# 6. Mahler and Transcendence

## Effective Constructions in Transcendental Number Theory

Yuri Nesterenko

We discuss Mahler results connected to effective constructions in transcendental number theory, and try to explain the influence of Mahler ideas on the current state of this theory.

## 1 Introduction

Mahler's activity connected to transcendental numbers is very impressive. Several directions of this theory have roots in his works. In this chapter, we will discuss the results of the theory of transcendental numbers, proved with the help of effective methods of constructing Diophantine approximations to the numbers under study, and, of course, having the most direct relation to K. Mahler. We practically do not consider theorems that were obtained with the help of constructions based on the so-called "Siegel Lemma". This includes many of the fundamental results of C. L. Siegel, A. O. Gelfond, Th. Schneider, Mahler himself, A. Baker and their followers. But there are, of course, exceptions, their appearance is always motivated.

The article does not give definitions of the basic concepts of the theory of transcendental numbers, such as algebraic independence, measure of transcendence, irrationality exponent, $E$-functions and other. If necessary, we recommend using the book [4].

## 2 Exponential function

### 2.1 Hermite's identities

The exponential function is the simplest of transcendental functions. In addition, it has a number of important properties. For example, it satisfies a simple differential equation $y' = y$ and satisfies the addition theorem $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$. Apparently, this explains why it was the first function whose values began to be investigated from the arithmetic point of view. In 1873, Hermite proved the transcendence of the number $e$, and in 1882, Lindemann established that the value of the function $e^z$ will also be transcendental at any algebraic point different from zero (Lindemann's theorem). In particular, this proves the transcendence of the number $\pi$. Lindemann stated without proof an even stronger result, and indicated that it can be proved using the same ideas (see the end of [15]): *If $\alpha_0, \alpha_1, \ldots, \alpha_m$ are distinct algebraic numbers, then the numbers*

$$e^{\alpha_0}, e^{\alpha_1}, \ldots, e^{\alpha_m}$$

*are linearly independent over the field of algebraic numbers.*

In 1885, this theorem was proved by Weierstrass [27], who also simplified the proof of Lindemann's result. It is now customary to call this result the Lindemann–Weierstrass Theorem.

An important tool connecting the proofs of these results is Hermite's identity,

$$F(a,z)e^{-az} - F(b,z)e^{-bz} = z^M \int_a^b e^{-zx} f(x)dx, \tag{1}$$

where $f(x) \in \mathbb{C}[x]$, $M = \deg f(x)$, and

$$F(x,z) = \sum_{k=0}^{M} f^{(k)}(x) z^{M-k}.$$

With the correct choice of the polynomial $f(x)$ and points $a, b$, Hermite's identity gives the Hermite–Padé approximations of a set of exponential functions,

$$e^{\alpha_k z} F(0,z) - F(\alpha_k, z) = z^M e^{\alpha_k z} \int_0^{\alpha_k} e^{-zx} f(x)dx, \qquad 0 \le k \le m,$$

and good simultaneous approximations by algebraic numbers to the values $e^{\alpha_k}$ when we take $z = 1$.

Transcendence of $e$ means that for any polynomial $P(z) \in \mathbb{Z}[z], P \neq 0$, we have $|P(e)| > 0$. The following result gives a quantitative improvement of this property. Let be $m = \deg P$ and $H(P)$ be the maximum of absolute values of all coefficients of $P$. In 1932, Mahler [M11] used the identity (1) to prove

Theorem 1. *There exists an absolute constant $c > 0$ such that for any $P(z) \in \mathbb{Z}[x]$, $P \neq 0$, for which $H = H(P)$ is greater than a certain bound depending on $m = \deg P$, one has*

$$|P(e)| \ge H^{-m - \frac{cm^2 \log(m+1)}{\log\log m}}.$$

The first result of this kind was proved in 1899 by E. Borel. To this day, Theorem 1 still gives the best result when $H(P)$ is sufficiently large with respect to $\deg P$.

In 1893, Hermite stated another identity involving the exponential function (see [9], Vol.4, p. 357-377), which in some sense gives a construction of approximations that is dual to (1). Let $\alpha_1, \ldots, \alpha_m$ be different complex numbers, $n_1, \ldots, n_m$ be non-negative integers and set

$$N+1=\sum_{k=1}^{m}(n_k+1) \quad \text{and} \quad Q(x)=\prod_{k=1}^{m}(x-\alpha_k)^{n_k+1}. \tag{2}$$

Then, as Hermite first proved, the identity

$$\frac{1}{2\pi i}\int_C \frac{f(\zeta)}{Q(\zeta)}d\zeta=\sum_{k=1}^{m}\sum_{j=0}^{n_k}\frac{f^{(n_k-j)}(\alpha_k)}{(n_k-j)!j!}\left(\frac{d}{dx}\right)^j\left(\frac{(x-\alpha_k)^{n_k+1}}{Q(x)}\right)\Big|_{x=\alpha_k}, \tag{3}$$

is valid, where $C$ is a circle containing all points $\alpha_1, \ldots, \alpha_m$ inside, where $f(z)$ is a function that is analytic in a domain containing the circle $C$ together with its interior. The right-hand side of the identity (3) is the sum of the residues of the function under the integral at points $\alpha_1, \ldots, \alpha_m$. Substituting $f(\zeta)=e^{z\zeta}$ in the identity (3) for an arbitrary fixed complex $z$, we obtain Hermite's second identity for the exponential function,

$$R(z)=\frac{1}{2\pi i}\int_C \frac{e^{z\zeta}d\zeta}{(\zeta-\alpha_1)^{n_1+1}\cdots(\zeta-\alpha_m)^{n_m+1}}=\sum_{k=1}^{m}P_k(z)e^{\alpha_k z}, \tag{4}$$

where for any $k$ with $0 \le k \le m$, the coefficient $A_k(z)$ is a polynomial in $z$ of degree $n_k$.

In 1967, comparing the two identities of Hermite, Mahler [M164, p. 200] wrote about the second of them,

> "On putting again $z=1$, one obtains now a linear form $a_0+a_1e+\cdots+a_me^m$ of small absolute value and with small integral coefficients, from which again the transcendency of $e$ may be deduced. Surprisingly, Hermite himself never took this step, and I was seemingly the first[1] to use the polynomials $P_k(z)$ for this purpose".[2]

## 2.2 Interpolation series

In 1929, Gelfond studied integrals

$$I_n=\frac{1}{2\pi i}\int_{C_n}\frac{e^{\pi i z}dz}{(z-z_0)(z-z_1)\cdots(z-z_n)}, \qquad n \ge 0, \tag{5}$$

where $z_0, z_1, \ldots,$ are all Gaussian integer numbers $a+bi$, $a, b \in \mathbb{Z}$, ordered by the conditions

[1] In the article [M11].

[2] In (4) we use notation $P_k(z)$ instead of $A_k(z)$.

1) $|z_k| \le |z_{k+1}|$ and

2) $\arg z_k < \arg z_{k+1}$ in case $|z_k| = |z_{k+1}|$.

The path of integration $C_n$ is a circle containing inside all the points $z_k$ with $0 \le k \le n$. Calculation of $I_n$ as a sum of residues at the points $z_k$ proves that $I_n = P_n(e^\pi, e^{-\pi}, i)$, where $P_n(x_1, x_2, x_3)$ is a polynomial with rational coefficients. It can be shown that the integrals $I_n$ tend to zero with increasing $n$ so rapidly that the assumption of algebraicity of $e^\pi$ leads, with the help of Liouville's theorem, to the conclusion $P_n(e^\pi, e^{-\pi}, i) = 0$ for all sufficiently large $n$. On the other hand, it is not difficult to prove, that the function $e^{\pi i z}$ decomposes into an interpolation series converging in the whole complex plane

$$e^{\pi i z} = I_0 + \sum_{n=0}^{\infty} I_n \cdot (z - z_0)(z - z_1) \cdots (z - z_n), \tag{6}$$

and therefore $e^{\pi i z}$ is a polynomial, which is certainly false. The resulting contradiction proves the transcendence of $e^\pi$—a particular case of Hilbert's seventh problem. In 1930, Kuzmin [12] and Siegel (unpublished[3]) extended Gelfond's proof to the case of real quadratic fields.

