g_{ij}(x)$ is the unique polynomial in x of degree $i-1$ with $g_{ij}(x_k) = y_{kj}$ for $k = 1, \cdots, i$. We write $g_{ij}^{(\ell)}$ for $(\frac{d}{dx})^\ell g_{ij}$.

Lemma 2.1. *With the above definitions,*

$$\det(y_{ij}) = \frac{V(x_1, \cdots, x_n)}{1! \ldots (n-1)!} \det \left(g_{ij}^{(i-1)} \right) . \square$$

We apply the lemma choosing $y_{ij} = \phi_j(x_i)$, where $x_1, \cdots, x_n \in I$ are distinct points and $\phi_j \in C^{n-1}$. Then the mean value theorem shows that $g_{ij}^{(i-1)}$ is in the range of $\phi_j^{(i-1)}$ (see for example Swinnerton–Dyer [13], Lemma 1, p. 131, or Posse [7]). We therefore get the following proposition.

Proposition 2.1. *Suppose the functions $\phi_1, \phi_2, \ldots, \phi_n$ possess derivatives of order $n-1$ on an interval $I = [a, b]$ and that $x_1, \ldots, x_n$ are distinct points of I. Let $\Delta = \det(\phi_j(x_i))$ be the $n \times n$ determinant. Then there exist intermediate points ξ_{ij} such that*

$$\Delta = \frac{V(x_1, \ldots, x_n)}{\Lambda(n)} \det \left(\phi_j^{(i-1)}(\xi_{ij}) \right),$$

where

$$\Lambda(n) = \prod_{i=1}^{n} (i-1)!. \square$$

A version of this theorem in which the interpolation points are the same in each column was proved by H.A. Schwarz ([11], Zw. Bd., p. 300), and also by T. Stieltjes [12]. A proof is also outlined in Pólya and Szegö [7, volume II, part 5, problems 95, 96].

It is natural to ask under what circumstances we can interpolate Δ by an intermediate value of the Wronskian: that is, take all the ξ_{ij} to be the same. That this does not hold in general is shown by the two functions $\sin x, \cos x$. In section 4 we show that it does hold for functions $\exp(x\xi)$ for distinct real ξ.

3

2.2 A Non-Vanishing Criterion

The following generalization of Rolle's theorem is due to Pólya [6]. The version without multiplicities is also outlined in Pólya and Szegö [7, volume II, part 5, problem 99]. For functions $\phi_1, \phi_2, \ldots, \phi_n$ possessing $n-1$ derivatives on an interval,

$$W(\phi_1, \phi_2, \ldots, \phi_n)$$

will denote the $n \times n$ Wronskian determinant.

Proposition 2.2. *Suppose the functions $\phi_1, \phi_2, \ldots, \phi_n$ possess derivatives of order $n-1$ on an interval $I = [a, b]$ and satisfy:*

$$W(\phi_1) > 0, \ W(\phi_1, \phi_2) > 0, \ldots, \ W(\phi_1, \ldots, \phi_{n-1}) > 0$$

for $x \in I$. Suppose that $x_1, \ldots, x_h$ are distinct points of I, and $k_1, \ldots, k_h$ are positive integers satisfying:

$$a \le x_1 < x_2 < \ldots < x_h \le b, \quad k_1 + k_2 + \ldots + k_h = n.$$

Let Δ be the $n \times n$ determinant:

$$\det \begin{pmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_n(x_1) \\ \phi_1'(x_1) & \phi_2'(x_1) & \ldots & \phi_n'(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{(k_1-1)}(x_1) & \phi_2^{(k_1-1)}(x_1) & \ldots & \phi_n^{(k_1-1)}(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{(k_h-1)}(x_h) & \phi_2^{(k_h-1)}(x_h) & \ldots & \phi_n^{(k_h-1)}(x_h) \end{pmatrix}.$$

There exists an intermediate point x, with $x_1 < x < x_k$, such that the value assumed at x by the Wronskian $W(\phi_1, \ldots, \phi_n)$ is $<, =, > 0$ according as Δ is $<, =, > 0$. $\square$

We appeal to the following contrapositive of the proposition: if

$$W(\phi_1), \ W(\phi_1, \phi_2), \ldots, \ W(\phi_1, \ldots, \phi_n)$$

are non-vanishing on I, then Δ is not zero.

