CONTINUED FRACTIONS AND PELL'S EQUATION

ETHAN MADDOX

ABSTRACT. Though the Euclidean algorithm has been known for at least 2,000 years, it was not until the late 16th century that mathematicians began the development of the theory of continued fractions, initially for the purpose of approximating the roots of quadratic equations. In the 17th century, English mathematician John Wallis introduced the term "continued fraction" to describe these objects, and it has been used ever since. In this paper, we develop the theory of continued fractions, examine its usage in Diophantine approximation, and apply these results to find solutions for Pell's equation.

1. INTRODUCTION

An expression of the form

$$b_0 + \cfrac{a_0}{b_1 + \cfrac{a_1}{b_2 + \cfrac{a_2}{b_3 + \cfrac{a_3}{\ddots}}}},$$

which may be either finite or infinite in length, is called a *continued fraction* [3]. Unless stated otherwise, we will assume that $a_i = 1$ for all i and that b_i is an integer that is positive for $i \geq 1$. Continued fractions with these properties are called *simple* [3].

The story of continued fractions has its origins with the Euclidean algorithm [2]. The earliest extant description of the Euclidean algorithm appears in Euclid's *Elements*, though it likely predates him [2]. Being that the Euclidean algorithm is just a few steps away from being the algorithm for generating a simple continued fraction, it is somewhat surprising that their explicit use—at least of the non-ascending kind—did not begin until the late 16th century [2].

In 1579, Italian mathematician Rafael Bombelli published the second edition of his treatise on algebra, in which he described an algorithm for computing $\sqrt{13}$ [2]. With modern notation, Bombelli's algorithm can be summarized by the equation

$$\sqrt{13} = 3 + \cfrac{4}{6 + \cfrac{4}{6 + \cfrac{4}{\ddots}}}.$$

Pietro Cataldi, another Italian mathematician, went even further, introducing the first formal notation for continued fractions, proving some of their basic properties, and showing an algorithm for generating convergents [2]. It wasn't until 1655, however, that the term "continued fraction" was first used by the English mathematician John Wallis in his *Arithmetica infinitorum* [2].

In Section 2 of this paper, we will develop the theory of simple continued fractions by defining them and deriving their key properties. In Section 3, we will apply this theory to Diophantine approximation and obtain some results on periodic continued fractions. Finally, in Section 4, we will show how continued fractions can be used to find solutions for Pell's equation. Most of the results and proofs of the following sections are based on [4], restated and reorganized so as to reflect my own understanding. Exceptions to this will be cited explicitly.

2. Continued Fractions

2.1. **Finite Simple Continued Fractions.** Consider a rational number t_0/t_1 in lowest terms and with $t_1 > 0$. Using the Euclidean algorithm, we can write

$$\begin{aligned} t_0 &= t_1 a_0 + t_2, & 0 < t_2 < t_1,\\ t_1 &= t_2 a_1 + t_3, & 0 < t_3 < t_2,\\ &\vdots & \vdots\\ t_{k-1} &= t_k a_{k-1} + t_{k+1}, & 0 < t_{k+1} < t_k,\\ t_k &= t_{k+1} a_k. \end{aligned}$$

By the inequalities above, the t_i are non-negative and decreasing for $i \geq 1$. Thus the process generating these equations will terminate because eventually we will get a remainder of zero. For all integers i with $0 \leq i \leq k$, let $\gamma_i = t_i/t_{i+1}$. We can now write the first k equations above as

$$\gamma_i = a_i + \frac{1}{\gamma_{i+1}}$$

and the last as $\gamma_k = a_k$. This allows us to express t_0/t_1 as a continued fraction. We have

$$\begin{aligned} \frac{t_0}{t_1} = \gamma_0 &= a_0 + \frac{1}{\gamma_1}\\ &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\gamma_2}}\\ &\;\vdots\\ &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_k}}}}. \end{aligned}$$

The above is a continued fraction expansion of γ_0. The integers a_i are the *partial quotients* of the simple continued fraction. In order to make our notation more compact, we make the convention that

$$[x_0; x_1, \dots, x_k] = x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots + \cfrac{1}{x_k}}}}.$$

Example 1. Consider the irreducible fraction 23/17. We have

$$\begin{aligned} 23 &= 17 \cdot 1 + 6, \\ 17 &= 6 \cdot 2 + 5, \\ 6 &= 5 \cdot 1 + 1, \\ 5 &= 1 \cdot 5. \end{aligned}$$

We can write the first three equations as

$$\begin{aligned} \frac{23}{17} &= 1 + \frac{6}{17}, \\ \frac{17}{6} &= 2 + \frac{5}{6}, \\ \frac{6}{5} &= 1 + \frac{1}{5}. \end{aligned}$$

But then

$$\begin{aligned} \frac{23}{17} &= 1 + \frac{6}{17} \\ &= 1 + \cfrac{1}{2 + \cfrac{5}{6}} \\ &= 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{5}}} \\ &= [1; 2, 1, 5]. \end{aligned}$$

Remark 1. There are several observations we can quickly make about the continued fractions obtained in this way. One is that, although t_0 may be any integer, $t_1, \dots, t_k$ are positive integers by the inequalities $0 < t_{i+1} < t_i$ implied by the Euclidean algorithm. This then implies that $a_i \geq 1$ for $i \geq 1$. It is also useful to note the identities

$$[x_0; x_1, \dots, x_k] = x_0 + \frac{1}{[x_1; x_2, \dots, x_k]}$$

and

$$[x_0; x_1, \dots, x_k] = \left[x_0; x_1, \dots, x_{k-2}, x_{k-1} + \frac{1}{x_k}\right].$$

A natural question to ask is whether finite simple continued fractions are unique, i.e., is there a unique finite simple continued fraction for each rational number? As it turns out, there are at least two such representations for a given rational number, as we have

$$[a_0; a_1, \dots, a_k] = [a_0; a_1, \dots, a_{k-1}, a_k - 1, 1].$$

Fortunately, these are the only possibilities.

Theorem 2.1. *Let $[a_0; a_1, \dots, a_m]$ and $[b_0; b_1, \dots, b_n]$ be finite simple continued fractions. If $a_m, b_n > 1$ and $[a_0; a_1, \dots, a_m] = [b_0; b_1, \dots, b_n]$, then $m = n$ and $a_i = b_i$ for $0 \leq i \leq m = n$.*

Proof. Let $\gamma_i = [a_i; a_{i+1}, \dots, a_m]$ and $\delta_j = [b_j; b_{j+1}, \dots, b_n]$. We note that

$$\gamma_i = [a_i; a_{i+1}, \dots, a_m] = a_i + \frac{1}{[a_{i+1}; a_{i+2}, \dots, a_m]} = a_i + \frac{1}{\gamma_{i+1}},$$

as above. Thus $\gamma_i > a_i \geq 1$ for $i \in \{1, \dots, m-1\}$ and $\gamma_m = a_m > 1$, so $a_i = \lfloor \gamma_i \rfloor$ for all $i \in \{0, \dots, m\}$. Similarly, we have $b_j = \lfloor \delta_j \rfloor$ for all $j \in \{0, \dots, n\}$. By hypothesis, we have $\gamma_0 = \delta_0$, so $a_0 = \lfloor \gamma_0 \rfloor = \lfloor \delta_0 \rfloor = b_0$.

Suppose inductively that $\gamma_i = \delta_i$ and $a_i = b_i$. Then we have

$$\frac{1}{\gamma_{i+1}} = \gamma_i - a_i = \delta_i - b_i = \frac{1}{\delta_{i+1}},$$

so $\gamma_{i+1} = \delta_{i+1}$ and $a_{i+1} = \lfloor \gamma_{i+1} \rfloor = \lfloor \delta_{i+1} \rfloor = b_{i+1}$. Suppose $m < n$. Then $\gamma_m = \delta_m$ and $a_m = b_m$, but $\gamma_m = a_m$ and $\delta_m > b_m$, a contradiction. An analogous contradiction arises if we assume $n < m$, so we are done. $\square$

Corollary 2.1. *Every finite simple continued fraction is rational, and every rational number can be written as a finite simple continued fraction in exactly two ways.*

2.2. **Infinite Simple Continued Fractions.** As previously mentioned, continued fractions can be infinite in length, a concept we will make unambiguous with the following results. Let $\{a_i\}_{i=0}^{\infty}$ be a sequence of integers such that $a_i > 0$ for all positive integers i. We define $\{h_i\}_{i=-2}^{\infty}$ and $\{k_i\}_{i=-2}^{\infty}$ recursively as follows:

$$h_{-2} = 0, \quad h_{-1} = 1, \quad h_i = a_i h_{i-1} + h_{i-2} \text{ for all } i \geq 0,$$

and

$$k_{-2} = 1, \quad k_{-1} = 0, \quad k_i = a_i k_{i-1} + k_{i-2} \text{ for all } i \geq 0.$$

Then $k_0 = a_0 k_{-1} + k_{-2} = 1$ and $k_1 = a_1 k_0 + k_{-1} = a_1 \geq 1$. Since the a_i are positive for all positive i, our recursive definition of k_i implies that the k_i are positive for $i \geq 0$ and increasing for $i \geq 1$.

