Transcendence of e and π

Adapted from [1], p. 867–873

Constanze Liaw
Henning Arnór Úlfarsson*

April 7, 2006

1 Introduction

We begin with a few definitions. We say that a complex number α is *algebraic* if it is a root of a polynomial with integer coefficients. So 1, $\sqrt{2}$ and i are all algebraic because they are the roots of the polynomials $x - 1$, $x^2 - 2$ and $x^2 + 1$, respectively. If α fails to be algebraic then it is said to be *transcendental*.

When proving it is impossible to 'square' the circle by a ruler–and–compass construction we have to appeal to the theorem that π is transcendental. It is our goal to prove this theorem. Since the algebraic numbers are the roots of integer polynomials, they are countably many. Cantor's proof in 1874 of the uncountability of the real numbers guaranteed the existence of (uncountably many) transcendental numbers. Thirty years earlier Liouville had actually constructed the transcendental number

$$\sum_{n=0}^{+\infty} \frac{1}{10^{n!}},$$

called *Liouville's constant*. This number is proven to be transcendental using Liouville's approximation theorem, which states: *for any algebraic number α of degree $n \geq 2$, a rational approximation p/q to α must satisfy*

$$\left| \alpha - \frac{p}{q} \right| > \frac{1}{q^n}$$

for sufficiently large q. However, no naturally occurring number, such as e or π, had been proven to be transcendental until in 1873 Hermite disposed of e. Then in 1882 π was proven to be transcendental by Lindemann, using methods related to those of Hermite. In 1900 Hilbert proposed the problem:

*conni@math.brown.edu, henning@math.brown.edu

If α, β are algebraic and $\alpha \neq 0, 1$ and β is irrational, prove that α^β is transcendental.

This problem was solved independently in 1934 by Gelfond and Schneider, and this will follow as a corollary of our main theorem.

2 The Main Theorem

Recall that if K is an extension of a field k then the *transcendence degree* of K over k is the greatest cardinality of algebraically independent subsets of K over k.

To state the main theorem we will need some definitions from complex analysis. Let f be an *entire* function, i.e., f is holomorphic on the whole complex plane. For our purposes we say that f is of *order* $\leq \rho$ if there exists a number $C > 1$ such that for all large R we have

$$|f(z)| \leq C^{R^\rho}, \qquad \text{whenever } |z| \leq R.$$

A *meromorphic* function, i.e., a function holomorphic outside of a discrete set of poles, is said to be of *order* $\leq \rho$ if it is the quotient of two entire functions of order $\leq \rho$. Now we are ready to state our main theorem.

Theorem 1. *Let K be a finite extension of the rational numbers. Let $f_1, \ldots, f_N$ be meromorphic functions of order $\leq \rho$. Assume that the field $K(f_1, \ldots, f_N)$ has transcendence degree ≥ 2 over K and that the derivative $D = \mathrm{d}/\mathrm{d}z$ maps the ring $K[f_1, \ldots, f_N]$ into itself. Let $w_1, \ldots, w_m$ be distinct complex numbers not lying among the poles of the f_i, such that*

$$f_i(w_v) \in K, \qquad \text{for all } i = 1, \ldots, N \text{ and } v = 1, \ldots, m.$$

Then $m \leq 32\rho[K : \mathbf{Q}]$.

Corollary 1 (Hermite–Lindemann). *If α is algebraic (over $\mathbf{Q}$) and $\alpha \neq 0$, then e^α is transcendental. Hence e and π are transcendental.*

Proof. Suppose α and e^α are algebraic. Let $K = \mathbf{Q}(\alpha, e^\alpha)$. The functions z and e^z are algebraically independent over K since if e^z is the root of some polynomial $q(T)$ (in $K[z]$)

$$(e^z)^n + a_{n-1}(e^z)^{n-1} + \cdots + a_1 e^z + a_0 = 0,$$

then the term $(e^z)^n = e^{nz}$ on the left dominates all the other terms for large z, which contradicts the above equality. The ring $K[z, e^z]$ is obviously mapped into itself by the derivative. Now for any $m \geq 1$ we can set

$$w_1 := \alpha, \quad w_2 := 2\alpha, \quad \ldots, \quad w_m := m\alpha,$$

and our functions z and e^z take on algebraic values at all w_i. This implies that $m \leq 32[K : \mathbf{Q}]$ for any m, contradicting the assumption that K is an algebraic, (and thus a finite) extension of $\mathbf{Q}$. (Note that $\rho = 1$). Since $e^1 = e$ and $e^{2\pi i} = 1$, it follows that e and π are transcendental. $\square$

