Lambert and the irrationality of π (1761)

by Alain Juhel, *agrégé* in mathematics,
maths teacher in *classe préparatoire* MPSI
Lycée Faidherbe, Lille

THE CONTEXT

The text by Lambert (1761) we will examine here occupies a pivotal place in the history of irrationality and transcendence.

1. In the history of the number π, this first proof of irrationality is clearly crucial.
2. It extends the proof of the irrationality of *e* given by Euler (1737) to that of all its powers.
3. It marks the beginning of the precise formulation of the notion of transcendence, and sets out the corresponding conjecture about the two remarkable numbers *e* and π, problems that would be resolved by Hermite (1872) and Lindemann (1882) respectively.
4. It constitutes an essential milestone on the road to negating the problem of squaring the circle, catalysing the resolution of a problem that had stagnated since it was first exposed in the 5th century BCE.
5. In passing, Lambert defines what we now call hyperbolic functions, by justifying, with a supporting figure, the notion of hyperbolic trigonometry.

If there is one mathematician who full recognises the value of Lambert's text, it has to be Charles Hermite. After all, following his triumphant proof of the transcendence of e, he states in the introduction to a simplified demonstration of Lambert's result: "All I can do is do what Lambert has already done, but in another way …"

Glossary

Irrationality: A number is rational if it can be expressed as the quotient of two integers. If not, it is irrational.

Transcendence: A number is transcendent when it is not the root of an equation with integer coefficients of any degree.

Squaring the circle: Constructing a square with the same area as a unit circle using only a compass and a ruler. This amounts to constructing $\sqrt{\pi}$.

THE MAN

Johann Lambert was born in 1728 in Mulhouse (in what was then … Switzerland). A self-taught man, he was invited by Euler to become a fellow of the Academy of Berlin in 1764, and died in that city in 1777.

Lambert's most famous publications, besides the essay considered here, are about non-Euclidean geometry, cartography – his conformal projection is still one of the most widely used projections among geographers – and the study of perspective (including the construction of a mechanical device known as the "perspectograph").

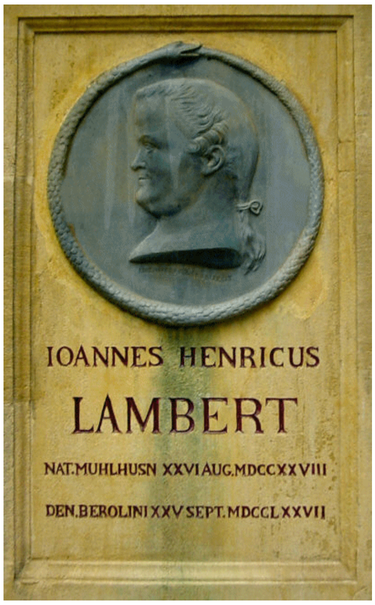


***Figure 1:* Lambert's column in Mulhouse**, *his home town (left). It is inscribed with a noon mark in memory of his astronomical work. Right, detail: Medallion in honour of Lambert at the base of the column.*

THE METHOD

The standard tool at the time of the pioneers was the theory of continued fractions.

Euler had used it first to prove the irrationality of *e* in 1737.[1] Nothing feels more natural once one has familiarised oneself with it. If the process of divisions does not terminate when the following algorithm is repeated[2]

$$y = x_0 - [x_0] \longrightarrow x_1 = \frac{1}{y} \longrightarrow x_0 = [x_0] + \frac{1}{x_1}$$

where $[x_0]$, depending on its usage, indicates the integer part of x_0, then the initial number x_0 is irrational, as the non-termination proves! All Euler had to do was obtain a nice infinite continued fraction with perfectly regularly terms – albeit a little mysteriously – to adduce the irrationality of e:

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \ldots}}}}}$$

Or, written more concisely: [2,1,2,1,1,4,1,1,6,...1,1,2n,1,1...]

Unfortunately, π proved itself a tougher nut to crack. As we can see in the example in the panel below, it is possible – using ever more precise decimal values – to obtain as many terms as one likes. But neither Euler, nor any other mathematician, has identified any regularity in this expansion … π therefore proves itself to be a "more complicated" number than e. The sequence of decimals, i.e. the series expansion

$$\sum_{n=0}^{\infty} \frac{d_n}{10^n}$$

is unpredictable for both *e* and π, but at least e establishes a regular form with another form of expansion. The same cannot be said of π.

