2. Continued Fractions

This important theory provides sharp information concerning the quantity

$$q\|q\alpha\|$$

where

$$\|\theta\| = \min_{n\in\mathbb{Z}} |\theta - n|$$

and gives a quick algorithm for finding the best rational approximations to α.

The expression $\|\theta\|$ satisfies the triangle inequality, and provides a metric for the space $\mathbb{R}/\mathbb{Z}$. With this notation we now have a succinct way of writing some of the properties of the continued fraction expansion for α.

A finite or terminating continued fraction is an expression of the kind

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

and is usually written $a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_n}$ or $[a_0; a_1, a_2, \ldots, a_n]$. We will suppose that $a_0 \in \mathbb{Z}$ and that $a_1, \ldots, a_n$ are positive integers. Often we will insist that $a_n > 1$. Obviously any finite continued fraction represents a rational number. Converesely, if $(b_0, r_0) = 1$, then we can write

$$\begin{aligned}
b_0 &= a_0 r_0 + r_1, \quad 0 \le r_1 < r_0,\\
r_0 &= a_1 r_1 + r_2, \quad 0 \le r_2 < r_1,\\
r_1 &= a_2 r_2 + r_3, \quad 0 \le r_3 < r_2,\\
&\ldots
\end{aligned}$$

and since the r_j are strictly decreasing the process must terminate with

$$r_{n-1} = a_n r_n.$$

Now

$$\begin{aligned}
\frac{b_0}{r_0} &= a_0 + \frac{r_1}{r_0} = a_0 + \frac{1}{\frac{r_1}{r_0}}\\
&= a_0 + \cfrac{1}{a_1 + \frac{r_2}{r_1}}\\
&= [a_0; a_1, a_2, \ldots, a_n].
\end{aligned}$$

This process gives a one-to-one correspondence between rational numbers b_0/r_0 and finite sequences $a_0, a_1, \ldots, a_n$ of integers with $a_1 \ge 1, \ldots, a_{n-1} \ge 1, a_n > 1$.

Put

$$\begin{aligned}
p_{-2} = 0,\ p_{-1} = 1,\ p_m = a_m p_{m-1} + p_{m-2} \quad (m \ge 0)\\
q_{-2} = 1,\ q_{-1} = 0,\ q_m = a_m q_{m-1} + q_{m-2} \quad (m \ge 0)
\end{aligned}$$

Then p_m and q_m are polynomials in $a_0, \dots, a_m$ with integral coefficients. We show by induction that these polynomials are interconnected by the identity

$$p_m q_{m-1} - p_{m-1} q_m = (-1)^{m-1} \quad (m \geq -1).$$

This is obvous for $m = -1$ or $m = 0$. Suppose that it holds for m replaced by $m-1$. Then

$$\begin{aligned} p_m q_{m-1} - p_{m-1} q_m &= (a_m p_{m-1} + p_{m-2} q_{m-1} - p_{m-1}(a_m q_{m-1} + q_{m-2}) \\ &= -(p_{m-1} q_{m-2} - p_{m-2} q_{m-1}) = -(-1)^{m-2} \\ &= (-1)^{m-1}. \end{aligned}$$

By induction we also show that

$$[a_0; a_1, \dots, a_m] = \frac{p_m}{q_m} \quad (m \geq 0).$$

This is clear for $m = 0$. Suppose this holds for m replaced by $m-1$. Then

$$\begin{aligned} [a_0; a_1, \dots, a_m] &= \left[a_0; a_1, \dots, a_{m-1} + \frac{1}{a_m} \right] \\ &= \frac{p_{m-1}\left(a_0, a_1, \dots, a_{m-1} + \frac{1}{a_m}\right)}{q_{m-1}\left(a_0, a_1, \dots, a_{m-1} + \frac{1}{a_m}\right)} \\ &= \frac{\left(a_{m-1} + \frac{1}{a_m}\right) p_{m-2} + p_{m-3}}{\left(a_{m-1} + \frac{1}{a_m}\right) q_{m-2} + q_{m-3}} \\ &= \frac{a_m(a_{m-1} p_{m-2} + p_{m-3}) + p_{m-2}}{a_m(a_{m-1} q_{m-2} + q_{m-3}) + q_{m-2}} \\ &= \frac{a_m p_{m-1} + p_{m-2}}{a_m q_{m-1} + q_{m-2}} \\ &= \frac{p_m}{q_m}. \end{aligned}$$

These are polynomial identities, but when the a_m are assigned integer values the p_m and q_m are also integers, and $(p_m, q_m) = 1$.

Let α be a real *irrational* number. Put $a_0 = [\alpha]$ and write $\alpha = a_0 + 1/\alpha_1$. Then $\alpha_1 > 1$. Put $a_1 = [\alpha_1]$ and write $\alpha_1 = a_1 + 1/\alpha_2$. Continuing in this manner, we find that $\alpha = [a_0; a_1, a_2, \dots, a_{n-1}, \alpha_n]$. This defines a seuence of integers $a_0, a_1 > 0, a_2 > 0, \dots$, real numbers $\alpha_1, \alpha_2, \dots$ greater than 1, and corresponding integers p_n, q_n. We now show that $p_n/q_n \to \alpha$. We begin be observing that

$$\begin{aligned} \frac{p_n}{q_n} &= a_0 + \sum_{m=1}^{n} \left(\frac{p_m}{q_m} - \frac{p_{m-1}}{q_{m-1}} \right) \\ &= a_0 + \sum_{m=1}^{n} \frac{(-1)^{m-1}}{q_{m-1} q_m}. \end{aligned}$$

The integers q_i are positive and strictly increasing for $i \geq 1$. Thus $q_i \to \infty$ (actually $q_i \geq f_i$, the i-th Fibonacci number, so that $q_i > c^i$ with $c > 1$). Hence, by the alternating series test the sum above converges as $n \to \infty$. Therefore p_n/q_n converges. Moreover

$$\alpha = \frac{p_{n+1}(a_0, \dots, a_n, \alpha_{n+1})}{q_{n+1}(a_0, \dots, a_n, \alpha_{n+1})} = \frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}}. \tag{1}$$

