Linear Independence Measures for Logarithms of Algebraic Numbers

Michel Waldschmidt

Institut de Mathématiques, Université Pierre et Marie Curie (Paris VI)

Let $\alpha_1, \ldots, \alpha_n$ be nonzero algebraic numbers and $b_1, \ldots, b_n$ rational integers. Assume $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$. According to Liouville's inequality (Proposition 1.13), the lower bound

$$\left|\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1\right| \geq e^{-cB}$$

holds with $B = \max\{|b_1|, \ldots, |b_n|\}$ and with a positive number c depending only on $\alpha_1, \ldots, \alpha_n$. A fundamental problem is to prove a sharper estimate.

Transcendence methods lead to linear independence measures, over the field of algebraic numbers, for logarithms of algebraic numbers. Such measures are nothing else than lower bounds for numbers of the form

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n,$$

where $\beta_0, \ldots, \beta_n$ are algebraic numbers, $\alpha_1, \ldots, \alpha_n$ are nonzero algebraic numbers, while $\log \alpha_1, \ldots, \log \alpha_n$ are logarithms of $\alpha_1, \ldots, \alpha_n$ respectively.

In the special case where $\beta_0 = 0$ and $\beta_1, \ldots, \beta_n$ are rational integers, writing b_i for β_i, we have

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n,$$

which is the so-called *homogeneous rational case.* The importance of this special case is due to the fact that for $|\Lambda| \leq 1/2$ we have

$$\frac{1}{2}|\Lambda| \leq \left|e^{\Lambda} - 1\right| \leq 2|\Lambda|$$

with

$$e^{\Lambda} - 1 = \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1.$$

Hence we are back to the problem of estimating from below the distance between 1 and a number of the form $\alpha_1^{b_1} \cdots \alpha_n^{b_n}$.

The first three lectures are devoted to the qualitative theory of transcendental numbers, the last three ones to the quantitative theory of Diophantine approximation.

D. Masser et al.: LNM 1819, F. Amoroso and U. Zannier (Eds.), pp. 249–344, 2003.

According to Hermite-Lindemann's Theorem, a number $\Lambda = \beta - \log \alpha$, with algebraic α and β, is zero only in the trivial case $\beta = \log \alpha = 0$. We start by assuming further that α and β are positive integers. It is a nontrivial fact that a positive integer b cannot be the logarithm of another positive integer a. In the first lecture we give two proofs of this result: the first one uses an auxiliary function, the second one uses an interpolation determinant together with a zero estimate. In the second lecture we complete the proof of Hermite-Lindemann's Theorem in the general case (with algebraic α and β), by means of the interpolation determinant method, but without a zero estimate: this is achieved thanks to an extrapolation argument.

In the third lecture we introduce Baker's Theorem on the linear independence of logarithms of algebraic numbers. After a brief survey of the available methods, we produce a proof by means of an interpolation determinant involving an extrapolation.

An introduction to Diophantine approximation is given in Section 4, where we address the question of estimating from below the distance between b and $\log a$, for a and b positive integers. A conjecture attributed to K. Mahler states that this distance should be at least a negative power of a:

$$|b - \log a| \overset{?}{\geq} a^{-c} \quad \text{for} \quad a \geq 2.$$

So far one does not know how to prove this result with a constant exponent, but only with exponent a constant times $\log \log a$ (K. Mahler; see (4.3) below):

$$|b - \log a| \geq a^{-c \log \log a} \quad \text{for} \quad a \geq 3.$$

We discuss a proof of this result by means of a method which is inspired by a recent work of M. Laurent and D. Roy [15].

The last two sections are devoted to Baker's method and to the question of measures of linear independence for an arbitrary number of logarithms of algebraic numbers. In the fifth lecture we survey available methods and in the last one we explain how to replace Matveev's auxiliary function by an interpolation determinant.

Notation. As a general rule we use the notation of [34]. In particular the absolute logarithmic height is denoted by h. The *length* of a polynomial $P \in \mathbb{C}[X_1, \ldots, X_n]$ (which is nothing else than the sum of the absolute values of its coefficients) is denoted by $\mathrm{L}(P)$.

For $\underline{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$, we set

$$|\underline{z}| = \max_{1 \leq i \leq n} |z_i| \quad \text{and} \quad \|\underline{z}\| = |z_1| + \cdots + |z_n|.$$

In Section 4.1, $\|\cdot\|_{\mathbb{Z}}$ denotes the distance of a real number to the nearest integer:

$$\|x\|_{\mathbb{Z}} = \min_{k \in \mathbb{Z}} |x - k|.$$

For $x \in \mathbb{R}$ we set

$$\log_+ x = \log\max\{1, x\} \quad \text{and} \quad |x|_+ = \max\{1, |x|\}.$$

The integral part of x is denoted by $[x]$:

$$[x] \in \mathbb{Z}, \quad 0 \le x - [x] < 1$$

while $\lceil x \rceil$ denotes the least integer $\ge x$:

$$\lceil x \rceil \in \mathbb{Z}, \quad 0 \le \lceil x \rceil - x < 1.$$

For $r \ge 0$ we denote by

$$B_n(0,r) = \left\{ \underline{z} \in \mathbb{C}^n \ ; \ |\underline{z}| \le r \right\}$$

the closed polydisk of $\mathbb{C}^n$ of center 0 and radius r, by $|f|_r$ the supremum norm of a continuous function $f\colon B_n(0,r) \to \mathbb{C}$ and by $\mathcal{H}_n(r)$ the set of continuous functions $f\colon B_n(0,r) \to \mathbb{C}$ which are holomorphic in the interior of $B_n(0,r)$.

Several differential operators will be used. For $\underline{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$ and F a function of n variables $z_1, \ldots, z_n$, $\mathcal{D}^{\underline{k}} F$ is the derivative

$$\left(\frac{\partial}{\partial z_1} \right)^{k_1} \cdots \left(\frac{\partial}{\partial z_n} \right)^{k_n} F.$$

For a function F of a single variable z we write $F^{(k)}$ in place of $(d/dz)^k F$.

The notation $\underline{k}!$ stands for $k_1! \cdots k_n!$ and $\underline{z}^{\underline{k}}$ for $z_1^{k_1} \cdots z_n^{k_n}$. In Section 3 for $\underline{x} = (x_0, \ldots, x_n) \in \mathbb{C}^{n+1}$ we shall introduce also the notation

$$\mathcal{D}_{\underline{x}} = x_0 \frac{\partial}{\partial z_0} + \cdots + x_n \frac{\partial}{\partial z_n} \cdot$$

In Section 4 we shall denote by D_b the following derivation, attached to a complex number b,

$$\frac{\partial}{\partial X} + bY \frac{\partial}{\partial Y},$$

on the ring $\mathbb{C}[X, Y]$.

For n and k rational integers, the binomial coefficient

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

is considered to be 0 unless $0 \le k \le n$.

The symmetric group on $\{1, \ldots, L\}$ will be denoted by $\mathfrak{S}_L$.

Acknowledgements. Many thanks to Stéphane Fischler who carefully checked a preliminary version of the first draft and made a lot of useful comments. It is also a great pleasure to thank Francesco Amoroso, Umberto Zannier and the Fondazione C.I.M.E. (Centro Internazionale Matematico Estivo) for their invitation to deliver these lectures in the superb surroundings provided by Grand Hotel San Michele in Cetraro (Cosenza).

1 First Lecture. Introduction to Transcendence Proofs

We shall provide two proofs of the following result.

Theorem 1.1. *Let a and b be two positive integers. Then $e^b \neq a$.*

This statement is a special case of Hermite-Lindemann's Theorem:

Theorem 1.2. *Let α and β be two nonzero algebraic numbers. Then $e^\beta \neq \alpha$.*

A proof of Theorem 1.2 will be given in Section 2.

1.1 Sketch of Proof

Here are the basic ideas of both proofs of Theorem 1.1. The guest star of these proofs is the exponential function e^z. It is a transcendental function: this means that the exponential monomials $z^\tau e^{tz}$ $(\tau \geq 0,\, t \geq 0)$ are linearly independent. Consider the values at a point $z = sb$ with $s \in \mathbb{N}$:

$$\bigl(z^\tau e^{tz}\bigr)(sb) = (sb)^\tau (e^b)^{ts}.$$

If both b and e^b are integers, then this number is also a rational integer.

We need to use a special property for the number e: Theorem 1.1 would not be true if e were replaced by 2 for instance! We take derivatives of our exponential monomials. For $\sigma = 0, 1, \ldots,$

$$\left(\frac{d}{dz}\right)^\sigma \bigl(z^\tau e^{tz}\bigr) = \sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!} z^{\tau-\kappa} t^{\sigma-\kappa} e^{tz}, \tag{1.1}$$

and

$$\left(\frac{d}{dz}\right)^\sigma \bigl(z^\tau e^{tz}\bigr)(sb) = \sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!} (sb)^{\tau-\kappa} t^{\sigma-\kappa} (e^b)^{ts}.$$

These numbers again belong to the ring $\mathbb{Z}[e^b, b]$.

Starting with these numbers, there are several ways of performing the proof. We indicate two of them.

The first one rests on the construction of an auxiliary polynomial (AP)[1]. Since the functions z and e^z are algebraically independent, if $P \in \mathbb{Z}[X, Y]$ is a nonzero polynomial, the exponential polynomial $F(z) = P(z, e^z)$ is not the zero function. Together with its derivatives, it takes values in $\mathbb{Z}[e^b, b]$ at all points sb, $s \in \mathbb{N}$. Assuming b and e^b are in $\mathbb{Z}$, one wants to construct such a nonzero polynomial P for which F is the zero function, which will be a contradiction. Now the numbers $F^{(\sigma)}(sb)$ are rational integers, hence

[1] With Masser's notation in [19].

have absolute value either 0 or at least 1; this is the lower bound (LB). If P is constructed so that many numbers $F^{(\sigma)}(sb)$ have absolute value < 1, then F will have a lot of zeroes, hence (by a rigidity principle for analytic functions, called *Schwarz' Lemma*) $|F|$ will be small on a rather large disk: this is the upper bound (UB). This will enable us to deduce that $F^{(\sigma)}(sb)$ vanishes for further values of (σ, s). Once we succeed in increasing the number of known zeroes of F, there is an alternative: either we proceed by induction and extrapolate until we get so many zeroes that F has to be the zero function (for instance if all derivatives of F vanish at one point), or else we prove an auxiliary result (the *zero estimate*, or non-vanishing condition (NV)) which yields the desired conclusion.

One should say a little bit more about the initial construction of P, for which many numbers $F^{(\sigma)}(sb)$ have absolute value < 1. One solution is to select P so that the first coefficients in the Taylor expansion at the origin of F have small absolute values (see [34], Section 4.5 and [26]). Another (more classical) way is to require that many numbers $F^{(\sigma)}(sb)$ vanish. Then the existence of $P \neq 0$ is clear as soon as the number of equations we consider is smaller than the number of unknowns (the unknowns are the coefficients of P), because the conditions are linear and homogeneous. In the special case we consider here with b a real positive number, it can be proved that when the number of unknowns is the same as the number of equations, then the determinant of the system is not zero (*Pólya's Lemma* 1.6). Therefore the extrapolation can be reduced to the minimum. On the other hand for the proof of Theorem 1.1 such an argument is not required.

The second method was suggested by M. Laurent [13]: instead of solving a system of homogeneous linear equations $F^{(\sigma)}(sb) = 0$ for several values of (σ, s), consider the matrix of this system. To be more precise the matrix one considers is the one which arises from the zero estimate: if the zero estimate shows, for a given set of pairs (σ, s), that no nonzero polynomial P (with suitable bounds for its degree) can satisfy all equations $F^{(\sigma)}(sb) = 0$, then the matrix of the associated linear system has maximal rank. Consider a maximal nonsingular submatrix and its nonzero determinant $\varDelta$. The main observation of M. Laurent is that a sharp upper bound for $|\varDelta|$ can be reached by means of Schwarz' Lemma. Since $\varDelta$ lies in the ring $\mathbb{Z}[e^b, b]$, as soon as the estimate $0 < |\varDelta| < 1$ is established one deduces that one at least of the two numbers b, e^b is not a rational integer.

1.2 Tools for the Auxiliary Function

We introduce four main tools for the proof of Theorem 1.1 by means of an auxiliary function: Liouville's inequality (LB), Schwarz' Lemma (UB), the Zero Estimate (NV) and Thue-Siegel's Lemma (AP).

We shall use here only a trivial case of Liouville's inequality (a more general statement is Proposition 1.13 below):

$$\textit{For any } n \in \mathbb{Z} \textit{ with } n \neq 0 \textit{ we have } |n| \geq 1. \tag{1.2}$$

The analytic upper bound for our auxiliary function is a consequence of the following Schwarz' Lemma (see [34] Exercise 4.3, [9] and also Lemma 1.12 for a quantitative refinement):

Lemma 1.3. *Let m, $\sigma_1, \ldots, \sigma_m$ be positive integers and r, R positive real numbers with $r \leq R$. Let $\zeta_1, \ldots, \zeta_m$ be distinct elements in the disk $|\zeta| \leq R$ and $F \in \mathcal{H}_1(R)$ an analytic function which vanishes at each ζ_i with multiplicity $\geq \sigma_i$ $(1 \leq i \leq m)$. Then*

$$|F|_r \leq |F|_R \prod_{i=1}^{m} \left(\frac{R^2 + r|\zeta_i|}{R(r + |\zeta_i|)} \right)^{-\sigma_i}.$$

We recall the definition: a function F vanishes at a point ζ with multiplicity $\geq \sigma$ if $F^{(k)}(\zeta) = 0$ for $0 \leq k < \sigma$.

In our applications, we shall introduce a parameter $E > 1$ such that $R \geq Er$ and $R \geq E|\zeta_i|$ for $1 \leq i \leq m$. The conclusion yields

$$|F|_r \leq |F|_R \left(\frac{E^2 + 1}{2E} \right)^{-N}$$

where $N = \sigma_1 + \cdots + \sigma_m$ is a lower bound for the number of zeroes of F in the disk $|z| \leq R/E$. In practice N will be large, $E \geq e$, and E^{-N} will be the main term in the right hand side. In particular $|F|_R$ will not be too large, and from the conclusion of Lemma 1.3 we shall infer that $|F|_r$ is quite small.

For our first transcendence proof the zero estimate is a very simple one:

- *If F is a nonzero analytic function near z_0, there exists $\sigma \in \mathbb{N}$ such that*

$$F^{(\sigma)}(z_0) \neq 0.$$

Our last tool is Thue-Siegel's Lemma:

Lemma 1.4. *Let m and n be positive integers with $n > m$ and a_{ij} $(1 \leq i \leq n$, $1 \leq j \leq m)$ rational integers. Define*

$$A = \max\bigl\{1, \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} |a_{ij}|\bigr\}.$$

There exist rational integers $x_1, \ldots, x_n$ which satisfy

$$0 < \max\{|x_1|, \ldots, |x_n|\} \leq (nA)^{m/(n-m)}$$

and

$$\sum_{i=1}^{n} a_{ij} x_i = 0 \quad \textit{for} \quad 1 \leq j \leq m.$$

For a proof of this result, we refer for instance to [2] Lemma 1, Chap. 2, [6], Theorem 6.1, Chap. 1, Section 6.1, [7], Theorem 1.10, Chap. 1, Section 4.1 or [11], Lemma 1, Chap. VII, Section 2.

1.3 Proof with an Auxiliary Function and without Zero Estimate

Here is a first proof of Theorem 1.1. Let b be a positive integer such that e^b is also a positive integer. We denote by T_0, T_1, S_0 and S_1 positive integers which we shall choose later: during the proof we shall introduce conditions on these parameters and at the end of the proof we shall check that it is possible to select the parameters so that these conditions are satisfied. Right now let us just say that these integers will be sufficiently large.

We want to deduce from Lemma 1.4 that there exists a nonzero polynomial $P \in \mathbb{Z}[X,Y]$, of degree $< T_0$ in X and degree $< T_1$ in Y, such that the exponential polynomial $F(z) = P(z, e^z)$ has a zero of multiplicity $\geq S_0$ at each point $0, b, 2b, \ldots, (S_1 - 1)b$. If this unknown polynomial P is

$$P(X,Y) = \sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} X^\tau Y^t,$$

then the conditions

$$F^{(\sigma)}(sb) = 0 \quad (0 \leq \sigma < S_0,\ 0 \leq s < S_1)$$

can be written

$$\sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} \left(\frac{d}{dz}\right)^\sigma \bigl(z^\tau e^{tz}\bigr)(sb) = 0 \quad (0 \leq \sigma < S_0,\ 0 \leq s < S_1).$$

Finding P amounts to solving a system of $S_0 S_1$ linear equations, with rational integers coefficients, in $T_0 T_1$ unknowns $c_{\tau t}$ $(0 \leq \tau < T_0,\ 0 \leq t < T_1)$. We are going to apply Lemma 1.4 with $n = T_0 T_1$ and $m = S_0 S_1$. In place of the condition $n > m$ we shall require $n \geq 2m$, so that the so-called "Dirichlet's exponent" $m/(n-m)$ is at most 1. This yields the first main condition on our parameters:

$$\boxed{T_0 T_1 \geq 2 S_0 S_1.}$$

We need an upper bound for the number A occurring in Lemma 1.4. Consider (1.1). We wish to estimate from above the modulus of the complex number

$$\sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!} z^{\tau-\kappa} t^{\sigma-\kappa}$$

for $|z| \leq R$ with $R > 0$, and for $0 \leq t < T_1$. A first upper bound is given by

$$\sum_{\kappa=0}^{\tau} \frac{\tau!}{\kappa!(\tau-\kappa)!} \sigma^\kappa R^{\tau-\kappa} T_1^{\sigma-\kappa} = T_1^\sigma \left(\frac{\sigma}{T_1} + R\right)^\tau$$

and another one is

$$\sum_{\kappa=0}^{\sigma} \frac{\sigma!}{\kappa!(\sigma-\kappa)!} \tau^\kappa R^{\tau-\kappa} T_1^{\sigma-\kappa} = R^\tau \left(\frac{\tau}{R} + T_1\right)^\sigma.$$

Putting these estimates together yields

$$\sup_{|z|\le R} \left| \sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!} z^{\tau-\kappa} t^{\sigma-\kappa} \right| \le R^\tau T_1^\sigma \min\left\{ \left(1+\frac{\sigma}{T_1R}\right)^\tau ; \left(1+\frac{\tau}{T_1R}\right)^\sigma \right\}. \qquad (1.3)$$

Similar estimates are known for more general exponential polynomials, also in several variables: see for instance [34] Lemmas 4.9 and 13.6.

Here we shall not use the full force of this estimate. Taking $z = bs$, $R = bS_1$ we deduce

$$A \le T_1^{S_0}(bS_1 + S_0)^{T_0} e^{bT_1(S_1-1)}.$$

Hence a nonzero polynomial P exists, satisfying the required conditions and with

$$\max_{\substack{0\le\tau<T_0\\0\le t<T_1}} |c_{\tau t}| \le T_0T_1A.$$

We need an upper bound for the length

$$\mathrm{L}(P) = \sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1} |c_{\tau t}|$$

of P. As soon as T_0, T_1, S_0, S_1 are sufficiently large, we have

$$bS_1 \le \frac{1}{4}bS_0S_1, \quad S_0 \le \frac{1}{4}bS_0S_1, \quad T_0^2 \le 2^{T_0}, \quad T_1^2 \le e^{bT_1},$$

hence

$$bS_1 + S_0 \le \frac{1}{2}bS_0S_1, \quad T_0^2(bS_1+S_0)^{T_0} \le (bS_0S_1)^{T_0}$$

and therefore

$$\mathrm{L}(P) \le T_0^2T_1^2A \le T_1^{S_0}(bS_0S_1)^{T_0}e^{bT_1S_1}.$$

Since the two functions z and e^z are algebraically independent, the function F is not the zero function: at each point sb with $0 \le s < S_1$ its vanishing order is finite (and $\ge S_0$, by construction). We denote by S_0' the minimum of these orders. In other terms S_0' is the largest integer such that the conditions

$$F^{(\sigma)}(sb) = 0 \quad (0 \le \sigma < S_0',\ 0 \le s < S_1)$$

hold and therefore there is an integer s' in the range $0 \le s' < S_1$ with

$$F^{(S_0')}(s'b) \ne 0.$$

An upper bound for S_0' follows from Lemma 1.6 below, namely:

$$S_0' S_1 \leq T_0 T_1,$$

but this estimate will not be used in the present proof: we need only the lower bound $S_0' \geq S_0$.

By assumption the number $F^{(S_0')}(s'b)$ is a nonzero rational integer, hence has absolute value ≥ 1. We shall deduce from Cauchy's inequalities and Schwarz' Lemma 1.3 an upper bound for this number in terms of the parameters T_0, T_1, S_0 and S_1. It will then suffice to check that the parameters can be selected so that this upper bound is less than 1 and the contradiction will follow.

Cauchy's inequalities yield

$$\left| F^{(S_0')}(s'b) \right| \leq S_0'! |F|_r$$

for any $r \geq s'b+1$. We take $r = 2bS_1$. Next we apply Lemma 1.3 with $R = Er$, where $E > 1$ is a new parameter which we are free to choose. As we shall see a suitable choice is $E = S_0'/b$; notice that E, which is selected at this stage of the proof, is allowed to depend on S_0', while S_0' in turn depends on T_0, T_1, S_0 and S_1.

Define $m = S_1$, $\sigma_1 = \cdots = \sigma_m = S_0'$ and $\zeta_i = (i-1)b \quad (1 \leq i \leq m)$. Since $\max_{1 \leq i \leq m} |\zeta_i| \leq r$, by Lemma 1.3 we have

$$|F|_r \leq \left(\frac{E^2+1}{2E} \right)^{-S_0' S_1} |F|_R.$$

It remains to bound $|F|_R$ from above:

$$\begin{aligned} |F|_R &\leq \sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} |c_{\tau t}| \sup_{|z|=R} \left| z^\tau e^{tz} \right| \\ &\leq (T/S) T_1^{S_0} (bS_0 S_1)^{T_0} e^{bT_1 S_1} R^{T_0} e^{T_1 R} \\ &\leq T_1^{S_0} (bES_0 S_1)^{2T_0} e^{3bET_1 S_1}. \end{aligned}$$

Hence

$$|F|_r \leq (E/2)^{-S_0' S_1} T_1^{S_0} (bES_0 S_1)^{2T_0} e^{3bET_1 S_1}.$$

This explains our second main condition on the parameters: taking into account the inequality $S_0'!|F|_r \geq 1$, we shall deduce the desired contradiction as soon as we are able to check

$$\boxed{S_0'! T_1^{S_0} (bES_0 S_1)^{2T_0} e^{3bET_1 S_1} < (E/2)^{S_0' S_1}.}$$

Here is an admissible choice for these parameters. Recall that b is a fixed positive integer. We start by selecting a sufficiently large, but fixed, positive integer S_1 and we set $T_1 = S_1$. Next let S_0 be an integer, which is much larger than S_1; the required estimates below are easy to check by letting $S_0 \to \infty$. Now define $T_0 = 3S_0$ and $E = S_0'/b$. With this choice we have

$$T_0T_1 > 2S_0S_1,$$

$$S_0'! < (S_0')^{S_0'} < (E/2)^{S_0'S_1/4}$$

because $S_0' < E^2 < (E/2)^{S_1/4}$,

$$T_1^{S_0} < (E/2)^{S_0'S_1/4}$$

because $T_1 < E/2 < (E/2)^{S_1/4}$ and $S_0 \le S_0'$,

$$bES_0S_1 < (S_0')^3 \quad \text{and} \quad (S_0')^{6T_0} < (E/2)^{S_0'S_1/4},$$

and finally

$$3bET_1S_1 = 3S_0'T_1S_1 < \frac{1}{4}S_0'S_1\log(E/2).$$

This completes the proof of Theorem 1.1. □

Remark. To a certain extent this proof of Theorem 1.1 involves an extrapolation: we get more and more derivatives of F vanishing at all points $0, b, \dots, (S_1-1)b$. It is only a matter of presentation: instead of defining S_0' as we did, it amounts to the same to check by induction on $S_0' \ge S_0$ that F has a zero of multiplicity at least S_0' at each point sb with $0 \le s < S_1$. At the end of the induction we get a contradiction.

We could also extrapolate on the points at the same time as on the derivatives. Here is a variant of the proof.

We may assume[2] $b \ge 3$. Fix a large[3] positive integer N and set

$$T_0(N) = 2N^2b[\log b], \qquad T_1(N) = N^2[\log b],$$
$$S_0(N) = N^3b[\log b], \qquad S_1(N) = N[\log b]$$

and $L(N) = T_0(N)T_1(N)$, so that

$$L(N) = 2N^4b[\log b]^2 = 2S_0(N)S_1(N).$$

The first step in the preceding proof yields a nonzero polynomial $P \in \mathbb{Z}[X,Y]$ of degree $< T_0(N)$ in X and $< T_1(N)$ in Y, of length bounded by

---

[2] For the proof of Theorem 1.1, this involves no loss of generality and the only reason for this assumption is that we prefer to write $\log b$ in place of $\log_+ b$. For the same reason when we shall need to introduce $\log\log b$ later we shall assume $b \ge 16$.

[3] We assume that N is larger than some absolute constant (independent of b); here one could assume as well than N is larger than some function of b, but it turns out not to be necessary. The relevance of this fact will appear in Section 4 only.

$$\begin{aligned}\mathrm{L}(P) &\leq T_1^{S_0}(bS_0S_1)^{T_0}e^{bT_1S_1}\\ &\leq \exp\{3L(N)(\log N)/N\}\\ &\leq \exp\{L(N)/\sqrt{N}\},\end{aligned}$$

such that the function $F(z) = P(z, e^z)$ satisfies

$$F^{(\sigma)}(sb) = 0 \qquad \text{for } \ 0 \leq \sigma < S_0(N) \ \text{ and } \ 0 \leq s < S_1(N).$$

The second step is an inductive argument: we prove that for any $M \geq N$ we have

$$F^{(\sigma)}(sb) = 0 \qquad \text{for } \ 0 \leq \sigma < S_0(M) \ \text{ and } \ 0 \leq s < S_1(M). \tag{1.4}$$

This is true by construction for $M = N$. Assuming (1.4) is true for M, we deduce it for $M+1$ as follows. Let $(\sigma', s') \in \mathbb{N}^2$ satisfy $0 \leq \sigma' < S_0(M+1)$ and $0 \leq s' < S_1(M+1)$. Combining the induction hypothesis with Lemma 1.3 where we choose

$$m = S_1(M), \quad \zeta_i = (i-1)b \quad (1 \leq i \leq m), \quad \sigma_1 = \cdots = \sigma_m = S_0(M),$$
$$r = 2bS_1(M+1) \ \text{ and } \ R = 2er,$$

we deduce

$$\begin{aligned}|F|_r &\leq e^{-S_0(M)S_1(M)}|F|_R\\ &\leq e^{-S_0(M)S_1(M)}\mathrm{L}(P)R^{T_0(N)}e^{T_1(N)R}\\ &\leq e^{-S_0(M)S_1(M)/2},\end{aligned}$$

because

$$\log \mathrm{L}(P) + T_0(N)\log R + T_1(N)R \leq \frac{1}{2}S_0(M)S_1(M).$$

Since

$$\log(S_0(M+1)!) \leq \frac{1}{4}S_0(M)S_1(M),$$

Cauchy's inequalities yield

$$\left|F^{(\sigma')}(s'b)\right| < 1$$

for $0 \leq \sigma' < S_0(M+1)$ and $0 \leq s' < S_1(M+1)$. Since the left hand side is a rational integer, we deduce $F^{(\sigma')}(s'b) = 0$ and the inductive argument follows.

Plainly we conclude $F = 0$, which completes this new proof of Theorem 1.1.

Remark. In this inductive argument from M to $M+1$, the first step (with $M = N$) is the hardest one: as soon as M is large with respect to N, the required estimates are easier to check.

1.4 Tools for the Interpolation Determinant Method

Some tools which have already been introduced above will be needed for the proof involving interpolation determinants. For instance Liouville's inequality is just the same (1.2). On the other hand in place of Schwarz' Lemma 1.3 we shall use M. Laurent's fundamental observation that interpolation determinants have a small absolute value (see [13], Section 6.3, Lemma 3 and [14], Section 6, Lemma 6). The following estimate ([34] Lemma 2.8) is a consequence of the case $m = 1$, $\zeta_1 = 0$ of Lemma 1.3.

Lemma 1.5. *Let $\varphi_1, \ldots, \varphi_L$ be entire functions in $\mathbb{C}$, $\zeta_1, \ldots, \zeta_L$ elements of $\mathbb{C}$, $\sigma_1, \ldots, \sigma_L$ nonnegative integers and $0 < r \leq R$ real numbers, with $|\zeta_\mu| \leq r$ $(1 \leq \mu \leq L)$. Then the absolute value of the determinant*

$$\Delta = \det\left(\varphi_\lambda^{(\sigma_\mu)}(\zeta_\mu) \right)_{1 \leq \lambda, \mu \leq L}$$

is bounded from above by

$$|\Delta| \leq \left(\frac{R}{r}\right)^{-(L(L-1)/2)+\sigma_1+\cdots+\sigma_L} L! \prod_{\lambda=1}^{L} \max_{1 \leq \mu \leq L} \sup_{|z|=R} \left| \varphi_\lambda^{(\sigma_\mu)}(z) \right|.$$

The zero estimate we need is the following result due to G. Pólya ([34] Corollary 2.3):

Lemma 1.6. *Let $w_1, \ldots, w_n$ be pairwise distinct real numbers, $x_1, \ldots, x_m$ also pairwise distinct real numbers and $\tau_1, \ldots, \tau_n$, $\sigma_1, \ldots, \sigma_m$ nonnegative integers, with*

$$\tau_1 + \cdots + \tau_n = \sigma_1 + \cdots + \sigma_m.$$

Choose any ordering for the pairs (τ, i) with $0 \leq \tau < \tau_i$ and $1 \leq i \leq n$ and any ordering for the pairs (σ, j) with $0 \leq \sigma < \sigma_j$ and $1 \leq j \leq m$. Then the square matrix

$$\left(\left(\frac{d}{dz}\right)^\sigma \left(z^\tau e^{w_i z}\right)(x_j) \right)_{\substack{(\tau,i) \\ (\sigma,j)}}$$

is nonsingular.

We call this result a *zero estimate* because it can be stated as follows: if $c_{\tau i}$ are complex numbers ($0 \leq \tau < \tau_i$, $1 \leq i \leq n$), not all of which are zero, then the exponential polynomial

$$f(z) = \sum_{i=1}^{n} \sum_{\tau=0}^{\tau_i - 1} c_{\tau i} z^\tau e^{w_i z}$$

cannot vanish at each x_j with multiplicity $\geq \sigma_j$ $(1 \leq j \leq m)$.

One main characteristic of Laurent's interpolation determinant method is that there is no need of Thue-Siegel's Lemma 1.4.

1.5 Proof with an Interpolation Determinant and a Zero Estimate

Here is another proof of Theorem 1.1.

We start with a positive real number $b \geq 1$, without any other assumption. We introduce auxiliary parameters T_0, T_1, S_0 and S_1, which are positive integers and $E > 1$ a real number. These parameters will be specified later, but it is convenient to assume T_0, T_1, S_0 and S_1 are all ≥ 2.

Consider the matrix

$$\mathtt{M} = \left(\left(\frac{d}{dz}\right)^{\sigma} \left(z^{\tau} e^{tz}\right)(sb) \right)_{\substack{0 \leq \tau < T_0,\ 0 \leq t < T_1 \\ 0 \leq \sigma < S_0,\ 0 \leq s < S_1}},$$

with $T_0 T_1$ rows labeled with (τ, t) and $S_0 S_1$ columns labeled with (σ, s). Here we shall work with a square matrix, which means that we require

$$\boxed{T_0 T_1 = S_0 S_1.}$$

We denote by L this number, so that $\mathtt{M}$ is a square $L \times L$ matrix. By Lemma 1.6 with $n = T_1$,

$$w_i = i - 1, \quad \tau_i = T_0 \qquad (1 \leq i \leq n),$$

$m = S_1$ and

$$x_j = (j-1)b, \quad \sigma_j = S_0 \qquad (1 \leq j \leq m),$$

it follows that $\mathtt{M}$ is nonsingular. Let Δ be the determinant of $\mathtt{M}$. By Lemma 1.5 with

$$\begin{aligned} \{\varphi_1, \ldots, \varphi_L\} &= \{z^{\tau} e^{tz}\ ;\ 0 \leq \tau < T_0,\ 0 \leq t < T_1\}, \\ \{(\sigma_1, \zeta_1), \ldots, (\sigma_L, \zeta_L)\} &= \{(\sigma, sb)\ ;\ 0 \leq \sigma < S_0,\ 0 \leq s < S_1\} \end{aligned}$$

and $r = bS_1$, $R = Er$, we have

$$|\Delta| \leq E^{-L(L-1-S_0)/2} L! \prod_{\tau=0}^{T_0-1} \prod_{t=0}^{T_1-1} \max_{0 \leq \sigma < S_0} \sup_{|z|=R} \left| \left(\frac{d}{dz}\right)^{\sigma} \left(z^{\tau} e^{tz}\right) \right|.$$

Since, for $0 \leq \sigma < S_0$ and $|z| = R$, we have by (1.3) and (1.1)

$$\begin{aligned} \left| \left(\frac{d}{dz}\right)^{\sigma} \left(z^{\tau} e^{tz}\right) \right| &\leq T_1^{S_0} (R + S_0)^{T_0} e^{T_1 R} \\ &\leq T_1^{S_0} (bES_0S_1)^{T_0} e^{bET_1S_1}, \end{aligned}$$

we deduce

$$\begin{aligned} |\Delta| &\leq E^{-L(L-1-S_0)/2} L! T_1^{LS_0} (bES_0S_1)^{LT_0} e^{bELT_1S_1} \\ &\leq E^{-L^2/2} L^L (ET_1)^{LS_0} (bES_0S_1)^{LT_0} e^{bELT_1S_1}. \end{aligned}$$

This estimate holds unconditionally. If we can select our parameters so that $|\Delta| < 1$, then this will prove that the nonzero number Δ cannot be a rational integer, hence one at least of b and e^b is not a rational integer. Therefore the proof of Theorem 1.1 will be completed if we show that our parameters may be selected so that

$$\boxed{E^{L/2} > L(ET_1)^{S_0}(bES_0S_1)^{T_0}e^{bET_1S_1}.}$$

Here is an admissible choice: let N be a sufficiently large positive integer (independent of b). Assuming $b \geq 3$, define $E = e$,

$$T_0 = N^2 b[\log b], \qquad T_1 = N^2[\log b],$$
$$S_0 = N^3 b[\log b], \qquad S_1 = N[\log b],$$

so that $L = N^4 b[\log b]^2$.

This completes the proof of Theorem 1.1. $\square$

1.6 Remarks

In this last proof of Theorem 1.1, we did not need to assume that b and e^b are integers: Liouville's inequality (1.2) is used at the very end and provides the conclusion. More precisely it is plain that the interpolation determinant method of Section 1.5 yields the following explicit result.