---

1. *De Fractionibus continuis Dissertatio*, 1737.
2. With this algorithm x_0 constructs x_1, then x_1 constructs x_2, and so on and so forth.

Continued fractions

All rational numbers can be written as $x = a + 1/y$, where a is an integer and y>1. We can therefore reiterate the method, which amounts to repeating Euclid's algorithm (corresponding to Euclidean division). So, to expand = 314159/100000, we write successively:

- 100000 = 7 x 14159 + 887
- 14159 = 15 x 887 + 854
- 887 = 1 x 854 + 33
- 854 = 25 x 33 + 29
- 33 = 1 x 29 + 4
- 29 = 7 x 4 + 1

Euclid's algorithm terminates, and the same is true for any rational algorithm. Therefore

$$x = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{25 + \cfrac{1}{1 + \cfrac{1}{7 + \cfrac{1}{1}}}}}}}$$

which, for convenience, can be written as:

$$x = 3 + \frac{1}{7+}\ \frac{1}{15+}\frac{1}{1+}\ \frac{1}{25+}\ \frac{1}{1+}\ \frac{1}{7+}\ \frac{1}{1} \text{ ou } [3,7,15,1,25,1,7,1]$$

The "intermediary" fractions

$$3 + \frac{1}{7} = \frac{22}{7}\ 3 + \cfrac{1}{7 + \cfrac{1}{15}} = \frac{333}{106}\ 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1}}} = \frac{355}{113}$$

are known as **reduced fractions.**

More generally, we can study fractions with any given terms,

$$x = \frac{a_1}{b_1 +}\ \frac{a_2}{b_2 +} \ldots \frac{a_k}{b_k +} \ldots$$

and whose successive reductions (fractions truncated to rank n) are calculated by the following algorithm (which is very interesting as it does not contain any division):

$$A_{n+1} = b_{n+1}A_n + a_{n+1}A_{n-1}$$

$$B_{n+1} = b_{n+1}B_n + a_{n+1}B_{n-1}$$

A more elaborate strategy of attack is required, and this is what makes Lambert's work so original. Here are the essential ingredients:

1. Abandoning regular continued fractions (with a numerator of 1) and replacing them with fractions with any numerator;
2. Formulating an associated criterion of irrationality;
3. Using them to develop a trigonometric function, the tangent – and no longer simply a number;
4. Operating by *reductio ad absurdum*: if $\pi/4$ were rational, $\tan(\pi/4) = 1$ would not be.

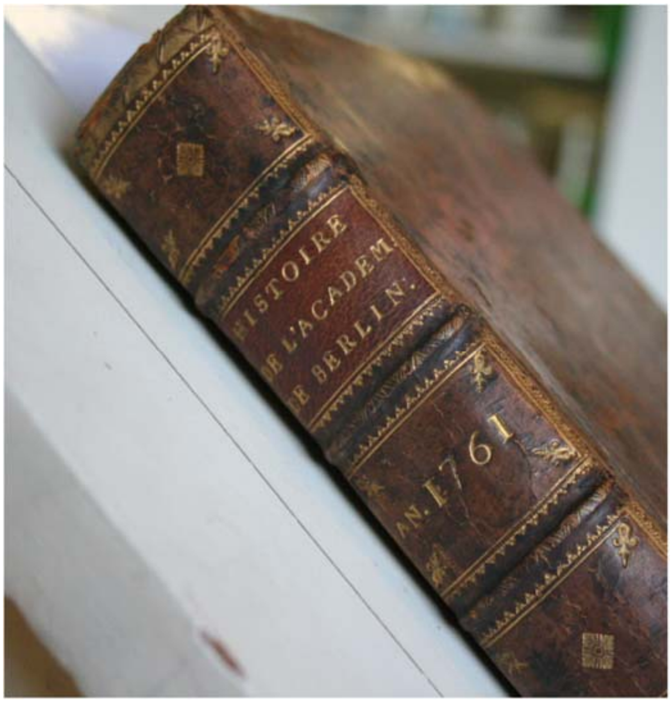
> *The purpose is to show that each time the arc of a given circle is commensurable with the radius, the tangent of that arc is incommensurable; & that reciprocally, any commensurable tangent is not that of a commensurable arc*. (§2)