Thus

$$\alpha - \frac{p_n}{q_n} = \frac{q_n p_{n-1} - p_n q_{n-1}}{\alpha_{n+1}q_n + q)_{n-1}} = \frac{(-1)^n}{q_n(\alpha_{n+1}q_n + q_{n-1}}$$

and α lies between the even and odd convergents. Hence $\lim_{n\to\infty} p_n/q_n = \alpha$ and we write

$$\alpha = [a_0; a_1, a_2 \dots].$$

Conversely, let $a_0, a_1, \dots$ be integers with $a_i \geq 1$ when $i \geq 1$. Put $\alpha = [a_0; a_1, \dots]$, $\alpha_1 = [a_1; a_2, \dots]$, $\alpha_2 = [a_2; a_3, \dots]$, and so on. The $\alpha_n = a_n + 1/\alpha_{n+1}$, $\alpha_{n+1} > 1$, and we have $a_n = [\alpha_n]$ for $n = 1, 2, 3, \dots$. Hence $a_0, a_1, \dots$ are the numbers that would have been generated if we had started from with α. Thus we have proved

Theorem 2.1. *The continued fraction expansion provides a one-to-one correspondence between real numbers α and the sequence $a_0, a_1, \dots$, of integers with $a_i \geq 1$ for $i > 0$. The sequence terminates if and only if α is rational and in this case we require that $\alpha = [a_1; a_1, \dots, a_n]$ with $a_n > 1$.*

Theorem 2.2. *We have $\|q_1\alpha\| > \|q_2\alpha\| > \dots$, and when $0 < q < q_n$ we have $\|q\alpha\| \geq \|q_{n-1}\alpha\|$.*

This characterises the q_n as the set of those q for which the value of $\|q\alpha\|$ is record breaking. Cassels takes this to be the definition of the q_n, and recovers the properties we started with (and Cassels also adopts a different convention with regard to indexing).

Proof. We have

$$\begin{aligned}\alpha = [a_0; a_1, \dots, a_n, \alpha_{n+1}] &= \frac{p_{n+1}(a_0, \dots, a_n, \alpha_{n+1})}{q_{n+1}(a_0, \dots, a_n, \alpha_{n+1})} \\ &= \frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}}.\end{aligned}$$

Hence

$$\begin{aligned}\alpha - \frac{p_n}{q_n} &= \frac{p_{n-1}q_n - p_n q_{n-1}}{q_n(\alpha_{n+1}q_n + q_{n-1})} \\ &= \frac{(-1)^n}{q_n(\alpha_{n+1}q_n + q_{n-1})}.\end{aligned}$$

Consequently

$$\|q_n\alpha\| = |q_n\alpha - p_n| = \frac{1}{\alpha_{n+1}q_n + q_{n-1}}.$$

The assertion that $\|q_n\alpha\|$ is strictly decreasing is equivalent to the assertion that $|\alpha_{n+1}q_n + q_{n-1}|$ is strictly increasing. As

$$\begin{aligned}\alpha_{n+1}q_n + q_{n-1} &\geq q_n + q_{n-1} = a_n q_{n-1} + q_{n-2} + q_{n-1} \\ &= (a_n + 1)q_{n-1} + q_{n-2} > \alpha_n q_{n-1} + q_{n-2}\end{aligned}$$

the first assertion is established.

Suppose now that $0 < q < q_n$. We first show that for any integers q and p there exist integers x and y such that

$$\begin{aligned}q &= q_n x + q_{n-1}y, \\ p &= p_n x + p_{n-1}y.\end{aligned}$$

Here the coefficient matrix has $\det = \pm 1$, so $(q,p) \in \mathbb{Z}^2 \iff (x,y) \in \mathbb{Z}^2$. If $x = 0$, then $q\alpha - p = y(q_{n-1}\alpha - p_{n-1})$, so that $|q\alpha - p| = |y|\|q_{n-1}\alpha\| \geq \|q_{n-1}\alpha\|$. If $y = 0$, then $q = q_n x \geq q_n$, contrary to the assumption that $0 < q < q_n$. Finally, suppose that neither x nor y is 0. From the relations $q = q_n x + q_{n-1}y$, $0 < q < q_n$ it follows that x and y are of opposite signs. We also know that $q_n\alpha - p_n$ and $q_{n-1}\alpha - p_{n-1}$ are of opposite sign. Hence $|q\alpha - p| = |x||q_n\alpha| + |y||q_{n-1}\alpha| \geq \|q_{n-1}\alpha\|$ and the proof is complete.

Since $\|q\alpha\| > \|q_n\alpha\|$ for all $q < q_n$, it follows that $|\alpha - p/q| > |\alpha - p_n/q_n|$ for $q < q_n$. We can also show that the rational approximations p_n/q_n include all good approximations in the following sense.

Theorem 2.3 (Legendre). *If $|\alpha - p/q| < 1/(2q^2)$ and $(p,q) = 1$, then $q = q_n$ for some n.*

Proof. Choose n so that $q_n \leq q < q_{n+1}$. Since $pq_n - qp_n = q(\alpha q_n - p_n) - q_n(q\alpha - p)$, by the triangle inequality,

$$|pq_n - qp_n| \leq q\|q_n\alpha\| + q_n\|q\alpha\|.$$

By Theorem 2.2 we know that $\|q_n\alpha\| \leq \|q\alpha\|$. Hence the above is

$$\leq q\|q\alpha\| + q\|q\alpha\| = 2q^2|\alpha - p/q| < 1.$$

Thus $pq_n - qp_n = 0$. As $(p_n, q_n) = (p,q) = 1$, it follows that $q = q_n$, and the proof is complete

A number α is *badly approximable* when have $\inf\{\beta : \beta = q\|q\alpha\|, q \in \mathbb{Z}\} > 0$. Put $M_n = \alpha_{n+1} + q_{n-1}/q_n$. By the first part of the proof of Theorem 2.2 we see that

$$\left|\alpha - \frac{p_n}{q_n}\right| = \frac{1}{M_n q_n^2}.$$

Since $a_{n+1} < M_n < a_{n+1} + 2$, we have at once

Theorem 2.4. *A real irrational number α is badly approximable if and only if its partial fraction coefficients are bounded.*

