CONTINUED FRACTIONS

HENRIK BÄÄRNHIELM

Abstract. The elementary arithmetic theory of continued fractions is presented, at an introductory level. Initially, the basic definitions and properties of finite and infinite continued fractions are stated and proved. Then the periodicity of the continued fractions for quadratic irrationals are investigated, and some results concerning approximation of real numbers by rationals are proved. Finally, some more advanced topics of continued fractions are briefly covered, and then the Pell equation and Wiener's attack on the RSA cryptosystem are presented, as nice applications of continued fractions.

Date: 2004-01-12.

Contents

1. Preface

This essay is about continued fractions: their properties and usage. It is written for those with at least some basic knowledge in number theory, but without any real earlier experience of continued fractions. Therefore, we prove the most important basic properties here, before advancing to the more sophisticated theorems about quadratic irrationals, approximation of real numbers and other topics. Finally, some applications of continued fractions are covered. This treatment of continued fractions is in no way complete, but is more of an overview, including the important results. For more substantial texts, see the bibliography, especially [HW79].

To be more precise, this essay is concerned with arithmetical theory of continued fractions. There is also a quite vast analytic theory, which has a somewhat different viewpoint, and it is not covered here. For the interested reader, [Wal73] is a starting point for the analytic theory.

2. Introduction

There are a few different ways in which continued fractions can be introduced. As the reader may know, a continued fraction is an expression like

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cdots}}}$$

and these can be defined as above without any motivation, with $a_i, b_i \in \mathbb{R}$, or possibly by giving "the continued fraction algorithm", an approach taken by [HW79]. Another approach is to define them using geometric means, which is done in [Sta78]. However, the author thinks that the best way to do it is by first showing how these (quite weird) expressions arise. This approach is also taken by [Dav99].

Thus, take $a, b = q_0 \in \mathbb{Z}$ and proceed with the normal Euclidean algorithm for computing $\gcd(a, b)$, as follows

$$\begin{aligned} a &= q_0 a_0 + q_1 \\ q_0 &= q_1 a_1 + q_2 \\ q_1 &= q_2 a_2 + q_3 \\ &\vdots \\ q_{n-1} &= q_n a_n + 1 \end{aligned} \tag{1}$$

Now, dividing each equality by corresponding q_i gives us

$$\begin{aligned} \frac{a}{b} &= a_0 + \frac{q_1}{q_0} \\ \frac{q_0}{q_1} &= a_1 + \frac{q_2}{q_1} \\ \frac{q_1}{q_2} &= a_2 + \frac{q_3}{q_2} \\ &\vdots \\ \frac{q_{n-1}}{q_n} &= a_n + \frac{q_{n+1}}{q_n} \end{aligned} \tag{2}$$

From these equations, we see that $\frac{a}{b}$ can be computed by substituting each line, inverted, into the preceding one, and this is just the continued fraction

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_n + \frac{q_{n+1}}{q_n}}}} \tag{3}$$

It is therefore evident that continued fractions arise quite naturally, and in view of this, it is obvious that every $q = \frac{a}{b} \in \mathbb{Q}$ can be expanded into a continued fraction (with $q_{n+1} = 1$ in this case, since we can assume $\gcd(a, b) = 1$), with $a_i > 0$ when $i > 0$ and that the expansion is unique. This follows from the Euclidean algorithm, and since we also know that the algorithm always terminates, the continued fraction is of finite length. Later, we will consider infinite continued fractions, which are connected to irrational numbers.

Note that the continued fractions determined by the Euclidean algorithm always have all "numerators" equal to 1, and we will mainly consider this case.

3. Finite continued fractions

After the rather informal introduction, we now proceed more formally.

Definition 3.1. Let $n \geq 0$. The function $\mathbb{R}^{n+1} \to \mathbb{R}$ in the variables $a_0, a_1, \ldots, a_n$ defined by

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_n}}} \tag{4}$$

is called a *finite continued fraction*. We will omit the word "finite" where it is not necessary, and we will denote the continued fraction (and its value) using the more concise expression

$$\langle a_0, a_1, \ldots, a_n \rangle \tag{5}$$

The numbers a_i are called the *partial quotients* of the continued fraction, and the values $\langle a_0, a_1, \ldots, a_k \rangle$ for $0 \leq k \leq n$ are called the *convergents* of the continued fraction. Two continued fractions are equal if they have the same partial quotients.

Note that, *a priori* it is not clear if different continued fractions necessarily evaluate to different numbers, and in fact this is not the case, as we will see later. Now, to begin with, we will state a simple recursive method to calculate the convergents.

Proposition 3.2. *Define* $\{p_i\}_{i=0}^N$, $\{q_i\}_{i=0}^N$ *as*

$$p_n = \begin{cases} a_0 & n = 0 \\ a_0 a_1 + 1 & n = 1 \\ a_n p_{n-1} + p_{n-2} & 2 \leq n \leq N \end{cases} \tag{6}$$

$$q_n = \begin{cases} 1 & n = 0 \\ a_1 & n = 1 \\ a_n q_{n-1} + q_{n-2} & 2 \leq n \leq N \end{cases} \tag{7}$$

then $\langle a_0, a_1, \ldots, a_n \rangle = \frac{p_n}{q_n}$ *for* $0 \leq n \leq N$.

Proof. The statement is trivial for $n = 0, 1$. Assume it holds for $n = k$ and consider

$$\begin{aligned}\langle a_0, a_1, \ldots, a_k, a_{k+1}\rangle &= \left\langle a_0, a_1, \ldots, a_k + \frac{1}{a_{k+1}} \right\rangle = \\ &= \frac{\left(a_k + \frac{1}{a_{k+1}}\right) p_{k-1} + p_{k-2}}{\left(a_k + \frac{1}{a_{k+1}}\right) q_{k-1} + q_{k-2}}\end{aligned} \tag{8}$$

where we have used the inductive assumption. Multiplying by a_{k+1} gives

$$\begin{aligned}\langle a_0, a_1, \ldots, a_k, a_{k+1}\rangle &= \frac{a_{k+1}(a_k+1)p_{k-1} + a_{k+1}p_{k-2}}{a_{k+1}(a_k+1)q_{k-1} + a_{k+1}q_{k-2}} = \\ &= \frac{a_{k+1}(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{a_{k+1}(a_k q_{k-1} + q_{k-2}) + q_{k-1}} = \\ &= \frac{a_{k+1}p_k + p_{k-1}}{a_{k+1}q_k + q_{k-1}}\end{aligned} \tag{9}$$

and the inductive step is proved. □

Proposition 3.3. *The numbers p_n and q_n, defined by* (6) *and* (7) *respectively, satisfy*

$$p_n q_{n+1} - q_n p_{n+1} = (-1)^{n+1} \tag{10}$$

$$p_n q_{n+2} - q_n p_{n+2} = (-1)^{n+1} a_{n+2} \tag{11}$$

Proof. We have that $p_0 q_1 - q_0 p_1 = a_0 a_1 - (a_0 a_1 + 1) = -1$. Assume that the statement holds for n and observe that

$$\begin{aligned}p_n q_{n+1} - q_n p_{n+1} &= \begin{vmatrix} p_n & q_n \\ p_{n+1} & q_{n+1}\end{vmatrix} = \begin{vmatrix} p_n & q_n \\ p_{n-1} + a_{n+1}p_n & q_{n-1} + a_{n+1}q_n\end{vmatrix} = \\ &= \begin{vmatrix} p_n & q_n \\ p_{n-1} & q_{n-1}\end{vmatrix} = -\begin{vmatrix} p_{n-1} & q_{n-1} \\ p_n & q_n\end{vmatrix} = \\ &= -\left(p_{n-1}q_n - q_{n-1}p_n\right) = -(-1)^n\end{aligned}$$

which proves the inductive step of (10). For (11), observe that

$$\begin{aligned}p_n q_{n+2} - q_n p_{n+2} &= \begin{vmatrix} p_n & q_n \\ p_{n+2} & q_{n+2}\end{vmatrix} = \begin{vmatrix} p_n & q_n \\ p_n + a_{n+2}p_{n+1} & q_n + a_{n+2}q_{n+1}\end{vmatrix} = \\ &= \begin{vmatrix} p_n & q_n \\ a_{n+2}p_{n+1} & a_{n+2}q_{n+1}\end{vmatrix} = a_{n+2}\begin{vmatrix} p_n & q_n \\ p_{n+1} & q_{n+1}\end{vmatrix} = \\ &= a_{n+2}(-1)^{n+1}\end{aligned}$$

where we in the last step used (10). □

Definition 3.4. A continued fraction $\langle a_0, a_1, \ldots, a_N\rangle$ is called *simple* if $a_i \in \mathbb{Z}$ for $0 \leq i \leq N$ and $a_i > 0$ for $0 < i \leq N$.