2.3 A Determinant of Vandermonde Type

We evaluate a determinant of Vandermonde type. The result is due to Schendel [10]. In the next subsection, this will be used in an extension of proposition 2.1 to the case where the rows occur with multiplicities. We begin with some notation. Let k_i be positive integers for $i = 1, \ldots, h$, with $k_1 + \ldots k_h = n$, and suppose that $x_i, i = 1, \ldots, h$ are indeterminates. We denote by

$$V(k_1, \ldots, k_h)$$

4

the determinant of the $n \times n$ matrix

$$
\begin{pmatrix}
1 & x_1 & x_1^2 & \cdots & & x_1^{k_1-1} & & \cdots & & x_1^{n-1} \\
 & 1 & 2x_1 & \cdots & & (k_1-1)x_1^{k_1-2} & & \cdots & & (n-1)x_1^{n-2} \\
 & & 2 & \cdots & & \cdots & & \cdots & & \\
 & & & \vdots & & \vdots & & \vdots & & \\
 & & & & & (k_1-1)! & & \cdots & & \\
1 & x_2 & x_2^2 & \cdots & x_2^{k_2-1} & & & \cdots & & x_2^{n-1} \\
 & 1 & 2x_2 & \cdots & (k_1-1)x_2^{k_2-2} & & & \cdots & & (n-1)x_2^{n-2} \\
 & & 2 & \cdots & \cdots & & & \cdots & & \\
 & & & \vdots & \vdots & & & \vdots & & \\
 & & & & (k_2-1)! & & & \cdots & & \\
 & & & \vdots & & \vdots & & \vdots & & \\
1 & x_h & x_h^2 & \cdots & & & x_h^{k_h-1} & \cdots & & x_h^{n-1} \\
 & & & \vdots & & & \vdots & & \vdots & \\
 & & & & & & (k_h-1)! & \cdots & &
\end{pmatrix}.
$$

Thus if $h = n$ and $k_i = 1$, then $V(1,1,\ldots,1)$ reduces to the ordinary Vandermonde determinant on $x_1,\ldots,x_n$; while if $h = 1$ and $k_1 = n$ we have

$$
V(n) = \prod_{i=1}^{n}(i-1)! = \Lambda(n).
$$

The case in which the k_i have a common value is given in Weihrauch [17]. The general case is due to Schendel [10], and we will refer to $V(k_1,\ldots,k_h)$ as a Schendel determinant. (The referee suggests the more descriptive term: confluent Vandermonde determinant.)

Lemma 2.3. *With the above definitions,*

$$
V(k_1,\ldots,k_h) = \prod_{i=1}^{h}\Lambda(k_i)\prod_{i<j}(x_j - x_i)^{k_i k_j}. \ \square
$$

2.4 A Further Mean Value Theorem

We now generalize lemma 2.1, and deduce a corresponding generalization of proposition 2.1. The extension we need replaces the Vandermonde determinant appearing in the statement with a Schendel determinant; however it eases notation to prove a version involving a completely arbitrary matrix. The proof procedure is identical to that given in [1].

We begin by letting

$$
M = (m_{ij}), \quad Y = (y_{ij}), \quad i,j = 1,\ldots,n
$$

be matrices of indeterminates. For a matrix A, we denote by

$$
A_i, \quad A^{*j}, \quad A^k, \quad A^{(a,b)}
$$

5

respectively the top left $i \times i$ submatrix of A, the rectangular matrix formed by removing column j of A, the column vector consisting of column k of A, and the (a, b) minor determinant of A. We further let X_i be the row vector $(1, x, \ldots, x^{i-1})$.

Now set

$$g_{ij}(x) = \frac{-1}{\det M_i} \det \begin{pmatrix} M_i & Y_i^j \\ X_i & 0 \end{pmatrix}.$$

We view $g_{ij}(x)$ as a polynomial in x, and denote by $g_{ij}^{(k)}(x)$ its derivative of order k.

