Fields Institute University of Ottawa, Ontario Canada June 30 – July 4, 2008
http://www.fields.utoronto.ca/programs/scientific/07-08/analytic/
updated: August 4, 2008

Summer School in Analytic Number Theory and Diophantine Approximation

An Introduction to Irrationality and Transcendence Methods.

Michel Waldschmidt

2 Historical introduction to transcendence

In 1873 C. Hermite [4] 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 in 1932 related the properties of the two types of Padé's approximants and used 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 [3] and [2] Chap. 2 § 3.

In the analogy with number theory, Padé approximants of type II are related with the simultaneous approximation of real numbers $\vartheta_1, \ldots, \vartheta_m$ 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 $a_0, \ldots, a_m$ are rational integers, not all of which are 0, in terms of the number $\max_{0\le i\le m} |a_i|$.

After his seminal work, Ch. Hermite wrote to C.A. Borchardt (see [5, 1]:

> *Tout ce que je puis, c'est de refaire ce qu'a déjà fait Lambert, seulement d'une autre manière.*

and

> *Je ne hasarderai point à la recherche d'une démonstration de la transcendance du nombre π. Que d'autres tentent l'entreprise; mais croyez m'en, mon cher ami, il ne laissera pas que de leur en coûter quelques efforts.*

F. Lindemann [6] was able to extend the argument and to prove the transcendence of π (hence he solved the old greek problem 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 [8]), 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 A criterion for linear independence

We first state a criterion for linear independence which will be used in § 2 for the proof by Hermite of the transcendence of e. This is a generalisation (from personal notes by Michel Laurent of a course he gave in Marseille) of one of Lemma 1.18. Most often in mathematics there is sort of an entropy: when a statement provides a necessary and sufficient condition, and when one of the two implications is easy while the other requires more work, then it is the difficult part which is most useful. Here we have a counterexample to this claim: in the Criterion 2.2 below, one of the implications is easy while the other is deeper; but it turns out that it is the easy one which is used in transcendence proofs.

Let $\vartheta_1, \dots, \vartheta_m$ be real numbers and a_0, $a_1, \dots, a_m$ rational integers, not all of which are 0. Our goal is to prove that the number

$$L = a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m$$

is not 0.

The idea is to approximate simultaneously $\vartheta_1, \dots, \vartheta_m$ by rational numbers $p_1/q, \dots, p_m/q$ with the same denominator $q > 0$.

Let $q, p_1, \dots, p_m$ be rational integers with $q > 0$. For $1 \le \mu \le m$ set

$$\epsilon_\mu = q\vartheta_\mu - p_\mu.$$

Then $qL = M + R$ with

$$M = a_0q + a_1p_1 + \cdots + a_mp_m \in \mathbb{Z} \quad \text{and} \quad R = a_1\epsilon_1 + \cdots + a_m\epsilon_m \in \mathbb{R}.$$

If $M \neq 0$ and $|R| < 1$ we deduce $L \neq 0$.

One of the main difficulties is often to check $M \neq 0$. This question gives rise to the so-called *zero estimates* or *non-vanishing lemmas*. In the present situation, we wish to find a $m+1$–tuple $(q, p_1, \dots, p_m)$ giving a simultaneous rational approximation to $(\vartheta_1, \dots, \vartheta_m)$, but we also require that it lies outside the hyperplane $a_0x_0 + a_1x_1 + \cdots + a_mx_m = 0$ of $\mathbb{Q}^{m+1}$. Since this needs to be checked for all hyperplanes, the solution is to construct not only one tuple $(q, p_1, \dots, p_m)$ in $\mathbb{Z}^{m+1} \setminus \{0\}$, but $m+1$ such tuples which are linearly independent. This yields $m+1$ pairs (M_k, R_k), $k = 0, \dots, m$ in place of a single pair (M, R). From $(a_0, \dots, a_m) \neq 0$ one deduces that one at least of $M_0, \dots, M_m$ is not 0.

It turns out that nothing is lossed by using such arguments: existence of linearly independent simultaneous rational approximations for $\vartheta_1, \dots, \vartheta_m$ are characteristic of linearly independent numbers $1, \vartheta_1, \dots, \vartheta_m$. As we just said earlier, we shall use only the easy part of the next Lemma 2.2.