As with other problems of this kind, this marvellous tool would later be sacrificed at the altar of simplifications. Euler's proof was supplanted by the proof attributed to Fourier (1815);[3] Liouville himself, in his second article on the construction of remarkable transcendental numbers (1844), put aside the continued fractions he had originally used;[4] while Hermite, in a letter to Borchardt (1873), provides proofs of the irrationality of π and π^2 that contain not a single continued fraction, even though, in their inspiration, they are indelibly marked by rational approximations constructed from continued fractions. Hermite himself felt some regret at this development:[5]

> *The expression used by Lambert, which I avoid, nevertheless remains a result of the highest order, one that leads the way to curious and interesting research.*

3. See the Stainville text on *BibNum*.
4. See the Liouville text on *BibNum*.
5. *Sur l'Irrationalité de la Base des Logarithmes Hyperboliques.* Report of the British Association for Advancement of Science, 43rd Meeting, 1873.

How the text unfolds



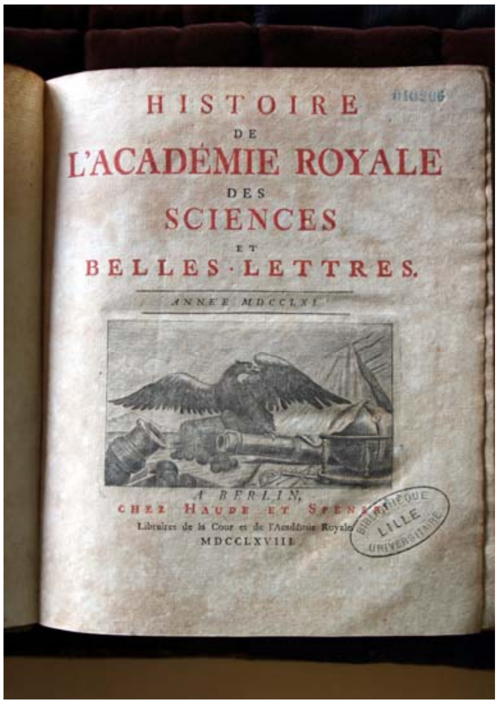


Figure 2: ***Histoire de l'Académie royale des sciences et belles-lettres de Berlin (1761),*** *which contains Lambert's text. In this era the minutes were written in French. (Copy conserved at the University of Lille's arts and humanities library.)*

The article is long, but – Lambert informs us at the bottom of the second paragraph – that is the price one has to pay if the reasoning is to be beyond reproach. And, once accepted, this principle seems a small sacrifice to make in the name of remarkable results! Indeed, it is these very results that make the text so much more than a straightforward proof of irrationality …

To avoid getting into a muddle at this first reading, let's map out the text's plan and key paragraphs.

Overview: §1–4

Lambert insists on the necessity of absolute rigour because of what is at stake: the question of squaring the circle. After outlining a classic series for π, complete with vague argumentation – as if to give the reader an example of all that he rejects – he presents the problem:

No matter how vague this reasoning is, there are nevertheless cases where nothing more is required. But these cases are not those of squaring the circle.

He goes on to recall the paradox of those obsessive "circle-squarers" (in 1775 the Academy of Sciences actually refused to examine any more 'solutions', in order to prevent squarers from going insane):

Most of those who seek [the solution] do so with an ardour that sometimes leads them to gravely doubt the most fundamental and established geometrical truths. Is it conceivable that they would be satisfied by what I have just said? Something completely different is needed.

The text would indeed prove exemplary in its rigour. In that it contrasts sharply with the writings of Lambert's near contemporaries, whether Euler before him or Legendre afterwards (1795), both of whom left uncertainties in terms of their sources of inspiration and the convergence of fractions they employ.

