The University of Arizona The Southwest Center for Arithmetic Geometry
2008 Arizona Winter School, March 15-19, 2008 *Special Functions and Transcendence*
http://swc.math.arizona.edu/aws/08/index.html *Updated:* February 26, 2008

An introduction to irrationality and transcendence methods.

Michel Waldschmidt

Lecture 2 [8]

2 Historical introduction to transcendence

In 1873 C. Hermite [6] proved that the number e is transcendental. In his paper he explains in a very clear way how he found his proof. He starts with an analogy between simultaneous diophantine approximation of real numbers on the one hand and analytic complex functions of one variable on the other. He first solves the analytic problem by constructing explicitly what is now called Padé approximants for the exponential function. In fact there are two types of such approximants, they are now called type I and type II, and what Hermite did in 1873 was to compute Padé approximants of type II. He also found those of type I in 1873 and studied them later in 1893. K. Mahler was the first in the mid's 1930 to relate the properties of the two types of Padé's approximants and to use those of type I in order to get a new proof of Hermite's transcendence Theorem (and also of the generalisation by Lindemann and Weierstraß as well as quantitative refinements). See [2] Chap. 2 § 3.

In the analogy with number theory, Padé approximants of type II are related with the simultaneous approximation of real numbers $\vartheta_1, \dots, \vartheta_m$ by rational numbers p_i/q with the same denominator q (one does not require that the fractions are irreducible), which means that we wish to bound from below

$$\max_{1\le i\le m}\left|\vartheta_i - \frac{p_i}{q}\right|$$

in terms of q, while type I is related with the study of lower bounds for linear combinations

$$|a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m|$$

when $a_0, \dots, a_m$ are rational integers, not all of which are 0, in terms of the number $\max_{0\le i\le m}|a_i|$.

After Hermite's seminal work, F. Lindemann was able to extend the argument and to prove the transcendence of π (hence he solved the old greek problem

[8] http://www.math.jussieu.fr/∼miw/articles/pdf/AWSLecture2.pdf

of the quadrature of the circle: *it is not possible using ruler and compass to draw a square and a circle having the same area*). This extension led to the so-called Hermite-Lindemann's Theorem:

Theorem 2.1 (Hermite–Lindemann)**.** *Let α be a non–zero complex algebraic number. Let $\log\alpha$ be any non-zero logarithm of α. Then $\log\alpha$ is transcendental.*

Equivalently, let β be a non-zero algebraic number. Then e^{β} is transcendental.

Recall that any non-zero complex number z has complex logarithms: these are the solutions $\ell \in \mathbb{C}$ of the equation $e^{\ell} = z$. If ℓ is one of them, then all solutions ℓ to this equation $e^{\ell} = z$ are $\ell + 2ik\pi$ with $k \in \mathbb{Z}$. The only non-zero complex of which 0 is a logarithm is 1.

The equivalence between both statements in Theorem 2.1 is easily seen by setting $e^{\beta} = \alpha$: one can phrase the result by saying that for any non-zero complex number β, one at least of the two numbers β, e^{β} is transcendental.

After the proofs by Hermite and Lindemann, a number of authors in the XIX-th century worked out variants of the argument. The main goal was apparently to get the shorter possible proof, and most often the reason for which it works is by no means so clear as in Hermite's original version. One can find in the literature such short proofs (see for instance [10]), the connection with Hermite's arguments are most often not so transparent. So we shall come back to the origin and try to explain what is going on.

We concentrate now on Hermite's proof for the transcendence of e. The goal is to prove that for any positive integer m, the numbers $1, e, e^2, \ldots, e^m$ are linearly independent over $\mathbb{Q}$.

2.1 Introduction to Hermite's work

The proofs given in subsection 1.2 of the irrationality of e^r for several rational values of r (namely $r \in \{1/a, 2/a, \sqrt{2}/a, \sqrt{3}/a \ ; \ a \in \mathbb{Z}, \ a \neq 0\}$) are similar: the idea is to start from the expansion of the exponential function, to truncate it and to deduce rational approximations to e^r. In terms of the exponential function this amounts to approximate e^z by a polynomial. The main idea, due to C. Hermite [6], is to approximate e^z by rational functions $A(z)/B(z)$. The word "approximate" has the following meaning (Hermite-Padé): an analytic function is *well approximated* by a rational function $A(z)/B(z)$ (where A and B are polynomial) if the difference $B(z)f(z) - A(z)$ has a zero at the origin of high multiplicity.

When we just truncate the series expansion of the exponential function, we approximate e^z by a polynomial in z with rational coefficients; when we substitute $z = a$ where a is a positive integer, this polynomial produces a rational number, but the denominator of this number is quite large (unless $a = \pm 1$). A trick gave the result also for $a = \pm 2$, but definitely for a a larger prime number for instance there is a problem: if we multiply by the denominator then the "remainder" is by no means small. As shown by Hermite, to produce a sufficiently large gap in the power expansion of $B(z)e^z$ will solve this problem.

Our first goal in this section is to prove Lambert's result on the irrationality of e^r when r is a non-zero rational number. Next we show how a slight modification implies the irrationality of π.

This proof serves as an introduction to Hermite's method. There are slightly different ways to present it: one is Hermite's original paper, another one is Siegel more algebraic point of view [12], and another was derived by Yu. V.Ñesterenko for [2] (unpublished manuscript).

2.1.1 Irrationality of e^r for $r \in \mathbb{Q}$: sketch of proof

If $r = a/b$ is a rational number such that e^r is also rational, then $e^{|a|}$ is also rational, and therefore the irrationality of e^r for any non-zero rational number r follows from the irrationality of e^a for any positive integer a. We shall approximate the exponential function e^z by a rational function $A(z)/B(z)$ and show that $A(a)/B(a)$ is a good rational approximation to e^a, sufficiently good in fact so that one may use Lemma 1.13.

Write

$$e^z = \sum_{k\geq 0} \frac{z^k}{k!}\cdot$$

We wish to multiply this series by a polynomial so that the Taylor expansion at the origin of the product $B(z)e^z$ has a large gap: the polynomial preceding the gap will be $A(z)$, the remainder $R(z) = B(z)e^z - A(z)$ will have a zero of high multiplicity at the origin.

In order to create such a gap, we shall use the differential equation of the exponential function - hence we introduce derivatives.

In Fourier's proof, we use for B a constant polynomial, of degree 0. For N sufficiently large set

$$B_N = N!, \quad A_N(z) = \sum_{n=0}^{N} \frac{N!}{n!} z^n, \quad R_N(z) = \sum_{n\geq N+1} \frac{N!}{n!} z^n.$$

Notice that the first term in the Taylor expansion of R_N is

$$\frac{1}{N+1} z^{N+1}.$$

This is sufficient for proving the irrationality of e, since for $z = 1$ we have

$$\lim_{N\to\infty} R_N(1) = 0.$$

But for $a > 1$ the sequence $(R_N(a))_{N\geq 1}$ tends to infinity.

Now take for B_N a degree 1 polynomial in $\mathbb{Z}[z]$ that we select so that the coefficient of z^N vanishes. It is easy to check that the solution is to take a multiple of $z - N$, and we take the product by $(N-1)!$ in order to have integral coefficients for A. So set

$$B_N(z) = (N-1)!z - N!, \quad A_N(z) = -N! - \sum_{n=1}^{N-1} \frac{(N-1)!}{n!}(N-n)z^n,$$

$$R_N(z)=\sum_{n\geq N+1}\frac{(N-1)!}{n!}(n-N)z^n$$

so that again $B_N(z)e^z=A_N(z)+R_N(z)$. Here the first term in the Taylor expansion of R_N is

$$\frac{1}{N(N+1)}z^{N+1}.$$

This is a tiny progress, since in the denominator we get a degree 2 polynomial in place of a degree 1 polynomial in N. But this is not sufficient to ensure that for fixed $a>1$ the sequence $(R_N(a))_{N\geq 1}$ tends to zero. So we shall take for B_N a polynomial of larger degree, depending on N.

2.1.2 First introduction to Hermite's proof

We first explain how to produce, from an analytic function whose Taylor development at the origin is

$$f(z)=\sum_{k\geq 0}a_kz^k, \tag{2.2}$$

another analytic function with one given Taylor coefficient, say the coefficient of z^m, is zero. The coefficient of z^m for f is $a_m=m!f^{(m)}(0)$. The same number a_m occurs when one computes the Taylor coefficient of z^{m-1} for the derivative f' of f. Writing

$$ma_m=m!(zf')^{(m)}(0),$$

we deduce that the coefficient of z^m in the Taylor development of $zf'(z)-mf(z)$ is 0, which is what we wanted.

It is the same thing to write

$$zf'(z)=\sum_{k\geq 0}ka_kz^k$$

so that

$$zf'(z)-mf(z)=\sum_{k\geq 0}(k-m)a_kz^k.$$

Now we want that several consecutive Taylor coefficients cancel. It will be convenient to introduce derivative operators.

We start with $D=d/dz$. As usual D^2 denotes $D\circ D$ and $D^m=D^{m-1}\circ D$ for $m\geq 2$. The derivation D and the multiplication by z do not commute:

$$D(zf)=f+zD(f),$$

relation which we write $Dz=1+zD$. From this relation it follows that the non-commutative ring generated by z and D over $\mathbb{C}$ is also the ring of polynomials in D with coefficients in $\mathbb{C}[z]$. In this ring $\mathbb{C}[z][D]$ there is an element which will be very useful for us, namely $\delta=zd/dz$. It satisfies $\delta(z^k)=kz^k$. To any polynomial $T\in\mathbb{C}[t]$ one associate the derivative operator $T(\delta)$.

By induction on m one checks $\delta^m z^k = k^m z^k$ for all $m \geq 0$. By linearity, one deduces that if T is a polynomial with complex coefficients, then

$$T(\delta) z^k = T(k) z^k.$$

For our function f with the Taylor development (2.2) we have

$$T(\delta) f(z) = \sum_{k \geq 0} a_k T(k) z^k.$$

Hence if we want a function with a Taylor expansion having 0 as coefficient of z^k, it suffices to consider $T(\delta) f(z)$ where T is a polynomial satisfying $T(k) = 0$. For instance if n_0 and n_1 are two non-negative integers and if we take

$$T(t) = (t - n_0 - 1)(t - n_0 - 2) \cdots (t - n_0 - n_1),$$

then the series $T(\delta) f(z)$ can be written $A(z) + R(z)$ with

$$A(z) = \sum_{k=0}^{n_0} T(k) a_k z^k$$

and

$$R(z) = \sum_{k \geq n_0 + n_1 + 1} T(k) a_k z^k.$$

This means that in the Taylor expansion at the origin of $T(\delta) f(z)$, all coefficients of $z^{n_0+1}, z^{n_0+2}, \ldots, z^{n_0+n_1}$ are 0.