Corollary 2 (Gelfond–Schneider). *If α is algebraic, $\alpha \neq 0, 1$ and if β is algebraic and irrational then $\alpha^\beta = e^{\beta \log \alpha}$ is transcendental.*

Proof. We proceed as in proving Corollary 1, but now we consider the functions $e^{\beta z}$ and e^z. If they are algebraically dependent then $e^{\beta z}$ and e^z would be the roots of a polynomial $q(T_1, T_2)$, so

$$0 = q(e^{\beta z}) = \sum_{i,j=0}^{N} b_{ij} (e^{\beta z})^i (e^z)^j = \sum_{i,j=0}^{N} b_{ij} e^{(i\beta + j)z}.$$

For this equation to hold we must have cancellations of two or more terms, i.e., for some i_1, i_2 and j_1, j_2 we have $i_1\beta + j_1 = i_2\beta + j_2$, or

$$(i_1 - i_2)\beta = j_2 - j_1.$$

This implies that either $i_1 = i_2$ and $j_1 = j_2$ or β is rational. Now let

$$w_1 := \log \alpha, \quad w_2 := 2 \log \alpha, \quad \ldots, \quad w_m := m \log \alpha,$$

so our functions e^z and $e^{\beta z}$ take on algebraic values at the w_i. This gives the desired contradiction on the degree of $K = \mathbf{Q}[\alpha, \alpha^\beta]$. □

3 The Lemmas

The first lemma is due to Siegel and is very important and useful both in algebraic number theory and algebraic geometry. It appears as one of the lemmas on the way in Falting's proof of the Mordell conjecture, which claims that Diophantine equations which give rise to surfaces with two or more holes have only finitely many solutions in Gaussian integers. Falting received the Fields medal in 1986 for his proof.

Lemma 1 (Siegel). *Let*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n &= 0 \end{aligned}$$

be a system of linear equations with integer coefficients a_{ij}, and $n > r$. Let A be a number such that $|a_{ij}| \leq A$ for all i, j. Then there exists an integral, non–trivial solution with

$$|x_j| \leq 2(3nA)^{r/(n-r)}.$$

Proof. If $A < 1$ then all the coefficients are 0 and we can take any solution we want. So assume that $A \geq 1$. We view our system of linear equations as a linear equation $L(X) = 0$, where L is a linear map, $L : \mathbf{Z}^{(n)} \to \mathbf{Z}^{(r)}$, determined by the matrix of coefficients. If B is a positive number, we denote by $\mathbf{Z}^{(n)}(B)$ the set of vectors X in $\mathbf{Z}^{(n)}$ such that $|X| \leq B$ (where $|X|$ is the maximum of the absolute values of the coefficients of X). For our purposes we will assume that $B \geq 1$. Then L maps $\mathbf{Z}^{(n)}(B)$ into $\mathbf{Z}^{(r)}(nBA)$. The number of elements in $\mathbf{Z}^{(n)}(B)$ is bounded below by B^n, (actually by $(2B-1)^n$ but for our purposes this 'worse' lower bound will be better suited) and by above by $(2B+1)^n$. We seek a value of B such that there will be two distinct elements in X, Y in $\mathbf{Z}^{(n)}(B)$ having the same image, $L(X) = L(Y)$. For this it will suffice that $B^n \geq (3nBA)^r$, since $3nBA > 2nBA + 1$. Thus it will suffice that

$$B = (3nA)^{r/(n-r)}.$$

We take $X - Y$ as the solution of our problem. $\square$

Let K be a finite extension of $\mathbf{Q}$, and let $\mathcal{O}_K$ be the *integral closure* of $\mathbf{Z}$ in K, i.e., the set of elements of K satisfying a monomial with coefficients in $\mathbf{Z}$. We call $\mathcal{O}_K$ the set of *algebraic integers*. The set $\mathcal{O}_K$ is a free module over $\mathbf{Z}$ of dimension $[K : \mathbf{Q}]$ (see either Exercise 5 of Chapter IX in [1] or Theorem 29 of Section 15.3 in [2]). We view K as contained in the complex numbers. If α is an element of K, a *conjugate* of α is an element $\sigma\alpha$ where σ is an embedding of K in $\mathbf{C}$. Since K is a finite extension of $\mathbf{Q}$, it is in particular algebraic. If α has a minimal polynomial q then the conjugates of α must also be roots of this polynomial, so there can be only finitely many conjugates of one element. We define the *size* of a set of elements of K to be the maximum of the absolute values of all conjugates of these elements. By the *size* of a vector $X = (x_1, \ldots, x_n)$ we shall mean the size of its coordinates.