Lemma 2.4. *With the above definitions,*

$$\det(Y) = \frac{\det M}{\Lambda(n)} \det \left(g_{ij}^{(i-1)} \right).$$

Proof. By evaluation of the right-hand side.

$$\frac{\det M_n}{\Lambda(n)} \det \left(g_{ij}^{(i-1)} \right)$$

$$= \frac{\det M_n}{\Lambda(n)} \det \left(\frac{-1}{\det M_i} \det \begin{pmatrix} M_i & Y_i^j \\ X_i^{(i-1)} & 0 \end{pmatrix} \right)$$

$$= \frac{\det M_n}{\det M_1 \ldots \det M_n} \det \left(\det \begin{pmatrix} M_i^{*i} & Y_i^j \end{pmatrix} \right)$$

$$= \frac{1}{\det M_1 \ldots \det M_{n-1}} \det \left(\sum_{k=1}^{i} (-1)^{k-i} M_i^{(k,i)} y_{kj} \right).$$

We express this last matrix as a product of a lower triangular matrix and the matrix Y to obtain the expression

$$\frac{1}{\det M_{n-1}} \det \begin{pmatrix} & & \bigcirc \\ & \ddots & \\ (-1)^{j-i} M_i^{(j,i)} & & \end{pmatrix} \det(y_{ij})$$

$$= \frac{1}{\det M_1 \ldots \det M_{n-1}} \cdot \prod_{k=1}^{n-1} \det M_k \cdot \det(y_{ij})$$

$$= \det(y_{ij}) \, . \, \square$$

We apply the lemma taking $M = V(k_1, \ldots, k_h)$, and

$$Y = \begin{pmatrix} y_{11}^0 & y_{12}^0 & \cdots & y_{1n}^0 \\ y_{11}^1 & y_{12}^1 & \cdots & y_{1n}^1 \\ \vdots & \vdots & & \vdots \\ y_{11}^{k_1-1} & y_{12}^{k_1-1} & \cdots & y_{1n}^{k_1-1} \\ y_{21}^0 & y_{22}^0 & \cdots & y_{2n}^0 \\ \vdots & \vdots & & \vdots \\ y_{h1}^{k_h-1} & y_{h2}^{k_h-1} & \cdots & y_{hn}^{k_h-1} \end{pmatrix}.$$

Note that, for an indeterminate y,

$$g_{ij}(x) = y - \frac{1}{\det M_i} \det \begin{pmatrix} M_i & Y_i^j \\ X_i & y \end{pmatrix},$$

so that $g_{ij}(x)$ is the unique polynomial in x of degree $i-1$ with $g_{ij}^{(p)}(x_\ell) = y_{kj}^p$ for $\ell = 1, \cdots; p = 1, \ldots, k_\ell$.

We choose $y_{ij}^p = \phi_j^{(p)}(x_i)$, where $x_1, \cdots, x_h$ are distinct points and $\phi_j \in C^{n-1}$. Then the mean value theorem shows that $g_{ij}^{(i-1)}$ is in the range of $\phi_j^{(i-1)}$. We thus get the following proposition.

Proposition 2.4. *Suppose the functions $\phi_1, \phi_2, \ldots, \phi_n$ possess derivatives of order $n-1$ on an interval $I = [a, b]$, that $x_1, \ldots, x_h$ are distinct points of I and that $k_1, \ldots, k_h$ are positive integers with $k_1 + \ldots + k_h = n$. Let Δ be the $n \times n$ determinant*

$$\det \begin{pmatrix} \phi_1(x_1) & \phi_2(x_1) & \ldots & \phi_n(x_1) \\ \phi_1'(x_1) & \phi_2'(x_1) & \ldots & \phi_n'(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{(k_1-1)}(x_1) & \phi_2^{(k_1-1)}(x_1) & \ldots & \phi_n^{(k_1-1)}(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ldots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{(k_h-1)}(x_h) & \phi_2^{(k_h-1)}(x_h) & \ldots & \phi_n^{(k_h-1)}(x_h) \end{pmatrix}.$$

Then there exist intermediate points ξ_{ij} such that

$$\Delta = \frac{V(k_1, \ldots, k_h)}{\Lambda(n)} \det \left(\phi_j^{(i-1)}(\xi_{ij}) \right). \quad \square$$

2.5 The Quantity $\Lambda(n)$

We have defined

$$\Lambda(n) = \prod_{i=1}^{n} (i-1)! = \prod_{i=1}^{n} (i-1)^{n-i+1}.$$

The size of this quantity is crucial to our arguments. Chiefly we need lower bounds, but in our proof of the Hermite-Lindemann theorem we require also upper bounds. We have

$$\log \Lambda(n) = \sum_{i=1}^{n-1} (n-i) \log i.$$

We obtain upper and lower bounds by comparing with the corresponding integrals.