Since $M_n > 1$ for all n, the pairs p_n, q_n give solutions of the inequality $|\alpha - p/q| < q^{-2}$. To obtain more refined results we will later examine M_n more closely. Before proceeding it is useful to note that

$$\frac{q_{n-1}}{q_n} = \frac{q_{n-1}}{a_n q_{n-1} + q_{n-2}} = \frac{1}{a_n + \frac{q_{n-2}}{q_{n-1}}}.$$

Hence $q_{n-1}/q_n = [0; a_n, a_{n-1}, \dots, a_1]$, and we have

$$\begin{aligned} M_n &= [a_{n+1}; a_{n+2}, a_{n+3}, \dots] + [0; a_n, a_{n-1}, \dots, a_1] \\ &= \alpha_{n+1} + \frac{1}{\beta_n} \end{aligned}$$

where

$$\beta_n = [a_n, a_{n-1}, \dots, a_1].$$

Theorem 2.5. *For each n at least one of the inequalities*

$$\left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| < \frac{1}{2q_{n-1}^2}, \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{2q_n^2}$$

holds.

Proof. Since $\alpha - p_n/q_n$ and $\alpha - p_{n-1}/q_{n-1}$ are of opposite signs, we have

$$\left|\frac{p_n}{q_n} - \alpha\right| + \left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{1}{q_n q_{n-1}}.$$

Since $q_n \neq q_{n-1}$, we have the trivial inequality

$$\frac{1}{q_n q_{n-1}} < \frac{1}{2q_n^2} + \frac{1}{2q_{n-1}^2}$$

and so at least on of the inequalities in question must hold.

Theorem 2.6 (Hurwitz, 1891). *The inequality*

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{\sqrt{5}q_n^2}$$

holds for at least one of any three consecutive values of n.

First proof. Suppose that, to the contrary,

$$\left|\alpha - \frac{p_n}{q_n}\right| \geq \frac{1}{\sqrt{5}q_n^2}, \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right| \geq \frac{1}{\sqrt{5}q_{n+1}^2}, \left|\alpha - \frac{p_{n+2}}{q_{n+2}}\right| \geq \frac{1}{\sqrt{5}q_{n+2}^2}.$$

Since

$$\frac{1}{q_n q_{n+1}} = \left|\alpha - \frac{p_n}{q_n}\right| + \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right|,$$

the first two inequalities imply that

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

We multiply each side of the equation by $\sqrt{5}q_{n+1}^2$. Thus $f(q_{n+1}/q_n) \leq 0$ where $f(u) = u^2 - \sqrt{5}u + 1$. The equation $f(u) = 0$ has the two roots $(\sqrt{5} \pm 1)/2$. Thus

$$q_{n+1} \leq cq_n \tag{2}$$

where $c = (1 + \sqrt{5})/2$. Similarly by the second and third inequalities we obtain

$$q_{n+2} \leq cq_{n+1}. \tag{3}$$

On the other hand, $q_{n+2} = a_{n+2}q_{n+1} + q_n \geq q_{n+1} + q_n$. We combine this (3) and find that $(c-1)q_{n+1} \geq q_n$. Hence, by (2),

$$\frac{q_n}{c-1} \leq q_{n+1} \leq cq_n.$$

We have $\frac{1}{c-1} = c$, so the above chain of inequalities implies that $q_{n+1} = cq_n$. But c is irrational whilst q_n and q_{n+1} are integers, so we have a contradiction and the theorem follows.

Second proof. We have $M_n = \alpha_{n+1} + q_{n-1}/q_n$ and

$$q_n/q_{n-1} = \beta_n = [a_n, a_{n-1}, \dots, a_1].$$

Thus

$$M_n = \alpha_{n+1} + \frac{1}{\beta_n} = \alpha_{n+1} + \frac{1}{a_n + \frac{1}{\beta_{n-1}}} = \frac{1}{\lambda} + \frac{1}{a+\mu}$$

where $\lambda = 1/\alpha_{n+1}$, $a = a_n$ and $\mu = 1/\beta_{n-1}$. Similarly

$$M_{n-1} = \alpha_n + \frac{1}{\beta_{n-1}} = a_n + \frac{1}{\alpha_{n+1}} + \frac{1}{\beta_{n-1}} = a + \lambda + \mu,$$

and

$$\begin{aligned} M_{n-2} &= \alpha_{n-1} + \frac{1}{\beta_{n-2}} = \left(a_{n-1} + \frac{1}{\alpha_n}\right) + (\beta_{n-1} - a_{n-1}) = \frac{1}{\alpha_n} + \beta_{n-1} \\ &= \frac{1}{a_n + \frac{1}{\alpha_{n+1}}} + \beta_{n-1} = \frac{1}{a+\lambda} + \frac{1}{\mu}. \end{aligned}$$

We note that $M_n = \frac{M_{n-1}}{\lambda(a+\mu)}$ and $M_{n-2} = \frac{M_{n-1}}{\mu(a+\lambda)}$. Thus

$$\begin{aligned} M_{n-1}\left(\frac{1}{M_n} + \frac{1}{M_{n-2}}\right) &= \lambda(a+\mu) + \mu(a+\lambda) \\ &= a(M_{n-1} - a) + 2\lambda\mu. \end{aligned}$$

We multiply both sides by 2, and note that $4\lambda\mu \le (\lambda+\mu)^2 = (M_{n-1}-a)^2$. Thus

$$2M_{n-1}\left(\frac{1}{M_n}+\frac{1}{M_{n-2}}\right) \le 2a(M_{n-1}-a)+(M_{n-1}-a)^2$$
$$= M_{n-1}^2 - a^2.$$