Lemma 2.2. *Let $\underline{\vartheta} = (\vartheta_1, \dots, \vartheta_m) \in \mathbb{R}^m$. Then the following conditions are equivalent.*
(i) The numbers $1, \vartheta_1, \dots, \vartheta_m$ are linearly independent over $\mathbb{Q}$.
(ii) For any $\epsilon > 0$ there exist $m+1$ linearly independent elements $\underline{b}_0, \underline{b}_1, \dots, \underline{b}_m$ in $\mathbb{Z}^{m+1}$, say

$$\underline{b}_i = (q_i, p_{1i}, \dots, p_{mi}), \quad (0 \le i \le m)$$

with $q_i > 0$, such that

$$\max_{1 \le \mu \le m} \left| \vartheta_\mu - \frac{p_{\mu i}}{q_i} \right| \le \frac{\epsilon}{q_i}, \quad (0 \le i \le m). \tag{2.3}$$

In (ii) there is no non-vanishing condition. For $m = 1$ this criterion is not identical to the irrationality criterion: in Lemma 1.10, we required for each ϵ one approximation p/q distinct from θ. Here, in case $m = 1$, we need two linearly independent approximations: hence, even if θ is rational, at least one of them is not the trivial one.

The condition on linear independence of the elements $\underline{b}_0, \underline{b}_1, \ldots, \underline{b}_m$ means that 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 0.

We shall prove a more explicit version of (ii)⇒(i): we check that *any tuple* $(q, p_1, \ldots, p_m) \in \mathbb{Z}^{m+1}$ *producing a tuple* $(p_1/q, \ldots, p_m/q) \in \mathbb{Q}^m$ *of sufficiently good rational approximations to* $\underline{\vartheta}$ *satisfies the same linear dependence relations as* $1, \vartheta_1, \ldots, \vartheta_m$.

Lemma 2.4. *Let* $\vartheta_1, \ldots, \vartheta_m$ *be real numbers. Assume that the numbers* $1, \vartheta_1, \ldots, \vartheta_m$ *are linearly dependent over* $\mathbb{Q}$*: let* $a_0, a_1, \ldots, a_m$ *be rational integers, not all of which are zero, satisfying*

$$a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m = 0.$$

Let $\epsilon > 0$ *satisfy* $\sum_{\mu=1}^{m} |a_\mu] < 1/\epsilon$. *Assume further that* $(q, p_1, \ldots, p_m) \in \mathbb{Z}^{m+1}$ *satisfies* $q > 0$ *and*

$$\max_{1 \le \mu \le m} |q\vartheta_\mu - p_\mu| \le \epsilon.$$

Then

$$a_0 q + a_1 p_1 + \cdots + a_m p_m = 0.$$

Proof. In the relation

$$q a_0 + \sum_{\mu=1}^{m} a_\mu p_\mu = -\sum_{\mu=1}^{m} a_\mu (q\vartheta_\mu - p_\mu),$$

the right hand side has absolute value less than 1 and the left hand side is a rational integer, so it is 0.

□

Proof of (ii)⇒(i) in Lemma 2.2. By assumption (ii) we have $m+1$ linearly independent elements $\underline{b}_i \in \mathbb{Z}^{m+1}$ such that the corresponding rational approximation satisfy the assumptions of Lemma 2.4. Consider a non–zero linear form L in $m+1$ variables $\underline{X} = (X_0, X_1, \ldots, X_m)$ with integer coefficients

$$L(\underline{X}) = a_0 X_0 + a_1 X_1 + \cdots + a_m X_m.$$

Since $L \neq 0$, one at least of the $L(\underline{b}_i)$ is not 0. For this $\underline{b}_i$ the conclusion of Lemma 2.4 is not satisfied, hence

$$a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m \neq 0.$$

□

Proof of (i)⇒(ii) in Lemma 2.2. Let $\epsilon > 0$. Assume (i) holds. By Dirichlet's box principle (Lemma 1.7), there exists $\underline{b} = (q, p_1, \ldots, p_m) \in \mathbb{Z}^{m+1}$ with $q > 0$ such that

$$\max_{1 \le \mu \le m} \left| \vartheta_\mu - \frac{p_\mu}{q} \right| \le \frac{\epsilon}{q}.$$

Consider the subset $E_\epsilon \subset \mathbb{Z}^{m+1}$ of these tuples. We show that the $\mathbb{Q}$-vector subspace V_ϵ of $\mathbb{Q}^{m+1}$ spanned by E_ϵ is $\mathbb{Q}^{m+1}$. It will follow that there are $m+1$ linearly independent elements in E_ϵ.

If $V_\epsilon \ne \mathbb{Q}^{m+1}$, then there is a hyperplane $a_0 z_0 + a_1 z_1 + \cdots + a_m z_m = 0$ containing E_ϵ. Any $\underline{b} = (q, p_1, \ldots, p_m)$ in E_ϵ has

$$a_0 q + a_1 p_1 + \cdots + a_m p_m = 0.$$

For each $n \ge 1/\epsilon$, let $\underline{b}_n = (q_n, p_{1n}, \ldots, p_{mn}) \in E_\epsilon$ satisfy

$$\max_{1 \le \mu \le m} \left| \vartheta_\mu - \frac{p_{\mu n}}{q_n} \right| \le \frac{1}{n q_n}.$$

Then

$$a_0 + a_1 \theta_1 + \cdots + a_m \theta_m = \sum_{\mu=1}^{m} a_\mu \left(\theta_\mu - \frac{p_{\mu n}}{q_n} \right).$$

Hence

$$|a_0 + a_1 \theta_1 + \cdots + a_m \theta_m| \le \frac{1}{n q_n} \sum_{\mu=1}^{m} |a_\mu|.$$

The right hand side tends to 0 as n tends to infinity, hence the left hand side vanishes, and $1, \vartheta_1, \ldots, \vartheta_m$ are $\mathbb{Q}$–linearly dependent, which contradicts (i).

