CONTINUED FRACTIONS AND PELL EQUATIONS

ARANYA KARIGHATTAM

Abstract. In this paper, we will study finite, infinite, and periodic continued fractions and discuss how they are used to find solutions to the Pell Equation $x^2 - dy^2 = \pm 1$ for any integer d. Specifically, we will look at interesting properties of convergents of continued fractions and prove Dirichlet's Theorem on Diophantine Approximation, which states that there are infinitely many rational numbers that are good approximations to an irrational number. We also prove an important result by Euler and Lagrange relating periodic continued fractions and quadratic irrationals. We will find the solutions to the Pell Equation using properties of the periodic continued fraction of $\sqrt{d}$. We follow the material in [1].

1. Introduction

The Pell Equation is a diophantine equation of the form $x^2 - dy^2 = \pm 1$ where x, y, d are positive integers. Can we find solutions to this equation? If $d = 2$, then $(1, 0)$ and $(3, 2)$ are solutions to $x^2 - 2y^2 = 1$ and $(1, 1)$ and $(7, 5)$ are solutions to $x^2 - 2y^2 = -1$. However, if $d \equiv 3 \pmod 4$, $x^2 - dy^2 = -1$ has no solutions. This can be seen by noticing that $x^2, y^2 \equiv 0$ or $1 \pmod 4$, so $x^2 - dy^2 \equiv 0, 1,$ or $2 \pmod 4$ but not $-1 \pmod 4$. If d is a perfect square, we can write d as m^2, so $x^2 - m^2y^2 = (x + my)(x - my) = 1$. In this case, $x + my$ and $x - my$ are either both 1 or -1. Since x, y, m are positive integers, $x + my$ cannot be equal to -1, so both terms must be equal to 1. Therefore the only solution is $(x, y) = (1, 0)$. The other equation is $x^2 - m^2y^2 = (x + my)(x - my) = -1$. One term must be equal to 1 and the other must be equal to -1. This implies $x + my = 1$ and $x - my = -1$. The only solution to this equation is $(x, y) = (0, 1)$ where $d = 1$.

In general, for any positive d that is not a perfect square, how many solutions does $x^2 - dy^2 = \pm 1$ have and how do we find them? In 1657, Fermat stated that there are infinitely many solutions to the Pell Equation $x^2 - dy^2 = 1$, and Wallis and Brouncker found that continued fractions can be used to find the solutions. Euler showed, in 1767, that if there is a fundamental solution to the Pell Equation, then there are infinitely many solutions. A proof by Lagrange in 1768 determines all the solutions to the Pell Equation.

In this paper, we describe many important properties of finite and infinite continued fractions such as convergents which are obtained by keeping only the first k terms of a continued fraction. We will prove Dirichlet's Theorem on Diophantine Approximation and a theorem, proved by Euler and Lagrange, that every infinite simple continued fraction of an irrational number is periodic if and only if that irrational number is a quadratic irrational. We will show how the period of the periodic simple continued fraction of $\sqrt{d}$ can be used to determine the solutions to the Pell Equation $x^2 - dy^2 = \pm 1$ and prove the following theorem.

Theorem 1.1. *Let d be a positive integer that is not a perfect square. Let n be the period of the periodic simple continued fraction of $\sqrt{d}$. If n is even, then there are infinitely many positive solutions to $x^2 - dy^2 = 1$ and no solutions to $x^2 - dy^2 = -1$. If n is odd, then there are infinitely many solutions to $x^2 - dy^2 = \pm 1$.*

2. Finite Continued Fractions

Definition 2.1 (Finite Continued Fraction)**.** For any real numbers $a_0, a_1, \ldots, a_n$ where $a_1, \ldots, a_n$ are positive, we define a *finite continued fraction* to be of the form

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

for some nonnegative integer n. We call $a_1, \ldots, a_n$ the coefficients of the finite continued fraction. If $a_1, \ldots, a_n$ are integers, then we consider the finite continued fraction to be *simple*.

For simplicity, we will denote the finite continued fraction as $[a_0; a_1, a_2, \ldots a_n]$.

Example 2.2. We will find the finite continued fraction representing 56/31. From the Euclidean algorithm, we see that

$$56 = 1 \cdot 31 + 25, \tag{1}$$
$$31 = 1 \cdot 25 + 6, \tag{2}$$
$$25 = 4 \cdot 6 + 1. \tag{3}$$

If we divide (1) by 31, (2) by 25, and (3) by 6, we get

$$\frac{56}{31} = 1 + \frac{25}{31} = 1 + \frac{1}{\frac{31}{25}}, \tag{4}$$
$$\frac{31}{25} = 1 + \frac{6}{25} = 1 + \frac{1}{\frac{25}{6}}, \tag{5}$$
$$\frac{25}{6} = 4 + \frac{1}{6}. \tag{6}$$

Combining (4), (5), and (6) gives

$$\frac{56}{31} = 1 + \frac{1}{1 + \frac{1}{4 + \frac{1}{6}}}.$$

In this case, 56/31 is represented by the finite simple continued fraction $[1; 1, 4, 6]$.

We could also start with a finite simple continued fraction and find the number it represents.

Example 2.3. The finite simple continued fraction

$$1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{4}}}$$

can be written as

$$\begin{aligned} 1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{4}}} &= 1 + \frac{1}{2 + \frac{4}{13}} \\ &= 1 + \frac{13}{30} \\ &= \frac{43}{30} \end{aligned}$$

which is a rational number.

The next step would be to check if all rational numbers are represented by finite simple continued fractions and all finite simple continued fractions represent rational numbers.

Theorem 2.4. *Every finite simple continued fraction represents a rational number.*

Proof. We will use induction on n to show that the finite continued fraction $[a_0; a_1, a_2, \ldots, a_n]$ represents a rational number. If $n = 0$, we see that $[a_0] = a_0$ is rational. Assume for n that $[a_0; a_1, a_2, \ldots, a_n] = r/s$ for some integers r, s. By induction, for $n + 1$,

$$[a_0; a_1, a_2, \ldots, a_n, a_{n+1}] = a_0 + \frac{1}{[a_1; a_2, a_3, \ldots, a_n, a_{n+1}]} = a_0 + \frac{y}{x} = \frac{a_0 x + y}{x}$$

for some integers x, y such that $[a_1; a_2, a_3, \ldots, a_n, a_{n+1}] = x/y$. Therefore $[a_0; a_1, a_2, \ldots, a_n, a_{n+1}]$ is rational. □

Theorem 2.5. *Every rational number can be represented by a finite simple continued fraction.*

The proof of this theorem follows from the Euclidean Algorithm. We refer the reader to [1] for the proof. We can also look at properties of the first k terms of a continued fraction.

Definition 2.6. Let C_n be the continued fraction $[a_0; a_1, a_2, \dots, a_n]$ for some nonnegative integer n. Let k be a positive integer and $k \leq n$. The continued fraction $C_k = [a_0; a_1, a_2, \dots, a_k]$ of the first k terms of C_n is called the kth *convergent*.

Let us find the first few convergents of C_n:

$$C_0 = [a_0] = a_0,$$

$$C_1 = [a_0; a_1] = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1},$$

$$C_2 = [a_0; a_1, a_2] = a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1} = \frac{a_2(a_0 a_1 + 1) + a_0}{a_1 a_2 + 1},$$

$$C_3 = [a_0; a_1, a_2, a_3] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3}}} = \frac{a_0 a_1 a_2 a_3 + a_0 a_1 + a_0 a_3 + a_2 a_3 + 1}{a_1 a_2 a_3 + a_1 + a_3} = \frac{a_3(a_2(a_0 a_1 + 1) + a_0) + (a_0 a_1 + 1)}{a_3(a_1 a_2 + 1) + a_1}.$$

Notice that there is an interesting pattern in the numerators and denominators of the convergents. If we let $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, and $q_1 = a_1$, then $C_0 = p_0/q_0$, $C_1 = p_1/q_1$, $C_2 = p_2/q_2 = (a_2 p_1 + p_0)/(a_2 q_1 + q_0)$, and $C_3 = p_3/q_3 = (a_3 p_2 + p_1)/(a_3 q_2 + q_1)$. We can now extend this pattern to C_k.