We divide both sides by M_{n-1}^2 and rearrange. Thus

$$\frac{2}{M_{n-1}M_n}+\frac{2}{M_{n-2}M_{n-1}}+\frac{a^2}{M_{n-1}^2} \le 1. \tag{4}$$

Since $a \ge 1$, the above inequality implies that we cannot have $M_{n-1} < \sqrt{5}$, $M_N < \sqrt{5}$, $M_{n-2} < \sqrt{5}$. Moreover, if $M_{n-1} \le \sqrt{5}$, $M_N \le \sqrt{5}$, $M_{n-2} \le \sqrt{5}$, then $a = 1$ and $M_{n-2} = M_{n-1} = M_n = \sqrt{5}$. But then $1 = \frac{M_{n-1}}{M_n} = \lambda(a+\mu) = \lambda(1+\mu)$ and $1 = \frac{M_{n-1}}{M_{n-2}} = \mu(a+\lambda) = \mu(1+\lambda)$. Hence $\lambda(1+\mu) = \mu(1+\lambda)$, which implies that $\lambda = \mu$. This is impossible since $\lambda = 1/\alpha_{n+1}$ is irrational, whilst $\mu = 1/\beta_{n-1}$ is rational. Therefore at least one of M_{n-2}, M_{n-1}, M_n exceeds $\sqrt{5}$.

The inequality (4) above contains further interesting information. If $a \ge 2$, then at least one of M_{n-2}, M_{n-1}, M_n is at least $\sqrt{8}$. The possibility that all three of these quantities is $\sqrt{8}$ can be eliminated much as above. Thus we find that if $\alpha = [a_0; a_1, a_2, \dots]$ and there are infinitely many m for which $a_m \ge 2$, then the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{8}q^2}$$

has infinitely many solutions. On the other hand, if $\alpha = \frac{1+\sqrt{5}}{2}$, then $\alpha^2 - \alpha - 1 = 0$, and hence $\alpha = 1 + \frac{1}{\alpha}$, so that $\alpha = [1; 1, 1, \dots]$. Since $M_n = \alpha_{n+1} + [0; 1, 1, \dots, 1] = \alpha + \left(\frac{p_n}{q_n} - 1\right) \sim 2\alpha - 1 = \sqrt{5}$, this gives a second proof of Theorem 1.4.

Let $M(\alpha) = \limsup_{n\to\infty} M_n$. Then we always have $M(\alpha) \ge \sqrt{5}$, and $M(\alpha) = \sqrt{5}$ if and only if $a_n = 1$ for all sufficiently large n. Otherwise $M(\alpha) \ge \sqrt{8}$. It can be shown that $M(\alpha) = \sqrt{8}$ if and only if $a_n = 2$ for all large n, that otherwise $M(\alpha) \ge \frac{\sqrt{221}}{5}$, that $M(\alpha) = \frac{\sqrt{221}}{5}$ if and only if the tail of the continued fraction for α is $\dots, 2, 2, 1, 1, 2, 2, 1, 1, 2, 2, 1, 1, \dots$, that otherwise $M(\alpha) \ge \dots$.

We call a positive integer a Markov number when there exist integers v, w such that $u^2 + v^2 + w^2 = 3uvw$. Then

$$M\left(\frac{1}{2u}\left(\sqrt{9u^2-4}+u+\frac{2v}{w}\right)\right) = \sqrt{9u^2-4}\,u.$$

There are infinitely many triples u, v, w. The complete set can be generated systematically. The first Markov numbers are 1, 2, 5, 13, 29, 34, 89, 169, 194,.... For indefinite binary quadratic forms $f(x,y) = ax^2+bxy+cy^2$ with real coefficients we have a corresponding family of results. Put

$$\mu(f) = \inf_{\substack{x,y\in\mathbb{Z}\\(x,y)\neq(0,0)}} \frac{|ax^2+bxy+cy^2|}{\sqrt{b^2-4ac}}.$$

Then $\mu(f) \le 1/\sqrt{5}$ always, $\mu(f) = 1/\sqrt{5}$ if and only if $f \sim x^2 + xy - y^2$, $\mu(f) \le 1/\sqrt{8}$ otherwise, $\mu(f) = 1/\sqrt{8}$ if and only if $f \sim x^2 - 2y^2$, and so on. The full details can be found in Chapter II of Cassel's Cambridge Tract on Diophantine Approximation. We can say something without too much elaboration. Let α be a root of $f(\gamma, 1) = 0$, so that $f(x, y) = a(x - \alpha y)^2 + (\sqrt{b^2 - 4c})(x - \alpha y)$, and consider the convergents p_n/q_n to the continued fraction for α. If we take $y = q_n$, $x = p_n$, then $|f(x, y)|/\sqrt{b^2 - 4ac} = \frac{1}{M_n} + O(q_n^{-2})$ and so $\mu(f) \le M(\alpha)^{-1}$. We have likewise for the other root β of $f(\gamma, 1) = 0$, so that

$$\mu(f) \le \min(M(\alpha)^{-1}, M(\beta)^{-1}).$$

When a, b, c are rational, then the values taken on by $f(x, y)$ for integral x, y are discrete, and hence so are those taken on by $f(x, y)/\sqrt{b^2 - 4ac}$. Thus the infimum must be attained. Moreover, it can be shown that there are arbitrarily large such integers with, in addition, $x - y\alpha$ arbitrarily small. It then follows that $\mu(f) \ge 1/M(\alpha)$ (and that $M(\alpha) = M(\beta)$.

