Continued Fractions and the Euclidean Algorithm

Lecture notes prepared for MATH 326, Spring 1997
Department of Mathematics and Statistics
University at Albany

William F. Hammond

Table of Contents

1 Introduction

Continued fractions offer a means of concrete representation for arbitrary real numbers. The continued fraction expansion of a real number is an alternative to the representation of such a number as a (possibly infinite) decimal.

The reasons for including this topic in the course on Classical Algebra are:

(i) The subject provides many applications of the method of recursion.
(ii) It is closely related to the Euclidean algorithm and, in particular, to "Bezout's Identity".
(iii) It provides an opportunity to introduce the subject of *group theory* via the 2-dimensional unimodular group $\mathrm{GL}_2(\mathbf{Z})$.

2 The continued fraction expansion of a real number

Every real number x is represented by a point on the real line and, as such, falls between two integers. For example, if n is an integer and

$$n \ \leq \ x \ < \ n+1 \ ,$$

x falls between n and $n+1$, and there is one and only one such integer n for any given real x. In the case where x itself is an integer, one has $n = x$. The integer n is sometimes called the *floor* of x, and one often introduces a notation for the floor of x such as

$$n \ = \ [x] \ .$$

Examples:

1.
$$-2 \ = \ [-1.5]$$

2.
$$3 \ = \ [\pi]$$

For any real x with $n = [x]$ the number $u = x - n$ falls in the *unit interval* I consisting of all real numbers u for which $0 \leq u < 1$.

Thus, for given real x there is a unique decomposition

$$x \ = \ n + u$$

where n is an integer and u is in the unit interval. Moreover, $u = 0$ if and only if x is an integer. This decomposition is sometimes called the *mod one decomposition* of a real number. It is the first step in the process of expanding x as a continued fraction.

The process of finding the continued fraction expansion of a real number is a recursive process that procedes one step at a time. Given x one begins with the mod one decomposition

$$x \ = \ n_1 + u_1 \ ,$$

where n_1 is an integer and $0 \leq u_1 < 1$.

If $u_1 = 0$, which happens if and only if x is an integer, the recursive process terminates with this first step. The idea is to obtain a sequence of integers that give a precise determination of x.

If $u_1 > 0$, then the reciprocal $1/u_1$ of u_1 satisfies $1/u_1 > 1$ since u_1 is in I and, therefore, $u_1 < 1$. In this case the second step in the recursive determination of the continued fraction expansion of x is to apply the mod one decomposition to $1/u_1$. One writes

$$1/u_1 \ = \ n_2 + u_2 \ ,$$

where n_2 is an integer and $0 \leq u_2 < 1$. Combining the equations that represent the first two steps, one may write

$$x \ = \ n_1 + \frac{1}{n_2 + u_2} \ .$$

Either $u_2 = 0$, in which case the process ends with the expansion

$$x = n_1 + \frac{1}{n_2} ,$$

or $u_2 > 0$. In the latter case one does to u_2 what had just been done to u_1 above under the assumption $u_1 > 0$. One writes

$$1/u_2 = n_3 + u_3 ,$$

where n_3 is an integer and $0 \leq u_3 < 1$. Then combining the equations that represent the first three steps, one may write

$$x = n_1 + \frac{1}{n_2 + \frac{1}{n_3+u_3}} .$$

After k steps, if the process has gone that far, one has integers $n_1, n_2, \ldots, n_k$ and real numbers $u_1, u_2, \ldots, u_k$ that are members of the unit interval I with $u_1, u_2, \ldots, u_{k-1}$ all positive. One may write

$$x = n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{\cdots + \frac{1}{n_k+u_k}}}} .$$

Alternatively, one may write

$$x = [n_1, n_2, n_3, \ldots, n_k + u_k] .$$

If $u_k = 0$, the process ends after k steps. Otherwise, the process continues at least one more step with

$$1/u_k = n_{k+1} + u_{k+1} .$$

In this way one associates with any real number x a sequence, which could be either finite or infinite, $n_1, n_2, \ldots$ of integers. This sequence is called the *continued fraction expansion of* x.

Convention. *When* $[n_1, n_2, ...]$ *is called a* ***continued fraction****, it is understood that all of the numbers* n_j *are integers and that* $n_j \geq 1$ *for* $j \geq 2$.