Theorem 2.7. *Let $C_n = [a_0; a_1, a_2, \dots, a_n]$ be a finite continued fraction. Define $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, and $q_1 = a_1$. For any positive integer k with $k \leq n$, let $p_k = a_k p_{k-1} + p_{k-2}$ and $q_k = a_k q_{k-1} + q_{k-2}$. Then, the kth convergent $C_k = [a_0; a_1, a_2 \dots, a_k]$ is equal to p_k/q_k.*

Proof. We will use induction on k. If $k = 0$, then

$$C_0 = [a_0] = a_0 = \frac{a_0}{1} = \frac{p_0}{q_0},$$

and if $k = 1$, then

$$C_1 = [a_0; a_1] = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{p_1}{q_1}.$$

Assume for k that

$$C_k = [a_0; a_1, a_2, \dots, a_{k-1}, a_k] = \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}} = \frac{p_k}{q_k}.$$

By induction, for $k + 1$,

$$\begin{aligned} C_{k+1} = [a_0; a_1, a_2, \dots, a_{k-1}, a_k, a_{k+1}] &= \left[a_0; a_1, a_2, \dots, a_{k-1}, a_k + \frac{1}{a_{k+1}}\right] \\ &= \frac{(a_k + \frac{1}{a_{k+1}})p_{k-1} + p_{k-2}}{(a_k + \frac{1}{a_{k+1}})q_{k-1} + q_{k-2}} \\ &= \frac{a_{k+1}(a_k p_{k-1} + p_{k-2} + \frac{1}{a_{k+1}} p_{k-1})}{a_{k+1}(a_k q k - 1 + q_{k-2} + \frac{1}{a_{k+1}} q_{k-1})} \\ &= \frac{a_{k+1} p_k + p_{k-1}}{a_{k+1} q_k + q_{k-1}} \\ &= \frac{p_{k+1}}{q_{k+1}} \end{aligned}$$

as desired.

□

Example 2.8. From Example (2.2), we found that $56/31 = [1; 1, 4, 6]$. Therefore we get

$$\begin{aligned} p_0 &= 1, \\ p_1 &= 1 \cdot 1 + 1 = 2, \\ p_2 &= 4 \cdot 2 + 1 = 9, \\ p_3 &= 6 \cdot 9 + 2 = 56, \end{aligned}$$

and

$$\begin{aligned}
q_0 &= 1, \\
q_1 &= 1, \\
q_2 &= 4 \cdot 1 + 1 = 5, \\
q_3 &= 6 \cdot 5 + 1 = 31.
\end{aligned}$$

This implies the convergents are

$$\begin{aligned}
C_0 &= \frac{p_0}{q_0} = \frac{1}{1} = 1, \\
C_1 &= \frac{p_1}{q_1} = \frac{2}{1} = 2, \\
C_2 &= \frac{p_2}{q_2} = \frac{9}{5}, \\
C_3 &= \frac{p_3}{q_3} = \frac{56}{31}.
\end{aligned}$$

These convergents seem to alternate where the convergents with an even index are smaller than the convergents with odd index. We would want to find out how close successive convergents really are.

Theorem 2.9. *Let $C_n = [a_0; a_1, a_1, \ldots a_n]$ be a finite simple continued fraction and let $C_k = p_k/q_k$ be the kth convergent for any positive integer k with $k \leq n$. Then*

$$C_k - C_{k-1} = \frac{(-1)^{k-1}}{q_k q_{k-1}}. \tag{7}$$

To prove this theorem, we need the following lemma.

Lemma 2.10. *Let $C_n = [a_0; a_1, a_2, \ldots, a_n]$ and $C_k = p_k/q_k$ be the kth convergent for any positive integer k with $k \leq n$. Then*

$$p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}.$$

Proof. We will use induction on k. Based on the definition of p_k and q_k in Theorem 2.7, if $k = 1$, we see that

$$p_1 q_0 - p_0 q_1 = (a_0 a_1 + 1) \cdot 1 - a_0 a_1 = 1 = (-1)^0.$$

Assume for k that

$$p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}.$$

Then for $k+1$,

$$\begin{aligned}
p_{k+1} q_k - p_k q_{k+1} &= (a_k p_k + p_{k-1}) q_k - p_k (a_k q_k + q_{k-1}) \\
&= a_k p_k q_k + p_{k-1} q_k - a_k p_k q_k - p_k q_{k-1} \\
&= p_{k-1} q_k - p_k q_{k-1} \\
&= -(p_k q_{k-1} - p_{k-1} q_k) \\
&= (-1)(-1)^{k-1} \\
&= (-1)^k
\end{aligned}$$

as desired. □

Proof of Theorem 2.9. Notice that

$$C_k - C_{k-1} = \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} = \frac{p_k q_{k-1} - q_k p_{k-1}}{q_k q_{k-1}} = \frac{(-1)^{k-1}}{q_k q_{k-1}},$$

so (7) holds. □

Similarly, we have the following result for convergents with index two apart.

Corollary 2.11. *Let $C_n = [a_0; a_1, a_1, \ldots a_n]$ be a finite simple continued fraction and let $C_k = p_k/q_k$ be the kth convergent for any positive integer k with $k \leq n$. Then*

$$C_k - C_{k-2} = \frac{a_k(-1)^k}{q_k q_{k-2}}. \tag{8}$$

Proof. We see that

$$\begin{aligned}
C_k - C_{k-2} &= \frac{p_k}{q_k} - \frac{p_{k-2}}{q_{k-2}} \\
&= \frac{p_k q_{k-2} - p_{k-2} q_k}{q_k q_{k-2}} \\
&= \frac{(a_k p_{k-1} + p_{k-2}) q_{k-2} - p_{k-2}(a_k q_{k-1} + q_{k-2})}{q_k q_{k-2}} \\
&= \frac{a_k p_{k-1} q_{k-2} + p_{k-2} q_{k-2} - a_k p_{k-2} q_{k-1} - p_{k-2} q_{k-2}}{q_k q_{k-2}} \\
&= \frac{a_k (p_{k-1} q_{k-2} - p_{k-2} q_{k-1})}{q_k q_{k-2}} \\
&= \frac{a_k(-1)^{k-2}}{q_k q_{k-2}} \\
&= \frac{a_k(-1)^k}{q_k q_{k-2}},
\end{aligned}$$

so (8) holds. □

We can apply Theorem 2.9 to the convergents of the finite simple continued fraction of 56/31 in Example 2.8. We see that

$$\begin{aligned}
C_1 - C_0 &= 2 - 1 = 1 = \frac{(-1)^0}{(q_1 q_0)}, \\
C_2 - C_1 &= \frac{9}{5} - 2 = -\frac{1}{5} = \frac{(-1)^1}{5 \cdot 1} = \frac{(-1)^1}{(q_2 q_1)}, \\
C_3 - C_2 &= \frac{56}{31} - \frac{9}{5} = \frac{1}{155} = \frac{(-1)^2}{31 \cdot 5} = \frac{(-1)^2}{q_3 q_2}.
\end{aligned}$$

Also, from Corollary 2.11, we have

$$C_3 - C_1 = \frac{56}{31} - 2 = -\frac{6}{31} = \frac{6 \cdot (-1)^3}{31 \cdot 1} = \frac{a_k(-1)^k}{q_k q_{k-2}}$$

since $a_3 = 6$.

Corollary 2.12. *Let p_k and q_k be as defined in Theorem 2.7. Then p_k and q_k are relatively prime.*

Proof. Let $m = \gcd(p_k, q_k)$. This implies m divides $p_k q_{k-1}$ and $p_{k-1} q_k$. By Lemma 2.10, m must divide $(-1)^{k-1}$ so $m = 1$. □

As we saw in Example 2.8, convergents of the form C_{2z}, with an even index, are smaller than convergents of the form C_{2w+1}, with an odd index. Using Theorem 2.9 and Corollary 2.11, the order of the convergents can be determined.

Theorem 2.13. *Let $C_n = [a_0; a_1, a_1, \ldots a_n]$ be a finite simple continued fraction and let $C_k = p_k/q_k$ be the kth convergent for any positive integer k with $k \leq n$. Then for all nonnegative integers w, z with $2w + 1, 2z \leq n$, $C_{2w+1} > C_{2w+3}$, $C_{2z} < C_{2z+2}$, and $C_{2w+1} > C_{2z}$.*

Proof. By (8) of Corollary 2.11,

$$C_{2w+3} - C_{2w+1} = \frac{a_{2w+3}(-1)^{2w+3}}{q_{2w+3} q_{2w+1}}.$$

Since $2w+3$ is odd and a_{2w+3}, q_{2w+3}, and q_{2w+1} are positive, $C_{2w+3}-C_{2w+1}$ is negative so $C_{2w+1} > C_{2w+3}$. Also,

$$C_{2z+2} - C_{2z} = \frac{a_{2z+2}(-1)^{2z+2}}{q_{2z+2}q_{2z}}.$$

Since $2z+2$ is even and a_{2z+2}, q_{2z+2}, and q_{2z} are positive, $C_{2z+2}-C_{2z}$ is positive so $C_{2z} < C_{2z+2}$. To show that $C_{2w+1} > C_{2z}$, we first note that by (7) of Theorem 2.9,

$$C_{2w+1} - C_{2w} = \frac{(-1)^{2w}}{q_{2w+1}q_{2w}}.$$

Since $2w+1$ is even and q_{2w+1} and q_{2w} are positive, $C_{2w+1}-C_{2w}$ is positive so $C_{2w+1} > C_{2w}$. This implies

$$C_{2w+1} > C_{2w+1+2z} > C_{2w+2z} > C_{2z}.$$

□

3. Infinite Continued Fractions

Infinite Continued Fractions are of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots}}$$

where $a_0, a_1, a_2, \dots$ are real numbers and $a_1, a_2, \dots$ are positive. As with finite simple continued fractions, infinite continued fractions are *simple* if the coefficients $a_0, a_1, \dots$ are integers. In Theorem 2.13, we found that the sequence of even-index convergents is increasing, the sequence of odd-index convergents is decreasing, and all odd-index convergents are greater than all even-index convergents. If we kept increaseing the value of n, would these two sequences eventually converge? We will now prove that we can define any infinite simple continued fraction as the limit of the convergents of the first k terms.