In 1930, Gelfond [7] used another sequence $z_0, z_1, \dots$ to prove Lindemann's theorem. Let $\alpha$ be a nonzero algebraic number, let $p > 1$ be an integer with $p \log 2 \ge |\alpha|$ and set $z_m = \lfloor \frac{m}{p} \rfloor$, for any integer $m \ge 0$. In the case $m = pn + q$ with $0 \le q < p$, we have $z_m = n$ and

$$(z - z_0)(z - z_1) \cdots (z - z_m) = (z(z-1) \cdots (z - n + 1))^p (z - n)^q.$$

One can prove that for any $p > \frac{|\alpha|}{\log 2}$ the function $e^{\alpha z}$ decomposes into the interpolation series, converging in the whole complex plane,

$$e^{\alpha z} = A_0 + \sum_{m=0}^{\infty} A_m \cdot (z - z_0)(z - z_1) \cdots (z - z_m), \tag{7}$$

where

$$A_m = \frac{1}{2\pi i} \int_{C_m} \frac{e^{\alpha z} dz}{(z(z-1) \cdots (z - n + 1))^p (z - n)^q}, \qquad m = pn + q, \qquad 0 \le q < p. \tag{8}$$

Since the function $e^{\alpha z}$ is not a polynomial, one can claim that there are infinitely many indexes $m$ such that $A_m \neq 0$. Calculating residues at points $0, 1, \dots, n$, it is easy to verify that the integral $A_m$ is a polynomial in $\alpha$ and $e^\alpha$ with rational coefficients; moreover, $p!((n+1)!)^p d_n^p A_m = P_m(\alpha, e^\alpha)$. Here $d_n = \operatorname{lcm}(1, 2, \dots, n)$ and $P_m(x, y) \in \mathbb{Z}[x, y]$. Assume that $e^\alpha$ is an algebraic

[3] See reference [19] of the article [6].

number. Applying Liouville's Theorem to algebraic numbers $\alpha$, $e^{\alpha}$ and the polynomial $P_m(x,y)$ with sufficiently large $m$ such that $A_m \neq 0$, we derive

$$p!((n+1)!)^p d_n^p |A_m| \geq e^{-c_1 pn}, \tag{9}$$

where $c_1$ is a positive constant depending only on $\alpha$ and $e^{\alpha}$. For the other side, one can estimate the integral in the representation (8) and prove that

$$p!((n+1)!)^p d_n^p |A_m| \leq e^{-pn(\log p - \log|\alpha| - 3)}. \tag{10}$$

Inequalities (9) and (10) contradict one another if $p$ is a fixed integer larger than a constant depending only on $\alpha$ and $n$ is sufficiently large with respect to $\alpha$ and $p$. This contradiction completes the proof of the Lindemann theorem proposed by Gelfond.

### 2.3 Lindemann–Weierstrass theorem

Interpolation series in the above proofs are only a means to establish the existence of an infinite sequence of nonzero integrals of a certain type. In proving the quantitative strengthening of the Lindemann–Weierstrass theorem, Mahler [M11] used the second Hermite identity and ideas of Siegel proposed to prove the algebraic independence of the values of $E$-functions, instead of the interpolation series. Let's go back to the identity (4). After easy calculations we derive

$$P_k(z) = \sum_{j=0}^{n_k} a_{kj} \frac{z^{n_k - j}}{(n_k - j)!} \tag{11}$$

and

$$a_{kj} = \frac{1}{j!}\left(\frac{d}{dx}\right)^j \left(\frac{(x-\alpha_k)^{n_k+1}}{Q(x)}\right)\Big|_{x=\alpha_k} = \\ (-1)^j \sum_{\mathbf{s}} \prod_{\substack{i=1 \\ i\neq k}}^{m} \binom{n_i + s_i}{n_i} (\alpha_k - \alpha_i)^{-1-n_i-s_i}, \tag{12}$$

where the summation is taken over all sets of nonnegative integers $s_1, \ldots, s_{k-1}$, $s_{k+1}, \ldots, s_m$ whose sum is equal to $j$. Besides

$$R(z) = \frac{z^N}{N!} + \ldots, \quad |R(z)| < \frac{|z|^N}{N!} e^{a|z|}, \quad \text{where } a = \max(|\alpha_1|, \ldots, |\alpha_m|).$$

If $\alpha_1, \ldots, \alpha_m$ are algebraic numbers and $\mathbb{K} = \mathbb{Q}(\alpha_1, \ldots, \alpha_m)$, then $R(1)$ is a linear form in the numbers $e^{\alpha_k}$ with coefficients in $\mathbb{K}$. If the integer number $q > 0$ is such that $q(\alpha_i - \alpha_j)^{-1} \in \mathbb{Z}_{\mathbb{K}},\ 1 \leq i < j \leq m$, then

$$n_k! q^N P_k(1) \in \mathbb{Z}_{\mathbb{K}}. \tag{13}$$

In addition, if $M = \max_{1\le i<j\le m} \lceil \alpha_i - \alpha_j \rceil^{-1}$, then

$$\lceil P_k(1) \rceil \le (2M+1)^N. \tag{14}$$

Such linear forms can be used to prove arithmetic results about the values of the exponential function.

Mahler [M11] first applied the identity (4) to prove a quantitative version of the Lindemann–Weierstrass Theorem and simultaneously to give another proof of this theorem.

Theorem 2. *Let $\theta_1, \dots, \theta_r$ be algebraic numbers that are linearly independent over $\mathbb{Q}$. There exists a positive constant $c$ depending only on the $\theta_i$ such that for any polynomial $P \in \mathbb{Z}[x_1, \dots, x_r]$, $M_i = \deg_{x_i} P \ge 1$, one has*

$$|P(e^{\theta_1}, \dots, e^{\theta_r})| > H^{-cM_1\cdots M_r},$$

*for any real $H$ that exceeds all absolute values of coefficients of $P$ and a certain bound that depends on $\theta_1, \dots, \theta_r$ and also on $M_1, ..., M_r$.*

The elimination part of the proof of this theorem makes use of a quantitative version of Siegel's argument in $E$-functions theory (1929). The inequality $|P(e^{\theta_1}, \dots, e^{\theta_r})| > 0$ for any polynomial $P \in \mathbb{Z}[x_1, \dots, x_r]$, $P \neq 0$, implies the algebraic independence of $e^{\theta_1}, \dots, e^{\theta_r}$ over the field of all algebraic numbers. This statement is equivalent to the formulation of the Lindemann–Weierstrass Theorem from the beginning of the chapter. Now we discuss main ideas of Mahler's proof of Theorem 2.

Denote by $\nu$ the degree of the field $\mathbb{Q}(\theta_1, \dots, \theta_r)$ and

$$\mu_k = \left[ \frac{M_k}{(1+\frac{1}{2\nu-1})^{1/r} - 1} \right], \qquad k = 1, 2, \dots, r.$$

Denote

$$\mu = (\mu_1 + 1)\cdots(\mu_r + 1), \qquad m = (M_1 + \mu_1 + 1)\cdots(M_r + \mu_r + 1),$$

and

$$\{\alpha_1, \dots, \alpha_m\} = \{\lambda_1\theta_1 + \dots + \lambda_r\theta_r : 0 \le \lambda_1 \le M_1 + \mu_1, \dots, 0 \le \lambda_r \le M_r + \mu_r\},$$

ordered in a certain way. Here $\lambda_1, \dots, \lambda_r$ are integers. All linear combinations from the right-hand side of the last equality are distinct since $\theta_1, \dots, \theta_r$ are linearly independent over $\mathbb{Q}$. It is easy to check that

$$m < \frac{2\nu}{2\nu-1}\mu, \quad n\mu - (n-1)m \ge \mu\frac{\nu}{2\nu-1}, \quad \mu \le (4r\nu)^r M_1 \dots M_r. \tag{15}$$

The linear forms that take small values at $(e^{\alpha_1}, \dots, e^{\alpha_m})$ are constructed using the identity (4), which gives simultaneous functional approximations for

$e^{\alpha_1 z}, \dots, e^{\alpha_m z}$. In order to get a full system of such forms, Mahler used a technique that goes back to Hermite (see [9], Vol. 4, p. 368). For large $n \in \mathbb{N}$ and for each $i = 1, \dots, m$, we choose $n_i$ in (4) as follows:

$$n_k = n - 1, \quad \text{if} \quad k \neq i, \qquad n_i = n.$$

Let $R_i(z)$ and $P_{i,k}(z)$ denote the function $R(z)$ and the polynomial $P_k(z)$ corresponding to this choice. Then

$$R_i(z) = \sum_{k=1}^{m} P_{i,k}(z) e^{\alpha_k z}, \qquad i = 1, \dots, m.$$

For any $i$ we have $N = mn$. It is easy to compute the determinant

$$\Delta(z) = \det \|P_{i,k}(z)\|_{1 \le i,k \le m}.$$

Namely, from the construction of $P_{i,k}(z)$ it follows that

$$\Delta(z) = Az^{mn} + Bz^{mn-1} + \dots, \qquad A \neq 0.$$

On the other hand, since $\mathrm{ord}_{z=0} R_i(z) \ge N = mn$, it easily follows that $\mathrm{ord}_{z=0} \Delta(z) \ge mn$. Thus, $\Delta(z) = Az^{mn}$, and $\Delta(1) = A \neq 0$.

