Matthijs J. Warrens

On Irrational and Transcendental Numbers

Bachelor thesis, 9 augustus 2012

Supervisor: Dr. Jan-Hendrik Evertse

Mathematical Institute, Leiden University

Contents

1 Introduction

1.1 History

Number theory is the branch of mathematics that is devoted to the study of integers, subsets of the integers like the prime numbers, and objects made out of the integers. An example of the latter are the rational numbers, the numbers that are fractions, ratios of integers. Examples are $\frac{1}{2}$ and $\frac{22}{7}$. The irrational numbers are those numbers that cannot be represented by fractions of integers. An example of an irrational number that was already known in Ancient Greece is $\sqrt{2}$. The irrationality of e, the base of the natural logarithms, was established by Euler in 1744. The irrationality of π, the ratio of the circumference to the diameter of a circle, was established by Lambert in 1761 (see Baker 1975).

A generalization of the integers are the algebraic numbers. A complex number α is called algebraic if there is a polynomial $f(x) \neq 0$ with integer coefficients such that $f(\alpha) = 0$. If no such polynomial exists α is called transcendental. The numbers $\sqrt{2}$ and i are algebraic numbers since they are zeros of the polynomials $x^2 - 2$ and $x^2 + 1$ respectively. The first to prove the existence of transcendental numbers was Liouville in 1844, using continued fractions. The so-called Liouville constant $\sum_{n=1}^{\infty} 10^{-n!} = 0.1100010...$ was the first decimal example of a transcendental number (see Burger, Tubbs 2004).

In 1873 Hermite proved that e is transcendental. This was the first number to be proved transcendental without having been specifically constructed for the purpose. Building on Hermite's result, Lindemann showed that π is transcendental in 1882. He thereby solved the ancient Greek problem of squaring the circle. The Greeks had sought to construct, with ruler and compass, a square with area equal to that of a given circle. If a unit length is prescribed this amounts to constructing two points in the plane at a distance $\sqrt{\pi}$ apart. In 1837 Wantzel showed that the constructible numbers are a subset of the algebraic numbers. Lindemann showed that $\sqrt{\pi}$ is however transcendental. For a historical overview, see Burger and Tubbs (2004), and Shidlovskii (1989).

In 1874 Cantor showed that the set of algebraic numbers is countably infinite. This follows from the fact that the polynomials with integer coefficients form a countable set and that each polynomial has a finite number of zeros. In the same paper Cantor also showed that the set of real numbers is uncountably infinite. Since the algebraic numbers are countable while the real numbers are uncountable, it follows that most real numbers are in fact transcendental (see Dunham 1990).

At the Second International Congress of Mathematicians in 1900, Hilbert posed a set of 23 problems "the study of which is likely to stimulate the further development of our science". In the 7th of these problems he conjectured that if α and β are algebraic numbers, $\alpha \neq 0, 1$ and β irrational, then α^β is transcendental. In 1934 both Gel'fond and Schneider independently and using different methods obtained a proof of Hilbert's conjecture. It follows from the Gel'fond-Schneider theorem that the numbers $2^{\sqrt{2}}$ and e^π are transcendental (see Shidlovskii 1989).

Since irrational and transcendental numbers are defined by what they are not, it may be difficult, despite their abundance, to show that a specific number is irrational or transcendental. For example, although e and π are irrational, it is unknown whether $e+\pi$, $e-\pi$, $e\pi$, 2^e, π^e or $\pi^{\sqrt{2}}$ are irrational.

1.2 The Riemann zeta function

The Riemann zeta function is defined as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

for a complex number s with $\operatorname{Re} s > 1$. For any positive even integer $2n$ we have the expression

$$\zeta(2n) = (-1)^{n+1} \frac{B_{2n}(2\pi)^{2n}}{2(2n)!},$$

where B_{2n} is the $2n$-th Bernoulli number (see Abramowitz, Stegun 1970, chapter 23). The first few Bernoulli numbers are $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$ and $B_8 = -1/30$. The odd Bernoulli numbers $B_3, B_5, \ldots$ are zero. The expression for $\zeta(2n)$ is due to Euler (Dunham 1990). It is unknown whether there is such a simple expression for odd positive integers.

Since π is a transcendental number, it follows from the above expression for even numbers that $\zeta(2n)$ is transcendental. In 1979 Apéry showed that the number $\zeta(3)$ is irrational. It is unknown if $\zeta(3)$ is also transcendental. Furthermore, it is unknown whether $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ and $\zeta(11)$ are all irrational, although Zudilin (2001) showed that at least one of them is irrational. Moreover, Rivoal (2000) showed that infinitely many of the numbers $\zeta(2n+1)$, where $2n+1$ is an odd integer, are irrational.

1.3 Outline

In this bachelor thesis we consider the proofs of some results on irrational and transcendental numbers. The thesis is organized as follows. In Section 2 we consider the irrationality of e, π and $\zeta(3)$. In Section 3 we prove the transcendence of the numbers e and π. We also consider the Lindemann-Weierstrass theorem in this section. In Section 4 we discuss the Gel'fond-Schneider theorem.

2 Irrational numbers

2.1 A theorem for the numbers e and π

In this subsection all integers are rational integers. Let $c \in \mathbb{R}_{>0}$. Suppose $f(x)$ is a function that is continuous on $[0, c]$ and positive on $(0, c)$. Furthermore, suppose there is associated with f an infinite sequence $\{f_i\}_{i=1}^{\infty}$ of anti-derivatives that are integer-valued at 0 and c and satisfy $f_1' = f$ and $f_i' = f_{i-1}$ for $i \geq 2$. Theorem 1 shows that if such a f exists for c, then the number c is irrational. The proof comes from Parks (1986). It is an extension of a simple proof by Niven (1947) that π is irrational.

Theorem 1. *Let c and $f(x)$ be as above. Then c is irrational.*

Proof: Suppose c is rational. Then there are $m, n \in \mathbb{Z}$ such that $c = m/n$.

First, let P_c be the set of polynomials $p(x) \in \mathbb{R}[x]$ such that $p(x)$ and all its derivatives are integer-valued at 0 and c. The set P_c is closed under addition. Furthermore, repeated application of the product rule shows that P_c is also closed under multiplication. Consider

$$p_0(x) = m - 2nx.$$

Since $p_0(0) = m$, $p_0(c) = -m$ and $p'(x) = -2n$ are all integers, we have $p_0(x) \in P_c$. Next, let $k \in \mathbb{Z}_{\geq 1}$ and let

$$p_k(x) = \frac{x^k(m - nx)^k}{k!}.$$

Using induction on k we will show that $p_k(x) \in P_c$. For $p_1(x) = x(m - nx)$ we have $p_1(0) = p_1(c) = 0$ and $p_1'(x) = p_0(x)$. Hence, $p_1(x) \in P_c$. Next, suppose $p_\ell(x) \in P_c$ and consider

$$p_{\ell+1}(x) = \frac{x^{\ell+1}(m - nx)^{\ell+1}}{(\ell + 1)!}.$$

We have $p_{\ell+1}(0) = p_{\ell+1}(c) = 0$. Furthermore, using the chain rule we have

$$p_{\ell+1}' = \frac{x^\ell(m - nx)^\ell}{\ell!}(m - 2nx) = p_\ell(x)p_0(x).$$

Since P_c is closed under multiplication, and since $p_0(x)$ and $p_\ell(x)$ are in P_c, it follows that $p_{\ell+1}(x) \in P_c$.

Next, since $f(x)$ is continuous on $[0, c]$, it attains a maximum on $[0, c]$. Let M denote this maximum. Furthermore, since $p_k(x)$ is a polynomial for all k it is continuous and differentiable on $[0, c]$. Hence, $p_k(x)$ attains a maximum on $[0, c]$, either in an endpoint or in the interior $(0, c)$ where $p_k'(x) = 0$. The derivative of $p_k(x)$ is $p_{k-1}(x)p_0(x)$. Since $p_{k-1}(x)$ is only zero at $x = 0$ and $x = c$, we must have $p_0(x) = 0$, or $x = m/2n$, in order to have $p'_k(x) = 0$. At $x = m/2n$ we have

$$p_k\left(\frac{m}{2n}\right) = \frac{\left(\frac{m^2}{4n}\right)^k}{k!}.$$

Replacing both $f(x)$ and $p_k(x)$ by their maxima, we obtain

$$\int_0^c f(x)p_k(x)dx \leq \frac{M\left(\frac{m^2}{4n}\right)^k}{k!}\int_0^c dx = \frac{Mc\left(\frac{m^2}{4n}\right)^k}{k!}.$$

The expression on the right-hand side of the inequality goes to 0 when $k \to \infty$. Hence, for sufficiently large k we have the strict inequality

$$\int_0^c f(x)p_k(x)dx < 1.$$

On the other hand, using integration by parts we obtain

$$\int_0^c f(x)p_k(x)dx = f_1(x)p_k(x)\Big|_{x=0}^{c} - \int_0^c f_1(x)p'_k(x)dx.$$

The first term on the right-hand side is an integer by hypothesis. By repeating integration by parts a number of times equal to the degree of $p(x)$, repeatedly integrating the '$f(x)$' part, while differentiating the '$p(x)$' part, we obtain a sum of integers. Hence, the integral $\int_0^c f(x)p_k(x)dx$ is an integer for all k.

Since $\int_0^c f(x)p_k(x)dx$ is an integer, since $f(x)$ is positive on $(0, c)$, and since $p_k(x)$ is positive at $c/2$ and equal to zero only at 0 and c for all k, it follows that $\int_0^c f(x)p_k(x)dx$ is a positive integer, that is,

$$\int_0^c f(x)p_k(x)dx \geq 1,$$

for all k. Hence, we have a contradiction, and we conclude that c is irrational. □

Corollary 2. *π is irrational.*

Proof: π is a positive real number, and $\sin(x)$ is continuous on $[0, \pi]$ and positive on $(0, \pi)$. As a sequence of anti-derivatives of $\sin x$ we may take $-\cos x$, $-\sin x$, $\cos x$, $\sin x$, etc., which all have values from $\{-1, 0, 1\}$ at $x = 0$ or $x = \pi$. □

Corollary 3. *Let $a \in \mathbb{R}_{>0}$, $a \neq 1$. If $\log a$ is rational, then a is irrational.*

Proof: Since $1/a$ is rational if and only if a is rational, and $\log(1/a) = -\log(a)$ is rational if and only if $\log a$ is rational, it suffices to prove the corollary for $a > 1$.