Theorem 2.2. *For any positive $x \in \mathbb{R}$, we have*

$$[a_0; a_1, \dots, a_{n-1}, x] = \frac{x h_{n-1} + h_{n-2}}{x k_{n-1} + k_{n-2}}.$$

Proof. If $n = 0$, we have

$$\frac{x h_{-1} + h_{-2}}{x k_{-1} + k_{-2}} = \frac{x}{1} = x = [x].$$

If $n = 1$, we have

$$\begin{aligned}\frac{xh_0 + h_{-1}}{xk_0 + k_{-1}} &= \frac{x(a_0h_{-1} + h_{-2}) + h_{-1}}{x(a_0k_{-1} + k_{-2}) + k_{-1}} \\ &= \frac{xa_0 + 1}{x} \\ &= a_0 + \frac{1}{x} \\ &= [a_0; x].\end{aligned}$$

Suppose inductively that

$$[a_0; a_1, \dots, a_{n-1}, x] = \frac{xh_{n-1} + h_{n-2}}{xk_{n-1} + k_{n-2}}.$$

Then we have

$$\begin{aligned}[a_0; a_1, \dots, a_n, x] &= \left[a_0; a_1, \dots, a_{n-1}, a_n + \frac{1}{x}\right] \\ &= \frac{(a_n + 1/x)h_{n-1} + h_{n-2}}{(a_n + 1/x)k_{n-1} + k_{n-2}} \\ &= \frac{a_nh_{n-1} + h_{n-1}/x + h_{n-2}}{a_nk_{n-1} + k_{n-1}/x + k_{n-2}} \\ &= \frac{x(a_nh_{n-1} + h_{n-2}) + h_{n-1}}{x(a_nk_{n-1} + k_{n-2}) + k_{n-1}} \\ &= \frac{xh_n + h_{n-1}}{xk_n + k_{n-1}}.\end{aligned}$$

□

Corollary 2.2. *Defining $p_n = [a_0; a_1, \dots, a_n]$ for all non-negative integers n, we have*

$$p_n = \frac{h_n}{k_n}.$$

Proof. By Theorem 2.2, we have

$$[a_0; a_1, \dots, a_{n-1}, a_n] = \frac{a_nh_{n-1} + h_{n-2}}{a_nk_{n-1} + k_{n-2}} = \frac{h_n}{k_n}.$$

□

Example 2. Again consider the fraction $23/17 = [1; 2, 1, 5]$ from Example 1. We have

$$\begin{aligned} h_0 &= 1 \cdot 1 + 0 = 1, & k_0 &= 1 \cdot 0 + 1 = 1, \\ h_1 &= 2 \cdot 1 + 1 = 3, & k_1 &= 2 \cdot 1 + 0 = 2, \\ h_2 &= 1 \cdot 3 + 1 = 4, & k_2 &= 1 \cdot 2 + 1 = 3, \\ h_3 &= 5 \cdot 4 + 3 = 23, & k_3 &= 5 \cdot 3 + 2 = 17,\end{aligned}$$

so the convergents of $23/17$ are

$$\frac{1}{1}, \quad \frac{3}{2}, \quad \frac{4}{3}, \quad \text{and} \quad \frac{23}{17}.$$

Theorem 2.3. *For all integers i such that $i \geq 1$, we have*

$$\begin{aligned} h_i k_{i-1} - h_{i-1} k_i &= (-1)^{i-1}, \\ p_i - p_{i-1} &= \frac{(-1)^{i-1}}{k_i k_{i-1}}, \\ h_i k_{i-2} - h_{i-2} k_i &= (-1)^i a_i, \\ p_{i+1} - p_{i-1} &= \frac{(-1)^{i+1} a_{i+1}}{k_{i+1} k_{i-1}}. \end{aligned}$$

It follows that the fractions h_i/k_i are in lowest terms.

Proof. If $i = -1$, we have

$$h_{-1} k_{-2} - h_{-2} k_{-1} = 1 \cdot 1 - 0 \cdot 0 = 1 = (-1)^{-2}.$$

Suppose inductively that $h_i k_{i-1} - h_{i-1} k_i = (-1)^{i-1}$. We have

$$\begin{aligned} h_{i+1} k_i - h_i k_{i+1} &= (a_{i+1} h_i + h_{i-1}) k_i - h_i (a_{i+1} k_i + k_{i-1}) \\ &= a_{i+1} h_i k_i + h_{i-1} k_i - a_{i+1} h_i k_i - h_i k_{i-1} \\ &= h_{i-1} k_i - h_i k_{i-1} \\ &= -(h_i k_{i-1} - h_{i-1} k_i) \\ &= -(-1)^{i-1} \\ &= (-1)^i. \end{aligned}$$

This proves the first identity. We then have

$$\begin{aligned} h_i k_{i-1} - h_{i-1} k_i &= (-1)^{i-1} \\ \frac{h_i}{k_i} - \frac{h_{i-1}}{k_{i-1}} &= \frac{(-1)^{i-1}}{k_{i-1} k_i} \\ p_i - p_{i-1} &= \frac{(-1)^{i-1}}{k_{i-1} k_i}, \end{aligned}$$

which proves the second identity. To prove the third identity, we have

$$\begin{aligned} h_i k_{i-2} - h_{i-2} k_i &= (a_i h_{i-1} + h_{i-2}) k_{i-2} - h_{i-2} (a_i k_{i-1} + k_{i-2}) \\ &= a_i h_{i-1} k_{i-2} + h_{i-2} k_{i-2} - a_i h_{i-2} k_{i-1} - h_{i-2} k_{i-2} \\ &= a_i h_{i-1} k_{i-2} - a_i h_{i-2} k_{i-1} \\ &= a_i (h_{i-1} k_{i-2} - h_{i-2} k_{i-1}) \\ &= (-1)^{i-2} a_i \\ &= (-1)^i a_i. \end{aligned}$$

We then see that

$$\begin{aligned} h_{i+1} k_{i-1} - h_{i-1} k_{i+1} &= (-1)^{i+1} a_{i+1} \\ \frac{h_{i+1}}{k_{i+1}} - \frac{h_{i-1}}{k_{i-1}} &= \frac{(-1)^{i+1} a_{i+1}}{k_{i+1} k_{i-1}} \\ p_{i+1} - p_{i-1} &= \frac{(-1)^{i+1} a_{i+1}}{k_{i+1} k_{i-1}}. \end{aligned}$$

Finally, we can see that the first identity shows that $\gcd(h_i, k_i) = 1$, as any factor of both must necessarily divide $(-1)^{i-1}$. $\square$

Theorem 2.4. *The p_n with even index form an increasing sequence, the p_n with odd index form a decreasing sequence, and every p_n of even index is less than every p_n of odd index. In symbols, we have*

$$p_0 < p_2 < p_4 < \cdots < p_5 < p_3 < p_1.$$

Furthermore, the limit

$$\lim_{n\to\infty} p_n$$

exists.

Proof. By Theorem 2.3, we have, for all $i \geq 0$,

$$p_{2i+2} - p_{2i} = \frac{(-1)^{2i+2} a_{2i+2}}{k_{2i+2} k_{2i}} = \frac{a_{2i+2}}{k_{2i+2} k_{2i}} > 0$$

because the k_i are positive for $i \geq 0$ and the a_i are positive for $i \geq 1$. Thus $p_{2i+2} > p_{2i}$ for $i \geq 0$, so the p_i of even index form an increasing sequence.

Using the same identity, for $i \geq 0$ we have

$$p_{2i+3} - p_{2i+1} = \frac{(-1)^{2i+3} a_{2i+3}}{k_{2i+3} k_{2i+1}} = -\frac{a_{2i+3}}{k_{2i+3} k_{2i+1}} < 0$$

for the same reasons as before. Thus $p_{2i+3} < p_{2i+1}$ for $i \geq 0$, so the p_i of odd index form a decreasing sequence.

By Theorem 2.3 again, for all $i \geq 0$ we have

$$p_{2i+1} - p_{2i} = \frac{(-1)^{2i}}{k_{2i+1} k_{2i}} = \frac{1}{k_{2i+1} k_{2i}} > 0$$

because the k_i are positive for $i \geq 0$. Thus $p_{2i+1} > p_{2i}$ for all $i \geq 0$.

Hence, for all $i, j \geq 0$, we have

$$p_{2i} \leq p_{2i+2j} < p_{2i+2j+1} \leq p_{2j+1},$$

so $p_{2i} < p_{2j+1}$ for all $i, j \geq 0$.