Proposition 1.7. *Let b be a positive real number. Let T_0, T_1, S_0, S_1 and L be positive integers satisfying*

$$L = T_0T_1 = S_0S_1.$$

Let E be a positive number, $E > 1$. Then there exists a polynomial $f \in \mathbb{Z}[Z_1, Z_2]$, of degree $< LT_1S_1$ in Z_1 and $< LT_0$ in Z_2, of length bounded by

$$\mathrm{L}(f) \leq L!T_1^{LS_0}(S_0S_1)^{LT_0},$$

such that

$$0 < |f(e^b, b)| \leq E^{-L^2/2}L!(ET_1)^{LS_0}(|b|_+ES_0S_1)^{LT_0}e^{|b|ELT_1S_1}.$$

We now explain how to modify the first proof (in Section 1.3) involving an auxiliary function and deduce the following variant of Proposition 1.7.

Proposition 1.8. *Let b be a positive real number. Let T_0, T_1, S_0, S_1 and L be positive integers such that*

$$L = T_0T_1 = S_0S_1.$$

Let E, U, V, W be positive real numbers satisfying

$$E \geq e, \quad W \geq 12\log E,$$

$$U \geq \log L + T_0 \log(|b|_+ES_1) + |b|_+ET_1S_1 \tag{1.5}$$

and

$$4(U+V+W)^2 \leq LW \log E. \tag{1.6}$$

There exists a nonzero polynomial $f \in \mathbb{Z}[Z_1, Z_2]$ *of degree* $< T_1S_1$ *in* Z_1 *and* $< T_0$ *in* Z_2*, of length*

$$\mathrm{L}(f) \leq Le^W T_1^{S_0}(S_0+S_1)^{T_0},$$

such that

$$0 < |f(e^b, b)| \leq S_0! e^{-V}.$$

Remark. It is interesting to compare the two estimates provided by Propositions 1.7 and 1.8. Condition (1.6) is satisfied with[4]

$$U = V = W = \frac{1}{36} L \log E.$$

Up to terms of smaller order (when T_0, T_1, S_0 and S_1 are all sufficiently large), the estimates one deduces from Proposition 1.7 for the degrees, the logarithm of the length and the logarithm of the absolute value are L times the corresponding ones in Proposition 1.8.

For all practical purposes, Proposition 1.8, which is obtained by the auxiliary function method, is much sharper that Proposition 1.7. This fact has been an obstacle during a while to develop the interpolation determinant method. For instance it took several years before proofs of algebraic independence results could be achieved by means of Laurent's interpolation determinant method. A nice solution has been provided by M. Laurent and D. Roy in [15], who point out that the polynomial f given by the proof of Proposition 1.7 has a further quite interesting property: its first derivatives

$$\left(\frac{\partial}{\partial Z_1}\right)^{k_1} \left(\frac{\partial}{\partial Z_2}\right)^{k_2} f$$

with $(k_1, k_2) \in \mathbb{N}^2$ satisfying, say, $k_1 + k_2 < L/2$, also have a small absolute value at the point (e^b, b). We shall develop this argument later (see Theorem 4.5).

Our proof of Proposition 1.8 uses an auxiliary function. Since we do not assume b and e^b are integers, we cannot apply Lemma 1.4 as we did in Section 1.3. Instead of solving linear equations, we select (again by means of Dirichlet's box principle) the coefficients $c_{\tau t}$ of the auxiliary polynomial P so that a set of inequalities is satisfied. There are several possibilities. Here we shall use the following auxiliary function.

---

[4] This choice yields a weak upper bound for the length of f. From this point of view a better choice is for instance $U = V = (1/20N)L \log E$ and $W = U/N$ with $N \geq 5$.

Lemma 1.9. *Let L be a positive integer, U, V, W, R, r positive real numbers and $\varphi_1,\dots,\varphi_L$ functions in $\mathcal{H}_1(R)$. Assume*

$$U+V+W\ge 12,\quad e\le \frac{R}{r}\le e^{(U+V+W)/6},\quad \sum_{\lambda=1}^{L}|\varphi_\lambda|_R\le e^U$$

and

$$4(U+V+W)^2\le LW\log(R/r).$$

Then there exist rational integers $p_1,\dots,p_L$, with

$$0<\max_{1\le\lambda\le L}|p_\lambda|\le e^W,$$

such that the function $F=p_1\varphi_1+\cdots+p_L\varphi_L$ satisfies

$$|F|_r\le e^{-V}.$$

We do not give the proof of Lemma 1.9 (see [34] Proposition 4.10, which provides a similar statement in several variables). It suffices to say that it combines Dirichlet's box principle (Lemma 1.10) with an interpolation formula (Lemma 1.11).

Here is Lemma 4.12 of [34].

Lemma 1.10. *Let ν, μ, X be positive integers, U, V positive real numbers and u_{ij} $(1\le i\le\nu,\ 1\le j\le\mu)$ complex numbers. Assume*

$$\sum_{i=1}^{\nu}|u_{ij}|\le e^U,\qquad (1\le j\le\mu)$$

and

$$\left(\sqrt{2}Xe^{U+V}+1\right)^{2\mu}\le (X+1)^\nu.$$

Then there exists $(\xi_1,\dots,\xi_\nu)\in\mathbb{Z}^\nu$ satisfying

$$0<\max_{1\le i\le\nu}|\xi_i|\le X$$

and

$$\max_{1\le j\le\mu}\left|\sum_{i=1}^{\nu}u_{ij}\xi_i\right|\le e^{-V}.$$

The next result is Lemma 4.13 of [34] (in case $r=0$ we agree that $r^{\|\underline{k}\|}=1$ for $\underline{k}=\underline{0}$). For the proof of Lemma 1.9, the case $n=1$ suffices, but we shall need the general case in Section 4.4.

Lemma 1.11. *Let n, K be positive integers, r and R real numbers satisfying $0\le r<R$ and F an entire function in $\mathbb{C}^n$. Then*

$$|F|_r\le(1+\sqrt{K})\left(\frac{r}{R}\right)^K|F|_R+\sum_{\|\underline{k}\|<K}|\mathcal{D}^{\underline{k}}F(0)|\frac{r^{\|\underline{k}\|}}{\underline{k}!}\cdot$$

Proof of Proposition 1.8. We apply Lemma 1.9 to the functions

$$\{\varphi_1,\dots,\varphi_L\}=\{z^\tau e^{tz}\ ;\ 0\le\tau<T_0,\ 0\le t<T_1\}$$

with

$$r=|b|_+S_1,\quad R=Er.$$

From hypothesis (1.5) we derive

$$\sum_{\lambda=1}^{L}|\varphi_\lambda|_R=\sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1}\sup_{|z|=R}\bigl|z^\tau e^{tz}\bigr|\le LR^{T_0}e^{T_1R}$$
$$\le L(|b|_+ES_1)^{T_0}e^{|b|_+ET_1S_1}\le e^U.$$

We deduce the existence of a nonzero polynomial

$$P(X,Y)=\sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1}c_{\tau t}X^\tau Y^t\in\mathbb{Z}[X,Y],$$

of degree $< T_0$ in X and $< T_1$ in Y, with integer coefficients bounded in absolute value by e^W, such that the function $F(z)=P(z,e^z)$ satisfies

$$|F|_r\le e^{-V}.$$

By Lemma 1.6 there is a nonzero element γ in the set

$$\bigl\{F^{(\sigma)}(sb)\ ;\ 0\le\sigma<S_0,\ 0\le s<S_1\bigr\}.$$

From Cauchy's inequality, and since $r\ge s|b|+1$, we deduce the upper bound

$$|\gamma|\le S_0!|F|_r\le S_0!e^{-V}.$$

Writing

$$\gamma=F^{(\sigma)}(sb),$$

define $f\in\mathbb{Z}[Z_1,Z_2]$ by

$$\sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1}c_{\tau t}\sum_{\kappa=0}^{\min\{\tau,\sigma\}}\frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!}s^{\tau-\kappa}t^{\sigma-\kappa}Z_1^{ts}Z_2^{\tau-\kappa}$$

so that, using (1.1), we can write

$$\gamma=f(e^b,b).$$

The degrees of f plainly satisfy the required conditions in the conclusion of Proposition 1.8, and finally the length of f is bounded thanks to (1.3):

$$\mathrm{L}(f)\le\sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1}|c_{\tau t}|\sum_{\kappa=0}^{\min\{\tau,\sigma\}}\frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!}s^{\tau-\kappa}t^{\sigma-\kappa}$$
$$\le Le^WT_1^{S_0}(S_0+S_1)^{T_0}.$$

□

One can prove a variant of Proposition 1.8 by constructing the auxiliary function $F(z) = P(z, e^z)$ in a slightly different way. One applies Lemma 1.10 again, but now we require that many values at points sb of the function F and of its first derivatives have a small absolute value[5]. A rigidity principle for analytic functions (Lemma 1.12) enables us to deduce that $|F|_r$ is rather small for a suitable parameter r. We are back to the situation of our first proof: we invoke Pólya's Lemma 1.6 and produce a nonzero value of F (or of one of its derivatives). This nonzero number is the value at the point (e^b, b) of a polynomial $f \in \mathbb{Z}[Z_1, Z_2]$ which satisfies the desired conclusion.

Lemma 1.11 is quite simple, since only one point $z = 0$ is involved. For functions of a single variable one can consider an arbitrary finite set of points[6]. Here is Lemma 5.1 of [26].

Lemma 1.12. *Let ℓ be a positive integer, $w_1, \ldots, w_\ell$ pairwise distinct complex numbers and $m_1, \ldots, m_\ell$ positive integers. Put*

$$L = \sum_{j=1}^{\ell} m_j, \qquad \varrho = \max_{1 \le j \le \ell} \max\bigl\{1, |w_j|\bigr\},$$

and

$$\delta_1 = \min_{1 \le j \le \ell} \prod_{\substack{1 \le j' \le \ell \\ j' \ne j}} |w_j - w_{j'}|^{m_{j'}/L},$$

$$\delta_2 = \min\bigl\{1, \min_{\substack{1 \le j, j' \le \ell \\ j \ne j'}} |w_j - w_{j'}|^{m_{j'}/L}\bigr\},$$

with the convention that $\delta_1 = \delta_2 = 1$ when $\ell = 1$. Then, for any pair of real numbers r and R with $R \ge 2r$ and $r \ge 2\varrho$ and for any function $F \in \mathcal{H}_1(R)$, we have

$$|F|_r \le \left(\frac{6r}{\delta_1 \delta_2}\right)^L \max_{\substack{1 \le j \le \ell \\ 0 \le \kappa < m_j}} \frac{1}{\kappa!} \left|F^{(\kappa)}(w_j)\right| + \left(\frac{6r}{R}\right)^L |F|_R.$$

Each of the two Propositions 1.7 and 1.8 yields the real case of Hermite-Lindemann's Theorem 1.2: in place of the trivial Liouville's inequality (1.2) we have used so far, it suffices to invoke the next result, which is Proposition 3.14 of [34]:

Proposition 1.13. (Liouville's Inequality). *Let $\mathbb{K}$ be a number field of degree D, v an Archimedean absolute value of $\mathbb{K}$ and $\nu_1, \ldots, \nu_\ell$ positive integers. For*

---

[5] Lemma 1.9 is proved in [34] by constructing P so that the first Taylor coefficients of F at the origin have a small absolute value; hence it may be considered as a variant of this approach, which consists in taking only $s = 0$ at this stage of the proof - the number b does not occur in this case.

[6] Such a statement is called an "approximate Schwarz' Principle" in [23], Section 3.a.

$1 \le i \le \ell$, let $\gamma_{i1}, \ldots, \gamma_{i\nu_i}$ be elements of $\mathbb{K}$. Further, let f be a polynomial in $\nu_1 + \cdots + \nu_\ell$ variables, with coefficients in $\mathbb{Z}$, which does not vanish at the point $\underline{\gamma} = \bigl(\gamma_{ij}\bigr)_{1\le j\le \nu_i, 1\le i\le \ell}$. Assume f has total degree at most N_i with respect to the ν_i variables corresponding to $\gamma_{i1}, \ldots, \gamma_{i\nu_i}$. Then

$$\log|f(\underline{\gamma})|_v \ge -(D-1)\log \mathrm{L}(f) - D\sum_{i=1}^{\ell} N_i \mathrm{h}(1\colon \gamma_{i1}\colon \,\cdots\, \colon \gamma_{i\nu_i}).$$

Finally the complex case of Hermite-Lindemann's Theorem 1.2 can also be proved easily by the same arguments, either with an auxiliary function or with an interpolation determinant. The only new feature is to replace Pólya's Lemma 1.6 by another zero estimate, for instance Lemma 4.3. We refer to [2], Chap. 1, Section 3, [6], Chap. 1, Section 9, [7], Chap. 2, Section 2 and [30], Section 3.1 for proofs of the Hermite-Lindemann's Theorem by means of an auxiliary function and to [34] Chap. 2 for the interpolation determinant method with a zero estimate.

In the next section we provide a new proof of Hermite-Lindemann's Theorem 1.2 by means of an interpolation determinant but without any zero estimate: we shall extrapolate like in Section 1.3.

2 Second Lecture. Extrapolation with Interpolation Determinants

The proof given in Section 1.3 rests on an auxiliary function and involves an extrapolation; this extrapolation enabled us to conclude without using the zero estimate Lemma 1.6. We explain here how to perform an extrapolation by means of the interpolation determinant method of Section 1.5.

2.1 Upper Bound for a Determinant in a Single Variable

We are looking for an upper bound for an interpolation determinant. Lemma 1.5 is proved by M. Laurent in [14], Section 6 (also in [34] Lemma 2.8) by means of Schwarz' Lemma 1.3 for the function

$$\varPhi : z \longmapsto \det\Bigl(\varphi_\lambda^{(\sigma_\mu)}(z\zeta_\mu)\Bigr)_{1\le\lambda,\mu\le L}$$

which has a zero at the origin of multiplicity at least

$$0 + 1 + \cdots + (L-1) - (\sigma_1 + \cdots + \sigma_L) = \frac{L(L-1)}{2} - \sum_{\mu=1}^{L}\sigma_\mu.$$

Example 2.1. (See Masser's Lecture 1 in [19]). Define

$$\varphi(z) = z + z^2 + z^4 + z^8 + \cdots = \sum_{m=0}^{\infty} z^{2^m}.$$

Set $L = 6$ and take for $\varphi_1, \ldots, \varphi_6$ the functions

$$1,\ z,\ \varphi(z),\ z\varphi(z),\ \varphi^2(z),\ z\varphi^2(z).$$

Further set

$$\sigma_1 = 0,\ \sigma_2 = 1,\ \sigma_3 = 2,\ \sigma_4 = 3,\ \sigma_5 = 4,\ \sigma_6 = 0$$

and

$$\zeta_1 = \zeta_2 = \zeta_3 = \zeta_4 = \zeta_5 = 0,\ \zeta_6 = 1.$$

The function

$$\begin{aligned}
\Phi(z) \prod_{\mu=1}^{6} \frac{1}{\sigma_\mu !} &= \det\left(\frac{1}{\sigma_\mu !} \varphi_\lambda^{(\sigma_\mu)}(z\zeta_\mu) \right)_{1 \le \lambda,\mu \le L} \\
&= \det \begin{pmatrix} 1\ 0\ 0\ 0\ 0 & 1 \\ 0\ 1\ 0\ 0\ 0 & z \\ 0\ 1\ 1\ 0\ 1 & \varphi(z) \\ 0\ 0\ 1\ 1\ 0 & z\varphi(z) \\ 0\ 0\ 1\ 2\ 1 & \varphi^2(z) \\ 0\ 0\ 0\ 1\ 2 & z\varphi^2(z) \end{pmatrix} \\
&= 2z\varphi^2(z) + 4z\varphi(z) - 3\varphi^2(z) + z - \varphi(z)
\end{aligned}$$

has a zero of multiplicity ≥ 5 at the origin; as pointed out by D.W. Masser [19], actually the multiplicity is 6.

Back to the general case, we need to take into account further zeroes. Such an upper bound is given in Corollary 2.4 of [33]; the proof relies on a Schwarz Lemma for Cartesian products (see [33] Proposition 2.3; see also [9] for a general discussion of this issue). Philippon ([23] Lemme 4) also gave upper estimates for interpolation determinants and he does not need to deal with Cartesian products: he uses a much more simple inductive argument which suffices for interpolation determinants (but does not seem to extend to Cartesian products). Here we follow his approach.

We first combine Schwarz' Lemma 1.3 with Cauchy's inequalities.

Lemma 2.2. *Let R, r, ϱ and E be positive real numbers, F an element of $\mathcal{H}_1(R)$, m a positive integer, $\zeta_1, \ldots, \zeta_m$ pairwise distinct complex numbers, ξ a complex number and κ, $\sigma_1, \ldots, \sigma_m$ nonnegative integers. Set*

$$N = \sigma_1 + \cdots + \sigma_m.$$

Assume

$$R \geq \max\{r,\, |\xi| + \varrho\}, \quad r \geq \max_{1 \leq i \leq m} |\zeta_i| \quad \text{and} \quad 1 \leq E \leq \frac{R^2 + r(|\xi| + \varrho)}{R(r + |\xi| + \varrho)}.$$

Assume also that F satisfies

$$F^{(\sigma)}(\zeta_i) = 0 \quad \text{for} \quad 0 \leq \sigma < \sigma_i \quad \text{and} \quad 1 \leq i \leq m.$$

Then

$$\left| F^{(\kappa)}(\xi) \right| \leq \kappa! \varrho^{-\kappa} E^{-N} |F|_R.$$

Proof. By Cauchy's inequalities

$$\left| F^{(\kappa)}(\xi) \right| \leq \kappa! \varrho^{-\kappa} |F|_{|\xi|+\varrho}.$$

Since F has at least N zeroes (counting multiplicities) in the disk $B_1(0, r)$, we deduce from Schwarz' Lemma 1.3:

$$|F|_{|\xi|+\varrho} \leq E^{-N} |F|_R.$$

Hence the result. $\square$

The next result is an extension of Corollary 2.4 of [33] where we include multiplicities.

Proposition 2.3. *Let R be a positive real number, $\varphi_1, \ldots, \varphi_L$ elements of $\mathcal{H}_1(R)$ with $L \geq 1$, $\xi_1, \ldots, \xi_L$ complex numbers in $B_1(0, R)$ and $\kappa_1, \ldots, \kappa_L$ nonnegative integers. Consider the determinant*

$$\Delta = \det\left(\varphi_\lambda^{(\kappa_\mu)}(\xi_\mu) \right)_{1 \leq \lambda, \mu \leq L}.$$

Further let $m_1, \ldots, m_L$ be nonnegative integers and, for $1 \leq \mu \leq L$ and $1 \leq i \leq m_\mu$, let $\zeta_{\mu i}$ be a complex number and $\sigma_{\mu i}$ a nonnegative integer. We assume that for each $\mu = 1, \ldots, L$, the m_μ numbers $\zeta_{\mu 1}, \ldots, \zeta_{\mu m_\mu}$ are pairwise distinct. Set

$$N_\mu = \sum_{i=1}^{m_\mu} \sigma_{\mu i} \quad (1 \leq \mu \leq L).$$

For $1 \leq \mu \leq L$, let r_μ, R_μ, ϱ_μ and E_μ be positive real numbers satisfying

$$R \geq R_\mu \geq \max\{r_\mu,\, |\xi_\mu| + \varrho_\mu\}, \quad r_\mu \geq \max_{1 \leq i \leq m_\mu} |\zeta_{\mu i}|$$

and

$$1 \leq E_\mu \leq \frac{R_\mu^2 + r_\mu(|\xi_\mu| + \varrho_\mu)}{R_\mu(r_\mu + |\xi_\mu| + \varrho_\mu)}.$$

Denote by Φ the analytic mapping

$$(\varphi_1, \ldots, \varphi_L)\colon B_1(0,R) \to \mathbb{C}^L.$$

Assume that for any $(\mu, i, \kappa) \in \mathbb{N}^3$ satisfying $1 \le \mu \le L$, $1 \le i \le m_\mu$ and $0 \le \kappa < \sigma_{\mu i}$, the μ vectors

$$\Phi^{(\kappa_1)}(\xi_1), \ldots, \Phi^{(\kappa_{\mu-1})}(\xi_{\mu-1}),\ \Phi^{(\kappa)}(\zeta_{\mu i}) \tag{2.1}$$

in $\mathbb{C}^L$ are linearly dependent. Then

$$|\Delta| \le L! \max_{\tau \in \mathfrak{S}_L} \prod_{\mu=1}^{L} \Bigl(\kappa_\mu! \varrho_\mu^{-\kappa_\mu} E_\mu^{-N_\mu} |\varphi_{\tau(\mu)}|_{R_\mu} \Bigr).$$

Proof. We prove Proposition 2.3 by induction on L. For $L = 1$ we have $\Phi = \varphi_1$,

$$\Delta = \Phi^{(\kappa_1)}(\xi_1), \qquad 1 \le E_1 \le \frac{R_1^2 + r_1(|\xi_1| + \varrho_1)}{R_1(r_1 + |\xi_1| + \varrho_1)}$$

and hypothesis (2.1) reads

$$\Phi^{(\kappa)}(\zeta_{1i}) = 0 \quad \text{for} \quad 0 \le \kappa < \sigma_{1i} \quad \text{and} \quad 1 \le i \le m_1.$$

From Lemma 2.2 we deduce

$$|\Delta| \le \kappa_1! \varrho_1^{-\kappa_1} E_1^{-N_1} |\varphi_1|_{R_1}.$$

Hence Proposition 2.3 is true in case $L = 1$.

Assume now that the conclusion is true for L replaced by $L - 1$. Define $F \in B_1(0, R)$ by

$$F(z) = \det\Bigl(\Phi^{(\kappa_1)}(\xi_1), \ldots, \Phi^{(\kappa_{L-1})}(\xi_{L-1}),\ \Phi(z)\Bigr).$$

By assumption (2.1) with $\mu = L$, for $1 \le i \le m_L$ and $0 \le \kappa < \sigma_{Li}$, we have

$$F^{(\kappa)}(\zeta_{Li}) = 0.$$

Since

$$\Delta = F^{(\kappa_L)}(\xi_L),$$

we deduce from Lemma 2.2

$$|\Delta| \le \kappa_L! \varrho_L^{-\kappa_L} E_L^{-N_L} |F|_{R_L}.$$

We expand the determinant F with respect to the last column: define, for $1 \le \lambda \le L$,

$$\Phi_\lambda = \bigl(\varphi_1, \ldots, \varphi_{\lambda-1}, \varphi_{\lambda+1}, \ldots, \varphi_L\bigr)\colon B_1(0,R) \to \mathbb{C}^{L-1}$$

and let Δ_λ denote the determinant of the $(L-1)\times(L-1)$ matrix

$$\left(\Phi_\lambda^{(\kappa_1)}(\xi_1),\ldots,\Phi_\lambda^{(\kappa_{L-1})}(\xi_{L-1})\right),$$

so that

$$F(z)=\sum_{\lambda=1}^{L}(-1)^{L-\lambda}\varphi_\lambda(z)\Delta_\lambda.$$

Hence

$$|F|_{R_L}\le L\max_{1\le\lambda\le L}|\varphi_\lambda|_{R_L}|\Delta_\lambda|$$

and therefore

$$|\Delta|\le\kappa_L!L\varrho_L^{-\kappa_L}E_L^{-N_L}\max_{1\le\lambda\le L}|\varphi_\lambda|_{R_L}|\Delta_\lambda|.$$

We fix an index $\lambda^0\in\{1,\ldots,L\}$ such that

$$|\Delta|\le\kappa_L!L\varrho_L^{-\kappa_L}E_L^{-N_L}|\varphi_{\lambda^0}|_{R_L}|\Delta_{\lambda^0}|.$$

Using the induction hypothesis, we deduce that there exists a bijective map t from $\{1,\ldots,L-1\}$ to $\{1,\ldots,\lambda^0-1,\lambda^0+1,\ldots,L\}$ such that

$$|\Delta_{\lambda^0}|\le(L-1)!\prod_{\mu=1}^{L-1}\left(\kappa_\mu!\varrho_\mu^{-\kappa_\mu}E_\mu^{-N_\mu}|\varphi_{t(\mu)}|_{R_\mu}\right).$$

Define $\tau\in\mathfrak{S}_L$ by $\tau(\mu)=t(\mu)$ for $1\le\mu<L$ and $\tau(L)=\lambda^0$. Proposition 2.3 follows. □

We shall use a special case of Proposition 2.3.

We consider a finite sequence $(\zeta_0,\ldots,\zeta_N)$ of complex numbers, which are not supposed to be pairwise distinct. We define the associated *multiplicity sequence* $(\sigma_0,\ldots,\sigma_N)$ as follows:

$$\sigma_\nu=\mathrm{Card}\{i;\ 0\le i<\nu\,,\zeta_i=\zeta_\nu\}\quad(0\le\nu\le N).$$

If $\zeta_0,\ldots,\zeta_N$ are pairwise distinct then $\sigma_0=\cdots=\sigma_N=0$. In general, $(\zeta_0,\ldots,\zeta_N)$ consists of ℓ distinct complex numbers $w_1,\ldots,w_\ell$, where w_j is repeated m_j times $(1\le j\le\ell)$, so that

$$\prod_{\nu=0}^{N}(z-\zeta_\nu)=\prod_{j=1}^{\ell}(z-w_j)^{m_j}\quad\text{and}\quad m_1+\cdots+m_\ell=N+1.$$

Then for an analytic function F the $N+1$ equations

$$F^{(\sigma_\nu)}(\zeta_\nu)=0\quad\text{for}\quad 0\le\nu\le N$$

are nothing else than

$$F^{(\kappa)}(w_j) = 0 \quad \text{for} \quad 0 \le \kappa < m_j \quad \text{and} \quad 1 \le j \le \ell.$$

The pairs (σ_ν, ζ_ν) $(0 \le \nu \le N)$ are pairwise distinct and for each $\nu = 0, \ldots, N$ and each $\sigma = 0, \ldots, \sigma_\nu$ there is an index μ in the range $0 \le \mu \le \nu$ with $(\sigma_\mu, \zeta_\mu) = (\sigma, \zeta_\nu)$.

Corollary 2.4. *Let L, N be integers with $1 \le L \le N+1$, R a positive real number and $(\zeta_0, \ldots, \zeta_N)$ a sequence of $N+1$ elements in $B_1(0,R)$. Denote by $(\sigma_0, \ldots, \sigma_N)$ the associated multiplicity sequence. Let $0 = \nu_0 \le \nu_1 < \cdots < \nu_L \le N$ be integers and $\varphi_1, \ldots, \varphi_L$ elements of $\mathcal{H}_1(R)$. Consider the determinant*

$$\Delta = \det\left(\varphi_\lambda^{(\sigma_{\nu_\mu})}(\zeta_{\nu_\mu})\right)_{1 \le \lambda, \mu \le L}.$$

For $1 \le \mu \le L$, let r_μ, R_μ, ϱ_μ and E_μ be positive real numbers satisfying

$$R \ge R_\mu \ge \max\{r_\mu, |\zeta_{\nu_\mu}| + \varrho_\mu\}, \quad r_\mu \ge \max_{0 \le \nu < \nu_\mu} |\zeta_\nu|$$

and

$$1 \le E_\mu \le \frac{R_\mu^2 + r_\mu(|\zeta_{\nu_\mu}| + \varrho_\mu)}{R_\mu(r_\mu + |\zeta_{\nu_\mu}| + \varrho_\mu)}.$$

Denote by Φ the analytic mapping

$$(\varphi_1, \ldots, \varphi_L) \colon B_1(0,R) \to \mathbb{C}^L.$$

Assume that for any $(\mu,\nu) \in \mathbb{N}^2$ satisfying $1 \le \mu \le L$ and $\nu_{\mu-1} \le \nu < \nu_\mu$, the μ vectors

$$\Phi^{(\sigma_{\nu_1})}(\zeta_{\nu_1}), \ldots, \Phi^{(\sigma_{\nu_{\mu-1}})}(\zeta_{\nu_{\mu-1}}),\ \Phi^{(\sigma_\nu)}(\zeta_\nu)$$

in $\mathbb{C}^L$ are linearly dependent. Then

$$|\Delta| \le L! \max_{\tau \in \mathfrak{S}_L} \prod_{\mu=1}^{L} \left(\sigma_{\nu_\mu}! \varrho_\mu^{-\sigma_{\nu_\mu}} E_\mu^{-\nu_\mu} |\varphi_{\tau(\mu)}|_{R_\mu}\right).$$

Proof. We apply Proposition 2.3 with

$$\xi_\mu = \zeta_{\nu_\mu} \quad \text{and} \quad \kappa_\mu = \sigma_{\nu_\mu} \quad (1 \le \mu \le L).$$

We define m_μ, $\zeta_{\mu i}$ and $\sigma_{\mu i}$ as follows: for $1 \le \mu \le L$, we denote by m_μ the number of distinct elements in the sequence $(\zeta_0, \ldots, \zeta_{\nu_\mu - 1})$, by $\zeta_{\mu i}$ these distinct elements and by $\sigma_{\mu i}$ the number of ν in the range $0 \le \nu < \nu_\mu$ such that $\zeta_\nu = \zeta_{\mu i}$. Therefore

$$N_\mu = \sigma_{\mu 1} + \cdots + \sigma_{\mu m_\mu} = \nu_\mu.$$

□

Remark. Corollary 2.4 includes Corollary 2.4 of [33]: if some σ_{ν_μ} is zero, then $\varrho_\mu^{-\sigma_{\nu_\mu}} = 1$ even if we replace ϱ_μ by 0 in the definitions of R_μ and E_μ. Another special case of Corollary 2.4 is related to Lemma 2.5 of [34], which is nothing else than the case $\sigma_1 = \cdots = \sigma_L = 0$ of Lemma 1.5. Indeed we can take in Corollary 2.4

$$\zeta_0, \ldots, \zeta_{L-1} \ \text{pairwise distinct}, \qquad N = L-1, \qquad \sigma_0 = \cdots = \sigma_{L-1} = 0,$$
$$\nu_\mu = \mu - 1, \quad R_\mu = R, \quad r_\mu = r, \quad E_\mu = (R^2 + r^2)/2rR \quad (1 \le \mu \le L),$$

with
$$R \ge r \ge \max\{|\zeta_0|, \ldots, |\zeta_{L-1}|\}.$$

Since $\nu_1 + \cdots + \nu_L = L(L-1)/2$, we deduce from Corollary 2.4

$$\left| \det\Bigl(\varphi_\lambda(\zeta_\mu)\Bigr)_{1 \le \lambda, \mu \le L} \right| \le \left(\frac{R^2 + r^2}{2rR} \right)^{-L(L-1)/2} L! \prod_{\lambda=1}^{L} \sup_{|z|=R} |\varphi_\lambda(z)| \, .$$

Apart from the quantity $(R^2+r^2)/2rR$ which replaces R/r, this is the estimate of Lemma 2.5 of [34].

It is not clear to me whether Proposition 2.3 contains the general case of Lemma 1.5 (without the restriction $\sigma_1 = \cdots = \sigma_L = 0$).

2.2 Proof of Hermite-Lindemann's Theorem with an Interpolation Determinant and without Zero Estimate

Thanks to Corollary 2.4, one can modify the proof of Section 1.5 involving an interpolation determinant so that Lemma 1.6 (zero estimate) is not required any more. In this section we explain how to extrapolate and to increase either the number of derivatives, or the number of points, or both.

Proof of Hermite-Lindemann's Theorem 1.2. Let α and β be two complex numbers with $\beta \neq 0$ and $\alpha = e^\beta$.

`Step 1. Introducing the Parameters`

Consider two nondecreasing sequences $(S_0(N))_{N \ge 0}$ and $(S_1(N))_{N \ge 0}$ of nonnegative integers with $S_0(0) = S_1(0) = 0$ and such that the sequence $(S_0(N)S_1(N))_{N \ge 0}$ is increasing.

We construct a sequence $\zeta_0, \zeta_1, \ldots$ as follows. For each $N \ge 0$, the sequence

$$\bigl(\zeta_{S_0(N)S_1(N)}, \ldots, \zeta_{S_0(N+1)S_1(N+1)-1}\bigr)$$

consists of

- each element $s\beta$ with $0 \le s < S_1(N)$ repeated $S_0(N+1) - S_0(N)$ times,

and

- each element $s\beta$ with $S_1(N) \le s < S_1(N+1)$ repeated $S_0(N+1)$ times.

Denote by $(\sigma_0, \sigma_1, \ldots, \sigma_\nu, \ldots)$ the associated multiplicity sequence.

For each $\nu \geq 0$ denote by N_ν the least integer $N \geq 1$ for which $\nu < S_0(N)S_1(N)$. Hence we have $N_0 = 1$ and for $\nu \geq 0$

$$\begin{gathered} S_0(N_\nu - 1)S_1(N_\nu - 1) \leq \nu < S_0(N_\nu)S_1(N_\nu), \\ \sigma_\nu < S_0(N_\nu) \ \text{ and } \ |\zeta_\nu| < S_1(N_\nu)|\beta|. \end{gathered} \tag{2.2}$$

We also introduce two sufficiently large integers T_0 and T_1 and we set $L = T_0T_1$.

`Step 2. The Matrix M and the Determinant` Δ

Consider the matrix with L rows and infinitely many columns

$$\mathtt{M} = \Big(\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_\nu, \ldots\Big),$$

where $\mathcal{C}_\nu$ is the column vector (with $L = T_0T_1$ rows)

$$\left(\left(\frac{d}{dz}\right)^{\sigma_\nu} \big(z^\tau e^{tz}\big)(\zeta_\nu)\right)_{0 \leq \tau < T_0,\ 0 \leq t < T_1}.$$

We claim that the rank of $\mathtt{M}$ is L. Indeed a linear relation between the rows

$$\sum_{\tau_0=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} \Big(\frac{d}{dz}\Big)^{\sigma_\nu} \big(z^\tau e^{tz}\big)(\zeta_\nu) = 0 \ \text{ for } \ \nu \geq 0$$

would mean that the exponential polynomial

$$F(z) = \sum_{\tau_0=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} z^\tau e^{tz}$$

satisfies

$$F^{(\sigma_\nu)}(\zeta_\nu) = 0 \ \text{ for } \ \nu \geq 0.$$

These relations can also be written

$$F^{(\sigma)}(sb) = 0 \ \text{ for } \ \sigma \geq 0 \ \text{ and } \ s \geq 0,$$

and they plainly imply $c_{\tau t} = 0$ for $0 \leq \tau < T_0$ and $0 \leq t < T_1$.

We select L columns of $\mathtt{M}$ as the minimal ones in the lexicographic ordering such that we obtain a nonsingular matrix. Concretely we define $\nu_1, \ldots, \nu_L$ as follows:

$$\nu_1 = \min\{\nu \geq 0,\ \mathcal{C}_\nu \neq 0\},$$

and for $2 \leq \mu \leq L$

$$\nu_\mu = \min\{\nu > \nu_{\mu-1}\,;\ \mathcal{C}_{\nu_1}, \ldots, \mathcal{C}_{\nu_{\mu-1}}, \mathcal{C}_\nu \text{ are linearly independent}\}.$$

Hence we have $0 = \nu_1 < \nu_2 < \cdots < \nu_L$ and the matrix

$$\Big(\mathcal{C}_{\nu_1}, \ldots, \mathcal{C}_{\nu_L}\Big)$$

is nonsingular. We denote its determinant by Δ.