The second part of the following result may not seem so obvious at first sight, but is in fact trivial.

Corollary 3.5. *Every convergent $\frac{p_n}{q_n}$ to a simple continued fraction is rational, and* $\gcd(p_n, q_n) = 1$.

Proof. The first statement follows immediately from proposition 3.2. Observe then that if $k|p_n$ and $k|q_n$ then k divides the left hand side of (10), and thus $k|1$. □

Corollary 3.6. *In the continued fraction $\langle a_0, a_1, \ldots, a_n\rangle$, if the partial quotients satisfy $a_i > 0$ for $i > 0$, then the numbers q_n are all positive and form a strictly increasing sequence. Moreover, if the continued fraction is simple, then $q_n \geq n$ for $n \geq 0$.*

Proof. The first statement follows immediately from (7). Assume that the continued fraction is simple, and observe that $q_0 = 1 > 0$ and $q_1 = a_1 \geq 1$. If $q_j \geq j$ for $0 \leq j \leq k$ we have

$$q_{k+1} = a_{k+1}q_k + q_{k-1} \geq k + k - 1 = 2k - 1$$

and if $k \geq 2$ we have $2k - 1 \geq k + 1$, which proves the inductive step. □

Lemma 3.7. *In the simple continued fraction $\langle a_0, a_1, \ldots, a_n\rangle$, if $a_n \geq 2$ then*

$$\langle a_0, a_1 \ldots, a_{n-1}, a_n\rangle = \langle a_0, a_1, \ldots, a_{n-1}, a_n - 1, 1\rangle \tag{12}$$

and if $a_n = 1$ then

$$\langle a_0, a_1 \ldots, a_{n-1}, a_n\rangle = \langle a_0, a_1, \ldots, a_{n-1} + 1\rangle \tag{13}$$

Proof. We use proposition 3.2. In the first case, the final convergent of the right hand side is

$$p_{n+1} \ = p_n + p_{n-1} = p_{n-1}(a_n - 1) + p_{n-2} + p_{n-1} = p_{n-1}a_n + p_{n-2} \tag{14}$$

which is p_n of the left hand side, and since the same is true for the q_n the values of the continued fractions are equal.

In the second case, the final convergent of the left hand side is

$$p_n = p_{n-1} + p_{n-2} = a_{n-1}p_{n-2} + p_{n-3} \ + p_{n-2} = (a_{n-1} + 1)p_{n-2} + p_{n-3} \tag{15}$$

which is p_{n-1} of the right hand side. □

From what we discovered in section 2, we get the following result.

Theorem 3.8. *For every $q \in \mathbb{Q}$, there exists a (finite) simple continued fraction $\langle a_0, a_1, \ldots, a_n\rangle = q$. Moreover, except for the ambiguity shown in lemma 3.7, the continued fraction is unique.*

Proof. Obvious from the Euclidean algorithm. □

As we also found in section 2, the continued fraction (that is, the partial quotients) of a rational number can be found by a small twist of the Euclidean algorithm. Indeed, in (2) the numbers a_i are exactly the partial quotients, and those numbers appear naturally in each step in the Euclidean algorithm as the quotients in the division algorithm. However, there is a more direct method of computing the partial quotients, which is normally referred to as "the continued fraction algorithm".

Theorem 3.9 (Continued fraction algorithm)**.** *Let $q \in \mathbb{Q}$. The partial quotients of the simple continued fraction $\langle a_0, a_1, \ldots, a_N\rangle = q$ can be found using the recurrences*

$$\theta_n = \begin{cases} q & n = 0 \\ \frac{1}{\theta_{n-1} - a_{n-1}} & 0 < n \leq N \end{cases} \tag{16}$$

$$a_n = \lfloor \theta_n \rfloor \tag{17}$$

where it is understood that N is the smallest number such that $\theta_N = a_N$.

Proof. It is obvious from (4) that the algorithm produces a valid continued fraction that evaluates to q. To be sure that it is in fact simple, note that by definition, every a_n is obviously an integer and $0 < \theta_n - a_n < 1$, so that $\frac{1}{\theta_n - a_n} > 1$, which implies $a_n > 0$ for $n > 0$.

Note that with this algorithm, the computed continued fraction always has the shortest possible number of partial quotients. □

Thus, we are justified in talking about *the continued fraction expansion* of a rational number. Since this expansion is always finite, the continued fraction algorithm sets up a correspondence between rational numbers and finite sequences of integers, and it is therefore possible to view a finite simple continued fraction as a representation of a rational number. This representation has the advantage over the normal decimal representation of always being of finite length. On the contrary, the decimal representation of a rational number is not always finite, even though it is periodic if it is not finite.

As we shall see, there is an analogue between this and the continued fraction expansion of irrational numbers. The decimal expansion of a rational number can be of infinite length but it is periodic, but the decimal expansion of an irrational number is not periodic. On the other hand, the continued fraction expansion of a rational number is of finite length, but the expansion of an irrational number is infinite and, quite remarkably, periodic.

To conclude this section on finite continued fractions, we state some results that are important when defining infinite continued fractions.

Proposition 3.10. *Let $x = \langle a_0, a_1, \ldots, a_N \rangle$ be a (not necessarily simple) continued fraction, such that $a_i > 0$ when $i > 0$. Let $\{x_{2k}\}_{k=0}^{\lfloor N/2 \rfloor}$ be the even convergents, and let $\{y_{2k+1}\}_{k=0}^{\lfloor (N-1)/2 \rfloor}$ be the odd convergents. Then*

1. *$x_i < x_j$ when $i < j$ and $y_i > y_j$ when $i < j$*
2. *$x_i < y_j$ for every i, j.*
3. *$x > x_i$ for every i and $x < y_j$ for every j, except that $x = x_N$ or $x = y_N$, depending on whether N is even or odd.*

Proof. (1) Write (11) in the form

$$\frac{p_n}{q_n} - \frac{p_{n+2}}{q_{n+2}} = \frac{(-1)^{n+1} a_{n+2}}{q_n q_{n+2}} \tag{18}$$

Since $a_{n+2} > 0$, using corollary 3.6 we see that the left hand side has the same sign as $(-1)^{n+1}$. In the case of even convergents, this implies that $x_n < x_{n+2}$, and in the case of odd convergents, this implies that $y_n > y_{n+2}$, so the result follows.

(2) Write (10) in the form

$$\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} = \frac{(-1)^{n+1}}{q_n q_{n+1}} \tag{19}$$

The numbers q_i are always positive, so the left hand side has the same sign as $(-1)^{n+1}$. This implies that when n is even, $\frac{p_n}{q_n} < \frac{p_{n+1}}{q_{n+1}}$, and when n is odd, $\frac{p_n}{q_n} > \frac{p_{n+1}}{q_{n+1}}$, so in any case an odd convergent is always greater than the adjacent even convergent.

Assume that there exist r, s such that $x_r \geq y_s$. If $r > s$ then $r-1 \geq s$ and $x_r \geq y_s \geq y_{r-1}$ since the odd convergents are decreasing, by the previous result. This contradicts the fact shown in the last paragraph, that an odd convergent is strictly greater than the following even convergent. If $r < s$ then $r \leq s-1$ and $y_s \leq x_r \leq x_{s-1}$ since the even convergents are increasing, and this is again a contradiction.

(3) If N is even, then $x = x_N > x_i$ for every $i < N$, by the first result, and by the second result, $x = x_N < y_j$ for every j. If N is odd, then $x = y_N < y_i$ for every $i < N$, by the first result, and by the second result, $x = y_N > x_j$ for every j.

