10

CONTINUED FRACTIONS OF ALGEBRAIC NUMBERS

Enrico Bombieri
Alfred J. van der Poorten

School of Mathematics
Institute for Advanced Study
Princeton, New Jersey 08540 USA
eb@math.ias.edu

CeNTRe for Number Theory Research
Macquarie University NSW 2109 Australia
alf@mpce.mq.edu.au

1 INTRODUCTION

Everyone knows that BC — before calculators, π was $22/7$ and AD — after decimals, π became $= 3.14159265\ldots$. In other words, π is quite well approximated[1] by the vulgar fraction $22/7$; and some of us know that $355/113$ does a yet better job since it yields as many as seven correct decimal digits. The 'why this is so' of the matter is this: It happens that

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \ddots}}}}}$$

For brevity, a common flat notation for such a *continued fraction* expansion is

$$[3\,,7\,,15\,,1\,,292\,,1\,,\ldots]\,.$$

[1] That $\pi \neq 22/7$ follows from

$$0 \neq \int_0^1 \frac{t^4(1-t)^4}{1+t^2}\,dt = \frac{22}{7} - \pi\,.$$

The entries 3, 7, 15, $\ldots$ are known as the *partial quotients*, and the truncations, for example

$$[3\ ,\ 7] = 3 + \frac{1}{7} = \frac{22}{7}\,,$$

or

$$[3\ ,\ 7\ ,\ 15\ ,\ 1] = \frac{355}{113}$$

are known as *convergents* of π.

The important truth is that the convergents p_h/q_h, $h = 0, 1, 2, \ldots$ yield good rational approximations, indeed excellent ones relative to the size of the denominator q_h. In the present example

$$\left|\pi - \frac{22}{7}\right| < \frac{1}{15\cdot 7^2} \text{ and } \left|\pi - \frac{355}{113}\right| < \frac{1}{292\cdot 113^2}\,,$$

instancing the general result that

$$\left|\pi - \frac{p_h}{q_h}\right| < \frac{1}{c_{h+1}\cdot q_h^2}\,,$$

where c_{h+1} is the next (as yet unused) partial quotient. In particular $22/7$ and $355/113$ yield unusually good approximations to π because the subsequent partial quotients, respectively 15 and 292, are relatively large.

1.1 Or consider the following example: Apéry's proof, see [8] of the irrationality of $\zeta(3)$ alerts one to simple combinatorial proofs for the pair of formulæ

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \zeta(2) = 3\sum_{n=1}^{\infty}\frac{1}{n^2\binom{2n}{n}}$$

and

$$\sum_{n=1}^{\infty}\frac{1}{n^3} = \zeta(3) = \tfrac{5}{2}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}}\,.$$

The proofs do not appear to generalise, but it seems natural to experiment and to ask about the constant c in

$$\sum_{n=1}^{\infty}\frac{1}{n^4} = \zeta(4) = c\sum_{n=1}^{\infty}\frac{1}{n^4\binom{2n}{n}}\,.$$

Since, of course, $\zeta(4) = \pi^4/90$, and since the series on the right converges quite rapidly, it was not a big thing to make a programmable calculator reveal that

$$c \approx 2.1176470588\,.$$

The question is: 'Have we computed a rational number?' This appears to raise a metaphysical problem until continued fractions come to the rescue. We have

$$c = 2 + \cfrac{1}{8 + \cfrac{1}{2}}$$

up to calculator accuracy (the next partial quotient is huge!). Thus, plainly, one should guess that $c = 36/17$ and that it is indeed rational[2].

1.2 In any event, continued fractions are a good thing and their properties should be better known. Accordingly we provide a crash introduction to the subject below. Our object is to describe a congenial method for finding the continued fraction expansion of *algebraic numbers*: zeros of polynomials with rational integer coefficients. An amusing feature turns out to be that the algorithm is proved to work by virtue of deep results from the theory of diophantine approximation.

If $\gamma = a/b$ is rational — algebraic of degree 1 — then we set $\gamma = \gamma_0 = a_0/b_0$ and obtain its sequence of complete quotients (γ_h) by

$$\gamma_h = \frac{a_h}{b_h} = c_h + \frac{b_{h+1}}{b_h} \text{ with } c_h \in \mathbb{Z} \text{ and } b_{h+1} < b_h\,;$$

next $a_{h+1} = b_h$. The continued fraction of γ terminates and its expansion is equivalent to applying the Euclidean algorithm to the pair of integers a and b consisting of the numerator and denominator of γ; that, incidentally, may explain the term 'partial quotient'. If γ is a real quadratic irrational then the algorithm is again easy to describe explicitly. In fact, there is a positive algebraic integer α, with conjugate denoted by $\overline{\alpha}$, and rational integers P and Q so that we may set

$$\gamma = \frac{\alpha + P}{Q} \text{ with } Q \mid (\alpha + P)(\overline{\alpha} + P)\,.$$

[2] This was discovered experimentally in just the way described here; and verified rather later: AvdP noticed, in the introduction to Lewin's book [6], that

$$\int_0^{\pi/3} x\Big(\log(2\sin\tfrac{1}{2}x)\Big)^2\,dx = \frac{17\pi^4}{6480}\,.$$

This must of course be the same 17. Because

$$2(\sin^{-1}x)^2 = \sum_{h\geq 1}(2x)^{2h} \Big/ \binom{2h}{h}\,,$$

integration by parts indeed shows this to be the formula we were attempting to establish; see [9].