Theorem 3.1. *Let $[a_0; a_1, a_2, \dots]$ be an infinite simple continued fraction and let $C_k = [a_0; a_1, a_2, \dots, a_k]$ for any positive integer k. Then there exists a real number α such that $[a_0; a_1, a_2, \dots]$ converges to α. In other words,*

$$\lim_{k\to\infty} C_k = \alpha.$$

Proof. For any nonnegative integer n, from Theorem 2.13, we have $C_{2w+1} > C_{2w+3}$, $C_{2z} < C_{2z+2}$, and $C_{2w+1} > C_{2z}$ for all nonnegative integers w, z with $2w+1, 2z \le n$. Since these inequalities hold for any n, we see that $C_{2w+1} > C_{2z}$ for all nonnegative integers w, z, so the sequence of convergents C_{2k+1} are bounded below by C_{2z} for any nonnegative integer z and the sequence of convergents C_{2k} are bounded above by C_{2w+1} for any nonnegative integer w. By the Monotone Convergence Theorem, both sequences have a limit, so

$$\lim_{k\to\infty} C_{2k+1} = \alpha_1$$

and

$$\lim_{k\to\infty} C_{2k} = \alpha_2.$$

We will now show that the limits are equal by proving that

$$\lim_{k\to\infty} (C_{2k+1} - C_{2k}) = 0,$$

since

$$\lim_{k\to\infty} (C_{2k+1} - C_{2k}) = \lim_{k\to\infty} C_{2k+1} - \lim_{k\to\infty} C_{2k}.$$

By (7) of Theorem 2.9,

$$C_{2k+1} - C_{2k} = \frac{(-1)^{2k}}{q_{2k+1}q_{2k}} = \frac{1}{q_{2k+1}q_{2k}}.$$

Since $q_k \ge 1$ for all nonnegative integers k, $q_k \ge q_{k-1} + q_{k-2} \ge q_{k-1} + 1 \ge k$ for all $k \ge 2$, so

$$0 < \frac{1}{q_{2k+1}q_{2k}} \le \frac{1}{(2k+1)(2k)}.$$

We know that

$$\lim_{k\to\infty} \frac{1}{(2k+1)(2k)} = 0,$$

so

$$\lim_{k\to\infty} (C_{2k+1} - C_{2k}) = \lim_{k\to\infty} \frac{1}{q_{2k+1}q_{2k}} = 0.$$

Therefore $\alpha = \alpha_1 = \alpha_2$ and

$$\lim_{k\to\infty} C_k = \alpha.$$

□

We proved earlier that finite simple continued fractions represent rational numbers. It turns out that infinite simple continued fractions represent irrational numbers and that the converse is also true.

Theorem 3.2. *Every infinite simple continued fraction* $[a_0; a_1, a_2, \ldots]$ *is irrational.*

Proof. Let $\alpha = [a_0; a_1, a_2, \ldots]$ and let C_k be the kth convergent of α for any nonnegative integer k. By Theorem 3.1, $C_{2k} < \alpha < C_{2k+1}$. This implies

$$0 < \alpha - C_{2k} < C_{2k+1} - C_{2k}$$

so

$$0 < \alpha - \frac{p_{2k}}{q_{2k}} < \frac{1}{q_{2k+1}q_{2k}}$$

Assume, for the sake of contradiction, that α is rational. Then we can write $\alpha = a/b$ for some integers a, b. If we multiply all three sides of the inequality by $q_{2k}b$, we get

$$0 < aq_{2k} - bp_{2k} < \frac{b}{q_{2k+1}}$$

Since q_{2k} and p_{2k} are integers, $aq_{2k} - bp_{2k}$ is always an integer greater than 0. But $q_{2k+1} \geq k$ for all nonnegative integers k, so we can choose a positive integer j greater than b such that $q_{2j+1} > b$ which would imply that $b/q_{2j+1} < 1$. This is a contradiction since there are no integers between 0 and 1. Therefore α must be irrational. □

Now that we have proved that any infinite simple continued fraction represents an irrational number, we would also like to show that any irrational number is always represented by an infinite simple continued fraction. We will describe an algorithm to construct an infinite simple continued fraction representing any irrational number α. First, expand α as follows:

$$\begin{aligned}
\alpha &= \lfloor\alpha\rfloor + \{\alpha\} \\
&= \lfloor\alpha\rfloor + \frac{1}{(1/\{\alpha\})} \\
&= \lfloor\alpha\rfloor + \frac{1}{\lfloor 1/\{\alpha\}\rfloor + \{1/\{\alpha\}\}} \\
&= \lfloor\alpha\rfloor + \cfrac{1}{\lfloor 1/\{\alpha\}\rfloor + \cfrac{1}{(1/\{1/\{\alpha\}\})}} \\
&= \lfloor\alpha\rfloor + \cfrac{1}{\lfloor 1/\{\alpha\}\rfloor + \cfrac{1}{\lfloor 1/\{1/\{\alpha\}\}\rfloor + \{1/\{1/\{\alpha\}\}}} \qquad (9)
\end{aligned}$$

Finding the first three coefficients of the expansion of α will help us find a recursive formula for the kth coefficient for any nonnegative integer k. From (9), the first three coefficients of the expansion of α are $\lfloor\alpha\rfloor$, $\lfloor 1/\{\alpha\}\rfloor$, and $\lfloor 1/\{1/\{\alpha\}\}\rfloor$. Let $\alpha_0 = \alpha$, $\alpha_1 = 1/\{\alpha_0\}$, and $\alpha_2 = 1/\{\alpha_1\}$. If $a_0 = \lfloor\alpha_0\rfloor$, $a_1 = \lfloor\alpha_1\rfloor$, and $a_2 = \lfloor\alpha_2\rfloor$, then a_0, a_1, and a_2 are integer coefficients of the expansion of α. This implies the kth coefficient is given by $a_k = \lfloor\alpha_k\rfloor$ where $\alpha_{k+1} = 1/\{\alpha_k\}$ and $\alpha_0 = \alpha$ for all nonnegative integers k. Since we want to show that $\alpha = [a_0; a_1, a_2, \ldots]$, by Theorem 3.1, we need to prove that as k approaches infinity, the finite simple continued fraction C_k equal to $[a_0; a_1, a_2, \ldots, a_k, \alpha_{k+1}]$ approaches α.

Theorem 3.3. *Let α be any irrational number. Define the integers $a_0, a_1, a_2, \ldots$ recursively as follows. Let $\alpha_0 = \alpha$, $\alpha_{k+1} = 1/\{\alpha_k\}$, and $a_k = \lfloor \alpha_k \rfloor$ for all positive integers k. If $C_k = [a_0; a_1, a_2, \ldots, a_k, \alpha_{k+1}] = \alpha$, then*

$$\lim_{k\to\infty} C_k = \alpha.$$

Proof. We can instead show that

$$\alpha - \lim_{k\to\infty} C_k = 0.$$

By Theorem 2.13, if k is odd, then $C_{k+1} < \alpha < C_k$. Subtracting C_k from all sides of the inequality, we get $C_{k+1} - C_k < \alpha - C_k < 0$, and by (7) of Theorem 2.9,

$$\alpha - C_k < C_k - C_{k+1} = \frac{1}{q_{k+1}q_k}. \tag{10}$$

If k is even, then $C_k < \alpha < C_{k+1}$, and subtracting C_k from all sides of the inequality gives $0 < \alpha - C_k < C_{k+1} - C_k$. By (7) of Theorem 2.9,

$$\alpha - C_k < C_{k+1} - C_k = -\frac{1}{q_{k+1}q_k}. \tag{11}$$

Based on (10) and (11), we see that

$$|\alpha - C_k| < \frac{1}{q_{k+1}q_k}. \tag{12}$$

As we found in the proof of Theorem 3.1,

$$\lim_{k\to\infty} \frac{1}{q_{k+1}q_k} = 0,$$

so

$$\lim_{k\to\infty} |C_k - \alpha| = 0.$$

Hence

$$\lim_{k\to\infty} C_k = \alpha.$$

□

Therefore, we can indeed write α as the infinite simple continued fraction $[a_0; a_1, a_2, \ldots]$. It immediately follows that this representation must be unique. Suppose the infinite simple continued fraction $[b_0; b_1, b_2, \ldots]$ is also equal to α. Then, using our definition of α, we can see by induction on k that $b_k = \lfloor \alpha_k \rfloor = a_k$ for all nonnegative integers k.

