Updated: *February 18, 2016*

These are informal notes (205 pages) of my course given in April – June 2010 at IMPA (*Instituto Nacional de Matematica Pura e Aplicada*), Rio de Janeiro, Brazil([1]).

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°1, *April 14, 2010*

1 Introduction

1.1 Irrationality of $\sqrt{2}$

We first give a geometrical proof of the irrationality of the number

$$\sqrt{2} = 1,414\,213\,562\,373\,095\,048\,801\,688\,724\,209\ \ldots$$

Starting with a rectangle having sides 1 and $1+\sqrt{2}$, we split it into two unit squares and a smaller rectangle. The length of this second rectangle is 1, its width is $\sqrt{2}-1$, hence its proportion is

$$\frac{1}{\sqrt{2}-1} = 1+\sqrt{2}.$$

Therefore the first and second rectangles have the same proportion. Now, if we repeat the process and split the small rectangle into two squares (of sides $\sqrt{2}-1$) and a third tiny rectangle, the proportions of this third rectangle will again be $1+\sqrt{2}$. This means that the process will not end, each time we shall get two squares and a remaining smaller rectangle having the same proportion.

[1] This text is available on the internet at the address
http://www.math.jussieu.fr/∼miw/articles/pdf/IMPA2010.pdf

On the other hand, if we start with a rectangle having integer side–lengths, if we split it into several squares and if a small rectangle remains, then clearly the small rectangle while have integer side–lengths([2]). Therefore the process will not continue forever, it will stop when there is no remaining small rectangle. This proves the irrationality of $\sqrt{2}$.

In algebraic terms, the number $x = 1 + \sqrt{2}$ satisfies

$$x = 2 + \frac{1}{x},$$

hence also

$$x = 2 + \frac{1}{2 + \dfrac{1}{x}} = 2 + \frac{1}{2 + \dfrac{1}{2 + \dfrac{1}{x}}} = \cdots,$$

which yields the *continued fraction expansion* of $1 + \sqrt{2}$.

1.2 Continued fractions

Here is the definition of the continued fraction expansion of a real number.

Given a real number x, the Euclidean division in $\mathbf{R}$ of x by 1 yields a quotient $\lfloor x \rfloor \in \mathbf{Z}$ (the *integral part of* x) and a remainder $\{x\}$ in the interval ¡ù($[0, 1)$ (the *fractional part of* x) satisfying

$$x = \lfloor x \rfloor + \{x\}.$$

Set $a_0 = \lfloor x \rfloor$. Hence $a_0 \in \mathbf{Z}$. If x is an integer then $x = \lfloor x \rfloor = a_0$ and $\{x\} = 0$. In this case we just write $x = a_0$ with $a_0 \in \mathbf{Z}$. Otherwise we have $\{x\} > 0$ and we set $x_1 = 1/\{x\}$ and $a_1 = \lfloor x_1 \rfloor$. Since $\{x\} < 1$ we have $x_1 > 1$ and $a_1 \geq 1$. Also

$$x = a_0 + \frac{1}{a_1 + \{x_1\}}.$$

Again, we consider two cases: if $x_1 \in \mathbf{Z}$ then $\{x_1\} = 0$, $x_1 = a_1$ and

$$x = a_0 + \frac{1}{a_1}$$

[2]Starting with a rectangle of side–lengths a and b, the process stops when a square of side–length d is reached, where d is the gcd of a and b: also d is the largest positive integer such that the initial rectangle can be covered with square tiles of side length d.

with two integers a_0 and a_1, with $a_1 \geq 2$ (recall $x_1 > 1$). Otherwise we can define $x_2 = 1/\{x_1\}$, $a_2 = \lfloor x_2 \rfloor$ and go one step further:

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \{x_2\}}}.$$

Inductively one obtains a relation

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \cfrac{}{a_{n-1} + \cfrac{1}{a_n + \{x_n\}}}}}}$$

with $0 \leq \{x_n\} < 1$. The connexion with the geometric proof of irrationality of $\sqrt{2}$ by means of rectangles and squares is now obvious: start with a positive real number x and consider a rectangle of sides 1 and x. Divide this rectangle into unit squares and a second rectangle. Then a_0 is the number of unit squares which occur, while the sides of the second rectangle are 1 and $\{x\}$. If x is not an integer, meaning $\{x\} > 0$, then we split the second rectangle into squares of sides $\{x\}$ plus a third rectangle. The number of squares is now a_1 and the third rectangle has sides $\{x\}$ and $1 - a_1\{x\}$. Going one in the same way, one checks that the number of squares we get at the n-th step is a_n.

This geometric point of view shows that the process stops after finitely many steps (meaning that some $\{x_n\}$ is zero, or equivalently that x_n is in $\mathbf{Z}$) if and only if x is rational.

For simplicity of notation, when $x_0, x_1, \ldots, x_n$ are real numbers with $x_1, \ldots, x_n$ positive, we write

$$x = [x_0,\ x_1, \ldots, x_n] \quad \text{for} \quad x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots \cfrac{}{x_{n-1} + \cfrac{1}{x_n}}}}}.$$

When $a_0, a_1, \ldots, a_n$ are integers with $a_1, \ldots, a_n$ positive, then $[a_0,\ a_1, \ldots, a_n]$ is a rational number. Conversely, given a rational number x, the previous algorithm produces a finite continued fraction $[a_0,\ a_1, \ldots, a_n]$ where $a_0 = \lfloor x \rfloor$ and $a_i > 0$ $(1 \leq i \leq n)$ are integers. If x is a rational integer, then $n = 0$,

$a_0 = x$ and the continued fraction which is produced by this algorithm is $x = [a_0]$. If x is not an integer, then $n \geq 1$ and $a_n \geq 2$. For any rational number, there are exactly two finite continued fractions equal to x: one, say $[a_0,\ a_1, \ldots, a_{n-1}, a_n]$, is given by the previous algorithm, the other one is $[a_0,\ a_1, \ldots, a_{n-1}, a_n - 1, 1]$. For instance if x is an integer the continued fraction produced by the algorithm is $[x]$, as we just saw, while the other continued fraction equal to x is $[x-1, 1]$. The two continued fractions equal to 1 are $[1]$ and $[0, 1]$, while any positive rational number distinct from 1 has one continued fraction expansion with the last term $a_n \geq 2$ and one with the last term 1.

When x is irrational, we write the continued fraction as $[a_0,\ a_1, \ldots, a_n, \ldots]$. We shall check later that when $a_0, a_1, \ldots, a_n, \ldots$ are integers with $a_1, \ldots, a_n, \ldots$ positive, the limit of $[a_0,\ a_1, \ldots, a_n]$ exists and is equal to x.

We need a further notation for ultimately periodic continued fraction. Assume that x is irrational and that for some integers n_0 and $r > 0$ its continued fraction expansion $[a_0,\ a_1, \ldots, a_n, \ldots]$ satisfies

$$a_{n+r} = a_n \quad \text{for any } n \geq n_0.$$

Then we write

$$x = [a_0,\ a_1, \ldots, a_{n_0-1}, \overline{a_{n_0},\ a_{n_0+1}, \ldots, a_{n_0+r-1}}].$$

For instance

$$\sqrt{2} = [1,\ 2,\ 2,\ 2, \ldots] = [1,\ \overline{2}]$$

and

$$\sqrt{3} = [1,\ 1,\ 2,\ 1,\ 2,\ 1,\ 2, \ldots] = [1,\ \overline{1, 2}].$$

References on continued fractions are [11, 31, 19, 23, 4]. An interesting remark [29] on the continued fraction expansion of $\sqrt{2}$ is to relate the A4 paper format 21×29.7 to the fraction expansion

$$\frac{297}{210} = \frac{99}{70} = [1,\ 2,\ 2,\ 2,\ 2,\ 2].$$

There is nothing special with the square root of 2: most of the previous argument extend to the proof of irrationality of $\sqrt{n}$ when n is a positive integer which is not the square of an integer. For instance, a proof of the irrationality of $\sqrt{n}$ when n is not the square of an integer runs as follows. Write $\sqrt{n} = a/b$ where b is the smallest positive integer such that $b\sqrt{n}$ is an integer. Further, denote by m the integral part of $\sqrt{n}$: this means that m is the positive integer such that $m < \sqrt{n} < m + 1$. The strict inequality

$m < \sqrt{n}$ is the assumption that n is not a square. From $0 < \sqrt{n} - m < 1$ one deduces

$$0 < (\sqrt{n} - m)b < b.$$

Now the number

$$b' := (\sqrt{n} - m)b = a - mb$$

is a positive rational integer, the product

$$b'\sqrt{n} = bn - am$$

is an integer and $b' < b$, which contradicts the choice of b minimal.

The irrationality of $\sqrt{5}$ is equivalent to the irrationality of the *Golden ratio* $\Phi = (1 + \sqrt{5})/2$, root of the polynomial $X^2 - X - 1$, whose continued fraction expansion is

$$\Phi = [1,\ 1,\ 1,\ 1 \ldots] = [\overline{1}].$$

This continued fraction expansion follows from the relation

$$\Phi = 1 + \frac{1}{\Phi}.$$

The geometric irrationality proof using rectangles that we described above for $1 + \sqrt{2}$ works in a similar way for the Golden ratio: a rectangle of sides Φ and 1 splits into a square and a small rectangle of sides 1 and $\Phi - 1$, hence the first and the second rectangles have the same proportion, namely

$$\Phi = \frac{1}{\Phi - 1}. \tag{1}$$

Therefore the process continues forever with one square and one smaller rectangle with the same proportion. Hence Φ and $\sqrt{5}$ are irrational numbers.

Exercise 1. *(a) Check that, in the geometric construction of splitting a rectangle of sides* 1 *and* x *into squares and rectangles, the number of successive squares is the sequence of integers* $(a_n)_{n \geq 0}$ *in the continued fraction expansion of* x*.*
(b) Start with a unit square. Put on top of it another unit square: you get a rectangle with sides 1 *and* 2*. Next put on the right a square of sides* 2*, which produces a rectangle with sides* 2 *and* 3*. Continue the process as follows: when you reach a rectangle of small side* a *and large side* b*, complete it with a square of sides* b*, so that you get a rectangle with sides* b *and* $a + b$*. Which is the sequence of sides of the rectangles you obtain with this process?*

Generalizing this idea, given positive integers $a_0,\ a_1, \ldots, a_k$, devise a geometrical construction of the positive rational number having the continued fraction expansion

$$[a_0,\ a_1, \ldots, a_k].$$

Another proof of the irrationality of Φ is to deduce from the equation (1) that a relation $\Phi = a/b$ with $0 < b < a$ yields

$$\Phi = \frac{b}{a-b},$$

hence a/b is not a rational fraction with minimal denominator.

1.3 Irrational numbers

If k is a positive integer and n a positive integer which is not the k-th power of a rational integer, then the number $n^{1/k}$ is irrational. This follows, for instance, from the fact that the roots of $X^k - n$ are algebraic integers, and algebraic integers which are rational numbers are rational integers.

Other numbers for which it is easy to prove the irrationality are quotients of logarithms: if m and n are positive integers such that $(\log m)/(\log n)$ is rational, say a/b, then $m^b = n^a$, which means that *m and n are multiplicatively dependent*. Recall that elements $x_1, \ldots, x_r$ in an additive group are *linearly independent* if a relation $a_1x_1 + \cdots + a_rx_r = 0$ with rational integers $a_1, \ldots, a_r$ implies $a_1 = \cdots = a_r = 0$. Similarly, elements $x_1, \ldots, x_r$ in a multiplicative group are *multiplicatively independent* if a relation $x_1^{a_1} \cdots x_r^{a_r} = 1$ with rational integers $a_1, \ldots, a_r$ implies $a_1 = \cdots = a_r = 0$. Therefore a quotient like $(\log 2)/\log 3$, and more generally $(\log m)/\log n$ where m and n are multiplicatively independent positive rational numbers, is irrational.

We have seen that *a real number is rational if and only if its continued fraction expansion is finite*. There is another criterion of irrationality using the b-adic expansion when b is an integer ≥ 2 (for $b = 10$ this is the decimal expansion, for $b = 2$ it is the diadic expansion). Indeed any real number x can be written

$$x = \lfloor x \rfloor + d_1b^{-1} + d_2b^{-2} + \cdots + d_nb^{-n} + \cdots$$

where the integers d_n (the digits of x) are in the range $0 \leq d_n < b$. There is unicity of such an expansion, unless x is an integral multiple of some b^{-n} with $n \geq 0$, in which case x has two expansions: one where all sufficiently large digits vanish, and one for which all sufficiently large digits are $b - 1$.

This is due to the equation

$$b^{-n} = \sum_{k=0}^{n} (b-1) b^{-n-k-1}.$$

Here is the irrationality criterion using such expansions: fix an integer $b \geq 2$. Then *the real number x is rational if and only if the sequence of digits $(d_n)_{n\geq 1}$ of x in basis b is ultimately periodic.*

Exercise 2. *Let $b \geq 2$ be an integer.*
(a) Show that a real number x is rational if and only if the sequence $(d_n)_{n\geq 1}$ of digits of x in the expansion in basis b

$$x = \lfloor x \rfloor + d_1 b^{-1} + d_2 b^{-2} + \cdots + d_n b^{-n} + \cdots \qquad (0 \leq d_n < b)$$

is ultimately periodic.
(b) Let $(u_n)_{n\geq 0}$ be an increasing sequence of positive integers. Assume there exists $c > 0$ such that, for all sufficiently large n,

$$u_n - u_{n-1} \geq cn.$$

Deduce from (a) that the number

$$\vartheta = \sum_{n\geq 0} b^{-u_n}$$

is irrational.

One might be tempted to conclude that it should be easy to decide whether a given real number is rational or not. However this is not the case with many constants from analysis, because most often one does not know any expansion, either in continued fraction or in any basis $b \geq 2$. And the fact is that for many such constants the answer is not known. For instance, one does not know whether the *Euler–Mascheroni constant*

$$\begin{aligned} \gamma &= \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n \right) \\ &= 0,577\,215\,664\,901\,532\,860\,606\,512\,090\,082\ldots \end{aligned}$$

is rational or not: one expects that it is an irrational number (and even a transcendental number - see later). Other formulas for the same number are

$$\begin{aligned}\gamma &= \sum_{k=1}^{\infty}\left(\frac{1}{k}-\log\left(1+\frac{1}{k}\right)\right)\\ &= \int_1^{\infty}\left(\frac{1}{\lfloor x\rfloor}-\frac{1}{x}\right)dx\\ &= -\int_0^1\int_0^1\frac{(1-x)dxdy}{(1-xy)\log(xy)}.\end{aligned}$$

J. Sondow uses (a generalization of) the last double integral in [35], he was inspired by F. Beukers' work on Apéry's proof of the irrationality of

$$\zeta(3)=\sum_{n\geq 1}\frac{1}{n^3}=1,202\,056\,903\,159\,594\,285\,399\,738\,161\,511\ldots$$

in 1978. Recall that the values of the *Riemann zeta function*

$$\zeta(s)=\sum_{n\geq 1}n^{-s}$$

was considered by Euler for real s and by Riemann for complex s, the series being convergent for the real part of s greater than 1. Euler proved that the values $\zeta(2k)$ of this function at the even positive integers ($k\in\mathbf{Z}$, $k\geq 1$) are rational multiples of π^{2k}. For instance, $\zeta(2)=\pi^2/6$. It is interesting to notice that Euler's proof relates the values $\zeta(2k)$ at the positive even integers with the values of the same function at the odd negative integers, namely $\zeta(1-2k)$. For Euler this involved divergent series, while Riemann defined $\zeta(s)$ for $s\in\mathbf{C}$, $s\neq 1$, by analytic continuation.

One might be tempted to guess that $\zeta(2k+1)/\pi^{2k+1}$ is a rational number when $k\geq 1$ is a positive integer. However the folklore conjecture is that this is not the case. In fact there are good reasons to conjecture that for any $k\geq 1$ and any non-zero polynomial $P\in\mathbf{Z}[X_0,X_1,\ldots,X_k]$, the number $P(\pi,\zeta(3),\zeta(5),\ldots,\zeta(2k+1))$ is not 0. But one does not know whether

$$\zeta(5)=\sum_{n\geq 1}\frac{1}{n^5}=1,036\,927\,755\,143\,369\,926\,331\,365\,486\,457\ldots$$

is irrational or not. And there is no proof so far that $\zeta(3)/\pi^3$ is irrational. According to T. Rivoal, among the numbers $\zeta(2n+1)$ with $n\geq 2$, infinitely

many are irrational. And W. Zudilin proved that one at least of the four numbers

$$\zeta(5), \zeta(7), \zeta(9), \zeta(11)$$

is irrational. References with more information on this topic are given in the Bourbaki talk [14] by S. Fischler.

A related open question is the arithmetic nature of *Catalan's constant*

$$G = \sum_{n\geq 1} \frac{(-1)^n}{(2n+1)^2} = 0,915\,965\,594\,177\,219\,015\,0\ldots$$

Other open questions can be asked on the values of *Euler's Gamma fonction*

$$\Gamma(z) = e^{-\gamma z} z^{-1} \prod_{n=1}^{\infty} \left(1+\frac{z}{n}\right)^{-1} e^{z/n} = \int_0^{\infty} e^{-t} t^z \cdot \frac{dt}{t}\cdot$$

As an example we do not know how to prove that the number

$$\Gamma(1/5) = 4,590\,843\,711\,998\,803\,053\,204\,758\,275\,929\,152\,0\ldots$$

is irrational.

The only rational values of z for which the answer is known (and in fact one knows the transcendence of the Gamma value in these cases) are

$$r \in \left\{\frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}\right\} \qquad (\text{mod } 1).$$

The number $\Gamma(1/n)$ appears when one computes *periods* of the Fermat curve $X^n + Y^n = Z^n$, and this curve is *simpler* (in technical terms it has genus ≤ 1) for $n = 2, 3, 4$ and 6. For $n = 5$ the genus is 2 and this is related with the fact that one is not able so far to give the answer for $\Gamma(1/5)$.

The list of similar open problems is endless. For instance, is the number

$$e + \pi = 5,859\,874\,482\,048\,838\,473\,822\,930\,854\,632\ldots$$

rational or not? The answer is not yet known. And the same is true for any number in the following list

$$\log \pi,\ 2^{\pi},\ 2^{e},\ \pi^{e},\ e^{e}.$$

1.4 History of irrationality

The history of irrationality is closely connected with the history of continued fractions (see[2, 3]). (Even the first examples of transcendental numbers produced by Liouville in 1844 involved continued fractions, before he considered series).

The question of the irrationality of π was raised in India by Nīlakaṇṭha Somayājī, who was born around 1444 AD. In his comments on the work of Āryabhaṭa, (b. 476 AD) who stated that an approximation for π is $\pi \sim 3.1416$, Somayājī asks([3]):

> *Why then has an approximate value been mentioned here leaving behind the actual value? Because it (exact value) cannot be expressed.*

In 1767, H. Lambert [20] proved that for x rational and non–zero, the number $\tan x$ cannot be rational. Since $\tan \pi/4 = 1$ it follows that π is irrational. Then he produced a continued fraction expansion for e^x and deduced that e^r is irrational when r is a non–zero rational number. This is equivalent to the fact that non–zero positive rational numbers have an irrational logarithm. A detailed description of Lambert's proof is given in [12].

Euler gave continued fractions expansions not only for e and e^2:

$$e = [2;\ \overline{1, 2j,\ 1}]_{j\geq 1} = [2;\ 1,\ 2,\ 1,\ 1,\ 4,\ 1,\ 1,\ 6,\ 1, \dots],$$
$$e^2 = [7;\ \overline{3j-1,\ 1,\ 1,\ 3j,\ 12j+6}]_{j\geq 1} = [7;\ 2,\ 1,\ 1,\ 3,\ 18,\ 5,\ 1,\ 1,\ 6,\ 30,\ 8, \dots],$$

but also for $(e+1)/(e-1)$, for $(e^2+1)/(e^2-1)$, for $e^{1/n}$ with $n > 1$, for $e^{2/n}$ with odd $n > 1$ and Hurwitz (1896) for $2e$ and $(e+1)/3$:

$$\frac{e+1}{e-1} = [\overline{2(2j+1)}]_{j\geq 0} = [2;\ 6,\ 10,\ 14, \dots],$$
$$\frac{e^2+1}{e^2-1} = [\overline{2j+1}]_{j\geq 0} = [1;\ 3,\ 5,\ 7, \dots],$$
$$e^{1/n} = [\overline{1,\ (2j+1)n-1,\ 1}]_{j\geq 0} \quad \text{for} \quad n \geq 2,$$
$$e^{2/n} = [\overline{1,\ (n-1)/2+3jn,\ 6n+12jn,\ (5n-1)/2+3jn,\ 1}]_{j\geq 0} \quad \text{for odd} \quad n \geq 3,$$
$$2e = [5,\ 2,\ \overline{3,\ 2j,\ 3,\ 1,\ 2j,\ 1}]_{j\geq 1},$$
$$\frac{e+1}{3} =$$
$$[1,\ 4,\ \overline{5,\ 4j-3,\ 1,\ 1,\ 36j-16,\ 1,\ 1,\ 4j-2,\ 1,\ 1,\ 36j-4,\ 1,\ 1,\ 4j-1,\ 1,\ 5,\ 4j,\ 1}]_{j\geq 1}.$$

[3] K. Ramasubramanian, *The Notion of Proof in Indian Science*, 13th World Sanskrit Conference, 2006. http://www.iitb.ac.in/campus/diary/2006/august/tday2.htm

Hermite proved the irrationality of π and π^2 (see [3] p. 207 and p. 247). Furthermore, A.M. Legendre proved, in 1794, by a modification of Lambert's proof, that π^2 is also an irrational number (see [3] p. 14).

There are not so many numbers for which one knows the irrationality but we don't know whether they are algebraic or transcendental ([4]). A notable exception is $\zeta(3)$, known to be irrational (Apéry, 1978) and expected to be transcendental.

1.5 Variation on a proof by Fourier (1815)

That e is not quadratic follows from the fact that the continued fraction expansion of e, which was known by L. Euler in 1737 [11, 7, 32, 36], is not periodic:

$$e = [2,\ 1,\ 2,\ 1,\ 1,\ 4,\ 1,\ 1,\ 6,\ 1,\ 1,\ 8, \dots]$$

Since this expansion is infinite we deduce that e is irrational. The fact that it is not ultimately periodic implies also that e is not a quadratic irrationality, as shown by Lagrange in 1770 – Euler knew already in 1737 that a number with an ultimately periodic continued fraction expansion is quadratic (see [11, 4, 31]).

The following easier and well known proof of the irrationality of e was given by J. Fourier in his course at the École Polytechnique in 1815. Later, in 1872 , C. Hermite proved that e is transcendental, while the work of F. Lindemann a dozen of years later led to a proof of the so-called Hermite–Lindemann Theorem: *for any nonzero algebraic number α the number e^α is transcendental.* However for this first section we study only weaker statements which are very easy to prove. We also show that Fourier's argument can be pushed a little bit further than what is usually done, as pointed out by J. Liouville in 1840.

1.5.1 Irrationality of e

We truncate the exponential series giving the value of e at some point N:

$$N!\ e - \sum_{n=0}^{N} \frac{N!}{n!} = \sum_{k\geq 1} \frac{N!}{(N+k)!}. \tag{2}$$

[4] Unless one considers complex numbers of the form ix where x is a real number expected to be transcendental, but for which no proof of irrationality is known: there are plenty of them!

The right hand side of (2) is a sum of positive numbers, hence is positive (not zero). From the lower bound (for the binomial coefficient)

$$\frac{(N+k)!}{N!k!} \geq N+1 \quad \text{for } k \geq 1,$$

one deduces

$$\sum_{k\geq 1} \frac{N!}{(N+k)!} \leq \frac{1}{N+1} \sum_{k\geq 1} \frac{1}{k!} = \frac{e-1}{N+1}.$$

Therefore the right hand side of (2) tends to 0 when N tends to infinity. In the left hand side, $\sum_{n=0}^{N} N!/n!$ is an integer. It follows that for any integer $N \geq 1$ the number $N!e$ is not an integer, hence e is an irrational number.

1.5.2 Irrationality of e^{-1}, following C.L. Siegel

In 1949, in his book on transcendental numbers [34], C.L. Siegel simplified the proof by Fourier: considering e^{-1} instead of e yields alternating series, hence it is no more necessary to estimate the remainder term.

The sequence $(1/n!)_{n\geq 0}$ is decreasing and tends to 0, hence for odd N,

$$1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{1}{(N-1)!} - \frac{1}{N!} < e^{-1} < 1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{1}{(N+1)!}.$$

Multiply by $N!$; the left hand side becomes

$$a_N := N! - \frac{N!}{1!} + \frac{N!}{2!} - \cdots + \frac{N!}{(N-1)!} - \frac{N!}{N!} \in \mathbf{Z},$$

while the right hand side becomes

$$a_N + \frac{1}{N+1} < a_N + 1.$$

Hence $0 < N!e^{-1} - a_N < 1$, and therefore $N!e^{-1}$ is not an integer.

1.5.3 The number e is not quadratic

.

The fact that e is not a rational number implies that for each $m \geq 1$ the number $e^{1/m}$ is not rational. To prove that e^2, for instance, is also irrational is not so easy (see the comment on this point in [1]).

The proof below is essentially the one given by J. Liouville in 1840 [25] which is quoted by Ch. Hermite [17] ("ces travaux de l'illustre géomètre").

To prove that e does not satisfy a quadratic relation ae^2+be+c with a, b and c rational integers, not all zero, requires some new trick. Indeed if we just mimic the same argument we get

$$cN!+\sum_{n=0}^{N}\left(2^n a+b\right)\frac{N!}{n!}=-\sum_{k\geq 0}\left(2^{N+1+k}a+b\right)\frac{N!}{(N+1+k)!}\cdot$$

The left hand side is a rational integer, but the right hand side tends to infinity (and not 0) with N, so we draw no conclusion.

Instead of this approach, Liouville writes the quadratic relation as $ae+b+ce^{-1}=0$. This time it works:

$$bN!+\sum_{n=0}^{N}\left(a+(-1)^n c\right)\frac{N!}{n!}=-\sum_{k\geq 0}\left(a+(-1)^{N+1+k}c\right)\frac{N!}{(N+1+k)!}\cdot$$

Again the left hand side is a rational integer, but now the right hand side tends to 0 when N tends to infinity, which is what we expected. However we need a little more work to conclude: we do not yet get the desired conclusion; we only deduce that both sides vanish. Now let us look more closely to the series in the right hand side. Write the two first terms A_N for $k=0$ and B_N for $k=1$:

$$\sum_{k\geq 0}\left(a+(-1)^{N+1+k}c\right)\frac{N!}{(N+1+k)!}=A_N+B_N+C_N$$

with

$$A_N=\left(a-(-1)^N c\right)\frac{1}{N+1},\qquad B_N=\left(a+(-1)^N c\right)\frac{1}{(N+1)(N+2)}$$

and

$$C_N=\sum_{k\geq 2}\left(a+(-1)^{N+1+k}c\right)\frac{N!}{(N+1+k)!}\cdot$$

The above proof that the sum $A_N+B_N+C_N$ tends to zero as N tends to infinity shows more: each of the three sequences

$$A_N,\quad (N+1)B_N,\quad (N+1)(N+2)C_N$$

tends to 0 as N tends to infinity. Hence, from the fact that the sum $A_N+B_N+C_N$ vanishes for sufficiently large N, it easily follows that for sufficiently large N, each of the three terms A_N, B_N and C_N vanishes, hence $a-(-1)^N c$ and $a+(-1)^N c$ vanish, therefore $a=c=0$, and finally $b=0$.

Exercise 3. *Let $(a_n)_{n\geq 0}$ be a bounded sequence of rational integers. Prove that the following conditions are equivalent:*
(i) *The number*

$$\vartheta_1 = \sum_{n\geq 0} \frac{a_n}{n!}$$

is rational.
(ii) *There exists $N_0 > 0$ such that $a_n = 0$ for all $n \geq N_0$.*

1.5.4 The number e^2 is not quadratic

The proof below is the one given by J. Liouville in 1840 [24] . See also [8].

We saw in § 1.5.3 that there was a difficulty to prove that e is not a quadratic number if we were to follow too closely Fourier's initial idea. Considering e^{-1} provided the clue. Now we prove that e^2 is not a quadratic number by truncating the series at carefully selected places. Consider a relation $ae^4 + be^2 + c = 0$ with rational integer coefficients a, b and c. Write $ae^2 + b + ce^{-2} = 0$. Hence

$$\frac{N!b}{2^{N-1}} + \sum_{n=0}^{N} \left(a + (-1)^n c\right) \frac{N!}{2^{N-n-1} n!} = -\sum_{k\geq 0} \left(a + (-1)^{N+1+k} c\right) \frac{2^k N!}{(N+1+k)!}.$$

Like in § 1.5.3, the right hand side tends to 0 as N tends to infinity, and if the two first terms of the series vanish for some value of N, then we conclude $a = c = 0$. What remains to be proved is that the numbers

$$\frac{N!}{2^{N-n-1} n!}, \quad (0 \leq n \leq N)$$

are integers. For $n = 0$ this is the coefficient of b, namely $2^{-N+1} N!$. The fact that these numbers are integers is not true for all values of N, it is not true even for all sufficiently large N; but we do not need so much, it suffices that they are integers for infinitely many N, and that much is true.

The exponent $v_p(N!)$ of p in the prime decomposition of $N!$ is given by the (finite) sum (see, for instance, [16])

$$v_p(N!) = \sum_{j\geq 1} \left\lfloor \frac{N}{p^j} \right\rfloor. \tag{3}$$

Using the trivial upper bound $\lfloor m/p^j \rfloor \leq m/p^j$ we deduce the upper bound

$$v_p(n!) \leq \frac{n}{p-1}$$

for all $n \geq 0$. In particular $v_2(n!) \leq n$. On the other hand, when N is a power of p, say $N = p^t$, then (3) yields

$$v_p(N!) = p^{t-1} + p^{t-2} + \cdots + p + 1 = \frac{p^t - 1}{p - 1} = \frac{N - 1}{p - 1}.$$

Therefore when N is a power of 2 the number $N!$ is divisible by 2^{N-1} and we have, for $0 \leq m \leq N$,

$$v_2(N!/n!) \geq N - n - 1,$$

which means that the numbers $N!/2^{N-n-1}n!$ are integers.

Exercise 4. *(Continuation of exercise 3). Let $(a_n)_{n\geq 0}$ be a bounded sequence of rational integers. Prove that these properties are also equivalent to*
(iii) *The number*

$$\vartheta_2 = \sum_{n\geq 0} \frac{a_n 2^n}{n!}$$

is rational.

Exercise 5. *Prove that $e^{\sqrt{2}}$ is an irrational number.*
`Hint.` Prove the stronger result that $e^{\sqrt{2}} + e^{-\sqrt{2}}$ is irrational.
Prove also the irrationality of $e^{\sqrt{3}}$.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°2, *April 19, 2010*

2 Irrationality Criteria

2.1 Statement of a criterion

Proposition 4. *Let ϑ be a real number. The following conditions are equivalent:*
(i) *ϑ is irrational.*
(ii) *For any $\epsilon > 0$, there exists $(p, q) \in \mathbf{Z}^2$ such that $q > 0$ and*

$$0 < |q\vartheta - p| < \epsilon.$$

(iii) *For any $\epsilon > 0$, there exist two linearly independent linear forms in two variables*

$$L_0(X_0, X_1) = a_0X_0 + b_0X_1 \quad \textit{and} \quad L_1(X_0, X_1) = a_1X_0 + b_1X_1,$$

with rational integer coefficients, such that

$$\max\left\{|L_0(1, \vartheta)|\ ,\ |L_1(1, \vartheta)|\right\} < \epsilon.$$

(iv) *For any real number $Q > 1$, there exists an integer q in the range $1 \leq q < Q$ and a rational integer p such that*

$$0 < |q\vartheta - p| < \frac{1}{Q}\cdot$$

(v) *There exist infinitely many $p/q \in \mathbf{Q}$ such that*

$$\left|\vartheta - \frac{p}{q}\right| < \frac{1}{q^2}\cdot$$

(vi) *There exist infinitely many $p/q \in \mathbf{Q}$ such that*

$$\left|\vartheta - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}\cdot$$

The implication (vi) $\Rightarrow$ (v) is trivial. We shall prove (i) $\Rightarrow$ (vi) later (in the section on continued fractions). We now prove the equivalence between the other conditions of Proposition 4 as follows:

$$\text{(iv)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(i)} \Rightarrow \text{(iv)} \Rightarrow \text{(v) and (v)} \Rightarrow \text{(ii)}.$$

Notice that given a positive integer q, there is at most one value of p such that $|q\vartheta - p| < 1/2$, namely the nearest integer to $q\vartheta$. Hence, when we approximate ϑ by a rational number p/q, we have only one free parameter in $\mathbf{Z}_{>0}$, namely q.

In condition (v), there is no need to assume that the left hand side is not 0: if one $p/q \in \mathbf{Q}$ produces 0, then all other ones do not, and there are again infinitely many of them.

Proof of (iv) $\Rightarrow$ (ii). Using (iv) with Q satisfying $Q > 1$ and $Q \geq 1/\epsilon$, we get (ii). □

Proof of (v) $\Rightarrow$ (ii). According to (v), there is an infinite sequence of distinct rational numbers $(p_i/q_i)_{i\geq 0}$ with $q_i > 0$ such that

$$\left|\vartheta - \frac{p_i}{q_i}\right| < \frac{1}{\sqrt{5}q_i^2}.$$

For each q_i, there is a single value for the numerator p_i for which this inequality is satisfied. Hence the set of q_i is unbounded. Taking $q_i \geq 1/\epsilon$ yields (ii). □

Proof of (ii) $\Rightarrow$ (iii). Let $\epsilon > 0$. From (ii) we deduce the existence of $(p, q) \in \mathbf{Z} \times \mathbf{Z}$ with $q > 0$ and $\gcd(p, q) = 1$ such that

$$0 < |q\vartheta - p| < \epsilon.$$

We use (ii) once more with ϵ replaced by $|q\vartheta - p|$. There exists $(p', q') \in \mathbf{Z} \times \mathbf{Z}$ with $q' > 0$ such that

$$0 < |q'\vartheta - p'| < |q\vartheta - p|. \tag{5}$$

Define $L_0(X_0, X_1) = pX_0 - qX_1$ and $L_1(X_0, X_1) = p'X_0 - q'X_1$. It only remains to check that $L_0(X_0, X_1)$ and $L_1(X_0, X_1)$ are linearly independent. Otherwise, there exists $(s, t) \in \mathbf{Z}^2 \setminus (0, 0)$ such that $sL_0 = tL_1$. Hence $sp = tp'$, $sq = tq'$, and $p/q = p'/q'$. Since $\gcd(p, q) = 1$, we deduce $t = 1$, $p' = sp$, $q' = sq$ and $q'\vartheta - p' = s(q\vartheta - p)$. This is not compatible with (5). □

Proof of (iii) $\Rightarrow$ (i). Assume $\vartheta \in \mathbf{Q}$, say $\vartheta = a/b$ with $\gcd(a,b) = 1$ and $b > 0$. For any non–zero linear form $L \in \mathbf{Z}X_0 + \mathbf{Z}X_1$, the condition $L(1,\vartheta) \neq 0$ implies $|L(1,\vartheta)| \geq 1/b$, hence for $\epsilon = 1/b$ condition (iii) does not hold.

□

Proof of (i) $\Rightarrow$ (iv) *using Dirichlet's box principle.* Let $Q > 1$ be a given real number. Define $N = \lceil Q \rceil$: this means that N is the integer such that $N - 1 < Q \leq N$. Since $Q > 1$, we have $N \geq 2$.

Let $\vartheta \in \mathbf{R} \setminus \mathbf{Q}$. Consider the subset E of the unit interval $[0,1]$ which consists of the $N+1$ elements

$$0,\ \{\vartheta\},\ \{2\vartheta\},\ \{3\vartheta\},\ \ldots, \{(N-1)\vartheta\},\ 1.$$

Since ϑ is irrational, these $N+1$ elements are pairwise distinct. Split the interval $[0,1]$ into N intervals

$$I_j = \left[\frac{j}{N}\,,\,\frac{j+1}{N}\right] \quad (0 \leq j \leq N-1).$$

One at least of these N intervals, say I_{j_0}, contains at least two elements of E. Apart from 0 and 1, all elements $\{q\vartheta\}$ in E with $1 \leq q \leq N-1$ are irrational, hence belong to the union of the *open* intervals $(j/N,\ (j+1)/N)$ with $0 \leq j \leq N-1$.

If $j_0 = N-1$, then the interval

$$I_{j_0} = I_{N-1} = \left[1 - \frac{1}{N}\,,\ 1\right]$$

contains 1 as well as another element of E of the form $\{q\vartheta\}$ with $1 \leq q \leq N-1$. Set $p = \lfloor q\vartheta \rfloor + 1$. Then we have $1 \leq q \leq N-1 < Q$ and

$$p - q\vartheta = \lfloor q\vartheta \rfloor + 1 - \lfloor q\vartheta \rfloor - \{q\vartheta\} = 1 - \{q\vartheta\}, \quad \text{hence} \quad 0 < p - q\vartheta < \frac{1}{N} \leq \frac{1}{Q}.$$

Otherwise we have $0 \leq j_0 \leq N-2$ and I_{j_0} contains two elements $\{q_1\vartheta\}$ and $\{q_2\vartheta\}$ with $0 \leq q_1 < q_2 \leq N-1$. Set

$$q = q_2 - q_1, \quad p = \lfloor q_2\vartheta \rfloor - \lfloor q_1\vartheta \rfloor.$$

Then we have $0 < q = q_2 - q_1 \leq N-1 < Q$ and

$$|q\vartheta - p| = |\{q_2\vartheta\} - \{q_1\vartheta\}| < 1/N \leq 1/Q.$$

□

Remark. Theorem 1.A in Chap. II of [31] states that for any real number ϑ, for any real number $Q > 1$, there exists an integer q in the range $1 \leq q < Q$ and a rational integer p such that

$$\left|\vartheta - \frac{p}{q}\right| \leq \frac{1}{qQ}.$$

The proof given there yields strict inequality $|q\vartheta - p| < 1/Q$ in case Q is not an integer. In the case where Q is an integer and ϑ is rational, the result does not hold with a strict inequality in general. For instance, if $\vartheta = a/b$ with $\gcd(a, b) = 1$ and $b \geq 2$, there is a solution p/q to this problem with strict inequality for $Q = b + 1$, but not for $Q = b$.

However, when Q is an integer and ϑ is irrational, the number $|q\vartheta - p|$ is irrational (recall that $q > 0$), hence not equal to $1/Q$.

Proof of (iv) $\Rightarrow$ (v). Assume (iv). We already know that (iv) $\Rightarrow$ (i), hence ϑ is irrational.

Let $\{q_1, \ldots, q_N\}$ be a finite set of positive integers. We are going to show that there exists a positive integer $q \notin \{q_1, \ldots, q_N\}$ satisfying the condition (v). Denote by $\|\cdot\|$ the distance to the nearest integer: for $x \in \mathbf{R}$,

$$\|x\| = \min_{a \in \mathbf{Z}} |x - a|.$$

Since ϑ is irrational, it follows that for $1 \leq j \leq N$, the number $\|q_j\vartheta\|$ is non–zero. Let $Q > 1$ satisfy

$$Q > \left(\min_{1 \leq j \leq N} \|q_j\vartheta\|\right)^{-1}.$$

From (iv) we deduce that there exists an integer q in the range $1 \leq q < Q$ such that

$$0 < \|q\vartheta_i\| \leq \frac{1}{Q}.$$

The right hand side is $< 1/q$, and the choice of Q implies $q \notin \{q_1, \ldots, q_N\}$. □

In the next section, we give another proof of (i) $\Rightarrow$ (iv) which rests on *Minkowski geometry of numbers*.

2.2 Geometry of numbers

Recall that a discrete subgroup of $\mathbf{R}^n$ of maximal rank n is called a *lattice* of $\mathbf{R}^n$.

Let G be a lattice in $\mathbf{R}^n$. For each basis $\mathbf{e} = \{e_1, \ldots, e_n\}$ of G the parallelogram

$$P_{\mathbf{e}} = \{x_1 e_1 + \cdots + x_n e_n \; ; \; 0 \leq x_i < 1 \; (1 \leq i \leq n)\}$$

is a *fundamental domain* for G, which means a complete system of representative of classes modulo G. We get a partition of $\mathbf{R}^n$ as

$$\mathbf{R}^n = \bigcup_{g \in G} (P_{\mathbf{e}} + g) \tag{6}$$

A change of bases of G is obtained with a matrix with integer coefficients having determinant ± 1, hence the Lebesgue measure $\mu(P_{\mathbf{e}})$ of $P_{\mathbf{e}}$ does not depend on $\mathbf{e}$: this number is called the *volume* of the lattice G and denoted by $v(G)$.

Here is an example of results obtained by H. Minkowski in the XIX–th century as an application of his *geometry of numbers.*

Theorem 7 (Minkowski)**.** *Let G be a lattice in $\mathbf{R}^n$ and B a measurable subset of $\mathbf{R}^n$. Assume $\mu(B) > v(G)$. Then there exist $x \neq y$ in B such that $x - y \in G$.*

Proof. From (6) we deduce that B is the disjoint union of the $B \cap (P_{\mathbf{e}} + g)$ with g running over G. Hence

$$\mu(B) = \sum_{g \in G} \mu\left(B \cap (P_{\mathbf{e}} + g)\right).$$

Since Lebesgue measure is invariant under translation

$$\mu\left(B \cap (P_{\mathbf{e}} + g)\right) = \mu\left((-g + B) \cap P_{\mathbf{e}}\right).$$

The sets $(-g+B) \cap P_{\mathbf{e}}$ are all contained in $P_{\mathbf{e}}$ and the sum of their measures is $\mu(B) > \mu(P_{\mathbf{e}})$. Therefore they are not all pairwise disjoint – this is one of the versions of the *Dirichlet box principle.* There exists $g \neq g'$ in G such that

$$(-g + B) \cap (-g' + B) \neq \emptyset.$$

Let x and y in B satisfy $-g + x = -g' + y$. Then $x - y = g - g' \in G \setminus \{0\}$. □

From Theorem 7 we deduce Minkowski's convex body Theorem (Theorem 2B, Chapter II of [31]).

Corollary 8. *Let G be a lattice in $\mathbf{R}^n$ and let B be a measurable subset of $\mathbf{R}^n$, convex and symmetric with respect to the origin, such that $\mu(B) > 2^n v(G)$. Then $B \cap G \neq \{0\}$.*

Proof. We use Theorem 7 with the set

$$B' = \frac{1}{2}B = \{x \in \mathbf{R}^n \ ; \ 2x \in B\}.$$

We have $\mu(B') = 2^{-n}\mu(B) > v(G)$, hence by Theorem 7 there exists $x \neq y$ in B' such that $x - y \in G$. Now $2x$ and $2y$ are in B, and since B is symmetric $-2y \in B$. Finally B is convex, hence $(2x - 2y)/2 = x - y \in G \cap B \setminus \{0\}$. □

Corollary 9. *With the notations of Corollary 8, if B is also compact in $\mathbf{R}^n$, then the weaker inequality $\mu(B) \geq 2^n v(G)$ suffices to reach the conclusion.*

Proof. Assume $\mu(B) = 2^n v(G)$. For $\epsilon > 0$, set $B_\epsilon = (1+\epsilon)B = \{(1+\epsilon)t \ ; \ t \in B\}$. Since $\mu(B_\epsilon) > 2^n v(G)$, we deduce from Corollary 8 $B_\epsilon \cap G \neq \{0\}$. Since B_ϵ is compact and G discrete, $B_\epsilon \cap G \setminus \{0\}$ is a finite non–empty set. Also

$$B_{\epsilon'} \cap G \subset B_\epsilon \cap G$$

for $\epsilon' < \epsilon$. Hence there exists $t \in G \setminus \{0\}$ such that $t \in B_\epsilon$ for all $\epsilon > 0$. Define $t_\epsilon \in B$ by $t = (1 + \epsilon)t_\epsilon$. Since B is compact, there is a sequence $\epsilon_n \to 0$ such that t_{ϵ_n} has a limit in B. But $\lim_{\epsilon \to 0} t_\epsilon = t$. Hence $t \in B$. □

Remark. The example of $G = \mathbf{Z}^n$ and $B = \left\{(x_1, \ldots, x_n) \in \mathbf{R}^n \ ; \ |x_i| < 1\right\}$ shows how sharp are Corollaries 8 and 9.

Minkowski's Linear Forms Theorem (see, for instance, [31] Chap. II § 2 Th. 2C) is the following result.

Theorem 10 (Minkowski's Linear Forms Theorem)**.** *Suppose that ϑ_{ij} $(1 \leq i, j \leq n)$ are real numbers with determinant ± 1 . Suppose that $A_1, \ldots, A_n$ are positive numbers with $A_1 \cdots A_n = 1$. Then there exists an integer point $\underline{x} = (x_1, \ldots, x_n) \neq 0$ such that*

$$|\vartheta_{i1}x_1 + \cdots + \vartheta_{in}x_n| < A_i \qquad (1 \leq i \leq n-1)$$

and

$$|\vartheta_{n1}x_1 + \cdots + \vartheta_{nn}x_n| \leq A_n.$$

Proof. We apply Corollary 8 with A_n replaced with $A_n + \epsilon$ for a sequence of ϵ which tends to 0. □

Here is a consequence of Theorem 10

Corollary 11. *Let $\vartheta_1, \ldots, \vartheta_m$ be real numbers. For any real number $Q > 1$, there exist $p_1, \ldots, p_m, q$ in $\mathbf{Z}$ such that $1 \leq q < Q$ and*

$$\max_{1 \leq i \leq m} \left| \vartheta_i - \frac{p_i}{q} \right| \leq \frac{1}{qQ^{1/m}}.$$

Proof of Corollary 11. We apply Theorem 10 to the $n \times n$ matrix (with $n = m + 1$)

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\vartheta_1 & 1 & 0 & \cdots & 0 \\ -\vartheta_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\vartheta_m & 0 & 0 & \cdots & 1 \end{pmatrix}$$

corresponding to the linear forms X_0 and $-\vartheta_i X_0 + X_i$ $(1 \leq i \leq m)$, and with $A_0 = Q$, $A_1 = \cdots = A_m = Q^{-1/m}$.

□

Proof of (i) ⇒ (iv) *in Proposition 4 using Minkowski's geometry of numbers.* Let $\epsilon > 0$. The subset

$$\mathcal{C}_\epsilon = \left\{ (x_0, x_1) \in \mathbf{R}^2 \ ; \ |x_0| < Q, \ |x_0\vartheta - x_1| < (1/Q) + \epsilon \right\}$$

of $\mathbf{R}^2$ is convex, symmetric and has volume > 4. By Minkowski's Convex Body Theorem (Corollary 8 below), it contains a non–zero element in $\mathbf{Z}^2$. Since $\mathcal{C}_\epsilon$ is also bounded, the intersection $\mathcal{C}_\epsilon \cap \mathbf{Z}^2$ is finite. Consider a non–zero element (x_0, x_1) in this intersection with $|x_0\vartheta - x_1|$ minimal. Then $(x_0, x_1) \in \mathcal{C}_\epsilon$ for all $\epsilon > 0$, hence $|x_0\vartheta - x_1| \leq 1/Q + \epsilon$ for all $\epsilon > 0$. Since this is true for all $\epsilon > 0$, we deduce $|x_0\vartheta - x_1| \leq 1/Q$. Finally, since ϑ is irrational, we also have $|x_0\vartheta - x_1| \neq 1/Q$. □

2.3 Irrationality of at least one number

Proposition 12. *Let $\vartheta_1, \ldots, \vartheta_m$ be real numbers. The following conditions are equivalent:*
(i) *One at least of $\vartheta_1, \ldots, \vartheta_m$ is irrational.*
(ii) *For any $\epsilon > 0$, there exist $p_1, \ldots, p_m, q$ in $\mathbf{Z}$ with $q > 0$ such that*

$$0 < \max_{1 \leq i \leq m} |q\vartheta_i - p_i| < \epsilon.$$

(iii) *For any* $\epsilon > 0$, *there exist* $m+1$ *linearly independent linear forms* $L_0, \dots, L_m$ *in* $m+1$ *variables with coefficients in* $\mathbf{Z}$ *in* $m+1$ *variables* $X_0, \dots, X_m$, *such that*

$$\max_{0 \le k \le m} |L_k(1, \vartheta_1, \dots, \vartheta_m)| < \epsilon.$$

(iv) *For any real number* $Q > 1$, *there exists* $p_1, \dots, p_m, q$ *in* $\mathbf{Z}$ *such that* $1 \le q < Q$ *and*

$$0 < \max_{1 \le i \le m} |q\vartheta_i - p_i| \le \frac{1}{Q^{1/m}}.$$

(v) *There is an infinite set of* $q \in \mathbf{Z}$, $q > 0$, *for which there exist* $p_1, \dots, p_m$ *in* $\mathbf{Z}$ *satisfying*

$$0 < \max_{1 \le i \le m} \left| \vartheta_i - \frac{p_i}{q} \right| < \frac{1}{q^{1+1/m}}.$$

We shall prove Proposition 12 in the following way:

$$\begin{array}{ccccc} \text{(i)} & \Rightarrow & \text{(iv)} & & \\ & & & \searrow & \\ \Uparrow & & & & \text{(v)} \\ & & & \swarrow & \\ \text{(iii)} & \Leftarrow & \text{(ii)} & & \end{array}$$

Proof of (iv) $\Rightarrow$ (v). We first deduce (i) from (iv). Indeed, if (i) does not hold and $\vartheta_i = a_i/b \in \mathbf{Q}$ for $1 \le i \le m$, then the condition

$$\max_{1 \le i \le m} |q\vartheta_i - p_i| < \frac{1}{b}$$

implies $q\vartheta_i - p_i = 0$ for $1 \le i \le m$, hence (iv) does not hold as soon as $Q > b^m$.