One has $\gamma = [c_0\,, c_1\,, c_2\,, \ldots]$, where the partial quotients are obtained by setting $\gamma = \gamma_0 = (\alpha + P_0)/Q_0$ and

$$\gamma_h = (\alpha + P_h)/Q_h = c_h - (\overline{\alpha} + P_{h+1}))/Q_h \text{ with } -1 < (\overline{\alpha} + P_{h+1}))/Q_h < 0\,;$$

so $P_h + P_{h+1} + (\alpha + \overline{\alpha}) = c_h Q_h$. The next complete quotient

$$\gamma_{h+1} = (\alpha + P_{h+1})/Q_{h+1}$$

is given by $Q_h Q_{h+1} = -(\alpha+P_{h+1})(\overline{\alpha}+P_{h+1})$. It is easy to confirm by induction that the P_h are integers bounded by $-\alpha < P_{h+1} < -\overline{\alpha}$ and that the Q_h are positive integers bounded above by $\alpha - \overline{\alpha}$. It follows immediately that there are only finitely many possibilities for the pairs (P_h, Q_h) and that therefore the continued fraction expansion must be eventually periodic.

1.3 Higher matters, that is the continued fraction expansions of algebraic numbers of higher degree, are less well understood. Suppose then that γ is presented as a real zero of a polynomial

$$f(X) = a_0 X^r + a_1 X^{r-1} + \cdots + a_r$$

with $a_0, a_1, \ldots, a_r \in \mathbb{Z}$ and $\gcd(a_0, \ldots, a_r) = 1$. We can suppose that $a_0 > 0$, and that γ is a simple zero of f since otherwise we deal with the greatest common divisor of f and f'. Thus we may locate γ by the fact that f suffers a change of sign in an interval containing γ.

Suppose that we have found that the complete quotient γ_h is the unique positive zero of the polynomial

$$f_h(X) = a_{h,0} X^r + a_{h,1} X^{r-1} + \cdots + a_{h,r}\,,$$

with rational integers $a_{h,i}$ and $a_{h,0} > 0$. We shall see, surprisingly perhaps, that this situation is generic — the phenomenon of *reduction* discussed below entails that our polynomials have just one positive zero. Then we may search for the integer c_h so that $f_h(c_h) < 0$ but $f_h(c_h) > 0$. Having found the partial quotient c_h we then define

$$f_{h+1}(X) = -X^r f_h(X^{-1} + c_h)\,,$$

and find that γ_{h+1} is the unique real positive zero of f_{h+1}.

It turns out, however, that a simple-minded search for c_h, sequentially trying each positive integer, can be very slow because the partial quotients may be surprisingly large. For example, with $\alpha = \sqrt[3]{2}$ one has, see [5],

$$c_{35} = 534\,, \quad c_{41} = 121\,, \quad \ldots\,, \quad c_{571} = 7451\,, \quad c_{619} = 4941\,, \quad \ldots$$

and yet more strikingly

$$\sqrt[3]{5} = [1\,,\,1\,,\,2\,,\,4\,,\,3\,,\,3\,,\,1\,,\,5\,,\,1\,,\,1\,,\,4\,,\,10\,,\\ 17\,,\,1\,,\,14\,,\,1\,,\,1\,,\,3052\,,\,1\,,\,1\,,\,1\,,\,\ldots]\,.$$

The continued fraction expansion of the real zero of $X^3 - 8X - 10$ is

$$[3\,,\,3\,,\,7\,,\,4\,,\,2\,,\,30\,,\,1\,,\,\ldots\,,\,c_h\,,\,\ldots]$$

with, inter alia

$$\begin{aligned} c_{17} &= 22986 \\ c_{33} &= 1501790 \\ c_{59} &= 35657 \\ c_{81} &= 49405 \\ c_{103} &= 53460 \\ c_{121} &= 16467250 \\ c_{139} &= 48120 \\ c_{161} &= 325927 \end{aligned}$$

These examples are admittedly quite exceptional. There is no reason to believe that the continued fraction expansions of nonquadratic algebraic irrationals generally do anything other than to faithfully follow Khintchine's Law as detailed below. Indeed experiment suggests that this is even true for parts, short relative to the length of the period, of the expansions of quadratic irrationals. Large partial quotients are statistical accident and warrant the comment that exception attracts. Mind you, the final example is, as it were, too accident prone to just be dismissed as statistical fluctuation. Its large partial quotients observed by Brillhart are more than just happenstance. This is beautifully explained by Stark [12]. After the initial excitement we detail above, the expansion settles down to normalcy.

