Chapter 1: Introduction and History

In February, 1657, Pierre de Fermat (1601-1665) issued a challenge to the European math community. "Find a number y which will make dy^2+1 a perfect square, where d is a positive integer which is not a square. If a general rule cannot be obtained, find the smallest values of y which will satisfy the equations $61y^2+1=x^2$ or $109y^2+1=x^2$." The challenge is quickly translated into integral solutions for the equation $x^2-dy^2=1$. Perhaps unknown to Fermat and the European community, was that this was not the first time mathematicians had studied this equation. There is evidence that Greek mathematicians around 400 B.C. and mathematicians in India during the 7th century A.D., had studied specific cases of the problem. We will see more on this later. One mathematician who decided to face Fermat's challenge was William Brouncker.

Brouncker, at first, thought rational numbers were permitted as solutions so he quickly found the general solution of $x=\frac{r^2+d}{r^2-d}$, $y=\frac{2r}{r^2-d}$ where r is an arbitrary rational number such that $r \neq d$. Fermat, however, was not happy with this solution. He stated, "Solutions in fractions, which can be given at once from the merest elements of arithmetic, do not satisfy me." Brouncker then teamed up with another mathematician John Wallis. They were successful in finding a method for solving the equation but were unable to give a proof. It was not for over a hundred years that a general solution was to be found.

Euler, in 1759, discovered a way to find the first solution of $x^2 - dy^2 = 1$ using infinite simple continued fractions but never gave proof that it led to anything other than the trivial solution $x = 1, y = 0$. In 1768 Lagrange, however, completed Euler's work and published a proof that all solutions can be found in the continued fraction expansion of $\sqrt{d}$.

So why is the equation $x^2 - dy^2 = 1$ called "Pell's Equation"? This was due to a mishap by Euler in confusing Brouncker's work with that of another mathematician John Pell, who had little or nothing to do with solving the equation.

Today many textbooks refer to a Pell equation as $x^2 - dy^2 = N$, where N is an integer. This paper will focus on the classic equation proposed by Fermat, $x^2 - dy^2 = 1$, as well as the case where $x^2 - dy^2 = -1$. Further, we will only discuss positive solutions because x, y is a solution if and only if $\pm x, \pm y$ is a solution for any combination you like. First, however, some background on infinite simple continued fractions is needed.

Chapter 2: Infinite Simple Continued Fractions

Definition: An infinite simple continued fraction is a fraction of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

where $\{a_n\}$ is a sequence of integers all positive except possibly a_0. Typically infinite

simple continued fractions are denoted by $[a_0; a_1, a_2, a_3, ...]$ since they are fairly space consuming to write otherwise. Further, I will usually refer to these fractions as just continued fractions throughout the paper but keep in mind that they are all infinite and simple unless otherwise stated. One very important aspect of continued fractions for us is the idea of convergents.

Definition: The k^{th} convergent of the continued fraction $[a_0; a_1, a_2, a_3, ...]$ is denoted by C_k where

$$C_k = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ a_{k-1} + \cfrac{1}{a_k}}}}$$

Notice here that all convergents are rational numbers so, throughout this paper, we will typically refer to $C_k = \frac{p_k}{q_k}$ where p_k and q_k are positive integers such that the $\gcd(p_k, q_k) = 1$. We can see then that $\lim_{k\to\infty} C_k = [a_0; a_1, a_2, ...]$. It is also helpful to notice, though we will not prove, that the sequence $\{C_k\}$ is an alternating sequence. For an example we will find the first few convergents for the continued fraction $[1;1,2,3,\ldots\ldots]$.

$$C_0 = a_0 = 1$$

$$C_1 = a_0 + \frac{1}{a_1} = 1 + \frac{1}{1} = 2 \quad (p_1 = 2 \text{ and } q_1 = 1)$$

$$C_2 = a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = 1 + \frac{1}{1 + \frac{1}{2}} = \frac{5}{3} \quad (p_2 = 5 \text{ and } q_2 = 3)$$

$$C_3 = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3}}} = 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{3}}} = \frac{17}{10} \quad (p_3 = 17 \text{ and } q_3 = 10)$$

Another fact we will take for granted is that the value of any infinite continued fraction is irrational. Further, if we are given an irrational number z, we can find a continued fraction whose value is z. Here is a method that works rather well.

Define $a_0 = \lfloor z \rfloor$, $z_1 = \frac{1}{z - a_0}$, $a_1 = \lfloor z_1 \rfloor$. Further, define $a_i = \lfloor z_i \rfloor$ and $z_{i+1} = \frac{1}{z_i - a_i}$.

To help clarify how this method works, let's take a closer look.

$$[a_0; z_1] = a_0 + \frac{1}{z_1} = a_0 + \frac{1}{\frac{1}{z - a_0}} = z$$

$$[a_0; a_1, z_2] = a_0 + \frac{1}{a_1 + \frac{1}{z_2}} = a_0 + \frac{1}{z_1} = z$$

We can see that this pattern will continue and so if

$[a_0; a_1, ..., a_{k-1}, z_k] = z$ for any k, then $\lim_{k \to \infty} [a_0; a_1, ..., a_{k-1}, z_k] = z$. Moreover,

$\lim_{k \to \infty} [a_0; a_1, ..., a_{k-1}, z_k] = [a_0; a_1, a_2, ...]$. So $[a_0; a_1, a_2, ...] = z$. Further, since z is irrational, we are guaranteed that the continued fraction will be infinite. Let's take a look at an example.