Lemma 2.5. *For $n \geq 2$, we have*

$$\frac{n^2}{2} \log n - 3n^2 \leq \log \Lambda(n) \leq \frac{n^2}{2} \log n - n^2. \quad \square$$

3. Transcendence Proofs

3.1 The Six Exponentials Theorem

Here we use the non-vanishing of the determinant

$$\Delta = \det\big(\exp(x_i\xi_j)\big)$$

whenever $\{x_1,\ldots,x_m\}$ and $\{\xi_1,\ldots,\xi_m\}$ are sets of distinct real numbers. This follows from proposition 2.2, since the corresponding Wronskian determinants do not vanish: this is because the functions comprising them form a basis of solutions to a linear ordinary differential equation. Another proof is given in Pólya and Szegö [7, vol. II, part 5, problem 76]. They go on to show the non-vanishing of a determinant $\det(F(x_i\xi_j))$ when F has a non-terminating Taylor expansion with non-negative coefficients (problem 86). This is in connection with an investigation into Descarte's rule of signs.

Theorem 3.1. *Suppose that x_1, x_2, x_3 are real numbers, linearly independent over the rational numbers, and suppose that ξ_1, ξ_2 are real numbers, linearly independent over the rational numbers. Then one at least of the six numbers*

$$\exp(x_i\xi_j)$$

is transcendental.

This theorem is given in Lang [2, Ch.2, thm 1], but was known to Siegel and others, as remarked by Lang. The result was also discovered independently by Ramachandra [8]. A stronger statement, the strong six exponentials theorem, has been proven by Waldschmidt [14]. A discussion is given there of related results and conjectures.

Proof. The proof is by contradiction: we suppose the numbers to be algebraic, and to generate a finite extension L of $\mathbf{Q}$ of degree m. Let S be the set of embeddings of L in $\mathbf{C}$. We suppose without loss of generality that $x_1, x_2, x_3, \xi_1, \xi_2$ are positive, and bounded by B. We choose T so that

$$|\exp(x_i\xi_j)^\sigma| \le T, \quad i = 1, 2, 3; \ j = 1, 2; \ \sigma \in S,$$

and positive integers e_{ij} such that

$$e_{ij}\exp(x_i\xi_j)$$

are algebraic integers. We let $N = \prod e_{ij}$.

We now take $n \ge 2$ to be an integer, later to be sent to infinity, and let Δ be the $n^6 \times n^6$ determinant with entries

$$\exp\big((i_1x_1 + i_2x_2 + i_3x_3)(j_1\xi_1 + j_2\xi_2)\big),$$

where $1 \le i_1, i_2, i_3 \le n^2$ index rows, and $1 \le j_1, j_2 \le n^3$ index columns.

8

As remarked above, Δ is not zero. We now estimate Δ as an algebraic number. Since $i_\nu j_\mu \leq n^5$, and Δ is a $n^6 \times n^6$ determinant, we find that, for any $\sigma \in S$,

$$|\Delta^\sigma| \leq n^6! (T^{6n^5})^{n^6}.$$

Thus

$$\log |\Delta^\sigma| = \mathrm{o}(n^{12}).$$

By multiplying each row of Δ^σ by N^{n^5}, one obtains an algebraic integer. Thus

$$\Omega = N^{mn^{11}} \prod_{\sigma \in S} \Delta^\sigma$$

is a non-vanishing rational integer. Thus

$$1 \leq |\Omega|.$$

We now seek an upper bound for $|\Omega|$. For $|\Delta^\sigma|$ we will use the estimate above; but for Δ itself we will use the mean value theorem (proposition 2.1). We find that

$$\Delta = \frac{V(\{i_1 x_1 + i_2 x_2 + i_3 x_3\})}{\Lambda(n^6)} \det(M),$$

with

$$M = \left((j_1 \xi_1 + j_2 \xi_2)^{i-1} \exp((j_1 \xi_1 + j_2 \xi_2) \eta_{i(j_1, j_2)}) \right)$$

where rows are indexed by i, and $\eta_{i(j_1, j_2)}$ is an appropriate intermediate point of the $i_1 x_1 + i_2 x_2 + i_3 x_3$.