1.4 The transformation T on $\alpha\colon\ 0 < \alpha < 1$ so that $T\alpha = \alpha^{-1} - \lfloor \alpha^{-1} \rfloor$ yields the complete quotients of α. It can be seen that generally

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} f(T^n\alpha) = \int_0^1 f(x)\,d\mu, \text{ with } d\mu = \frac{1}{\log 2}\,\frac{dx}{1+x}\,,$$

so on taking f as the characteristic function of the interval $(\frac{1}{k+1}, \frac{1}{k}]$ we find that for a random real $\alpha = [c_0\,,\,c_1\,,\,\ldots]$ the probability that some given partial quotient c_h equals k is

$$\frac{1}{\log 2}\int_{1/(k+1)}^{1/k}\frac{dx}{1+x}=\frac{1}{\log 2}\big(\log\frac{k+1}{k}-\log\frac{k+2}{k+1}\big),\quad k=1,2,\ldots$$

Thus, for example, almost all real α have some 41% of their partial quotients equal to 1. The interested reader will find an extended discussion of this and related matters in Knuth [K], §s 4.5.2–3. Remarkably, Gauß had already guessed the correct form of the invariant measure for the transformation T accounting for Khintchine's Law as just described.

Of course, we do not know whether algebraic numbers of degree greater than 2 behave as do almost all real numbers. We believe that to be so, but there are neither theorems nor examples. Thus it is an open question whether all, or indeed any, algebraic numbers of degree 3 or more have unbounded partial quotients.

2 AN INTRODUCTION TO CONTINUED FRACTIONS

As we have already remarked, a continued fraction is an object of the shape

$$c_0+\cfrac{1}{c_1+\cfrac{1}{c_2+\cfrac{1}{c_3+\ddots}}}$$

which we abbreviate as

$$[c_0\,,\,c_1\,,\,c_2\,,\,c_3\,,\,\ldots]\,.$$

2.1 Virtually all principles of the subject are revealed by the following correspondence:

If a sequence $c_0, c_1, c_2, \ldots$ defines the sequences (p_h) and (q_h) by

$$(1)\quad \begin{pmatrix}c_0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}c_1 & 1\\ 1 & 0\end{pmatrix}\cdots\cdots\begin{pmatrix}c_h & 1\\ 1 & 0\end{pmatrix}=\begin{pmatrix}p_h & p_{h-1}\\ q_h & q_{h-1}\end{pmatrix}\quad \text{for } h=0,\,1,\,2,\,\ldots$$

then

$$\frac{p_h}{q_h}=[c_0\,,\,c_1\,,\,\ldots\,,\,c_h]\quad \text{for } h=0,\ 1,\ 2,\ \ldots\ .$$

Conversely the sequence of convergents p_h/q_h define matrices

$$\begin{pmatrix} p_h & p_{h-1} \\ q_h & q_{h-1} \end{pmatrix}$$

which decompose as above as a product of matrices displaying the sequence of partial quotients (c_h).

Since a quotient p/q leaves the pair (p,q) ill-defined we are to interpret that last remark sympathetically in the sense that the claim is true for some choice of p and q (in practice, with the two relatively prime and q positive).

The indicated correspondence between continued fractions and special products of 2×2 matrices is readily established by an inductive argument. Notice firstly that the sequence of partial quotients (c_h) defines the sequences (p_h) and (q_h) appearing in the first column of the matrix product. Since the empty product of 2×2 matrices is the identity matrix, we are committed to

$$\begin{pmatrix} p_{-1} & p_{-2} \\ q_{-1} & q_{-2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{2}$$

We may then readily verify by induction on h that the second column of the product indeed has the alleged entries. Thus we have the recursive formulae

$$\begin{aligned} p_{h+1} &= c_{h+1}p_h + p_{h-1} \\ q_{h+1} &= c_{h+1}q_h + q_{h-1} . \end{aligned} \tag{3}$$

We verify the principal claim by induction on the *number* $h+1$ of matrices appearing on the left in the product. The claim is easily seen true for $h=0$ since, indeed $p_0 = c_0$ and $q_0 = 1$. Accordingly, we suppose that

$$\begin{pmatrix} c_1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} c_2 & 1 \\ 1 & 0 \end{pmatrix}\cdots\cdots\begin{pmatrix} c_h & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix} \text{ for } h = 0,\, 1,\, 2,\, \ldots$$

if and only if

$$\frac{x_h}{y_h} = [c_1\,, c_2\,, \ldots\,, c_h] \text{ for } h = 0,\, 1,\, 2,\, \ldots \quad ,$$

noting that this is a case of just h matrices.

But

$$\begin{pmatrix} p_h & p_{h-1} \\ q_h & q_{h-1} \end{pmatrix} = \begin{pmatrix} c_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_h & x_{h-1} \\ y_h & y_{h-1} \end{pmatrix} = \begin{pmatrix} c_0x_h + y_h & c_0x_{h-1} + y_{h-1} \\ x_h & x_{h-1} \end{pmatrix}$$

entails

$$\frac{p_h}{q_h} = c_0 + \frac{y_h}{x_h} = c_0 + \frac{1}{[c_1\,,\,\ldots\,,\,c_h]} = [c_0\,,\,c_1\,,\,\ldots\,,\,c_h]\,, \tag{4}$$

verifying the claim by induction.

