ON THE SOLUTIONS OF PELL EQUATIONS

SHIVEN BHARGAVA

Abstract. In this paper, we use multiple techniques to find the complete set of solutions to Pell equations. We prove the existence of a nontrivial solution using Dirichlet's Approximation Theorem and then find all solutions from the minimal positive solution. We briefly mention the generalized Pell equation and how to generate infinitely many solutions from just one. In addition, we utilize the method of continued fractions as developed by Joseph-Louis Lagrange to represent solutions as convergents to the continued fraction of $\sqrt{d}$. We provide the proof of multiple obscure theorems related to continued fractions, such as the Law of Best Approximations, Lagrange's Theorem, Galois' Theorem, and eventually, the solution to Pell equations. We end with resources on generalized and negative Pell equations, and also prove a theorem that solves negative Pell equations using continued fractions. Ultimately, this paper thoroughly describes the solutions of Pell equations and an in-depth introduction to continued fractions.

1. Introduction

Pell equations are equations of the form $x^2 - dy^2 = 1$ for a fixed $d \in \mathbb{N}$, where $\mathbb{N}$ represents the set of positive integers. Generalized Pell equations are of the form $x^2 - dy^2 = n$ for any $n \in \mathbb{Z}$, and negative Pell equations are of the form $x^2 - dy^2 = -1$. The focus of this paper is to find integer pairs (x, y) that satisfy the first equation, which is always the one we are analyzing unless stated otherwise. Hence, whenever we discuss a solution to a Pell equation, we are referring to an integer one. We sometimes refer to a solutions to Pell equations as Pell solutions.

Now we discuss the trivial case. Notice that if d is a perfect square, say $d = c^2$, then we want to solve $x^2 - (cy)^2 = 1$, implying $(x - cy)(x + cy) = 1$. Thus $cy = 0$, so $(x, y) = (\pm 1, 0)$ are the only solutions in this case. The solutions for x, y that we just found are valid for any d, square or not. Hence, these are referred to as the *trivial* solutions to a Pell equation. Other solutions are called *nontrivial*, and if a solution satisfies the additional constraint that $x, y \in \mathbb{N}$, then we refer to it as a *positive* solution. Observe that all nontrivial solutions must have nonzero x, y. Because we have already handled the trivial case, we assume d is not a perfect square in all of our theorems, even ones relating to generalized or negative Pell equations.

1.1. History of the Equation.

History of the Equation. Pell equations have a remarkable history. The first known occurrence of a Pell equation is in the "Cattle Problem of Archimedes" in around 250 BCE, where one is trying to solve the equation $x^2 - 4729494y^2 = 1$. A generalized Pell equation appears in a work of Diophantus approximately 500 years later. In the 7th century, Brahmagupta, an Indian mathematician, came up with preliminary techniques to solve the generalized equation, making a considerable amount of progress and setting the basis for the work of Bhaskara II in India in the 12th century CE. Bhaskara II gave an algorithmic method to solve the generalized equation and even derived gargantuan solutions for some

Date: July 12, 2026.

specific values of d and n (such as $d = 61, n = 1$, whose minimal positive solution for y is 226153980).

Centuries later, in Europe, Fermat attempted to solve Pell equations, and sent this as a problem to mathematicians around Europe. In the 1600s, John Wallis and Lord William Brouncker confronted these equations, but it was Lagrange who successfully solved them with continued fractions. However, these equations are known today as Pell equations, the reason being that Euler accidentally attributed work on the equation to English mathematician John Pell (1611-1685), who had a very remote connection to these equations.

1.2. **Main Results.** We now describe the findings of this paper and the results of each section. We start out by discussing a few preliminaries needed to formally discuss Pell equations. We leave some theorems unproved, but the proofs of those can be found in the references. Then we move on to proving that any Pell equation has a nontrivial solution. In Section 4, we discuss the composition principle and its use in finding all solutions to a Pell equation from the minimal positive one. We shift gears in the next section and start expanding on the continued fraction theorems mentioned in Section 2. After introducing the basics, we analyze periodic and purely periodic continued fractions and prove Lagrange's Theorem and Galois' Theorem. Finally, in Section 7, we give all positive Pell solutions in terms of convergents to the simple continued fraction of $\sqrt{d}$. We conclude with an interesting theorem (and proof) on negative Pell equations and give a few resources for the reader to continue learning about this field.

2. Preliminaries

Before we dive into solving Pell equations, we need to introduce some notation and a few definitions and lemmas about continued fractions.

In this paper, we will use $\mathbb{Z}[\sqrt{d}]$ to denote the set $\{x + y\sqrt{d} \mid x, y \in \mathbb{Z}\}$. Note that $\mathbb{Z}[\sqrt{d}]$ is closed under multiplication. This is because $(x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d}) = (x_1x_2 + dy_1y_2) + (x_1y_2 + y_1x_2)\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$. Also, observe that for any $\alpha \in \mathbb{Z}[\sqrt{d}]$, there exist *unique* integers x, y such that $\alpha = x + y\sqrt{d}$. This follows simply from the fact that $\sqrt{d}$ is irrational. Furthermore, each element in $\mathbb{Z}[\sqrt{d}]$ has a unique conjugate:

Definition 2.1. The conjugate of an element $\alpha = x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ is $\overline{\alpha} = x - y\sqrt{d}$.

Moreover, we can define the norm of an element:

Definition 2.2. The norm of an element $\alpha \in \mathbb{Z}[\sqrt{d}]$ is defined as

$$N(\alpha) = \alpha\overline{\alpha}.$$

If $\alpha = x + y\sqrt{d}$, then

$$N\left(x + y\sqrt{d}\right) = \left(x + y\sqrt{d}\right)\left(x - y\sqrt{d}\right) = x^2 - dy^2.$$

Notice that solutions to $x^2 - dy^2 = 1$ are equivalent to elements of $\mathbb{Z}[\sqrt{d}]$ that have norm 1. We use (x, y) and $x + y\sqrt{d}$ interchangeably. We now prove a lemma about the norm.

Lemma 2.3. *The norm is multiplicative, i.e. for any $\alpha, \beta \in \mathbb{Z}[\sqrt{d}]$, we have*

$$N(\alpha\beta) = N(\alpha)\,N(\beta).$$

Proof. Let $\alpha = x_1 + y_1\sqrt{d}, \beta = x_2 + y_2\sqrt{d}$. Then

$$\begin{aligned}\overline{\alpha\beta} &= \overline{\left(x_1 + y_1\sqrt{d}\right)\left(x_2 + y_2\sqrt{d}\right)} \\ &= \overline{(x_1x_2 + dy_1y_2) + (x_1y_2 + y_1x_2)\sqrt{d}} \\ &= (x_1x_2 + dy_1y_2) - (x_1y_2 + y_1x_2)\sqrt{d} \\ &= \left(x_1 - y_1\sqrt{d}\right)\left(x_2 - y_2\sqrt{d}\right) \\ &= \overline{\alpha}\cdot\overline{\beta}\end{aligned}$$

Therefore,

$$N(\alpha\beta) = \alpha\beta\overline{\alpha\beta} = \alpha\overline{\alpha}\cdot\beta\overline{\beta} = N(\alpha)N(\beta). \qquad \blacksquare$$

Now that we have the notation setup, we can move onto continued fractions. Although they are not very well-known, continued fractions are extremely powerful tools that pop up in the most unexpected places. The results here are presented without proof, and a more thorough explanation can be found at [Yan08]. The more complicated theorems will be discussed in Sections 5 and 6.

Definition 2.4. Let $a_0, a_1, a_2, a_3, \ldots, a_m$ be a sequence of real numbers. Then

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{\cdots + \cfrac{1}{a_m}}}}}}$$

is referred to as a **finite continued fraction**, and is shorthanded by $[a_0; a_1, a_2, \ldots, a_m]$. If the fraction never ends, it is referred to as an **infinite continued fraction** and is written as $[a_0; a_1, a_2, \ldots]$. A **simple continued fraction** is a continued fraction under the constraint that a_0 is a nonzero integer and $a_i \in \mathbb{N}$ for all $i > 0$.

From now on, we always assume $[a_0; a_1, a_2, \ldots, a_m]$ is a continued fraction and that the values a_i are defined as above.