Let $n_0 \geq 0$, $n_1 \geq 0$ be two integers. Define $N = n_0 + n_1$ and

$$T(t) = (t - n_0 - 1)(t - n_0 - 2) \cdots (t - N).$$

Since T is monic of degree n_1 with integer coefficients, it follows from the differential equation of the exponential function

$$\delta(e^z) = z e^z$$

that there is a polynomial $B \in \mathbb{Z}[z]$, which is monic of degree n_1, such that $T(\delta) e^z = B(z) e^z$.

Set

$$A(z) = \sum_{k=0}^{n_0} T(k) \frac{z^k}{k!} \quad \text{and} \quad R(z) = \sum_{k \geq N+1} T(k) \frac{z^k}{k!}\cdot$$

Then

$$B(z) e^z = A(z) + R(z),$$

where A is a polynomial with rational coefficients of degree n_0 and leading coefficient

$$\frac{T(n_0)}{n_0!} = (-1)^{n_1} \frac{n_1!}{n_0!}\cdot$$

Also the analytic function R has a zero of multiplicity $\geq N+1$ at the origin.

We can explicit these formulae for A and R. For $0 \leq k \leq n_0$ we have

$$\begin{aligned} T(k) &= (k-n_0-1)(k-n_0-2)\cdots(k-N) \\ &= (-1)^{n_1}(N-k)\cdots(n_0+2-k)(n_0+1-k) \\ &= (-1)^{n_1}\frac{(N-k)!}{(n_0-k)!}\cdot \end{aligned}$$

For $k \geq N+1$ we write in a similar way

$$T(k) = (k-n_0-1)(k-n_0-2)\cdots(k-N) = \frac{(k-n_0-1)!}{(k-N-1)!}\cdot$$

Hence we have proved:

Proposition 2.3 (Hermite's formulae for the exponential function)**.** *Let $n_0 \geq 0$, $n_1 \geq 0$ be two integers. Define $N = n_0 + n_1$. Set*

$$A(z) = (-1)^{n_1}\sum_{k=0}^{n_0} \frac{(N-k)!}{(n_0-k)!k!}\cdot z^k \quad \textit{and} \quad R(z) = \sum_{k\geq N+1} \frac{(k-n_0-1)!}{(k-N-1)!k!}\cdot z^k.$$

Finally, define $B \in \mathbb{Z}[z]$ by the condition

$$(\delta-n_0+1)(\delta-n_0+2)\cdots(\delta-N)e^z = B(z)e^z.$$

Then

$$B(z)e^z = A(z) + R(z).$$

Further, B is a monic polynomial with integer coefficients of degree n_1, A is a polynomial with rational coefficients of degree n_0 and leading coefficient $(-1)^{n_1}n_1!/n_0!$, and the analytic function R has a zero of multiplicity $N+1$ at the origin.
Furthermore, the polynomial $(n_0!/n_1!)A$ has integer coefficients. In particular, if $n_1 \geq n_0$, then the coefficients of A itself are integers.

Proof. It remains only to check the last assertion on the integrality of the coefficients of A for $n_1 \geq n_0$. Indeed when $n_1 \geq n_0$ each coefficient of the polynomial A is an integral multiple of a binomial coefficient:

$$\frac{(N-k)!}{(n_0-k)!k!} = (N-k)(N-k-1)\cdots(n_0+1)\cdot\frac{n_0!}{(n_0-k)!k!}$$

for $0 \leq k \leq n_0$. Hence $A \in \mathbb{Z}[z]$. □

We now restrict to the case $n_0 = n_1$ and we set $n = n_0 = n_1$. We write also

$$T_n(z) = (z-n-1)(z-n-2)\cdots(z-2n)$$

and we denote by A_n, B_n and R_n the Hermite polynomials and the remainder in Hermite's Proposition 2.3.

Remark. *For $n_1 < n_0$ the leading coefficient $(-1)^{n_1} n_1!/n_0!$ of A is not an integer.*

Lemma 2.4. *Let $z \in \mathbb{C}$. Then*

$$|R_n(z)| \leq \frac{|z|^{2n+1}}{n!} e^{|z|}.$$

In particular the sequence $(R_n(z))_{n\geq 0}$ tends to 0 as n tends to infinity.

Proof. We have

$$R_n(z) = \sum_{k\geq 2n+1} \frac{(k-n-1)!}{(k-2n-1)!k!} \cdot z^k = \sum_{\ell\geq 0} \frac{(\ell+n)!}{(\ell+2n+1)!} \cdot \frac{z^{\ell+2n+1}}{\ell!}.$$

The trivial estimates

$$\frac{(\ell+2n+1)!}{(\ell+n)!} = (\ell+2n+1)(\ell+2n)(\ell+2n-1)\cdots(\ell+n+1)ge(n+1)! \geq n!$$

yields

$$|R_n(z)| \leq \frac{|z|^{2n+1}}{n!} \sum_{\ell\geq 0} \frac{|z|^\ell}{\ell!}.$$

Lemma 2.4 follows. □

2.1.3 Second introduction to Hermite's proof

In [12] C.L. Siegel introduces an algebraic point of view which yields the following:

Theorem 2.5. *Given two integers $n_0 \geq 0$, $n_1 \geq 0$, there exist two polynomials A and B in $\mathbb{C}[z]$ with A of degree $\leq n_0$ and $B \neq 0$ of degree $\leq n_1$ such that the function $R(z) = B(z)e^z - A(z)$ has a zero at the origin of multiplicity $\geq N+1$ with $N = n_0+n_1$. This solution (A, B, R) is unique if we require B to be monic. Moreover A has degree n_0, B has degree n_1 and R has multiplicity $N+1$ at the origin.*

We denote by D the derivation d/dz. When f is a complex valued function of one complex variable z, we shall sometimes write $D\bigl(f(z)\bigr)$ in place of Df. For instance $D(zf) = f + zDf$. We write as usual D^2 for $D \circ D$ and $D^\ell = D \circ D^{\ell-1}$. The Taylor expansion at the origin of an analytic function f is

$$f(z) = \sum_{\ell\geq 0} \frac{1}{\ell!} D^\ell f(0) z^\ell.$$

Proof. We first prove the existence of a non-trivial solution (A, B, R). For $n \geq 0$ denote by $\mathbb{C}[z]_{\leq n}$ the $\mathbb{C}$–vector space of polynomials of degree $\leq n$. Its dimension is $n+1$. Consider the linear mapping

$$\begin{array}{cccc} \mathcal{L}: & \mathbb{C}[z]_{\leq n_1} & \longrightarrow & \mathbb{C}^{n_1} \\ & B(z) & \longmapsto & \Bigl(D^\ell\bigl(B(z)e^z\bigr)_{z=0}\Bigr)_{n_0<\ell\leq N} \end{array}$$

This map is not injective, its kernel has dimension ≥ 1. Let $B \in \ker \mathcal{L}$. Define

$$A(z) = \sum_{\ell=0}^{n_0} D^\ell \bigl(B(z)e^z\bigr)_{z=0} \frac{z^\ell}{\ell!}$$

and

$$R(z) = \sum_{\ell \geq N+1} D^\ell \bigl(B(z)e^z\bigr)_{z=0} \frac{z^\ell}{\ell!}.$$

Then (A, B, R) is a solution to the problem:

$$B(z)e^z = A(z) + R(z). \tag{2.6}$$

There is an alternative proof of the existence as follows [12]. Consider the linear mapping

$$\begin{array}{ccc} \mathbb{C}[z]_{\leq n_0} \times \mathbb{C}[z]_{\leq n_1} & \longrightarrow & \mathbb{C}^{N+1} \\ \bigl(A(z), B(z)\bigr) & \longmapsto & \Bigl(D^\ell \bigl(B(z)e^z\bigr)_{z=0}\Bigr)_{0 \leq \ell \leq N} \end{array}$$

This map is not injective, its kernel has dimension ≥ 1. If (A, B) is a non-zero element in the kernel, then $B \neq 0$.

We now check that the kernel of $\mathcal{L}$ has dimension 1. Let $B \in \ker \mathcal{L}$, $B \neq 0$ and let (A, B, R) be the corresponding solution to (2.6).

Since A has degree $\leq n_0$, the $(n_0 + 1)$-th derivative of R is

$$D^{n_0+1} R = D^{n_0+1}(B(z)e^z),$$

hence it is the product of e^z with a polynomial of the same degree as the degree of B and same leading coefficient. Now R has a zero at the origin of multiplicity $\geq n_0 + n_1 + 1$, hence $D^{n_0+1}R(z)$ has a zero of multiplicity $\geq n_1$ at the origin. Therefore

$$D^{n_0+1}R = cz^{n_1}e^z \tag{2.7}$$

where c is the leading coefficient of B; it follows also that B has degree n_1. This proves that $\ker \mathcal{L}$ has dimension 1.

Since $D^{n_0+1}R$ has a zero of multiplicity exactly n_1, it follows that R has a zero at the origin of multiplicity exactly $N+1$, so that R is the unique function satisfying $D^{n_0+1}R = cz^{n_1}e^z$ with a zero of multiplicity n_0 at 0.

It remains to check that A has degree n_0. Multiplying (2.6) by e^{-z}, we deduce

$$A(z)e^{-z} = B(z) - R(z)e^{-z}.$$

We replace z by $-z$:

$$A(-z)e^z = B(-z) - R(-z)e^z. \tag{2.8}$$

It follows that $\bigl(B(-z), A(-z), -R(-z)e^z\bigr)$ is a solution to the Padé problem (2.6) for the parameters (n_1, n_0), hence A has degree n_0.

□

Following [12], we give formulae for A, B and R.

Consider the operator J defined by

$$J(\varphi) = \int_0^z \varphi(t)dt.$$

It satisfies

$$DJ\varphi = \varphi \quad \text{and} \quad JDf = f(z) - f(0).$$

Hence the restriction of the operator of D to the functions vanishing at the origin is a one-to-one map with inverse J.