Let $\{q_1, \dots, q_N\}$ be a finite set of positive integers. Using (iv) again, we are going to show that there exists a positive integer $q \notin \{q_1, \dots, q_N\}$ satisfying the condition (v). Recall that $\|\cdot\|$ denotes the distance to the nearest integer. From (i) it follows that for $1 \le j \le N$, the number $\max_{1 \le i \le m} \|q_j \vartheta_i\|$ is non–zero. Let $Q > 1$ be sufficiently large such that

$$Q^{-1/m} < \min_{1 \le j \le N} \max_{1 \le i \le m} \|q_j \vartheta_i\|.$$

We use (iv): there exists an integer q in the range $1 \le q < Q$ such that

$$0 < \max_{1 \le i \le m} \|q\vartheta_i\| \le Q^{-1/m}.$$

The right hand side is $< q^{-1/m}$, and the choice of Q implies $q \notin \{q_1, \dots, q_N\}$. □

Proof of (v) ⇒ (ii). Given $\epsilon > 0$, there is a positive integer $q > \max\{1, 1/\epsilon^m\}$ satisfying the conclusion of (v). Then (ii) follows. □

Proof of (ii) ⇒ (iii). Let $\epsilon > 0$. From (ii) we deduce the existence of $(p_1, \dots, p_m, q)$ in $\mathbf{Z}^{m+1}$ with $q > 0$ such that

$$0 < \max_{1 \le i \le m} |q\vartheta_i - p_i| < \epsilon.$$

Without loss of generality we may assume $\gcd(p_1, \dots, p_m, q) = 1$. Define $L_1, \dots, L_m$ by $L_i(X_0, \dots, X_m) = p_i X_0 - qX_i$ for $1 \le i \le m$. Then $L_1, \dots, L_m$ are m linearly independent linear forms in $m+1$ variables with rational integer coefficients satisfying

$$0 < \max_{1 \le i \le m} |L_i(1, \vartheta_1, \dots, \vartheta_m)| < \epsilon.$$

We use (ii) once more with ϵ replaced by

$$\max_{1 \le i \le m} |L_i(1, \vartheta_1, \dots, \vartheta_m)| = \max_{1 \le i \le m} |q\vartheta_i - p_i|.$$

Hence there exists $p'_1, \dots, p'_m, q'$ in $\mathbf{Z}$ with $q' > 0$ such that

$$0 < \max_{1 \le i \le m} |q'\vartheta_i - p'_i| < \max_{1 \le i \le m} |q\vartheta_i - p_i|. \tag{13}$$

It remains to check that one at least of the m linear forms

$$L'_i(X_0, \dots, X_m) = p'_i X_0 - q' X_i$$

for $1 \le i \le m$ is linearly independent of $L_1, \dots, L_m$. Otherwise, for $1 \le i \le m$, there exist rational integers $s_i, t_{i1}, \dots, t_{im}$, with $s_i \ne 0$, such that

$$\begin{aligned} s_i(p'_i X_0 - q' X_i) &= t_{i1}L_1 + \cdots + t_{im}L_m \\ &= (t_{i1}p_1 + \cdots + t_{im}p_m)X_0 - q(t_{i1}X_1 + \cdots + t_{im}X_m). \end{aligned}$$

These relations imply, for $1 \le i \le m$,

$$s_i q' = q t_{ii}, \quad t_{ki} = 0 \quad \text{and} \quad s_i p'_i = p_i t_{ii} \quad \text{for } 1 \le k \le m, \quad k \ne i,$$

meaning that the two projective points $(p_1 : \cdots : p_m : q)$ and $(p'_1 : \cdots : p'_m : q')$ are the same. Since $\gcd(p_1, \dots, p_m, q) = 1$, it follows that $(p'_1, \dots, p'_m, q')$ is an integer multiple of $(p_1, \dots, p_m, q)$. This is not compatible with (13). □

Proof of (iii) $\Rightarrow$ (i). We proceed by contradiction. Assume (i) is not true: there exists $(a_1, \ldots, a_m, b) \in \mathbf{Z}^{m+1}$ with $b > 0$ such that $\vartheta_k = a_k/b$ for $1 \le k \le m$. Use (iii) with $\epsilon = 1/b$: we get $m+1$ linearly independent linear forms $L_0, \ldots, L_m$ in $\mathbf{Z}X_0 + \cdots + \mathbf{Z}X_m$. One at least of them, say L_k, does not vanish at $(1, \vartheta_1, \ldots, \vartheta_m)$. Then we have

$$0 < |L_k(b, a_1, \ldots, a_m)| = b|L_k(1, \vartheta_1, \ldots, \vartheta_m)| < b\epsilon = 1.$$

Since $L_k(b, a_1, \ldots, a_m)$ is a rational integer, we obtain a contradiction.

□

Proof of (i) $\Rightarrow$ (iv). Use Corollary 11. From the assumption (i) we deduce

$$\max_{1 \le i \le m} |q\vartheta_i - p_i| \neq 0.$$

□

Remark. This proof of the implication (i) $\Rightarrow$ (iv) in Proposition 12 (compare with [31] Chap. II § 2 p. 35) relies on Minkowski's linear form Theorem. Another proof of (i) $\Rightarrow$ (iv) in the special case where $Q^{1/m}$ is an integer, by means of Dirichlet's box principle, can be found in [31] Chap. II Th. 1E p. 28. A third proof (using again the geometry of numbers, but based on a result by Blichfeldt) is given in [31] Chap. II § 2 p. 32.

3 Criteria for linear independence

3.1 Hermite's method

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

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

is not 0.

Hermite's idea (see [18] and [13] Chap. 2 § 1.3) is to approximate simultaneously $\vartheta_1, \ldots, \vartheta_m$ by rational numbers $p_1/q, \ldots, p_m/q$ with the same denominator $q > 0$.

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

$$\epsilon_k = q\vartheta_k - p_k.$$

Then $qL = M + R$ with

$$M = a_0q + a_1p_1 + \cdots + a_mp_m \in \mathbf{Z}$$

and

$$R = a_1\epsilon_1 + \cdots + a_m\epsilon_m \in \mathbf{R}.$$

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 $(q, p_1, \ldots, p_m)$ such that $(p_1/q, \ldots, p_m/q)$ is a simultaneous rational approximation to $(\vartheta_1, \ldots, \vartheta_m)$, but we also require that it lies outside the hyperplane $a_0X_0 + a_1X_1 + \cdots + a_mX_m = 0$ of $\mathbf{Q}^{m+1}$. Our goal is to prove the linear independence over $\mathbf{Q}$ of $1, \vartheta_1, \ldots, \vartheta_m$; hence this needs to be checked for all hyperplanes. The solution to this problem is to construct not only one tuple $(q, p_1, \ldots, p_m)$ in $\mathbf{Z}^{m+1} \setminus \{0\}$, but $m+1$ such tuples which are linearly independent. This yields $m+1$ pairs (M_k, R_k) $(k = 0, \ldots, m)$ in place of a single pair (M, R). From $(a_0, \ldots, a_m) \neq (0, \ldots, 0)$, one deduces that one at least of $M_0, \ldots, M_m$ is not 0.

It turns out (Proposition 14 below) that nothing is lost by using such arguments: existence of linearly independent simultaneous rational approximations for $\vartheta_1, \ldots, \vartheta_m$ are characteristic of linearly independent real numbers $1, \vartheta_1, \ldots, \vartheta_m$.

3.2 Rational approximations

The following criterion is due to M. Laurent [22].

Proposition 14. *Let $\underline{\vartheta} = (\vartheta_1, \ldots, \vartheta_m) \in \mathbf{R}^m$. Then the following conditions are equivalent:*
(i) *The numbers $1, \vartheta_1, \ldots, \vartheta_m$ are linearly independent over $\mathbf{Q}$.*
(ii) *For any $\epsilon > 0$, there exist $m+1$ linearly independent elements $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m$ in $\mathbf{Z}^{m+1}$, say*

$$\mathbf{u}_i = (q_i, p_{1i}, \ldots, p_{mi}) \quad (0 \leq i \leq m)$$

with $q_i > 0$, such that

$$\max_{1\leq k\leq m} \left|\vartheta_k - \frac{p_{ki}}{q_i}\right| \leq \frac{\epsilon}{q_i} \quad (0 \leq i \leq m). \tag{15}$$

The condition of linear independence on the elements $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m$ 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.

For $0 \le i \le m$, set

$$\underline{r}_i = \left(\frac{p_{1i}}{q_i}, \ldots, \frac{p_{mi}}{q_i} \right) \in \mathbf{Q}^m.$$

Further define, for $\underline{x} = (x_1, \ldots, x_m) \in \mathbf{R}^m$,

$$|\underline{x}| = \max_{1 \le i \le m} |x_i|.$$

Also for $\underline{x} = (x_1, \ldots, x_m) \in \mathbf{R}^m$ and $\underline{y} = (y_1, \ldots, y_m) \in \mathbf{R}^m$ set

$$\underline{x} - \underline{y} = (x_1 - y_1, \ldots, x_m - y_m),$$

so that

$$|\underline{x} - \underline{y}| = \max_{1 \le i \le m} |x_i - y_i|.$$

Then the relation (15) in Proposition 14 can be written

$$|\underline{\vartheta} - \underline{r}_i| \le \frac{\epsilon}{q_i}, \quad (0 \le i \le m).$$

The easy implication (which is also the useful one for Diophantine applications: linear independence, transcendence and algebraic independence) is (ii) $\Rightarrow$ (i) . We shall prove a more explicit version of it by checking that *any tuple* $(q, p_1, \ldots, p_m) \in \mathbf{Z}^{m+1}$, *with* $q > 0$, *producing a tuple* $(p_1/q, \ldots, p_m/q) \in \mathbf{Q}^m$ *of sufficiently good rational approximations to* $\underline{\vartheta}$ *satisfies the same linear dependence relations as* $1, \vartheta_1, \ldots, \vartheta_m$.

Lemma 16. *Let* $\vartheta_1, \ldots, \vartheta_m$ *be real numbers. Assume that the numbers* $1, \vartheta_1, \ldots, \vartheta_m$ *are linearly dependent over* $\mathbf{Q}$*: let* $a, b_1, \ldots, b_m$ *be rational integers, not all of which are zero, satisfying*

$$a + b_1\vartheta_1 + \cdots + b_m\vartheta_m = 0.$$

Let ϵ *be a real number satisfying*

$$0 < \epsilon < \left(\sum_{k=1}^{m} |b_k| \right)^{-1}.$$

Assume further that $(q, p_1, \ldots, p_m) \in \mathbf{Z}^{m+1}$ *satisfies* $q > 0$ *and*

$$\max_{1 \le k \le m} |q\vartheta_k - p_k| \le \epsilon.$$

Then

$$aq + b_1p_1 + \cdots + b_mp_m = 0.$$

Proof. In the relation

$$qa + \sum_{k=1}^{m} b_k p_k = \sum_{k=1}^{m} b_k (p_k - q\vartheta_k),$$

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 Proposition 14.* Let

$$aX_0 + b_1 X_1 + \cdots + b_m X_m$$

be a non–zero linear form with integer coefficients. For sufficiently small ϵ, assumption (ii) show that there exist $m+1$ linearly independent elements $\mathbf{u}_i \in \mathbf{Z}^{m+1}$ such that the corresponding rational approximation satisfy the assumptions of Lemma 16. Since $\mathbf{u}_0, \ldots, \mathbf{u}_m$ is a basis of $\mathbf{Q}^{m+1}$, one at least of the $L(\mathbf{u}_i)$ is not 0. Hence Lemma 16 implies

$$a + b_1\vartheta_1 + \cdots + b_m\vartheta_m \neq 0.$$

□

Proof of (i) ⇒ (ii) *in Proposition 14.* Let $\epsilon > 0$. By Corollary 11, there exists $\mathbf{u} = (q, p_1, \ldots, p_m) \in \mathbf{Z}^{m+1}$ with $q > 0$ such that

$$\max_{1\leq k\leq m} \left|\vartheta_k - \frac{p_k}{q}\right| \leq \frac{\epsilon}{q}.$$

Consider the subset $E_\epsilon \subset \mathbf{Z}^{m+1}$ of these tuples. Let V_ϵ be the $\mathbf{Q}$-vector subspace of $\mathbf{Q}^{m+1}$ spanned by E_ϵ.

If $V_\epsilon \neq \mathbf{Q}^{m+1}$, then there is a hyperplane $a_0x_0 + a_1x_1 + \cdots + a_mx_m = 0$ containing E_ϵ. Any $\mathbf{u} = (q, p_1, \ldots, p_m)$ in E_ϵ has

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

For each $n \geq 1/\epsilon$, let $\mathbf{u} = (q_n, p_{1n}, \ldots, p_{mn}) \in E_\epsilon$ satisfy

$$\max_{1\leq k\leq m} \left|\vartheta_k - \frac{p_{kn}}{q_n}\right| \leq \frac{1}{nq_n}.$$

Then

$$a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m = \sum_{k=1}^{m} a_k \left(\vartheta_k - \frac{p_{kn}}{q_n}\right).$$

Hence

$$|a_0 + a_1\vartheta_1 + \cdots + a_m\vartheta_m| \leq \frac{1}{nq_n}\sum_{k=1}^{m}|a_k|.$$

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 $\mathbf{Q}$–linearly dependent, which means that (i) does not hold.

Therefore, if (i) holds, then $V_\epsilon = \mathbf{Q}^{m+1}$, hence there are $m+1$ linearly independent elements in E_ϵ.

□

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°3, *April 26, 2010*

3.3 Linear forms

3.3.1 Siegel's method: $m+1$ linear forms

For proving linear independence of real numbers, Hermite [18] considered simultaneous approximation to these numbers by algebraic numbers. The point of view introduced by Siegel in 1929 [33] is dual (duality in the sense of convex bodies): he considers simultaneous approximation by means of independent linear forms.

We define the *height* of a linear form $L = a_0X_0 + \cdots + a_mX_m$ with complex coefficients by

$$H(L) = \max\{|a_0|, \ldots, |a_m|\}.$$

Lemma 17. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers. Assume that, for any $\epsilon > 0$, there exists $m+1$ linearly independent linear forms $L_0, \ldots, L_m$ in $m+1$ variables, with coefficients in $\mathbf{Z}$, such that*

$$\max_{0\le k\le m} |L_k(1, \vartheta_1, \ldots, \vartheta_m)| < \frac{\epsilon}{H^{m-1}} \quad \text{where} \quad H = \max_{0\le k\le m} H(L_k).$$

Then $1, \vartheta_1, \ldots, \vartheta_m$ are linearly independent over $\mathbf{Q}$.

The proof is given by C.L. Siegel in [33]; see also [13] Chap. 2 § 1.4 and [6]. We sketch the argument here, and we expand it below.

Assume $1, \vartheta_1, \ldots, \vartheta_m$ are linearly dependent over $\mathbf{Q}$: let $\Lambda_0 \in \mathbf{Z}X_0 + \mathbf{Z}X_1 + \cdots + \mathbf{Z}X_m$ be a non–zero linear form in $m+1$ variables which vanishes at the point $(1, \vartheta_1, \ldots, \vartheta_m)$. Denote by A the maximum of the absolute values of the coefficients of Λ_0 and use the assumption with $\epsilon = 1/m!mA$. Among the $m+1$ linearly independent linear forms which are given by the assumption of Lemma 17, select m of them, say $\Lambda_1, \ldots, \Lambda_m$, which form with Λ_0 a set of $m+1$ linearly independent linear forms. The $(m+1)\times(m+1)$ matrix of coefficients of these forms is regular; using the inverse matrix, one expresses its determinant Δ as a linear combination with integer coefficients

of $\Lambda_k(1, \vartheta_1, \ldots, \vartheta_m)|$, $1 \leq k \leq m$. The choice of ϵ yields the contradiction $|\Delta| < 1$.

We develop this idea and deduce the following more precise statement.

Proposition 18. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers and $L_0, \ldots, L_m$ be $m+1$ linearly independent linear forms in $m+1$ variables with coefficients in $\mathbf{Z}$. Then*

$$\max_{0 \leq k \leq m} \frac{|L_k(1, \vartheta_1, \ldots, \vartheta_m)|}{H(L_k)} \geq \frac{1}{(m+1)! H(L_0) \cdots H(L_m)}.$$

Proof. For $0 \leq k \leq m$, write

$$L_k(X_0, \ldots, X_m) = \sum_{i=0}^{m} \ell_{ki} X_i \quad \text{and set} \quad \lambda_k = L_k(1, \vartheta_1, \ldots, \vartheta_m).$$

Define $\vartheta_0 = 1$. Let $\underline{\mathrm{L}}$ be the regular $(m+1) \times (m+1)$ matrix $\left(\ell_{ki}\right)_{0 \leq k, i \leq m}$. Using the relation

$$\begin{pmatrix} \vartheta_0 \\ \vdots \\ \vartheta_m \end{pmatrix} = \underline{\mathrm{L}}^{-1} \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_m \end{pmatrix},$$

one can write the product of $\vartheta_0 = 1$ by $\det(\underline{\mathrm{L}})$ as a linear combination of $\lambda_0, \ldots, \lambda_m$ with rational integer coefficients. In this linear combination, the absolute value of the coefficient of λ_k is $\leq m! H(L_0) \cdots H(L_m)/H(L_k)$. We deduce

$$1 \leq |\det(\underline{\mathrm{L}})| \leq m! \sum_{k=0}^{m} H(L_0) \cdots H(L_m) \frac{|\lambda_k|}{H(L_k)}.$$

Proposition 18 follows.

□

An straightforward consequence of Proposition 18 is the following:

Corollary 19. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers, H be a positive real number and $L_0, \ldots, L_m$ be $m+1$ linearly independent linear forms in $m+1$ variables with coefficients in $\mathbf{Z}$ of height $\leq H$. Then*

$$\max_{0 \leq k \leq m} |L_k(1, \vartheta_1, \ldots, \vartheta_m)| \geq \frac{1}{(m+1)! H^m}.$$

Using either Proposition 18 or Corollary 19, we deduce the following result (compare with [27] Lemma 2.4):

Corollary 20. *Let $\vartheta_1, \dots, \vartheta_m$ be complex numbers and $\kappa \geq 0$ be a real number. Assume that, for any $\epsilon > 0$, there exists $m+1$ linearly independent linear forms $L_0, \dots, L_m$ in $m+1$ variables, with coefficients in $\mathbf{Z}$, such that*

$$\max_{0 \leq k \leq m} |L_k(1, \vartheta_1, \dots, \vartheta_m)| < \frac{\epsilon}{H^\kappa} \quad \text{where} \quad H = \max_{0 \leq k \leq m} H(L_k).$$

Denote by $r+1$ the dimension of the $\mathbf{Q}$–vector space spanned by $1, \vartheta_1, \dots, \vartheta_m$. Then $r > \kappa$.

Under the assumptions of Corollary 20, since $r \leq m$, we deduce $\kappa < m$, which is a plain consequence of Corollary 19.

We recover Lemma 17 by taking $\kappa = m - 1$.

Also we recover the implication (iii) $\Rightarrow$ (i) from Proposition 12 by taking $\kappa = 0$.

Proof. We give two slightly different proofs of Corollary 20. For the first one, we use Proposition 18 as follows: consider $m - r$ linearly independent linear relations among $1, \vartheta_1, \dots, \vartheta_m$. Denote by $\widetilde{L}_{r+1}, \dots, \widetilde{L}_m$ these linear forms and by c their maximal height. Take $0 < \epsilon < 1/((m+1)!c^{m-r})$. Select $r+1$ linear forms $\widetilde{L}_0, \dots, \widetilde{L}_r$ among $L_0, \dots, L_m$ to get a maximal system of $m+1$ linearly independent linear forms $\widetilde{L}_0, \dots, \widetilde{L}_m$. From Proposition 18 one deduces

$$\begin{aligned}
\frac{1}{(m+1)!c^{m-r}H(\widetilde{L}_0)\cdots H(\widetilde{L}_r)} &\leq \frac{1}{(m+1)!H(\widetilde{L}_0)\cdots H(\widetilde{L}_m)} \\
&\leq \max_{0 \leq k \leq m} \frac{|\widetilde{L}_k(1, \vartheta_1, \dots, \vartheta_m)|}{H(\widetilde{L}_k)} \\
&\leq \max_{0 \leq k \leq r} \frac{|\widetilde{L}_k(1, \vartheta_1, \dots, \vartheta_m)|}{H(\widetilde{L}_k)} \\
&\leq \max_{0 \leq k \leq m} \frac{|L_k(1, \vartheta_1, \dots, \vartheta_m)|}{H(L_k)}.
\end{aligned}$$

From the choice of ϵ, one concludes $H^\kappa < H^r$, hence $r > \kappa$.

Our second proof of Corollary 20 rests on Corollary 19. Let $1, \xi_1, \dots, \xi_r$ be a basis of the $\mathbf{Q}$–vector space spanned by $1, \vartheta_1, \dots, \vartheta_m$. Define $\xi_0 = \vartheta_0 = 1$ and write

$$\vartheta_h = \sum_{j=0}^{r} a_{hj}\xi_j \quad (0 \leq h \leq m).$$

In particular $a_{00} = 1$ and $a_{0j} = 0$ for $1 \le j \le m$. Define

$$c = \max_{0 \le j \le r} \sum_{h=0}^{m} |a_{hj}|$$

and let ϵ satisfy $0 < \epsilon < 1/(r+1)!c^r$. Let $L_0, \dots, L_m$ be the $m+1$ linearly independent linear forms in $m+1$ variables with integer coefficients given by the assumption of Corollary 20. Write

$$L_k(X_0, \dots, X_m) = \sum_{h=0}^{m} \ell_{kh} X_h \quad (0 \le k \le m).$$

By assumption $\max_{0 \le k,h \le m} |\ell_{kh}| \le H$. Consider the $m+1$ linear forms $\Lambda_0, \dots, \Lambda_m$ in $r+1$ variables $Y_0, \dots, Y_r$ defined by

$$\Lambda_k(Y_0, \dots, Y_r) = \lambda_{k0} Y_0 + \cdots + \lambda_{kr} Y_r \quad (0 \le k \le m)$$

with

$$\lambda_{kj} = \sum_{h=0}^{m} \ell_{kh} a_{hj}.$$

The connexion between the linear forms $L_0, \dots, L_m$ in $\mathbf{Z}X_0 + \cdots + \mathbf{Z}X_m$ on the one side and and $\Lambda_0, \dots, \Lambda_m$ in $\mathbf{Z}Y_0 + \cdots + \mathbf{Z}Y_r$ on the other side is

$$\Lambda_k(Y_0, \dots, Y_r) = L_k \left(\sum_{j=0}^{r} a_{0j} Y_j, \dots, \sum_{j=0}^{r} a_{mj} Y_j \right) \qquad (0 \le k \le m).$$

Since $1, \xi_1, \dots, \xi_r$ are $\mathbf{Q}$–linearly independent, the $r+1$ columns of the $(m+1) \times (r+1)$ matrix $\big(a_{hj}\big)_{\substack{0 \le h \le m \\ 0 \le j \le r}}$ are linearly independent in $\mathbf{Q}^{m+1}$, hence this matrix has rank $r+1$, and therefore the rank of the set of $m+1$ linear forms $\Lambda_0, \dots, \Lambda_m$ is $r+1$. By construction

$$\Lambda_k(1, \xi_1, \dots, \xi_r) = L_k(1, \vartheta_1, \dots, \vartheta_m) \quad (0 \le k \le m).$$

Applying Corollary 19 to the point $(1, \xi_1, \dots, \xi_r)$ with $r+1$ independent linear forms among $\Lambda_0, \dots, \Lambda_m$, we deduce

$$\max_{0 \le k \le m} |\Lambda_k(1, \xi_1, \dots, \xi_r)| \ge \frac{1}{(r+1)! \widetilde{H}^r}$$

with

$$\widetilde{H} = \max_{0 \le k \le m} H(\Lambda_k) = \max_{\substack{0 \le k \le m \\ 0 \le j \le r}} |\lambda_{kj}| \le cH.$$

Again, from the choice of ϵ, one concludes $H^\kappa < H^r$, hence $r > \kappa$.

Corollary 20 follows.

□

3.3.2 Nesterenko's Criterion for linear independence

In 1985, Yu.V. Nesterenko [26], obtained a variant of Proposition 18 (Siegel's linear independence criterion). There are two main differences: on the one hand, Nesterenko does not need $m+1$ linearly independent forms, but he needs only one; at the same time he does not only assume an upper bound for the value of this linear form at the point $(1, \vartheta_1, \ldots, \vartheta_m)$, but also a lower bound. On the other hand, for Nesterenko it is not sufficient to have infinitely many linear forms as in Siegel's Proposition 18, but he needs a sequence of such forms (for all sufficiently large n, and not only for infinitely many n). A simplification of the original proof by Nesterenko was proposed by F. Amoroso and worked out by P. Colmez. A new approach, which at the same time simplifies further the argument and yields refinements, is due to S. Fischler and W. Zudilin [15].

The main reference for this section is [6].

Theorem 21 (Nesterenko linear independence criterion)**.** *Let c_1, c_2, τ_1, τ_2 be positive real numbers and $\sigma(n)$ a non–decreasing positive function such that*

$$\lim_{n\to\infty} \sigma(n) = \infty \quad \text{and} \quad \limsup_{n\to\infty} \frac{\sigma(n+1)}{\sigma(n)} = 1.$$

Let $\underline{\vartheta} = (\vartheta_1, \ldots, \vartheta_m) \in \mathbf{R}^m$. Assume that, for all sufficiently large integers n, there exists a linear form with integer coefficients in $m+1$ variables

$$L_n(\underline{X}) = \ell_{0n}X_0 + \ell_{1n}X_1 + \cdots + \ell_{mn}X_m,$$

which satisfies the conditions

$$H(L_n) \le e^{\sigma(n)} \quad \text{and} \quad c_1 e^{-\tau_1\sigma(n)} \le |L_n(1, \underline{\vartheta})| \le c_2 e^{-\tau_2\sigma(n)}.$$

Then $\dim_{\mathbf{Q}}(\mathbf{Q} + \mathbf{Q}\vartheta_1 + \cdots + \mathbf{Q}\vartheta_m) \ge (1+\tau_1)/(1+\tau_1-\tau_2)$.

The main result of [6], which relies on the arguments in [15], is the following.

Theorem 22. *Let $\underline{\xi} = (\xi_i)_{i\ge 0}$ be a sequence of real numbers with $\xi_0 = 1$, $(r_n)_{n\ge 0}$ a non–decreasing sequence of positive integers, $(Q_n)_{n\ge 0}$, $(A_n)_{n\ge 0}$ and $(B_n)_{n\ge 0}$ sequences of positive real numbers such that $\lim_{n\to\infty} A_n^{1/r_n} = \infty$ and, for all sufficiently large integers n,*

$$Q_n B_n \le Q_{n+1} B_{n+1}.$$

Assume that, for any sufficiently large integer n, there exists a linear form with integer coefficients in $r_n + 1$ variables

$$L_n(\underline{X}) = \ell_{0n}X_0 + \ell_{1n}X_1 + \cdots + \ell_{r_n n}X_{r_n}$$

such that

$$\sum_{i=0}^{r_n} |\ell_{in}| \le Q_n, \quad 0 < |L_n(\underline{\xi})| \le \frac{1}{A_n} \quad \text{and} \quad \frac{|L_{n-1}(\underline{\xi})|}{|L_n(\underline{\xi})|} \le B_n.$$

Then $A_n \le 2^{r_n+1}(B_nQ_n)^{r_n}$ for all sufficiently large integers n.

One deduces from Theorem 22 a slight refinement of Theorem 21 where the condition $\limsup_{n\to\infty} \frac{\sigma(n+1)}{\sigma(n)} = 1$ is relaxed, the cost being to replace $\sigma(n)$ by $\sigma(n+1)$ in the upper bound for $|L_n(1, \underline{\vartheta})|$.

Corollary 23. *Let τ_1, τ_2 be positive real numbers and $\sigma(n)$ a non–decreasing positive function such that $\lim_{n\to\infty} \sigma(n) = \infty$. Let $\underline{\vartheta} = (\vartheta_1, \ldots, \vartheta_m) \in \mathbf{R}^m$. Assume that, for all sufficiently large integers n, there exists a linear form with integer coefficients in $m + 1$ variables*

$$L_n(\underline{X}) = \ell_{0n}X_0 + \ell_{1n}X_1 + \cdots + \ell_{mn}X_m$$

which satisfies the conditions

$$H(L_n) \le e^{\sigma(n)} \quad \text{and} \quad e^{-(\tau_1+o(1))\sigma(n)} \le |L_n(1, \underline{\vartheta})| \le e^{-(\tau_2+o(1))\sigma(n+1)}.$$

Then $\dim_{\mathbf{Q}}(\mathbf{Q} + \mathbf{Q}\vartheta_1 + \cdots + \mathbf{Q}\vartheta_m) \ge (1 + \tau_1)/(1 + \tau_1 - \tau_2)$.

Further consequences of Theorem 22 are given in [6]. See also Corollary 33 below.

4 Criteria for transcendence

The main Diophantine tool for proving transcendence results is Liouville's inequality.

4.1 Liouville's inequality

Recall that the ring $\mathbf{Z}[X]$ is factorial, its irreducible elements of positive degree are the non-constant polynomials with integer coefficients which are irreducible in $\mathbf{Q}[X]$ (i.e., not a product of two non-constant polynomials

in $\mathbf{Q}[X]$) and have content 1. The *content* of a polynomial in $\mathbf{Z}[X]$ is the greatest common divisor of its coefficients.

The *minimal polynomial* of an algebraic number α is the unique irreducible polynomial $P \in \mathbf{Z}[X]$ which vanishes at α and has a positive leading coefficient.

The next lemma is one of many variants of Liouville's inequality (see, for instance, [21, 31, 37, 28, 27]), which is close to the original one of 1844.

Lemma 24. *Let α be an algebraic number of degree $d \geq 2$ and minimal polynomial $P \in \mathbf{Z}[X]$. Define $c = |P'(\alpha)|$. Let $\epsilon > 0$. Then there exists an integer q_0 such that, for any $p/q \in \mathbf{Q}$ with $q \geq q_0$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{(c+\epsilon)q^d}.$$

Proof. The result is trivial if α is not real: an admissible value for q_0 is

$$q_0 = (c|\Im m(\alpha)|)^{-1/d}.$$

Assume now α is real. Let q be a sufficiently large positive integer and let p be the nearest integer to $q\alpha$. In particular,

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2q}.$$

Denote by a_0 the leading coefficient of P and by $\alpha_1, \ldots, \alpha_d$ the roots with $\alpha_1 = \alpha$. Hence

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

and

$$q^d P(p/q) = a_0 q^d \prod_{i=1}^{d} \left(\frac{p}{q} - \alpha_i\right). \tag{25}$$

Also

$$P'(\alpha) = a_0 \prod_{i=2}^{d} (\alpha - \alpha_i).$$

The left hand side of (25) is a rational integer. It is not zero because P is irreducible of degree ≥ 2. For $i \geq 2$ we use the estimate

$$\left|\alpha_i - \frac{p}{q}\right| \leq |\alpha_i - \alpha| + \frac{1}{2q}.$$

We deduce

$$1 \le q^d a_0 \left|\alpha - \frac{p}{q}\right| \prod_{i=2}^{d} \big(|\alpha_i - \alpha| + \frac{1}{2q}\big).$$

For sufficiently large q the right hand side is bounded from above by

$$q^d \left|\alpha - \frac{p}{q}\right| (|P'(\alpha)| + \epsilon).$$

□

The same proof yields the next result.

Define the height $H(P)$ of a polynomial P with complex coefficients (any number of variables) as the maximum modulus of its coefficients.

Proposition 26 (Liouville's inequality)**.** *Let $\alpha_1, \ldots, \alpha_m$ be algebraic numbers. There exists a constant $c = c(\alpha_1, \ldots, \alpha_m) > 0$ such that, for any polynomial $P \in \mathbf{Z}[X_1, \ldots, X_m]$ satisfying $P(\alpha_1, \ldots, \alpha_m) \neq 0$, the inequality*

$$|P(\alpha_1, \ldots, \alpha_m)| \ge H^{-c} e^{-cd}$$

holds with $H = \max\{2, H(P)\}$ and d the total degree of P.

The constant c can be explicitly computed (see, for instance, [13, 38]), but this is not relevant here.

The corollary below (which is [27] Prop. 3.1) is useful for proving transcendence results.

Corollary 27. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers $\mathbf{C}$. Let $\sigma(n)$ and $\lambda(n)$ be two non–decreasing positive real functions with $\lim_{n\to\infty} \sigma(n) = \infty$ and $\lim_{n\to\infty} \lambda(n)/\sigma(n) = \infty$. Assume that there exists a sequence $(P_n)_{n\ge 0}$ of polynomials in $\mathbf{Z}[X_1, \ldots, X_m]$, with P_n of degree $\le \sigma(n)$ and height $H(P_n) \le e^{\sigma(n)}$, such that, for infinitely many n,*

$$0 < |P_n(\vartheta_1, \ldots, \vartheta_m)| \le e^{-\lambda(n)}.$$

Then one at least of the numbers $\vartheta_1, \ldots, \vartheta_m$ is transcendental.

4.2 Transcendence criterion of A. Durand

Liouville's result is not a necessary and sufficient condition for transcendence. One way of extending the irrationality criterion of Proposition 4 into a transcendence criterion is to replace rational approximation by approximation by algebraic numbers. For instance, given an integer d, one gets a

criterion for ϑ not being algebraic of degree $\leq d$ by considering algebraic approximation of ϑ by algebraic numbers of degree $\leq d$. One may also let d vary and get a transcendence criterion as follows.

Define the height of a $H(\alpha)$ of an algebraic number α as the height of its irreducible polynomial in $\mathbf{Z}[X]$, and the size $s(\alpha)$ as

$$s(\alpha) := [\mathbf{Q}(\alpha) : \mathbf{Q}] + \log H(\alpha).$$

The following result (we shall not use it and we do not include a proof) is due to A. Durand [9, 10].

Proposition 28. *Let ϑ be a complex number. The following conditions are equivalent:*
(i) *ϑ is transcendental.*
(ii) *For any $\kappa > 0$ there exists an algebraic number α such that*

$$0 < |\vartheta - \alpha| < e^{-\kappa s(\alpha)}.$$

(iii) *There exists a sequence $(\alpha_n)_{n\geq 0}$ of pairwise distinct algebraic numbers such that*

$$\lim_{n\to\infty} \frac{\log |\vartheta - \alpha_n|}{s(\alpha_n)} = -\infty.$$

Another way of getting transcendence criteria for a number ϑ (resp. criteria for ϑ not being of degree $\leq d$) is to consider polynomial approximations $|P(\vartheta)|$ by polynomials in $\mathbf{Z}[X]$ (resp. by polynomials of degree $\leq d$).

5 Criteria for algebraic independence

5.1 Small transcendence degree: Gel'fond's criterion

Gel'fond's criterion (see, for instance, [21, 37, 28, 27]) is a powerful tool to prove the algebraic independence of at least two numbers.

A slightly refined version (due to A. Chantanasiri) is the following one.

Define the size $t(P)$ of a polynomial $P \in \mathbf{C}[X]$ as

$$t(P) := \log H(P) + (\log 2) \deg P.$$

Theorem 29 (Gel'fond's Transcendence Criterion)**.** *Let $\vartheta \in \mathbf{C}$ and let γ be a real number with $\gamma > 1$. Let $(d_n)_{n=1}^{\infty}$ and $(t_n)_{n=1}^{\infty}$ be two non–decreasing sequences of real numbers with $\lim_{n\to\infty} t_n = \infty$. Assume that there exists a*

sequence $(P_n)_{n\geq 0}$ *of polynomials in* $\mathbf{Z}[X]$ *with* P_n *of degree* $\leq d_n$ *and size* $t(P_n) \leq t_n$ *such that, for all sufficiently large integer* n*,*

$$|P_n(\vartheta)| \leq e^{-\gamma(d_n t_n + d_{n+1} t_n + d_n t_{n+1})}.$$

Then ϑ *is algebraic and* $P_n(\vartheta) = 0$ *for all sufficiently large* n*.*

A consequence of Theorem 29 is the following variant of Gel'fond's Criterion (Lemma 3.5 of [27]):

Corollary 30. *Let* $\vartheta \in \mathbf{C}$ *and let* $\sigma(n)$ *be a non–decreasing unbounded positive real function. Assume that there exists a sequence* $(P_n)_{n\geq 0}$ *of polynomials in* $\mathbf{Z}[X]$ *with* P_n *of size* $t(P_n) \leq \sigma(n)$ *such that, for all sufficiently large integer* n*,*

$$|P_n(\vartheta)| \leq e^{-5\sigma(n+1)^2}.$$

Then ϑ *is algebraic and* $P_n(\vartheta) = 0$ *for all sufficiently large* n*.*

This result is useful to prove that in some given set of specific numbers, at least two numbers are algebraically independent ([27] § 3.3 Prop. 3.3).

Corollary 31. *Let* $\vartheta_1, \ldots, \vartheta_m$ *be complex numbers. Let* $\sigma(n)$ *and* $\lambda(n)$ *be two non–decreasing positive real function with* $\lim_{n\to\infty} \sigma(n) = \infty$ *and* $\lim_{n\to\infty} \lambda(n)/\sigma(n+1)^2 = \infty$*. Assume that there exists a sequence* $(P_n)_{n\geq 0}$ *of polynomials in* $\mathbf{Z}[X_1, \ldots, X_m]$*, with* P_n *of degree* $\leq \sigma(n)$ *and height* $H(P_n) \leq e^{\sigma(n)}$*, such that, for all sufficiently large* n*,*

$$0 < |P_n(\vartheta_1, \ldots, \vartheta_m)| \leq e^{-\lambda(n)}.$$

Then at least two of the numbers $\vartheta_1, \ldots, \vartheta_m$ *are algebraically independent.*

One should stress the following differences with Corollary 27: the conclusion of Theorem 29 is that the transcendence degree of the field $\mathbf{Q}(\vartheta_1, \ldots, \vartheta_m)$ is at least 2, while Liouville's argument shows only that it is at least 1. There is a price for that. On the one hand, the assumption

$$\lim_{n\to\infty} \lambda(n)/\sigma(n+1)^2 = \infty$$

is stronger than the assumption

$$\lim_{n\to\infty} \lambda(n)/\sigma(n) = \infty$$

in Corollary 27 (what is important is the square, not the $n+1$ in place of n). On the other hand, Liouville's assumption is assumed to be satisfied for infinitely many n, while Gel'fond requires it for all sufficiently large n.

5.2 Large transcendence degree

It took some time before Gel'fond's transcendence criterion could be extended into a criterion for large transcendence degree. One approach suggested by S. Lang [21] involves his so-called *transcendence type* (see [27] § 7.3): this is an assumption which amounts to avoid Liouville type numbers. The idea is to prove algebraic independence by induction, but the results which are obtained in this way are comparatively weak.

One might hope that assuming $\lim_{n\to\infty} \lambda(n)/\sigma(n+1)^k = \infty$ in Corollary 31 would suffice to prove that the transcendence degree of the field $\mathbf{Q}(\vartheta_1, \ldots, \vartheta_m)$ is at least k. However this is not the case, as an example from Khinchine (reproduced in Cassels's book on Diophantine approximation [5]) shows. The first one to obtain a criterion for large transcendence degree was G.V. Chudnovskii in 1976. The original criterion was not sharp, the estimate for the transcendence degree was the logarithm of the expected one. A few years later Philippon reached the optimal exponent.

One of the main tools, in Nesterenko's proof of his main result (Theorem 4.2 in [27]), is this criterion for algebraic independence due to Philippon ([27] Chap. 6). Here is Corollary 6.2 of [27]. See also [30, 28].

Theorem 32. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers, $\sigma(n)$ and $S(n)$ be two non–decreasing positive real functions and k be a real number in the range $1 \le k \le m$. Assume that the functions*

$$\sigma(n) \quad \text{and} \quad \frac{S(n-1)}{\sigma(n)^k}$$

are non–decreasing and unbounded. Assume, further, that there exists a constant c_0 and a sequence $(P_n)_{n\ge 0}$ of polynomials in $\mathbf{Z}[X]$ with P_n of size $t(P_n) \le \sigma(n)$ such that, for all sufficiently large n,

$$e^{-c_0 S(n-1)} < |P_n(\vartheta_1, \ldots, \vartheta_m)| \le e^{-S(n)}.$$

Then the transcendence degree over $\mathbf{Q}$ of the field $\mathbf{Q}(\vartheta_1, \ldots, \vartheta_m)$ is $> k-1$.

The special case $k = 1$ of this result is close to (but weaker than) Corollary 27, the special case $k = 2$ of this result is close to (but weaker than) Theorem 29 (where no lower bound was requested).

It is interesting to compare with the following criterion for algebraic independence (Corollary 3.6 of [6]), which is a corollary of Theorem 22.

Corollary 33. *Let $\vartheta_1, \ldots, \vartheta_t$ be real numbers and $(\tau_d)_{d\geq 1}$, $(\eta_d)_{d\geq 1}$ two sequences of positive real numbers satisfying*

$$\frac{\tau_d}{d^{t-1}(1+\eta_d)} \longrightarrow +\infty.$$

Further, let $\sigma(n)$ be a non–decreasing unbounded positive real function. Assume that for all sufficiently large d, there is a sequence $(P_n)_{n\geq n_0(d)}$ of polynomials in $\mathbf{Z}[X_1, \ldots, X_t]$, where P_n has degree $\leq d$ and length $\leq e^{\sigma(n)}$, such that, for $n \geq n_0(d)$,

$$e^{-(\tau_d+\eta_d)\sigma(n)} \leq |P_n(\vartheta_1, \ldots, \vartheta_t)| \leq e^{-\tau_d\sigma(n+1)}.$$

Then $\vartheta_1, \ldots, \vartheta_t$ are algebraically independent.

The proof of Corollary 33 is much easier than the proof of Theorem 32, since it relies on linear elimination instead of polynomial elimination. Unfortunately, Corollary 33 does not seem to suffice for the proof of Nesterenko's algebraic independence Theorem on q, $P(q)$, $Q(q)$ and $R(q)$ (Theorem 4.2 of [27]).

Exercise. *Let $\vartheta_1, \ldots, \vartheta_m$ be complex numbers and d a positive integer. Check that the following conditions are equivalent:*
(i) *There exists a non–zero polynomial $A \in \mathbf{Q}[X_1, \ldots, X_m]$ of degree $\leq d$ such that $A(\vartheta_1, \ldots, \vartheta_m) = 0$.*
(ii) *The dimension of the $\mathbf{Q}$–vector space spanned by the numbers*

$$\vartheta_1^{i_1} \cdots \vartheta_m^{i_m} \quad (i_1 + \cdots + i_m \leq n)$$

is bounded from above by

$$d\frac{n^{m-1}}{(m-1)!} + \mathrm{O}(n^{m-1})$$

as $n \to \infty$.

Appendix: the resultant of two polynomials in one variable

The main tool for the proof of Gel'fond's criterion is the resultant of two polynomials in one variable.

Given two linear equations in two unknowns

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2, \end{cases}$$

in order to compute y, one eliminates x. This amounts to find the projection on the y axis of the intersection point (x, y) of two lines in the plane. More generally, linear algebra enables one to find the intersection point (unique in general) of n hyperplanes in dimension n by means of a determinant.

Given two plane curves

$$f(x, y) = 0 \quad \text{and} \quad g(x, y) = 0$$

without common components, there are only finitely many intersection points; the values y of the coordinates (x, y) of these points are roots of a polynomial R in $K_0[Y]$, where K_0 is the base field. This polynomial is computed by eliminating x between the two equations $f(x, y) = 0$ and $g(x, y) = 0$. The ideal of $K_0[Y]$ which is the intersection of $K_0[Y]$ with the ideal of $K_0[X, Y]$ generated by f and g is principal, and R is a generator: there is a pair (U, V) of polynomials in $K_0[X, Y]$ such that $R = Uf + Vg$. If (U, V) satisfies this *Bézout condition*, then so does $(U - Wg, V + Wf)$ for any W in $K_0[X, Y]$. By Euclidean division in the ring $K_0[Y][X]$ of U by g, one gets a solution (U, V) with $\deg U < \deg g$, and then $\deg V < \deg f$. When f and g have no common factor, such a pair (U, V) is unique up to a multiplicative constant. When f and g have their coefficients in a domain A_0 in place of a field K_0, one takes for K_0 the quotient field of A_0 and one multiplies by a denominator, so that U and V can be taken as polynomials in $A_0[X, Y]$, and then $R \in A_0$.

The multiplicities of intersection of the two curves are reflected by the multiplicities of zeros of the roots of R as a polynomial in Y.

It is useful to work with a ring A more general than $A_0[Y]$. Let A be a commutative ring with unity. Denote by S the ring $A[X]$ of polynomials in one variable with coefficients in A. For d a non-negative integer, let S_d be the A–module of elements in S of degree $\leq d$. Then S_d is a free A–module of rank $d + 1$ with a basis $1, X, \ldots, X^d$.

Let P and Q be polynomials of degrees p and q respectively:

$$P(X) = a_0 + a_1X + \cdots + a_pX^p, \qquad Q(X) = b_0 + b_1X + \cdots + b_qX^q.$$

The homomorphism of A–modules

$$\begin{array}{ccc} S_{q-1} \times S_{p-1} & \longrightarrow & S_{p+q-1} \\ (U, V) & \longmapsto & UP + VQ \end{array}$$

has the following matrix in the given bases: for q larger than p,

$$\begin{pmatrix}
a_0 & 0 & \cdot & \cdot & \cdot & 0 & b_0 & 0 & \cdots & 0 \\
a_1 & a_0 & \cdot & \cdot & \cdot & 0 & b_1 & b_0 & \cdots & 0 \\
\vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\
a_{p-1} & a_{p-2} & \cdot & \cdot & \cdot & 0 & b_{p-1} & b_{p-2} & \cdots & b_0 \\
a_p & a_{p-1} & \cdot & \cdot & \cdot & 0 & b_p & b_{p-1} & \cdots & b_1 \\
0 & a_p & \cdot & \cdot & \cdot & 0 & b_{p+1} & b_p & \cdots & b_2 \\
\vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdot & \cdot & \cdot & a_0 & b_{q-1} & b_{q-2} & \cdots & b_{q-p} \\
0 & 0 & \cdot & \cdot & \cdot & a_1 & b_q & b_{q-1} & \cdots & b_{q-p+1} \\
0 & 0 & \cdot & \cdot & \cdot & a_2 & 0 & b_q & \cdots & b_{q-p+2} \\
\vdots & \vdots & \cdot & \cdot & \cdot & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdot & \cdot & \cdot & a_p & 0 & 0 & \cdots & b_q
\end{pmatrix}$$

and for p larger than q,

$$\begin{pmatrix}
a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdot & \cdot & \cdot & 0 \\
a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdot & \cdot & \cdot & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdot & \cdot & \cdot & \vdots \\
a_{q-1} & a_{q-2} & \cdots & a_0 & b_{q-1} & b_{q-2} & \cdot & \cdot & \cdot & 0 \\
a_q & a_{q-1} & \cdots & a_1 & b_q & b_{q-1} & \cdot & \cdot & \cdot & 0 \\
a_{q+1} & a_q & \cdots & a_2 & 0 & b_q & \cdot & \cdot & \cdot & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdot & \cdot & \cdot & \vdots \\
a_{p-1} & a_{p-2} & \cdots & a_{p-q} & 0 & 0 & \cdot & \cdot & \cdot & b_0 \\
a_p & a_{p-1} & \cdots & a_{p-q-1} & 0 & 0 & \cdot & \cdot & \cdot & b_1 \\
0 & a_p & \cdots & a_{p-q-2} & 0 & 0 & \cdot & \cdot & \cdot & b_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdot & \cdot & \cdot & \vdots \\
0 & 0 & \cdots & a_p & 0 & 0 & \cdot & \cdot & \cdot & b_q
\end{pmatrix}$$

The q first columns are the components, in the basis $(1, X, \ldots, X^{p+q-1})$, of $P, XP, \ldots, X^{q-1}P$, while the p last columns are the components, in the same basis, of $Q, XQ, \ldots, X^{p-1}Q$. The main diagonal is $(a_0, \ldots, a_0, b_q, \ldots, b_q)$.