Set

$$L_i(\overline{x}) = n! q^N \sum_{k=1}^{m} P_{i,k}(1) x_k, \qquad i = 1, ..., m,$$

where $q \in \mathbb{Z}, q > 0$, is such that $q(\alpha_i - \alpha_j)^{-1} \in \mathbb{Z}_{\mathbb{K}}$, $\mathbb{K} = \mathbb{Q}(\theta_1, ..., \theta_r)$, $1 \le i < j \le m$. It follows from above that these linear forms $L_i(\overline{x})$ are linearly independent. It can be shown that the linear forms $L_i(\overline{x})$ have coefficients that are integers of the field $\mathbb{K}$, and that these linear forms take rather small values at the point $(e^{\alpha_1}, \dots, e^{\alpha_m})$.

As usual in $E$-functions theory, another set of linearly independent linear forms in $(e^{\alpha_1}, \dots, e^{\alpha_m})$ with coefficients in $\mathbb{Z}$ can be constructed as

$$e^{k_1\theta_1 + \dots + k_r\theta_r} P(e^{\theta_1}, \dots, e^{\theta_r}), \qquad 0 \le k_1 \le \mu_1, \dots, 0 \le k_r \le \mu_r.$$

To these linear forms one can apply Siegel's arguments [24], and due to (15), derive the bound for the value of the polynomial $P$ in Theorem 2.

The quantitative improvement of the Lindemann–Weierstrass Theorem in its linear version—see the beginning of the chapter—with an estimate depending on each coefficient was first proved by A. Baker in 1965.

Theorem 3. *Let $\alpha_1, \dots, \alpha_m$ be distinct nonzero rational numbers. Then for any nonzero integers $h_1, \dots, h_m$ one has*

$$|h_1 e^{\alpha_1} + \dots + h_m e^{\alpha_m}| \ge c|h_1 \cdots h_m|^{-1} h^{1-\varepsilon(h)}$$

*where $h = \max_{1 \le i \le m} |h_i|$ and $\varepsilon(h) = \nu(\log\log h)^{-1/2}$.*

Discussion of this theorem and generalisations for other functions can be found in [4, Ch. 5,§6].

## 2.4 Lindemann's Theorem for $p$-adic numbers

For every prime number $p$ denote by $\mathbb{Q}_p$ the field of $p$-adic numbers. Let $\mathbb{C}_p$ be the completion of the algebraic closure of $\mathbb{Q}_p$. The field $\mathbb{C}_p$ is complete and algebraically closed. We denote by $|\cdot|_p$ the $p$-adic absolute value extended to the field $\mathbb{C}_p$ and satisfying $|p|_p = p^{-1}$. In what follows, we shall call any element of the field $\mathbb{C}_p$ algebraic over $\mathbb{Q}$ a $p$-adic algebraic number.

The $p$-adic exponential function $e^z = \exp_p(z)$ is determined by the usual series

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

As is known, the domain of convergence of this series in $\mathbb{C}_p$ consists of numbers $z$ satisfying the inequality $|z|_p < p^{-\frac{1}{p-1}}$, see [10]. The number $n!$ is very large in complex absolute value, but the same number is rather small in $p$-adic absolute value. This explains difficulties of the $p$-adic case in connection to exponential functions. Nevertheless, in 1932, Mahler [M14] proved the $p$-adic analogue of Lindemann's theorem.

Theorem 4. *If $\alpha \in \mathbb{C}_p$ is a non-zero $p$-adic algebraic number, $|\alpha|_p < p^{-\frac{1}{p-1}}$, then $e^\alpha \in \mathbb{C}_p$ is transcendental.*

The main ideas of the proof are connected to the identity (4), where we take special values of parameters

$$n_1 = \cdots = n_m = n, \qquad \alpha_k = k-1, \qquad k = 1, \dots, m.$$

Expanding the function $e^{z\zeta}$ in a power series in $z$, substituting this expansion in the integral (4) and integrating term by term, we find

$$R(z) = \sum_{k=N}^{\infty} \frac{c_k}{k!} z^k = \sum_{k=1}^{m} P_k(z) e^{(k-1)z} = Q(z, e^z), \tag{16}$$

where $Q(x,y) = \sum_{k=1}^m \sum_{j=0}^n \frac{a_{k,j}}{(n-j)!} x^{n-j} y^{k-1}$ and

$$c_k = \sum_{s_1+\cdots+s_m=k-N} \binom{s_1+n}{n} \cdots \binom{s_m+n}{n} \alpha_1^{s_1} \cdots \alpha_m^{s_m} \in \mathbb{Z}. \tag{17}$$

Identity (16) is valid for any complex $z$.

Let $f(z)$ and $g(z)$ be two power series from $\mathbb{C}_p[[z]]$ and $h(z) = P(f(z), g(z)) \in \mathbb{C}_p[[z]]$ with $P(x,y) \in \mathbb{Q}[x,y]$. If $f(z)$ and $g(z)$ converge at the point $\zeta \in \mathbb{C}_p$, then $h(z)$ also converges at this point, and $h(\zeta) = P(f(\zeta), g(\zeta))$. To prove this assertion, we can use induction on the number of monomials and factors in $P(x,y)$. This reduces the proof to the cases $P = x + y$ or $P = xy$. In these cases, the assertion can be proved directly. In our situation $f(z) = z$, $g(z) = e^z$ and we have

$$\frac{\zeta^N}{N!}\left(1 + \sum_{j=1}^{\infty} \frac{c_{N+j}}{(N+1)\cdots(N+j)}\zeta^j\right) = Q(\zeta, e^\zeta) = \sum_{k=1}^{m} P_k(\zeta)e^{(k-1)\zeta} \tag{18}$$

for any $\zeta \in \mathbb{C}_p$ such that $|\zeta|_p < p^{-\frac{1}{p-1}}$.

Assume now that $e^\alpha \in \mathbb{C}_p$ is an algebraic number and denote by $\nu$ the degree of the field $\mathbb{F} = \mathbb{Q}(\alpha, e^\alpha) \subset \mathbb{C}_p$ over $\mathbb{Q}$. Mahler uses very delicate choices of parameters: $\zeta = p^\lambda \alpha$, where $\lambda$ is rather large natural number, dependig on $\nu$ and $p$ [4],

$$n = p^r, \quad \varphi = \varphi(n+1), \quad m = \frac{p^{r+\mu\varphi}+1}{n+1}, \quad N = p^{r+\mu\varphi}.$$

Here, by choosing sufficiently large positive integers $r$ and $\mu$, conditions $n > n_0(\zeta)$ and $m > m_0(\zeta, n)$ are provided. With this choice, one can prove that the $p$-adic absolute value of the expression in brackets on the left side of (18) is equal to 1. Hence

$$0 < |Q(\zeta, e^\zeta)|_p = \left|\frac{\zeta^N}{N!}\right|_p. \tag{19}$$

This equality gives an upper bound for $|Q(\zeta, e^\zeta)|_p$. From the other side, due to the precise expression for $Q(x, y)$, one can estimate the common denominator, maximum of all coefficients of $Q$ and degrees of $Q$, then according to Liouville's theorem find a lower bound for $|Q(\zeta, e^\zeta)|_p$. But this lower bound contradicts the upper bound following from (19). This contradiction implies that the number $e^\alpha$ is transcendental.

In 1935, Mahler proved [M30, M30a], that *for* $\alpha, \beta \in \mathbb{Q}_p$ *algebraic over* $\mathbb{Q}$ *with* $0 < |\alpha - 1|_p \leq p^{-1}$, $0 < |\beta - 1|_p \leq p^{-1}$ *the fraction* $\log_p \alpha / \log_p \beta$ *is rational or it is transcendental.* In case $p = 2$ one should additionally exclude $\beta = -1$. This was the $p$-adic analogue of the Hilbert's seventh problem proved one year before by A. O. Gelfond and Th. Schneider. Mahler's proof basically follows Gelfond's method, but instead of complex integrals, Mahler applies the following statement.

Lemma 1. *Let* $f(x) = a_0 + a_1x + a_2x^2 + \cdots$ *be a power series with* $a_k \in \mathbb{C}_p$, $|a_k|_p \leq 1$, $\lim_{k\to\infty} |a_k|_p = 0$, *and suppose that the function* $f(x)$ *defined by this series in the circle* $|x| \leq 1$ *has zeros at points* $x_0, x_1, \ldots, x_{\nu-1}$ *with multiplicities* $d_0, d_1, \ldots, d_{\nu-1}$, *respectively. Suppose*

$$|x_\lambda|_p \leq \frac{1}{p}, \qquad \lambda = 0, 1, \ldots, \nu - 1,$$

[4] $\varphi(n)$ is the Euler totient function.

*and $x$ satisfies the inequality*

$$|x|_p \leq \frac{1}{p}.$$

*Then*

$$|f^{(k)}(x)|_p \leq p^{-(d_0+\cdots+d_{\nu-1}-k)}, \qquad k = 0, 1, 2, \ldots.$$

Another useful tool proved in [M30] is a $p$-adic analogue of Liouville's theorem for lower bounds for polynomials from $\mathbb{Z}[x_1, \ldots, x_m]$ at algebraic points.