Expansion of tan x – Act 1:
The pleasure of discovery: §5–14

Lambert starts with two classic series giving sin υ and cos υ, then forms their quotient and "attempts division". Let's imagine the following:

$$\tan\ \upsilon = \frac{\upsilon - \upsilon^3/6 + R_1}{1 - \upsilon^2/2 + R_2} \text{ and } F(\upsilon) = \frac{\upsilon - \upsilon^3/6}{1 - \upsilon^2/2}$$

In the case of *F* (obtained by ignoring the remainders), one could very easily proceed with a view to obtaining a limited continued fraction expansion. One would simply need to replace the numbers by polynomials in υ, and the Euclidean division of integers by polynomial division, to obtain the start of the fraction:

$$F(\upsilon) = \cfrac{\upsilon}{1 - \cfrac{\upsilon^2}{3 - \rho(\upsilon)}}$$

Going up to the power of 5 in the numerator and the power of 4 in the denominator would confirm the 1 and 3 obtained and reveal a new term, etc. With this intuition to guide him, Lambert adds a careful calculation of the successive remainders. Below is the expansion he obtains – but does not demonstrate – for, as

Lambert was well aware (far more acutely than Euler), saying "and so on *ad infinitum*" is not enough.

The calculation in detail

$$F(\upsilon) = \frac{\upsilon - \upsilon^3/6}{1 - \upsilon^2/2} = \upsilon \times \frac{1 - \upsilon^2/6}{1 - \upsilon^2/2} = \frac{\upsilon}{G(\upsilon)} \text{ avec } G(\upsilon) = \frac{1 - \upsilon^2/2}{1 - \upsilon^2/6}$$

Operating by division by increasing powers:

$$1 - \upsilon^2/2 = \left(1 - \upsilon^2/6\right) \times 1 - \upsilon^2/3$$

$$G(\upsilon) = 1 - \frac{\upsilon^2/3}{1 - \upsilon^2/6} = 1 - \frac{\upsilon^2}{3 - \upsilon^2/2} \text{ donc } F(\upsilon) = \frac{\upsilon}{G(\upsilon)} = \frac{\upsilon}{1 - \frac{\upsilon^2}{3 - \upsilon^2/2}}$$

What we should be able to do with it … §15–16

The reader will no doubt have noticed that we have used υ whereas Lambert seems to further complicate things by performing his operations with $w = 1/\upsilon$. The reason for this is explained here:

> *The problem posed by Euclid is to find the largest common divisor of two integers … This last supposition occurs each time* $1/\upsilon$ *is an integer.*

In this case, then, here is Lambert's proof of irrationality: he used only integers, the algorithm does not terminate, and so the number tan υ

> *will be an irrational quantity each time arc* υ *is an aliquot part of the radius.*

At the end of §16 he concludes:

> *This therefore is the use to which we can confine Euclid's proposition.*

Reasoning by the absurd, as he envisages, would suppose that $\pi/4$ is rational (commensurable to the radius) but not necessarily that it has a numerator of 1 (aliquot part).

It is now a question of extending it to all cases where the arc is commensurable to the radius.

Indeed, Lambert had forewarned us of this as early as §3:

But it should be noted that while Euclid applies it only to integers and rational numbers, I must use it another way …

Expansion of tan x – Act 2: Forming reduced fractions: §17–22

Now our hero simply has to resume his attack. The advantage of his heuristics is that they give him a continued fraction to consider *a priori* when he sets himself the task of proving that the fraction converges, and converges towards tan υ:

Yet, by carrying over the quotients w, 3w, 5w, &c. as much as one likes, one would simply have to reduce them, to have fractions that express the tangent of v all the more exactly than if one had carried over a larger number of quotients.

He presents the reduced fractions: it is an inverse calculation of our simplified expansions. He then demonstrates the (classic) formula of separately calculating the numerators and denominators of the reduced fraction A_n/B_n, which in modern notation we would write as:

$$A_{n+1} = (2n+1).w\ A_n - A_{n-1}$$
$$B_{n+1} = (2n+1).w\ B_n - B_{n-1}$$

The algorithm is shown in a table calculating the first partial numerators and denominators.