Definition 2.5. Let $0 \le n \le m$ be nonnegative integers. Let $x = [a_0, a_1, a_2, \ldots, a_{n-1}, a_n, \ldots, a_m]$ be a continued fraction. The ***n*th convergent** to x is defined as the continued fraction $[a_0, a_1, \ldots, a_n]$. Define two sequences of real numbers (p_n) and (q_n) in relation to the continued fraction of x where

$$p_{-1} = 1, p_0 = a_0, \text{ and } p_n = a_np_{n-1} + p_{n-2},$$

and analogously,

$$q_{-1} = 0, q_0 = 1, \text{ and } q_n = a_nq_{n-1} + q_{n-2}.$$

From now on, p_n, q_n will always be used as defined here. Although the definitions of p_n and q_n may look arbitrary, they are not. One can prove the following by induction:

Lemma 2.6. *$\frac{p_n}{q_n}$ is the nth convergent to $[a_0; a_1, a_2, a_3, \ldots, a_m]$.*

We also have the following relations between (p_n) and (q_n).

Lemma 2.7. *The following is true for all $1 \leq n \leq m$:*

(1) $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$
(2) $p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n$.

Corollary 2.8. *If $[a_0; a_1, a_2, \ldots, a_m]$ is a simple continued fraction, then p_n, q_n are relatively prime integers for all $0 \leq n \leq m$.*

We end this section with two definitions.

Definition 2.9. Let α be a real number. We define two recursive sequences, (a_n) and (α_n), as follows:

(1) $\alpha_0 = \alpha$
(2) $a_n = \lfloor \alpha_n \rfloor$
(3) $\alpha_n = a_n + \frac{1}{\alpha_{n+1}}$.

The sequences stop once $\alpha_n = a_n$. If this never happens, then the sequences are infinite.

Remark 2.10. Observe that $[a_0; a_1, a_2, \ldots]$ is the simple continued fraction expansion of α. In fact, one can show that the simple continued fraction expansion of any real number is unique by using this definition. Hence, from now on, we are referring to the simple continued fraction expansion of a real number if we refer to its continued fraction unless stated otherwise.

Furthermore, if α is a rational number, say $\alpha = \frac{p}{q}$ where $\gcd(p, q) = 1$, then notice that the last condition is equivalent to performing the Euclidean algorithm on p, q. This is because subtracting $a_n = \lfloor \alpha_n \rfloor$ is equivalent to taking the remainder when dividing, and reciprocating is switching the roles of the modulus and the remainder. It follows that the continued fraction of any rational number is finite since the Euclidean algorithm always terminates. The converse also holds: if α is irrational, its continued fraction must be infinite.

Additionally, note that $\alpha_n - a_n < 1$ for all $n \geq 0$, so $\alpha_{n+1} = \frac{1}{\alpha_n - a_n} > 1$, giving $a_{n+1} = \lfloor \alpha_{n+1} \rfloor \geq 1$, which implies that $a_n, \alpha_n \geq 1$ for any $n > 0$. Finally, observe that $\alpha = [a_0; a_1, a_2, \ldots, a_{m-1}, \alpha_m]$ and $\alpha_m = [a_m, a_{m+1}, a_{m+2}, \ldots,]$ for any positive m. These results can be proved through a simple induction.

Definition 2.11. An infinite continued fraction $[a_0, a_1, ...]$ is referred to as periodic if, for all sufficiently large n, $a_n = a_{n+l}$ for a fixed positive integer l called the period. The minimal period is the smallest positive integer l' such that this holds. The continued fraction is purely periodic if for any n we have $a_n = a_{n+l}$ for a fixed positive integer l. A periodic continued fraction is represented by putting a bar over the repeating part, i.e. $[1; \overline{1, 2}] = [1; 1, 2, 1, 2, 1, 2, \ldots]$.

3. Existence of a Solution

We begin this section by proving a famous result known as Dirichlet's Approximation Theorem. We then prove that there always exists a nontrivial solution to a Pell equation.

Lemma 3.1 (Dirichlet's Approximation Theorem)**.** *For any $\alpha \in \mathbb{R}$ and integer $N > 1$, there exists an integer p and a positive integer q such that $q \leq N$ and*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qN},$$

or equivalently,

$$|q\alpha - p| < \frac{1}{N} \leq \frac{1}{q}.$$

If $\alpha > 1$, then we can also impose the condition that p must be positive.

Proof. This proof follows that of [RS23]. Consider the $N+1$ numbers $a_i = i\alpha - \lfloor i\alpha \rfloor$, where $0 \leq i \leq N$. By properties of the floor function, $0 \leq a_i < 1$. Define N intervals $I_k = [\frac{k-1}{N}, \frac{k}{N})$, where $1 \leq k \leq N$. All these intervals are disjoint, and their union is $[0, 1)$. Hence, by the pigeonhole principle, there are two a_i's in the same interval, say I_k. Then we have

$$\frac{k-1}{N} \leq a_i, a_j < \frac{k}{N} \implies |a_i - a_j| < \frac{1}{N}.$$

Without loss of generality, we assume $i > j$. Substituting the definitions of a_i, we obtain

$$a_i - a_j = (i - j)\,\alpha + (\lfloor j\alpha \rfloor - \lfloor i\alpha \rfloor)\,.$$

Let $p = \lfloor i\alpha \rfloor - \lfloor j\alpha \rfloor$ and $q = i - j$. Note that $1 \leq q \leq N$. Hence,

$$|a_i - a_j| = |q\alpha - p| < \frac{1}{N},$$

as desired. If $\alpha > 1$, then $i\alpha > j\alpha + 1$, so $\lfloor i\alpha \rfloor > \lfloor j\alpha \rfloor$. Hence, $p = \lfloor i\alpha \rfloor - \lfloor j\alpha \rfloor > 0$ if $\alpha > 1$. ■

Theorem 3.2. *Any Pell equation has a nontrivial solution. Furthermore, it has a positive solution.*

Proof. To prove this theorem, we will show that there exists $M \in \mathbb{Z}$ where $|M| < 1 + 2\sqrt{d}$ such that the equation $x^2 - dy^2 = M$ has infinitely many solutions. We then use modular arithmetic and some clever algebraic manipulations to conclude the proof. This proof is taken from [Conb].

Since $\sqrt{d} > 1$, there exist infinitely many positive integer pairs (x, y) such that $|x - y\sqrt{d}| < \frac{1}{y}$ by Lemma 3.1. Hence,

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

Thus, we have $x + y\sqrt{d} < 1 + 2y\sqrt{d}$. We compute

$$\left|x^2 - dy^2\right| = \left(x + y\sqrt{d}\right)\left|x - y\sqrt{d}\right| < \left(1 + 2y\sqrt{d}\right)\frac{1}{y} \leq 1 + 2\sqrt{d}.$$

This inequality implies that $|x^2 - dy^2| < 1 + 2\sqrt{d}$ for infinitely many positive integer pairs (x, y). Hence, by the pigeonhole principle, there exists some $M \in \mathbb{Z}$ where $|M| < 1 + 2\sqrt{d}$ and

$$N(x + y\sqrt{d}) = x^2 - dy^2 = M \tag{3.1}$$

for infinitely many pairs of positive integers (x, y). Observe that, since $\sqrt{d}$ is irrational, M cannot be equal to 0. We apply the pigeonhole principle again, but this time, we also utilize modular arithmetic. Hence, there exist infinitely many positive integer pairs (x, y) that satisfy (3.1) and the following conditions for some fixed positive integer pair (x_0, y_0): $x \equiv x_0 \bmod |M|$, $y \equiv y_0 \bmod |M|$, $N(x + y\sqrt{d}) = x^2 - dy^2 = M$, and $(x, y) \neq (x_0, y_0)$. Thus, let (x_1, y_1) and (x_2, y_2) be two of these pairs. By the conditions, we have $x_1 \equiv x_2 \bmod |M|$ and $y_1 \equiv y_2 \bmod |M|$, so write $x_1 = x_2 + Mu$, $y_1 = y_2 + Mv$ for $u, v \in \mathbb{Z}$. Then

$$x_1 + y_1\sqrt{d} = x_2 + y_2\sqrt{d} + M\left(u + v\sqrt{d}\right),$$

$$x_1 - y_1\sqrt{d} = x_2 - y_2\sqrt{d} + M\left(u - v\sqrt{d}\right).$$

Plugging in $M = N(x_2 + y_2\sqrt{d}) = (x_2 + y_2\sqrt{d})(x_2 - y_2\sqrt{d})$, we compute

$$x_1 + y_1\sqrt{d} = \left(x_2 + y_2\sqrt{d}\right)\left(1 + \left(x_2 - y_2\sqrt{d}\right)\left(u + v\sqrt{d}\right)\right). \tag{3.2}$$