The proof of the equivalent form of the Gelfond–Schneider Theorem about transcendence of numbers $\alpha^\beta = e^{\beta \log_p \alpha}$ with algebraic $\alpha, \beta \in \mathbb{C}_p$ was published in 1940 by Veldkamp [25]. He used Schneider's approach. Other results about transcendence and algebraic independence of numbers connected to values of the $p$-adic exponential function can be found in the article [20], which also contains the proof of the following theorem ("half of the Lindemann–Weierstrass Theorem" in $p$-adic case).

Theorem 5. *Let $\alpha_1, \ldots, \alpha_m \in \mathbb{C}_p$ be algebraic numbers that form a basis of a finite extension of $\mathbb{Q}$ and satisfy the inequalities $|\alpha_j|_p < p^{-\frac{1}{p-1}}$. Then among the numbers*

$$e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_m},$$

*there are at least $\frac{m}{2}$ algebraically independent over $\mathbb{Q}$.*

To prove the analogue of the Lindemann–Weierstrass theorem for $p$-adic numbers is still an open problem.

### 2.5 Lower bounds for polynomials and polynomials with two-sided bounds

In 1932, Mahler [M11, M13] used the functional identity (3) to prove a bound for the transcendence measure of $\pi$ and the transcendence measure of the logarithm of a rational number.

Theorem 6. *Let $\omega = \pi$ or else $\omega = \log(a/b) \in \mathbb{R}$, where $a, b \in \mathbb{N}$. Then there exist constants $c = c(a, b) > 0$ and $\gamma = \gamma(a, b, d) > 0$ such that*

$$|P(\omega)| \geq \gamma \cdot H^{-c^d}$$

*for any $P(x) \in \mathbb{Z}[x]$ with $P \not\equiv 0$, $\deg P \leq d$ and $H(P) \leq H$.*

In order to prove theorem 6, one has to show that there exists a sequence of polynomials $Q_m(x) \in \mathbb{Z}[x]$ of fixed degree such that

$$0 < H(Q_m) < C^m, \qquad C^{-\lambda_1 m} \leq |Q_m(\omega)| \leq C^{-\lambda_2 m}, \tag{20}$$

where $C > 1$, $\lambda_1$ and $\lambda_2$ are constants with $\lambda_1 > \lambda_2 > 0$, see [M13, Section 20].

Mahler constructed the required sequence of polynomials using the functional identity (3) with $n_1 = \cdots = n_m = n$ and $\alpha_k = k-1$. If $\theta \neq 1$ is a positive rational number, then for fixed $n$ and increasing $m$ the sequence of expressions in (3) with $z = \log\theta$ will be a sequence of polynomials of degree $n$ in $z = \log\theta$ with rational coefficients. The bounds (20) can be proved for these polynomials after multiplication in (3) of some factors killing denominators of $\theta^k$ and coefficients of $P_k(x)$. To prove the two-sided estimate for this sequence of polynomials for positive real $z$ Mahler represents the integral

$$R_m(z) = \frac{1}{2\pi i}\int_C \frac{e^{z\zeta}d\zeta}{\zeta^{n+1}(\zeta-1)^{n+1}\cdots(\zeta-m+1)^{n+1}} = \\ = \frac{1}{2\pi i}\int_C \left(\frac{\Gamma(\zeta-m+1)}{\Gamma(\zeta+1)}\right)^{n+1} e^{z\zeta}d\zeta, \tag{21}$$

in the form of a multiple real integral

$$R(z) = z^N \int_0^1 \cdots \int_0^1 U(x_1,\ldots,x_{m-1})e^{zV(x_1,\ldots,x_{m-1})}dx_1\cdots dx_{m-1}, \tag{22}$$

where $N = -1 + \sum_{k=1}^m (n_k+1)$,

$$\begin{aligned} V(x_1,\cdots,x_{m-1}) =& (\alpha_1-\alpha_2)x_1\cdots x_{m-1} + \cdots \\ &+ (\alpha_{m-2}-\alpha_{m-1})x_{m-2}x_{m-1} + \alpha_{m-1}x_{m-1}, \\ U(x_1,\ldots,x_{m-1}) =& (1-x_1)^{n_2}\cdots(1-x_{m-1})^{n_m} \\ &\times x_1^{n_1}x_2^{n_1+n_2+1}\cdots x_{m-1}^{n_1+\cdots+n_{m-1}+m-2}. \end{aligned}$$

The choice of parameters $n_k$ and $\alpha_k$ was indicated earlier. Hermite actually found a similar identity as early as 1873. In the case $\alpha_m = 0$, he published this representation for the reminder $R(z)$ in a functional approximation (see [9], Vol.3, p.146-149).

Mahler applies the mean value theorem to the integral (22) and obtains the bounds

$$\frac{z^N}{N!} \le R_m(z) \le e^{mz}\frac{z^N}{N!},$$

where $N = m(n+1)-1$.

To estimate polynomials at the point $\pi$, Mahler considered the sequence $R_m(2\pi i)$ of polynomials in $\pi$. In the complex case, the mean value theorem is not applicable. Mahler transformed the integration path of the complex integral into a straight line that goes perpendicular to the real axis through a point on the real axis (Barne's type of integral) and finds the asymptotic behaviour of $R_m(2\pi i)$ in the form (21) as $m \to \infty$. As a result, he again obtains the two-sided estimate (20). It seems this was the first application of the saddle point method in transcendence theory!

Mahler's application of the two-sided sequence of polynomials was also used by Koksma and Popken [11]. In 1932, they used Gelfond's analytic method to construct a sequence of polynomials $A_n$ with two-sided bounds for their values at the point $e^\pi$ and proved by this way a lower estimate for the measure of transcendence of $e^\pi$. Since that time, in connection with the estimates of irrationality measures, transcendence measures, linear and algebraic independence of numbers, many auxiliary assertions about the sequences of polynomials with two-sided estimates have been proved. We note here the results of Gelfond, Danilov, Chudnovskii, Hata, and the results of many mathematicians in connection with the proofs of the algebraic independence of numbers. A review of this trend of transcendental number theory can be found in the book [4].

We present here a statement of this kind, see [17], that was used in the proof of Rivoal's Theorem (see Theorem 9 below) about zeta-values.

Lemma 2. *Let $N_0, c_1, c_2, \tau_1$ and $\tau_2 > \tau_1$ denote positive numbers, and let $\sigma(t)$ be a monotone-increasing function for all $t \geq N_0$, satisfying*

$$\lim_{t\to\infty} \sigma(t) = \infty, \qquad \overline{\lim}_{t\to\infty} \frac{\sigma(t+1)}{\sigma(t)} = 1.$$

*Let $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m$, $\theta \neq 0$, and assume that for each natural number $N > N_0$ there exists a linear form $L_N(\overline{x}) = a_{N,1}x_1 + \cdots + a_{N,m}x_m$ with integer coefficients such that*

$$\log \|L_N\| < \sigma(N), \qquad c_1 e^{-\tau_1\sigma(N)} \leq |L_N(\overline{\theta})| \leq c_2 e^{-\tau_2\sigma(N)},$$

*where $\|L_N\|$ denotes the length of the vector $(a_{N,1}, \dots, a_{N,m})$. Then the number of $\mathbb{Q}$-linearly independent elements of $\theta_1, \dots, \theta_m$ is at least $\frac{\tau_1+1}{1+\tau_1-\tau_2}$. In particular, if $(m-1)\tau_2 - (m-2)(1+\tau_1) > 0$, then the numbers $\theta_1, \dots, \theta_m$ are linearly independent over $\mathbb{Q}$.*

## 3 Variations

### 3.1 Variation 1

In 1953, Mahler [M118, M119] proved the inequalities

$$\left|\pi - \frac{p}{q}\right| > q^{-30}, \quad q \geq q_0, \quad \text{and} \quad \left|\pi - \frac{p}{q}\right| > q^{-42}, \quad q \geq 2. \tag{23}$$

The first inequality (23) can be expressed as a bound for the exponent of irrationality of $\pi$, namely, $\mu(\pi) \leq 30$. The second inequality is a famous result of Mahler. The proof of both inequalities gives evidence that the constant $q_0$ can be precisely computed and that Mahler's approach is effective. These inequalities started a line of improvements.