□

Corollary 3.11. *Let $x = \langle a_0, a_1, \ldots, a_N \rangle$ be a (not necessarily simple) continued fraction, such that $a_i > 0$ for $i > 0$. Then for $0 \leq n < N$ we have*

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} \tag{20}$$

Proof. By the last statement of proposition 3.10 we have that x is strictly between adjacent convergents, so indeed

$$\left| x - \frac{p_n}{q_n} \right| < \left| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right| = \left| \frac{p_n q_{n+1} - p_{n+1} q_n}{q_n q_{n+1}} \right| = \frac{1}{|q_n q_{n+1}|} \tag{21}$$

where we used (10). Using corollary 3.6 we get our result. □

4. Infinite continued fractions

Our treatment of finite continued fractions was really only an introduction to the more interesting topic of infinite continued fractions. Informally, we say that $x = \langle a_0, a_1, a_2, \ldots \rangle$ is an infinite continued fraction, but it is not clear that this is well-defined. We need the following convergence result.

Lemma 4.1. *If $\{a_i\}_{i=0}^{\infty}$ is a sequence of positive integers, then the sequence $\{x_n\}_{n=0}^{\infty}$ defined as*

$$x_n = \frac{p_n}{q_n} = \langle a_0, a_1, \ldots, a_n \rangle \tag{22}$$

is convergent.

Proof. From proposition 3.10 we see that $x_{2k} < x_1$ for every $k \geq 1$, so the even convergents are bounded above, and by the same proposition they form a monotonically increasing sequence. Hence they converge to a limit x_L.

On the other hand, the odd convergents are monotonically decreasing, and bounded below by x_0, so they converge to a limit x_U, where $x_U \geq x_L$.

To prove that $x_U = x_L$ we consider the differences between even and odd convergents. Following (21) and using corollary 3.6 (note that x_n is the convergent of a *simple* continued fraction) we get

$$|x_{2n} - x_{2n+1}| = \left| \frac{p_{2n}}{q_{2n}} - \frac{p_{2n+1}}{q_{2n+1}} \right| = \frac{1}{|q_{2n} q_{2n+1}|} \leq \frac{1}{2n(2n+1)} \to 0 \tag{23}$$

as $n \to \infty$, so indeed $x_L = x_U$. □

Given this, we are justified in making the following definition.

Definition 4.2. Given a sequence $\{a_i\}_{i=0}^{\infty}$ of positive integers, we say that $x = \langle a_0, a_1, a_2, \ldots \rangle$ is the *infinite continued fraction* determined by the sequence, and it is defined as

$$x = \lim_{n \to \infty} \langle a_0, a_1, \ldots, a_n \rangle \tag{24}$$

As in the finite case, the numbers a_i are called the *partial quotients*, and the numbers $x_n = \langle a_0, a_1, \ldots, a_n \rangle$ are called the *convergents* of the continued fraction.

The results in section 3 generalise in the obvious way to infinite continued fractions, except for theorem 3.8, which requires more work. As was stated earlier, infinite continued fractions represent irrational numbers, and we now make this precise.

Corollary 4.3. *The value of an infinite continued fraction is an irrational number.*

Proof. Let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R}$. If $x \in \mathbb{Q}$, then we can expand x into a finite continued fraction, $x = \langle b_0, b_1, \ldots, b_N \rangle$. But since this continued fraction is unique, and the continued fraction algorithm then terminates, it is impossible to also expand x into an infinite continued fraction. Thus $x \notin \mathbb{Q}$. □

Theorem 4.4. *For every irrational number x, there exists an infinite continued fraction $\langle a_0, a_1, \ldots \rangle = x$. Moreover, the continued fraction is unique.*

Proof. We apply the continued fraction algorithm to x. If there is $N \in \mathbb{N}$ such that the algorithm terminates after computing a_N, then we have $\theta_N = a_N$ and $x = \langle a_0, a_1, a_2, \ldots, a_N \rangle \in \mathbb{Q}$ which is impossible. Thus the algorithm cannot terminate, and the continued fraction must be infinite.

On the other hand, assume that $\langle b_0, b_1, \ldots \rangle = x$, and once again apply the continued fraction algorithm to $x = \theta_0$. Since $0 < \langle b_0, b_1, \ldots \rangle - b_0 < 1$ we have $b_0 = \lfloor x \rfloor = \lfloor \theta_0 \rfloor$ which implies that $b_0 = a_0$ in the continued fraction algorithm. If we let $\theta_1 = \frac{1}{x - b_0}$ we see that $0 < \theta_1 - b_1 < 1$ and thus $b_1 = a_1$ in the continued fraction algorithm. Continuing in this way it is evident that the continued fraction $\langle b_0, b_1, \ldots \rangle$ is the same as the continued fraction computed by the algorithm. Since the algorithm is deterministic, the continued fraction of x is unique. □

In the light of these results we are justified in stating that the continued fraction algorithm sets up a bijection between the irrational numbers and all infinite sequences of integers, $a_0, a_1, a_2, \ldots$ where $a_1, a_2, \ldots$ are positive. To conclude this section, we state another result about the difference between the value of the continued fraction and its convergents.

Proposition 4.5. *Let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R} \backslash \mathbb{Q}$, and let $\left\{ \frac{p_n}{q_n} \right\}_{n=0}^{\infty}$ be its convergents. Then we have that the sequence*

$$\left\{ \left| x - \frac{p_n}{q_n} \right| \right\}_{n=1}^{\infty} \tag{25}$$

is strictly decreasing.

Proof. We use a little trick that let us regard x as a finite (non-simple) continued fraction. For $N \geq 1$, let $\alpha_{N+1} = \langle a_{N+1}, a_{N+2}, \ldots \rangle$, which implies that $x = \langle a_0, a_1, \ldots, a_N, \alpha_{N+1} \rangle$. By proposition 3.2 we get

$$x = \frac{p_{N+1}}{q_{N+1}} = \frac{\alpha_{N+1} p_N + p_{N-1}}{\alpha_N q_N + q_{N-1}} \tag{26}$$

from which it follows that

$$x - \frac{p_N}{q_N} = \frac{\alpha_{N+1}p_N + p_{N-1}}{\alpha_{N+1}q_N + q_{N-1}} - \frac{p_N}{q_N} = \frac{p_{N-1}q_N - p_N q_{N-1}}{q_N(\alpha_{N+1}q_N + q_{N-1})} = \\ = \frac{(-1)^N}{q_N(\alpha_{N+1}q_N + q_{N-1})} \tag{27}$$

where we have used (10). Now we observe that $\alpha_{N+1} = a_{N+1} + \frac{1}{\alpha_{N+2}}$ and since $\alpha_{N+2} > 1$ it follows that

$$a_{N+1} < \alpha_{N+1} < a_{N+1} + 1 \tag{28}$$

Moreover, using proposition 3.2 with (28) we get the inequalities

$$\alpha_{N+1}q_N + q_{N-1} < a_{N+1}q_N + q_{N-1} + q_N = q_{N+1} + q_N \leq q_{N+2} \tag{29}$$

$$\alpha_{N+1}q_N + q_{N-1} > a_{N+1}q_N + q_{N-1} = q_{N+1} \tag{30}$$

and by applying them to (27) we arrive at

$$\frac{1}{q_{N+2}^2} \leq \frac{1}{q_N q_{N+2}} < \left| x - \frac{p_N}{q_N} \right| < \frac{1}{q_N q_{N+1}} \leq \frac{1}{q_N^2} \tag{31}$$

Since the q_i are strictly increasing, the proof is completed. □

Remark 1. The essence of proposition 4.5, that the difference between the value of the continued fraction and its convergents decreases steadily, also holds for finite continued fractions.

4.1. **Geometric interpretaion.** Continued fractions can also be considered from a geometric point of view, and this approach is heavily developed in [Sta78]. We will not dwelve too deeply into the geometric aspect, but it is worth saying a few words about it.

In this context, we consider the first quadrant of the real plane $\mathbb{R}^2$ (including the two half-axes). The non-negative rational numbers then correspond to the lattice of points with integer coordinates in the plane, where $\frac{a}{b} \in \mathbb{Q}$ corresponds to the point $(a, b) \in \mathbb{R}^2$. Imagine the lattice as holes in the plane, where sticks are inserted.