Similarly, we have

$$x_1 - y_1\sqrt{d} = (x_2 - y_2\sqrt{d})(1 + (x_2 + y_2\sqrt{d})(u - v\sqrt{d}). \tag{3.3}$$

In order to simplify the RHS of (3.2), we define

$$x' + y'\sqrt{d} = 1 + (x_2 - y_2\sqrt{d})(u + v\sqrt{d}) = (1 + x_2u - dy_2v) + (x_2v - y_2u)\sqrt{d}.$$

Notice that the second factor on the RHS of (3.3) is equivalent to $x - y\sqrt{d}$, giving

$$x_1 + y_1\sqrt{d} = \left(x_2 + y_2\sqrt{d}\right)\left(x' + y'\sqrt{d}\right),$$

$$x_1 - y_1\sqrt{d} = \left(x_2 - y_2\sqrt{d}\right)\left(x' - y'\sqrt{d}\right).$$

Multiplying the equalities, we see that

$$N\left(x_1 + y_1\sqrt{d}\right) = N\left(x_2 + y_2\sqrt{d}\right)N\left(x' + y'\sqrt{d}\right) \implies N\left(x' + y'\sqrt{d}\right) = 1$$

since $N(x_1 + y_1\sqrt{d}) = N(x_2 + y_2\sqrt{d}) = M$. Hence $x'^2 - dy'^2 = 1$. If $(x', y') = (1, 0)$ then $x_1 + y_1\sqrt{d} = x_2 + y_2\sqrt{d}$, implying $x_1 = x_2, y_1 = y_2$, a contradiction. If $(x', y') = (-1, 0)$ then $x_1 + y_1\sqrt{d} = -x_2 - y_2\sqrt{d}$, implying $x_1 = -x_2, y_1 = -y_2$, contradicting the fact that all four numbers are positive. Therefore, (x', y') is nontrivial, implying that x', y' are nonzero. If x' or y' are negative, then we can alter the signs without affecting the value of $x'^2 - dy'^2$. Hence, there is always a positive solution to $x^2 - dy^2 = 1$. ∎

4. All Solutions

In this section, our goal is to prove the following theorem.

Theorem 4.1. *Let (x_1, y_1) be a positive solution to the equation $x^2 - dy^2 = 1$ such that y_1 is minimal. Then all solutions to $x^2 - dy^2 = 1$ where $x, y \in \mathbb{Z}$ are, up to sign, generated by taking integral powers of $x_1 + y_1\sqrt{d}$, i.e.*

$$x + y\sqrt{d} = \pm\left(x_1 + y_1\sqrt{d}\right)^k \tag{4.1}$$

where $k \in \mathbb{Z}$. The solutions where x, y are positive are generated by taking $k > 0$ and the $+$ sign out front.

We divide the proof into two cases, namely proving that the condition for (x, y) to be a solution is sufficient and necessary.

4.1. **The Condition is Sufficient.** To prove sufficiency, we start with a simple but powerful theorem.

Theorem 4.2 (Composition Principle)**.** *Let $x_1, y_1, x_2,$ and y_2 be integers such that $x_1^2 - dy_1^2 = n_1$ and $x_2^2 - dy_2^2 = n_2$. Define $x_3 = x_1x_2 + dy_1y_2$ and $y_3 = x_1y_2 + y_1x_2$. Then*

$$x_3^2 - dy_3^2 = n_1n_2 = \left(x_1^2 - dy_1^2\right)\left(x_2^2 - dy_2^2\right).$$

Additionally, x_3, y_3 are the unique integers that satisfy $x_3 + y_3\sqrt{d} = (x_1 + y_1\sqrt{d})(x_2 + y_2\sqrt{d})$.

Proof. We translate this theorem into a statement about norms. Notice that

$$\begin{aligned} N\left(x_1 + y_1\sqrt{d}\right) N\left(x_2 + y_2\sqrt{d}\right) &= N\left(\left(x_1 + y_1\sqrt{d}\right)\left(x_2 + y_2\sqrt{d}\right)\right) \\ &= N\left((x_1x_2 + dy_1y_2) + (x_1y_2 + y_1x_2)\sqrt{d}\right), \end{aligned}$$

which is what we desire. ■

Corollary 4.3. *If $(x, y) = (x_1, y_1)$ is a positive solution to $x^2 - dy^2 = 1$, then the coefficients of $\pm(x_1 + y_1\sqrt{d})^k$ are a solution for all $k \in \mathbb{Z}$. The trivial solutions have $k = 0$ and the positive ones have $k > 0$ and $+$ out front. In particular, the equation $x^2 - dy^2 = 1$ has infinitely many nontrivial solutions.*

Proof. Notice that if (x, y) is a solution, so is $(-x, -y)$. Hence, we only need to prove that $(x_1 + y_1\sqrt{d})^k$ is a solution for all $k \in \mathbb{Z}$. Notice that, by the composition principle and repeated multiplication, the expression is a solution for all $k > 0$. Additionally, $k = 0$ gives $(x, y) = (\pm 1, 0)$, the trivial solutions. That leaves the case when $k < 0$. Let $K = -k$ and $(x_1 + y_1\sqrt{d})^K = x_K + y_K\sqrt{d}$ where $x_K, y_K \in \mathbb{Z}$. Then, since $x_K^2 - dy_K^2 = 1$, we have

$$\begin{aligned} \left(x_1 + y_1\sqrt{d}\right)^{-K} &= \frac{1}{\left(x_1 + y_1\sqrt{d}\right)^K} \\ &= \frac{1}{x_K + y_K\sqrt{d}} \\ &= \frac{x_K - y_K\sqrt{d}}{\left(x_K + y_K\sqrt{d}\right)\left(x_K - y_K\sqrt{d}\right)} \\ &= \frac{x_K - y_K\sqrt{d}}{x_K^2 - dy_K^2} \\ &= x_K - y_K\sqrt{d}. \end{aligned}$$

Since (x_K, y_K) is a Pell solution, so is $(x_K, -y_K)$, as desired. Additionally, all (x, y) generated by the expression are different, since otherwise there would be some integer $k \neq 0$ for which $(x + y\sqrt{d})^k = \pm 1$, a contradiction. Hence, there are infinitely many nontrivial Pell solutions. Lastly, if $k < 0$, the signs of x and y are always different, and if $k > 0$, then x, y are positive iff we take the $+$ out front. ■

We now briefly digress and discuss a result relating to generalized Pell equations.

Corollary 4.4. *If the equation $x^2 - dy^2 = n \neq 0$ has a solution, then it has infinitely many solutions.*

Proof. By Corollary 4.3, we know that $x^2 - dy^2 = 1$ has infinitely many nontrivial solutions, so let $(x, y) = (x_1, y_1)$ be one of these solutions. Let $(x, y) = (x_2, y_2)$ be a solution to $x^2 - dy^2 = n$. Then we know that

$$(x_3, y_3) = (x_1x_2 + dy_1y_2, x_1y_2 + y_1x_2) \tag{4.2}$$

is another solution to $x^2 - dy^2 = n$. Notice that if each (x_1, y_1) would generate a unique (x_3, y_3), then the equation $x^2 - dy^2 = n$ would have infinitely many solutions. To prove that all the pairs (x_3, y_3) are distinct, we let $(x', y') \neq (x_1, y_1)$ be a Pell solution such that the pairs generated by (4.2) are equal. Then, by comparing the values of x_3, we have

$$x_1x_2 + dy_1y_2 = x'x_2 + dy'y_2,$$

implying

$$(x_1 - x')\, x_2 = dy_2\, (y' - y_1)\,. \tag{4.3}$$

Similarly, by comparing the values of y_3, we compute

$$(x_1 - x')\, y_2 = x_2\, (y' - y_1)\,. \tag{4.4}$$

We split into three cases.

Case 1 ($\boldsymbol{x_2 = 0}$): Consider both 4.3 and 4.4. If $x_2 = 0$ then $y_2 \neq 0$, since otherwise $x^2 - dy^2 = 0$, a contradiction. Hence, since $d \neq 0$, $x_1 = x'$ and $y_1 = y'$, a contradiction.

Case 2 ($\boldsymbol{y_1 = y'}$): Consider 4.3. If $y_1 = y'$, then either $x_1 = x'$ or $x_2 = 0$. Since $(x_1, y_1) \neq (x', y')$, $x_1 \neq x'$, so $x_2 = 0$. This reduces to Case 1.