To explain the proof of (23) let us take $\alpha_k = k-1$, $1 \leq k \leq m$ and $n_1 = n_2 = \cdots = n_m = n$ in the identity (4). Denote

$$z = \log x, \quad -\pi < \arg x < \pi, \quad x^\zeta = e^{\zeta \log x}.$$

Then the identity (4) can be written in the form

$$F(x)=\frac{1}{2\pi i}\int_C\frac{x^{\zeta}d\zeta}{(\zeta(\zeta-1)\cdots(\zeta-m+1))^n}$$
$$=\sum_{k=1}^{m}P_k(\log x)x^{k-1}=\sum_{j=0}^{n}A_j(x)(\log x)^j.$$

This means the function $F(x)$ can be represented as a polynomial in $x$ and $\log x$ or as a linear form in $\log x$ with coefficients in $\mathbb{Q}[x]$. We can reduce the power of the zeta and consider the integrals

$$F_h(x)=\frac{1}{2\pi i}\int_C\frac{\zeta^h x^{\zeta}d\zeta}{(\zeta(\zeta-1)\cdots(\zeta-m+1))^{n+1}}=\sum_{j=0}^{n}A_{h,j}(x)(\log x)^j,$$
$$A_{h,j}(x)\in\mathbb{Q}[x],\quad \deg A_{h,j}(x)\le m-1,\quad h=0,1,\dots,n.$$

Let us take $n=10$, $x=i$ and $\log x=\frac{\pi i}{2}$. One can prove by the same way as in Section 2.3 that $\Delta(i)=\det\|A_{h,j}(i)\|\neq 0$. Consequently, there exists an index $h$ such that the polynomial $S(y)=\sum_{j=0}^{n}A_{h,j}(x)y^j$ is distinct from 0 at the point $pi/2q$. Both polynomials $S(y)$ and $2qy-pi$ are very small at the point $\pi i/2$, the first one if $m$ is sufficiently large, and the second one if we assume that (23) is wrong. This implies that the resultant $S(pi/2q)$ of these two polynomials should be small too. This number can be estimated from below, since it belongs to $\mathbb{Q}(i)$, is distinct from 0 and its denominator can be easily bounded from above. With the correct choice of $m$, the upper and lower bounds for the resultant contradict one another. Fractions, for which the second inequality is violated are convergents to $\pi$. Moreover, since $q_0$ can be explicitly computed in the proof of the first inequality, it is possible to specify exactly the finite set of "bad" convergents. The question is then solved by a small number of checks. This is a sketch of the proof of (23).

In 1974 Mignotte, with $n=5$, improved Mahler's result. His bound is $\mu(\pi)\le 20$. G. Chudnovskii with the same construction and $n=5$, but with more precise calculations proved $\mu(\pi)\le 19.889999\ldots$. The following essential improvement of this estimate belongs to M. Hata, in 1993. In the identity (4), he used the parameters

$$\alpha_k=k-1,\quad 1\le k\le m,\quad m=r+3s+1,\qquad s=\left[\frac{3r}{31}\right],$$

and $R(z)$ equal to

$$\frac{1}{2\pi i}\int_C\frac{e^{z\zeta}d\zeta}{(\zeta-r-3s)_{r+3s+1}(\zeta-r-2s)_{r+s+1}(\zeta-r-s)_{r-s+1}(\zeta-r)_{r-3s+1}},$$

where $(\lambda+1)_\nu=(\lambda+1)\cdots(\lambda+\nu)$ if $\nu\ge 1$, and $(\lambda+1)_0=1$ for any complex $\lambda$. Hata's bound is $\mu(\pi)\le 13.398$. The present record value is $\mu(\pi)\le$

7.606308..., which belongs to V. Salikhov [23]. Such results were proved not only for $\pi = \frac{1}{2i}\log 1$, but also for $\log 2$, $\log 3$ and logarithms of some other integer and rational numbers.

### 3.2 Variation 2

Let $r_1, \ldots, r_m$ be positive integers and let $w_1, \ldots, w_m$ be complex numbers such that $w_j - w_k \notin \mathbb{Z}$ when $j \neq k$. In (4), we let $\sigma = r_1 + \cdots + r_n$ instead of $m$, $n_1 = \cdots = n_\sigma = 0$ and $\alpha_1, \ldots, \alpha_\sigma$ be the numbers

$$w_j + h, \quad 1 \le j \le m, \quad 0 \le h < r_j, \quad h \in \mathbb{Z},$$

ordered in some way, and set

$$Q(\zeta) = \prod_{j=1}^{m} \prod_{h=0}^{r_j-1} (\zeta - w_j - h).$$

If we further take $z = \log(1-x)$ with $x \neq 1$, then (4) becomes

$$E(x) = \frac{1}{2\pi i} \int_C \frac{(1-x)^\zeta d\zeta}{Q(\zeta)} = \sum_{j=1}^{m} Q_j(x)(1-x)^{w_j}, \tag{24}$$

where $Q_j(x) \in \mathbb{C}[x]$, $\deg Q_j(x) \le r_j - 1$ and $C$ is a circle of sufficiently large radius. Since $\operatorname{ord}_{x=0} E(x) = \sigma - 1$, the formula (24) gives the Hermite–Padé approximation for binomials $(1-x)^{w_j}$ at the point $x = 0$. Mahler used $w_j = \frac{j-1}{n}$, $1 \le j \le m$ and $r_1 = \ldots = r_m = r$ to prove the following result.

Theorem 7. *Let $\xi = (a/b)^{1/n}$ be an algebraic number of degree $n \ge 3$, where $a$ and $b$ are natural numbers, let $\varepsilon$ be an arbitrary positive number and let $s$ be an arbitrary natural number. Then the inequality*

$$\left|\xi - \frac{p}{q}\right| < q^{1-s-(n/s)-\varepsilon}$$

*has only finitely many rational solutions $p/q$.*

This is an analogue of a general Thue–Siegel non-effective theorem in the special case $\xi = (a/b)^{1/n}$. We will neither be concerned with ineffective results on the approximation of algebraic numbers by rationals, proven in the line from Thue to Roth, nor effective estimates obtained using linear forms in the logarithms of algebraic numbers (the direction of research initiated by A. Baker). The binomial $(1-x)^w = F(-w, 1, 1; x)$, and, as any Gaussian hypergeometric function with one of the first two parameters equal to 1, admits a decomposition into an irregular continued fraction, all of whose elements can be written explicitly. Apparently, Thue, in 1908, was the first to use this continued fraction to construct rational approximations to irrational numbers of the form $(a/b)^{1/n}$.

These approximations allowed him, in 1918, to prove the first effective results on the solutions of Diophantine equations of a special type. This method was subsequently developed by many mathematicians. Such remarkable scientists as Siegel and Baker participated in this activity. For example, Baker in 1964 proved that

$$\left|\sqrt[3]{2}-\frac{p}{q}\right| > 10^{-6}q^{-2.955}.$$

The approach proposed by Mahler, using the Hermite–Padé approximations, turned out to be less popular. Nevertheless, with its help, interesting new results were obtained. Among the effective results related to the use of the hypergeometric functions for constructing Diophantine approximations, we only touch upon the following theorem of M. A. Bennett [3] whose proof uses Mahler's approach.

Theorem 8. *If $a, b$ and $n$ are integers with $ab \neq 0$ and $n \geq 3$, then the equation*

$$|ax^n - by^n| = 1$$

*has at most one solution in positive integers $(x, y)$.*

### 3.3 Variation 3

The exponential function is not the only function that can be placed into Hermite's identity (3). For any meromorphic function $f(z)$ satisfying a linear differential equation with coefficients from $\mathbb{C}(z)$, the right side of (3) contains not more than $md$ terms, where $d$ is the order of the differential equation. In many applications the generalised hypergeometric functions are used. These functions are solutions of linear differential equations with coefficients from $\mathbb{C}(z)$. They have representations as real multiple integrals (Euler representation, like (22)), as complex integrals (Barne's type integrals, like (21)) and power series (like (16) and (17)). The reason, apparently, is that the exponential function is the so-called confluent case of a Gaussian hypergeometric function. The choice of representation depends on convenience of work, for example, as in Theorem 6 above. In 1976, A. Galochkin [5] applied Mahler's technique to the function

$$\varphi(z) = \sum_{\nu=0}^{\infty} \frac{1}{(\lambda_1+1)_\nu \cdots (\lambda_m+1)_\nu}\left(\frac{z}{m}\right)^{m\nu}, \qquad \lambda_i \neq -1, -2, \ldots .$$

He constructed Hermite–Padé approximations for functions $\varphi_j(z) = \delta^{j-1}\varphi(z)$, $j = 1, \ldots, m$, where $\delta = z\frac{d}{dz}$ by the formulas

$$\begin{aligned} F_i(z) = \sum_{\nu=mn-1+i}^{\infty} &\frac{\nu(\nu-1)\cdots(\nu-mn+2-i)}{(\lambda_1+1)_\nu \cdots (\lambda_m+1)_\nu} z^\nu \\ &= P_{i,1}(z)\varphi_1(z) + \cdots + P_{i,m}(z)\varphi_m(z), \qquad i = 1, \ldots, m. \end{aligned}$$

These linear forms $F_i(z)$ are linearly independent. Under special arithmetical conditions on the numbers $\lambda_i$, Galochkin proved very precise lower bounds for linear forms with integer coefficients in the numbers $\varphi_j(1/b)$, $j = 1, \ldots, m$, where $b > 0$ is an integer. Another broader set of functions, to which the Mahler approach is applicable, is the set of so-called Meyer's $G$-functions. An example of the use of these functions to prove the irrationality of the number $\zeta(3)$ can be found in the article [19].