Definition. The *resultant* of P and Q is the determinant of this matrix. We denote it by $\mathrm{Res}(P, Q)$. The *universal resultant* is the resultant of the two polynomials

$$U_0 + U_1 X + \cdots + U_p X^p \quad \text{and} \quad V_0 + V_1 X + \cdots + V_q X^q,$$

in the ring $A_{pq} = \mathbf{Z}[U_0, U_1, \ldots, U_p, V_0, V_1, \ldots, V_q]$ of polynomials with coefficients in $\mathbf{Z}$ in $p+q+2$ variables. One deduces the resultant of P and Q by *specialisation*, i.e., as the image under the canonical homomorphism from A_{pq} to A which maps U_i to a_i and V_j to b_j. When the characteristic is 0, this canonical homomorphism is injective.

From the above expression of the resultant as a determinant, one deduces:

Proposition 34. *The universal resultant is a polynomial in*

$$U_0, U_1, \ldots, U_p, V_0, V_1, \ldots, V_q$$

which is homogeneous of degree q in $U_0, \ldots, U_p$, and homogeneous of degree p in $V_0, \ldots, V_q$.

Proposition 35. *There exist two polynomials U and V in $A[X]$, of degrees $< q$ and $< p$ respectively, such that the resultant $R = \mathrm{Res}(P, Q)$ of P and Q can be written $R = UP + VQ$.*

It follows that if P and Q have a common zero in some field containing A, then $\mathrm{Res}(P, Q) = 0$. The converse is true. It uses the following easy property, whose proof is left as an exercise.

Proposition 36. *Let A_0 be a ring, $A = A_0[Y_1, \ldots, Y_n]$ the ring of polynomials in n variables with coefficients in A_0, and P, Q elements in $A_0[Y_0, \ldots, Y_n]$, homogeneous of degrees p and q respectively. Consider P and Q as elements in $A[Y_0]$ and denote by $R = \mathrm{Res}_{Y_0}(P, Q) \in A$ their resultant with respect to Y_0. Then R is homogeneous of degree pq in $Y_1, \ldots, Y_n$.*

From these properties we deduce:

Proposition 37. *If*

$$P(X) = a_0 \prod_{i=1}^{p} (X - \alpha_i) \quad \textit{and} \quad Q(X) = b_0 \prod_{j=1}^{q} (X - \beta_j),$$

then

$$\begin{aligned} \mathrm{Res}(P, Q) &= a_0^q b_0^p \textstyle\prod_{i=1}^{p} \prod_{j=1}^{q} (\alpha_i - \beta_j) \\ &= (-1)^{pq} b_0^p \textstyle\prod_{j=1}^{q} P(\beta_j) \\ &= a_0^q \textstyle\prod_{i=1}^{p} Q(\alpha_i). \end{aligned}$$

Proof. Without loss of generality, one may assume that A is the ring of polynomials with coefficients in $\mathbf{Z}$ in the variables $a_0, b_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$. In this factorial ring, $\alpha_i - \beta_j$ is an irreducible element which divides $R = \mathrm{Res}(P, Q)$ (indeed, if one specializes $\alpha_i = \beta_j$, then the resultant vanishes). Now

$$a_0^q b_0^p \prod_{i=1}^{p} \prod_{j=1}^{q} (\alpha_i - \beta_j)$$

is homogeneous of degree q in the coefficients of P and of degree p in the coefficients of Q. Therefore it can be written cR with some $c \in \mathbf{Z}$. Finally the coefficient of the monomial $a_0^p b_0^q$ is 1, hence $c = 1$. □

Corollary 38. *Let K be a field containing A in which P and Q completely split in factors of degree* 1. *Then the resultant* $\mathrm{Res}(P, Q)$ *is zero if and only if P and Q have a common zero in K.*

Corollary 39. *If the ring A is factorial, then* $\mathrm{Res}(P, Q) = 0$ *if and only if P and Q have a common irreducible factor.*

References

[1] M. Aigner and G. M. Ziegler, *Proofs from The Book*, Springer-Verlag, Berlin, fourth ed., 2010.

[2] C. Brezinski, *The long history of continued fractions and Padé approximants*, in Padé approximation and its applications (Amsterdam, 1980), vol. 888 of Lecture Notes in Math., Springer, Berlin, 1981, pp. 1–27.

[3] ——, *History of continued fractions and Padé approximants*, vol. 12 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1991.
`http://www.emis.de/cgi-bin/MATH-item?0714.01001`.

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

[5] J. W. S. Cassels, *An introduction to Diophantine approximation*, Hafner Publishing Co., New York, 1972. Facsimile reprint of the 1957 edition, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45.

[6] A. Chantanasiri, *On the criteria for linear independence of Nesterenko, Fischler and Zudilin.* Chamchuri Journal of Math., to appear, 13 p.
arXiv:0912.4904v1 (math.NT).

[7] H. Cohn, *A short proof of the simple continued fraction expansion of e*, Amer. Math. Monthly, 113 (2006), pp. 57–62.
arXiv:math/0601660.

[8] J. Cosgrave, *New proofs of the irrationality of e^2 and e^4.*
http://staff.spd.dcu.ie/johnbcos/transcendental_numbers.htm
http://staff.spd.dcu.ie/johnbcos/esquared.htm.

[9] A. Durand, *Un critère de transcendance*, in Séminaire Delange-Pisot-Poitou (15e année: 1973/74), Théorie des nombres, Fasc. 2, Exp. No. G11, Secrétariat Mathématique, Paris, 1975, p. 9.
http://www.numdam.org/numdam-bin/fitem?id=SDPP_1973-1974__15_2_A7_0.

[10] ——, *Indépendance algébrique de nombres complexes et critère de transcendance*, Compositio Math., 35 (1977), pp. 259–267.
http://www.numdam.org/numdam-bin/fitem?id=CM_1977__35_3_259_0.

[11] L. Euler, *De fractionibus continuis dissertatio*, Commentarii Acad. Sci. Petropolitanae, 9 (1737), pp. 98–137. Opera Omnia Ser. I vol. 14, Commentationes Analyticae, p. 187–215.
Classification Enseström E71 – Archive Euler
www.math.dartmouth.edu/∼euler/pages/E071.html.

[12] P. Eymard and J.-P. Lafon, *The number π*, American Mathematical Society, Providence, RI, 2004. Translated from the 1999 French original by Stephen S. Wilson.

[13] N. I. Fel′dman and Y. V. Nesterenko, *Transcendental numbers*, in Number Theory, IV, vol. 44 of Encyclopaedia Math. Sci., Springer, Berlin, 1998.

[14] S. Fischler, *Irrationalité de valeurs de zêta (d'après Apéry, Rivoal,...)*, Astérisque, 294 (2004), pp. 27–62. Séminaire Bourbaki, Vol. 2002/2003.

[15] S. Fischler and W. Zudilin, *A refinement of Nesterenko's linear independence criterion with applications to zeta values.* Math. Annalen, to appear.
http://www.mpim-bonn.mpg.de/preprints/send?bid=4020.

[16] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, Oxford, sixth ed., 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[17] C. Hermite, *Sur la fonction exponentielle*, C. R. Acad. Sci. Paris, 77 (1873), pp. 18–24, 74–79, 226–233, 285–293. Œuvres de Charles Hermite, Paris: Gauthier-Villars, (1905), III, 150–181. See also *Oeuvres* III, 127–130, 146–149, and *Correspondance Hermite-Stieltjes*, II, lettre 363, 291–295. University of Michigan Historical Math Collection `http://name.umdl.umich.edu/AAS7821.0001.001`.

[18] ——, *Sur la fonction exponentielle*, C. R. Acad. Sci. Paris, 77 (1873), pp. 18–24, 74–79, 226–233, 285–293. Œuvres de Charles Hermite, Paris: Gauthier-Villars, 1905-1917. University of Michigan Historical Math Collection `http://name.umdl.umich.edu/AAS7821.0001.001`.

[19] A. Y. Khinchin, *Continued fractions*, Dover Publications Inc., Mineola, NY, russian ed., 1997. With a preface by B. V. Gnedenko, Reprint of the 1964 translation.

[20] 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., 1767.

[21] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966. Collected papers. Vol. I (1952–1970), Springer–Verlag 2000, p. 396–506.

[22] M. Laurent, *Cours de DEA à l'Université de Marseille, IML (Institut de Mathématiques de Luminy).* Unpublished manuscript notes, 2007.

[23] H. W. Lenstra, Jr., *Solving the Pell equation*, Notices Amer. Math. Soc., 49 (2002), pp. 182–192. `http://www.ams.org/notices/200202/fea-lenstra.pdf`.

[24] J. Liouville, *Addition à la note sur l'irrationalité du nombre e*, J. Math. Pures Appl., 1 (1840), pp. 193–194. `http://portail.mathdoc.fr/JMPA/`.

[25] ——, *Sur l'irrationalité du nombre* $e = 2,718\ldots$, J. Math. Pures Appl., 1 (1840), p. 192. `http://gallica.bnf.fr/Catalogue/noticesInd/FRBNF34348784.htm`.

[26] Y. V. Nesterenko, *Linear independence of numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh., (1985), pp. 46–49, 108.

[27] ——, *Algebraic independence*, Published for the Tata Institute of Fundamental Research, Bombay, 2009.

[28] Y. V. Nesterenko and P. Philippon, eds., *Introduction to algebraic independence theory*, vol. 1752 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2001.

[29] B. Rittaud, *Le fabuleux destin de* $\sqrt{2}$. Éditions Le Pommier, 2006.

[30] D. Roy, *Philippon's criterion for algebraic independence (lectures 3 and 4)*. Summer School in Analytic Number Theory and Diophantine Approximation University of Ottawa, Ontario Canada, June 30-July 11, 2008
http://aix1.uottawa.ca/∼droy/summer-school-2008/droy_lecture3-4.pdf.

[31] W. M. Schmidt, *Diophantine approximation*, vol. **785**, Lecture Notes in Mathematics. Berlin-Heidelberg-New York: Springer-Verlag, 1980, new ed. 2001.

[32] S. Shirali, *Continued fraction for e*. Resonance, vol. **5** N°1, Jan. 2000, 14–28.
http://www.ias.ac.in/resonance/.

[33] C. L. Siegel, *Über einige Anwendungen diophantischer Approximationen*, Abhandlungen Akad. Berlin, Nr. 1 (1929), p. 70 S.

[34] ——, *Transcendental numbers*, Princeton, N. J.: Princeton University Press, 1949.

[35] J. Sondow, *Criteria for irrationality of Euler's constant*, Proc. Amer. Math. Soc., 131 (2003), pp. 3335–3344 (electronic). Proc. Amer. Math. Soc. **131** (2003), 3335–3344
http://xxx.lanl.gov/pdf/math.NT/0209070
http://home.earthlink.net/∼jsondow/.

[36] B. Sury, *Bessels contain continued fractions of progressions*. Resonance, vol. **10** N°3, March 2005, 80–87.
http://www.ias.ac.in/resonance/.

[37] M. Waldschmidt, *Nombres transcendants*, Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 402
http://www.springerlink.com/content/110312/ .

[38] ——, *Diophantine approximation on linear algebraic groups*, vol. 326 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2000. Transcendence properties of the exponential function in several variables.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°4, *April 28, 2010*

This course was devoted to Liouville's inequality (§ 4.1).
The present notes consist of

- Pages 65–85 of [38] (beginning of Chapter 3: Heights).
- Liouville's inequality for quadratic numbers.
- A short historical survey on Diophantine Approximation.

4.1.2 Liouville's inequality for quadratic numbers

Consider Lemma 24 in the special case $d = 2$ where α is a quadratic algebraic number. Write its minimal polynomial $f(X) = aX^2 + bX + c$ and let $\Delta := b^2 - 4ac$ be its discriminant. Since we are interested in the approximation of α by rational numbers, we assume $\Delta > 0$. If $\alpha = (-b \pm \sqrt{\Delta})/2a$, then the other root is $\alpha' = (-b \mp \sqrt{\Delta})/2a$ and

$$f'(\alpha) = a(\alpha - \alpha') = \pm\sqrt{\Delta}.$$

Lemma 40. *Let α be an algebraic number of degree 2 and minimal polynomial $P \in \mathbf{Z}[X]$. Define Let $\epsilon > 0$. Then there exists an integer q_0 such that, for any $p/q \in \mathbf{Q}$ with $q \geq q_0$,*

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{1}{(\sqrt{\Delta} + \epsilon)q^2}.$$

The smallest positive discriminant of an irreducible quadratic polynomial with coefficients in $\mathbf{Z}$ is 5, which is the value of the discriminant of $X^2 - X - 1$, with roots Φ and $-\Phi^{-1}$ where $\Phi = 1.6180339887499\ldots$ denotes the Golden ratio.

The next result deals with the Fibonacci sequence $(F_n)_{n\geq 0}$:

$$F_0 = 0,\ F_1 = 1,\ F_n = F_{n-1} + F_{n-2} \quad (n \geq 2).$$

Lemma 41. *For any $q \geq 1$ and any $p \in \mathbf{Z}$,*

$$\left|\Phi - \frac{p}{q}\right| > \frac{1}{\sqrt{5}q^2 + (q/2)}.$$

On the other hand

$$\lim_{n\to\infty} F_{n-1}^2 \left|\Phi - \frac{F_n}{F_{n-1}}\right| = \frac{1}{\sqrt{5}}.$$

Proof. It suffices to prove the lower bound when p is the nearest integer to $q\Phi$. From $X^2 - X - 1 = (X - \Phi)(X + \Phi^{-1})$ we deduce

$$p^2 - pq - q^2 = q^2 \left(\frac{p}{q} - \Phi\right)\left(\frac{p}{q} + \Phi^{-1}\right).$$

The left hand side is a non-zero rational integer, hence has absolute value at least 1. We now bound the absolute value of the right hand side from above. Since $p < q\Phi + (1/2)$ and $\Phi + \Phi^{-1} = \sqrt{5}$ we have

$$\frac{p}{q} + \Phi^{-1} < \sqrt{5} + \frac{1}{2q}.$$

Hence

$$1 < q^2 \left| \frac{p}{q} - \Phi \right| \left(\sqrt{5} + \frac{1}{2q} \right)$$

The first part of Lemma 41 follows.

The real vector space of sequences $(v_n)_{n\geq 0}$ satisfying $v_n = v_{n-1}+v_{n-2}$ has dimension 2, a basis is given by the two sequences $(\Phi^n)_{n\geq 0}$ and $((-\Phi^{-1})^n)_{n\geq 0}$. From this one easily deduces the formula

$$F_n = \frac{1}{\sqrt{5}}(\Phi^n - (-1)^n \Phi^{-n})$$

due to A. De Moivre (1730), L. Euler (1765) and J.P.M. Binet (1843). It follows that F_n is the nearest integer to

$$\frac{1}{\sqrt{5}}\Phi^n,$$

hence the sequence $(u_n)_{n\geq 2}$ of quotients of Fibonacci numbers

$$u_n = F_n/F_{n-1}$$

satisfies $\lim_{n\to\infty} u_n = \Phi$.

By induction one easily checks

$$F_n^2 - F_n F_{n-1} - F_{n-1}^2 = (-1)^{n-1}$$

for $n \geq 1$. The left hand side is $F_{n-1}^2(u_n - \Phi)(u_n + \Phi^{-1})$, as we already saw. Hence

$$F_{n-1}^2 |\Phi - u_n| = \frac{1}{\Phi^{-1} + u_n},$$

and the limit of the right hand side is $1/(\Phi + \Phi^{-1}) = 1/\sqrt{5}$. The result follows.

□

Remark. *The sequence $u_n = F_n/F_{n-1}$ is also defined by*

$$u_2 = 2,\ u_n = 1 + \frac{1}{u_{n-1}}, \quad (n \geq 3).$$

Hence

$$u_n = 1 + \frac{1}{1 + \dfrac{1}{u_{n-2}}} = 1 + \frac{1}{1 + \dfrac{1}{1 + \dfrac{1}{u_{n-3}}}} = \cdots$$

Remark. *It is known (see for instance [31] p. 25) that if k is a positive integer, if an irrational real number ϑ has a continued fraction expansion $[a_0; a_1, a_2, \dots]$ with $a_n \geq k$ for infinitely many n, then*

$$\liminf_{q\to\infty} q^2 \left|\vartheta - \frac{p}{q}\right| \leq \frac{1}{\sqrt{4+k^2}}.$$

4.1.3 Diophantine Approximation: historical survey

References for this section are [2, 31, 13, 1].

Definition Given a real irrational number ϑ, a function $\varphi = \mathbf{N} \to \mathbf{R}_{>0}$ is an irrationality measure for ϑ if there exists an integer $q_0 > 0$ such that, for any $p/q \in \mathbf{Q}$ 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) $\Rightarrow$ (iv) in Proposition 4) it follows that any irrationality exponent κ satisfies $\kappa \geq 2$. Irrational quadratic numbers have irrationality exponent 2. It is known (see for instance [31] 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* $(a_0, a_1, \dots)$ in the continued fraction expansion of ϑ is bounded: these are called the *badly approximable numbers*.

From Liouville's inequality in Lemma 24 it follows that any irrational algebraic real number α of degree d 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 connexion with the triadic Cantor set (see [3]) 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 42 (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. We start with a special example. Liouville's estimate for the rational Diophantine approximation of $\sqrt[3]{2}$ is

$$\left|\sqrt[3]{2} - \frac{p}{q}\right| > \frac{1}{9q^3}$$

for sufficiently large q (use Lemma 24 with $P(X) = X^3 - 2$, $c = 3\sqrt[3]{2} < 9$). Thue was the first to achieve an improvement of the exponent 3. An explicit estimate was then obtained by A. Baker, namely

$$\left|\sqrt[3]{2} - \frac{p}{q}\right| > \frac{1}{10^6 q^{2.955}},$$

and refined by Chudnovskii, Easton, Rickert, Voutier and others, until 1997 when M. Bennett proved that *for any $p/q \in \mathbf{Q}$,*

$$\left|\sqrt[3]{2} - \frac{p}{q}\right| \geq \frac{1}{4\ q^{2,5}}.$$

From his own result, Thue deduced that *for any fixed $k \in \mathbf{Z} \setminus \{0\}$, there are only finitely many $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ satisfying the Diophantine equation $x^3 - 2y^3 = k$.* The result of Baker shows more precisely that if $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ is a solution to $x^3 - 2y^3 = k$, then

$$|x| \leq 10^{137} |k|^{23}.$$

M. Bennett gave the sharper estimate: *for any $(x, y) \in \mathbf{Z}^2$ with $x > 0$,*

$$|x^3 - 2y^3| \geq \sqrt{x}.$$

The connexion between Diophantine approximation to $\sqrt[3]{2}$ and the Diophantine equation $x^3 - 2y^3 = k$ is explained in the next lemma.

Lemma 43. *Let η be a positive real number. The two following properties are equivalent:*
(i) *There exists a constant $c_1 > 0$ such that, for any $p/q \in \mathbf{Q}$ with $q > 0$,*

$$\left|\sqrt[3]{2} - \frac{p}{q}\right| > \frac{c_1}{q^\eta}.$$

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

$$|x^3 - 2y^3| \geq c_2 x^{3-\eta}.$$

Properties (i) and (ii) are true but uninteresting with $\eta \geq 3$. They are true with $\eta = 3$ ((i) is Liouville's estimate while (ii) is trivial), they are true also for any $\eta > 2$ by Roth's Theorem. They are not true with $\eta < 2$. It is expected that they are not true with $\eta = 2$. The constants are explicit for $\eta \geq 2.5$ by Bennett's result, but not yet for η in the range $2 < \eta < 2.5$.

Proof. We assume $\eta < 3$, otherwise the result is trivial. Set $\alpha = \sqrt[3]{2}$.

Assume (i) and let $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ have $x > 0$. Set $k = x^3 - 2y^3$. Since 2 is not the cube of a rational number we have $k \neq 0$. If $y = 0$ assertion (ii) plainly holds. So assume $y \neq 0$.

Write

$$x^3 - 2y^3 = (x - \alpha y)(x^2 + \alpha xy + \alpha^2 y^2).$$

The polynomial $X^2 + \alpha X + \alpha^2$ has negative discriminant $-3\alpha^2$, hence has a positive minimum $c_0 = 3\alpha^2/4$. Hence the value at (x, y) of the quadratic form $X^2 + \alpha XY + \alpha^2 Y^2$ is bounded form below by $c_0 y^2$. From (i) we deduce

$$|k| = |y|^3 \left| \sqrt[3]{2} - \frac{x}{y} \right| (x^2 + \alpha xy + \alpha^2 y^2) \geq \frac{c_1 c_0 |y|^3}{|y|^\eta} = c_3 |y|^{3-\eta}.$$

This gives an upper bound for $|y|$:

$$|y| \leq c_4 |k|^{1/(3-\eta)}, \quad \text{hence} \quad |y^3| \leq c_4 |k|^{3/(3-\eta)}.$$

We want an upper bound for x: we use $x^3 = k + 2y^3$ and we bound $|k|$ by $|k|^{3/(3-\eta)}$ since $3/(3-\eta) > 1$. Hence

$$x^3 \leq c_5 |k|^{3/(3-\eta)} \quad \text{and} \quad x^{3-\eta} \leq c_6 |k|.$$

Conversely, assume (ii). Let p/q be a rational number. If p is not the nearest integer to $q\alpha$, then $|q\alpha - p| > 1/2$ and the estimate (i) is trivial. So we assume $|q\alpha - p| \leq 1/2$. We need only the weaker estimate $c_7 q < p < c_8 q$ with some positive constants c_7 and c_8, showing that we may replace p by q or q by p in our estimates, provided that we adjust the constants. From

$$p^3 - 2q^3 = (p - \alpha q)(p^2 + \alpha pq + \alpha^2 q^2),$$

using (ii), we deduce

$$c_2 p^{3-\eta} \leq c_{10} q^3 \left| \alpha - \frac{p}{q} \right|,$$

and (i) easily follows.

□

Here is the most general result of Thue on Diophantine equations.

Theorem 44 (Thue)**.** *Let $f \in \mathbf{Z}[X]$ 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 $(x,y) \in \mathbf{Z} \times \mathbf{Z}$.*

The equation $F(x,y) = m$ in Proposition 44 is called *Thue equation*. The connexion between Thue equation and Liouville's inequality has been explained in Lemma 43 in the special case $\sqrt[3]{2}$; the general case is similar.

Lemma 45. *Let α be an algebraic number of degree $d \geq 3$ and minimal polynomial $f \in \mathbf{Z}[X]$, let $F(X,Y) = Y^d f(X/Y) \in \mathbf{Z}[X,Y]$ 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 $p/q \in \mathbf{Q}$,*

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

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

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

In 1921 C.L. Siegel sharpened Thue's result 42 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 42 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 46 (A. Thue, C.L. Siegel, F. Dyson, K.F. Roth 1955)**.** *For any real algebraic number α, for any $\epsilon > 0$, the set of $p/q \in \mathbf{Q}$ 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!

If one restricts the denominators q of the rational approximations p/q by requesting that their prime factor belong to a given finite set, then the exponent 2 can be replaced by 1. This has been proved by D. Ridout in 1957.

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 47 (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 $p/q \in \mathbf{Q}$, with q a S–integer and $|\alpha - p/q| < q^{-1-\epsilon}$, is finite.*

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 do not state it in full generality but we give only two special cases.

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

$$|\mathbf{x}| = \max\{|x_1|, \ldots, |x_m|\}.$$

Theorem 48 (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*

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

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

Thue–Siegel–Roth's Theorem 46 follows from Theorem 48 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 $\mathbf{Q}$-vector subspace of $\mathbf{Q}^2$ which is not $\{0\}$ not $\mathbf{Q}^2$ (that is *a proper subspace*) generated by an element $(p_0, q_0) \in \mathbf{Q}^2$. There is one such subspace with $q_0 = 0$, namely $\mathbf{Q} \times \{0\}$ 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 48 says that there is only a finite set of exceptions p/q in Roth's Theorem.

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

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

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

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

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

$$|\mathbf{x}| = \max\{|x_1|_p, \ldots, |x_m|_p\}.$$

Theorem 49 (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 $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbf{Z}^m$ 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 $\mathbf{Q}^m$.

Ridout's Theorem 47 is a corollary of Schmidt's Subspace Theorem: in Theorem 49 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

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

and

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

since b is a S–integer.

Further references

References

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

[2] 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.

[3] M. Waldschmidt – *Report on some recent progress in Diophantine approximation*. To appear
http://www.math.jussieu.fr/∼miw/articles/pdf/miwLangMemorialVolume.pdf
Number Theory Math arXiv: http://fr.arxiv.org/abs/0908.3973
Archives Ouvertes http://hal.archives-ouvertes.fr/hal-00407199/fr/

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°5, *May 3, 2010*

6 Continued fractions

We first consider generalized continued fractions of the form

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{\ddots}}},$$

which we denote by[5]

$$a_0 + \frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \frac{b_3|}{\ddots}.$$

Next we restrict to the special case where $b_1 = b_2 = \cdots = 1$, which yields the simple continued fractions

$$a_0 + \frac{1\,|}{|a_1} + \frac{1\,|}{|a_2} + \cdots = [a_0, a_1, a_2, \dots],$$

already considered in section § 1.1.

6.1 Generalized continued fractions

To start with, $a_0, \dots, a_n, \dots$ and $b_1, \dots, b_n, \dots$ will be independent variables. Later, we shall specialize to positive integers (apart from a_0 which may be negative).

[5] Another notation for $a_0 + \frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \cdots + \frac{b_n|}{a_n}$ introduced by Th. Muir and used by Perron in [7] Chap. 1 is

$$K\begin{pmatrix} b_1, \dots, b_n \\ a_0, a_1, \dots, a_n \end{pmatrix}$$

Consider the three rational fractions

$$a_0, \qquad a_0 + \frac{b_1}{a_1} \quad \text{and} \quad a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2}}\cdot$$

We write them as

$$\frac{A_0}{B_0}, \qquad \frac{A_1}{B_1} \quad \text{and} \quad \frac{A_2}{B_2}$$

with

$$\begin{array}{lll} A_0 = a_0, & A_1 = a_0a_1 + b_1, & A_2 = a_0a_1a_2 + a_0b_2 + a_2b_1, \\ B_0 = 1, & B_1 = a_1, & B_2 = a_1a_2 + b_2. \end{array}$$

Observe that

$$A_2 = a_2A_1 + b_2A_0, \quad B_2 = a_2B_1 + b_2B_0.$$

Write these relations as

$$\begin{pmatrix} A_2 & A_1 \\ B_2 & B_1 \end{pmatrix} = \begin{pmatrix} A_1 & A_0 \\ B_1 & B_0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ b_2 & 0 \end{pmatrix}.$$

Define inductively two sequences of polynomials with positive rational coefficients A_n and B_n for $n \geq 3$ by

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} A_{n-1} & A_{n-2} \\ B_{n-1} & B_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}. \tag{50}$$

This means

$$A_n = a_nA_{n-1} + b_nA_{n-2}, \qquad B_n = a_nB_{n-1} + b_nB_{n-2}.$$

This recurrence relation holds for $n \geq 2$. It will also hold for $n = 1$ if we set $A_{-1} = 1$ and $B_{-1} = 0$:

$$\begin{pmatrix} A_1 & A_0 \\ B_1 & B_0 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix}$$

and it will hold also for $n = 0$ if we set $b_0 = 1$, $A_{-2} = 0$ and $B_{-2} = 1$:

$$\begin{pmatrix} A_0 & A_{-1} \\ B_0 & B_{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ b_0 & 0 \end{pmatrix}.$$

Obviously, an equivalent definition is

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ b_0 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ b_{n-1} & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}. \tag{51}$$

These relations (51) hold for $n \geq -1$, with the empty product (for $n = -1$) being the identity matrix, as always.

Hence $A_n \in \mathbf{Z}[a_0, \ldots, a_n, b_1, \ldots, b_n]$ is a polynomial in $2n+1$ variables, while $B_n \in \mathbf{Z}[a_1 \ldots, a_n, b_2, \ldots, b_n]$ is a polynomial in $2n-1$ variables.

Exercise 6. *Check, for $n \geq -1$,*

$$B_n(a_1, \ldots, a_n, b_2, \ldots, b_n) = A_{n-1}(a_1, \ldots, a_n, b_2, \ldots, b_n).$$

Lemma 52. *For $n \geq 0$,*

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_n|}{|a_n} = \frac{A_n}{B_n}.$$

Proof. By induction. We have checked the result for $n = 0$, $n = 1$ and $n = 2$. Assume the formula holds with $n-1$ where $n \geq 3$. We write

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|x}$$

with

$$x = a_{n-1} + \frac{b_n}{a_n}.$$

We have, by induction hypothesis and by the definition (50),

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} = \frac{A_{n-1}}{B_{n-1}} = \frac{a_{n-1}A_{n-2} + b_{n-1}A_{n-3}}{a_{n-1}B_{n-2} + b_{n-1}B_{n-3}}.$$

Since A_{n-2}, A_{n-3}, B_{n-2} and B_{n-3} do not depend on the variable a_{n-1}, we deduce

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|x} = \frac{xA_{n-2} + b_{n-1}A_{n-3}}{xB_{n-2} + b_{n-1}B_{n-3}}.$$

The product of the numerator by a_n is

$$\begin{aligned}(a_n a_{n-1} + b_n)A_{n-2} + a_n b_{n-1}A_{n-3} &= a_n(a_{n-1}A_{n-2} + b_{n-1}A_{n-3}) + b_n A_{n-2} \\ &= a_n A_{n-1} + b_n A_{n-2} = A_n\end{aligned}$$

and similarly, the product of the denominator by a_n is

$$\begin{aligned}(a_n a_{n-1} + b_n)B_{n-2} + a_n b_{n-1}B_{n-3} &= a_n(a_{n-1}B_{n-2} + b_{n-1}B_{n-3}) + b_n B_{n-2} \\ &= a_n B_{n-1} + b_n B_{n-2} = B_n.\end{aligned}$$

□

From (51), taking the determinant, we deduce, for $n \geq -1$,

$$A_n B_{n-1} - A_{n-1} B_n = (-1)^{n+1} b_0 \cdots b_n. \tag{53}$$

which can be written, for $n \geq 1$,

$$\frac{A_n}{B_n} - \frac{A_{n-1}}{B_{n-1}} = \frac{(-1)^{n+1} b_0 \cdots b_n}{B_{n-1} B_n}. \tag{54}$$

Adding the telescoping sum, we get, for $n \geq 0$,

$$\frac{A_n}{B_n} = A_0 + \sum_{k=1}^{n} \frac{(-1)^{k+1} b_0 \cdots b_k}{B_{k-1} B_k}. \tag{55}$$

We now substitute for $a_0, a_1, \ldots$ and $b_1, b_2, \ldots$ rational integers, all of which are ≥ 1, apart from a_0 which may be ≤ 0. We denote by p_n (resp. q_n) the value of A_n (resp. B_n) for these special values. Hence p_n and q_n are rational integers, with $q_n > 0$ for $n \geq 0$. A consequence of Lemma 52 is

$$\frac{p_n}{q_n} = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_n|}{|a_n} \quad \text{for} \quad n \geq 0.$$

We deduce from (50),

$$p_n = a_n p_{n-1} + b_n p_{n-2}, \qquad q_n = a_n q_{n-1} + b_n q_{n-2} \quad \text{for} \quad n \geq 0,$$

and from (53),

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} b_0 \cdots b_n \quad \text{for} \quad n \geq -1,$$

which can be written, for $n \geq 1$,

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n+1} b_0 \cdots b_n}{q_{n-1} q_n}. \tag{56}$$

Adding the telescoping sum (or using (55)), we get the alternating sum

$$\frac{p_n}{q_n} = a_0 + \sum_{k=1}^{n} \frac{(-1)^{k+1} b_0 \cdots b_k}{q_{k-1} q_k}. \tag{57}$$

Recall that for real numbers a, b, c, d, with b and d positive, we have

$$\frac{a}{b} < \frac{c}{d} \implies \frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}. \tag{58}$$

Since a_n and b_n are positive for $n \geq 0$, we deduce that for $n \geq 2$, the rational number

$$\frac{p_n}{q_n} = \frac{a_n p_{n-1} + b_n p_{n-2}}{a_n q_{n-1} + b_n q_{n-2}}$$

lies between p_{n-1}/q_{n-1} and p_{n-2}/q_{n-2}. Therefore we have

$$\frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots < \frac{p_{2n}}{q_{2n}} < \cdots < \frac{p_{2m+1}}{q_{2m+1}} < \cdots < \frac{p_3}{q_3} < \frac{p_1}{q_1}. \tag{59}$$

From (56), we deduce, for $n \geq 3$, $q_{n-1} > q_{n-2}$, hence $q_n > (a_n + b_n)q_{n-2}$.

The previous discussion was valid without any restriction, now we assume $a_n \geq b_n$ for all sufficiently large n, say $n \geq n_0$. Then for $n > n_0$, using $q_n > 2b_n q_{n-2}$, we get

$$\left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{b_0 \cdots b_n}{q_{n-1}q_n} < \frac{b_n \cdots b_0}{2^{n-n_0} b_n b_{n-1} \cdots b_{n_0+1} q_{n_0} q_{n_0-1}} = \frac{b_{n_0} \cdots b_0}{2^{n-n_0} q_{n_0} q_{n_0-1}}$$

and the right hand side tends to 0 as n tends to infinity. Hence the sequence $(p_n/q_n)_{n\geq 0}$ has a limit, which we denote by

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

From (57), it follows that x is also given by an alternating series

$$x = a_0 + \sum_{k=1}^{\infty} \frac{(-1)^{k+1} b_0 \cdots b_k}{q_{k-1}q_k}.$$

We now prove that x is irrational. Define, for $n \geq 0$,

$$x_n = a_n + \frac{b_{n+1}|}{|a_{n+1}} + \cdots$$

so that $x = x_0$ and, for all $n \geq 0$,

$$x_n = a_n + \frac{b_{n+1}}{x_{n+1}}, \qquad x_{n+1} = \frac{b_{n+1}}{x_n - a_n}$$

and $a_n < x_n < a_n + 1$. Hence for $n \geq 0$, x_n is rational if and only if x_{n+1} is rational, and therefore, if x is rational, then all x_n for $n \geq 0$ are also rational. Assume x is rational. Consider the rational numbers x_n with $n \geq n_0$ and select a value of n for which the denominator v of x_n is minimal, say $x_n = u/v$. From

$$x_{n+1} = \frac{b_{n+1}}{x_n - a_n} = \frac{b_{n+1} v}{u - a_n v} \quad \text{with} \quad 0 < u - a_n v < v,$$

it follows that x_{n+1} has a denominator strictly less than v, which is a contradiction. Hence x is irrational.

Conversely, given an irrational number x and a sequence $b_1, b_2, \ldots$ of positive integers, there is a unique integer a_0 and a unique sequence $a_1, \ldots, a_n, \ldots$ of positive integers satisfying $a_n \geq b_n$ for all $n \geq 1$, such that

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

Indeed, the unique solution is given inductively as follows: $a_0 = \lfloor x \rfloor$, $x_1 = b_1/\{x\}$, and once $a_0, \ldots, a_{n-1}$ and $x_1, \ldots, x_n$ are known, then a_n and x_{n+1} are given by

$$a_n = \lfloor x_n \rfloor, \quad x_{n+1} = b_{n+1}/\{x_n\},$$

so that for $n \geq 1$ we have $0 < x_n - a_n < 1$ and

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|x_n}.$$

Here is what we have proved.

Proposition 60. *Given a rational integer a_0 and two sequences $a_0, a_1, \ldots$ and $b_1, b_2, \ldots$ of positive rational integers with $a_n \geq b_n$ for all sufficiently large n, the infinite continued fraction*

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

exists and is an irrational number.
Conversely, given an irrational number x and a sequence $b_1, b_2, \ldots$ of positive integers, there is a unique $a_0 \in \mathbf{Z}$ and a unique sequence $a_1, \ldots, a_n, \ldots$ of positive integers satisfying $a_n \geq b_n$ for all $n \geq 1$ such that

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n} + \cdots$$

These results are useful for proving the irrationality of π and e^r when r is a non–zero rational number, following the proof by Lambert. See for instance Chapter 7 (Lambert's Irrationality Proofs) of David Angell's course on Irrationality and Transcendence([6]) at the University of New South Wales:

[6] I found this reference from the website of John Cosgrave
`http://staff.spd.dcu.ie/johnbcos/transcendental_numbers.htm`.

http://www.maths.unsw.edu.au/ angell/5535/

The following example is related with Lambert's proof [20]:

$$\tanh z = \frac{z|}{|1} + \frac{z^2|}{|\,3} + \frac{z^2|}{|\,5} + \cdots + \frac{z^2\ \ |}{|2n+1} + \cdots$$

Here, z is a complex number and the right hand side is a complex valued function. Here are other examples (see Sloane's Encyclopaedia of Integer Sequences([7]))

$$\frac{1}{\sqrt{e}-1} = 1 + \frac{2|}{|3} + \frac{4|}{|5} + \frac{6|}{|7} + \frac{8|}{|9} + \cdots = 1.541\,494\,082\ldots \qquad \text{(A113011)}$$

$$\frac{1}{e-1} = \frac{1|}{|1} + \frac{2|}{|2} + \frac{3|}{|3} + \frac{4|}{|4} + \cdots = 0.581\,976\,706\ldots \qquad \text{(A073333)}$$

Remark. *A variant of the algorithm of simple continued fractions is the following. Given two sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ of elements in a field K and an element x in K, one defines a sequence (possibly finite) $(x_n)_{n\geq 1}$ of elements in K as follows. If $x = a_0$, the sequence is empty. Otherwise x_1 is defined by $x = a_0 + (b_1/x_1)$. Inductively, once $x_1, \ldots, x_n$ are defined, there are two cases:*

- *If $x_n = a_n$, the algorithm stops.*
- *Otherwise, x_{n+1} is defined by*

$$x_{n+1} = \frac{b_{n+1}}{x_n - a_n}, \quad \text{so that} \quad x_n = a_n + \frac{b_{n+1}}{x_{n+1}}.$$

If the algorithm does not stop, then for any $n \geq 1$, one has

$$x = a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|x_n}.$$

In the special case where $a_0 = a_1 = \cdots = b_1 = b_2 = \cdots = 1$, the set of x such that the algorithm stops after finitely many steps is the set $(F_{n+1}/F_n)_{n\geq 1}$ of quotients of consecutive Fibonacci numbers. In this special case, the limit of

$$a_0 + \frac{b_1|}{|a_1} + \cdots + \frac{b_{n-1}|}{|a_{n-1}} + \frac{b_n|}{|a_n}$$

is the Golden ratio, which is independent of x, of course!

[7] http://www.research.att.com/∼njas/sequences/

6.2 Simple continued fractions

We restrict now the discussion of § 6.1 to the case where $b_1 = b_2 = \cdots = b_n = \cdots = 1$. We keep the notations A_n and B_n which are now polynomials in $\mathbf{Z}[a_0, a_1, \ldots, a_n]$ and $\mathbf{Z}[a_1, \ldots, a_n]$ respectively, and when we specialize to integers $a_0, a_1, \ldots, a_n \ldots$ with $a_n \geq 1$ for $n \geq 1$ we use the notations p_n and q_n for the values of A_n and B_n.

The recurrence relations (50) are now, for $n \geq 0$,

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} A_{n-1} & A_{n-2} \\ B_{n-1} & B_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}, \tag{61}$$

while (51) becomes, for $n \geq -1$,

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}. \tag{62}$$

From Lemma 52 one deduces, for $n \geq 0$,

$$[a_0, \ldots, a_n] = \frac{A_n}{B_n}.$$

Taking the determinant in (62), we deduce the following special case of (53)

$$A_n B_{n-1} - A_{n-1} B_n = (-1)^{n+1}. \tag{63}$$

The specialization of these relations to integral values of $a_0, a_1, a_2 \ldots$ yields

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{for } n \geq 0, \tag{64}$$

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{for } n \geq -1, \tag{65}$$

$$[a_0, \ldots, a_n] = \frac{p_n}{q_n} \qquad \text{for } n \geq 0 \tag{66}$$

and

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} \qquad \text{for } n \geq -1. \tag{67}$$

From (67), it follows that for $n \geq 0$, the fraction p_n/q_n is in lowest terms: $\gcd(p_n, q_n) = 1$.

Transposing (65) yields, for $n \geq -1$,

$$\begin{pmatrix} p_n & q_n \\ p_{n-1} & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}$$

from which we deduce, for $n \geq 1$,

$$[a_n, \ldots, a_0] = \frac{p_n}{p_{n-1}} \quad \text{and} \quad [a_n, \ldots, a_1] = \frac{q_n}{q_{n-1}}$$

Lemma 68. *For $n \geq 0$,*

$$p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n.$$

Proof. We multiply both sides of (64) on the left by the inverse of the matrix

$$\begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \quad \text{which is} \quad (-1)^n \begin{pmatrix} q_{n-2} & -p_{n-2} \\ -q_{n-1} & p_{n-1} \end{pmatrix}.$$

We get

$$(-1)^n \begin{pmatrix} p_n q_{n-2} - p_{n-2} q_n & p_{n-1} q_{n-2} - p_{n-2} q_{n-1} \\ -p_n q_{n-1} + p_{n-1} q_n & 0 \end{pmatrix} = \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}$$

□

6.2.1 Finite simple continued fraction of a rational number

Let u_0 and u_1 be two integers with u_1 positive. The first step in Euclid's algorithm to find the gcd of u_0 and u_1 consists in dividing u_0 by u_1:

$$u_0 = a_0 u_1 + u_2$$

with $a_0 \in \mathbf{Z}$ and $0 \leq u_2 < u_1$. This means

$$\frac{u_0}{u_1} = a_0 + \frac{u_2}{u_1},$$

which amonts to dividing the rational number $x_0 = u_0/u_1$ by 1 with quotient a_0 and remainder $u_2/u_1 < 1$. This algorithms continues with

$$u_m = a_m u_{m+1} + u_{m+2},$$

where a_m is the integral part of $x_m = u_m/u_{m+1}$ and $0 \leq u_{m+2} < u_{m+1}$, until some $u_{\ell+2}$ is 0, in which case the algorithms stops with

$$u_\ell = a_\ell u_{\ell+1}.$$

Since the gcd of u_m and u_{m+1} is the same as the gcd of u_{m+1} and u_{m+2}, it follows that the gcd of u_0 and u_1 is $u_{\ell+1}$. This is how one gets the regular continued fraction expansion $x_0 = [a_0, a_1, \ldots, a_\ell]$, where $\ell = 0$ in case x_0 is a rational integer, while $a_\ell \geq 2$ if x_0 is a rational number which is not an integer.

Exercise 7. *Compare with the geometrical construction of the continued fraction given in § 1.1.*
Give a variant of this geometrical construction where rectangles are replaced by segments.

Repeating what was already said in § 1.2, we can state

Proposition 69. *Any finite regular continued fraction*

$$[a_0, a_1, \ldots, a_n],$$

where $a_0, a_1, \ldots, a_n$ are rational numbers with $a_i \geq 2$ for $1 \leq i \leq n$ and $n \geq 0$, represents a rational number. Conversely, any rational number x has two representations as a continued fraction, the first one, given by Euclid's algorithm, is

$$x = [a_0, a_1, \ldots, a_n]$$

and the second one is

$$x = [a_0, a_1, \ldots, a_{n-1}, a_n - 1, 1].$$

If $x \in \mathbf{Z}$, then $n = 0$ and the two simple continued fractions representations of x are $[x]$ and $[x-1, 1]$, while if x is not an integer, then $n \geq 1$ and $a_n \geq 2$.

We shall use later (in the proof of Lemma 81 in § 6.3.7) the fact that any rational number has one simple continued fraction expansion with an odd number of terms and one with an even number of terms.

6.2.2 Infinite simple continued fraction of an irrational number

Given a rational integer a_0 and an infinite sequence of positive integers $a_1, a_2, \ldots$, the continued fraction

$$[a_0, a_1, \ldots, a_n, \ldots]$$

represents an irrational number. Conversely, given an irrational number x, there is a unique representation of x as an infinite simple continued fraction

$$x = [a_0, a_1, \ldots, a_n, \ldots]$$

Definitions The numbers a_n are the *partial quotients*, the rational numbers

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$$

are the *convergents* (in French *réduites*), and the numbers

$$x_n = [a_n, a_{n+1}, \ldots]$$

are the *complete quotients.*

From these definitions we deduce, for $n \geq 0$,

$$x = [a_0, a_1, \ldots, a_n, x_{n+1}] = \frac{x_{n+1}p_n + p_{n-1}}{x_{n+1}q_n + q_{n-1}}. \tag{70}$$

Lemma 71. *For $n \geq 0$,*

$$q_n x - p_n = \frac{(-1)^n}{x_{n+1}q_n + q_{n-1}}.$$

Proof. From (70) one deduces

$$x - \frac{p_n}{q_n} = \frac{x_{n+1}p_n + p_{n-1}}{x_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{(x_{n+1}q_n + q_{n-1})q_n}.$$

□

Corollary 72. *For $n \geq 0$,*

$$\frac{1}{q_{n+1} + q_n} < |q_n x - p_n| < \frac{1}{q_{n+1}}.$$

Proof. Since a_{n+1} is the integral part of x_{n+1}, we have

$$a_{n+1} < x_{n+1} < a_{n+1} + 1.$$

Using the recurrence relation $q_{n+1} = a_{n+1}q_n + q_{n-1}$, we deduce

$$q_{n+1} < x_{n+1}q_n + q_{n-1} < a_{n+1}q_n + q_{n-1} + q_n = q_{n+1} + q_n.$$

□

In particular, since $x_{n+1} > a_{n+1}$ and $q_{n-1} > 0$, one deduces from Lemma 71

$$\frac{1}{(a_{n+1} + 2)q_n^2} < \left| x - \frac{p_n}{q_n} \right| < \frac{1}{a_{n+1}q_n^2}. \tag{73}$$

Therefore any convergent p/q of x satisfies $|x - p/q| < 1/q^2$ (compare with (i) $\Rightarrow$ (v) in Proposition 4). Moreover, if a_{n+1} is large, then the approximation p_n/q_n is sharp. Hence, large partial quotients yield good rational approximations by truncating the continued fraction expansion just before the given partial quotient.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°6, *May 5, 2010*

6.3 Pell's equation

Let D be a positive integer which is not the square of an integer. It follows that $\sqrt{D}$ is an irrational number. The Diophantine equation

$$x^2 - Dy^2 = \pm 1, \tag{74}$$

where the unknowns x and y are in $\mathbf{Z}$, is called *Pell's equation*.

An introduction to the subject has been given in the colloquium lecture on April 15. We refer to

http://seminarios.impa.br/cgi-bin/SEMINAR_palestra.cgi?id=4752

http://www.math.jussieu.fr/∼ miw/articles/pdf/PellFermatEn2010.pdf

and

http://www.math.jussieu.fr/∼ miw/articles/pdf/PellFermatEn2010VI.pdf

Here we supply complete proofs of the results introduced in that lecture.

6.3.1 Examples

The three first examples below are special cases of results initiated by O. Perron and related with real quadratic fields of Richaud-Degert type.

Example 1. Take $D = a^2b^2 + 2b$ where a and b are positive integers. A solution to

$$x^2 - (a^2b^2 + 2b)y^2 = 1$$

is $(x, y) = (a^2b + 1, a)$. As we shall see, this is related with the continued fraction expansion of $\sqrt{D}$ which is

$$\sqrt{a^2b^2 + 2b} = [ab, \overline{a, 2ab}]$$

since

$$t = \sqrt{a^2b^2 + 2b} \Longleftrightarrow t = ab + \cfrac{1}{a + \cfrac{1}{t + ab}}\cdot$$

This includes the examples $D = a^2 + 2$ (take $b = 1$) and $D = b^2 + 2b$ (take $a = 1$). For $a = 1$ and $b = c - 1$ his includes the example $D = c^2 - 1$.

Example 2. Take $D = a^2b^2 + b$ where a and b are positive integers. A solution to

$$x^2 - (a^2b^2 + b)y^2 = 1$$

is $(x, y) = (2a^2b + 1, 2a)$. The continued fraction expansion of $\sqrt{D}$ is

$$\sqrt{a^2b^2 + b} = [ab, \overline{2a, 2ab}]$$

since

$$t = \sqrt{a^2b^2 + b} \Longleftrightarrow t = ab + \frac{1}{2a + \dfrac{1}{t + ab}}\cdot$$

This includes the example $D = b^2 + b$ (take $a = 1$).

The case $b = 1$, $D = a^2 + 1$ is special: there is an integer solution to

$$x^2 - (a^2 + 1)y^2 = -1,$$

namely $(x, y) = (a, 1)$. The continued fraction expansion of $\sqrt{D}$ is

$$\sqrt{a^2 + 1} = [a, \overline{2a}]$$

since

$$t = \sqrt{a^2 + 1} \Longleftrightarrow t = a + \frac{1}{t + a}\cdot$$

Example 3. Let a and b be two positive integers such that $b^2 + 1$ divides $2ab + 1$. For instance $b = 2$ and $a \equiv 1 \pmod 5$. Write $2ab + 1 = k(b^2 + 1)$ and take $D = a^2 + k$. The continued fraction expansion of $\sqrt{D}$ is

$$[a, \overline{b, b, 2a}]$$

since $t = \sqrt{D}$ satisfies

$$t = a + \frac{1}{b + \dfrac{1}{b + \dfrac{1}{a + t}}} = [a, b, b, a + z].$$

A solution to $x^2 - Dy^2 = -1$ is $x = ab^2 + a + b$, $y = b^2 + 1$.