A notable result, proved by Dirichlet, states that there are infinitely many rational numbers that are good approximations of any irrational number α.

Theorem 3.4 (Dirichlet's Theorem on Diophantine Approximation)**.** *Let α be an irrational number. Then there are infinitely many rational numbers r/s such that*

$$\left|\alpha - \frac{r}{s}\right| < \frac{1}{s^2}.$$

Proof. If p_k/q_k is the kth convergent of the infinite simple continued fraction of α for any nonnegative integer k, then by (12),

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_{k+1}q_k}.$$

Since $q_{k+1} > q_k$, we see that $1/q_{k+1}q_k < 1/q_k^2$. Therefore

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^2}.$$

There are infinitely many convergents p_k/q_k and hence this theorem holds. □

We now show that the convergents p_k/q_k of the infinite simple continued fraction of α are the best approximations of α compared to any other rational number r/s where $s < q_k$.

Theorem 3.5. *Let α be an irrational number and let p_k/q_k be the kth convergent of the infinite simple continued fraction representing α for any nonnegative integer k. If r/s is a rational number with $s > 0$ such that*

$$\left|\alpha - \frac{r}{s}\right| < \left|\alpha - \frac{p_k}{q_k}\right|$$

then $s > q_k$.

To prove this theorem, the following lemma is needed.

Lemma 3.6. *Let α be an irrational number and let p_k/q_k be the kth convergent of the infinite simple continued fraction representing α for any nonnegative integer k. If r and s are integers such that*

$$|s\alpha - r| < |q_k\alpha - p_k|,$$

then $s \geq q_{k+1}$.

Proof. Assume, for the sake of contradiction, that $1 \leq s < q_{k+1}$. There exists integers x, y that are solutions to the system of equations

$$p_k x + p_{k+1} y = r \tag{13}$$

$$q_k x + q_{k+1} y = s. \tag{14}$$

To solve for y, we first multiply (13) by q_k and (14) by p_k to get

$$p_k q_k x + p_{k+1} q_k y = r q_k \tag{15}$$

$$q_k p_k x + q_{k+1} p_k y = s p_k, \tag{16}$$

and then subtract (16) from (15), which implies

$$(p_{k+1} q_k - p_k q_{k+1}) y = r q_k - s p_k.$$

By Lemma 2.10, we see that

$$y = (-1)^k (r q_k - s p_k). \tag{17}$$

Similarly, to solve for x, we multiply (13) by q_{k+1} and (14) by p_{k+1} to get

$$p_k q_{k+1} x + p_{k+1} q_{k+1} y = r q_{k+1} \tag{18}$$

$$q_k p_{k+1} x + q_{k+1} p_{k+1} y = s p_{k+1}, \tag{19}$$

and then subtract (19) from (18) which gives

$$(p_k q_{k+1} - q_k p_{k+1}) x = r q_{k+1} - s p_{k+1}.$$

By Lemma 2.10, we see that

$$x = (-1)^k (r q_{k+1} - s p_{k+1}). \tag{20}$$

We will show that any value of x and y leads to a contradiction. First, if $x = 0$, then by (20), $s p_{k+1} = r q_{k+1}$. Since p_{k+1} and q_{k+1} are relatively prime by Corollary 2.12, $q_{k+1} \leq s$. But this contradicts our assumption, so $x \neq 0$. Next, if $y = 0$, then from (13) and (14), we see that $r = p_k x$ and $s = q_k x$, which implies

$$|s\alpha - r| = |q_k x \alpha - p_k x| = |x||q_k\alpha - p_k| \geq |q_k\alpha - p_k|$$

since $|x| \geq 1$. But this contradicts the statement of this theorem, so $y \neq 0$. The remaining case is where x, y are nonzero. We will first show that $y < 0$ implies $x > 0$, and $y > 0$ implies $x < 0$. From (14), we know that

$$x = \frac{s - q_{k+1} y}{q_k},$$

so if $y < 0$, since $s, q_k, q_{k+1} > 0$, then $x > 0$. If $y > 0$, then $q_{k+1} y \geq q_{k+1} > s$, which implies $q_k x = s - q_{k+1} y < 0$, so $x < 0$. Next, based on (13) and (14),

$$\begin{aligned} |s\alpha - r| &= |(q_k x + q_{k+1} y)\alpha - (p_k x + p_{k+1} y)| \\ &= |(q_k\alpha - p_k)x + (q_{k+1}\alpha - p_{k+1})y|. \end{aligned}$$

Note that by Theorem 2.13, either

$$p_k/q_k < \alpha < p_{k+1}/q_{k+1} \tag{21}$$

or

$$p_{k+1}/q_{k+1} < \alpha < p_k/q_k. \tag{22}$$

If (21) is true, then $0 < q_k\alpha - p_k$ and $q_{k+1}\alpha - p_{k+1} < 0$, and if (22) is true, then $q_k\alpha - p_k > 0$ and $0 < q_{k+1}\alpha - p_{k+1}$. This implies $(q_k\alpha - p_k)x$ and $(q_{k+1}\alpha - p_{k+1})y$ are either both positive or negative. Therefore, we get

$$\begin{aligned} |s\alpha - r| &= |x||q_k\alpha - p_k| + |y||q_{k+1}\alpha - p_{k+1}| \\ &\geq |x||q_k\alpha - p_k \\ &\geq |q_k\alpha - p_k|, \end{aligned}$$

which contradicts the statement of the theorem. Hence, $s \geq q_{k+1}$. □

Now we can prove Theorem 3.5.

Proof of Theorem 3.5. Assume, for the sake of contradiction, that $s \leq q_k$. Multiplying this inequality by $|\alpha - r/s| < |\alpha - p_k/q_k|$, we obtain

$$s\left|\alpha - \frac{r}{s}\right| < q_k\left|\alpha - \frac{p_k}{q_k}\right|,$$

which is equivalent to

$$|s\alpha - r| < |q_k\alpha - p_k|.$$

But by Lemma 3.6, this inequality only holds for $s \geq q_{k+1}$, so our original assumption must be incorrect. Therefore $s > q_k$. □

We will show that the convergents of the infinite simple continued fraction of α are the only best approximations of α.

Theorem 3.7. *Let α be an irrational number and let p_k/q_k be the kth convergent of the infinite simple continued fraction representing α for any nonnegative integer k. If r, s are integers with $\gcd(r, s) = 1$ and $s > 0$ such that*

$$\left|\alpha - \frac{r}{s}\right| < \frac{1}{2s^2}$$

then $r/s = p_k/q_k$ for some nonnegative integer k.

Proof. Assume, for the sake of contradiction, that r/s is not a convergent of the infinite simple continued fraction representing α. Then there exists convergents p_k/q_k and p_{k+1}/q_{k+1} such that $q_k \leq s < q_{k+1}$. By Lemma 3.6, $|q_k\alpha - p_k| \leq |s\alpha - r|$. Since $|s\alpha - r| = s|\alpha - r/s| < 1/2s$, we see that

$$|q_k\alpha - p_k| < \frac{1}{2s}.$$

If we divide the inequality by q_k, we get

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{2sq_k}.$$

Since $r/s \neq p_k/q_k$, then $sp_k - rq_k \neq 0$, so $1 \leq |sp_k - rq_k|$. This implies that

$$\begin{aligned} \frac{1}{sq_k} &\leq \frac{|sp_k - rq_k|}{sq_k} \\ &= \left|\frac{p_k}{q_k} - \frac{r}{s}\right| \\ &\leq \left|\alpha - \frac{p_k}{q_k}\right| + \left|\alpha - \frac{r}{s}\right| \\ &< \frac{1}{2sq_k} + \frac{1}{2s^2}. \end{aligned}$$

Hence,

$$\frac{1}{2sq_k} < \frac{1}{2s^2},$$

so $q_k > s$ which contradicts our assumption that $q_k \leq s$. Therefore r/s must be a convergent of the infinite simple continued fraction of α. $\square$

4. Periodic Continued Fractions

Let's find the infinite simple continued fraction of $\sqrt{3}$. We can see that

$$\sqrt{3} = 1 + (\sqrt{3} - 1) = 1 + \frac{1}{\frac{\sqrt{3}+1}{2}} = 1 + \frac{1}{1 + \frac{\sqrt{3}-1}{2}} = 1 + \frac{1}{1 + \frac{1}{1+\sqrt{3}}} = 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2 + \cdots}}}}.$$

There appears to be a repeating pattern in the coefficients of this infinite simple continued fraction giving $\sqrt{3} = [1; 1, 2, 1, 2, \ldots]$. A continued fraction with this property is called a *periodic continued fraction.* We now give the precise definition.