For an irrational number $\alpha \in \mathbb{R}$, we now consider the line $y = \alpha x$. The line does not pass through any point of the lattice, since if there is a point (a, b) with $a, b \in \mathbb{N}$ such that the line passes through it, then $\alpha = \frac{b}{a} \in \mathbb{Q}$. Imagine the line as a long string through the plane, with endpoint in the origin.

Now if the string is moved by moving its endpoint at the origin, then it will hit the sticks in the lattice points. Either it will hit the sticks "below" the line or the sticks "above" the line, and the lattice points with these sticks are exactly the even and odd convergents of the continued fraction for α, respectively.

Our results can now be formulated using these concepts.

4.2. **Quadratic irrationals.** We will now prove a most remarkable fact about quadratic irrational numbers, namely that they are characterised by the property that the partial quotients in their continued fractions are ultimately periodic.

Definition 4.6. A real number $x \in \mathbb{R}$ is a *quadratic irrational* if there exist $a, b, c \in \mathbb{Z}$ such that $ax^2 + bx + c = 0$ and $d = b^2 - 4ac > 0$ and d is not a perfect square.

Definition 4.7. An (infinite) continued fraction $\langle a_0, a_1, a_2, \ldots \rangle$ is *periodic* if there exist integers $N \geq 0$, $k \geq 1$ such that for all $n \geq N$ we have $a_{n+k} = a_n$. In this case we write the continued fraction as $\langle a_0, a_1, \ldots, a_{N-1}, \overline{a_N, a_{N+1}, \ldots, a_{N+k}} \rangle$.

We prove the two parts of the characterisation separately, starting with the easier implication.

Theorem 4.8. *If the continued fraction of $x \in \mathbb{R}$ is periodic then x is a quadratic irrational.*

Proof. Assume that there exist integers $k \geq 1$ and $N \geq 0$ such that

$$x = \langle a_0, a_1, \ldots, a_{N-1}, \overline{a_N, a_{N+1}, \ldots, a_{N+k}} \rangle \tag{32}$$

Observe that if we define $\alpha = \langle \overline{a_N, a_{N+1}, \ldots, a_{N+k}} \rangle$ then we can write x and α as finite (but not simple!) continued fractions as

$$x = \langle a_0, a_1, \ldots, a_{N-1}, \alpha \rangle \tag{33}$$

$$\alpha = \langle a_N, a_{N+1}, \ldots, a_{N+k}, \alpha \rangle \tag{34}$$

In the case where $N \geq 2$ and $k \geq 2$, we can use proposition 3.2 to get

$$x = \frac{\alpha p_{N-1} + p_{N-2}}{\alpha q_{N-1} + q_{N-2}} \tag{35}$$

$$\alpha = \frac{\alpha p_{N+k-1} + p_{N+k-2}}{\alpha q_{N+k-1} + q_{N+k-2}} \tag{36}$$

Now we can solve for α in these expressions and arrive at, respectively

$$0 = \alpha^2 q_{N+k-1} + \alpha(q_{N+k-2} - p_{N+k-1}) - p_{N+k-2} \tag{37}$$

$$\alpha = \frac{x q_{N-2} - p_{N-2}}{p_{N-1} - x q_{N-1}} \tag{38}$$

and by substituting (38) into (37) we arrive at a quadratic equation $ax^2 + bx + c = 0$. Since the continued fraction for x is infinite, we have that x is irrational, and it follows that $d = b^2 - 4ac$ is not a perfect square and $d > 0$.

In the case where $k = 1$, we have $a_N = a_{N+k}$ for every $k \in \mathbb{N}$, so that

$$\alpha = a_N + \frac{1}{\alpha} \tag{39}$$

and thus $\alpha^2 - a_N \alpha - 1 = 0$. If we use with this equation instead of (37), the theorem follows.

If $N = 0$, then $x = \alpha$ and the theorem follows from (37), or (39) if $k = 1$.

If $N = 1$, then $x = \frac{a_0 \alpha + 1}{\alpha}$ and $\alpha = \frac{1}{x - a_0}$ and if we substitute this into (37), or (39) if $k = 1$, we again arrive at a quadratic equation for x. □

Theorem 4.9. *If $x \in \mathbb{R}$ is a quadratic irrational, then the continued fraction of x is periodic.*

Proof. Assume that there exist $a, b, c \in \mathbb{Z}$ such that $ax^2 + bx + c = 0$ and consider the continued fraction $x = \langle a_0, a_1, a_2, \ldots \rangle$. Define $\alpha_n = \langle a_n, a_{n+1}, \ldots \rangle$, so that for every $n \in \mathbb{N}$ we have $x = \langle a_0, a_1, \ldots, a_n, \alpha_{n+1} \rangle$. By proposition 3.2 it follows that

$$x = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}} \tag{40}$$

If $\{\alpha_n\}_{n=0}^{\infty}$ is a finite set, then there exist $i < j$ such that $\alpha_i = \alpha_j$. This implies that the continued fraction for x is periodic, since if $k = \min_j \{j \in \mathbb{N} \mid i < j, \alpha_i = \alpha_j\}$ then $\alpha_i = \langle a_i, a_{i+1}, \ldots, a_{k-1}, \alpha_i \rangle = \langle \overline{a_i, a_{i+1}, \ldots, a_{k-1}} \rangle$ and

$$x = \langle a_0, a_1, \ldots, a_{i-1}, \overline{a_i, a_{i+1}, \ldots, a_{k-1}} \rangle \tag{41}$$

so that for all $n \geq i$ we have $a_{n+k-i} = a_n$ and x is therefore periodic. Thus, it is sufficient to show that $\{\alpha_n\}_{n=0}^{\infty}$ is a finite.

Consider the quadratic form

$$f(u,v) = au^2 + buv + cv^2 \tag{42}$$

and observe that $f(x,1) = 0$ by our assumption. If we make the substitution

$$\begin{aligned} u &= p_n u' + p_{n-1} v' \\ v &= q_n u' + q_{n-1} v' \end{aligned} \tag{43}$$

where $\frac{p_n}{q_n}$ are the convergents of the continued fraction for x, then for each $n \in \mathbb{N}$ we get a new quadratic form $f_n(u', v')$, as follows

$$\begin{aligned} f_n(u',v') &= a(p_n u' + p_{n-1}v')^2 + b(p_n u' + p_{n-1}v')(q_n u' + q_{n-1}v') + \\ &+ c(p_n u' + p_{n-1}v')^2 \\ &= (u')^2(ap_n^2 + bp_n q_n + cq_n^2) + \\ &+ u'v'(2ap_n p_{n-1} + 2cq_n q_{n-1} + b(p_n q_{n-1} + p_{n-1}q_n)) + \\ &+ (y')^2(ap_{n-1}^2 + bp_{n-1}q_{n-1} + cq_{n-1}^2) \end{aligned} \tag{44}$$

Thus, $f_n(u', v') = a'_n(u')^2 + b'_n u'v' + c'_n(v')^2$ where $a'_n = f(p_n, q_n)$ and $c'_n = a'_{n-1}$. A straightforward but cumbersome calculation also shows that $(b'_n)^2 - 4a'_n c'_n = (b^2 - 4ac)(p_n q_{n-1} - p_{n-1}q_n)^2 = b^2 - 4ac$ when applying (10).

Now, using (43) and (40) we see that

$$\begin{aligned} \frac{f_n(\alpha_{n+1}, 1)}{(\alpha_{n+1}q_n + q_{n-1})^2} &= \frac{f(\alpha_{n+1}p_n + p_{n-1}, \alpha_{n+1}q_n + q_{n-1})}{(\alpha_{n+1}q_n + q_{n-1})^2} = \\ &= a\frac{(\alpha_{n+1}p_n + p_{n-1})^2}{(\alpha_{n+1}q_n + q_{n-1})^2} + b\frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}} + c^2 = f(x,1) = 0 \end{aligned} \tag{45}$$

and this implies that $f_n(\alpha_{n+1}, 1) = 0$. Thus, for each $n \in \mathbb{N}$ we have that α_{n+1} is a root of $f_n(z) = f_n(z, 1)$ and if we can show that the number of quadratic forms $f_n(u', v')$ are finite, then the number of α_n must also be finite, since each quadratic form has two roots, and the theorem will follow.