Since the p_n of even index form an increasing sequence bounded above by p_1, it has a least upper bound. Similarly, the p_n of odd index form a decreasing sequence bounded below by p_0, so they have a greatest lower bound. Since

$$\lim_{i\to\infty} (p_i - p_{i-1}) = \lim_{i\to\infty} \frac{(-1)^{i-1}}{k_i k_{i-1}} = 0,$$

these bounds must be equal, so the limit

$$\lim_{n\to\infty} p_n$$

exists. □

The preceding results allow us to make the following definition unambiguously.

Definition 2.1. An infinite sequence of integers $\{a_i\}_{i=0}^{\infty}$ with $a_i > 0$ for $i > 0$ can be formed into an infinite simple continued fraction $[a_0; a_1, a_2, \dots,]$. We define

$$[a_0; a_1, a_2, \dots] = \lim_{n\to\infty} [a_0; a_1, \dots, a_n],$$

with the number $[a_0; a_1, \dots, a_n]$ being called the nth convergent of $[a_0; a_1, a_2, \dots]$.

We saw in the previous section that the simple continued fractions of finite length correspond exactly to the rational numbers. In an example of the elegance of continued fractions, it turns out that the infinite simple continued fractions correspond exactly to the irrational numbers. The next theorem is the first direction of this statement.

Theorem 2.5. *Every infinite simple continued fraction is irrational.*

Proof. Let $\xi = [a_0; a_1, a_2, \dots]$ be an infinite simple continued fraction. Suppose ξ is rational, so that $\xi = p/q$ for some integers p and q with $\gcd(p, q) = 1$ and $q > 0$. Let p_n denote the nth convergent of ξ. By Theorem 2.4, ξ is strictly between p_n and p_{n+1}, so we have

$$\begin{aligned}0 &< |\xi - p_n| < |p_{n+1} - p_n| \\ 0 &< \left|\frac{p}{q} - \frac{h_n}{k_n}\right| < \left|\frac{(-1)^n}{k_{n+1}k_n}\right| \\ 0 &< |k_n p - h_n q| < \frac{q}{k_{n+1}}.\end{aligned}$$

The k_n are increasing in n, so we can choose n such that $q < k_{n+1}$. But then the integer $|k_n p - h_n q|$ would lie strictly between 0 and 1, a contradiction. □

Remark 2. We now know that any finite simple continued fraction is rational, and that any infinite simple continued fraction is irrational. In the case of finite simple continued fractions, we saw that every rational number has exactly two representations. It is then natural to ask whether two distinct infinite simple continued fractions can have the same value.

Lemma 2.1. *Let $\xi_0 = [a_0; a_1, a_2, \dots]$ be a simple continued fraction. Then $a_0 = \lfloor \xi_0 \rfloor$ and—letting $\xi_1 = [a_1; a_2, a_3, \dots]$—we have $\xi_0 = a_0 + 1/\xi_1$.*

Proof. By Theorem 2.4, $p_0 < \xi_0 < p_1$, i.e., $a_0 < \xi_0 < a_0 + 1/a_1$. Since $a_1 \geq 0$, we have $a_0 < \xi_0 < a_0 + 1$, so $\lfloor \xi_0 \rfloor = a_0$. Finally, we have

$$\begin{aligned}\xi_0 &= \lim_{n\to\infty} [a_0; a_1, \dots, a_n] \\ &= \lim_{n\to\infty} \left(a_0 + \frac{1}{[a_1; a_2, \dots, a_n]}\right) \\ &= a_0 + \frac{1}{\lim\limits_{n\to\infty} [a_1; a_2, \dots, a_n]} \\ &= a_0 + \frac{1}{\xi_1}.\end{aligned}$$

□

Theorem 2.6. *Distinct infinite simple continued fractions have distinct values.*

Proof. Let $\xi = [a_0; a_1, a_2, \dots] = [b_0; b_1, b_2, \dots]$ and let $\xi_i = [a_i; a_{i+1}, a_{i+2}, \dots]$ for all $i \geq 1$. By Theorem 2.1, we have $\lfloor \xi \rfloor = a_0 = b_0$. Suppose inductively that $a_i = b_i$ for integers $i \in [0, k]$. Note that

$$[a_0; a_1, a_2, \dots] = [a_0; a_1, \dots, a_k, [a_{k+1}; a_{k+2}, \dots]]$$

and

$$[b_0; b_1, b_2, \dots] = [b_0; b_1, \dots, b_k, [b_{k+1}; b_{k+2}, \dots]].$$

Since these are both equal to ξ and have the same first $k+1$ convergents, we must have

$$[a_{k+1}; a_{k+2}, a_{k+3}, \dots] = [b_{k+1}; b_{k+2}, b_{k+3}, \dots].$$

But then, by Theorem 2.1, $a_{k+1} = b_{k+1}$. Hence $a_i = b_i$ for all non-negative integers i by mathematical induction. $\square$

We have seen that every infinite simple continued fraction is an irrational number. We now consider the converse: whether every irrational number can be represented as an infinite simple continued fraction.

Let ξ be an irrational number. Letting $\xi_0 = \xi$ and $a_0 = \lfloor \xi_0 \rfloor$, we define

$$\xi_{i+1} = \frac{1}{\xi_i - a_i} \quad \text{and} \quad a_{i+1} = \lfloor \xi_{i+1} \rfloor. \tag{1}$$

The a_i are integers by definition. $\xi_0 = \xi$ is irrational by definition, and the recursive formula

$$\xi_{i+1} = \frac{1}{\xi_i - a_i}$$

shows by induction that the ξ_i are all irrational. We then have

$$\begin{gathered}
\lfloor \xi_{i-1} \rfloor < \xi_{i-1} < \lfloor \xi_{i-1} \rfloor + 1 \\
a_{i-1} < \xi_{i-1} < a_{i-1} + 1 \\
0 < \xi_{i-1} - a_{i-1} < 1 \\
1 < \frac{1}{\xi_{i-1} - a_{i-1}} = \xi_i,
\end{gathered}$$

so $a_i = \lfloor \xi_i \rfloor \geq 1$ for $i \geq 1$. Next we have

$$\xi = \xi_0 = a_0 + \frac{1}{\xi_1} = [a_0; \xi_1].$$

If $\xi = [a_0; a_1, \dots, a_{n-1}, \xi_n]$, then we have

$$\xi = [a_0; a_1, \dots, a_{n-1}, \xi_n] = \left[a_0; a_1, \dots, a_{n-1}, a_n + \frac{1}{\xi_{n+1}}\right] = [a_0; a_1, \dots, a_n, \xi_{n+1}].$$

Hence by induction we have

$$\xi = [a_0; a_1, \dots, a_{n-1}, \xi_n]$$

for all integers $n \geq 0$.

By Theorem 2.2, we now have

$$\xi = [a_0; a_1, \dots, a_{n-1}, \xi_n] = \frac{\xi_n h_{n-1} + h_{n-2}}{\xi_n k_{n-1} + k_{n-2}}.$$

This then implies

$$\begin{aligned}
\xi - \frac{h_{n-1}}{k_{n-1}} &= \frac{\xi_n h_{n-1} + h_{n-2}}{\xi_n k_{n-1} + k_{n-2}} - \frac{h_{n-1}}{k_{n-1}} \\
&= \frac{k_{n-1}(\xi_n h_{n-1} + h_{n-2})}{k_{n-1}(\xi_n k_{n-1} + k_{n-2})} - \frac{h_{n-1}(\xi_n k_{n-1} + k_{n-2})}{k_{n-1}(\xi_n k_{n-1} + k_{n-2})} \\
&= \frac{\xi_n h_{n-1} k_{n-1} + h_{n-2} k_{n-1} - \xi_n h_{n-1} k_{n-1} - h_{n-1} k_{n-2}}{k_{n-1}(\xi_n k_{n-1} + k_{n-2})} \\
&= \frac{-(h_{n-1} k_{n-2} - h_{n-2} k_{n-1})}{k_{n-1}(\xi_n k_{n-1} + k_{n-2})}
\end{aligned}$$

$$\xi - \frac{h_{n-1}}{k_{n-1}} = \frac{(-1)^{n-1}}{k_{n-1}(\xi_n k_{n-1} + k_{n-2})}. \tag{2}$$

Since $k_n \to \infty$ as $i \to \infty$ and the ξ_n are positive for $n \geq 1$, the above fraction tends to 0 as $n \to \infty$. Hence we have

$$\xi = \lim_{n\to\infty} \frac{h_n}{k_n} = \lim_{n\to\infty} [a_0; a_1, \dots, a_n] = [a_0; a_1, a_2, \dots].$$

We can summarize the preceding discussion in the following theorem.