Definition 4.1 (Periodic Continued Fractions)**.** Let $\alpha = [a_0; a_1, a_2, \ldots]$ be an infinite simple continued fraction. If there exists nonnegative integers k, N such that whenever $n \geq N$, $a_n = a_{n+k}$, then α is a periodic simple continued fraction. We write

$$\alpha = [a_0; , a_1, \ldots, a_{N-1}, \overline{a_N, a_{N+1}, \ldots, a_{N+k-1}}],$$

where $a_N, a_{N+1}, \ldots, a_{N+k-1}$ are the repeating coefficients and k is the period.

Periodic simple continued fractions are special types of infinite simple continued fractions, so they only represent irrational numbers. We can express the irrational number $\sqrt{3}$ as a periodic simple continued fraction, but can we do the same for $3 - \sqrt{2}$, e, or π? Later in this section, we will state and prove a very important result, proved by Euler and Lagrange, that the infinite simple continued fraction of every irrational number is periodic if and only if that irrational number is a quadratic irrational. Let us start by defining a quadratic irrational.

Definition 4.2. A *quadratic irrational* α is an irrational number that is a root of a quadratic equation $a\alpha^2 + b\alpha + c = 0$ for some integers a, b, c where $a \neq 0$ (i.e., α is an irrational number of the form $(x + \sqrt{d})/y$ for some integers x, y, d with $d > 0$, where d is not a perfect square, and $y \neq 0$).

Definition 4.3. Let α be a quadratic irrational equal to $(a + \sqrt{b})/c$. Then the *conjugate* of α is α', which is equal to $\alpha' = (a - \sqrt{b})/c$.

If α is a root of the polynomial $ax^2 + bx + c = 0$ where α is equal to $(-b + \sqrt{b^2 - 4ac})/(2a)$, then based on the definition above, the conjugate of α is $\alpha' = (-b - \sqrt{b^2 - 4ac})/(2a)$. The following lemma is necessary for finding conjugates of larger expressions.

Lemma 4.4. *If α and β are rational or quadratic irrationals with $\alpha = (a + b\sqrt{d})/c$ and $\beta = (x + y\sqrt{d})/z$, then*

(i) $(\alpha + \beta)' = \alpha' + \beta'$
(ii) $(\alpha - \beta)' = \alpha' - \beta'$
(iii) $(\alpha\beta)' = \alpha'\beta'$
(iv) $(\frac{\alpha}{\beta})' = \frac{\alpha'}{\beta'}$.

Proof. Note that $\alpha' = (a - b\sqrt{d})/c$ and $\beta' = (x - y\sqrt{d})/z$.
(i) We see that

$$\begin{aligned}
(\alpha+\beta)' &= \left(\frac{a+b\sqrt{d}}{c} + \frac{x+y\sqrt{d}}{z}\right)' \\
&= \left(\frac{az+bz\sqrt{d}+xc+yc\sqrt{d}}{cz}\right)' \\
&= \left(\frac{(az+xc)+(bz+yc)\sqrt{d}}{cz}\right)' \\
&= \frac{(az+xc)-(bz+yc)\sqrt{d}}{cz} \\
&= \frac{a-b\sqrt{d}}{c} + \frac{x-y\sqrt{d}}{z} \\
&= \alpha' + \beta'.
\end{aligned}$$

(ii) Similarly,

$$\begin{aligned}
(\alpha-\beta)' &= \left(\frac{a+b\sqrt{d}}{c} - \frac{x+y\sqrt{d}}{z}\right)' \\
&= \left(\frac{az+bz\sqrt{d}-xc-yc\sqrt{d}}{cz}\right)' \\
&= \left(\frac{(az-xc)+(bz-yc)\sqrt{d}}{cz}\right)' \\
&= \frac{(az-xc)-(bz-yc)\sqrt{d}}{cz} \\
&= \frac{a-b\sqrt{d}}{c} - \frac{x-y\sqrt{d}}{z} \\
&= \alpha' - \beta'.
\end{aligned}$$

(iii) Notice that

$$\begin{aligned}
(\alpha\cdot\beta)' &= \left(\frac{a+b\sqrt{d}}{c} \cdot \frac{x+y\sqrt{d}}{z}\right)' \\
&= \left(\frac{(ax+byd)+(ay+bx)\sqrt{d}}{cz}\right)' \\
&= \frac{(ax+byd)-(ay+bx)\sqrt{d}}{cz} \\
&= \frac{ax-ay\sqrt{d}-bx\sqrt{d}+byd}{cz} \\
&= \left(\frac{a-b\sqrt{d}}{c}\right)\left(\frac{x-y\sqrt{d}}{z}\right) \\
&= \alpha'\beta'.
\end{aligned}$$

(iv) Observe that

$$\begin{aligned}
\left(\frac{\alpha}{\beta}\right)' &= \left(\frac{a+b\sqrt{d}}{c}\cdot\frac{z}{x+y\sqrt{d}}\right)' \\
&= \left(\frac{z(a+b\sqrt{d})(x-y\sqrt{d})}{c(x+y\sqrt{d})(x-y\sqrt{d})}\right)' \\
&= \left(\frac{(axz-byzd)+(bxz-ayz)\sqrt{d}}{cx^2-cdy^2}\right)' \\
&= \frac{(axz-byzd)-(bxz-ayz)\sqrt{d}}{c(x^2-dy^2)} \\
&= \frac{z(ax+ay\sqrt{d}-bx\sqrt{d}-byd)}{c(x^2-dy^2)} \\
&= \frac{z(a-b\sqrt{d})(x+y\sqrt{d}}{c(x+y\sqrt{d})(x-y\sqrt{d}} \\
&= \left(\frac{a-b\sqrt{d}}{c}\right)\left(\frac{z}{x-y\sqrt{d}}\right) \\
&= \frac{\alpha'}{\beta'}.
\end{aligned}$$

□

In the previous section, we showed an algorithm for writing any irrational number as an infinite simple continued fraction. Using the definition of quadratic irrationals and the following lemma, we can define an algorithm for writing any quadratic irrational as an infinite simple continued fraction.

Lemma 4.5. *Let α be a quadratic irrational. Then there exists integers P, Q, d with $Q \neq 0$ and $d > 0$ where d is not a perfect square such that*

$$\alpha = \frac{P+\sqrt{d}}{Q},$$

and $Q|(d-P^2)$.

Proof. We can write α as

$$\alpha = \frac{a+\sqrt{b}}{c},$$

for some integers a, b, c with $b > 0, c \neq 0$. Multiplying this equation by $|c|$ in the numerator and denominator gives

$$\alpha = \frac{a|c|+|c|\sqrt{b}}{c|c|} = \frac{a|c|+\sqrt{bc^2}}{c|c|}.$$

Let $P = a|c|$, $Q = c|c|$, and $d = bc^2$. Then P, Q, d satisfy the conditions of this lemma since $Q \neq 0$, $d > 0$, d is not a perfect square, and $Q|(d-P^2)$ as $d - P^2 = bc^2 - ac^2 = c^2(b-a) = \pm Q(b-a)$. □

We now state and prove the algorithm.