Lemma 2.9. *For $n \geq 0$,*

$$J^{n+1}\varphi = \frac{1}{n!}\int_0^z (z-t)^n \varphi(t)dt.$$

Proof. The formula is valid for $n = 0$. We first check it for $n = 1$. The derivative of the function

$$\int_0^z (z-t)\varphi(t)dt = z\int_0^z \varphi(t)dt - \int_0^z t\varphi(t)dt$$

is

$$\int_0^z \varphi(t)dt + z\varphi(z) - z\varphi(z) = \int_0^z \varphi(t)dt.$$

We now proceed by induction. For $n \geq 1$ the derivative of the function of z

$$\frac{1}{n!}\int_0^z (z-t)^n\varphi(t)dt = \sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} \cdot z^k \int_0^z t^{n-k}\varphi(t)dt$$

is

$$\sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!}\left(kz^{k-1}\int_0^z t^{n-k}\varphi(t)dt + z^n\varphi(z)\right). \tag{2.10}$$

Since $n \geq 1$, we have

$$\sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} = 0$$

and (2.10) is nothing else than

$$\sum_{k=1}^n \frac{(-1)^{n-k}}{(k-1)!(n-k)!} \cdot z^{k-1}\int_0^z t^{n-k}\varphi(t)dt = \frac{1}{(n-1)!}\int_0^z (z-t)^{n-1}\varphi(t)dt.$$

□

From (2.7) with $c = 1$ and Lemma 2.9 it plainly follows:

Lemma 2.11. *The remainder $R(z)$ in Hermite's fomula with parameters n_0 and n_1 (and B monic) is given by*

$$R(z) = \frac{1}{n_0!}\int_0^z (z-t)^{n_0} t^{n_1} e^t dt.$$

Replacing t by tz yields

$$R(z) = \frac{z^{N+1}}{n_0!}\int_0^1 (1-t)^{n_0} t^{n_1} e^{tz} dt.$$

Hence:

Lemma 2.12. *Let $z \in \mathbb{C}$. Then*

$$|R(z)| \leq \frac{|z|^{N+1}}{n_0!} e^{|z|}.$$

In particular for $n_0 = n_1 = n$, if we denote R by R_n, then the sequence $(R_n(z))_{n\geq 0}$ tends to 0 as n tends to infinity.

Remark. *This is the estimate for B monic. When $n_1 < n_0$ the coefficients of the associated polynomial A are not integers. For instance in case $n_1 = 0$ (hence $n_0 = N$) the polynomial B is 1 and A (which is the head of the Taylor expansion of the exponential function) has denominator $N!$. In case $n_1 = 1$ we need to multiply by $(N-1)!$, as explained above, to get integer coefficients. More generally in case $n_1 < n_0$ we need to multiply by $n_0!/n_1!$ in order to get integer coefficients, so the remainder in this case is bounded by*

$$\frac{n_0!}{n_1!}|R(z)| \leq \frac{|z|^{N+1}}{n_1!} e^{|z|}.$$

If we want to have a small remainder we need to take n_1 at least a constant times $N/\log N$. The choice $n_1 = n_0 = N/2$ is the most natural one.

We now give formulae for A and B in Theorem 2.5.

When $S \in \mathbb{C}[[t]]$ is a power series, say

$$S(t) = \sum_{i\geq 0} s_i t^i,$$

and f an analytic complex valued function, we define

$$S(D)f = \sum_{i\geq 0} s_i D^i f,$$

and we shall use this notation only when the sum is finite: either S is a polynomial in $\mathbb{C}[t]$ or f is a polynomial in $\mathbb{C}[z]$.

We reproduce [12], Chap.I § 1: for two powers series S_1 and S_2 and an analytic function f we have

$$\big(S_1(D) + S_2(D)\big)f = S_1(D)f + S_2(D)f$$

and

$$\bigl(S_1(D)S_2(D)\bigr)f = S_1(D)\bigl(S_2(D)\bigr)f.$$

Also if $s_0 \neq 0$ then the series S has an inverse in the ring $\mathbb{C}[[t]]$

$$S^{-1}(t) = \sum_{i\geq 0} t_i t^i, \quad (t_0 = 1/s_0)$$

and

$$S^{-1}(D)\bigl(S(D)f\bigr) = f.$$

If the power series S and the polynomial f have integer coefficients, then $S(D)f$ is also a polynomial with integer coefficients. The same holds also for $S^{-1}(D)f$ if, further, $s_0 = \pm 1$.

For $\lambda \in \mathbb{C}$ and $P \in \mathbb{C}[z]$, we have

$$D(e^{\lambda z}P) = e^{\lambda z}(\lambda + D)P.$$

Hence for $n \geq 1$,

$$D^n(e^{\lambda z}P) = e^{\lambda z}(\lambda + D)^n P$$

and $(\lambda + D)^n P$ is again a polynomial; further it has the same degree as P when $\lambda \neq 0$. Conversely, assuming $\lambda \neq 0$, given a polynomial $Q \in \mathbb{C}[z]$, the unique solution $P \in \mathbb{C}[z]$ to the differential equation

$$(\lambda + D)^n P = Q$$

is

$$P = (\lambda + D)^{-n} Q.$$

In the case $\lambda = \pm 1$, when Q has integer coefficients, then so does P.

We come back now to the solution (A, B, R) to the Padé problem (2.6) in Theorem 2.5, where $B \in \mathbb{C}[z]$ is monic of degree n_1 and $A \in \mathbb{C}[z]$ has degree n_0, while $R \in \mathbb{C}[[z]]$ has a zero of multiplicity $N + 1$ at 0.

From

$$D^{n_0+1}\bigl(B(z)e^z\bigr) = z^{n_1}e^z$$

we deduce

$$B(z) = (1 + D)^{-n_0-1} z^{n_1}.$$

From this formula it follows that B has integer coefficients. It is easy to explicit the polynomial B. From

$$(1 + D)^{-n_0-1} = \sum_{\ell\geq 0} (-1)^\ell \binom{n_0 + \ell}{\ell} D^\ell,$$

we deduce

$$B(z) = \sum_{\ell=0}^{n_1} (-1)^\ell \binom{n_0 + \ell}{\ell} \frac{n_1!}{(n_1 - \ell)!} z^{n_1 - \ell},$$

which can be written also as

$$B(z)=(-1)^{n_1}\frac{n_1!}{n_0!}\sum_{k=0}^{n_1}(-1)^k\frac{(N-k)!}{(n_1-k)!k!}z^k. \tag{2.13}$$

One checks that B is monic of degree n_1.

If B is monic then $c=1$ in (2.7) and it follows that the coefficient of z^{N+1} in R is

$$\frac{n_1!}{(N+1)!}\cdot$$

In the proof of Theorem 2.5, we found a link between the Padé solution with parameters (n_0,n_1) and the solution with parameters (n_1,n_0). We explicit this link. For that we denote by $(A_{n_0,n_1},B_{n_0,n_1},R_{n_0,n_1})$ the solution of (2.6) for the parameters (n_0,n_1). From (2.8) we infer

$$\begin{aligned}
A_{n_0,n_1}(z)&=(-1)^N\frac{n_1!}{n_0!}B_{n_1,n_0}(-z),\\
B_{n_0,n_1}(z)&=(-1)^N\frac{n_1!}{n_0!}A_{n_1,n_0}(-z),\\
R_{n_0,n_1}(z)&=(-1)^N\frac{n_1!}{n_0!}R_{n_1,n_0}(-z)e^{-z}.
\end{aligned}$$

Hence

$$A(z)=(-1)^{n_1}\sum_{k=0}^{n_0}\frac{(N-k)!}{(n_0-k)!k!}\cdot z^k. \tag{2.14}$$

The leading coefficient of A is $(-1)^{n_1}n_1!/n_0!$. It follows also from (2.14) that $(n_0!/n_1!)A$ has integer coefficients. In particular if $n_1\ge n_0$, then A is in $\mathbb{Z}[z]$.

We can also check this formula (2.14) starting from

$$D^{n_1+1}\bigl(A(z)e^{-z}\bigr)=-D^{n_1+1}\bigl(R(z)e^{-z}\bigr),$$

where the left hand side is the product of e^{-z} with a polynomial of degree $\le n_0$, while the right hand side has a multiplicity $\ge n_0$ at the origin. We deduce

$$D^{n_1+1}\bigl(A(z)e^{-z}\bigr)=az^{n_0}e^{-z}$$

where a is the leading coefficient of a. From

$$D^{n_1+1}\bigl(A(z)e^{-z}\bigr)=e^{-z}(-1+D)^{n_1+1}A(z)$$

we deduce

$$(-1+D)^{n_1+1}A(z)=-az^{n_0}$$

and

$$A(z)=-a(-1+D)^{-n_1-1}z^{n_0}.$$

We shall give another proof of the formulae (2.13) and (2.14), which provide also the next formula for R:

$$R(z) = \sum_{k\geq N+1} \frac{(k-n_0-1)!}{(k-N-1)!k!} \cdot z^k. \tag{2.15}$$

Lemma 2.12 also follows from (2.15). Indeed we have

$$R_n(z) = \sum_{k\geq 2n+1} \frac{(k-n-1)!}{(k-2n-1)!k!} \cdot z^k = \sum_{\ell\geq 0} \frac{(\ell+n)!}{(\ell+2n+1)!} \cdot \frac{|z|^{\ell+2n+1}}{\ell!}\cdot$$

The trivial upper bound

$$\prod_{j=n+1}^{n+\ell} j \leq \prod_{j=n+1}^{n+\ell} (j+n+1)$$

is equivalent to

$$\frac{(\ell+n)!}{(\ell+2n+1)!} \leq \frac{n!}{(2n+1)!},$$

hence

$$|R_n(z)| \leq \frac{n!|z|^{2n+1}}{(2n+1)!} \sum_{\ell\geq 0} \frac{|z|^\ell}{\ell!}\cdot$$

We bound $n!/(2n+1)!$ by $n!$: Lemma 2.12 follows.

2.1.4 Irrationality of e^r: end of the proof

We are now able to complete the proof of the irrationality of e^r for $\in \mathbb{Q}$, $r \neq 0$.