In the case $a = 1$ and $b = 2$ (so $k = 1$), the continued fraction has period length 1 only:

$$\sqrt{5} = [1, \overline{2}].$$

Example 4. Integers which are *Polygonal numbers* in two ways are given by the solutions to quadratic equations.

Triangular numbers are numbers of the form

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \quad \text{for } n \geq 1;$$

their sequence starts with

1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, 136, 153, 171, ...

`http://www.research.att.com/~njas/sequences/A000217`.

Square numbers are numbers of the form

$$1 + 3 + 5 + \cdots + (2n+1) = n^2 \quad \text{for } n \geq 1;$$

their sequence starts with

1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121, 144, 169, 196, 225, 256, 289, ...

`http://www.research.att.com/~njas/sequences/A000290`.

Pentagonal numbers are numbers of the form

$$1 + 4 + 7 + \cdots + (3n+1) = \frac{n(3n-1)}{2} \quad \text{for } n \geq 1;$$

their sequence starts with

1, 5, 12, 22, 35, 51, 70, 92, 117, 145, 176, 210, 247, 287, 330, 376, 425, ...

`http://www.research.att.com/~njas/sequences/A000326`.

Hexagonal numbers are numbers of the form

$$1 + 5 + 9 + \cdots + (4n+1) = n(2n-1) \quad \text{for } n \geq 1;$$

their sequence starts with

1, 6, 15, 28, 45, 66, 91, 120, 153, 190, 231, 276, 325, 378, 435, 496, 561, ...

`http://www.research.att.com/~njas/sequences/A000384`.

For instance, numbers which are at the same time triangular and squares are the numbers y^2 where (x, y) is a solution to Pell's equation with $D = 8$. Their list starts with

$$0, 1, 36, 1225, 41616, 1413721, 48024900, 1631432881, 55420693056, \ldots$$

See `http://www.research.att.com/~njas/sequences/A001110`.

Example 5. Integer rectangle triangles having sides of the right angle as consecutive integers a and $a+1$ have an hypothenuse c which satisfies $a^2+(a+1)^2=c^2$. The admissible values for the hypothenuse is the set of positive integer solutions y to Pell's equation $x^2-2y^2=-1$. The list of these hypothenuses starts with

$$1, 5, 29, 169, 985, 5741, 33461, 195025, 1136689, 6625109, 38613965,$$

See `http://www.research.att.com/~njas/sequences/A001653`.

6.3.2 Existence of integer solutions

Let D be a positive integer which is not a square. We show that Pell's equation (74) has a non–trivial solution $(x,y)\in\mathbf{Z}\times\mathbf{Z}$, that is a solution $\neq(\pm1,0)$.

Proposition 75. *Given a positive integer D which is not a square, there exists $(x,y)\in\mathbf{Z}^2$ with $x>0$ and $y>0$ such that $x^2-Dy^2=1$.*

Proof. The first step of the proof is to show that there exists a non–zero integer k such that the Diophantine equation $x^2-Dy^2=k$ has infinitely many solutions $(x,y)\in\mathbf{Z}\times\mathbf{Z}$. The main idea behind the proof, which will be made explicit in Lemmas 77, 78 and Corollary 79 below, is to relate the integer solutions of such a Diophantine equation with rational approximations x/y of $\sqrt{D}$.

Using the implication (i) $\Rightarrow$ (v) of the irrationality criterion 4 and the fact that $\sqrt{D}$ is irrational, we deduce that there are infinitely many $(x,y)\in\mathbf{Z}\times\mathbf{Z}$ with $y>0$ (and hence $x>0$) satisfying

$$\left|\sqrt{D}-\frac{x}{y}\right|<\frac{1}{y^2}.$$

For such a (x,y), we have $0<x<y\sqrt{D}+1<y(\sqrt{D}+1)$, hence

$$0<|x^2-Dy^2|=|x-y\sqrt{D}|\cdot|x+y\sqrt{D}|<2\sqrt{D}+1.$$

Since there are only finitely integers $k\neq0$ in the range

$$-(2\sqrt{D}+1)<k<2\sqrt{D}+1,$$

one at least of them is of the form x^2-Dy^2 for infinitely many (x,y).

The second step is to notice that, since the subset of $(x, y) \pmod k$ in $(\mathbf{Z}/k\mathbf{Z})^2$ is finite, there is an infinite subset $E \subset \mathbf{Z} \times \mathbf{Z}$ of these solutions to $x^2 - Dy^2 = k$ having the same $(x \pmod k, y \pmod k)$.

Let (u_1, v_1) and (u_2, v_2) be two distinct elements in E. Define $(x, y) \in \mathbf{Q}^2$ by

$$x + y\sqrt{D} = \frac{u_1 + v_1\sqrt{D}}{u_2 + v_2\sqrt{D}}.$$

From $u_2^2 - Dv_2^2 = k$, one deduces

$$x + y\sqrt{D} = \frac{1}{k}(u_1 + v_1\sqrt{D})(u_2 - v_2\sqrt{D}),$$

hence

$$x = \frac{u_1u_2 - Dv_1v_2}{k}, \quad y = \frac{-u_1v_2 + u_2v_1}{k}.$$

From $u_1 \equiv u_2 \pmod k$, $v_1 \equiv v_2 \pmod k$ and

$$u_1^2 - Dv_1^2 = k, \quad u_2^2 - Dv_2^2 = k,$$

we deduce

$$u_1u_2 - Dv_1v_2 \equiv u_1^2 - Dv_1^2 \equiv 0 \pmod k$$

and

$$-u_1v_2 + u_2v_1 \equiv -u_1v_1 + u_1v_1 \equiv 0 \pmod k,$$

hence x and y are in $\mathbf{Z}$. Further,

$$\begin{aligned} x^2 - Dy^2 &= (x + y\sqrt{D})(x - y\sqrt{D}) \\ &= \frac{(u_1 + v_1\sqrt{D})(u_1 - v_1\sqrt{D})}{(u_2 + v_2\sqrt{D})(u_2 - v_2\sqrt{D})} \\ &= \frac{u_1^2 - Dv_1^2}{u_2^2 - Dv_2^2} = 1. \end{aligned}$$

It remains to check that $y \neq 0$. If $y = 0$ then $x = \pm 1$, $u_1v_2 = u_2v_1$, $u_1u_2 - Dv_1v_2 = \pm 1$, and

$$ku_1 = \pm u_1(u_1u_2 - Dv_1v_2) = \pm u_2(u_1^2 - Dv_1^2) = \pm ku_2,$$

which implies $(u_1, u_2) = (v_1, v_2)$, a contradiction.

Finally, if $x < 0$ (resp. $y < 0$) we replace x by $-x$ (resp. y by $-y$).

□

Once we have a non–trivial integer solution (x, y) to Pell's equation, we have infinitely many of them, obtained by considering the powers of $x + y\sqrt{D}$.

6.3.3 All integer solutions

There is a natural order for the positive integer solutions to Pell's equation: we can order them by increasing values of x, or increasing values of y, or increasing values of $x+y\sqrt{D}$ - it is easily checked that the order is the same.

It follows that there is a minimal positive integer solution[8] (x_1, y_1), which is called *the fundamental solution to Pell's equation* $x^2 - Dy^2 = \pm 1$. In the same way, there is a fundamental solution to Pell's equations $x^2 - Dy^2 = 1$. Furthermore, when the equation $x^2 - Dy^2 = -1$ has an integer solution, then there is also a fundamental solution.

Proposition 76. *Denote by (x_1, y_1) the fundamental solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Then the set of all positive integer solutions to this equation is the sequence $(x_n, y_n)_{n\geq 1}$, where x_n and y_n are given by*

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n, \quad (n \in \mathbf{Z}, \quad n \geq 1).$$

In other terms, x_n and y_n are defined by the recurrence formulae

$$x_{n+1} = x_n x_1 + D y_n y_1 \qquad \text{and} \qquad y_{n+1} = x_1 y_n + x_n y_1, \qquad (n \geq 1).$$

More explicitly:
• *If $x_1^2 - Dy_1^2 = 1$, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = 1$, and there is no integer solution to Pell's equation $x^2 - Dy^2 = -1$.*
• *If $x_1^2 - Dy_1^2 = -1$, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = -1$, and the fundamental solution to Pell's equation $x^2 - Dy^2 = 1$ is (x_2, y_2). The set of positive integer solutions to Pell's equation $x^2 - Dy^2 = 1$ is $\{(x_n, y_n)\ ;\ n \geq 2 \ \ even\}$, while the set of positive integer solutions to Pell's equation $x^2 - Dy^2 = -1$ is $\{(x_n, y_n)\ ;\ n \geq 1 \ \ odd\}$. The set of all solutions $(x, y) \in \mathbf{Z} \times \mathbf{Z}$ to Pell's equation $x^2 - Dy^2 = \pm 1$ is the set $(\pm x_n, y_n)_{n\in\mathbf{Z}}$, where x_n and y_n are given by the same formula*

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n, \quad (n \in \mathbf{Z}).$$

The trivial solution $(1, 0)$ is (x_0, y_0), the solution $(-1, 0)$ is a torsion element of order 2 in the group of units of the ring $\mathbf{Z}[\sqrt{D}]$.

Proof. Let (x, y) be a positive integer solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Denote by $n \geq 0$ the largest integer such that

$$(x_1 + y_1\sqrt{D})^n \leq x + y\sqrt{D}.$$

[8] We use the letter x_1, which should not be confused with the first complete quotient in the section § 6.2.2 on continued fractions

Hence $x + y\sqrt{D} < (x_1 + y_1\sqrt{D})^{n+1}$. Define $(u, v) \in \mathbf{Z} \times \mathbf{Z}$ by

$$u + v\sqrt{D} = (x + y\sqrt{D})(x_1 - y_1\sqrt{D})^n.$$

From

$$u^2 - Dv^2 = \pm 1 \quad \text{and} \quad 1 \leq u + v\sqrt{D} < x_1 + y_1\sqrt{D},$$

we deduce $u = 1$ and $v = 0$, hence $x = x_n$, $y = y_n$.

□

6.3.4 On the group of units of $\mathbf{Z}[\sqrt{D}]$

Let D be a positive integer which is not a square. The ring $\mathbf{Z}[\sqrt{D}]$ is the subring of $\mathbf{R}$ generated by $\sqrt{D}$. The map $\sigma : z = x + y\sqrt{D} \longmapsto x - y\sqrt{D}$ is the *Galois automorphism* of this ring. The *norm* $N : \mathbf{Z}[\sqrt{D}] \longrightarrow \mathbf{Z}$ is defined by $N(z) = z\sigma(z)$. Hence

$$N(x + y\sqrt{D}) = x^2 - Dy^2.$$

The restriction of N to the group of unit $\mathbf{Z}[\sqrt{D}]^\times$ of the ring $\mathbf{Z}[\sqrt{D}]$ is a homomorphism from the multiplicative group $\mathbf{Z}[\sqrt{D}]^\times$ to the group of units $\mathbf{Z}^\times$ of $\mathbf{Z}$. Since $\mathbf{Z}^\times = \{\pm 1\}$, it follows that

$$\mathbf{Z}[\sqrt{D}]^\times = \{z \in \mathbf{Z}[\sqrt{D}] \ ; \ N(z) = \pm 1\},$$

hence $\mathbf{Z}[\sqrt{D}]^\times$ is nothing else than the set of $x + y\sqrt{D}$ when (x, y) runs over the set of integer solutions to Pell's equation $x^2 - Dy^2 = \pm 1$.

Proposition 75 means that $\mathbf{Z}[\sqrt{D}]^\times$ is not reduced to the torsion subgroup ± 1, while Proposition 76 gives the more precise information that this group $\mathbf{Z}[\sqrt{D}]^\times$ is a (multiplicative) abelian group of rank 1: there exists a so–called *fundamental unit* $u \in \mathbf{Z}[\sqrt{D}]^\times$ such that

$$\mathbf{Z}[\sqrt{D}]^\times = \{\pm u^n \ ; \ n \in \mathbf{Z}\}.$$

The fundamental unit $u > 1$ is $x_1 + y_1\sqrt{D}$, where (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Pell's equation $x^2 - Dy^2 = \pm 1$ has integer solutions if and only if the fundamental unit has norm -1.

That the rank of $\mathbf{Z}[\sqrt{D}]^\times$ is at most 1 also follows from the fact that the image of the map

$$\begin{array}{ccc} \mathbf{Z}[\sqrt{D}]^\times & \longrightarrow & \mathbf{R}^2 \\ z & \longmapsto & \big(\log|z|, \ \log|z'|\big) \end{array}$$

is discrete in $\mathbf{R}^2$ and contained in the line $t_1 + t_2 = 0$ of $\mathbf{R}^2$. This proof is not really different from the proof we gave of Proposition 76: the proof that the discrete subgroups of $\mathbf{R}$ have rank ≤ 1 relies on Euclid's division.

6.3.5 Connection with rational approximation

Lemma 77. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = 1$.
(ii) $0 < \dfrac{x}{y} - \sqrt{D} < \dfrac{1}{2y^2\sqrt{D}}$.
(iii) $0 < \dfrac{x}{y} - \sqrt{D} < \dfrac{1}{y^2\sqrt{D}+1}$.

Proof. We have $\dfrac{1}{2y^2\sqrt{D}} < \dfrac{1}{y^2\sqrt{D}+1}$, hence (ii) implies (iii).
(i) implies $x^2 > Dy^2$, hence $x > y\sqrt{D}$, and consequently

$$0 < \frac{x}{y} - \sqrt{D} = \frac{1}{y(x + y\sqrt{D})} < \frac{1}{2y^2\sqrt{D}}.$$

(iii) implies

$$x < y\sqrt{D} + \frac{1}{y\sqrt{D}} < y\sqrt{D} + \frac{2}{y},$$

and

$$y(x + y\sqrt{D}) < 2y^2\sqrt{D} + 2,$$

hence

$$0 < x^2 - Dy^2 = y\left(\frac{x}{y} - \sqrt{D}\right)(x + y\sqrt{D}) < 2.$$

Since $x^2 - Dy^2$ is an integer, it is equal to 1. □

The next variant will also be useful.

Lemma 78. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = -1$.
(ii) $0 < \sqrt{D} - \dfrac{x}{y} < \dfrac{1}{2y^2\sqrt{D}-1}$.
(iii) $0 < \sqrt{D} - \dfrac{x}{y} < \dfrac{1}{y^2\sqrt{D}}$.

Proof. We have $\dfrac{1}{2y^2\sqrt{D}-1} < \dfrac{1}{y^2\sqrt{D}}$, hence (ii) implies (iii).
The condition (i) implies $y\sqrt{D} > x$. We use the trivial estimate

$$2\sqrt{D} > 1 + 1/y^2$$

and write

$$x^2 = Dy^2 - 1 > Dy^2 - 2\sqrt{D} + 1/y^2 = (y\sqrt{D} - 1/y)^2,$$

hence $xy > y^2\sqrt{D} - 1$. From (i) one deduces

$$\begin{aligned} 1 = Dy^2 - x^2 &= (y\sqrt{D} - x)(y\sqrt{D} + x) \\ &> \left(\sqrt{D} - \frac{x}{y}\right)(y^2\sqrt{D} + xy) \\ &> \left(\sqrt{D} - \frac{x}{y}\right)(2y^2\sqrt{D} - 1). \end{aligned}$$

(iii) implies $x < y\sqrt{D}$ and

$$y(y\sqrt{D} + x) < 2y^2\sqrt{D},$$

hence

$$0 < Dy^2 - x^2 = y\left(\sqrt{D} - \frac{x}{y}\right)(y\sqrt{D} + x) < 2.$$

Since $Dy^2 - x^2$ is an integer, it is 1. □

From these two lemmas one deduces:

Corollary 79. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. The following conditions are equivalent:*
(i) $x^2 - Dy^2 = \pm 1$.
(ii) $\left|\sqrt{D} - \frac{x}{y}\right| < \frac{1}{2y^2\sqrt{D} - 1}$.
(iii) $\left|\sqrt{D} - \frac{x}{y}\right| < \frac{1}{y^2\sqrt{D} + 1}$.

Proof. If $y > 1$ or $D > 3$ we have $2y^2\sqrt{D} - 1 > y^2\sqrt{D} + 1$, which means that (ii) implies trivially (iii), and we may apply Lemmas 77 and 78.

If $D = 2$ and $y = 1$, then each of the conditions (i), (ii) and (iii) is satisfied if and only if $x = 1$. This follows from

$$2 - \sqrt{2} > \frac{1}{2\sqrt{2} - 1} > \frac{1}{\sqrt{2} + 1} > \sqrt{2} - 1.$$

If $D = 3$ and $y = 1$, then each of the conditions (i), (ii) and (iii) is satisfied if and only if $x = 2$. This follows from

$$3 - \sqrt{3} > \sqrt{3} - 1 > \frac{1}{2\sqrt{3} - 1} > \frac{1}{\sqrt{3} + 1} > 2 - \sqrt{3}.$$

□

It is instructive to compare with Liouville's inequality (see § 5.2).

Lemma 80. *Let D be a positive integer which is not a square. Let x and y be positive rational integers. Then*

$$\left|\sqrt{D} - \frac{x}{y}\right| > \frac{1}{2y^2\sqrt{D}+1}.$$

Proof. If $x/y < \sqrt{D}$, then $x \le y\sqrt{D}$ and from

$$1 \le Dy^2 - x^2 = (y\sqrt{D}+x)(y\sqrt{D}-x) \le 2y\sqrt{D}(y\sqrt{D}-x),$$

one deduces

$$\sqrt{D} - \frac{x}{y} > \frac{1}{2y^2\sqrt{D}}.$$

We claim that if $x/y > \sqrt{D}$, then

$$\frac{x}{y} - \sqrt{D} > \frac{1}{2y^2\sqrt{D}+1}.$$

Indeed, this estimate is true if $x - y\sqrt{D} \ge 1/y$, so we may assume $x - y\sqrt{D} < 1/y$. Our claim then follows from

$$1 \le x^2 - Dy^2 = (x+y\sqrt{D})(x-y\sqrt{D}) \le (2y\sqrt{D}+1/y)(x-y\sqrt{D}).$$

□

This shows that a rational approximation x/y to $\sqrt{D}$, which is only slightly weaker than the limit given by Liouville's inequality, will produce a solution to Pell's equation $x^2 - Dy^2 = \pm 1$. The distance $|\sqrt{D} - x/y|$ cannot be smaller than $1/(2y^2\sqrt{D}+1)$, but it can be as small as $1/(2y^2\sqrt{D}-1)$, and for that it suffices that it is less than $1/(y^2\sqrt{D}+1)$

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°7, *May 10, 2010*

6.3.6 The main lemma

The theory which follows is well–known (a classical reference is the book [7] by O. Perron), but the point of view which we develop here is slightly different from most classical texts on the subject. We follow [2, 3, 9]. An important role in our presentation of the subject is the following result (Lemma 4.1 in [8]).

Lemma 81. *Let $\epsilon = \pm 1$ and let a, b, c, d be rational integers satisfying*

$$ad - bc = \epsilon$$

and $d \geq 1$. Then there is a unique finite sequence of rational integers $a_0, \ldots, a_s$ with $s \geq 1$ and $a_1, \ldots, a_{s-1}$ positive, such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix} \tag{82}$$

These integers are also characterized by

$$\frac{b}{d} = [a_0, a_1, \ldots, a_{s-1}], \quad \frac{c}{d} = [a_s, \ldots, a_1], \quad (-1)^{s+1} = \epsilon. \tag{83}$$

For instance, when $d = 1$, for b and c rational integers,

$$\begin{pmatrix} bc+1 & b \\ c & 1 \end{pmatrix} = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} bc-1 & b \\ c & 1 \end{pmatrix} = \begin{pmatrix} b-1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c-1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proof. We start with unicity. If $a_0, \ldots, a_s$ satisfy the conclusion of Lemma 81, then by using (82), we find $b/d = [a_0, a_1, \ldots, a_{s-1}]$. Taking the transpose, we also find $c/d = [a_s, \ldots, a_1]$. Next, taking the determinant, we

obtain $(-1)^{s+1} = \epsilon$. The last equality fixes the parity of s, and each of the rational numbers b/d, c/d has a unique continued fraction expansion whose length has a given parity (cf. Proposition 69). This proves the unicity of the factorisation when it exists.

For the existence, we consider the simple continued fraction expansion of c/d with length of parity given by the last condition in (83), say $c/d = [a_s, \dots, a_1]$. Let a_0 be a rational integer such that the distance between b/d and $[a_0, a_1, \dots, a_{s-1}]$ is $\leq 1/2$. Define a', b', c', d' by

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix}.$$

We have

$$d' > 0, \quad a'd' - b'c' = \epsilon, \quad \frac{c'}{d'} = [a_s, \dots, a_1] = \frac{c}{d}$$

and

$$\frac{b'}{d'} = [a_0, a_1, \dots, a_{s-1}], \quad \left| \frac{b'}{d'} - \frac{b}{d} \right| \leq \frac{1}{2}.$$

From $\gcd(c, d) = \gcd(c', d') = 1$, $c/d = c'/d'$ and $d > 0$, $d' > 0$ we deduce $c' = c$, $d' = d$. From the equality between the determinants we deduce $a' = a + kc$, $b' = b + kd$ for some $k \in \mathbf{Z}$, and from

$$\frac{b'}{d'} - \frac{b}{d} = k$$

we conclude $k = 0$, $(a', b', c', d') = (a, b, c, d)$. Hence (82) follows.

□

Corollary 84. *Assume the hypotheses of Lemma 81 are satisfied.*
(a) If $c > d$, then $a_s \geq 1$ and

$$\frac{a}{c} = [a_0, a_1, \dots, a_s].$$

(b) If $b > d$, then $a_0 \geq 1$ and

$$\frac{a}{b} = [a_s, \dots, a_1, a_0].$$

The following examples show that the hypotheses of the corollary are not superfluous:

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} b-1 & b \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} b-1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} c-1 & 1 \\ c & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c-1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Proof of Corollary 84. The assumption $c > d$ implies $a_s > 0$. This proves part (a), and part (b) follows by transposition (or repeating the proof). □

Another consequence of Lemma 81 is the following classical result (Satz 13 p. 47 of [7]).

Corollary 85. *Let a, b, c, d be rational integers with $ad - bc = \pm 1$ and $c > d > 0$. Let x and y be two irrational numbers satisfying $y > 1$ and*

$$x = \frac{ay + b}{cy + d}.$$

Let $x = [a_0, a_1, \ldots]$ be the simple continued fraction expansion of x. Then there exists $s \geq 1$ such that

$$a = p_s, \quad b = p_{s-1}, \quad c = q_s, \quad r = q_{s-1}, \quad y = x_{s+1}.$$

Proof. Using lemma 81, we write

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a'_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a'_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a'_s & 1 \\ 1 & 0 \end{pmatrix}$$

with $a'_1, \ldots, a'_{s-1}$ positive and

$$\frac{b}{d} = [a'_0, a'_1, \ldots, a'_{s-1}], \quad \frac{c}{d} = [a'_s, \ldots, a'_1].$$

From $c > d$ and corollary 84, we deduce $a'_s > 0$ and

$$\frac{a}{c} = [a'_0, a'_1, \ldots, a'_s] = \frac{p'_s}{q'_s}, \quad x = \frac{p'_s y + p'_{s-1}}{q'_s y + q'_{s-1}} = [a'_0, a'_1, \ldots, a'_s, y].$$

Since $y > 1$, it follows that $a'_i = a_i$, $p'_i = p_i$, $q'_i = q_i$ for $0 \leq i \leq s$ and $y = x_{s+1}$.

□

6.3.7 Simple Continued fraction of $\sqrt{D}$

An infinite sequence $(a_n)_{n\geq 1}$ is *periodic* if there exists a positive integer s such that

$$a_{n+s} = a_n \quad \text{for all } n \geq 1. \tag{86}$$

In this case, the finite sequence $(a_1, \ldots, a_s)$ is called a *period* of the original sequence. For the sake of notation, we write

$$(a_1, a_2, \ldots) = (\overline{a_1, \ldots, a_s}).$$

If s_0 is the smallest positive integer satisfying (86), then the set of s satisfying (86) is the set of positive multiples of s_0. In this case $(a_1, \ldots, a_{s_0})$ is called *the fundamental period* of the original sequence.

Theorem 87. *Let D be a positive integer which is not a square. Write the simple continued fraction of $\sqrt{D}$ as $[a_0, a_1, \ldots]$ with $a_0 = \lfloor \sqrt{D} \rfloor$.*
(a) The sequence $(a_1, a_2, \ldots)$ is periodic.
(b) Let (x, y) be a positive integer solution to Pell's equation $x^2 - Dy^2 = \pm 1$. Then there exists $s \geq 1$ such that $x/y = [a_0, \ldots, a_{s-1}]$ and

$$(a_1, a_2, \ldots, a_{s-1}, 2a_0)$$

is a period of the sequence $(a_1, a_2, \ldots)$. Further, $a_{s-i} = a_i$ for $1 \leq i \leq s-1$ [9].
(c) Let $(a_1, a_2, \ldots, a_{s-1}, 2a_0)$ be a period of the sequence $(a_1, a_2, \ldots)$. Set $x/y = [a_0, \ldots, a_{s-1}]$. Then $x^2 - Dy^2 = (-1)^s$.
(d) Let s_0 be the length of the fundamental period. Then for $i \geq 0$ not multiple of s_0, we have $a_i \leq a_0$.

If $(a_1, a_2, \ldots, a_{s-1}, 2a_0)$ is a period of the sequence $(a_1, a_2, \ldots)$, then

$$\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}] = [a_0, a_1, \ldots, a_{s-1}, a_0 + \sqrt{D}].$$

[9] Note (2016). As kindly pointed out to me by Yoishi Motohashi, the fact that the word $a_1, \ldots, a_{s-1}$ is a palindrom is proved in 'Essai sur la théorie des nombres' by Legendre (1798).
In his first paper published at the age of 17 by Evariste Galois, it is proved that if the expansion of a quadratic irrational α is purely periodic, then the same is true for the conjugate α' of α, and the continued fraction of α' is obtained by reversing the order of the continued fraction of α. Besides, this continued fraction is a palindrom if and only if $\alpha\alpha' = -1$.
É. Galois, *Démonstration d'un théorème sur les fractions continues périodiques.*
Annales de Mathématiques Pures et Appliquées, **19** (1828-1829), p. 294-301.
`http://archive.numdam.org/article/AMPA_1828-1829__19__294_0.pdf`
For more information on these contributions by Galois, see
`https://www.bibnum.education.fr/mathematiques/algebre/demonstration-d-un-theoreme-sur-les-fractions-continues-periodiques`

Consider the fundamental period $(a_1, a_2, \ldots, a_{s_0-1}, a_{s_0})$ of the sequence $(a_1, a_2, \ldots)$. By part (b) of Theorem 87 we have $a_{s_0} = 2a_0$, and by part (d), it follows that s_0 is the smallest index i such that $a_i > a_0$.

From (b) and (c) in Theorem 87, it follows that the fundamental solution (x_1, y_1) to Pell's equation $x^2 - Dy^2 = \pm 1$ is given by $x_1/y_1 = [a_0, \ldots, a_{s_0-1}]$, and that $x_1^2 - Dy_1^2 = (-1)^{s_0}$. Therefore, if s_0 is even, then there is no solution to the Pell's equation $x^2 - Dy^2 = -1$. If s_0 is odd, then (x_1, y_1) is the fundamental solution to Pell's equation $x^2 - Dy^2 = -1$, while the fundamental solution (x_2, y_2) to Pell's equation $x^2 - Dy^2 = 1$ is given by $x_2/y_2 = [a_0, \ldots, a_{2s-1}]$.

It follows also from Theorem 87 that the $(ns_0 - 1)$-th convergent

$$x_n/y_n = [a_0, \ldots, a_{ns_0-1}]$$

satisfies

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n. \tag{88}$$

We shall check this relation directly (Lemma 92).

Proof. Start with a positive solution (x, y) to Pell's equation $x^2 - Dy^2 = \pm 1$, which exists according to Proposition 75. Since $Dy \geq x$ and $x > y$, we may use lemma 81 and corollary 84 with

$$a = Dy, \quad b = c = x, \quad d = y$$

and write

$$\begin{pmatrix} Dy & x \\ x & y \end{pmatrix} = \begin{pmatrix} a'_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a'_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a'_s & 1 \\ 1 & 0 \end{pmatrix} \tag{89}$$

with positive integers $a'_0, \ldots, a'_s$ and with $a'_0 = \lfloor \sqrt{D} \rfloor$. Then the continued fraction expansion of Dy/x is $[a'_0, \ldots, a'_s]$ and the continued fraction expansion of x/y is $[a'_0, \ldots, a'_{s-1}]$.

Since the matrix on the left hand side of (89) is symmetric, the word $a'_0, \ldots, a'_s$ is a palindrome. In particular $a'_s = a'_0$.

Consider the periodic continued fraction

$$\delta = [a'_0, \overline{a'_1, \ldots, a'_{s-1}, 2a'_0}].$$

This number δ satisfies

$$\delta = [a'_0, a'_1, \ldots, a'_{s-1}, a'_0 + \delta].$$

Using the inverse of the matrix

$$\begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \quad \text{which is} \quad \begin{pmatrix} 0 & 1 \\ 1 & -a_0' \end{pmatrix},$$

we write

$$\begin{pmatrix} a_0' + \delta & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix}$$

Hence the product of matrices associated with the continued fraction of δ

$$\begin{pmatrix} a_0' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1' & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s-1}' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0' + \delta & 1 \\ 1 & 0 \end{pmatrix}$$

is

$$\begin{pmatrix} Dy & x \\ x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta & 1 \end{pmatrix} = \begin{pmatrix} Dy + \delta x & x \\ x + \delta y & y \end{pmatrix}.$$

It follows that

$$\delta = \frac{Dy + \delta x}{x + \delta y},$$

hence $\delta^2 = D$. As a consequence, $a_i' = a_i$ for $0 \leq i \leq s-1$ while $a_s' = a_0$, $a_s = 2a_0$.

This proves that if (x, y) is a non–trivial solution to Pell's equation $x^2 - Dy^2 = \pm 1$, then the continued fraction expansion of $\sqrt{D}$ is of the form

$$\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}] \tag{90}$$

with $a_1, \ldots, a_{s-1}$ a palindrome, and x/y is given by the convergent

$$x/y = [a_0, a_1, \ldots, a_{s-1}]. \tag{91}$$

Consider a convergent $p_n/q_n = [a_0, a_1, \ldots, a_n]$. If $a_{n+1} = 2a_0$, then (73) with $x = \sqrt{D}$ implies the upper bound

$$\left| \sqrt{D} - \frac{p_n}{q_n} \right| \leq \frac{1}{2a_0 q_n^2},$$

and it follows from Corollary 79 that (p_n, q_n) is a solution to Pell's equation $p_n^2 - Dq_n^2 = \pm 1$. This already shows that $a_i < 2a_0$ when $i+1$ is not the length of a period. We refine this estimate to $a_i \leq a_0$.

Assume $a_{n+1} \geq a_0 + 1$. Since the sequence $(a_m)_{m \geq 1}$ is periodic of period length s_0, for any m congruent to n modulo s_0, we have $a_{m+1} > a_0$. For these m we have

$$\left| \sqrt{D} - \frac{p_m}{q_m} \right| \leq \frac{1}{(a_0 + 1) q_m^2}.$$

For sufficiently large m congruent to n modulo s we have

$$(a_0+1)q_m^2 > q_m^2\sqrt{D}+1.$$

Corollary 79 implies that (p_m, q_m) is a solution to Pell's equation $p_m^2 - Dq_m^2 = \pm 1$. Finally, Corollary 84 implies that $m+1$ is a multiple of s_0, hence $n+1$ also.

□

6.3.8 Connection between the two formulae for the n-th positive solution to Pell's equation

Lemma 92. *Let D be a positive integer which is not a square. Consider the simple continued fraction expansion $\sqrt{D} = [a_0, \overline{a_1, \ldots, a_{s_0-1}, 2a_0}]$ where s_0 is the length of the fundamental period. Then the fundamental solution (x_1, y_1) to Pell's equation $x^2 - Dy^2 = \pm 1$ is given by the continued fraction expansion $x_1/y_1 = [a_0, a_1, \ldots, a_{s_0-1}]$. Let $n \geq 1$ be a positive integer. Define (x_n, y_n) by $x_n/y_n = [a_0, a_1, \ldots, a_{ns_0-1}]$. Then $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$.*

This result is a consequence of the two formulae we gave for the n-th solution (x_n, y_n) to Pell's equation $x^2 - Dy^2 = \pm 1$. We check this result directly.

Proof. From Lemma 81 and relation (89), one deduces

$$\begin{pmatrix} Dy_n & x_n \\ x_n & y_n \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{ns_0-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Since

$$\begin{pmatrix} Dy_n & x_n \\ x_n & y_n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -a_0 \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix},$$

we obtain

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{ns_0-1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix}. \qquad (93)$$

Notice that the determinant is $(-1)^{ns_0} = x_n^2 - Dy_n^2$. Formula (93) for $n+1$ and the periodicity of the sequence $(a_1, \ldots, a_n, \ldots)$ with $a_{s_0} = 2a_0$ give :

$$\begin{pmatrix} x_{n+1} & Dy_{n+1} - a_0x_{n+1} \\ y_{n+1} & x_{n+1} - a_0y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0x_n \\ y_n & x_n - a_0y_n \end{pmatrix} \begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s_0-1} & 1 \\ 1 & 0 \end{pmatrix}.$$

Take first $n = 1$ in (93) and multiply on the left by

$$\begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -a_0 \end{pmatrix} = \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix}.$$

Since

$$\begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 & Dy_1 - a_0 x_1 \\ y_1 & x_1 - a_0 y_1 \end{pmatrix} = \begin{pmatrix} x_1 + a_0 y_1 & (D - a_0^2) y_1 \\ y_1 & x_1 - a_0 y_1 \end{pmatrix}.$$

we deduce

$$\begin{pmatrix} 2a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s_0-1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_1 + a_0 y_1 & (D - a_0^2) y_1 \\ y_1 & x_1 - a_0 y_1 \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} x_{n+1} & Dy_{n+1} - a_0 x_{n+1} \\ y_{n+1} & x_{n+1} - a_0 y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n & Dy_n - a_0 x_n \\ y_n & x_n - a_0 y_n \end{pmatrix} \begin{pmatrix} x_1 + a_0 y_1 & (D - a_0^2) y_1 \\ y_1 & x_1 - a_0 y_1 \end{pmatrix}.$$

The first column gives

$$x_{n+1} = x_n x_1 + D y_n y_1 \qquad \text{and} \qquad y_{n+1} = x_1 y_n + x_n y_1,$$

which was to be proved.

□

6.3.9 Records

For large D, Pell's equation may obviously have small integer solutions. Examples are

For $D = m^2 - 1$ with $m \geq 2$ the numbers $x = m$, $y = 1$ satisfy $x^2 - Dy^2 = 1$,

for $D = m^2 + 1$ with $m \geq 1$ the numbers $x = m$, $y = 1$ satisfy $x^2 - Dy^2 = -1$,

for $D = m^2 \pm m$ with $m \geq 2$ the numbers $x = 2m \pm 1$ satisfy $y = 2$, $x^2 - Dy^2 = 1$,

for $D = t^2 m^2 + 2m$ with $m \geq 1$ and $t \geq 1$ the numbers $x = t^2 m + 1$, $y = t$ satisfy $x^2 - Dy^2 = 1$.

On the other hand, relatively small values of D may lead to large fundamental solutions. Tables are available on the internet[10].

[10] For instance:
Tomás Oliveira e Silva: Record-Holder Solutions of Pell's Equation
`http://www.ieeta.pt/~tos/pell.html`.

For D a positive integer which is not a square, denote by $S(D)$ the base 10 logarithm of x_1, when (x_1, y_1) is the fundamental solution to $x^2 - Dy^2 = 1$. The integral part of $S(D)$ is the number of digits of the fundamental solution x_1. For instance, when $D = 61$, the fundamental solution (x_1, y_1) is

$$x_1 = 1\,766\,319\,049, \quad y_1 = 226\,153\,980$$

and $S(61) = \log_{10} x_1 = 9.247\,069\ldots$

An integer D is a *record holder* for S if $S(D') < S(D)$ for all $D' < D$.

Here are the record holders up to 1021:

D	2	5	10	13	29	46	53	61	109
$S(D)$	0.477	0.954	1.278	2.812	3.991	4.386	4.821	9.247	14.198

D	181	277	397	409	421	541	661	1021
$S(D)$	18.392	20.201	20.923	22.398	33.588	36.569	37.215	47.298

Some further records with number of digits successive powers of 10:

D	3061	169789	12765349	1021948981	85489307341
$S(D)$	104.051	1001.282	10191.729	100681.340	1003270.151

6.3.10 A criterion for the existence of a solution to the negative Pell equation

Here is a recent result on the existence of a solution to Pell's equation $x^2 - Dy^2 = -1$

Proposition 94 (R.A. Mollin, A. Srinivasan[11])**.** *Let d be a positive integer which is not a square. Let (x_0, y_0) be the fundamental solution to Pell's equation $x^2 - dy^2 = 1$. Then the equation $x^2 - dy^2 = -1$ has a solution if and only if $x_0 \equiv -1 \pmod{2d}$.*

Proof. If $a^2 - db^2 = -1$ is the fundamental solution to $x^2 - dy^2 = -1$, then $x_0 + y_0\sqrt{d} = (a + b\sqrt{d})^2$, hence

$$x_0 = a^2 + db^2 = 2db^2 - 1 \equiv -1 \pmod{2d}.$$

Conversely, if $x_0 = 2dk - 1$, then $x_0^2 = 4d^2k^2 - 4dk + 1 = dy_0^2 + 1$, hence $4dk^2 - 4k = y_0^2$. Therefore y_0 is even, $y_0 = 2z$, and $k(dk - 1) = z^2$. Since k and $dk - 1$ are relatively prime, both are squares, $k = b^2$ and $dk - 1 = a^2$, which gives $a^2 - db^2 = -1$. □

[11] Pell equation: non-principal Lagrange criteria and central norms; Canadian Math. Bull., to appear

6.3.11 Arithmetic varieties

Let D be a positive integer which is not a square. Define $\mathcal{G} = \{(x, y) \in \mathbf{R}^2 \; ; \; x^2 - Dy^2 = 1\}$.

The map

$$\begin{array}{ccc} \mathcal{G} & \longrightarrow & \mathbf{R}^\times \\ (x, y) & \longmapsto & t = x + y\sqrt{D} \end{array}$$

is bijective: the inverse of that map is obtained by writing $u = 1/t$, $2x = t + u$, $2y\sqrt{D} = t - u$, so that $t = x + y\sqrt{D}$ and $u = x - y\sqrt{D}$. By transfer of structure, this endows $\mathcal{G}$ with a multiplicative group structure, which is isomorphic to $\mathbf{R}^\times$, for which

$$\begin{array}{ccc} \mathcal{G} & \longrightarrow & \mathrm{GL}_2(\mathbf{R}) \\ (x, y) & \longmapsto & \begin{pmatrix} x & Dy \\ y & x \end{pmatrix}. \end{array}$$

is an injective group homomorphism. Let $G(\mathbf{R})$ be its image, which is therefore isomorphic to $\mathbf{R}^\times$.

A matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ respects the quadratic form $x^2 - Dy^2$ if and only if

$$(ax + by)^2 - D(cx + dy)^2 = x^2 - Dy^2,$$

which can be written

$$a^2 - Dc^2 = 1, \quad b^2 - Dd^2 = D, \quad ab = cdD.$$

Hence the group of matrices of determinant 1 with coefficients in $\mathbf{Z}$ which respect the quadratic form $x^2 - Dy^2$ is the group

$$G(\mathbf{Z}) = \left\{ \begin{pmatrix} a & Dc \\ c & a \end{pmatrix} \in \mathrm{GL}_2(\mathbf{Z}) \right\}.$$

According to the work of Siegel, Harish–Chandra, Borel and Godement, the quotient of $G(\mathbf{R})$ by $G(\mathbf{Z})$ is compact. Hence $G(\mathbf{Z})$ is infinite (of rank 1 over $\mathbf{Z}$), which means that there are infinitely many solutions to the equation $a^2 - Dc^2 = 1$.

This is not a new proof of Proposition 75, but an interpretation and a generalization. Such results are valid for *arithmetic varieties*[12].

[12] See for instance Nicolas Bergeron, "Sur la forme de certains espaces provenant de constructions arithmétiques", Images des Mathématiques, (2004). `http://people.math.jussieu.fr/~bergeron/Recherche_files/Images.pdf`.

References

[1] E. J. Barbeau, *Pell's equation*, Problem Books in Mathematics, Springer-Verlag, New York, 2003.

[2] E. Bombieri, *Continued fractions and the Markoff tree*, Expo. Math., 25 (2007), pp. 187–213.

[3] E. Bombieri and A. J. van der Poorten, *Continued fractions of algebraic numbers*, in Computational algebra and number theory (Sydney, 1992), vol. 325 of Math. Appl., Kluwer Acad. Publ., Dordrecht, 1995, pp. 137–152.

[4] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, Oxford, sixth ed., 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[5] M. J. Jacobson, Jr. and H. C. Williams, *Solving the Pell equation*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2009.

[6] H. W. Lenstra, Jr., *Solving the Pell equation*, Notices Amer. Math. Soc., 49 (2002), pp. 182–192.

[7] O. Perron, *Die Lehre von den Kettenbrüchen. Dritte, verbesserte und erweiterte Aufl. Bd. II. Analytisch-funktionentheoretische Kettenbrüche*, B. G. Teubner Verlagsgesellschaft, Stuttgart, 1957.

[8] D. Roy, *On the continued fraction expansion of a class of numbers*, in Diophantine approximation, vol. 16 of Dev. Math., SpringerWien-NewYork, Vienna, 2008, pp. 347–361.
`http://arxiv.org/abs/math/0409233`.

[9] A. J. van der Poorten, *An introduction to continued fractions*, in Diophantine analysis (Kensington, 1985), vol. 109 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1986, pp. 99–138.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°8, *May 12, 2010*

Addition to Lemma 81.
In [1], § 4, there is a variant of the matrix formula (64) for the simple continued fraction of a real number.

Given integers $a_0, a_1, \ldots$ with $a_i > 0$ for $i \geq 1$ and writing, for $n \geq 0$, as usual, $p_n/q_n = [a_0, a_1, \ldots, a_n]$, one checks, by induction on n, the two formulae

$$\left.\begin{array}{ll} \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_n \\ 0 & 1 \end{pmatrix} & = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \quad \text{if } n \text{ is even} \\ \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ a_n & 1 \end{pmatrix} & = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \quad \text{if } n \text{ is odd} \end{array}\right\} \tag{95}$$

Define two matrices U (up) and L (low) in $\mathrm{GL}_2(\mathbf{R})$ of determinant $+1$ by

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad L = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

For p and q in $\mathbf{Z}$, we have

$$U^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad L^q = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix},$$

so that these formulae (95) are

$$U^{a_0} L^{a_1} \cdots U^{a_n} = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} \quad \text{if } n \text{ is even}$$

and

$$U^{a_0} L^{a_1} \cdots L^{a_n} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \quad \text{if } n \text{ is odd.}$$

The connexion with Euclid's algorithm is

$$U^{-p} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a - pc & b - pd \\ c & d \end{pmatrix} \quad \text{and} \quad L^{-q} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c - qa & d - qb \end{pmatrix}.$$

The corresponding variant of Lemma 81 is also given in [1], § 4: *If a, b, c, d are rational integers satisfying $b > a > 0$, $d > c \geq 0$ and $ad - bc = 1$, then there exist rational integers $a_0, \ldots, a_n$ with n even and $a_1, \ldots, a_n$ positive, such that*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & a_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_n \\ 0 & 1 \end{pmatrix}$$

These integers are uniquely determined by $b/d = [a_0, \ldots, a_n]$ with n even.

6.3.12 Periodic continued fractions

An infinite sequence $(a_n)_{n \geq 0}$ is said to be *ultimately periodic* if there exists $n_0 \geq 0$ and $s \geq 1$ such that

$$a_{n+s} = a_n \quad \text{for all } n \geq n_0. \tag{96}$$

The set of s satisfying this property (6.3.12) is the set of positive multiples of an integer s_0, and $(a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+s_0-1})$ is called *the fundamental period*.

A continued fraction with a sequence of partial quotients satisfying (96) will be written

$$[a_0, a_1, \ldots, a_{n_0-1}, \overline{a_{n_0}, \ldots, a_{n_0+s-1}}].$$

Example. For D a positive integer which is not a square, setting $a_0 = \lfloor \sqrt{D} \rfloor$, we have by Theorem 87

$$a_0 + \sqrt{D} = [\overline{2a_0, a_1, \ldots, a_{s-1}}] \quad \text{and} \quad \frac{1}{\sqrt{D} - a_0} = [\overline{a_1, \ldots, a_{s-1}, 2a_0}].$$

Lemma 97 (Euler 1737)**.** *If an infinite continued fraction*

$$x = [a_0, a_1, \ldots, a_n, \ldots]$$

is ultimately periodic, then x is a quadratic irrational number.

Proof. Since the continued fraction of x is infinite, x is irrational. Assume first that the continued fraction is periodic, namely that (96) holds with $n_0 = 0$:

$$x = [\overline{a_0, \ldots, a_{s-1}}].$$

This can be written

$$x = [a_0, \ldots, a_{s-1}, x].$$

Hence

$$x = \frac{p_{s-1}x + p_{s-2}}{q_{s-1}x + q_{s-2}}.$$

It follows that

$$q_{s-1}X^2 + (q_{s-2} - p_{s-1})X - p_{s-2}$$

is a non–zero quadratic polynomial with integer coefficients having x as a root. Since x is irrational, this polynomial is irreducible and x is quadratic.

In the general case where (96) holds with $n_0 > 0$, we write

$$x = [a_0, a_1, \ldots, a_{n_0-1}, \overline{a_{n_0}, \ldots, a_{n_0+s-1}}] = [a_0, a_1, \ldots, a_{n_0-1}, y],$$

where $y = [\overline{a_{n_0}, \ldots, a_{n_0+s-1}}]$ is a periodic continued fraction, hence is quadratic. But

$$x = \frac{p_{n_0-1}y + p_{n_0-2}}{q_{n_0-1}y + q_{n_0-2}},$$

hence $x \in \mathbf{Q}(y)$ is also quadratic irrational.

□

Lemma 98 (Lagrange, 1770)**.** *If x is a quadratic irrational number, then its continued fraction*

$$x = [a_0, a_1, \ldots, a_n, \ldots]$$

is ultimately periodic.

Proof. For $n \geq 0$, define $d_n = q_n x - p_n$. According to Corollary 72, we have $|d_n| < 1/q_{n+1}$.

Let $AX^2 + BX + C$ with $A > 0$ be an irreducible quadratic polynomial having x as a root. For each $n \geq 2$, we deduce from (70) that the convergent x_n is a root of a quadratic polynomial $A_nX^2 + B_nX + C_n$, with

$$\begin{aligned}
A_n &= Ap_{n-1}^2 + Bp_{n-1}q_{n-1} + Cq_{n-1}^2, \\
B_n &= 2Ap_{n-1}p_{n-2} + B(p_{n-1}q_{n-2} + p_{n-2}q_{n-1}) + 2Cq_{n-1}q_{n-2}, \\
C_n &= A_{n-1}.
\end{aligned}$$

Using $Ax^2 + Bx + C = 0$, we deduce

$$\begin{aligned}
A_n &= (2Ax + B)d_{n-1}q_{n-1} + Ad_{n-1}^2, \\
B_n &= (2Ax + B)(d_{n-1}q_{n-2} + d_{n-2}q_{n-1}) + 2Ad_{n-1}d_{n-2}.
\end{aligned}$$

There are similar formulae expressing A, B, C as homogeneous linear combinations of A_n, B_n, C_n, and since $(A, B, C) \neq (0, 0, 0)$, it follows that $(A_n, B_n, C_n) \neq (0, 0, 0)$. Since x_n is irrational, one deduces $A_n \neq 0$.

From the inequalities

$$q_{n-1}|d_{n-2}| < 1, \quad q_{n-2}|d_{n-1}| < 1, \quad q_{n-1} < q_n, \quad |d_{n-1}d_{n-2}| < 1,$$

one deduces

$$\max\{|A_n|,\ |B_n|/2, |C_n|\} < A + |2Ax + B|.$$

This shows that $|A_n|$, $|B_n|$ and $|C_n|$ are bounded independently of n. Therefore there exists $n_0 \geq 0$ and $s > 0$ such that $x_{n_0} = x_{n_0+s}$. From this we deduce that the continued fraction of x_{n_0} is purely periodic, hence the continued fraction of x is ultimately periodic. □

A *reduced quadratic irrational number* is an irrational number $x > 1$ which is a root of a degree 2 polynomial $ax^2 + bx + c$ with rational integer coefficients, such that the other root x' of this polynomial, which is the *Galois conjugate of* x, satisfies $-1 < x' < 0$. If x is reduced, then so is $-1/x'$.