Expansion of tan x – Act 3: The demands of rigour: §23–34

In this table, Lambert discovers the general explicit form of A_n and B_n. But one only ever finds what one is looking for ... In view of this study of convergence, what Lambert is hoping to find is an A_n close to the sine series – the numerator of tan υ – and a B_n close to the cosine series – the denominator of tan υ. With this beacon to guide him he can safely conclude the determination, whereas a "human calculator" – even an experienced one – would get lost in the operational fog. And, as it is out of the question to generalise without proof, for each *n*:

It is advisable, in order, here again, to avoid any form of induction, to give and demonstrate its general expression.

This he manages at §29 for the denominator, and §30 for the numerator. For example:

$$\frac{1}{1\times 3\times 5\ldots\times(2n-1)}A_n=\sum_{k=1}^{\infty}a_{k,n}\frac{(-1)^k}{(2k-1)!}\upsilon^{2k-1}$$

This is probably the only lacuna in his proof, in that he extends n towards infinite, invoking

$$\lim_{n\longrightarrow\infty}a_{k,n}=1$$

to conclude. However, Lebesgue[6] later showed that this point warranted much lengthier discussion. In addition, he concludes by emphasising the very high convergence quality of the approximants. This can be fully appreciated today – and more easily than in Lebesgue's time – with the help of a computer-plotted image. He describes the difference between tan υ and the reduced fraction A_n / B_n in the following terms:

All these sequences are more convergent than any decreasing geometric progression.

It would be inexact to say that this was the first continued fraction expansion of a function (and not a mere number). Euler had expanded *exp (x)* using differential equations. But, if we compare this expansion with our breakdown of Lambert's proof into three stages, we see that Euler had completed only the first stage. This is therefore a remarkable turning point in the approximation of functions.

6. Henri Lebesgue, *Leçons sur les constructions géométriques*, Gauthier-Villars, 1949 (rep. J. Gabay 1987).

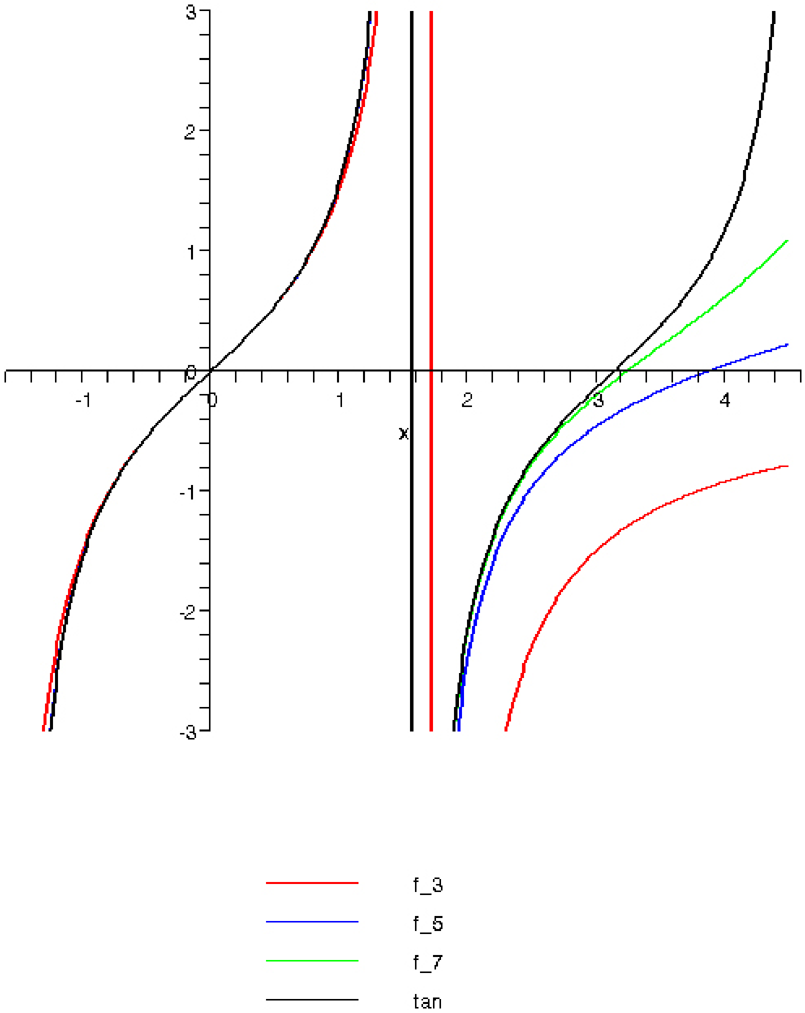


***Figure 3:* Approximants of the tangent function used by Lambert.**
The notation f_n underneath the curves indicates the reduced function, taking into account the partial denominators up to $\upsilon^2/2n$.