2.2 Taking determinants in the correspondence immediately yields the fundamental formula

$$p_nq_{n-1} - p_{n-1}q_n = (-1)^{n+1} \quad \text{or} \quad \frac{p_n}{q_n} = \frac{p_{n-1}}{q_{n-1}} + (-1)^{n-1}\frac{1}{q_{n-1}q_n}\,. \tag{5}$$

It is then immediate that

$$\frac{p_n}{q_n} = c_0 + \frac{1}{q_0q_1} - \frac{1}{q_1q_2} + \cdots + (-1)^{n-1}\frac{1}{q_{n-1}q_n}\,. \tag{6}$$

Almost invariably, but not always, in the sequel the c_i are positive integers — excepting c_0 which may have any sign; indeed those are the criteria for partial quotients to be *admissible*. However, our description is formal and the actual nature of the partial quotients is thus of no matter in much of our description, the next remark being an exception.

It follows from what we have said that one can make sense of nonterminating continued fractions

$$\gamma = [c_0\,,\,c_1\,,\,\ldots]\,,$$

for evidently,

$$\gamma = c_0 + \frac{1}{q_0q_1} - \frac{1}{q_1q_2} + \cdots = c_0 + \sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{q_{n-1}q_n} \tag{7}$$

and, this being an alternating series of terms with decreasing size, the series converges to some real number γ.

2.3 In this context, we recall that the terminating continued fractions

$$\frac{p_h}{q_h} = [c_0\,,\,c_1\,,\,\ldots\,,\,c_h] \qquad h = 0,\,1,\,2,\,\ldots\ldots$$

are called *convergents* of γ and the tails

$$\gamma_{h+1} = [c_{h+1} \, , \, c_{h+2} \, , \, \ldots] \tag{8}$$

are known as its *complete quotients*. Note that we have, formally,

$$\gamma = [c_0 \, , \, c_1 \, , \, \ldots \, , \, c_h, \gamma_{h+1}] \qquad\qquad h = 0, 1, 2, \ldots \, . \tag{9}$$

These remarks immediately yield the approximation properties of the convergents. For we have

$$\gamma - \frac{p_h}{q_h} = (-1)^h \Big(\frac{1}{q_h q_{h+1}} - \frac{1}{q_{h+1} q_{h+2}} + \cdots \Big) \, . \tag{10}$$

This shows that the sequence $(q_h\gamma - p_h)$ alternates in sign and that, in absolute value, it converges monotonically to zero. Less precisely, we see that

$$\left| \gamma - \frac{p_h}{q_h} \right| < \frac{1}{q_h q_{h+1}}$$

and, recalling (3) : $q_{h+1} = c_{h+1} q_h + q_{h-1}$ implies yet less accurately that

$$\left| \gamma - \frac{p_h}{q_h} \right| < \frac{1}{c_{h+1} q_h^2} \, .$$

Thus a convergent yields an exceptionally sharp approximation when the *next* partial quotient is exceptionally large. This is amply illustrated by the example

$$\pi = [3 \, , \, 7 \, , \, 15 \, , \, 1 \, , \, 292 \, , \, 1 \, , \, \ldots] \, ,$$

already cited in our introduction, which with

$$[3 \, , \, 7] = 22/7 \qquad\qquad [3 \, , \, 7 \, , \, 15 \, , \, 1] = 355/113$$

entails

$$\left| \pi - \frac{22}{7} \right| < \frac{1}{15.7^2} \qquad\qquad \left| \pi - \frac{355}{113} \right| < \frac{1}{292.113^2} < 10^{-6}$$

making appropriate the popularity of those rational approximations to π.

2.4 We now return to the beginning. Noting that

$$\gamma = [c_0 \, , \, c_1 \, , \, \ldots] = c_0 + \frac{1}{[c_1 \, , \, c_2 \, , \, \ldots]}$$

we see that

$$c_0 = \lfloor \gamma \rfloor$$

and

$$\gamma_1 = [c_1\,,\,c_2\,,\,\ldots] = (\gamma - c_0)^{-1}\,.$$

The general step in the continued fraction algorithm is

$$c_h = \lfloor \gamma_h \rfloor \text{ and } \gamma_{h+1} = (\gamma_h - c_h)^{-1} \qquad h = 0,\,1,\,2,\,\ldots$$

An infinite partial quotient terminates the expansion. Since

$$[c_0\,,\,c_1\,,\,\ldots\,,\,c_h]$$

is rational it is evident that if the continued fraction of some γ terminates then that γ is rational. Conversely, since, as is plain from (5), p_h and q_h are relatively prime, and, since by (3) the sequences $(|p_h|)$ and (q_h) are both monotonic increasing, it follows that if γ is rational then its continued fraction does terminate. Indeed, for a rational $\gamma = a/b$, the continued fraction algorithm is just the Euclidean algorithm. That is (setting $a_{h+1} = b_h$):

$$\begin{aligned}
a &= c_0 b \;\; + b_1 &\qquad& 0 \le b_1 < b \\
a_1 &= c_1 b_1 \; + b_2 && 0 \le b_2 < b_2 \\
a_2 &= c_2 b_2 \; + b_3 && 0 \le b_3 < b_2 \\
&\;\;\vdots \\
a_h &= c_h b_h
\end{aligned}$$

corresponds to

$$\frac{a}{b} = [c_0\,,\,a_1\,,\,\ldots\,,\,c_h] \text{ and } \gcd(a,b) = d = c_h$$

and as we have mentioned, explains the term ‘partial quotient’. Since $a/b = p_h/q_h$ with $\gcd(p_h,q_h) = 1$ we must have $dp_h = a$ and $dq_h = b$. Moreover, by (5)

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} \text{ so } aq_{n-1} - bp_{n-1} = (-1)^{n-1} d\,,$$

and this displays the greatest common divisor as a $\mathbb{Z}$-linear combination of a and b. By $|p_{n-1}| < |p_n|$ and $q_{n-1} < q_n$ it follows that this combination is minimal.