3 First examples

$$\begin{aligned}
\frac{15}{11} &= 1 + \frac{4}{11} \\
&= 1 + \frac{1}{\frac{11}{4}} \\
&= 1 + \frac{1}{2 + \frac{3}{4}} \\
&= 1 + \frac{1}{2 + \frac{1}{\frac{4}{3}}} \\
&= 1 + \frac{1}{2 + \frac{1}{1+\frac{1}{3}}} \\
&= [1, 2, 1, 3] .
\end{aligned}$$

$$\begin{aligned}
\sqrt{10} &= 3 + \frac{1}{\frac{1}{\sqrt{10}-3}} \\
&= 3 + \frac{1}{\sqrt{10}+3} \\
&= 3 + \frac{1}{6 + \frac{1}{\frac{1}{\frac{1}{\sqrt{10}-3}}}} \\
&= 3 + \frac{1}{6 + \frac{1}{\sqrt{10}+3}} \\
&= 3 + \frac{1}{6 + \frac{1}{6+\frac{1}{\cdots}}} \\
&= [3, 6, 6, 6, \ldots] \; .
\end{aligned}$$

$$\begin{aligned}
[2,3,5,2] &= 2 + \frac{1}{[3,5,2]} \\
&= 2 + \frac{1}{3 + \frac{1}{[5,2]}} \\
&= 2 + \frac{1}{3 + \frac{1}{5+\frac{1}{2}}} \\
&= 2 + \frac{1}{3 + \frac{1}{\frac{11}{2}}} \\
&= 2 + \frac{1}{3 + \frac{2}{11}} \\
&= 2 + \frac{1}{\frac{35}{11}} \\
&= 2 + \frac{11}{35} \\
&= \frac{81}{35} \; .
\end{aligned}$$

Let

$$x \;=\; 1 + \frac{1}{2 + \frac{1}{3+\frac{1}{2+\frac{1}{3+\frac{1}{2+\ldots}}}}} \quad .$$

In this case one finds that

$$x \;=\; 1 + \frac{1}{y} \; ,$$

where

$$y \;=\; 2 + \frac{1}{3 + \frac{1}{2+\frac{1}{3+\frac{1}{2+\ldots}}}} \quad .$$

Further reflection shows that the continued fraction structure for y is self-similar:

$$y \;=\; 2 + \frac{1}{3 + \frac{1}{y}} \quad .$$

This simplifies to

$$y \;=\; \frac{7y+2}{3y+1}$$

and leads to the quadratic equation

$$3y^2 - 6y - 2 \;=\; 0$$

with discriminant 60. Since $y > 2$, one of the two roots of the quadratic equation cannot be y, and, therefore,

$$y \;=\; \frac{3+\sqrt{15}}{3} \quad .$$

Finally,

$$x \;=\; \frac{\sqrt{15}-1}{2} \quad .$$

The idea of the calculation above leads to the conclusion that any continued fraction $[n_1, n_2, \ldots]$ that eventually repeats is the solution of a quadratic equation with positive discriminant and integer coefficients. The converse of this statement is also true, but a proof requires further consideration.

4 The case of a rational number

The process of finding the continued fraction expansion of a rational number is essentially identical to the process of applying the Euclidean algorithm to the pair of integers given by its numerator and denominator.

Let $x = a/b,\ b > 0,$ be a representation of a rational number x as a quotient of integers a and b. The mod one decomposition

$$\frac{a}{b} \;=\; n_1 + u_1\,, \quad u_1 \;=\; \frac{a - n_1 b}{b}$$

shows that $u_1 = r_1/b$, where r_1 is the remainder for division of a by b. The case where $u_1 = 0$ is the case where x is an integer. Otherwise $u_1 > 0$, and the mod one decomposition of $1/u_1$ gives

$$\frac{b}{r_1} \;=\; n_2 + u_2\,, \quad u_2 \;=\; \frac{b - n_2 r_1}{r_1} \quad .$$

This shows that $u_2 = r_2/r_1$, where r_2 is the remainder for division of b by r_1. Thus, the successive quotients in Euclid's algorithm are the integers $n_1, n_2, \ldots$ occurring in the continued fraction. Euclid's algorithm terminates after a finite number of steps with the appearance of a zero remainder. Hence, the continued fraction expansion of every rational number is finite.