Irrationality … At last! §35–51

Here it is question of adapting any given continued fractions to the criterion Euler applied to regular continued fractions. Again, this is the "Use it, but differently" maxim from §3 and §16. Lambert therefore notes the expansion and the reduced fractions for $\upsilon = \Phi/\omega$, Φ and ω being integers. The idea is simple: if tan Φ/ω is rational – and also written as the quotient M/P of two integers – the partial quotients produced by the expansion of tan υ will correspond to the application of Euclid's algorithm to M and P. On one hand, the algorithm should therefore terminate; on the other, we will be left with an infinite and strictly decreasing sequence of integers, hence the contradiction. This infinitely decreasing sequence derives from the fact that, after a certain rank, all the partial quotients are inferior to 1: Φ and ω are fixed, while the denominator is successively multiplied by 3, 5, 7, 9, etc. Lambert's proof is long, whereas Lebesgue's (*op. cit*.) runs to less than one page. To avoid losing sight of the bigger picture, readers are advised to skip over this passage at a first reading. It is simpler to accept the following lemma:[7]

$$\text{Let } x = \frac{a_1}{b_1 +}\ \frac{a_2}{b_2 +} \cdots \frac{a_k}{b_k +} \cdots \text{ a continued fraction such as } \forall k,\ 1 + |a_k| < |b_k|,$$

with a_k and b_k integers ≥ 1, so $|x| < 1$ and x is irrational.

It should also be noted that though the generalised continued fraction expansion of π had been widely known since Lord Brouncker (1655) –

$$\Pi = \frac{1}{1+}\ \frac{1^2}{2+}\ \frac{3^2}{2+}\ \frac{5^2}{2+}\ \frac{7^2}{2+} \cdots,$$

– the former criterion did not apply: the fraction did not converge rapidly enough. Lambert's method was therefore truly groundbreaking!

Paragraphs 52–72 are of no great interest to the modern-day reader, so we shall skip over them at this first reading.

From circular functions to hyperbolic functions: §73–80

As early as §4, just after noting the series expansions of cos and sin, Lambert informs the reader:

> *[I]n what follows I will give two sequences for the hyperbola that will differ from these two only in that all the signs are positive …*

He therefore sets out two series, which he adds up in exponential terms, for example:

---

7. Details can be found in Lebesgue's book, *op. cit.*

$$\frac{e^{\upsilon}-e^{-\upsilon}}{2}=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}\upsilon^{2k+1}$$

He then expands their quotient by observing that the only difference is the replacement of all the – signs with + signs:

$$\frac{e^{\upsilon}-e^{-\upsilon}}{e^{\upsilon}+e^{-\upsilon}}=\cfrac{\upsilon}{1+\cfrac{\upsilon^2}{3+\cfrac{\upsilon^2}{5+\ldots}}}$$

The modern-day notations (cosh, sinh, tanh) are not yet present here, but that it is beside the point. Not content with observing a transition from one type to the other by changing υ to $i\upsilon$ – Euler had got there before him! – Lambert is keen to provide a geometric interpretation. In the figure below, he shows the hyperbola $x^2 - y^2 = 1$ touching the circle $x^2 + y^2 = 1$: it is tangent with the circle at its apex A (1,0).