Theorem 2.7. *Every irrational number ξ has a unique representation as an infinite simple continued fraction. Letting $\xi_0 = \xi$ and using the recursive definitions of Equation (1), we find that $\xi = [a_0; a_1, a_2, \dots]$*

Example 3. Consider computing the continued fraction expansion of π. We have

$$\begin{aligned}
\xi_0 &= \pi, & a_0 &= \lfloor \pi \rfloor = 3, \\
\xi_1 &= \frac{1}{\pi - 3} = 7.062\dots, & a_1 &= \lfloor 7.062\dots \rfloor = 7, \\
\xi_2 &= \frac{1}{7.062\dots - 7} = 15.996\dots, & a_2 &= \lfloor 15.996\dots \rfloor = 15.
\end{aligned}$$

Thus the continued fraction expansion of π begins $[3; 7, 15, \dots]$ and the first three convergents of π are

$$3, \quad \frac{22}{7}, \quad \text{and} \quad \frac{333}{106}.$$

Before moving on, we quickly note the following result, which can easily be proved by induction.

Theorem 2.8. *Let x be a real number greater than 1. The ith convergent of $1/x$ is the reciprocal of the $(i-1)$th convergent of x.*

3. Approximating Irrational Numbers

We now show that that the convergents to an irrational number form an "efficient" sequence of rational approximations in the sense that the errors are small compared to the size of the denominators.

Theorem 3.1. *For all integers $n \geq 0$, we have*

$$\left| \xi - \frac{h_n}{k_n} \right| < \frac{1}{k_n k_{n+1}}$$

and

$$|\xi k_n - h_n| < \frac{1}{k_{n+1}}.$$

Proof. By Equation (2), we have

$$\begin{aligned}\left|\xi - \frac{h_n}{k_n}\right| &= \left|\frac{(-1)^n}{k_n(\xi_{n+1}k_n + k_{n-1})}\right| \\ &= \frac{1}{k_n(\xi_{n+1}k_n + k_{n-1})} \\ &< \frac{1}{k_n(a_{n+1}k_n + k_{n-1})} \\ &= \frac{1}{k_n k_{n+1}}.\end{aligned}$$

Hence

$$\left|\xi - \frac{h_n}{k_n}\right| < \frac{1}{k_n k_{n+1}},$$

which immediately implies

$$|\xi k_n - h_n| < \frac{1}{k_{n+1}}$$

via multiplication by k_n. □

Since $k_n \leq k_{n+1}$ for all $n \geq 0$, we have the weaker inequality

$$\left|\xi - \frac{h_n}{k_n}\right| < \frac{1}{k_n^2}.$$

Since the h_n/k_n are infinitely many rationals satisfying the above equation, we have the following result.

Corollary 3.1. *For every irrational number ξ, there exists infinitely many rationals p/q with $q > 0$ such that*

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Theorem 3.2. *For all integers $n \geq 0$, we have*

$$|\xi k_{n+1} - h_{n+1}| < |\xi k_n - h_n|$$

and

$$\left|\xi - \frac{h_{n+1}}{k_{n+1}}\right| < \left|\xi - \frac{h_n}{k_n}\right|.$$

Proof. Since the ξ_n are irrational, $\xi_n < \lfloor \xi_n \rfloor + 1 = a_n + 1$, so

$$\begin{aligned}\xi_{n+1}k_n + k_{n-1} &< (a_{n+1} + 1)k_n + k_{n-1} \\ &= a_{n+1}k_n + k_{n-1} + k_n \\ &= k_{n+1} + k_n \\ &\leq a_{n+2}k_{n+1} + k_n \\ &= k_{n+2}.\end{aligned}$$

We now have, by Equation (2),

$$\left|\xi - \frac{h_n}{k_n}\right| = \frac{1}{k_n(\xi_{n+1}k_n + k_{n-1})} > \frac{1}{k_n k_{n+2}}.$$

Multiplying by k_n and using Theorem 3.1, we have

$$|\xi k_n - h_n| > \frac{1}{k_{n+2}} > |\xi k_{n+1} - h_{n+1}|,$$

which proves the first inequality. We then have

$$\begin{aligned}\left|\xi - \frac{h_{n+1}}{k_{n+1}}\right| &= \frac{1}{k_{n+1}}|\xi k_{n+1} - h_{n+1}| \\ &< \frac{1}{k_{n+1}}|\xi k_n - h_n| \\ &\leq \frac{1}{k_n}|\xi k_n - h_n| \\ &= \left|\xi - \frac{h_n}{k_n}\right|,\end{aligned}$$

which proves the second inequality. □

Theorem 3.3. *Let p/q be a rational number with $q > 0$. If $|\xi q - p| < |\xi k_n - h_n|$ for some $n \geq 0$, then $q \geq k_{n+1}$. Furthermore, if $|\xi - p/q| < |\xi - h_n/k_n|$ for some $n \geq 1$, then $q > k_n$.*

Proof. Suppose $|\xi q - p| < |\xi k_n - h_n|$ for some $n \geq 0$ and $q < k_{n+1}$. Consider the linear system

$$\begin{Bmatrix} xh_n + yh_{n+1} = p, \\ xk_n + yk_{n+1} = q. \end{Bmatrix}$$

Noting that

$$\det \begin{bmatrix} h_n & h_{n+1} \\ k_n & k_{n+1} \end{bmatrix} = h_n k_{n+1} - h_{n+1} k_n = (-1)^{n+1}$$

by Theorem 2.3, Cramer's Rule implies that our system has the unique solution

$$\begin{aligned}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} h_n & h_{n+1} \\ k_n & k_{n+1} \end{bmatrix}^{-1} \begin{bmatrix} p \\ q \end{bmatrix} \\ &= \frac{1}{h_n k_{n+1} - h_{n+1} k_n} \begin{bmatrix} k_{n+1} & -h_{n+1} \\ -k_n & h_n \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} \\ &= (-1)^{n+1} \begin{bmatrix} k_{n+1} & -h_{n+1} \\ -k_n & h_n \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix},\end{aligned}$$

which will be an integer solution. We claim that x and y are non-zero. If $x = 0$, then $q = yk_{n+1}$, so $y \geq 1$ and $q \geq k_{n+1}$, contradicting our assumption that $q < k_{n+1}$. If $y = 0$, then $p = xh_n$, $q = xk_n$, and

$$|\xi q - p| = |\xi x k_n - x h_n| = |x||\xi k_n - h_n| \geq |\xi k_n - h_n|,$$

contradicting our assumption that $|\xi q - p| < |\xi k_n - h_n|$.

We can further prove that x and y have opposite signs. If $y < 0$, then $xk_n = q - yk_{n+1} > 0$, so $x > 0$. If $y > 0$, then our assumption that $q < k_{n+1}$ implies that $q < yk_{n+1}$. Then $xk_n = q - yk_{n+1} < 0$, so $x < 0$. By Theorem 2.7, $\xi k_n - h_n$ and $\xi k_{n+1} - h_{n+1}$ have opposite signs, so $x(\xi k_n - h_n)$ and $y(\xi k_{n+1} - h_{n+1})$ have the same signs. By the equations defining x and y, we get

$$\begin{aligned}\xi q - p &= \xi(xk_n + yk_{n+1}) - (xh_n + yh_{n+1}) \\ &= x\xi k_n + y\xi k_{n+1} - xh_n - yh_{n+1} \\ &= x(\xi k_n - h_n) + y(\xi k_{n+1} - h_{n+1}).\end{aligned}$$

Since the two terms on the right have the same signs, the absolue value of their sum equals the sum of their absolute values. Hence we have

$$\begin{aligned}|\xi q - p| &= |x(\xi k_n - h_n) + y(\xi k_{n+1} - h_{n+1})| \\ &= |x(\xi k_n - h_n)| + |y(\xi k_{n+1} - h_{n+1})| \\ &> |x(\xi k_n - h_n)| \\ &= |x||\xi k_n - h_n| \\ &\geq |\xi k_n - h_n|.\end{aligned}$$

But this contradicts our assumption that $|\xi q - p| < |\xi k_n - h_n|$. Hence we have proven the first implication by contradiction.

We now prove the second implication again by contradiction. Suppose there exists a rational p/q with $q > 0$ such that $|\xi - p/q| < |\xi - h_n/k_n|$ and $q \leq k_n$. Multiplying these two inequalities, we get $|\xi q - p| < |\xi k_n - h_n|$, which by the statement just proven implies that $q \geq k_{n+1} > k_n$, contradicting our assumption that $q \leq k_n$. This proves our second implication. □

Theorem 3.4. *Let ξ be an irrational number. If p/q is a rational number with $q > 0$ such that*

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{2q^2},$$

then p/q is a convergent of the simple continued fraction expansion of ξ.