Remark. We may assume $\zeta_0 = 0$ and $\sigma_0 = 0$; in this case the first column has T_1 components 1 (those with index $(0, t)$ such that $0 \le t < T_1$) and $(T_0 - 1)T_1$ components 0 (the other ones, with index (τ, t) such that $1 \le \tau < T_0$ and $0 \le t < T_1$).

`Step 3. Upper Bound for` $|\Delta|$

We apply Corollary 2.4 with $\varrho_\mu = 1$, $r_\mu = |\beta|_+ S_1(N_{\nu_\mu})$, $R_\mu = 3er_\mu$, $E_\mu = e$ and

$$\{\varphi_1, \ldots, \varphi_L\} = \bigl\{z^\tau e^{tz} \; ; \; 0 \le \tau < T_0,\, 0 \le t < T_1\bigr\}.$$

We deduce

$$|\Delta| \le L! \prod_{\mu=1}^{L} \bigl(e^{-\nu_\mu} \sigma_{\nu_\mu}! R_\mu^{T_0} e^{T_1 R_\mu}\bigr) \le L! \exp\left(\sum_{\mu=1}^{L} \rho_\mu\right).$$

where

$$\rho_\mu = -\nu_\mu + \log\bigl(S_0(N_{\nu_\mu})!\bigr) + T_0 \log\bigl(3e|\beta|_+ S_1(N_{\nu_\mu})\bigr) + 3e|\beta|_+ T_1 S_1(N_{\nu_\mu}).$$

`Step 4. Lower Bound for` $|\Delta|$

The number Δ is not zero and lies in the ring $\mathbb{Z}[\alpha, \beta]$: there is a polynomial $f \in \mathbb{Z}[Z_1, Z_2]$ such that

$$\Delta = f(\alpha, \beta),$$

the degree of f in Z_1 and Z_2 respectively is at most

$$\sum_{\mu=1}^{L} T_1 S_1(N_{\nu_\mu}) \quad \text{and} \quad LT_0,$$

while the length of f is bounded by

$$\mathrm{L}(f) \le L! \prod_{\mu=1}^{L} T_1^{S_0(N_{\nu_\mu})} \bigl(S_0(N_{\nu_\mu}) + S_1(N_{\nu_\mu})\bigr)^{T_0}.$$

Assume now that α and β are both algebraic. Then we may use Liouville's inequality Proposition 1.13:

$$|\Delta| \ge L!^{-D+1} e^{-LDT_0 \mathrm{h}(\beta)}$$
$$\times \prod_{\mu=1}^{L} \Bigl(T_1^{-(D-1)S_0(N_{\nu_\mu})} \bigl(S_0(N_{\nu_\mu}) + S_1(N_{\nu_\mu})\bigr)^{-(D-1)T_0} e^{-DT_1 S_1(N_{\nu_\mu}) \mathrm{h}(\alpha)}\Bigr).$$

`Step 5. Choice of parameters`

Define, for $N \ge 0$,

$$S_0(N) = N^3 \quad \text{and} \quad S_1(N) = N.$$

By (2.2) we have for $\nu \geq 1$

$$(N_\nu - 1)^4 \leq \nu < N_\nu^4,$$

hence

$$S_0(N_\nu) \leq (\nu^{1/4}+1)^3 \leq 8\nu^{3/4}, \quad S_1(N_\nu) \leq \nu^{1/4}+1 \leq 2\nu^{1/4}.$$

Fix a sufficiently large integer T_0 (larger than some constant depending only on α and β) and define $T_1 = T_0$, so that $L = T_0^2$.

For $1 \leq \mu \leq L$ the following estimates are plain:

$$\begin{aligned} D\log L + DT_0\mathrm{h}(\beta) + (D-1)T_0\log\bigl(S_0(N_{\nu_\mu})S_1(N_{\nu_\mu})\bigr) \\ + T_0\log\bigl(3e|\beta|_+S_1(N_{\nu_\mu})\bigr) \leq (c_1 + c_2\log\nu_\mu)T_0 \end{aligned}$$

and

$$\begin{aligned} (D-1)S_0(N_{\nu_\mu})\log\bigl(T_1S_0(N_{\nu_\mu})\bigr) &\leq c_3\nu_\mu^{3/4}\log(T_0\nu_\mu), \\ T_1S_1(N_{\nu_\mu})\bigl(D\mathrm{h}(\alpha)+3e|\beta|_+\bigr) &\leq c_4T_0\nu_\mu^{1/4}, \end{aligned}$$

where $c_1, \ldots, c_4$ are positive real numbers which depend only on α and β. Therefore we have

$$\begin{aligned} D\log L + DT_0\mathrm{h}(\beta) + (D-1)T_0\log\bigl(S_0(N_{\nu_\mu})S_1(N_{\nu_\mu})\bigr) \\ + T_0\log\bigl(3e|\beta|_+S_1(N_{\nu_\mu})\bigr) + (D-1)S_0(N_{\nu_\mu})\log T_1 \\ + S_0(N_{\nu_\mu})\log\bigl(S_0(N_{\nu_\mu})\bigr) + T_1S_1(N_{\nu_\mu})\bigl(D\mathrm{h}(\alpha)+3e|\beta|_+\bigr) \leq Q_\mu, \end{aligned}$$

where

$$Q_\mu = c_5T_0\nu_\mu^{1/4} + c_6\nu_\mu^{3/4}\log(T_0\nu_\mu)$$

and again c_5, c_6 depend only on α and β.

`Step 6. Conclusion`

We claim

$$\sum_{\mu=1}^{L}(\nu_\mu - Q_\mu) > 0. \tag{2.3}$$

Indeed, since $\mu \mapsto Q_\mu$ is increasing, we have

$$\sum_{1\leq\mu\leq L/2} Q_\mu \leq \sum_{L/2<\mu\leq L} Q_\mu.$$

The estimate

$$\nu_\mu \geq 0 \quad \text{for} \quad 1 \leq \mu \leq L/2$$

is trivial, while for $L/2 < \mu \leq L$ the lower bound

$$\nu_\mu \geq \mu - 1 \geq (L-1)/2 = (T_0^2-1)/2$$

implies

$$\nu_\mu^{3/4} > 4c_5 T_0 \quad \text{and} \quad \nu_\mu^{1/4} > 4c_6 \log(T_0\nu_\mu),$$

hence

$$\nu_\mu > 2Q_\mu \quad \text{for} \quad L/2 < \mu \le L.$$

Therefore our claim (2.3) is vindicated.

According to steps 5 and 6, the conclusions of steps 3 and 4 are not compatible, hence one at least of the two numbers α, β is transcendental.

This completes the proof of Hermite-Lindemann's Theorem 1.2. □

3 Third Lecture. Linear Independence of Logarithms of Algebraic Numbers

The main result, due to A. Baker ([1] and [2] Th. 1.2), is the following:

Theorem 3.1. *Let $\alpha_1, \dots, \alpha_n$ be nonzero algebraic numbers. For each $i = 1, \dots, n$, let $\lambda_i \in \mathbb{C}$ satisfy $e^{\lambda_i} = \alpha_i$. Assume the n numbers $\lambda_1, \dots, \lambda_n$ are linearly independent over $\mathbb{Q}$. Then the $n+1$ numbers $1, \lambda_1, \dots, \lambda_n$ are linearly independent over the field $\overline{\mathbb{Q}}$ of algebraic numbers.*

We shall use the notation $\log \alpha_i$ in place of λ_i. One should keep in mind that this notation may be troublesome: for instance Theorem 3.1 can be applied with

$$\alpha_1 = \alpha_2 = 2, \quad \lambda_1 = \log 2, \quad \lambda_2 = \log 2 + 2i\pi,$$

and the conclusion shows that the three numbers $1, \log 2, \pi$ are linearly independent over $\overline{\mathbb{Q}}$. However the same conclusion can be obtained by taking $\alpha_1 = 2$ and $\alpha_2 = -1$ for instance.

By the way, when $\alpha_1, \dots, \alpha_n$ are nonzero complex numbers, for any choice $(\lambda_1, \dots, \lambda_n) \in \mathbb{C}^n$ with $e^{\lambda_i} = \alpha_i \quad (1 \le i \le n)$, the following conditions are clearly equivalent:

(i) The numbers $\alpha_1, \dots, \alpha_n$ are multiplicatively independent, which means that any relation
$$\alpha_1^{a_1} \cdots \alpha_n^{a_n} = 1$$
with $(a_1, \dots, a_n) \in \mathbb{Z}^n$ implies $a_1 = \cdots = a_n = 0$.

(ii) The $n+1$ complex numbers $2\pi i, \lambda_1, \dots, \lambda_n$ are linearly independent over $\mathbb{Q}$.

Hence, given complex numbers $\lambda_1, \dots, \lambda_n$, the multiplicative subgroup of $\mathbb{C}^\times$ generated by $e^{\lambda_1}, \dots, e^{\lambda_n}$ has rank (as a $\mathbb{Z}$-module) equal to $r-1$, where r is the dimension of the $\mathbb{Q}$-vector space spanned by $2\pi i, \lambda_1, \dots, \lambda_n$. In particular, if $\lambda_1, \dots, \lambda_n$ are $\mathbb{Q}$-linearly independent, then this rank is

$$\begin{cases} n & \text{if } \; 2\pi i, \lambda_1, \dots, \lambda_n \; \text{ are linearly independent over } \; \mathbb{Q}, \\ n-1 & \text{if } \; 2\pi i, \lambda_1, \dots, \lambda_n \; \text{ are } \; \mathbb{Q}-\text{linearly dependent.} \end{cases}$$

Proofs of Baker's Theorem 3.1 on linear independence of logarithms are given in [2] Chap. 2, [6] Chap. 10, Section 1, [7], Chap. 4, Section 1.3, [30], Chap. 8 and [34], Section 10.1. A "dual" argument (extension of Schneider's method, while Baker's method is an extension of Gel'fond's method) is worked out in [34] Chap. 6 (for the homogeneous case) and Section 9.1 (for the non-homogeneous case). See also [34], Section 4.2 for a proof, following Bertrand and Masser, which rests on Schneider-Lang's Criterion for Cartesian products (involving again Gel'fond's method).

Here we consider only Baker's method. In Section 3.1 we explain why Baker's method can be introduced by means of functions of either one or several variables. A proof of Theorem 3.1 by means of Baker's method, involving an auxiliary function with an extrapolation argument, is given in Section 3.2, which includes also a sketch of proof with an interpolation determinant but without any extrapolation. For both proofs the zero estimate which is used there is due to Philippon [22]. In Section 3.3 we show how to replace this zero estimate by a much simpler one, due to R. Tijdeman, by means of a further extrapolation with the auxiliary function. Our ultimate goal in this third lecture is to extrapolate with an interpolation determinant (Section 3.5), in order to complete the proof of Theorem 3.1 without any auxiliary function, and with Tijdeman's zero estimate in place of Philippon's one. This is achieved thanks to a generalization (in Section 3.4) in several variables of the results of Section 2.1 giving upper bounds for interpolation determinants.

3.1 Introduction to Baker's Method

We explain the basic ideas of the proof of Theorem 3.1 by means of Baker's method with an auxiliary function involving an extrapolation.

Assume $\alpha_1, \ldots, \alpha_n$ are nonzero algebraic numbers, $\log\alpha_1, \ldots, \log\alpha_n$ are linearly independent over $\mathbb{Q}$, $\beta_0, \ldots, \beta_{n-1}$ are algebraic numbers, and

$$\log\alpha_n = \beta_0 + \beta_1 \log\alpha_1 + \cdots + \beta_{n-1}\log\alpha_{n-1}. \tag{3.1}$$

We shall eventually reach a contradiction.

From now on α^z stands for $\exp(z\log\alpha)$, which has a meaning as soon as a complex number $\lambda = \log\alpha$ has been selected with $e^\lambda = \alpha$.

Hence relation (3.1) implies

$$\alpha_n = e^{\beta_0}\alpha_1^{\beta_1}\cdots\alpha_{n-1}^{\beta_{n-1}},$$

and more generally for $z \in \mathbb{C}$

$$\alpha_n^z = e^{\beta_0 z}\alpha_1^{\beta_1 z}\cdots\alpha_{n-1}^{\beta_{n-1} z}.$$

To each polynomial $P \in \mathbb{Z}[Y_0, Y_1, \ldots, Y_n]$ we associate analytic functions of 1, n and $n+1$ complex variables. The proof of Baker's qualitative Theorem 3.1 on linear independence of logarithms of algebraic numbers requires a Schwarz'

Lemma, while the quantitative refinements (measures of linear independence; see Section 5) will require an approximate Schwarz' Lemma. A fundamental fact is that one needs such auxiliary results only for functions of a single variable: even if we introduce functions of several variables, we shall consider only the values of our functions at multiples of a single point (but derivatives are taken in several directions); therefore it would be possible to avoid completely the introduction of several variables, but we use them only to explain the role of certain differential operators.

a) Using a Single Variable

Consider the entire functions

$$z, \alpha_1^z, \ldots, \alpha_n^z.$$

To the auxiliary polynomial $P \in \mathbb{Z}[\underline{Y}]$ is attached the exponential polynomial

$$G(z) = P(z, \alpha_1^z, \ldots, \alpha_n^z),$$

which can be written also

$$G(z) = P\big(z, \alpha_1^z, \ldots, \alpha_{n-1}^z, e^{\beta_0 z} \alpha_1^{\beta_1 z} \cdots \alpha_{n-1}^{\beta_{n-1} z}\big).$$

In order to take (3.1) into account, we consider derivatives of G. We avoid difficulties (related with Liouville's inequality) arising from the unwanted transcendental numbers $\log \alpha_i \quad (1 \le i < n)$ by writing the derivatives of G as polynomials in $\log \alpha_1, \ldots, \log \alpha_{n-1}$, and the coefficients of these polynomials are themselves exponential polynomials with algebraic coefficients.

We start with the first derivative $G'(z) = (d/dz)G(z)$ of G: this is the value, at $(z, \alpha_1^z, \ldots, \alpha_n^z)$, of the polynomial

$$\left(\frac{\partial}{\partial Y_0} + \sum_{k=1}^{n} (\log \alpha_k) \frac{\partial}{\partial Y_k}\right) P = \big(\partial_0 + \partial_1 \log \alpha_1 + \cdots + \partial_{n-1} \log \alpha_{n-1}\big) P,$$

where $\partial_0, \ldots, \partial_{n-1}$ are the differential operators

$$\partial_0 = \frac{\partial}{\partial Y_0} + \beta_0 Y_n \frac{\partial}{\partial Y_n}, \qquad \partial_k = Y_k \frac{\partial}{\partial Y_k} + \beta_k Y_n \frac{\partial}{\partial Y_n} \quad (1 \le k \le n-1)$$

on the ring $\mathbb{C}[\underline{Y}]$

We now take higher derivatives. For $\underline{\sigma} = (\sigma_0, \ldots, \sigma_{n-1}) \in \mathbb{N}^n$ we write $\partial^{\underline{\sigma}}$ in place of

$$\partial_0^{\sigma_0} \cdots \partial_{n-1}^{\sigma_{n-1}}.$$

Since

$$\bigl(\partial_0 + \partial_1 \log\alpha_1 + \cdots + \partial_{n-1}\log\alpha_{n-1}\bigr)^k = \sum_{\|\underline{\sigma}\|=k} \frac{k!}{\underline{\sigma}!} (\log\alpha_1)^{\sigma_1}\cdots(\log\alpha_{n-1})^{\sigma_{n-1}}\partial^{\underline{\sigma}},$$

we have

$$G^{(k)}(z) = \sum_{\|\underline{\sigma}\|=k} \frac{k!}{\underline{\sigma}!} (\log\alpha_1)^{\sigma_1}\cdots(\log\alpha_{n-1})^{\sigma_{n-1}} G_{\underline{\sigma}}(z),$$

where

$$G_{\underline{\sigma}}(z) = P_{\underline{\sigma}}(z, \alpha_1^z, \ldots, \alpha_n^z) \quad\text{and}\quad P_{\underline{\sigma}} = \partial^{\underline{\sigma}} P \in \mathbb{C}[\underline{Y}].$$

Now if we compose derivations we easily deduce for $\underline{\sigma}$ and $\underline{\lambda}$ in $\mathbb{N}^n$

$$\partial^{\underline{\sigma}+\underline{\lambda}} = \partial^{\underline{\sigma}} \circ \partial^{\underline{\lambda}},$$

which yields the fundamental relation for $\underline{\sigma} \in \mathbb{N}^n$ and $\ell \in \mathbb{N}$

$$G_{\underline{\sigma}}^{(\ell)}(z) = \sum_{\|\underline{\lambda}\|=\ell} \frac{\ell!}{\underline{\lambda}!} (\log\alpha_1)^{\lambda_1}\cdots(\log\alpha_{n-1})^{\lambda_{n-1}} G_{\underline{\sigma}+\underline{\lambda}}(z). \tag{3.2}$$

As a consequence, if S_0 and S_1 are positive integers for which

$$G_{\underline{\sigma}}(s) = 0 \quad\text{for}\quad \|\underline{\sigma}\| < S_0 \quad\text{and}\quad 0 \le s < S_1, \tag{3.3}$$

then for any $\underline{\sigma} \in \mathbb{N}^n$ with $\|\underline{\sigma}\| < S_0$ the function $G_{\underline{\sigma}}$ has a zero at $s = 0, \ldots, S_1 - 1$ of multiplicity $\ge S_0 - \|\underline{\sigma}\|$.

This is really the main point in Baker's method [1], I, p.212, which has no counterpart in the dual method of [34] Chap. 6 (there is no efficient extrapolation so far when one deals with functions of several variables). By means of the one dimensional approximate Schwarz' Lemma 1.12, one deduces from (3.3) a sharp upper bound for $|G_{\underline{\sigma}}|_r$ and gets more equations like (3.3): this is the extrapolation.

Remark. If we were to replace (3.1) by an algebraic relation between logarithms of algebraic numbers, for instance

$$\log\alpha_n = A(\log\alpha_1, \ldots, \log\alpha_{n-1})$$

where A is a polynomial of total degree > 1, then one could also write the derivatives of G as polynomials in $(\log\alpha_1, \ldots, \log\alpha_{n-1})$, but there are no nice relations like (3.2) between the corresponding exponential polynomials replacing the $G_{\underline{\sigma}}$.

b) Introducing n Variables

As we said, from a strict logical point of view, introducing functions of several variables is not required. But it may help to understand better the meaning of the differential operators $\partial_k \quad (0 \le k < n)$.

To the auxiliary polynomial P is also associated an analytic function of n complex variables

$$\Phi(z_0, z_1 \ldots, z_{n-1}) = P\bigl(z_0, e^{z_1}, \ldots, e^{z_{n-1}}, e^{\beta_0 z_0 + \beta_1 z_1 + \cdots + \beta_{n-1} z_{n-1}}\bigr).$$

We take derivatives of Φ with respect to the n variables, and consider the values of these derivatives at the point

$$\underline{v} = (1, \log \alpha_1, \ldots, \log \alpha_{n-1}) \in \mathbb{C}^n.$$

Obviously we have

$$\Phi(z\underline{v}) = G(z)$$

for $z \in \mathbb{C}$, but what is more interesting is the connection between the derivatives. From the definition of the differential operators $\partial_0, \ldots, \partial_{n-1}$, it is plain that for $0 \le k \le n-1$ we have

$$\frac{\partial}{\partial z_k} \Phi(z_0, z_1 \ldots, z_{n-1}) = (\partial_k P)\bigl(z_0, e^{z_1}, \ldots, e^{z_{n-1}}, e^{\beta_0 z_0 + \beta_1 z_1 + \cdots + \beta_{n-1} z_{n-1}}\bigr).$$

Hence for $\underline{\sigma} = (\sigma_0, \ldots, \sigma_{n-1}) \in \mathbb{N}^n$ and $z \in \mathbb{C}$ we have

$$G_{\underline{\sigma}}(z) = (\mathcal{D}^{\underline{\sigma}} \Phi)(z\underline{v})$$

where

$$\mathcal{D}^{\underline{\sigma}} = \left(\frac{\partial}{\partial z_0}\right)^{\sigma_0} \cdots \left(\frac{\partial}{\partial z_{n-1}}\right)^{\sigma_{n-1}}.$$

c) Introducing $n+1$ Variables

Instead of working with n variables it is sometimes convenient (for instance for the zero estimate) to consider $n+1$ variables: define

$$F(z_0, z_1 \ldots, z_n) = P\bigl(z_0, e^{z_1}, \ldots, e^{z_n}\bigr).$$

The point

$$\underline{u} = (1, \log \alpha_1, \ldots, \log \alpha_n) \in \mathbb{C}^{n+1}$$

lies in the hyperplane W of equation

$$z_n = \beta_0 z_0 + \beta_1 z_1 + \cdots + \beta_{n-1} z_{n-1}.$$

A basis of W is $\mathbf{w} = (\underline{w}_0, \ldots, \underline{w}_{n-1})$ where

$$\underline{w}_k = (\delta_{k0}, \dots, \delta_{k,n-1}, \beta_k) \quad (0 \le k \le n-1).$$

These elements $\underline{w}_0, \dots, \underline{w}_{n-1}$ are the n column vectors of the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \beta_0 & \beta_1 & \cdots & \beta_{n-1} \end{pmatrix} = \begin{pmatrix} & & \\ & \mathtt{I}_n & \\ & & \\ \beta_0 & \cdots & \beta_{n-1} \end{pmatrix}.$$

Since

$$z_0\underline{w}_0 + \cdots + z_{n-1}\underline{w}_{n-1} = (z_0, \dots, z_{n-1}, \beta_0 z_0 + \cdots + \beta_{n-1}z_{n-1}),$$

we have

$$\Phi(z_0, \dots, z_{n-1}) = F(z_0\underline{w}_0 + \cdots + z_{n-1}\underline{w}_{n-1}),$$

and one may view Φ as the restriction of F to the hyperplane W, equipped with the basis $\mathbf{w}$, by means of the isomorphism

$$\begin{array}{ccc} \mathbb{C}^n & \longrightarrow & W \\ (z_0, \dots, z_{n-1}) & \longmapsto & z_0\underline{w}_0 + \cdots + z_{n-1}\underline{w}_{n-1}. \end{array}$$

To take the derivatives of Φ in all n directions amounts to taking the derivatives of F in the directions of W. More precisely, for $\underline{x} = (x_0, \dots, x_n) \in \mathbb{C}^{n+1}$, define

$$\mathcal{D}_{\underline{x}} = x_0 \frac{\partial}{\partial z_0} + \cdots + x_n \frac{\partial}{\partial z_n}\cdot$$

For instance

$$\mathcal{D}_{\underline{w}_k} = \frac{\partial}{\partial z_k} + \beta_k \frac{\partial}{\partial z_n} \quad \text{for} \quad 0 \le k \le n-1,$$

hence for $0 \le k \le n-1$

$$\mathcal{D}_{\underline{w}_k} F(z_0, \dots, z_n) = (\partial_k P)(z_0, e^{z_1}, \dots, e^{z_n}).$$

Define also, for $\underline{\sigma} = (\sigma_0, \dots, \sigma_{n-1}) \in \mathbb{N}^n$,

$$\mathcal{D}_{\mathbf{w}}^{\underline{\sigma}} = \mathcal{D}_{\underline{w}_0}^{\sigma_0} \cdots \mathcal{D}_{\underline{w}_{n-1}}^{\sigma_{n-1}}.$$

Then

$$\mathcal{D}_{\mathbf{w}}^{\underline{\sigma}} F(z_0, \dots, z_n) = (\partial^{\underline{\sigma}} P)(z_0, e^{z_1}, \dots, e^{z_n})$$

and

$$\mathcal{D}_{\mathbf{w}}^{\underline{\sigma}} F(z_0, \dots, z_{n-1}, \beta_0 z_0 + \cdots + \beta_{n-1}z_{n-1}) \\ = \left(\frac{\partial}{\partial z_0}\right)^{\sigma_0} \cdots \left(\frac{\partial}{\partial z_{n-1}}\right)^{\sigma_{n-1}} \Phi(z_0, \dots, z_{n-1}).$$

In particular

$$G_{\underline{\sigma}}(z) = \mathcal{D}_{\mathbf{w}}^{\underline{\sigma}} F(z\underline{u}).$$

3.2 Proof of Baker's Theorem

We prove Theorem 3.1 following basically [30] Chap. 8. One main difference is that we shall use a sharper zero estimate than Theorem 6.1.1 of [30], and therefore we do not need a long extrapolation like in [30]: here a single step will be sufficient. This explains why the interpolation determinant method could easily be used in [34], Section 10.1. In Section 3.3 we shall explain how a longer extrapolation enables one to use a weaker zero estimate.

Denote by $\mathbb{K}$ the number field

$$\mathbb{K} = \mathbb{Q}\big(\alpha_1, \dots, \alpha_n, \beta_0, \dots, \beta_{n-1}\big).$$

Let $T_0, T_1, \dots, T_n$, S_0 and S_1 be sufficiently large positive integers. Explicit conditions on these parameters will occur along the proof, and we shall discuss them later, but it may help the reader to know that a suitable choice is

$$\begin{aligned} T_0 &= 2[\mathbb{K}:\mathbb{Q}]N^{2n+1}, & T_1 &= \cdots = T_n = N^{2n-1}, \\ S_0 &= N^{2n+1}, & S_1 &= N. \end{aligned} \tag{3.4}$$

We also set $L = T_0 T_1 \cdots T_n$ and $T = T_1 + \cdots + T_n$.

a) Construction of the Auxiliary Polynomial

By means of Thue-Siegel's Lemma 1.4, we show the existence of a nonzero auxiliary polynomial $P \in \mathbb{Z}[Y_0, \dots, Y_n]$, with degree $< T_i$ in Y_i $(0 \le i \le n)$, such that the equations (3.3) hold.

These conditions amount to a homogeneous linear system of equations with coefficients in $\mathbb{K}$, where the unknowns are the coefficients of P.

Our first condition on the parameters will be

$$\boxed{2[\mathbb{K}:\mathbb{Q}]\binom{S_0+n-1}{n} S_1 \le L,}$$

so that the number of equations is at most half the number of unknowns[7].

The coefficients of the linear system are the numbers

$$\gamma_{\tau \underline{t}}^{(\underline{\sigma}, s)} = \partial^{\underline{\sigma}}\big(Y_0^{\tau} Y_1^{t_1} \cdots Y_n^{t_n}\big)\big(s, \alpha_1^s, \dots, \alpha_n^s\big) \tag{3.5}$$

with $0 \le \tau < T_0$, $0 \le t_i < T_i$ $(1 \le i \le n)$ and $\underline{\sigma} \in \mathbb{N}^n$, $\|\underline{\sigma}\| < S_0$, $0 \le s < S_1$.

We write them explicitly by computing the derivatives of

$$z_0^{\tau} e^{t_1 z_1} \cdots e^{t_{n-1} z_{n-1}} e^{t_n(\beta_0 z_0 + \cdots + \beta_{n-1} z_{n-1})}.$$

One obtains easily

[7] One could construct P with coefficients in the field $\mathbb{K}$, and then omit the factor $[\mathbb{K}:\mathbb{Q}]$.

$$\gamma_{\tau\underline{t}}^{(\underline{\sigma},s)} = \sum_{\kappa=0}^{\min\{\tau,\sigma_0\}} \frac{\sigma_0!\tau!}{\kappa!(\sigma_0-\kappa)!(\tau-\kappa)!} s^{\tau-\kappa}(\beta_0 t_n)^{\sigma_0-\kappa} \prod_{i=1}^{n-1}(t_i+t_n\beta_i)^{\sigma_i} \prod_{j=1}^{n} \alpha_j^{t_j s}.$$

One should use a variant of Lemma 1.4 taking into account the fact that the coefficients of our linear system are in $\mathbb{K}$ rather than in $\mathbb{Z}$, but anyway a rough estimate shows that we end up with a nonzero polynomial P of length at most

$$\mathrm{L}(P) \le \exp\{c_1(T_0+S_0)\log L + c_2 T S_1\}.$$

Here and below, $c_1, \dots, c_{16}$ denote positive numbers which do not depend on T_0, $T_1, \dots, T_n$, S_0, S_1. For instance with our choice (3.4) we get

$$\mathrm{L}(P) \le \exp\{c_3 N^{2n+1} \log N\}.$$

Remark. The whole point in this argument is that (3.1) allows us to consider values of polynomials in $n-1$ variables at the point

$$(\log \alpha_1, \dots, \log \alpha_{n-1})$$

in place of values of polynomials in n variables at the point

$$(\log \alpha_1, \dots, \log \alpha_n).$$

Without (3.1) it would be necessary to replace

$$\binom{S_0+n-1}{n} \quad \text{with} \quad \binom{S_0+n}{n+1}$$

and the only difference with the present proof is that no choice of parameters would be admissible!

b) Extrapolation on Integral Points

We introduce further parameters S_0' and S_1' which are positive integers with $S_0' < S_0$ and $S_1' > S_1$, and we are going to prove

$$G_{\underline{\sigma}}(s) = 0 \quad \text{for} \quad (\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N} \quad \text{with} \quad \|\underline{\sigma}\| < S_0' \quad \text{and} \quad 0 \le s < S_1'. \tag{3.6}$$

With the choice of parameters (3.4) we shall take

$$S_0' = [S_0/2], \quad S_1' = N^2.$$

Fix $\underline{\sigma} \in \mathbb{N}^n$ with $\|\underline{\sigma}\| < S_0'$. The function

$$G_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}} P(z, \alpha_1^z, \dots, \alpha_n^z)$$

has a zero at each point $s = 0, \ldots, S_1 - 1$ of multiplicity $\geq S_0 - S_0'$. The one variable Schwarz' Lemma 1.3 with $r = S_1'$, $E = e$, $R = 2er$ provides the following upper bound:

$$|G_{\underline{\sigma}}|_r \leq e^{-(S_0 - S_0')S_1} |G_{\underline{\sigma}}|_R.$$

This yields an upper bound for $|G_{\underline{\sigma}}(s)|$ with $s \in \mathbb{Z}$ in the range $0 \leq s < S_1'$ which is not compatible with Liouville's lower bound provided that

$$\boxed{c_4(T_0 + S_0)\log L + c_5 T S_1' < (S_0 - S_0')S_1.}$$

Hence (3.6).

c) Using Philippon's Zero Estimate

The next auxiliary Lemma is a very special case of Philippon's zero estimate [22] (see also Chap. 8 of [34] by D. Roy).

Proposition 3.2. *Let $\alpha_1, \ldots, \alpha_n$ be nonzero complex numbers which generate a multiplicative subgroup of $\mathbb{C}^\times$ of rank $\geq n-1$ and let $\beta_0, \ldots, \beta_{n-1}$ be complex numbers. Assume that $1, \beta_1, \ldots, \beta_{n-1}$ are linearly independent over $\mathbb{Q}$. Let T_0, $T_1, \ldots, T_n$, S_0, S_1 and L be positive integers satisfying the following conditions:*

$$L = T_0 T_1 \cdots T_n, \qquad T_1 \leq T_2 \leq \cdots \leq T_n,$$

$$S_0 > (n+1)T_n, \qquad S_0 S_1 > \frac{1}{2} n!(n+1)! \max\{T_0,\ 2T_n\},$$

and

$$\binom{S_0 + n - 1}{n} S_1 > (n+1)! T_0 T_1 \cdots T_n. \tag{3.7}$$

Assume also

- *either $\beta_0 \neq 0$*
- *or else $n \geq 2$ and*

$$\binom{S_0 + n - 2}{n-1} S_1 > (n+1)! T_1 \cdots T_n. \tag{3.8}$$

For $\tau \in \mathbb{N}$, $\underline{t} \in \mathbb{N}^n$, $\underline{\sigma} \in \mathbb{N}^n$ and $s \in \mathbb{N}$, consider the number $\gamma_{\tau \underline{t}}^{(\underline{\sigma} s)}$ given by (3.5) and build up the matrix:

$$\mathtt{M} = \left(\gamma_{\tau \underline{t}}^{(\underline{\sigma} s)} \right)_{\substack{(\tau, \underline{t}) \\ (\underline{\sigma}, s)}}$$

where the index of rows $(\tau, \underline{t})$ runs over the elements in $\mathbb{N} \times \mathbb{N}^n$ with $0 \leq \tau < T_0$, $0 \leq t_i < T_i \quad (1 \leq i \leq n)$, while the index of columns $(\underline{\sigma}, s)$ runs over the elements in $\mathbb{N}^n \times \mathbb{N}$ with $\|\underline{\sigma}\| < (n+1)S_0$ and $0 \leq s < (n+1)S_1$. Then $\mathtt{M}$ has rank L.

Remark. Using (3.7) it is easily checked that (3.8) can be replaced by

$$nT_0 > S_0 + n - 1.$$

Remark. Given $\mathbb{Q}$-linearly independent complex numbers $\lambda_1, \dots, \lambda_n$, the n numbers $\alpha_i = e^{\lambda_i}$ $(1 \le i \le n)$ generate a multiplicative subgroup of $\mathbb{C}^\times$ of rank $\ge n-1$. Conversely, if $\alpha_1, \dots, \alpha_n$ are nonzero elements of $\mathbb{C}$ which generate a multiplicative subgroup of $\mathbb{C}^\times$ of rank $\ge n-1$, then there exist $\mathbb{Q}$-linearly independent complex numbers $\lambda_1, \dots, \lambda_n$ such that $\alpha_i = e^{\lambda_i}$ for $1 \le i \le n$.

Proposition 3.2 is essentially Proposition 10.2 of [34], with a few differences:

- We do not assume $\beta_0 \neq 0$ here. At the same time our points are $(s, \alpha_1^s, \dots, \alpha_n^s)$ in place[8] of $(s\beta_0, \alpha_1^s, \dots, \alpha_n^s)$.
- We work with polynomials in $Y_0, \dots, Y_n$ of degree $< T_i$ in Y_i $(0 \le i \le n)$, while in [34] we considered polynomials in $X_0, X_1^{\pm 1}, \dots, X_n^{\pm 1}$ of degree $\le T_0$ in X_0 and degree $\le T_1$ in each of the variables $X_i^{\pm 1}$. Also here we consider nonnegative integers s with $0 \le s < (n+1)S_1$, while in [34] we had $s \in \mathbb{Z}$ with $|s| \le (n+1)S_1$. Also here we use strict inequalities for $\|\underline{\sigma}\|$.

These changes introduce few modifications in the proof of Proposition 10.2 of [34], but for the convenience of the reader we provide the details.