2.5 The entire matter of continued fractions of real numbers could have been introduced using the following:

A rational p'/q' with $\gcd(p',q')=1$ is a convergent of γ if and only if

$$|q'\gamma - p'| < |q\gamma - p| \text{ for all integers } q < q' \text{ and } p\,.$$

To see this suppose that n is chosen so that $q_{n-1} < q < q_n$. Then, by the unimodularity of the matrix

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$$

there are integers a and b so that

$$\begin{aligned} ap_{n-1} + bp_n =& p \\ aq_{n-1} + bq_n =& q \end{aligned}$$

and, necessarily, $ab < 0$. Multiplying by γ and subtracting yields

$$q\gamma - p = a(q_{n-1}\gamma - p_{n-1}) + b(q_n\gamma - p_n)\,.$$

But, by (10), we have $(q_{n-1}\gamma - p_{n-1})(q_n\gamma - p_n) < 0$. Hence

$$|q\gamma - p| = |a||q_{n-1}\gamma - p_{n-1}| + |b||q_n\gamma - p_n|\,,$$

and plainly $|q\gamma - p| > |q_n\gamma - p_n|$ as asserted.

The preceding proposition asserts that the convergents of γ are exactly those quantities yielding the *locally best approximations* to γ. One can develop the entire theory, working backwards in the present program, from the notion of locally best approximation; once again, the formula (5) plays the fundamental role.

Moreover, one has the following useful criterion due to Lagrange:

If

$$|q\gamma - p| < \frac{1}{2q}$$

then p/q is a convergent of γ. Note that this condition is sufficient but not necessary.

By our previous remark it suffices to show that $|q\gamma - p|$ is a locally best approximation. To see that is so take integers r, s with $0 < s < q$ and notice that

$$1 \le |qr - ps| = |s(q\gamma - p) - q(s\gamma - r)| \le s|q\gamma - p| + q|s\gamma - r| \\ \le \frac{s}{2q} + q|s\gamma - r|\,.$$

So certainly $q|s\gamma - r| \ge 1 - s/2q > 1/2$ and it follows that $|q\gamma - p| < |s\gamma - r|$ as claimed.

Incidentally, this argument shows that when, for example, dealing with continued fractions of formal power series, one already has that p/q is a convergent of f if and only if $|qf - p| < |q^{-1}|$.

2.6 We conclude by applying the matrix correspondence to develop a formulaire: From

$$\gamma = [c_0\,,\,c_1\,,\,\ldots\,,\,c_h\,,\,\gamma_{h+1}] \longleftrightarrow \begin{pmatrix} p_h & p_{h-1} \\ q_h & q_{h_1} \end{pmatrix} \begin{pmatrix} \gamma_{h+1} & 1 \\ 1 & 0 \end{pmatrix}$$

we have

$$\gamma = \frac{\gamma_{h+1}p_h + p_{h-1}}{\gamma_{h+1}q_h + q_{h-1}} \text{ and } \gamma_{h+1} = -\frac{q_{h-1}\gamma - p_{h-1}}{q_h\gamma - p_h}\,.$$

Hence

$$\text{(11)} \qquad \gamma - \frac{p_h}{q_h} = \frac{\gamma_{h+1}p_h + p_{h-1}}{\gamma_{h+1}q_h + q_{h-1}} - \frac{p_h}{q_h} = (-1)^{h-1}\frac{1}{q_h(\gamma_{h+1}q_h + q_{h-1})}\,.$$

Transposition of

$$\begin{pmatrix} c_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots\cdots \begin{pmatrix} c_h & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_h & p_{h-1} \\ q_h & q_{h-1} \end{pmatrix}$$

yields

$$\frac{p_h}{p_{h-1}} = [c_h\,,\,c_{h-1}\,,\,\ldots\,,\,c_0] \text{ and } \frac{q_h}{q_{h-1}} = [c_h\,,\,c_{h-1}\,,\,\ldots\,,\,c_1]\,.$$

Hence

$$\text{(12)} \quad -\gamma_{h+1} = \frac{-\gamma q_{h-1} + p_{h-1}}{\gamma q_h + p_h} \longleftrightarrow -\gamma_{h+1} = [0\,,\,c_h\,,\,c_{h-1}\,,\,\ldots\,,\,c_0 - \gamma]\,.$$

3 CONTINUED FRACTIONS OF ALGEBRAIC NUMBERS

We suppose that γ is presented as a real zero of a polynomial

$$f(X) = a_0X^r + a_1X^{r-1} + \cdots + a_r$$

with $a_0, a_1, \ldots, a_r \in \mathbb{Z}$ and $\gcd(a_0, \ldots, a_r) = 1$, and that to distinguish it from other zeros of f we are told that

$$\gamma = [c_0\,,\, c_1\,,\, \ldots\,,\, c_m\,,\, \ldots]\,.$$

We can suppose that $a_0 > 0$, and that γ is a simple zero of f since otherwise we will choose to deal with the gcd of f and f', a polynomial of smaller degree. Thus we may locate γ by the fact that f suffers a change of sign in an interval containing γ.