Theorem 1. *The continued fraction expansion of a real number is finite if and only if the real number is rational.*

Proof. It has just been shown that if x is rational, then the continued fraction expansion of x is finite because its calculation is given by application of the Euclidean algorithm to the numerator and denominator of x. The converse statement is the statement that every finite continued fraction represents a rational number. That statement will be demonstrated in the following section.

5 The symbol $[t_1,\, t_2,\, \ldots,\, t_r]$

For arbitrary real numbers $t_1, t_2, \ldots, t_r$ with each $t_j \geq 1$ for $j \geq 2$ the symbol $[t_1, t_2, \ldots, t_r]$ is defined recursively by:

$$[t_1] \;=\; t_1$$

$$[t_1, t_2, \ldots, t_r] \;=\; t_1 + \frac{1}{[t_2, \ldots, t_r]}\;. \tag{1}$$

In order for this definition to make sense one needs to know that the denominator in the right-hand side of (1) is non-zero. The condition $t_j \geq 1$ for $j \geq 2$ guarantees, in fact, that $[t_2, \ldots, t_r] > 0$, as one may prove using induction.

It is an easy consequence of mathematical induction that the symbol $[t_1, t_2, \ldots, t_r]$ is a rational number if each t_j is rational. In particular, each finite continued fraction is a rational number. (Note that the symbol $[t_1, t_2, \ldots, t_r]$ is to be called a continued fraction, according to the convention of the first section, only when each t_j is an integer.)

Observe that the recursive nature of the symbol $[t_1, \ldots, t_r]$ suggests that the symbol should be computed in a particular case working from right to left. Consider again, for example, the computation above showing that $[2, 3, 5, 2] \;=\; 81/35$. Working from right to left one has:

$$\begin{aligned}
[2] &= 2 \\
[5, 2] &= 5 + \frac{1}{[2]} = 5 + \frac{1}{2} = \frac{11}{2} \\
[3, 5, 2] &= 3 + \frac{1}{[5, 2]} = 3 + \frac{2}{11} = \frac{35}{11} \\
[2, 3, 5, 2] &= 2 + \frac{1}{[3, 5, 2]} = 2 + \frac{11}{35} = \frac{81}{35}
\end{aligned}$$

There is, however, another approach to computing $[t_1, t_2, \ldots, t_r]$. Let, in fact, $t_1, t_2, \ldots$ be any (finite or infinite) sequence of real numbers. One uses the double recursion

$$p_j \;=\; t_j p_{j-1} + p_{j-2}\,, \;\; j \geq 1\,, \;\; p_0 \;=\; 1\,, \;\; p_{-1} \;=\; 0 \tag{2}$$

to define the sequence $\{p_j\}\,,\; j \geq -1$. The double recursion, differently initialized,

$$q_j \;=\; t_j q_{j-1} + q_{j-2}\,, \;\; j \geq 1\,, \;\; q_0 \;=\; 0\,, \;\; q_{-1} \;=\; 1 \tag{3}$$

defines the sequence $\{q_j\}\,,\; j \geq -1$. Note that $p_1 \;=\; t_1$, $p_2 \;=\; t_1 t_2 + 1$, $\ldots$ and $q_1 \;=\; 1$, $q_2 \;=\; t_2$, $q_3 \;=\; t_2 t_3 + 1$, $\ldots$.

One now forms the matrix

$$M_j \;=\; \begin{pmatrix} p_j & q_j \\ p_{j-1} & q_{j-1} \end{pmatrix} \text{ for } j \geq 0\;. \tag{4}$$

Thus, for example,

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } M_1 = \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix}.$$

It is easy to see that the matrices M_j satisfy the double recursion

$$M_j = \begin{pmatrix} t_j & 1 \\ 1 & 0 \end{pmatrix} M_{j-1}\,, \ j \geq 1 \tag{5}$$

as a consequence of the double recursion formulas for the p_j and q_j. Hence, a simple argument by mathematical induction shows that

$$M_r = \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} t_2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix}, \ r \geq 1\,. \tag{6}$$

This is summarized by:

Proposition 1. *For any sequence $\{t_j\}$, $j \geq 1$ of real numbers, if $\{p_j\}$ and $\{q_j\}$ are the sequences defined by the double recursions (2) and (3), then one has the matrix identity*

$$\begin{pmatrix} p_r & q_r \\ p_{r-1} & q_{r-1} \end{pmatrix} = \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} t_2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix} \tag{7}$$

for each integer $r \geq 1$.