We say that α is equivalent to α', and write $\alpha \sim \alpha'$, when there exist integers a, b, c, d with $ad - bc = \pm 1$, $\alpha = \frac{a\alpha' + b}{c\alpha' + d}$. These maps are composed by matrix multiplication, and hence this is an equivalence relation in the usual sense. A real number is expanded as a continued fraction by repeatedly applying the two transformations $\alpha \to \alpha - 1$, $\alpha \to 1/\alpha$. These are of the required type and so $\alpha \sim \alpha_n$ for all n. More explicitly, by (1) above,

$$\alpha = \frac{p_{n-1}\alpha_n + p_{n-2}}{q_{n-1}\alpha_n + q_{n-2}}.$$

In the converse direction we prove

Theorem 2.7. *Suppose that $\alpha = \frac{a\alpha' + b}{c\alpha' + d}$ where a, b, c, d are integers, $ad - bc = \pm 1$, $\alpha' > 1$, $c > d > 0$. Then for some n we have $\alpha' = \alpha_n$, $a = p_{n-1}$, $b = p_{n-2}$, $c = q_{n-1}$, $d = q_{n-2}$ where $\frac{p_{n-2}}{q_{n-2}}$ and $\frac{p_{n-1}}{q_{n-1}}$ are successive convergents to α.*

Proof. Write $a/c = [b_0; b_1, b_2, \dots, b_{m-1}]$, and let r_j/s_j denote the typical convergents of this continued fraction, so that $r_{m-1} = a$, $s_{m-1} = c$. Each rational number has two expansions, one with last coefficient 1 and one with $b_{m-1} > 1$. By choosing between these we may control the parity of m, and thus we may arrange that $r_{m-1}s_{m-2} - s_{m-1}r_{m-2} = ad - bc$. That is, $r_{m-1}s_{m-2} - s_{m-1}r_{m-2} = r_{m-1}d - s_{m-1}b$, so that $r_{m-1}(s_{m-2} - d) = s_{m-1}(r_{m-2} - b)$. Hence $d \equiv s_{m-2} \pmod{s_{m-1}}$. Moreover $0 < s_{m-2} < s_{m-1}$ and $0 < d < c = s_{m-1}$, so that $d = s_{m-2}$, and thus $r_{m-2} = b$. Therefore

$$\alpha = \frac{a\alpha' + b}{c\alpha' + d} = \frac{r_{m-1}\alpha' + r_{m-2}}{s_{m-1}\alpha' + s_{m-2}} = [b_0; b_1, b_2, \dots, b_{m-1}, \alpha'].$$

Since $\alpha' > 1$, it has continued fraction $\alpha' = [c_0; c_1, c_2, \dots]$ with $c_0 \ge 1$. Thus $\alpha = [b_0; b_1, \dots, b_{m-1}, c_0, c_1, \dots]$. By the uniqueness of the continued fraction expansion for α it now follows that $b_j = a_j$, $\alpha' = \alpha_m$, $r_j = p_j$ $(j = 0, \dots, m-1)$, $s_j = q_j$ $(j = 0, \dots, m-1)$.

Theorem 2.8. *Let $\alpha = [a_0; a_1, a_2, \dots]$, $\alpha' = [a'_0; a'_1, a'_2, \dots]$. If $\alpha \sim \alpha'$, then there is an integer k (possibly negative) such that $a_{n+k} = a'_n$ for all sufficiently large n, and vice versa.*

Proof. We suppose furst of all that $\alpha \sim \alpha'$. We have $\alpha' = \frac{a\alpha+b}{c\alpha+d}$ where a, b, c, d are integers with $ad - bc = \pm 1$. If $c\alpha + d < 0$, then multiply each a, b, c, d by -1. Thus we may assume that $c\alpha + d > 0$. We have, for any k,

$$\alpha = \frac{\alpha_{k+1}p_k + p_{k-1}}{\alpha_{k+1}q_k + q_{k-1}},$$

so that

$$\begin{aligned}\alpha' &= \frac{a(\alpha_{k+1}p_k + p_{k-1} + b(\alpha_{k+1}q_k + q_{k-1})}{c(\alpha_{k+1}p_k + p_{k-1} + d(\alpha_{k+1}q_k + q_{k-1})} \\ &= \frac{A\alpha_{k+1} + B}{C\alpha_{k+1} + D},\end{aligned}$$

say. Here $D = cp_{k-1} + dq_{k-1} = q_{k-1}\left(c\frac{p_{k-1}}{q_{k-1}} + d\right) > 0$ when k is large, since $c\frac{p_{k-1}}{q_{k-1}} + d \to c\alpha + d > 0$ as $k \to \infty$. As $C = cp_k + d_k$, we see that

$$\begin{aligned}C - D &= c(p_k - p_{k-1}) + d(q_k - q_{k-1}) \\ &= c((a_k - 1)p_{k-1} + p_{k-2}) + d((a_k - 1)q_{k-1} + q_{k-2})) \\ &= (a_k - 1)(cp_{k-1} + dq_{k-1}) + (cp_{k-2} + dq_{k-2}) \\ &> 0\end{aligned}$$

for all sufficiently large k. Moreover, $\alpha_{k+1} > 1$. We also have

$$\det\begin{pmatrix} A & B \\ C & D\end{pmatrix} = \det\left(\begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1}\end{pmatrix}\right) = \det\begin{pmatrix} a & b \\ c & d\end{pmatrix}\det\begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1}\end{pmatrix}$$

and so $AD - BC = \pm 1$. Thus the hypotheses of Theorem 2.7 are satisfied, so that $\alpha_{k+1} = \alpha'_{n+1}$ for some n. Then $a_{m+k} = a'_{m+n}$ for $m = 1, 2, 3, \dots$, i.e. $a_{m+(k-n)} = a'_m$ for all sufficiently large m.

We now consider the converse. Suppose that for every integer non-negative integer $n \geq n_0$ we have $a_{n+k} = a'_n$. We suppose $k \geq 0$, for otherwise we could interchange α and α' in the following argument. First suppose that $n_0 = 0$. Then $\alpha' = [a_k; a_{k+1}, \dots]$ and so $\alpha = [a_0; a_1, \dots, a_{k-1}, \alpha'] = \frac{\alpha' p_{k-1} + p_{k-2}}{\alpha' q_{k-1} + q_{k-2}}$. Thus, by one of the identities connecting the q_n and p_n we see that $\alpha \sim \alpha'$. Now suppose that $n_0 > 0$. Let $\alpha'' = [a_{n_0+k}; a_{n_0+k+1}, \dots]$. Then $\alpha \sim \alpha'' \sim \alpha'$ by the previous argument.