Suppose a is rational. Then there are $m, n \in \mathbb{Z}$ such that $a = m/n$. Since $a > 1$, we have $\log a > 0$. Let $c = \log a$ and apply Theorem 1 with $f(x) = ne^x$. Then we may take the anti-derivatives of f all equal to f. We have $f(0) = n$ and

$$f(c) = f\left(\log \frac{m}{n}\right) = m,$$

which are both integers. It follows from Theorem 1 that $\log a$ is irrational. This contradicts the hypothesis. Hence, we conclude that a is irrational. □

Corollary 4. *e is irrational.*

Proof: e is a real number, $e \neq 1$. Since $\log e = 1$ is a rational number, it follows from Corollary 3 that e is irrational. □

2.2 Auxiliary results for the irrationality of $\zeta(3)$

In the next subsection we show that

$$\zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3} = 1 + \frac{1}{8} + \frac{1}{27} + \cdots$$

is irrational. We give a proof by Beukers (1979). We first prove some lemmas.

Lemma 5. *If $f(x) \in \mathbb{Z}[x]$, then for any $j \in \mathbb{Z}_{\geq 0}$ all the coefficients of the j-th derivative $f^{(j)}(x)$ are divisible by $j!$.*

Proof: Since differentiation is a linear operation, it suffices to prove the lemma for the polynomial x^k for $k > 0$. The j-th derivative is 0 if $j > k$ and if $j \in \{1, 2, \ldots, k\}$ then it is equal to

$$\frac{k!}{(k-j)!} x^{k-j} = j! \binom{k}{j} x^{k-j},$$

in which $\binom{k}{j}$ is an integer. □

Lemma 6. *Let $\epsilon > 0$. Then there is an N_ϵ such that if $n \geq N_\epsilon$, then*

$$d_n := \operatorname{lcm}(1, 2, \ldots, n) < e^{(1+\epsilon)n}.$$

Proof: Let p be a positive prime number and $r \in \mathbb{R}_{>0}$. If p^r divides a number in the set $\{1, 2, \ldots, n\}$, then $p^r \leq n$, and we have $r \leq \log n / \log p$. On the other hand, $p^{[\log n/\log p]}$ does divide one such number, namely itself. Thus,

$$d_n = \prod_{p \le n} p^{[\log n/\log p]}.$$

Let $\pi(n)$ be the prime-counting function that gives the number of primes less than or equal to n. The prime number theorem states that

$$\lim_{n\to\infty} \frac{\pi(n)\log n}{n} = 1.$$

Hence, for n sufficiently large, we have

$$d_n = \prod_{p\le n} p^{[\log n/\log p]} \le \exp\left(\sum_{p\le n} \log n\right) = e^{\pi(n)\log n} < e^{(1+\epsilon)n}.$$

□

Lemma 7. *Let $r, s \in \mathbb{Z}_{>0}$. If $r > s$, then*

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s dxdy \tag{1}$$

is a rational number whose denominator when reduced divides d_r^3. If $r = s$ we have

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s dxdy = 2\left(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3}\right).$$

Proof: Using the identity

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots, \qquad |x| < 1,$$

we obtain that (1) is equal to

$$-\int_0^1 \int_0^1 \sum_{k=0}^{\infty} \log(xy) x^{r+k} y^{s+k} dxdy. \tag{2}$$

Since $x, y \in [0,1]$, the series

$$\sum_{k=0}^{\infty} x^{r+k} y^{s+k}$$

is convergent, and it follows that

$$\int_0^1 \sum_{k=0}^{\infty} \left|\log(xy) x^{r+k} y^{s+k}\right| dx < \infty.$$

Hence, applying Fubini's theorem we obtain that (2) is equal to

$$-\int_0^1 \left(\sum_{k=0}^{\infty} \int_0^1 \log(xy) x^{r+k} y^{s+k} dx \right) dy. \tag{3}$$

Let $k \geq 0$. Using integrating by parts we obtain

$$\begin{aligned}
\int_0^1 (\log x) x^{r+k} dx &= \lim_{\epsilon \to 0} \int_\epsilon^1 (\log x) x^{r+k} dx \\
&= \lim_{\epsilon \to 0} \log x \frac{x^{r+k+1}}{r+k+1} \Bigg|_{x=\epsilon}^1 - \lim_{\epsilon \to 0} \int_\epsilon^1 \frac{x^{r+k}}{r+k+1} dx \\
&= 0 - \lim_{\epsilon \to 0} \frac{x^{r+k}}{(r+k+1)^2} \Bigg|_{x=\epsilon}^1 \\
&= \frac{-1}{(r+k+1)^2}.
\end{aligned}$$

Using this identity and $\log(xy) = \log x + \log y$ in (3), we obtain that the expression in (3) and hence (1) is equal to

$$-\sum_{k=0}^{\infty} \int_0^1 \left(\frac{y^{s+k} \log y}{r+k+1} - \frac{y^{s+k}}{(r+k+1)^2} \right) dy.$$

Integrating next with respect to y we obtain, in a similar fashion,

$$\sum_{k=0}^{\infty} \left(\frac{1}{(r+k+1)(s+k+1)^2} + \frac{1}{(r+k+1)^2(s+k+1)} \right). \tag{4}$$

For $r > s$, we have

$$\begin{aligned}
&\frac{r-s}{(r+k+1)(s+k+1)^2} + \frac{r-s}{(r+k+1)^2(s+k+1)} \\
&\qquad = \frac{r-s}{(r+k+1)(s+k+1)} \left(\frac{1}{s+k+1} + \frac{1}{r+k+1} \right) \\
&\qquad = \left(\frac{1}{s+k+1} - \frac{1}{r+k+1} \right) \left(\frac{1}{s+k+1} + \frac{1}{r+k+1} \right) \\
&\qquad = \frac{1}{(s+k+1)^2} - \frac{1}{(r+k+1)^2}.
\end{aligned}$$

Hence, if $r > s$, (4) and hence (1) can be written as

$$\begin{aligned}
\frac{1}{r-s} \sum_{k=0}^{\infty} \left(\frac{1}{(s+k+1)^2} - \frac{1}{(r+k+1)^2} \right) &= \frac{1}{r-s} \sum_{k=1}^{\infty} \left(\frac{1}{(s+k)^2} - \frac{1}{(r+k)^2} \right) \\
&= \frac{1}{r-s} \sum_{k=1}^{r-s} \frac{1}{(s+k)^2}.
\end{aligned}$$

The least common multiple of $(r-s)(s+1)^2, (r-s)(s+2)^2, \ldots, (r-s)r^2$ is a divisor of d_r^3, which completes the first part of the lemma.

Finally, if $r = s$ (4) and hence (1) becomes

$$2\sum_{k=0}^{\infty}\frac{1}{(r+k+1)^3} = 2\sum_{k=1}^{\infty}\frac{1}{(r+k)^3} = 2\left(\zeta(3) - \sum_{k=1}^{r}\frac{1}{k^3}\right).$$

□

Lemma 8. *Let $D = \{(u,v,w) : u,v,w \in (0,1)\}$. Then the function f given by*

$$f(u,v,w) = \left(u, v, \frac{1-w}{1-(1-uv)w}\right)$$

is a bijection from D to D. Furthermore, its Jacobian determinant is

$$\frac{\partial f(u,v,w)}{\partial(u,v,w)} = \frac{-uv}{(1-(1-uv)w)^2}.$$

Proof: Note that f is defined on D. We first show that $f(D) \subset D$. Let $(u,v,w) \in D$. Since $0 < 1-uv < 1$, we have $0 < 1-w < 1-(1-uv)w < 1$, or

$$0 < \frac{1-w}{1-(1-uv)w} < 1,$$

and hence $f(u,v,w) \in D$, and it follows that f is well-defined.

Next, let $f^2 = f \circ f$ denote the two times iteration of f. We have

$$\begin{aligned} f^2(u,v,w) &= f\left(u, v, \frac{1-w}{1-(1-uv)w}\right) = \left(u, v, \frac{1-\frac{1-w}{1-(1-uv)w}}{1-(1-uv)\frac{1-w}{1-(1-uv)w}}\right) \\ &= \left(u, v, \frac{1-(1-uv)w-(1-w)}{1-(1-uv)w-(1-uv)(1-w)}\right) \\ &= (u,v,w), \end{aligned}$$

that is, f is self-inverse. In particular, f is bijective.

Finally, if we denote $f(u,v,w) = (x,y,z)$, then we have

$$\frac{\partial z}{\partial w} = \frac{-uv}{(1-(1-uv)w)^2},$$

and the Jacobian determinant equals

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{pmatrix} = \frac{\partial z}{\partial w} = \frac{-uv}{(1-(1-uv)w)^2}.$$

□

Lemma 9. *In the region $D = \{(u, v, w) : u, v, w \in (0, 1)\}$, the function*

$$f(u, v, w) = \frac{u(1-u)v(1-v)w(1-w)}{1-(1-uv)w}$$

is bounded from above by $1/27$.