Corollary 1. *One has the identity $p_r q_{r-1} - q_r p_{r-1} = (-1)^r$ for each integer $r \geq 1$.*

Proof. The number $p_r q_{r-1} - q_r p_{r-1}$ is the determinant of the matrix M_r. From the formula (6) the matrix M_r is the product of r matrix factors, each of which has determinant -1. Since the determinant of the product of matrices is the product of the determinants of the factors, it is clear that $\det(M_r) = (-1)^r$.

Corollary 2. *One has the vector identity*

$$\begin{pmatrix} p_r \\ q_r \end{pmatrix} = \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{8}$$

for each integer $r \geq 1$.

Proof. First recall (i) that the product of a matrix and a (column) vector is defined by the relation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix},$$

(ii) that the *transpose* of a matrix is the matrix whose rows are the columns of the given matrix, and (iii) that the transpose operation reverses matrix multiplication. One tranposes both sides of the relation (7) to obtain:

$$\begin{pmatrix} p_r & p_{r-1} \\ q_r & q_{r-1} \end{pmatrix} = \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} t_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix}. \tag{9}$$

To this relation one applies the principle that the first column of any 2×2 matrix is the product of that matrix with the column

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

in order to obtain the column identity (8).

Theorem 2. *For any sequence* $\{t_j\}$, $j \geq 1$ *of real numbers, if* $\{p_j\}$ *and* $\{q_j\}$ *are the sequences defined by the double recursions (2) and (3), and if* $t_j \geq 1$ *for* $j \geq 2$, *then the value of the symbol* $[t_1, \ldots, t_r]$ *is given by the formula*

$$[t_1,\, t_2,\, \ldots,\, t_r] \;=\; \frac{p_r}{q_r} \;\; \textit{for} \;\; r \geq 1 \;. \tag{10}$$

Proof. What is slightly strange about this important result is that while the $\{p_r\}$ and the $\{q_r\}$ are defined by the *front end* recursions, albeit double recursions, (2) and (3), the symbol $[t_1, \ldots, t_r]$ is defined by the *back end* recursion (1). The proof begins with the comment that the right-hand side of (10) does not make sense unless one can be sure that the denominator $q_r \neq 0$. One can show easily by induction on r that $q_r \geq 1$ for $r \geq 1$ under the hypothesis $t_j \geq 1$ for $j \geq 2$.

The proof proceeds by induction on r. If $r = 1$, the assertion of the theorem is simply the statement $t_1 = p_1/q_1$, and, as noted above, $p_1 = t_1$ and $q_1 = 1$. Assume now that $r \geq 2$. By induction we may assume the correctness of the statement (10) for symbols of length $r - 1$, and, therefore, for the symbol $[t_2, \ldots, t_r]$. That case of the statement says that $[t_2, \ldots, t_r]$ must be equal to a/c, where by corollary 2

$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

with

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} t_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} t_r & 1 \\ 1 & 0 \end{pmatrix} .$$

Now by (1)

$$[t_1,\, t_2,\, \ldots,\, t_r] \;=\; t_1 + \frac{1}{a/c} \;=\; t_1 + \frac{c}{a} \;=\; \frac{at_1 + c}{a} \;.$$

But by corollary 2 again

$$\begin{pmatrix} p_r \\ q_r \end{pmatrix} = \begin{pmatrix} t_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} at_1 + c & bt_1 + d \\ a & b \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} at_1 + c \\ a \end{pmatrix} .$$

Hence,

$$\frac{p_r}{q_r} \;=\; \frac{at_1 + c}{a} \;=\; [t_1,\, t_2,\, \ldots,\, t_r] \;\; .$$

6 Application to Continued Fractions

Recall that $[n_1, n_2, \ldots]$ is called a continued fraction only when each n_j is an integer and $n_j \geq 1$ for $j \geq 2$. The sequence $n_1, n_2, \ldots$ may be finite or infinite. The symbol $c_r = [n_1, n_2, \ldots, n_r]$ formed with the first r terms of the sequence, is called the r^{th} *convergent* of the continued fraction. Associated with a given sequence $n_1, n_2, \ldots$ are two sequences $p_1, p_2, \ldots$ and $q_1, q_2, \ldots$ that are given, according to the double recursions (2), (3) of the previous section with $t_j = n_j$.