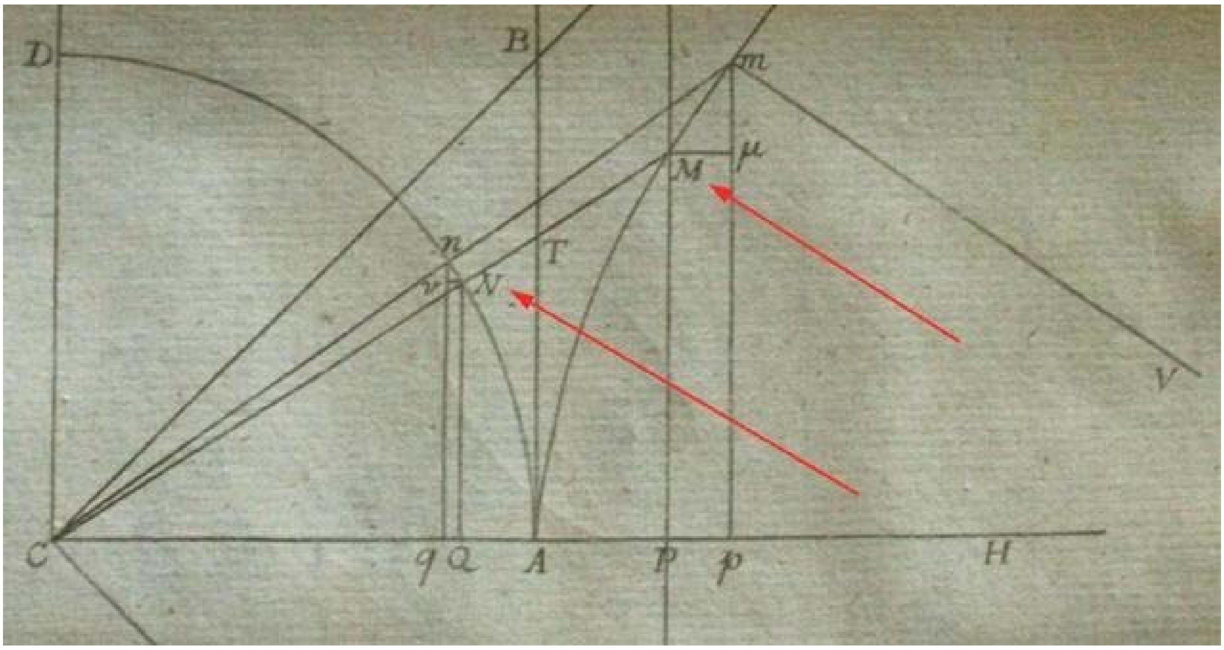


***Figure 4:* Plate at the end of Lambert's article (detail).** *We can see the line of the quadrant AD and the branch of the hyperbola AM. Arrows show the common (intersection) points M (on the hyperbola) and N (on the circle) (photograph: A. Juhel).*

A common secant from the centre C of the two conics cuts the circle in N(x,y) and the hyperbola in M(ξ,η). Introducing the parameter u as double the area of the hyperbolic section AMCA, Lambert demonstrates that

$$\xi = \cosh(u),\ \eta = \sinh(u),\ y/x = \tan(\varphi) = \tanh(u) = \xi / \eta$$

Just as the functions cos (φ) and sin (φ) define the parameters of the circle, the functions *cosh (u)* and *sinh (u)* define the parameters of the hyperbola. This is the birth of hyperbolic functions, complete with detailed notations.

Beyond Euler: The irrationality of exp(n) and exp(1/n): §81–88

Though Euler had the expansion of *tanh (u)* at his disposal, he was anxious to identify regular continued fraction expansions, and for this reason could not go as far as the irrationality of e^2. Lambert's method goes further: *u* and *exp(u)* cannot be simultaneously rational (§81):

> *[E]very rational hyperbolic logarithm is that of an irrational number … Every rational number has an irrational hyperbolic logarithm.*

And, when *u* is an integer or the inverse of an integer, he concludes:

> *These fractions demonstrate the extent to which the irrationality of the number e = 2,718281828... is transcendental, in that none of its [powers[8]] nor roots is rational …*

That grand word – transcendence – is finally unleashed, yet its meaning here remains vague: irrational beyond anything one could imagine. This meaning was already present in Lambert's announcement at the end of §2:

> [This statement] *again demonstrates the extent to which transcendental circular quantities are transcendental & beyond all commensurability.*

But the three final paragraphs go much further, showing the way forward both for Wantzel in his research on constructability, and Hermite in his demonstration of the transcendence of number *e*.