Now since $x_i \leq B$ and $i_\nu \leq n^2$ we have

$$|V(\{i_1 x_1 + i_2 x_2 + i_3 x_3\})| \leq (3n^2 B)^{\frac{n^{12}}{2}},$$

so that

$$\log |V| \leq n^{12} \log n + \text{smaller terms},$$

while according to 2.5

$$\Lambda(n^6) \geq \exp(3n^{12} \log n - 3n^{12}).$$

Considering M, we have $j_1 \xi_1 + j_2 \xi_2 \leq 2n^3 B$, and

$$\exp\left((j_1 \xi_1 + j_2 \xi_2) \eta_{i(j_1, j_2)}\right) \leq \exp(6n^5 T);$$

thus an entry in row i is bounded by

$$(2n^3 B)^{i-1} \exp(6n^5 T).$$

We conclude that

$$|\det M| \leq n^6! \exp(6n^{11} T)(2n^3 B)^{\sum_{i=1}^{n^6} i - 1} \leq n^6! \exp(6n^{11} T)(2n^3 B)^{\frac{n^{12}}{2}},$$

from which we see that

$$\log |\det M| \leq \frac{3}{2} n^{12} \log n + \text{smaller terms}.$$

Combining, we find that

$$|\Omega| \leq \exp(-\frac{1}{2} n^{12} \log n + \text{smaller terms});$$

this contradicts $1 \leq |\Omega|$ when n is sufficiently large. $\square$

3.2 The Gelfond-Schneider Theorem

Here we use the non-vanishing of the determinant

$$\Delta = \det\left(x_i^{j_1}\exp(\lambda j_2 x_i)\right)$$

when the x_i are distinct real numbers, λ is real and not zero, $j_1 = 0, 1, \ldots,$ and j_2 are distinct. Again this follows from proposition 2.2, since the corresponding Wronskian determinants do not vanish: as before this is because the functions comprising them form a basis of solutions to a linear ordinary differential equation.

Theorem 3.2. *Suppose that α is a real positive algebraic number with $\alpha \neq 1$, and suppose β is a real algebraic irrational number. Then α^β is transcendental.*

Proof. Let $\gamma = \alpha^\beta$. The proof is by contradiction: we suppose γ to be algebraic, and that α, β, γ generate a finite extension L of $\mathbf{Q}$ of degree m, with embeddings S. We suppose without loss of generality that β is positive. We choose T so that

$$|\alpha^\sigma|, 1 + |\beta^\sigma|, |\gamma^\sigma| \leq T, \quad \sigma \in S,$$

and positive integers a, b, c such that

$$a\alpha, \quad b\beta, \quad c\gamma$$

are algebraic integers. We let $N = abc$.

We now take $n \geq 2$ to be an integer, and let Δ be the $n^4 \times n^4$ determinant with entries

$$\exp\left((i_1 + i_2\beta)(j_2\log\alpha)\right)(i_1 + i_2\beta)^{j_1 - 1},$$

where $1 \leq i_1, i_2 \leq n^2$ index rows, and $1 \leq j_1 \leq n^3, \; j_2 \leq n$ index columns.

Thus $\Delta \in L$ is non-zero. We estimate Δ as an algebraic number. The entries have the form:

$$\sigma(\alpha)^{i_1 j_2}\sigma(\gamma)^{i_2 j_2}\sigma(i_1 + i_2\beta)^{j_1},$$

and these have absolute value bounded by

$$T^{n^3}T^{n^3}(n^2 T)^{n^3}.$$

Thus

$$|\Delta^\sigma| \leq n^4! T^{3n^7}n^{2n^7}.$$

Multiplying each row of Δ^σ by N^{n^3} yields an algebraic integer. Thus

$$\Omega = N^{mn^7}\prod_{\sigma \in S}\Delta^\sigma$$

is a non-vanishing rational integer. Thus $1 \leq |\Omega|$.

10

We now pursue an upper bound for $|\Omega|$. For $|\Delta^\sigma|$ we will use the estimate above; but for Δ itself we will use proposition 2.1. We find that

$$\Delta = \frac{V(\{i_1 + i_2\beta\})}{\Lambda(n^4)} \det(M),$$

with

$$M = \left((x^{j_1}\alpha^{j_2 x})^{(i-1)}\Big|_{x=\eta_{i(j_1,j_2)}}\right)$$

where rows are indexed by i, superscript $(i-1)$ means differentiation, and $\eta_{i(j_1,j_2)}$ is an appropriate intermediate point of the $i_1 + i_2\beta$.