For any $\alpha \in K$ we define its *trace* to be

$$\mathrm{Tr}(\alpha) = \sum_{\sigma} \sigma\alpha,$$

where the sum is taken over *distinct* conjugates of α. Let $\omega_1, \ldots, \omega_M$ be a basis of $\mathcal{O}_K$ over $\mathbf{Z}$. Let $\alpha \in \mathcal{O}_K$ and write

$$\alpha = a_1\omega_1 + \cdots + a_M\omega_M.$$

Let $\omega'_1, \ldots, \omega'_M$ be the dual basis of $\omega_1, \ldots, \omega_M$ with respect to the trace. Then we can express the coefficients of a_j as a trace,

$$a_j = \mathrm{Tr}(\alpha\omega'_j).$$

This is possible since the trace is a non–degenerate bilinear form on $\mathcal{O}_K \times \mathcal{O}_K$, and we just choose the dual basis such that $\mathrm{Tr}(\omega_i\omega'_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta symbol.

The trace is a sum over the conjugates. Hence the size of these coefficients is bounded by the size of α times a fixed constant depending on the size of the elements ω'_j.

Lemma 2 (Siegel). *Let K be a finite extension of $\mathbf{Q}$. Let*

$$\begin{aligned} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n &= 0 \\ &\vdots \\ \alpha_{r1}x_1 + \cdots + \alpha_{rn}x_n &= 0 \end{aligned}$$

be a system of linear equations with coefficients in $\mathcal{O}_K$, and $n > r$. Let A be a number such that $\mathrm{size}(\alpha_{ij}) \le A$ for all i, j. Then there exists a non–trivial solution X in $\mathcal{O}_K$ such that

$$\mathrm{size}(X) \le C_1(C_2 nA)^{r/(n-r)},$$

where C_1, C_2 are constants depending only on K.

Proof. Let $\omega_1, \dots, \omega_M$ be a basis of $\mathcal{O}_K$ over $\mathbf{Z}$. Each x_j can be written

$$x_j = \xi_{j1}\omega_1 + \cdots + \xi_{jM}\omega_M$$

with unknowns $\xi_{j\lambda} \in \mathbf{Z}$. Each α_{ij} can be written

$$\alpha_{ij} = a_{ij1}\omega_1 + \cdots + a_{ijM}\omega_M$$

with integers $a_{ij\lambda} \in \mathbf{Z}$. If we multiply out (in K) the $\alpha_{ij}x_j$ we find that our linear equations with coefficients in $\mathcal{O}_K$ are equivalent to a system of rM linear equations in the nM unknowns $\xi_{j\lambda}$

$$\alpha_{ij}x_j = \sum_{k,l=1}^{M} a_{ijk}\omega_k\omega_l\xi_{jl} = \sum_{l=1}^{M}(\sum_{k=1}^{M} a_{ijk}\omega_k\omega_l)\xi_{jl},$$

and from the trace estimate above a_{ijk} is bounded by $C\mathrm{size}(\alpha_{ij})$ where C is a constant depending on the size of the elements ω_j', (i.e., depending only on K). This implies that

$$\sum_{k=1}^{M} a_{ijk}\omega_k\omega_l \le C'A,$$

where C' is a constant depending on C and the size of the products $\omega_k\omega_l$. Therefore the linear system in terms of the $\xi_{j\lambda}$ will have integer coefficents bounded by $C''A$, where C'' is a constant depending on C' and the size of the elements ω_λ. In other words C'' depends only on K. Applying Lemma 1 we obtain a solution in terms of the $\xi_{j\lambda}$ bounded by

$$2(3nMC''A)^{rM/(nM-rM)} = 2(3nMC''A)^{r/(n-r)},$$

and hence a solution $X = (x_j)$ in $\mathcal{O}_K$ that satisfies

$$\mathrm{size}(x_j) \le 2L(3nMC''A)^{r/(n-r)},$$

where L is a constant depending on the sizes of the elements ω_λ. So setting $C_1 = 2L$ and $C_2 = 3nMC''$ we obtain our desired bound. $\square$

The next lemma has to do with estimates of derivatives. By the *size* of a polynomial with coefficients in K we shall mean the size of its set of coefficients. A *denominator* for a set of elements of K will be any positive integer whose product with every element of the set is an *algebraic integer*, i.e., an element of $\mathcal{O}_K$. We define in a similar way a denominator for a polynomial with coefficients in K. We abbreviate 'denominator' by den.

