Updated: *June 1, 2010*

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°13, *May 31, 2010*

These are informal notes of my course given in April – June 2010 at IMPA (*Instituto Nacional de Matematica Pura e Aplicada*), Rio de Janeiro, Brazil.

8 Hermite's method

The proofs given in subsection 1.5 of the irrationality of e^r for several rational values of r (namely $r \in \{1, 2, \sqrt{2}, \sqrt{3}\}$) 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 [3], is to approximate e^z by rational functions $A(z)/B(z)$. The word "approximate" has the following meaning (Hermite-Padé): in a loose sense, an analytic function is *well approximated* by a rational function $A(z)/B(z)$ (where A and B are polynomial) if *the first* coefficients of the Taylor expansion of $f(z)$ and $A(z)/B(z)$ at the origin are the same. When $B(0) \neq 0$, this amounts to asking that 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 (section § 8.1) is to give, following Hermite, a new proof of 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 [5], and another was derived by Yu. V.Ñesterenko for [2] (*A simple proof of the irrationality of* π. Russ. J. Math. Phys. 13 (2006), no. 4, 473). See also ROBERT BREUSCH, *A Proof of the Irrationality of* π, The American Mathematical Monthly, Vol. **61**, No. 9 (Nov., 1954), pp. 631-632.

8.1 Irrationality of e^r and π

8.1.1 Irrationality of e^r for $r \in \mathbf{Q}$

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 the usual irrationality criterion (Proposition 4).

Write

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

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, namely at least the degree of A plus the length of the gap.

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

8.1.2 Derivative operators

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

$$f(z) = \sum_{k\geq 0} a_k z^k, \tag{126}$$

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

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

so that

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

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

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

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

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 $\mathbf{C}$ is also the ring of polynomials in D with coefficients in $\mathbf{C}[z]$. In this ring $\mathbf{C}[z][D]$ 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 $T \in \mathbf{C}[t]$ we 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

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

Recalling our function f with the Taylor development (126), 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 Taylor coefficient of z^k at the origin, 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

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

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

Set

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

Then

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

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

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

Also the analytic function R has a zero of multiplicity $N + 1$ at the origin with leading term $T(N+1)z^{N+1}/(N+1)!$.

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

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

Hence

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

Since

$$\frac{n_0!(n_0+n_1-k)!}{n_1!(n_0-k)!k!} \in \mathbf{Z},$$

we deduce $(n_0!/n_1!)A(z) \in \mathbf{Z}[z]$.

For $k \ge N+1$ 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 127 (Hermite's formulae for the exponential function)**.** *Let $n_0 \ge 0$, $n_1 \ge 0$ be two integers. Define $N = n_0 + n_1$. Set*

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

Finally, define $B \in \mathbf{Z}[z]$ 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.

Remark. *For $n_1 < n_0$ the leading coefficient $(-1)^{n_1}n_1!/n_0!$ of A is not an integer, but for $n_1 \ge n_0$ the coefficients of A are integers.*

We check the following elementary estimate for the remainder.

Lemma 128. *Let $z \in \mathbf{C}$. Then*

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

Proof. We have

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

The trivial estimate

$$\frac{(\ell+N+1)!}{(\ell+n_1)!} = (\ell+n_0+n_1+1)(\ell+n_0+n_1)\cdots(\ell+n_1+1) \geq n_0!$$

yields the conclusion of Lemma 128. □

We are now able to complete the proof of the irrationality of e^a for a a positive integer (hence, for e^r when $r \in \mathbf{Q}$, $r \neq 0$). We take a large positive integer n and we select $n_0 = n_1 = n$. 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 127. for $n_0 = n_1 = n$.

Replace z by a in the previous formulae; we deduce

$$B_n(a)e^a - A_n(a) = R_n(a).$$

All coefficients in R_n are positive, hence $R_n(a) > 0$. Therefore $B_n(a)e^a - A_n(a) \neq 0$. Lemma 128 shows that $R_n(a)$ tends to 0 when n tends to infinity. Since $B_n(a)$ and $A_n(a)$ are rational integers, we may use the implication (ii)⇒(i) in (Proposition 4): we deduce that the number e^a is irrational.