Let $r = a/b$ be a non-zero rational number. Assume first r is positive. Set $s = e^r$ and replace z by $a = br$ in the previous formulae; we deduce

$$B_n(a)s^b - A_n(a) = R_n(a).$$

All coefficients in R_n are positive, hence $R_n(a) > 0$. Therefore $B_n(a)s^b - A_n(a) \neq 0$. Since $R_n(a)$ tends to 0 when n tends to infinity and since $B_n(a)$ and $A_n(a)$ are rational integers, we may use the implication (ii)⇒(i) in Lemma 1.13: we deduce that the number s^b is irrational. As we already saw this readily implies that $s = e^r$ and $s^{-1} = e^{-r}$ are irrational.

2.1.5 Irrationality of π

This proof of the irrationality of $\log s$ for s a positive rational number given in § 2.1.4 can be extended to the case $s = -1$ in such a way that one deduces Lambert's result (see § 1.1.2) on the irrationality of the number π.

Assume π is a rational number, $\pi = a/b$. Substitute $z = ia = i\pi b$ in the previous formulae. Notice that $e^z = (-1)^b$:

$$B_n(ia)(-1)^b - A_n(ia) = R_n(ia),$$

and that the two complex numbers $A_n(ia)$ and $B_n(ia)$ are in $\mathbb{Z}[i]$. The left hand side is in $\mathbb{Z}[i]$, the right hand side tends to 0 as n tends to infinity, hence both sides are 0.

In the proof of § 2.1.1 we used the positivity of the coefficients of R_n and we deduced that $R_n(a)$ was not 0 (this is the so-called "zero estimate" in transcendental number theory). Here we need another argument.

The last step of the proof of the irrationality of π is achieved by using two consecutive indices n and $n+1$. We eliminate e^z among the two relations

$$B_n(z)e^z - A_n(z) = R_n(z) \quad \text{and} \quad B_{n+1}(z)e^z - A_{n+1}(z) = R_{n+1}(z).$$

We deduce that the polynomial

$$\Delta_n = B_n A_{n+1} - B_{n+1} A_n \tag{2.16}$$

can be written

$$\Delta_n = -B_n R_{n+1} + B_{n+1} R_n. \tag{2.17}$$

As we have seen, the polynomial B_n is monic of degree n; the polynomial A_n also has degree n, its highest degree term is $(-1)^n z^n$. It follows from (2.16) that Δ_n is a polynomial of degree $2n+1$ and highest degree term $(-1)^n 2z^{2n+1}$. On the other hand since R_n has a zero of multiplicity at least $2n+1$, the relation (2.17) shows that it is the same for Δ_n. Consequently

$$\Delta_n(z) = (-1)^n 2z^{2n+1}.$$

We deduce that Δ_n does not vanish outside 0. From (2.17) we deduce that R_n and R_{n+1} have no common zero apart from 0. This completes the proof of the irrationality of π.

2.2 Transcendence of e, following Hermite

2.2.1 Padé approximants

Henri Eugène Padé (1863–1953), who was a student of Charles Hermite (1822–1901), gave his name to the following objects which he studied thoroughly in his thesis in 1892.

Lemma 2.18. *Let $f_1, \dots, f_m$ be analytic functions of one complex variable near the origin. Let $n_0, n_1, \dots, n_m$ be non-negative integers. Set*

$$N = n_0 + n_1 + \cdots + n_m.$$

Then there exists a tuple $(Q, P_1, \dots, P_m)$ of polynomials in $\mathbb{C}[X]$ satisfying the following properties:
(i) The polynomial Q is not zero, it has degree $\le N - n_0$.
(ii) For $1 \le \mu \le m$, the polynomial P_μ has degree $\le N - n_\mu$.
(iii) For $1 \le \mu \le m$, the function $x \mapsto Q(x)f_\mu(x) - P_\mu(x)$ has a zero at the origin of multiplicity $\ge N+1$.

Definition. *A tuple $(Q, P_1, \dots, P_m)$ of polynomials in $\mathbb{C}[X]$ satisfying the condition of Lemma 2.18 is called a* Padé system of the second type for $(f_1, \dots, f_m)$ attached to the parameters $n_0, n_1, \dots, n_m$.

Proof. The polynomial Q of Lemma 2.18 should have degree $\leq N - n_0$, so we have to find (or rather to prove the existence of) its $N - n_0 + 1$ coefficients, not all being zero. We consider these coefficients as unknowns. The property we require is that for $1 \leq \mu \leq m$, the Taylor expansion at the origin of $Q(x)f_\mu(x)$ has zero coefficients for $x^{N-n_\mu+1}, x^{N-n_\mu+1}, \dots, x^N$. If this property holds for $1 \leq \mu \leq m$, we shall define P_μ by truncating the Taylor series at the origin of $Q(x)f_\mu(x)$ at the rank x^{N-n_μ}, hence P_μ will have degree $\leq N - n_\mu$, while the remainder $Q(x)f_\mu(x) - P_\mu(x)$ will have a mutiplicity $\geq N + 1$ at the origin.

Now for each given μ the condition we stated amounts to require that our unknowns (the coefficients of Q) satisfy n_μ homogeneous linear relations, namely

$$\left(\frac{d}{dx}\right)^k [Q(x)f_\mu(x)]_{x=0} = 0 \quad \text{for} \quad N - n_\mu < k \leq N.$$

Therefore altogether we get $n_1 + \cdots + n_m = N - n_0$ homogeneous linear equations, and since the number $N - n_0 + 1$ of unknowns (the coefficients of Q) is larger, linear algebra tells us that a non-trivial solution exists.

□

There is no unicity, because of the homogeneity of the problem: the set of solutions (together with the trivial solution 0) is a vector space over $\mathbb{C}$, and Lemma 2.18 tells us that it has positive dimension. In the case where this dimension is 1 (which means that there is unicity up to a multiplicative factor), the system of approximants is called *perfect*. An example is with $m = 1$ and $f(x) = e^x$, as shown by Hermite's work (see § 5.3).

We discuss briefly *Padé approximants of type I* in § 4.1.

Most often it is not easy to find explicit solutions: we only know their existence. As we are going to show, Hermite succeeded to produce explicit solutions for the systems of Padé approximants of the functions $(e^x, e^{2x}, \dots, e^{mx})$.

2.2.2 Hermite's proof of the transcendence of e

Hermite gave explicit formulae for solving the Padé problem for the exponential function, and he deduced the transcendence of e as follows. The next formula is one of the many disguises of what is called *Hermite's identity.*

Lemma 2.19. *Let f be a polynomial of degree $\leq N$. Define*

$$F = f + Df + D^2 + \cdots + D^N f.$$

Then for $z \in \mathbb{C}$

$$\int_0^z e^{-t} f(t)dt = F(0) - e^{-z}F(z).$$

We can also write the definition of F as

$$F = (1-D)^{-1}f \quad \text{where} \quad (1-D)^{-1} = \sum_{k\geq 0} D^k.$$

The series in the right hand side is infinite, but when we apply the operator to a polynomial only finitely many $D^k f$ are not 0: when f is a polynomial of degree $\leq N$ then $D^k f = 0$ for $k > N$.

Proof. More generally, if f is a complex function which is analytic at the origin and N is a positive integer, if we set

$$F = f + Df + D^2 + \cdots + D^N f,$$

then the derivative of $e^{-t}F(t)$ is $-e^{-t}f(t) + e^{-t}D^{N+1}f(t)$. □

Let f be a polynomial. Hermite's Lemma 2.19 gives a formula for

$$\int_0^z e^{-t}f(t)dt$$

for $z \in \mathbb{C}$. A change of variables leads to a formula for

$$\int_0^u e^{-xt}f(t)dt$$

when x and u are complex numbers. Here, in place of using Lemma 2.19, we repeat the proof. Integrate by part $e^{-xt}f(t)$ between 0 and u:

$$\int_0^u e^{-xt}f(t)dt = -\left[\frac{1}{x}e^{-xt}f(t)\right]_0^u + \frac{1}{x}\int_0^u e^{-xt}f'(t)dt.$$

By induction we deduce

$$\int_0^u e^{-xt}f(t)dt = -\sum_{k=0}^{m}\left[\frac{1}{x^{k+1}}e^{-xt}D^k f(t)\right]_0^u + \frac{1}{x^{m+1}}\int_0^u e^{-xt}D^{m+1}f(t)dt.$$

Let N be an upper bound for the degree of f. For $m = N$ the last integral vanishes and

$$\begin{aligned}\int_0^u e^{-xt}f(t)dt &= -\sum_{k=0}^{N}\left[\frac{1}{x^{k+1}}e^{-xt}D^k f(t)\right]_0^u\\ &= \sum_{k=0}^{N}\frac{1}{x^{k+1}}D^k f(0) - e^{-xu}\sum_{k=0}^{N}\frac{1}{x^{k+1}}D^k f(u).\end{aligned}$$

Multipling by $x^{N+1}e^{ux}$ yields:

Lemma 2.20. *Let f be a polynomial of degree $\leq N$ and let x, u be complex numbers. Then*

$$e^{xu}\sum_{k=0}^{N} x^{N-k}D^k f(0) = \sum_{k=0}^{N} x^{N-k}D^k f(u) + x^{N+1}e^{xu}\int_0^u e^{-xt}f(t)dt.$$

With the notation of Lemma 2.20, the function

$$x \mapsto \int_0^u e^{-xt}f(t)dt$$

is analytic at $x = 0$, hence its product with x^{N+1} has a mutiplicity $\geq N+1$ at the origin. Moreover

$$Q(x) = \sum_{k=0}^{N} x^{N-k}D^k f(0) \quad \text{and} \quad P(x) = \sum_{k=0}^{N} x^{N-k}D^k f(u)$$

are polynomials in x.

If the polynomial f has a zero of multiplicity $\geq n_0$ at the origin, then Q has degree $\leq N - n_0$. If the polynomial f has a zero of multiplicity $\geq n_1$ at u, then P has degree $\leq N - n_1$.

For instance in the case $u = 1$, $N = n_0 + n_1$, $f(t) = t^{n_0}(t-1)^{n_1}$, the two polynomials

$$Q(x) = \sum_{k=n_0}^{N} x^{N-k}D^k f(0) \quad \text{and} \quad P(x) = \sum_{k=n_1}^{N} x^{N-k}D^k f(1)$$

satisfy the properties which were required in section §2.1.1 (see Proposition 2.3), namely $R(z) = Q(z)e^z - P(z)$ has a zero of multiplicity $> n_0 + n_1$ at the origin, P has degree $\leq n_0$ and Q has degree $\leq n_1$.

Lemma 2.20 is a powerful tool to go much further.