Proposition 2. *If* $[n_1, n_2, \ldots]$ *is a continued fraction, then the integers* p_r *and* q_r *are coprime for each* $r \geq 1$.

Proof. By Corollary 1 of the previous section $p_r q_{r-1} - q_r p_{r-1} = (-1)^r$. Hence, any positive divisor of both p_r and q_r must divide the left-hand side of this relation, and, therefore, must also divide $(-1)^r$.

Proposition 3. *The difference between successive convergents of the continued fraction* $[n_1, n_2, \dots]$ *is given by the formula*

$$c_r - c_{r-1} \;=\; \frac{(-1)^r}{q_r q_{r-1}} \quad \textit{for} \;\; r \geq 2\,. \tag{11}$$

Proof. According to the theorem (formula 10) at the end of the last section the convergent c_r is given by

$$c_r \;=\; \frac{p_r}{q_r} \quad .$$

Hence,

$$\begin{aligned} c_r - c_{r-1} &= \frac{p_r}{q_r} - \frac{p_{r-1}}{q_{r-1}} \\ &= \frac{p_r q_{r-1} - p_{r-1} q_r}{q_r q_{r-1}} \\ &= \frac{(-1)^r}{q_r q_{r-1}} \quad . \end{aligned}$$

(The last step is by Corollary 1 above.)

Remark 1. *The formula (11) remains true if* $c_r = [t_1, \dots, t_r]$ *where the* t_j *are real numbers subject to the assumption* $t_j \geq 1$ *for* $j \geq 1$*.*

Lemma. *The sequence* $\{q_j\}$ *is a strictly increasing sequence for* $j \geq 2$*.*

Proof. This is easily proved by induction from the recursive definition (3) of the sequence.

Theorem 3. *If* $[n_1, n_2, \dots]$ *is an infinite continued fraction, then the limit*

$$\lim_{r \to \infty} \frac{p_r}{q_r}$$

always exists.

Proof. As one plots the convergents c_r on the line of real numbers, one moves alternately right and left. The formula (11) for the difference between successive convergents elucidates not only the fact of alternate right and left movement but also the fact that each successive movement is smaller than the one preceding. Therefore, one has

$$c_1 < c_3 < c_5 < \dots < c_6 < c_4 < c_2 \quad .$$

Since any strictly increasing sequence of positive integers must have infinite limit, the seqence $q_j q_{j-1}$ has infinite limit, and so the sequence of reciprocals $1/q_j q_{j-1}$ must converge to zero. Hence, the sequences of odd- and even-indexed convergents must have the same limit, which is the limit of the sequence of all convergents.

Definition 1. *The limit of the sequence of convergents of an infinite continued fraction is called the* value *of that continued fraction.*

Theorem 4. *If* $[n_1, n_2, \ldots]$ *is the continued fraction expansion of an irrational number* x, *then*

$$x \;=\; \lim_{r\to\infty} \frac{p_r}{q_r} \;;$$

that is, the value of the continued fraction expansion of a real number is that real number.

Proof. For each $r \geq 1$ the continued fraction expansion $[n_1, n_2, \ldots]$ of x is characterized by the identity

$$(12) \qquad x \;=\; [n_1,\, n_2,\, \ldots,\, n_r + u_r] \;,$$

where u_r is a real number with $0 \leq u_r < 1$. The sequences of p's and q's for the symbol $[n_1, n_2, \ldots, n_r + u_r]$ agree with those for the symbol $[n_1, n_2, \ldots, n_r]$ except for the r^{th} terms. One has by (10)

$$[n_1,\, n_2,\, \ldots,\, n_r + u_r] \;=\; \frac{P_r}{Q_r} \;,$$

where by (3)

$$\begin{aligned} q_r \;&=\; n_r q_{r-1} \;+\; q_{r-2} \\ Q_r \;&=\; (n_r + u_r) q_{r-1} \;+\; q_{r-2} \end{aligned}$$