8.1.3 Irrationality of π

The irrationality of e^r for $r \in \mathbf{Q} \setminus \{0\}$ is equivalent to the irrationality of $\log s$ for $s \in \mathbf{Q}_{>0}$. We extend this proof to $s = -1$ (so to speak) and get the irrationality of 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 $\mathbf{Z}[i]$. The left hand side is in $\mathbf{Z}[i]$, the right hand side tends to 0 as n tends to infinity, hence both sides are 0.

In the proof of § 8.1.1, we used the positivity of the coefficients of R_n and we deduced that $R_n(a)$ was not 0 (this is a simple example of 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

$$\Delta_n = B_n A_{n+1} - B_{n+1} A_n \tag{129}$$

can be written

$$\Delta_n = -B_n R_{n+1} + B_{n+1} R_n. \tag{130}$$

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 (129) 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 (130) 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 (130) 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.

8.2 Padé approximation to the exponential function

For $h \geq 0$, the h-th derivative $D^h R(z)$ of the remainder in Proposition 145 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

$$D^{n_0+1} R = \sum_{k \geq N+1} \frac{z^{k-n_0-1}}{(k-N-1)!} = z^{n_1} e^z. \tag{131}$$

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

8.2.1 Siegel's point of view

Theorem 132. *Given two integers $n_0 \geq 0$, $n_1 \geq 0$, there exist two polynomials A and B in $\mathbf{C}[z]$ 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. Further, A has degree n_0, B has degree n_1 and R has multiplicity $N+1$ at the origin. Furthermore, when B is monic, we have $D^{n_0+1}R = z^{n_1}e^z$.*

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

$$\begin{array}{rccc} \mathcal{L}: & \mathbf{C}[z]_{\leq n_1} & \longrightarrow & \mathbf{C}^{n_1} \\ & B(z) & \longmapsto & \Big(D^\ell\big(B(z)e^z\big)_{z=0}\Big)_{n_0<\ell\leq N} \end{array}$$

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

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

and

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

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

$$B(z)e^z = A(z) + R(z). \tag{133}$$

There is an alternative proof of the existence as follows [5]. Consider the linear mapping

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

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 (133).

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

$$D^{n_0+1}R = cz^{n_1}e^z \tag{134}$$

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 (133) by e^{-z}, we deduce

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

We replace z by $-z$:

$$A(-z)e^{z} = B(-z) - R(-z)e^{z}. \tag{135}$$

It follows that $\big(B(-z), A(-z), -R(-z)e^z\big)$ is a solution to the Padé problem (133) for the parameters (n_1, n_0). Therefore A has degree n_0.

□

Denote by $(A_{n_0,n_1},\ B_{n_0,n_1},\ R_{n_0,n_1})$ the solution to the Padé problem (133) for the parameters (n_0, n_1): the polynomial A has degree n_0 and leading term $n_1!/n_0!$, the polynomial B is monic of degree n_1, and R has a zero of multiplicity $N+1$ at the origin with leading term $n_1!z^{N+1}/(N+1)!$. As before $N = n_0 + n_1$. Then we have

$$\begin{aligned} A_{n_1,n_0}(z) &= (-1)^N \frac{n_0!}{n_1} B_{n_0,n_1}(-z), \\ B_{n_1,n_0}(z) &= (-1)^N \frac{n_0!}{n_1} A_{n_0,n_1}(-z), \\ R_{n_1,n_0}(z) &= (-1)^{N+1} \frac{n_0!}{n_1} R_{n_0,n_1}(-z)e^z. \end{aligned} \tag{136}$$

Following [5], 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 137. *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

$$\sum_{k=0}^{n} \frac{(-1)^{n-k}}{k!(n-k)!}\left(kz^{k-1}\int_0^z t^{n-k}\varphi(t)dt + z^n\varphi(z)\right). \tag{138}$$

Since $n \geq 1$, we have

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

and equation (138) is nothing else than

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

□