Theorem 4.6. *Let α be a quadratic irrational. We can write α as*

$$\alpha = \frac{P_0+\sqrt{d}}{Q_0},$$

for some integers P_0, Q_0, d with $Q_0 \neq 0$ and $d > 0$ where d is not a perfect square and $Q_0 | (d - P_0^2)$. For all nonnegative integers k, define

$$
\begin{aligned}
\alpha_k &= \frac{P_k + \sqrt{d}}{Q_k}, \\
a_k &= \lfloor \alpha_k \rfloor, \\
P_{k+1} &= a_k Q_k - P_k, \\
Q_{k+1} &= \frac{d - P_{k+1}^2}{Q_k},
\end{aligned}
\tag{23}
$$

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

Proof. We will use induction on k to prove that P_k and Q_k are integers with $Q_k \neq 0$ and $Q_k | (d - P_k^2)$. For $k = 0$, we know that P_0, Q_0 are integers with $Q_0 \neq 0$ and $Q_0 | (d - P_0^2)$. Assume for k that P_k, Q_k are integers with $Q_k \neq 0$ and $Q_0 | (d - P_0^2)$. By induction, for $k+1$, since a_k, P_k, Q_k are integers, P_{k+1} is an integer. Also

$$
\begin{aligned}
Q_{k+1} &= \frac{d - P_{k+1}^2}{Q_k} \\
&= \frac{d - (a_k Q_k - P_k)^2}{Q_k} \\
&= \frac{d - (a_k Q_k)^2 + 2 a_k Q_k P_k - P_k^2}{Q_k} \\
&= \frac{d - P_k^2}{Q_k} + 2 a_k P_k - a_k^2 Q_k.
\end{aligned}
$$

Since Q_k divides $d - P_k^2$ and $2a_k P_k - a_k^2 Q_k$ is an integer, Q_{k+1} is an integer. Next, we know that $Q_k \neq 0$ and d is not a perfect square, so $d \neq P^2$ which implies that $(d - P_{k+1}^2)/Q_k$ and $Q_{k+1} \neq 0$. Finally, multiplying (23) on both sides by Q_k / Q_{k+1}, we get

$$
Q_k = \frac{d - P_{k+1}^2}{Q_{k+1}},
$$

and hence $Q_{k+1} | (d - P_{k+1}^2)$. Now we need to check that $[a_0; a_1, a_2, \ldots]$ is the infinite simple continued fraction representing α. We know that by Theorem 3.3, if

$$
\alpha_{k+1} = \frac{1}{\{\alpha_k\}} = \frac{1}{\alpha_k - a_k}
$$

for all nonnegative integers k, then $[a_0; a_1, a_2, \ldots]$ is the infinite simple continued fraction representing α. Since

$$\begin{aligned}
\alpha_k - a_k &= \frac{P_k + \sqrt{d}}{Q_k} - a_k \\
&= \frac{\sqrt{d} - a_k Q_k + P_k}{Q_k} \\
&= \frac{\sqrt{d} - (a_k Q_k - P_k)}{Q_k} \\
&= \frac{\sqrt{d} - P_{k+1}}{Q_k} \\
&= \frac{(\sqrt{d} - P_{k+1})(\sqrt{d} + P_{k+1})}{Q_k(\sqrt{d} + P_{k+1})} \\
&= \frac{d - P_{k+1}^2}{Q_k(\sqrt{d} + P_{k+1})} \\
&= \frac{Q_{k+1}}{\sqrt{d} + P_{k+1}} \\
&= \frac{1}{\alpha_{k+1}}.
\end{aligned}$$

Therefore $\alpha_{k+1} = 1/(\alpha_k - a_k)$, so $[a_0; a_1, a_2, \ldots]$ is the infinite simple continued fraction representing α. □

This algorithm will help us prove that the infinite simple continued fraction of a quadratic irrational is periodic, which is one direction of the following important result.

Theorem 4.7 (Euler, Lagrange)**.** *The infinite simple continued fraction representing an irrational number α is periodic if and only if α is a quadratic irrational.*

Proof. We will first prove that if the infinite simple continued fraction representing α is periodic, then α is a quadratic irrational. Since α is periodic, we can write α as

$$\alpha = [a_0; a_1, a_2, \ldots, a_{N-1}, \overline{a_N, \ldots, a_{N+k}}].$$

Let

$$\beta = [\overline{a_N, \ldots, a_{N+k}}].$$

This implies

$$\beta = [a_N, \ldots, a_{N+k}, \beta] \tag{24}$$

and

$$\alpha = [a_0; a_1, a_2, \ldots, a_{N-1}, \beta]. \tag{25}$$

Let p_k/q_k and p_{k-1}/q_{k-1} be convergents of the simple continued fraction $[a_N, \ldots, a_{N+k}]$ where $p_0/q_0 = [a_N]$. By (24),

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

so

$$q_k \beta^2 + (q_{k-1} - p_k)\beta - p_{k-1} = 0.$$

and hence β is a quadratic irrational. Similarly, let x_{N-1}/y_{N-1} and x_{N-2}/y_{N-2} be convergents of the simple continued fraction $[a_0; a_1, a_2, \ldots, a_{N-1}]$ where $x_0/y_0 = [a_0]$. By (25),

$$\alpha = \frac{\beta x_{N-1} + x_{N-2}}{\beta y_{N-1} + y_{N-2}}.$$

Rationalizing the denominator, we find that α is a quadratic irrational.

We now prove the other direction of the theorem. We will give the motivation behind the proof that if α is a quadratic irrational, then the infinite simple continued fraction representing α is periodic. We need to show that there exists nonnegative integers i, j with $i < j$ such that for all nonnegative integers m, $a_{i+m} = a_{j+m}$. If this holds, then by Theorem 4.6, $\alpha_{i+m} = \alpha_{j+m}$. Alternatively, we can show that $P_{i+m} = P_{j+m}$ and $Q_{i+m} = Q_{j+m}$ since α_k can be written as

$$\alpha_k = \frac{P_k + \sqrt{d}}{Q_k},$$

where P_k, Q_k are as defined in Theorem 4.6. This can be proved by showing that there exists an integer N such that for all $k \geq N$, P_k and Q_k are bounded above and below.

Since α is a quadratic irrational, by Lemma 4.5, we can write α as

$$\alpha = \frac{P_0 + \sqrt{d}}{Q_0},$$

for some $Q_0 \neq 0$, $d > 0$ where d is not a perfect square, and $Q_0 | (d - Pk + 1^2)$. Using the algorithm in Theorem 4.6, we see that $\alpha = [a_0; a_1, a_2 \ldots]$. Since $\alpha = [a_0; a_1, a_2, \ldots, a_{k-1}, \alpha_k]$, then

$$\alpha = \frac{\alpha_k p_{k-1} + p_{k-2}}{\alpha_k q_{k-1} + q_{k-2}}.$$

From Lemma 4.4, the conjugate of α is

$$\alpha' = \left(\frac{\alpha_k p_{k-1} + p_{k-2}}{\alpha_k q_{k-1} + q_{k-2}} \right)' = \frac{\alpha'_k p_{k-1} + p_{k-2}}{\alpha'_k q_{k-1} + q_{k-2}}.$$

To solve for α'_k, note that

$$\alpha'(\alpha'_k q_{k-1} + q_{k-2}) - (\alpha'_k p_{k-1} + p_{k-2}) = 0,$$

which implies that

$$\alpha'_k = \frac{p_{k-2} - \alpha' q_{k-2}}{\alpha' q_{k-1} - p_{k-1}} = \left(-\frac{q_{k-2}}{q_{k-1}} \right) \left(\frac{\alpha' - (p_{k-2}/q_{k-2})}{\alpha' - (p_{k-1}/q_{k-1})} \right).$$

From Theorem 3.1, we know that p_{k-2}/q_{k-2} and p_{k-1}/q_{k-1} converges to α as k approaches infinity, and

$$\left(\frac{\alpha' - (p_{k-2}/q_{k-2})}{\alpha' - (p_{k-1}/q_{k-1})} \right)$$

converges to 1. Therefore, there must exist a nonnegative integer N such that whenever $k \geq N$, then α'_k is negative. By Theorem 3.3, for all $k > 1$, α_k is positive, so for all $k \geq N$,

$$\alpha_k - \alpha'_k = \frac{P_k + \sqrt{d}}{Q_k} - \frac{P_k - \sqrt{d}}{Q_k} = \frac{2\sqrt{d}}{Q_k} > 0,$$

and hence Q_k must be positive. From (23) of Theorem 4.6, we see that $Q_k Q_{k+1} = d - P_{k+1}^2$, so for all $k \leq N$,

$$Q_k \leq Q_k Q_{k+1} = d - P_{k+1}^2 \leq d$$

and $P_{k+1}^2 = d - Q_k Q_{k+1} \leq d$. This implies

$$-\sqrt{d} \leq P_{k+1} \leq \sqrt{d}$$

and

$$0 \leq Q_k \leq d$$

for all $k \geq N$.