Remark. Denominators always exist: It suffices to look at a single element α of K. Since K is algebraic over $\mathbf{Q}$, α satisfies some polynomial with integer coefficients, $q(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. We set $\mathrm{den}\alpha = a_n$ so $\mathrm{den}\alpha \cdot \alpha = a_n\alpha$ which satisfies a monomial

$$\begin{aligned}
&(a_n\alpha)^n + a_{n-1}(a_n\alpha)^{n-1} + \cdots + a_n^{n-2}a_1(a_n\alpha) + a_{n-1}^n a_0 \\
=&a_n^n\alpha^n + a_n^{n-1}a_{n-1}\alpha^{n-1} + \cdots + a_n^{n-1}a_1\alpha + a_n^{n-1}a_0 \\
=&a_n^{n-1}(a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0) \\
=&0.
\end{aligned}$$

Let

$$P(T_1, \dots, T_N) = \sum_{(v)} \alpha_{(v)} M_{(v)}(T)$$

be a polynomial with complex coefficients, and let

$$Q(T_1, \dots, T_N) = \sum_{(v)} \beta_{(v)} M_{(v)}(T)$$

be a polynomial with non–negative real coefficients. Here (v) is a *multi–index*, which is an N–tuple $(v_1, \dots, v_N)$ of integers. So in the two sums above, where we sum over all N–tuples $(v) = (v_1, \dots, v_N)$, we must require that $\alpha_{(v)}$ and $\beta_{(v)}$ are nonzero for only finitely many (v). Note also that the symbol $M_{(v)}(T)$ represents the monomial $T_1^{v_1}T_2^{v_2}\cdots T_N^{v_N}$.

In this setup we say that Q *dominates* P if $\left|\alpha_{(v)}\right| \le \beta_{(v)}$ for all (v). It is then immediately verified that dominance is preserved under addition, multiplication, and taking partial derivatives with respect to the variables $T_1, \dots, T_N$.

Lemma 3. *Let K be a finite extension of $\mathbf{Q}$. Let $f_1, \dots, f_N$ be functions, holomorphic on a neighborhood of a point $w \in \mathbf{C}$, and assume that $D = \mathrm{d}/\mathrm{d}z$ maps the ring $K[f_1, \dots, f_N]$ into itself. Assume that $f_i(w) \in K$ for all i. Then there exists a number C having the following property. Let $P(T_1, \dots, T_N)$ be any polynomial with coefficients in K, of degree $\le r$. If we set $f := P(f_1, \dots, f_N)$, then we have for all positive integers k,*

$$\mathrm{size}(D^k f(w)) \le \mathrm{size}(P) r^k k! C^{k+r}.$$

Furthermore, there is a denominator for $D^k f(w)$ bounded by $\mathrm{den}(P)C^{k+r}$. [1]

[1] Note that r must satisfy $r \ge h$ where h is an integer depending only on the f_i. See the proof for the details.

Before the proof we need a definition. A *derivation* on the polynomial ring $K[T_1, \ldots, T_N]$ is an additive homomorphism

$$\overline{D} : K[T_1, \ldots, T_N] \to K[T_1, \ldots, T_N],$$
$$\overline{D}(P+Q) = \overline{D}(P) + \overline{D}(Q),$$

also satisfying a *Leibniz condition*

$$\overline{D}(PQ) = \overline{D}(P)Q + P\overline{D}(Q).$$

Proof. There exist polynomials $P_i(T_1, \ldots, T_N)$ with coefficients in K such that

$$Df_i = P_i(f_1, \ldots, f_N).$$

Let h be the maximum of their degrees. There exists a unique derivation $\overline{D}$ on $K[T_1, \ldots, T_N]$ such that $\overline{D}T_i = P_i(T_1, \ldots, T_N)$. For any polynomial P we have