Proof. Consider the algebraic groups $G_0 = \mathbb{G}_{\mathrm{a}}$, $G_1 = \mathbb{G}_{\mathrm{m}}^n$, $G = G_0 \times G_1$, of dimensions $d_0 = 1$, $d_1 = n$ and $d = n+1$ respectively. Let W be the hyperplane in $\mathbb{C}^{n+1}$ of equation

$$\beta_0 z_0 + \beta_1 z_1 + \cdots + \beta_{n-1} z_{n-1} = z_n.$$

Introduce also the set

$$\Sigma = \bigl\{(s, \alpha_1^s, \dots, \alpha_n^s)\,;\, s \in \mathbb{N},\, 0 \le s < S_1\bigr\} \subset G(\mathbb{C}) = \mathbb{C} \times (\mathbb{C}^\times)^n.$$

If the rank of the matrix $\mathtt{M}$ is less than L, then there exists a nonzero polynomial P in $\mathbb{C}[Y_0, Y_1, \dots, Y_n]$, of degree $< T_i$ in Y_i $(0 \le i \le n)$, for which the functions $G_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}} P\bigl(z, \alpha_1^z, \dots, \alpha_n^z\bigr)$ satisfy

$$\begin{aligned} G_{\underline{\sigma}}(s) = 0 \quad &\text{for} \quad (\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N} \\ &\text{with} \quad \|\underline{\sigma}\| < (n+1)S_0 \quad \text{and} \quad 0 \le s < (n+1)S_1. \end{aligned}$$

According to Philippon's zero estimate there exists an algebraic subgroup G^* of G of dimension $d^* \le n$ and codimension $d' = n+1-d^*$ such that

$$\binom{S_0 + \ell_0 - 1}{\ell_0} \mathrm{Card}\left(\frac{\Sigma + G^*}{G^*}\right) \mathcal{H}(G^*; \underline{T}) \le \mathcal{H}(G; \underline{T}), \tag{3.9}$$

[8] This was an oversight in Proposition 10.2 of [34]!

where

$$\ell_0 = \begin{cases} d' - 1 & \text{if } T_e(G^*) \subset W, \\ d' & \text{otherwise.} \end{cases}$$

The notation $\mathcal{H}(G^*; \underline{T})$ stands for a multihomogeneous Hilbert-Samuel polynomial (see [34], Section 5.2.3); for instance $\mathcal{H}(G; \underline{T}) = (n+1)!L$.

We first check that this inequality (3.9) is not satisfied with $G^* = \{e\}$: indeed when $G^* = \{e\}$ we have

$$d^* = 0, \quad \ell_0 = n, \quad \mathrm{Card}\left(\frac{\Sigma + G^*}{G^*}\right) = \mathrm{Card}(\Sigma) = S_1, \quad \mathcal{H}(G^*; \underline{T}) = 1,$$

so that by (3.7)

$$\binom{S_0 + \ell_0 - 1}{\ell_0} \mathrm{Card}\left(\frac{\Sigma + G^*}{G^*}\right) \mathcal{H}(G^*; \underline{T}) > (n+1)!L.$$

Therefore $d^* \geq 1$.

Write $G^* = G_0^* \times G_1^*$ where G_0^* is an algebraic subgroup of G_0 and G_1^* an algebraic subgroup of G_1. Denote by d_0^* and d_1^* the dimensions of G_0^* and G_1^* respectively, and by d_0' and d_1' their codimensions:

$$d_0^* + d_0' = d_0 = 1, \quad d_1^* + d_1' = d_1 = n.$$

Assume first $T_e(G^*) \subset W$. Since $1, \beta_1, \ldots, \beta_{n-1}$ are linearly independent over $\mathbb{Q}$, the hyperplane of $\mathbb{C}^n$ of equation

$$\beta_1 z_1 + \cdots + \beta_{n-1} z_{n-1} = z_n$$

does not contain any nonzero element of $\mathbb{Q}^n$. Since $T_e(G_1^*)$ is a subspace of $\mathbb{C}^n$ which is rational over $\mathbb{Q}$, we deduce $G_1^* = \{e\}$, hence $G^* = \mathbb{G}_{\mathrm{a}} \times \{e\}$, $d^* = 1$, $d' = n$, $\ell_0 = n - 1$ and $\mathcal{H}(G^*; \underline{T}) = T_0$. Now the condition $T_e(G^*) \subset W$ implies $\beta_0 = 0$, hence (3.8) gives $n \geq 2$ and

$$\binom{S_0 + n - 2}{n - 1} S_1 > (n+1)! T_1 \cdots T_n.$$

Since $n \geq 2$ and since $\alpha_1, \ldots, \alpha_n$ generate a subgroup of $\mathbb{C}^\times$ of rank $\geq n-1$, we have

$$\mathrm{Card}\left(\frac{\Sigma + G^*}{G^*}\right) = \mathrm{Card}\big\{(\alpha_1^s, \ldots, \alpha_n^s)\,;\, s \in \mathbb{N},\, 0 \leq s < S_1\big\} = S_1.$$

Hence (3.9) does not hold and we get a contradiction.

Therefore $T_e(G^*) \not\subset W$ and $\ell_0 = d'$.

Consider the case $G_0^* = \{0\}$. We have $d_0^* = 0$, $d^* = d_1^*$, $d' = n + 1 - d_1^*$,

$$\mathcal{H}(G^*; \underline{T}) \geq (d_1^* + 1)! T_1 \cdots T_{d_1^*} \quad \text{and} \quad \mathrm{Card}\left(\frac{\Sigma + G^*}{G^*}\right) = S_1.$$

Therefore (3.9) implies

$$\binom{S_0+d'-1}{d'}S_1 \le \frac{(n+1)!}{(n+2-d')!}T_0T_{d_1^*+1}\cdots T_n,$$

hence

$$S_0^{d'}S_1 \le \frac{(n+1)!d'!}{(n+2-d')!}T_0T_{d_1^*+1}\cdots T_n.$$

However we have $S_0 > T_n$, $1 \le d' \le n$ and

$$\frac{d'!}{(n+2-d')!} \le \frac{1}{2}n!,$$

hence we get a contradiction with the inequality

$$S_0S_1 > \frac{1}{2}n!(n+1)!T_0.$$

So we have $d_0^* = 1$, $d^* = d_1^* + 1$, $d' = n - d_1^*$ and

$$\mathcal{H}(G^*;\underline{T}) \ge (d_1^*+1)!T_0T_1\cdots T_{d_1^*}.$$

Now (3.9) gives

$$\binom{S_0+d'-1}{d'}\mathrm{Card}\left(\frac{\Sigma+G^*}{G^*}\right) \le \frac{(n+1)!}{(n+1-d')!}T_{d_1^*+1}\cdots T_n$$

from which we deduce

$$S_0^{d'}\mathrm{Card}\left(\frac{\Sigma+G^*}{G^*}\right) \le \frac{(n+1)!d'!}{(n+1-d')!}T_{d_1^*+1}\cdots T_n.$$

Using the estimates

$$S_0 > T_n, \quad 1 \le d' \le n, \quad \frac{d'!}{(n+1-d')!} \le n!$$

and

$$S_0S_1 > n!(n+1)!T_n$$

we obtain

$$\mathrm{Card}\left(\frac{\Sigma+G^*}{G^*}\right) < S_1,$$

which means $\Sigma \cap G^* \ne \{e\}$. The assumption on the rank of the subgroup of $\mathbb{C}^*$ generated by $\alpha_1,\ldots,\alpha_n$ then implies $d_1^* = n-1$, $d' = 1$ and we get the estimate

$$S_0 \le (n+1)T_n$$

which is not compatible with our assumptions. □

d) End of the Proof of Baker's Theorem 3.1

We apply Proposition 3.2 with S_0 and S_1 replaced respectively by

$$\lceil S_0'/(n+1)\rceil \quad\text{and}\quad \lceil S_1'/(n+1)\rceil.$$

From (3.4) and (3.6), using the estimates

$$S_0' > (n+1)^2 T, \quad S_0'S_1' > n!(n+1)!(n+1)^2 \max\left\{\frac{1}{2}T_0,\ T\right\},$$

$$nT_0 > S_0 + n - 1 \quad\text{and}\quad {S_0'}^n S_1' > n!(n+1)!L,$$

we deduce a contradiction, which completes the proof of Baker's Theorem 3.1. □

Remark. The basic ideas for a proof of Baker's Theorem with an interpolation determinant in [34], Section 10.1.4 are essentially the same. Instead of using Dirichlet's pigeonhole principle to solve a system of linear equations, we only consider the matrix of this linear system. More precisely the relevant matrix $\mathtt{M}$ is the one occurring in the zero estimate (Proposition 3.2): it has maximal rank, and enables one to start with a nonzero determinant Δ. As usual the required lower bound for $|\Delta|$ is given by Liouville's estimate (Proposition 1.13). On the other hand the argument occurring above turns out to be perfectly adaptable to yield an upper bound for $|\Delta|$ which gives just what we need.

The difference between the proof with an auxiliary function and the proof with an interpolation determinant is that in the latter Dirichlet's box principle is not required. However there is a substitute to the auxiliary function, which is the (explicit) exponential polynomial given by a determinant

$$\det\Big(\mathcal{C}_1,\ \ldots,\ \mathcal{C}_{L-1},\ \Phi(z)\Big),$$

where $\mathcal{C}_1,\ldots,\mathcal{C}_{L-1}$ are $L-1$ vector columns of $\mathtt{M}$, while the last column vector $\Phi : \mathbb{C} \to \mathbb{C}^L$ is given by

$$\big(z^\tau \alpha_1^{t_1 z} \cdots \alpha_n^{t_n z}\big)_{0\le\tau<T_0,\ 0\le t_i<T_i\ (1\le i\le n)}.$$

3.3 Further Extrapolation with the Auxiliary Function

In this section we shall explain how to replace, in the previous proof, Philippon's zero estimate (Proposition 3.2) by a simpler one. The idea is to extrapolate further and to prove by induction on $j = 0, 1, \ldots,$

$$G_{\underline{\sigma}}(s) = 0 \quad\text{for}\quad (\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N}$$
$$\text{with}\quad \|\underline{\sigma}\| < S_0/2^j \quad\text{and}\quad 0 \le s < S_1^{(j)},$$

with $S_1^{(0)} = S_1$ and $S_1^{(1)} = S_1'$. One cannot continue such an induction forever (in any case one needs $2^j \le S_0$). On the other hand, obviously, when the number of equations increases, it is easier to derive a contradiction by means of a zero estimate. This is not only of historical interest: our motivation is related with the problem of linear independence measures, where a short extrapolation yields weaker estimates than a longer one (see Section 5).

Baker used a variety of arguments for concluding his proofs, including clever non-vanishing results for certain determinants. In the real case one could just appeal to Pólya's Lemma 1.6. Dealing with the general case of complex algebraic numbers α_i and β_j, we shall use Tijdeman's zero estimate for exponential polynomials in one variable ([29]; see also [6], Chap. 9, Section 4, Lemma 8.9, [2], Chap. 12, Section 2, Lemma 6 and [30], Chap. 6).

Lemma 3.3. *Let $a_1, \dots, a_n$ be polynomials in $\mathbb{C}[z]$, not all of which are zero, of degrees $d_1, \dots, d_n$. Let $w_1, \dots, w_n$ be pairwise distinct complex numbers. Define*

$$\Omega = \max\{|w_1|, \dots, |w_n|\}.$$

Then for $R > 0$ the number of zeroes (counting multiplicities) of the function

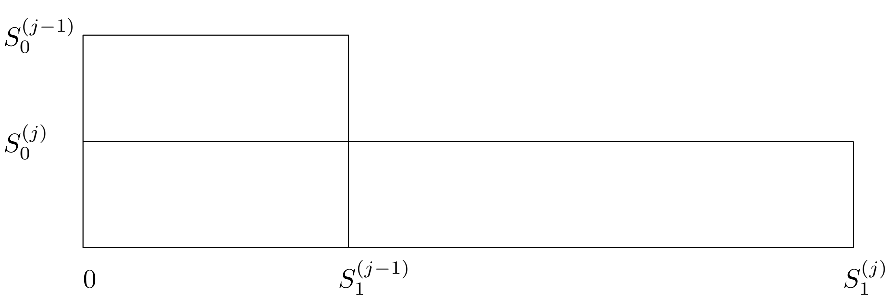
$$F(z) = \sum_{i=1}^{n} a_i(z) e^{w_i z}$$

in the disk $B_1(0, R)$ is at most $2(d_1 + \cdots + d_n + n - 1) + 5R\Omega$.

We complete the proof of Baker's Theorem 3.1 as follows. We repeat the argument of Section 3.2 b) and perform an induction on j with $1 \le j \le J$. We introduce further parameters $S_0^{(j)}$ and $S_1^{(j)}$ which are positive integers with

$$S_0^{(0)} = S_0, \quad S_1^{(0)} = S_1, \qquad S_0^{(1)} = S_0', \quad S_1^{(1)} = S_1',$$
$$S_0^{(j)} < S_0^{(j-1)} \ \text{ and } \ S_1^{(j)} > S_1^{(j-1)} \quad (1 \le j \le J).$$

One may keep in mind the following picture:



We want to prove, for $0 \le j \le J$,

$$G_{\underline{\sigma}}(s) = 0 \ \text{ for } \ \|\underline{\sigma}\| < S_0^{(j)} \ \text{ and } \ 0 \le s < S_1^{(j)}. \tag{3.10}$$

This is true by construction (Section 3.2, a) for $j = 0$, and by the first extrapolation (Section 3.2, b) for $j = 1$. Relations (3.10) for $j + 1$ follow from those for j provided that

$$\boxed{c_6(T_0 + S_0)\log\bigl(LS_1^{(j+1)}\bigr) + c_7TS_1^{(j+1)} < \bigl(S_0^{(j)} - S_0^{(j+1)}\bigr)S_1^{(j)}.}$$

Assuming (3.4), let us choose

$$J = 2n^2, \quad S_0^{(j)} = \bigl[S_0/2^j\bigr], \quad S_1^{(j)} = N^{j+1} \quad (1 \le j \le J). \tag{3.11}$$

At the end of the induction ($j = J$) we get an exponential polynomial G with a zero at each point s with $0 \le s < S_1^{(J)}$ of multiplicity $\ge S_0^{(J)}$. Since

$$\boxed{S_0^{(J)}S_1^{(J)} > 2L + c_8TS_1^{(J)},}$$

we get a contradiction with Lemma 3.3.

3.4 Upper Bound for a Determinant in Several Variables

Let $F \in \mathcal{H}_n(R)$ be a function with a zero of multiplicity $\ge S_0$ at S_1 distinct points of $B_n(0,r)\cap\mathbb{C}\underline{v}$ with $r \le R$. Then $G(z) = F(\underline{v}z)$ (which may be viewed as the restriction of F to $\mathbb{C}\underline{v}$) is a function of a single variable with S_0S_1 zeroes in $B_1(0, r/|\underline{v}|)$[9]. From Schwarz' Lemma 1.3 we deduce

$$\begin{aligned}\sup_{\underline{z}\in B_n(0,r)\cap\mathbb{C}\underline{v}} |F(\underline{z})| = |G|_{r/|\underline{v}|} &\le \left(\frac{R^2+r^2}{2rR}\right)^{-S_0S_1} |G|_{R/|\underline{v}|}\\ &\le \left(\frac{R^2+r^2}{2rR}\right)^{-S_0S_1} |F|_R.\end{aligned}$$

Unfortunately there is no similar upper bound for $|F|_r$ and we cannot use Cauchy's inequalities to bound derivatives of F in other directions than $\mathbb{C}\underline{v}$. For instance taking $n = 2$, $\underline{v} = (1,0)$, $F(z_1, z_2) = z_2^L$, we have $G(z) = F(z,0) = 0$, but

$$\left(\frac{\partial}{\partial z_2}\right)^L F(z,0) = L!.$$

According to Baker's remark (see Section 3.1, a)), for $\|\underline{\sigma}\| < S_0$, the one variable function $G_{\underline{\sigma}}(z) = \mathcal{D}^{\underline{\sigma}}F(\underline{v}z)$ has at least $(S_0 - \|\underline{\sigma}\|)S_1$ zeroes in $B_1(0, r/|\underline{v}|)$. Hence

[9] Here we do not use all the information: a zero of multiplicity $\ge S_0$ for F involves $\binom{S_0+n-1}{n}$ conditions, while for G it involves only S_0 conditions.

$$\sup_{\underline{z}\in B_n(0,r)\cap\mathbb{C}\underline{v}} |\mathcal{D}^{\underline{\sigma}}F(\underline{z})| \le \left(\frac{R^2+r^2}{2rR}\right)^{-(S_0-\|\underline{\sigma}\|)S_1} |\mathcal{D}^{\underline{\sigma}}F|_R.$$

This is the key point which explains why, in Baker's extrapolation argument, the order of derivation needs to decrease; one compensates by increasing the set of points.

Assuming for simplicity $\|\underline{\sigma}\| \le S_0/2$, the exponent of $2rR/(R^2+r^2)$ is $\ge S_0S_1/2$. For a function of n variables with S_1 zeroes of multiplicity $\ge S_0$, one should expect only an exponent $S_0S_1^{1/n}$ (up to a small absolute multiplicative constant), but the point here is that these zeroes lie on a complex line, and this explains why the exponent can be as large as a constant multiple of S_0S_1. It is interesting to compare with the interpolation determinant method: the exponent which arises more naturally (see [34], Section 10.1.4) is $S_0S_1^{1/n}$ and this is sufficient to achieve nontrivial estimates, but a refinement can be included (see [34] Proposition 10.5), so that one reaches the same exponent (namely S_0S_1, up to a constant) as with the auxiliary function.

We extend Corollary 2.4 to the situation arising in Baker's method. We deal with derivatives of functions of several variables, but the points we consider lie on a complex line $V = \mathbb{C}\underline{v} \subset \mathbb{C}^n$.

We first give a variant of Lemma 2.2 for functions of several variables.

Lemma 3.4. *Let $n \ge 1$ be an integer, V a complex subspace of dimension 1 of $\mathbb{C}^n$, R, r, ϱ and E be positive real numbers, F an element of $\mathcal{H}_n(R)$, $\underline{\zeta}_1,\dots,\underline{\zeta}_m$ pairwise distinct elements of V, $\underline{v}$ an element of V, $\sigma_1,\dots,\sigma_m$ nonnegative integers and $\underline{\kappa}$ an element of $\mathbb{N}^n$. Set*

$$M = \sum_{i=1}^m \max\{0, \sigma_i - \|\underline{\kappa}\|\}.$$

Assume

$$R \ge \varrho + \max\{r,\, |\underline{v}|\}, \quad r \ge \max_{1\le i\le m} |\underline{\zeta}_i| \;\; \text{and} \;\; 1 \le E \le \frac{(R-\varrho)^2 + r|\underline{v}|}{(R-\varrho)(r+|\underline{v}|)}.$$

Assume also that F satisfies

$$\mathcal{D}^{\underline{\sigma}}F(\underline{\zeta}_i) = 0 \;\; \text{for} \;\; \underline{\sigma}\in\mathbb{N}^n \;\; \text{with} \;\; \|\underline{\sigma}\| < \sigma_i \;\; \text{and} \;\; 1\le i\le m.$$

Then

$$|\mathcal{D}^{\underline{\kappa}}F(\underline{v})| \le \underline{\kappa}!\varrho^{-\|\underline{\kappa}\|}E^{-M}|F|_R.$$

Proof. Let $\underline{v}_0 \in V$ satisfy $|\underline{v}_0| = 1$. Define a function $G \in \mathcal{H}_1(R)$ of a single variable z by

$$G(z) = \mathcal{D}^{\underline{\kappa}}F(\underline{v}_0 z).$$

Define also $\zeta_1,\dots,\zeta_m$ in $B_1(0,r)$ by $\underline{\zeta}_i = \underline{v}_0\zeta_i$ $(1\le i\le m)$. The formula

$$G^{(k)}(z) = \sum_{\|\underline{\tau}\|=k} \frac{k!}{\underline{\tau}!} \underline{v}_0^{\underline{\tau}} \mathcal{D}^{\underline{\kappa}+\underline{\tau}} F(\underline{v}_0 z) \qquad (k \geq 0)$$

shows that G has a zero of multiplicity $\geq \max\{0, \sigma_i - \|\underline{\kappa}\|\}$ at ζ_i for $1 \leq i \leq m$. According to Lemma 1.3 (with r replaced by $|\underline{v}|$ and R by $R - \varrho$) we have

$$|G|_{|\underline{v}|} \leq E^{-M} |G|_{R-\varrho}.$$

We derive the conclusion from Cauchy's inequalities:

$$|G|_{R-\varrho} = \sup_{|z|=R-\varrho} |\mathcal{D}^{\underline{\kappa}} F(\underline{v}_0 z)| \leq \underline{\kappa}! \varrho^{-\|\underline{\kappa}\|} |F|_R.$$

□

Remark. For $n = 1$ Lemma 3.4 contains Lemma 1.3, but not Lemma 2.2 when $\kappa > 0$.

The next result is a variant of Proposition 2.3 for functions of several variables.

Proposition 3.5. *Let $n \geq 1$ be an integer, V a complex line in $\mathbb{C}^n$, R a positive real number, $\varphi_1, \ldots, \varphi_L$ elements of $\mathcal{H}_n(R)$, $\underline{\xi}_1, \ldots, \underline{\xi}_L$ elements of $V \cap B_n(0, R)$ and $\underline{\kappa}_1, \ldots, \underline{\kappa}_L$ elements of $\mathbb{N}^n$. Consider the determinant*

$$\Delta = \det \Big(\mathcal{D}^{\underline{\kappa}_\mu} \varphi_\lambda (\underline{\xi}_\mu) \Big)_{1 \leq \lambda, \mu \leq L}.$$

Let $m_1, \ldots, m_L$ be nonnegative integers and, for $1 \leq \mu \leq L$ and $1 \leq i \leq m_\mu$, let $\underline{\zeta}_{\mu i}$ be an element of V and $\sigma_{\mu i}$ a nonnegative integer. We assume that for each $\mu = 1, \ldots, L$, the m_μ elements $\underline{\zeta}_{\mu 1}, \ldots, \underline{\zeta}_{\mu m_\mu}$ are pairwise distinct. Set

$$M_\mu = \sum_{i=1}^{m_\mu} \max\{0, \sigma_{\mu i} - \|\underline{\kappa}_\mu\|\} \quad (1 \leq \mu \leq L).$$

For $1 \leq \mu \leq L$, let r_μ, R_μ, ϱ_μ and E_μ be positive real numbers satisfying

$$R \geq R_\mu \geq \varrho_\mu + \max\{r_\mu, |\underline{\xi}_\mu|\}, \quad r_\mu \geq \max_{1 \leq i \leq m_\mu} |\underline{\zeta}_{\mu i}|$$

and

$$1 \leq E_\mu \leq \frac{(R_\mu - \varrho_\mu)^2 + r_\mu |\underline{\xi}_\mu|}{(R_\mu - \varrho_\mu)(r_\mu + |\underline{\xi}_\mu|)}.$$

Denote by Φ the analytic mapping

$$(\varphi_1, \ldots, \varphi_L) \colon B_n(0, R) \to \mathbb{C}^L.$$

Assume that for any $(\mu, i, \underline{\kappa}) \in \mathbb{N}^{2+n}$ satisfying $1 \leq \mu \leq L$, $1 \leq i \leq m_\mu$ and $\|\underline{\kappa}\| < \sigma_{\mu i}$, the μ vectors

$$\mathcal{D}^{\underline{\kappa}_1}\Phi(\underline{\xi}_1),\ldots,\mathcal{D}^{\underline{\kappa}_{\mu-1}}\Phi(\underline{\xi}_{\mu-1}),\ \mathcal{D}^{\underline{\kappa}}\Phi(\underline{\zeta}_{\mu i}) \tag{3.12}$$

in $\mathbb{C}^L$ *are linearly dependent. Then*

$$|\Delta| \le L!\max_{\tau\in\mathfrak{S}_L}\prod_{\mu=1}^{L}\left(\underline{\kappa}_\mu!\varrho_\mu^{-\|\underline{\kappa}_\mu\|}E_\mu^{-M_\mu}|\varphi_{\tau(\mu)}|_{R_\mu}\right).$$

Proof. We prove Proposition 3.5 by induction on L. For $L=1$ we have $\Phi=\varphi_1$,

$$\Delta=\mathcal{D}^{\underline{\kappa}_1}\Phi(\underline{\xi}_1),\qquad 1\le E_1\le\frac{(R_1-\varrho_1)^2+r_1|\underline{\xi}_1|}{(R_1-\varrho_1)(r_1+|\underline{\xi}_1|)}$$

and hypothesis (3.12) reads

$$\mathcal{D}^{\underline{\kappa}}\Phi(\underline{\zeta}_{1i})=0\quad\text{for}\quad\|\underline{\kappa}\|<\sigma_{1i}\quad\text{and}\quad 1\le i\le m_1.$$

From Lemma 3.4 we deduce

$$|\Delta|\le\underline{\kappa}_1!\varrho_1^{-\|\underline{\kappa}_1\|}E_1^{-M_1}|\varphi_1|_{R_1}.$$

Hence Proposition 3.5 is true in case $L=1$.

Assume now that the conclusion is true for L replaced by $L-1$. Define $F\in B_n(0,R)$ by

$$F(\underline{z})=\det\left(\mathcal{D}^{\underline{\kappa}_1}\Phi(\underline{\xi}_1),\ldots,\mathcal{D}^{\underline{\kappa}_{L-1}}\Phi(\underline{\xi}_{L-1}),\ \Phi(\underline{z})\right). \tag{3.13}$$

By assumption (3.12) with $\mu=L$, for $1\le i\le m_L$ and $\|\underline{\kappa}\|<\sigma_{Li}$, we have

$$\mathcal{D}^{\underline{\kappa}}F(\underline{\zeta}_{Li})=0.$$

Since

$$\Delta=\mathcal{D}^{\underline{\kappa}_L}F(\underline{\xi}_L),$$

we deduce from Lemma 3.4

$$|\Delta|\le\underline{\kappa}_L!\varrho_L^{-\|\underline{\kappa}_L\|}E_L^{-M_L}|F|_{R_L}.$$

We expand the determinant in the right hand side of (3.13) with respect to the last column: define, for $1\le\lambda\le L$,

$$\Phi_\lambda=\left(\varphi_1,\ldots,\varphi_{\lambda-1},\varphi_{\lambda+1},\ldots,\varphi_L\right)\colon B_n(0,R)\to\mathbb{C}^{L-1}$$

and let Δ_λ denote the determinant of the $(L-1)\times(L-1)$ matrix

$$\left(\mathcal{D}^{\underline{\kappa}_1}\Phi_\lambda(\underline{\xi}_1),\ldots,\mathcal{D}^{\underline{\kappa}_{L-1}}\Phi_\lambda(\underline{\xi}_{L-1})\right).$$

We have

$$F(\underline{z}) = \sum_{\lambda=1}^{L} (-1)^{L-\lambda} \varphi_\lambda(\underline{z}) \Delta_\lambda,$$

hence

$$|F|_{R_L} \le L \max_{1\le\lambda\le L} |\varphi_\lambda|_{R_L} |\Delta_\lambda|$$

and therefore

$$|\Delta| \le \underline{\kappa}_L! L \varrho_L^{-\|\underline{\kappa}_L\|} E_L^{-M_L} \max_{1\le\lambda\le L} |\varphi_\lambda|_{R_L} |\Delta_\lambda|.$$

We fix an index $\lambda^0 \in \{1, \dots, L\}$ such that

$$|\Delta| \le \underline{\kappa}_L! L \varrho_L^{-\|\underline{\kappa}_L\|} E_L^{-M_L} |\varphi_{\lambda^0}|_{R_L} |\Delta_{\lambda^0}|.$$

Using the induction hypothesis, we deduce that there exists a bijective map t from $\{1, \dots, L-1\}$ to $\{1, \dots, \lambda^0 - 1, \lambda^0 + 1, \dots, L\}$ such that

$$|\Delta_{\lambda^0}| \le (L-1)! \prod_{\mu=1}^{L-1} \left(\underline{\kappa}_\mu! \varrho_\mu^{-\|\underline{\kappa}_\mu\|} E_\mu^{-M_\mu} |\varphi_{t(\mu)}|_{R_\mu} \right).$$

Define $\tau \in \mathfrak{S}_L$ by $\tau(\mu) = t(\mu)$ for $1 \le \lambda < L$ and $\tau(L) = \lambda^0$. Proposition 3.5 follows. $\square$

We shall use a special case of Proposition 3.5.

Given a sequence $(\underline{\sigma}_\nu, \zeta_\nu)_{0\le\nu\le N}$ of elements in $\mathbb{N}^n \times \mathbb{C}$, an index ν in the range $0 \le \nu \le N$ and a complex number ζ, we define the *weight* $w_\nu(\zeta)$ *of index* ν *of* ζ *in this sequence* as follows: $w_\nu(\zeta) = 0$ if $\zeta_i \ne \zeta$ for $0 \le i \le \nu$ and otherwise

$$w_\nu(\zeta) = \max\bigl\{ \|\underline{\sigma}\| \; ; \; \exists i \text{ with } 0 \le i \le \nu \text{ such that } (\underline{\sigma}_i, \zeta_i) = (\underline{\sigma}, \zeta) \bigr\}. \quad (3.14)$$

We consider a sequence $(\underline{\sigma}_\nu, \zeta_\nu)_{0\le\nu\le N}$ which satisfies the following property:

$$\begin{aligned} &\text{For } 0 \le \nu \le N \text{ and for any } \underline{\sigma} \in \mathbb{N}^n \text{ satisfying } \|\underline{\sigma}\| < \|\underline{\sigma}_\nu\|, \\ &\qquad\qquad \text{there exists } i \text{ with } 0 \le i < \nu \text{ such that } (\underline{\sigma}_i, \zeta_i) = (\underline{\sigma}, \zeta_\nu). \end{aligned} \quad (3.15)$$

Such a sequence will be called *admissible.*

In the case $n = 1$ the sequence $(\sigma_\nu, \zeta_\nu)_{0\le\nu\le N}$ which occurred in the hypotheses of Corollary 2.4 is admissible and satisfies

$$w_\nu(\zeta_\nu) = \sigma_\nu \quad \text{for} \quad 0 \le \nu \le N.$$

For an admissible sequence $(\underline{\sigma}_\nu, \zeta_\nu)_{0\le\nu\le N}$, for an analytic function F of n variables and for $\underline{v} \in \mathbb{C}^n$, the $N+1$ conditions

$$\mathcal{D}^{\underline{\sigma}_\nu} F(\underline{v}\zeta_\nu) = 0 \quad \text{for} \quad 0 \le \nu \le N \quad (3.16)$$

imply that the one-variable function $G(z)=F(\underline{v}z)$ has a zero of multiplicity $\geq w_N(\zeta)$ at each point $\zeta\in\mathbb{C}$. Moreover, for any $\underline{\tau}\in\mathbb{N}^n$, the same conditions (3.16) imply that the function $G_{\underline{\tau}}(z)=\mathcal{D}^{\underline{\tau}}F(\underline{v}z)$ has a zero at ζ of multiplicity

$$\geq \max\bigl\{0, w_N(\zeta)-\|\underline{\tau}\|\bigr\}.$$

Corollary 3.6. *Let n and L be positive integers, $\varphi_1,\dots,\varphi_L$ elements of $\mathcal{H}_n(R)$, $0=\nu_0\leq\nu_1<\cdots<\nu_L\leq N$ nonnegative integers, $\underline{\sigma}_0,\dots,\underline{\sigma}_N$ elements of $\mathbb{N}^n$, $\zeta_0,\dots,\zeta_N$ complex numbers and $\underline{v}\in\mathbb{C}^n\setminus\{0\}$. Consider the determinant*

$$\Delta=\det\left(\mathcal{D}^{\underline{\sigma}_{\nu_\mu}}\varphi_\lambda(\underline{v}\zeta_{\nu_\mu})\right)_{1\leq\lambda,\mu\leq L}.$$

Assume $(\underline{\sigma}_\nu,\zeta_\nu)_{0\leq\nu\leq N}$ is an admissible sequence. For $0\leq\nu\leq N$ define

$$M_\nu=\sum_{\zeta\in B_1(0,R/|\underline{v}|)}\max\bigl\{0,w_\nu(\zeta)-\|\underline{\sigma}_\nu\|\bigr\}.$$

For $1\leq\mu\leq L$, let r_μ, R_μ and ϱ_μ be positive real numbers satisfying

$$R\geq R_\mu\geq\varrho_\mu+\max\{r_\mu,\,|\underline{v}\zeta_{\nu_\mu}|\}\quad\textit{and}\quad r_\mu\geq\max_{0\leq\nu<\nu_\mu}|\underline{v}\zeta_\nu|.$$

Let E_μ satisfy

$$1\leq E_\mu\leq\frac{(R_\mu-\varrho_\mu)^2+r_\mu|\underline{v}\zeta_{\nu_\mu}|}{(R_\mu-\varrho_\mu)(r_\mu+|\underline{v}\zeta_{\nu_\mu}|)}.$$

Denote by Φ the analytic mapping

$$(\varphi_1,\dots,\varphi_L)\colon B_n(0,R)\to\mathbb{C}^L.$$

Assume that for $0\leq\mu\leq L-1$ and for $\nu_\mu\leq\nu<\nu_{\mu+1}$ the system of $\mu+1$ vectors

$$\mathcal{D}^{\underline{\sigma}_{\nu_1}}\Phi(\underline{v}\zeta_{\nu_1}),\dots,\mathcal{D}^{\underline{\sigma}_{\nu_\mu}}\Phi(\underline{v}\zeta_{\nu_\mu}),\ \mathcal{D}^{\underline{\sigma}_\nu}\Phi(\underline{v}\zeta_\nu)$$

in $\mathbb{C}^L$ is linearly dependent. Then

$$|\Delta|\leq L!\max_{\tau\in\mathfrak{S}_L}\prod_{\mu=1}^{L}\left(\underline{\sigma}_{\nu_\mu}!\varrho_\mu^{-\|\underline{\sigma}_{\nu_\mu}\|}E_\mu^{-M_{\nu_\mu}}|\varphi_{\tau(\mu)}|_{R_\mu}\right).$$

Remark. This statement does not contain Corollary 2.4, because the exponent M_ν is usually smaller than

$$\sum_{\zeta\in B_1(0,R/|\underline{v}|)}w_\nu(\zeta).$$

It does not seem clear how to get a result containing both the one variable Proposition 2.3 and Corollary 3.6. It might be easier to combine Corollary 3.6 with Proposition 10.5 of [34].

Proof. We apply Proposition 3.5 with $V = \mathbb{C}\underline{v}$,

$$\underline{\xi}_\mu = \underline{v}\zeta_{\nu_\mu} \quad \text{and} \quad \underline{\kappa}_\mu = \underline{\sigma}_{\nu_\mu} \quad (1 \le \mu \le L).$$

We define m_μ, $\underline{\zeta}_{\mu i}$ and $\sigma_{\mu i}$ as follows: for $1 \le \mu \le L$, we consider the sequence $(\zeta_0, \dots, \zeta_{\nu_\mu})$ given by the hypotheses of Corollary 3.6, we denote by m_μ the number of distinct elements in this sequence, by $\zeta_{\mu i}$ these distinct elements, by $\sigma_{\mu i}$ the weight of index ν_μ of $\zeta_{\mu i}$ in this sequence, and we set $\underline{\zeta}_{\mu i} = \underline{v}\zeta_{\mu i}$. Therefore

$$\sum_{\zeta \in B_1(0, R/|\underline{v}|)} \max\{0, w_{\nu_\mu}(\zeta) - \|\underline{\sigma}_{\nu_\mu}\|\} \ge \sum_{i=1}^{m_\mu} \max\{0, \sigma_{\mu i} - \|\underline{\kappa}_\mu\|\} \quad (1 \le \mu \le L).$$

Corollary 3.6 follows. □

3.5 Extrapolation with an Interpolation Determinant

We give a proof of Baker's Theorem 3.1 by means of an interpolation determinant and an extrapolation; the zero estimate which will enable us to get the conclusion is Tijdeman's Lemma 3.3.