Proposition 2.21. *Let m be a positive integer, $n_0, \ldots, n_m$ be non-negative integers. Set $N = n_0 + \cdots + n_m$. Define the polynomial $f \in \mathbb{Z}[t]$ of degree N by*

$$f(t) = t^{n_0}(t-1)^{n_1}\cdots(t-m)^{n_m}.$$

Further set, for $1 \leq \mu \leq m$,

$$Q(x) = \sum_{k=n_0}^{N} x^{N-k}D^k f(0), \quad P_\mu(x) = \sum_{k=n_\mu}^{N} x^{N-k}D^k f(\mu)$$

and

$$R_\mu(x) = x^{N+1}e^{x\mu}\int_0^\mu e^{-xt}f(t)dt.$$

Then the polynomial Q has exact degree $N - n_0$, while P_μ has exact degree $N - n_\mu$, and R_μ is an analytic function having at the origin a multiplicity $\geq N + 1$. Further, for $1 \leq \mu \leq m$,

$$Q(x)e^{\mu x} - P_\mu(x) = R_\mu(x).$$

Hence $(Q, P_1, \ldots, P_m)$ is a Padé system of the second type for the m-tuple of functions $(e^x, e^{2x}, \ldots, e^{mx})$, attached to the parameters $n_0, n_1, \ldots, n_m$. Furthermore, the polynomials $(1/n_0!)Q$ and $(1/n_\mu!)P_\mu$ for $1 \leq \mu \leq m$ have integral coefficients.

These polynomials $Q, P_1, \ldots, P_m$ are called the *Hermite-Padé polynomials attached to the parameters $n_0, n_1, \ldots, n_m$.*

For $h \geq 0$, the h-th derivative $D^h R(z)$ of the remainder in Proposition 2.21 is given by

$$D^h R(z) = \sum_{k \geq N+1} \frac{(k - n_0 - 1)!}{(k - N - 1)!} \cdot \frac{z^{k-h}}{(k-h)!}.$$

In particular for $h = n_0 + 1$ the formula becomes

$$D^{n_0+1} R = \sum_{k \geq N+1} \frac{z^{k-n_0-1}}{(k - N - 1)!} = z^{n_1} e^z. \tag{2.22}$$

This relations determines R since R has a zero of multiplicity $\geq n_0 + 1$ at the origin.

Proof. The coefficient of x^{N-n_0} in the polynomial Q is $D^{n_0} f(0)$, so it is not zero since f has mutiplicity exactly n_0 at the origin. Similarly for $1 \leq \mu \leq m$ the coefficient of x^{N-n_μ} in P_μ is $D^{n_0} f(\mu) \neq 0$.

The assertion on the integrality of the coefficients follows from the next lemma.

Lemma 2.23. *Let f be a polynomial with integer coefficients and let k be a non-negative integer. Then the polynomial $(1/k!)D^k f$ has integer coefficients.*

Proof. If $f(X) = \sum_{n \geq 0} a_n X^n$ then

$$\frac{1}{k!} D^k f = \sum_{n \geq 0} a_n \binom{n}{k} X^n \quad \text{with} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!},$$

and the binomial coefficients are rational integers. □

From Lemma 2.23 it follows that for any polynomial $f \in \mathbb{Z}[X]$ and for any integers k and n with $n \geq k$, the polynomial $(1/k!)D^n f$ also belongs to $\mathbb{Z}[X]$. This completes the proof of Proposition 2.21.

□

In order to complete the proof of the transcendence of e, we shall substitute 1 to x in the relations

$$Q(x)e^{\mu x} = P_\mu(x) + R_\mu(x)$$

and deduce simultaneous rational approximations $(p_1/q, p_2/q, \dots, p_m/q)$ to the numbers $e, e^2, \dots, e^m$. In order to use Lemma 1.29, we need to have independent such approximations. This is a subtle point which Hermite did not find easy to overcome, according to his owns comments in [6]. The following approach is due to K. Mahler, we can view it as an extension of the simple non-vanishing argument used in § 2.1.5 for the irrationality of π.

We fix integers $n_0, \dots, n_1$, all ≥ 1. For $j = 0, 1, \dots, m$ we denote by $Q_j, P_{j1}, \dots, P_{jm}$ the Hermite-Padé polynomials attached to the parameters

$$n_0 - \delta_{j0}, n_1 - \delta_{j1}, \dots, n_m - \delta_{jm},$$

where δ_{ji} is Kronecker's symbol

$$\delta_{ji} = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{if } j \neq i. \end{cases}$$

These parameters are said to be *contiguous* to $n_0, n_1, \dots, n_m$. They are the rows of the matrix

$$\begin{pmatrix} n_0 - 1 & n_1 & n_2 & \cdots & n_m \\ n_0 & n_1 - 1 & n_2 & \cdots & n_m \\ \vdots & \vdots & \ddots & \vdots & \\ n_0 & n_1 & n_2 & \cdots & n_m - 1 \end{pmatrix}.$$

Proposition 2.24. *There exists a non-zero constant c such that the determinant*

$$\Delta(x) = \begin{vmatrix} Q_0(x) & P_{10}(x) & \cdots & P_{m0}(x) \\ \vdots & \vdots & \ddots & \vdots \\ Q_m(x) & P_{1m}(x) & \cdots & P_{mm}(x) \end{vmatrix}$$

is the monomial cx^{mN}.

Proof. The matrix of degrees of the entries in the determinant defining Δ is

$$\begin{pmatrix} N - n_0 & N - n_1 - 1 & \cdots & N - n_m - 1 \\ N - n_0 - 1 & N - n_1 & \cdots & N - n_m - 1 \\ \vdots & \vdots & \ddots & \vdots \\ N - n_0 - 1 & N - n_1 - 1 & \cdots & N - n_m \end{pmatrix}.$$

Therefore Δ is a polynomial of exact degree $N - n_0 + N - n_1 + \cdots + N - n_m = mN$, the leading coefficient arising from the diagonal. This leading coefficient is $c = c_0 c_1 \cdots c_m$, where c_0 is the leading coefficient of Q_0 and c_μ is the leading coefficient of $P_{\mu\mu}$, $1 \leq \mu \leq m$.

It remains to check that Δ has a multiplicity at least mN at the origin. Linear combinations of the columns yield

$$\Delta(x) = \begin{vmatrix} Q_0(x) & P_{10}(x) - e^x Q_0(x) & \cdots & P_{m0}(x) - e^{mx} Q_0(x) \\ \vdots & \vdots & \ddots & \vdots \\ Q_m(x) & P_{1m}(x) - e^x Q_m(x) & \cdots & P_{mm}(x) - e^{mx} Q_m(x) \end{vmatrix}.$$

Each $P_{\mu j}(x) - e^{\mu x} Q_j(x)$, $1 \leq \mu \leq m$, $0 \leq j \leq m$, has multiplicity at least N at the origin, because for each contiguous triple $(1 \leq j \leq m)$ we have

$$\sum_{i=0}^{m} (n_i - \delta_{ji}) = n_0 + n_1 + \cdots + n_m - 1 = N - 1.$$

Looking at the multiplicity at the origin, we can write

$$\Delta(x) = \begin{vmatrix} Q_0(x) & \mathcal{O}(x^N) & \cdots & \mathcal{O}(x^N) \\ \vdots & \vdots & \ddots & \vdots \\ Q_m(x) & \mathcal{O}(x^N) & \cdots & \mathcal{O}(x^N) \end{vmatrix}.$$

This completes the proof of Proposition 2.24. □

Now we fix a sufficiently large integer n and we use the previous results for $n_0 = n_1 = \cdots = n_m = n$ with $N = (m+1)n$. We define, for $0 \leq j \leq m$, the integers $q_j, p_{1j}, \ldots, p_{nj}$ by

$$(n-1)! q_j = Q_j(1), \ (n-1)! p_{\mu j} = P_{\mu j}(1), \quad (1 \leq \mu \leq m).$$

Proposition 2.25. *There exists a constant $\kappa > 0$ independent on n such that for $1 \leq \mu \leq m$ and $0 \leq j \leq m$,*

$$|q_i e^{\mu} - p_{\mu j}| \leq \frac{\kappa^n}{n!}.$$

Further, the determinant

$$\begin{vmatrix} q_0 & p_{10} & \cdots & p_{m0} \\ \vdots & \vdots & \ddots & \vdots \\ q_m & p_{1m} & \cdots & p_{mm} \end{vmatrix}$$

is not zero.

Proof. Recall Hermite's formulae in Proposition 2.21:

$$Q_j(x) e^{\mu x} - P_{\mu j}(X) = x^{mn} e^{\mu x} \int_0^{\mu} e^{-xt} f_j(t) dt, \quad (1 \leq \mu \leq m, \ 0 \leq j \leq m),$$

where

$$\begin{aligned} f_j(t) &= (t-j)^{-1} \big(t(t-1) \cdots (t-m) \big)^n \\ &= (t-j)^{n-1} \prod_{\substack{1 \leq i \leq m \\ i \neq j}} (t-i)^n. \end{aligned}$$

We substitute 1 to x and we divide by $(n-1)!$:

$$q_j e^\mu - p_{\mu j} = \frac{1}{(n-1)!}\bigl(Q_j(1)e^\mu - P_{\mu j}(1)\bigr) = \frac{e^\mu}{(n-1)!}\int_0^\mu e^{-t} f_j(t)dt.$$

Now the integral is bounded from above by

$$\int_0^\mu e^{-t}|f_j(t)|dt \le m \sup_{0\le t\le m} |f_j(t)| \le m^{1+(m+1)n}.$$

Finally the determinant in the statement of Proposition 2.25 is $\Delta(1)/n!^{m+1}$, where Δ is the determinant of Proposition 2.24. Hence it does not vanish since $\Delta(1) \neq 0$.

□

Since $\kappa^n/n!$ tends to 0 as n tends to infinity, we may apply the criterion for linear independence Lemma 1.29. Therefore the numbers $1, e, e^2, \dots, e^m$ are linearly independent, and since this is true for all integers m, Hermite's Theorem on the transcendence of e follows.

Exercise 2.26. *Using Hermite's method as explained in § 2, prove that for any non-zero $r \in \mathbb{Q}(i)$, the number e^r is transcendental.*

Exercise 2.27. *Let m be a positive integer and $\epsilon > 0$ a real number. Show that there exists $q_0 > 0$ such that, for any $q \ge q_0$ and for any tuple $(q, p_1, \dots, p_m)$ of rational integers with $q > q_0$,*

$$\max_{1\le\mu\le m}\left| e^\mu - \frac{p_\mu}{q}\right| \ge \frac{1}{q^{1+(1/m)+\epsilon}}.$$

Is it possible to improve the exponent by replacing $1 + (1/m)$ with a smaller number?