Proof. Let p/q be a rational number satisfying the hypotheses of the theorem. Without loss of generality, we may take $\gcd(p, q) = 1$. Let h_i/k_i be the ith convergent of the simple continued fraction expansion of ξ for all $i \geq 0$. Suppose that p/q is not one of these convergents. There exists a unique integer n such that $k_n \leq q < k_{n+1}$. By Theorem 3.3, we must then have

$$\begin{aligned}|\xi k_n - h_n| \leq |\xi q - p| < \frac{1}{2q} \\ \left|\xi - \frac{h_n}{k_n}\right| < \frac{1}{2qk_n}.\end{aligned}$$

By assumption, $p/q \neq h_n/k_n$, so we have

$$\begin{aligned} \frac{1}{qk_n} &\leq \frac{|qh_n - pk_n|}{qk_n} \\ &= \left|\frac{h_n}{k_n} - \frac{p}{q}\right| \\ &\leq \left|\frac{h_n}{k_n} - \xi\right| + \left|\xi - \frac{p}{q}\right| \\ \frac{1}{qk_n} &< \frac{1}{2qk_n} + \frac{1}{2q^2} \\ \frac{1}{2qk_n} &< \frac{1}{2q^2} \\ \frac{1}{k_n} &< \frac{1}{q} \\ q &< k_n, \end{aligned}$$

contradicting the definition of k_n. $\square$

The next two results establish a stronger result than the one in Corollary 3.1.

Lemma 3.1. *Let $x \in [1, \infty)$. If $x + x^{-1} < \sqrt{5}$, then $x < (\sqrt{5}+1)/2$ and $x^{-1} > (\sqrt{5}-1)/2$.*

Proof. If $x \in (1, \infty)$, then we have

$$\begin{aligned} x^2 &> 1 \\ 1 &> \frac{1}{x^2} \\ 1 - \frac{1}{x^2} &> 0 \\ \frac{d}{dx}\left(x + x^{-1}\right) &> 0, \end{aligned}$$

so $x + x^{-1}$ is increasing for $x \geq 1$. Since $x + x^{-1} = \sqrt{5}$ for $x = (\sqrt{5}+1)/2$, $x + x^{-1} < \sqrt{5}$ implies $x < (\sqrt{5}+1)/2$. We then have

$$\begin{aligned} x &< \frac{\sqrt{5}+1}{2} \\ \frac{2}{\sqrt{5}+1} &< \frac{1}{x} \\ \frac{\sqrt{5}-1}{2} &< \frac{1}{x}. \end{aligned}$$

$\square$

Theorem 3.5 (Hurwitz's Theorem)**.** *Let ξ be an irrational number. There exists infinitely many rational numbers p/q such that*

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

Proof. We will show that at least one of every three consecutive convergents of the simple continued fraction expansion of ξ satisfies the given inequality. This then immediately implies the theorem.

Let $u_i = k_i/k_{i-1}$. We first prove that

$$u_j + u_j^{-1} < \sqrt{5}$$

if

$$\left|\xi - \frac{h_i}{k_i}\right| \geq \frac{1}{\sqrt{5}k_i^2}$$

for $i = j-1$ and $i = j$. Suppose the above does hold for $i = j-1$ and $i = j$. Then we have

$$\left|\xi - \frac{h_{j-1}}{k_{j-1}}\right| + \left|\xi - \frac{h_j}{k_j}\right| \geq \frac{1}{\sqrt{5}k_{j-1}^2} + \frac{1}{\sqrt{5}k_j^2}.$$

Since ξ is between h_{j-1}/k_{j-1} and h_j/k_j by Theorem 2.3, we have

$$\left|\xi - \frac{h_{j-1}}{k_{j-1}}\right| + \left|\xi - \frac{h_j}{k_j}\right| = \left|\frac{h_{j-1}}{k_{j-1}} - \frac{h_j}{k_j}\right| = \frac{1}{k_{j-1}k_j}.$$

Combining these, we get

$$\begin{aligned} \frac{1}{\sqrt{5}k_{j-1}^2} + \frac{1}{\sqrt{5}k_j^2} &\leq \frac{1}{k_{j-1}k_j} \\ \frac{k_j}{k_{j-1}} + \frac{k_{j-1}}{k_j} &\leq \sqrt{5} \\ u_j + u_j^{-1} &\leq \sqrt{5}. \end{aligned}$$

Since the left side is rational and the right side is irrational, this is actually a strict inequality, as claimed.

Now suppose that

$$\left|\xi - \frac{h_i}{k_i}\right| \geq \frac{1}{\sqrt{5}k_i^2}$$

for $i = j-1, j, j+1$. We then have $u_i + u_i^{-1} < \sqrt{5}$ for $i = j, j+1$. By Lemma 3.1, this implies, in particular, that $u_j^{-1} > (\sqrt{5}-1)/2$ and $u_{j+1} < (\sqrt{5}+1)/2$. Thus we have

$$\begin{aligned} \frac{\sqrt{5}+1}{2} &> u_{j+1} \\ &= \frac{k_{j+1}}{k_j} \\ &= \frac{a_{j+1}k_j + k_{j-1}}{k_j} \\ &= a_{j+1} + u_j^{-1} \\ &> a_{j+1} + \frac{\sqrt{5}-1}{2} \\ &\geq 1 + \frac{\sqrt{5}-1}{2} \\ &= \frac{\sqrt{5}+1}{2}, \end{aligned}$$

a contradiction. $\square$

Theorem 3.6. *The constant $\sqrt{5}$ in Hurwitz's Theorem is the best possible, i.e., there exists an irrational number ξ such that, for all $\epsilon > 0$, there exists only finitely many rationals p/q such that*

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{(\sqrt{5}+\epsilon)q^2}.$$

Proof. Let $\xi = [1; 1, 1, \dots]$. We have

$$\begin{aligned} \xi &= 1 + \frac{1}{[1; 1, 1, \dots]} \\ \xi &= 1 + \frac{1}{\xi} \\ \xi^2 &= \xi + 1 \\ \xi^2 - \xi - 1 &= 0. \end{aligned}$$

Thus

$$\xi = \frac{1 \pm \sqrt{5}}{2},$$

but since $\xi > 1$, we in fact have $\xi = (1+\sqrt{5})/2$. Letting $\xi_0 = \xi$ and $a_0 = \lfloor \xi_0 \rfloor$, define the sequences $\{a_i\}_{i=0}^{\infty}$ and $\{\xi_i\}_{i=0}^{\infty}$ as in Equation (1). Suppose inductively that $\xi_i = (1+\sqrt{5})/2$. We then have

$$\xi_{i+1} = \frac{1}{\xi_i - a_i} = \frac{1}{(\sqrt{5}+1)/2 - 1} = \frac{1}{(\sqrt{5}-1)/2} = \frac{\sqrt{5}+1}{2},$$

so $\xi_i = \xi$ for all $i \geq 0$. Noting that $a_i = 1$ for all $i \geq 0$, we can rewrite $h_i = h_{i-1} + h_{i-2}$ and $k_i = k_{i-1} + k_{i-2}$, so

$$k_1 = k_0 + k_{-1} = k_{-1} + k_{-2} + k_{-1} = 1 = h_{-1} + h_{-2} = h_0.$$

If $k_{i+1} = h_i$ for all integers $i \in [0, n-1]$, then we have

$$k_{n+1} = k_n + k_{n-1} = h_{n-1} + h_{n-2} = h_n,$$

so by induction $k_{i+1} = h_i$ for all integers $i \geq 0$. Then

$$\begin{aligned} \lim_{n\to\infty} \left(\xi_{n+1} + \frac{k_{n-1}}{k_n}\right) &= \lim_{n\to\infty} \xi_{n+1} + \lim_{n\to\infty} \frac{k_{n-1}}{k_n} \\ &= \xi + \lim_{n\to\infty} \frac{k_{n-1}}{h_{n-1}} \\ &= \xi + \frac{1}{\xi} \\ &= \sqrt{5}. \end{aligned}$$

This means that $\xi_{n+1} + k_{n-1}/k_n$ is eventually less than $\sqrt{5}+\epsilon$ for every $\epsilon > 0$, so

$$\xi_{n+1} + \frac{k_{n-1}}{k_n} > \sqrt{5} + \epsilon$$

holds for only finitely many integers n. Hence there are only finitely many i such that

$$\left|\xi - \frac{h_i}{k_i}\right| = \frac{1}{k_i(\xi_{i+1}k_i + k_{i-1})} = \frac{1}{k_i^2(\xi_{i+1} + k_{i-1}/k_i)} < \frac{1}{(\sqrt{5}+\epsilon)k_i^2}.$$

Since any rational number p/q satisfying $|\xi - p/q| < 1/((\sqrt{5}+\epsilon)q^2) < 1/2q^2$ is one of the convergents of the simple continued fraction of ξ by Theorem 3.4, this proves the result. $\square$

3.1. **Quadratic Irrationals.** Let $[a_0; a_1, a_2, \dots]$ be an infinite simple continued fraction. $[a_0; a_1, a_2, \dots]$ is *periodic* if there exists an integer n such that $a_i = a_{i+n}$ for i sufficiently large. We can write these in the form

$$[a_0; a_1, \dots, a_j, \overline{a_{j+1}, \dots, a_{j+n}}],$$

where the bar above $a_{j+1}, \dots, a_{j+n}$ indicates that this block of partial quotients is repeated indefinitely.