Observe that since $f(x,1) = 0$ we have

$$\frac{a'_n}{q_n^2} = \frac{f(p_n, q_n)}{q_n^2} = f(\frac{p_n}{q_n}, 1) - f(x,1) = a\left[\left(\frac{p_n}{q_n}\right)^2 - x^2\right] + b\left(\frac{p_n}{q_n} - x\right) \tag{46}$$

Since the q_n are integers, we have $q_n q_{n+1} > q_n^2$, and then using corollary 3.11 we get

$$\left|\left(\frac{p_n}{q_n}\right)^2 - x^2\right| = \left|\left(x - \frac{p_n}{q_n}\right)\left(x + \frac{p_n}{q_n}\right)\right| < \frac{\left(x + \frac{p_n}{q_n}\right)}{q_n^2} < \frac{2|x| + 1}{q_n^2} \tag{47}$$

If we use this in (46), we arrive at

$$\left|\frac{a'_n}{{q_n}^2}\right| < \left|a\frac{(2|x|+1)}{q_n^2}\right| + \left|b\frac{1}{{q_n}^2}\right| \tag{48}$$

and thus $a'_n < (2|x| + 1)(|a| + |b|)$. This means that there are only finitely many possible values of a'_n, since it is bounded independently of n, and the same is true of b'_n and c'_n since $c'_n = a'_{n-1}$ and $(b'_n)^2 = b^2 - 4ac + 4a'c'$. Hence, there are only

finitely many possible quadratic forms $f_n(u', v')$ and we have finally proved the theorem. □

4.3. **Approximations of irrational numbers.** We turn now to the issue of using continued fractions as a method of getting rational approximations to irrational numbers. If we use the continued fraction algorithm to expand $x \in \mathbb{R} \setminus \mathbb{Q}$, and at some point halt the algorithm, then we get a finite simple continued fraction $y = \langle a_0, a_1, \ldots, a_n \rangle \in \mathbb{Q}$, and the question is if the rational number y, the last convergent of continued fraction that we calculate, can be used as an approximation of x. That is, can we be sure that $|y - x|$ is small, and can we get it smaller, a better approximation, if we calculate a longer continued fraction approximation?

The answers to both questions are "yes", and are really already answered by corollary 3.11, for if $y = \frac{p_n}{q_n}$ then clearly

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2} \tag{49}$$

so indeed we can get arbitrary good approximations, since the q_i forms a strictly increasing sequence.

Now, other questions arise, as follows:

(1) Is the convergent $\frac{p_n}{q_n}$ given by the continued fraction algorithm the *best* approximation to x?
(2) Is the upper bound in (49) the best possible, or can we get a better bound that holds for all $x \in \mathbb{R}$?
(3) If we have some rational $\frac{p}{q}$ such that $\left| x - \frac{p}{q} \right| < 1/q^2$, is it necessarily true that $\frac{p}{q}$ is a convergent in the continued fraction for x?

These questions will now be made precise, and they will all be answered.

Theorem 4.10. *Let $x = \langle a_0, a_1, a_2, \ldots \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let $\frac{p_n}{q_n} \in \mathbb{Q}$ be the n-th convergent, where $n > 1$. For every $\frac{p}{q} \in \mathbb{Q}$ such that $0 < q \leq q_n$ and $\frac{p}{q} \neq \frac{p_n}{q_n}$ we have*

$$\left| x - \frac{p_n}{q_n} \right| < \left| x - \frac{p}{q} \right| \tag{50}$$

Proof. The proof is by induction on n, so first let $n = 2$. However, the argument for this basis is just a special case of the induction, so we skip it. Assume instead that the result holds for $n \in \{1, \ldots, k\}$ and consider $n = k + 1$. Observe that if $0 < q \leq q_k$ then

$$\left| x - \frac{p_{k+1}}{q_{k+1}} \right| < \left| x - \frac{p_k}{q_k} \right| < \left| x - \frac{p}{q} \right| \tag{51}$$

where we have used proposition 4.5 and the inductive assumption. Clearly, we can therefore assume that $q_k < q \leq q_{k+1}$.

Now we first consider the case when $q = q_{k+1} = q_n$. From corollary 3.11 and corollary 3.6 we get the inequality

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} \leq \frac{1}{2q_n} \tag{52}$$

where we have used the fact that $q_{n+1} \geq 2$ since $n \geq 2$. Observe that since $p \neq p_n$, we have the following inequality

$$\frac{1}{q_n} \leq \left|\frac{p_n}{q_n} - \frac{p}{q}\right| = \left|\frac{p_n}{q_n} - x + x - \frac{p}{q_n}\right| \leq \left|x - \frac{p_n}{q_n}\right| + \left|x - \frac{p}{q_n}\right| < \frac{1}{2q_n} + \left|x - \frac{p}{q_n}\right| \tag{53}$$

which implies that $\left|x - \frac{p}{q_n}\right| > \frac{1}{2q_n}$ and comparing this with (52), we that the result follows.

Now consider the case when $q_k < q < q_{k+1} = q_n$. Notice that this inequality implies that our result will follow from $|p - qx| > |p_n - q_n x|$, so we prove this instead. Choose $a, b \in \mathbb{R}$ such that

$$\begin{aligned} ap_n + bp_{n-1} &= p \\ aq_n + bq_{n-1} &= q \end{aligned}$$

and observe that these linear equations can be solved uniquely for a and b, since the determinant is $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1}$. We get the solutions

$$\begin{aligned} a &= (-1)^{n+1}(pq_{n-1} - qp_{n-1}) \\ b &= (-1)^{n}(pq_n - qp_n) \end{aligned}$$

which shows that in fact $a, b \in \mathbb{Z}$, and by our assumption they are non-zero. Moreover, by proposition 3.10, $x - \frac{p_n}{q_n}$ and $x - \frac{p_{n-1}}{q_{n-1}}$ have opposite sign, and since $q_i \geq 0$ for $i \geq 0$, this is also true for $p_n - q_n x$ and $p_{n-1} - q_{n-1} x$.

Since $0 < q = aq_n + bq_{n-1} < q_n$ we must have that a and b have different sign, and thus $a(p_n - q_n x)$ and $b(p_{n-1} - q_{n-1} x)$ have the same sign. By our definition of a and b we have

$$p - qx = a(p_n - q_n x) + b(p_{n-1} - q_{n-1} x) \tag{54}$$

In the case the terms on the right hand side are both positive, we immediately get $p - qx > a(p_n - q_n x)$ and thus $|p - qx| > |p_n - q_n x|$ since $a \neq 0$. In the case they are both negative, we get

$$\begin{aligned} |p - qx| &> |a(p_n - q_n x) + b(p_{n-1} - q_{n-1} x)| \geq \\ &\geq |a(p_n - q_n x)| - |b(p_{n-1} - q_{n-1} x)| > |a(p_n - q_n x)| \end{aligned} \tag{55}$$

Thus, the proof is completed. □

Remark 2. The essence of theorem 4.10, namely that the convergents are the best approximations to the value of the continued fraction, also holds for finite contineud fractions, but the formulation is of course slightly different.

The preceding theorem answers the first of our questions. Indeed, in the sense given in the theorem, the convergents of the continued fraction are the best approximations to an irrational numbers. The second question must be partly reformulated to be interesting, since by the well-known Dirichlet's Theorem, for any $x \in \mathbb{R}$ and any $N \in \mathbb{N}$ we can always find $p, q \in \mathbb{Z}$ such that $\left|x - \frac{p}{q}\right| < \frac{1}{Nq}$. What we want is an upper bound that holds for *infinitely* many rational numbers, and so we have to restrict ourselves to approximations of irrational numbers. In this setting we can actually find better upper bounds, as shown by the following theorem.