Lemma 99. *A continued fraction*

$$x = [a_0, a_1, \ldots, a_n \ldots]$$

is purely periodic if and only if x is a reduced quadratic irrational number. In this case, if $x = [\overline{a_0, a_1, \ldots, a_{s-1}}]$ and if x' is the Galois conjugate of x, then

$$-1/x' = [\overline{a_{s-1}, \ldots, a_1, a_0}]$$

Proof. Assume first that the continued fraction of x is purely periodic:

$$x = [\overline{a_0, a_1, \ldots, a_{s-1}}].$$

From $a_s = a_0$ we deduce $a_0 > 0$, hence $x > 1$. From $x = [a_0, a_1, \ldots, a_{s-1}, x]$ and the unicity of the continued fraction expansion, we deduce

$$x = \frac{p_{s-1}x + p_{s-2}}{q_{s-1}x + q_{s-2}} \quad \text{and} \quad x = x_s.$$

Therefore x is a root of the quadratic polynomial

$$P_s(X) = q_{s-1}X^2 + (q_{s-2} - p_{s-1})X - p_{s-2}.$$

This polynomial P_s has a positive root, namely $x > 1$, and a negative root x', with the product $xx' = -p_{s-2}/q_{s-1}$. We transpose the relation

$$\begin{pmatrix} p_{s-1} & p_{s-2} \\ q_{s-1} & q_{s-2} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{s-1} & 1 \\ 1 & 0 \end{pmatrix}$$

and obtain

$$\begin{pmatrix} p_{s-1} & q_{s-1} \\ p_{s-2} & q_{s-2} \end{pmatrix} = \begin{pmatrix} a_{s-1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Define

$$y = [\overline{a_{s-1}, \ldots, a_1, a_0}],$$

so that $y > 1$,

$$y = [a_{s-1}, \ldots, a_1, a_0, y] = \frac{p_{s-1}y + q_{s-1}}{p_{s-2}y + q_{s-2}}$$

and y is the positive root of the polynomial

$$Q_s(X) = p_{s-2}X^2 + (q_{s-2} - p_{s-1})X - q_{s-1}.$$

The polynomials P_s and Q_s are related by $Q_s(X) = -X^2 P_s(-1/X)$. Hence $y = -1/x'$.

For the converse, assume $x > 1$ and $-1 < x' < 0$. Let $(x_n)_{n \geq 1}$ be the sequence of complete quotients of x. For $n \geq 1$, define x'_n as the Galois conjugate of x_n. One deduces by induction that $x'_n = a_n + 1/x'_{n+1}$, that $-1 < x'_n < 0$ (hence x_n is reduced), and that a_n is the integral part of $-1/x'_{n+1}$.

If the continued fraction expansion of x were not purely periodic, we would have

$$x = [a_0, \ldots, a_{h-1}, \overline{a_h, \ldots, a_{h+s-1}}]$$

with $a_{h-1} \neq a_{h+s-1}$. By periodicity we have $x_h = [a_h, \ldots, a_{h+s-1}, x_h]$, hence $x_h = x_{h+s}$, $x'_h = x'_{h+s}$. From $x'_h = x'_{h+s}$, taking integral parts, we deduce $a_{h-1} = a_{h+s-1}$, a contradiction. □

Corollary 100. *If $r > 1$ is a rational number which is not a square, then the continued fraction expansion of $\sqrt{r}$ is of the form*

$$\sqrt{r} = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$$

with $a_1, \ldots, a_{s-1}$ a palindrome and $a_0 = [\sqrt{r}]$.
Conversely, if the continued fraction expansion of an irrational number $t > 1$ is of the form

$$t = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$$

with $a_1, \ldots, a_{s-1}$ a palindrome, then t^2 is a rational number.

Proof. If $t^2 = r$ is rational > 1, then for and $a_0 = [\sqrt{t}]$ the number $x = t + a_0$ is reduced. Since $t' + t = 0$, we have

$$-\frac{1}{x'} = \frac{1}{x - 2a_0}.$$

Hence

$$x = [\overline{2a_0, a_1, \ldots, a_{s-1}}], \quad -\frac{1}{x'} = [\overline{a_{s-1}, \ldots, a_1, 2a_0}]$$

and $a_1, \ldots, a_{s-1}$ a palindrome.

Conversely, if $t = [a_0, \overline{a_1, \ldots, a_{s-1}, 2a_0}]$ with $a_1, \ldots, a_{s-1}$ a palindrome, then $x = t + a_0$ is periodic, hence reduced, and its Galois conjugate x' satisfies

$$-\frac{1}{x'} = [\overline{a_1, \ldots, a_{s-1}, 2a_0}] = \frac{1}{x - 2a_0},$$

which means $t + t' = 0$, hence $t^2 \in \mathbf{Q}$. □

Lemma 101 (Serret, 1878)**.** *Let x and y be two irrational numbers with continued fractions*

$$x = [a_0, a_1, \ldots, a_n \ldots] \quad \text{and} \quad y = [b_0, b_1, \ldots, b_m \ldots]$$

respectively. Then the two following properties are equivalent.

(i) *There exists a matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *with rational integer coefficients and determinant* ± 1 *such that*

$$y = \frac{ax + b}{cx + d}.$$

(ii) *There exists $n_0 \geq 0$ and $m_0 \geq 0$ such that $a_{n_0+k} = b_{m_0+k}$ for all $k \geq 0$.*

Condition (i) means that x *and* y *are equivalent modulo the action of* $\mathrm{GL}_2(\mathbf{Z})$ *by homographies.*

Condition (ii) means that there exists integers n_0, m_0 and a real number $t > 1$ such that

$$x = [a_0, a_1, \ldots, a_{n_0-1}, t] \quad \text{and} \quad y = [b_0, b_1, \ldots, b_{m_0-1}, t].$$

Example.

$$\text{If } x = [a_0, a_1, x_2], \text{ then } \quad -x = \begin{cases} [-a_0 - 1, 1, a_1 - 1, x_2] & \text{if } a_1 \geq 2, \\ [-a_0 - 1, 1 + x_2] & \text{if } a_1 = 1. \end{cases} \tag{102}$$

Proof. We already know by (70) that if x_n is a complete quotient of x, then x and x_n are equivalent modulo $\mathrm{GL}_2(\mathbf{Z})$. Condition (ii) means that there is a partial quotient of x and a partial quotient of y which are equal. By transitivity of the $\mathrm{GL}_2(\mathbf{Z})$ equivalence, (ii) implies (i).

Conversely, assume (i):

$$y = \frac{ax+b}{cx+d}.$$

Let n be a sufficiently large number. From

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} u_n & u_{n-1} \\ v_n & v_{n-1} \end{pmatrix}$$

with

$$\begin{aligned} u_n &= ap_n + bq_n, \quad u_{n-1} = ap_{n-1} + bq_{n-1}, \\ v_n &= cp_n + dq_n, \quad v_{n-1} = cp_{n-1} + dq_{n-1}, \end{aligned}$$

we deduce

$$y = \frac{u_n x_{n+1} + u_{n-1}}{v_n x_{n+1} + v_{n-1}}.$$

We have $v_n = (cx+d)q_n + c\delta_n$ with $\delta_n = p_n - q_n x$. We have $q_n \to \infty$, $q_n \geq q_{n-1} + 1$ and $\delta_n \to 0$ as $n \to \infty$. Hence, for sufficiently large n, we have $v_n > v_{n-1} > 0$. From part 1 of Corollary 84, we deduce

$$\begin{pmatrix} u_n & u_{n-1} \\ v_n & v_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_s & 1 \\ 1 & 0 \end{pmatrix}$$

with $a_0, \ldots, a_s$ in $\mathbf{Z}$ and $a_1, \ldots, a_s$ positive. Hence

$$y = [a_0, a_1, \ldots, a_s, x_{n+1}].$$

□

A computational proof of (i) ⇒ (ii). Another proof is given by Bombieri [2] (Theorem A.1 p. 209). He uses the fact that $\mathrm{GL}_2(\mathbf{Z})$ is generated by the two matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The associated fractional linear transformations are K and J defined by

$$K(x) = x + 1 \quad \text{and} \quad J(x) = 1/x.$$

We have $J^2 = 1$ and

$$K([a_0, t]) = [a_0 + 1, t], \quad K^{-1}([a_0, t]) = [a_0 - 1, t].$$

Also $J([a_0,t]) = [0,a_0,t]$ if $a_0 > 0$ and $J([0,t]) = [t]$. According to (102), the continued fractions of x and $-x$ differ only by the first terms. This completes the proof. [13]

□

6.4 Diophantine approximation and simple continued fractions

Lemma 103 (Lagrange, 1770)**.** *The sequence* $(|q_n x - p_n|)_{n\geq 0}$ *is strictly decreasing: for* $n \geq 1$ *we have*

$$|q_n x - p_n| < |q_{n-1}x - p_{n-1}|.$$

Proof. We use Lemma 71 twice: on the one hand

$$|q_n x - p_n| = \frac{1}{x_{n+1}q_n + q_{n-1}} < \frac{1}{q_n + q_{n-1}}$$

because $x_{n+1} > 1$, on the other hand

$$|q_{n-1}x - p_{n-1}| = \frac{1}{x_n q_{n-1} + q_{n-2}} > \frac{1}{(a_n+1)q_{n-1} + q_{n-2}} = \frac{1}{q_n + q_{n-1}}$$

because $x_n < a_n + 1$. □

Corollary 104. *The sequence* $(|x - p_n/q_n|)_{n\geq 0}$ *is strictly decreasing: for* $n \geq 1$ *we have*

$$\left|x - \frac{p_n}{q_n}\right| < \left|x - \frac{p_{n-1}}{q_{n-1}}\right|.$$

Proof. For $n \geq 1$, since $q_{n-1} < q_n$, we have

$$\left|x - \frac{p_n}{q_n}\right| = \frac{1}{q_n}|q_n x - p_n| < \frac{1}{q_n}|q_{n-1}x - p_{n-1}| = \frac{q_{n-1}}{q_n}\left|x - \frac{p_{n-1}}{q_{n-1}}\right| < \left|x - \frac{p_{n-1}}{q_{n-1}}\right|.$$

□

Here is the *law of best approximation* of the simple continued fraction.

[13] Bombieri in [2] gives formulae for $J([a_0,t])$ when $a_0 \leq -1$. He distinguishes eight cases, namely four cases when $a_0 = -1$ ($a_1 > 2$, $a_1 = 2$, $a_1 = 1$ and $a_3 > 1$, $a_1 = a_3 = 1$), two cases when $a_0 = -2$ ($a_1 > 1$, $a_1 = 1$) and two cases when $a_0 \leq -3$ ($a_1 > 1$, $a_1 = 1$). Here, (102) enables us to simplify his proof by reducing to the case $a_0 \geq 0$.

Lemma 105. *Let $n \geq 0$ and $(p, q) \in \mathbf{Z} \times \mathbf{Z}$ with $q > 0$ satisfy*

$$|qx - p| < |q_n x - p_n|.$$

Then $q \geq q_{n+1}$.

Proof. The system of two linear equations in two unknowns u, v

$$\begin{cases} p_n u + p_{n+1} v &= p \\ q_n u + q_{n+1} v &= q \end{cases} \tag{106}$$

has determinant ± 1, hence there is a solution $(u, v) \in \mathbf{Z} \times \mathbf{Z}$.

Since $p/q \neq p_n/q_n$, we have $v \neq 0$.

If $u = 0$, then $v = q/q_{n+1} > 0$, hence $v \geq 1$ and $q \geq q_{n+1}$.

We now assume $uv \neq 0$.

Since q, q_n and q_{n+1} are > 0, it is not possible for u and v to be both negative. In case u and v are positive, the desired result follows from the second relation of (106). Hence one may suppose u and v of opposite signs. Since $q_n x - p_n$ and $q_{n+1} x - p_{n+1}$ also have opposite signs, the numbers $u(q_n x - p_n)$ and $v(q_{n+1} x - p_{n+1})$ have same sign, and therefore

$$|q_n x - p_n| \leq |u(q_n x - p_n)| + |v(q_{n+1} x - p_{n+1})| = |qx - p| < |q_n x - p_n|,$$

which is a contradiction.

□

A consequence of Lemma 105 is that the sequence of p_n/q_n produces the best rational approximations to x in the following sense: any rational number p/q with denominator $q < q_n$ has $|qx - p| > |q_n x - p_n|$. This is sometimes referred to as *best rational approximations of type* 0.

Corollary 107. *The sequence $(q_n)_{n \geq 0}$ of denominators of the convergents of a real irrational number x is the increasing sequence of positive integers for which*

$$\|q_n x\| < \|qx\| \quad \textit{for} \quad 1 \leq q < q_n.$$

As a consequence,

$$\|q_n x\| = \min_{1 \leq q \leq q_n} \|qx\|.$$

The theory of continued fractions is developed starting from Corollary 107 as a definition of the sequence $(q_n)_{n \geq 0}$ in Cassels's book [5].

Corollary 108. *Let $n \geq 0$ and $p/q \in \mathbf{Q}$ with $q > 0$ satisfy*

$$\left| x - \frac{p}{q} \right| < \left| x - \frac{p_n}{q_n} \right|.$$

Then $q > q_n$.

Proof. For $q \leq q_n$ we have

$$\left| x - \frac{p}{q} \right| = \frac{1}{q}|qx - p| > \frac{1}{q}|q_n x - p_n| \frac{q_n}{q} \left| x - \frac{p_n}{q_n} \right| \geq \left| x - \frac{p_n}{q_n} \right|.$$

□

Corollary 108 shows that the denominators q_n of the convergents are also among the *best rational approximations of type* 1 in the sense that

$$\left| x - \frac{p}{q} \right| > \left| x - \frac{p_n}{q_n} \right| \quad \text{for} \quad 1 \leq q < q_n,$$

but they do not produce the full list of them: to get the complete set, one needs to consider also some of the rational fractions of the form

$$\frac{p_{n-1} + ap_n}{q_{n-1} + aq_n}$$

with $0 \leq a \leq a_{n+1}$ (*semi–convergents*) – see for instance [7], Chap. II, § 16.

Lemma 109 (Vahlen, 1895)**.** *Among two consecutive convergents p_n/q_n and p_{n+1}/q_{n+1}, one at least satisfies $|x - p/q| < 1/2q^2$.*

Proof. Since $x - p_n/q_n$ and $x - p_{n-1}/q_{n-1}$ have opposite signs,

$$\left| x - \frac{p_n}{q_n} \right| + \left| x - \frac{p_{n-1}}{q_{n-1}} \right| = \left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = \frac{1}{q_n q_{n-1}} < \frac{1}{2q_n^2} + \frac{1}{2q_{n-1}^2}.$$

The last inequality is $ab < (a^2 + b^2)/2$ for $a \neq b$ with $a = 1/q_n$ and $b = 1/q_{n-1}$. Therefore,

$$\text{either} \quad \left| x - \frac{p_n}{q_n} \right| < \frac{1}{2q_n^2} \quad \text{or} \quad \left| x - \frac{p_{n-1}}{q_{n-1}} \right| < \frac{1}{2q_{n-1}^2}.$$

□

Lemma 110 (É. Borel, 1903)**.** *Among three consecutive convergents p_{n-1}/q_{n-1}, p_n/q_n and p_{n+1}/q_{n+1}, one at least satisfies $|x - p/q| < 1/\sqrt{5}q^2$.*

This completes the proof of the irrationality criterion Proposition 4 including (i) ⇒ (vi) in § 2.1.

The fact that the constant $\sqrt{5}$ cannot be replaced by a larger one was proved in Lemma 41. This is true for any number with a continued fraction expansion having all but finitely many partial quotients equal to 1 (which means the Golden number Φ and all rational numbers which are equivalent to Φ modulo $\mathrm{GL}_2(\mathbf{Z})$).

Proof. Recall Lemma 71: for $n \geq 0$,

$$q_n x - p_n = \frac{(-1)^n}{x_{n+1}q_n + q_{n-1}}.$$

Therefore $|q_n x - p_n| < 1/\sqrt{5}q_n$ if and only if $|x_{n+1}q_n + q_{n-1}| > \sqrt{5}q_n$. Define $r_n = q_{n-1}/q_n$. Then this condition is equivalent to $|x_{n+1} + r_n| > \sqrt{5}$.

Recall the inductive definition of the convergents:

$$x_{n+1} = a_{n+1} + \frac{1}{x_{n+2}}.$$

Also, using the definitions of r_n, r_{n+1}, and the inductive relation $q_{n+1} = a_{n+1}q_n + q_{n-1}$, we can write

$$\frac{1}{r_{n+1}} = a_{n+1} + r_n.$$

Eliminate a_{n+1}:

$$\frac{1}{x_{n+2}} + \frac{1}{r_{n+1}} = x_{n+1} + r_n.$$

Assume now

$$|x_{n+1} + r_n| \leq \sqrt{5} \quad \text{and} \quad |x_{n+2} + r_{n+1}| \leq \sqrt{5}.$$

We deduce

$$\frac{1}{\sqrt{5} - r_{n+1}} + \frac{1}{r_{n+1}} \leq \frac{1}{x_{n+2}} + \frac{1}{r_{n+1}} = x_{n+1} + r_n \leq \sqrt{5},$$

which yields

$$r_{n+1}^2 - \sqrt{5}r_{n+1} + 1 \leq 0.$$

The roots of the polynomial $X^2 - \sqrt{5}X + 1$ are $\Phi = (1+\sqrt{5})/2$ and $\Phi^{-1} = (\sqrt{5}-1)/2$. Hence $r_{n+1} > \Phi^{-1}$ (the strict inequality is a consequence of the irrationality of the Golden ratio). .

This estimate follows from the hypotheses $|q_n x - p_n| < 1/\sqrt{5}q_n$ and $|q_{n+1}x - p_{n+1}| < 1/\sqrt{5}q_{n+1}$. If we also had $|q_{n+2}x - p_{n+2}| < 1/\sqrt{5}q_{n+2}$, we would deduce in the same way $r_{n+2} > \Phi^{-1}$. This would give

$$1 = (a_{n+2} + r_{n+1})r_{n+2} > (1 + \Phi^{-1})\Phi^{-1} = 1,$$

which is impossible.

□

Lemma 111 (Legendre, 1798)**.** *If $p/q \in \mathbf{Q}$ satisfies $|x - p/q| \leq 1/2q^2$, then p/q is a convergent of x.*

Proof. Let r and s in $\mathbf{Z}$ satisfy $1 \leq s < q$. From

$$1 \leq |qr - ps| = |s(qx-p) - q(sx-r)| \leq s|qx-p| + q|sx-r| \leq \frac{s}{2q} + q|sx-r|$$

one deduces

$$q|sx - r| \geq 1 - \frac{s}{2q} > \frac{1}{2} \geq q|qx - p|.$$

Hence $|sx - r| > |qx - p|$ and therefore Lemma 105 implies that p/q is a convergent of x.

□

References

[1] P. Flajolet, B. Vallée, I. Vardi, *Continued fractions from Euclid to the present day*, 44p.
http://www.lix.polytechnique.fr/Labo/Ilan.Vardi/continued_fractions.ps

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°9, *May 17, 2010*

This course was devoted to

- Proposition 94 of Mollin and Srinivasan on the negative Pell's equation $x^2 - Dy^2 = -1$.
- The proof of Legendre's Theorem 111 according to which an approximation p/q of an irrational number x satisfying $|x - p/q| \leq 1/q^2$ is a convergent of x.
- The proof of Corollary 100 on the continued fraction expansion of the square root of a rational number.
- An introduction to number fields and the connexion between Pell's equation and Dirichlet's unit Theorem.

Dirichlet's unit Theorem

A number field is a finite algebraic extension of $\mathbf{Q}$, which means a field containing $\mathbf{Q}$ as a subfield and which is a $\mathbf{Q}$–vector space of finite dimension.

In a finite extension, any element is algebraic.

An example of a number field is $\mathbf{Q}(\alpha)$ (the smallest field containing α, or the field generated by α), when α is an algebraic number. In this case $\mathbf{Q}(\alpha) = \mathbf{Q}[\alpha]$, which means that the ring $\mathbf{Q}[\alpha]$ generated by α over $\mathbf{Q}$ is a field. According to the *Theorem of the primitive element*, any number field can be written $\mathbf{Q}(\alpha)$ for some algebraic number α.

Let $f \in \mathbf{Q}[X]$ be the (monic) irreducible polynomial of α. The degree d of f is the dimension of the $\mathbf{Q}$–vector space $\mathbf{Q}(\alpha)$, it is called the *degree of α over $\mathbf{Q}$* and also the *degree of the extension* $\mathbf{Q}(\alpha)/\mathbf{Q}$, it is denoted by $[\mathbf{Q}(\alpha) : \mathbf{Q}]$.

When we factorize the polynomial f over $\mathbf{R}$ into a product of irreducible polynomials, we get a certain number, say r_1, of degree 1 polynomials, and a certain number, say r_2, of degree 2 polynomials with negative discriminant.

Hence $0 \leq r_1 \leq d$, $0 \leq r_2 \leq d/2$ and $r_1 + 2r_2 = d$. In $\mathbf{C}$, f has d distinct roots, r_1 of which are real, say $\alpha_1, \ldots, \alpha_{r_1}$, and $2r_2$ of which are not real and pairwise complex conjugates, say $\alpha_{r_1+1}, \ldots, \alpha_{r_1+r_2}$, $\overline{\alpha}_{r_1+1}, \ldots, \overline{\alpha}_{r_1+r_2}$. There are exactly d fields homomorphisms (also called *embeddings*) $\sigma_i : \mathbf{Q}(\alpha) \longrightarrow \mathbf{C}$, where, for $1 \leq i \leq d$, σ_i is uniquely determined by $\sigma_i(\alpha) = \alpha_i$. For γ in $\mathbf{Q}(\alpha)$, the elements $\sigma_i(\gamma)$ are the conjugates of γ (that means the complex roots of the irreducible polynomial of γ), n of them are distinct, where $n = [\mathbf{Q}(\gamma) : \mathbf{Q}]$ divides d, say $d = nk$, and

$$\prod_{i=1}^{d}(X - \sigma_i(\gamma))$$

is the k–th power of the irreducible polynomial of γ.

Let k be a number field. The norm $\mathrm{N}_{k/\mathbf{Q}}$ is the homomorphism between the multiplicative groups $k^\times = k \setminus \{0\} \longrightarrow \mathbf{Q}^\times$ defined by

$$\mathrm{N}_{k/\mathbf{Q}}(\gamma) = \sigma_1(\gamma) \cdots \sigma_d(\gamma).$$

The *canonical embedding* of k is $\underline{\sigma} = (\sigma_1, \ldots, \sigma_{r_1+r_2}) : k \longrightarrow \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$.

An algebraic number α is called an *algebraic integer* is it satisfies the following equivalent conditions.
(i) The irreducible (monic) polynomial of α in $\mathbf{Q}[X]$ has its coefficients in $\mathbf{Z}$.
(ii) There exists a monic polynomial with rational integer coefficients having α as a root.
(iii) The subring $\mathbf{Z}[\alpha]$ of $\mathbf{C}$ generated by α is a finitely generated $\mathbf{Z}$–module.
(iii) There exists a ring which contains $\mathbf{Z}[\alpha]$ as a subring and which is a finitely generated $\mathbf{Z}$–module.

For instance, the algebraic integers in $\mathbf{Q}$ are the rational integers.

The set of algebraic integers is a subring of $\mathbf{C}$. Its intersection with a number field k is the *ring of integers of* k, which we denote by $\mathbf{Z}_k$. For instance, when $k = \mathbf{Q}(\sqrt{D})$, where D is a rational integer which is not a square,

$$\mathbf{Z}_k = \begin{cases} \mathbf{Z}[\sqrt{D}] & \text{if } D \equiv 2 \text{ or } 3 \pmod 4, \\ \mathbf{Z}[(1+\sqrt{D})/2] & \text{if } D \equiv 1 \pmod 4. \end{cases}$$

It is easy to check that the image $\underline{\sigma}(\mathbf{Z}_k)$ of the ring of integers of k under the canonical embedding is discrete in $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$.

The group of units $\mathbf{Z}_k^\times$ of $\mathbf{Z}_k$ is also called *the group of units* of the number field k (this terminology is standard but should not yield to a confusion: recall that the units in a field k are the non–zero elements of k!). An integer

in k is a unit if and only if it has norm ± 1. The torsion elements of $\mathbf{Z}_k^\times$ are the roots of unity in k, it is easy to check that they form a finite cyclic group $k_{\text{tors}}^\times$.

The *logarithmic embedding* is the map $\lambda : k^\times \longrightarrow \mathbf{R}^{r_1+r_2}$ obtained by composing the restriction of $\underline{\sigma}$ to $k^\times$ with the map

$$(z_n)_{1\leq n\leq r_1+r_2} \longmapsto (\log|z_n|)_{1\leq n\leq r_1+r_2}$$

from $(\mathbf{R}^\times)^{r_1} \times (\mathbf{C}^\times)^{r_2}$ to $\mathbf{R}^{r_1+r_2}$:

$$\lambda(\alpha) = (\log|\sigma_n(\alpha)|)_{1\leq n\leq r_1+r_2}.$$

The image $\lambda(\mathbf{Z}_k^\times)$ of the group of units of k is a subgroup of the additive group $\mathbf{R}^{r_1+r_2}$, it is contained in the hyperplane H of equation

$$x_1 + \cdots + x_{r_1+r_2} = 0,$$

and $\lambda(\mathbf{Z}_k^\times)$ is discrete in H. From these properties , one easily deduces that as a $\mathbf{Z}$–module, $\mathbf{Z}_k^\times$ is finitely generated of rank $\leq r$, where $r = r_1 + r_2 - 1$ is the dimension of H as a $\mathbf{R}$–vector space.

Dirichlet's units Theorem states:

Theorem. *The group of units of an algebraic number field k of degree d with r_1 real embeddings and $2r_2$ conjugate complex embeddings is a finitely generated group of rank $r := r_1 + r_2 - 1$.*

In other terms, there exists a system of fundamental units $(u_1, \ldots, u_r)$ in $\mathbf{Z}_k^\times$, such that any unit $u \in \mathbf{Z}_k^\times$ can be written in a unique way as $\zeta u_1^{m_1} \ldots u_r^{m_r}$, where $\zeta \in k$ is a root of unity and $m_1, \ldots, m_r$ are rational integers:

$$\mathbf{Z}_k^\times \simeq k_{\text{tors}}^\times \times \mathbf{Z}^r.$$

In the special case of a real quadratic field $\mathbf{Q}(\sqrt{D})$ with $D \equiv 2$ or $3 \pmod 4$, the fact that the group of units is a finitely generated group of rank 1 means that the set of solution of Pell's equation $X^2 - Dy^2 = \pm 1$ is the set of $\pm(x_m, y_m)$, $m \in \mathbf{Z}$, where x_m and y_m are defined by $x_m + y_m\sqrt{D} = (x_1 + y_1\sqrt{D})^m$, where (x_1, y_1) denotes the fundamental solution of Pell's equation.

The proof of the existence of a system of r fundamental units rests on Minkowski's geometry of numbers.

There are plenty of references on this subject. Lists of *online number theory lecture notes and teaching materials* are available on the internet. For instance

`http://www.numbertheory.org/ntw/lecture_notes.html`

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°10, *May 19, 2010*

This course was devoted to Markoff's equation - see

On the Markoff equation $x^2 + y^2 + z^2 = 3xyz$*;*

`http://www.math.jussieu.fr/~miw/articles/pdf/MarkoffEn2011.pdf`

and

`http://www.math.jussieu.fr/~miw/articles/pdf/MarkoffEn2011VI.pdf`

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°11, *May 24, 2010*

Recall Hurwitz's Theorem, which is the implication (i)$\Longrightarrow$(vi) of Proposition 4.

Lemma 112. *Let ϑ be a real number. The following conditions are equivalent:*
(i) *ϑ is irrational.*
(ii) *There exist infinitely many $p/q \in \mathbf{Q}$ such that*

$$0 < \left|\vartheta - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}\cdot$$

We proved it by using continued fractions, as a consequence of Borel's Lemma 110: *among three consecutive convergents of the continued fraction of an irrational number ϑ, one at least satisfies property* (ii) *of Lemma 112.*

We give two further proofs of Lemma 112: the first one rests on Farey's series, the last one does not involve continued fractions nor Farey series (but the ideas are very similar). The last proof yields a new irrationality criterion (Lemma 120).

6.5 Farey series

6.5.1 Definition and properties

For $n \geq 1$, the *Farey series* $\mathcal{F}_n$ of order n is the finite increasing sequence of rational numbers in the range $[0, 1]$ having denominators $\leq n$. Each of them starts with 0 and ends with 1. Here are the first ones

$$\mathcal{F}_1 = \{0, 1\}$$

$$\mathcal{F}_2 = \left\{0, \frac{1}{2}, 1\right\}$$

$$\mathcal{F}_3 = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\right\}$$

$$\mathcal{F}_4 = \left\{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1\right\}$$
$$\mathcal{F}_5 = \left\{0, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1\right\}$$
$$\mathcal{F}_6 = \left\{0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, 1\right\}$$
$$\mathcal{F}_7 = \left\{0, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, 1\right\}$$
$$\mathcal{F}_8 = \left\{0, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{3}{8}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{5}{8}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, 1\right\}$$

The number of elements in $\mathcal{F}_n$ is given by the inductive relation

$$|\mathcal{F}_n| = |\mathcal{F}_{n-1}| + \varphi(n),$$

with $|\mathcal{F}_1| = 2$, where $\varphi(n)$ is Euler's function ($\varphi(n)$ is the number of integers in the range $1, \ldots, n$ which are relativey prime to n). Hence

$$|\mathcal{F}_n| = 1 + \sum_{m=1}^{n} \varphi(m).$$

One can deduce the estimate

$$|\mathcal{F}_n| \sim \frac{3n^2}{\pi^2}.$$

Proposition 113. *If $h/k < h'/k'$ are successive terms in a Farey series $\mathcal{F}_n$, then $h'k - hk' = 1$.*

For the proof, we follow § I.2 of [2]. Other proofs are given in [1], Chap. 3.

Lemma 114. *Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be two elements of $\mathbf{Z}^2$. The following conditions are equivalent:*
(i) *(x, y) is a basis of $\mathbf{Z}^2$ over $\mathbf{Z}$.*
(ii) *$x_1y_2 - x_2y_1 = \pm 1$.*
(iii) *x and y are are linearly independent over $\mathbf{R}$, and the closed parallelogram*

$$\mathcal{P} = \left\{\lambda x + \mu y \ ; \ \lambda \in \mathbf{R}, \ \mu \in \mathbf{R}, \ 0 \leq \lambda \leq 1, \ 0 \leq \mu \leq 1\right\}$$

with vertices 0, x, y and $x + y$ does not contain integer points in $\mathbf{Z}^2$ but its vertices.
(iv) *x and y are are linearly independent over $\mathbf{R}$, and the closed triangle*

$$\mathcal{T} = \left\{\lambda x + \mu y \ ; \ \lambda \in \mathbf{R}_{\geq 0}, \ \mu \in \mathbf{R}_{\geq 0}, \ \lambda + \mu \leq 1\right\}$$

with vertices 0, x and y, does not contain integer points in $\mathbf{Z}^2$ but its vertices.

Proof. A change of basis for $\mathbf{Z}^2$ has a invertible matrix with determinant a unit in $\mathbf{Z}$, hence (i) $\Longleftrightarrow$ (ii).

Assume (i). Any element z in $\mathbf{Z}^2$ can be written in a unique way as $\lambda x + \mu y$ with λ and μ in $\mathbf{R}$, and these numbers λ and μ are in $\mathbf{Z}$. Hence, when $z \in \mathcal{P}$, we have $0 \leq \lambda \leq 1$, $0 \leq \mu \leq 1$, and therefore each of λ, μ is 0 or 1. This proves (iii).

Conversely, assume (iii). Let $u \in \mathbf{Z}^2$. Since (x, y) is a basis of $\mathbf{R}^2$ over $\mathbf{R}$, we can write $u = tx + t'y$ with t and t' in $\mathbf{R}$. Define two integers a and a' by $a = \lfloor t \rfloor$ and $a' = \lfloor t' \rfloor$. From $0 \leq t - a < 1$ and $0 \leq t' - a' < 1$ we deduce $u - ax - a'y \in \mathbf{Z}^2 \cap \mathcal{P}$ with $u - ax - a'y \notin \{x,\ y\}$, hence $u = ax + a'y$. This proves (i).

Since $\mathcal{P}$ contains $\mathcal{T}$, (iii) implies (iv).

Finally, assume (iv). If $z \in \mathcal{P} \cap \mathbf{Z}^2$ is distinct from 0, x and y, then $z \notin \mathcal{T}$, from which we deduce that $x + y - z \in \mathcal{T} \cap \mathbf{Z}^2$. From $z \neq x$ and $z \neq y$ we deduce $x + y - z = 0$, hence $z = x + y$. Therefore $\mathcal{P} \cap \mathbf{Z}^2 = \{0,\ x,\ y,\ x + y\}$, which is (iii).

□

Proof of Proposition 113. Let $h/k < h'/k'$ be successive terms in the Farey series $\mathcal{F}_n$. From $(h, k) \neq (h', k')$ and $\gcd(h, k) = \gcd(h', k') = 1$, we deduce that the two vectors (h, k) and (h', k') of $\mathbf{R}^2$ are linearly independent. Since $h'k - hk' > 0$, using Lemma 114, it suffices to check that the triangle $\mathcal{T}$ with vertices 0, x and y does not contain any element of $\mathbf{Z}^2$ but the vertices. Assume $z = (h'', k'') \in \mathcal{T} \cap \mathbf{Z}^2$ with $z \notin \{0,\ x,\ y\}$. We have $z = \lambda x + \mu y$ with $\lambda \geq 0$, $\mu \geq 0$, $0 < \lambda + \mu \leq 1$, $(\lambda, \mu) \notin \{(0, 1)\ ;\ (1, 0)\}$. Then $k'' = \lambda k + \mu k' \leq n$ and $h/k < h''/k'' < h'/k'$, which contradicts the assumption that there is no element between h/k and h'/k' in $\mathcal{F}_n$.

□

Corollary 115. *if $h/k < h''/k'' < h'/k'$ are successive elements in a Farey series $\mathcal{F}_n$, then*

$$\frac{h''}{k''} = \frac{h + h'}{k + k'}.$$

Proof. From Proposition 113 we deduce $h''k - hk'' = 1$, $h'k'' - h''k' = 1$, hence $h''(k + k') = k''(h + h')$. □

Examples in $\mathcal{F}_5$ of $\mathcal{F}_6$ are $1/3 < 2/5 < 1/2 < 2/5$: the fraction $(h + h')/(k + k')$ may or may not be in reduced form.

Here is our second proof of Lemma 112.

Proposition 116. *Let $h/k < h'/k'$ be successive elements in a Farey series $\mathcal{F}_n$. Define $h'' = h + h'$, $k'' = k + k'$. Then h''/k'' is in reduced form, and for any α in the interval $h/k \leq \alpha \leq h'/k'$, at least one of the following inequalities hold:*

$$\alpha - \frac{h}{k} < \frac{1}{\sqrt{5}k^2}, \qquad \left|\alpha - \frac{h''}{k''}\right| < \frac{1}{\sqrt{5}k''^2}, \qquad \frac{h'}{k'} - \alpha < \frac{1}{\sqrt{5}k'^2}.$$

Proof. From $h(k+k') - (h+h')k = 1$, we deduce that $k+k'$ and $h+h'$ are relatively prime.

By symmetry we may assume $h''/k'' < \alpha < h'/k'$. If none of the inequalities hold, then

$$\alpha - \frac{h}{k} \geq \frac{1}{\sqrt{5}k^2}, \qquad \alpha - \frac{h''}{k''} \geq \frac{1}{\sqrt{5}k''^2}, \qquad \frac{h'}{k'} - \alpha \geq \frac{1}{\sqrt{5}k'^2}.$$

Using $h'k - hk' = 1$ and $h'k'' - h''k' = 1$, we deduce

$$\frac{1}{kk'} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{k^2} + \frac{1}{k'^2}\right)$$

and

$$\frac{1}{k'k''} \geq \frac{1}{\sqrt{5}}\left(\frac{1}{k'^2} + \frac{1}{k''^2}\right).$$

We deduce

$$\sqrt{5}kk' \geq k^2 + k'^2 \quad \text{and} \quad \sqrt{5}k'k'' \geq k'^2 + k''^2,$$

which means that the numbers $x = k/k'$ and $y = k'/k''$ satisfy

$$x^2 - \sqrt{5}x + 1 \leq 0 \quad \text{and} \quad y^2 - \sqrt{5}y + 1 \leq 0.$$

Since the roots of $X^2 - \sqrt{5}X + 1$ are Φ and $1/\Phi$, it follows that x and y lie in the intervall $(1/\Phi, \Phi)$. From $k'' = k + k'$ we deduce $1/y = x + 1$, hence:

$$\frac{1}{\Phi} + 1 \leq x + 1 = \frac{1}{y} \leq \Phi.$$

Since x and y are rational numbers, this is not compatible with the irrationality of Φ. □

Notice that the end of the proof is the same as the proof of Borel's Lemma 110.

We conclude this section by some further remark on Farey sequences, which we do not plan to use, but which may be interesting to know.

The converse of Corollary 115 is true: *If h, k, h', k' are positive integers with $0 < h/k < h'/k' < 1$ which satisfy $h'k - kh' = 1$, then h/k and h'/k' are consecutive elements in the Farey series $\mathcal{F}_n$ with $n = \max\{k, k'\}$.*

Here is the proof. Suppose first $k \geq k'$. Denote by h''/k'' the successor of h/k in $\mathcal{F}_k$. Then $h''k - k''h = 1$ and $1 \leq k'' \leq k$, hence $(h''-h')k = (k''-k')h$, which shows that k' and k'' are congruent modulo k. Since they both lie in the interval $[1, k]$, we deduce $k' = k''$, hence $h' = h''$.

Similarly, if $k < k'$, we denote by h''/k'' the predecessor of h'/k' in $\mathcal{F}_{k'}$. The same argument gives $h/k = h''/k''$. □

It follows that *if $h/k < h'/k'$ are consecutive in the Farey series $\mathcal{F}_n$, then the smallest $m > n$ such that there is an element h''/k'' of $\mathcal{F}_m$ in the interval $h/k < h''/k'' < h'/k'$ is $m = k + k'$, this element h''/k'' is unique and $h'' = h + h'$, $k'' = k + k' = m$.*

Indeed, by definition of m, we have $m = k''$. From the inequalities

$$\frac{h}{k} < \frac{h+h'}{k+k'} < \frac{h'}{k'},$$

it follows that $m \leq k + k'$. The unicity of an element of $\mathcal{F}_m$ in this interval follows from the fact that two distinct rational numbers with denominator m are at distance $\geq 1/m$, while Proposition 113 yields

$$\frac{h'}{k'} - \frac{h}{k} = \frac{1}{kk'} < \frac{1}{m}.$$

We have seen in Proposition 116 that $(h + h')/(k + k')$ is in reduced form. Finally Corollary 115 shows that $h''/k'' = (h+h')/(k+k')$, hence $k'' = k+k'$.

Here is a connection with continued fractions: *let p/q be an irreducible fraction with $q \geq 2$; write the continued fraction of p/q which ends with $a_n \geq 2$ as $p/q = [0, a_1, \ldots, a_n]$. Then the predecessors and successors of p/q in the Farey series $\mathcal{F}_q$ have continued fractions* $[0, a_1, \ldots, a_n - 1]$ and $[0, a_1, \ldots, a_{n-1}]$:

$$[0, a_1, \ldots, a_{n-1}] < \frac{p}{q} = [0, a_1, \ldots, a_n] < [0, a_1, \ldots, a_n - 1] \quad \text{if } n \text{ is odd},$$

$$[0, a_1, \ldots, a_n - 1] < \frac{p}{q} = [0, a_1, \ldots, a_n] < [0, a_1, \ldots, a_{n-1}] \quad \text{if } n \text{ is odd}.$$

Indeed, using the other continued fraction $p/q = [0, a_1, \ldots, a_n - 1, 1] = p_{n+1}/q_{n+1}$, we write as in (64)

$$\begin{pmatrix} p & p_n \\ q & q_n \end{pmatrix} = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \begin{pmatrix} a_n - 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

where $p_n/q_n = [0, a_1, \ldots, a_n - 1]$ and $p_{n-1}/q_{n-1} = [0, a_1, \ldots, a_{n-1}]$, and we have $q_{n-1} < q_n < q$, $pq_n - p_n q = (-1)^n$, $pq_{n-1} - p_{n-1}q = (-1)^{n-1}$ (recall Lemma 68 with n replaced by $n+1$ and $a_{n+1} = 1$). Hence the result follows from the previous remarks.

6.5.2 Hurwitz Theorem

Here is the third proof of of Hurwitz's Lemma 112.

We start with the next auxiliary result, which also follows from the results we proved on continued fractions (take for p/q and r/s two consecutive convergents of ϑ) or on Farey series (take two consecutive elements of a Farey series such that ϑ is in their interval).

Lemma 117. *Let ϑ be a real irrational number. Then there exist infinitely many pairs $(p/q, r/s)$ of irreducible fractions such that*

$$\frac{p}{q} < \vartheta < \frac{r}{s} \quad \textit{and} \quad qr - ps = 1.$$

In this statement and the next ones, it is sufficient to prove inequalities $\leq$ in place of $<$: this follows from the irrationality of ϑ.

Proof. Let H be a positive integer. Among the irreducible rational fractions a/b with $1 \leq b \leq H$, select one for which $|\vartheta - a/b|$ is minimal. If $a/b < \vartheta$ rename a/b as p/q, while if $a/b > \vartheta$, then rename a/b as r/s.

First consider the case where $a/b < \vartheta$, hence $a/b = p/q$. Since $\gcd(p, q) = 1$, using Euclidean's algorithm, one deduces (Bézout's Theorem) that there exist $(r, s) \in \mathbf{Z}^2$ such that $qr - sp = 1$ with $1 \leq s < q$ and $|r| < |p|$. Since $1 \leq s < q \leq H$, from the choice of a/b it follows that

$$\left| \vartheta - \frac{p}{q} \right| \leq \left| \vartheta - \frac{r}{s} \right|$$

hence r/s does not belong to the interval $[p/q, \vartheta]$. Since $qr - sp > 0$ we also have $p/q < r/s$, hence $\vartheta < r/s$.

In the second case where $a/b > \vartheta$ and $r/s = a/b$ we solve $qr - sp = 1$ by Euclidean algorithm with $1 \leq q < s$ and $|p| < r$, and the argument is similar.

We now complete the proof of the existence of infinitely many such pairs. Once we have a finite set of such pairs $(p/q, r/s)$, we use the fact that there is a rational number m/n closer to ϑ than any of these rational fractions. We use the previous argument with $H \geq n$. This way we produce a new pair $(p/q, r/s)$ of rational numbers which is none of the previous ones (because one at least of the two rational numbers p/q, r/s is a better approximation than the previous ones). Hence this construction yields infinitely many pairs, as claimed.

□

Lemma 118. *Let ϑ be a real irrational number. Assume $(p/q, r/s)$ are irreducible fractions such that*

$$\frac{p}{q} < \vartheta < \frac{r}{s} \quad \textit{and} \quad qr - ps = 1.$$

Then

$$\min\left\{q^2\left(\vartheta - \frac{p}{q}\right)\ ,\ s^2\left(\frac{r}{s} - \vartheta\right)\right\} < \frac{1}{2}.$$

Proof. Define

$$\delta = \min\left\{q^2\left(\vartheta - \frac{p}{q}\right)\ ,\ s^2\left(\frac{r}{s} - \vartheta\right)\right\}.$$

From

$$\frac{\delta}{q^2} \leq \vartheta - \frac{p}{q} \quad \text{and} \quad \frac{\delta}{s^2} \leq \frac{r}{s} - \vartheta$$

with $qr - ps = 1$ one deduces that the number $t = s/q$ satisfies

$$t + \frac{1}{t} \leq \frac{1}{\delta}.$$

Since the minimum of the function $t \mapsto t + 1/t$ is 2 and since $t \neq 1$, we deduce $\delta < 1/2$.

□

Remark. *The inequality $t + (1/t) \geq 2$ for all $t > 0$ with equality if and only if $t = 1$ is equivalent to the arithmetico-geometric inequality*

$$\sqrt{xy} \leq \frac{x+y}{2},$$

when x and y are positive real numbers, with equality if and only if $x = y$. The correspondance between both estimates is $t = \sqrt{x/y}$.

From Lemmas 117 and 118 it follows that for $\vartheta \in \mathbf{R} \setminus \mathbf{Q}$, there exist infinitely many $p/q \in \mathbf{Q}$ such that

$$0 < \left| \vartheta - \frac{p}{q} \right| < \frac{1}{2q^2}.$$

A further step is required in order to complete the proof of Lemma 112.

Lemma 119. *Let ϑ be a real irrational number. Assume $(p/q, r/s)$ are irreducible fractions such that*

$$\frac{p}{q} < \vartheta < \frac{r}{s} \quad \textit{and} \quad qr - ps = 1.$$

Define $u = p + r$ and $v = q + s$. Then

$$\min \left\{ q^2 \left(\vartheta - \frac{p}{q} \right) \ , \ s^2 \left(\frac{r}{s} - \vartheta \right) \ , \ v^2 \left| \vartheta - \frac{u}{v} \right| \right\} < \frac{1}{\sqrt{5}}.$$

Proof. First notice that $qu - pv = 1$ and $rv - su = 1$. Hence

$$\frac{p}{q} < \frac{u}{v} < \frac{r}{s}.$$

We repeat the proof of lemma 118 ; we distinguish two cases according to whether u/v is larger or smaller than ϑ. Since both cases are quite similar, let us assume $\vartheta < u/v$. The proof of lemma 118 shows that

$$\frac{s}{q} + \frac{q}{s} \le \frac{1}{\delta} \quad \text{and} \quad \frac{v}{q} + \frac{q}{v} \le \frac{1}{\delta}.$$

Hence each of the four numbers s/q, q/s, v/q, q/v satisfies $t + 1/t \le 1/\delta$. Now the function $t \mapsto t + 1/t$ is decreasing on the interval $(0, 1)$ and increasing on the interval $(1, +\infty)$. It follows that our four numbers all lie in the interval $(1/x, x)$, where x is the root > 1 of the equation $x + 1/x = 1/\delta$. The two roots x and $1/x$ of the quadratic polynomial $X^2 - (1/\delta)X + 1$ are at a mutual distance equal to the square root of the discriminant $\Delta = (1/\delta)^2 - 4$ of this polynomial. Now

$$\frac{v}{q} - \frac{s}{q} = 1,$$

hence the length $\sqrt{\Delta}$ of the interval $(1/x, x)$ is ≥ 1 and therefore $\delta \le 1/\sqrt{5}$. This completes the proof of Lemma 119. □

Remark. In the three proofs of Hurwitz's Theorem, the number $\sqrt{5}$ occurs as follows: for any $x > 1$,

$$\max\left\{x + \frac{1}{x}\ ,\ \frac{1+x}{x} + \frac{x}{1+x}\right\} \geq \sqrt{5},$$

with equality if and only if $x = \Phi$ (the Golden ratio). Indeed, for $x > 1$ we have

$$x + \frac{1}{x} > \sqrt{5} \Longleftrightarrow x > \Phi$$

and, with $t = (x+1)/x$,

$$t + \frac{1}{t} > \sqrt{5} \Longleftrightarrow t > \Phi \Longleftrightarrow x + \frac{1}{x} > \sqrt{5} \Longleftrightarrow x < \Phi.$$

6.5.3 A further irrationality criterion

Lemma 120. *Let ϑ be a real number. The following conditions are equivalent:*
(i) *ϑ is irrational.*
(ii) *For any $\epsilon > 0$ there exists p/q and r/s in $\mathbf{Q}$ such that*

$$\frac{p}{q} < \vartheta < \frac{r}{s}, \quad qr - ps = 1$$

and

$$\max\{q\vartheta - p\ ;\ r - s\vartheta\} < \epsilon.$$

(iii) *There exist infinitely many pairs $(p/q, r/s)$ of rational numbers such that*

$$\frac{p}{q} < \vartheta < \frac{r}{s}, \quad qr - ps = 1$$

and

$$\max\{q(q\vartheta - p)\ ;\ s(r - s\vartheta)\} < 1.$$

Proof. The implications (iii) $\Longrightarrow$ (ii) $\Longrightarrow$ (i) are easy. For (i)$\Longrightarrow$(iii) we are going to combine the arguments in the proof of Lemma 117 with results from the theory of continued fractions.