Our method is, in principle, to compute a sequence of polynomials

$$f_h(X) = a_{h,0}X^r + a_{h,1}X^{r-1} + \cdots + a_{h,r}\,,\ \ h = 0,\ 1,\ \ldots\,,$$

with rational integer coefficients $a_{h,i}$ and $a_{h,0} > 0$ having the complete quotients γ_h as a zero. Indeed, we will have, sequentially,

$$f_{h+1}(X) = \pm X^r f_h(X^{-1} + c_h)\,.$$

A core observation is that we will eventually obtain a *reduced* polynomial (and then the minus sign is always appropriate):

Proposition 1 *The zeros β_h (say) of f_h distinct from γ_h all satisfy $|\beta_h| < 1$ and $-1 < \operatorname{Re}\beta_h < 0$ for h sufficiently large.*

Proof. Suppose firstly that $\operatorname{Re}\beta_m < 0$. Evidently $\operatorname{Re}(\beta_m - c_m) < -1$ since $c_m \geq 1$, so $-1 < \operatorname{Re}\beta_{m+1} < 0$ and $|\beta_{m+1}| < 1$. Secondly, if $\operatorname{Re}\beta_{m-1} < c_{m-1}$ then clearly $\operatorname{Re}\beta_m < 0$. Thirdly, if $\operatorname{Re}\beta_{m-2} > c_{m-2} + 1$, then evidently $\operatorname{Re}\beta_{m-1} < 1 < c_{m-1}$. Finally, suppose that $c_n < \operatorname{Re}\beta_n < c_n + 1$. Then β shares its first n partial quotients with γ and that entails $|\gamma - \beta| < q_{n-1}^{-2}$. But that is eventually absurd because β is distinct from γ.

Note. Given f, there evidently is an effective upper bound for m, as described, in terms of the degree r and the height of the given polynomial f. In practice, the data $\gamma = [c_0\,,\, c_1\,,\, \ldots\,,\, c_s\,,\, \ldots]$ required to identify the zero γ uniquely may well essentially suffice to yield a reduced polynomial.

In the sequel we assume, as we evidently may, that f is reduced from the outset; thus for each $h = 0,\ 1,\ \ldots$ $\gamma = \gamma_0,\ \gamma_1,\ \ldots$ is inter alia the unique zero > 1 of $f = f_0,\ f_1,\ \ldots\,$.

Having obtained the complete quotient γ_{h+1}, thus having

$$\gamma = [c_0\,,\, c_1\,,\, \ldots\,,\, c_h\,,\, \gamma_{h+1}]\,,$$

we now turn to the problem of finding the next partial quotient c_{h+1}. Recalling that γ_{h+1} is a zero of

$$f_{h+1}(X) = a_{h+1,0}X^r + a_{h+1,1}X^{r-1} + \cdots + a_{h+1,r}\,,$$

we may use the fact that f_{h+1} is reduced to estimate $\gamma_{h+1} \approx -a_{h+1,1}/a_{h+1,0}$, knowing that this cannot involve an error greater than $r-1$. But we can do much better:

Proposition 2 *To avoid notational clutter set* $q_h = q$, $p_h/q_h = x$, *and* $q_{h-1} = q'$. *Then*

$$\gamma_{h+1} = \frac{(-1)^{h+1}}{q^2}\frac{f'(x)}{f(x)} - \frac{q'}{q} + \frac{(-1)^h}{q^2}\sum_{\substack{f(\beta)=0\\ \beta\neq\gamma}}\frac{1}{x-\beta}\,.$$

Proof. Trivially

$$\gamma_{h+1} = \sum_{f_{h+1}(\alpha_{h+1})=0}\alpha_{h+1} - \sum_{\substack{f_{h+1}(\beta_{h+1})=0\\ \beta_{h+1}\neq\gamma_{h+1}}}\beta_{h+1}\,.$$

But by (11), whereby $\alpha - x = (-1)^h/q(q\alpha_{h+1}+q')$, we obtain

$$\alpha_{h+1} = \frac{(-1)^{h+1}}{q^2}\frac{1}{x-\alpha} - q'/q\,.$$

Then (as is readily seen by taking the logarithmic derivative)

$$\frac{f'(x)}{f(x)} = \sum_{f(\alpha)=0}\frac{1}{x-\alpha}$$

yields the assertion.