$$\overline{D}(P(T_1, \ldots, T_N)) = \sum_{i=1}^{N} (D_i P)(T_1, \ldots, T_N) \cdot P_i(T_1, \ldots, T_N),$$

where $D_1, \ldots, D_N$ are the partial derivatives. The polynomial P is dominated by

$$\mathrm{size}(P)(1 + T_1 + \cdots + T_N)^r,$$

and each P_i is dominated by $\mathrm{size}(P_i)(1 + T_1 + \cdots + T_N)^h$. Thus $\overline{D}P$ is dominated by

$$\mathrm{size}(P) C_2 r (1 + T_1 + \cdots + T_N)^{r+h},$$

where $C_2 := N \max_i(\mathrm{size}(P_i))$. Now if we differentiate again we find that $\overline{D}^2 P$ is dominated by

$$\begin{aligned} & \sum_{i=1}^{N} \mathrm{size}(P) C_2 r (r+h)(1 + T_1 + \cdots + T_N)^{r+h} P_i (1 + T_1 + \cdots + T_N) \\ \leq \quad & \mathrm{size}(P) C_2^2 r^2 2 (1 + T_1 + \cdots + T_N)^{r+2h}, \end{aligned}$$

and as we can assume that $r \geq h$ we replaced $r + h$ by $2r$. Proceeding inductively, one sees that $\overline{D}^k P$ is dominated by

$$\mathrm{size}(P) C_2^k r^k k! (1 + T_1 + \cdots + T_N)^{r+kh}.$$

Substituting the values $f_i(w)$ for T_i, we obtain the desired bound on $D^k f(w)$:

$$\begin{aligned} \mathrm{size}(D^k f(w)) &= \mathrm{size}(\overline{D}^k P(w)) \\ &\leq \mathrm{size}(P) C_2^k r^k k! (1 + f_1(w) + \cdots + f_N(w))^{r+kh} \\ &= \mathrm{size}(P) r^k k! C^{k+r}, \end{aligned}$$

where we have collected together some consants in C. Note that the first equality above follows from the chain rule. The second assertion in the theorem concerning denominators is also proved by induction. $\square$

Exercise 1. *Prove the existence and uniqueness of the derivation $\overline{D}$ in the proof above.*

4 Proving the Main Theorem

This proof is a prime example of methods for analyzing Diophantine equations.

Proof of the main theorem. Let K be a finite extension of $\mathbf{Q}$. Let $f_1, \ldots, f_N$ be meromorphic functions of order $\leq \rho$. Assume that the field $K(f_1, \ldots, f_N)$ has transcendence degree ≥ 2 and that the derivative $D = \mathrm{d}/\mathrm{d}z$ maps the ring $K[f_1, \ldots, f_N]$ to itself. Let $w_1, \ldots, w_m$ be distinct complex numbers not lying among the poles of the f_i, such that

$$f_i(w_v) \in K, \qquad \text{for all } i = 1, \ldots, N \text{ and } v = 1, \ldots, m.$$

We need to show that $m \leq 32\rho[K : \mathbf{Q}]$.

Let g and h be two functions among $f_1, \ldots, f_N$ which are algebraically independent over K, i.e., for all non–zero polynomials $p(x, y) \in K[x, y]$ we have $p(g(z), h(z)) \neq 0$ for some $z \in \mathbf{C}$. Let t be a positive integer divisible by $2m$. We shall let t tend to infinity at the end of the proof. Define

$$f(g, h) := \sum_{i,j=1}^{t} b_{ij} g^i h^j$$

with $b_{ij} \in K$. Let $l = t^2/2m$. Consider the linear system

$$D^k f(w_v) = 0, \qquad k = 0, \ldots, l-1 \text{ and } v = 1, \ldots, m \tag{1}$$

of lm equations and $2lm$ unknowns b_{ij} with coefficients $D^k g^i h^j|_{w_v} \in K$. Let b_{ij} denote a particular (such that Lemma 2 applies later) non–trivial solution. We multiply these equations by the denominator for the coefficients (without changing notation). Now $D^k g^i h^j|_{w_v} \in \mathcal{O}_K$ and $b_{ij} \in \mathcal{O}_K$.