Theorem 4.11. *Let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let $\frac{p_n}{q_n}$ be its convergents. For every $n \geq 1$, we have*

$$\left| x - \tfrac{p_n}{q_n} \right| < \tfrac{1}{2q_n^2} \quad \text{or} \quad \left| x - \tfrac{p_{n+1}}{q_{n+1}} \right| < \tfrac{1}{2q_{n+1}^2} \tag{56}$$

Proof. For contradiction, assume that the result does not hold, which implies

$$\left| x - \frac{p_n}{q_n} \right| + \left| x - \frac{p_{n+1}}{q_{n+1}} \right| \geq \frac{1}{2q_n^2} + \frac{1}{2q_{n+1}^2} \tag{57}$$

From proposition 3.10, $x - \frac{p_n}{q_n}$ and $x - \frac{p_{n+1}}{q_{n+1}}$ have opposite signs, and we assume without loss of generality that $x - \frac{p_n}{q_n} < 0$. Hence,

$$\left| x - \frac{p_n}{q_n} \right| + \left| x - \frac{p_{n+1}}{q_{n+1}} \right| = \left| \frac{p_n}{q_n} - x + x - \frac{p_{n+1}}{q_{n+1}} \right| = \left| \frac{p_n q_{n+1} - p_{n+1} q_n}{q_n q_{n+1}} \right| = \frac{1}{q_n q_{n+1}} \tag{58}$$

where we have used (10). Combining (57) and (58) we get

$$\frac{1}{q_n q_{n+1}} \geq \frac{1}{2q_n^2} + \frac{1}{2q_{n+1}^2} \tag{59}$$

or equivalently

$$2q_n q_{n+1} \geq q_n^2 + q_{n+1}^2 = (q_n - q_{n+1})^2 + 2q_n q_{n+1} \tag{60}$$

which is absurd, since it implies $q_n = q_{n+1}$. □

Now one wonders if we can get even better upper bounds, and in fact we can, as we will see later. But first, consider our third question above. As it is formulated, the answer is clearly negative, since

$$\left| \sqrt{3} - \frac{12}{7} \right| < \frac{1}{7^2} \tag{61}$$

but $\frac{12}{7}$ is not a convergent of the continued fraction of $\sqrt{3}$. However, we have the following remarkable and extremely important theorem.

Theorem 4.12. *Let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let $\frac{p_n}{q_n}$ be its convergents. For every $\frac{p}{q} \in \mathbb{Q}$ such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2} \tag{62}$$

we have that $\frac{p}{q} = \frac{p_n}{q_n}$ for some $n \in \mathbb{N}$.

Proof. Define $n \in \mathbb{N}$ such that $q_n \leq q < q_{n+1}$. Observe that

$$\left| \frac{p}{q} - \frac{p_n}{q_n} \right| \leq \left| x - \frac{p}{q} \right| + \left| x - \frac{p_n}{q_n} \right| < 2 \left| x - \frac{p}{q} \right| < \frac{1}{q^2} \tag{63}$$

where we have used theorem 4.10 and our assumption. Moreover, we have the inequality

$$\left| \frac{p}{q} - \frac{p_n}{q_n} \right| = \left| \frac{pq_n - qp_n}{qq_n} \right| \geq \frac{1}{q^2} \left| pq_n - qp_n \right| \tag{64}$$

Comparing the last two equations, we see that we must have $|pq_n - qp_n| < 1$ which implies $\frac{p}{q} = \frac{p_n}{q_n}$. □

Remark 3. The result of theorem 4.12 remains true for finite continued fractions and rational numbers.

We now state the best possible upper bound that holds for all irrational numbers.

Theorem 4.13 (Hurwitz's theorem)**.** *Let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let $\frac{p_n}{q_n}$ be its convergents. For every $n \geq 1$ we have that at least one of $\frac{p_n}{q_n}$, $\frac{p_{n+1}}{q_{n+1}}$ and $\frac{p_{n+2}}{q_{n+2}}$ satisfy the inequality*

$$\left| x - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2} \tag{65}$$

Proof. Assume for contradiction that the statement is false. Then the analogues of (59) are

$$\frac{1}{q_n q_{n+1}} \geq \frac{1}{\sqrt{5}q_n^2} + \frac{1}{\sqrt{5}q_{n+1}^2} \tag{66}$$

$$\frac{1}{q_{n+2} q_{n+1}} \geq \frac{1}{\sqrt{5}q_{n+2}^2} + \frac{1}{\sqrt{5}q_{n+1}^2} \tag{67}$$

which simplify to

$$\frac{q_n}{q_{n+1}} + \frac{q_{n+1}}{q_n} \leq \sqrt{5} \tag{68}$$

$$\frac{q_{n+2}}{q_{n+1}} + \frac{q_{n+1}}{q_{n+2}} \leq \sqrt{5} \tag{69}$$

and if we define $\alpha = \frac{q_n}{q_{n+1}}$ and $\beta = \frac{q_{n+1}}{q_{n+2}}$ we see that $\alpha + 1/\alpha \leq \sqrt{5}$ and $\beta + 1/\beta \leq \sqrt{5}$. Since $\alpha, \beta \in \mathbb{Q}^+$, the inequalities are strict, and thus

$$\alpha^2 + 1 - \sqrt{5}\alpha = \left(\alpha - \frac{1+\sqrt{5}}{2} \right) \left(\alpha + \frac{1-\sqrt{5}}{2} \right) < 0 \tag{70}$$

$$\beta^2 + 1 - \sqrt{5}\beta = \left(\beta - \frac{1+\sqrt{5}}{2} \right) \left(\beta + \frac{1-\sqrt{5}}{2} \right) < 0 \tag{71}$$

This implies that the parantheses in each equation have opposite sign. If the first is positive and the second is negative, we get $\alpha, \beta > \frac{1}{2}(\sqrt{5}+1)$ and $\alpha, \beta < \frac{1}{2}(\sqrt{5}-1)$ which is impossible. Thus $\alpha, \beta < \frac{1}{2}(1+\sqrt{5})$. From (7) we have $q_{n+2} = a_{n+2}q_{n+1} + q_n$ or equivalently $\frac{q_{n+2}}{q_{n+1}} = a_{n+2} + \frac{q_n}{q_{n+1}}$, and it follows that $\beta \geq 1 + 1/\alpha$. Observe that

$$1 + \frac{1}{\alpha} > 1 - \frac{1-\sqrt{5}}{2} = \frac{1+\sqrt{5}}{2} > \beta \tag{72}$$

which is a contradiction. □

Corollary 4.14. *For every $x \in \mathbb{R} \setminus \mathbb{Q}$, there exist infinitely many $\frac{p}{q} \in \mathbb{Q}$ such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2} \tag{73}$$

Proof. Immediate from theorem 4.13. □

The upper bound with $\sqrt{5}$ is the best possible in the sense that the preceding corollary is not true for any tighter bound. To see this, it is enough to find an

irrational number x such that the inequality $\left|x - \frac{p}{q}\right| < \frac{1}{Aq^2}$ has only finitely many solutions $\frac{p}{q} \in \mathbb{Q}$ when $A > \sqrt{5}$. The easiest example is the golden ratio, $x = \frac{\sqrt{5}-1}{2}$.

Assume that there are inifinitely many $\frac{p}{q} \in \mathbb{Q}$ such that $x - \frac{p}{q} = \frac{\epsilon}{q^2}$ where $|\epsilon| < \frac{1}{A}$. Then $\frac{\epsilon}{q} = qx - p$ and thus

$$\frac{\epsilon}{q} - \frac{\sqrt{5}q}{2} = -\frac{1}{2}q - p \tag{74}$$

We see then that

$$\left(\frac{\epsilon}{q} - \frac{\sqrt{5}q}{2}\right)^2 = \frac{\epsilon^2}{q^2} + \frac{5}{4}q^2 - \epsilon\sqrt{5} = \frac{1}{4}q^2 + p^2 + pq \tag{75}$$

and also

$$\frac{\epsilon^2}{q^2} - \epsilon\sqrt{5} = p^2 + pq - q^2 \tag{76}$$

Now we see that $0 \leq \frac{\epsilon^2}{q^2} - \epsilon\sqrt{5} < 1$ if and only if $0 \leq \frac{\epsilon}{q^2} < \frac{1}{\epsilon} + \sqrt{5}$ if and only if $\left|\epsilon^2\right| < \frac{1}{2}\left|q^2\right|$ which is true when q is large, and this happens since there are infinitely many rational approximations $\frac{p}{q}$. Thus, when q is large, the left hand side of (76) is less than 1 in absolute value while the right hand side is an integer, and this implies that $p^2 - q^2 + pq = 0$ or, by completing the square, $(2p + q)^2 = 5q^2$, which is impossible.

5. Other topics

We will here say a few words about some more advanced topics in the theory of continued fractions, without going into any detail.