Since ϑ is irrational, there are infinitely many p/q such that

$$\left|\vartheta - \frac{p}{q}\right| < \frac{1}{2q^2}.$$

This is a weak form of Hurwitz Lemma 112. According to Legendre's Lemma 111, such a p/q is a convergent of ϑ. The best approximation property of the convergents (Lemma 105) implies that for any $a/b \in \mathbf{Q}$ with $1 \leq b \leq q$ and $a/b \neq p/q$, we have

$$\left|\vartheta - \frac{a}{b}\right| > \left|\vartheta - \frac{p}{q}\right|.$$

Assume first $p/q < \vartheta$. Let r/s be defined by $qr - ps = 1$ and $1 \leq s < q$, $|r| < |p|$. We have

$$0 < \frac{r}{s} - \vartheta < \frac{r}{s} - \frac{p}{q} = \frac{1}{qs} \leq \frac{1}{s^2}.$$

Next assume $p/q > \vartheta$. In this case rename it r/s and define p/q by $qr - ps = 1$ and $1 \leq q < s$, $|p| < |r|$.

Finally, repeat the argument in the proof of Lemma 117 to get an infinite set of approximations. Lemma 120 follows. □

References

[1] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Oxford University Press, Oxford, sixth ed., 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[2] W. M. Schmidt, *Diophantine approximation*, vol. **785**, Lecture Notes in Mathematics. Berlin-Heidelberg-New York: Springer-Verlag, 1980, new ed. 2001.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°12, *May 26, 2010*

7 Approximation of functions

We give Lambert's proof of the irrationality of π and e^r for $r \in \mathbf{Q} \setminus \{0\}$, involving continued fractions of analytic functions. Then we give a very short introduction to generalized hypergeometric functions.

7.1 Lambert's proof of the irrationality of π and e^r for $r \in \mathbf{Q} \setminus \{0\}$

The fundamental result of Lambert's paper [3] is:

Theorem 121 (Lambert, 1761)**.** *For any $r \in \mathbf{Q} \setminus \{0\}$, the numbers $\tan r$ and e^r are irrational. In particular the number π is irrational.*

The main tool is continued fractions, and the first goal of Lambert is to develop $\tan x = \sin x / \cos x$ and $(e^x - e^{-x})/(e^x + e^{-x})$ into continued fractions.

Proposition 122. *The functions $\tan x$ and $(e^x - e^{-x})/(e^x + e^{-x})$ can be represented as a continued fraction*

$$\tan x = \frac{x|}{|1} + \frac{-x^2|}{|\ 3} + \frac{-x^2|}{|\ 5} + \cdots + \frac{-x^2\ |}{|2k-1} + \cdots$$

and

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{x|}{|1} + \frac{x^2|}{|\ 3} + \frac{x^2|}{|\ 5} + \cdots + \frac{x^2\ |}{|2k-1} + \cdots$$

Each of these continued fractions converges uniformly to the function in the left hand side on any compact subset of $\mathbf{C}$ on which this function is bounded.

These two formulae are related by

$$\tan t = \frac{1}{i} \cdot \frac{e^{it} - e^{-it}}{e^{it} + e^{-it}}.$$

The next tool is a criterion for irrationality, by means of such irregular continued fractions. Here is Proposition 1, § 4.3.3, of [1].

Proposition 123. *Let $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ be two sequences of rational integers. Assume that the continued fraction*

$$\frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \frac{b_3|}{|a_3} + \cdots + \frac{b_n|}{|a_n} + \cdots$$

converges to some real number x. Assume also that there exists a positive integer n_0 such that, for all $n \geq n_0$, we have $0 < |b_n| < |a_n|$. Then for each $n \geq 1$ the continued fraction

$$\frac{b_n|}{|a_n} + \frac{b_{n+1}|}{|a_{n+1}} + \frac{b_{n+2}|}{|a_{n+2}} + \cdots + \frac{b_{n+m}|}{|a_{n+m}} + \cdots$$

converges to a limit x_n. Further, we have $|x_n| \leq 1$ for all $n \geq n_0$. Furthermore, if $x_n \neq \pm 1$ for all $n \geq n_0$, then x is irrational.

From

$$\frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \frac{b_3|}{|a_3} + \cdots + \frac{b_n \quad |}{|a_n + x_{n+1}},$$

using (51), we deduce

$$x = \frac{A_{n-1} + x_n A_{n-2}}{B_{n-1} + x_n B_{n-2}}.$$

This is an analog of (70) but for generalized continued fractions and with x_n replaced by $1/x_n$. Therefore, x is rational if and only if x_n is rational for at least one $n \geq 1$, if and only if x_n is rational for all $n \geq 1$.

We assume these two propositions and we complete the proof of the irrationality of $\tan r$ for $r \in \mathbf{Q}$ non–zero.

We shall use several times the following lemma, which means, in short terms

$$a_0 + \frac{b_1|}{|a_1} + \frac{b_2|}{|a_2} + \cdots + \frac{b_n|}{|a_n} = a_0 + \frac{\lambda_1 b_1|}{|\lambda_1 a_1} + \frac{\lambda_1\lambda_2 b_2|}{|\ \lambda_2 a_2} + + \cdots + \frac{\lambda_{n-1}\lambda_n b_n|}{|\quad \lambda_n a_n}.$$

Lemma 124. *Consider a generalized finite continued fraction and define, as usual (cf. (51))*

$$\begin{pmatrix} A_n & A_{n-1} \\ B_n & B_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ b_{n-1} & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ b_n & 0 \end{pmatrix}.$$

Let $\lambda_1, \ldots, \lambda_n$ be further variables. Define, for $n \geq 0$, $a'_n = \lambda_n a_n$ and, for $n \geq 1$, $b'_n = \lambda_{n-1}\lambda_n b_n$, with $\lambda_0 = 1$. Then the polynomials A'_n and B'_n defined by

$$\begin{pmatrix} A'_n & A'_{n-1} \\ B'_n & B'_{n-1} \end{pmatrix} = \begin{pmatrix} a'_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a'_1 & 1 \\ b'_1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a'_{n-1} & 1 \\ b'_{n-1} & 0 \end{pmatrix} \begin{pmatrix} a'_n & 1 \\ b'_n & 0 \end{pmatrix}.$$

are

$$A'_n = \lambda_1 \cdots \lambda_n A_n \quad \textit{and} \quad B'_n = \lambda_1 \cdots \lambda_n B_n.$$

In particular

$$\frac{A'_n}{B'_n} = \frac{A_n}{B_n}.$$

Proof. This is true for $n = 0$ and $n = 1$, and by induction this follows from the recurrence formulae satisfied by A_n, B_n, A'_n and B'_n:

$$A'_n = a'_n A'_{n-1} + b'_n A'_{n-2}, \qquad B'_n = a'_n B'_{n-1} + b'_n B'_{n-2}.$$

□

Proof of Lambert's irrationality result on $\tan r$ *for* $r \in \mathbf{Q} \setminus \{0\}$. Write $r = p/q$ with $q \geq 1$ and $p \neq 0$ integers. From proposition 122 we deduce

$$\tan p/q = \frac{p/q|}{|\ 1} + \frac{-p^2/q^2|}{|\quad 3} + \frac{-p^2/q^2|}{|\quad 5} + \cdots + \frac{-p^2/q^2|}{|\, 2n+1} + \cdots$$

Lemma 124 with $a_0 = 0$, $a_n = 2n - 1$ for $n \geq 1$, $b_1 = p/q$, $b_n = -p^2/q^2$ for $n \geq 2$, $\lambda_n = q$ for $n \geq 1$, yields

$$\tan p/q = \frac{p|}{|q} + \frac{-p^2|}{|\ 3q} + \frac{-p^2|}{|\ 5q} + \cdots + \frac{-p^2\quad|}{|(2n+1)q} + \cdots$$

For $n > \max\{3, p^2/2q\}$, set

$$y_n = \frac{-p^2\quad|}{|(2n+1)q} + \frac{-p^2\quad|}{|(2n+3)q} + \cdots + \frac{-p^2\quad|}{|(2n+m)q} + \cdots$$

so that

$$y_n = -\frac{p^2}{(2n+1)q + y_{n-1}}.$$

One deduces from Proposition 123 that $|y_n| \leq 1$. From the estimate

$$|y_n| = \frac{p^2}{(2n+1)q - |y_{n-1}|} \leq \frac{p^2}{2nq} < 1,$$

it follows that $|y_n| < 1$. Therefore $y_n \neq \pm 1$ for all sufficiently large n, hence again we can apply Proposition 123 and conclude.

□

The proof of Proposition 123 is similar to the proof of Proposition 60, the main difference being that we do not assume the numbers a_n and b_n to be positive - but here we assume the strict inequality $|a_n| > |b_n|$.

Proof of Proposition 123. We start with the following remark. *Let a, b and x be real numbers satisfying $|a| \geq |b| + 1$, $|b| \geq 1$ and $|x| < 1$. Then $a + x$ has the sign of a and*

$$\left| \frac{b}{a+x} \right| < 1.$$

When a and b are rational integers, the hypotheses on a and b hold as soon as $|a| > |b| > 0$.

From this observation and the assumption $0 < |b_n| < |a_n|$, $0 < |b_{n+1}| < |a_{n+1}|$, we deduce that for all $n \geq n_0$,

$$\frac{b_n|}{|a_n} + \frac{b_{n+1}|}{|a_{n+1}} = \frac{b_n}{a_n + \dfrac{b_{n+1}}{a_{n+1}}}$$

has the same sign as b_n/a_n and has modulus < 1. By induction, one finds that, for all $m \geq 0$,

$$\frac{b_n|}{|a_n} + \frac{b_{n+1}|}{|a_{n+1}} + \cdots + \frac{b_{n+m}|}{|a_{n+m}}$$

has the same sign as b_n/a_n and has modulus < 1. Since the continued fraction (of x, hence of x_n) is convergent, it follows that for all $n \geq n_0$, x_n has the same sign as a_{n_0}/b_{n_0} and $|x_n| \leq 1$.

Assume now that $|x_n| < 1$ for all $n \geq n_0$ and that x is rational. By induction, x_n is rational for all $n \geq 1$; write $x_n = u_n/v_n$ with $|u_n| < v_n$ for $n \geq n_0$. From $x_n = b_n/(a_n + x_{n+1})$ it follows that

$$x_{n+1} = -a_n + \frac{b_n}{x_n} = \frac{-a_n u_n + b_n v_n}{u_n}$$

is a rational number of modulus < 1 and denominator $|u_n|$ smaller than the denominator v_n of x_n. By infinite descent we reach a contradiction. □

Remark. *Assume the assumptions of Proposition 123 are satisfied, but $x_n = \pm 1$ for some $n \geq n_0$. Once some x_n is rational, all x_n are rational, therefore $x_n = \pm 1$ for all sufficiently large n. Since the x_n with $n \geq n_0$ have constant sign, we have $x_n = x_{n+1}$, and from $x_n = b_n/(a_n + b_{n+1})$ with $|a_n| > |b_n| > 0$ we deduce $x_n = -1$ and $a_n = b_n - 1 \leq -2$. An example is*

$$1 = \frac{-1\,|}{|\,-2} + \frac{-1\,|}{|\,-2} + \cdots + \frac{-1\,|}{|\,-2} + \cdots = [0, 2, -2, 2, -2, \ldots].$$

It remains to prove Proposition 122.

Proof of Proposition 122. Lambert starts with the power series expansions of sin and cos:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$$

and

$$\cos x = 1 - x^2 + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$$

Divide sin by cos and write $\tan x = \sin x / \cos x = x/(1 + A_1)$. The power series A_1 starts with $-x^2/3$. Next write $A_1 = -x^2/(3 + A_2)$, so that

$$\tan x = \frac{x}{1 + A_1} = \frac{x}{1 + \dfrac{-x^2}{3 + A_2}}.$$

The first term of A_2 is $-x^2/5$. For $A_2 = -x^2/(5 + A_3)$ we have

$$\tan x = \frac{x}{1 + \dfrac{-x^2}{3 + \dfrac{-x^2}{5 + A_3}}} = \frac{x|}{|1} + \frac{-x^2|}{|\ 3} + \frac{-x^2\ |}{|5 + A_3}.$$

The closed formulae for A_1, A_2 and A_3 are given in [1]. Here is the formula for A_k which is computed from

$$\tan x = \frac{x|}{|1} + \frac{-x^2|}{|\ 3} + \frac{-x^2|}{|\ 5} + \cdots + \frac{-x^2\quad|}{|2k - 1 + A_k},$$

namely

$$A_k = \frac{\displaystyle\sum_{n=0}^{\infty} (-1)^{n+1} x^{2n+2} \frac{(2n+2)(2n+4)\cdots(2n+2k)}{(2n+2k+1)!}}{\displaystyle\sum_{n=0}^{\infty} (-1)^{n} x^{2n} \frac{(2n+2)(2n+4)\cdots(2n+2k-2)}{(2n+2k-1)!}}.$$

One can write also the coefficients respectively

$$\frac{(2n+2)(2n+4)\cdots(2n+2k)}{(2n+2k+1)!} = \frac{2^k (n+k)!}{n!(2n+2k+1)!}$$

and

$$\frac{(2n+2)(2n+4)\cdots(2n+2k-2)}{(2n+2k-1)!} = \frac{2^{k-1}(n+k-1)!}{n!(2n+2k-1)!}.$$

The proof of the convergence of the continued fraction requires to compute the convergents, which is something done by Lambert. He writes

$$\frac{x|}{|1} + \frac{-x^2|}{|\ 3} + \frac{-x^2|}{5} + \cdots + \frac{-x^2\ |}{|2n-1} = \frac{P_n}{Q_n}$$

where

$$P_{n+1} = (2n+1)P_n - x^2 P_{n-1}, \qquad Q_{n+1} = (2n+1)Q_n - x^2 Q_{n-1}$$

for $n \geq 2$, with the initial conditions $P_1 = x$, $Q_1 = 1$, $P_2 = 3x$, $Q_2 = 3 - x^2$. By induction, it follows that the polynomial P_n is odd, of degree n if n is odd and $n-1$ is n is even, while Q_n is an even polynomial, of degree n if n is even and $n-1$ is n is odd. The explicit formulae are

$$P_n = c_n p_n, \quad Q_n = c_n q_n, \quad c_n = 1 \cdot 3 \cdot 5 \cdots (2n-1) = \frac{(2n)!}{2^n n!},$$

with

$$p_n = \sum_{1 \leq k \leq (n+1)/2} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} \cdot \frac{(2n-2k)(2n-2k-2)\cdots(2n-4k+4)}{(2n-1)(2n-3)\cdots(2n-2k+3)}$$

and

$$q_n = \sum_{0 \leq k \leq n/2} (-1)^k \frac{x^{2k}}{(2k)!} \cdot \frac{(2n-2k)(2n-2k-2)\cdots(2n-4k+2)}{(2n-1)(2n-3)\cdots(2n-2k+1)}.$$

As n tends to infinity, p_n and q_n converge uniformly on any compact subset of $\mathbf{C}$ to sin and cos: the difference between the sums of the first k terms in the Taylor expansion at the origin of p_n and sin (respectively of q_n and cos) is bounded above by

$$\frac{|x|^{2k+1}}{(2k+1)!} + \frac{|x|^{2k+2}}{(2k+2)!} + \frac{|x|^{2k+3}}{(2k+3)!} + \cdots$$

and therefore p_n/q_n converge to $\tan x$ uniformly on any compact subset of $\mathbf{C}$ where $|\tan x|$ is bounded.

□

Remark. *In the proof of Theorem 121, we may replace the Lambert's irrationality criterion (Proposition 123) for continued fractions by our standard criterion (Proposition 4) involving rational approximations, as follows.*

Writing the function $f(z) = (1/z)\tan z$ *as a continued fraction and using (54), we obtain, for* $n > 0$,

$$f(z) = \frac{P_n(z)}{Q_n(z)} + \sum_{m>n}\left(\frac{P_m(z)}{Q_m(z)} - \frac{P_{m+1}(z)}{Q_{m+1}(z)}\right) = \frac{P_n(z)}{Q_n(z)} + \sum_{m\geq n}\frac{z^{2m}}{Q_{m-1}Q_m(z)}.$$

The polynomials P_n *and* Q_n *have integral coefficients and degrees* $\leq n$*; for* n *tending to infinity,* $Q_n(p/q)$ *grows like* $2^n n!$*. One checks that the rational approximation given by* $P_n(p/q)/Q_n(p/q)$ *is too sharp for* $f(p/q)$ *to be a rational number.*

From Lemma 124, it follows that the continued fraction for $(e^x - e^{-x})/(e^x + e^{-x})$ given in Proposition 122 can be written

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} = [0, 1/x, 3/x, 5/x, \ldots, (2k-1)/x, \ldots].$$

For $x = 1/2$ this gives

$$\frac{e+1}{e-1} = [2,\ 6,\ 10,\ 14,\ \ldots,\ 4k+2,\ \ldots]. = \overline{[4k+2]}_{k\geq 0}.$$

Let us deduce Euler's continued fraction expansion for e (see § 1.4)

$$e = [2,\ 1,\ 2,\ 1,\ 1,\ 4,\ 1,\ 1\ldots] = [2, \overline{1, 2k, 1}]_{k\geq 1}.$$

Define p_k/q_k as the k–th convergent of $x = [2, 6, \ldots, 4k+2, \ldots]$ and r_k/s_k as the k–th convergent of $y = [1, 1, 2, 1, 1, 4, \ldots, 1, 2k, 1, \ldots]$. We eliminate the indices which are not congruent to 1 modulo 3 among the 5 relations involving 7 symbols

$$\begin{aligned} r_{3k-3} &= r_{3k-4} + r_{3k-5}, \\ r_{3k-2} &= r_{3k-3} + r_{3k-4}, \\ r_{3k-1} &= 2k r_{3k-2} + r_{3k-3}, \\ r_{3k} &= r_{3k-1} + r_{3k-2}, \\ r_{3k+1} &= r_{3k} + r_{3k-1} \end{aligned}$$

and deduce

$$r_{3k+1} = (4k+2)r_{3k-2} + r_{3k-5}.$$

We do the same for s_k and get

$$\begin{pmatrix} r_{3k+1} & r_{3k-2} \\ s_{3k+1} & s_{3k-2} \end{pmatrix} = \begin{pmatrix} r_{3k-2} & r_{3k-5} \\ s_{3k-2} & s_{3k-5} \end{pmatrix} \begin{pmatrix} 4k+2 & 1 \\ 1 & 0 \end{pmatrix}.$$

These are the same recurrence relations which are satisfied by p_k and q_k. Since

$$p_{-2} = 0, \quad p_{-1} = 1, \quad p_0 = 2, \qquad q_{-2} = 1, \quad q_{-1} = 0, \quad q_0 = 1$$

and

$$r_{-2} = 0 = 2q_{-1}, \quad r_1 = 2 = 2q_0, \qquad s_{-2} = 1 = p_{-1} - q_{-1}, \quad s_1 = 1 = p_0 - q_0,$$

we deduce $r_{3k+1} = 2q_k$ and $s_{3k+1} = p_k - q_k$ for all k. From $y = \lim_{k\to\infty} r_{3k}/s_{3k}$ we deduce $y = 2/(x-1)$. Since $x = (e+1)/(e-1)$, we get $y = e-1$.

The same argument starting from

$$\frac{e^2+1}{e^2-1} = [\overline{2j+1}]_{j\geq 0} = [1;\ 3,\ 5,\ 7, \dots],$$

yields Euler's continued fraction expansion for e^2 (see § 1.4)

$$e^2 = [7;\ \overline{3j-1,\ 1,\ 1,\ 3j,\ 12j+6}]_{j\geq 1} = [7;\ 2,\ 1,\ 1,\ 3,\ 18,\ 5,\ 1,\ 1,\ 6,\ 30,\ 8, \dots],$$

7.2 Hypergeometric functions

A (generalized) *hypergeometric series* is a power series

$$1 + \alpha_1 z + \alpha_2 z^2/2 + \cdots + \alpha_n z^n/n! + \cdots$$

such that there exists a rational fraction $A \in \mathbf{C}(T)$ satisfying, for all $n \geq 0$,

$$\alpha_{n+1} = \alpha_n A(n).$$

Write this rational fraction as

$$A(T) = c\frac{(a_1+T)\cdots(a_p+T)}{(b_1+T)\cdots(b_q+T)}.$$

We assume that A has no pole on $\mathbf{Z}_{\geq 0}$, which means $b_j \notin \mathbf{Z}_{\leq 0}$ for $1 \leq j \leq q$, so that $A(n)$ is defined for all $n \geq 0$. Then

$$\alpha_{n+1} = c\frac{(a_1+n)\cdots(a_p+n)}{(b_1+n)\cdots(b_q+n)}\alpha_n$$

and therefore

$$\alpha_n = c^n \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n},$$

where $(a)_n$ denotes the *Pochhammer symbol*

$$(a)_n = a(a+1)\cdots(a+n-1) \quad \text{for } n \geq 1 \text{ and } (a)_0 = 1.$$

It is also called *raising factorial*: notice that $(1)_n = n!$ and satisfies an number of relations, among which

$$(a)_{k+m} = (a)_k (a+k)_m.$$

For each $n \geq 0$, we have

$$\lim_{a\to\infty} \frac{(a)_n}{a^n} = 1$$

and for each $a \in \mathbf{C} \setminus \mathbf{Z}_{<0}$, we have

$$\lim_{n\to\infty} \frac{(a)_n}{n!} = 1.$$

For p and q non–negative integers, we define

$${}_pF_q\left(\begin{matrix} a_1 & a_2 & \cdots & a_p \\ b_1 & b_2 & \cdots & b_q \end{matrix}\middle| z\right) = \sum_{n\geq 0} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \cdot \frac{z^n}{n!}.$$

We shall use also the notation

$${}_pF_q\big(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots, b_q; z\big).$$

In the case where some a_i is in $\mathbf{Z}_{\leq 0}$, then ${}_pF_q$ is a polynomial. Otherwise, this power series has a radius of convergence which is infinite when $q \geq p$, finite if $q = p-1$, and 0 if $q < p-1$.

For $a_p = b_q = c$ we have

$${}_pF_q\left(\begin{matrix} a_1 & a_2 & \cdots & a_{p-1} & c \\ b_1 & b_2 & \cdots & b_{q-1} & c \end{matrix}\middle| z\right) = {}_{p-1}F_{q-1}\left(\begin{matrix} a_1 & a_2 & \cdots & a_{p-1} \\ b_1 & b_2 & \cdots & b_{q-1} \end{matrix}\middle| z\right)$$

Examples. The basic example is ${}_0F_0(z) = e^z$. Other examples are

$${}_1F_0\big(a; z\big) = \sum_{n\geq 0} \frac{a(a+1)\cdots(a+n-1)}{n!} \cdot z^n = (1-z)^{-a}$$

and

$$ {}_2F_1\big(1,1;2;z\big) = \sum_{n\geq 0} \frac{z^n}{n+1} = -\frac{1}{z}\log(1-z). $$

We consider the special case $p=0$, $q=1$ of Gauss hypergeometric series:

$$ {}_0F_1(c;z) = \sum_{n\geq 0} \frac{z^n}{(c)_n n!}. $$

We denote this function by $f(c;z)$.

Since

$$ \left(\frac{1}{2}\right)_n = \left(\frac{1}{2}\right)\left(\frac{1}{2}+1\right)\cdots\left(\frac{1}{2}+n-1\right) = \frac{(2n)!}{2^{2n}n!} $$

and

$$ \left(\frac{3}{2}\right)_n = \left(\frac{3}{2}\right)\left(\frac{3}{2}+1\right)\cdots\left(\frac{3}{2}+n-1\right) = \frac{(2n+1)!}{2^{2n}n!}, $$

special cases are

$$ f(1/2;z^2) = \sum_{n\geq 0} \frac{(2z)^{2n}}{(2n)!} = \cosh(2z) $$

and

$$ f(3/2;z^2) = \sum_{n\geq 0} \frac{(2z)^{2n}}{(2n+1)!} = \frac{1}{2z}\sinh(2z). $$

From

$$ \frac{1}{(c)_n} = \frac{c+n}{(c)_{n+1}} = \frac{1}{(c+1)_n} + \frac{n}{(c)_{n+1}} $$

one deduces

$$ f(c;z) = \sum_{n\geq 0} \frac{z^n}{(c+1)_n n!} + \sum_{n\geq 1} \frac{nz^n}{(c)_{n+1}n!}. $$

The first series is $f(c+1;z)$, the second is

$$ \sum_{n\geq 0} \frac{z^{n+1}}{(c)_{n+2}n!} = \frac{z}{c(c+1)} \sum_{n\geq 0} \frac{z^n}{(c+2)_n n!} = \frac{z}{c(c+1)} f(c+2;z). $$

This is the functional equation relating $f(c;z)$, $f(c+1;z)$ and $f(c+2;z)$:

$$ f(c;z) = f(c+1;z) + \frac{z}{c(c+1)} f(c+2;z). $$

Hence the function $g(c;z) = f(c;z)/f(c+1;z)$ satisfies

$$g(c,z) = 1 + \frac{z}{c(c+1)} \cdot \frac{1}{g(c+1;z)}.$$

Next define $h(c;z) = (c/z)g(c;z^2)$: we get

$$h(c;z) = \frac{c}{z} + \frac{1}{h(c+1;z)}.$$

Therefore, for $k \geq 1$,

$$h(c;z) = \left[\frac{c}{z}, \frac{c+1}{z}, \ldots, \frac{c+k-1}{z}, h(c+k;z)\right].$$

Replacing h by its value in terms of f yields

$$\frac{c}{z} \cdot \frac{f(c;z^2)}{f(c+1;z^2)} = \left[\frac{c}{z}, \frac{c+1}{z}, \ldots, \frac{c+k-1}{z}, \frac{(c+k)}{z} \cdot \frac{f(c+k;z^2)}{f(c+k+1;z^2)}\right].$$

We now take the limit on k:

Lemma 125. *For c and z positive real numbers, the infinite continued fraction converges and we have*

$$\frac{c}{z} \cdot \frac{f(c;z^2)}{f(c+1;z^2)} = \left[\frac{c}{z}, \frac{c+1}{z}, \ldots, \frac{c+k}{z}, \cdots\right].$$

Proof. We first check the following auxiliary result:

> *Let $(a_n)_{n\geq 0}$ be a sequence of real numbers, all ≥ 1. Let x be a real number. Assume that for all $n \geq 1$, there exists a real number $x_n \geq 1$ such that*
>
> $$x = [a_0, a_1, \ldots, a_{n-1}, x_n].$$
>
> *Then the infinite continued fraction $[a_0, a_1, \ldots, a_n, \ldots]$ converges to x.*

We already proved this result when the a_n are integers, the proof in the general case is the same: we write

$$[a_0, a_1, \ldots, a_n] = \frac{A_n}{B_n}$$

with $A_n = a_n A_{n-1} + A_{n-2}$, $B_n = a_n B_{n-1} + B_{n-2}$, so that

$$x = \frac{x_{n+1} A_n + A_{n-1}}{x_{n+1} B_n + B_{n-1}},$$

we note that $B_n \geq B_{n-1} + B_{n-2}$, which implies that B_n tends to infinity, and we conclude with the estimate

$$\left| x - \frac{A_n}{B_n} \right| = \frac{1}{B_n (x_{n+1} B_n + B_{n-1})} \leq \frac{1}{B_n^2}.$$

To complete the proof of Lemma 125, we notice that for c and z positive, we have

$$f(c+k+1; z^2) < f(c+k; z^2) \quad \text{and} \quad \frac{c+k}{z} \geq 1$$

for sufficiently large k.

□

In the special cases $c = 1/2$, this provides another proof of the continued fraction expansion from Proposition 122:

$$\frac{e^z - e^{-z}}{e^z + e^{-z}} = [0, 1/z, 3/z, \ldots, (2k-1)/z, \ldots] = \frac{z|}{|1} + \frac{z^2|}{|\,3} + \frac{z^2|}{|\,5} + \cdots + \frac{z^2\ \ |}{|2k-1} + \cdots$$

References

[1] P. Eymard and J.-P. Lafon, *The number π*, American Mathematical Society, Providence, RI, 2004. Translated from the 1999 French original by Stephen S. Wilson.

[2] C. Hermite, *Sur la fonction exponentielle*, C. R. Acad. Sci. Paris, 77 (1873), pp. 18–24, 74–79, 226–233, 285–293. Œuvres de Charles Hermite, Paris: Gauthier-Villars, (1905), III, 150–181. See also *Oeuvres* III, 127–130, 146–149, and *Correspondance Hermite-Stieltjes*, II, lettre 363, 291–295. University of Michigan Historical Math Collection `http://name.umdl.umich.edu/AAS7821.0001.001`.

[3] 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., 1767.

[4] S. Shirali, *Continued fraction for e.* Resonance, vol. **5** N°1, Jan. 2000, 14–28.
`http://www.ias.ac.in/resonance/`.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°13, *May 31, 2010*

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

$$m a_m = m! (z f')^{(m)}(0),$$

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 $0 \leq k \leq n_0$ we have

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

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

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

Hence we have proved:

Proposition 127 (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 $B \in \mathbf{Z}[z]$ 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 $n_1 \geq n_0$ 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 146 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

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

□

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°14, *June 2, 2010*

From (134) with $c = 1$ and Lemma 137 we deduce that 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:

Lemma 139. *The remainder $R(z)$ in Hermite's fomula with parameters n_0 and n_1 (and B monic) is given by*

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

An easy consequence of Lemma 139 is the estimate for the remainder term given in Lemma 128.

We now recover the explicit formulae for A and B which we derived in Proposition 127 in the context of Theorem 132.

When $S \in \mathbf{C}[[t]]$ 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 $\mathbf{C}[t]$ or f is a polynomial in $\mathbf{C}[z]$.

We reproduce [5], Chap.I § 1: for two power series S_1 and S_2 and an analytic function f we have

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

and

$$(S_1S_2)(D)f = S_1(D)\big(S_2(D)f\big).$$

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

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

and

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

For instance with $S(t) = 1 - t$ and $S^{-1}(t) = 1 + t + t^2 + \cdots$,

$$(1 - D)\sum_{n\geq 0} D^n f = \sum_{n\geq 0} D^n(1 - D)f = f.$$

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 $\lambda \in \mathbf{C}$ and $P \in \mathbf{C}[z]$, 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 $Q \in \mathbf{C}[z]$, the unique solution $P \in \mathbf{C}[z]$ to the differential equation

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

is

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

and this solution P is a polynomial of the same degree as 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 (133) in Theorem 132, where $B \in \mathbf{C}[z]$ is monic of degree n_1 and $A \in \mathbf{C}[z]$ has degree n_0, while $R \in \mathbf{C}[[z]]$ has a zero of multiplicity $N + 1$ at 0.

From

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

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

$$B(z) = (-1)^{n_1} \frac{n_1!}{n_0!} \sum_{k=0}^{n_1} (-1)^k \frac{(N-k)!}{(n_1-k)!k!} z^k. \tag{140}$$

One checks that B is monic of degree n_1. This formula matches with Proposition 127 and the duality (136) between (n_0, n_1) and (n_1, n_0).

We can also check the formula for A starting from

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

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

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

where a is the leading coefficient of A. From

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

we deduce

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

and

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

Hence the same computation as was done before for B will give the formula for A.

Thanks to these explicit formulae, we can express A and B in terms of hypergeometric series:

Lemma 141. *The numerator A_{n_0,n_1} and the denominator B_{n_0,n_1} of the Padé approximant of index (n_0, n_1) for the exponential function are given by hypergeometric polynomials*

$$A_{n_0,n_1}(z) = (-1)^{n_1} \frac{N!}{n_0!} \, {}_1F_1(-n_0; -N; z)$$

and

$$B_{n_0,n_1}(z) = (-1)^{n_1} \frac{N!}{n_0!} \, {}_1F_1(-n_1; -N; -z).$$

Proof. The proofs for both formulae are similar – in fact (136) shows that they are equivalent. Consider

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

and write

$$(-n_0)_k = (-1)^k \frac{n_0!}{(n_0-k)!} \quad \text{and} \quad (-N)_k = (-1)^k \frac{N!}{(N-k)!}.$$

Then

$$A_{n_0,n_1}(z) = (-1)^{n_1} \frac{N!}{n_0!} \sum_{k=0}^{n_0} \frac{(-n_0)_k}{(-N)_k k!} \cdot z^k = (-1)^{n_1} \frac{N!}{n_0!} \, {}_1F_1(-n_0; -N; z).$$

□

One can find the explicit values of these polynomials on the internet by looking for *Padé table for the exponential function*. Here is the table for B_{n_0,n_1} – the table for A_{n_0,n_1} is easy to deduce from (136).

n_0 \ n_1	0	1	2	3
0	1	$z-1$	z^2-2z+2	z^3-3z^2+6z-6
1	1	$z-2$	z^2-4z+6	$z^3-6z^2+18z-24$
2	1	$z-3$	$z^2-6z+12$	$z^3-9z^2+36z-60$
3	1	$z-4$	$z^2-8z+20$	$z^3-12z^2+60z-120$

These polynomials are also useful for giving continued fractions expressions for the exponential function.

8.3 Hermite's transcendence proof

In 1873 C. Hermite [3] 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 $\vartheta_1, \ldots, \vartheta_m$ 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 estimate

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

in terms of q, while type I is related with the study of estimates 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 $\max_{0\le i\le m} |a_i|$.

We explained Hermite's strategy in § 3.1: in order to apply the criterion for linear independence Proposition 14 and obtain the linear independence over $\mathbf{Q}$ of $1, e, e^2, \ldots$ (and therefore the transcendence of e), Hermite first "approximates" simultaneously the functions $e^z, e^{2z}, \ldots$ by rational fractions P_1/Q, P_m/Q, and then substitutes $z = 1$.

8.3.1 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 (for a complete historical survey of the theory, see [1]).

Lemma 142. *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 $\mathbf{C}[X]$ satisfying the following properties:
(i) The polynomial Q is not zero, it has degree $\leq N - n_0$.
(ii) For $1 \leq \mu \leq m$, the polynomial P_μ has degree $\leq N - n_\mu$.
(iii) For $1 \leq \mu \leq m$, the function $x \mapsto Q(x)f_\mu(x) - P_\mu(x)$ has a zero at the origin of multiplicity $\geq N + 1$.

Definition. *A tuple $(Q, P_1, \dots, P_m)$ of polynomials in $\mathbf{C}[X]$ satisfying the condition of Lemma 142 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 142 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 $\mathbf{C}$, and Lemma 142 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.

Here is the definition of the *Padé approximants of type I*:

Lemma 143. *Let $f_1, \dots, f_m$ be analytic functions of one complex variable near the origin. Let $d_0, d_1, \dots, d_m$ be non-negative integers. Set*

$$M = d_0 + d_1 + \cdots + d_m + m.$$

Then there exists a tuple $(A_0, \dots, A_m)$ of polynomials in $\mathbf{C}[X]$, not all of which are zero, where A_i has degree $\leq d_i$, such that the function

$$A_0 + A_1 f_1 + \cdots + A_m f_m$$

has a zero at the origin of multiplicity $\geq M$.

Definition. *A tuple $(A_0, A_1, \dots, A_m)$ of polynomials in $\mathbf{C}[X]$ satisfying the condition of Lemma 143 is called a* Padé system of the first type for $(f_1, \dots, f_m)$ attached to the parameters $n_0, n_1, \dots, n_m$.

Proof. The map from the product of linear spaces $\mathbf{C}[z]_{\leq n_0} \times \cdots \mathbf{C}[z]_{\leq n_m}$ to $\mathbf{C}^M$ which sends a tuple $(A_0, \dots, A_m)$ to

$$\big(D^j(A_0 + A_1 f_1 + \cdots + A_m f_m)(0)\big)_{0 \leq j < M}$$

is not injective, and any non–zero element in the kernel satisfies the required property. □

In the case $m = 1$, the notions of Padé approximants of type I and II coincide – and an explicit solution has been given in the previous courses when $f_1(x) = e^x$.

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 type II for the functions $(e^x, e^{2x}, \dots, e^{mx})$.

8.3.2 Hermite's identity

FromLemma 139 we deduce the value of the integral

$$\int_0^1 (1-t)^{n_0} t^{n_1} e^{tz} dt.$$

One can compute similar more general integrals, where $f(t) = (1-t)^{n_0} t^{n_1}$ is replaced by any polynomial. We start with a simple example.

Lemma 144. *Let f be a polynomial of degree $\leq N$. Define*

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

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

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

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 $-e^{-t} f(t) + e^{-t} D^{N+1} f(t)$. □

A change of variables in Lemma 144 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 144, 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 145. *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 145, 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 §8.1.1 (see Proposition 127), 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 145 is a powerful tool to go much further.

Proposition 146. *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 $f \in \mathbf{Z}[t]$ 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 $n_0, n_1, \ldots, n_m$.*

Remark. If one wants to compare the formulae of § 8.1 with the special case $m = 1$ of Proposition 146, one should be aware that we shifted somewhat the notations: in § 8.1 we worked with $f(t) = t^{n_1}(1-t)^{n_0}$, while in Proposition 146 with $m = 1$ the polynomial which occurs is $f(t) = t^{n_0}(t-1)^{n_1}$.

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 147. *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 147 it follows that for any polynomial $f \in \mathbf{Z}[X]$ and for any integers k and n with $n \geq k$, the polynomial $(1/k!)D^n f$ also belongs to $\mathbf{Z}[X]$. This completes the proof of Proposition 146.

□

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°15, *June 7, 2010*

We complete the proof of the transcendence of e, following Hermite.
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 Proposition 14, 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: we quote from p. 77 of [3]

Mais une autre voie conduira à une démonstration plus rigoureuse

The following approach is due to K. Mahler, we can view it as an extension of the simple non-vanishing argument used in § 8.1.3 for the irrationality of π.

We fix integers $n_0, \dots, n_1$, all ≥ 1. We set $N = n_0 + \cdots + n_m$. 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

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

We are going to use the previous results, but one should notice that the sum of the parameters on each row is now $N' = N - 1$, not N as before.

Proposition 148. *There exists a non-zero constant c such that the determinant*

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

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 $P_{\mu\mu}$, $1 \leq \mu \leq m$.

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 148. □

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 149. *There exists a constant $\kappa > 0$ independent on n such that*

$$\max_{1\le\mu\le m} \max_{0\le j\le m} |q_i e^{\mu} - p_{\mu j}| \le \frac{\kappa^n}{n!}.$$

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 146:

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

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

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

Finally the determinant in the statement of Proposition 149 is

$$\frac{\Delta(1)}{(n-1)!^{m+1}},$$

where Δ is the determinant of Proposition 148. 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 Proposition 14. 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 8. Using Hermite's method as explained in § 8.3, prove that for any non-zero $r \in \mathbf{Q}(i)$, the number e^r is transcendental.

Exercise 9. Let m be a positive integer and $\epsilon > 0$ a real number. Show that there exists $q_0 > 0$ such that, for any tuple $(q, p_1, \ldots, p_m)$ of rational integers with $q > q_0$,

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

Check that it is not possible to replace the exponent $1 + (1/m)$ by a smaller number.

`Hint.` Consider Hermite's proof of the transcendence of e (§ 8.3.2), especially Proposition 149. 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 § 8.3 with n satisfying

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

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*, vol. 12 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1991.
`http://www.emis.de/cgi-bin/MATH-item?0714.01001`.

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

[3] C. Hermite, *Sur la fonction exponentielle*, C. R. Acad. Sci. Paris, 77 (1873), pp. 18–24, 74–79, 226–233, 285–293. Œuvres de Charles Hermite, Paris: Gauthier-Villars, (1905), III, 150–181. See also *Oeuvres* III, 127–130, 146–149, and *Correspondance Hermite-Stieltjes*, II, lettre 363, 291–295. University of Michigan Historical Math Collection `http://name.umdl.umich.edu/AAS7821.0001.001`.

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

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

9 Interpolation

9.1 Weierstraß question

Weierstraß (see [3]) initiated the question of investigating the set of algebraic numbers where a given transcendental entire function f takes algebraic values.

Denote by $\overline{\mathbf{Q}}$ the *field of algebraic numbers* (algebraic closure of $\mathbf{Q}$ in $\mathbf{C}$). For an entire function f, we define the *exceptional set* S_f of f as the set of algebraic numbers α such that $f(\alpha)$ is algebraic:

$$S_f := \{\alpha \in \overline{\mathbf{Q}}\ ;\ f(\alpha) \in \overline{\mathbf{Q}}\}.$$

For instance, the Hermite–Lindemann's Theorem on the transcendence of $\log \alpha$ and e^β for α and β algebraic numbers is the fact that the exceptional set of the function e^z is $\{0\}$. Also, the exceptional set of $e^z + e^{1+z}$ is empty, by the Theorem of Lindemann–Weierstrass. The exceptional set of functions like 2^z or $e^{i\pi z}$ is $\mathbf{Q}$, as shown by the Theorem of Gel'fond and Schneider.

The exceptional set of a polynomial is $\overline{\mathbf{Q}}$ if the polynomial has algebraic coefficients, otherwise it is finite. Also, any finite set of algebraic numbers is the exceptional set of some entire function: for $s \geq 1$ the set $\{\alpha_1, \ldots, \alpha_s\}$ is the exceptional set of the polynomial $\pi(z-\alpha_1)\cdots(z-\alpha_s) \in \mathbf{C}[z]$ and also of the transcendental entire function $(z-\alpha_2)\cdots(z-\alpha_s)e^{z-\alpha_1}$. Assuming Schanuel's conjecture, further explicit examples of exceptional sets for entire functions can be produced, for instance $\mathbf{Z}_{\geq 0}$ or $\mathbf{Z}$.

The study of exceptional sets started in 1886 with a letter of Weierstrass to Strauss. This study was later developed by Strauss, Stäckel, Faber – see [3]. Further results are due to van der Poorten, Gramain, Surroca and others (see [1, 5]).

Among the results which were obtained, a typical one is the following: *if A is a countable subset of $\mathbf{C}$ and if E is a dense subset of $\mathbf{C}$, there exist transcendental entire functions f mapping A into E.*

Also, van der Poorten noticed in [4] that there are transcendental entire functions f such that $D^k f(\alpha) \in \mathbf{Q}(\alpha)$ for all $k \geq 0$ and all algebraic α.

The question of possible sets S_f has been solved in [2]: *any set of algebraic numbers is the exceptional set of some transcendental entire function.* Also multiplicities can be included, as follows: define the *exceptional set with multiplicity* of a transcendental entire function f as the subset of $(\alpha, t) \in \overline{\mathbf{Q}} \times \mathbf{Z}_{\geq 0}$ such that $f^{(t)}(\alpha) \in \overline{\mathbf{Q}}$. Here, $f^{(t)}$ stands for the t-th derivative of f, which we denote also by $D^t f$.

Then *any subset of* $\overline{\mathbf{Q}} \times \mathbf{Z}_{\geq 0}$ *is the exceptional set with multiplicities of some transcendental entire function* f. More generally, the main result of [2] is the following:

> *Let* A *be a countable subset of* $\mathbf{C}$. *For each pair* (α, s) *with* $\alpha \in A$, *and* $s \in \mathbf{Z}_{\geq 0}$, *let* $E_{\alpha,s}$ *be a dense subset of* $\mathbf{C}$. *Then there exists a transcendental entire function* f *such that*
>
> $$\left(\frac{d}{dz}\right)^s f(\alpha) \in E_{\alpha,s} \tag{150}$$
>
> *for all* $(\alpha, s) \in A \times \mathbf{Z}_{\geq 0}$.

One may replace $\mathbf{C}$ by $\mathbf{R}$: it means that one may take for the sets $E_{\alpha,s}$ dense subsets of $\mathbf{R}$, provided that one requires A to be a countable subset of $\mathbf{R}$.

The proof is a construction of an interpolation series on a sequence where each w occurs infinitely often. The coefficients of the interpolation series are selected recursively to be sufficiently small (and nonzero), so that the sum f of the series is a transcendental entire function.

This process yields uncountably many such functions. Further, one may also require that they are algebraically independent over $\mathbf{C}(z)$ together with their derivatives. Furthermore, at the same time, one may request further restrictions on each of these functions f. For instance, given any transcendental function g with $g(0) \neq 0$, one may require $|f|_R \leq |g|_R$ for all $R \geq 0$.

As a very special case of 150 (selecting A to be the set $\overline{\mathbf{Q}}$ of algebraic numbers and each $E_{\alpha,s}$ to be either $\overline{\mathbf{Q}}$ or its complement in $\mathbf{C}$), one deduces the existence of uncountably many algebraic independent transcendental entire functions f such that any Taylor coefficient at any algebraic point α takes a prescribed value, either algebraic or transcendental.

Exercise 10. . Check that a consequence of the main result (150) of [2] is the following.
Let A *be a countable subset of* $\mathbf{C}$. *For any non negative integer* s *and any* $\alpha \in A$, *let* $E_{\alpha s}$ *be a dense subset in* $\mathbf{C}$. *Let* g *be a transcendental entire function with* $g(0) \neq 0$. *Then there exists a set* $\{f_i \ i \in I\}$ *of entire functions, with* I *a set having the power of continuum, with the following properties.*

- *For any* $i \in I$, *any* $\alpha \in A$ *and any integer* $s \geq 0$, $f_i^{(s)}(\alpha) \in E_{\alpha s}$.
- *For any* $i \in I$ *and any real number* $r \geq 0$, $|f_i|_r \leq |g|_r$.
- *The functions* $f_i^{(s)}$, *(*$i \in I$, $s \geq 0$*) are algebraically independent over* $\mathbf{C}(z)$.

`Hint.` Use (150) with A replaced by $A \cup \{z_1, z_2\}$, where z_1, z_2 are two algebraically independent complex numbers which do not belong to A. For $s \geq 0$, set $E_{z_1,s} = \overline{\mathbf{Q}}$. If there is a non–trivial relation of algebraic dependence among some of the functions $f_i^{(s)}$, then there is such a relation with coefficients in $\overline{\mathbf{Q}}(z_1)$. Next select a set of numbers $x_{i,s}$, $i \in I$, $s \geq 0$, having the power of continuum, which are algebraically independent over $\mathbf{Q}(z_1, z_2)$ – it is easy to give explicit examples with Liouville numbers. To produce f_i, set $E_{z_2,s} = \overline{\mathbf{Q}} x_{i,s} \setminus \{0\}$.

References

[1] F. Gramain, *Fonctions entières arithmétiques*, in Séminaire d'analyse 1985–1986 (Clermont-Ferrand, 1985–1986), Univ. Clermont-Ferrand II, Clermont, 1986, pp. 9, Exp. No. 9.

[2] J. Huang, D. Marques, and M. Mereb, *Algebraic values of transcendental functions at algebraic points.* `http://arxiv.org/abs/0808.2766`, 2008.

[3] K. Mahler, *Lectures on transcendental numbers*, Springer-Verlag, Berlin, 1976. Lecture Notes in Mathematics, Vol. 546.

[4] A. J. Van der Poorten, *Transcendental entire functions mapping every algebraic number field into itself*, J. Austral. Math. Soc., 8 (1968), pp. 192—193.

[5] A. Surroca, *Valeurs algébriques de fonctions transcendantes*, Int. Math. Res. Not., Art. ID 16834 (2006), p. 31.

[6] M. Waldschmidt, *Auxiliary functions in transcendental number theory.*, Ramanujan J., 20 (2009), pp. 341–373.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°16, *June 9, 2010*

9.2 Integer valued entire functions

We have seen in § 9.1 that there is no hope to prove a general transcendence theorem on the values of entire functions. One needs to be less ambitious, and the most natural thing to do is to put restrictions on the functions. For instance the functions produced in § 9.1 with large exceptional sets do not satisfy differential equations (more precisely, as we have seen, it is possible to produce such functions which do not satisfy differential equations – it is another challenge to prove that none of them satisfies a differential equation!). We shall see, with the Schneider–Lang Theorem, that general transcendence results can be proved for functions satisfying some differential equations.

However, one of the earliest progresses in the theory came from adding restrictions not on the functions, but on the numbers. We were considering in § 9.1 algebraic values of transcendental functions at algebraic points. A much more restricted question is to investigate integer values at integral points. This is the story that we are telling now. We even start with a more specific topic by looking at zero values. Next we consider Pólya's pioneer work on integer valued entire functions, we pursue with Gel'fond's extension to Gaussian integers, and then with his proof of the transcendence of e^π.

When f is a complex function which is bounded on a disc $|z| \le r$, we set

$$|f|_r = \sup_{|z|=r} |f(z)|.$$

9.2.1 Weierstraß canonical products

Recall that if f is an analytic function on a simply connected open subset D of $\mathbf{C}$ without zero in D, then there exists a analytic function g in D such that $f = e^g$. If f has only finitely many zeros, then $f(z) = A(z)e^{g(z)}$, where A is a polynomial (having the same zeros as f) and g is analytic in D. We are interested in having a similar decomposition when f has infinitely many zeroes - recall that if f is not the zero function, then the zeroes are isolated.

We assume $D = \mathbf{C}$ (hence f is an *entire* function). Its zeros form a discrete set, one can order them by non–decreasing modulus: let $(\alpha_0, \alpha_1, \ldots)$ be this sequence of zeros of f, counting multiplicities. It will be convenient to assume $f(0) \neq 0$, hence $|\alpha_0| > 0$.