`Hint.` *Consider Hermite's proof of the transcendence of e (§ 2.1.3), especially Proposition 2.25. First check (for instance using Cauchy's formulae)*

$$\max_{0\le j\le m} \frac{1}{k!}|D^k f_j(\mu)| \le c_1^n,$$

where c_1 is a positive real number which does not depend on n. Next, check that the numbers p_j and $q_{\mu j}$ satisfy

$$\max\{q_j, |p_{\mu j}|\} \le (n!)^m c_2^m$$

for $1 \le \mu \le m$ and $0 \le j \le n$, where again $c_2 > 0$ does not depend on n. Then repeat the proof of Hermite in § 2 with n satisfying

$$(n!)^m c_3^{-2mn} \le q < \bigl((n+1)!\bigr)^m c_3^{-2m(n+1)},$$

where $c_3 > 0$ is a suitable constant independent on n. One does not need to compute c_1, c_2 and c_3 in terms of m, one only needs to show their existence so that the proof yields the desired estimate.

2.3 Transcendental numbers: historical survey

We already stated Hermite's Theorem on the transcendence of e, Lindemann's Theorem on the transcendence of π and Hermite-Lindemann's Theorem on the transcendence of $\log\alpha$ and e^β for non-zero algebraic numbers α and β (with the proviso $\log\alpha \neq 0$) – see Theorem 2.1.

We give a brief review of the theory of Diophantine Approximation, next we complete the history of the theory in the XIX-th century, and then discuss the development in the XX-th century.

References are [3] and [2].

2.3.1 Diophantine Approximation: historical survey

References for this section are [3, 21, 12, 6].

The next corollary of Lemma 1.24 was proved by J. Liouville in 1844: this his how he constructed the first examples of transcendental numbers. His first explicit examples were given by continued fractions, next he gave further examples with series like

$$\vartheta = \sum_{n\geq 0} g^{-n!} \tag{2.28}$$

for any integer $g \geq 2$.

Lemma 2.29. *For any algebraic number α of degree d, there exist a constant c such that, for any rational number $p/q \neq \alpha$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^d}\cdot$$

Exercise 2.30. *Denote by $P \in \mathbb{Z}[X]$ the minimal polynomial of α.*
a) Prove this result with κ given by

$$\kappa = \max\bigl\{1\;;\;\max_{|t-\alpha|\leq 1}|P'(t)|\bigr\}.$$

b) Check also that the same estimate is true with κ given by

$$\kappa = a_0\prod_{i=2}^{d}(|\alpha_j - \alpha| + 1),$$

where a_0 is the leading coefficient and $\alpha_1,\ldots,\alpha_d$ the roots of P with $\alpha_1 = \alpha$:

$$P(X) = a_0(X-\alpha_1)(X-\alpha_2)\cdots(X-\alpha_d).$$

`Hint:` *For both parts of this exercise one may distinguish two cases, whether $|\alpha - p/q|$ is ≥ 1 or < 1.*

Definition. *A real number θ is a* Liouville number *if for any $\kappa > 0$ there exists $p/q \in \mathbb{Q}$ with $q \geq 2$ and*

$$0 < \left|\alpha - \frac{p}{q}\right| \leq \frac{c}{q^\kappa}\cdot$$

It follows from Lemma 2.29 that Liouville numbers are transcendental. In dynamical systems one says that an irrational real number *satisfies a Diophantine condition* if is not Liouville: this means that there exists a constant $\kappa > 0$ such that, for any $p/q \in \mathbb{Q}$ with sufficiently large q,

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^\kappa}\cdot$$

Let us check that the numbers (2.28) are Liouville numbers. More generally, let $g \geq 2$ be an integer, $(a_n)_{n\geq 0}$ a bounded sequence of rational numbers with infinite support (the *support* is the set of n with $a_n \neq 0$) and $\kappa > 0$ a real number. We show that

$$\vartheta = \sum_{n\geq 0} a_n g^{-n!}$$

is a Liouville number. Let A be an upper bound for $|a_n|$ ($n \geq 0$). Let N be a sufficiently large integer such that $a_{N+1} \neq 0$. Set

$$q = g^{N!}, \quad p = \sum_{n=0}^{N} a_n g^{N!-n!}, \quad r_1 = \frac{A}{g^{N!N}} \quad \text{and} \quad r_2 = \sum_{k\geq 2} \frac{A}{g^{(N+k)!-N!}}$$

so that

$$\vartheta - \frac{p}{q} = r_1 + r_2.$$

For $k \geq 2$ we use the crude estimates

$$(N+k)! - N! \geq N!N(N+k-1) \quad \text{and} \quad \sum_{k\geq 1} g^{-(k-1)} \leq 2,$$

which yield, for sufficiently large N,

$$|r_2| < \frac{1}{g^{N!N}} \leq |r_1| < \frac{A}{q^N}, \quad \text{hence} \quad 0 < \left|\vartheta - \frac{p}{q}\right| < \frac{2A}{q^N}\cdot$$

Definition. *Given a real irrational number ϑ, a function $\varphi = \mathbb{N} \to \mathbb{R}_{>0}$ is an irrationality measure for ϑ if there exists an integer $q_0 > 0$ such that, for any $p/q \in \mathbb{Q}$ with $q \geq q_0$,*

$$\left|\vartheta - \frac{p}{q}\right| \geq \varphi(q).$$

Further, a real number κ is an irrationality exponent for ϑ if there exists a positive constant c such that the function c/q^κ is an irrationality measure for ϑ.

From Dirichlet's box principle (see (i)⇒(iv) in Lemma 1.13) it follows that any irrationality exponent κ satisfies $\kappa \geq 2$. Irrational quadratic numbers have irrationality exponent 2. It is known (see for instance [21] Th. 5F p. 22) that 2 is an irrationality exponent for an irrational real number ϑ if and only if the sequence of *partial quotients* $(a_0, a_1, \ldots)$ in the continued fraction expansion of ϑ is bounded: these are called the *badly approximable numbers*.

From Liouville's inequality in Lemma 2.29 it follows that any irrational algebraic real number α has a finite irrationality exponent $\leq d$. Liouville numbers are by definition exactly the irrational real numbers which have no finite irrationality exponent.

For any $\kappa \geq 2$, there are irrational real numbers ϑ for which κ is an irrationality exponent and is the best: no positive number less than κ is an irrationality exponent for ϑ. Examples due to Y. Bugeaud in connection with the triadic Cantor set (see [15]) are

$$\sum_{n=0}^{\infty} 3^{-\lceil \lambda\kappa \rceil^n}$$

where λ is any positive real number.

The first significant improvement to Liouville's inequality is due to the Norwegian mathematician Axel Thue who proved in 1909:

Theorem 2.31 (A. Thue, 1909)**.** *Let α be a real algebraic number of degree $d \geq 3$. Then any $\kappa > (d/2) + 1$ is an irrationality exponent for α.*

The fact that the irrationality exponent is $< d$ has very important corollaries in the theory of Diophantine equations:

Theorem 2.32 (Thue)**.** *Let $f \in \mathbb{Z}[X]$ be an irreducible polynomial of degree $d \geq 3$ and m a non-zero rational integer. Define $F(X, Y) = Y^d f(X/Y)$. Then the Diophantine equation $F(x, y) = m$ has only finitely many solutions $(x, y) \in \mathbb{Z} \times \mathbb{Z}$.*

The equation $F(x, y) = m$ in Proposition 2.32 is called *Thue equation*. The connection between Thue equation and Diophantine approximation is the following:

Lemma 2.33. *Let α be an algebraic number of degree $d \geq 3$ and minimal polynomial $f \in \mathbb{Z}[X]$, let $F(X, Y) = Y^d f(X/Y) \in \mathbb{Z}[X, Y]$ be the associated homogeneous polynomial. Let $0 < \kappa \leq d$. The following conditions are equivalent:*
(i) There exists $c_1 > 0$ such that, for any $p/q \in \mathbb{Q}$,

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c_1}{q^\kappa}\cdot$$

(ii) There exists $c_2 > 0$ such that, for any $(x, y) \in \mathbb{Z}^2$ with $x > 0$,

$$|F(x, y)| \geq c_2\ x^{d-\kappa}.$$

In 1921 C.L. Siegel sharpened Thue's result 2.31 by showing that any real number

$$\kappa > \min_{1\leq j\leq d}\left(\frac{d}{j+1} + j\right)$$

is an irrationality exponent for α. With $j = [\sqrt{d}]$ it follows that $2\sqrt{d}$ is an irrationality exponent for α. Dyson and Gel'fond in 1947 independently refined

Siegel's estimate and replaced the hypothesis in Thue's Theorem 2.31 by $\kappa > \sqrt{2d}$. The essentially best possible estimate has been achieved by K.F. Roth in 1955: any $\kappa > 2$ is an irrationality exponent for a real irrational algebraic number α.

Theorem 2.34 (A. Thue, C.L. Siegel, F. Dyson, K.F. Roth 1955)**.** *For any real algebraic number α, for any $\epsilon > 0$, the set of $p/q \in \mathbb{Q}$ with $|\alpha - p/q| < q^{-2-\epsilon}$ is finite.*

It is expected that the result is not true with $\epsilon = 0$ as soon as the degree of α is ≥ 3, which means that it is expected no real algebraic number of degree at least 3 is badly approximable, but essentially nothing is known on the continued fraction of such numbers: we do not know whether there exists an irrational algebraic number which is not quadratic and has bounded partial quotient in its continued fraction expansion, but we do not know either whether there exists a real algebraic number of degree at least 3 whose sequence of partial quotients is not bounded!

Here is an example of application of Diophantine approximation to transcendental number theory. Let $(u_n)_{n\geq 0}$ be an increasing sequence of integers and let g be a rational integer, $g \geq 2$. We wish to prove that the number

$$\vartheta = \sum_{n\geq 0} g^{-u_n} \tag{2.35}$$

is transcendental. A conjecture of Borel (1950 – see [13] and § 5.8) states that *the digits in the binary expansion of a real algebraic irrational number should be uniformly equidistributed*; in particular the sequence of 1's should not be lacunary.