5.1. **The Markoff Chain.** As was shown above, the Hurwitz approximation theorem 4.13 is in a sense the best possible approximation that holds for all irrational numbers. However, if one restricts attention by removing certain "bad" numbers that cannot be approximated accurately, then one can get better approximation bounds. The crucial definition is the following:

Definition 5.1. Two real numbers x, y are called *equivalent* if there exist $a, b, c, d \in \mathbb{Z}$ such that $ad - bc = \pm 1$ and

$$y = \frac{ax + b}{cx + d} \tag{77}$$

It is an easy exercise to show that this determines an equivalence relation on $\mathbb{R}$, where all rational numbers are equivalent. Moreover, two irrational numbers x, y are equivalent if and only if $x = \langle a_0, \ldots, a_n, c_0, c_1, \ldots\rangle$ and $y = \langle b_0, \ldots, b_m, c_0, c_1, \ldots\rangle$, so that at some point, the tails of the partial quotients of the continued fraction expansions are equal.

Now one show that for all irrational numbers that are not equivalent to the golden ratio $\frac{\sqrt{5}-1}{2}$, the constant $\sqrt{5}$ in Hurwits theorem can be strengthened to $\sqrt{8}$, and if one removes other equivalence classes of irrational numbers, even higher constants are possible. In fact, there is an infinite sequence of such approximation results, and it is called the *Markoff chain*.

5.2. **Kinchin's constant.** An amazing fact about continued fractions appear if one considers the geometric mean of the partial quotients. That is, let $x = \langle a_0, a_1, \ldots \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let $G_n(x) = (a_0 a_1 \cdots a_n)^{1/n}$. Then, for almost every x in the sense of the Lebesgue measure, the limit of $G_n(x)$ is independent of x

$$K = \lim_{n\to\infty} G_n(x) = 2.68545\ldots \tag{78}$$

and K is called *Kinchin's constant*. However, nobody has proved that $G_n(x)$ approaches K for any particular irrational number x, but some numbers are known that do not saisfy this, a simple example being e.

5.3. **Purely periodic continued fractions.** The theory of quadratic irrationals can be somewhat extended, to give a characterisation of all quadratic irrationals $x \in \mathbb{R} \setminus \mathbb{Q}$ such that $x = \langle \overline{a_0, a_1, \ldots, a_m} \rangle$, that is, x has a *purely periodic* continued fraction. Since all irrational numbers x that have periodic continued fractions are solutions of quadratic equations, we can define the *conjugate* of x as the other root of its associated quadratic equation.

The characterisation of purely periodic continued fractions is then contained in the following theorem. The proof is not difficult, but is left out for space reasons.

Theorem 5.2. *Let $x = \langle a_0, a_1, \ldots, a_k, \overline{a_{k+1}, \ldots, a_{k+m}} \rangle \in \mathbb{R} \setminus \mathbb{Q}$ and let x' be the conjugate of x. Then $k = 0$ if and only if $x > 1$ and $-1 < x' < 0$.*

Corollary 5.3. *For any $d \in \mathbb{N}$ that is not a perfect square, we have $\sqrt{d} = \langle a_0, \overline{a_1, \ldots, a_m} \rangle$ for some $m \geq 1$.*

Proof. Since $\sqrt{d} > 1$ we have $x = \sqrt{d} + \lfloor \sqrt{d} \rfloor > 1$ and x satisfies the quadratic equation $(x - \sqrt{d} - \lfloor \sqrt{d} \rfloor)(x + \sqrt{d} - \lfloor \sqrt{d} \rfloor) = 0$ which have integer coefficients, so the conjugate of x is $x' = -\sqrt{d} + \lfloor \sqrt{d} \rfloor$ which obviosly satisfies $-1 < x' < 0$. By theorem 5.2, x has a purely periodic continued fraction, and since $\lfloor x \rfloor = 2\lfloor \sqrt{d} \rfloor$ is the first partial quotient, we have $x = \left\langle \overline{2\lfloor \sqrt{d} \rfloor, a_1, \ldots, a_m} \right\rangle$ for some $m \geq 0$, and this implies that $\sqrt{d} = \left\langle \lfloor \sqrt{d} \rfloor, \overline{a_1, \ldots, a_m, 2\lfloor \sqrt{d} \rfloor} \right\rangle$. □

6. Applications

Now we will show some applications of continued fractions. They are all based on theorem 4.12.

6.1. **Pell's equation.** The *Pell equation* is the following Diophantine equation:

$$x^2 - Ny^2 = 1 \tag{79}$$

where $N \in \mathbb{N}$ is not a perfect square. The original question is whether there exist a non-trivial solution $x, y \in \mathbb{Z}$ for a particular N, and if the solution is unique. Obviously, there is always a trivial solution $x = 1$, $y = 0$. The equation was a challenge by Fermat to his English collegues, one of them being Pell. According to most sources, Pell himself did not work with the problem, but for some reason, his name was attached to the equation by Euler.

To begin with, we see that any solution in negative integers can always be easily found from the corresponding solution in positive integers, and each such positive solution must satisfy x, y satisfies

$$x - \sqrt{N}y = \frac{1}{x + \sqrt{N}y} > 0 \tag{80}$$

which implies that $x > \sqrt{N}y$. This together with (80) implies $0 < x - \sqrt{N}y < \frac{1}{2y\sqrt{N}}$ and dividing by y gives us

$$\left|\frac{x}{y} - \sqrt{N}\right| < \frac{1}{2y^2\sqrt{N}} < \frac{1}{2y^2} \tag{81}$$

Now it follows from theorem 4.12 that any solution, in positive integers, to (79) with positive right hand side must be a convergent to the continued fraction of $\sqrt{N}$. By corollary 5.3 $\sqrt{N} = \langle a_0, \overline{a_1, a_2, \ldots, a_m}\rangle$ for some $m \geq 1$. If we denote the convergents of this continued fraction by $\frac{p_k}{q_k} \in \mathbb{Q}$ for $k \geq 1$, we then have that $x = p_n$, $y = q_n$ for some $n \geq 1$. By writing $\sqrt{N} = \langle a_0, a_1, \ldots, a_n, \alpha_{n+1}\rangle$ where $\alpha_{n+1} = \langle a_{n+1}, a_{n+1}, \ldots\rangle$ it follows from proposition 3.2 that

$$\sqrt{N} = \frac{p_n\alpha_{n+1} + p_{n-1}}{q_n\alpha_{n+1} + q_{n-1}} \tag{82}$$

and using (10) we arrive at

$$q_n\sqrt{N} - p_n = \frac{\alpha_{n+1}p_nq_n + p_{n-1}q_n - \alpha_{n+1}p_nq_n - p_nq_{n-1}}{q_n\alpha_{n+1} + q_{n-1}} = \frac{(-1)^n}{q_n\alpha_{n+1} + q_{n-1}} \tag{83}$$

From this we see that when n is even, $q_n\sqrt{N} - p_n > 0$, and from (80) it follows that $x = p_n$, $y = q_n$ can not be a solution to (79), so we conclude that n must be odd.

Some manipulations of (82) leads to $(p_n - \sqrt{N}q_n)\alpha_{n+1} = q_{n-1}\sqrt{N} - p_{n-1}$ and hence

$$\begin{aligned}(p_n^2 - Nq_n^2)\alpha_{n+1} &= (q_{n-1}\sqrt{N} - p_{n-1})(p_n + \sqrt{N}q_n) = \\ &= Nq_nq_{n-1} + \sqrt{N}(p_nq_{n-1} - q_np_{n-1}) - p_np_{n-1} = \\ &= (-1)^n\sqrt{N} + Nq_nq_{n-1} - p_np_{n-1}\end{aligned} \tag{84}$$

By assumption, $x = p_n$ and $y = q_n$ is a solution to (79), and the above argument shows that n must be odd, so the previous equation simplifies to

$$\alpha_{n+1} = \sqrt{N} + Nq_nq_{n-1} - p_np_{n-1} = a_0 + \frac{1}{\alpha_1} + Nq_nq_{n-1} - p_np_{n-1} \tag{85}$$

However, we also have $\alpha_{n+1} = a_{n+1} + \frac{1}{\alpha_{n+2}}$ and by the uniqueness of the continued fraction expansion we get $a_0 + Nq_nq_{n-1} - p_np_{n-1} = a_{n+1}$ and $\alpha_1 = \alpha_{n+2}$. We conclude that the period m divides $n + 1$.