Remark. It is not difficult to obtain increasingly precise estimates for the final *error term* in (13). Indeed,

$$\lim_{X\to\gamma}\left(\frac{f'(X)}{f(X)} - \frac{1}{X-\gamma}\right) = \frac{f''(\gamma)}{2f'(\gamma)}$$

entails that in

$$\sum_{\substack{f(\beta)=0\\ \beta\neq\gamma}}\frac{1}{x-\beta} = \frac{f''(x)}{2f'(x)} + \left(\frac{f'(x)}{f(x)} - \frac{1}{x-\gamma} - \frac{f''(x)}{2f'(x)}\right)\,,$$

the error term on the right is of order $O(x-\gamma)=O(q^{-2})$. One can iterate this idea[3] to obtain an expansion

$$\sum_{\substack{f(\beta)=0\\ \beta\neq\gamma}}\frac{1}{x-\beta}=\frac{f''(x)}{2f'(x)}+(x-\gamma)\frac{3\big(f''(x)\big)^2+2f'(x)f'''(x)}{12\big(f'(x)\big)^2}+O\big((x-\gamma)^2\big)$$

allowing one, in principle, to have an arbitrarily small error in one's estimate for γ_{h+1} in terms of p_h/q_h, q_h and q_{h-1}. However, in practice the error term is not needed at all, at any rate once h is large.

Indeed, if the error term were to affect the true value of $\lfloor\gamma_{h+1}\rfloor$ then γ_{h+1} is extremely close to an integer and evidently either $c_{h+2}=O(q_h^2)$ or $c_{h+2}=1$ and $c_{h+3}=O(q_h^2)$. In either case we obtain a rational p/q approximating γ to an accuracy $O(q^{-4})$. By Roth's theorem [11] this is not possible for arbitrarily large q. Of course Roth's theorem is ineffective. However for cubic irrationals γ, thus when $r=3$, we have an effective guarantee from Liouville's theorem and for higher degree irrationals we may, if necessary, choose to estimate the error term to appropriate accuracy, as proposed above, to obtain a definite value for $\lfloor\gamma_{h+1}\rfloor$. In summary:

Theorem 3 *Define the* $(r+1)\times(r+1)$ *matrices* $M(c)$ *by*

$$M(c)=-\left\|\binom{r-j}{i}c^{r-j-i}\right\|_{0\le i,j\le r}$$

and denote by $a(h)$ *the column* $(a_{h,0},\dots,a_{h,r})'$ *of coefficients of the polynomial*

$$f_h(X)=a_{h,0}X^r+a_{h,1}X^{r-1}+\cdots+a_{h,r}$$

given by

$$M(c_h)a(h)=a(h+1)\qquad h=0,1,\dots\,.$$

If $\gamma>1$ *is the unique positive zero of the polynomial* $f(X)=f_0(X)$ *and the positive integers* c_h *satisfy* $f(c_h)<0<f(c_h+1)$ *then* γ *has the continued fraction expansion*

$$\gamma=[c_0\,,\,c_1\,,\,c_2\,,\,\dots]$$

[3] As was pointed out to one of us by David Cantor.

and the partial quotients c_h *are given by* $c_{h+1} = \lfloor \gamma_{h+1} \rfloor$ *where*

$$\begin{aligned}(13) \quad \gamma_{h+1} &= -\frac{a_{h,1}}{a_{h,0}} + (r-1)\frac{q_{h-1}}{q_h} + \frac{(-1)^h}{q_h^2} \sum_{\substack{f(\beta)=0 \\ \beta\neq\gamma}} \frac{1}{\frac{p_h}{q_h} - \beta} \\ &= \frac{(-1)^{h+1}}{q_h^2} \frac{f'(\frac{p_h}{q_h})}{f(\frac{p_h}{q_h})} - \frac{q_{h-1}}{q_h} + \frac{(-1)^h}{q_h^2} \sum_{\substack{f(\beta)=0 \\ \beta\neq\gamma}} \frac{1}{\frac{p_h}{q_h} - \beta} .\end{aligned}$$

Example. It is easy to confirm that $X^3 - 2$ is reduced after the initial step of the algorithm and that indeed $\sqrt[3]{2} = [1 \ , c_1 \ , c_2 \ , c_3 \ , \ldots]$, with

$$c_{h+1} = \left\lfloor \frac{(-1)^{h+1}}{q_h} \frac{3p_h^2}{p_h^3 - 2q_h^3} - \frac{q_{h-1}}{q_h} \right\rfloor \qquad h = 0, 1, 2, \ldots ,$$

where

$$\begin{pmatrix} p_{h+1} & p_h \\ q_{h+1} & q_h \end{pmatrix} = \begin{pmatrix} c_{h+1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p_h & p_{h-1} \\ q_h & q_{h-1} \end{pmatrix} \text{ and } \begin{pmatrix} p_0 & p_{-1} \\ q_0 & q_{-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} .$$

Thus, contrary to received wisdom, there may well be 'a formula' for the continued fraction expansion of an algebraic number of degree greater than 2. However, such a formula does not necessarily usefully increase our understanding of the nature of the partial quotients of such a number. To make that manifest one notes that the notorious $3x + 1$ problem is readily recast as a similar, but simpler formula

$$x_{h+1} = (1 + x_h - 2\lfloor x_h/2 \rfloor) x_h - \lfloor x_h/2 \rfloor \text{ with } x_0 \in \mathbb{Z}.$$

John Conway has shown that it is undecidable whether general games of this sort are unbounded or cycle.