Let T_0, $T_1, \dots, T_n$, J, $S_0^{(j)}$ and $S_1^{(j)}$ $(0 \le j \le J)$ be positive integers. Assume that the sequence $(S_0^{(j)})_{0 \le j \le J}$ is decreasing and that the sequence $(S_1^{(j)})_{0 \le j \le J}$ increases. We write S_0 and S_1 for $S_0^{(0)}$ and $S_1^{(0)}$ respectively and we set $L = T_0 \cdots T_n$, $T = T_1 + \cdots + T_n$.

We denote by $\{\varphi_1, \dots, \varphi_L\}$ the L exponential monomials

$$z^\tau \alpha_1^{t_1 z} \cdots \alpha_n^{t_n z} \quad \text{for} \quad 0 \le \tau < T_0 \quad \text{and} \quad 0 \le t_i < T_i \ (1 \le i \le n).$$

For $0 \le j \le J$, define $\mathcal{S}_j$ as the set of $(\underline{\sigma}, s)$ in $\mathbb{N}^n \times \mathbb{N}$ satisfying $\|\underline{\sigma}\| < S_0^{(j)}$ and $0 \le s < S_1^{(j)}$. We need to choose an ordering $(\underline{\sigma}_\nu, s_\nu)_{0 \le \nu < N_J}$ on the union $\mathcal{S}$ of these sets. Define

$$N_{-1} = 0, \quad N_0 = \binom{S_0 + n - 1}{n} S_1$$

and, for $1 \le j \le J$,

$$N_j = N_{j-1} + \binom{S_0^{(j)} + n - 1}{n} (S_1^{(j)} - S_1^{(j-1)}).$$

We choose an ordering so that

$$\bigl\{(\underline{\sigma}_\nu, s_\nu)\ ;\ 0 \le \nu < N_0\bigr\} = \bigl\{(\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N}\ ;\ \|\underline{\sigma}\| < S_0 \ \text{ and } \ 0 \le s < S_1\bigr\}$$

and, for $1 \le j \le J$,

$$\begin{aligned}\bigl\{(\underline{\sigma}_\nu, s_\nu)\ ;\ N_{j-1} \le \nu < N_j\bigr\}\\ = \bigl\{(\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N}\ ;\ \|\underline{\sigma}\| < S_0^{(j)} \quad\text{and}\quad S_1^{(j-1)} \le s < S_1^{(j)}\bigr\}.\end{aligned}$$

It remains to specify the order inside each interval $N_{j-1} \le \nu < N_j$ $(0 \le j \le J)$. There is no complete rigidity, but it suffices to say that we choose an order for which the resulting sequence is admissible (see (3.15)).

For each τ, $\underline{t}$, $\underline{\sigma}$ and s, consider the algebraic number $\gamma_{\tau\underline{t}}^{(\underline{\sigma}s)}$ given by (3.5) and build up the matrix

$$\mathtt{M} = \Bigl(\gamma_{\tau\underline{t}}^{(\underline{\sigma}_\nu s_\nu)}\Bigr)_{\substack{(\tau,\underline{t})\\ 0\le\nu<N_J}}$$

with L rows indexed by $(\tau, \underline{t})$ with $0 \le \tau < T_0$ and $0 \le t_i < T_i$ $(1 \le i \le n)$, and N_J columns indexed by ν with $0 \le \nu < N_J$.

We first deduce from Lemma 3.3 that $\mathtt{M}$ has rank L. Indeed otherwise there exists a nonzero polynomial $P \in \mathbb{C}[Y_0, \dots, Y_n]$ for which

$$\partial^{\underline{\sigma}_\nu} P\bigl(s_\nu, \alpha_1^{s_\nu}, \dots, \alpha_n^{s_\nu}\bigr) = 0 \quad\text{for}\quad 0 \le \nu < N_J,$$

which means

$$\partial^{\underline{\sigma}} P\bigl(s, \alpha_1^{s}, \dots, \alpha_n^{s}\bigr) = 0 \quad\text{for}\quad (\underline{\sigma}, s) \in \mathcal{S}.$$

In particular the function $G(z) = P(z, \alpha_1^z, \dots, \alpha_n^z)$ satisfies

$$G^{(k)}(s) = 0 \quad\text{for}\quad 0 \le k < S_0^{(J)} \quad\text{and}\quad 0 \le s < S_1^{(J)}.$$

In order to apply Lemma 3.3, we assume

$$\boxed{S_0^{(J)} S_1^{(J)} > 2L + c_8 T S_1^{(J)}.}$$

Since $\mathtt{M}$ has maximal rank L, one can select L columns which produce a nonzero determinant Δ. We select them minimal as in the proof of Section 2.2 and we write

$$\Delta = \det\Bigl(\gamma_{\tau\underline{t}}^{(\underline{\sigma}_{\nu_\mu} s_{\nu_\mu})}\Bigr)_{\substack{(\tau,\underline{t})\\ 1\le\mu\le L}}.$$

We want to derive from Corollary 3.6 an upper bound for $|\Delta|$. Set

$$\underline{v} = (1, \log\alpha_1, \dots, \log\alpha_{n-1}).$$

We need first to estimate the weights (3.14) related to our sequence

$$(\underline{\sigma}_\nu, s_\nu)_{0\le\nu<N_J}.$$

Let ν be an index with $N_0 \le \nu < N_J$; define j in the range $1 \le j \le J$ by $N_{j-1} \le \nu < N_j$. Thanks to the construction of the sequence $(\underline{\sigma}_\nu, s_\nu)_{0\le\nu<N_J}$, we have

$$w_\nu(s) = S_0^{(j-1)} - 1 \quad \text{for} \quad 0 \le s < S_1^{(j-1)}.$$

Since $\|\underline{\sigma}_\nu\| < S_0^{(j)}$, it follows that the number

$$M_\nu = \sum_{\zeta \in B_1(0, R/|\underline{v}|)} \max\bigl\{0, w_\nu(\zeta) - \|\underline{\sigma}_\nu\|\bigr\}$$

satisfies

$$M_\nu \ge \bigl(S_0^{(j-1)} - S_0^{(j)}\bigr) S_1^{(j-1)}. \tag{3.17}$$

For each μ with $1 \le \mu \le L$ we define $j_\mu \in \{0, \ldots, J\}$ by

$$N_{j_\mu - 1} \le \nu_\mu < N_{j_\mu}.$$

Since $N_{-1} = 0$, for $\nu_\mu < N_0$ we have $j_\mu = 0$. Hence for any $\mu = 1, \ldots, L$, we have

$$\|\underline{\sigma}_{\nu_\mu}\| < S_0^{(j_\mu)} \quad \text{and} \quad 0 \le s_{\nu_\mu} < S_1^{(j_\mu)}.$$

We want to use (3.17) for each ν_μ with μ in the range $L/2 < \mu \le L$, so we need to check $\nu_\mu \ge N_0$ for these μ. For this reason we require our parameters to satisfy

$$\boxed{2\binom{S_0 + n - 1}{n} S_1 \le L.} \tag{3.18}$$

Since

$$\Delta = \det\bigl(\mathcal{D}^{\underline{\sigma}_{\nu_\mu}} \varphi_\lambda(\underline{v} s_{\nu_\mu})\bigr)_{1 \le \lambda, \mu \le L},$$

we may apply Corollary 3.6 with

$$r_\mu = c_9 S_1^{(j_\mu)}, \quad \varrho_\mu = 1, \quad R_\mu = 3 e r_\mu, \quad E_\mu = e.$$

We deduce

$$\begin{aligned} |\Delta| &\le L! \prod_{\mu=1}^{L} \Bigl(\underline{\sigma}_{\nu_\mu}! e^{-M_{\nu_\mu}} R_\mu^{T_0} e^{c_{10} T R_\mu}\Bigr) \\ &\le L! \exp\left\{\sum_{\mu=1}^{L} \Bigl(-M_{\nu_\mu} + c_{11}\bigl(T_0 + S_0\bigr) \log\bigl(L S_1^{(j_\mu)}\bigr) + c_{12} T S_1^{(j_\mu)}\Bigr)\right\} \end{aligned}$$

(recall that $S_0^{(j)} \le S_0$ for $0 \le j \le J$).

Since Δ is a nonzero number in the field $\mathbb{K}$, Liouville's Proposition 1.13 produces a lower bound for $|\Delta|$:

$$|\Delta| \ge L!^{-D+1} \exp\left\{\sum_{\mu=1}^{L} \Bigl(-c_{13}(T_0 + S_0) \log\bigl(L S_1^{(j_\mu)}\bigr) - c_{14} T S_1^{(j_\mu)}\Bigr)\right\}.$$

Define, for $1 \le \mu \le L$,

$$Q_\mu = c_{15}(T_0 + S_0)\log\bigl(LS_1^{(j_\mu)}\bigr) + c_{16}TS_1^{(j_\mu)}.$$

We assume our parameters are selected so that

$$\boxed{2Q_\mu < \left(S_0^{(j_\mu-1)} - S_0^{(j_\mu)}\right) S_1^{(j_\mu-1)}.}$$

Then we have

$$M_{\nu_\mu} \ge 0 \quad \text{for} \quad 1 \le \mu < \frac{L}{2} - 1,$$

$$\sum_{1\le\mu\le L/2} Q_\mu \le \sum_{L/2<\mu\le L} Q_\mu$$

and by (3.17)

$$M_{\nu_\mu} > 2Q_\mu \quad \text{for} \quad \frac{L}{2} - 1 \le \mu \le L.$$

Therefore

$$\sum_{\mu=1}^{L}(M_{\nu_\mu} - Q_\mu) > 0,$$

and the contradiction follows.

It remains to select our parameters. We take the same values as in Section 3.3, namely (3.4) and (3.11):

$$T_0 = 2[\mathbb{K}:\mathbb{Q}]N^{2n+1}, \quad T_1 = \cdots = T_n = N^{2n-1}, \quad S_0 = N^{2n+1}, \quad S_1 = N$$

and

$$J = 2n^2, \qquad S_0^{(j)} = \bigl[S_0/2^j\bigr], \qquad S_1^{(j)} = N^{j+1} \quad (1 \le j \le J).$$

4 Fourth Lecture. Introduction to Diophantine Approximation

4.1 On a Conjecture of Mahler

Consider the successive powers of e, namely

$$e,\ e^2,\ e^3,\ e^4,\ \dots,$$

and their distances to the nearest integer

$$\|e\|_{\mathbb{Z}},\ \|e^2\|_{\mathbb{Z}},\ \|e^3\|_{\mathbb{Z}},\ \|e^4\|_{\mathbb{Z}},\ \dots$$

It is an open problem to prove that these numbers are equidistributed[10] in the interval $(0, 1/2)$. We are interested in estimating $\Psi(B) = \min_{1\le b\le B} \|e^b\|_{\mathbb{Z}}$ as $B \to \infty$ from below, but let us first say one word on upper bounds. Essentially nothing is known: it is not yet proved that $\Psi(B)$ tends to 0 when $B \to \infty$:

(?) *For any $\epsilon > 0$ there exists $b \in \mathbb{N}$ such that $\|e^b\|_{\mathbb{Z}} < \epsilon$,*

but it is expected that $B\Psi(B)$ tends to 0 when $B \to \infty$:

(?) *For any $\epsilon > 0$ there exists $b \in \mathbb{N}$ such that $\|e^b\|_{\mathbb{Z}} < \dfrac{\epsilon}{b}\cdot$*

We come back to lower bounds for $\Psi(B)$. In [17] p. 397, K. Mahler says:

- *... one can easily show that*

$$|\log a - b| < \frac{1}{a}$$

for an infinite increasing sequence of positive integers a and suitable integers b.

Indeed, this inequality holds for each pair (a, b) of positive integers for which $\log a < b < \log(a + 1)$.

For a given positive integer B, if the B numbers

$$\|e^b\|_{\mathbb{Z}} \quad (1 \le b \le B)$$

were evenly distributed in the interval $(0, 1/2)$, the smallest of them would not be less than a constant times $1/B$. This is likely to be too optimistic for a conjecture, but a more reasonable question is:

- *Is there an absolute constant $\kappa \ge 1$ such that*

$$|e^b - a| \ge \frac{1}{b^\kappa} \tag{4.1}$$

for any positive integers a and b with $b \ge 2$?

A straightforward computation shows that this inequality holds with $\kappa = 2.25$ for $2 \le b \le 100$.

If $|e^b - a|$ is small, then b is close to $\log a$, hence

$$e^b - a = a\bigl(e^{b-\log a} - 1\bigr)$$

is close to $a(b - \log a)$. One easily deduces that (4.1) is equivalent to:

[10] A result of Koksma ([10], Chap. VIII, Section 3, N° 12 Satz 16) states that for almost all real $\theta \ge 1$ (in the sense of Lebesgue measure) the sequence of fractional parts of θ^k is uniformly distributed modulo 1. On the other hand there there is no known example of a transcendental number which satisfies this property.

- *Is there an absolute constant $\kappa \geq 1$ such that*
$$|b - \log a| \geq \frac{1}{a(\log a)^\kappa} \tag{4.2}$$
for any positive integers a and b with $a \geq 3$?

A symmetric refinement of both (4.1) and (4.2) could be formulated by restricting to large values of a and b, say $a \geq a_0$ and $b \geq b_0$, and by introducing functions $x(\log x)(\log_2 x)\cdots(\log_{k-1} x)(\log_k x)^\kappa$, where $\log_2 = \log\log$ and $\log_k$ is the k-th iterated logarithm; but since (4.1) and (4.2) are not yet known, it seems more reasonable to consider weaker statements rather than stronger ones!

A weaker estimate than (4.1) and (4.2), which is still an open question, is usually attributed to Mahler (for instance in [34] Chap. 14, open problems and [35]) but is not explicit in [17] nor in [18].

Conjecture 4.1. Does there exist a positive absolute constant c such that, for any positive rational integers a and b with $a \geq 2$,
$$|e^b - a| \geq a^{-c} \ ?$$

The same argument as above shows that Conjecture 4.1 amounts to the existence of $c' > 1$ with
$$|b - \log a| \overset{?}{\geq} e^{-c'b}.$$
If we restrict to sufficiently large a and b, then c' is essentially $c+1$.

The best known estimate in the direction of Conjecture 4.1 is due to Mahler himself [17]:
$$|e^b - a| \geq a^{-c\log\log a} \quad \text{and} \quad |b - \log a| \geq b^{-cb} \tag{4.3}$$
for $a \geq 3$ and $b \geq 2$, with an absolute constant $c > 0$. Assuming a is sufficiently large, K. Mahler in [17] gave a sharp explicit numerical value for c, namely $c = 40$ (for both estimates), which he refined in [18], getting $c = 33$. A further refinement is due to F. Wielonsky [36]: for sufficiently large a, Mahler's estimates (4.3) hold with $c = 20$.

Fel'dman proved several lower bounds for $|e^\beta - \alpha|$ and $|\beta - \log\alpha|$ when α and β are algebraic numbers. References are given in [21], where a refinement including most previously known results was established.

The following result is a slight refinement of the main estimate of [21] (apart from the fact that our constant c is not explicit).

Theorem 4.2. *There is an absolute constant $c > 0$ with the following property. Let α, β be two nonzero algebraic numbers and λ a logarithm of α. Define $\mathbb{K} = \mathbb{Q}(\alpha, \beta)$ and $D = [\mathbb{K} : \mathbb{Q}]$. Let A, B and E be positive real numbers satisfying $E \geq e$,*

$$\log A \geq \max\bigl\{\mathrm{h}(\alpha)\,,\; D^{-1}\log E\,,\; E|\lambda|/D\bigr\}$$

and[11]

$$\log B \geq \mathrm{h}(\beta) + \log(D\log A) + \log D + \frac{1}{D}\log E. \tag{4.4}$$

Then

$$|\beta-\lambda| \geq \exp\Bigl\{-cD^2(\log A)(\log B)\bigl(D\log D+\log E\bigr)(\log E)^{-2}\Bigr\}.$$

According to Theorem 5 of [21], if we replace the condition (4.4) on B by

$$\log B \geq \mathrm{h}(\beta) + \log_+\log A + \log D + \log E,$$

then $c = 105\,500$ is an admissible value. It is not a big challenge to produce a smaller numerical value; for instance, taking $D=1$ and $E=e$, one deduces from [21] that for any positive rational numbers a and b,

$$|e^b-a| \geq \exp\Bigl\{-1000\bigl(\mathrm{h}(b)+\log\log A+12\bigr)\bigl(\log A+12\bigr)\Bigr\},$$

where $\log A = \max\{1, \mathrm{h}(a)\}$. In particular we recover Mahler's results (4.3) on $|e^b-a|$ and $|b-\log a|$ as a consequence of Theorem 4.2.

On the other hand, any further improvement of Theorem 4.2 in terms of either A, B, D or E seems to require a new idea.

For the proof of Theorem 4.2, a refinement of Propositions 1.7 and 1.8 is needed. To begin with, assume for simplicity that $\alpha=a$ and $\beta=b$ are positive integers. As we shall see, the condition $f(e^b,b)\neq 0$, which occurs in both Propositions 1.7 and 1.8, does not suffice any more, but we need $f(a,b)\neq 0$. This is achieved by the following zero estimate due to Yu.V. Nesterenko.
Let b be a complex number. The derivation $\mathrm{D}_b = (\partial/\partial X) + bY(\partial/\partial Y)$, on the ring of polynomials in two variables X and Y, has the following remarkable property: for a polynomial $P\in\mathbb{C}[X,Y]$, the derivative of the complex function $P(z,e^{bz})$ is the exponential polynomial $(\mathrm{D}_bP)(z,e^{bz})$.

Here is Lemma 2 of [21] (see also [34] Prop. 2.14).

Lemma 4.3. *Assume $b\neq 0$. Let T_0, T_1, S_0 and S_1 be positive integers satisfying*

$$S_0S_1 > (T_0+S_1-1)T_1.$$

Let $(\xi_0,\eta_0),\dots,(\xi_{S_1-1},\eta_{S_1-1})$ be elements in $\mathbb{C}\times\mathbb{C}^\times$ with $\xi_0,\dots,\xi_{S_1-1}$ pairwise distinct. Then there is no nonzero polynomial $P\in\mathbb{C}[X,Y]$, of degree $< T_0$ in X and of degree $< T_1$ in Y, which satisfies

$$\mathrm{D}_b^\sigma P(\xi_s,\eta_s) = 0 \quad \textit{for} \quad 0\leq s<S_1 \quad \textit{and} \quad 0\leq\sigma<S_0.$$

[11] It does not make any difference if one omits the summand $\log(D\log A)$ in (4.4), provided that one replaces the factor $(\log B)$ in the conclusion by $(\log B + \log(D\log A))$.

Remark. An obvious necessary condition for the nonexistence of P is

$$S_0S_1 \geq T_0T_1.$$

Indeed if $S_0S_1 < T_0T_1$, then the homogeneous linear system of equations given by the conditions $\mathrm{D}_b^\sigma P(\xi_\mu, \eta_\mu) = 0$ has a nontrivial solution, because the number of unknowns (which are the coefficients of P) is larger than the number of equations. On the other hand, if $\eta_0 = \cdots = \eta_{S_1-1}$ and $S_0 < T_1$, then

$$(Y - \eta_0)^{S_0}$$

is a nontrivial solution.

Hence one cannot replace $T_0 + S_1 - 1$ in the hypothesis of Lemma 4.3 by a smaller number than $\max\{T_0\,, S_1\}$.

Sketch of proof of Theorem 4.2 . We first consider the special case where $\alpha = a$ and $\beta = b$ are positive integers with, say, $b \geq 3$. Our goal is to check

$$|b - \log a| \geq \exp\{-cb(\log b)\} \tag{4.5}$$

for some absolute constant $c > 0$ (which we shall not compute explicitly). Recall that when b and $\log a$ are close together, then $b(\log b)$ is close to $(\log a)(\log b)$.

Here the parameter E of Theorem 4.2 is not helpful and we shall just take $E = e$, but we shall keep it for a while to show why we cannot get a better result by taking a large quotient R/r in Schwarz' Lemma.

If we were following the proof of Proposition 1.8 in Section 1.6, we would get $b(\log b)^2$ instead of $b \log b$ in (4.5). Let us explain why.

As we shall see in (4.10), the conclusion will have the shape

$$|b - \log a| \geq E^{-L}.$$

Set $U = V = W = (1/36)L \log E$, so that (1.6) holds (for simplicity we forget the upper bound for the length of f). In order to check (1.5) we need the number $L \log E$ to be larger than bET_1S_1 and than $T_0 \log(bE)$. Since $L = T_0T_1$, we get the conditions

$$T_0 \log E > bES_1 \quad \text{and} \quad T_1 \log E > \log b.$$

Writing now $S_0 = T_0T_1/S_1$, we obtain

$$S_0 > b(\log b)E(\log E)^{-2}.$$

In particular this explains our choice $E = e$.

The conclusion of Proposition 1.8 involves $S_0!$, so we shall require $L > \log(S_0!)$, hence $S_1 > \log S_0$ (we discard lower order terms), and therefore $L = S_0S_1$ should satisfy $L > b(\log b)^2$.

From this point of view the choice of parameters in Section 1.5 yields the smallest possible value for L in terms of b.

The main term which is responsible for $(\log b)^2$ is $S_0!$. If we could forget it, we would get only a constant times $b(\log b)(\log\log b)$ for L by taking

$$T_0 = N^2 b, \quad T_1 = N^2[\log b], \quad S_0 = N^3 b[\log b], \quad \text{and} \quad S_1 = N$$

with a suitable (sufficiently large) integer N. The quantity $b(\log b)(\log\log b)$ arises from the term $S_0 \log T_1$ which appears in the estimate for the length of f in Proposition 1.8.

We perform a very simple change of variables[12] which will allow us to take a smaller value for S_0. In place of evaluating the functions $z^\tau e^{tz}$ at the points $z = sb$, we consider rather $z^\tau e^{btz}$ at the points $z = s$. This means that we replace

$$\left(\frac{d}{dz}\right)^\sigma \big(z^\tau e^{tz}\big)(sb) = \sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!}(sb)^{\tau-\kappa} t^{\sigma-\kappa} e^{tsb}$$

by

$$\left(\frac{d}{dz}\right)^\sigma \big(z^\tau e^{btz}\big)(s) = \sum_{\kappa=0}^{\min\{\tau,\sigma\}} \frac{\sigma!\tau!}{\kappa!(\sigma-\kappa)!(\tau-\kappa)!} s^{\tau-\kappa} (tb)^{\sigma-\kappa} e^{tsb}.$$

While the former number is a positive real number bounded from above by

$$(bS_1)^\tau T_1^\sigma \min\left\{\left(1+\frac{\sigma}{bT_1S_1}\right)^\tau ; \left(1+\frac{\tau}{bT_1S_1}\right)^\sigma\right\} e^{bT_1S_1},$$

(see (1.3)), the latter is at most

$$S_1^\tau (bT_1)^\sigma \min\left\{\left(1+\frac{\sigma}{bT_1S_1}\right)^\tau ; \left(1+\frac{\tau}{bT_1S_1}\right)^\sigma\right\} e^{bT_1S_1}.$$

Thanks to this change of variables a variant of Proposition 1.8 can be deduced, where the upper bound for the length of f is

$$Le^W S_1^{T_0}(T_0+T_1)^{S_0},$$

the degree of f in Z_2 is $< S_0$ and (1.5) is replaced by

$$U \geq \log L + T_0 \log(ES_1) + |b|_+ ET_1S_1.$$

A suitable choice for the parameters is then

---

[12] Compare with the "duality" of [34], Section 13.7.

$$T_0 = N^2 b[\log b], \qquad T_1 = N^2[\log\log b],$$
$$S_0 = N^3 b[\log\log b], \qquad S_1 = N[\log b],$$

which may be used to prove (4.5) with

$$\exp\{-cb(\log b)\} \quad \text{replaced by} \quad \exp\{-cb(\log b)(\log\log b)\}$$

(for $b \geq 16$, say).

In order to remove the extra factor $\log\log b$ and to reach the desired estimate (4.5) involving only $\exp\{-cb(\log b)\}$, we need to get rid of the term $S_1^{T_0}$ in the upper bound for $\mathrm{L}(f)$ in Proposition 1.8, so that one could take

$$T_0 = N^2 b[\log b], \quad T_1 = N^2, \quad S_0 = N^3 b, \quad \text{and} \quad S_1 = N[\log b].$$

This is achieved by means of the so-called Fel'dman's polynomials. We use here a variant due to E.M. Matveev (see [34] Lemma 9.8): we replace z^τ by the polynomials[13] $\triangle(z;\tau,T_0^\sharp)$ which we now define.

4.2 Fel'dman's Polynomials

For r a nonnegative integer and $z \in \mathbb{C}$, consider the binomial (or Fel'dman's) polynomial

$$\triangle(z;r) = \frac{z(z+1)\cdots(z+r-1)}{r!}$$

if $r > 0$ and $\triangle(z;0) = 1$.

Let $\tau \geq 0$ and $T_0^\sharp > 0$ be two integers. Following Matveev [20], I, Section 7, we define a polynomial $\triangle(z;\tau,T_0^\sharp) \in \mathbb{Q}[z]$ of degree τ by

$$\triangle(z;\tau,T_0^\sharp) = \bigl(\triangle(z;T_0^\sharp)\bigr)^q \cdot \triangle(z;r),$$

where q and r are the quotient and remainder of the division of τ by $T_0^\sharp$:

$$\tau = T_0^\sharp q + r, \quad 0 \leq r < T_0^\sharp.$$

For $\sigma \geq 0$, define

$$\triangle(z;\tau,T_0^\sharp,\sigma) = \left(\frac{d}{dz}\right)^\sigma \triangle\,(z;\tau,T_0^\sharp).$$

For any positive integer n, denote by $\nu(n)$ the least common multiple of $1, 2, \ldots, n$. The estimate

$$\nu(n) \leq e^{107n/103}$$

[13] The polynomial denoted here by $\triangle(z;\tau,T_0^\sharp)$ is denoted $\delta_{T_0^\sharp}(z;\tau)$ in [34], Section 9.2.1, while our notation $\triangle(z;\tau,T_0^\sharp,\sigma)$ corresponds to $\delta_{T_0^\sharp}(z;\tau,\sigma)$ in [34], Section 9.2.1. Here we follow Matveev's notation in [20], I, Section 7.

can be deduced from the prime number Theorem (see for instance [34], Section 9.2.1).

The next result is due to Matveev [20], I, Lemma 7.1 (see also [34] Lemma 9.8).

Lemma 4.4. *Let $T_0^\sharp > 0$, $\tau > 0$ and $\sigma \geq 0$ be rational integers. For any integer κ in the interval $0 \leq \kappa \leq \sigma$ and any rational integer $m \in \mathbb{N}$, the number*

$$\nu(T_0^\sharp)^\sigma \cdot \frac{1}{\kappa!} \triangle(m; \tau, T_0^\sharp, \kappa)$$

is a nonnegative rational integer. Moreover, for any complex number z, we have

$$\sum_{\kappa=0}^{\sigma} \binom{\sigma}{\kappa} \left|\triangle(z; \tau, T_0^\sharp, \kappa)\right| \leq \sigma! e^{\tau + T_0^\sharp} \left(\frac{|z|}{T_0^\sharp} + 1\right)^\tau.$$

4.3 Output of the Transcendence Argument

Here is the basic estimate which follows from the transcendence proof (by means of the interpolation determinant method) and will enable us to deduce Theorem 4.2. It is worth to notice that Theorem 4.5 does not involve any arithmetic assumption (the proof does not use Liouville's inequality).

Theorem 4.5. *Let b be a nonzero complex number. Let T_0, $T_0^\sharp$, T_1, S_0, S_1, K be positive integers and E a real number with $E \geq e$. Set $L = T_0 T_1$ and assume $0 \leq K < L$ and*

$$S_0 S_1 > (T_0 + S_1 - 1) T_1.$$

Then there exists a set $\{f_1, \ldots, f_M\}$ of polynomials in the ring $\mathbb{Z}[Z_1, Z_2]$, each of degree $< LT_1S_1$ in Z_1 and degree $< LS_0$ in Z_2, of length at most

$$L! \nu(T_0^\sharp)^{LS_0} (S_0!)^L e^{(T_0^\sharp + T_0)L} \left(\frac{S_1}{T_0^\sharp} + 1\right)^{T_0 L} T_1^{S_0 L},$$

such that the polynomials $\{f_1(Z_1, b), \ldots, f_M(Z_1, b)\}$ in $\mathbb{C}[Z_1]$ have no common zero in $\mathbb{C}^\times$ and such that

$$\begin{aligned}\frac{1}{L} \max_{\substack{\underline{k} \in \mathbb{N}^2 \\ \|\underline{k}\| < K}} \log\left|\frac{1}{\underline{k}!} \mathcal{D}^{\underline{k}} f_i(e^b, b)\right| &\leq -\frac{1}{2L}(L-K)(L-K-1)\log E + \log(8L) \\ &+ S_0 \log\bigl(\nu(T_0^\sharp) S_0 E\bigr) + T_0^\sharp + T_0 + T_0 \log\left(\frac{S_1 E}{T_0^\sharp} + 1\right) \\ &+ S_0 \log(2T_1 |b|_+) + T_1 S_1 E |b|\end{aligned}$$

for $1 \leq i \leq M$.

In the conclusion $\mathcal{D}^{\underline{k}}$ stands for $(\partial/\partial Z_1)^{k_1}(\partial/\partial Z_2)^{k_2}$.

For the proof of Theorem 4.5 we shall need the following refinement of Lemma 1.5:

Lemma 4.6. *Let L and L' be positive integers with $L' \le L$ and $\varphi_1, \dots, \varphi_{L'}$ entire functions in $\mathbb{C}$. Let also $\zeta_1, \dots, \zeta_L$ be complex numbers and $\sigma_1, \dots, \sigma_L$ nonnegative integers. Let S_0 satisfy $S_0 \ge \max_{1\le\mu\le L} \sigma_\mu$. Furthermore, for $L'+1 \le \lambda \le L$ and $1 \le \mu \le L$ let $\delta_{\lambda\mu}$ be a complex number. For $1 \le \lambda \le L'$ and $1 \le \mu \le L$ we define*

$$\delta_{\lambda\mu} = \varphi_\lambda^{(\sigma_\mu)}(\zeta_\mu).$$

Finally, let $E > 1$ and $Q_1, \dots, Q_L$ be positive real numbers satisfying

$$Q_\lambda \ge \log \sup_{|z|=E} \max_{1\le\mu\le L} |\varphi_\lambda^{(\sigma_\mu)}(z\zeta_\mu)| \qquad (1 \le \lambda \le L'),$$

$$Q_\lambda \ge \log \max_{1\le\mu\le L} |\delta_{\lambda,\mu}| \qquad (L'+1 \le \lambda \le L).$$

We consider the determinant

$$\Delta = \det\Big(\delta_{\lambda\mu}\Big)_{1\le\lambda,\mu\le L}.$$

Then we have

$$\log|\Delta| \le -\frac{1}{2}L'(L'-1)\log E + L'S_0\log E + \log(L!) + Q_1 + \cdots + Q_L.$$

Proof. Lemma 4.6 is essentially the case $n = 1$ of Lemma 7.5 in [34] but it includes derivatives like in Proposition 9.13 of [34].

For $1 \le \mu \le L$, we define functions $d_{1\mu}(z), \dots, d_{L\mu}(z)$ by

$$d_{\lambda\mu}(z) = \begin{cases} \varphi_\lambda^{(\sigma_\mu)}(\zeta_\mu z) & \text{for } \ 1 \le \lambda \le L', \\ \\ \delta_{\lambda\mu} & \text{for } \ L' < \lambda \le L. \end{cases}$$

From Lemma 9.2 of [34] we deduce that the function

$$D(z) = \det\Big(d_{\lambda\mu}(z)\Big)_{1\le\lambda,\mu\le L}$$

has a zero at the origin of multiplicity $\ge (1/2)L'(L'-1) - L'S_0$. As $D(1) = \Delta$, we conclude the proof of Lemma 4.6 by using Schwarz' Lemma 1.3. $\square$

Proof of Theorem 4.5. For each $a \in \mathbb{C}^\times$ we shall construct a polynomial $f \in \mathbb{Z}[Z_1, Z_2]$ (depending on a) satisfying the required conditions and such that

$$f(a, b) \ne 0.$$

Since the degrees and lengths of these polynomials will be bounded (independently of a), this will mean that we end up with a finite set of polynomials

$\{f_1,\ldots,f_M\}$ in $\mathbb{Z}[Z_1,Z_2]$, such that the only possible common zero in $\mathbb{C}$ of the polynomials $f_i(Z_1,b)\ \ (i=1,\ldots,M)$ is $Z_1=0$.

Define, for $0\le\tau<T_0$ and $0\le t<T_1$,

$$\varphi_{\tau t}(z)=\triangle(z;\tau,T_0^\sharp)e^{tbz}.$$

For $\sigma=0,1,\ldots,S_0-1$, we have

$$\varphi_{\tau t}^{(\sigma)}(z)=\sum_{\kappa=0}^{\sigma}\binom{\sigma}{\kappa}\triangle\,(z;\tau,T_0^\sharp,\kappa)(tb)^{\sigma-\kappa}e^{tbz}.$$

Define $\widetilde{f}_{\tau t}^{(\sigma s)}\in\mathbb{Q}[Z_1,Z_2]$ by

$$\widetilde{f}_{\tau t}^{(\sigma s)}(Z_1,Z_2)=\sum_{\kappa=0}^{\sigma}\binom{\sigma}{\kappa}\triangle\,(s;\tau,T_0^\sharp,\kappa)(tZ_2)^{\sigma-\kappa}Z_1^{ts},$$

so that[14]

$$\widetilde{f}_{\tau t}^{(\sigma s)}(e^b,b)=\varphi_{\tau t}^{(\sigma)}(s)\tag{4.6}$$

and

$$\widetilde{f}_{\tau t}^{(\sigma s)}(a,b)=\mathrm{D}_b^\sigma\bigl(\triangle(X;\tau,T_0^\sharp)Y^t\bigr)(s,a^s).$$

This polynomial $\widetilde{f}_{\tau t}^{(\sigma s)}$ has rational coefficients. In order to get a polynomial with integer coefficients, we multiply it by a denominator: using Lemma 4.4, define $f_{\tau t}^{(\sigma s)}\in\mathbb{Z}[Z_1,Z_2]$ by

$$f_{\tau t}^{(\sigma s)}(Z_1,Z_2)=\nu(T_0^\sharp)^\sigma\widetilde{f}_{\tau t}^{(\sigma s)}(Z_1,Z_2).$$

In place of the exponent σ for $\nu(T_0^\sharp)$ we could as well put the exponent $\min\{\sigma,\tau\}$, because $\triangle(z;\tau,T_0^\sharp,\kappa)=0$ for $\kappa>\tau$. Notice also that we use only the fact that

$$\nu(T_0^\sharp)^\sigma\,\triangle\,(s;\tau,T_0^\sharp,\kappa)$$

is an integer; indeed, according to Lemma 4.4, it is a multiple of $\kappa!$, but we do not use this fact[15].

The polynomial $f_{\tau t}^{(\sigma s)}$ has degree $\le ts$ in Z_1, degree $\le\sigma$ in Z_2 and length at most

$$\nu(T_0^\sharp)^\sigma\sum_{\kappa=0}^{\sigma}\binom{\sigma}{\kappa}\triangle\,(s;\tau,T_0^\sharp,\kappa)t^{\sigma-\kappa}\le\nu(T_0^\sharp)^{S_0}S_0!e^{T_0^\sharp+T_0}\left(\frac{S_1}{T_0^\sharp}+1\right)^{T_0}T_1^{S_0}.$$

Consider the matrix

[14] The upper index (σs) in the notation $\widetilde{f}_{\tau t}^{(\sigma s)}$ is not a derivation, while $\varphi_{\tau t}^{(\sigma)}=(d/dz)^\sigma\varphi_{\tau t}$.