Thus, $n + 1 = rm$ for some $r \geq 1$ and since $n + 1$ must be even, r must be even when m is odd.

On the other hand, it is in fact the case that all convergents $\frac{p_n}{q_n}$ to $\sqrt{N}$ where $m|n + 1$ are solutions to (79). Since the continued fraction of $\sqrt{N}$ is periodic, we have $\alpha_1 = \alpha_{n+2}$ for all odd n such that $n + 1 = rm$ for some $r \geq 1$. We can thus write (82) as

$$\sqrt{N} = \frac{p_{n+1}\alpha_1 + p_n}{q_{n+1}\alpha_1 + q_n} \tag{86}$$

but we also have $\sqrt{N} = a_0 + 1/\alpha_1$, and this implies

$$\sqrt{N} = \frac{p_{n+1} + p_n(\sqrt{N} - a_0)}{q_{n+1} + q_n(\sqrt{N} - a_0)} \tag{87}$$

from which it follows that

$$\sqrt{d}(q_{n+1} - q_n a_0 - p_n) + dq_n - p_{n+1} + p_n a_0 = 0 \tag{88}$$

Since $\sqrt{d}$ is irrational, the coeffient of $\sqrt{d}$ as well as the integer part must both be 0, and hence

$$dq_n - p_{n+1} = p_n a_0 \tag{89}$$

$$q_{n+1} - p_n = q_n a_0 \tag{90}$$

If we eliminate a_0 we arrive at

$$p_n^2 - dq_n^2 = q_{n+1}p_n - q_n p_{n+1} = (-1)^{n+1} \tag{91}$$

and n is odd by assumption, so $p_n^2 - dq_n^2 = 1$ and thus $p_n = x$, $q_n = y$ is a (positive) solution to the Pell equation.

The conclusion is that the Pell equation always has infinitely many solutions, since all convergents $\frac{p_n}{q_n}$ of the continued fraction expansion of $\sqrt{N}$ satisfying that n is odd and $m|n+1$, where m is the period of the continued fraction, are solutions of the equation. Moreover, the solutions can be found using continued fractions.

6.2. **RSA and Wiener's attack.** We will now present an application of continued fractions that originate in theoretical computer science, namely from cryptography. Certain computational notions, such as "efficient algorithm" and "difficult problem", will be used here without further explanation, but the concepts are probably quite clear intuitively. In the case of any confusion, a good introductory reference to complexity theory is [Pap94].

The RSA cryptosystem is famous and widely used, and it is defined as follows: let $p, q \in \mathbb{N}$ be (large) primes, and call $n = pq$ the *modulus*. Choose relatively prime numbers $1 < e, d < (p-1)(q-1)$ such that $ed \equiv 1 \pmod{(p-1)(q-1)}$. Publish the pair (n, e) as the *public key*, where e is called the *public exponent*, and keep the *private exponent* d secret. The pair (n, d) is called the *private key*.

The ciphertext x_C of a given plaintext x_P is then computed as $x_C \equiv x_P^e \pmod{n}$ and the plaintext y_P of a given ciphertext y_C is computed as $y_P \equiv y_C^d \pmod{n}$. According to the following famous theorem, these are inverse functions to each other.

Theorem 6.1 (RSA)**.** *For all* $x \in \{0, \dots, n\}$

$$(x^d)^e \equiv (x^e)^d \equiv x \pmod{n} \tag{92}$$

Proof. See [RSA78]. The proof is essentially an application of Fermat's Little Theorem. □

As with most public-key cryptosystems, in RSA everyone can send a message to a given recipient using the public key, but since the decryption uses the private key, only the recipient can then decrypt the message.

The RSA cryptosystem has been very successful in practice, since there are efficient algorithms for all operations involved, including finding large primes, computing e, d and computing exponents modulo n, but there are no *known* efficient algorithms for computing the plaintext from a ciphertext without the knowledge of d, and it can be shown that this problem is equivalent to the well-known difficult problem of factorising n into its prime factors p, q.

However, in the special case where we know that the private exponent d is small, then n can be factorised efficiently using continued fractions, and this is known as the *Wiener attack* on RSA, after its author.

For some $k \in \mathbb{N}$ we have $ed = 1 + k(p-1)(q-1)$ and if we divide by dn and perform some simple manipulations we arrive at

$$\frac{e}{n} - \frac{k}{d} = \frac{1 + k(1-p-q)}{dn} \tag{93}$$

Now, if $\left|\frac{e}{n} - \frac{k}{d}\right| < \frac{1}{2d^2}$ then by theorem 4.12 and remark 3 the rational number $\frac{k}{d}$ must be a convergent of the continued fraction expansion of $\frac{e}{n}$, and the latter consists of public information only. Using (93) we see that the inequality is true when

$$2d < \frac{n}{1 + k(1-p-q)} \tag{94}$$

Thus, if this inequality is satisfied, then we will find k and d from the continued fraction expansion, but as an attacker on the cryptosystem, we don't know in advance if the inequality holds, and we must in either case have a method of determining which convergent of the continued fraction expansion that is $\frac{k}{d}$. This test involves computing guesses of p and q, so we will also get a factorisation of n when we find the correct convergent.

We proceed as follows: given a convergent $\frac{a}{b}$ of $\frac{e}{n}$, compute a guess m of $(p-1)(q-1)$ using (93)

$$m = \frac{eb - 1}{a} \tag{95}$$

and since

$$\frac{p+q}{2} = \frac{n + 1 - (p-1)(q-1)}{2} \tag{96}$$

we then compute a guess u of $\frac{p+q}{2}$ as

$$u = \frac{n - m + 1}{2} \tag{97}$$

Using the following identity

$$\left(\frac{p+q}{2}\right)^2 - n = \left(\frac{p-q}{2}\right)^2 \tag{98}$$

we compute a guess v of $\left(\frac{p-q}{2}\right)^2$, $v = u^2 - n$, and if our guesses are correct we then have that

$$p = u + \sqrt{v} \tag{99}$$

$$q = -(\sqrt{v} - u) \tag{100}$$

and this can be checked, since $n = pq$ is public information.

Thus, we have an efficient method of determining when a convergent $\frac{a}{b} = \frac{k}{d}$, and when we have found the right one, we also get the factorisation of n which completely breaks the cryptosystem. To make the test even more efficient, we can do intermediate checks to make sure that $m > 0$, that u is in fact an integer and that v is a perfect square.

Also, it is important to that the number of possible convergents in the continued fraction is not too large, since otherwise the whole algorithm would not be efficient.

However, one can prove that the number of convergents is not too large, so we can therefore conclude that when the private exponent d satisfies (94), then RSA can be easily cracked using continued fractions.

References

[Bak84] Alan Baker, *A concise introduction to the theory of numbers*, ch. 6 and 8, Cambridge University Press, 1984.

[Dav99] H. Davenport, *The higher arithmetic*, seventh ed., ch. 4, Cambridge University Press, 1999.

[HW79] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, fifth ed., ch. 10–11, Oxford University Press, 1979.

[Pap94] Christos H. Papadimitriou, *Computational complexity*, 1 ed., Addison-Wesley, January 1994.

[RO03] Edmund F. Robertson and John J. O'Connor, *The MacTutor History of Mathematics archive*, 2003, http://www-history.mcs.st-andrews.ac.uk/history/index.html.

[RSA78] Ronald Rivest, Adi Shamir, and Leonard Adleman, *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM **21** (1978), no. 2, 158–164.

[Sta78] Harold M. Stark, *An introduction to number theory*, ch. 7, MIT Press, 1978.

[Wal73] H. S. Wall, *Analytic theory of continued fractions*, Chelsea publishing company, 1973.

[Wei03] Eric W. Weisstein, *Eric weisstein's world of Mathematics*, 2003, http://mathworld.wolfram.com/.

[Wie90] Michael J. Wiener, *Cryptanalysis of short RSA secret exponents*, IEEE Transactions on Information Theory **36** (1990), no. 3, 553–558.

E-mail address: henrik.baarnhielm@imperial.ac.uk

URL: http://www.ma.imperial.ac.uk/~hb103/