We have

$$|V(\{i_1 + i_2\beta\})| \leq (n^2 T)^{\frac{n^8}{2}},$$

while

$$\Lambda(n^4) \geq \exp(2n^8 \log n - 3n^8).$$

Considering M, we first note $(\log \alpha)\eta_{i(j_1,j_2)} \leq n^2 T$. Now

$$\left(\frac{d}{dx}\right)^\ell (x^{j_1}\alpha^{j_2 x}) = \sum_{\nu+\mu=\ell} \frac{\ell!}{\nu!\mu!}\left(\frac{d}{dx}\right)^\nu x^{j_1}\left(\frac{d}{dx}\right)^\mu \alpha^{j_2 x}.$$

It suffices to use the estimate $\sum_{\nu+\mu=\ell} \frac{\ell!}{\nu!\mu!} \leq 2^\ell$; but we will be more frugal, and note that $\left(\frac{d}{dx}\right)^\nu x^{j_1}$ vanishes if $\nu \geq j_1 + 1$, and that $j_1 \leq n^3 - 1$. Thus an entry in row i of M is bounded by

$$n^3(i-1)^{n^3} n^3! (n^2 T)^{n^3} (Tn^2)^n (n \log \alpha)^{i-1},$$

the last term being the dominant and significant one. We conclude that

$$|\det(M)| \leq n^4! \left[n^3(n^4)^{n^3} n^3! (n^2 T)^{n^3+n}\right]^{n^4} (n \log \alpha)^{\Sigma_{i=1}^{n^4} i - 1}.$$

Since $\sum_{i=1}^{n^4} i - 1 \leq \frac{1}{2}n^8$, combining our estimates yields

$$|\Omega| \leq \exp(-\frac{1}{2}n^8 \log n + \text{smaller terms});$$

this contradicts $1 \leq |\Omega|$ when n is sufficiently large. $\square$

3.2 The Hermite-Lindemann Theorem

Here we use the non-vanishing of the determinant

$$\Delta = \det\left((x_{i_1}^{j_1} \exp(j_2 x_{i_1}))^{(i_2-1)}\right)$$

when the x_i are distinct real numbers: indeed the Wronskians are the same as in the previous subsection. The proof proceeds as before, except that we appeal to proposition 2.4 instead of proposition 2.1 since we now have multiplicities in the rows.

Theorem 3.3. *Suppose that α is a non-zero real algebraic number. Then e^α is transcendental.*

Proof. We can assume that α is positive. Suppose that α and e^α are algebraic and generate a finite extension L of $\mathbf{Q}$ of degree m, with embeddings S. We choose T so that

$$|\alpha^\sigma|, |(e^\alpha)^\sigma| \leq T, \ \ \sigma \in S,$$

and positive integers a, b such that

$$a\alpha, \ \ be^\alpha$$

are algebraic integers. We let $N = ab$.

We now take $n \geq 2$ to be an integer, and let Δ be the $n^4 \times n^4$ determinant with entries

$$\left(\exp(j_2 x) x^{j_1 - 1} \right)^{(i_1 - 1)} \Big|_{x = (i_2 - 1)\alpha},$$

where $1 \leq i_1, i_2 \leq n^2$ index rows, and $1 \leq j_1 \leq n^3$, and $1 \leq j_2 \leq n$ index columns.

Thus $\Delta \in L$ is non-zero. We estimate Δ as an algebraic number. Since each entry is bounded by

$$2^{n^2} n^3! (n^2 T)^{n^3} n^{n^2} T^{n^3}$$

we conclude that for any $\sigma \in S$

$$|\Delta^\sigma| \leq n^4! 2^{n^6} (n^3!)^{n^4} (n^2 T^2)^{n^7} n^{n^6}.$$

Further,

$$\Omega = N^{mn^7} \prod_{\sigma \in S} \Delta^\sigma$$

is a non-vanishing rational integer, whence $1 \leq |\Omega|$.

By proposition 2.4

$$\Delta = \frac{\Lambda(n^2)^{n^2} V(\{(i_2 - 1)\alpha\})^{n^4}}{\Lambda(n^4)} \det(M),$$

with

$$M = \left((x^{j_1} e^{j_2 x})^{(i-1)} \Big|_{x = \eta_{i(j_1, j_2)}} \right)$$

where rows are indexed by i, superscript $(i - 1)$ means differentiation, and $\eta_{i(j_1, j_2)}$ is an appropriate intermediate point of the $(i_2 - 1)\alpha$.