Proof: Let $(u, v, w) \in D$. Using the arithmetic-geometric means inequality we obtain the inequality

$$1-(1-uv)w = (1-w) + uvw \geq 2\sqrt{1-w}\sqrt{uvw}.$$

Hence, we have

$$f(u, v, w) \leq \frac{u(1-u)v(1-v)w(1-w)}{2\sqrt{1-w}\sqrt{uvw}} = \frac{1}{2}\sqrt{u}(1-u)\sqrt{v}(1-v)\sqrt{w(1-w)}.$$

For $t \in [0, 1]$, the maximum of $\sqrt{t}(1-t)$ occurs at $t = 1/3$ and the maximum of $\sqrt{t(1-t)}$ occurs at $t = 1/2$. Hence, we have

$$f(u, v, w) \leq \frac{1}{2} \cdot \frac{1}{\sqrt{3}}\left(1-\frac{1}{3}\right) \cdot \frac{1}{\sqrt{3}}\left(1-\frac{1}{3}\right) \cdot \sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)} = \frac{1}{27}.$$

□

2.3 The irrationality of $\zeta(3)$

Theorem 10. *The number $\zeta(3)$ is irrational.*

Proof: The n-th shifted Legendre polynomial is given by

$$P_n(x) = \frac{1}{n!}\frac{d^n}{dx^n}\left(x^n(1-x)^n\right).$$

The first three polynomials are

$$\begin{aligned} P_1(x) &= 1-2x \\ P_2(x) &= 1-6x+6x^2 \\ P_3(x) &= 1-12x+30x^2-20x^3. \end{aligned}$$

Consider the double integral

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)dxdy.$$

It follows from Lemma 5 that $P_n(x) \in \mathbb{Z}[x]$. Since $P_n(x)$ is of degree n, the quantity $P_n(x)P_n(y)$ is a sum of terms of the form $a_{ij}x^iy^j$ where

$i, j \in \{0, 1, \ldots, n\}$, and $a_{ij} \in \mathbb{Z}$. Since a_{ii} is a square for each i, we have $a_{ii} > 0$ for each i. Note that the double integral can be written as a sum of double integrals of the form in Lemma 7. It follows from Lemma 7 that the double integral is a sum of rational numbers whose denominators divide d_n^3 plus a positive integer multiple of $\zeta(3)$. Hence, there exists integers A_n and $B_n > 0$ such that the double integral equals $(A_n + B_n\zeta(3))/d_n^3$.

Next, we find a second expression for the double integral. Since

$$-\frac{\log(xy)}{1-xy} = -\frac{\log(1-(1-xy)z)}{1-xy}\bigg|_{z=0}^{1} = \int_0^1 \frac{1}{1-(1-xy)z}dz,$$

the double integral becomes

$$\int_0^1\int_0^1\int_0^1 \frac{P_n(x)P_n(y)}{1-(1-xy)z}dxdydz. \tag{5}$$

For $k \in \{0, 1, \ldots, n-1\}$ the multiple derivative $(d^k)/(dx^k)\,(x^n(1-x)^n)$ can be expressed as a sum of terms each having both x and $1-x$ as a factor. Switching order of integration and integrating by parts repeatedly, the triple integral (5) becomes

$$\begin{aligned}
&\frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)\frac{\frac{d^n}{dx^n}\left(x^n(1-x)^n\right)}{1-(1-xy)z}dxdydz \\
&\quad= \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)\frac{1}{1-(1-xy)z}d\left(\frac{d^{n-1}}{dx^{n-1}}\left(x^n(1-x)^n\right)\right)dydz \\
&\quad= \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)yz\frac{\frac{d^{n-1}}{dx^{n-1}}\left(x^n(1-x)^n\right)}{(1-(1-xy)z)^2}dxdydz \\
&\quad= \cdots = \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)n!(yz)^n\frac{x^n(1-x)^n}{(1-(1-xy)z)^{n+1}}dxdydz \\
&\quad= \int_0^1\int_0^1\int_0^1 \frac{x^ny^nz^n(1-x)^nP_n(y)}{(1-(1-xy)z)^{n+1}}dxdydz.
\end{aligned} \tag{6}$$

Applying the transformation of Lemma 8 we have $u = x$, $v = y$,

$$z^n = \frac{(1-w)^n}{(1-(1-uv)w)^n}$$

and

$$(1-(1-xy)z)^{n+1} = \left(1-(1-uv)\frac{1-w}{1-(1-uv)w}\right)^{n+1} = \frac{(uv)^{n+1}}{(1-(1-uv)w)^{n+1}}.$$

The triple integral then becomes

$$\begin{aligned}
&\int_0^1\int_0^1\int_0^1 \frac{u^nv^n(1-u)^n(1-w)^nP_n(v)\,(1-(1-uv)w)^{n+1}}{(1-(1-uv)w)^n\,(uv)^{n+1}}\cdot\frac{uv}{(1-(1-uv)w)^2}\,dudvdw \\
&\quad= \int_0^1\int_0^1\int_0^1 (1-u)^n(1-w)^n\frac{P_n(v)}{1-(1-uv)w}\,dudvdw.
\end{aligned}$$

With the same arguments we used to show that the triple integral in (5) is equal to the integral in (6), but now with respect to v instead of x, we finally obtain the identity

$$\begin{aligned}&\int_0^1\int_0^1 -\frac{\log(xy)}{1-xy}P_n(x)P_n(y)dxdy\\&\quad=\int_0^1\int_0^1\int_0^1 u^n(1-u)^nv^n(1-v)^nw^n(1-w)^n\frac{dudvdw}{(1-(1-uv)w)^{n+1}}.\end{aligned}$$

Applying Lemma 9 and Lemma 7 (with $r=s=0$) we obtain

$$\begin{aligned}0<\int_0^1\int_0^1 -\frac{\log(xy)}{1-xy}P_n(x)P_n(y)dxdy &\le \left(\frac{1}{27}\right)^n\int_0^1\int_0^1\int_0^1\frac{dudvdw}{1-(1-uv)w}\\&=\left(\frac{1}{27}\right)^n\int_0^1\int_0^1 -\frac{\log(uv)}{1-uv}dudv\\&=2\zeta(3)\left(\frac{1}{27}\right)^n.\end{aligned}$$

For a positive integer n and integers A_n and B_n we have

$$0<\frac{|A_n+B_n\zeta(3)|}{d_n^3}<2\zeta(3)\left(\frac{1}{27}\right)^n.$$

Assume now that $\zeta(3)=a/b$ for some integers a,b with $b>0$. By Lemma 6, we have, for sufficiently large n,

$$\begin{aligned}0&<|bA_n+aB_n|\le 2\zeta(3)\left(\frac{1}{27}\right)^n d_n^3 b\\&<2\zeta(3)\left(\frac{1}{27}\right)^n(2.8)^{3n}b=2\zeta(3)\left(\frac{(2.8)^3}{27}\right)^n b<2\zeta(3)(0.9)^n b.\end{aligned}$$

Since bA_n+aB_n is an integer, we obtain a contradiction for sufficiently large n. Hence, $\zeta(3)$ is irrational. □

3 The Hermite-Lindemann approach

In this section we prove the transcendence of the numbers e and π. We also present the Lindemann-Weierstrass theorem. In our proof we follow Baker (1975) and Shidlovskii (1989). We first prove a lemma.

3.1 Hermite's identity

Lemma 11. *Let* $f \in \mathbb{C}[x]$ *with* $\deg f = m$, $u \in \mathbb{C}$, *and let*

$$I(u; f) = \int_0^u e^{u-t} f(t) dt \tag{7}$$

be the integral along the line segment from 0 *to* u. *Then*

$$I(u; f) = e^u \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(u). \tag{8}$$

Proof: Using integration by parts we obtain the relation

$$\begin{aligned} I(u; f) &= -e^{u-t} f(t) \Big|_{t=0}^{u} + \int_0^u e^{u-t} f'(t) dt \\ &= e^u f(0) - f(u) + \int_0^u e^{u-t} f'(t) dt. \end{aligned}$$

If we repeat this process $m - 1$ times we obtain identity (8). □

Identity (8) is also called Hermite's identity (Shidlovskii 1989).

3.2 The number e

The proof of Theorem 12 is a simplified version of the original proof by Hermite. This version can be found in Baker (1975) and Shidlovskii (1989).

Theorem 12. *e is transcendental.*

Proof: Suppose e is algebraic. Then there are $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ with $a_0 \neq 0$ such that

$$\sum_{k=0}^{n} a_k e^k = a_0 + a_1 e + \cdots + a_n e^n = 0. \tag{9}$$

Let p be a prime number with $p > \max\{n, |a_0|\}$ and define

$$f(x) = x^{p-1}(x-1)^p \cdots (x-n)^p. \tag{10}$$

Using this f with $\deg f = m = (n+1)p - 1$ and $I(u; f)$ in (7), we define the quantity

$$J = \sum_{k=0}^{n} a_k I(k; f) = a_0 I(0; f) + a_1 I(1; f) + \cdots + a_n I(n; f)$$

We first derive an algebraic lower bound for $|J|$. Since (9) holds, the contribution to the first summand on the right-hand side of (8) to J is 0, and we have

$$J = -\sum_{k=0}^{n} \sum_{j=0}^{m} a_k f^{(j)}(k).$$

The polynomial $f(x)$ in (10) has 0 as a root of multiplicity $p-1$ and $1, 2, \ldots, n$ as roots of multiplicity p. Hence, we have

$$J = -\sum_{j=p-1}^{m} a_0 f^{(j)}(k) + \sum_{j=p}^{m} \sum_{k=1}^{n} a_k f^{(j)}(k). \tag{11}$$

Since $f(x)$ in (10) can be written as

$$f(x) = x^{p-1} \left((-1)(-2) \cdots (-n) + b_1 x + b_2 x^2 + \cdots + b_n x^n \right)^p$$

for some $b_1, \ldots, b_n \in \mathbb{Z}$, we have

$$f^{p-1}(0) = (p-1)!(-1)^{np}(n!)^p.$$

Due to Lemma 5 each term on the right-hand side of (11) is divisible by $p!$, except for $f^{p-1}(0)$ since $p > n$. Furthermore, since $p > |a_0|$, it follows that J is an integer which is divisible by $(p-1)!$ but not by p. Hence, J is an integer with $|J| \geq (p-1)!$.