Theorem 2.9. *If $\alpha \sim \alpha'$, then $M(\alpha) = M(\alpha')$.*

Proof. By the previous theorem there is an integer k such that $a_{n+k} = a'_n$ for all $n \geq n_0$. If necessary by interchanging α and α', we may suppose that $k \geq 0$. It then suffices to show that, for any m, $M(\alpha) = M(\alpha_m)$, for then, by the same token, $M(\alpha') = M(\alpha_{n_0+k})$ and the conclusion follows. Now $M_n = \alpha_{n+1} + q_{n-1}/q_n$. Moreover $\alpha_{n+1} = (\alpha_m)_{n-m+1}$, provided that n is large. We also know

that $q_{n-1}/q_n = [0; a_n, \dots, a_1]$. Let Q_n, Q'_n be positive integers with $(Q_n, Q'_n) = 1$ so that

$$[0; a_n, \dots, a_{m+1}] = \frac{Q_n}{Q'_n}.$$

It suffices to show that

$$q_{n-1}/q_n \sim Q_n/Q'_n \tag{5}$$

as $n \to \infty$. Choose P, P' likewise so that

$$[0; a_n, \dots, a_{m+2}] = \frac{P}{P'}.$$

Then, by the standard identities for finite continued fractions, we have

$$\begin{aligned}\frac{q_{n-1}}{q_n} &= [0; a_n, \dots, a_{m+1}, q_m/q_{m-1}] \\ &= \frac{\frac{q_m}{q_{m-1}} Q_n + P}{\frac{q_m}{q_{m-1}} Q'_n + P'} \\ &= \frac{q_m Q_n + q_{m-1} P}{q_m Q'_n + q_{m-1} P'}.\end{aligned}$$

Both $q_m Q_n + q_{m-1} P$ and $q_m Q'_n + q_{m-1} P'$ divide $q_{m-1}(PQ'_n - Q_n P') = \pm q_{m-1}$ and $q_m(PQ'_n - Q_n P') = \pm q_m$ and so are coprime. Thus $q_{n-1} = q_m Q_n + q_{m-1} P$ and $q_n = q_m Q'_n + q_{m-1} P'$. Thus $q_{n-1} Q'_n - q_n Q_n = \pm q_{m-1}$, and so

$$\frac{q_{n-1}}{q_n} = \frac{Q_n}{Q'_n} \pm \frac{q_{m-1}}{q_n Q'_n}.$$

This gives (5) and completes the proof of the theorem.

The proof of Theorem generalises to arbitrary quadratic irrationals, and so we know that the continued fraction coefficients of a quadratic irrational are bounded. We can now prove more.

Theorem 2.10 (Lagrange). *The continued fraction coefficients of a real quadratic irrational number are ultimately periodic, and conversely.*

Proof. Suppose the coefficients are ultimately periodic. When the coefficients are purely periodic we have $\alpha = [a_0; a_1, \dots, a_k, a_0, a_1, \dots]$. Thus $\alpha_{k+1} = \alpha$ and we have

$$\alpha = \frac{\alpha_{k+1} p_k + p_{k-1}}{\alpha_{k+1} q_k + q_{k-1}} = \frac{\alpha p_k + p_{k-1}}{\alpha q_k + q_{k-1}}.$$

On solving for α we see that α is a quadratic irrational. For a general periodic α, choose n so that α_n is purely periodic. Then $\alpha \sim \alpha_n$ and so α is also a quadratic irrational.

Conversely, suppose that $A\alpha^2 + B\alpha + C = 0$ where A, B, C are integers with $A > 0$ and $D > 0$ where $D = B^2 - 4AC$ and D is not a perfect square. We have $\alpha = \frac{\alpha_n p_{n-1} + p_{n-2}}{\alpha_n q_{n-1} + q_{n-2}}$ and an alternative way of writing this is, in terms of vectors and matrices,

$$(\alpha, 1) = (\alpha_n q_{n-1} + q_{n-2})^{-1} (\alpha_n, 1) P_n,$$

where

$$P_n = \begin{pmatrix} p_{n-1} & q_{n-1} \\ p_{n-2} & q_{n-2} \end{pmatrix}$$

Therefore

$$(\alpha_n, 1) P_n R P_n^T (\alpha_n, 1)^T = 0,$$

where

$$R = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$$

and so

$$A_n \alpha_n^2 + B_n \alpha_n + C_n = 0$$

where

$$A_n = A p_{n-1}^2 + B p_{n-1} q_{n-1} + C q_{n-1}^2,$$

$$C_n = A p_{n-2}^2 + B p_{n-2} q_{n-2} + C q_{n-2}^2,$$

and

$$B_n^2 - 4 A_n C_n = -4 \det(P_n R P_n^T) = D.$$

Now

$$A_n q_{n-1}^{-2} = -A \left(\alpha - \frac{p_{n-1}}{q_{n-1}} \right) \left(\alpha + \frac{p_{n-1}}{q_{n-1}} \right) - B \left(\alpha - \frac{p_{n-1}}{q_{n-1}} \right)$$

and so A_n is uniformly bounded in n. Likewise so is C_n. Therefore so is B_n. Thus we have an infinite number of bounded triples and so there must be infinitley many n for which the triples are identical. Moreover at most two possible values of α_n are associated with each one. Thus there are infinitely many n for which the α_n are identical. In particular there are at least two different values of n for which α_n is identical, and so the continued fraction for α is periodic from some point onwards.

The length of the period can be shown to be $\ll \sqrt{D} \log D$ and numerical evidence suggests that it is 'usually' of order $\sqrt{D}$.