Hence,

$$Q_r \;=\; q_r \;+\; u_r q_{r-1} \;\;.$$

Therefore, the displacement from c_{r-1} to x is by (11)

$$\frac{(-1)^r}{Q_r q_{r-1}} \;=\; \frac{(-1)^r}{(q_r q_{r-1} + u_r q_{r-1}^2)} \;,$$

which is in the same direction but of smaller magnitude than the displacement from c_{r-1} to c_r. Therefore, x must be larger than every odd-indexed convergent and smaller than every even-indexed convergent. But since all convergents have the same limit, that limit must be x.

7 Bezout's Identity and the double recursion

It has already been observed that the process of finding the continued fraction expansion of a rational number a/b $(b > 0)$, involves the same series of long divisions that are used in the application of the Euclidean algorithm to the pair of integers a and b. Recall that at each stage in the Euclidean algorithm the divisor for the current stage is the remainder from the previous stage and the dividend for the current stage is the divisor from the previous stage, or, equivalently, the dividend for the current stage is the remainder from the second previous stage. The Euclidean algorithm may thus be viewed as a double recursion that is used to construct the sequence of remainders. One starts the double recursion with

$$r_{-1} \;=\; a \;\text{ and }\; r_0 \;=\; b \;\;.$$

At the j^{th} stage one performs long division of r_{j-2} by r_{j-1} to obtain the integer quotient n_j and the integer remainder r_j that satisfies $0 \leq r_j < r_{j-1}$. Thus,

$$(13) \qquad r_j \;=\; r_{j-2} - n_j r_{j-1} \;\;.$$

The Euclidean algorithm admits an additional stage if $r_j > 0$. Since

$$0 \leq r_j < r_{j-1} < \ldots < r_2 < r_1 < r_0 \ = \ b \ ,$$

there can be at most b stages.

One may use the sequence of successive quotients n_j $(j \geq 1)$ to form sequences $\{p_j\}$ and $\{q_j\}$, as in the previous section, according to the double recursions:

$$p_j \ = \ n_j p_{j-1} + p_{j-2} \ , \ j \geq 1 \ ; \ \ p_0 \ = \ 1 \ , \ \ p_{-1} \ = \ 0 \ . \tag{14}$$

$$q_j \ = \ n_j q_{j-1} + q_{j-2} \ , \ j \geq 1 \ ; \ \ q_0 \ = \ 0 \ , \ \ q_{-1} \ = \ 1 \ . \tag{15}$$

It has already been observed that $q_j \geq 1$ for $j \geq 1$ and

$$[n_1, n_2, \ldots, n_j] \ = \ \frac{p_j}{q_j} \ , \ \ j \geq 1 \ .$$

Bezout's Identity says not only that the greatest common divisor of a and b is an integer linear combination of them but that the coefficents in that integer linear combination may be taken, up to a sign, as q and p.

Theorem 5. *If the application of the Euclidean algorithm to a and b $(b > 0)$ ends with the m^{th} long division, i.e., $r_m = 0$, then*

$$r_j \ = \ (-1)^{j-1} \left(q_j a - p_j b\right) \ , \ \ 1 \leq j \leq m \ . \tag{16}$$

Proof. One uses induction on j. For $j = 1$ the statement is $r_1 = q_1 a - p_1 b$. Since by (14, 15) $q_1 = 1$ and $p_1 = n_1$, this statement is simply the case $j = 1$ in (13). Assume $j \geq 2$, and that the formula (16) has been established for indices smaller than j. By (13) one has

$$r_j \ = \ r_{j-2} - n_j r_{j-1} \ .$$

In this equation one may use (16) to expand the terms r_{j-2} and r_{j-1} to obtain:

$$\begin{aligned} r_j &= \{(-1)^{j-3}(q_{j-2}a - p_{j-2}b)\} \ - \ n_j\{(-1)^{j-2}(q_{j-1}a - p_{j-1}b)\} \\ &= \{(-1)^{j-1}(q_{j-2}a - p_{j-2}b)\} \ + \ n_j\{(-1)^{j-1}(q_{j-1}a - p_{j-1}b)\} \\ &= (-1)^{j-1}\{(q_{j-2}a - p_{j-2}b) \ + \ n_j(q_{j-1}a - p_{j-1}b)\} \\ &= (-1)^{j-1}\{(q_{j-2} + n_j q_{j-1})a - (p_{j-2} + n_j p_{j-1})b\} \\ &= (-1)^{j-1}\{q_j a - p_j b\} \ . \end{aligned}$$