[15] Maybe one should, in order to remove the extra $\log\log A$ which arises in $\log B+\log\log A$...

$$\mathtt{M} = \Bigl(f_{\tau t}^{(\sigma s)}(a,b) \Bigr)_{\substack{(\tau,t)\\ (\sigma,s)}}$$

with L rows indexed by (τ, t), $(0 \le \tau < T_0,\ 0 \le t < T_1)$ and S_0S_1 columns indexed by (σ, s) with $0 \le \sigma < S_0$, $0 \le s < S_1$.

We claim that $\mathtt{M}$ has rank $L = T_0T_1$. Indeed consider a linear dependence relation between the rows:

$$\sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} f_{\tau t}^{(\sigma s)}(a,b) = 0 \quad \text{for} \quad 0 \le \sigma < S_0 \quad \text{and} \quad 0 \le s < S_1.$$

Define

$$P(X,Y) = \sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} \bigtriangleup (X; \tau, T_0^\sharp) Y^t.$$

Then we have

$$\mathrm{D}_b^\sigma P(s, a^s) = \sum_{\tau=0}^{T_0-1} \sum_{t=0}^{T_1-1} c_{\tau t} \widetilde{f}_{\tau t}^{(\sigma s)}(a,b),$$

hence

$$\mathrm{D}_b^\sigma P(s, a^s) = 0 \quad \text{for} \quad 0 \le \sigma < S_0 \quad \text{and} \quad 0 \le s < S_1.$$

According to Lemma 4.3 with

$$\xi_s = s, \quad \eta_s = a^s \quad (0 \le s < S_1)$$

and thanks to the assumption

$$S_0S_1 > (T_0 + S_1 - 1)T_1,$$

we deduce $c_{\tau t} = 0$ for $0 \le \tau < T_0$ and $0 \le t < T_1$.

Hence $\mathtt{M}$ has rank $L = T_0T_1$. Select L columns of $\mathtt{M}$ with indices (σ_μ, s_μ) $(1 \le \mu \le L)$, which are linearly independent. This means that the polynomial

$$f(Z_1, Z_2) = \det \Bigl(f_{\tau t}^{(\sigma_\mu s_\mu)}(Z_1, Z_2) \Bigr)_{\substack{(\tau,t)\\ 1\le\mu\le L}}$$

does not vanish at the point (a, b). The degree of f in Z_1 is $< LT_1S_1$ and in Z_2 is $< LS_0$. We bound from above the length of f as follows: write

$$f_{\tau t}^{(\sigma_\mu s_\mu)}(Z_1, Z_2) = Z_1^{ts_\mu} \sum_{j=0}^{\sigma_\mu} q_{j\tau t\mu} Z_2^j$$

with

$$q_{j\tau t\mu} = \nu(T_0^\sharp)^{\sigma_\mu} \binom{\sigma_\mu}{j} \bigtriangleup (s_\mu; \tau, T_0^\sharp, \sigma_\mu - j) t^j.$$

By Lemma 4.4 we have

$$
\begin{aligned}
\mathrm{L}(f_{\tau t}^{(\sigma_\mu s_\mu)}) &= \sum_{j=0}^{\sigma_\mu} |q_{j\tau t\mu}| \\
&\le \nu(T_0^\sharp)^{S_0} \sum_{j=0}^{\sigma_\mu} \binom{\sigma_\mu}{j} \triangle\,(s_\mu;\tau,T_0^\sharp,\sigma_\mu-j)t^j \\
&\le \nu(T_0^\sharp)^{S_0} S_0! e^{T_0^\sharp+T_0}\left(\frac{S_1}{T_0^\sharp}+1\right)^{T_0} T_1^{S_0}.
\end{aligned}
$$

Therefore

$$
\mathrm{L}(f) \le L!\nu(T_0^\sharp)^{LS_0} S_0!^L e^{L(T_0^\sharp+T_0)}\left(\frac{S_1}{T_0^\sharp}+1\right)^{LT_0} T_1^{LS_0}.
$$

These crude estimates could be slightly improved, but this is irrelevant for our purpose, since we do not pay too much attention to the absolute constants.

We want to estimate from above the number

$$
\max_{\substack{\underline{k}\in\mathbb{N}^2\\ \|\underline{k}\|<K}} \log\left|\frac{1}{\underline{k}!}\mathcal{D}^{\underline{k}} f(e^b,b)\right|.
$$

Fix $\underline{k}\in\mathbb{N}^2$ with $\|\underline{k}\|<K$. We have

$$
\frac{1}{\underline{k}!}\mathcal{D}^{\underline{k}} f(e^b,b) = \sum_{\underline{\underline{\kappa}}} \varDelta_{\underline{\underline{\kappa}}} \quad\text{with}\quad \varDelta_{\underline{\underline{\kappa}}} = \det\Big(\frac{1}{\underline{\kappa}_{\tau t}!}\mathcal{D}^{\underline{\kappa}_{\tau t}} f_{\tau t}^{(\sigma_\mu s_\mu)}(e^b,b)\Big)_{\substack{(\tau,t)\\ 1\le\mu\le L}},
$$

where $\underline{\underline{\kappa}}$ ranges over the set of elements

$$
(\underline{\kappa}_{\tau t})_{\substack{0\le\tau<T_0\\ 0\le t<T_1}} \in (\mathbb{N}^2)^L \quad\text{with}\quad \sum_{\tau=0}^{T_0-1}\sum_{t=0}^{T_1-1} \underline{\kappa}_{\tau t} = \underline{k}. \tag{4.7}
$$

For $\underline{k}=(k_1,k_2)$ the number of such $\underline{\underline{\kappa}}$ is

$$
\binom{k_1+L-1}{L-1}\binom{k_2+L-1}{L-1} < 2^{K+2L} < 2^{3L}.
$$

Fix $\underline{\underline{\kappa}}$ satisfying (4.7) and define

$$
\widetilde{\varDelta}_{\underline{\underline{\kappa}}} = \det\Big(\frac{1}{\underline{\kappa}_{\tau t}!}\mathcal{D}^{\underline{\kappa}_{\tau t}} \widetilde{f}_{\tau t}^{(\sigma_\mu s_\mu)}(e^b,b)\Big)_{\substack{(\tau,t)\\ 1\le\mu\le L}},
$$

so that

$$
\varDelta_{\underline{\underline{\kappa}}} = \widetilde{\varDelta}_{\underline{\underline{\kappa}}}\,\nu(T_0^\sharp)^{\sigma_1+\cdots+\sigma_L}.
$$

It remains to check the hypotheses of Lemma 4.6 with $\zeta_\mu = s_\mu$, $L' = L - K$ and

$$Q_{\tau t} = \log(S_0!) + T_0 + T_0^\sharp + T_0 \log\left(\frac{S_1 E}{T_0^\sharp} + 1\right) + S_0 \log(2T_1|b|_+) + T_1 S_1 E|b|$$

for $0 \le \tau < T_0$ and $0 \le t < T_1$.

From (4.7) it follows that there are at least $L - K$ indices (τ, t) with $0 \le \tau < T_0$ and $0 \le t < T_1$ for which $\underline{\kappa}_{\tau t} = (0,0)$; for these indices we use (4.6). We conclude by means of the estimates

$$\max_{1\le\mu\le L} \sup_{|z|=E} \left|\varphi_{\tau t}^{(\sigma_\mu)}(s_\mu z)\right| \le e^{Q_{\tau t}}$$

and, for $\underline{\kappa}_{\tau t} = (\kappa_1, \kappa_2)$,

$$\begin{aligned}
&\left|\frac{1}{\underline{\kappa}_{\tau t}!}\mathcal{D}^{\underline{\kappa}_{\tau t}} \widetilde{f}_{\tau t}^{(\sigma_\mu s_\mu)}(e^b, b)\right| \\
&\le \sum_{\kappa=0}^{\sigma_\mu} \binom{\sigma_\mu}{\kappa} \triangle\,(s_\mu; \tau, T_0^\sharp, \kappa) t^{\sigma_\mu - \kappa} \binom{\sigma_\mu - \kappa}{\kappa_2} |b|_+^{\sigma_\mu - \kappa - \kappa_2} \binom{ts_\mu}{\kappa_1} e^{|b|(ts_\mu - \kappa_1)},
\end{aligned}$$

which implies

$$\begin{aligned}
&\max_{1\le\mu\le L} \left|\frac{1}{\underline{\kappa}_{\tau t}!}\mathcal{D}^{\underline{\kappa}_{\tau t}} \widetilde{f}_{\tau t}^{(\sigma_\mu s_\mu)}(e^b, b)\right| \\
&\qquad\qquad \le S_0! e^{T_0 + T_0^\sharp} \left(\frac{S_1}{T_0^\sharp} + 1\right)^{T_0} (2|b|_+ T_1)^{S_0} (2e^{|b|})^{T_1 S_1} \le e^{Q_{\tau t}}.
\end{aligned}$$

□

4.4 From Polynomial Approximation to Algebraic Approximation

Let $\underline{\theta} \in \mathbb{C}^n$ and $\underline{\gamma} \in \overline{\mathbb{Q}}^n$. We want to estimate from below $|\underline{\theta} - \underline{\gamma}|$. Our strategy is as follows (see [34] Proposition 15.3 and Exercise 15.3). Assume there is a polynomial $f \in \mathbb{Z}[\underline{X}]$ such that $f(\underline{\gamma}) \neq 0$ while $|f(\underline{\theta})|$ is sufficiently small. Liouville's inequality (Proposition 1.13) yields a lower bound for $|f(\underline{\gamma})|$ (depending on the degrees and length of f, as well as the degrees and heights of the components of γ). Since $|f(\underline{\theta}) - f(\underline{\gamma})|$ can be estimated from above in terms of $|\underline{\theta} - \underline{\gamma}|$, we deduce the desired lower bound for $|\underline{\theta} - \underline{\gamma}|$.

A refined estimate can be obtained when not only $f(\underline{\theta})$, but also the values of several derivatives of f at $\underline{\theta}$, have a small absolute value. The required upper bound for $|f(\underline{\theta}) - f(\underline{\gamma})|$ in terms of $|\underline{\theta} - \underline{\gamma}|$ will be a consequence of the interpolation formula (Lemma 1.11).

Lemma 4.7. *Let ℓ, K, $\nu_1, \dots, \nu_\ell$, $N_1, \dots, N_\ell$ be positive integers. Define*

$$m = \nu_1 + \cdots + \nu_\ell.$$

Let $\mathbb{K}$ be a number field of degree $D = [\mathbb{K} \colon \mathbb{Q}]$. For $1 \le i \le \ell$, write

$$\underline{\theta}_i = (\theta_{i1}, \dots, \theta_{i\nu_i}) \in \mathbb{C}^{\nu_i} \quad \text{and} \quad \underline{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{i\nu_i}) \in \mathbb{K}^{\nu_i}.$$

Set

$$\underline{\theta} = (\underline{\theta}_1, \dots, \underline{\theta}_\ell) \in \mathbb{C}^m \quad \text{and} \quad \underline{\gamma} = (\underline{\gamma}_1, \dots, \underline{\gamma}_\ell) \in \mathbb{K}^m$$

and assume

$$|\underline{\theta} - \underline{\gamma}| \le \frac{1}{2m}.$$

Let f be a polynomial with integer coefficients in m variables X_{ij} $(1 \le j \le \nu_i$, $1 \le i \le \ell)$, of total degree $\le N_i$ with respect to the variables $X_{i1}, \dots, X_{i\nu_i}$, $(1 \le i \le \ell)$, such that

$$f(\underline{\gamma}) \neq 0.$$

Define

$$\epsilon = \frac{1}{2} \mathrm{L}(f)^{-D} \exp\left\{ -D \sum_{i=1}^{\ell} N_i \mathrm{h}\bigl(1 \colon \gamma_{i1} \colon \cdots \colon \gamma_{i\nu_i}\bigr) \right\}$$

and assume

$$\max_{\substack{\underline{k} \in \mathbb{N}^m \\ \|\underline{k}\| < K}} \frac{1}{\underline{k}!} |\mathcal{D}^{\underline{k}} f(\underline{\theta})| \le \frac{1}{2} \epsilon \mathrm{L}(f). \tag{4.8}$$

Then

$$|\underline{\theta} - \underline{\gamma}|^K \ge \epsilon (1 + \sqrt{K})^{-1} \prod_{i=1}^{\ell} (1 + |\underline{\theta}_i|)^{-N_i}.$$

Proof. We use the interpolation formula of Lemma 1.11 with $F(\underline{z}) = f(\underline{\theta} + \underline{z})$, $r = |\underline{\theta} - \underline{\gamma}|$, $R = 1$. We estimate $|F|_R$ from above as follows:

$$|F|_R \le \mathrm{L}(f) \prod_{i=1}^{\ell} (1 + |\underline{\theta}_i|)^{N_i}.$$

Since

$$\sum_{\substack{\underline{k} \in \mathbb{N}^m \\ \|\underline{k}\| < K}} r^{\|\underline{k}\|} \le \sum_{\underline{k} \in \mathbb{N}^m} r^{\|\underline{k}\|} = (1 - r)^{-m}$$

and since

$$\left(1 - \frac{1}{2m}\right)^{-m} \le 2,$$

by the assumption $r \le 1/(2m)$ we have:

$$\sum_{\substack{\underline{k}\in\mathbb{N}^m \\ \|\underline{k}\|<K}} r^{\|\underline{k}\|} \leq 2.$$

Therefore

$$|f(\underline{\gamma})| \leq |\underline{\theta}-\underline{\gamma}|^K(1+\sqrt{K})\mathrm{L}(f)\prod_{i=1}^{\ell}(1+|\underline{\theta}_i|)^{N_i} + 2\max_{\|\underline{k}\|<K}\frac{1}{\underline{k}!}|\mathcal{D}^{\underline{k}}f(\underline{\theta})|.$$

By Liouville's inequality (Proposition 1.13), we deduce from the assumption $f(\underline{\gamma}) \neq 0$:

$$|f(\underline{\gamma})| \geq \mathrm{L}(f)^{1-D}\prod_{i=1}^{\ell} e^{-DN_i\mathrm{h}(1:\,\gamma_{i1}:\,\cdots:\,\gamma_{i\nu_i})}.$$

Hence

$$\begin{aligned}\mathrm{L}(f)^{1-D}&\prod_{i=1}^{\ell} e^{-DN_i\mathrm{h}(1:\,\gamma_{i1}:\,\cdots:\,\gamma_{i\nu_i})}\\ &\leq |\underline{\theta}-\underline{\gamma}|^K(1+\sqrt{K})\mathrm{L}(f)\prod_{i=1}^{\ell}(1+|\underline{\theta}_i|)^{N_i} + 2\max_{\|\underline{k}\|<K}\frac{1}{\underline{k}!}|\mathcal{D}^{\underline{k}}f(\underline{\theta})|. \qquad (4.9)\end{aligned}$$

We now use the definition of ϵ and the assumption (4.8): the left hand side of (4.9) is $\geq 2\epsilon\mathrm{L}(f)$ and the second term on the right hand side is at most $\epsilon\mathrm{L}(f)$, hence

$$|\underline{\theta}-\underline{\gamma}|^K(1+\sqrt{K})\prod_{i=1}^{\ell}(1+|\underline{\theta}_i|)^{N_i} \geq \epsilon.$$

□

Remark. In the special case $m = 1$ (hence $\ell = \nu_1 = 1$, $\theta \in \mathbb{C}$, and, say, $N = N_1$), here is a slightly different estimate, suggested by Exercise 15.3.a of [34]:

$$|\underline{\theta}-\underline{\gamma}|^K \geq \frac{\epsilon}{K\binom{N}{K}(1+|\theta|)^{N-K}}\cdot$$

The proof is quite similar, but one replaces Lemma 1.11 by Taylor's expansion

$$f(\gamma) = \sum_{k=0}^{K-1}\frac{1}{k!}f^{(k)}(\theta)(\gamma-\theta)^k + \frac{1}{(K-1)!}\int_{\theta}^{\gamma} f^{(K)}(t)(\gamma-t)^{K-1}dt.$$

The conclusion follows from the estimates

$$\sum_{k=0}^{K-1}\frac{1}{k!}\left|f^{(k)}(\theta)\right||\gamma-\theta|^k \leq \epsilon\mathrm{L}(f), \qquad |f(\gamma)| \geq 2\epsilon\mathrm{L}(f)$$

and

$$\frac{1}{(K-1)!}\sup\left\{\left|f^{(K)}(t)\right| \; ; \; t=\theta(1-u)+\gamma u,\ 0\le u\le 1\right\}$$
$$\le K\binom{N}{K}\mathrm{L}(f)(1+|\theta|)^{N-K}.$$

4.5 Proof of Theorem 4.2

We start by selecting a sufficiently large absolute constant N. Next we assume the hypotheses of Theorem 4.2 are satisfied. From the assumptions we deduce $\log B \ge (1/D)+\log D \ge 1$, hence

$$\log\max\{B,N\}\le(\log N)(\log B).$$

Replacing if necessary c by $c\log N$, we may assume $B\ge N$.

Next from Liouville's inequality (see [34] Exercise 3.7) we deduce

$$|\beta|\ge e^{-D\mathrm{h}(\beta)},$$

hence there is no loss of generality to assume $|\beta|\le 2|\lambda|$ (and consequently $\lambda\neq 0$).

Define first T_1, S_1 and $T_0^\sharp$ by

$$T_1=\left[N^2\frac{D\log D+\log E}{\log E}\right],\qquad S_1=\left[N\frac{D\log B}{\log E}\right],\qquad T_0^\sharp=[\log B],$$

and then T_0 and S_0 by

$$T_0=\left[N\frac{D\log A}{\log E}\right]S_1,\qquad S_0=2\left[N\frac{D\log A}{\log E}\right]T_1.$$

Plainly the number $L=T_0T_1$ is also equal to $(1/2)S_0S_1$. Up to terms of lower order[16], $L\log E$ is

$$N^4D^2(\log A)(\log B)\bigl(D\log D+\log E\bigr)(\log E)^{-2},$$

which is the main term in the final estimate of the conclusion of Theorem 4.2.

Denoting by $c_1,\ldots,c_{12}$ positive absolute constants, the following estimates are plain:

$$T_1S_1\le\frac{c_1L\log E}{ND\log A},\qquad S_0\le\frac{c_2L\log E}{ND\log B},$$
$$T_0^\sharp S_0\le\frac{c_3}{ND}L\log E,\qquad \log S_0\le\frac{c_4}{ND}S_1\log E,$$
$$\log T_1\le\frac{c_5}{ND}S_1\log E\quad\text{and}\quad T_1S_1E|\beta|\le\frac{c_6}{N}L\log E.$$

[16] Beware of integral parts!

We apply Theorem 4.5 with $b = \beta$ and $K = [L/2]$. Since the polynomials $f_i(Z_1, \beta)$ $(1 \le i \le n)$ have no common zero in $\mathbb{C}^\times$, one of $f_1, \dots, f_M$, say f, has $f(\alpha, \beta) \neq 0$. This polynomial has degree at most

$$N_1 = \frac{c_7 L^2 \log E}{ND \log A} \quad \text{and} \quad N_2 = \frac{c_8 L^2 \log E}{ND \log B}$$

in Z_1 and Z_2 respectively, and length at most

$$\exp\left(\frac{c_9 L^2 \log E}{ND}\right).$$

Moreover, if we set $\ell = m = 2$, $\nu_1 = \nu_2 = 1$, $(\theta_1, \theta_2) = (e^\beta, \beta)$ and $(\gamma_1, \gamma_2) = (\alpha, \beta)$, then the left hand side of (4.8) is at most

$$E^{-c_{10}L^2}.$$

Since

$$\epsilon \ge E^{-c_{11}L^2/N}$$

we deduce from Lemma 4.7

$$|e^\beta - \alpha|^{L/2} \ge E^{-c_{12}L^2/N}.$$

Therefore

$$|\beta - \log \alpha| \ge E^{-L} \tag{4.10}$$

and the conclusion of Theorem 4.2 follows. □

5 Fifth Lecture. Measures of Linear Independence of Logarithms of Algebraic Numbers

5.1 Introduction

The last two lectures are devoted to the question of measures of linear independence for logarithms of algebraic numbers. Here are a few references on this topic.

A simple proof for a homogeneous measure of linear independence of an arbitrary number of logarithms is given in Chap. 7 of [32], using an extension of Schneider's method. A refined estimate (relying on the same ideas) is given in Chap. 7 of [34]. The sharpest known estimate (in the general case, homogeneous or not) arising from this method is established in Chap. 9 of [34].

Baker's method is explained in [2] Chap. 3, [11] Chap. 8, 10 and 11, [6] Chap. 10, [28] Chap. 3, [7] Chap. 4 and [34], Section 10.1 (see also the introduction of [20], I, for a historical survey). While these proofs involve an auxiliary function, a measure of linear independence for logarithms of algebraic numbers is obtained in Section 10.2 of [34] by means of an interpolation determinant (without extrapolation).

A comparative discussion of these methods can be found in [34] (see Section 14.4), where a more general estimate is established (the so-called *quantitative version of the linear subgroup Theorem*).

The state of the art including references to the sharpest known measures of $\mathbb{Q}$-linear independence for logarithms of algebraic numbers is given in [34], Section 10.4.6. Now one should add to this picture Matveev's recent result in [20], II, (see Theorem 6.1; we refer also to Nesterenko's lectures).

Here is the main result for the rational case.

Theorem 5.1. *For each positive integer* n *there exists a positive constant* $C(n)$ *with the following property. Let* $\alpha_1, \ldots, \alpha_n$ *be nonzero algebraic numbers and* $\log\alpha_1, \ldots, \log\alpha_n$ *logarithms of* $\alpha_1, \ldots, \alpha_n$ *respectively. Assume that the numbers* $\log\alpha_1, \ldots, \log\alpha_n$ *are* $\mathbb{Q}$*-linearly independent. Let* $b_1, \ldots, b_n$ *be rational integers. Denote by* D *the degree of the number field* $\mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ *over* $\mathbb{Q}$*. Further, let* W*,* E*,* E^* *be positive real numbers, each* $\geq e$ *and let* $V_1, \ldots, V_n$ *be positive real numbers. Assume*

$$V_j \geq \max\left\{\mathrm{h}(\alpha_j),\ \frac{E|\log\alpha_j|}{D},\ \frac{\log E}{D}\right\} \quad (1 \leq j \leq n),$$

$$\log E^* \geq \max\left\{\frac{1}{D}\log E,\ \log\left(\frac{D}{\log E}\right)\right\}$$

and $W \geq \log E^*$*. Further, assume* $b_n \neq 0$ *and*

$$e^W \geq \max_{1\leq j\leq n-1}\left\{\frac{|b_n|}{V_j} + \frac{|b_j|}{V_n}\right\} \cdot \frac{\log E}{D}.$$

Then the number

$$\Lambda = b_1\log\alpha_1 + \cdots + b_n\log\alpha_n$$

has absolute value bounded from below by

$$|\Lambda| > \exp\{-C(n)D^{n+2}WV_1\cdots V_n(\log E^*)(\log E)^{-n-1}\}.$$

For several applications (especially for solving explicitly Diophantine equations) it is quite important to produce an estimate with a small numerical constant $C(n)$, even to the cost of relaxing the dependence in some parameters. A striking example arises with the case $n = 2$ where a very small value of $C(2)$ can be reached, provided that W is replaced by W^2 (see Theorem 5.10 below). We shall explain in an appendix to this Section 5 how such an estimate can be used also for the case $n \geq 2$ in order to get a lower bound for $|\Lambda|$ which involves a small constant – we need to replace W by $e^{\kappa W}$ for some $\kappa > 0$, but the point is that an admissible value for κ is < 1, hence the result is not trivial.

Apart from the explicit value for C, Theorem 5.1 includes essentially all known estimates on this number $|\Lambda|$.

Admissible numerical values for C are given in [34], Section 10.4.6. Assuming $E = e$, Matveev proved recently in [20], II, that $C(n)$ can be replaced by C^n where now C is a positive absolute constant (see Section 6).

Our goal is to explain some of the main ideas of the proof of Theorem 5.1 by means of Baker's method. In this lecture we discuss the classical approach following Baker's method with an auxiliary function and in the last one we shall show how to replace the auxiliary function by an interpolation determinant.

5.2 Baker's Method with an Auxiliary Function

Our main goal is to introduce the strategy of the proof, with an emphasis on the ideas and tools. We do not produce exact estimates, but we consider only the dominating terms which will explain the choice of parameters and at the same time provide some explanation for the limitation of the present method.

a) Main Conditions on the Parameters

Assume the hypotheses of Theorem 5.1 are satisfied: $\alpha_1, \dots, \alpha_n$ are nonzero algebraic numbers, $\log \alpha_1, \dots, \log \alpha_n$ are $\mathbb{Q}$-linearly independent, $b_1, \dots, b_n$ are rational integers, $b_n \neq 0$ and

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n.$$

As usual we need to introduce parameters: let $T_0, T_1, \dots, T_n$, $T_0^\sharp$, S_0, S_1 and L be positive integers with $L = T_0 \cdots T_n$. We also introduce another important parameter, U, which is a positive real number and will play the main role: assuming

$$0 < |\Lambda| \leq e^{-U},$$

we plan to derive a contradiction as soon as U is sufficiently large. Of course all the point is to be explicit on this condition that U is large enough.

The proof will have a lot in common with the transcendence proof of Theorem 3.1 in Section 3 (where the assumption was $\Lambda = 0$). Instead of setting $\beta_k = -b_k/b_n$ and $\beta_0 = 0$, it is slightly more convenient to change the definitions of $\partial_0, \dots, \partial_{n-1}$ and to set now

$$\partial_0 = \frac{\partial}{\partial Y_0}, \qquad \partial_k = b_n Y_k \frac{\partial}{\partial Y_k} - b_k Y_n \frac{\partial}{\partial Y_n} \qquad (1 \leq k \leq n-1).$$

The auxiliary polynomial will have the shape

$$P(Y_0, \dots, Y_n) = \sum_{(\tau, \underline{t})} c_{\tau, \underline{t}} \bigtriangleup (Y_0; \tau, T_0^\sharp) Y_1^{t_1} \cdots Y_n^{t_n}.$$

In the sum, $(\tau, \underline{t})$ ranges over the set of all elements in $\mathbb{N} \times \mathbb{N}^n$ with $0 \le \tau < T_0$ and $0 \le t_j < T_j \quad (1 \le j \le n)$.

We are interested in the algebraic numbers

$$\partial^{\underline{\sigma}} P(s, \alpha_1^s, \dots, \alpha_n^s)$$

for $(\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N}$ satisfying $\|\underline{\sigma}\| < S_0$ and $0 \le s < S_1$. They can be explicitly written down as

$$\sum_{(\tau, \underline{t})} c_{\tau, \underline{t}} \gamma_{\tau \underline{t}}^{(\underline{\sigma} s)},$$

where

$$\gamma_{\tau \underline{t}}^{(\underline{\sigma} s)} = \partial^{\underline{\sigma}} \big(\triangle(Y_0; \tau, T_0^\sharp) Y_1^{t_1} \cdots Y_n^{t_n} \big)(s, \alpha_1^s, \dots, \alpha_n^s)$$

and

$$\partial^{\underline{\sigma}} \big(\triangle(Y_0; \tau, T_0^\sharp) Y_1^{t_1} \cdots Y_n^{t_n} \big) = \triangle(Y_0; \tau, T_0^\sharp, \sigma_0) \prod_{k=1}^{n-1} (b_n t_k - b_k t_n)^{\sigma_k} \cdot Y_1^{t_1} \cdots Y_n^{t_n}.$$

The auxiliary polynomial P is selected so that equations (3.3) hold.

Our first condition on the parameters will be written

$$\boxed{\frac{1}{n!} S_0^n S_1 < L.} \tag{5.1}$$

This is a lousy way of requiring that the number of equations is less than the number of unknowns (namely the coefficients $c_{\tau \underline{t}}$). Here we do not pay attention to the exact value of the absolute constants which come into the picture, but we are only interested with the main constraints. For the same reason we do not look at the estimate for the coefficients $c_{\tau \underline{t}}$.

For $\underline{\sigma} \in \mathbb{N}^n$ define

$$G_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}} P(z, \alpha_1^z, \dots, \alpha_n^z).$$

According to their definitions, the operators $\partial^{\underline{\sigma}}$ are related to derivatives along the hyperplane W of equation $b_1 z_1 + \dots + b_n z_n = 0$ in $\mathbb{C}^{n+1}$. The point $\underline{v} = (1, \log \alpha_1, \dots, \log \alpha_n)$ does not lie in this hyperplane (because $\Lambda \neq 0$), but it is not far from it: to a certain extent $|\Lambda|$ measures the "distance" between $\underline{v}$ and W. If the point $\underline{v}$ were on the hyperplane W, the conditions (3.3) would imply that the one variable entire function $G_{\underline{\sigma}}$ has a zero of multiplicity $\ge S_0 - \|\underline{\sigma}\|$ as each integer s with $0 \le s < S_1$. When $|\Lambda|$ is small, the first $S_0 - \|\underline{\sigma}\|$ derivatives of $G_{\underline{\sigma}}$ at these points have a small absolute value. More precisely define

$$g_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}} P\big(z, \alpha_1^z, \dots, \alpha_{n-1}^z, (\alpha_1^{b_1} \cdots \alpha_{n-1}^{b_{n-1}})^{-z/b_n}\big).$$

Then

$$G_{\underline{\sigma}}(z) - g_{\underline{\sigma}}(z)$$
$$= \sum_{(\tau,\underline{t})} c_{\tau,\underline{t}} \triangle (z;\tau,T_0^\sharp,\sigma_0) \prod_{k=1}^{n-1} (b_n t_k - b_k t_n)^{\sigma_k} \cdot \prod_{i=1}^{n} \alpha_i^{t_i z} \cdot \bigl(1 - e^{-t_n z \Lambda/b_n}\bigr).$$

From the inequality $|e^w - 1| \leq |w| e^{|w|}$ which is valid for any $w \in \mathbb{C}$ one deduces

$$\left|1 - e^{-t_n z \Lambda/b_n}\right| \leq |t_n z \Lambda/b_n| e^{|t_n z \Lambda/b_n|},$$

hence $|G_{\underline{\sigma}}(z) - g_{\underline{\sigma}}(z)|$ is quite small for $|z|$ not too large:

$$|G_{\underline{\sigma}}(z) - g_{\underline{\sigma}}(z)| \leq |\Lambda|^{1/2} = e^{-U/2}$$

for all relevant $\underline{\sigma}$ and z. In particular from the relations $G_{\underline{\sigma}}(s) = 0$ for $\|\underline{\sigma}\| < S_0$ and $0 \leq s < S_1$ one deduces

$$\left| g_{\underline{\sigma}}^{(\ell)}(s) \right| \leq e^{-U/3}$$

for $\|\underline{\sigma}\| < S_0 - S_0'$ and $0 \leq s < S_1$. Using an approximate Schwarz' Lemma like Lemma 1.3 one deduces

$$|g_{\underline{\sigma}}(s)| \leq e^{-2U'}$$

for the same $\underline{\sigma}$, s, with

$$U' = \min\{U, (S_0 - S_0')S_1 \log E\}.$$

Hence

$$|G_{\underline{\sigma}}(s)| \leq e^{-U'}$$

for the same $\underline{\sigma}$ and s. Combining this estimate with Liouville's inequality (Proposition 1.13) we deduce that $G_{\underline{\sigma}}$ satisfies (3.6). This conclusion can be reached only if the parameters satisfy certain conditions. In the estimate arising from Liouville's inequality (Proposition 1.13), the dominating terms are the following ones:

$$e^{-DWS_0}(E^*)^{-DT_0} \quad \text{arising from} \quad \triangle (s; T_0^\sharp, \tau, \sigma_0)$$

and

$$\prod_{j=1}^{n} e^{-DT_j V_j S_1'} \quad \text{arising from} \quad \prod_{j=1}^{n} \alpha_j^{-t_j s}.$$

The exact conditions which motivate the definitions of $V_1, \ldots, V_n$, W and E^* are explained in [34] Chap. 9 and 10. In particular the difference between the "homogeneous rational case" considered here and the "general case" (where the rational coefficients b_i's are replaced by algebraic numbers β_i, and there may be also an extra β_0) occurs here: the factor

$$(b_n t_k - b_k t_n)^{\sigma_k}$$

like it stands would require stronger assumptions on W, namely

$$W \geq \max_{1\leq k\leq n} \log V_k \quad\text{and}\quad W \geq \max_{1\leq k\leq n} \log |b_k|,$$

so that $DS_0 \log T_i$ is bounded by U' for $1 \leq i \leq n$. In order to deal with our weaker condition on W, the idea is to replace this factor by

$$\triangle(b_n t_k - b_k t_n, \sigma_k).$$

See [34] Lemma 9.11.

Therefore it is reasonable to require

$$T_0 \leq \frac{U'}{D\log E^*}, \quad S_0 \leq \frac{U'}{DW} \quad\text{and}\quad T_j \leq \frac{U'}{nDS_1'V_j} \quad (1\leq j\leq n).$$

Let us take

$$T_0 = \left[c_1\frac{U}{D\log E^*}\right], \quad S_0 = \left[c_2\frac{U}{DW}\right], \quad T_j = \left[c_3\frac{U}{nDS_1'V_j}\right] \quad (1\leq j\leq n), \tag{5.2}$$

where c_1, c_2, c_3 (as well as $c_4,\ldots,c_{11}$ below) denote absolute positive constants, which we are not interested in.

From condition (5.1) we deduce

$$U > c_4 S_1 (S_1'/W)^n D V_1 \cdots V_n (\log E^*). \tag{5.3}$$

At the end of this first extrapolation step we derive (3.6), provided that

$$(S_0 - S_0')S_1 \log E \geq U.$$

As a first try, let us take $S_0' = [S_0/2]$ and then

$$S_1 = \left[c_5\frac{DW}{\log E}\right]. \tag{5.4}$$

We started with $\binom{S_0+n}{n}S_1$ equations and we end up with $\binom{S_0'+n}{n}S_1'$ new equations. It is conceivable that no real progress has been achieved unless the number of new equations exceeds the number of old equations, which means essentially that ${S_0'}^n S_1'$ should not be smaller than $S_0^n S_1$. Hence we need at least $S_1' \geq 2^n S_1$.