In the following result, *quadratic irrationals* are those irrational numbers which are the solutions to quadratic equations with integer coefficients.

Theorem 3.7. *Let $\xi = [a_0; a_1, a_2, \dots,]$. $[a_0; a_1, a_2, \dots]$ is periodic if and only if ξ is a quadratic irrational.*

Proof. Suppose ξ is periodic so that $\xi = [a_0; a_1, \dots, a_j, \overline{a_{j+1}, \dots, a_{j+n}}]$. Letting

$$\theta = [\overline{a_{j+1}, \dots, a_{j+n}}] = [a_{j+1}; a_{j+2}, \dots, a_{j+n}, \theta]$$

we have

$$\theta = \frac{\theta h_{n-1} + h_{n-2}}{\theta k_{n-1} + k_{n-2}},$$

which is a quadratic equation in θ with integer coefficients, so θ is either a rational or quadratic irrational. Since θ is necessarily irrational, θ is a quadratic irrational. Now we have

$$\xi = [a_0; a_1, \dots, a_j, \theta] = \frac{\theta p + p'}{\theta q + q'},$$

where p/q and p'/q' are the last two convergents of $[a_0; a_1, \dots, a_j]$. Hence ξ is also a quadratic irrational.

Suppose ξ is a quadratic irrational, so it can be represented as $\xi = \xi_0 = (a+\sqrt{b})/c$ for some integers a, b, c with b positive and c non-zero. Since ξ is irrational, b is not a perfect square. We then have

$$\xi_0 = \frac{a+\sqrt{b}}{c} = \frac{a|c| + |c|\sqrt{b}}{c|c|} = \frac{a|c| + \sqrt{bc^2}}{c|c|},$$

so we can write

$$\xi_0 = \frac{m_0 + \sqrt{d}}{q_0}$$

for some integers m_0, q_0, d such that $q_0 \neq 0$, $q_0 \mid (d - m_0^2)$, and d is not a perfect square. Defining

$$a_i = \lfloor \xi_i \rfloor, \qquad m_{i+1} = a_i q_i - m_i,$$
$$q_{i+1} = \frac{d - m_{i+1}^2}{q_i}, \qquad \xi_i = \frac{m_i + \sqrt{d}}{q_i}, \tag{3}$$

we claim that $\xi = [a_0; a_1, a_2, \dots]$.

We first show that m_i and q_i are integers such that $q_i \neq 0$ and $q_i \mid (d - m_i^2)$. Having already established this for $i = 0$, suppose inductively that this is true for i. Then $m_{i+1} = a_i q_i - m_i$ is an integer since a_i is an integer by definition, and

$$q_{i+1} = \frac{d - m_{i+1}^2}{q_i} = \frac{d - (a_i^2 q_i^2 - 2a_i m_i q_i + m_i^2)}{q_i} = \frac{d - m_i^2}{q_i} + 2a_i m_i - a_i^2 q_i$$

is an integer because $q_i \mid (d - m_i^2)$. If $q_{i+1} = 0$, then $d = m_{i+1}^2$, contradicting the fact that d is not a perfect square; thus $q_{i+1} \neq 0$. Since $q_i q_{i+1} = d - m_{i+1}^2$, $q_{i+1} \mid (d - m_{i+1}^2)$, so we are done.

Noting that

$$\begin{aligned}
\xi_i - a_i &= \frac{m_i + \sqrt{d}}{q_i} - a_i \\
&= \frac{-(a_i q_i - m_i) + \sqrt{d}}{q_i} \\
&= \frac{\sqrt{d} - m_{i+1}}{q_i} \\
&= \frac{d - m_{i+1}^2}{q_i(\sqrt{d} + m_{i+1})} \\
&= \frac{q_{i+1}}{\sqrt{d} + m_{i+1}} \\
&= \frac{1}{\xi_{i+1}} \\
\xi_i &= a_i + \frac{1}{\xi_{i+1}},
\end{aligned}$$

we have proven that $\xi = [a_0; a_1, a_2, \ldots]$ by Theorem 2.7.

Let $\xi_i' = (m_i - \sqrt{d})/q_i$ for $i \geq 0$ (ξ_i' is the *conjugate* of ξ_i). By rules of conjugation and Theorem 2.2, we then have

$$\begin{aligned}
\xi_0' &= \frac{\xi_i' h_{i-1} + h_{i-2}}{\xi_i' k_{i-1} + k_{i-2}} \\
\xi_0' \xi_i' k_{i-1} + \xi_0' k_{i-2} &= \xi_i' h_{i-1} + h_{i-2} \\
\xi_i'(\xi_0' k_{i-1} - h_{i-1}) &= -(\xi_0' k_{i-2} - h_{i-2}) \\
\xi_i' &= -\frac{k_{i-2}}{k_{i-1}} \left(\frac{\xi_0' - h_{i-2}/k_{i-2}}{\xi_0' - h_{i-1}/k_{i-1}} \right).
\end{aligned}$$

$h_{i-2}/k_{i-2}, h_{i-1}/k_{i-1} \to \xi_0$ as $i \to \infty$, and since $\xi_0 \neq \xi_0'$, the fraction in parenthesis goes to 1. Thus ξ_i' is negative for $i > N$ for some fixed N. Since ξ_i is positive for $i \geq 1$, this implies $\xi_i - \xi_i' > 0$ for $i > N$. We thus have

$$\frac{2\sqrt{d}}{q_i} = \frac{m_i + \sqrt{d}}{q_i} - \frac{m_i - \sqrt{d}}{q_i} > 0,$$

so $q_i > 0$ for $i > N$. We also have

$$\begin{aligned} q_i \leq q_i q_{i+1} &= d - m_{i+1}^2 \leq d, \\ m_{i+1}^2 < m_{i+1}^2 + q_i q_{i+1} &= d \\ |m_{i+1}| < \sqrt{d}, \end{aligned}$$

for $i > N$. Since d is fixed, this implies that there are only finitely many values for q_i and m_i and thus only finitely many distinct pairs (m_i, q_i). Hence we can choose integers r, s with $r < s$ such that $m_r = m_s$ and $q_r = q_s$. But then

$$\xi_r = \frac{m_r + \sqrt{d}}{q_r} = \frac{m_s + \sqrt{d}}{q_s} = \xi_s,$$

so $a_r = a_s$. Suppose inductively that $m_{r+i} = m_{s+i}$ and $q_{r+i} = q_{s+i}$. Then $\xi_{r+i} = \xi_{s+i}$, so $a_{r+i} = a_{s+i}$. But then we have

$$m_{r+i+1} = a_{r+i} q_{r+i} - m_{r+i} = a_{s+i} q_{s+i} - m_{s+i} = m_{s+i+1}$$

and

$$q_{r+i+1} = \frac{d - m_{r+i+1}^2}{q_{r+i}} = \frac{d - m_{s+i+1}^2}{q_{s+i}} = q_{s+i+1}.$$

Hence by induction

$$a_{r+i} = \lfloor \xi_{r+i} \rfloor = \left\lfloor \frac{m_{r+i} + \sqrt{d}}{q_{r+i}} \right\rfloor = \left\lfloor \frac{m_{s+i} + \sqrt{d}}{q_{s+i}} \right\rfloor = \lfloor \xi_{s+i} \rfloor = a_{s+i}$$

for all $i \geq 0$. Thus $[a_0; a_1, a_2, \dots]$ is periodic. $\square$

We say that a simple continued fraction is *purely periodic* if it is of the form $[\overline{a_0; a_1, \dots, a_n}]$.