Case 3 ($\boldsymbol{y_2 = 0}$): Consider both 4.3 and 4.4. If $x_2 = 0$, then this reduces to Case 1. Otherwise, $(x_1, y_1) = (x', y')$, a contradiction.

Thus, we can divide (4.3) by $x_2(y' - y_1)$ and (4.4) by $y_2(y' - y_1)$ to get

$$\frac{dy_2}{x_2} = \frac{x_1 - x'}{y' - y_1} = \frac{x_2}{y_2} \implies x_2^2 - dy^2 = 0,$$

contradicting the fact that $n \neq 0$. Thus, no two (x_3, y_3) are the same, and there are infinitely many solutions. ∎

4.2. **The Condition is Necessary.** To finish the proof of Theorem 4.1, we prove multiple inequalities specific to Pell solutions which would not hold otherwise. We then combine these inequalities to derive the final proof.

Lemma 4.5. *If (x, y) is a Pell solution such that $x + y\sqrt{d} > 1$, then $x \geq 2$ and $y \geq 1$.*

Note that $2^2 - 3 \cdot 1^2 = 1$, so the bounds can be achieved when $d = 3$.

Proof. Since $x^2 - dy^2 = 1$, $x - y\sqrt{d} = \frac{1}{x+y\sqrt{d}}$. Hence

$$x + y\sqrt{d} > 1 > x - y\sqrt{d} > 0 \implies 2y\sqrt{d} > 0.$$

Thus $y > 0$, implying $y \geq 1$. Hence, $x > y\sqrt{d} > y \geq 1$, so $x \geq 2$. ∎

Lemma 4.6. *Let (a, b) and (u, v) be two solutions to the equation $x^2 - dy^2 = 1$ such that $a, b, u, v \in \mathbb{Z}_{\geq 0}$. Then*

$$a + b\sqrt{d} < u + v\sqrt{d} \iff a < u \text{ and } b < v \iff a < u \text{ or } b < v.$$

Proof. The first ($\Leftarrow$) and the second ($\Rightarrow$), are obvious. To show the first ($\Rightarrow$), observe that $a \geq 1$ since $a \neq 0$. Similarly, $u \geq 1$. Thus $a + b\sqrt{d} \geq 1$, and similarly, $u + v\sqrt{d} \geq 1$. Since $x - y\sqrt{d} = \frac{1}{x+y\sqrt{d}}$ for any Pell solution (x, y), we have

$$a + b\sqrt{d} < u + v\sqrt{d}$$

and

$$u - v\sqrt{d} < a - b\sqrt{d}.$$

Hence $(a + u) + (b - v)\sqrt{d} < (a + u) + (v - b)\sqrt{d}$, implying $b < v$. Since a and u are nonnegative, $a^2 = 1 + db^2 < 1 + dv^2 = u^2$ implies $a < u$. Now we prove the second ($\Leftarrow$). Note that since $(a, b) \neq (u, v)$, $a + b\sqrt{d} \neq u + v\sqrt{d}$. Thus, since $a < u$ or $b < v$, we deduce that $a < u$ and $b < v$ from the first ($\Rightarrow$). ∎

We can now fully prove Theorem 4.1.

Proof of Theorem 4.1. This proof follows that of [Cona]. By Corollary 4.3, we know that all of the pairs generated by (4.1) satisfy the equation $x^2 - dy^2 = 1$. We just wish to prove that all Pell solutions are of that form. Let (x, y) be a positive Pell solution. Then $x + y\sqrt{d} > 1$. We will prove that $x + y\sqrt{d} = (x_1 + y_1\sqrt{d})^k$ for some $k > 0$. We will finish by using this result to prove the other cases.

Since $x_1 + y_1\sqrt{d} > 1$, the sequence $a_n = (x_1 + y_1\sqrt{d})^n$ for $n \geq 0$ forms a strictly increasing sequence starting at $a_0 = 1$. Hence, there exists a unique integer $k \geq 0$ such that

$$\left(x_1 + y_1\sqrt{d}\right)^k \leq x + y\sqrt{d} < \left(x_1 + y_1\sqrt{d}\right)^{k+1}. \tag{4.5}$$

Thus, dividing (4.5) by $(x_1 + y_1\sqrt{d})^k$ on both sides, we obtain

$$1 \leq \left(x + y\sqrt{d}\right)\left(x_1 + y_1\sqrt{d}\right)^{-k} < x_1 + y_1\sqrt{d}.$$

Note that $(x + y\sqrt{d})(x_1 + y_1\sqrt{d})^{-k}$ has coefficients that are a solution of a Pell equation by Theorem 4.2. Setting $a + b\sqrt{d} = (x + y\sqrt{d})(x_1 + y_1\sqrt{d})^{-k}$ for $a, b \in \mathbb{Z}$, we have $a^2 - db^2 = 1$ and

$$1 \leq a + b\sqrt{d} < x + y\sqrt{d}.$$

If $a + b\sqrt{d} = 1$, then $x + y\sqrt{d} = (x_1 + y_1\sqrt{d})^k$, as desired. Now assume $a + b\sqrt{d} > 1$. Then, by Lemma 4.5, we have $a, b \in \mathbb{N}$, so $0 < b < y_1$ by Lemma 4.6, contradicting the fact that y_1 is minimal.

If x and y are not both positive, then $\alpha \coloneqq x+y\sqrt{d}$ is not in the range $(1, \infty)$ by Lemma 4.5. If $\alpha = \pm 1$, i.e. (x, y) is a trivial solution, then we can take $k = 0$. Otherwise, α is in one of the ranges $(-\infty, -1), (-1, 0),$ or $(0, 1)$, so exactly one of the numbers

$$\frac{1}{\alpha} = x - y\sqrt{d}, \ \frac{-1}{\alpha} = -x + y\sqrt{d}, \text{ or } -\alpha = -x - y\sqrt{d}$$

is in the range $(1, \infty)$. Observe that all of these are Pell solutions as we can freely switch the signs on (x, y). Hence $\pm\alpha^{\pm 1} = (x_1 + y_1\sqrt{d})^k$ for some positive integer k, as desired. ∎

5. Elementary Theorems on Continued Fractions

In this section, we prove some elementary theorems about continued fractions and begin to develop tools needed to solve Pell equations. This section follows that of [Yan08].

Lemma 5.1. *Let α be irrational, and let $\frac{p_n}{q_n}$ be the nth convergent to its continued fraction. Then*

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

Proof. By Remark 2.10, we know that $\alpha = [a_0; a_1, a_2, \ldots, a_n, \alpha_{n+1}]$. We take this as the continued fraction of α. Hence, we have

$$\begin{aligned}
\alpha - \frac{p_n}{q_n} &= \frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} \\
&= \frac{\alpha_{n+1}p_n + p_{n-1}}{\alpha_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} \\
&= \frac{\alpha_{n+1}p_nq_n + p_{n-1}q_n - \alpha_{n+1}p_nq_n - p_nq_{n-1}}{q_n\left(\alpha_{n+1}q_n + q_{n-1}\right)} \\
&= \frac{p_{n-1}q_n - p_nq_{n-1}}{q_n\left(\alpha_{n+1}q_n + q_{n-1}\right)} \\
&= \frac{(-1)^{n-1}}{q_n\left(\alpha_{n+1}q_n + q_{n-1}\right)},
\end{aligned}$$

and the desired conclusion follows. ■

Remark 5.2. Since $\alpha - \frac{p_n}{q_n} = \frac{(-1)^{n-1}}{q_n(\alpha_{n+1}q_n + q_{n-1})}$, $\alpha - \frac{p_n}{q_n}$ and $\alpha - \frac{p_{n+1}}{q_{n+1}}$ have opposite signs for all $n \geq 0$.

The next result is a powerful inequality.