We have

$$|V|^{n^4} \leq (n^2 T)^{\frac{n^8}{2}};$$

$$\Lambda(n^2)^{n^2} \leq \exp(n^6 \log n);$$

$$\Lambda(n^4) \geq \exp(2n^8 \log n - 3n^8).$$

The derivatives of x^{j_1} vanishing after order $j_1 \leq n^3$, we find that an entry in row i of M is bounded by

$$n^3 (i-1)^{n^3} n^3! (n^2 T)^{n^3} T^{n^3} n^{i-1}.$$

Therefore

$$|\det(M)| \leq n^4! \left[n^3 (n^4)^{n^3} n^3! (n^2 T^2)^{n^3} \right]^{n^4} (n)^{\sum_{i=1}^{n^4} i - 1}.$$

Since $\sum_{i=1}^{n^4} i - 1 \leq \frac{1}{2} n^8$, combining our estimates yields

$$|\Omega| \leq \exp(-\frac{1}{2} n^8 \log n + \text{smaller terms});$$

contrary to $1 \leq |\Omega|$ for n sufficiently large. □

4. A Determinant

We consider here the determinant

$$\Delta = \det \left(\exp(x_i \xi_j) \right)$$

where x_i and ξ_j are distinct complex variables. It is divisible (as an entire function) by the Vandermonde determinants in each system of variables, V_x and V_ξ. The quotient

$$\frac{\Delta}{V_x \, V_\xi}$$

is entire: the value taken when arguments are repeated being the quotient by the appropriate Schendel determinants of the determinant in which repeated arguments yield differentiated rows or columns.

Our main conclusion is that this quotient, as a function of $2n$ real variables, is strictly increasing in each of them. It follows that if $A \leq x_i \leq B$ and $C \leq \xi_j \leq D$ then

$$\frac{\exp(nAC)}{\Lambda(n)} \leq \frac{\Delta}{V_x \, V_\xi} \leq \frac{\exp(nBD)}{\Lambda(n)}.$$

We will prove this via a completely formal argument. We let

$$F(z) = \sum_{i=0}^{\infty} a_i z^i$$

be a formal power series with indeterminate coefficients. Further let x_i, ξ_j, $i, j = 1, \ldots, n$ be indeterminates and put

$$\Delta = \det \left(F(x_i \xi_j) \right).$$

Proposition 4.1 *The quotient*

$$\frac{\Delta}{V_x \, V_\xi}$$

is a formal power series in $\mathbf{Z}[a_i][[x_i, \xi_j]]$ whose coefficients are all non-negative integers. Moreover, the term corresponding to any selection of n distinct a_i is non-vanishing.

13

Proof. Expanding the determinant Δ, we see that all the terms have the form

$$a_{k_1} x_1^{k_1} \xi_{j_1}^{k_1} \cdots a_{k_n} x_n^{k_n} \xi_{j_n}^{k_n}$$

where j_n gives a permutation of $1, \ldots, n$. We will consider the terms corresponding to a particular choice of n of the a_i. Consider a selection of n distinct a_i,

$$a_{k_1}, a_{k_2}, \ldots, a_{k_n}; \quad k_1 < k_2 \cdots k_n.$$

The corresponding part of the determinant, comprising all terms whose a_i lie among the selected set, is given by

$$\det \Big(\sum_{m=1}^{n} a_{k_m} x_i^{k_m} \xi_j^{k_m} \Big).$$

We can factor this matrix to obtain the expression

$$\det \big(a_{k_j} x_i^{k_j} \big) \ \det \big(\xi_j^{k_i} \big)$$

$$= a_{k_1} \cdots a_{k_n} \ \det \big(x_i^{k_j} \big) \ \det \big(\xi_j^{k_i} \big).$$

We thus see that all the surviving terms in the expansion of Δ involve some selection of n distinct a_i. The assertion of the lemma is now reduced to showing that for $k_1 < k_2 < \cdots < k_n$ the quotient of determinants

$$\frac{\det \big(x_i^{k_j} \big)}{\det \big(x_i^{j-1} \big)}$$

is a symmetric function with non-negative integral coefficients. This is easily shown by induction on n: subtracting the first row of the matrix in the numerator from the subsequent rows, dividing through row $i \neq 1$ by $x_i - x_1$, and using the column linearity of the determinant expresses the desired quotient as a positive integral sum of like determinants of order $n - 1$. $\square$

Corollary. *Suppose that $F(z) = \Sigma a_i z^i$ is a power series with non-negative real coefficients, of which at least n are not zero. Suppose the power series converges for $|z| \leq r, r > 0$. Then the determinantal quotient of the proposition is analytic for $|x_i \xi_j| < r$, and, as a function of $2n$ real variables, is strictly increasing in each of them for all positive values.* $\square$

The determinantal quotient formed using the exponential function is strictly increasing for *all* real arguments, positive or negative, because

$$\det \big(\exp((x_i + h)\xi_j) = \exp(h\xi_1 + \cdots h\xi_n) \ \det \big(\exp(x_i \xi_j) \big),$$

$$\det \big(\exp(x_i(\xi_j + k)) = \exp(kx_1 + \cdots kx_n) \ \det \big(\exp(x_i \xi_j) \big)$$

while the Vandermonde (or Schendel) determinants are translation invariant.