### 3.4 Variation 4

In 1979, E. M. Nikishin [21] constructed Hermite–Padé approximations for 1 and the polylogarithm functions $L_1(1/z), L_2(1/z), \ldots, L_m(1/z)$ in a neighbourhood of $\infty$. Here

$$L_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k}, \quad k = 1, 2, \ldots, m.$$

Define the rational function

$$R(\zeta) = \frac{\zeta(\zeta-1)\cdots(\zeta-nm-q+2)}{(\zeta+1)^m\cdots(\zeta+n)^m(\zeta+n+1)^q} \tag{25}$$

with integers $n, m, q$, satisfying inequalities $n \geq 1$ and $0 \leq q \leq m$. Nikishin proved the identity

$$\begin{aligned} H_q(z) &= \frac{1}{2iz}\int_C R(t)\left(-\frac{1}{z}\right)^{\zeta}\frac{d\zeta}{\sin(\pi\zeta)} \\ &= A_0(z) + A_1(z)L_1(1/z) + \cdots + A_m(z)L_m(1/z), \end{aligned} \tag{26}$$

where the contour $C$ is the line $\Re\zeta = -\frac{1}{2}$ traversed in the upward direction and the coefficients $c_{k,j}$ of polynomials

$$A_k(z) = \sum_{j=0}^{n} c_{k,j} z^j, \qquad k = 1, \ldots, m$$

are defined by the partial fraction expansion

$$R(t) = \sum_{j=0}^{n-1}\sum_{k=1}^{m} \frac{c_{k,j}}{(t+j+1)^k} + \sum_{k=1}^{q} \frac{c_{k,n}}{(t+n+1)^k}.$$

In particular,

$$\deg A_0(z) \leq n-1, \quad \deg A_k(z) \leq \begin{cases} n & \text{for } \ 1 \leq k \leq q, \\ n-1 & \text{for } \ q+1 \leq k \leq m. \end{cases}$$

The integral converges for $|z| > 1$ and $\operatorname{ord}_{z=\infty} H_q(z) \geq nm+q$. Note that both sides of (26) are linear in $R(t)$. That's why the case (25) can be reduced to simplest fractions $R(t) = \frac{1}{(t+\ell)^k}$, $k, \ell \geq 1$. In this last case, the identity can be easily checked by the computations of residues. Moreover, by the same way, the identity (26) can be proved for any rational function $R(t) \in \mathbb{Q}(t)$ having a zero of the second order at infinity and whose poles are negative integers. By this functional construction, Nikishin proved *the linear independence over* $\mathbb{Q}$ *of the numbers* $1, L_1(a/b), L_2(a/b), \ldots, L_m(a/b)$ *at rational points* $a/b$ *with integers* $a$ *and* $b$ *satisfying some conditions.* In particular, the number $a/b$ should be very small. For the proof, Nikishin worked with the functional determinant as at the end of Subsection 2.3; a complete set of small linear forms in these numbers was constructed from the values $H_q(a/b)$ for $q = 0, 1, \ldots, m$.

### 3.5 Variation 5

Following Rivoal, let us take

$$R(s) = (n!)^{m-2r} \frac{(s-rn)\cdots(s-1)(s+n+1)\cdots(s+(r+1)n)}{(s(s+1)\cdots(s+n))^m}, \tag{27}$$

where $n$ is even, $m = 2d + 1 \geq 3$ is an odd number and where $r = \lfloor \frac{m}{\log^2 m} \rfloor$. It is easy to check that, under these assumptions, the rational function (27) satisfies the identity $R(-n-s) = -R(s)$. This leads to the equalities $A_k(1) = (-1)^{k+1} A_k(1)$ for the coefficients $A_k(z)$ in the linear form (26) constructed for the function (27). Consequently,

$$A_{2l}(1) = 0.$$

Since $L_k(1) = \zeta(k)$, $k \geq 2$, we derive

$$\sum_{\nu=1}^{\infty} R(\nu) = A_0(1) + A_3(1)\zeta(3) + A_5(1)\zeta(5) + \cdots + A_{2d+1}(1)\zeta(2d+1),$$

a linear form in values of the zeta-function at odd points with rational coefficients. Rivoal proved with this construction that there are infinitely many numbers $\zeta(2k+1)$ which together with 1 are linearly independent over $\mathbb{Q}$. So in particular there are infinitely many irrationals among $\zeta(2k+1)$.

The symmetry of the function $R(t)$, of course, does not allow one to easily construct a complete system of linearly independent linear forms, as in the previous case, but in this situation Rivoal uses a lemma on the sequence of linear forms with two-sided estimates, see Lemma 2 in Section 2 above. In this way he obtains an estimate that grows in accordance with $m$, and this proves the following theorem of Rivoal [22]; see also Ball and Rivoal [1].

Theorem 9. *The linear space over the field of rational numbers* $\mathbb{Q}$*, generated by the values of the Riemann zeta function*

$$\zeta(2k+1), \qquad k = 1, 2, 3, \ldots$$

*has infinite dimension. In particular, among these numbers there are infinitely many irrationals.*

# 4 Modular functions

## 4.1 Mahler and Popken

Apparently, the first work devoted to the transcendence problems of the values of modular functions was the joint article of K. Mahler and J. Popken [M33], published in 1935. Another version of the reasoning is documented in Mahler's book [M200]. In this section, we shortly discuss these expositions. Definitions and properties of normalised Eisenstein series[5] $E_{2k}(\tau)$, theta-functions and theta-constants $\theta_2(\tau)$, $\theta_3(\tau)$, $\theta_4(\tau)$, Weierstrass $\wp$ and sigma-functions can be found, for example, in the books [13], [14] or [28].

Theorem 10. *The following hold.*

1) *For any Weierstrass elliptic function $\wp(z)$ with invariants $g_2$ and $g_3$, at least one of three numbers*

$$\frac{\omega}{\pi}\cdot\frac{\eta}{\pi}, \quad \left(\frac{\omega}{\pi}\right)^4 g_2, \quad \left(\frac{\omega}{\pi}\right)^6 g_3,$$

*is transcendental. Here $\omega$ is one half of a primitive period of $\wp(z)$ and $\eta$ is corresponding quasiperiod.*

2) *For any $\xi \in \mathbb{C}$ with $\Im\xi > 0$, at least one of the three values of the Eisenstein series*

$$E_2(\xi), \quad E_4(\xi), \quad E_6(\xi) \tag{28}$$

*is transcendental.*

3) *For any $\xi \in \mathbb{C}$ with $\Im\xi > 0$, at least one of values of the logarithmic derivatives of theta-constants*

$$\psi_2(\xi), \quad \psi_3(\xi), \quad \psi_4(\xi).$$

*is transcendental. Here $\psi_k(\tau) = \frac{\delta\theta_k}{\theta_k}(\tau)$, $k = 2, 3, 4$ and $\delta = \frac{1}{\pi i}\frac{\partial}{\partial\tau} = q\frac{\partial}{\partial q}$.*

Since the identities

$$E_2\left(\frac{\omega'}{\omega}\right) = 3\cdot\frac{\omega}{\pi}\cdot\frac{\eta}{\pi}, \quad E_4\left(\frac{\omega'}{\omega}\right) = \frac{3}{4}\left(\frac{\omega}{\pi}\right)^4 g_2, \quad E_6\left(\frac{\omega'}{\omega}\right) = \frac{27}{8}\left(\frac{\omega}{\pi}\right)^6 g_3, \tag{29}$$

hold, where $2\omega$ and $2\omega'$ generate the period lattice and $\Im(\omega'/\omega) > 0$, statements 1) and 2) of Theorem 10 are equivalent. Mahler and Popken gave a direct proof of the second statement. Here is a sketch of this proof. Let $\sigma(u)$ be a Weierstrass sigma-function and let $f(x)$ be defined by the identity

$$e^{-\frac{\eta u}{2\omega}}\sigma(u) = \frac{\omega}{\pi i}xf(x), \quad \text{where} \quad x = 2i\sin\frac{\pi u}{2\omega}.$$

[5] For brevity, the word "normalised" will be omitted later.

Another way to write this function is

$$xf(x) = 2i\frac{\theta_1(\frac{\pi u}{2\omega},\tau)}{\theta_1'(0)}.$$

Then

$$f(x) = \prod_{n=1}^{\infty}\left(1 - \frac{q^{2n}x^2}{(1-q^{2n})^2}\right) = \sum_{k=0}^{\infty} a_k x^{2k},$$

where $q = e^{\frac{\pi i\omega'}{\omega}}$. The product allows one to give very good upper bounds for the Taylor coefficients; in particular, $|a_h| \leq e^{-c_1h^2}$. In a previous work, Popken proved a general result about lower bounds for the moduli of non-zero coefficients of the Taylor series of a function satisfying an algebraic differential equation with coefficients in $\mathbb{Q}$, provided all the coefficients of the function are algebraic numbers. The function $f(x)$ satisfies a differential equation of the suitable type and Popken's lower bound contradicts the upper bound for $|a_h|$ indicated above. This implies that at least one of coefficients $a_h$ should be transcendental. Due to algebraic relations between the $a_h$ and Eisenstein series, one can claim that at least one of the Eisenstein series has transcendental value at the point $\tau$. And this proves the second statement of the Theorem 10.