Corollary 3. *The greatest common divisor d of a and b is given by the formula*

$$d \ = \ (-1)^m (q_{m-1}a \ - \ p_{m-1}b) \ , \tag{17}$$

where m is the number of divisions required to obtain zero remainder in the Euclidean algorithm.

Proof. One knows that d is the last non-zero remainder r_{m-1} in the Euclidean algorithm. This formula for d is the case $j = m - 1$ in (16).

Corollary 4.

$$p_m \;=\; \frac{a}{d}\,, \quad q_m \;=\; \frac{b}{d}\,. \tag{18}$$

Proof. The last remainder $r_m = 0$. The case $j = m$ in (16) shows that $a/b = p_m/q_m$. Since, by the first proposition of the preceding section, p_m and q_m have no common factor, this corollary is evident.

8 The action of $\mathbf{GL}_2(\mathbf{Z})$ on the projective line

If a, b, c, d are real numbers with $ad - bc \neq 0$ and

$$M \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is the matrix with entries a, b, c, and d, then $M \cdot z$, for z real, will denote the expression

$$M \cdot z \;=\; \frac{az \;+\; b}{cz \;+\; d}\,. \tag{19}$$

One calls $M \cdot z$ the *action* of M on z.

$M \cdot z$ is a perfectly good function of z except for the case $z = -d/c$ where the denominator $cz + d$ vanishes. If it were also true that $az + b = 0$ for the same z, then one would have $-b/a = -d/c$, in contradiction of the assumption $ad - bc \neq 0$. Thus, when $z = -d/c$, the value of $|M \cdot w|$ increases beyond all bounds as w approaches z, and it is convenient to say that

$$M \cdot \left(-\frac{d}{c}\right) \;=\; \infty$$

where ∞ is regarded as large and signless. If further it is agreed to define

$$M \cdot \infty \;=\; \frac{a}{c}\,,$$

which is the limiting value of $M \cdot w$ as $|w|$ increases without bound, then one may regard the expression $M \cdot z$ as being defined always for all real z and for ∞. The set consisting of all real numbers and also the object (not a number) ∞ is called the *projective line.* The projective line is therefore the union of the (ordinary) affine line with a single point ∞.

Proposition 4. *If* $[n_1, n_2, \ldots]$ *is any continued fraction, then*

$$[n_1, n_2, \ldots, n_r, n_{r+1}, \ldots] \;=\; M \cdot [n_{r+1}, \ldots]\;. \tag{20}$$

where

$$M \;=\; \begin{pmatrix} n_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} n_r & 1 \\ 1 & 0 \end{pmatrix}\;.$$

Proof. Let $z = [n_{r+1}, \ldots]$. Then

$$[n_1, n_2, \ldots, n_r, n_{r+1}, \ldots] \;=\; [n_1, n_2, \ldots, n_r, z]\;.$$

The statement of the proposition now becomes

$$[n_1, n_2, \ldots, n_r, z] \quad = \quad M \cdot z \ .$$

This may be seen to follow by multiplying both sides in formula (9), after replacing t_j with n_j, by the column

$$\begin{pmatrix} z \\ 1 \end{pmatrix} \ .$$

The matrix M in the preceding proposition is an integer matrix with determinant ±1. The notation $\mathrm{GL}_2(\mathbf{Z})$ denotes the set of all such matrices. (The 2 indicates the size of the matrices, and the $\mathbf{Z}$ indicates that the entries in such matrices are numbers in the set $\mathbf{Z}$ of integers.) It is easy to check that the product of two members of $\mathrm{GL}_2(\mathbf{Z})$ is a member of $\mathrm{GL}_2(\mathbf{Z})$ and that the matrix inverse of a member of $\mathrm{GL}_2(\mathbf{Z})$ is a member of $\mathrm{GL}_2(\mathbf{Z})$. Thus, $\mathrm{GL}_2(\mathbf{Z})$ forms what is called a *group.* The formula (19) defines what is called the *action* of $\mathrm{GL}_2(\mathbf{Z})$ on the projective line.