Next, we derive an analytic upper bound for $|J|$. On the interval $x \in [0, n]$ each of the factors $x - k$ for $k \in \{0, 1, \ldots, n\}$ is bounded by n. Thus,

$$|f(x)| = |x^{p-1}(x-1)^p \cdots (x-n)^p| \leq n^{(n+1)p-1} \leq \left(n^{n+1}\right)^p,$$

for $x \in [0, r]$. Moreover, we have

$$|I(k; f)| \leq \int_0^k |e^{k-t} f(t)| dt \leq \left(\int_0^k dt \right) e^k \max_{t \in [0,k]} |f(t)| \leq k e^k \left(n^{n+1}\right)^p$$

for $k \in \{0, 1, \ldots, n\}$ and, using the triangle inequality,

$$|J| \leq \sum_{k=0}^{n} |a_k| |I(k; f)| \leq \sum_{k=0}^{n} |a_k| k e^k \left(n^{n+1}\right)^p \leq c_1 c_2^p,$$

for some constants c_1 and c_2 that are independent of p. Since we also have $|J| \geq (p-1)!$, we obtain a contradiction for sufficiently large p. The contradiction proves the theorem. $\square$

3.3 Algebraic integers and the house of an algebraic number

Recall that a complex number α is called algebraic if there is a non-zero polynomial f with integer coefficients such that $f(\alpha) = 0$. There is a unique polynomial $F_\alpha \in \mathbb{Z}[x]$ such that $F_\alpha(\alpha) = 0$, F_α is irreducible in $\mathbb{Q}[x]$, the leading coefficient of F_α is positive, and the coefficients of F_α have greatest common divisor 1. This polynomial F_α is called the minimum polynomial of α. The other zeros in $\mathbb{C}$ of the minimum polynomial of α are called the conjugates of α.

An algebraic number α is said to be an algebraic integer if its minimum polynomial has leading coefficient 1. The algebraic integers form a subring of $\mathbb{C}$. If α is algebraic we have

$$b_n\alpha^n + b_{n-1}\alpha^{n-1} + \cdots + b_1\alpha + b_0 = 0$$

for certain $b_0, \ldots, b_n \in \mathbb{Z}$ with $b_n \neq 0$. If we multiply this equation by b_n^{n-1} we obtain

$$(b_n\alpha)^n + b_{n-1}(b_n\alpha)^{n-1} + \cdots + b_n^{n-2}b_1(b_n\alpha) + b_n^{n-1}b_0 = 0.$$

Hence, if α is an algebraic number and b_n is the leading coefficient of its minimal polynomial, then $b_n\alpha$ is an algebraic integer.

Let $\alpha_1 \in \mathbb{C}$ be an algebraic number and let α_i for $i \in \{2, 3, \ldots, n\}$ denote the conjugates of α_1 in $\mathbb{C}$. The house of α_1 denoted by $\overline{|\alpha_1|}$ is defined as

$$\overline{|\alpha_1|} = \max\{|\alpha_1|, |\alpha_2|, \ldots, |\alpha_n|\}.$$

The following lemma will be used in the proof of the Gel'fond Schneider theorem in Section 4.

Lemma 13. *Let $\alpha_1 \in \mathbb{C}$, $\alpha_1 \neq 0$ be algebraic and* $\deg \alpha_1 = n$. *Let $T \in \mathbb{Z}$, $T > 0$ be such that $T\alpha_1$ is an algebraic integer. Then*

$$|\alpha_1| \geq \frac{1}{T^n \overline{|\alpha_1|}^{n-1}}.$$

Proof: Let α_i for $i \in \{2, 3, \ldots, n\}$ denote the conjugates of α_1. Since the numbers $T\alpha_i$ for $i \in \{1, 2, \ldots, n\}$ are algebraic integers, the number $T\alpha_1 T\alpha_2 \cdots T\alpha_n = T^n\alpha_1 \cdots \alpha_n$ is an algebraic integer. Since the minimum polynomial of $T\alpha_1$ is given by

$$(x - T\alpha_1)(x - T\alpha_2) \cdots (x - T\alpha_n) \in \mathbb{Z}[x],$$

it follows that $T^n\alpha_1 \cdots \alpha_n \in \mathbb{Z}$, and thus that $|T^n\alpha_1 \cdots \alpha_n| \geq 1$. Hence,

$$|\alpha_1| \geq \frac{|\alpha_1 \cdots \alpha_n|}{\overline{|\alpha_1|}^{n-1}} = \frac{|T^n\alpha_1 \cdots \alpha_n|}{T^n\overline{|\alpha_1|}^{n-1}} \geq \frac{1}{T^n\overline{|\alpha_1|}^{n-1}}.$$

□

From here on we make a distinction between rational integers, which are simply elements of $\mathbb{Z}$, and algebraic integers.

3.4 The number π

The proof of Theorem 14 is a simplified version of the original proof by Lindemann. This version can be found in Baker (1975) and Shidlovskii (1989).

Theorem 14. *π is transcendental.*

Proof: Suppose π is algebraic. Then πi is also algebraic. Let $\alpha_1 = \pi i$ with $\deg \alpha_1 = d$, and let $\alpha_2, \ldots, \alpha_d$ be the conjugates of α_1. Since $1 + e^{\pi i} = 0$, we obtain

$$\prod_{\ell=1}^{d} (1 + e^{\alpha_\ell}) = (1 + e^{\alpha_1}) \cdots (1 + e^{\alpha_d}) = 0.$$

If we expand this product, we obtain

$$\prod_{\ell=1}^{d} (1 + e^{\alpha_\ell}) = \sum_{\epsilon_1=0}^{1} \cdots \sum_{\epsilon_d=0}^{1} e^{\epsilon_1\alpha_1+\cdots+\epsilon_d\alpha_d}$$

The exponents inside the multiple sum include some which are non-zero, for example, $\epsilon_1 = 1$ and $\epsilon_2 = \cdots = \epsilon_d = 0$, and also some which are zero, for example, $\epsilon_1 = \cdots = \epsilon_d = 0$. Call the exponents $\theta_1, \theta_2, \ldots, \theta_{2^d}$ and let the first n be the non-zero ones. We have $n < 2^d$, and

$$2^d - n + e^{\theta_1} + e^{\theta_2} + \cdots + e^{\theta_n} = 0. \qquad (12)$$

It turns out that the numbers $\theta_1, \ldots, \theta_n$ are the zeros of a polynomial $g(x) \in \mathbb{Z}[x]$ of degree n. We have the polynomial

$$h(x) = \prod_{\epsilon_1=0}^{1} \cdots \prod_{\epsilon_d=0}^{1} (x - (\epsilon_1\alpha_1 + \cdots + \epsilon_d\alpha_d))$$

with $\deg h = 2^d$. If we consider $h(x)$ as a polynomial in $\alpha_1, \ldots, \alpha_d$, then $h(x)$ is symmetric in $\alpha_1, \ldots, \alpha_d$. Since $\alpha_1, \ldots, \alpha_d$ are a complete set of conjugates, it follows from the theory of elementary symmetric functions that $h(x) \in \mathbb{Q}[x]$. The zeros of $h(x)$ are $\theta_1, \ldots, \theta_n$, and 0 with multiplicity $2^d - n$. Hence, the polynomial $h(x)/x^{2^d-n} \in \mathbb{Q}[x]$ of degree n has precisely the numbers $\theta_1, \ldots, \theta_n$ as its zeros. If we let r be the least common denominator of the coefficients of $h(x)/x^{2^d-n}$, then the polynomial

$$g(x) = \frac{r}{x^{2^d-n}} h(x) \in \mathbb{Z}[x]$$

has also precisely $\theta_1, \ldots, \theta_n$ as its zeros.

Next, let p be a prime number, let b be the leading coefficient of $g(x)$, and define

$$f(x) = b^{(n-1)p} x^{p-1} g^p(x) = b^{np} x^{p-1} (x - \theta_1)^p \cdots (x - \theta_n)^p$$

with $\deg f = m = (n+1)p - 1$. Furthermore, using $I(u; f)$ in (7) we define

$$J = \sum_{k=1}^{n} I(\theta_k; f) = I(\theta_1; f) + I(\theta_2; f) + \cdots + I(\theta_n; f).$$

We first derive an algebraic lower bound for $|J|$. Using (8) and (12) we can write J as

$$J = -\left(2^d - n\right) \sum_{j=p-1}^{m} f^{(j)}(0) - \sum_{j=p}^{m} \sum_{k=1}^{n} f^{(j)}(\theta_k). \tag{13}$$

It turns out that the inner sum over k is a rational integer. Indeed, first note that since $b\alpha_\ell$ for $\ell \in \{1, 2, \ldots, d\}$ is an algebraic integer, $b\theta_k$ for $k \in \{1, 2, \ldots, n\}$ is also an algebraic integer. Furthermore, since $g(x) \in \mathbb{Z}[x]$ we have that $f(x) \in \mathbb{Z}[x]$. Hence, since the sum over k is a symmetric polynomial in $b\theta_1, \ldots, b\theta_n$ with coefficients in $\mathbb{Z}$ and thus a symmetric polynomial with rational integer coefficients in the 2^d numbers $b(\epsilon_1\alpha_1 + \cdots + \epsilon_d\alpha_d)$, it follows from the theory of elementary symmetric functions that the sum over k is a rational integer.

Since $f^{(j)}(\theta_k) = 0$ for $j < p$, it follows from Lemma 5 that the double sum in (13) is a rational integer divisible by $p!$. Furthermore, we have $f^{(j)}(0) = 0$ for $j < p - 1$ and $f^{(j)}(0)$ is divisible by $p!$ for $j \geq p$ due to Lemma 5. It follows from the theory of elementary symmetric functions that

$$f^{(p-1)}(0) = b^{np}(p-1)!(-1)^{np}(\theta_1\theta_2 \cdots \theta_n)^p,$$

is divisible by $(p-1)!$. However, if p is sufficiently large $f^{(p-1)}(0)$ is not divisible by $p!$. Hence, if $p > 2^d - n$ it follows that $|J| \geq (p-1)!$.

Similar to the proof of Theorem 12 we can derive that $|J| \leq c_1 c_2^p$ where c_1 and c_2 are constants that are independent of p. We get a contradiction, which completes the proof. $\square$

3.5 The Lindemann-Weierstrass theorem

Theorems 12 and 14 on the transcendence of e and π are special cases of a more general result which Lindemann sketched in 1882. The result was later rigorously demonstrated by Weierstrass in 1885 (see Baker 1975). The proof of Theorem 15 comes from Baker (1975).