Next we estimate size(b_{ij}). We are going to apply Lemma 3 with $P(g, h) := g^i h^j|_{w_v}$. Then size$(P) = 1$. So for all $1 \leq i, j \leq t$ and for all $k = 0, \ldots, l-1$

$$\text{size}\left(D^k \left[g^i h^j\right]|_{w_v}\right) \leq t^k k! C^{k+t} \leq l^{(l-1)/2}(l-1)! C^{l-1+\sqrt{l}}.$$

Then an application of Lemma 2 with $\frac{r}{n-r} = \frac{lm}{2lm - lm} = 1$ yields for large l

$$\text{size}(b_{ij}) \leq C\left(l \cdot l^{(l-1)/2}(l-1)! C^{l+\sqrt{l}}\right)^1 \leq O\left(l^{3l}\right), \tag{2}$$

because the exponent of l satisfies $1 + \frac{l}{2} - \frac{1}{2} + l - 1 + l + \sqrt{l} \leq 3l$.

Since g, h are algebraically independent over K, the function f is not identically zero. Let s be the smallest integer such that all derivatives of f up to order $s-1$ vanish at all points $w_1, \ldots, w_m$, but $D^s f$ does not vanish at one of the w, say w_1. Then $l \leq s$. Define

$$\gamma := D^s f(w_1) \in K \backslash \{0\}.$$

Let b be the denominator of γ. Recall that $b_{ij} \in \mathcal{O}_K$. So $\mathrm{den}(f) = 1$. By the last statement of Lemma 3 for large l

$$\mathrm{size}(b) \le C^{s-1+\sqrt{s}}\mathrm{den}(f) \le O\left(C^{3s/2}\right).$$

We introduce the notion of a norm $N_{\mathbf{Q}}^K(\beta)$ of $\beta \in K$ from K to $\mathbf{Q}$. Let $\sigma_1, \dots, \sigma_{[K:\mathbf{Q}]}$ be the linearly independent embeddings of K into an algebraic closure $\bar{\mathbf{Q}}$ of $\mathbf{Q}$. Define

$$N_{\mathbf{Q}}^K(\beta) := \prod_{\mu=1}^{[K:\mathbf{Q}]} \sigma_\mu \beta.$$

With this $N_{\mathbf{Q}}^K(b\gamma)$ is non–zero (because $b\gamma \neq 0$) and an integer. Indeed, we know that $b\gamma$ is an algebraic integer. So the $\sigma_\mu b\gamma$ are algebraic integers (they solve the same polynomial). As a product of algebraic integers, $N_{\mathbf{Q}}^K(b\gamma)$ is an algebraic integer. Since the norm is invariant under $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, it is an algebraic integer in $\mathbf{Q}$. Finally, a rational number $q = \frac{u}{v}$ (reduced to lowest terms with $v > 0$) has minimal polynomial $m_q(x) = vx - u$, so q is an algebraic integer if and only if $m_q(x)$ is monic, i.e., $v = 1$. This is equivalent to $q \in \mathbf{Z}$. Since the definition of the size of an element is the maximum of the absolute values of all conjugates, each conjugate of $b\gamma$ is bounded by

$$C\mathrm{size}(b)\mathrm{size}\left(D^k g^i h^j|_{w_v}\right)\mathrm{size}(b_{ij}) \le O\left(s^{8s}\right),$$

where $k = 0, \dots, s$. Consequently, we get

$$1 \le |N_{\mathbf{Q}}^K(b\gamma)| \le O\left(s^{8s}\right)^{[K:\mathbf{Q}]-1}|\gamma|.$$

It remains to establish the estimate

$$|\gamma| \le \frac{s^{4s}C^s}{s^{ms/(4\rho)}} \tag{3}$$

using global arguments. Indeed, when we let t tend to infinity, then l and s tend to infinity (recall that $l \le s$). Combining the last two inequalities, for large t

$$1 \le O\left(s^{8s}\right)^{[K:\mathbf{Q}]-1}\frac{s^{4s}C^s}{s^{ms/(4\rho)}} = O\left(s^{8s[K:\mathbf{Q}]-3s-ms/(4\rho)}\right).$$