There are several possibilities now. The easiest one is to apply immediately a zero estimate. This is possible only if $({S_0'}^n/n!)S_1'$ is somewhat larger than $L = T_0 T_1 \cdots T_n$ (which is the opposite of (5.1) when S_0 and S_1 are replaced by S_0' and S_1'). So if we wish to conclude immediately by means of the zero estimate we need to require

$$\binom{S_0' + n - 1}{n} S_1' > c_{ZE} L, \tag{5.5}$$

where $c_{ZE} = (n+1)!$ is the loss arising from Proposition 3.2 (this c_{ZE} is a notation of [34], Section 13.5). This gives rise to the condition

$$\boxed{{S_0'}^n S_1' > n!^2 L.}$$

Comparing with (5.1) gives $S_1' > n! S_1$, hence (5.3) yields the condition

$$U > c_6^n (n!)^n D^{n+2} W V_1 \cdots V_n (\log E^*)(\log E)^{-n-1}. \tag{5.6}$$

The estimate (5.6) corresponds to the sharpest available result with respect to $V_1, \ldots, V_n$, W, E and E^* (but not with respect to n), and the choice of parameters is given by (5.2) and (5.4).

b) End of the Proof

What we have shown so far is only that the above scheme of proof cannot reach a better estimate than $|\Lambda| \geq e^{-U}$ with U satisfying (5.6). It is a different issue to prove that one can indeed reach such an estimate by means of these arguments. However let us now say that there is mainly a single serious difficulty in doing so: (5.5) is only a necessary condition for applying the zero estimate (see (3.7)). More precisely the zero estimate enables one to conclude the proof, unless there is a linear dependence relation between the rational integers $b_1, \ldots, b_n$ with "small" coefficients. Indeed, starting from $0 < |\Lambda| \leq e^{-U}$, the transcendence machinery produces a tuple $(t_1, \ldots, t_n) \in \mathbb{Z}^n \setminus \{0\}$ such that $t_1 b_1 + \cdots + t_n b_n = 0$ together with a sharp upper bound for $\max_{1 \leq i \leq n} |t_i|$. This small linear dependence relation is the explanation for the fact that the system of equations (3.6) is somehow degenerate.

At this stage, the idea is to eliminate one b_i thanks to this relation and to work with a linear combination of $n-1$ logarithms instead of n. This is not the most efficient way; one may proceed by induction indeed, but it is better to repeat the transcendence argument and to take these relations into account. In [24] (and also [25] for a more general situation dealing with commutative algebraic groups), the induction is done as follows. The zero estimate produces an algebraic subgroup of $\mathbb{G}_{\mathrm{a}} \times \mathbb{G}_{\mathrm{m}}^n$ and the strategy is to start the proof from scratch (construction of the auxiliary polynomial) with such an obstructing subgroup. Thanks to this obstructing subgroup one has a better control on the rank of the system of linear equations to which we apply Thue-Siegel's Lemma. After the extrapolation, one produces a new set of relations (3.6) to which one applies again the zero estimate; the extremality of the initial obstructing subgroup enables one to conclude (by contradiction).

A different way of performing this induction is used by [5], but the underlying ideas are basically the same. This induction is somewhat technical but it is well under control now, and we shall not tell more about it. We assume implicitly that there is no "small" linear dependence relation between the b_i's, in which case the zero estimate shows that there is no nonzero polynomial $P \in \mathbb{C}[Y_0, \ldots, Y_n]$ for which a set of equalities (3.6) holds for $(\underline{\sigma}, s)$ ranging over a set with slightly more than $T_0 T_1 \cdots T_n$ elements (compare with condition (5.1)).

So our goal is to get more than $T_0 T_1 \cdots T_n$ equations (3.6) for P.

c) Baker's Method with an Interpolation Determinant and without Extrapolation

This Section 5.2 is devoted to Baker's method with an auxiliary function, but we make a small digression to point out that the arguments described in Section 5.2, a) and Section 5.2, b) involving an auxiliary function work out perfectly well for the interpolation determinant method. This is the topic of Chap. 10 of [34].

The basic scheme of proof is the following: Philippon's zero estimate enables us to produce a nonzero determinant ([34], Proposition 10.9) which is an algebraic number. Liouville's estimate provides a lower bound for its absolute value. The upper bound is obtained by analytic means ([34], Proposition 10.5). In this analytic argument, one takes into account the order of vanishing of an interpolation determinant at the origin only: extrapolation like in Section 3.5 is not necessary.

The role of the obstructing subgroup in connection with interpolation determinants is explained in [34], Section 10.2.3.

d) Dependence on n

We consider the dependence on n now.

The factor $(n!)^n$ in (5.6) is very large (it is comparable with the estimate which occurs in the interpolation determinant method of [34] Chap. 10 — see Section 14.4.3).

Even if one were to replace the condition $S_1' > n!S_1$ by the weaker $S_1' > 2^n S_1$, one would get a large "constant" in terms of n, involving 2^{n^2}.

In order to refine condition (5.6), an obvious solution would be to improve the zero estimate. However if one could replace $(n+1)!$ in (3.7) by, say, n^{c_7}, then one would still end up with $n^{c_8 n}$ in place of $c_6^n (n!)^n$ in the right hand side of (5.6)[17].

As soon as we wish to increase S_1'/S_1 by a factor $\lambda > 1$, we get a factor λ^n in U. Hence we should not take S_1'/S_1 larger than an absolute constant if we do not wish to introduce some n^n. We shall take $S_1' = 2S_1$. Now, since

[17] This is a significative difference with the "dual" method in Chap. 9 of [34], dealing with the algebraic group $\mathbb{G}_{\mathrm{a}}^n \times \mathbb{G}_{\mathrm{m}}$, where any improvement of the constant in the zero estimate immediately applies to the final estimate of Theorem 5.1.

we want to end up with at least as many equations as we started with, we need that ${S'_0}^n S'_1$ is not less than $S_0^n S_1$. It is natural to require that S'_0/S_0 is comparable with $2^{-1/n}$, which is not far from $1-(c_9/n)$. Let us take

$$S'_0 = S_0\left(1-\frac{1}{2n}\right).$$

We need to replace (5.4) by

$$S_1 = \left[c_{10}n\frac{DW}{\log E}\right].$$

From (5.3) we obtain the condition

$$U > c_{11}^n n^n D^{n+2} W V_1 \cdots V_n (\log E^*)(\log E)^{-n-1}. \tag{5.7}$$

We shall see later (Section 5.2, h)) how to remove the coefficient n^n in the right hand side of (5.7), but let us continue.

It seems we have not earned much: we started with roughly $(S_0^n/n!)S_1$ equations and we got $({S'_0}^n/n!)S'_1$ new ones, which is about the same. The only improvement concerns the term $S_0S_1\log E$, that we can replace at the end of this extrapolation by $S'_0S'_1\log E$, which is essentially twice as large.

This procedure may be repeated: we do it n times (this seems to be an optimal choice) with

$$S_0^{(j)} = S_0\left(1-\frac{j}{2n}\right) \quad \text{and} \quad S_1^{(j)} = 2^j S_1.$$

There is a small cost: at the end we shall get $|\Lambda| \geq e^{-2^nU}$ in place of $|\Lambda| \geq e^{-U}$. If one wishes to keep the conclusion $|\Lambda| \geq e^{-U}$ then one should replace everywhere else U by $U/2^n$. Indeed this is just the choice of parameters in [31].

At the end of the n extrapolation steps one has $S_0^{(n)} = S_0/2$ and $S_1^{(n)} = 2^nS_1$. The number of equations is not large enough to apply immediately the zero estimate 3.2.

Since we have been unable to increase the number of equations enough, we shall follow another strategy which originates in works by Baker and Stark and relies on arguments arising from Kummer's theory. The idea is to introduce division points and prove

$$\partial^{\underline{\sigma}} P(s/p, \alpha_1^{s/p}, \ldots, \alpha_n^{s/p}) = 0 \tag{5.8}$$

for various $\underline{\sigma} \in \mathbb{N}^n$, $s \in \mathbb{Z}$ and p. At an early stage of the theory, p was selected as a sufficiently large prime number. Later, another argument, which we explain in Section 5.2, f), enabled us to work with a smaller value for p, until it was realized that $p = 2$ also works!

We refer to [11] for several proofs involving (5.8):

- In Chap. 8, Section 2 and Section 6, for many values (s, p) with $\underline{\sigma} = \underline{0}$.
- In Chap. 10, Section 2 and Section 5, for a single sufficiently large prime number p and several values $(\underline{\sigma}, s)$.
- In Chap. 11, Section 2 and Section 4, (5.8) is applied with $p = 2$ only.

In recent works involving Kummer's theory only $p = 2$ is used; however it may be instructive to recall briefly what was done earlier with a large prime p.

e) Kummer's Theory

In this section we denote by n a positive integer, $\mathbb{K}$ a number field and $\alpha_1, \dots, \alpha_n$ nonzero elements of $\mathbb{K}$.

We first quote Lemma 3 of [4].

Lemma 5.2. *Let p be a prime. For $1 \leq j \leq n$ denote by $\alpha_j^{1/p}$ any p-th root of α_j. For $1 \leq r \leq n$ set*

$$\mathbb{K}_r = \mathbb{K}\bigl(\alpha_1^{1/p}, \dots, \alpha_r^{1/p}\bigr).$$

Assume

$$\bigl[\mathbb{K}_n : \mathbb{K}_{n-1}\bigr] < p.$$

Then there exist an element $\gamma \in \mathbb{K}^\times$ and integers $j_1, \dots, j_{n-1}$ satisfying $0 \leq j_\ell < p \quad (1 \leq \ell \leq n-1)$, such that

$$\alpha_n = \alpha_1^{j_1} \cdots \alpha_{n-1}^{j_{n-1}} \gamma^p. \tag{5.9}$$

Remark. Lemma 5.2 is proved by induction in [4]. When $\mathbb{K}$ contains the p-th roots of unity it can also be proved by using arguments from Kummer's theory as follows. Define

$$G_0 = \bigl\{x^p \; ; \; x \in \mathbb{K}^\times\bigr\}$$

and, for $1 \leq r \leq n$ let G_r denote the multiplicative subgroup of $\mathbb{K}^\times$ spanned by G_0, $\alpha_1, \dots, \alpha_r$. According to [12] Chap. 6, Section 8 Th. 8.1, the field $\mathbb{K}_r$ is an abelian extension of $\mathbb{K}$ and the degree of this extension is the index of G_0 in G_r. Assume p is such that $\bigl[\mathbb{K}_n : \mathbb{K}_{n-1}\bigr] < p$. Then $(G_n : G_{n-1}) < p$, hence there is a relation

$$\alpha_n^{a_n} = \alpha_1^{a_1} \cdots \alpha_{n-1}^{a_{n-1}} \beta^p,$$

with some $\beta \in \mathbb{K}^\times$ and integers $a_1, \dots, a_n$ satisfying $0 \leq a_\ell < p \quad (1 \leq \ell \leq n-1)$ and $1 \leq a_n < p$. Writing Bézout's relation $a_n u + pv = 1$ with rational integers u, v, one deduces the desired relation (5.9).

Lemma 5.3. *For $1 \leq i \leq n$ select a complex logarithm $\log \alpha_j$ of α_j. Assume that the numbers $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over $\mathbb{Q}$ and denote by $\mathcal{G}$ the set of $\lambda \in \mathbb{C}$ such that*

$$e^\lambda \in \mathbb{K}^\times \quad \textit{and} \quad \lambda, \log\alpha_1, \ldots, \log\alpha_n \textit{ are linearly dependent over } \mathbb{Q}.$$

Then $\mathcal{G}$ is a free $\mathbb{Z}$-module of rank n, containing $\mathbb{Z}\log\alpha_1 + \cdots + \mathbb{Z}\log\alpha_n$ as a subgroup of finite index.

Further, if $\log\theta_1, \ldots, \log\theta_n$ is a basis of $\mathcal{G}$ over $\mathbb{Z}$, then for any prime p for which $\mathbb{K}$ contains a primitive p-th root of unity, we have

$$\big[\mathbb{K}\big(\theta_1^{1/p}, \ldots, \theta_n^{1/p}\big) \colon \mathbb{K}\big] = p^n,$$

where $\theta_i^{1/p}$ stands for $\exp\{(1/p)\log\theta_i\}$ $(1 \le i \le n)$.

Remark. The assumption in the last sentence that $\mathbb{K}$ contains a primitive p-th root of unity cannot be omitted. Here is an example: take $\mathbb{K} = \mathbb{Q}$, $n = 1$, $\alpha_1 = 1$ and $\log\alpha_1 = 2i\pi$. In this case $\log\theta_1 = \pm 2i\pi$ and for each odd prime p we have

$$\big[\mathbb{K}(\theta_1^{1/p}) \colon \mathbb{K}\big] = p - 1.$$

Proof. It will be more convenient to work with

$$\Lambda = \big\{(r_1, \ldots, r_n) \in \mathbb{Q}^n \,;\, r_1\log\alpha_1 + \cdots + r_n\log\alpha_n \in \mathcal{G}\big\},$$

which is a subgroup of $\mathbb{R}^n$ isomorphic to $\mathcal{G}$ under the mapping

$$\begin{aligned} \Lambda &\longrightarrow \qquad\qquad G \\ \underline{r} &\longmapsto r_1\log\alpha_1 + \cdots + r_n\log\alpha_n. \end{aligned}$$

We have $\mathbb{Z}^n \subset \Lambda \subset \mathbb{Q}^n$. We first want to check that Λ is discrete in $\mathbb{R}^n$. For $X > 0$ and for $\underline{r} \in \Lambda$ satisfying $|\underline{r}| \le X$ we have

$$\mathrm{h}(\alpha_1^{r_1} \cdots \alpha_n^{r_n}) \le X \sum_{i=1}^n \mathrm{h}(\alpha_i),$$

hence $\alpha_1^{r_1} \cdots \alpha_n^{r_n}$ belongs to a finite subset of $\mathbb{K}^\times$ as $\underline{r}$ ranges over the elements of $\Lambda \cap B_n(0, X)$. This means that the image of $\Lambda \cap B_n(0, X)$ under the mapping

$$\begin{aligned} \Lambda &\longrightarrow \quad K^\times \\ \underline{r} &\longmapsto \alpha_1^{r_1} \cdots \alpha_n^{r_n} \end{aligned}$$

is finite. Hence the proof that $\Lambda \cap B_n(0, X)$ is finite will be completed if we check that for each $\underline{r}^0 \in \mathbb{Q}^n \cap B_n(0, X)$, the set of $\underline{r} \in \mathbb{Q}^n \cap B_n(0, X)$ such that

$$\alpha_1^{r_1} \cdots \alpha_n^{r_n} = \alpha_1^{r_1^0} \cdots \alpha_n^{r_n^0}$$

is finite. Indeed, since $\log\alpha_1, \ldots, \log\alpha_n$ are linearly independent over $\mathbb{Q}$, there exists $\underline{s} \in \mathbb{Q}^n$ such that

$$\big\{\underline{r} \in \mathbb{Q}^n \,;\, \alpha_1^{r_1} \cdots \alpha_n^{r_n} = 1\big\} = \mathbb{Z}\underline{s};$$

we deduce that the set

$$\left\{\underline{r} \in \mathbb{Q}^n \cap B_n(0, 2X)\,;\; \alpha_1^{r_1} \cdots \alpha_n^{r_n} = 1\right\}$$

is finite.

This shows that Λ is a lattice (discrete subgroup of rank n) in $\mathbb{R}^n$, hence a free $\mathbb{Z}$-module, and therefore $\mathcal{G}$ is also a free $\mathbb{Z}$-module of rank n.

Let $\log\theta_1, \ldots, \log\theta_n$ be a basis of $\mathcal{G}$ over $\mathbb{Z}$. Assume

$$\left[\mathbb{K}\big(\theta_1^{1/p}, \ldots, \theta_n^{1/p}\big) : \mathbb{K}\right] < p^n$$

for some prime p. Let m be minimal with $1 \le m \le n$ such that

$$\left[\mathbb{K}\big(\theta_1^{1/p}, \ldots, \theta_m^{1/p}\big) : \mathbb{K}\right] < p^m.$$

We use Lemma 5.2:

$$\theta_m = \theta_1^{j_1} \cdots \theta_{m-1}^{j_{m-1}} \gamma^p,$$

for some $\gamma \in \mathbb{K}^\times$ and $0 \le j_\ell < p$, $(1 \le \ell \le m-1)$. Define

$$\lambda = \frac{1}{p}\log\theta_m - \sum_{\ell=1}^{m-1} \frac{j_\ell}{p}\log\theta_\ell.$$

Since $\mathbb{K}$ contains the p-th roots of unity, from $(e^\lambda/\gamma)^p \in \mathbb{K}^\times$ we deduce $e^\lambda \in \mathbb{K}^\times$, hence $\lambda \in \mathcal{G}$, and therefore $\lambda \in \mathbb{Z}\log\theta_1 + \cdots + \mathbb{Z}\log\theta_n$, which is clearly a contradiction. $\square$

In the transcendence proof following Baker's method, Lemma 5.3 is applied as follows: when we need estimates for the height, we use the algebraic numbers α_i, while when we want to apply Kummer's condition, we use the numbers θ_i. This does not make a difference when using interpolation determinants; if we use an auxiliary function, the systems of equations are equivalent, but in order to investigate small values we need estimates for the transition matrix. The index N of $\mathbb{Z}^n$ in Λ plays an important role in Matveev's paper [20], II.

f) Kummer's Theory with a Large Prime *p*

The results in this subsection are no more used in recent papers dealing with linear independence measures for logarithms of algebraic numbers, but they keep their independent interest.

We denote by n a positive integer and by $\alpha_1, \ldots, \alpha_n$ nonzero algebraic numbers.

Lemma 5.4. *Let $\mathbb{K}$ be a number field containing $\alpha_1, \ldots, \alpha_n$ and let p be a sufficiently large prime.*
a) Assume $\alpha_1, \ldots, \alpha_n$ are multiplicatively independent. Denote by $\alpha_i^{1/p}$ any p-th root of α_i $(1 \le i \le n)$. Then

$$\left[\mathbb{K}\left(\alpha_1^{1/p},\ldots,\alpha_n^{1/p}\right)\colon\mathbb{K}\right]=p^n.$$

b) Assume α_1 is a primitive q-th root of unity for some positive integer q, while $\alpha_2,\ldots,\alpha_n$ are multiplicatively independent. Denote by $\alpha_1^{1/p}$ a primitive pq-th root of unity and by $\alpha_i^{1/p}$ any p-th root of α_i $(2\le i\le n)$. Then

$$\left[\mathbb{K}\left(\alpha_1^{1/p},\ldots,\alpha_n^{1/p}\right)\colon\mathbb{K}\right]=(p-1)p^{n-1}.$$

Proof. a) To start with, consider the case $n=1$. The assumption is that $\alpha=\alpha_1$ is not a root of unity, and the conclusion amounts to say that if α is a p-th power in $\mathbb{K}$, then p is bounded (depending on α and $\mathbb{K}$). We prove this claim by using heights: if $\alpha=\beta^p$ for some $\beta\in\mathbb{K}$, then $\mathrm{h}(\beta)=(1/p)\mathrm{h}(\alpha)$. Moreover β has degree $\le[\mathbb{K}:\mathbb{Q}]$ and is not a root of unity. From a result of Kronecker's (see for instance [34], Section 3.6) $\mathrm{h}(\beta)$ is bounded from below by a positive constant c depending only on $[\mathbb{K}:\mathbb{Q}]$. Hence $p\le c^{-1}\mathrm{h}(\alpha)$.

More generally, assume $\alpha_1,\ldots,\alpha_n$ are multiplicatively independent. Define

$$\mathbb{K}_r=\mathbb{K}\left(\alpha_1^{1/p},\ldots,\alpha_r^{1/p}\right)\quad(1\le r\le n)$$

and $\mathbb{K}_0=\mathbb{K}$. Let p be a prime number such that $[\mathbb{K}_n\colon\mathbb{K}]<p^n$. Denote by r the least integer such that $[\mathbb{K}_r\colon\mathbb{K}]<p^r$. We have $1\le r\le n$, $[\mathbb{K}_{r-1}\colon\mathbb{K}]=p^{r-1}$ and $\left[\mathbb{K}_r\colon\mathbb{K}_{r-1}\right]<p$. By Lemma 5.2 there exists a nontrivial relation

$$\alpha_r=\alpha_1^{j_1}\cdots\alpha_{r-1}^{j_{r-1}}\gamma^p \tag{5.10}$$

with $j_\ell\in\mathbb{Z}$, $0\le j_\ell<p$ $(1\le\ell<r)$ and $\gamma\in\mathbb{K}^\times$. From the properties of the height (see for instance [34] Chap. 3) we deduce

$$\mathrm{h}(\gamma)\le\mathrm{h}(\alpha_1)+\cdots+\mathrm{h}(\alpha_r),$$

hence γ belongs to a finite subset of $\mathbb{K}^\times$ which depends only on $\alpha_1,\ldots,\alpha_n$. For each fixed γ there is a unique multiplicative dependence relation like (5.10), hence p is bounded. Moreover explicit upper bounds for the exponents in such a multiplicative dependence relation are known, depending only on $\alpha_1,\ldots,\alpha_r$ and $[\mathbb{K}:\mathbb{Q}]$ (see for instance [34] Lemma 7.19).

b) Assume now α_1 is a primitive q-th root of unity, while $\alpha_1^{1/p}$ is a primitive pq-th root of unity. For any prime p which does not divide q the number α_1 is a p-th power of an element in $\mathbb{K}$: indeed writing $up+vq=1$ we get $\alpha_1=(\alpha_1^u)^p$. Hence the field $\mathbb{K}_1=\mathbb{K}\left(\alpha_1^{1/p}\right)$ has degree $<p$ over $\mathbb{K}$. Moreover, since α_1^q is a primitive p-th root of unity, the field $\mathbb{Q}(\alpha_1^q)$ is the cyclotomic field of degree $p-1$ and discriminant $\pm p^{p-2}$. Hence, as soon as p does not divide the discriminant of $\mathbb{K}$, the field $\mathbb{K}_1$ has degree $p-1$ over $\mathbb{K}$.

Next we apply part a) of this Lemma 5.4 to the numbers $\alpha_2,\ldots,\alpha_n$ and the field $\mathbb{K}$: we deduce

$$\left[\mathbb{K}\left(\alpha_2^{1/p},\ldots,\alpha_n^{1/p}\right)\colon\mathbb{K}\right]=p^{n-1}.$$

Finally, $\mathbb{K}\left(\alpha_1^{1/p},\ldots,\alpha_n^{1/p}\right)$ is the compositum of $\mathbb{K}_1$ and $\mathbb{K}\left(\alpha_2^{1/p},\ldots,\alpha_n^{1/p}\right)$, hence has degree $(p-1)p^{n-1}$ over $\mathbb{K}$. $\square$

Lemma 5.5. *Assume $\alpha_1,\dots,\alpha_n$ are multiplicatively independent. Then there exists a positive integer D with the following property: if*

$$Q\in\mathbb{K}[X_1,\dots,X_n]$$

is a nonzero polynomial of degree $<D$ in each variable X_i $(1\le i\le n)$ and p a prime with $p\ge D$, then

$$Q\bigl(\alpha_1^{1/p},\dots,\alpha_n^{1/p}\bigr)\neq 0.$$

Proof. Using Lemma 5.4, take D sufficiently large so that for $p\ge D$,

$$\bigl[\mathbb{K}\bigl(\alpha_1^{1/p},\dots,\alpha_n^{1/p}\bigr)\colon\mathbb{K}\bigr]=p^n.$$

For such a p we prove the result by induction on n. For $n=1$, since $\alpha_1^{1/p}$ has degree $p\ge D>\deg Q$ over $\mathbb{K}$, and since Q has coefficients in $\mathbb{K}$, it follows that $\alpha_1^{1/p}$ is not a root of Q.

If Lemma 5.5 holds for $n-1$, then the polynomial

$$P(X)=Q\bigl(\alpha_1^{1/p},\dots,\alpha_{n-1}^{1/p},X\bigr)\in\mathbb{K}_{n-1}[X]$$

is nonzero, has degree $<D$ and coefficients in the field

$$\mathbb{K}_{n-1}=\mathbb{K}\bigl(\alpha_1^{1/p},\dots,\alpha_{n-1}^{1/p}\bigr),$$

while $\alpha_n^{1/p}$ is algebraic of degree $p\ge D$ over this field $\mathbb{K}_{n-1}$. Hence $\alpha_n^{1/p}$ is not a root of P. $\square$

Lemma 5.6. *Assume $\alpha_1,\dots,\alpha_n$ are multiplicatively independent. Then there exists a positive integer D with the following property. Let*

$$Q\in\mathbb{K}[X_0,X_1,\dots,X_n]$$

be a nonzero polynomial of degree $<D_0$ in X_0 and $<D$ in each variable X_i $(1\le i\le n)$. Let p be a prime with $p>D$ and $\mathcal{S}$ a subset of $\mathbb{Z}$ with at least D_0 elements such that $(s,p)=1$ for any $s\in\mathcal{S}$. Then one at least of the numbers

$$Q\bigl(s/p,\alpha_1^{s/p},\dots,\alpha_n^{s/p}\bigr)\quad(s\in\mathcal{S})$$

is not 0.

Proof. For $(s,p)=1$ Bézout's relations imply

$$\mathbb{Q}(\alpha,\alpha^{1/p})=\mathbb{Q}(\alpha,\alpha^{s/p}).$$

Therefore the same arguments as for Lemma 5.5 yield Lemma 5.6. $\square$

Lemmas 5.5 and 5.6 occurred in Baker's method at an earlier stage of the theory (see [11] pp. 176–177 and 185). Later, similar results taking into account several values of $\underline{\sigma}$ were required (see [11], pp. 223–226 and 237–238).

g) Kummer's Theory with $p = 2$

Let p be a prime for which the field generated over $\mathbb{K} = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ by the p-th roots $\alpha_i^{1/p} = \exp((1/p)\log\alpha_i)$ with $1 \le i \le n$ has maximal degree p^n:

$$[\mathbb{K}(\alpha_1^{1/p}, \ldots, \alpha_n^{1/p}) : \mathbb{K}] = p^n. \tag{5.11}$$

This condition (5.11) yields a decomposition of each relation (5.8)

$$\sum_{(\tau,\underline{t})} c_{\tau,\underline{t}} \triangle (s/p; \tau, T_0^\sharp, \sigma_0) \prod_{k=1}^{n-1} (b_n t_k - b_k t_n)^{\sigma_k} \cdot \alpha_1^{t_1 s/p} \cdots \alpha_n^{t_n s/p} = 0$$

for which s is prime to p into p^n equations, where the sum over $(\tau, \underline{t})$ is restricted to the elements with $t_i \equiv t_i^0 \pmod p$. For one at least of these equations, say corresponding to some $\underline{t}^0 \in \mathbb{N}^n$ with $0 \le t_i^0 < p$, the coefficients $c_{\tau,\underline{t}}$ with

$$t_i \equiv t_i^0 \pmod p \quad (1 \le i \le n)$$

do not all vanish. We reduce the number of coefficients of P as follows: fix such a $\underline{t}^0$, write $t_i = t_i^0 + pt_i'$ $(1 \le i \le n)$ and consider now the new auxiliary polynomial obtained as a chunk of P

$$\sum_{(\tau,\underline{t}')} c_{\tau,\underline{t}^0+p\underline{t}'} \triangle (Y_0; \tau, T_0^\sharp) Y_1^{t_1'} \cdots Y_n^{t_n'}.$$

In this process the number of coefficients[18] of the auxiliary polynomial P decreases (the upper bound T_i for the degree in Y_i with $1 \le i \le n$ is divided by p). It turns out that the number of relations (3.6) will be essentially fixed along the inductive process, but at the end of the extrapolation the number of relations will be higher than the number of coefficients of the last auxiliary polynomial P, which is what we were looking for.

Each time Kummer's condition (5.11) is used with, say, $p = 2$, we replace T_i by $T_i/2$. If one performs this extrapolation sufficiently far, one ends up (after the double induction) with some T_j, say T_n, replaced by 0 and the system of equations one gets in this case is easily shown (see Lemma 5.7) to have no nontrivial solution: there is no need to appeal to the zero estimate 3.2.

If one wishes to get $T_n/2^J < 1$ (after J steps), we need $2^J > T_n$. It turns out that such a long extrapolation procedure has a cost: there is another factor $\log V_{n-1}$ (assuming $V_1 \le \cdots \le V_{n-1} \le V_n$) in the final estimate. This is exactly the main result in [31], which improves earlier results by Baker in [3]. See also [20], I, as well as [37] for the p-adic case.

Now Proposition 3.2 enables us to perform a shorter extrapolation, which avoids this cost of $\log V_{n-1}$. This is done in [24] as well as in [5] and [20], II. See also [38], I, for the p-adic case.

[18] By "the number of coefficients" of P we mean the number $T_0 \cdots T_n$ where T_i is the known upper bound for the degree of P with respect to Y_i $(1 \le i \le n)$. This is a loose way of speaking, but the values of $T_0, \ldots, T_n$ should be clear from the context.

h) A Simple Zero Estimate

As we saw in Section 5.2, f), combining the extrapolation with (5.11), one reduces the set of coefficients of the auxiliary polynomial. The next result shows that we reach a contradiction as soon as all coefficients $c_{\tau\underline{t}}$ with $1 \leq t_n < T_n$ vanish.

Lemma 5.7. *Let $T_0, \ldots, T_{n-1}$, S_0, S_1 be positive integers and $\alpha_1, \ldots, \alpha_{n-1}$ nonzero complex numbers. Assume*

$$S_0 S_1 \geq T_0 \quad \textit{and} \quad S_0 \geq T_j \quad \textit{for} \quad 1 \leq j < n.$$

Then the matrix

$$\mathtt{M} = \Big(\partial^{\underline{\sigma}}\big(Y_0^{\tau} Y_1^{t_1} \cdots Y_{n-1}^{t_{n-1}}\big)(s, \alpha_1^s, \ldots, \alpha_{n-1}^s)\Big)_{\substack{(\tau,\underline{t})\\(\underline{\sigma},s)}},$$

where the rows are indexed by $(\tau, \underline{t})$ and the columns by $(\underline{\sigma}, s)$ with

$$0 \leq \tau < T_0, \quad 0 \leq t_i < T_i \quad (1 \leq i \leq n-1) \quad \textit{and} \quad \|\underline{\sigma}\| < S_0, \quad 0 \leq s < S_1,$$

has rank $T_0 T_1 \cdots T_{n-1}$.

Remark. Because of (5.2) and (5.4), the conditions $S_0 S_1 \geq T_0$ and $S_0 \geq T_j$ are responsible for the requirements $D \log E^* \geq \log E$ and $V_j \geq \log E$ in the hypotheses of Theorem 5.1.

Proof. On the subring $\mathbb{C}[Y_0, Y_1, \ldots, Y_{n-1}]$ of $\mathbb{C}[Y_0, Y_1, \ldots, Y_n]$, we have

$$\partial^{\underline{\sigma}} = \left(\frac{\partial}{\partial Y_0}\right)^{\sigma_0} \cdots \left(\frac{\partial}{\partial Y_{n-1}}\right)^{\sigma_{n-1}}.$$

Hence the problem is reduced to a Cartesian product situation which one deals with by induction as follows.

Consider a relation between the rows of $\mathtt{M}$:

$$\sum_{\tau=0}^{T_0-1} \sum_{t_1=0}^{T_1-1} \cdots \sum_{t_{n-1}=0}^{T_{n-1}-1} c_{\tau\underline{t}} \partial^{\underline{\sigma}}\big(Y_0^{\tau} Y_1^{t_1} \cdots Y_{n-1}^{t_{n-1}}\big)(s, \alpha_1^s, \ldots, \alpha_{n-1}^s) = 0$$

for any $(\underline{\sigma}, s)$ satisfying $\|\underline{\sigma}\| < S_0$ and $0 \leq s < S_1$. This means that the polynomial

$$P(\underline{Y}) = \sum_{\tau=0}^{T_0-1} \sum_{t_1=0}^{T_1-1} \cdots \sum_{t_{n-1}=0}^{T_{n-1}-1} c_{\tau\underline{t}} Y_0^{\tau} Y_1^{t_1} \cdots Y_{n-1}^{t_{n-1}}$$

satisfies

$$\partial^{\underline{\sigma}} P(s, \alpha_1^s, \ldots, \alpha_{n-1}^s) = 0 \quad \text{for} \quad \|\underline{\sigma}\| < S_0 \quad \text{and} \quad 0 \leq s < S_1.$$

For $(\sigma_0,\dots,\sigma_{n-2})\in\mathbb{N}^{n-1}$ with $\sigma_0+\cdots+\sigma_{n-2}<S_0$ and for $0\le s<S_1$, the polynomial

$$\left(\frac{\partial}{\partial Y_0}\right)^{\sigma_0}\cdots\left(\frac{\partial}{\partial Y_{n-2}}\right)^{\sigma_{n-2}} P(s,\alpha_1^s,\dots,\alpha_{n-2}^s,Y_{n-1})\in\mathbb{C}[Y_{n-1}]$$

has degree $<T_{n-1}\le S_0$ and a zero at the point α_{n-1}^s of multiplicity $\ge S_0$, hence this polynomial is 0. By induction one shows in the same way that for $1\le k<n$, $(\sigma_0,\dots,\sigma_{k-1})\in\mathbb{N}^k$ with $\sigma_0+\cdots+\sigma_{k-1}<S_0$ and $0\le s<S_1$, the polynomial

$$\left(\frac{\partial}{\partial Y_0}\right)^{\sigma_0}\cdots\left(\frac{\partial}{\partial Y_{k-1}}\right)^{\sigma_{k-1}} P(s,\alpha_1^s,\dots,\alpha_{k-1}^s,Y_k,\dots,Y_{n-1})\in\mathbb{C}[Y_k,\dots,Y_{n-1}]$$

is zero. Hence for $k=1$ we get

$$\left(\frac{\partial}{\partial Y_0}\right)^{\sigma_0} P(s,Y_1,\dots,Y_{n-1})=0\quad\text{for}\quad 0\le\sigma_0<S_0\quad\text{and}\quad 0\le s<S_1.$$

Since $S_0S_1>T_0$ we deduce $P=0$, and therefore $\mathtt{M}$ has rank $T_0T_1\cdots T_{n-1}$. □

Remark. There is nothing special with the field $\mathbb{C}$: the result is valid for any field of zero characteristic. Also the same result holds if we select another basis than $Y_0^\tau Y_1^{t_1}\cdots Y_{n-1}^{t_{n-1}}$ for the space $\mathbb{C}[Y_0,\dots,Y_{n-1}]$, for instance $\triangle(Y_0;\tau,T_0^\sharp)Y_1^{t_1}\cdots Y_{n-1}^{t_{n-1}}$.

i) Removing n^n under Kummer's Condition

In [20], I, E.M. Matveev succeeds to remove the factor n^n in (5.7) under Kummer's condition (5.11) for $p=2$. The idea is the following. He restricts the set of exponents $(\tau,\underline{t})$ of his auxiliary polynomial to a subset satisfying

$$\bigl|t_1\log\alpha_1+\cdots+t_n\log\alpha_n\bigr|\le\frac{U}{MES_1},\tag{5.12}$$

where $M\ge1$ is a new parameter. This allows him to take a larger radius R for the disk where he will apply a Schwarz Lemma (or an approximate Schwarz Lemma): in place of $R=Er$, he may take (almost without extra cost at this place) $R=EMr$. In the conclusion, in place of $(\log E)^{-n-1}$ appears now $\bigl(\log(ME)\bigr)^{-n-1}$. On the other hand, for application of Thue-Siegel's Lemma in the construction of P, it is necessary to estimate from below the number of such elements $(\tau,\underline{t})$. Matveev in [20], I, uses arguments from geometry of numbers. Yu Kunrui [38], II, worked out simpler arguments, using only Dirichlet's box principle, which suffice (the final numerical estimate may be not as sharp, but the argument works as well for the p-adic case; see also [34], Section 9.3 for another adaptation of this argument). Essentially, requiring (5.12) divides the number of tuples $(\tau, \underline{t})$ by M. So in the final result the new parameter U is the old one multiplied by $M\big(\log(ME)\big)^{-n-1}$. Assuming $E = e$ and taking $M = c^n$, one gets rid of the unwanted term n^n in (5.7).