4 COMMENTS

It is not of course a new thought that one should develop the continued fraction of an algebraic number by finding the sequence of defining polynomials of its complete quotients. Indeed all we say, other than the 'explicit' formulae for the partial quotients, is applied in the computations of Lang and Trotter [5] and presumably in earlier work [1], [7] and [10]. Our observations were in part motivated by our desire not to inconvenience a HP-41 calculator in guessing the partial quotients. In the interim, during the 'maturation' of our thoughts, [3] reports apparently related formulas.

Nonetheless, it does not seem to be widely appreciated — presumably because the cited authors commence with a reduced polynomial without commenting extensively on that simplification — that the algorithm leads to a reduced polynomial generalising the phenomenon well known in the quadratic case, where reduction signals the commencement of periodicity. Mind you, that the algorithm leads to a reduced polynomial could be better known[4]. In essence, it dates back at least to Vincent's Theorem of 1836. In effect, Vincent shows that applying our algorithm with arbitrary positive integer 'partial quotients' eventually yields a polynomial with at most one positive zero. Of this result Uspensky [13] writes: 'This remarkable theorem was published by Vincent in 1836 in the first issue of *Liouville's Journal*, but later [was] so completely forgotten that no mention of it is found even in such a capital work as the *Enzyclopädie der mathematischen Wissenschaften*. Yet Vincent's theorem is the basis of the very efficient method for separating real roots ... '. The issue of identifying the particular real zeros to be expanded, with which we deal just cursorily, is considered in extenso in [2].

Proposition 2 and the subsequent remark yields very precise rational approximations for the complete quotient. It is immediate that a substantial sequence of the partial quotients of such a rational coincides with the partial quotients of γ. Generally speaking, using just Proposition 2, one can expect to be able to double the number of partial quotients already obtained by expanding the rational approximation. Whether a partial quotient x obtained in this way is indeed a partial quotient of γ can be checked by observing that

$$f(x) \approx (x-\gamma)f'(\gamma)\,,$$

so that x certainly is a partial quotient whenever $f(x)$ is surprisingly smaller than $f'(\gamma)$ relative to the height of x. However, since the difficulty in applying the continued fraction algorithm is the exponential growth of the p_n and q_n, it is not clear whether the ability to leap through the expansion is all that material. Moreover, the rational approximation is of the same quality as the approximation yielded by the naïve use of Newton approximation.

[4] We are indebted for this information to Emery Thomas.

REFERENCES

[1] A. D. Bryuno, ‘Continued fraction expansion of algebraic numbers’, *Zh. Vychisl. Mat. i. Mat. Fiz.* **4** (1964), 211–221. English translation, *USSR Comput. Math. and Math. Phys.* **4** (1964), 1–15.

[2] David G. Cantor, Paul H. Galyean and Horst Günter Zimmer, ‘A continued fraction algorithm for real algebraic numbers’, *Math. Comp.* **26** (1972), 785–791.

[3] E. A. Dzenskevich and A. P. Shapiro, ‘On the expansion of irrationalities of third and fourth degree into a continued fraction’, *Algorithmic and numerical problems in algebra and number theory* (Russian), 3–9, 87, *Akad. Nauk. SSSR*, Dal’nevostochn. Otdel., Vladivostok, 1987.

[4] Donald E. Knuth, *Seminumerical algorithms* $=$ *The art of computer programming* Volume **2** Addison-Wesley.

[5] Serge Lang and Hale Trotter, ‘Continued fractions for some algebraic numbers’, *J. für Math.* **255** (1972), 112–134; Addendum: *ibid.* 219–220.

[6] L. Lewin, *Dilogarithms and associated functions*, Macdonald, London (1958).

[7] J. von Neumann and B. Tuckerman, ‘Continued fraction expansion of $2^{\frac{1}{3}}$ ’, *Math. Tables Aids Comput.* **9** (1955), 23–24.

[8] Alfred van der Poorten, ‘A proof that Euler missed … Apéry’s proof of the irrationality of $\zeta(3)$; An informal report’, *The Mathematical Intelligencer* **1** (1979), 195-203.

[9] A. J. van der Poorten, ‘Some wonderful formulas … an introduction to polylogarithms’, *Queen’s Papers in Pure and Applied Mathematics* **54** (1980), 269–286.

[10] R. D. Richtmyer, M. Devaney and N. Metropolis, ‘Continued fraction expansions of algebraic numbers’, *Numer. Math.* **4** (1962), 68–84.

[11] K. F. Roth, ‘Rational approximations to algebraic numbers’, *Mathematika* **2** (1955), 1–20; Corrigendum, 168.

[12] H. M. Stark, ‘An explanation of some exotic continued fractions found by Brillhart’, in A. O. L. Atkin and B. J. Birch eds., *Computers in Number Theory*, Academic Press, London and New York, 1971, 21–35.

[13] J. V. Uspensky, *Theory of Equations*, McGraw-Hill Book Company, 1948.

Acknowledgements

The work of the second author was supported in part by grants from the Australian Research Council and by a research agreement with Digital Equipment Corporation.

1991 Mathematics subject classification: 11A55, 11Y65, 11J70.