Theorem 15. *For any distinct numbers $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{Q}}$, and non-zero numbers $\beta_1, \dots, \beta_n \in \overline{\mathbb{Q}}$, we have $\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n} \neq 0$.*

Proof: Suppose

$$\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n} = 0. \tag{14}$$

We can assume that the β_i are rational integers. If this is not the case, we consider the product of all the expressions formed by substituting for one or more of the β_j one of its conjugates. Suppose β_j has degree m_j, let its m_j conjugates be denoted by $\beta_j(i_j)$ for $i_j \in \{1, 2, \dots, m_j\}$, and put

$$M = \prod_{j=1}^{n} m_j.$$

The product is given by

$$\prod_{i_1=1}^{m_1} \cdots \prod_{i_n=1}^{m_n} \left(\beta_1(i_1)e^{\alpha_1} + \cdots + \beta_n(i_n)e^{\alpha_n}\right) = \sum_{j_1,\dots,j_n} \beta(j_1, \dots, j_n) e^{j_1\alpha_1 + \cdots + j_n\alpha_n},$$

where the latter sum is taken over all tuples of non-negative integers $(j_1, \dots, j_n)$ with $j_1 + \cdots + j_n = M$ and $\beta(j_1, \dots, j_n)$ is a polynomial expression in $b_1(1), \dots, \beta_n(m_n)$ which has rational integer coefficients and which is invariant under any permutation of $(\beta_i(1), \dots, \beta_i(m_i))$ for $i \in \{1, 2, \dots, n\}$. Hence, all $\beta(j_1, \dots, j_n) \in \mathbb{Q}$. Let $\gamma_1, \dots, \gamma_t$ be the distinct numbers among the $j_1\alpha_1 + \cdots + j_n\alpha_n$. Then the product becomes

$$\delta_1 e^{\gamma_1} + \cdots + \delta_t e^{\gamma_t},$$

where each δ_i is the sum of some of the terms $\beta(j_1, \dots, j_n)$. Hence, $\delta_1, \dots, \delta_t \in \mathbb{Q}$. To complete, we multiply the rational numbers by a common denominator.

We now show that at least one of the new coefficients δ_j is non-zero. To this end, we define on $\mathbb{C}$ a lexicographic ordering $\prec$ such that $\zeta \prec \eta$ if $\operatorname{Re}\zeta < \operatorname{Re}\eta$ or $\operatorname{Re}\zeta = \operatorname{Re}\eta$ and $\operatorname{Im}\zeta < \operatorname{Im}\eta$. If $\zeta_1, \dots \zeta_r, \eta_1, \dots \eta_r$ are complex numbers with $\zeta_1 \prec \eta_1, \dots, \zeta_r \prec \eta_r$, then it holds that $\zeta_1 + \cdots + \zeta_r \prec \eta_1 + \cdots + \eta_r$. We assume without loss of generality that $\alpha_1 \prec \cdots \prec \alpha_n$ and $\gamma_1 \prec \cdots \prec \gamma_t$. Hence, we have $\gamma_t = M\alpha_n$ and $j_1\alpha_1 + \cdots + j_n\alpha_n < \gamma_t$ for $(j_1, \dots, j_n) \neq (0, \dots, M)$, and thus $\delta_t = (\beta_n(1) \cdots \beta_n(m_n))^{m_1 \cdots m_n} \neq 0$.

Next, we can assume that the set $\{\alpha_1, \dots, \alpha_n\}$ is closed under conjugation, that is, it contains all conjugates of each element occurring in it, and moreover, for any two indices j and k such that α_j and α_k are conjugates, we have $\beta_j = \beta_k$.

If this is not the case, let K be any finite normal extension of $\mathbb{Q}$ containing $\alpha_1, \dots, \alpha_n$, and let $\{\sigma_1, \dots, \sigma_m\}$ be the Galois group of $K/\mathbb{Q}$. Then clearly,

$$\prod_{i=1}^{m} (\beta_1 e^{\sigma_i(\alpha_1)} + \cdots + \beta_n e^{\sigma_i(\alpha_n)}) = 0.$$

By expanding the product on the left-hand side, we get

$$\sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \beta_{i_1} \cdots \beta_{i_m} \exp(\sigma_1(\alpha_{i_1}) + \cdots + \sigma_m(\alpha_{i_m})) = 0.$$

By grouping together those terms for which the exponents $\sigma_1(\alpha_{i_1}) + \cdots + \sigma_m(\alpha_{i_m})$ have equal values we obtain an identity of the form

$$\delta_1 e^{\gamma_1} + \cdots + \delta_t e^{\gamma_t} = 0,$$

where $\gamma_1, \dots, \gamma_t$ are the distinct numbers among the exponents $\sigma_1(\alpha_{i_1}) + \cdots + \sigma_m(\alpha_{i_m})$. Clearly, $\{\gamma_1, \dots, \gamma_t\}$ is closed under conjugation, and $\delta_j = \delta_k$ whenever γ_j and γ_k are conjugate to one another.

It remains to show that at least one of the numbers δ_k is non-zero, and for this, we use the argument from above. For $i \in \{1, 2, \dots, m\}$, let j_i be the index j for which $\sigma_i(\alpha_j)$ is the largest among $\sigma_i(\alpha_1), \dots, \sigma_i(\alpha_n)$ in the lexicographic ordering. This index j_i is unique since $\alpha_1, \dots, \alpha_n$ are distinct. Then $\sigma_1(\alpha_{j_1}) + \cdots + \sigma_m(\alpha_{j_m}) = \gamma_k$ is in the lexicographic ordering larger than all other exponents $\sigma_1(\alpha_{i_1}) + \cdots + \sigma_m(\alpha_{i_m})$ and thus, the coefficient $\delta_k = \beta_{i_1} \cdots \beta_{i_m} \neq 0$.

For the remainder of the proof we can now assume that

$$\beta_1 e^{\alpha_1} + \beta_2 e^{\alpha_2} + \cdots + \beta_n e^{\alpha_n} = 0, \tag{15}$$

where $\alpha_1, \dots, \alpha_n$ are distinct and the β_i are rational integers, and that there are integers $0 = n_0 < n_1 < \cdots < n_r$ such that $\alpha_{n_t+1}, \dots, \alpha_{n_{t+1}}$ is a complete set of conjugates for each t, and

$$\beta_{n_t+1} = \beta_{n_t+2} = \cdots = \beta_{n_{t+1}}.$$

Since the $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ are algebraic, we can choose a non-zero rational integer b such that $b\alpha_1, \dots, b\alpha_n$ and $b\beta_1, \dots, b\beta_n$ are algebraic integers. Let p be a prime number and define for $i \in \{1, 2, \dots, n\}$ the functions

$$f_i(x) = b^{np} \frac{[(x - \alpha_1) \cdots (x - \alpha_n)]^p}{(x - \alpha_i)}$$

with $\deg f_i = m = np - 1$. Using these $f_i(x)$ and $I(u; f)$ in (7) we define for $i \in \{1, 2, \dots, n\}$ the quantities

$$J_i = \sum_{k=1}^{n} \beta_k I_i(\alpha_k; f_i) = \beta_1 I_i(\alpha_1; f_i) + \cdots + \beta_n I_i(\alpha_n; f_i)$$

We first derive an algebraic lower bound for $|J_1 \cdots J_n|$. Using (8) and (15) we obtain

$$J_i = -\sum_{j=0}^{m} \sum_{k=1}^{n} \beta_k f_i^{(j)}(\alpha_k).$$

Using a modification of Lemma 5, we find that $f_i^{(j)}(\alpha_k)$ is $p!$ times an algebraic integer unless $j = p - 1$ and $k = i$. In this particular case we have

$$f_i^{(p-1)}(\alpha_i) = b^{np}(p-1)! \prod_{k=1, k \neq j}^{n} (\alpha_i - \alpha_k)^p.$$

Hence, $f_i^{(p-1)}(\alpha_i)$ is an algebraic integer divisible by $(p-1)!$ but not by $p!$ if p is sufficiently large. It then follows that J_i is an algebraic integer that is divisible by $(p-1)!$.

Next, we show that $J_i \neq 0$. For sufficiently large p, the number J_i can be written as

$$J_i = -\sum_{j=0}^{m} \sum_{t=0}^{r-1} \beta_{n_{t+1}} \left[f_i^{(j)}(\alpha_{n_t+1}) + \cdots + f_i^{(j)}(\alpha_{n_{t+1}}) \right].$$

Note that by construction, $f_i(x)$ can be written as a polynomial whose coefficients are polynomials in the α_i, with rational integer coefficients independent of the α_i. Thus, noting that the α_i form a complete set of conjugates and using the fundamental theorem on symmetric polynomials as in the previous proof, we see that the product of the J_i is in fact a rational number. Since it is an algebraic integer, it is an integer. Thus, $J_1 \cdots J_n$ is a rational integer, and it is divisible by $((p-1)!)^n$. Thus, $|J_1 \cdots J_n| \geq [(p-1)!]^n$.

Finally, using the triangle inequality we have, for each i,

$$|J_i| \leq \sum_{k=1}^{n} |\beta_k| |I_i(\alpha_k; f_i)|.$$

Hence, similar to the proofs of Theorems 12 and 14 we can derive that $|J| \leq c_1 c_2^p$ where c_1 and c_2 are constants that are independent of p. We get a contradiction, which completes the proof. □

The transcendence of e and π follows directly from Theorem 15. We also have the following corollaries.