Since there are finitely many integers P_k, Q_k satisfying the inequalities above and infinitely many integers k such that $k \geq N$, there must exist integers i, j with $i < j$ such that $P_i = P_j$ and $Q_i = Q_j$. Hence $\alpha_i = \alpha_j$ which implies $a_{i+m} = a_{j+m}$ for all nonnegative integers m. Therefore, we see that α is represented by the periodic simple continued fraction

$$\alpha = [a_0; a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_{j-1}, \ldots] = [a_0; a_1, a_2, \ldots, a_{i-1}, \overline{a_i, a_{i+1}, \ldots, a_{j-1}}]$$

□

There are also quadratic irrationals that are represented by periodic simple continued fractions of the form $[\overline{a_0; a_1, a_2, \ldots, a_n}]$ where n is the period and there are no terms before the repeating part. For example,

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{1}{(1+\sqrt{5})/2} = 1 + \cfrac{1}{1 + \cfrac{1}{(1+\sqrt{5})/2}} = 1 + \cfrac{1}{1 + \cfrac{1}{1+\ldots}},$$

so the periodic simple continued fraction of $(1+\sqrt{5})/2$ is $[1; 1, 1, 1, \ldots]$ or $[\overline{1}]$. We called such continued fractions *purely periodic*.

Definition 4.8. Let $\alpha = [a_0; a_1, a_2, \ldots]$ be a periodic simple continued fraction. Then α is *purely periodic* if there exists an integer n such that $a_k = a_{k+n}$ for all nonnegative integers k. In that case, we can write α as

$$[a_0; a_1, a_2, \ldots] = [\overline{a_0; a_1, a_2, \ldots, a_{n-1}}]$$

We call the quadratic irrational α *reduced* if $\alpha > 1$ and $-1 < \alpha' < 0$. As an example, if α is equal to $1+\sqrt{3}$, then α' is equal to $1-\sqrt{3}$. We can see that α is reduced since $1+\sqrt{3} > 1$ and $-1 < 1-\sqrt{3} < 0$.

Theorem 4.9. *Let α be a quadratic irrational. Then the periodic simple continued fraction of α is purely periodic if and only if α is reduced. If α is reduced and $\alpha = [\overline{a_0; a_1, a_2, \ldots, a_{n-1}}]$, then the periodic simple continued fraction of $-1/\alpha'$ is $[\overline{a_{n-1}; a_{n-2}, \ldots, a_0}]$.*

Proof. Assume that α is reduced. We will prove that the periodic simple continued fraction of α is purely periodic. By Theorem 3.3, we know that $\alpha_0 = \alpha$, $a_k = \lfloor \alpha_k \rfloor$, and $\alpha_{k+1} = 1/\alpha_k = 1/(\alpha_k - a_k)$ for all nonnegative integers k. This implies that

$$1/\alpha_{k+1} = (\alpha_k - a_k) \tag{26}$$

and by Lemma 4.4,

$$1/\alpha'_{k+1} = (\alpha'_k - a_k).$$

We will prove by induction on k, that

$$-1 < \alpha'_k < 0. \tag{27}$$

If $k = 0$, then by the definition of reduced, $-1 < \alpha'_0 = \alpha' < 0$. Assume for k that $-1 < \alpha'_k < 0$. Note that $a_k \geq 1$ for all nonnegative integers k since $\alpha > 1$ and a_k for all positive integers k is positive by definition. By induction,

$$\frac{1}{\alpha'_{k+1}} = \alpha'_k - a_k < -1,$$

which implies that $-1 < 1/\alpha'_{k+1} < 0$ as desired. We know that

$$\alpha'_k = a_k + \frac{1}{\alpha'_{k+1}},$$

so by (27), we see that

$$-1 < \alpha'_k = a_k + \frac{1}{\alpha'_{k+1}} < 0.$$

This implies that

$$-1 - \frac{1}{\alpha'_{k+1}} < a_k < -\frac{1}{\alpha'_{k+1}},$$

so $a_k = \lfloor -1/\alpha'_{k+1} \rfloor$. Since α is a quadratic irrational, α can be represented by a periodic simple continued fraction. By the proof of Theorem 4.7, there exists nonnegative integers i, j with $i < j$ such that for all nonnegative integers m, $a_{i+m} = a_{j+m}$. This implies that $\alpha_{i+m} = \alpha_{j+m}$, so $1/\alpha'_{i+m} = 1/\alpha'_{j+m}$ and $-1/\alpha'_{i+m} = -1/\alpha'_{j+m}$. Since $a_{(i+m)-1} = \lfloor -1/\alpha'_{i+m} \rfloor$ and $a_{(j+m)-1} = \lfloor -1/\alpha'_{j+m} \rfloor$, then $a_{(i+m)-1} = a_{(j+m)-1}$. Similarly, we can see that $a_{(i+m)-2} = a_{(j+m)-2}$. If we continue this process, we will find that $a_0 = a_{j-i}$. Therefore we can write α as the purely periodic simple continued fraction

$$\alpha = [\overline{a_0; a_1, \ldots, a_{j-i-1}}].$$

We refer to [1] for a proof of the other direction that if the periodic simple continued fraction of α is purely periodic, then α is reduced. □

The following theorem is useful for finding the solutions to the Pell Equation.

Theorem 4.10. *Let d be a positive integer that is not a perfect square. For all nonnegative integers k, define $\alpha_0 = \sqrt{d}$ and*

$$\begin{aligned}\alpha_k &= \frac{P_k + \sqrt{d}}{Q_k},\\ a_k &= \lfloor \alpha_k \rfloor,\\ P_{k+1} &= a_k Q_k - P_k,\\ Q_{k+1} &= \frac{d - P_{k+1}^2}{Q_k}.\end{aligned}$$

Let p_k/q_k be the kth convergent of the periodic simple continued fraction of $\sqrt{d}$. Then,

$$p_k^2 - dq_k^2 = (-1)^{k-1} Q_{k+1}.$$

Proof. Since $\sqrt{d} = \alpha_0 = [a_0; a_1, a_2, \ldots, a_k, \alpha_{k+1}]$, we see that

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

As defined in the statement of the theorem, $\alpha_{k+1} = (P_{k+1} + \sqrt{d})/Q_{k+1}$, so

$$\sqrt{d} = \frac{(P_{k+1} + \sqrt{d}) p_k + Q_{k+1} p_{k-1}}{(P_{k+1} + \sqrt{d}) q_k + Q_{k+1} q_{k-1}},$$

which implies

$$((P_{k+1} + \sqrt{d}) q_k + Q_{k+1} q_{k-1}) \sqrt{d} = (P_{k+1} + \sqrt{d}) p_k + Q_{k+1} p_{k-1},$$

and therefore

$$dq_k + (P_{k+1} q_k + Q_{k+1} q_{k-1}) \sqrt{d} = (P_{k+1} p_k + Q_{k+1} p_{k-1}) + p_k \sqrt{d}.$$

Note that if r, s, t, u are rational numbers and $r + s\sqrt{d} = t + u\sqrt{d}$, then $r = t$ and $s = u$. Since q_k, p_k, Q_k, P_k are rational numbers,

$$dq_k = P_{k+1} p_k + Q_{k+1} p_{k-1} \tag{28}$$

and

$$p_k = P_{k+1} q_k + Q_{k+1} q_{k-1}. \tag{29}$$

If we multiply (28) by q_k and (29) by p_k, we get

$$dq_k^2 = P_{k+1} p_k q_k + Q_{k+1} p_{k-1} q_k \tag{30}$$

and

$$p_k^2 = P_{k+1} q_k p_k + Q_{k+1} q_{k-1} p_k. \tag{31}$$

Subtracting (30) from (31) gives

$$\begin{aligned}p_k^2 - dq_k^2 &= (P_{k+1} q_k p_k + Q_{k+1} q_{k-1} p_k) - (P_{k+1} p_k q_k + Q_{k+1} p_{k-1} q_k)\\ &= Q_{k+1} q_{k-1} p_k - Q_{k+1} p_{k-1} q_k\\ &= (q_{k-1} p_k - p_{k-1} q_k) Q_{k+1}\\ &= (-1)^{k-1} Q_{k+1}\end{aligned}$$

as desired. □

5. Pell Equations

The solutions to the Pell Equation $x^2 - dy^2 = \pm 1$ can be found using properties of the periodic simple continued fraction of $\sqrt{d}$. We will show that any positive solution to the Pell Equation must be a convergent the periodic simple continued fraction of $\sqrt{d}$.

Theorem 5.1. *Let d, n be integers such that $n \leq 0$, $d > 0$ where d is not a perfect square, and $|n| < \sqrt{d}$. If $x^2 - dy^2 = n$ for positive integers x, y, then x/y is a convergent of the periodic simple continued fraction of $\sqrt{d}$.*

Before we prove this theorem, we need the following lemma.

Lemma 5.2. *Let d be a positive integer that is not a perfect square such that $\sqrt{d} > 1$. The kth convergent of the periodic simple continued fraction of $1/\sqrt{d}$ is the reciprocal of the $(k-1)$th convergent of the periodic simple continued fraction of $\sqrt{d}$.*

Now we can prove Theorem 5.1.