We further assume that f has finite order of growth ϱ, namely (cf [10] Chap. X § 3):

$$\varrho = \limsup_{r\to\infty} \frac{\log\log|f|_r}{\log r}.$$

Recall the Taylor expansion at the origin of $\log(1-z)$:

$$\log(1-z) = -z - \frac{z^2}{2} - \frac{z^3}{3} - \cdots \frac{z^m}{m} - \cdots$$

For $m \geq 0$, one defines the *Weierstraß factor* ([10] Chap. X § 2) as:

$$E(z,m) = (1-z)e^{z+z^2/2+z^3/3+\cdots+z^m/m};$$

in particular $E(z,0) = 1-z$. This function is very close to 1 (especially when m is large) for $|z|$ not too large: according to [10] Chap. X § 2 Lemma 2.2, for $|z| \leq 1/2$ one has $|\log E(z,m)| \leq 2|z|^{m+1}$.

A classical result (see [10] Chap. X § 3 Th. 3.5) is that there exist an integer $m \leq \varrho$ and a polynomial P of degree $\leq \varrho$ such that

$$f(z) = e^{P(z)} \prod_{n\geq 0} E(z/\alpha_n, m).$$

The integer m is the integral part of ϱ if ϱ is not an integer, it is ϱ or $\varrho - 1$ if ϱ is an integer.

Conversely, given a discrete sequence of non–zero complex numbers $(\alpha_n)_{n\geq 0}$, ordered with non–decreasing modulus, there exists a sequence of non–negative numbers $(m_n)_{n\geq 0}$ such that the product

$$\prod_{n\geq 0} E(z/\alpha_n, m_n)$$

is normally convergent over any compact subset of $\mathbf{C}$ (see [16] Chap. VII § 7.6 and [10] Chap. X § 2 Th. 2.3). When this property is true for a constant sequence $m_n = m$, $(n \geq 0)$, and when m is the smallest integer such that the product

$$\prod_{n\geq 0} E(z/\alpha_n, m)$$

is convergent, then this product is called the *canonical product of Weierstraß associated with the sequence* $(\alpha_n)_{n\geq 0}$.

Examples.

• (See [16] Chap. XII and [10] Chap. XII § 2).
The canonical product of Weierstraß associated with the non–negative integers $\mathbf{Z}_{\geq 0}$ is

$$z\prod_{n\geq 1}\left(1-\frac{z}{n}\right)e^{z/n}=-\frac{e^{\gamma z}}{\Gamma(-z)}\cdot$$

• (see [16] Chap. XII § 12.4 and [10] Chap. X § 2).
The canonical product of Weierstraß associated with the rational integers $\mathbf{Z}$, is

$$z\prod_{n\in\mathbf{Z}\setminus\{0\}}\left(1-\frac{z}{n}\right)e^{z/n}=\pi^{-1}\sin(\pi z)=\frac{-z}{\Gamma(z)\Gamma(1-z)}\cdot$$

• (see [16] Chap. XX and [10] Chap. XI § 4). and [1, 9, 14]. Let $\Omega=\mathbf{Z}\omega_1+\mathbf{Z}\omega_2$ be a lattice in $\mathbf{C}$. The *Weierstraß canonical product* attached to Ω is the *Weierstraß sigma function* σ_Ω defined by

$$\sigma_\Omega(z)=z\prod_{\omega\in\Omega\setminus\{0\}}\left(1-\frac{z}{\omega}\right)e^{\frac{z}{\omega}+\frac{z^2}{2\omega^2}}\cdot$$

Exercise 11. *Show that the function*

$$g(z)=\sum_{n\geq 0}(-1)^n\frac{\pi^{2n}}{2^{2n}(2n)!}z^n$$

has the infinite product expansion

$$g(z)=\prod_{n\in\mathbf{Z}}\left(1-\frac{z}{(2n+1)^2}\right)\cdot$$

`Hint:` *Check* $g(t^2)=\cos(\pi t/2)$.

An entire function f is said to be of finite exponential type if the number

$$\alpha=\limsup_{r\to\infty}\frac{\log|f|_r}{r}$$

is finite. In this case f is said to be of exponential type α. Notice that a function of finite exponential type has order ≤ 1; if the order is <1, then the type α is zero.

Lemma 151. *A function of exponential type < 1 which vanishes for all $n = 0, 1, \ldots$ is the zero function.*

The proof relies on the following auxiliary result:

Lemma 152 (Jensen's Formula)**.** *If g is an analytic function in an open set containing the closed disk disc $|z| \leq r$ with zeros $(a_j)_{1 \leq j \leq k}$ in this disc and if $g(0) \neq 0$, then*

$$\log |g(0)| + \sum_{j=1}^{k} \log \frac{r}{|a_j|} = \frac{1}{2\pi} \int_0^{2\pi} \log |g(re^{i\theta})| d\theta.$$

Sketch of proof of Jensen's Formula 152. Assume first that g has no zero in the closed disk disc $|z| \leq r$. Then there is an open disk containing this closed disk, where g has no zero, and therefore there is an analytic function h in a neighborhood of the disc $|z| \leq r$ such that $g = e^h$. Since $|g| = e^{\Re e h}$, the formula follows by taking the real part of

$$h(0) = \frac{1}{2i\pi} \int_{|z|=r} h(z) \frac{dz}{z} = \frac{1}{2\pi} \int_0^{2\pi} h(re^{i\theta}) d\theta.$$

In the general case, one can write $g(z) = (z - a_1) \cdots (z - a_k) e^{h(z)}$, where h is analytic. By multiplicativity of both sides of the conclusion of Lemma 152, the formula reduces to the following one: *for any complex number α,*

$$\int_0^1 \log |e^{2i\pi t} - \alpha| dt = \log \max\{1, |\alpha|\}.$$

(See for instance [11], pp. 5–6, or [10] Chap. IX Th. 1.3).

□

Proof of Lemma 151. Assume f is not the zero function and vanishes at all the non–negative integers $n = 0, 1, \ldots$ Since the zeroes of f are isolated, there exists $z_0 \in (0, 1)$ such that $f(z_0) \neq 0$. Use Jensen's formula 152 for the function $g(z) = f(z_0 + z)$ with $r = N - z_0$, where N is a large integer. The set of zeroes of g in the disc $|z| \leq r$ contains the elements $n - z_0$, $1 \leq n \leq N - z_0$. For $1 \leq n \leq N - z_0$ we have $(N - z_0)/(n - z_0) \geq N/n$. For the other zeros we use the trivial estimate $\log(r/|a_j|) \geq 0$. Also $|g|_r \leq |f|_N$. We deduce an upper bound of the right hand side of Jensen's Formula by using the assumption: there exists $c > 0$ and $\lambda < 1$ such that $|f|_N \leq ce^{\lambda N}$:

$$\frac{1}{2\pi} \int_0^{2\pi} \log |g(re^{i\theta})| d\theta \leq \log |g|_r \leq \log |f|_N \leq \lambda N + \log c.$$

Hence

$$\log|f(z_0)| \leq \lambda N - \sum_{n=1}^{N} \log(N/n) + \log c = \lambda N - N\log N + \log N! + \log c.$$

Since $\lambda < 1$, it follows from Stirling's formula:

$$N! \simeq N^N e^{-N}\sqrt{2\pi N} \tag{153}$$

that $\lambda N - N\log N + \log N!$ tends to $-\infty$ as N tends to infinity, which contradicts $f(z_0) \neq 0$. □

Remark on Jensen's Formula. In many situations, one can replace Jensen's formula (Lemma 152) by Schwarz's Lemma (see § 10.4), which gives an upper bound for $|f|_r$ when f has N zeroes (counting multiplicities) in $|z| \leq r$: for $R > r$ one has

$$|f|_r \leq \left(\frac{R^2 + r^2}{2rR}\right)^{-N} |f|_R. \tag{154}$$

However, here, it would give a weaker result: in order to reach the conclusion of Lemma 151, using (154), one needs to assume that f has exponential type $\leq \gamma$ where γ satisfies

$$\gamma < \sup_{\lambda>1} \frac{1}{\lambda}\log\left(\frac{\lambda^2+1}{2\lambda}\right) < \frac{1}{5}.$$

Remark on Stirling's Formula (153). We needed only a weak form of Stirling's formula. Asymptotic expansions (see the definition in Chap. VIII of [16]) for the logarithm of the Gamma function are known:

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \frac{1}{2}\log(2\pi) - \int_0^{+\infty} \frac{P_1(t)}{z+t}dt$$

for

$$-\pi + \delta < \arg z < \pi + \delta \quad \text{with} \quad 0 < \delta < \pi,$$

where $P_1(t) = t - \lfloor t \rfloor - 1/2$. Denote by $(B_n)_{n\geq 0}$ the sequence of Bernoulli numbers, which are defined by ([16] § 7.1)

$$\frac{x}{e^x - 1} = \sum_{n\geq 0} B_n \frac{x^n}{n!}.$$

The first non–zero values are

$$B_0 = 1,\ B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ B_4 = -\frac{1}{30},\ B_6 = \frac{1}{42},\ B_8 = -\frac{1}{30},\ B_{10} = \frac{5}{66}.$$

For z with argument $\leq (\pi/2) - \delta$ with $\delta > 0$, we have (see Chap. XII § 12.33 of [16]):

$$\log\Gamma(z) = \left(z - \frac{1}{2}\right)\log z - z + \frac{1}{2}\log(2\pi) + \frac{B_2}{1\cdot 2\cdot z} + \frac{B_4}{3\cdot 4\cdot z^2} + \frac{B_6}{5\cdot 6\cdot z^3} + \cdots$$

9.2.2 Pólya and 2^z

Satz I in [12] is the following result.

Theorem 155 (Pólya)**.** *If an entire function f satisfies $f(n) \in \mathbf{Z}$ for all $n = 0, 1, \ldots,$ and*

$$\lim_{r\to\infty} \frac{r^{1/2}|f|_r}{2^r} = 0,$$

then f is a polynomial.

A consequence of Pólya's Theorem 155 is that an entire function of exponential type $< \log 2$ is a polynomial. In loose terms, it means that the function 2^z is the transcendental function mapping $\mathbf{Z}_{\geq 0}$ to $\mathbf{Z}$ which grows the least rapidly.

In his 1929 paper [12], Pólya also considered entire functions mapping $\mathbf{Z}$ to $\mathbf{Z}$: he proved that the smallest such transcendental function is

$$\frac{1}{\sqrt{5}}\left(\left(\frac{3+\sqrt{5}}{2}\right)^z - \left(\frac{3-\sqrt{5}}{2}\right)^z\right).$$

After Pólya's work, a number of papers have been written on the subject. In particular Ch. Pisot used the Laplace–Borel transform to prove that an entire function mapping $\mathbf{Z}_{\geq 0}$ to $\mathbf{Z}$ of exponential type $\leq \log 2 = 0.69314718\ldots$ is of the form $A(z) + B(z)2^z$, where A and B are polynomials. See [6, 7].

Pólya's proof involves the calculus of finite differences [4] which we now introduce.

9.2.3 Calculus of finite differences

Given a function f and points $x_0, x_1, \ldots, x_m$ where f is analytic, one defines inductively analytic functions $f_1, f_2, \ldots$ as follows:

$$f_1(z) = \frac{f(z) - f(z_0)}{z - z_0}, \quad f_2(z) = \frac{f_1(z) - f_1(z_1)}{z - z_1}, \quad f_3(z) = \frac{f_2(z) - f_2(z_2)}{z - z_2}, \ldots$$

so that

$$\begin{aligned} f(z) &= f(z_0) + (z - z_0) f_1(z), \\ f_1(z) &= f_1(z_1) + (z - z_1) f_2(z), \\ f_2(z) &= f_2(z_2) + (z - z_2) f_3(z), \ldots \end{aligned}$$

This gives the expansion

$$\begin{aligned} f(z) = c_0 + c_1(z - z_0) + c_2(z - z_0)(z - z_1) + c_3(z - z_0)(z - z_1)(z - z_2) + \cdots \\ + c_m(z - z_0)(z - z_1) \cdots (z - z_{m-1}) + (z - z_0)(z - z_1) \cdots (z - z_m) f_{m+1}(z) \end{aligned}$$

with $c_0 = f(z_0)$, $c_1 = f_1(z_1)$, $\ldots$, $c_m = f_m(z_m)$.

Here is a first set of formulae for the coefficients $c_0, c_1, \ldots c_m$. For simplicity we assume that the points $x_0, x_1, \ldots, x_m$ are pairwise distinct. Define first

$$[x_0] = f(x_0), \ [x_1] = f(x_1), \ \ldots, [x_m] = f(x_m),$$

and next set

$$[x_0, x_1] = \frac{[x_0] - [x_1]}{x_0 - x_1}, \ [x_1, x_2] = \frac{[x_1] - [x_2]}{x_1 - x_2}, \ldots, [x_{m-1}, x_m] = \frac{[x_{m-1}] - [x_m]}{x_{m-1} - x_m},$$

$$[x_0, x_1, x_2] = \frac{[x_0, x_1] - [x_1, x_2]}{x_0 - x_2}, \ [x_1, x_2, x_3] = \frac{[x_1, x_2] - [x_2, x_3]}{x_1 - x_3}, \ldots,$$

and so on, up to

$$[x_0, x_1, \ldots, x_m] = \frac{[x_0, x_1, \ldots, x_{m-1}] - [x_1, x_2, \ldots, x_m]}{x_0 - x_m}.$$

Then

$$c_0 = [x_0], \quad c_1 = [x_0, x_1], \quad \ldots, \quad c_m = [x_0, \ldots, x_m].$$

We now explain another way of getting such an expansion, by means of an identity due to Ch. Hermite (see [13]):

$$\frac{1}{x - z} = \frac{1}{x - x_0} + \frac{z - x_0}{x - x_0} \cdot \frac{1}{x - z}.$$

We replace the last factor $1/(x-z)$ by repeating the same formula with x_0 replaced by x_1:

$$\frac{1}{x-z} = \frac{1}{x-x_0} + \frac{z-x_0}{x-x_0} \cdot \left(\frac{1}{x-x_1} + \frac{z-x_1}{x-x_1} \cdot \frac{1}{x-z} \right).$$

Inductively we deduce

$$\begin{aligned} \frac{1}{x-z} = & \sum_{j=0}^{m} \frac{(z-x_0)(z-x_1)\cdots(z-x_{j-1})}{(x-x_0)(x-x_1)\cdots(x-x_j)} \\ & + \frac{(z-x_0)(z-x_1)\cdots(z-x_m)}{(x-x_0)(x-x_1)\cdots(x-x_m)} \cdot \frac{1}{x-z}. \end{aligned}$$

Now we multiply by $(1/2i\pi)f(x)$ and integrate along a simple contour $\mathcal{C}$ which contains all the x_i as well as z: this produces *Newton interpolation expansion*

$$f(z) = \sum_{j=0}^{m} c_j (z-x_0)\cdots(z-x_{j-1}) + R_m(z)$$

with

$$c_j = \frac{1}{2i\pi} \int_{\mathcal{C}} \frac{f(x)dx}{(x-x_0)(x-x_1)\cdots(x-x_j)} \quad (0 \le j \le m-1)$$

and

$$R_m(z) = (z-x_0)(z-x_1)\cdots(z-x_m) \cdot \frac{1}{2i\pi} \int_{\mathcal{C}} \frac{f(x)dx}{(x-x_0)(x-x_1)\cdots(x-x_m)(x-z)}.$$

Similar formulae exist when the points x_i are not distinct: when one repeats m times the same x_i, one considers the values $f^{(s)}(x_i)$ of the successive derivatives of f at x_i, for $s = 0, \ldots, m-1$. See § 9.2.8 and [10] Chap. IX § 2.

9.2.4 Proof of Pólya's Theorem

Proof. The Newton's interpolation series introduced in § 9.2.3 associated with the function f and the points $x_j = j$ for $j \ge 0$ is the formal series

$$F(z) = \sum_{n \ge 0} c_n z(z-1)\cdots(z-n+1),$$

where, for $n \geq 0$,

$$c_n = \sum_{k=0}^{n} \frac{f(k)}{\prod\limits_{\substack{0 \leq j \leq n \\ j \neq k}} (k-j)}.$$

Since

$$\prod_{\substack{0 \leq j \leq n \\ j \neq k}} (k-j) = k(k-1)\cdots 2 \cdot 1 \cdot (-1)(-2)\cdots(k-n) = (-1)^{n-k} k!(n-k)!,$$

we deduce

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

Hence c_n is a rational number and more precisely $n!c_n$ is a rational integer. We are going to prove that c_n vanishes for sufficiently large n. In order to do so, we produce an upper bound for $|c_n|$ by using the hypothesis of Theorem 155, namely

$$|f|_r = \epsilon(r) r^{-1/2} 2^r$$

where $\epsilon(r) \to 0$ as $r \to \infty$. From the integral formula

$$c_n = \frac{1}{2i\pi} \int_{|z|=r_n} \frac{f(z)dz}{z(z-1)\cdots(z-n)}$$

which is valid for any $r_n > n$, we deduce

$$|c_n| \leq \epsilon(r_n) r_n^{-1/2} 2^{r_n} \frac{1}{(r_n-1)(r_n-2)\cdots(r_n-n)}.$$

The best choice [12] is $r_n = 2n$. Using Stirling's formula (153) we obtain

$$\begin{aligned} n!|c_n| &\leq \frac{\epsilon(2n)}{\sqrt{2n}} 2^{2n} \frac{n!(n-1)!}{(2n-1)!} \\ &= \frac{\epsilon(2n)}{\sqrt{2n}} 2^{2n+1} \frac{n!^2}{(2n)!} \\ &\sim \frac{\epsilon(2n)}{\sqrt{2n}} 2^{2n+1} \frac{(n^n e^{-n}\sqrt{2\pi n})^2}{(2n)^{2n} e^{-2n}\sqrt{4\pi n}} = \epsilon(2n)\sqrt{2\pi}. \end{aligned}$$

Hence $|c_n| < 1/n!$ for sufficiently large n, and therefore $c_n = 0$ for sufficiently large n, which means that the interpolation series F is a polynomial. Since $f - F$ vanishes for all $n = 0, 1, \ldots$ (by the construction of the interpolation series) and has exponential type $< \log 2 < 1$, it follows from Lemma 151 that $f = F$, hence f is a polynomial. □

Remark. In [12], Pólya explains his choice of $r_n = 2n$ by letting $r_n = n/\xi$ with $0 < \xi < 1$, and by performing all the details of the computation with ξ. He shows that the optimal value for ξ is obtained when the function $\xi^\xi(1-\xi)^{1-\xi}$ assumes its minimal value, which is at $\xi = 1/2$, and then he completes the proof with this choice.

9.2.5 Integer valued entire functions on Gaussian integers

In 1926, S. Fukasawa extended Pólya's result to the Gaussian integers: he proved that if f is an entire function mapping $\mathbf{Z}[i]$ to $\mathbf{Z}[i]$ and if, for any $\epsilon > 0$, there exists $\theta_\epsilon > 0$ such that

$$|f|_r \le e^{\theta_\epsilon r^{\sigma-\epsilon}} \quad \text{with} \quad \sigma = \frac{1440}{919 + 27\sqrt{5}} = 1.470\ldots,$$

then f is a polynomial. In 1929, A.O. Gel'fond [3] refined the result and obtained the right exponent 2 in place of $\sigma - \epsilon$: he proved that an entire function f mapping $\mathbf{Z}[i]$ to $\mathbf{Z}[i]$ and satisfying

$$|f|_r \le e^{\gamma r^2} \quad \text{with} \quad \gamma < \frac{\pi}{2(1+e^{164/\pi})^2} \simeq 0.7 \cdot 10^{-45}$$

is a polynomial.

The proofs by Fukasawa and Gel'fond rely on Newton's interpolation series at the points in $\mathbf{Z}[i]$.

That the exponent 2 cannot be improved is shown by the Weierstraß sigma function associated to $\mathbf{Z}[i]$. Gel'fond wrote that his estimate for the constant γ is not the right limit for the problem. In 1980, D.W. Masser showed that the result cannot hold with γ replaced by a constant larger than $\pi/(2e)$. In 1981, F. Gramain [5] proved that the result holds with $\pi/(2e)$, which is therefore best possible:

> *If f is an entire function which is not a polynomial and maps $\mathbf{Z}[i]$ to $\mathbf{Z}[i]$, then*
>
> $$\limsup_{r\to\infty} \frac{1}{r^2} \log |f|_r \ge \frac{\pi}{2e}\cdot$$

9.2.6 The constant of Gramain–Weber

The work by Masser and Gramain on entire functions mapping $\mathbf{Z}[i]$ to $\mathbf{Z}[i]$ gave rise to the following problem, which is still unsolved. For each integer

$k \geq 2$, let r_k be the minimal radius of a closed disk in $\mathbf{R}^2$ containing at least k points of $\mathbf{Z}^2$, and for $n \geq 2$ define

$$\delta_n = -\log n + \sum_{k=2}^{n} \frac{1}{\pi r_k^2}.$$

The limit $\delta = \lim_{n\to\infty} \delta_n$ exists (it is an analogue in dimension 2 of the Euler constant), and the best known estimates for it are [8]

$$1.811\cdots < \delta < 1.897\ldots$$

(see also [2]). F. Gramain conjectures that

$$\delta = 1 + \frac{4}{\pi}\big(\gamma L(1) + L'(1)\big),$$

where γ is Euler's constant and

$$L(s) = \sum_{n\geq 0} (-1)^n (2n+1)^{-s}$$

is the L function of the quadratic field $\mathbf{Q}(i)$ (Dirichlet beta function). Since $L(1) = \pi/4$ and

$$L'(1) = \sum_{n\geq 0} (-1)^{n+1} \cdot \frac{\log(2n+1)}{2n+1} = \frac{\pi}{4}\big(3\log\pi + 2\log 2 + \gamma - 4\log\Gamma(1/4)\big),$$

Gramain's conjecture is equivalent to

$$\delta = 1 + 3\log\pi + 2\log 2 + 2\gamma - 4\log\Gamma(1/4) = 1.822825\ldots$$

Other problems related to the lattice $\mathbf{Z}[i]$ are described in the section "*On the borders of geometry and arithmetic*" of [15].

References

[1] K. Chandrasekharan, *Elliptic functions*, vol. 281 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1985.

[2] S.R. Finch *Mathematical Constants*, Encyclopedia of Mathematics and its Applications, **94**. Cambridge University Press, Cambridge, 2003.

[3] A.O. Gel'fond, *Sur les propriétés arithmétiques des fonctions entières*, Tôhoku Math. Journ., **30** (1929), pp. 280–285.
http://www.journalarchive.jst.go.jp/jnlpdf.php?cdjournal=tmj1911&
cdvol=30& noissue=0& startpage=280& lang=en& from=jnltoc.

[4] ——, *Calculus of finite differences*, Hindustan Publishing Corp., Delhi, 1971. Translated from the Russian, International Monographs on Advanced Mathematics and Physics.

[5] F. Gramain *Sur le théorème de Fukasawa–Gel'fond*, Invent. Math. **63** N° 3 (1981), 495–506.
http://www.springerlink.com/content/j546xr85w52742w7/fulltext.pdf

[6] ——, *Fonctions entières arithmétiques*, in Séminaire Pierre Lelong-Henri Skoda (Analyse), Année 1976/77, vol. **694** of Lecture Notes in Math., Springer, Berlin, 1978, pp. 96–125.
http://www.springerlink.com/content/fulltext.pdf?id=doi:10.1007/BFb0063245.

[7] ——, *Fonctions entières arithmétiques*, in Séminaire Delange-Pisot-Poitou, 19e année: 1977/78, Théorie des nombres, Fasc. 1, Secrétariat Math., Paris, 1978, pp. Exp. No. 8, 14 p.
http://www.numdam.org/numdam-bin/fitem?id=SDPP_1977-1978__19_1_A6_0.

[8] F. Gramain and M. Weber, *Computing an arithmetic constant related to the ring of Gaussian integers*, Math. Comp. **44** N° 169 (1985), 241–250.
http://www.ams.org/journals/mcom/1985-44-169/S0025-5718-1985-0771043-3/S0025-5718-1985-0771043-3.pdf
Corrigendum: Math. Comp. **48** N° 178 (1987), 854.
http://www.ams.org/journals/mcom/1987-48-178/S0025-5718-1987-0878711-5/S0025-5718-1987-0878711-5.pdf

[9] S. Lang, *Elliptic curves: Diophantine analysis*, vol. **231** of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1978.

[10] ——, *Complex analysis*, vol. **103** of Graduate Texts in Mathematics, Springer-Verlag, New York, fourth ed., 1999.

[11] K. Mahler, *Lectures on transcendental numbers*, Springer-Verlag, Berlin, 1976. Lecture Notes in Mathematics, Vol. **546**.

[12] G. Pólya, *Über ganzwertige ganze Funktionen*, Rend. Circ. Mat. Palermo, **40** (1915), pp. 1–16.
http://springerlink.com/content/7q0r434816514656/fulltext.pdf

[13] T. Rivoal, *Applications arithmétiques de l'interpolation Lagrangienne*, Intern. J. Number Th., **5** (2009), pp. 185–208.
http://www.worldscinet.com/ijnt/05/preserved-docs/0502/S1793042109001992.pdf

[14] A. Robert, *Elliptic curves*, Lecture Notes in Mathematics, Vol. **326**, Springer-Verlag, Berlin, 1973. Notes from postgraduate lectures given in Lausanne 1971/72.

[15] W. Sierpiński, *A selection of problems in the theory of numbers*, Translated from the Polish by A. Sharma. Popular lectures in mathematics, **11**. A Pergamon Press Book The Macmillan Co., New York 1964

[16] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1996. An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, Reprint of the fourth (1927) edition.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°17, *June 14, 2010*

The work by Fukasawa on integer valued entire functions at the points of $\mathbf{Z}[i]$ requires estimates on the number of points of $\mathbf{Z}[i]$ into a disc. More generally, Fukasawa showed that if A is a domain bounded by finitely many curves of finite length, if we set

$$A = \int\int_{(D)} dxdy, \qquad B = \int\int_{(D)} \log\sqrt{x^2+y^2}dxdy,$$

then the number of points in $Dt \cap \mathbf{Z}[i]$ satisfies

$$At^2 \log t + Bt^2 + O(t \log t) \quad \text{as} \quad t \to \infty.$$

For the unit disc $D = \{z \in \mathbf{C} \ ; \ |z| \le 1\}$, one has $A = \pi$ and $B = -\pi/2$. One deduces

$$\log \prod_{\substack{0\neq\omega\in\mathbf{Z}[i]\\ |\omega|\le t}} |\omega| = \sum_{\substack{0\neq\omega\in\mathbf{Z}[i]\\ |\omega|\le t}} \log|\omega| = \pi r^2 \log r - \frac{\pi}{2} r^2 + o(r^2).$$

This yields

Lemma 156. *An entire function f satisfying $f(\mathbf{Z}[i]) = \{0\}$ and, for all sufficiently large r,*

$$|f|_r \le e^{\kappa r^2}$$

with $\kappa < \pi/2$, is a polynomial.

Proof. Like in the proof of Lemma 151, this follows from Jensen's formula, but here one replaces Stirling's formula by the estimates

$$\sum_{|\omega|\le r} 1 = \pi r^2 + o(r^2)$$

and

$$\sum_{\substack{0\neq\omega\in\mathbf{Z}[i]\\ |\omega|\le t}} \log(|\omega|/r) = \pi r^2 \log r - \frac{\pi}{2} r^2 - \pi r^2 \log r + o(r^2) = -\frac{\pi}{2} r^2 + o(r^2).$$

□

9.2.7 Transcendence of e^π

In [2], just after his paper [1] on integer valued entire functions on $\mathbf{Z}[i]$, A.O. Gel'fond extended his proof and obtained the following outstanding result:

Theorem 157 (Ge'lfond)**.** *The number*

$$e^\pi = 23,140\,692\,632\,779\,269\,005\,729\,086\,367\ldots$$

is transcendental.

This was the first step towards a solution of the seventh of the 23 problems raised by D. Hilbert at the International Congress of Mathematicians in Paris in 1900: *for algebraic α and β with $\alpha \neq 0$, $\alpha \neq 1$ and β irrational, the number α^β is transcendental.*

The number α^β is defined as $\alpha^\beta = \exp(\beta \log \alpha)$, where $\log \alpha$ is any logarithm of α. The condition $\alpha \neq 1$ may be replaced by $\log \alpha \neq 0$, both statements are equivalent.

Taking $\alpha = -1$, $\log \alpha = i\pi$, $\beta = -i$ gives $\alpha^\beta = e^\pi$.

Proof of Theorem 157. . Gel'fond starts by ordering $\mathbf{Z}[i]$ by non–decreasing modulus, and for those of the same modulus by increasing arguments in $[0, 2\pi)$:

$$\mathbf{Z}[i] = \{x_0, x_1, x_2, \ldots, x_n, \ldots\}$$

with $x_0 = 0$. Hence

$$\{x_0, x_1, x_2, \ldots\} = \{0,\ 1,\ i,\ -1,\ -i,\ 1+i,\ -1+i,\ -1-i,\ 2,\ 2i, \ldots\}.$$

If the disc $|z| \leq r_n$ contains the points x_i for $0 \leq i \leq n$, then the number $n+1$ of these points is

$$n + 1 = \pi r_n^2 + \alpha r_n + o(r_n)$$

with $\alpha < 2\sqrt{2}\pi$, hence $|x_n| = \sqrt{n/\pi} + o(\sqrt{n})$.

For $n \geq 1$, define $P_n(z) = z(z - x_1)\cdots(z - x_{n-1})$. Gel'fond expands the function $e^{\pi z}$ into a series of P_n:

$$e^{\pi z} = \sum_{k=0}^{n} A_k P_k(z) + R_n(z),$$

where, following 9.2.3,

$$A_k = \frac{1}{2i\pi} \int_{|\zeta|=n} \frac{e^{\pi\zeta} d\zeta}{P_{k+1}(\zeta)} \quad \text{and} \quad R_n(z) = \frac{P_{n+1}(z)}{2i\pi} \int_{|\zeta|=n} \frac{e^{\pi\zeta}}{P_{k+1}(\zeta)} \cdot \frac{d\zeta}{\zeta - z}.$$

Since the zeroes of P_{k+1} are simple, the residue formula gives, for $n \geq 0$,

$$A_n = \sum_{k=0}^{n} \frac{e^{\pi x_k}}{\omega_{n,k}}, \quad \text{with} \quad \omega_{n,k} = \prod_{\substack{0 \leq j \leq n \\ j \neq k}} (x_k - x_j).$$

The number $e^{\pi x_k}$ is $\pm e^{\pi \Re e(x_k)}$ and $\Re e(x_k)$ is a rational integer of absolute value $\leq \sqrt{n/\pi} + o(\sqrt{n})$. Hence A_n is a polynomial in e^{π} and $e^{-\pi}$ of degree $\leq \sqrt{n/\pi} + o(\sqrt{n})$ and coefficients in $\mathbf{Q}(i)$. The integral over the circle $|\zeta| = n$ yields the upper bound

$$|A_n| \leq \frac{e^{\pi n}}{\prod_{0 \leq j \leq n} (n - |x_j|)} \leq e^{-n \log n + \pi n + O(\sqrt{n})}.$$

In his previous work [1], Gel'fond proved that the least common multiple Ω_n of the numbers $\omega_{n,k}$ for $0 \leq k \leq n$ (which is also the least common denominator of the numbers $1/\omega_{n,k}$ for $0 \leq k \leq n$) satisfies

$$\Omega_n \leq e^{\frac{1}{2} n \log n + 163n + o(n)}.$$

The product $\Omega_n A_n$ is in $\mathbf{Z}[i][e^{\pi}, e^{-\pi}]$:

$$\Omega_n A_n = \sum_{k=0}^{n} B_{kn} e^{\pi x_k} \quad \text{with} \quad B_{kn} = \Omega_n / \omega_{n,k} \in \mathbf{Z}[i]$$

and

$$\max_{0 \leq k \leq n} |B_{kn}| \leq e^{\frac{1}{2} n \log n + 163n - \frac{1}{2} n \log n + 3\pi n + o(n)} \leq e^{173n + o(n)}.$$

Assuming e^{π} is algebraic, Liouville's inequality (Lemma 26) implies $A_n = 0$ for all sufficiently large n, and therefore the interpolation series

$$F(z) = \sum_{n \geq 0} A_n P_n(z)$$

is a polynomial. This polynomial F, by construction, takes the value $e^{\pi x_k}$ at $z = x_k$, which means that the entire function $e^{\pi z} - F(z)$ vanishes on $\mathbf{Z}[i]$. But this function has exponential type π, hence order 1, and Lemma 156 implies that this function is the zero function. This is a contradiction with the fact that $e^{\pi z}$ is a transcendental function.

□

9.2.8 Interpolation formulae

In the easiest case where there are no multiplicities, the interpolation problem is to find a function f taking given values at distinct points. When x_i and y_i are m given points ($0 \leq i \leq m-1$), with x_i pairwise distinct, there is a unique polynomial P of degree $< m$ satisfying $P(x_i) = y_i$ for $0 \leq i \leq m-1$. This polynomial is

$$f(z) = \sum_{j=0}^{m-1} y_j f_j(z),$$

where f_j is the solution of the same problem for the special case where $y_i = \delta_{ij}$ (Kronecker symbol, which is 1 for $i = j$ and 0 otherwise). Explicitly,

$$f_j(z) = \prod_{\substack{0 \leq i \leq m-1 \\ i \neq j}} \frac{z - x_i}{x_j - x_i}.$$

Similar formulae exist when the x_i may be repeated. As a simple example, if $x_i = x_0$ for $0 \leq i \leq m$, then the condition on f becomes $f^{(j)}(x_0) = y_j$ ($0 \leq j < m$), and the solution is given by the Taylor's expansion

$$f(z) = \sum_{j=0}^{m-1} y_j f_j(z) \quad \text{with} \quad f_j(z) = \frac{1}{j!}(z - x_0)^j.$$

In the very general case, one way to produce such formulae is to introduce integral formulae.

Let $Q(z)$ be a monic polynomial with roots $z_1, \ldots, z_n$, and for $1 \leq i \leq n$ let $m_i \geq 1$ be the multiplicity of z_i as a root of Q:

$$Q(z) = \prod_{i=1}^{n} (z - z_i)^{m_i}.$$

Let R be a real number with $R > \max_{1 \leq i \leq n} |z_i]$, so that the disc $|z| < R$ contains all points z_i. We denote by Γ the circle $|z| = R$. Further, for $1 \leq i \leq n$, let r_i be a real number in the range

$$0 < r_i < \min_{\substack{1 \leq k \leq n \\ k \neq i}} |z_i - z_k|.$$

We denote by Γ_i the circle $|z_i| \leq r_i$: it contains z_i, but no z_k for $k \neq i$. The following formula is due to Hermite: *for f analytic in an open domain*

containing the disc $|z| \leq R$ and for z in the open disc $|z| < R$ distinct from all z_i,

$$\frac{f(z)}{Q(z)} = \frac{1}{2i\pi}\int_\Gamma \frac{f(\zeta)}{Q(\zeta)}\cdot\frac{d\zeta}{\zeta - z} - \frac{1}{2i\pi}\sum_{i=1}^{n}\sum_{j=0}^{m_i-1}\frac{f^{(j)}(z_i)}{j!}\int_{\Gamma_i}\frac{(\zeta - z_i)^j}{Q(\zeta)}\cdot\frac{d\zeta}{\zeta - z}.$$

The proof is a simple application of the residue formula (see for instance [3] Chap. IX § 2): the first integral divided by $2i\pi$ is the sum of the residues of the function

$$\varphi(\zeta) = \frac{f(\zeta)}{Q(\zeta)}\cdot\frac{1}{\zeta - z}$$

at the poles in $|z| < R$. The pole $\zeta = z$ is simple, and the residue is $f(z)/Q(z)$, which gives the left hand side. Also, each sum

$$\sum_{j=0}^{m_i-1}\frac{f^{(j)}(z_i)}{j!}\int_{\Gamma_j}\frac{(\zeta - z_i)^j}{Q(\zeta)}\cdot\frac{d\zeta}{\zeta - z}$$

in the right hand side is $2i\pi$ times the residue at $\zeta = z_i$ of $\varphi(\zeta)$. Hence the formula drops out.

If f is a polynomial of degree $< M$ where $M = m_1 + \cdots + m_n$, then the first integral vanishes.

For $1 \leq i_0 \leq n$ and $0 \leq j_0 < m_i$, define the function $f_{i_0,j_0}(z)$ on the open set $|z - z_{i_0}| > r_{i_0}$ by

$$f_{i_0,j_0}(z) = -\frac{1}{j!}\cdot\frac{1}{2i\pi}Q(z)\int_{|\zeta - z_{i_0}| = r_{i_0}}\frac{(\zeta - z_{i_0})^{j_0}}{Q(\zeta)}\cdot\frac{d\zeta}{\zeta - z}.$$

Here, r_{i_0} is any number satisfying $0 < r_{i_0} < \min_{i \neq i_0}|z_i - z_{i_0}|$. Computing the integral by means of the residue Theorem shows that the integral extends to a meromorphic function in $\mathbf{C}$ with a single pole at $z = z_{i_0}$ of order $\leq m_i$. Also, letting $|z|$ tend to infinity shows that $f_{i_0,j_0}(z)$ is a polynomial of degree $< M$. Hence f_{i_0,j_0} is the unique polynomial of degree $< M$ satisfying

$$f^{(j)}_{i_0,j_0}(z_i) = \delta_{(i_0,j_0),(i,j)} \quad \text{where} \quad \delta_{(i_0,j_0),(i,j)} = \begin{cases} 1 & \text{if } i = i_0 \text{ and } j = j_0, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that, given distinct points $z_1, \ldots, z_n$, positive integers $m_1, \ldots, m_n$ and complex numbers y_{ij} $(1 \leq i \leq n,\ 0 \leq j \leq m_i - 1)$, there is a unique polynomial of degree $< M$, where $M = m_1 + \cdots + m_n$, satisfying the M conditions $f^{(j)}(z_i) = y_{ij}$ for $1 \leq i \leq n$ and $0 \leq j \leq m_i - 1$. This polynomial is given by

$$\sum_{i=1}^{n}\sum_{j=0}^{m_i-1} y_{ij} f_{ij}.$$

9.2.9 Rational interpolation

We just mention another kind of interpolation formula, which was introduced by René Lagrange in 1935, and used more recently by Tanguy Rivoal [4] for producing Diophantine results, including a new proof of Apéry's theorem on the irrationality of $\zeta(3)$.

One starts with the formula

$$\frac{1}{x-z} = \frac{\alpha-\beta}{(x-\alpha)(x-\beta)} + \frac{x-\beta}{x-\alpha}\cdot\frac{z-\alpha}{z-\beta}\cdot\frac{1}{x-z}\cdot$$

Iterating and integrating yields

$$f(z) = \sum_{n=0}^{N-1} B_n \frac{(z-\alpha_1)\cdots(z-\alpha_n)}{(z-\beta_1)\cdots(z-\beta_n)} + \tilde{R}_N(z).$$

This is an expansion of f into rational fractions, with given zeroes and poles.

References

[1] A.O. Gel'fond, *Sur les propriétés arithmétiques des fonctions entières*, Tôhoku Math. Journ., **30** (1929), pp. 280–285.
http://www.journalarchive.jst.go.jp

[2] ——, *Sur les nombres transcendants.*, C. R. 189, 1224-1226, (1929).
http://gallica.bnf.fr/ark:/12148/bpt6k3142j

[3] S. Lang, *Complex analysis*, vol. 103 of Graduate Texts in Mathematics, Springer-Verlag, New York, fourth ed., 1999.

[4] T. Rivoal, *Applications arithmétiques de l'interpolation Lagrangienne*, Intern. J. Number Th., 5 (2009), pp. 185–208.
http://www.worldscinet.com/ijnt/05/preserved-docs/0502/S1793042109001992.pdf

10 The Schneider–Lang Theorem

The Theorem of Schneider-Lang is a general statement dealing with values of meromorphic functions of one or several complex variables, satisfying differential equations.

The first general result dealing with analytic or meromorphic functions of one variable and containing the solution to Hilbert's seventh problem

appears in [4]. In fact one can deduce the transcendence of α^β (Gel'fond-Schneider Theorem 1.4) from this theorem, either by using the two functions z and α^z without derivatives (Schneider's method), or else e^z and $e^{\beta z}$ with derivatives (Gel'fond's method). The statement is rather complicated, and Th. Schneider made successful attempts to simplify it [5]. Schneider's criteria in [5], Chap. II, § 3, Th.12 and 13 deal only with Gel'fond's method, i.e. involve derivatives. Further simplifications have been introduced by S. Lang later: either for Schneider's method (see [1], Chap. III, § 1, Th.1), or else for Gel'fond's method and functions satisfying differential equations (see [1], Chap. III, § 1, Th.1 and [3], Appendix 1). This last result is known as the *Theorem of Schneider-Lang*.

10.1 Statement and first corollaries

`Content of the course:` Theorem of Schneider–Lang, corollaries: theorem of Hermite–Lindemann, Theorem of Gel'fond–Schneider.
Outline of the proof.
`References:` [6] (Chap. 3, § 3.7) and [7] (§ 2.2).
See also [5] (Chap. II, § 3, Th.12 and 13); [1] (Chap. III, § 1, Th.1); [3] (Appendix 1).
There is also a proof in [2] (Chap. IX § 3) for the special case where the number field is $\mathbf{Q}$: this allows to avoid any use of algebraic number theory.

References

[1] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966.

[2] ——, *Complex analysis*, vol. 103 of Graduate Texts in Mathematics, Springer-Verlag, New York, fourth ed., 1999.

[3] ——, *Algebra*, vol. 211 of Graduate Texts in Mathematics, Springer-Verlag, New York, third ed., 2002.

[4] T. Schneider, *Ein Satz über ganzwertige Funktionen als Prinzip für Transzendenzbeweise.*, Math. Ann., 121 (1949), pp. 131–140.
`http://www.springerlink.com/content/t4556743mv342614/fulltext.pdf`

[5] ——, *Einführung in die transzendenten Zahlen.* Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957. *Introduction aux nombres transcendants.* Traduit de l'allemand par P. Eymard. Gauthier-Villars, Paris 1959.

[6] M. Waldschmidt, *Nombres transcendants*, Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 402.
http://www.springerlink.com/content/110312/

[7] ——, *Transcendence methods*, vol. 52 of Queen's Papers in Pure and Applied Mathematics, Queen's University, Kingston, Ont., 1979.
http://www.math.jussieu.fr/ miw/articles/pdf/QueensPaper52.pdf

[8] ——, *Nombres transcendants et groupes algébriques*, Astérisque, (1987), p. 218. With appendices by Daniel Bertrand and Jean-Pierre Serre.

[9] ——, *Elliptic functions and transcendence.* Alladi, Krishnaswami (ed.), Surveys in number theory. New York, NY: Springer. Developments in Mathematics 17, 1-46 (2008)., 2008.
http://hal.archives-ouvertes.fr/hal-00407231/fr/

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°18, *June 16, 2010*

10.2 Siegel's Lemma

`References:` [2] Chap. 1 Lemme 1.3.1 and [3] § 1.2.

10.3 Liouville's inequality

`Reference:` [2] Chap. 1 § 1.2. See also Proposition 26.

10.4 Schwarz's Lemma

See (154).
`References:` [3] § 1.3 and [4] Chap. 7.

10.5 Differential equations

`Reference:` Lemma 2.2.5 of [3].

10.6 Proof of the Schneider–Lang Theorem

`Reference:` [2] Chap. 3. See also [1] Chap. III.

References

[1] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966.

[2] M. Waldschmidt, *Nombres transcendants*, Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 402.
`http://www.springerlink.com/content/110312/`

[3] ——, *Transcendence methods*, vol. 52 of Queen's Papers in Pure and Applied Mathematics, Queen's University, Kingston, Ont., 1979.
`http://www.math.jussieu.fr/ miw/articles/pdf/QueensPaper52.pdf`

[4] ——, *Nombres transcendants et groupes algébriques*, Astérisque, (1987), p. 218. With appendices by Daniel Bertrand and Jean-Pierre Serre.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°19, *June 21, 2010*

10.7 Elliptic functions

10.7.1 Introduction to elliptic functions

Among many references for this section are the books by Chandrasekharan [4], Chap. 1–6; by S. Lang [16], Chap. 1–6 and [14], § 1–6; by Alain Robert, [20], Chap I; by J. Silverman [23, 24], and by M. Hindry and J. Silverman [9].

The text below is taken from [29] § 2 and § 3.

An elliptic curve may be defined as

- $y^2 = C(x)$ for a squarefree cubic polynomial $C(x)$,
- a connected compact Lie group of dimension 1,
- a complex torus $\mathbf{C}/\Omega$ where Ω is a lattice in $\mathbf{C}$,
- a Riemann surface of genus 1,
- a non-singular cubic in $\mathbf{P}_2(\mathbf{C})$ (together with a point at infinity),
- an algebraic group of dimension 1, with underlying projective algebraic variety.

We shall use the Weierstraß form

$$E = \left\{(t : x : y)\ ;\ y^2t = 4x^3 - g_2xt^2 - g_3t^3\right\} \subset \mathbf{P}_2.$$

Here g_2 and g_3 are complex numbers, with the only assumption $g_2^3 \neq 27g_3^2$, which means that the discriminant of the polynomial $4X^3 - g_2X - g_3$ does not vanish.

An analytic parametrization of the complex points $E(\mathbf{C})$ of E is given by means of *the Weierstraß elliptic function* $\wp$, which satisfies the differential equation

$$\wp'^2 = 4\wp^3 - g_2\wp - g_3. \tag{158}$$

It has a double pole at the origin with principal part $1/z^2$ and also satisfies an addition formula

$$\wp(z_1+z_2) = -\wp(z_1) - \wp(z_2) + \frac{1}{4}\cdot\left(\frac{\wp'(z_1)-\wp'(z_2)}{\wp(z_1)-\wp(z_2)}\right)^2. \tag{159}$$

The exponential map of the Lie group $E(\mathbf{C})$ is

$$\begin{array}{rccc} \exp_E: & \mathbf{C} & \to & E(\mathbf{C}) \\ & z & \mapsto & \big(1:\wp(z):\wp'(z)\big). \end{array}$$

The kernel of this map is a *lattice* in $\mathbf{C}$ (that is a discrete rank 2 subgroup),

$$\Omega = \ker\exp_E = \{\omega\in\mathbf{C}\ ;\ \wp(z+\omega)=\wp(z)\} = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2.$$

Hence $\exp_E$ induces an isomorphism between the quotient additive group $\mathbf{C}/\Omega$ and $E(\mathbf{C})$ with the law given by (159). The elements of Ω are the *periods* of $\wp$. A pair (ω_1,ω_2) of fundamental periods is given by (cf. [30] § 20.32 Example 1)

$$\omega_i = 2\int_{e_i}^{\infty}\frac{dx}{\sqrt{4x^3-g_2x-g_3}}, \qquad (i=1,2),$$

where

$$4x^3-g_2x-g_3 = 4(x-e_1)(x-e_2)(x-e_3).$$

Indeed, since $\wp'$ is periodic and odd, it vanishes at $\omega_1/2$, $\omega_2/2$ and $(\omega_1+\omega_2)/2$, hence the values of $\wp$ at these points are the three distinct complex numbers e_1, e_2 and e_3 (recall that the discriminant of $4x^3-g_2x-g_3$ is not 0).

Conversely, given a lattice Ω, there is a unique Weierstraß elliptic function $\wp_\Omega$ whose period lattice is Ω (see § 10.7.5). We denote its invariants in the differential equation (158) by $g_2(\Omega)$ and $g_3(\Omega)$.

We shall be interested mainly (but not only) with elliptic curves which are defined over the field of algebraic numbers: they have a Weierstraß equation with algebraic g_2 and g_3. However we shall also use the Weierstraß elliptic function associated with the lattice $\lambda\Omega$ where $\lambda\in\mathbf{C}^\times$ may be transcendental; the relations are

$$\wp_{\lambda\Omega}(\lambda z) = \lambda^{-2}\wp_\Omega(z), \qquad g_2(\lambda\Omega)=\lambda^{-4}g_2(\Omega), \qquad g_3(\lambda\Omega)=\lambda^{-6}g_3(\Omega). \tag{160}$$

The lattice $\Omega = \mathbf{Z} + \mathbf{Z}\tau$, where τ is a complex number with positive imaginary part, satisfies

$$g_2(\mathbf{Z} + \mathbf{Z}\tau) = 60G_2(\tau) \quad \text{and} \quad g_3(\mathbf{Z} + \mathbf{Z}\tau) = 140G_3(\tau),$$

where, for $G_k(\tau)$ (with $k \geq 2$) are the Eisenstein series (see, for instance, [22] Chap. VII, § 2.3, [11] Chap. III § 2 or [23] Chap. VI § 3— the normalization in [31] p. 240 is different):

$$G_k(\tau) = \sum_{(m,n)\in\mathbf{Z}^2\setminus\{(0,0)\}} (m + n\tau)^{-2k}. \tag{161}$$

10.7.2 Morphisms between elliptic curves. The modular invariant

If Ω and Ω' are two lattices in $\mathbf{C}$ and if $f : \mathbf{C}/\Omega \to \mathbf{C}/\Omega'$ is an analytic homomorphism, then the map $\mathbf{C} \to \mathbf{C}/\Omega \to \mathbf{C}/\Omega'$ factors through a homothecy $\mathbf{C} \to \mathbf{C}$ given by some $\lambda \in \mathbf{C}$ such that $\lambda\Omega \subset \Omega'$:

$$\begin{array}{ccc} \mathbf{C} & \xrightarrow{\lambda} & \mathbf{C} \\ \downarrow & & \downarrow \\ \mathbf{C}/\Omega & \xrightarrow[f]{} & \mathbf{C}/\Omega' \end{array}$$

If $f \neq 0$, then $\lambda \in \mathbf{C}^\times$ and f is surjective.