Corollary 16. *If $\alpha \neq 0$ is algebraic, then e^{α} is transcendental.*

Proof: If $e^{\alpha} = \beta$ is algebraic, then we have $e^{\alpha} - \beta e^{0} = 0$, which contradicts Theorem 15. □

Corollary 17. *If $\alpha \neq 0$ is algebraic, then $\sin \alpha$ and $\cos \alpha$ are transcendental.*

Proof: We have

$$\sin \alpha = \frac{e^{i\alpha} - e^{-i\alpha}}{2i}, \quad \text{and} \quad \cos \alpha = \frac{e^{i\alpha} + e^{-i\alpha}}{2}.$$

If $\sin \alpha = \beta$ is algebraic, then $e^{i\alpha} - e^{-i\alpha} - 2i\beta e^{0} = 0$, which contradicts Theorem 15. □

Corollary 18. *If $\alpha \in \mathbb{C} \setminus \{0, 1\}$ is algebraic, then $\log \alpha$ is transcendental for every branch of the logarithm.*

Proof: If $\log \alpha = \beta$, then $e^{\beta} = \alpha$. By Corollary 16, since α is algebraic, β must be transcendental. □

4 The Gel'fond-Schneider theorem

In this section we prove the Gel'fond-Schneider theorem. We first prove some analytic lemmas. Before presenting the lemmas we introduce the following notation.

Let $w \in \mathbb{C}$, $R \in \mathbb{R}_{>0}$, and let

$$D(R, w) = \{z \in \mathbb{C} : |z - w| < R\}$$

and

$$\overline{D}(R, w) = \{z \in \mathbb{C} : |z - w| \leq R\} .$$

If $w = 0$ we write $D(R)$ and $\overline{D}(R)$. Furthermore, let the maximum of $|f(z)|$ on $\overline{D}(R, w)$ be denoted by $M(R, w, f)$. If $w = 0$ we write $M(R, f)$. If $f(z)$ is analytic on $D(R)$ and continuous on $\overline{D}(R)$, then it follows from the maximum modulus principle that $|f(z)|$ attains its maximum on $|z| = R$. If $f(z)$ is analytic on $\overline{D}(R, w)$, then $N(R, w, f)$ will be used to denote the number of zeros of $f(z)$ in $\overline{D}(R, w)$.

4.1 Some auxiliary results

Lemma 19. *Let $a_1(t), \ldots, a_n(t)$ be non-zero polynomials in $\mathbb{R}[t]$ of degrees $d_1, \ldots, d_n$ respectively. Let $w_1, \ldots, w_n$ be pairwise distinct real numbers. Then*

$$f(t) = \sum_{j=1}^{n} a_j(t) e^{w_j t}$$

has at most $n - 1 + \sum_{j=1}^{n} d_j$ real zeros.

Proof: By multiplying through by $e^{-w_n t}$ if necessary, we may suppose that $w_n = 0$ and $w_j \neq 0$ for $j \in \{1, 2, \ldots, n-1\}$. Let $E = n + \sum_{j=1}^{n} d_j$. We proceed by induction on E.

If $E = 1$, then $n = 1$ and $d_1 = 0$. In this case there are no zeros, that is, there are at most $E - 1 = 0$ zeros.

Next, suppose the lemma holds for $\ell \in \{2, 3, \ldots, E-1\}$ and consider $\ell = E$. We have the first derivative

$$f'(t) = \sum_{j=1}^{n-1} \left[a_j'(t) + w_j a_j(t)\right] e^{w_j t} + a_n'(t).$$

Since the w_j are pairwise distinct, and since $w_j \neq 0$ for $j \in \{1, 2, \ldots, n-1\}$, $a_j'(t) + w_j a_j(t)$ has exactly degree d_j for $j \in \{1, 2, \ldots, n-1\}$. Furthermore, since we may suppose that $w_n = 0$, the derivative $a_n'(t)$ has degree $d_n - 1$. It follows from the induction hypothesis that $f'(t)$ has at most $(n-2) + \sum_{j=1}^{n} d_j$ real zeros.

Finally, let N denote the number of real zeros of $f(t)$, and let $b_1 < b_2 < \ldots < b_N$ denote these zeros. Since $f(t)$ is continuous on the intervals $[b_i, b_{i+1}]$ and differentiable on (b_i, b_{i+1}) for $i \in \{1, 2, \ldots, N-1\}$, it follows from Rolle's theorem that $f'(t)$ has at least $N-1$ real zeros. Hence, $N - 1 \leq (n-2) + \sum_{j=1}^{n} d_j$, or $N \leq (n-1) + \sum_{j=1}^{n} d_j$. □

Lemma 20. *Let $r, R \in \mathbb{R}$ with $1 \leq r \leq R$. Let $f_1(z), f_2(z), \ldots, f_m(z)$ be analytic in $D(R)$ and continuous on $\overline{D}(R)$. Let $y_1, y_2, \ldots, y_m \in \mathbb{C}$ with $|y_i| \leq r$ for $i \in \{1, 2, \ldots, m\}$. Then the determinant*

$$\Delta = \det \begin{pmatrix} f_1(y_1) & \cdots & f_m(y_1) \\ \vdots & \ddots & \vdots \\ f_1(y_m) & \cdots & f_m(y_m) \end{pmatrix}$$

satisfies the inequality

$$|\Delta| \leq \left(\frac{R}{r}\right)^{-m(m-1)/2} m! \prod_{j=1}^{m} M(R, f_j).$$

Proof: Consider the determinant

$$h(z) = \det(f_j(y_i z)) = \det \begin{pmatrix} f_1(y_1 z) & \cdots & f_m(y_1 z) \\ \vdots & \ddots & \vdots \\ f_1(y_m z) & \cdots & f_m(y_m z) \end{pmatrix}.$$

Since the y_i satisfy $|y_i| \leq r$, the functions $f_j(y_i z)$ are analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$. Since it is a sum of products of the $f_j(y_i z)$, the determinant $h(z)$ itself is analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$.

Next, let $K = m(m-1)/2$. Since the $f_j(y_i z)$ are analytic functions on $D(R/r)$ they can be expanded into power series on $D(R/r)$. It follows that

$$f_j(y_i z) = \sum_{k=0}^{K-1} b_k(j) y_i^k z^k + z^K g_{ij}(z),$$

where $b_k(j) \in \mathbb{C}$ for each k and $g_{ij}(z)$ is analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$. Since the determinant is linear in each of its columns, we can view $h(z)$ as z^K times an analytic function on $D(R/r)$ plus terms involving the factor

$$z^{n_1+n_2+\cdots+n_m} \det\left(y_i^{n_j}\right) = z^{n_1+n_2+\cdots+n_m} \det \begin{pmatrix} y_1^{n_1} & \cdots & y_1^{n_m} \\ \vdots & \ddots & \vdots \\ y_m^{n_1} & \cdots & y_m^{n_m} \end{pmatrix},$$

where $n_1, n_2, \ldots, n_m \in \mathbb{Z}_{\geq 1}$ and $n_j \in \{0, 1, \ldots, K-1\}$. The determinant in the last expression is zero if two of the n_j are identical. Therefore, the non-zero terms of this form satisfy

$$n_1 + n_2 + \cdots + n_m \geq 0 + 1 + \cdots + (m-1) = \frac{m(m-1)}{2} = K.$$

Hence, we deduce that $h(z)$ is divisible by z^K.

Finally, since $h(z)$ is analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$, and since $h(z)$ is divisible by z^K, it follows that $h(z)/z^K$ is analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$. Since $h(z)/z^K$ is analytic in $D(R/r)$ and continuous on $\overline{D}(R/r)$, it follows from the maximum modulus principle that $h(z)/z^K$ attains its maximum value on the boundary $\partial\overline{D}(R/r)$. Hence, for $w \in \overline{D}(R/r)$, we have the inequality

$$\left|\frac{h(w)}{w^K}\right| \leq M\left(\frac{R}{r}, \frac{h(z)}{z^K}\right) = \left(\frac{r}{R}\right)^K M\left(R/r, h(z)\right).$$

For $|z| = R/r$ we have $|y_i z| \leq R$. The determinant of a $m \times m$ matrix is the sum of $m!$ products, where each product consists of m entries, such that for each row and column only one entry is part of a product. For each row index j we have $|f_j(y_i z)| \leq M(R, f_j)$ for $i \in \{1, 2, \ldots, m\}$. Thus,

$$M(R/r, h(z)) \leq m! \prod_{j=1}^{m} M(R, f_j).$$

Since $|\Delta| = h(1)$ and $1 \leq R/r \leq R$ we obtain

$$|\Delta| \leq \left(\frac{r}{R}\right)^K M(R/r, h(z)) \leq \left(\frac{r}{R}\right)^K m! \prod_{j=1}^{m} M(R, f_j),$$

from which the desired inequality follows. □

4.2 The case $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$

We first present a proof of the Gel'fond-Schneider theorem for $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. The proof comes from course notes by Filaseta (2011). The proof is based on the method of interpolation determinants developed by Laurent (1994).

Theorem 21. *If $\alpha, \beta \in \overline{\mathbb{Q}} \cap \mathbb{R}$ with $\alpha > 0$ and $\alpha \neq 1$, and $\beta \notin \mathbb{Q}$, then α^β is transcendental.*

An equivalent formulation of Theorem 21 is the following. Assume that $\alpha, \beta, \alpha^\beta \in \overline{\mathbb{Q}} \cap \mathbb{R}$ and $\alpha > 0$. Then $\beta \in \mathbb{Q}$.

Proof: Part of our arguments will be needed also in the proof of the general Gel'fond-Schneider theorem (Theorem 25), where the condition $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$ is not needed. It is only when we apply Lemma 19 above that we have to assume $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$. For the moment we assume $\alpha, \beta, \alpha^\beta \in \overline{\mathbb{Q}}$ with $\alpha \neq 0, 1$, where $\alpha^\beta = e^{\beta \log \alpha}$ is any choice of the branch of the logarithm. When we are at the point to apply Lemma 19, we use the assumption $\alpha, \beta, \alpha^\beta \in \mathbb{R}$ and deduce that $\beta \in \mathbb{Q}$.