Theorem 3.8. *Let $\xi = [a_0; a_1, a_2, \dots,]$ be a quadratic irrational. Then $[a_0; a_1, a_2, \dots]$ is purely periodic if and only if $\xi > 1$ and $-1 < \xi' < 0$, where ξ' is the conjugate of ξ.*

Proof. Suppose $\xi > 1$ and $-1 < \xi' < 0$. Defining ξ_i for $i \geq 0$ as in Equation (1), by rules of conjugation we have

$$\frac{1}{\xi_{i+1}'} = \xi_i' - a_i.$$

Since $\xi_0 = \xi > 1$, $a_i \geq 1$ for all $i \geq 0$. Thus if $\xi_i' < 0$, by the above equation we must have $1/\xi_{i+1}' < -1$, so that $-1 < \xi_{i+1}' < 0$. Since $-1 < \xi' = \xi_0' < 0$, this proves by induction that $-1 < \xi_i' < 0$ for all $i \geq 0$. Hence

$$0 < -\xi_i' = -a_i - \frac{1}{\xi_{i+1}'} < 1,$$

so $\lfloor -1/\xi_{i+1}' \rfloor = a_i$. Since ξ is a quadratic irrational, we must have $\xi_r = \xi_s$ for some integers $0 < r < s$. Then $\xi_r' = \xi_s'$ and we have

$$\begin{aligned} a_{r-1} &= \left\lfloor -\frac{1}{\xi_r'} \right\rfloor = \left\lfloor -\frac{1}{\xi_s'} \right\rfloor = a_{s-1}, \\ \xi_{r-1} &= a_{r-1} + \frac{1}{\xi_r} = a_{s-1} + \frac{1}{\xi_s} = \xi_{s-1}. \end{aligned}$$

By iterating this argument, we then have $\xi_0 = \xi_{s-r}$. Thus

$$\xi = \xi_0 = [a_0; a_1, \dots, a_{r-s-1}, \xi_{r-s}] = [a_0; a_1, \dots, a_{r-s-1}, \xi_0] = [\overline{a_0; a_1, \dots, a_{r-s-1}}].$$

Suppose conversely that ξ is purely periodic, i.e., $\xi = [\overline{a_0; a_1, \dots, a_{n-1}}]$. Since $a_0 = a_n \geq 1$, $a_i \geq 1$ for all $i \geq 0$. Then $\xi > \lfloor \xi \rfloor = a_0 \geq 1$. We also have

$$\xi = [a_0; a_1, \dots, a_{n-1}, \xi] = \frac{\xi h_{n-1} + h_{n-2}}{\xi k_{n-1} + k_{n-2}},$$

so ξ satisfies the quadratic equation

$$f(x) = k_{n-1}x^2 + (k_{n-2} - h_{n-1})x - h_{n-2} = 0.$$

ξ and ξ' are the only solutions to this equation, so it suffices to show that f has a root between -1 and 0 by showing that $f(-1)$ and $f(0)$ have opposite signs. We have $f(0) = -h_{n-2} < 0$ since the a_i are positive for $i \geq 0$. Further, we have

$$\begin{aligned} f(-1) &= k_{n-1} - k_{n-2} + h_{n-1} - h_{n-2} \\ &= a_{n-1}k_{n-2} + k_{n-3} - k_{n-2} + a_{n-1}h_{n-2} + h_{n-3} - h_{n-2} \\ &= (k_{n-2} + h_{n-2})(a_{n-1} - 1) + k_{n-3} + h_{n-3} \\ &\geq k_{n-3} + h_{n-3} \\ &> 0 \end{aligned}$$

for $n \geq 1$. $\square$

Theorem 3.9. *Let d be a positive integer that is not a perfect square. The simple continued fraction expansion of $\sqrt{d}$ has the form*

$$\sqrt{d} = [a_0; \overline{a_1, \dots, a_{r-1}, 2a_0}].$$

Furthermore, setting $\xi_0 = \sqrt{d}$, $q_0 = 1$, and $m_0 = 0$ and letting r be the length of the shortest period in the expansion of $\sqrt{d}$, we have $q_i = 1$ if and only if $r \mid i$ and $q_i \neq -1$ for all $i \geq 0$.

Proof. Consider the number $\sqrt{d} + \lfloor \sqrt{d} \rfloor$. We have $\sqrt{d} + \lfloor \sqrt{d} \rfloor \geq \sqrt{2} + 1 > 1$ and

$$\lfloor \sqrt{d} \rfloor < \sqrt{d} < \lfloor \sqrt{d} \rfloor + 1 \implies -1 < -\sqrt{d} + \lfloor \sqrt{d} \rfloor < 0,$$

so by Theorem 3.8 $\sqrt{d} + \lfloor \sqrt{d} \rfloor$ is purely periodic, i.e.,

$$\sqrt{d} + \lfloor \sqrt{d} \rfloor = [\overline{a_0; a_1, \dots, a_{r-1}}] = [a_0; \overline{a_1, \dots, a_{r-1}, a_0}],$$

where r is chosen as small as possible. Note that $\xi_i = [a_i; a_{i+1}, a_{i+2}, \dots]$ is purely periodic for all i and $\xi_0 = \xi_{jr}$ for all $j \geq 0$. We must also have $\xi_1, \dots, \xi_{r-1}$ all different from ξ_0, for otherwise would imply that r is not minimal. Thus $\xi_i = \xi_0$ if and only if $i = jr$ for some $j \geq 0$.

Let $\xi_0 = \sqrt{d} + \lfloor \sqrt{d} \rfloor$, $q_0 = 1$, and $m_0 = \lfloor \sqrt{d} \rfloor$. Then $m_0, q_0 \in \mathbb{Z}$, $q_0 \neq 0$, and $q_0 \mid (d - m_0^2)$. Thus we have, for all $j \geq 0$,

$$\begin{aligned} \xi_{jr} &= \xi_0 \\ \frac{m_{jr} + \sqrt{d}}{q_{jr}} &= \frac{m_0 + \sqrt{d}}{q_0} \\ \frac{m_{jr} + \sqrt{d}}{q_{jr}} &= \sqrt{d} + \lfloor \sqrt{d} \rfloor \\ m_{jr} - q_{jr}\lfloor \sqrt{d} \rfloor &= (q_{jr} - 1)\sqrt{d}. \end{aligned}$$

Since the left side is an integer and the right side is irrational when $q_{jr} \neq 1$, we must have $q_{jr} = 1$. Suppose $q_i = 1$. Then $\xi_i = m_i + \sqrt{d}$, so by Theorem 3.8 we have

$$-1 < m_i - \sqrt{d} < 0 \implies \sqrt{d} - 1 < m_i < \sqrt{d},$$

so $m_i = \lfloor \sqrt{d} \rfloor$. Then $\xi_i = \xi_0$, so $i = jr$ for some $j \geq 0$.

Suppose $q_i = -1$. Then $\xi_i = -m_i - \sqrt{d}$, so by Theorem 3.8 we have $-m_i - \sqrt{d} > 1$ and $-1 < -m_i + \sqrt{d} < 0$. But this implies $\sqrt{d} < m_i < -\sqrt{d} - 1$, a contradiction.

Seeing that $a_0 = \lfloor \sqrt{d} + \lfloor \sqrt{d} \rfloor \rfloor = 2\lfloor \sqrt{d} \rfloor$, we now have

$$\begin{aligned}
\sqrt{d} &= -\lfloor \sqrt{d} \rfloor + (\sqrt{d} + \lfloor \sqrt{d} \rfloor) \\
&= -\lfloor \sqrt{d} \rfloor + [a_0; \overline{a_1, \dots, a_{r-1}, a_0}] \\
&= -\lfloor \sqrt{d} \rfloor + \left[2\lfloor \sqrt{d} \rfloor; \overline{a_1, \dots, a_{r-1}, 2\lfloor \sqrt{d} \rfloor}\right] \\
&= \left[\lfloor \sqrt{d} \rfloor; \overline{a_1, \dots, a_{r-1}, 2\lfloor \sqrt{d} \rfloor}\right].
\end{aligned}$$

Consider applying our recursive definitions to the starting values $\xi_0 = \sqrt{d}$, $q_0 = 1$, and $m_0 = 0$. We have

$$\begin{aligned}
a_0 &= \lfloor \xi_0 \rfloor = \lfloor \sqrt{d} \rfloor, \\
m_1 &= a_0 q_0 - m_0 = \lfloor \sqrt{d} \rfloor, \\
q_1 &= \frac{d - m_1^2}{q_0} = d - \lfloor \sqrt{d} \rfloor^2.
\end{aligned}$$

But the values of m_1 and q_1 are the same when we start with $\xi_0 = \sqrt{d} + \lfloor \sqrt{d} \rfloor$, $m_0 = \lfloor \sqrt{d} \rfloor$, and $q_0 = 1$, so the partial quotients a_i are the same for $\sqrt{d}$ and $\sqrt{d} + \lfloor \sqrt{d} \rfloor$ for $i \geq 1$. □

4. Pell's Equation

Let d and N be integers, with $d > 0$ and not a perfect square. The Diophantine equation $x^2 - dy^2 = N$ is called Pell's equation. It shares its name with English mathematician John Pell due to a mistaken attribution by Euler. In the following section, we explore how our results on continued fractions can be applied to Pell's equation.