Corollary 5.3. *For any irrational α and positive integer n,*

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

Proof. By Lemma 5.1, we have

$$\left|\alpha - \frac{p_n}{q_n}\right| = \frac{1}{q_n\left(\alpha_{n+1}q_n + q_{n-1}\right)} \implies \left|q_n\alpha - p_n\right| = \frac{1}{\alpha_{n+1}q_n + q_{n-1}}.$$

Similarly,

$$\left|q_{n-1}\alpha - p_{n-1}\right| = \frac{1}{\alpha_nq_{n-1} + q_{n-2}}.$$

Hence, we wish to prove the following:

$$\begin{aligned}
\frac{1}{\alpha_{n+1}q_n + q_{n-1}} &< \frac{1}{\alpha_n q_{n-1} + q_{n-2}} \\
&= \frac{1}{\left(a_n + \frac{1}{\alpha_{n+1}}\right) q_{n-1} + q_{n-2}} \\
&= \frac{1}{a_n q_{n-1} + \frac{q_{n-1}}{\alpha_{n+1}} + q_{n-2}} \\
&= \frac{\alpha_{n+1}}{a_n q_{n-1}\alpha_{n+1} + q_{n-1} + q_{n-2}\alpha_{n+1}} \\
&= \frac{\alpha_{n+1}}{\alpha_{n+1}\left(a_n q_{n-1} + q_{n-2}\right) + q_{n-1}} \\
&= \frac{\alpha_{n+1}}{\alpha_{n+1}q_n + q_{n-1}}.
\end{aligned}$$

Since $\alpha_{n+1}q_n + q_{n-1} > 0$, it suffices to prove $\alpha_{n+1} > 1$, which is true by Remark 2.10. ■

To finish off this subsection, we will prove two beautiful approximation theorems.

Theorem 5.4 (The Law of Best Approximations)**.** *Let α be any real number with convergent $\frac{p_n}{q_n}$ for some $n \geq 2$. If integers p, q satisfy $0 < q \leq q_n$ and $\frac{p}{q} \neq \frac{p_n}{q_n}$, then*

$$|q_n\alpha - p_n| < |q\alpha - p|.$$

Additionally, if integers p', q' with $q' \geq q_2$ satisfy the above inequality for all $0 < q \leq q'$ and $\frac{p}{q} \neq \frac{p'}{q'}$, then $\frac{p'}{q'}$ is a convergent.

Proof. Notice that if $\frac{p}{q}$ is a convergent to α, then the result follows from Corollary 5.3. In the rest of the proof suppose $\frac{p}{q}$ is not a convergent. Assume $q = q_n$. Then, since $p \neq p_n$,

$$\left|\frac{p}{q} - \frac{p_n}{q_n}\right| = \left|\frac{p - p_n}{q_n}\right| \geq \frac{1}{q_n}.$$

Notice that, because (q_n) is strictly increasing for $n \geq 1$ and $q_1 \geq q_0 = 1$ (this can be seen through the recursive formula), $q_{n+1} \geq 3$ for $n \geq 2$. Hence,

$$\begin{aligned}
\left|\alpha - \frac{p_n}{q_n}\right| &= \frac{1}{q_n\left(\alpha_{n+1}q_n + q_{n-1}\right)} \\
&< \frac{1}{q_n\left(a_{n+1}q_n + q_{n-1}\right)} \\
&= \frac{1}{q_n q_{n+1}} \\
&< \frac{1}{2q_n}.
\end{aligned}$$

Therefore, by the triangle inequality, we have

$$\begin{aligned}\left|\alpha - \frac{p}{q}\right| &\geq \left|\frac{p_n}{q_n} - \frac{p}{q}\right| - \left|\alpha - \frac{p_n}{q_n}\right| \\ &> \frac{1}{q_n} - \frac{1}{2q_n} \\ &= \frac{1}{2q_n} \\ &> \left|\alpha - \frac{p_n}{q_n}\right|,\end{aligned}$$

and we get the result by multiplying by $q = q_n$. Now suppose $0 < q < q_n$. Let (x, y) be the solution to the system of equations

$$p_n X + p_{n-1} Y = p$$

$$q_n X + q_{n-1} Y = q \tag{5.1}$$

Then we compute

$$x = \frac{pq_{n-1} - qp_{n-1}}{p_n q_{n-1} - q_n p_{n-1}} = \pm \left(pq_{n-1} - qp_{n-1}\right),$$

$$y = \frac{pq_n - qp_n}{p_n q_{n-1} - q_n p_{n-1}} = \pm \left(pq_n - qp_n\right).$$

Since $\frac{p}{q}$ is not a convergent, x, y are nonzero integers. Hence, we conclude that x and y have opposite signs by (5.1) and the assumption that $q_n > q$. Therefore, we have

$$q\alpha - p = \left(q_n x + q_{n-1} y\right)\alpha - \left(p_n x + p_{n-1} y\right) = x\left(q_n\alpha - p_n\right) + y\left(q_{n-1}\alpha - p_{n-1}\right).$$

Observe that $q_n\alpha - p_n$ and $q_{n-1}\alpha - p_{n-1}$ have opposite signs by Remark 5.2, so $x(q_n\alpha - p_n)$ and $y(q_{n-1}\alpha - p_{n-1})$ have the same sign. Hence,

$$|q\alpha - p| = |x\left(q_n\alpha - p_n\right)| + |y\left(q_{n-1}\alpha - p_{n-1}\right)| \geq |q_{n-1}\alpha - p_{n-1}| > |q_n\alpha - p_n|.$$

Thus, we have proved the first part of the theorem. For the second part, notice that for any $q_{n-1} < q < q_n$, this inequality holds, so $\frac{p'}{q'} = \frac{p}{q}$ cannot satisfy the inequality, as desired. ■

Theorem 5.5. *Given any two consecutive convergents to a real number α, at least one of them satisfies*

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

Moreover, any fraction $\frac{p}{q}$ satisfying the above inequality is a convergent.

Proof. For the first part, suppose that there are two convergents, say the nth and the $n+1$th, that do not satisfy the inequality. Then, since the sign of $\alpha - \frac{p_n}{q_n}$ alternates by Remark 5.2, we have

$$\begin{aligned}\frac{1}{q_nq_{n+1}} &= \left|\frac{p_nq_{n+1}-q_np_{n+1}}{q_nq_{n+1}}\right| \\ &= \left|\frac{p_n}{q_n}-\frac{p_{n+1}}{q_{n+1}}\right| \\ &= \left|\alpha-\frac{p_n}{q_n}\right|+\left|\alpha-\frac{p_{n+1}}{q_{n+1}}\right| \\ &\geq \frac{1}{2q_n^2}+\frac{1}{2q_{n+1}^2} \\ &= \frac{q_n^2+q_{n+1}^2}{2q_n^2q_{n+1}^2}.\end{aligned}$$

Multiplying, we see that

$$2q_nq_{n+1} \geq q_n^2+q_{n+1}^2 \implies (q_n-q_{n+1})^2 \leq 0.$$

However, (q_n) is strictly increasing for positive n, so the only possible case in which this holds is if $q_0 = q_1 = a_1 = 1$ and $n = 0$. Then, since $\alpha > \frac{p_0}{q_0}$,

$$0 < \frac{p_1}{q_1} - \alpha = a_1a_0 + 1 - \alpha = a_0 + 1 - [a_0; 1, a_2, a_3, \ldots] = 1 - [0; 1, a_2, a_3, \ldots].$$

Let $x = [a_3; a_4, a_5, \ldots]$. Then

$$a_2 + \frac{1}{x} > a_2,$$

$$1 + \cfrac{1}{a_2 + \cfrac{1}{x}} < 1 + \frac{1}{a_2},$$

$$\cfrac{1}{1 + \cfrac{1}{a_2 + \cfrac{1}{x}}} > \cfrac{1}{1 + \cfrac{1}{a_2}}.$$

Hence, we compute

$$1 - [0; 1, a_2, a_3, \ldots] < 1 - \cfrac{1}{1 + \cfrac{1}{a_2}} = 1 - \frac{a_2}{a_2+1} \leq \frac{1}{2} = \frac{1}{2q_1^2},$$

so $|\alpha - \frac{p_1}{q_1}| < \frac{1}{2q_1^2}$. Therefore, we have exhausted all possible cases.

Now we prove the second part of the theorem. By Theorem 5.4, we want to show that $\frac{p}{q}$ is a best approximation to α if it satisfies the inequality. Let integers P, Q satisfy $\frac{P}{Q} \neq \frac{p}{q}$, $|Q\alpha - P| \leq |q\alpha - p| < \frac{1}{2q}$, where the last inequality follows from the assumption. Notice that

$$\begin{aligned}\frac{1}{qQ} &\le \left|\frac{Pq-pQ}{qQ}\right| \\ &= \left|\frac{P}{Q}-\frac{p}{q}\right| \\ &\le \left|\alpha-\frac{P}{Q}\right| + \left|\alpha-\frac{p}{q}\right| \\ &< \frac{1}{2qQ}+\frac{1}{2q^2} \\ &= \frac{q+Q}{2q^2Q}.\end{aligned}$$

Hence, we have

$$2q < q+Q \implies q < Q.$$

Therefore, $\frac{p}{q}$ is a best approximation, and thus a convergent. ■

6. Lagrange and Galois

In this section, we prove two famous theorems relating to continued fractions and quadratic irrationals. Theorems about periodic continued fractions are taken from [Ike24] and ones relating to purely periodic continued fractions are taken from [Ike19].