Let $L_0, L_1, S \in \mathbb{Z}_{\geq 2}$, $L = (L_0+1)(L_1+1)$, and $K \in \mathbb{R}$. K, L_0, L_1 and L will be increasing functions of S, and repeatedly we will choose S sufficiently large so that K is sufficiently large. The functions K, L_0 and L_1 and S must be chosen such that the inequalities

$$KL_0 \log S \leq L, \qquad KL_1 S \leq L \quad \text{and} \quad L \leq (2S-1)^2$$

hold. This can be done by taking, for example, S large, $L_0 = \lfloor S \log S \rfloor$, $L_1 = \lfloor S/\log S \rfloor$ and $K = \log\log S$. If we combine the first two inequalities we obtain

$$KL\,(L_0 \log S + L_1 S) \leq 2L^2. \tag{16}$$

Next, consider some arrangement

$$(s_1(i), s_2(i))_{i=1}^{(2S-1)^2}$$

of the pairs $(s_1, s_2) \in \{0, \ldots, S-1\} \times \{0, \ldots, S-1\}$. Furthermore, let

$$(u(j), v(j))_{j=1}^{L}$$

be an arrangement of the pairs $(u, v) \in \{0, \ldots, L_0\} \times \{0, \ldots, L_1\}$. We define the $(2S-1)^2 \times L$ matrix

$$\mathcal{M} = \left((s_1(i) + s_2(i)\beta)^{u(j)} \left(\alpha^{s_1(i)+s_2(i)\beta} \right)^{v(j)} \right).$$

Let

$$f_j(z) = z^{u(j)} \alpha^{v(j)z} = z^{u(j)} e^{v(j)z \log \alpha}$$

for $j \in \{1, 2, \ldots, L\}$ be functions in the complex variable z, let

$$y_i = s_1(i) + s_2(i)\beta$$

for $i \in \{1, 2, \ldots, L\}$, and consider the determinant $\Delta = \det(f_j(y_i))$ of an arbitrary $L \times L$ submatrix $(f_j(y_i))$. We will show that all $L \times L$ submatrices of $\mathcal{M}$ have determinant zero. Under the assumption that $\Delta \neq 0$ we derive an analytic upper bound and an algebraic lower bound for $\log|\Delta|$, from which we derive a contradiction.

We first derive an upper bound. We have $\alpha^{v(j)z} = \exp(v(j)z\log\alpha)$, and $f_j(z)$ represents an entire function for each $j \in \{1, 2, \ldots, L\}$. For $z_1, z_2 \in \mathbb{C}$ we have

$$|e^{z_1 z_2}| = e^{\mathrm{Re}(z_1 z_2)} \leq e^{|z_1 z_2|} = e^{|z_1||z_2|}.$$

Hence, for any $R \in \mathbb{R}_{>0}$, we have

$$M(R, f_j) = M(R, z^{u(j)} e^{v(j)z\log\alpha}) \leq R^{u(j)} e^{v(j)R|\log\alpha|}. \tag{17}$$

Taking the log on both sides of (17) we obtain the inequality

$$\log M(R, f_j) \leq u(j)\log R + v(j)R|\log\alpha| \leq L_0 \log R + L_1 R|\log\alpha|. \tag{18}$$

Next, applying Lemma 20 to Δ with $r = S(1 + |\beta|)$ and $R = e^2 r$ we obtain the inequality

$$|\Delta| \leq e^{-L(L-1)} L! \prod_{j=1}^{L} M(R, f_j). \tag{19}$$

Taking the log on both sides of (19), and using inequality (18), we obtain

$$\begin{aligned} \log |\Delta| &\leq -L(L-1) + \log L! + \sum_{j=1}^{L} \log M(R, f_j) \\ &\leq -L^2 + L + L \log L + L \max_{1 \leq j \leq L} \{\log M(R, f_j)\} \\ &\leq -L^2 + L\left(1 + \log L + L_0 \log R + L_1 R |\log \alpha|\right) \end{aligned}$$

or

$$\log |\Delta| \leq -L^2 + c_1 L \left(L_0 \log S + L_1 S\right) \tag{20}$$

for some absolute constant $c_1 \in \mathbb{R}$ independent of S. If we choose S such that $K \geq 4c_1$, inequality (16) becomes

$$c_1 L \left(L_0 \log S + L_1 S\right) \leq \frac{L^2}{2}.$$

Combining this inequality with (20) we obtain

$$\log |\Delta| \leq -\frac{L^2}{2}, \tag{21}$$

which specifies an upper bound for $\log |\Delta|$.

Next, we derive a lower bound for $\log |\Delta|$ under the assumption that $\Delta \neq 0$. Fix $T \in \mathbb{Z}_{>0}$ such that $T\alpha$, $T\beta$ and $T\alpha^{\beta}$ are algebraic integers. Then $T^{L_0+2L_1S}$ has the property that $T^{L_0+2L_1S}$ times any element of $\mathcal{M}$, and hence $T^{L_0+2L_1S}$ times any element of the matrix describing Δ is an algebraic integer. Hence, $T^{L(L_0+2L_1S)}\Delta$ is an algebraic integer in $\mathbb{Q}\left(\alpha, \beta, \alpha^{\beta}\right)$. Thus $T^{L(L_0+2L_1S)}\Delta$ is a zero of a monic polynomial of degree N, where N is at most the product of the degrees of the minimal polynomials of α, β and α^{β}.

The house $\overline{|\Delta|}$ is the maximum of the absolute values of Δ and its conjugates. We have the upper bound

$$\overline{|\Delta|} \leq L! S^{L_0 L} \left(1 + \overline{|\beta|}\right)^{L_0 L} \left(1 + \overline{|\alpha|}\right)^{L_1 L S} \left(1 + \overline{|\alpha^{\beta}|}\right)^{L_1 L S}.$$

In the latter inequality we have taken into consideration that $\overline{|\beta|}$, $\overline{|\alpha|}$ and $\overline{|\alpha^{\beta}|}$ may be smaller than 1. If $\Delta \neq 0$, it follows from Lemma 13 that

$$|\Delta| \geq T^{-NL(L_0+2L_1S)} \overline{|\Delta|}^{1-N} \geq T^{-NL(L_0+2L_1S)} \overline{|\Delta|}^{-N}.$$

Combining these two inequalities we obtain

$$|\Delta| \geq T^{-NL(L_0+2L_1S)} (L!)^{-N} S^{-NL_0L} \left(1 + \overline{|\beta|}\right)^{-NL_0L} \left(1 + \overline{|\alpha|}\right)^{-NL_1LS} \left(1 + \overline{|\alpha^{\beta}|}\right)^{-NL_1LS}.$$

Since $N \log L! \leq NL \log L$, we obtain, after taking the log on both sides, the inequality

$$\log|\Delta| \geq -NL(L_0 + 2L_1 S)\log T - NL\log L - NL_0 L \log S \\ - NL_0 L \log\left(1 + \lceil\beta\rceil\right) - NL_1 LS \log\left(1 + \lceil\alpha\rceil\right) - NL_1 LS \log\left(1 + \lceil\alpha^\beta\rceil\right).$$

Since N, T, $\log\left(1+\lceil\beta\rceil\right)$, $\log\left(1+\lceil\alpha\rceil\right)$ and $\log\left(1+\lceil\alpha^\beta\rceil\right)$ are constants that only depend on α and β, there is an absolute constant $c_2 \in \mathbb{R}$ independent of S for which

$$\log|\Delta| \geq -c_2 L\left(L_0 + \log L + L_0 \log S + L_1 S\right).$$

If we choose S sufficiently large we obtain the inequality

$$\log|\Delta| \geq -c_3 L(L_0 \log S + L_1 S), \tag{22}$$

for some absolute constant $c_3 \in \mathbb{R}$ independent of c. Furthermore, if we choose S such that $K \geq 6c_3$, inequality (16) becomes

$$c_3 L\left(L_0 \log S + L_1 S\right) \leq \frac{L^2}{3}.$$

Combining this inequality with (22) we obtain

$$\log|\Delta| \geq -\frac{L^2}{3}, \tag{23}$$

which specifies a lower bound for $\log|\Delta|$. We now get a contradiction between the upper bound in (21) and the lower bound in (23). Since Δ was an arbitrary submatrix, this shows indeed that all $L \times L$ submatrices of $\mathcal{M}$ have determinant zero.

Since $\Delta = \det(f_j(y_i)) = 0$ for any sub-determinant Δ, it follows that the columns of the matrix $(f_j(y_i))$ are linearly dependent over $\mathbb{R}$. Hence, there exists $b_1, b_2, \ldots, b_L \in \mathbb{R}$, not all 0, such that

$$\sum_{j=1}^{L} b_j f_j(y_i) = 0, \quad \text{for } i \in \left\{1, 2, \ldots, (2S-1)^2\right\}. \tag{24}$$

Since $f_j(y_i) = y_i^{u(j)}\alpha^{v(j)y_i}$, identity (24) is equal to

$$\sum_{j=1}^{L} b_j y_i^{u(j)}\alpha^{v(j)y_i} = 0, \quad \text{for } i \in \left\{1, 2, \ldots, (2S-1)^2\right\}. \tag{25}$$

If we consider identity (25) for all pairs (u, v) with $u \in \{0, 1, \ldots, L_0\}$ and $v \in \{0, 1, \ldots, L_1\}$, we obtain

$$\sum_{v=0}^{L_1}\left(\sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} y_i^u\right)\alpha^{vy_i} = 0, \quad \text{for } i \in \left\{1, 2, \ldots, (2S-1)^2\right\}. \tag{26}$$

Choosing

$$a_v(t) = \sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} t^u, \quad w_v = v \log \alpha, \quad \text{and} \quad t = y_i = s_1(i) + s_2(i)\beta,$$

we can write the left-hand side of (26) as

$$\sum_{v=0}^{L_1} \left(\sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} y_i^u \right) \alpha^{v y_i} = \sum_{v=0}^{L_1} a_v(t) e^{w_v t}.$$

Each of the L values of y_i is a zero of $\sum_{v=0}^{L_1} a_v(t) e^{w_v t}$. Note that this sum consists of $L_1 + 1$ polynomials, each of degree L_0. We now at last use our assumption $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$, and apply Lemma 19. By that lemma, there are at most

$$L_0(L_1 + 1) + (L_1 + 1) - 1 = L - 1$$

distinct zeros. Since $L - 1 < L \leq (2S - 1)^2$, two of the y_i must be the same, and we have

$$s_1(i) + s_2(i)\beta = s_1(i') + s_2(i')\beta \quad \text{for some } i, i' \text{ with } 1 \leq i < i' \leq (2S-1)^2.$$

However, since the pairs $(s_1(i), s_2(i))$ and $(s_1(i'), s_2(i'))$ are distinct, it follows that

$$\beta = \frac{s_1(i') - s_1(i)}{s_2(i) - s_2(i')} \in \mathbb{Q}.$$

□

4.3 The general case

In this subsection we consider the Gel'fond-Schneider theorem for the complex case. In the proof for the real case in the previous subsection, only in Lemma 19 we used the assumptions $\alpha, \beta \in \mathbb{R}$ and $\alpha > 0$. The idea is to replace Lemma 19 by Proposition 22 for complex numbers. The following result comes from Tijdeman (1971).

