Introduction to Diophantine methods *Michel Waldschmidt*
http://www.math.jussieu.fr/∼miw/coursHCMUNS2007.html

Fifth course: september 19, 2007. [7]

2.1.3 Hermite's identity

Let us come back to the problem which was considered in § 1.4.1 and solved by Hermite (Proposition 1.20):

Given two integers $n_0 \geq 0$, $n_1 \geq 0$, find two polynomials A and B with A of degree $\leq n_0$ and B 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 $N = n_0 + n_1$.

From § 1.4.3 one easily deduces that *there is a non-trivial solution, and it is unique if one requires B to be monic. Moreover B has degree n_1 and R has multiplicity exactly $N+1$ at the origin.*

Indeed, 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 $D^{n_0+1}R = cz^{n_1}e^z$ where c is the leading coefficient of B. 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 $N+1$. Finally R is the unique function satisfying $D^{n_0+1}R = cz^{n_1}e^z$ with a zero of multiplicity $\geq n_0$ at 0. According to Lemma 1.25, this implies that the unique solution R for which $c = 1$ is given by the formula of Lemma 1.26:

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

Hence Padé system for the exponential function is perfect.

Our goal is to generalize these results.

Let f be a polynomial. Hermite's Lemma 1.27 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

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

[7]Updated: October 12, 2007

when x and u are complex numbers. Here, in place of using Lemma 1.27, 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

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

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

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

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*

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

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

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

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

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

satisfy the properties which were required in section §1.4.1 (see Proposition 1.20), 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.

$\square$

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 2.2, 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 [4]. The following approach is due to K. Mahler, we can view it as an extension of the simple non-vanishing argument used in § 1.4.2 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 2.10. *There exists a non-zero constant c such that the determinant*

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

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 $c = c_0c_1\cdots c_m$, where c_0 is the leading coefficient of Q_0 and c_μ is the leading coefficient of $P_{\mu\mu}$, $1 \le \mu \le m$.

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

$$\Delta(x) = \begin{vmatrix} Q_0 & P_{10} - e^xQ_0 & \cdots & P_{m0} - e^{mx}Q_0 \\ \vdots & \vdots & \ddots & \vdots \\ Q_m & P_{1m} - e^xQ_m & \cdots & P_{mm} - e^{mx}Q_m \end{vmatrix}.$$

Each $P_{\mu j} - e^{\mu x}Q_j$, $1 \le \mu \le m$, $0 \le j \le m$, 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

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

This completes the proof of Proposition 2.10. $\square$

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

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

Proposition 2.11. *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 $n!$:

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

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 2.11 is $\Delta(1)/n!^{m+1}$, where Δ is the determinant of Proposition 2.10. Hence it does not vanish since $\Delta(1)\ne 0$.

□

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.

2.2 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 $\log\alpha\ne 0$) – see Theorem 2.1. 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].

2.2.1 Transcendental numbers before 1900: Liouville, Hermite, Lindemann, Weierstraß

The next corollary of Lemma 1.13 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

$$\theta_a = \sum_{n\ge 0} a^{-n!} \tag{2.12}$$

for any integer $a\ge 2$.

Lemma 2.13. *For any algebraic number α, there exist two constants c and d such that, for any rational number $p/q \neq \alpha$,*

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

It follows also from Lemma 1.13 that in Lemma 2.13, one can take for d the degree of α (that is the degree of the minimal polynomial of α).

Exercise 2.14. *Compute an explicit value for c in Lemma 2.13 when d the degree of α.*

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 2.13 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 (2.12) are Liouville numbers: let $a \geq 2$ be an integer and $\kappa > 0$ a real number. For sufficiently large N, set

$$q = a^{N!}, \quad p = \sum_{n=0}^{N} a^{N!-n!}.$$

Then we have

$$0 < \theta_a - \frac{p}{q} = \sum_{k \geq 1} \frac{1}{a^{(N+k)!-N!}}\cdot$$

For $k \geq 1$ we use the crude estimate

$$(N+k)! - N! \geq N!N(N+1)\cdot(N+k-1) \geq N!\bigl(N+(k-1)!\bigr),$$

which yields

$$0 < \theta_a - \frac{p}{q} \leq \frac{e}{q^N}\cdot$$

We shall discuss the development of this topic in the next subsection.

After the contributions of Ch. Hermite in 1873, F. Lindemann in 1882 and the Theorem of Hermite Lindemann 2.1, K. Weierstraß completed in 1888 the proof of a claim by Lindemann:

Theorem 2.15 (Lindemann–Weierstraß – first form)**.** *Let $\alpha_1, \dots, \alpha_m$ be algebraic numbers which are pairwise distinct: $\alpha_i \neq \alpha_j$ for $i \neq j$. Then the numbers $e^{\alpha_1}, \dots, e^{\alpha_m}$ are linearly independent over $\mathbb{Q}$.*

It is easy to checked that Theorem 2.15 is equivalent to the next statement:

Theorem 2.16 (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 2.15 is also equivalent to the next statement:

Theorem 2.17 (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 $\overline{\mathbb{Q}}$.*

This does not cover all the history of transcendental numbers in the XIX-th Century. In particular the work of Cantor is another main contribution which gave rise to many development in the XX-th Century.

References

[1] L. Euler – *De fractionibus continuis dissertatio*, Commentarii Acad. Sci. Petropolitanae, 9 (1737), 1744, p. 98–137; Opera Omnia Ser. I vol. 14, Commentationes Analyticae, p. 187–215.

[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.