One says that two points z and w of the projective line are *rationally equivalent* if there is a matrix M in $\mathrm{GL}_2(\mathbf{Z})$ for which $w = M \cdot z$. Since (i) $\mathrm{GL}_2(\mathbf{Z})$ is a group, (ii) $M_1 \cdot (M_2 \cdot z) = (M_1 M_2) \cdot z$, and (iii) $w = M \cdot z$ if and only $z = M^{-1} \cdot w$, it is easy to see that every point of the projective line belongs to one and only one rational equivalence class and that two points rationally equivalent to a third must be rationally equivalent to each other.

Terminology. *The rational equivalence of points on the projective line is said to be the equivalence relation on the projective line defined by the action of* $GL_2(\boldsymbol{Z})$.

Example 1. *The set of real numbers rationally equivalent to the point* ∞ *is precisely the set of rational numbers.*

Example 2. *The proposition above shows that any continued fraction is rationally equivalent to each of its tails. It follows that all tails of a continued fraction are rationally equivalent to each other.*

9 Periodic continued fractions

In one of the first examples of a continued fraction expansion, it was shown that $\sqrt{10} = [3, 6, 6, 6, \ldots]$. This is an example of a *periodic* continued fraction. After a finite number of terms the sequence of integers repeats cyclically. If a cyclic pattern is present from the very first term, then the continued fraction is called *purely periodic.* Evidently, $[6, 6, 6, \ldots] = \sqrt{10} - 3$ is an example of a purely periodic continued fraction.

Note that a periodic continued fraction cannot represent a rational number since the continued fraction expansion of a rational number is finite.

Theorem 6. *Every periodic continued fraction is the continued fraction expansion of a real quadratic irrational number.*

Proof. For clarity: it is being asserted that every periodic continued fraction represent a number of the form

$$\frac{a + b\sqrt{m}}{c}$$

where $a, \quad b, \quad c, \quad$ and m are all integers with $m > 0, \quad c \neq 0, \quad$ and m not a perfect square.

Numbers of this form with fixed m but varying integers $a, \quad b, \quad$ and $c \neq 0$ may be added, subtracted, multiplied, and divided without leaving the class of such numbers. (The statement here about division becomes clear if one remembers always to *rationalize* denominators.) Consequently, for M in $\mathrm{GL}_2(\mathbf{Z})$ the number $M \cdot z$ will be a number of this form or ∞ if and only if z is in the same class.

Since a periodic continued fraction is rationally equivalent to a purely periodic continued fraction, the question of whether any periodic continued fraction is a quadratic irrationality reduces to the question of whether a purely periodic continued fraction is such. Let

$$x \quad = \quad [n_1, \ldots, n_r, n_1, \ldots, n_r, n_1, \ldots, n_r, \ldots]$$

be a purely periodic continued fraction. By the proposition of the preceding section, $x = M \cdot x$ where M is notationally identical to the M in (20). Ignoring the computation (9) of M in terms of convergents, let

$$M \quad = \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad .$$

Then

$$x \quad = \quad \frac{ax + b}{cx + d} \, ,$$

or, otherwise said, x is a solution of the quadratic equation

$$cx^2 - (a - d)x - b \quad = \quad 0 \quad .$$

Remark 2. *It is conversely true that the continued fraction expansion of every real quadratic irrationality is periodic.*

This converse will not be proved here.

References

[1] G. Chrystal, *Algebra: An Elementary Textbook* (2 vols.), Chelsea.

[2] G. Hardy & E. Wright, *An Introduction to the Theory of Numbers*, Oxford Univ. Press.

[3] S. Lang, *Introduction to Diophantine Approximations*, Addison-Wesley.

[4] O. Perron, *Die Lehre von den Kettenbrüchen*, 2nd ed., Chelsea.

[Processed from GELLMU to LaTeX on Wed 14 Mar 2001 at 15:48:50 EST]