Proposition 22. *Let $a_1(z), \ldots, a_n(z)$ be non-zero polynomials in $\mathbb{C}[z]$ of degrees $d_1, \ldots, d_n$ respectively, let $w_1, \ldots, w_n$ be pairwise distinct complex numbers, let*

$$f(z) = \sum_{k=1}^{n} a_k(z) e^{w_k z},$$

and put

$$E = n + \sum_{k=1}^{n} d_k, \quad \text{and} \quad m = \max_k |w_k|.$$

Furthermore, let $R, s, t \in \mathbb{R}_{>0}$, $s > 1$, and let $y \in \mathbb{C}$. Then

$$N(R, y, f) \leq \frac{1}{\log s}\left((E-1)\log\frac{st+s+t}{t} + (st+s+2t)Rm + \frac{1}{s}\right). \quad (27)$$

We first prove the following result of Tijdeman (1971, Lemma 1).

Lemma 23. *Let $R, s, t \in \mathbb{R}_{>0}$, $s > 1$, and let $f \neq 0$ be analytic on $\overline{D}((st+s+t)R)$. Then*

$$N(R, f) \leq \frac{1}{\log s}\log\frac{M((st+s+t)R, f)}{M(tR, f)}.$$

Proof: Let $w \in \overline{D}(tR)$ such that $|f(w)| = M(tR, f)$. It then follows that

$$\overline{D}(R) \subset \overline{D}((1+t)R, w) \quad (28)$$

and

$$\overline{D}((st+s)R, w) \subset \overline{D}((st+s+t)R). \quad (29)$$

By Jensen's formula (Greene, Krantz 2006, p. 279) we have

$$\int_0^{sR}\frac{N(r, w, f)}{r}dr = \frac{1}{2\pi}\log\left|f\left(w + sRe^{i\theta}\right)\right| d\theta - \log|f(w)|.$$

We also have

$$\int_0^{sR}\frac{N(r, w, f)}{r}dr \geq \int_R^{sR}\frac{N(R, w, f)}{r}dr = N(R, w, f)\log s.$$

Combining the two previous formulas we obtain

$$N(R, w, f) \leq \frac{1}{\log s}M\left(sR, w, \log\frac{|f|}{|f(w)|}\right).$$

This inequality, together with the inclusions (28) and (29), implies that

$$\begin{aligned} N(R, f) \leq N((1+t)R, w, f) &\leq \frac{1}{\log s}M\left(sR(1+t), w, \log\frac{|f|}{|f(w)|}\right) \\ &\leq \frac{1}{\log s}\log\frac{M((st+s+t)R, f)}{M(tR, f)}. \end{aligned}$$

□

Lemma 23 is used in the proof of Proposition 22. The following result from Balkema and Tijdeman (1973, Theorem 2) is also used in the proof of Proposition 22.

Lemma 24. *Let $a_1(z), \ldots, a_n(z)$ be non-zero polynomials in $\mathbb{C}[z]$ of degrees $d_1, \ldots, d_n$ respectively, let $w_1, \ldots, w_n$ be pairwise distinct complex numbers, let*

$$f(z) = \sum_{k=1}^{n} a_k(z) e^{w_k z},$$

and put

$$E = n + \sum_{k=1}^{n} d_k, \quad \text{and} \quad m = \max_k |w_k|.$$

Furthermore, let $R, \gamma \in \mathbb{R}_{>0}$, $\gamma > 1$. Then

$$M(\gamma R, f) \leq \frac{\gamma^E - 1}{\gamma - 1}\, e^{Rm(\gamma+1)} M(R, f).$$

We are now ready to present the proof of Proposition 22.

Proof of Proposition 22: Let $\gamma \in \mathbb{R}_{>1}$. Using Lemma 24 we obtain the inequality

$$M(\gamma t R, f) \leq \frac{\gamma^E - 1}{\gamma - 1}\, e^{tRm(\gamma+1)} M(tR, f).$$

Taking $\gamma = (st + s + t)/t$ we have

$$\frac{\gamma^E - 1}{\gamma - 1} \leq \frac{t}{st + s} \left(\frac{st + s + t}{t} \right)^E \leq \left(1 + \frac{1}{s} \right) \left(\frac{st + s + t}{t} \right)^{E-1}.$$

Combining the previous two inequalities we obtain

$$M((st + s + t)R, f) \leq \left(1 + \frac{1}{s} \right) \left(\frac{st + s + t}{t} \right)^{E-1} e^{(st+s+2t)Rm} M(tR, f).$$

Combining this inequality with Lemma 23 we obtain the desired inequality. □

In the complex case, the Gel'fond-Schneider theorem holds for every branch of the complex logarithm. By replacing Lemma 19 with Proposition 22 in the proof of Theorem 21 we obtain the following result.

Theorem 25. *Let $\alpha, \beta \in \overline{\mathbb{Q}}$ and $\alpha \neq 0, 1$, and $\beta \notin \mathbb{Q}$. For any branch of $\log z$ we have that $\alpha^\beta = e^{\beta \log \alpha}$ is transcendental.*

Proof: Following the same set up and arguments as in the proof of Theorem 21, we find at the end that each of the $(2S-1)^2$ values of y_i is a zero of $f(z) = \sum_{v=0}^{L_1} a_v(z) e^{w_v z}$. Let E, R and m be as defined in Proposition 22, and let 0 be the center of the disc $\overline{D}(R)$. Taking $s = 5$ and $t = \frac{1}{5}$ in inequality (27), and using that

$$\frac{\log 31}{\log 5} < 2.2 \quad \text{and} \quad \frac{32}{5 \log 5} < 3.9,$$

we obtain

$$N(R, f) \leq 3(E - 1) + 4Rm. \tag{30}$$

The sum $f(z)$ consists of $L_1 + 1$ polynomials, each of degree L_0. Hence,

$$E = L_1 + 1 + L_0(L_1 + 1) = L.$$

The complex numbers are of the form $y_i = s_1(i) + s_2(i)\beta$ where $s_1, s_2 \in \mathbb{Z}$ with $|s_1|, |s_2| < S$. Hence, since we consider the disc with center 0, we have the upper bound

$$R \leq S\left(1 + \lceil \beta \rceil\right).$$

Finally, we have

$$m = \max_{v \in \{0,\ldots,L_1\}} |w_v| = \max_{v \in \{0,\ldots,L_1\}} |v \log \alpha| = L_1 |\log \alpha|.$$

Using the values of E and m and the upper bound for R in inequality (30) we obtain that the number of zeros of f satisfy

$$N(f) \leq 3(L - 1) + 4SL_1\left(1 + \lceil \beta \rceil\right)|\log \alpha|.$$

Using here the specific definitions $L_0 = \lfloor S \log S \rfloor$ and $L_1 = \lfloor S / \log S \rfloor$, we obtain for sufficiently large S that

$$\begin{aligned} N(f) &\leq 3\left(S^2 + S \log S + \frac{S}{\log S}\right) + \frac{4S^2}{\log S}\left(1 + \lceil \beta \rceil\right)|\log \alpha| \\ &< 4S^2 - 4S + 1 = (2S - 1)^2. \end{aligned}$$

Hence, at least two of the y_i must be the same, which completes the proof. $\square$

References

Abramowitz M, Stegun IA (1970) *Handbook of Mathematical Functions (with Formulas, Graphs and Mathematical Tables)*. Dover Publications, New York.

Apéry R (1979) Irrationalité de $\zeta(2)$ et $\zeta(3)$. *Astérisque, 61,* 11-13.

Baker A (1975) *Transcendental Number Theory.* Cambridge University Press, Cambridge.

Balkema AA, Tijdeman R (1973) Some estimates in the theory of exponential sums. *Acta Mathematica Academiae Scientiarum Hungaricae, 24,* 115-133.

Beukers F (1979) A note on the irrationality of $\zeta(2)$ and $\zeta(3)$. *Bulletin of the London Mathematical Society, 11,* 268-272.

Burger EB, Tubbs R (2004) *Making Transcendence Transparent.* Springer, New York.

Cantor GFLP (1874) Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. *Journal für die reine und angewandte Mathematik, 77,* 258-262.

Dunham W (1990) *Journey through Genius: The Great Theorems of Mathematics.* John Wiley & Sons.

Filaseta M (2011) *Transcendental Number Theory.* Course notes, University of South Carolina.

Greene RE, Krantz SG (2006) Function Theory of One Complex Variable. American Mathematical Society, Rhode Island.

Laurent M (1994) Linear forms in two logarithms and interpolation determinants. *Acta Arithmetica, 66,* 181-199.

Niven I (1947) A simple proof that π is irrational. *Bulletin of the American Mathematical Society, 53,* 509.

Parks AE (1986) Pi, e and other irrational numbers. *The American Mathematical Monthly, 93,* 722-723.

Rivoal T (2000) La fonction zeta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs. *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique, 331,* 267270.

Shidlovskii AB (1989) *Transcendental Numbers.* Walter de Gruijter, Berlin.

Tijdeman R (1971) On the number of zeros of general exponential polynomials. *Indagationes Mathematicae, 74,* 1-7.

Zudilin W (2001) One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational. *Russian Mathematical Surveys, 56,* 774-776.