So the powers of this estimate satisfy

$$0 \le 8s[K:\mathbf{Q}] - 3s - ms/(4\rho) \le 8s[K:\mathbf{Q}] - ms/(4\rho),$$

which is equivalent to

$$m \le 32\rho[K:\mathbf{Q}].$$

For the proof of inequality (3), let θ be an entire function of order $\leq \rho$ such that θg and θh are entire (of order $\leq 2\rho$) and $\theta(w_1) \neq 0$. Then $\theta^{2t} f$ is entire. We consider the entire function

$$H(z) := \frac{\theta(z)^{2t} f(z)}{\prod_{v=1}^{m} (z - w_v)^s}.$$

By the maximum modulus principle, the absolute value of $H(w_1)$ is bounded by the maximum of $|H(z)|$ on a large circle of radius R. If we take R large, then the factors $z - w_v$ have approximately the same absolute value as R. By the definition of the order of an entire function, for $|z| = R$ and large R we have

$$|\theta(z)g(z)|^i \leq C^{iR^{2\rho}} \qquad \text{and} \qquad |\theta(z)h(z)|^j \leq C^{jR^{2\rho}}.$$

Combination of the latter two statements with estimate (2) on the size of the b_{ij} yields

$$|H(z)| \leq \frac{s^{3s} C^{2tR^{2\rho}}}{R^{ms}} \leq \frac{s^{3s} C^{2t\sqrt{s}}}{s^{ms/(4\rho)}} \qquad \text{for} \qquad |z| = R := s^{1/(4\rho)},\ s \text{ large}.$$

Since f satisfies system (1), we have $D^k f(w_1) = 0$ for $k = 0, \ldots, s-1$, and by Taylor expansion

$$f(z) = \frac{(z - w_1)^s}{s!} D^s f(w_1) + O\big((z - w_1)^{s+1}\big)$$

for z near w_1. We obtain

$$H(w_1) = \frac{\theta(w_1)^{2t}}{C^s s!} D^s f(w_1).$$

So for large t, taking $t \sim \sqrt{s}$ yields

$$|\gamma| = |D^s f(w_1)| \leq C^s s! \frac{s^{3s} C^{2t\sqrt{s}}}{s^{ms/(4\rho)}} \leq \frac{s^{4s} C^s}{s^{ms/(4\rho)}},$$

which completes the proof. $\square$

5 Conclusion

We have given a classical proof of Hilbert's problem:

If α, β are algebraic and $\alpha \neq 0, 1$ and β is irrational, prove that α^β is transcendental.

We used the techniques of Gelfond and Schneider. This allowed us to prove the transcendence of π and e, along with numbers such as $p^{\sqrt{p}}$ for p prime. Note however that we can only get countably many transcendentals in this way. So there are still uncountably many out there! The following construction, shown to us by Steven J. Miller, gives an explicit formula for uncountably many transcendental numbers. Let α be an irrational number in $[0, 1]$ with binary expansion

$$\alpha = \sum_{n=1}^{\infty} \frac{a_n(\alpha)}{2^n},$$

where $a_n(\alpha) \in \{0, 1\}$. As α is irrational, infinitely many $a_n(\alpha)$ equal 1. Define the number

$$\chi(\alpha) = \sum_{n=1}^{\infty} \frac{1}{10^{-(a_n(\alpha)+1)(2n)!}}.$$

Now, $\chi(\alpha)$ is too well approximated by rational numbers, and so is transcendental by Liouville's theorem. Since there are uncountably many irrational numbers in $[0, 1]$ we get an uncountable collection of transcendental numbers by this construction.

Some famous numbers which are still not known to be transcendental are

- Apéry's constant $\zeta(3) = \sum \frac{1}{n^3}$,
- πe and $\pi + e$, although it is known that they cannot both be algebraic,
- e^e, π^π, π^e.

A conjecture made by Schanuel is:

If $\lambda_1, \ldots, \lambda_n$ are complex numbers, linearly independent over the rationals, then

$$\mathbf{Q}(\lambda_1, \ldots, \lambda_n, e^{\lambda_1}, \ldots, e^{\lambda_n})$$

has transcendence degree at least n.

If this conjecture is true then it follows that e and π are algebraically independent (set $\lambda_1 = 1$ and $\lambda_2 = 2\pi i$), which would imply that both $e + \pi$ and $e\pi$ are transcendental.

References

[1] S. Lang, *Algebra*, Revised Third Edition, Springer–Verlag (2002).

[2] D. S. Dummit and R. M. Foote, *Abstract Algebra*, Third Edition, John Wiley and Sons (2004).