Appendix. From 2 to n Logarithms

In this appendix we develop a remark which originates in Gel'fond's work and has been also used by Bombieri, Bilu and Bugeaud (see [34], Theorem 1.9 and Corollary 10.18). The goal is to deduce, from a nontrivial irrationality measure for the quotient of two logarithms of algebraic numbers, a nontrivial measure of linear independence for n logarithms. The idea is to write the coefficients b_i as $\widetilde{b}q_i + r_i$ with integers $\widetilde{b}$, $q_1, \ldots, q_n$, $r_1, \ldots, r_n$, so that the linear form

$$b_1X_1 + \cdots + b_nX_n$$

has the same value at the point $(\lambda_1, \ldots, \lambda_n)$ than the binary form

$$\widetilde{b}Y_1 + Y_2$$

at the point

$$\big(q_1\lambda_1 + \cdots + q_n\lambda_n, r_1\lambda_1 + \cdots + r_n\lambda_n\big).$$

We need to select $\widetilde{b}$ so that the "remainders" $r_i \in \mathbb{Z}$ have comparatively small absolute values.

We expand this argument in the following lemma.

Lemma 5.8. *Let K be a number field of degree D, $\lambda_1, \ldots, \lambda_n$ elements of $\mathcal{L}$ such that the algebraic numbers $\alpha_i = e^{\lambda_i}$ are in K. Define $V_1, \ldots, V_n$ and V by*

$$V_i = \max\left\{h(\alpha_i), \frac{|\lambda_i|}{D}, \frac{1}{D}\right\} \quad (i = 1, \ldots, n) \quad \textit{and} \quad V = \max\{V_1, \ldots, V_n\}.$$

Further, let $b_1, \ldots, b_n$ be rational integers and let B be a positive integer satisfying $B \geq \max\{|b_1|, \ldots, |b_n|\}$.

Then there exist $\widetilde{\lambda}_1, \widetilde{\lambda}_2$ in $\mathcal{L}$ and $\widetilde{b}$ in $\mathbb{Z}$ such that $\widetilde{\alpha}_1 = e^{\widetilde{\lambda}_1}$ and $\widetilde{\alpha}_2 = e^{\widetilde{\lambda}_2}$ are in K,

$$b_1\lambda_1 + \cdots + b_n\lambda_n = \widetilde{b}\;\widetilde{\lambda}_1 + \widetilde{\lambda}_2,$$

$1 \leq \widetilde{b} \leq B$, and such that the numbers

$$\widetilde{V}_i = \max\left\{h(\widetilde{\alpha}_i), \frac{|\widetilde{\lambda}_i|}{D}, \frac{i}{D}\right\} \qquad (i = 1, 2)$$

satisfy

$$\widetilde{V}_1\widetilde{V}_2 \leq 2n^2B^{1-\kappa_n}V^2 \quad \textit{with} \quad \kappa_n = \frac{1}{2n-1}\cdot$$

For the proof of Lemma 5.8, we shall use Minkowski's Linear Forms Theorem (see for instance [27], Chap. II, S 1, Theorem 2C)[19].

Lemma 5.9. *Let* $\big(x_{ij}\big)_{1\le i,j\le n}$ *be a* $n\times n$ *matrix with determinant* ± 1. *Let* $A_1,\dots,A_n$ *be positive real numbers with product* $A_1\cdots A_n=1$. *Then there exists* $(q_1,\dots,q_n)\in\mathbb{Z}^n\setminus\{0\}$ *such that*

$$|q_1x_{i1}+\cdots+q_nx_{in}|<A_i\quad(1\le i<n)$$

and

$$|q_1x_{n1}+\cdots+q_nx_{nn}|\le A_n.$$

Proof of Lemma 5.8. In the trivial case $b_1=\cdots=b_n=0$ we set $\widetilde{b}=1$, $\widetilde{\lambda}_1=\widetilde{\lambda}_2=0$ and the conclusion is satisfied. So we may assume $(b_1,\dots,b_n)\neq 0$. By symmetry, we may assume $b_1\ge |b_i|$ for $1\le i\le n$.

Further, if $n=1$, we set

$$\widetilde{b}=b_1,\quad \widetilde{\lambda}_1=\lambda_1,\quad \widetilde{\lambda}_2=0,$$

so that $\widetilde{V}_1=V_1$, $\widetilde{V}_2=2/D$ and again the conclusion is satisfied. Hence we shall assume $n\ge 2$.

Define

$$Q=B^{(n-1)\kappa_n}.$$

Notice that the exponent is

$$(n-1)\kappa_n=\frac{1}{2}(1-\kappa_n).$$

Using Lemma 5.9, we deduce that there exist rational integers $q_1,\dots,q_n$, not all of which are zero, satisfying

$$\left|\frac{b_i}{b_1}q_1-q_i\right|<Q^{-1/(n-1)}\quad\text{for}\quad 2\le i\le n\quad\text{and}\quad |q_1|\le Q.$$

Since

$$|q_i|<\left|\frac{b_i}{b_1}\cdot q_1\right|+1\le |q_1|+1,$$

we have $|q_i|\le|q_1|$. In particular $q_1\neq 0$. Replacing, if necessary, all q_i by $-q_i$, we may assume $q_1>0$.

We define

$$\widetilde{b}=\big[b_1/q_1\big].$$

Clearly the inequalities $1\le\widetilde{b}\le B$ are satisfied. Further, set

[19] If one applies Dirichlet's box principle in place of Minkowski's Theorem, one deduces a weaker estimate for $\widetilde{V}_1\widetilde{V}_2$, where $2n^2B^{1-\kappa_n}V^2$ is replaced by $4n^2B^{1-\kappa_n}V^2$.

$$r_i = b_i - \widetilde{b}q_i \quad (1 \le i \le n),$$
$$\widetilde{\lambda}_1 = \sum_{i=1}^n q_i\lambda_i \;\text{ and }\; \widetilde{\lambda}_2 = \sum_{i=1}^n r_i\lambda_i,$$

so that

$$\widetilde{b}\widetilde{\lambda}_1 + \widetilde{\lambda}_2 = \sum_{i=1}^n (\widetilde{b}q_i + r_i)\lambda_i = \sum_{i=1}^n b_i\lambda_i.$$

It remains to estimate $\widetilde{V}_1$ and $\widetilde{V}_2$. Define $R = \max_{1\le i\le n}|r_i|$. We have

$$\widetilde{V}_1 \le nq_1V, \quad \widetilde{V}_2 \le nRV,$$

hence

$$\widetilde{V}_1\widetilde{V}_2 \le n^2q_1RV^2.$$

Finally from the inequalities

$$|b_i - \widetilde{b}q_i| \le \frac{1}{q_1}|b_iq_1 - b_1q_i| + |q_i|\left|\frac{b_1}{q_1} - \widetilde{b}\right| \le q_1^{-1}B^{1-\kappa_n} + q_1$$

and

$$q_1^2 \le Q^2 \le B^{1-\kappa_n}$$

we deduce

$$q_1R \le B^{1-\kappa_n} + q_1^2 \le 2B^{1-\kappa_n}.$$

□

We shall combine Lemma 5.8 with the following sharp estimate for two logarithms (Corollary 1 in [16]).

Theorem 5.10. *Let λ_1, λ_2 be two elements in $\mathcal{L}$ and b_1, b_2 two nonzero rational integers. Define*

$$\alpha_1 = e^{\lambda_1}, \quad \alpha_2 = e^{\lambda_2} \quad \textit{and} \quad D = [\mathbb{Q}(\alpha_1,\alpha_2):\mathbb{Q}].$$

Assume α_1 and α_2 are multiplicatively independent. Let V_1, V_2 and W be positive real numbers satisfying

$$V_i = \max\left\{h(\alpha_i), \frac{|\lambda_i|}{D}, \frac{1}{D}\right\} \qquad (i = 1, 2)$$

and

$$W \ge 1, \quad W \ge \frac{21}{D}, \quad e^W \ge \frac{|b_2|}{DV_1} + \frac{|b_1|}{DV_2}.$$

Then

$$|b_1\lambda_1 + b_2\lambda_2| \ge \exp\{-31D^4V_1V_2W^2\}.$$

Corollary 5.11. *Let $\alpha_1,\dots,\alpha_n$ be nonzero multiplicatively independent algebraic numbers. Under the assumptions of Theorem 5.1, if we set $V=\max\{V_1,\dots,V_n\}$, $E=e$,*

$$B=\max\bigl\{e,\ e^{21/D},\ |b_1|,\dots,|b_n|\bigr\}$$

and $\kappa_n=1/(2n-1)$, then we have

$$|\Lambda|>\exp\{-62n^2D^4B^{1-\kappa_n}(\log B)^2V^2\}.$$

Further similar estimates are easy to produce.

Remark. Such an argument is used in [8] in order to reduce a linear form in logarithms from 5 to 2 terms and to show that all integer solutions to the Diophantine equation in 3 variables

$$\begin{aligned}x^5+(z-1)^2x^4y-(2z^3+4z+4)x^3y^2+(z^4+z^3+2z^2+4z-3)x^2y^3\\ +(z^3+z^2+5z+3)xy^4+y^5=\pm1\end{aligned}$$

are the trivial ones given by

$$\pm(x,y)=\begin{cases}(1,0),\ (0,1), & z\neq-1,\ 0,\\ (1,0),\ (0,1),\ (\pm1,1),\ (-2,1), & z\in\{-1,\ 0\}.\end{cases}$$

6 Sixth Lecture. Matveev's Theorem with Interpolation Determinants

We refer to Nesterenko's lectures for an introduction to Matveev's proof of the following Theorem:

Theorem 6.1. *There exists an absolute positive constant C with the following property. Let n be a positive integer, $\alpha_1,\dots,\alpha_n$ nonzero algebraic numbers and $\log\alpha_1,\dots,\log\alpha_n$ logarithms of $\alpha_1,\dots,\alpha_n$ respectively. Assume that the numbers $\log\alpha_1,\dots,\log\alpha_n$ are $\mathbb{Q}$-linearly independent. Let $b_1,\dots,b_n$ be rational integers, not all of which are zero. Denote by D the degree of the number field $\mathbb{Q}(\alpha_1,\dots,\alpha_n)$ over $\mathbb{Q}$. Further, let W, $V_1,\dots,V_n$ be positive real numbers. Assume*

$$V_j\geq\max\left\{\mathrm{h}(\alpha_j),\ \frac{e|\log\alpha_j|}{D}\right\}\quad(1\leq j\leq n)\quad\textit{and}\quad W\geq1+\log D.$$

Further, assume $b_n\neq0$ and

$$e^W\geq\frac{1}{D}\max_{1\leq j\leq n-1}\left\{\frac{|b_n|}{V_j}+\frac{|b_j|}{V_n}\right\}.$$

Then the number

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_n \log \alpha_n$$

has absolute value bounded from below by

$$|\Lambda| > \exp\{-C^n D^{n+2}(1+\log D) W V_1 \cdots V_n\}.$$

Our goal is to explain how to prove Matveev's Theorem 6.1 by means of interpolation determinants and to avoid the construction of an auxiliary function involving Thue-Siegel-Bombieri-Vaaler's Lemma in [20], II. We aim at obtaining the conclusion with an unspecified (but effectively computable) value for the absolute constant C.

6.1 First Extrapolation

Matveev uses several auxiliary polynomials which are of the form

$$P(\underline{Y}) = \sum_{(\tau,\underline{t})\in\mathcal{L}} c_{\tau,\underline{t}} \bigtriangleup (Y_0;\tau,T_0^\sharp) Y_1^{t_1}\cdots Y_n^{t_n} \in \mathbb{C}[Y_0,\ldots,Y_n],$$

where $\mathcal{L}$ is a suitable set of $(\tau,\underline{t}) \in \mathbb{N}\times\mathbb{N}^n$ with $0 \le \tau < T_0$ and $0 \le t_j < T_j$ $(1 \le j \le n)$. He starts the construction by means of Bombieri-Vaaler's version of Thue-Siegel's Lemma and solves a system of equations (3.3). Next a first extrapolation enables him to get more relations like (3.6), say

$$\begin{aligned} \partial^{\underline{\sigma}} P(s,\alpha_1^s,\ldots,\alpha_n^s) = 0 \quad &\text{for} \quad \|\underline{\sigma}\| < S_0^{(j)} \\ &\text{and} \quad 0 \le s < S_1^{(j)} \quad (0 \le j \le J). \end{aligned} \tag{6.1}$$

At this stage of his proof we consider the matrix $\mathtt{M}$ of the linear system given by equations (6.1). We write this matrix as follows. Consider the column vector with $|\mathcal{L}|$ rows and entries in $\mathbb{Q}[\underline{Y}]$:

$$\mathcal{P}(\underline{Y}) = \Big(\bigtriangleup(Y_0;\tau,T_0^\sharp) Y_1^{t_1}\cdots Y_n^{t_n}\Big)_{(\tau,\underline{t})\in\mathcal{L}}.$$

For each $\underline{\sigma} \in \mathbb{N}^n$, define

$$\mathcal{G}_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}} \mathcal{P}(z,\alpha_1^z,\ldots,\alpha_n^z)$$

and set

$$\mathtt{M} = \Big(\mathcal{G}_{\underline{\sigma}}(s)\Big)_{(\underline{\sigma},s)},$$

where $(\underline{\sigma},s)$ ranges over the set of elements in $\mathbb{N}^n\times\mathbb{N}$ such that $\|\underline{\sigma}\| < S_0^{(j)}$ and $0 \le s < S_1^{(j)}$ for at least one j in the range $0 \le j \le J$.

We consider two cases. The first one is when the rank of $\mathtt{M}$ is maximal, equal to $|\mathcal{L}|$. In this case the proof is very short: we select a maximal nonvanishing determinant

$$\Delta = \det\Big(\mathcal{G}_{\underline{\sigma}_\mu}(s_\mu)\Big)_{1\le\mu\le|\mathcal{L}|}$$

by taking the first $|\mathcal{L}|$ columns in minimal lexicographic ordering like in Section 2.2 and Section 3.5, we bound its absolute value from below by means of Liouville's inequality (Proposition 1.13), and from above thanks to analytic arguments. Corollary 3.6 is not quite sufficient for this purpose: it corresponds to a Schwarz' Lemma like Lemma 1.3 for a function with many zeroes, while we need a statement corresponding to an approximated Schwarz' Principle like Lemma 1.12 for a function with many small values. These estimates provide the required conclusion.

Now we assume $\mathtt{M}$ has rank $< |\mathcal{L}|$. Denote by $L_1 - 1$ this rank. We select a subset $\mathcal{L}_1$ of $\mathcal{L}$ with L_1 elements such that, if we set

$$\mathcal{P}_1(\underline{Y}) = \Big(\triangle(Y_0;\tau,T_0^\sharp)Y_1^{t_1}\cdots Y_n^{t_n}\Big)_{(\tau,\underline{t})\in\mathcal{L}_1}$$

and

$$\mathcal{G}^{(1)}_{\underline{\sigma}}(z) = \partial^{\underline{\sigma}}\mathcal{P}_1(z,\alpha_1^z,\dots,\alpha_n^z),$$

then the associated truncated matrix with L_1 rows only

$$\mathtt{M}_1 = \Big(\mathcal{G}^{(1)}_{\underline{\sigma}}(s)\Big)_{(\underline{\sigma},s)}$$

has rank $L_1 - 1$. Thanks to Lemma 5.7 we may assume that all $(\tau,\underline{t}) \in \mathcal{L}$ with $t_n = 0$ belong to $\mathcal{L}_1$.

Again we select the first $L_1 - 1$ columns of $\mathtt{M}_1$ which are linearly independent like in Section 2.2 and Section 3.5, say

$$\mathcal{G}^{(1)}_{\underline{\sigma}_1}(s_1),\dots,\mathcal{G}^{(1)}_{\underline{\sigma}_{L_1-1}}(s_{L_1-1}),$$

and we consider the polynomial

$$P_1(\underline{Y}) = \det\Big(\mathcal{G}^{(1)}_{\underline{\sigma}_1}(s_1),\dots,\mathcal{G}^{(1)}_{\underline{\sigma}_{L_1-1}}(s_{L_1-1}),\mathcal{P}_1(\underline{Y})\Big).$$

From the construction it follows that P_1 is not zero. We extrapolate on the division points as follows: for $\|\underline{\sigma}\| < S_0^{(J+1)}$ and $0 \le s < S_1^{(J+1)}$ we prove

$$\partial^{\underline{\sigma}}P_1(s/2,\alpha_1^{s/2},\dots,\alpha_n^{s/2}) = 0.$$

Thanks to Lemma 5.3, for odd s each such equation decomposes into 2^n equations. We explain now how to use this fact.

6.2 Using Kummer's Condition

Lemma 6.2. *Let ℓ, $L_1,\dots,L_\ell$, M, N be positive integers, $\mathbb{L}$ a field, $\mathbb{K}$ a subfield of $\mathbb{L}$, $t_1,\dots,t_\ell$ elements in $\mathbb{L}$ which are linearly independent over $\mathbb{K}$,*

$A_1,\dots,A_\ell$, $B_1,\dots,B_\ell$ matrices with entries in $\mathbb{K}$, where A_ν has size $L_\nu\times M$ and B_ν size $L_\nu\times N$ $(1\le\nu\le\ell)$. Define

$$A=\begin{pmatrix}A_1\\ \vdots\\ A_\ell\end{pmatrix}\quad\text{and}\quad C=\begin{pmatrix}A_1 & t_1B_1\\ \vdots & \vdots\\ A_\ell & t_\ell B_\ell\end{pmatrix}$$

and assume

$$\operatorname{rank}(C)=\operatorname{rank}(A)<L_1+\cdots+L_\ell.$$

Then for at least one index ν in the range $1\le\nu\le\ell$ we have

$$\operatorname{rank}(B_\nu)<L_\nu.$$

Proof. For $1\le\nu\le\ell$, write

$$B_\nu=\Bigl(b_{\lambda j}^{(\nu)}\Bigr)_{\substack{1\le\lambda\le L_\nu\\ 1\le j\le N}}=\Bigl(\underline{b}_1^{(\nu)},\dots,\underline{b}_N^{(\nu)}\Bigr),$$

where $\underline{b}_1^{(\nu)},\dots,\underline{b}_N^{(\nu)}$ are the column vectors in $\mathbb{K}^{L_\nu}$. Let r be the rank of the matrix A. Since C has also rank r, each of the column vectors

$$\begin{pmatrix}t_1\underline{b}_j^{(1)}\\ \vdots\\ t_\ell\underline{b}_j^{(\ell)}\end{pmatrix}\quad\text{of}\quad\begin{pmatrix}t_1B_1\\ \vdots\\ t_\ell B_\ell\end{pmatrix}$$

$(1\le j\le N)$ belongs to the space spanned by the column vectors of A. We write the N relations

$$\operatorname{rank}\begin{pmatrix}A_1 & t_1\underline{b}_j^{(1)}\\ \vdots & \vdots\\ A_\ell & t_\ell\underline{b}_j^{(\ell)}\end{pmatrix}=r\qquad(1\le j\le N)$$

expanding by minors:

$$\sum_{\nu=1}^{\ell}\sum_{\lambda=1}^{L_\nu}\alpha_{\nu\lambda}^{(\varrho)}t_\nu b_{\lambda j}^{(\nu)}=0\quad\text{for}\quad 1\le\varrho\le R\quad\text{and}\quad 1\le j\le N,$$

where the coefficients $\alpha_{\nu\lambda}^{(\varrho)}$ are in $\mathbb{K}$ (independent of j) and are not all zero (because A has rank r). Since $t_1,\dots,t_\ell$ are linearly independent over $\mathbb{K}$, we deduce

$$\sum_{\lambda=1}^{L_\nu}\alpha_{\nu\lambda}^{(\varrho)}b_{\lambda j}^{(\nu)}=0\quad\text{for}\quad 1\le\varrho\le R,\quad 1\le j\le N\quad\text{and}\quad 1\le\nu\le\ell.\tag{6.2}$$

There is at least one ν in the range $1\le\nu\le\ell$ such that the coefficients $\alpha_{\nu\lambda}^{(\varrho)}$ with $1\le\varrho\le R$ and $1\le j\le N$ are not all zero. For such a ν the relations (6.2) yield a nontrivial dependence relation between the row vectors of B_ν, hence B_ν has rank $<L_\nu$. $\square$

6.3 Second Extrapolation

At this stage of the proof, after the first extrapolation including division points, we get a new matrix

$$\mathtt{M}_2 = \left(\mathcal{G}^{(2)}_{\underline{\sigma}}(s) \right)_{(\underline{\sigma},s)}$$

built with

$$\mathcal{G}^{(2)}_{\underline{\sigma}}(s) = \partial^{\underline{\sigma}} \mathcal{P}_2(s, \alpha_1^s, \ldots, \alpha_n^s)$$

and

$$\mathcal{P}_2(\underline{Y}) = \left(\triangle(Y_0; \tau, T_0^\sharp) Y_1^{t_1} \cdots Y_n^{t_n} \right)_{(\tau,\underline{t}) \in \mathcal{L}_2}.$$

Now the set $\mathcal{L}_1$ has been reduced to $\mathcal{L}_2$, with

$$|t_i| < T_i/2 \quad \text{for} \quad (\tau, \underline{t}) \in \mathcal{L}_2 \quad \text{and} \quad 1 \le i \le n,$$

while the set of $(\underline{\sigma}, s)$ is bigger than for $\mathtt{M}_1$, say

$$\|\underline{\sigma}\| < S_0^{(2,j)} \quad \text{and} \quad 0 \le s < S_1^{(2,j)} \quad (0 \le j \le J_2). \tag{6.3}$$

The upper index $(2, j)$ corresponds to the j-th step in the second extrapolation (hence we may set $S_0^{(1,j)} = S_0^{(j)}$ and $S_1^{(1,j)} = S_1^{(j)}$).

Before starting our second induction let us have a look at Matveev's arguments in [20], II: his second auxiliary polynomial P_2 is a linear combination of the components of $\mathcal{P}_2(\underline{Y})$; he shows by induction on j the relations

$$\partial^{\underline{\sigma}} P_2(s, \alpha_1^s, \ldots, \alpha_n^s) = 0$$

for all $(\underline{\sigma}, s) \in \mathbb{N}^n \times \mathbb{N}$ in (6.3).

As far as we are concerned we repeat the argument of Section 5.1. However it is necessary to be slightly careful with the construction of the next matrix: for an efficient application of the analytic argument we need to introduce a condition like (3.18). This difficulty did not arise in the first extrapolation (Section 6.1), because the condition

$$|\mathcal{L}| > 2 \binom{S_0 + n - 1}{n} S_1$$

was in force: Matveev needed it to solve his system of linear equations and produce his first auxiliary polynomial. For the second extrapolation this condition is not there and we proceed as follows.

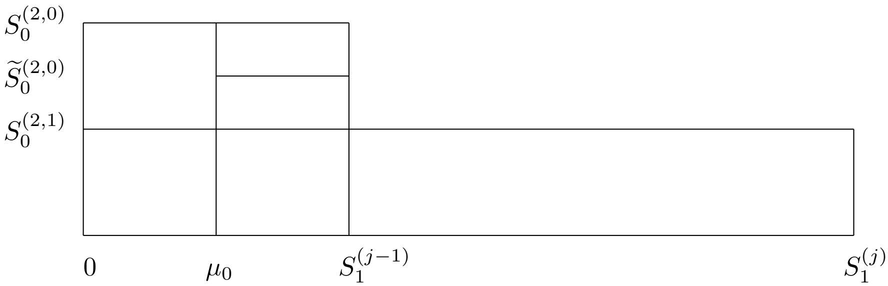
Let $L_2 - 1$ be the rank of $\mathtt{M}_2$. Define $\mu_0 = L_2/2^{n+2}$. We select the first μ_0 columns which are linearly independent. Next we consider the matrix $\widetilde{\mathtt{M}}_2$ which consists of these μ_0 columns together with all other columns of $\mathtt{M}_2$ for which

$$\|\underline{\sigma}\| < \widetilde{S}_0^{(2,j)} \quad \text{and} \quad 0 \le s < S_1^{(2,j)} \quad (0 \le j \le J_2),$$

where

$$\begin{cases} \widetilde{S}_0^{(2,0)} = \frac{1}{2}\big(S_0^{(2,0)} + S_0^{(2,1)}\big) \\ \widetilde{S}_0^{(2,j)} = S_0^{(2,j)} & \text{for } \ 1 \le j \le J_2. \end{cases}$$

When μ_0 is large, there is no point to distinguish between $\mathtt{M}_2$ and $\widetilde{\mathtt{M}}_2$, but otherwise the picture is the following:



Denote by $\widetilde{L}_2$ the rank of $\widetilde{\mathtt{M}}_2$. We already selected μ_0 columns, we select $\widetilde{L}_2 - \mu_0$ further columns in the usual way in order to get a maximal number of independent columns.

If $\widetilde{L}_2 - 1 = |\mathcal{L}_2|$, we easily conclude the proof by usual means: $\widetilde{\mathtt{M}}_2$ is a nonsingular square matrix with nonzero determinant, say Δ_2. We bound the absolute value of Δ_2 from below by arithmetic means and from above by analytic means.

If $\widetilde{L}_2 < \mu_0 = L_2/2^{n+2}$, then the second extrapolation is complete: our goal is precisely to produce a matrix with such a low rank.

Assume now $L_2/2^{n+2} \le \widetilde{L}_2 \le |\mathcal{L}_2|$. We select $\widetilde{L}_2$ rows of $\widetilde{\mathtt{M}}_2$, corresponding to a subset $\widetilde{\mathcal{L}}_2$ of $\mathcal{L}_2$ with $\widetilde{L}_2$ elements. Our new auxiliary polynomial is the determinant $\widetilde{P}_2(\underline{Y})$ of the square matrix

$$\Big(\widetilde{\mathcal{G}}^{(2)}_{\underline{\sigma}_1}(s_1), \ldots, \widetilde{\mathcal{G}}^{(2)}_{\underline{\sigma}_{\widetilde{L}_2-1}}(s_{\widetilde{L}_2-1}), \widetilde{\mathcal{P}}_2(\underline{Y})\Big).$$

We extrapolate on division points and complete the second induction as we did for the first one.

6.4 An Approximate Schwarz Lemma for Interpolation Determinants

The results of Section 3.4 deal with zeroes of interpolation determinants and involve Schwarz' Lemma. They are sufficient for the proof of qualitative statements like Baker's Theorem 3.1.

In Section 4, we were able to prove a quantitative result by means of Lemma 4.6, which is also a Schwarz' Lemma; however only the zero at the origin is used, and there is no extrapolation; this explains why the required approximate Schwarz' Lemma in the proof of Lemma 4.7 is just the interpolation formula of Lemma 1.11 (or just Taylor's expansion as in the remark in Section 4.4).

Such an argument does not seem to suffice when we want to extrapolate. Indeed the determinant

$$\Big(\partial^{\underline{\sigma}}\mathcal{P}\big(s,\alpha_1^s,\ldots,\alpha_{n-1}^s,(\alpha_1^{b_1}\cdots\alpha_{n-1}^{b_{n-1}})^{-s/b_n}\big)(s_\mu)\Big)$$

does not satisfy the assumption (3.12) of Proposition 3.5.

Another attempt is to follow the method of Section 7.3 in [34]: in this case we return to the results of Section 3.4 but we replace in Proposition 3.5 $\mathcal{D}^{\underline{\kappa}_\mu}\varphi_\lambda(\underline{\xi}_\mu)$ by $\mathcal{D}^{\underline{\kappa}_\mu}\varphi_\lambda(\underline{\xi}_\mu)+\epsilon_{\lambda\mu}$ with sufficiently small complex numbers $\epsilon_{\lambda\mu}$. Unfortunately this does not seem to work either: the main hypothesis (3.12) of Proposition 3.5 is not satisfied for a matrix where some entries are replaced by constants.

Finally the only way so far is to combine Lemmas 1.12 and 3.4 by proving an approximate Schwarz Lemma for interpolation determinants.

Details will appear elsewhere.

References

1. Baker, A. – Linear forms in the logarithms of algebraic numbers. I, II, III, IV. Mathematika **13** (1966), 204–216; ibid. **14** (1967), 102–107; ibid. **14** (1967), 220–228; ibid. **15** (1968), 204–216.
2. Baker, A. – *Transcendental number theory.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1975. Second edition, 1990.
3. Baker, A. – The theory of linear forms in logarithms. *Transcendence theory: advances and applications* (Proc. Conf., Univ. Cambridge, Cambridge, 1976), pp. 1–27. Academic Press, London, 1977.
4. Baker, A.; Stark, H. M. – On a fundamental inequality in number theory. Ann. of Math. (2) **94** (1971), 190–199.
5. Baker, A.; Wüstholz, G. – Logarithmic forms and group varieties. J. reine angew. Math. **442** (1993), 19–62.
6. Fel'dman, N. I. – *Hilbert's seventh problem.* Moskov. Gos. Univ., Moscow, 1982.
7. Fel'dman, N. I.; Nesterenko, Y. V. – *Number theory. IV. Transcendental Numbers.* Encyclopaedia of Mathematical Sciences **44**. Springer-Verlag, Berlin, 1998.
8. Gaál, I.; Lettl, G. – A parametric family of quintic Thue equations II. Monatsh. Math. **131**, No. 1 (2000), 29–35.
9. Gramain, F. – Lemme de Schwarz pour des produits cartésiens. Ann. Math. Blaise Pascal **8**, No. 2 (2001), 67-75.
10. Koksma, J. F.– *Diophantische Approximationen.* Reprint. Springer-Verlag, Berlin-New York, 1974.
11. Lang, S. – *Elliptic curves: Diophantine analysis.* Grundlehren der Mathematischen Wissenschaften **231**. Springer-Verlag, Berlin-New York, 1978.
12. Lang, S. – *Algebra.* Third edition. Addison-Wesley Publishing Co., Reading, Mass., 1993.
13. Laurent, M. – Sur quelques résultats récents de transcendance. *Journées Arithmétiques, 1989* (Luminy, 1989). Astérisque No. **198–200** (1991), 209–230.
14. Laurent, M. – Linear forms in two logarithms and interpolation determinants. Acta Arith. **66**, No. 2 (1994), 181–199.

15. Laurent, M.; Roy, D. – Criteria of algebraic independence with multiplicities and approximation by hypersurfaces; J. reine angew. Math. **536** (2001), 65–114.
16. Laurent, M.; Mignotte, M.; Nesterenko, Y. V. – Formes linéaires en deux logarithmes et déterminants d'interpolation. J. Number Theory **55**, No. 2 (1995), 285–321.
17. Mahler, K. – On the approximation of logarithms of algebraic numbers. Philos. Trans. Roy. Soc. London. Ser. A. **245** (1953), 371–398.
18. Mahler, K. – Applications of some formulae by Hermite to the approximation of exponentials and logarithms. Math. Ann. **168** (1967), 200–227.
19. Masser, D.W. – Heights, transcendence, and linear independence on commutative group varieties. This volume.
20. Matveev, E. M. – An explicit lower bound for a homogeneous rational linear forms in the logarithms of algebraic numbers. I and II. Izv. Akad. Nauk SSSR. Ser. Mat. **62** No. 4 (1998), 81–136; ibid. **64** No. 6 (2000), 125–180; Engl. transl.: Izv. Math. **62** No. 4 (1998), 723–772; ibid. **64** No. 6 (2000), 1217–1269.
21. Nesterenko, Y. V.; Waldschmidt, M. – On the approximation of the values of exponential function and logarithm by algebraic numbers. (Russian) *Diophantine approximations, Proceedings of papers dedicated to the memory of Prof. N. I. Fel'dman*, ed. Yu. V. Nesterenko, Centre for applied research under Mech.-Math. Faculty of MSU, Moscow (1996), 23–42.
`Number Theory xxx Math Archives : http://arXiv.org/abs/math/0002047`
22. Philippon, P. – Nouveaux lemmes de zéros dans les groupes algébriques commutatifs. *Symposium on Diophantine Problems* (Boulder, CO, 1994). Rocky Mountain J. Math. **26** No. 3 (1996), 1069–1088.
23. Philippon, P. – Une approche méthodique pour la transcendance et l'indépendance algébrique de valeurs de fonctions analytiques. J. Number Theory **64** No. 2 (1997), 291–338.
24. Philippon, P. and Waldschmidt, M. – Lower bounds for linear forms in logarithms. *New advances in transcendence theory* (Durham, 1986), 280–312, Cambridge Univ. Press, Cambridge-New York, 1988.
25. Philippon, P. and Waldschmidt, M. – Formes linéaires de logarithmes sur les groupes algébriques commutatifs. Illinois J. Math. **32** No. 2 (1988), 281–314.
26. Roy, D. – Interpolation formulas and auxiliary functions. J. Number Theory **94** No. 2 (2002), 248-285.
27. Schmidt, W. M.– *Diophantine approximation.* Lecture Notes in Mathematics, **785**. Springer, Berlin, 1980.
28. Sprindžuk, V. G. – *Classical Diophantine equations.* Translated from the 1982 Russian original (Nauka, Moscow). Translation edited by Ross Talent and Alf van der Poorten. Lecture Notes in Mathematics **1559**. Springer-Verlag, Berlin, 1993.
29. Tijdeman, R. – On the number of zeros of general exponential polynomials. Nederl. Akad. Wetensch. Proc. Ser. A **74** = Indag. Math. **33** (1971), 1–7.
30. Waldschmidt, M. – *Nombres transcendants.* Lecture Notes in Mathematics **402**. Springer-Verlag, Berlin-New York, 1974.
31. Waldschmidt, M. – A lower bound for linear forms in logarithms. Acta Arith. **37** (1980), 257–283.
32. Waldschmidt, M. – *Linear independence of logarithms of algebraic numbers.* The Institute of Mathematical Sciences, Madras, IMSc Report No. **116** (1992), 168 pp.
`http://www.math.jussieu.fr/~miw/articles/IMSc.Rpt.116.html`

33. Waldschmidt, M. – Integer valued functions on products. J. Ramanujan Math. Soc. **12** No. 1 (1997), 1–24.
34. Waldschmidt, M. – *Diophantine Approximation on Linear Algebraic Groups. Transcendence Properties of the Exponential Function in Several Variables.* Grundlehren der Mathematischen Wissenschaften **326**, Springer-Verlag, Berlin-Heidelberg, 2000.
35. Waldschmidt, M. – On a Problem of Mahler Concerning the Approximation of Exponentials and Logarithms. Publ. Math. Debrecen **56** No. 3-4 (2000), 713–738.
36. Wielonsky, F. – Hermite-Padé approximants to exponential functions and an inequality of Mahler. J. Number Theory **74** No. 2 (1999), 230–249.
37. Yu, Kun Rui – Linear forms in p-adic logarithms. Acta Arith. **53** (1989), No. 2, 107–186. II. Compositio Math. **74**, No. 1 (1990), 15–113 and **76** (1990), No. 1-2, 307. III. Compositio Math. **91**, No. 3 (1994), 241–276.
38. Yu, Kun Rui – p-adic logarithmic forms and Group Varieties I. J. reine angew. Math. **502** (1998), 29–92. II. Acta Arith. **89** No. 4 (1999), 337–378.