One can check that the Eisenstein series can be expressed as polynomials with rational coefficients in logarithmic derivatives of theta-constants. It is not difficult to verify, using these identities, that statements 2) and 3) are equivalent. Now we give a sketch of the proof of the last statement; it is essentially shorter than the original Mahler–Popken proof of the statement 2). Define polynomials

$$S_n(x) = \prod_{k=0}^{n-1}(x-k^2), \qquad n \geq 1, \qquad S_0(x) = 1.$$

The first step is to prove that, *for any integer* $n \geq 0$ *there exists a polynomial* $B_n[x_2, x_3, x_4] \in \mathbb{Z}[x_2, x_3, x_4]$ *such that*

$$\frac{S_n(\delta)\theta_3}{\theta_3} = B_n(\psi_2, \psi_3, \psi_4), \tag{30}$$

*wherein*

$$B_n = O\left(n!(n+1)!(1 + x_2 + x_3 + x_4)^n\right). \tag{31}$$

The proof of this assertion uses induction on the order of differentiation and is based on the fact that the functions $\psi_2$, $\psi_3$ and $\psi_4$ satisfy the system of differential equations [8]

$$\begin{aligned} D\psi_2 &= 2(\psi_2\psi_3 + \psi_2\psi_4 - \psi_3\psi_4),\\ D\psi_3 &= 2(\psi_3\psi_2 + \psi_3\psi_4 - \psi_2\psi_4),\\ D\psi_4 &= 2(\psi_4\psi_2 + \psi_4\psi_3 - \psi_2\psi_3), \end{aligned} \tag{32}$$

where

$$D = 2(x_2x_3 + x_2x_4 - x_3x_4)\frac{\partial}{\partial x_2} + 2(x_3x_2 + x_3x_4 - x_2x_4)\frac{\partial}{\partial x_3} + 2(x_4x_2 + x_4x_3 - x_2x_3)\frac{\partial}{\partial x_4}.$$

For convienece, set $\overline{x} = (x_2, x_3, x_4)$ We define

$$B_{n+1}(\overline{x}) = DB_n(\overline{x}) + (x_3 - n^2)\cdot B_n(\overline{x}), \quad n \geq 0, \qquad B_0(\overline{x}) = 1.$$

Note that functions $\psi_2(\tau), \psi_3(\tau), \psi_4(\tau)$ are algebraically independent over the field $\mathbb{C}(\tau)$; apply [M171, Theorem 1] to the functions $e^{\pi i\tau}$ and $j(\tau)$.

According to (30), for sufficiently large $n$, we have

$$B_n(\psi_2, \psi_3, \psi_4) = \frac{S_n(\delta)\theta_3}{\theta_3} = \frac{1}{\theta_3}\sum_{j=n}^{\infty}\prod_{k=0}^{n-1}(j^2 - k^2)q^{j^2}, \tag{33}$$

$$\left|B_n(\psi_2, \psi_3, \psi_4) - \frac{1}{\theta_3}n(2n-1)!q^{n^2}\right| < \frac{1}{|\theta_3|}n(2n-1)!|q|^{n^2}(3^{-1} + 3^{-2} + \cdots) = \frac{1}{2|\theta_3|}n(2n-1)!|q|^{n^2}$$

and the bound

$$\frac{1}{2|\theta_3|}n(2n-1)!|q|^{n^2} < |B_n(\psi_2(\omega), \psi_3(\omega), \psi_4(\omega))| < \frac{3}{2|\theta_3|}n(2n-1)!|q|^{n^2}. \tag{34}$$

If we assume that all three values $\psi_2(\omega), \psi_3(\omega), \psi_4(\omega)$ are algebraic numbers and apply Liouville's Theorem to the polynomial $B_n$ and these numbers, then according to (31), we derive the lower bound

$$|B_n(\overline{\psi})| > e^{-cn\ln n},$$

where $c$ is a positive constant. This contradicts the right inequality in (34). This contradiction proves the third assertion of Theorem 10.

### 4.2 Eisenstein series and modular invariants

In 1936–1937, Schneider extended to elliptic functions the method by which he succeeded in solving Hilbert's seventh problem. As corollaries of his general theorems he proved the following statements.

1) *Assume that the invariants $g_2$ and $g_3$ of $\wp(z)$ are algebraic. Then any non-zero period $\omega$ and any quasi-period of $\wp(z)$ are transcendental numbers.*

2) *Let $\wp(z)$ be a Weierstrass elliptic function with algebraic invariants $g_2$ and $g_3$. Let $\alpha$ be a non-zero algebraic number. Then $\alpha$ is not a pole of $\wp(z)$ and $\wp(\alpha)$ is transcendental.*

3) *If $\xi$ is algebraic number of degree larger than 2, then the value of the modular invariant $j(\xi)$ is transcendental.*

The last statement describes a property of the modular function. It seems strange that the method of proof uses elliptic functions. At the time, Schneider posed the problem of finding a proof of this assertion based solely on the properties of modular functions—this problem is still open.

In 1948, Gelfond proved that some finite sets of numbers associated to the exponential function contain at least two algebraically independent numbers. Chudnovsky, in 1976–1979, applied this method to elliptic functions, and proved, in particular, the following statements.

Theorem 11. *The following hold.*

1) *Let $\omega$ be a non-zero period of $\wp(z)$ and $\eta$ be a corresponding quasi-period. If $g_2$ and $g_3$ are algebraic, then the two numbers $\frac{\omega}{\pi}$ and $\frac{\eta}{\pi}$ are algebraically independent.*

2) *Assume that $g_2$ and $g_3$ are algebraic and the elliptic function $\wp(z)$ has complex multiplication. Let $\omega$ be a non-zero period of $\wp(z)$. Then the two numbers $\omega$ and $\pi$ are algebraically independent.*

3) *The numbers $\pi$ and $\Gamma(\frac{1}{4})$ are algebraically independent. Also the numbers $\pi$ and $\Gamma(\frac{1}{3})$ are algebraically independent.*

The first statement of this theorem improves the first statement of Theorem 10 in the case $g_2$ and $g_3$ are algebraic, and due to (29), it proves the following proposition, which is essentially stronger than the second statement of the Theorem 10.

4) *For any complex number $\xi$ with positive imaginary part at least two of the three values of the Eisenstein functions*

$$E_2(\xi), \quad E_4(\xi), \quad E_6(\xi)$$

*are algebraically independent.*

The following theorem was proved in 1996, see [18].

Theorem 12. *For any complex number $\xi$ with positive imaginary part at least three of the four values*

$$e^{2\pi i\xi}, \quad E_2(\xi), \quad E_4(\xi), \quad E_6(\xi)$$

*are algebraically independent.*

The proof of this theorem is based on a number of properties of Eisenstein functions proved in previous papers. In particular,

A) Eisenstein functions can be presented by Fourier series

$$P(e^{2\pi i\tau}) = E_2(\tau), \quad Q(e^{2\pi i\tau}) = E_4(\tau), \quad R(e^{2\pi i\tau}) = E_6(\tau),$$

where $P(z)$, $Q(z)$ and $R(z)$ are power series having integer coefficients that converge in the circle $|z| < 1$.

B) S. Ramanujan established in 1916 that these functions satisfy the differential equations

$$z\frac{dP}{dz} = \frac{1}{12}(P^2 - Q), \quad z\frac{dQ}{dz} = \frac{1}{3}(PQ - R), \quad z\frac{dR}{dz} = \frac{1}{2}(PR - Q^2). \quad (35)$$

C) In 1969, Mahler [M171] proved, as a special case of more general theorem, that the five functions

$$\tau, e^{\pi i\tau}, j(\tau), j'(\tau), j''(\tau) \quad (36)$$

are algebraically independent over $\mathbb{C}$. It is easy to prove that the field generated over $\mathbb{Q}$ by Eisenstein functions coincides with the field generated by the last three functions from (36). In particular, this implies a property that is very important for the proof of the Theorem 12: Ramanujan's functions $P(z)$, $Q(z)$ and $R(z)$ are algebraically independent over the field $\mathbb{C}(z)$.

As a corollary of Theorem 12, one can deduce for $\xi = i$ that the numbers $\pi$, $e^\pi$ and $\Gamma(\frac{1}{4})$ are algebraically independent. There are many other concrete corollaries following from this theorem, for example see [26].

### 4.3 A conjecture of Mahler

Finally, we shortly discuss a conjecture of Mahler formulated at the end of his article [M170].

> "Let
> $$F(z) = j\left(\frac{\log z}{2\pi i}\right) - z^{-1},$$
> where $j(\tau)$ is Weber's modular function of level 1. By the transformation theory of $j(\tau)$, $F(z)$ satisfies an algebraic functional equation. However, a decision as to whether $F(z_0)$ is transcendental for algebraic $z_0$, where $0 < |z_0| < 1$, does not seem to be easy."