Theorem 2.11 (H. J. S. Smith). *Suppose that $\alpha > 0$, and $[a_0; a_1, a_2, \dots]$ is the continued fraction for α. Let $\mathcal{C}_e$ be the piecewise linear path with vertices $(p_{-2}, q_{-2}), (p_0, q_0), (p_2, q_2), \dots$ and let $\mathcal{C}_o$ be the piecewise linear path with vertices $(p_{-1}, q_{-1}), (p_1, q_1), \dots$. Then the region between $\mathcal{C}_o$ and $\mathcal{C}_e$ (which contains the ray $y = \alpha x$) in the first quadrant contains no lattice points.*

Proof. We show that the triangle with vertices $(0,0), (p_{n-1}, q_{n-1}), (p_{n+1}, q_{n+1})$ contains no lattice point in its interior. This triangle has area

$$\left| \frac{1}{2} \det \begin{bmatrix} 1 & 0 & 0 \\ 1 & p_{n-1} & q_{n-1} \\ 1 & p_{n+1} & q_{n+1} \end{bmatrix} \right| = \frac{1}{2} a_n.$$

The edge from (p_{n-1}, q_{n-1}) to (p_{n+1}, q_{n+1}) contains $a_n - 1$ interior lattice points, namely the points $(p_{n-1} + k p_n, q_{n-1} + k q_n)$, $k = 1, 2, \dots, a_n - 1$. Thus this triangle can be broken in to a_n triangles each with area $\frac{1}{2}$. Moreover a triangle with lattice points for vertices and having area $\frac{1}{2}$ can contain no interior lattice point.

We have already seen how rational approximations to real numbers are relevant to Pell's equation. Thus it is no surprise that continued fractions can be applied to that equation, and indeed to general indefinite quadratic forms. We know that when d is not a perfect square, then all solutions of the equation $x^2 - dy^2 = 1$ with $y > 0$, are obtained by taking the one with y minimal, the fundamental one, call it (x_0, y_0), and taking $(x, y) = (x_n, y_n)$ where $x_n + y_n\sqrt{d} = (x_0 + y_0\sqrt{d})^{(n-1)}$. The equation $x^2 - dy^2 = -1$ may have no solution in integers, but when it does it again has a fundamental solution (x_1, y_1) with $x_1 > 0$, $y_1 > 0$, y_1 minimal and (x_n, y_n) given by $x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$ gives all solutions in positive integers x, y to $x^2 - dy^2 = \pm 1$. In particular $x_0 = x_1^2 + dy_1^2$ and $y_0 = 2x_1y_1$. Our knowledge concerning the distribution of the numbers y_0 is incomplete, but it seems that for most d we have $\log y_0$ about order of magnitude $\sqrt{d}$. Thus a crude search for y_0 can be prohibitively expensive. However we have the following theorem.

Theorem 2.12. *Suppose that d is a positive integer but not a perfect square, that $x > 0$, $y > 0$, $(x, y) = 1$ and $|x^2 - dy^2| < \sqrt{d}$. Then $x = p_n$, $y = q_n$ for some n where p_n/q_n is a convergent to the continued fraction for $\sqrt{d}$.*

Proof. Let $N = x^2 - dy^2$ and suppose first that $0 < N < \sqrt{d}$. Then

$$\left|\frac{x}{y} - \sqrt{d}\right| = \frac{1}{y}|x - y\sqrt{d}| = \frac{N}{y(x + y\sqrt{d})} < \frac{\sqrt{d}}{y(x + y\sqrt{d})}.$$

Since $N > 0$ it follows that $x - y\sqrt{d} > 0$, and hence that $x + y\sqrt{d} > 2y\sqrt{d}$. Thus the expression displayed on the right above is $< 1/(2y^2)$, and the result follows from Legendres theorem (Theorem 2.3).

Suppose now that $-\sqrt{d} < N < 0$. Let $d' = 1/d$ and $N' = -N/d$. Then $y^2 - d'x^2 = N'$ and $0 < N' < \sqrt{d'}$. The above argument makes no use of the integrality of d and N, and when repeated here shows that y/x is a convergent to $\sqrt{d'}$, and hence x/y is a convergent to $\sqrt{d}$ (observe that if $\sqrt{d'} = [a'_0; a'_1, \dots, a'_n, \alpha'_{n+1}]$ and $p'_n/q'_n = [a'_0; a'_1, \dots, a'_n]$ is the convergent, then $a'_0 = 0$ and $\sqrt{d} = [a'_1; a'_2 \dots, a'_n, \alpha'_{n+1}]$ and $q'_n/p'_n = [a'_1; a'_2 \dots, a'_n]$).

As an example, suppose that $d = 34$. Then

$$\sqrt{34} = [5; 1, 4, 1, 10, 1, 4, 10, 1, \dots],$$

and hence

n	a_n	p_n	q_n	$p_n^2 - 34q_n^2$
-2		0	1	-34
-1		1	0	1
0	5	5	1	-9
1	1	6	1	2
2	4	29	5	-9
3	1	35	6	1
4	10	379	65	-9

Since the fundamental solution $(35, 6)$ of $x^2 - 34y^2 = 1$ is encountered without finding a solution of $x^2 - 34y^2 = -1$, it follows that this latter equation has no integral solution. It is interesting to note, however, that the congruence

$$x^2 - 34y^2 \equiv -1 \pmod{m}$$

has a solution modulo m for every m. This can be seen simply by choosing m_1, m_2 so that $m_1 m_2 = m$, $(m_1, m_2) = 1$, $3 \nmid m_1$, $5 \nmid m_2$ and then choosing n so that $3n \equiv 1 \pmod{m_1}$. We observe that $5^2 - 34 = -3^2$. Thus the congruence $x^2 - 34y^2 \equiv -1 \pmod{m_1}$ has the solution $x = 5n$, $y = n$. Similarly $3^2 - 34 = -5^2$, so that if we choose p so that $5p \equiv 1 \pmod{m_2}$, then $x^2 - 34y^2 \equiv -1 \pmod{m_2}$ has the solution $x = 3p$, $y = p$. Since the congruence has solutions $\pmod{m_j}$, $j = 1, 2$ and $(m_1, m_2) = 1$. it follows by the Chinese Remainder Theorem that the congruence has a solution $\pmod{m}$.

Since the equation $x^2 - 34y^2$ also has real solutions, it follows that the equation is everywhere locally soluble, but not globally soluble. This does not contradict the Hasse-Minkowski principle, which asserts less. In the weak form it asserts that $x^2 - 34y^2 + z^2 = 0$ has a non-trivial integral solution, and in the strong form it asserts that $x^2 - 34y^2 = -1$ has a rational solution.