Definition 6.1. An irrational number α is referred to as a quadratic irrational if it is the root of a degree two polynomial with integer coefficients. The other root β of the polynomial is its conjugate. Moreover, we say α is reduced if $\alpha > 1$ and $-1 < \beta < 0$.

Remark 6.2. The conjugate of a quadratic irrational as defined here is equivalent to one defined in Definition 6.9.

To prove Lagrange's Theorem, there are multiple facts about quadratic irrationals that are very simple to prove, so we present them here without proof.

Proposition 6.3. *A real number α is a quadratic irrational iff it can be represented in the form $\frac{p+\sqrt{q}}{r}$ where $p, q, r \in \mathbb{Z}$, $r \neq 0$, and q is positive and not a perfect square. Furthermore, if x is a quadratic irrational, then so is $a_0 + \frac{1}{x}$ for $a_0 \in \mathbb{Z}$.*

Lemma 6.4. *If x is a quadratic irrational, then so is $[a_0; a_1, a_2, \ldots, a_n, x]$ where $a_i \in \mathbb{Z}$ for all $0 \le i \le n$.*

Proof. If $n = 0$, this is true by Proposition 6.3. Assume the statement for $n = k$. We will prove it for $n = k+1$. Notice that

$$[a_0; a_1, \ldots, a_{k+1}, x] = [a_0; a_1, \ldots, a_k, [a_k+1; x]]\,.$$

By the base case, $[a_{k+1}; x]$ is a quadratic irrational, so the desired conclusion follows from the inductive hypothesis. ■

We include another proposition without proof that follows from basic algebra and can be proven using induction.

Proposition 6.5. *For integers a_i, the continued fraction $[a_0, a_1, \ldots, a_n, x]$ can be represented as*

$$\frac{ax+b}{cx+d}$$

for integers a, b, c, d.

Lemma 6.6. *If the simple continued fraction of x is $[\overline{a_0; a_1, a_2, \ldots, a_n}]$, then x is a quadratic irrational.*

Proof. Notice that $x = [a_0; a_1, \ldots, a_n, x]$, implying that $x = \frac{ax+b}{cx+d}$. Thus,

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

Since the continued fraction of x is infinite, x must be irrational, and so x is a quadratic irrational. ∎

Proposition 6.7. *Every quadratic irrational x can be written as $\frac{m+\sqrt{d}}{s}$, where*

1. *(1) $s \mid d - m^2$*
2. *(2) $m, s, d \in \mathbb{Z}$*
3. *(3) $s \neq 0$*
4. *(4) $d > 0$ and d is not a perfect square.*

Proof. Let $x = \frac{p+\sqrt{q}}{r}$ be written in the form mentioned in Proposition 6.3. Let $d = qs^2$, $s = s|s|$, and $m = m|s|$. Then $\frac{m+\sqrt{d}}{s} = \frac{p+\sqrt{q}}{r}$ and the desired conditions hold. ∎

Lemma 6.8. *Let $x = \frac{m+\sqrt{d}}{s}$ be a quadratic irrational satisfying the conditions in Proposition 6.7. Then there exist infinite integer sequences $(m_k), (s_k), (a_k)$ and an infinite sequence of irrationals (x_k) such that:*

$$m_0 = m, s_0 = s, x_0 = x,$$

$$a_i = \lfloor x_i \rfloor \text{ for all } i \geq 0,$$

$$m_{k+1} = a_k s_k - m_k \text{ for all } k \geq 0,$$

$$s_{k+1} = \frac{d - m_{k+1}^2}{s_k} \text{ for all } k \geq 0,$$

$$x_{k+1} = \frac{m_{k+1} + \sqrt{d}}{s_{k+1}} \text{ for all } k \geq 0.$$

These sequences satisfy

1. *(1) $s_k \neq 0$, $s_k \mid d - m_k^2$, and $s_k \mid d - m_{k+1}^2$ for all $k \geq 0$,*
2. *(2) The simple continued fraction of x is $x = [a_0, a_1, a_2, \ldots]$.*

Proof. We first prove that $m_k, a_k,$ and s_k are integers sequences as well as condition (1). Note that a_0, m_0, and $s_0 = s \neq 0$ are all integers. Observe that

$$d - m_1^2 = d - (a_0 s_0 - m_0)^2 = d - (a_0 s - m)^2 = (d - m^2) - s(a_0^2 s - 2a_0 m).$$

Hence $s_0 = s \mid d - m_1^2$. Since $s_0 = s \mid d - m^2 = d - m_0^2$, the claims are true for $k = 0$. Assume the claims for $n = k-1$. Then a_k, $m_k = a_{k-1}s_{k-1} - m_{k-1}$, and $s_k = \frac{d-m_k^2}{s_{k-1}}$ are obviously integers. Also, x_k is irrational. Furthermore, if $s_k = 0$, then $d = m_k^2$, a contradiction as d is not a perfect square. Hence $s_k \neq 0$. Notice that

$$\begin{aligned} d - m_{k+1}^2 &= d - (a_k s_k - m_k)^2 \\ &= (d - m_k^2) - a_k^2 s_k^2 + 2a_k s_k m_k \\ &= s_k s_{k-1} - s_k(a_k^2 s_k - 2a_k m_k). \end{aligned}$$

Thus $s_k \mid d - m_{k+1}^2$, and since $s_k s_{k+1} = d - m_{k+1}^2$, we obtain $s_{k+1} \mid d - m_{k+1}^2$. We are done by induction.

Now we prove condition (2). We have

$$\begin{aligned} x_k - a_k &= \frac{m_k + \sqrt{d}}{s_k} - a_k \\ &= \frac{(m_k - a_k s_k) + \sqrt{d}}{s_k} \\ &= \frac{\sqrt{d} - m_{k+1}}{s_k} \\ &= \frac{\sqrt{d} - m_{k+1}}{s_k} \cdot \frac{\sqrt{d} + m_{k+1}}{\sqrt{d} + m_{k+1}} \\ &= \frac{d - m_{k+1}^2}{s_k(\sqrt{d} + m_{k+1})} \\ &= \frac{s_{k+1}}{\sqrt{d} + m_{k+1}} \\ &= \frac{1}{x_{k+1}}. \end{aligned}$$

Hence, since $a_k = \lfloor x_k \rfloor$, this is the algorithm defined in Definition 2.9. Thus $[a_0, a_1, a_2, \ldots]$ is the simple continued fraction expansion of x. ■

We present a definition and a few properties whose proofs are left to the reader.

Definition 6.9. The conjugate of a quadratic irrational $x = \frac{p+\sqrt{q}}{r}$ in the form of Proposition 6.3 is $\overline{x} = \frac{p-\sqrt{q}}{r}$. The following properties hold:

$$\overline{x+y} = \overline{x} + \overline{y},\ \overline{xy} = \overline{x} \cdot \overline{y},\ \overline{\left(\frac{x}{y}\right)} = \frac{\overline{x}}{\overline{y}}.$$

Additionally, the conjugate of a rational number is itself.

We can now prove Lagrange's Theorem.

Theorem 6.10 (Lagrange's Theorem)**.** *A real number α is a quadratic irrational iff its continued fraction is periodic.*

Proof. For the backwards direction, we are done by Proposition 6.5 and Lemma 6.6. The proof for the forwards direction has 3 steps. First, we prove that $s_n > 0$ for all sufficiently large n. Then, we show that s_n and m_n both assume only finitely many values. We conclude by considering x_n.