---

8. [Translators' note] Lambert uses the term *dignités*, which corresponds to the modern mathematical term *powers*.

Conjecturing about transcendence: Towards non constructability: §89–91

From the outset, Lambert's tone is prophetic:

> *Everything I have just shown of circular & logarithmic transcendental quantities seems to be founded on much more universal principles, but which are not as yet sufficiently developed. Here, however, is some idea.*

And lo and behold, there follows a modern definition of an algebraic number! Lambert not only provides a list of explicit examples suggesting an arbitrary "accretion" of radicals of all orders and algebraic operations:

> *there is an infinitude of other [quantities] known as* algebraic *[quantities]: & such are all* radical irrational quantities, *like* $\sqrt{2}$, $\sqrt{3}$, $\sqrt[3]{4}$, *& c*, $\sqrt{2+\sqrt{3}}$, *&c. & all the irrational roots of algebraic equations …*

More significantly, it is in these terms that he defines the roots of equations of any order with integer coefficients, though without an explicit formula:

> *& all the* irrational roots of algebraic equations *such as, for example, those of the equations* $x^3 - 5x + 1$, *&c.*

The *&c.* (etc.) clearly suggests an indeterminate degree, and therefore $x^5 - 5x + 1$, for which there is no known solution (Abel–Ruffini theorem), and more generally $x^n - 5x + 1$. He concludes with the following:

> *& here is the theorem, which I believe can be demonstrated.*

> *I therefore state that no transcendental circular & logarithmic quantity can be expressed by a given radical irrational quantity, that refers to the same unity & contains no transcendental quantity.*

It is useful to specify the nature of the coefficients: e is the root of the equation x – e = 0, but the coefficients are not rational (*referr[ing] to the same unity*) and, if e were the coefficient, there would be *a transcendental quantity*. The theorem therefore states that, among others, *e* and π are not the roots of any equation with rational coefficients, or – and this amounts to the same thing once reduced to the same denominator – of any integer equation.

The concluding paragraphs consider the problems of constructability and squaring the circle. Though he provides no further demonstrations, Lambert's vision is perfectly clear:

> *Once this theorem is demonstrated in all its universality, it will follow that, since the circumference of a circle cannot be expressed by any radical quantity, nor by any rational quantity, there will be no way of determining it by any geometric construction.*

Squaring the circle would entail constructing $\sqrt{\pi}$ – and therefore π – with a ruler and compass, yet

> *anything that can be constructed geometrically involves rational and radical quantities.*

Here Lambert no doubt shows himself to be an informed reader of Descartes' *Geometry*, in which the latter had shown how to construct square sums, products and roots. Equally, he is not oblivious to the fact that for 2,000 years, mathematicians had been trying (and failing) to solve a problem contemporary with that of squaring the circle: doubling the cube – or the Delian problem – which would equate to constructing $\sqrt[3]{a}$, *a* being a given rational number. Drawing on this practical failure as conjectural evidence of a theoretical impossibility, he states:

> *& the possibility that these can be indifferently constructed appears remote to say the least.*

It would be another three quarters of a century, however, before Wantzel (1837) could solve the question once and for all. Yet Lambert's intuitions about non-constructability are there in embryonic form in the conclusion to his article:

> *It is clear that this is true for the arcs of all circles whose length or two outermost points are given either by rational quantities or by radical quantities. For, if the length of the arc is given, one would need to find its outermost points by using the chord, sine or tangent, or any other straight line, which, in order to be constructed, would be dependent on or reducible to one of the lines I have just mentioned. But as the length of the arc would be given by rational or radical quantities, these lines would be transcendental, & therefore irreducible to any rational or radical quantity. The same would be true if the two outermost points of the arc were given, and by that I mean by rational or radical quantities. For, in this case, the length of the arc would be a*

transcendental quantity, which means irreducible to any rational or radical quantity, & in that, admitting no geometric construction.

■

(February 2009)

(version V2 slightly modified, March 2015)

(translated by Helen Tomlinson, published March 2015)