Theorem 4.1. *Let d be a positive integer not a perfect square. Setting $\xi_0 = \sqrt{d}$, $q_0 = 1$, and $m_0 = 0$, expand $\sqrt{d}$ into a simple continued fraction and let h_i/k_i be the ith convergent of $\sqrt{d}$. Then we have $h_i^2 - dk_i^2 = (-1)^{i-1} q_{i+1}$ for all integers $i \geq 0$.*

Proof. We have

$$\sqrt{d} = \xi_0 = [a_0; a_1, \dots, a_i, \xi_{i+1}] = \frac{\xi_{i+1} h_i + h_{i-1}}{\xi_{i+1} k_i + k_{i-1}} = \frac{(m_{i+1} + \sqrt{d}) h_i + q_{i+1} h_{i-1}}{(m_{i+1} + \sqrt{d}) k_i + q_{i+1} k_{i-1}},$$

and so

$$\begin{aligned}
\sqrt{d} &= \frac{(m_{i+1}+\sqrt{d})h_i + q_{i+1}h_{i-1}}{(m_{i+1}+\sqrt{d})k_i + q_{i+1}k_{i-1}}\\
\sqrt{d} &= \frac{m_{i+1}h_i + h_i\sqrt{d} + q_{i+1}h_{i-1}}{m_{i+1}k_i + k_i\sqrt{d} + q_{i+1}k_{i-1}}\\
m_{i+1}k_i\sqrt{d} + dk_i + q_{i+1}k_{i-1}\sqrt{d} &= m_{i+1}h_i + h_i\sqrt{d} + q_{i+1}h_{i-1}\\
(m_{i+1}k_i + q_{i+1}k_{i-1} - h_i)\sqrt{d} &= m_{i+1}h_i + q_{i+1}h_{i-1} - dk_i.
\end{aligned}$$

If $m_{i+1}k_i + q_{i+1}k_{i-1} - h_i \neq 0$, then the left side is irrational and the right is rational, a contradiction, so both sides must equal zero. This then immediately implies

$$m_{i+1}h_ik_i + q_{i+1}h_ik_{i-1} - h_i^2 = 0$$

and

$$m_{i+1}h_ik_i + q_{i+1}h_{i-1}k_i - dk_i^2 = 0,$$

so we have

$$\begin{aligned}
m_{i+1}h_ik_i + q_{i+1}h_ik_{i-1} - h_i^2 &= m_{i+1}h_ik_i + q_{i+1}h_{i-1}k_i - dk_i^2\\
h_i^2 - dk_i^2 &= (h_ik_{i-1} - h_{i-1}k_i)q_{i+1}\\
h_i^2 - dk_i^2 &= (-1)^{i-1}q_{i+1}.
\end{aligned}$$

□

Thus Pell's equation is guaranteed to be solvable if N can be found amongst the integers $(-1)^{i-1}q_{i+1}$. The periodicity of $\sqrt{d}$ then implies that the convergents to $\sqrt{d}$ in fact provide infinitely many solutions.

Corollary 4.1. *Let d be a positive integer not a perfect square, let h_i/k_i be the ith convergent of $\sqrt{d}$, and let r be the length of the shortest period of the expansion of $\sqrt{d}$. Then, for all integers $i \geq 0$, we have*

$$h_{ir-1}^2 - dk_{ir-1}^2 = (-1)^{ir-2}q_{ir} = (-1)^{ir}.$$

Theorem 4.2. *Let d be a positive integer not a perfect square, let h_i/k_i be the ith convergent to $\sqrt{d}$, and let N be an integer with $|N| < \sqrt{d}$. If s and t are positive integers such that $s^2 - dt^2 = N$ and $\gcd(s,t) = 1$, then $s = h_i$ and $t = k_i$ for some integer $i \geq 1$.*

Proof. Let E and M be positive integers such that $\gcd(E,M) = 1$ and $E^2 - \rho M^2 = \sigma$, where ρ and σ are real numbers with $\sqrt{\rho}$ irrational and $0 < \sigma < \sqrt{\rho}$. Then

$$\frac{E}{M} - \sqrt{\rho} = \frac{E - M\sqrt{\rho}}{M} = \frac{E^2 - \rho M^2}{M(E + M\sqrt{\rho})} = \frac{\sigma}{M(E + M\sqrt{\rho})},$$

so we have

$$0 < \frac{\sigma}{M(E+M\sqrt{\rho})} = \frac{E}{M} - \sqrt{\rho} < \frac{\sqrt{\rho}}{M(E+M\sqrt{\rho})} = \frac{1}{M^2(E/(M\sqrt{\rho})+1)}.$$

Since $0 < E/M - \sqrt{\rho}$, we have $E/(M\sqrt{\rho}) > 1$, so the above implies

$$\left|\frac{E}{M} - \sqrt{\rho}\right| < \frac{1}{2M^2}.$$

By Theorem 3.4, E/M must be a convergent of the continued fraction expansion of $\sqrt{\rho}$.

Suppose $N > 0$. Letting $\sigma = N$, $\rho = d$, $E = s$, and $M = t$, the argument above shows that the theorem holds in this case.

Suppose $N < 0$. Since $t^2 - s^2/d = -N/d$, letting $\sigma = -N/d$, $\rho = 1/d$, $E = t$, and $M = s$, we see that $E/M = t/s$ must be a convergent of $\sqrt{\rho} = 1/\sqrt{d}$. By Theorem 2.8, this shows that s/t is a convergent of $\sqrt{d}$. $\square$

Theorem 4.3. *Let d be a positive integer not a perfect square and let r be the length of the shortest period in the expansion of $\sqrt{d}$. If r is even, then $x^2-dy^2 = -1$ has no solution, and all positive solutions of $x^2 - dy^2 = 1$ are $x = h_{ir-1}$ and $y = k_{ir-1}$ for integers $i \geq 1$. If r is odd, then $x = h_{ir-1}$ and $y = k_{ir-1}$ give all positive solutions of $x^2 - dy^2 = -1$ for positive odd integers i and all positive solutions of $x^2 - dy^2 = 1$ for positive even integers.*

Proof. Since d is a positive integer not a perfect square, $\sqrt{d} \geq \sqrt{2} > 1$, so by Theorem 4.2, any positive solution of $x^2 - dy^2 = \pm 1$ is of the form $x = h_i$ and $y = k_i$ for some integer $i \geq 1$, where h_i/k_i is the ith convergent of $\sqrt{d}$. By Theorem 4.1, we have

$$x^2 - dy^2 = h_i^2 - dk_i^2 = (-1)^{i-1} q_{i+1}$$

for all integers $i \geq 0$. By Theorem 3.9, $q_{i+1} = 1$ if and only if $r \mid i+1$ and $q_{i+1} \neq -1$ for all $i \geq -1$. Thus the solutions of $x^2 - dy^2 = \pm 1$ can only occur at $x = h_{ir-1}$ and $y = k_{ir-1}$ for integers $i \geq 1$.

Suppose r is odd. Then $x = h_{ir-1}$ and $y = k_{ir-1}$ is a solution of $x^2 - dy^2 = -1$ when i is an odd positive integer and a solution of $x^2 - dy^2 = 1$ when i is an even positive integer.

Suppose r is even. Then $x = h_{ir-1}$ and $y = k_{ir-1}$ is a solution of $x^2 - dy^2 = 1$ and thus not a solution of $x^2 - dy^2 = -1$ for all integers $i \geq 1$. $\square$

We have thus completely solved Pell's equation for the case $N = \pm 1$. A more general result turns out to be not so simple. Suppose $x^2 - dy^2 = N$ has a positive solution (x, y). Let (u_i, v_i) be the positive solutions of $x^2 - dy^2 = 1$. Then

$$\left(u_i + v_i\sqrt{d}\right)\left(x + y\sqrt{d}\right) = (u_i x + v_i y d) + (u_i y + v_i x)\sqrt{d}$$

and

$$\begin{aligned}
(u_i x + v_i y d)^2 - d(u_i y + v_i x)^2 &= u_i^2 x^2 + 2u_i v_i x y d + v_i^2 y^2 d^2 - u_i^2 y^2 d - 2u_i v_i x y d - v_i^2 x^2 d \\
&= u_i^2 x^2 + v_i^2 y^2 d^2 - u_i^2 y^2 d - v_i^2 x^2 d \\
&= x^2\left(u_i^2 - dv_i^2\right) - dy^2\left(u_i^2 - dv_i^2\right) \\
&= x^2 - dy^2 \\
&= N.
\end{aligned}$$

Thus, given any initial solution of $x^2 - dy^2 = N$, we can generate infinitely many solutions [1]. It can also be shown that there exists a finite set of solutions (x_i, y_i) of $x^2-dy^2 = N$ such that every solution is either in the set or of the form $x+y\sqrt{d} = (x_i + y_i\sqrt{d})(u_j + v_j\sqrt{d})$ [1].

References

[1] Titu Andreescu and Dorin Andrica. *Quadratic Diophantine Equations.* Springer, 2015. ISBN: 9780387351568.

[2] Claude Brezinski. *History of Continued Fractions and Padé Approximants.* Springer-Verlag, 1991. ISBN: 9783540152866.

[3] George Chrystal. *Algebra: An Elementary Text-book for the Higher Classes of Secondary Schools and for Colleges: Part II.* AMS Chelsea Publishing, 1900. ISBN: 9780821816493.

[4] Ivan Niven, Herbert S. Zuckerman, and Hugh L. Montgomery. *An Introduction to the Theory of Numbers.* John Wiley Sons, Inc., 1991. ISBN: 9780471625469.