14

We conclude that, given $x_1 \leq \cdots \leq x_n$ and $\xi_1 \leq \cdots \leq \xi_n$ (allowing multiplicities), there exist intermediate points x and ξ such that

$$\frac{\Delta}{V_x \ V_\xi} = \frac{\exp(nx\xi)}{\Lambda(n)}.$$

The above enables us to simplify the estimates for the determinants Δ used in the transcendence proofs. However, we have preferred to prove those theorems using only the mean value theorems valid for arbitrary functions.

To get lower bounds for Δ we must have lower bounds for the corresponding Schendel determinants. These can be evaluated explicitly in the Hermite-Lindemann case, since the x_i are integral multiples of α, and the ξ_j are integers. In the Gelfond Schneider case, a lower bound for the Schendel determinant may be obtained using (say) the Liouville estimate for rational approximations to an irrational algebraic number. The lower bound can then be used to get a transcendence measure, as remarked in the introduction.

References.

[1] BOMBIERI, E., AND PILA, J., The Number of Integral Points on Arcs and Ovals, *Duke Math. J.* **59** (1989), 337–357.

[2] LANG, S., *Introduction to Transcendental Numbers,* Addison-Wesley, Palo Alto, 1966.

[3] LAURENT, M., Sur quelques resultats recents de transcendence, to appear in *Journeés arithmétiques de Luminy /89,* Astérisque.

[4] LAURENT, M., letter dated 12 June 1991.

[5] PILA, J., Geometric Postulation of a Smooth Function and the Number of Rational Points, *Duke Math. J.* **63** (1991), 449–463.

[6] PÓLYA, G., On the Mean Value Theorem corresponding to a given Linear Homogeneous Differential Equation, *Trans. Amer. Math. Soc.* **24** (1922), 312–324. Also *Collected Papers*, vol. 3, M.I.T. Press, Cambridge, 1984, pp.96-108.

[7] PÓLYA, G., AND SZEGÖ, G., *Problems and Theorems in Analysis,* Springer-Verlag, Berlin, 1976.

[8] POSSE, C., Sur le terme complémentaire de la formule de M. Tchebychef donnant l'expression approchée d'une intégrale définie par d'autres prises entre les mêmes limites, *Bull. Sci. Math.* (2) **7** (1883), 214–224.

[9] RAMACHANDRA, K., Contributions to the theory of Transcendental Numbers, *Acta Arith.,* **14** (1968), 65-68.

[10] SCHENDEL, L., Mathematische Miscellen. III. Das alternierende Exponential- differenzproduct, *Zeitschrift fur Math. u. Phys.* **38** (1891), 84-87.

[11] SCHWARZ, H.A., Verallgemeinerung eines analytischen Fundamentalsatzes, *Annali di Mat.* (2) **10** (1880), 129–136. Also *Gesammelte Mathematische Abhandlungen.* Julius Springer, Berlin 1890, Bd. 2, 296–302.

[12] STIELTJES, T.J., Sur une généralisation de la formule des accroissements finis, *Bull. Soc. Math. de France,* **16** (1887), 100-113.

[13] SWINNERTON–DYER, H.P.F., The Number of Lattice Points on a Convex Curve, *J. Number Theory* **6** (1974), 128–135.

[14] WALDSCHMIDT, M., On the transcendence methods of Gelfond and Schneider in Several Variables, *New Advances in Transcendence Theory,* Baker, ed., Cambridge, 1988.

[15] WALDSCHMIDT, M., *Transcendence des valeurs de la fonction exponentielle,* course notes, Fall 1990.

[16] WALDSCHMIDT, M., *Transcendence problems in several variables,* plenary address, third conference of the Canadian number theory association, Queen's University, August 1991.

[17] WEIHRAUCH, K., Ueber eine algebraische Determinante mit eigenthümlichem Bildungsgesetz der Elemente, *Zeitschrift fur Math. u. Phys.* **36** (1889), 34-40.

August 27, 1991.

Department of Mathematics, Columbia University, New York, NY 10027.