To prove that $s_n > 0$ for sufficiently large n, notice that since $x_0 = [a_0; a_1, \ldots, a_{n-1}, x_n]$,

$$x_0 = \frac{p_n}{q_n} = \frac{p_{n-1}x_n + p_{n-2}}{q_{n-1}x_n + q_{n-2}}.$$

Observe that x_0, x_n are quadratic irrationals with the same radical: $\sqrt{d}$. Hence we can take the conjugate of both sides of the equation above, giving

$$\overline{x_0} = \overline{\left(\frac{p_{n-1}x_n + p_{n-2}}{q_{n-1}x_n + q_{n-2}}\right)} = \frac{\overline{p_{n-1}x_n + p_{n-2}}}{\overline{q_{n-1}x_n + q_{n-2}}} = \frac{p_{n-1}\overline{x_n} + p_{n-2}}{q_{n-1}\overline{x_n} + q_{n-2}}.$$

Now solve for $\overline{x_n}$:

$$\begin{aligned}
\overline{x_0}(q_{n-1}\overline{x_n} + q_{n-2}) &= p_{n-1}\overline{x_n} + p_{n-2} \\
\overline{x_0}q_{n-1}\overline{x_n} - p_{n-1}\overline{x_n} &= p_{n-2} - q_{n-2}\overline{x_0} \\
\overline{x_n} &= \frac{p_{n-2} - q_{n-2}\overline{x_0}}{q_{n-1}\overline{x_0} - p_{n-1}} \\
&= -\frac{q_{n-2}}{q_{n-1}} \cdot \frac{\frac{p_{n-2}}{q_{n-2}} - \overline{x_0}}{\frac{p_{n-1}}{q_{n-1}} - \overline{x_0}}.
\end{aligned}$$

Notice that, as n increases, $\frac{p_{n-2}}{q_{n-2}}$ and $\frac{p_{n-1}}{q_{n-1}}$ converge to $x = x_0$. Hence, since $x_0 \neq \overline{x_0}$ as x_0 is irrational,

$$\lim_{n\to\infty} \frac{\frac{p_{n-2}}{q_{n-2}} - \overline{x_0}}{\frac{p_{n-1}}{q_{n-1}} - \overline{x_0}} \to 1.$$

Therefore,

$$\overline{x_n} \approx -\frac{q_{n-2}}{q_{n-1}},$$

so $\overline{x_n} < 0$ for all sufficiently large n, or equivalently, $x_n - \overline{x_n} > 0$. This implies $\frac{m_n+\sqrt{d}}{s_n} - \frac{m_n-\sqrt{d}}{s_n} > 0$, so $\frac{2\sqrt{d}}{s_n} > 0$, giving $s_n > 0$ for all sufficiently large n.

For the second step, note that $s_n s_{n+1} = d - m_{n+1}^2$. Since s_n, s_{n+1} are positive for all sufficiently large n, we have $s_n s_{n+1} \leq d$. Thus $1 \leq s_n \leq d$, so s_n only assumes finitely many values for all sufficiently large n. This implies that $m_{n+1}^2 = d - s_n s_{n+1} < d$, so $-\sqrt{d} < m_n < \sqrt{d}$ for all sufficiently large n.

Combining the facts that both m_n, s_n assume only finitely many values for large enough n, we see that the pair (m_n, s_n) also assumes only finitely many values. Then $x_n = \frac{m_n+\sqrt{d}}{s_n}$ also assumes only finitely many values. Let $(m_i, s_i, x_i) = (m_j, s_j, x_j)$ for some $i < j$. Then we see that $(m_{i+1}, s_{i+1}, x_{i+1}) = (m_{j+1}, s_{j+1}, x_{j+1})$ and so on. This repetition also implies that $a_i = \lfloor x_i \rfloor$ repeats periodically. Hence the simple continued fraction of a quadratic irrational is periodic. ■

We now discuss purely periodic continued fractions.

Theorem 6.11 (Galois' Theorem)**.** *The simple continued fraction of a quadratic irrational $x = [a_0; a_1, \ldots]$ is purely periodic iff x is reduced.*

Proof. First we prove the backwards direction, so suppose $x > 1$ and $-1 < \overline{x} < 0$. Then, since $x_{n+1} = \frac{1}{x_n - a_n}$,

$$\frac{1}{\overline{x_{n+1}}} = \overline{x_n - a_n} = \overline{x_n} - a_n.$$

We know that $a_n \geq 1$ if $n \geq 1$. Additionally, $a_0 \geq 1$ since $x > 1$, so $a_n \geq 1$ for all n. We claim that $-1 < \overline{x_n} < 0$ for all n. Notice that this is true for $n = 0$. Assume the claim for $n = k$. Then $-1 < \overline{x_k} < 0$. Since $a_k \geq 1$, $-a_k \leq -1$, so $\overline{x_k} - a_k < -1$. Hence $\frac{1}{\overline{x_{k+1}}} < -1$, so $-1 < \overline{x_{k+1}} < 0$, and the claim is true by induction. Since $\overline{x_n} = \frac{1}{\overline{x_{n+1}}} + a_n$,

$$-1 < \overline{x_n} = \frac{1}{\overline{x_{n+1}}} + a_n < 0.$$

Therefore,

$$0 < \left(-\frac{1}{\overline{x_{n+1}}}\right) - a_n < 1 \implies a_n < -\frac{1}{\overline{x_{n+1}}} < a_n + 1 \implies a_n = \left\lfloor -\frac{1}{\overline{x_{n+1}}} \right\rfloor.$$

Since x is a quadratic irrational, there exist positive integers $i < j$ such that $x_i = x_j$, so

$$a_{i-1} = \left\lfloor -\frac{1}{\overline{x_i}} \right\rfloor = \left\lfloor -\frac{1}{\overline{x_j}} \right\rfloor = a_{j-1}.$$

Hence

$$a_{i-1} + \frac{1}{x_i} = a_{j-1} + \frac{1}{x_j} \implies x_{i-1} = x_{j-1}.$$

Hence, $x_i = x_j$ implies $x_{i-1} = x_{j-1}$, and continuing to reduce the indices, we end up with $a_n = a_{n+(j-i)}$ for all n.

For the forwards direction, notice that if $x = [\overline{a_0, a_1, \ldots, a_{n-1}}]$, then $x_0 = x_n$. Hence

$$x_0 = \frac{p_{n-1}x_n + p_{n-2}}{q_{n-1}x_n + q_{n-2}} = \frac{p_{n-1}x_0 + p_{n-2}}{q_{n-1}x_0 + q_{n-2}}.$$

Expanding, we obtain

$$q_{n-1}x^2 + (q_{n-2} - p_{n-1})x - p_{n-2} = 0.$$

Let $f(x)$ denote this quadratic. $f(x)$ has two roots, namely $x, \overline{x}$. We already know that $x > 1$ since $a_0 = a_n \geq 1$. Hence, we only need to show that $-1 < \overline{x} < 0$. This is equivalent to showing that f has a root between -1 and 0 since the other root, x, is not. Note that $f(0) = -p_{n-2} < 0$. Also,

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

Hence, since $f(0) < 0 < f(-1)$, we conclude the desired by the Intermediate Value Theorem.

■

Using the two theorems we proved in this section (Lagrange's and Galois'), we can derive the following.

Theorem 6.12. *The simple continued fraction of $x = \sqrt{d}$ when d is not a perfect square is*

$$x = \left[\left\lfloor \sqrt{d} \right\rfloor; \overline{a_1, a_2, \ldots, a_n, 2\left\lfloor \sqrt{d} \right\rfloor}\right]$$

for integers a_i. Particularly, x_n is purely periodic for all $n > 0$.

Proof. Consider irrationals $\alpha = \lfloor\sqrt{d}\rfloor + \sqrt{d}$, $\beta = \lfloor\sqrt{d}\rfloor - \sqrt{d}$. Observe that α, β are conjugates and that $-1 < \beta < 0, 1 < \alpha$. Moreover, α is reduced. Hence, the simple continued fraction of α is purely periodic. Since $\lfloor\alpha\rfloor = 2\left\lfloor\sqrt{d}\right\rfloor$, we have

$$\sqrt{d} + \left\lfloor\sqrt{d}\right\rfloor = \alpha = \left[\overline{2\left\lfloor\sqrt{d}\right\rfloor, a_1, a_2, \ldots, a_n}\right] = \left[2\left\lfloor\sqrt{d}\right\rfloor, \overline{a_1, a_2, \ldots, a_n, 2\left\lfloor\sqrt{d}\right\rfloor}\right].$$

Hence,

$$\sqrt{d} = \left[\left\lfloor\sqrt{d}\right\rfloor, \overline{a_1, a_2, \ldots, a_n, 2\left\lfloor\sqrt{d}\right\rfloor}\right]$$

■

7. Pell Solutions from Convergents

This section ends with a theorem detailing how to find all Pell solutions from convergents. We cite [Lag67] since it is the original text in which these results appeared.