Proof of Theorem 5.1. By Theorem 3.7, we need to show that

$$\left| \frac{x}{y} - \sqrt{d} \right| < \frac{1}{2y^2}.$$

We will check the following two cases for n. The first case is where $n > 0$. Since $n = (x + y\sqrt{d})(x - y\sqrt{d})$ and x, y are positive, then $x - y\sqrt{d} > 0$, so $x > y\sqrt{d}$ and $(x/y) - \sqrt{d} > 0$. Since $0 < n < \sqrt{d}$, it follows that

$$\frac{x}{y} - \sqrt{d} = \frac{x - y\sqrt{d}}{y} = \frac{x^2 - dy^2}{(x + y\sqrt{d})y} = \frac{n}{(x + y\sqrt{d})y} < \frac{\sqrt{d}}{2y^2\sqrt{d}} = \frac{1}{2y^2},$$

and hence x/y is a convergent of the periodic simple continued fraction of $\sqrt{d}$. The other case is where $n < 0$. We can divide both sides of the equation $x^2 - dy^2 = n$ by $-d$ to get

$$y^2 - \frac{1}{d}x^2 = -\frac{n}{d}.$$

Therefore

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

This implies that $y - (1/\sqrt{d})x > 0$, so $y\sqrt{d} > x$ and $(y/x) - (1/\sqrt{d}) > 0$. By Lemma 5.2, it is sufficient to show that

$$\left| \frac{y}{x} - \frac{1}{\sqrt{d}} \right| < \frac{1}{2x^2}.$$

Since $0 < -n < \sqrt{d}$, it follows that

$$\begin{aligned} \frac{y}{x} - \frac{1}{\sqrt{d}} &= \frac{y\sqrt{d} - x}{x\sqrt{d}} \\ &= \frac{(y\sqrt{d} - x)(y\sqrt{d} + x)}{(y\sqrt{d} + x)x\sqrt{d}} \\ &< \frac{dy^2 - x^2}{2x^2\sqrt{d}} \\ &= -\frac{n}{2x^2\sqrt{d}} \\ &< \frac{\sqrt{d}}{2x^2\sqrt{d}} \\ &= \frac{1}{2x^2}. \end{aligned}$$

Therefore, y/x is a convergent of the periodic simple continued fraction of $1/\sqrt{d}$, and by Lemma 5.2, x/y is a convergent of the periodic simple continued fraction of $\sqrt{d}$. □

We will prove Theorem 1.1. Here we also include the exact solutions. This theorem uses the period of the periodic simple continued fraction of $\sqrt{d}$ and other properties of $\sqrt{d}$ to find all the solutions to the Pell Equation $x^2 - dy^2 = \pm 1$.

Theorem 5.3. *Let d be a positive integer that is not a perfect square. We can write the periodic simple continued fraction of $\sqrt{d}$ as $C = [a_0; a_1, a_2, \ldots]$ where the nonnegative integer n is the period. Let p_k/q_k be the kth convergent of C for any nonnegative integer k. If n is even, then the solutions to $x^2 - dy^2 = 1$ are $x = \pm p_{jn-1}$ and $y = \pm q_{jn-1}$ for all positive integers j, and $x^2 - dy^2 = -1$ has no solutions. If n is odd, then the solutions to $x^2 - dy^2 = 1$ are $x = \pm p_{2jm-1}$ and $y = \pm q_{2jm-1}$, and the solutions to $x^2 - dy^2 = -1$ are $x = \pm p_{(2j-1)n-1}$ and $y = \pm q_{(2j-1)n-1}$ for all positive integers j.*

Proof. Since the quadratic irrational $\lfloor\sqrt{d}\rfloor + \sqrt{d}$ is equal to

$$\lfloor\sqrt{d}\rfloor + \sqrt{d} = [\lfloor\sqrt{d}\rfloor; a_1, a_2, \ldots] + \sqrt{d} = [2\lfloor\sqrt{d}\rfloor; a_1, a_2, \ldots],$$

then it is reduced because $\lfloor\sqrt{d}\rfloor + \sqrt{d} > 1$ and $-1 < \lfloor\sqrt{d}\rfloor - \sqrt{d} < 0$. We will let $\beta_0 = \lfloor\sqrt{d}\rfloor + \sqrt{d}$ and $b_0 = 2\lfloor\sqrt{d}\rfloor$. By Theorem 4.9, the periodic simple continued fraction representing $\lfloor\sqrt{d}\rfloor + \sqrt{d}$ is purely periodic. Since the period of the periodic simple continued fraction of $\sqrt{d}$ is n, we see that

$$\sqrt{d} = [\overline{2\lfloor\sqrt{d}\rfloor; a_1, a_2, \ldots, a_{n-1}}] - \lfloor\sqrt{d}\rfloor = [\lfloor\sqrt{d}\rfloor; \overline{a_1, a_2, \ldots, a_{n-1}, 2\lfloor\sqrt{d}\rfloor}].$$

This implies $Q_{jn} = Q_0 = 1$ for all positive integers j. From Theorem 4.10, we get

$$p_{jn-1}^2 - dq_{jn-1}^2 = (-1)^{jn} Q_{jn} = (-1)^{jn}.$$

Therefore we have two cases. If n is even, then $(x, y) = (p_{jn-1}, q_{jn-1})$ is a solution of $x^2 - dy^2 = 1$ for all positive integers j. If n is odd, then $(x, y) = (p_{2jn-1}, q_{2jn-1})$ is a solution to $x^2 - dy^2 = 1$ and $(x, y) = (p_{(2j-1)n-1}, q_{(2j-1)n-1})$ is a solution to $x^2 - dy^2 = -1$ for all positive integers j. We need to check that these are the only solutions to the Pell Equation. Since

$$p_k^2 - dq^2 = (-1)^{k-1} Q_{k+1},$$

for (p_k, q_k) to be a solution to $x^2 - dy^2 = \pm 1$, Q_{k+1} must be equal to $Q_{k+1} = Q_{jn} = 1$. Therefore we can show that if $Q_{k+1} = 1$, then $n|(k+1)$ and $Q_j \neq -1$. If $Q_{k+1} = 1$, then

$$\alpha_{k+1} = P_{k+1} + \sqrt{d}.$$

Since $\alpha_{k+1} = [a_{k+1}; a_{k+2}, \ldots]$ and $\sqrt{d} = [\lfloor\sqrt{d}\rfloor; \overline{a_1, a_2, \ldots, a_{n-1}, 2\lfloor\sqrt{d}\rfloor}]$, then the periodic simple continued fraction representing α_{k+1} is purely periodic. By Theorem 4.9, $\alpha_{k+1} = P_{k+1} + \sqrt{d} > 1$ and $-1 < \alpha'_{k+1} = P_{k+1} - \sqrt{d} < 0$. This implies $P_{k+1} = \lfloor\sqrt{d}\rfloor$. Since $\alpha_{k+1} = \lfloor\sqrt{d}\rfloor + \sqrt{d} = \beta_0$, then $a_{k+1} = b_0$, so $n|k+1$. Next, assume for the sake of contradiction, that $Q_j = -1$. Then $\alpha_j = -P_j - \sqrt{d}$. We know that the periodic simple continued fraction of α_j is purely periodic, so $\alpha_j = -P_j - \sqrt{d} > 1$ and $-1 < \alpha'_j = -P_j + \sqrt{d} < 0$. This implies that $P_j < -(1 + \sqrt{d})$ and $P_j > \sqrt{d}$ which is a contradiction. Hence $Q_j \neq -1$. Therefore, the solutions we found earlier are the only solutions to the Pell Equation $x^2 - dy^2 = \pm 1$. ☐

Using this theorem, we can find solutions to the Pell Equation for any value of d.

Example 5.4. We will find a few solutions to the Pell Equation $x^2 - 3y^2 = \pm 1$ using Theorem 5.3. Since the periodic continued fraction of $\sqrt{3}$ is $[1; \overline{1, 2}]$ with period 2, the positive solutions to $x^2 - 3y^2 = 1$ are $(x, y) = (p_2 j - 1, q_2 j - 1)$ for all positive integers j and there are no solutions to $x^2 - 3y^2 = -1$. Note that $p_0 = 1, p_1 = 2, q_0 = 1$, and $q_1 = 1$, so the first few convergents are $C_0 = 1, C_1 = 2, C_2 = 5/3, C_3 = 7/4$. This implies that $(x, y) = (7, 4)$ and $(2, 1)$ are two of the solutions to $x^2 - 3y^2 = 1$.

References

[1] Rosen, Kenneth H. Elementary Number Theory and Its Applications. New York: Pearson/Addison Wesley, 2005.