For sufficiently large n, define

$$q_n = g^{u_n}, \quad p_n = \sum_{k=0}^{n} g^{u_n - u_k} \quad \text{and} \quad r_n = \vartheta - \frac{p_n}{q_n}\cdot$$

Since the sequence $(u_n)_{n\geq 0}$ is increasing, we have $u_{n+h} - u_{n+1} \geq h - 1$ for any $h \geq 1$, hence

$$0 < r_n \leq \frac{1}{g^{u_{n+1}}} \sum_{h\geq 1} \frac{1}{g^{h-1}} = \frac{g}{2^{u_{n+1}}(g-1)} \leq \frac{2}{q_n^{u_{n+1}/u_n}}\cdot$$

Therefore if the sequence $(u_n)_{n\geq 0}$ satisfies

$$\limsup_{n\to\infty} \frac{u_{n+1}}{u_n} = +\infty$$

then ϑ is a Liouville number, and therefore is transcendental. For instance $u_n = n!$ satisfies this condition: hence the number $\sum_{n\geq 0} g^{-n!}$ is transcendental.

Exercise 2.36. *Let $g \geq 2$ be an integer, $(a_n)_{n\geq 0}$ be a bounded sequence of rational integers and $(u_n)_{n\geq 0}$ be an increasing sequence of integers satisfying*

$$\limsup_{n\to\infty} \frac{u_{n+1}}{u_n} = +\infty.$$

Assume that the support $\{n \geq 0\ ;\ a_n \neq 0\}$ of the sequence $(a_n)_{n\geq 0}$ is infinite.

Define

$$\vartheta = \sum_{n\geq 0} a_n g^{-u_n}.$$

Show that ϑ is a Liouville number.

Roth's Theorem 2.34 yields the transcendence of the number ϑ in (2.35) under the weaker hypothesis

$$\limsup_{n\to\infty} \frac{u_{n+1}}{u_n} > 2.$$

The sequence $u_n = [2^{\theta n}]$ satisfies this condition as soon as $\theta > 1$. For example the transcendence of the number

$$\sum_{n\geq 0} g^{-3^n}$$

follows from Theorem 2.34.

A stronger result follows from Ridout's Theorem 2.37 below, using the fact that the denominators g^{u_n} are powers of b.

Let S be a set of prime. A rational number is called *a S–integer* if it can be written u/v where all prime factors of the denominator v belong to S. For instance when a, b and m are rational integers with $b \neq 0$, the number a/b^m is a S–integer for S the set of prime divisors of b.

Theorem 2.37 (D. Ridout, 1957)**.** *Let S be a finite set of prime numbers. For any real algebraic number α, for any $\epsilon > 0$, the set of $p/q \in \mathbb{Q}$ with q a S–integer and $|\alpha - p/q| < q^{-1-\epsilon}$ is finite.*

Therefore the condition

$$\limsup_{n\to\infty} \frac{u_{n+1}}{u_n} > 1$$

suffices to imply the transcendence of the sum of the series (2.35). An example is the transcendence of the number

$$\sum_{n\geq 0} g^{-2^n}.$$

This result goes back to A. J. Kempner in 1916.

The theorems of Thue–Siegel–Roth and Ridout are very special cases of Schmidt's subspace Theorem (1972) together with its p-adic extension by H.P. Schlickewei (1976). We state do not state it in full generality but we give only two special cases.

For $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{Z}^m$, define $|\mathbf{x}| = \max\{|x_1|, \ldots, |x_m|\}$.

Theorem 2.38 (W.M. Schmidt (1970): simplified form)**.** *For $m \geq 2$ let $L_1, \ldots, L_m$ be independent linear forms in m variables with algebraic coefficients. Let $\epsilon > 0$. Then the set*

$$\{\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{Z}^m \; ; \; |L_1(\mathbf{x}) \cdots L_m(\mathbf{x})| \leq |\mathbf{x}|^{-\epsilon}\}$$

is contained in the union of finitely many proper subspaces of $\mathbb{Q}^m$.

Thue–Siegel–Roth's Theorem 2.34 follows from Theorem 2.38 by taking

$$m = 2, \quad L_1(x_1, x_2) = x_1, \quad L_2(x_1, x_2) = \alpha x_1 - x_2.$$

A $\mathbb{Q}$-vector subspace of $\mathbb{Q}^2$ which is not $\{0\}$ not $\mathbb{Q}^2$ (that is *a proper subspace* is of the generated by an element $(p_0, q_0) \in \mathbb{Q}^2$. There is one such subspace with $q_0 = 0$, namely $\mathbb{Q} \times \{0\}$ generated by $(1, 0)$, the other ones have $q_0 \neq 0$. Mapping such a rational subspace to the rational number p_0/q_0 yields a 1 to 1 correspondence. Hence Theorem 2.38 says that there is only a finite set of exceptions p/q in Roth's Theorem 2.34.

For x a non–zero rational number, write the decomposition of x into prime factors

$$x = \prod_p p^{v_p(x)},$$

where p runs over the set of prime numbers and $v_p(x) \in \mathbb{Z}$ (with only finitely many $v_p(x)$ distinct from 0), and set

$$|x|_p = p^{-v_p(x)}.$$

For $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{Z}^m$ and p a prime number, define $|\mathbf{x}| = \max\{|x_1|_p, \ldots, |x_m|_p\}$.

Theorem 2.39 (Schmidt's Subspace Theorem)**.** *Let $m \geq 2$ be a positive integer, S a finite set of prime numbers. Let $L_1, \ldots, L_m$ be independent linear forms in m variables with algebraic coefficients. Further, for each $p \in S$ let $L_{1,p}, \ldots, L_{m,p}$ be m independent linear forms in m variables with rational coefficients. Let $\epsilon > 0$. Then the set of $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{Z}^m$ such that*

$$|L_1(\mathbf{x}) \cdots L_m(\mathbf{x})| \prod_{p \in S} |L_{1,p}(\mathbf{x}) \cdots L_{m,p}(\mathbf{x})|_p \leq |\mathbf{x}|^{-\epsilon}$$

is contained in the union of finitely many proper subspaces of $\mathbb{Q}^m$.

Ridout's Theorem 2.37 is a corollary of Schmidt's subspace Theorem: in Theorem 2.39 take $m = 2$,

$$\begin{aligned} &L_1(x_1, x_2) = L_{1,p}(x_1, x_2) = x_1, \\ &L_2(x_1, x_2) = \alpha x_1 - x_2, \quad L_{2,p}(x_1, x_2) = x_2. \end{aligned}$$

For $(x_1, x_2) = (b, a)$ with b a S–integer and $p \in S$, we have

$$|L_1(x_1, x_2)| = b, \quad |L_2(x_1, x_2)| = |b\alpha - a|,$$
$$|L_{1p}(x_1, x_2)|_p = |b|_p, \quad |L_{2,p}(x_1, x_2)|_p = |a|_p \leq 1.$$

and

$$\prod_{p \in S} |b|_p = b^{-1}$$

since b is a S–integer.

An important issue is the one on *effectivity*. We have seen that Liouville's inequality in Lemma 2.29 is explicit. The improvements which provided by Thue, Siegel, Dyson, Roth in Theorem 2.34 and Schmidt in Theorem 2.39 are not effective: the proofs do not enable one to compute the constants. In some cases it is possible to produce explicit upper bounds for the number of exceptions. A lot of recent work is devoted to this question. A reference is [1].

2.3.2 Transcendental numbers after Liouville and before 1900: Cantor, Hermite, Lindemann, Weierstraß

Liouville gave the first explicit examples of transcendental numbers in 1844 (see § 2.3.1). In 1874 and 1891 G. Cantor produced another argument proving the existence of transcendental numbers. A detailed discussion of the constructive aspect of Cantor's work is [5].

In 2 we already quoted the contributions of Ch. Hermite in 1873 (transcendence of e), F. Lindemann in 1882 (transcendence of π) and the Theorem 2.1 of Hermite and Lindemann on the transcendence of non–zero logarithms of algebraic numbers: $\mathcal{L} \cap \overline{\mathbb{Q}} = \{0\}$.

In 1888, K. Weierstraß completed the proof of the following claim by F. Lindemann:

Theorem 2.40 (Lindemann–Weierstraß – first form)**.** *Let $\alpha_1, \ldots, \alpha_m$ be algebraic numbers which are pairwise distinct: $\alpha_i \neq \alpha_j$ for $i \neq j$. Then the numbers $e^{\alpha_1}, \ldots, e^{\alpha_m}$ are linearly independent over $\mathbb{Q}$.*

It is easy to checked that Theorem 2.40 is equivalent to the next statement:

Theorem 2.41 (Lindemann–Weierstraß – second form)**.** *Let $\beta_1, \ldots, \beta_n$ be algebraic numbers which are linearly independent over $\mathbb{Q}$. Then the numbers $e^{\beta_1}, \ldots, e^{\beta_n}$ are algebraically independent over $\mathbb{Q}$.*

Now the algebraic independence of complex numbers over $\mathbb{Q}$ is equivalent to the algebraic independence over the field $\overline{\mathbb{Q}}$ of algebraic numbers. Therefore Theorem 2.40 is also equivalent to the next statement:

Theorem 2.42 (Lindemann–Weierstraß – third form)**.** *Let $\alpha_1, \ldots, \alpha_m$ be algebraic numbers which are pairwise distinct. Then the numbers $e^{\alpha_1}, \ldots, e^{\alpha_m}$ are linearly independent over $\overline{\mathbb{Q}}$.*

2.3.3 Hilbert's seventh problem: Gel'fond, Schneider, Baker

The solution of Hilbert's seventh problem on the transcendence of α^β was obtained by Gel'fond and Schneider in 1934. (see [4, 11]).

Theorem 2.43 (Gel'fond-Schneider, 1934). *For α and β algebraic numbers with $\alpha \neq 0$ and $\beta \notin \mathbb{Q}$ and for any choice of $\log\alpha \neq 0$, the number $\alpha^\beta = \exp(\beta \log \alpha)$ is transcendental.*

This means that the two algebraically independent functions e^z and $e^{\beta z}$ cannot take algebraic values at the points $\log\alpha$ (A.O. Gel'fond) and also that the two algebraically independent functions z and $\alpha^z = e^{z\log\alpha}$ cannot take algebraic values at the points $m + n\beta$ with $(m,n) \in \mathbb{Z}^2$ (Th. Schneider).