Lemma 7.1. *Let $x = \sqrt{d}$ where d is not a perfect square. Then, in the notation of Lemma 6.8, we have*

$$p_{n-1}^2 - dq_{n-1}^2 = (-1)^n s_n \tag{7.1}$$

for all $n \geq 2$.

Proof. Observe that

$$\sqrt{d} = \frac{x_n p_{n-1} + p_{n-2}}{x_n q_{n-1} + q_{n-2}} = \frac{\left(m_n + \sqrt{d}\right) p_{n-1} + s_n p_{n-2}}{\left(m_n + \sqrt{d}\right) q_{n-1} + s_n q_{n-2}}.$$

Hence, we obtain

$$\sqrt{d}\left(\left(m_n + \sqrt{d}\right) q_{n-1} + s_n q_{n-2}\right) = \left(m_n + \sqrt{d}\right) p_{n-1} + s_n p_{n-2}$$
$$dq_{n-1} + (m_n q_{n-1} + s_n q_{n-2})\sqrt{d} = (m_n p_{n-1} + s_n p_{n-2}) + p_{n-1}\sqrt{d}$$

Since all the variables are integers other than $\sqrt{d}$ which is irrational, we have $p_{n-1} = m_n q_{n-1} + s_n q_{n-2}$, $dq_{n-1} = m_n p_{n-1} + s_n p_{n-2}$. Multiplying the first by p_{n-1}, the second by q_{n-1}, and subtracting, we obtain

$$p_{n-1}^2 - dq_{n-1}^2 = p_{n-1}q_{n-1}(m_n - m_n) + s_n(p_{n-1}q_{n-2} - p_{n-2}q_{n-1}) = (-1)^n s_n.$$

■

Theorem 7.2. *Let l be the minimal period of the continued fraction of $\sqrt{d}$. We use the notation of Lemma 6.8, the only difference being that we use α instead of x. Then all positive solutions to a Pell equation are as follows:*

$$(x, y) = \begin{cases} (p_{lm-1}, q_{lm-1}) & \text{if } l \text{ is even} \\ (p_{2lm-1}, q_{2lm-1}) & \text{if } l \text{ is odd} \end{cases}$$

where m is a positive integer.

Proof. First we show that all positive Pell solutions are convergents to the continued fraction of $\sqrt{d}$ (i.e. $x = p_n, y = q_n$). Let (x, y) be a Pell solution. Then $x > y$. Hence, since $\sqrt{d} > 1$, we obtain

$$\left|\sqrt{d} - \frac{x}{y}\right| = \left|\frac{y\sqrt{d} - x}{y}\right| = \left|\frac{\left(y\sqrt{d} - x\right)\left(y\sqrt{d} + x\right)}{y\left(y\sqrt{d} + x\right)}\right| = \left|\frac{-1}{y^2\left(\sqrt{d} + \frac{x}{y}\right)}\right| < \frac{1}{2y^2},$$

so $\frac{x}{y}$ is a convergent by Theorem 5.5. We will now show that $s_n = 1$ iff $l \mid n$ and that $s_n \neq 1$ for all n. To prove that $s_n = 1$ implies $l \mid n$, observe that α_1 is purely periodic with period l by Theorem 6.12, so $\alpha_1 = \alpha_{n+1}$ iff $l \mid n$. Let s_n for some n. We already know that $n = 0$ satisfies the claim, so assume $n > 0$. Then

$$\alpha_{n+1} = \frac{1}{\alpha_n - \lfloor \alpha_n \rfloor} = \frac{1}{(m_n + \sqrt{d}) - (m_n + \left\lfloor \sqrt{d} \right\rfloor)} = \frac{1}{\sqrt{d} - \left\lfloor \sqrt{d} \right\rfloor} = \frac{1}{\alpha_0 - \lfloor \alpha_0 \rfloor} = \alpha_1,$$

as desired.

We also have to prove that $l \mid n$ implies $s_n = 1$. However, we already proved that $l \mid n$ implies $\alpha_1 = \alpha_{n+1}$. Hence

$$\alpha_n - \lfloor \alpha_n \rfloor = \alpha_0 - \lfloor \alpha_0 \rfloor = \sqrt{d} - \left\lfloor \sqrt{d} \right\rfloor,$$

so $\alpha_n = x + \sqrt{d}$ for some integer x, implying that $s_n = 1$.

Finally, suppose for the sake of contradiction that $s_n = -1$. For this to happen, we know that $n > 0$. Then, by Theorem 6.12, α_n is purely periodic, so we can apply Galois' Theorem. Since $\alpha_n = -m_n - \sqrt{d}$, we obtain

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

a contradiction.

Since $s_n \neq 1$, we have that for (p_n, q_n) to be a Pell solution, $s_{n+1} = 1$. We also require $n + 1$ to be even by Lemma 7.1. Hence, n is 1 less than an even multiple of l, completing the proof. ■

8. Further Reading

In this section, we start by proving an interesting theorem about negative Pell equations and continued fractions. Then we will provide resources for the reader to continue learning about generalized Pell equations.

Theorem 8.1. *Let all the variables be defined as in Theorem 7.2. If l is even, there are no solutions to the equation $x^2 - dy^2 = -1$. If l is odd, then all positive solutions to the negative Pell equation $x^2 - dy^2 = -1$ are*

$$(x, y) = (p_{lm-1}, q_{lm-1})$$

where m is an odd positive integer.

Proof. Let (x, y) be a solution to the given negative Pell equation. Then notice that $x^2 = dy^2 - 1 \geq 2y^2 - 1 \geq y^2$, so $x \geq y$. Hence

$$\left|\sqrt{d} - \frac{x}{y}\right| = \left|\frac{y\sqrt{d} - x}{y}\right| = \left|\frac{\left(y\sqrt{d} - x\right)\left(y\sqrt{d} + x\right)}{y\left(y\sqrt{d} + x\right)}\right| = \left|\frac{1}{y^2\left(\sqrt{d} + \frac{x}{y}\right)}\right| < \frac{1}{2y^2},$$

so $\frac{x}{y}$ is a convergent to the simple continued fraction of $\sqrt{d}$. By Lemma 7.1, we know that if a convergent $\frac{p_n}{q_n}$ is a solution to $x^2 - dy^2 = -1$, then $s_{n+1} \pm 1$. However, by the proof of Theorem 7.2, $s_n \neq 1$ for any n, so $s_{n+1} = 1$ and $n + 1$ is odd. However, by the same proof, $s_{n+1} = 1$ iff $l \mid n + 1$, so n is 1 less than an odd multiple of l, as desired. ■

In this paper, we only briefly touched on generalized Pell equations. A good resource to learn more about this is [Conb], in which the author, Keith Conrad, bounds the minimal positive solution for the generalized equation and briefly extends the continued fraction method of Section 7. In [Cona], he provides a method known as modular contradiction to show that some generalized Pell equations have no integer solutions. He also mentions negative Pell equations, explaining their relation to Pell solutions and provides a formula similar to the one in Section 4 that generates solutions to $x^2 - dy^2 = \pm 1$ from the minimal negative solution.

If you would like to read about applications of Pell equations, then [Con08] would be a good place to start. These slides were also made my Keith Conrad.

Lastly, I would like to link the slides of a talk I gave on Pell equations: https://github.com/shivenbhargava/publications/blob/main/Pell_Equations_Talk.pdf.

Acknowledgments

The author would like to thank Arianna Ridgeway and Simon Rubinstein-Salzedo for helpful conversations.

References

[Cona] Keith Conrad. Pell's equation, I. https://kconrad.math.uconn.edu/blurbs/ugradnumthy/pelleqn1.pdf. Expository notes, University of Connecticut.

[Conb] Keith Conrad. Pell's equation, II. https://kconrad.math.uconn.edu/blurbs/ugradnumthy/pelleqn2.pdf. Expository notes, University of Connecticut.

[Con08] Keith Conrad. Applications of pell's equation. Lecture slides, University of Connecticut, August 2008.

[Ike19] Bruce Ikenaga. Purely periodic continued fractions. Number Theory course notes, Millersville University, 2019.

[Ike24] Bruce Ikenaga. Periodic continued fractions. Number Theory course notes, Millersville University, November 2024.

[Lag67] Joseph Louis Lagrange. *Œuvres de Lagrange*, volume 9. Gauthier-Villars, Paris, 1867.

[RS23] Simon Rubinstein-Salzedo. *Transition to proofs*. World Scientific, 2023.

[Yan08] Seung Hyun Yang. Continued fractions and pell's equation. August 2008. REU paper, University of Chicago.