□

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

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

Then there exists a tuple $(Q, P_1, \ldots, P_m)$ 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 2.5 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 of Lemma 2.5. The polynomial Q of Lemma 2.5 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.

In other terms, the linear map

$$\begin{array}{ccc} \mathbb{C}[X]_{\leq N-n_0} & \longrightarrow & \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_m} \\ Q & \longmapsto & \left(\left(\frac{d}{dz}\right)^k \big[Q(z)f_\mu(z)\big]_{z=0}\right)_{N-n_\mu<k\leq N,\ 1\leq\mu\leq m} \end{array}$$

from a space of dimension $N - n_0 + 1$ to a space of dimension $n_1 + \cdots + n_m = N - n_0$ is not injective. □

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 2.5 shows that it has positive dimension. In the case where this dimension is 1 (which means that there is unicity up to a multiplicative constant), the system of approximants is called *perfect*. An example is with $m = 1$ and $f(x) = e^x$, as shown by Theorem 1.23.

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 2.6. *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 2.6 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 2.6, 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 2.7. *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 2.7, 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 P and Q defined by

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

satisfy the properties which were required for A and B in section §1.4.1 (see Proposition 1.21), 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 2.7 is a powerful tool to go much further.

Proposition 2.8. *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 $n_0, n_1, \ldots, n_m$*.

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 2.9. *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 2.9 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 2.8.

□

Exercise 2.10. *Compare:*
- *the case $m = 1$ of Proposition 2.8,*
- § *1.4.2*
- § *1.4.3.*

(and fix the misprints!)

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 $(p_1/q, p_2/q, \ldots, p_m/q)$ to the numbers $e, e^2, \ldots, e^m$. In order to use Lemma 2.2, we need to have independent such approximations. This is a subtle point which Hermite did not find easy to overcome: we quote from p. 77 of [4]:

Mais une autre voie conduira à une démonstration plus rigoureuse

The following approach, due to K. Mahler, may be viewed as an extension of the simple non-vanishing argument used in § 1.4.5 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 2.11. *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 $P_{\mu j}(x) - e^{\mu x}Q_j(x)$, $1 \le \mu \le m$, $0 \le j \le m$, has multiplicity at least N at the origin, because for each contiguous triple ($1 \le j \le m$) 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 2.11. □

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 $0 \le j \le m$, the integers $q_j, p_{1j}, \ldots, p_{nj}$ by

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

Proposition 2.12. *There exists a constant $\kappa > 0$ independent on n such that for $1 \le \mu \le m$ and $0 \le j \le m$,*

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

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 2.8:

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

where

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

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

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

Now the integral is bounded from above by

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

Finally the determinant in the statement of Proposition 2.12 is $\Delta(1)/n!^{m+1}$, where Δ is the determinant of Proposition 2.11. 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 2.2. Therefore the numbers $1, e, e^2, \ldots, e^m$ are linearly independent, and since this is true for all integers m, Hermite's Theorem on the transcendence of e follows.

Exercise 2.13. *Using Hermite's method as explained in § 2.2, prove that for any non-zero $r \in \mathbb{Q}(i)$, the number e^r is transcendental.*

Exercise 2.14. *Let m be a positive integer and $\epsilon > 0$ a real number. Show that there exists $q_0 > 0$ such that, for any $q \geq q_0$ and for any tuple $(q, p_1, \ldots, p_m)$ of rational integers with $q > q_0$,*

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

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 (§ 1.4.3), especially Proposition 2.12. First check (for instance using Cauchy's formulae)*

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

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

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

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

$$(n!)^m c_3^{-2mn} \leq q < \big((n+1)!\big)^m c_3^{-2m(n+1)},$$

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.

References

[1] C. Brezinski – *History of continued fractions and Padé approximants.* Springer Series in Computational Mathematics, **12**. Springer-Verlag, Berlin, 1991.

[2] N. I. Fel′dman & Yu. V. Ñesterenko – *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*, Uspehi Mat. Nauk **22** (1967) no. 3 (135) 3–81; Engl. transl. in Russian Math. Surveys, **22** (1967), no. 3, 1–79.

[4] 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.

[5] M. Jacob – *La quadrature du cercle*, *Un problème à la mesure des Lumières*, Fayard, 2006.

[6] 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.

[7] 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.

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