Examples (quoted by D. Hilbert in 1900) of numbers whose transcendence follows from Theorem 2.43 are $2^{\sqrt{2}}$ and e^π (recall $e^{i\pi} = -1$). The transcendence of e^π had already been proved in 1929 by A.O. Gel'fond.

Here is an equivalent statement to Gel'fond-Schneider Theorem 2.43:

Theorem 2.44 (Gel'fond-Schneider, 1934). *Let $\log\alpha_1, \log\alpha_2$ be two non-zero logarithms of algebraic numbers. Assume that the quotient $(\log\alpha_1)/(\log\alpha_2)$ is irrational. Then this quotient is transcendental.*

The generalization of Theorem 2.44 to more than two logarithms, conjectured by A.O. Gel'fond [4], was proved by A. Baker in 1966. His results include not only Gel'fond-Scheider's Theorem 2.44 but also Hermite-Lindemann's Theorem 2.1.

Theorem 2.45 (Baker, 1966). *Let $\log\alpha_1, \ldots, \log\alpha_n$ be $\mathbb{Q}$–linearly independent logarithms of algebraic numbers. Then the numbers $1, \log\alpha_1, \ldots, \log\alpha_n$ are linearly independent over the field $\overline{\mathbb{Q}}$.*

2.3.4 The Six Exponentials Theorem

The next result, which does not follow from any of the previously mentioned results, was proved independently in the 1940's by C.L. Siegel (unpublished) and in the 1960's by S. Lang and K. Ramachandra; see also Problem 1 in [11] for the Four Exponentials Conjecture). As suggested by K. Ramachandra, Theorem 2.46 also follows from Schneider's criterion proved in 1949.

Theorem 2.46 (Six Exponentials Theorem). *Let $x_1, \ldots, x_d$ be $\mathbb{Q}$-linearly independent complex numbers and let $y_1, \ldots, y_\ell$ be $\mathbb{Q}$-linearly independent complex numbers. Assume $\ell d > \ell + d$. Then at least one of the ℓd numbers*

$$e^{x_i y_j}, \qquad (1 \leq i \leq d,\ 1 \leq j \leq \ell)$$

is transcendental.

Notice that the condition $\ell d > \ell + d$ can be written ($\ell \geq 2$ and $d \geq 3$) or ($\ell \geq 3$ and $d \geq 2$); it suffices to consider the case $\ell d = 6$ (hence the name of the result). Therefore, Theorem 2.46 can be stated in an equivalent form:

Theorem 2.47 (Six Exponentials Theorem - logarithmic form)**.** *Let*

$$M = \begin{pmatrix} \log\alpha_1 & \log\alpha_2 & \log\alpha_3 \\ \\ \log\beta_1 & \log\beta_2 & \log\beta_3 \end{pmatrix}$$

be a 2 *by* 3 *matrix whose entries are logarithms of algebraic numbers. Assume that the three columns are linearly independent over* $\mathbb{Q}$ *and the two rows are also linearly independent over* $\mathbb{Q}$*. Then the matrix* M *has rank* 2*.*

It is expected that the condition $d\ell > d+\ell$ in Theorem 2.46 is too restrictive and that the same conclusion holds in case $d = \ell = 2$.

We state this conjecture in the logarithmic form:

Conjecture 2.48 (Four Exponentials conjecture - logarithmic form)**.** *Let* M *be a* 2×2 *matrix whose entries are logarithms of algebraic numbers:*

$$M = \begin{pmatrix} \log\alpha_{11} & \log\alpha_{12} \\ \log\alpha_{21} & \log\alpha_{22} \end{pmatrix};$$

assume that the two rows of this matrix are linearly independent over $\mathbb{Q}$ *(in* $\mathbb{C}^2$*), and also that the two columns are linearly independent over* $\mathbb{Q}$*; then the rank of* M *is* 2*.*

This topic has been extensively studied (see [14]), giving rise to the five exponentials theorem, the strong six exponentials theorem and the strong four exponential conjecture. We investigate further open questions as well as related ones in § 5.1.

2.3.5 Algebraic independence of the values of the exponential function

We already stated the Theorem 2.41 of Lindemann–Weierstraß on the algebraic independence of exponentials of algebraic numbers. This is one of the few strong results of algebraic independence: the point is that it reduces to a statement on linear independence (Theorem 2.40), and linear independence is easier to deal with than algebraic independence.

In 1948 and 1949, A.O. Gel'fond extended his solution of Hilbert's seventh problem to a result of algebraic independence [4]. One of his theorems is that the two numbers $2^{\sqrt[3]{2}}$ and $2^{\sqrt[3]{4}}$ are algebraically independent. His general statements can be seen as extensions of Theorem 2.46 into a result of algebraic independence (in spite of the fact that the Six Exponentials Theorem 2.46 was stated and proved only several years later). In his original work, Gel'fond needed a stronger assumption, namely a measure of linear independence of the x_i's as well as of the y_j's. This assumption was removed in the early 1970's by R. Tijdeman.

Theorem 2.49. *Let* $x_1, \dots, x_d$ *be* $\mathbb{Q}$*-linearly independent complex numbers and let* $y_1, \dots, y_\ell$ *be* $\mathbb{Q}$*-linearly independent complex numbers.*
1. If $d\ell \geq 2(d+\ell)$*, then at least two of the* $d\ell$ *numbers*

$$e^{x_i y_j} \qquad (1 \leq i \leq d,\ 1 \leq j \leq \ell)$$

are algebraically independent.
2. If $d\ell \geq d+2\ell$, then at least two of the $d\ell+d$ numbers

$$x_i,\ e^{x_iy_j} \qquad (1 \leq i \leq d,\ 1 \leq j \leq \ell)$$

are algebraically independent.
3. If $d\ell > d+\ell$, then at least two of the $d\ell+d+\ell$ numbers

$$x_i,\ y_j,\ e^{x_iy_j} \qquad (1 \leq i \leq d,\ 1 \leq j \leq \ell)$$

are algebraically independent.
4. If $d=\ell=2$ and if the two numbers $e^{x_1y_1}$ and $e^{x_1y_2}$ are algebraic, then at least two of the 6 numbers

$$x_1,\ x_2,\ \ y_1,\ y_2,\ \ e^{x_2y_1},\ \ e^{x_2y_2}$$

are algebraically independent.

From the last part of Theorem 2.49, taking $x_1 = y_1 = i\pi$ and $x_2 = y_2 = 1$, one deduces that at least one of the two following statements is true:
(i) *The number e^{π^2} is transcendental.*
(ii) *The two numbers e and π are algebraically independent.*

One expects that both statements are true (se § 5).

If it were possible to prove that, under the assumptions of Theorem 2.49, at least two of the 8 numbers

$$x_1,\ x_2,\ \ y_1,\ y_2,\ \ e^{x_1y_1},\ \ e^{x_1y_2},\ \ e^{x_2y_1},\ \ e^{x_2y_2}$$

are algebraically independent, one would deduce the algebraic independence of the two numbers π and e^{π} (take $x_1 = 1$, $x_2 = i$, $y_1 = \pi$, $y_2 = i\pi$; see Corollary 3.48 below).

For results concerning *large transcendence degree*, see § 3.3.3 below; for conjectures see 5.1.

References

[1] Y. Bugeaud – *Approximation by algebraic numbers*, Cambridge Tracts in Mathematics, vol. 160, Cambridge University Press, Cambridge, 2004.

[2] N. I. Fel′dman & Yu. V. Nesterenko – *Transcendental numbers*, in *Number Theory, IV*, Encyclopaedia Math. Sci., vol. **44**, Springer, Berlin, 1998, p. 1–345.

[3] N.I. Fel’dman & A.B. Šidlovskiĭ – *The development and present state of the theory of transcendental numbers*, (Russian) Uspehi Mat. Nauk **22** (1967) no. 3 (135) 3–81; Engl. transl. in Russian Math. Surveys, **22** (1967), no. 3, 1–79.

[4] A. O. Gel′fond – *Transcendental and algebraic numbers*, Translated from the first Russian edition by Leo F. Boron, Dover Publications Inc., New York, 1960.

[5] R. Gray – *Georg Cantor and transcendental numbers*, Amer. Math. Monthly **101** (1994), no. 9, 819–832.

[6] C. Hermite – *Sur la fonction exponentielle*, C. R. Acad. Sci. Paris, **77** (1873), 18–24; 74–79; 226–233; 285–293; *Oeuvres*, Gauthier Villars (1905), III, 150–181. See also *Oeuvres* III, 127–130, 146–149, and *Correspondance Hermite-Stieltjes*, II, lettre 363, 291–295.

[7] H. Lambert – *Mémoire sur quelques propriétés remarquables des quantités transcendantes circulaires et logarithmiques*, Mémoires de l'Académie des Sciences de Berlin, 17 (1761), 1768, p. 265–322; lu en 1767; Math. Werke, t. II.

[8] F. Lindemann – Sur le rapport de la circonférence au diamètre, et sur les logarithmes népériens des nombres commensurables ou des irrationnelles algébriques. C.R. Acad. Sci. Paris, **95** (1882), 72–74.

[9] J. Liouville – Sur des classes très étendues de quantités dont la valeur n"est ni algébrique, ni même réductible à des irrationnelles algébriques. C.R. Acad. Sci. Paris, **18** (1844), 883–885 et 910–911. J. Math. Pures et Appl. (1) **16** (1851), 133–142.

[10] I. Niven – *Irrational numbers*, Carus Math. Monographs **11** (1956).

[11] Th. Schneider –*Einführung in die transzendenten Zahlen*, Springer-Verlag, Berlin, 1957.

[12] C.L. Siegel – *Transcendental Numbers*, Annals of Mathematics Studies, **16**. Princeton University Press, Princeton, N. J., 1949.

[13] M. Waldschmidt – *Words and Transcendence.* "Analytic Number Theory - Essays in Honour of Klaus Roth", Cambridge University Press, to appear. http://www.math.jussieu.fr/~miw/articles/pdf/WordsTranscendence.pdf

[14] M. Waldschmidt – *Diophantine Approximation on linear algebraic groups. Transcendence Properties of the Exponential Function in Several Variables.* Grundlehren der Mathematischen Wissenschaften **326**, Springer-Verlag, Berlin-Heidelberg, 2000.

[15] M. Waldschmidt – *Report on some recent progress in Diophantine approximation.* To appear
http://www.math.jussieu.fr/~miw/articles/pdf/miwLangMemorialVolume.pdf