This conjecture, in another form, was proved in 1995 by K. Barré-Sirieix, G. Diaz, F. Gramain and G. Philibert [2].

Theorem 13. *For any complex number $\xi$ with positive imaginary part, at least one of the two values*

$$e^{2\pi i\xi}, \quad j(\xi) \tag{37}$$

*is transcendental.*

Since

$$j(\xi) = 12^3 \frac{E_4(\xi)^3}{E_4(\xi)^3 - E_6(\xi)^2},$$

Theorem 13 is a consequence of Theorem 12.

The direct proof[6] of Theorem 13, in lieu of the system of differential equations (35), uses the so-called transformation polynomials $\Phi_n(x,y) \in \mathbb{Z}[x,y]$, $n \geq 2$, for which $\Phi_n(j(n\tau), j(\tau)) = 0$ holds. The assumption that both numbers in (37) are algebraic, due to transformation polynomials, implies the algebraicity of all numbers $j(n\xi)$, $n \geq 2$. To obtain a contradiction, we need to apply Liouville's Theorem to the algebraic numbers $e^{2\pi i\xi}, j(\xi), j(n\xi)$. Consequently, we need to know the size of integer coefficients of the transformation polynomial $\Phi_n(x,y)$ depending on $n$. The first such bound was proved in 1969 by Mahler [M188].

A hypothesis stronger than Mahler's, was published in 1971 by Yu. I. Manin [16]. There are two versions of this hypothesis: for complex numbers and for $p$-adic numbers. We will formulate below a complex version, in a style more appropriate to our exposition—it is equivalent to the original one.

Conjecture 1. *Let $a$ be a non-zero algebraic number and $\xi$ be a complex number with positive imaginary part. Then at least one of the two values*

$$a^{\xi}, \qquad j(\xi) \tag{38}$$

*is transcendental.*

In both cases, $p$-adic and complex, this conjecture is still open. The $p$-adic version of Mahler's conjecture was proved in [2].

Theorem 14. *For any $p$-adic number $q \in \mathbb{C}_p$ with $0 < |q|_p < 1$ at least one of the two $p$-adic numbers*

$$q, \qquad J(q) = 12^3 \frac{Q^3(q)}{Q(q)^3 - R(q)^2}$$

*is transcendental.*

References

[M11] K. Mahler, *Zur Approximation der Exponentialfunktion und des Logarithmus, I*, J. Reine Angew. Math. (Crelle) 166 (1931), 118–136.

[6] it was proved earlier than Theorem 12

[M13] K. Mahler, *Zur Approximation der Exponentialfunktion und des Logarithmus, II*, J. Reine Angew. Math. (Crelle) 166 (1932), 137–150.

[M14] K. Mahler, *Ein Beweis der Transzendenz der $P$-adischen Exponentialfunktion*, J. Reine Angew. Math. (Crelle) 169 (1932), 61–66.

[M30] K. Mahler, *Über transzendente $P$-adische Zahlen*, Compositio Math. 2 (1935), 259–275.

[M30a] K. Mahler, *A correction to "Über transzendente $P$-adische Zahlen"*, Compositio Math. 8 (1949), 112.

[M33] J. Popken and K. Mahler, *Ein neues Prinzip für Transzendenzbeweise*, Proc. Akad. Wet. Amsterdam 38 (1935), 864–871.

[M118] K. Mahler, *On the approximation of logarithms of algebraic numbers*, Philos. Trans. R. Soc. London, Ser. A 245 (1953), 371–398.

[M119] K. Mahler, *On the approximation of* $\pi$, Nederl. Akad. Wet., Proc., Ser. A 56 (1953), 30–42.

[M164] K. Mahler, *Applications of some formulae by Hermite to the approximation of exponentials and logarithms*, Math. Ann. 168 (1967), 200–227.

[M170] K. Mahler, *Remarks on a paper by W. Schwarz*, J. Number Theory 1 (1969), 512–521.

[M171] K. Mahler, *On algebraic differential equations satisfied by automorphic functions*, J. Aust. Math. Soc. 10 (1969), 445–450.

[M188] K. Mahler, *On the coefficients of transformation polynomials for the modular function*, Bull. Aust. Math. Soc. 10 (1974), 197–218.

[M200] K. Mahler, *Lectures on transcendental numbers, Edited and completed by B. Divis and W. J. LeVeque*, Springer, Berlin, 1976.

[1] K. Ball and T. Rivoal, *Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs*, Invent. Math. 146 (2001), no. 1, 193–207.

[2] K. Barré-Sirieix, G. Diaz, F. Gramain, and G. Philibert, *Une preuve de la conjecture de Mahler-Manin*, Invent. Math. 124 (1996), no. 1-3, 1–9.

[3] M. A. Bennett, *Rational approximation to algebraic numbers of small height: the Diophantine equation* $|ax^n - by^n| = 1$, J. Reine Angew. Math. 535 (2001), 1–49.

[4] N. I. Fel'dman and Yu. V. Nesterenko, *Transcendental numbers*, Number theory, IV, Encyclopaedia Math. Sci., vol. 44, Springer, Berlin, 1998, pp. 1–345.

[5] A. I. Galočkin, *Improving the estimates of certain linear forms*, Mat. Zametki 20 (1976), no. 1, 35–45.

[6] A. O. Gel'fond, *Essay on the history and current state of the theory of transcendental numbers*, Selected works, Izdat. "Nauka", Moscow, 1973, pp. 16–37.

[7] A. O. Gel'fond, *On Hilbert's 7th problem*, Selected works, vol. 2, Izdat. "Nauka", Moscow, 1973, pp. 48–50.

[8] G. Halphen, *Sur une système d'equations différentielles*, C. R. Acad. Sci. Paris 92 (1881), 1101–1103.

[9] Ch. Hermite, *Oeuvres*, Gauthier–Villar, Paris, 1917.

[10] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, Springer-Verlag, New York-Heidelberg, 1977, Graduate Texts in Mathematics, vol. 58.

[11] J. F. Koksma and J. Popken, *Zur Transzendenz von* $e^\pi$, J. Reine Angew. Math. 168 (1932), 211–230.

[12] R. O. Kuzmin, *Sur une nouvelle classe de nombres transcendants*, Bull. Acad. Sci. de l'URSS VII ser. 6 (1930), 585–597.

[13] S. Lang, *Elliptic functions*, Addison-Wesley Publishing Co., Inc., Reading, Mass.-London-Amsterdam, 1973, With an appendix by J. Tate.

[14] D. F. Lawden, *Elliptic functions and applications*, Applied Mathematical Sciences, vol. 80, Springer-Verlag, New York, 1989.

[15] F. Lindemann, *Ueber die Zahl* $\pi$, Math. Ann. 20 (1882), no. 2, 213–225.

[16] Ju. I. Manin, *Cyclotomic fields and modular curves*, Uspehi Mat. Nauk 26 (1971), no. 6(162), 7–71.

[17] Yu. V. Nesterenko, *Linear independence of numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1985), no. 1, 46–49, 108.

[18] Yu. V. Nesterenko, *Modular functions and transcendence questions*, Mat. Sb. 187 (1996), no. 9, 65–96.

[19] Yu. V. Nesterenko, *Some remarks on* $\zeta(3)$, Mat. Zametki 59 (1996), no. 6, 865–880, 960.

[20] Yu. V. Nesterenko, *Algebraic independence of p-adic numbers*, Izv. Ross. Akad. Nauk Ser. Mat. 72 (2008), no. 3, 159–174.

[21] E. M. Nikišin, *Irrationality of values of functions $F(x, s)$*, Mat. Sb. (N.S.) 109(151) (1979), no. 3, 410–417, 479.

[22] T. Rivoal, *La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs*, C. R. Acad. Sci. Paris Sér. I Math. 331 (2000), no. 4, 267–270.

[23] V. Kh. Salikhov, *On the measure of irrationality of $\pi$*, Mat. Zametki 88 (2010), no. 4, 583–593.

[24] C. L. Siegel, *Über einige Anwendungen diophantischer Approximationen.*, Abh. Preuß. Akad. Wiss., Phys.-Math. Kl. 1929 (1929), no. 1, 1–70.

[25] G. R. Veldkamp, *Ein Transzendenz-Satz für p-adische Zahlen*, J. London Math. Soc. 15 (1940), 183–192.

[26] M. Waldschmidt, *Elliptic functions and transcendence*, Surveys in number theory, Dev. Math., vol. 17, Springer, New York, 2008, pp. 143–188.

[27] C. Weierstrass, *Zu Lindemann's Abhandlung: "Über die Ludolph'sche Zahl"*, Berl. Ber. 1885 (1885), 1067–1086.

[28] E. T. Whittaker and G. N. Watson, *A course in modern analysis*, Cambridge University Press, 1927.

Yuri Nesterenko
Department of Number Theory
Faculty of Mechanics and Mathematics
Lomonosov Moscow State University
119991, Moscow
Leninskie Gory 1
Russian Federation
nester@mi.ras.ru