In order to make more precise statements about the solutions to Pell's equation it is useful to have the following theorem.

Theorem 2.13. *The continued fraction expansion of a real quadratic irrational number α is purely periodic if and only if $\alpha > 1$ and $-1 < \alpha' < 0$ where α' denotes the algebraic conjugate of α.*

Proof. To be more precise, by the algebraic conjugate we mean that if α is a root of $ax^2 + bx + c$ where a, b, c are integers with $a > 0$ and $b^2 - 4ac$ positive and not a perfect square, then α' is the other root. Suppose first of all that $\alpha > 1$ and $-1 < \alpha' < 0$. We show that the α_n', the algebraic conjugate of α_n' also satisifies $-1 < \alpha_n' < 0$. Since $\alpha_n \sim \alpha$, α_n is also a quadratic irrational. We have $\alpha_n = a_n + 1/\alpha_{n+1}$ and it follows by considering the quadratic equation satisfied by α_n and the consequential one satisfied by α_n, that $\alpha_n' = a_n + 1/\alpha_{n+1}'$. Another way to view these relationships is that they are all linear functions of of $\sqrt{b^2 - 4ac}$ with rational coefficients and the conjugates are obtained simply by changing the sign of $\sqrt{b^2 - 4ac}$. Now the inequality $-1 < \alpha_n' < 0$ implies that $1/\alpha_{n+1}' < -a_n \le -1$ and so $-1 < \alpha_{n+1}' < 0$. Hence, by induction $-1 < \alpha_n' < 0$ for every n.

Now $a_n = [-1/\alpha_{n+1}']$ for every n, and since α is a quadratic irrational we have $\alpha_m = \alpha_n$ for some pair m, n with $m < n$. Moreover we can suppose that m is minimal. Thus if $m > 0$, then we have $1/\alpha_m' = 1/\alpha_n'$, and so $a_{m-1} = a_{n-1}$. It follows that $\alpha_{m-1} = \alpha_{n-1}$ contradiciting the minimality of m, and so $m = 0$ as required.

Conversely, if α is purely periodic, then $\alpha > a_0 > 1$. Further, for some $n \ge 1$ we have

$$\alpha = \frac{\alpha p_n + p_{n-1}}{\alpha q_n + q_{n-1}}$$

and so α is a root of $q_n x^2 + (q_{n-1} - p_n)x - p_{n-1} = 0$. The function of x on the left here is $-p_{n-1}$ when $x = 0$ and has the value $p_n + q_n - (p_{n-1} + q_{n-1}) > 0$ when $x = -1$, so its other root lies between -1 and 0.

We have already seen in Theorem 2.12 that when d is a positive integer not a perfect square, then every solution to $x^2 - dy^2 = 1$ is a convergent to the continued fraction for $\sqrt{d}$. This continued fraction is not itself purely periodic, but it almost is. In fact $\alpha = \sqrt{d} + [\sqrt{d}]$ satisfies $\alpha > 1$ and $-1 < \alpha' < 0$ and so has a purely periodic expansion. Moreover, for the latter expansion we have $a_0 = 2[\sqrt{d}]$. Let m be the

length of the period. Then $\sqrt{d} = [[\sqrt{d}]; a_1, a_2, \dots, a_{m-1}, 2[\sqrt{d}], \dots]$. Suppose that $p_n^2 - dq_n^2 = 1$ where p_n/q_n, as usual, denote convergents to the continued fraction. We can now say something quite precise about n. Obviously $p_n - q_n\sqrt{d} > 0$. But $q_n\sqrt{d} - p_n = (-1)^n/(\alpha_{n+1}q_n + q_{n-1})$. Thus n is odd. Moreover

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

so that $(p_n - q_n\sqrt{d})\alpha_{n+1} = q_{n-1}\sqrt{d} - p_{n-1}$. Thus

$$(p_n^2 - dq_n^2)\alpha_{n+1} = (q_{n-1}\sqrt{-}p_{n-1})(q_n\sqrt{d} + p_n) = (-1)^{n-1}\sqrt{d} + c$$

where $c \in \mathbb{Z}$. But $p_n^2 - dq_n^2 = 1$ and n is odd. Hence $\alpha_{n+1} = \sqrt{d} + c$. Now $\sqrt{d} = a_0 + 1/\alpha_1$ and so α_1 is purely periodic. Thus $\alpha_1 > 1$, $\alpha_{n+2} > 1$, $a_{n+1} + 1/\alpha_{n+2} = \alpha_{n+1} = a_0 + c + 1/\alpha_1$, whence $a_{n+1} = a_0 + c$ and $\alpha_1 = \alpha_{n+2}$. It follows that $n+1$ is divisible by m. Thus n is of the form $lm - 1$ where $l = 1, 2, 3, \dots$ when m is even and $l = 2, 4, 6, \dots$ when m is odd. We complete our study by showing that every n of this form gives a solution. In view of the periodicity m the continued fraction of $\sqrt{d}$ we have $\alpha_1 = \alpha_{n+2}$ for every n of the above form. Thus

$$\sqrt{d} = \frac{\alpha_1 p_{n+1} + p_n}{\alpha_1 q_{n+1} + q_n}$$

and moreover $\sqrt{d} = a_0 + 1/\alpha_1$. Substituting for α_1 and using the fact that $\sqrt{d}$ is irrational we see that $p_n = q_{n+1} - a_0 q_n$ and $p_{n+1} - a_0 p_n = q_n d$. On eliminating a_0 we obtain $p_n^2 - dq_n^2 = p_n q_{n+1} - p_{n+1}q_n = 1$ since n is odd.

A similar analysis of the equation $x^2 - dy^2 = -1$ shows that it is insoluble when the period m is even, but when m is odd there are solutions and they are given by $x = p_n$, $y = q_n$ where $n = lm - 1$ with $l = 1, 3, 5, \dots$.