Conversely, if there exists $\lambda \in \mathbf{C}$ such that $\lambda\Omega \subset \Omega'$, then $f_\lambda(x + \Omega) = \lambda x + \Omega'$ defines an analytic surjective homomorphism $f_\lambda : \mathbf{C}/\Omega \to \mathbf{C}/\Omega'$. In this case $\lambda\Omega$ is a subgroup of finite index in Ω', hence the kernel of f_λ is finite and there exists $\mu \in \mathbf{C}^\times$ with $\mu\Omega' \subset \Omega$: the two elliptic curves $\mathbf{C}/\Omega$ and $\mathbf{C}/\Omega'$ are *isogeneous*.

If Ω and Ω^* are two lattices, $\wp$ and $\wp^*$ the associated Weierstraß elliptic functions and g_2, g_3 the invariants of $\wp$, the following statements are equivalent:
(i) There is a 2×2 matrix with rational coefficients which maps a basis of Ω to a basis of Ω^*.
(ii) There exists $\lambda \in \mathbf{Q}^\times$ such that $\lambda\Omega \subset \Omega^*$.
(iii) There exists $\lambda \in \mathbf{Z} \setminus \{0\}$ such that $\lambda\Omega \subset \Omega^*$.
(iv) The two functions $\wp$ and $\wp^*$ are algebraically dependent over the field $\mathbf{Q}(g_2, g_3)$.
(v) The two functions $\wp$ and $\wp^*$ are algebraically dependent over $\mathbf{C}$.

The map f_λ is an isomorphism if and only if $\lambda\Omega = \Omega'$.

The number

$$j = \frac{1728g_2^3}{g_2^3 - 27g_3^2}$$

is the *modular invariant* of the elliptic curve E. Two elliptic curves over $\mathbf{C}$ are isomorphic if and only if they have the same modular invariant.

Set $\tau = \omega_2/\omega_1$, $q = e^{2\pi i\tau}$ and $J(e^{2\pi i\tau}) = j(\tau)$. Then

$$\begin{aligned} J(q) &= q^{-1}\left(1 + 240\sum_{m=1}^{\infty} m^3 \frac{q^m}{1-q^m}\right)^3 \prod_{n=1}^{\infty}(1-q^n)^{-24} \\ &= \frac{1}{q} + 744 + 196884\ q + 21493760\ q^2 + \cdots \end{aligned}$$

— see [19] § 4.12 or [22] Chap. VII § 3.3 and § 4.

10.7.3 Endomorphisms of an elliptic curve; complex multiplications

Let Ω be a lattice in $\mathbf{C}$. The set of analytic endomorphisms of $\mathbf{C}/\Omega$ is the subring

$$\mathrm{End}(\mathbf{C}/\Omega) = \left\{ f_\lambda\ ;\ \lambda \in \mathbf{C}\ \text{ with }\ \lambda\Omega \subset \Omega \right\}$$

of $\mathbf{C}$. We also call it the ring of endomorphisms of the associated elliptic curve, or of the corresponding Weierstraß $\wp$ function and we identify it with the subring

$$\left\{ \lambda \in \mathbf{C}\ \ ;\ \lambda\Omega \subset \Omega \right\}$$

of $\mathbf{C}$. The *field of endomorphisms* is the quotient field $\mathrm{End}(\mathbf{C}/\Omega) \otimes_{\mathbf{Z}} \mathbf{Q}$ of this ring.

If $\lambda \in \mathbf{C}$ satisfies $\lambda\Omega \subset \Omega$, then λ is either a rational integer or else an algebraic integer in an imaginary quadratic field. For such a λ, $\wp_\Omega(\lambda z)$ is a rational function of $\wp_\Omega(z)$; the degree of the numerator is λ^2 if $\lambda \in \mathbf{Z}$ and $N(\lambda)$ otherwise (here, N is the norm of the imaginary quadratic field); the degree of the denominator is $\lambda^2 - 1$ if $\lambda \in \mathbf{Z}$ and $N(\lambda) - 1$ otherwise.

Let E be the elliptic curve attached to the Weierstraß $\wp$ function. The ring of endomorphisms $\mathrm{End}(E)$ of E is either $\mathbf{Z}$ or else an order in an imaginary quadratic field k. The latter case arises if and only if the quotient $\tau = \omega_2/\omega_1$ of a pair of fundamental periods is a quadratic number; in this case the field of endomorphisms of E is $k = \mathbf{Q}(\tau)$ and the curve E has *complex multiplications* – this is the so-called *CM case*. This means also that the two functions $\wp(z)$ and $\wp(\tau z)$ are algebraically dependent. In this case, the value $j(\tau)$ of the modular invariant j is an algebraic integer whose degree is the class number of the quadratic field $k = \mathbf{Q}(\tau)$.

Remark. *From Gel'fond–Schneider Theorem (§ 10.1) one deduces the transcendence of the number*

$$e^{\pi\sqrt{163}} = 262\ 537\ 412\ 640\ 768\ 743.999\ 999\ 999\ 999\ 250\,072\,59\ldots$$

If we set

$$\tau = \frac{1+i\sqrt{163}}{2}, \quad q = e^{2\pi i\tau} = -e^{-\pi\sqrt{163}},$$

then the class number of the imaginary quadratic field $\mathbf{Q}(\tau)$ *is* 1, *we have* $j(\tau) = -(640\ 320)^3$ *and*

$$\left| j(\tau) - \frac{1}{q} - 744 \right| < 10^{-12}.$$

Also ([6] § 2.4)

$$\left(e^{\pi\sqrt{163}} - 744\right)^{1/3} = 640\,319.999\,999\,999\,999\,999\,999\,999\,999\,390\,31\ldots$$

Let $\wp$ be a Weierstraß elliptic function with field of endomorphisms k. Hence $k = \mathbf{Q}$ if the associated elliptic curve has no complex multiplication, while in the other case k is an imaginary quadratic field, namely $k = \mathbf{Q}(\tau)$, where τ is the quotient of two linearly independent periods of $\wp$. Let $u_1, \ldots, u_d$ be non-zero complex numbers. Then the functions $\wp(u_1 z), \ldots, \wp(u_d z)$ are algebraically independent (over $\mathbf{C}$ or over $\mathbf{Q}(g_2, g_3)$, this is equivalent) if and only if the numbers $u_1, \ldots, u_d$ are linearly independent over k. This generalizes the fact that $\wp(z)$ and $\wp(\tau z)$ are algebraically dependent if and only if the elliptic curve has complex multiplications. Much more general and deeper results of algebraic independence of functions (exponential and elliptic functions, zeta functions...) were proved by W.D. Brownawell and K.K. Kubota [3].

If $\wp$ is a Weierstraß elliptic function with algebraic invariants g_2 and g_3, if E is the associated elliptic curve and if k denotes its field of endomorphisms, then the set

$$\mathcal{L}_E = \Omega \cup \left\{ u \in \mathbf{C} \setminus \Omega \ ; \ \wp(u) \in \overline{\mathbf{Q}} \right\}$$

is a k-vector subspace of $\mathbf{C}$: this is the set of *elliptic logarithms of algebraic points on* E. It plays a role with respect to E similar to the role of $\mathcal{L}$ for the multiplicative group $\mathbf{G}_m$.

Let $k = \mathbf{Q}(\sqrt{-d})$ be an imaginary quadratic field with class number $h(-d) = h$. There are h non-isomorphic elliptic curves $E_1, \ldots, E_h$ with ring of endomorphisms the ring of integers of k. The numbers $j(E_i)$ are conjugate algebraic integers of degree h; each of them generates the Hilbert class field H of k (maximal unramified abelian extension of k). The Galois group of H/k is isomorphic to the ideal class group of k.

Since the group of roots of units of an imaginary quadratic field is $\{-1, +1\}$ except for $\mathbf{Q}(i)$ and $\mathbf{Q}(\varrho)$, where $\varrho = e^{2\pi i/3}$, it follows that there

are exactly two elliptic curves over $\mathbf{Q}$ (up to isomorphism) having an automorphism group bigger than $\{-1,+1\}$. They correspond to Weierstraß elliptic functions $\wp$ for which there exists a complex number $\lambda \neq \pm 1$ with $\lambda^2 \wp(\lambda z) = \wp(z)$.

The first one has $g_3 = 0$ and $j = 1728$. An explicit value for a pair of fundamental periods of the elliptic curve

$$y^2 t = 4x^3 - 4xt^2$$

follows from computations by Legendre using Gauss's lemniscate function ([30] § 22.8) and yields (see [1], as well as Appendix 1 of [28])

$$\omega_1 = \int_1^\infty \frac{dx}{\sqrt{x^3 - x}} = \frac{1}{2} B(1/4, 1/2) = \frac{\Gamma(1/4)^2}{2^{3/2}\pi^{1/2}} \quad \text{and} \quad \omega_2 = i\omega_1. \quad (162)$$

The lattice $\mathbf{Z}[i]$ has $g_2 = 4\omega_1^4$, thus

$$\sum_{(m,n)\in\mathbf{Z}^2\setminus\{(0,0)\}} (m+ni)^{-4} = \frac{\Gamma(1/4)^8}{2^6 \cdot 3 \cdot 5 \cdot \pi^2}.$$

The second one has $g_2 = 0$ and $j = 0$. Again from computations by Legendre ([30] § 22.81 II) one deduces that a pair of fundamental periods of the elliptic curve

$$y^2 t = 4x^3 - 4t^3$$

is (see once more [1] and Appendice 1 of [28])

$$\omega_1 = \int_1^\infty \frac{dx}{\sqrt{x^3 - 1}} = \frac{1}{3} B(1/6, 1/2) = \frac{\Gamma(1/3)^3}{2^{4/3}\pi} \quad \text{and} \quad \omega_2 = \varrho\omega_1. \quad (163)$$

The lattice $\mathbf{Z}[\varrho]$ has $g_3 = 4\omega_1^6$, thus

$$\sum_{(m,n)\in\mathbf{Z}^2\setminus\{(0,0)\}} (m+n\varrho)^{-6} = \frac{\Gamma(1/3)^{18}}{2^8 \cdot 5 \cdot 7 \cdot \pi^6}.$$

These two examples involve special values of Euler's Gamma function

$$\Gamma(z) = \int_0^\infty e^{-t} t^z \cdot \frac{dt}{t} = e^{-\gamma z} z^{-1} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^{-1} e^{z/n}, \quad (164)$$

where

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right) = 0.577\,215\,664\,901\,532\,860\,606\,512\,09\ldots$$

is Euler's constant (§ 12.1 in [30]), while Euler's Beta function is

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 x^{a-1}(1-x)^{b-1}dx.$$

More generally, the formula of Chowla and Selberg (1966) [5] (see also [2, 7, 8, 10, 12, 26] for related results) expresses periods of elliptic curves with complex multiplications as products of Gamma values: *if k is an imaginary quadratic field and $\mathcal{O}$ an order in k, if E is an elliptic curve with complex multiplications by $\mathcal{O}$, then the corresponding lattice Ω determines a vector space $\Omega \otimes_{\mathbf{Z}} \mathbf{Q}$ which is invariant under the action of k and thus has the form $k \cdot \omega$ for some $\omega \in \mathbf{C}^\times$ defined up to elements in $k^\times$. In particular, if $\mathcal{O}$ is the ring of integers $\mathbf{Z}_k$ of k, then*

$$\omega = \alpha\sqrt{\pi} \prod_{\substack{0<a<d \\ (a,d)=1}} \Gamma(a/d)^{w\epsilon(a)/4h},$$

where α is a non-zero algebraic number, w is the number of roots of unity in k, h is the class number of k, ϵ is the Dirichlet character modulo the discriminant d of k.

10.7.4 Standard relations among Gamma values

Euler's Gamma function satisfies the following relations ([30] Chap. XII):
(Translation)

$$\Gamma(z+1) = z\Gamma(z);$$

(Reflection)

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)};$$

(Multiplication) For any positive integer n,

$$\prod_{k=0}^{n-1} \Gamma\left(z + \frac{k}{n}\right) = (2\pi)^{(n-1)/2} n^{-nz+(1/2)} \Gamma(nz).$$

D. Rohrlich conjectured that any multiplicative relation among Gamma values is a consequence of these standard relations, while S. Lang was more optimistic (see [15], [17] I Chap. 2 Appendix p. 66 and [2] Chap. 24):

Conjecture 165 (D. Rohrlich)**.** *Any multiplicative relation*

$$\pi^{b/2} \prod_{a \in \mathbf{Q}} \Gamma(a)^{m_a} \in \overline{\mathbf{Q}}$$

with b and m_a in $\mathbf{Z}$ is a consequence of the standard relations.

Conjecture 166 (S. Lang)**.** *Any algebraic dependence relation with algebraic coefficients among the numbers $(2\pi)^{-1/2}\Gamma(a)$ with $a \in \mathbf{Q}$ is in the ideal generated by the standard relations.*

10.7.5 Quasi-periods of elliptic curves and elliptic integrals of the second kind

Let $\Omega = \mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$ be a lattice in $\mathbf{C}$. The *Weierstraß canonical product* attached to this lattice is the entire function σ_Ω defined by ([30] § 20.42)

$$\sigma_\Omega(z) = z \prod_{\omega \in \Omega \setminus \{0\}} \left(1 - \frac{z}{\omega}\right) e^{\frac{z}{\omega} + \frac{z^2}{2\omega^2}}.$$

It has a simple zero at any point of Ω.

Hence the Weierstraß sigma function plays, for the lattice Ω, the role which is played by the function

$$z \prod_{n \geq 1} \left(1 - \frac{z}{n}\right) e^{z/n} = -e^{\gamma z}\Gamma(-z)^{-1}$$

for the set of positive integers $\mathbf{N} \setminus \{0\} = \{1, 2, \dots\}$ (see the infinite product (164) for Euler's Gamma function), and also by the function

$$\pi^{-1} \sin(\pi z) = z \prod_{n \in \mathbf{Z} \setminus \{0\}} \left(1 - \frac{z}{n}\right) e^{z/n}$$

for the set $\mathbf{Z}$ of rational integers ([4] Chap. IV § 2).

The Weierstraß sigma function σ associated with a lattice in $\mathbf{C}$ is an entire function of *order* 2:

$$\limsup_{r \to \infty} \frac{1}{\log r} \cdot \log \log \sup_{|z|=r} |\sigma(z)| = 2;$$

the product $\sigma^2\wp$ is also an entire function of order 2 (this can be checked by using infinite products, but it is easier to use the quasi-periodicity of σ, see formula (167) below).

The logarithmic derivative of the sigma function is *the Weierstraß zeta function* $\zeta = \sigma'/\sigma$ whose Laurent expansion at the origin is

$$\zeta(z) = \frac{1}{z} - \sum_{k\geq 2} s_k z^{2k-1},$$

where, for $k \in \mathbf{Z}$, $k \geq 2$,

$$s_k = s_k(\Omega) = \sum_{\substack{\omega\in\Omega \\ \omega\neq 0}} \omega^{-2k} = \omega_1^{-2k} G_k(\tau)$$

The derivative of ζ is $-\wp$. From

$$\wp'' = 6\wp^2 - (g_2/2)$$

one deduces that $s_k(\Omega)$ is a homogenous polynomial in $\mathbf{Q}[g_2, g_3]$ of weight $2k$ for the graduation of $\mathbf{Q}[g_2, g_3]$ determined by assigning to g_2 the degree 4 and to g_3 the degree 6.

As a side remark, we notice that for any $u \in \mathbf{C} \setminus \Omega$ we have

$$\mathbf{Q}(g_2, g_3) \subset \mathbf{Q}\big(\wp(u), \wp'(u), \wp''(u)\big).$$

Since its derivative is periodic, the function ζ is *quasi-periodic*: for each $\omega \in \Omega$ there is a complex number $\eta = \eta(\omega)$ such that

$$\zeta(z + \omega) = \zeta(z) + \eta.$$

These numbers η are the *quasi-periods* of the elliptic curve. If (ω_1, ω_2) is a pair of fundamental periods and if we set $\eta_1 = \eta(\omega_1)$ and $\eta_2 = \eta(\omega_2)$, then, for $(a, b) \in \mathbf{Z}^2$,

$$\eta(a\omega_1 + b\omega_2) = a\eta_1 + b\eta_2.$$

Coming back to the sigma function, one deduces that

$$\sigma(z + \omega_i) = -\sigma(z) \exp\Big(\eta_i\big(z + (\omega_i/2)\big)\Big) \qquad (i = 1, 2). \tag{167}$$

The zeta function also satisfies an addition formula:

$$\zeta(z_1 + z_2) = \zeta(z_1) + \zeta(z_2) + \frac{1}{2} \cdot \frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)}.$$

The Legendre relation relating the periods and the quasi-periods

$$\omega_2\eta_1 - \omega_1\eta_2 = 2\pi i,$$

when ω_2/ω_1 has positive imaginary part, can be obtained by integrating $\zeta(z)$ along the boundary of a fundamental parallelogram.

In the case of complex multiplication, if τ is the quotient of a pair of fundamental periods of $\wp$, then the function $\zeta(\tau z)$ is algebraic over the field $\mathbf{Q}\big(g_2, g_3, z, \wp(z), \zeta(z)\big)$.

Examples For the curve $y^2t = 4x^3 - 4xt^2$ the quasi-periods attached to the pair of fundamental periods (162) are

$$\eta_1 = \frac{\pi}{\omega_1} = \frac{(2\pi)^{3/2}}{\Gamma(1/4)^2}, \qquad \eta_2 = -i\eta_1; \tag{168}$$

it follows that the fields $\mathbf{Q}(\omega_1, \omega_2, \eta_1, \eta_2)$ and $\mathbf{Q}\big(\pi, \Gamma(1/4)\big)$ have the same algebraic closure over $\mathbf{Q}$, hence the same transcendence degree. For the curve $y^2t = 4x^3 - 4t^3$ with periods (163), they are

$$\eta_1 = \frac{2\pi}{\sqrt{3}\omega_1} = \frac{2^{7/3}\pi^2}{3^{1/2}\Gamma(1/3)^3}, \qquad \eta_2 = \varrho^2\eta_1. \tag{169}$$

In this case the fields $\mathbf{Q}(\omega_1, \omega_2, \eta_1, \eta_2)$ and $\mathbf{Q}\big(\pi, \Gamma(1/3)\big)$ have the same algebraic closure over $\mathbf{Q}$, hence the same transcendence degree.

10.7.6 Elliptic integrals

Let

$$\mathcal{E} = \{(t : x : y) \in \mathbf{P}_2; y^2t = 4x^3 - g_2xt^2 - g_3t^3\}$$

be an elliptic curve. The field of rational (meromorphic) functions on $\mathcal{E}$ over $\mathbf{C}$ is $\mathbf{C}(\mathcal{E}) = \mathbf{C}(\wp, \wp') = \mathbf{C}(x, y)$ where x and y are related by the cubic equation $y^2 = 4x^3 - g_2x - g_3$. Under the isomorphism $\mathbf{C}/\Omega \to \mathcal{E}(\mathbf{C})$ given by $(1 : \wp : \wp')$, the differential form dz is mapped to dx/y. The holomorphic differential forms on $\mathbf{C}/\Omega$ are λdz with $\lambda \in \mathbf{C}$.

The differential form $d\zeta = \zeta'/\zeta$ is mapped to $-xdx/y$. The differential forms of second kind on $\mathcal{E}(\mathbf{C})$ are $adz + bd\zeta + d\chi$, where a and b are complex numbers and $\chi \in \mathbf{C}(x, y)$ is a meromorphic function on $\mathcal{E}$.

Assume that the elliptic curve $\mathcal{E}$ is defined over $\overline{\mathbf{Q}}$: the invariants g_2 and g_3 are algebraic. We shall be interested with differential forms which are defined over $\overline{\mathbf{Q}}$. Those of second kind are $adz + bd\zeta + d\chi$, where a and b are algebraic numbers and $\chi \in \overline{\mathbf{Q}}(x, y)$.

An elliptic integral is an integral

$$\int R(x, y)dx$$

where R is a rational function of x and y, while y^2 is a polynomial in x of degree 3 or 4 without multiple roots, with the proviso that the integral cannot be integrated by means of elementary functions. One may transform this integral as follows: one reduces it to an integral of $dx/\sqrt{P(x)}$ where P is a polynomial of 3rd or 4th degree; in case P has degree 4 one replaces it with a degree 3 polynomial by sending one root to infinity; finally one reduces it to a Weierstraß equation by means of a birational transformation. The value of the integral is not modified.

For transcendence purposes, if the initial differential form is defined over $\overline{\mathbf{Q}}$, then all these transformations involve only algebraic numbers.

10.7.7 Transcendence results of numbers related with elliptic functions

The main references for this section are [13, 21, 27, 29].

The first transcendence result on periods of elliptic functions was proved by C.L. Siegel as early as 1932.

Theorem 170 (Siegel, 1932)**.** *Let $\wp$ be a Weierstraß elliptic function with period lattice $\mathbf{Z}\omega_1 + \mathbf{Z}\omega_2$. Assume that the invariants g_2 and g_3 of $\wp$ are algebraic. Then at least one of the two numbers ω_1, ω_2 is transcendental.*

In the case of complex multiplication, it follows from Theorem 170 that *any non-zero period of $\wp$ is transcendental.*

From formulae (162) and (163) it follows as a consequence of Siegel's 1932 result that both numbers $\Gamma(1/4)^4/\pi$ and $\Gamma(1/3)^3/\pi$ are transcendental.

Other consequences of Siegel's result concern the transcendence of the length of an arc of an ellipse [21]

$$2\int_{-b}^{b} \sqrt{1 + \frac{a^2x^2}{b^4 - b^2x^2}}\, dx$$

for algebraic a and b, as well as the transcendence of an arc of the lemniscate $(x^2+y^2)^2 = 2a^2(x^2-y^2)$ with a algebraic.

A further example of application of Siegel's Theorem is the transcendence of values of hypergeometric series related with elliptic integrals

$$\begin{aligned} K(z) &= \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-z^2x^2)}} \\ &= \frac{\pi}{2} \cdot {}_2F_1\left(1/2,\ 1/2\ ;\ 1 \,\middle|\, z^2\right), \end{aligned}$$

where ${}_2F_1$ denotes Gauss hypergeometric series

$$ {}_2F_1\left(a,\ b\ ;\ c\ \middle|\ z\right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \cdot \frac{z^n}{n!} $$

with $(a)_n = a(a+1)\cdots(a+n-1)$.

Further results on this topic were obtained by Th. Schneider in 1934 and in a joint work by K. Mahler and J. Popken in 1935 using Siegel's method. These results were superseded by Th. Schneider's work in 1936 where he proved a number of definitive results on the subject, including:

Theorem 171 (Schneider, 1936)**.** *Assume that the invariants g_2 and g_3 of $\wp$ are algebraic. Then for any non-zero period ω of $\wp$, the numbers ω and $\eta(\omega)$ are transcendental.*

It follows from Theorem 171 that any non-zero period of an elliptic integral of the first or second kind is transcendental:

Corollary 172. *Let $\mathcal{E}$ be an elliptic curve over $\overline{\mathbf{Q}}$, p_1 and p_2 two algebraic points on $\mathcal{E}(\overline{\mathbf{Q}})$, w a differential form of first or second kind on $\mathcal{E}$ which is defined over $\overline{\mathbf{Q}}$, holomorphic at p_1 and p_2 and which is not the differential of a rational function. Let γ be a path on $\mathcal{E}$ from p_1 to p_2. In case $p_1 = p_2$ one assumes that γ is not homologous to* 0*. Then the number*

$$ \int_\gamma w $$

is transcendental.

Examples: Using Corollary 172 and formulae (168) and (169), one deduces that the numbers

$$ \Gamma(1/4)^4/\pi^3 \quad \text{and} \quad \Gamma(1/3)^3/\pi^2 $$

are transcendental.

The main results of Schneider's 1936 paper are as follows (see [21]):

Theorem 173 (Schneider, 1936)**. 1.** *Let $\wp$ be a Weierstraß elliptic function with algebraic invariants g_2, g_3. Let β be a non-zero algebraic number. Then β is not a pole of $\wp$ and $\wp(\beta)$ is transcendental.*
More generally, if a and b are two algebraic numbers with $(a,b) \neq (0,0)$, then for any $u \in \mathbf{C} \setminus \Omega$ at least one of the two numbers $\wp(u)$, $au + b\zeta(u)$ is transcendental.
2. *Let $\wp$ and $\wp^*$ be two algebraically independent elliptic functions with algebraic invariants g_2, g_3, g_2^*, g_3^*. If $t \in \mathbf{C}$ is not a pole of $\wp$ or of $\wp^*$, then*

at least one of the two numbers $\wp(t)$ and $\wp^*(t)$ is transcendental.
3. *Let $\wp$ be a Weierstraß elliptic function with algebraic invariants g_2, g_3. Then for any $t \in \mathbf{C} \setminus \Omega$, at least one of the two numbers $\wp(t)$, e^t is transcendental.*

It follows from Theorem 173.2 that the quotient of an elliptic integral of the first kind (between algebraic points) by a non-zero period is either in the field of endomorphisms (hence a rational number, or a quadratic number in the field of complex multiplications), or a transcendental number.

Here is another important consequence of Theorem 173.2.

Corollary 174 (Schneider, 1936)**.** *Let $\tau \in \mathcal{H}$ be a complex number in the upper half plane $\Im m(\tau) > 0$ such that $j(\tau)$ is algebraic. Then τ is algebraic if and only if τ is imaginary quadratic.*

In this connection we quote Schneider's second problem in [21], which is still open (see papers by Wakabayashi

Conjecture 175 (Schneiders' second problem)**.** *Prove Corollary 174 without using elliptic functions.*

Sketch of proof of Corollary 174 as a consequence of part 2 of Theorem 173.

Assume that both $\tau \in \mathcal{H}$ and $j(\tau)$ are algebraic. There exists an elliptic function with algebraic invariants g_2, g_3 and periods ω_1, ω_2 such that

$$\tau = \frac{\omega_2}{\omega_1} \quad \text{and} \quad j(\tau) = \frac{1728 g_2^3}{g_2^3 - 27 g_3^2}.$$

Set $\wp^*(z) = \tau^2 \wp(\tau z)$. Then $\wp^*$ is a Weierstraß function with algebraic invariants g_2^*, g_3^*. For $u = \omega_1/2$ the two numbers $\wp(u)$ and $\wp^*(u)$ are algebraic. Hence the two functions $\wp(z)$ and $\wp^*(z)$ are algebraically dependent. It follows that the corresponding elliptic curve has non-trivial endomorphisms, therefore τ is quadratic. □

A quantitative refinement of Schneider's Theorem on the transcendence of $j(\tau)$ given by A. Faisant and G. Philibert in 1984 became useful 10 years later in connection with Nesterenko's result. (see § 11.3).

We will not review the results related with abelian integrals, but only quote the first result on this topic, which involves the Jacobian of a Fermat curve: in 1941 Schneider proved that *for a and b in $\mathbf{Q}$ with a, b and $a+b$ not in $\mathbf{Z}$, the number*

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

is transcendental. We notice that in his 1932 paper, C.L. Siegel had already announced partial results on the values of the Euler Gamma function.

Schneider's above mentioned results deal with elliptic (and abelian) integrals of the first or second kind. His method can be extended to deal with elliptic (and abelian) integrals of the third kind (this is Schneider's third problem in [21]).

As pointed out by J-P. Serre in 1979, it follows from the quasi-periodicity of the Weierstraß sigma function (167) that the function

$$F_u(z) = \frac{\sigma(z+u)}{\sigma(z)\sigma(u)} e^{-z\zeta(u)}$$

satisfies

$$F_u(z+\omega_i) = F_u(z) e^{\eta_i u - \omega_i \zeta(u)}.$$

Theorem 176. *Let u_1 and u_2 be two non-zero complex numbers. Assume that g_2, g_3, $\wp(u_1)$, $\wp(u_2)$, β are algebraic and $\mathbf{Z}u_1 \cap \Omega = \{0\}$. Then the number*

$$\frac{\sigma(u_1+u_2)}{\sigma(u_1)\sigma(u_2)} e^{\big(\beta - \zeta(u_1)\big)u_2}$$

is transcendental.

From the next corollary, one can deduce that non-zero periods of elliptic integrals of the third kind are transcendental.

Corollary 177. *For any non-zero period ω and for any $u \in \mathbf{C} \setminus \Omega$ the number $e^{\omega\zeta(u) - \eta u + \beta\omega}$ is transcendental.*

References

[1] M. Abramowitz & I. A. Stegun – *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, A Wiley-Interscience Publication, New York: John Wiley & Sons, Inc; Washington, D.C, 1984, Reprint of the 1972 ed.

[2] Y. André – *Une introduction aux motifs (motifs purs, motifs mixtes, périodes)*, Panoramas et Synthèses, vol. 17, Société Mathématique de France, Paris, 2004.

[3] W. D. Brownawell & K. K. Kubota – "The algebraic independence of Weierstrass functions and some related numbers", *Acta Arith.* **33** (1977), no. 2, p. 111–149.

[4] K. Chandrasekharan, *Elliptic functions*, vol. 281 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1985.

[5] S. Chowla & A. Selberg – "On Epstein's zeta-function", *J. reine angew. Math.* **227** (1967), p. 86–110.

[6] H. Cohen – "Elliptic curves", in *From Number Theory to Physics (Les Houches, 1989)*, Springer, Berlin, 1992, p. 212–237.

[7] B. H. Gross – "On the periods of abelian integrals and a formula of Chowla and Selberg", *Invent. Math.* **45** (1978), no. 2, p. 193–211, With an appendix by David E. Rohrlich.

[8] — , "On an identity of Chowla and Selberg", *J. Number Theory* **11** (1979), no. 3 S. Chowla Anniversary Issue, p. 344–348.

[9] M. Hindry & J. H. Silverman – *Diophantine geometry*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000, An introduction.

[10] N. Koblitz – "Gamma function identities and elliptic differentials on Fermat curves", *Duke Math. J.* **45** (1978), no. 1, p. 87–99.

[11] — , *Introduction to elliptic curves and modular forms*, second ed., Graduate Texts in Mathematics, vol. 97, Springer-Verlag, New York, 1993.

[12] N. Koblitz & D. Rohrlich – "Simple factors in the Jacobian of a Fermat curve", *Canad. J. Math.* **30** (1978), no. 6, p. 1183–1205.

[13] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966.

[14] — , *Elliptic curves: Diophantine analysis*, vol. **231** of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1978.

[15] — , "Relations de distributions et exemples classiques", in *Séminaire Delange-Pisot-Poitou, 19e année: 1977/78, Théorie des nombres, Fasc. 2*, Secrétariat Math., Paris, 1978, p. Exp. No. 40, 6 (= [18] p. 59–65).

[16] — , *Elliptic functions*, second ed., Graduate Texts in Mathematics, vol. 112, Springer-Verlag, New York, 1987, With an appendix by J. Tate.

[17] — , *Cyclotomic fields I and II*, second ed., Graduate Texts in Mathematics, vol. 121, Springer-Verlag, New York, 1990, With an appendix by Karl Rubin.

[18] — , *Collected papers. Vol. III*, Springer-Verlag, New York, 2000, 1978–1990.

[19] Y. Manin – "Cyclotomic fields and modular curves.", *Uspekhi Mat. Nauk* **26** (1971), no. 6, p. 7–71, Engl. Transl. Russ. Math. Surv. **26** (1971), No 6, 7-78.

[20] A. Robert, *Elliptic curves*, Lecture Notes in Mathematics, Vol. **326**, Springer-Verlag, Berlin, 1973. Notes from postgraduate lectures given in Lausanne 1971/72.

[21] T. Schneider, *Einführung in die transzendenten Zahlen.* Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957. *Introduction aux nombres transcendants.* Traduit de l'allemand par P. Eymard. Gauthier-Villars, Paris 1959.

[22] J.-P. Serre – *Cours d'arithmétique*, Collection SUP: "Le Mathématicien", vol. 2, Presses Universitaires de France, Paris, 1970, reprinted 1977. Engl. transl.: *A course in arithmetic*, Graduate Texts in Mathematics, Vol. 7. Springer-Verlag, New York, 1978.

[23] J. H. Silverman – *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1986.

[24] — , *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 151, Springer-Verlag, New York, 1994.

[25] M. Waldschmidt, *Nombres transcendants*, Springer-Verlag, Berlin, 1974. Lecture Notes in Mathematics, Vol. 402.
`http://www.springerlink.com/content/110312/`

[26] — , "Diophantine properties of the periods of the Fermat curve.", in *Number theory related to Fermat's last theorem*, Proc. Conf., Prog. Math. 26, 79-88 , 1982.

[27] — , *Nombres transcendants et groupes algébriques*, Astérisque, (1987), p. 218. With appendices by Daniel Bertrand and Jean-Pierre Serre.

[28] — , "Transcendance et indépendance algébrique de valeurs de fonctions modulaires", in *Number theory (Ottawa, ON, 1996)*, CRM Proc. Lecture Notes, vol. 19, Amer. Math. Soc., Providence, RI, 1999, p. 353–375.

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

[30] E. Whittaker & G. Watson – *A course of modern analysis. An introduction to the general theory on infinite processes and of analytic functions; with an account of the principal transcendental functions. 4th ed., reprinted*, Cambridge: At the University Press. 608 p. , 1962.

[31] D. Zagier – "Introduction to modular forms", in *From Number Theory to Physics (Les Houches, 1989)*, Springer, Berlin, 1992, p. 238–291.

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°20, *June 23, 2010*

Content of the course:

1. Algebraic independence of the two functions $\wp(z)$ and e^z.
Legendre's relation $\eta_2\omega_1 - \eta_1\omega_2 = 2i\pi$. Proof: integrate $\zeta(z)dz$ on a fundamental parallelogram.
Application: algebraic independence of the two functions $az + b\zeta(z)$ and $\wp(z)$.

2. Section § 10.7.2: Morphisms between elliptic curves. The modular invariant.

3. Section § 10.7.3: Endomorphisms of an elliptic curve; complex multiplications.
Algebraic independence of $\wp$ and $\wp^*$.
Schneider's Theorem on the transcendence of $j(\tau)$ (corollary 174).

11 Algebraic independence

11.1 Chudnovskii's results

`References:` [1], [3], Lecture 8. [5] § 5.2.

The text below is taken from [5] § 5.2.

In the 1970's G.V. Chudnovsky proved strong results of algebraic independence (small transcendence degree) related with elliptic functions. One of his most spectacular contributions was obtained in 1976:

Theorem 178 (G.V. Chudnovsky, 1976)**.** *Let $\wp$ be a Weierstraß elliptic function with invariants g_2, g_3. Let (ω_1, ω_2) be a basis of the lattice period of $\wp$ and $\eta_1 = \eta(\omega_1)$, $\eta_2 = \eta(\omega_2)$ the associated quasi-periods of the associated Weierstraß zeta function. Then at least two of the numbers g_2, g_3, ω_1, ω_2, η_1, η_2 are algebraically independent.*

A more precise result is that, for any non-zero period ω, at least two of the four numbers $g_2,\ g_3,\ \omega/\pi,\ \eta/\omega$ (with $\eta = \eta(\omega)$) are algebraically independent.

In the case where g_2 and g_3 are algebraic one deduces from Theorem 178 that two among the four numbers $\omega_1,\ \omega_2,\ \eta_1,\ \eta_2$ are algebraically independent; this statement is also a consequence of the next result:

Theorem 179 (G.V. Chudnovsky, 1981)**.** *Assume that g_2 and g_3 are algebraic. Let ω be a non-zero period of $\wp$, set $\eta = \eta(\omega)$ and let u be a complex number which is not a period such that u and ω are $\mathbf{Q}$-linearly independent: $u \notin \mathbf{Q}\omega \cup \Omega$. Assume $\wp(u) \in \overline{\mathbf{Q}}$. Then the two numbers*

$$\zeta(u) - \frac{\eta}{\omega}u, \quad \frac{\eta}{\omega}$$

are algebraically independent.

From Theorem 178 or Theorem 179 one deduces:

Corollary 180. *Let ω be a non-zero period of $\wp$ and $\eta = \eta(\omega)$. If g_2 and g_3 are algebraic, then the two numbers π/ω and η/ω are algebraically independent.*

The following consequence of Corollary 180 shows that in the CM case, Chudnovsky's results are sharp:

Corollary 181. *Assume that g_2 and g_3 are algebraic and the elliptic curve has complex multiplications. Let ω be a non-zero period of $\wp$. Then the two numbers ω and π are algebraically independent.*

As a consequence of formulae (162) and (163), one deduces:

Corollary 182. *The numbers π and $\Gamma(1/4)$ are algebraically independent. Also the numbers π and $\Gamma(1/3)$ are algebraically independent.*

References

[1] G. V. Chudnovsky –"Algebraic independence of values of exponential and elliptic functions", in *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)* (Helsinki), Acad. Sci. Fennica, 1980, p. 339–350.

[2] M. Waldschmidt, *Les travaux de G. V. Čudnovskiĭ sur les nombres transcendants*, in Séminaire Bourbaki, Vol. 1975/76, 28e année, Exp. No. 488, Springer, Berlin, 1977, pp. 274–292. Lecture Notes in Math., Vol. 567.
http://archive.numdam.org/article/SB_1975-1976__18__274_0.pdf

[3] — , *Transcendence methods*, vol. 52 of Queen's Papers in Pure and Applied Mathematics, Queen's University, Kingston, Ont., 1979.
http://www.math.jussieu.fr/ miw/articles/pdf/QueensPaper52.pdf

[4] — , *Elliptic curves and complex multiplication* English translation by Franz Lemmermeyer of notes by A. Faisant, R. Lardon and G. Philibert, Sém Arithm. Univ. St Etienne, 1981-82, N°4, 23 p.
http://www.math.jussieu.fr/ miw/articles/ps/eccm.ps

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

Diophantine approximation, irrationality and transcendence

Michel Waldschmidt

Course N°21, *June 28, 2010*

The text below is taken from [4] § 5.2.

11.2 Modular functions and Ramanujan functions

S. Ramanujan introduced the following functions

$$P(q) = 1-24\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n}, \quad Q(q) = 1+240\sum_{n=1}^{\infty}\frac{n^3q^n}{1-q^n}, \quad R(q) = 1-504\sum_{n=1}^{\infty}\frac{n^5q^n}{1-q^n}\cdot$$

They are special cases of Fourier expansions of Eisenstein series. Recall the Bernoulli numbers B_k defined by:

$$\frac{z}{e^z-1} = 1-\frac{z}{2}+\sum_{k=1}^{\infty}(-1)^{k+1}B_k\frac{z^{2k}}{(2k)!},$$

$$B_1 = 1/6, \quad B_2 = 1/30, \quad B_3 = 1/42.$$

For $k \geq 1$ the normalized Eisenstein series of weight k is

$$E_{2k}(q) = 1+(-1)^k\frac{4k}{B_k}\sum_{n=1}^{\infty}\frac{n^{2k-1}q^n}{1-q^n}\cdot$$

The connection with (161) is

$$E_{2k}(q) = \frac{1}{2\zeta(2k)}\cdot G_k(\tau),$$

for $k \geq 2$, where $q = e^{2\pi i\tau}$. In particular

$$G_2(\tau) = \frac{\pi^4}{3^2\cdot 5}\cdot E_4(q), \qquad G_3(\tau) = \frac{2\pi^6}{3^3\cdot 5\cdot 7}\cdot E_6(q).$$

With Ramanujan's notation we have

$$P(q) = E_2(q), \quad Q(q) = E_4(q), \quad R(q) = E_6(q).$$

The discriminant Δ and the modular invariant J are related with these functions by Jacobi's product formula

$$\Delta = \frac{(2\pi)^{12}}{12^3} \cdot (Q^3 - R^2) = (2\pi)^{12} q \prod_{n=1}^{\infty} (1-q^n)^{24} \quad \text{and} \quad J = \frac{(2\pi)^{12} Q^3}{\Delta} = \frac{(2^4 3^2 5 G_2)^3}{\Delta}.$$

Let q be a complex number, $0 < |q| < 1$. There exists τ in the upper half plane $\mathcal{H}$ such that $q = e^{2\pi i \tau}$. Select any twelfth root ω of $\Delta(q)$. The invariants g_2 and g_3 of the Weierstraß $\wp$ function attached to the lattice $(\mathbf{Z} + \mathbf{Z}\tau)\omega$ satisfy $g_2^3 - 27 g_3^2 = 1$ and

$$P(q) = 3\frac{\omega}{\pi} \cdot \frac{\eta}{\pi}, \quad Q(q) = \frac{3}{4}\left(\frac{\omega}{\pi}\right)^4 g_2, \quad R(q) = \frac{27}{8}\left(\frac{\omega}{\pi}\right)^6 g_3.$$

According to formulae (162) and (163), here are a few special values

- For $\tau = i$, $q = e^{-2\pi}$,

$$P(e^{-2\pi}) = \frac{3}{\pi}, \quad Q(e^{-2\pi}) = 3\left(\frac{\omega_1}{\pi}\right)^4, \quad R(e^{-2\pi}) = 0 \quad \text{and} \quad \Delta(e^{-2\pi}) = 2^6 \omega_1^{12}, \tag{183}$$

with

$$\omega_1 = \frac{\Gamma(1/4)^2}{\sqrt{8\pi}} = 2.6220575542\ldots$$

- For $\tau = \varrho$, $q = -e^{-\pi\sqrt{3}}$,

$$P(-e^{-\pi\sqrt{3}}) = \frac{2\sqrt{3}}{\pi},\ Q(-e^{-\pi\sqrt{3}}) = 0, \quad R(-e^{-\pi\sqrt{3}}) = \frac{27}{2}\left(\frac{\omega_1}{\pi}\right)^6,\ \Delta(-e^{-\pi\sqrt{3}}) = -2^4 3^3 \omega_1^{12}, \tag{184}$$

with

$$\omega_1 = \frac{\Gamma(1/3)^3}{2^{4/3}\pi} = 2.428650648\ldots$$

11.3 Nesterenko's result

In 1976, D. Bertrand pointed out that Schneider's Theorem 173 on the transcendence of ω/π implies:
For any $q \in \mathbf{C}$ with $0 < |q| < 1$, at least one of the two numbers $Q(q)$, $R(q)$ is transcendental.

He also proved the p-adic analog by means of a new version of the Schneider–Lang criterion for meromorphic functions (he allows one essential singularity) which he applied to Jacobi–Tate elliptic functions. Two years later he noticed that G.V. Chudnovsky's Theorem 178 yields:
For any $q \in \mathbf{C}$ with $0 < |q| < 1$, at least two of the numbers $P(q)$, $Q(q)$, $R(q)$ are algebraically independent.

The following result of Yu.V. Nesterenko goes one step further:

Theorem 185 (Nesterenko, 1996)**.** *For any $q \in \mathbf{C}$ with $0 < |q| < 1$, three of the four numbers q, $P(q)$, $Q(q)$, $R(q)$ are algebraically independent.*

Among the tools used by Nesterenko in his proof is the following result due to K. Mahler:

The functions P, Q, R are algebraically independent over $\mathbf{C}(q)$.

Also he uses the fact that they satisfy a system of differential equations for $D = q\, d/dq$ discovered by S. Ramanujan in 1916:

$$12\frac{DP}{P} = P - \frac{Q}{P}, \quad 3\frac{DQ}{Q} = P - \frac{R}{Q}, \quad 2\frac{DR}{R} = P - \frac{Q^2}{R}.$$

One of the main steps in his original proof is his following zero estimate:

Theorem 186 (Nesterenko's zero estimate)**.** *Let L_0 and L be positive integers, $A \in \mathbf{C}[q, X_1, X_2, X_3]$ a non-zero polynomial in four variables of degree $\le L_0$ in q and $\le L$ in each of the three other variables X_1, X_2, X_3. Then the multiplicity at the origin of the analytic function $A\big(q, P(q), Q(q), R(q)\big)$ is at most $2 \cdot 10^{45} L_0 L^3$.*

In the special case where $J(q)$ is algebraic, P. Philippon produced an alternative proof for Nesterenko's result where this zero estimate 186 is not used; instead of it, he used Philibert's measure of algebraic independence for ω/π and η/π. However Philibert's proof requires a zero estimate for algebraic groups.

Using (183) one deduces from Theorem 185

Corollary 187. *The three numbers π, e^π, $\Gamma(1/4)$ are algebraically independent.*

while using (184) one deduces

Corollary 188. *The three numbers* π, $e^{\pi\sqrt{3}}$, $\Gamma(1/3)$ *are algebraically independent.*

Consequences of Corollary 187 are the transcendence of the numbers

$$\sigma_{\mathbf{Z}[i]}(1/2) = 2^{5/4}\pi^{1/2}e^{\pi/8}\Gamma(1/4)^{-2}$$

and (P. Bundschuh)

$$\sum_{n=0}^{\infty}\frac{1}{n^2+1} = \frac{1}{2}+\frac{\pi}{2}\cdot\frac{e^{\pi}+e^{-\pi}}{e^{\pi}-e^{-\pi}}.$$

D. Duverney, K. and K. Nishioka and I. Shiokawa as well as D. Bertrand derived from Nesterenko's Theorem 185 a number of interesting corollaries, including the following ones

Corollary 189. *Rogers-Ramanujan continued fraction:*

$$RR(\alpha) = 1+\cfrac{\alpha}{1+\cfrac{\alpha^2}{1+\cfrac{\alpha^3}{1+\ddots}}}$$

is transcendental for any algebraic α *with* $0<|\alpha|<1$.

Corollary 190. *Let* $(F_n)_{n\geq 0}$ *be the Fibonacci sequence:* $F_0=0$, $F_1=1$, $F_n=F_{n-1}+F_{n-2}$. *Then the number*

$$\sum_{n=1}^{\infty}\frac{1}{F_n^2}$$

is transcendental.

Jacobi Theta Series are defined by

$$\theta_2(q) = 2q^{1/4}\sum_{n\geq 0}q^{n(n+1)} = 2q^{1/4}\prod_{n=1}^{\infty}(1-q^{4n})(1+q^{2n}),$$

$$\theta_3(q) = \sum_{n\in\mathbf{Z}}q^{n^2} = \prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1})^2,$$

$$\theta_4(q) = \theta_3(-q) = \sum_{n\in\mathbf{Z}}(-1)^n q^{n^2} = \prod_{n=1}^{\infty}(1-q^{2n})(1-q^{2n-1})^2.$$

Corollary 191. . *Let i, j and $k \in \{2,3,4\}$ with $i \neq j$. Let $q \in \mathbf{C}$ satisfy $0 < |q| < 1$. Then each of the two fields*

$$\mathbf{Q}\big(q, \theta_i(q), \theta_j(q), D\theta_k(q)\big) \quad \textit{and} \quad \mathbf{Q}\big(q, \theta_k(q), D\theta_k(q), D^2\theta_k(q)\big)$$

has transcendence degree ≥ 3 over $\mathbf{Q}$.

As an example, *for an algebraic number $q \in \mathbf{C}$ with $0 < |q| < 1$, the three numbers*

$$\sum_{n\geq 0} q^{n^2}, \quad \sum_{n\geq 1} n^2 q^{n^2}, \quad \sum_{n\geq 1} n^4 q^{n^2}$$

are algebraically independent. In particular *the number*

$$\theta_3(q) = \sum_{n\in\mathbf{Z}} q^{n^2}$$

is transcendental. The number $\theta_3(q)$ was explicitly considered by Liouville as far back as 1851.

The proof of Yu.V. Nesterenko is effective and yields quantitative refinements (measures of algebraic independence).

References

[1] *Introduction to algebraic independence theory*, vol. 1752 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2001. With contributions from F. Amoroso, D. Bertrand, W. D. Brownawell, G. Diaz, M. Laurent, Yuri V. Nesterenko, K. Nishioka, Patrice Philippon, G. Rémond, D. Roy and M. Waldschmidt, Edited by Nesterenko and Philippon.

[2] M. Waldschmidt, *Sur la nature arithmétique des valeurs de fonctions modulaires*, Astérisque, (1997), pp. Exp. No. 824, 3, 105–140. Séminaire Bourbaki, Vol. 1996/97.

[3] ——, *Transcendance et indépendance algébrique de valeurs de fonctions modulaires*, in Number theory (Ottawa, ON, 1996), vol. 19 of CRM Proc. Lecture Notes, Amer. Math. Soc., Providence, RI, 1999, pp. 353–375.

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