Example: Find a continued fraction whose value is $\sqrt{2}$.

By following what we did above, $a_0 = \lfloor \sqrt{2} \rfloor = 1$ and $z_1 = \frac{1}{\sqrt{2}-1} = \frac{\sqrt{2}+1}{2-1} = \sqrt{2}+1$.

Then, $a_1 = \lfloor z_1 \rfloor = \lfloor \sqrt{2}+1 \rfloor = 2$

$$z_2 = \frac{1}{\sqrt{2}+1-2} = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1.$$

So $a_2 = 2$ and the pattern will repeat itself. Therefore, $\sqrt{2} = [1;2,2,2,.....]$.

We can verify this is the case.

$$[1;2,2,....] = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{\ldots}}} = 1 + \cfrac{1}{1 + [1 + \cfrac{1}{2 + \cfrac{1}{\ldots}}}$$
$$\ldots \qquad\qquad \ldots]$$

$$\Rightarrow [1;2,2,....] = 1 + \frac{1}{1+[1;2,2,....]} \Rightarrow [1;2,2,....] + [1;2,2,....]^2 = 2 + [1;2,2,....]$$

$$\Rightarrow [1;2,2,....]^2 = 2$$

$$\Rightarrow [1;2,2,....] = \sqrt{2}$$

One theorem about continued fractions that is important in solving Pell's equation is,

Theorem 1: Every irrational number has a unique representation as an infinite continued fraction.

Proof: We have just seen that given an irrational number z, we can find an infinite simple continued fraction whose value is z. So it suffices to prove it is unique.

First we need to show if $[a_0; a_1, a_2, ...] = z$ then $C_0 < z < C_1$.

Well, $C_0 = a_0 < a_0 + \dfrac{1}{[a_1; a_2, a_3, ...]} = z$ since $a_i \geq 1$ for all $i \geq 1$. Similarly,

$C_1 = a_0 + \dfrac{1}{a_1} > a_0 + \dfrac{1}{a_1 + [a_2; a_3, a_4, ...]} = z$ since $a_1 < a_1 + [a_2; a_3, a_4, ...]$. Therefore,

$C_0 < z < C_1$.

Now suppose there exists $[a_0; a_1, a_2, ...] = z$ and $[b_0; b_1, b_2, ...] = z$. Well, from above,

$a_0 < z < a_0 + \dfrac{1}{a_1} \Rightarrow a_0 < z < a_0 + 1$ since $a_1 \geq 1$. Then $\lfloor z \rfloor = a_0$. Similarly,

$b_0 < z < b_0 + \dfrac{1}{b_1} \Rightarrow b_0 < z < b_0 + 1$ so $\lfloor z \rfloor = b_0$ and so $a_0 = b_0$. Then

$$a_0 + \frac{1}{[a_1; a_2, a_3, ...]} = z = b_0 + \frac{1}{[b_1; b_2, b_3, ...]}$$

$$\Rightarrow \frac{1}{[a_1; a_2, a_3, ...]} = \frac{1}{[b_1; b_2, b_3, ...]} \text{ since } a_0 = b_0$$

$$\Rightarrow [a_1; a_2, a_3, ...] = [b_1; b_2, b_3, ...].$$

Now lets assume it holds for some integer k, i.e. that for all $i \leq k$, $a_i = b_i$. Then

$[a_{k+1}; a_{k+2}, a_{k+3}, ...] = [b_{k+1}; b_{k+2}, b_{k+3}, ...] = z_0$ for some irrational number z_0. Then,

$a_{k+1} < z_0 < a_{k+1} + \dfrac{1}{a_{k+2}} \Rightarrow a_{k+1} < z < a_{k+1} + 1$, and so $\lfloor z_0 \rfloor = a_{k+1}$. Similarly,

$b_{k+1} < z_0 < b_{k+1} + \dfrac{1}{b_{k+2}} \Rightarrow b_{k+1} < z < b_{k+1} + 1$, so $\lfloor z_0 \rfloor = b_{k+1}$, and $a_{k+1} = b_{k+1}$.

Therefore, for all integers k, $a_k = b_k$, which proves the continued fraction is unique. If z is an irrational number we call the unique continued fraction representation of z the

continued fraction expansion of z. Another important concept involving continued fractions that we will use later is that of periodic continued fractions.

Definition: An infinite simple continued fraction $[a_0; a_1, a_2, a_3, ...]$ is said to be periodic if there is an integer r such that $a_k = a_{k+r}$ for all sufficiently large integers k. The smallest such r is referred to as the length of the period.

When a continued fraction is periodic it is usually denoted as such, $[a_0; a_1, a_2, ..., a_s, \overline{a_{s+1}, ..., a_{s+r}}]$. If a periodic continued fraction is of the form $[a_0; \overline{a_1, a_2, ..., a_r}]$ then we say the continued fraction is purely periodic. For example, the continued fraction $[1;3,5,8,7,6,7,6,7,6,.....]$,where the 7 and 6 continue to repeat, is denoted as $[1;3,5,8,\overline{7,6}]$ and has period length 2. The continued fraction$[2;3,9,7,3,9,7,.....]$ is the same as $[2;\overline{3,9,7,}]$ and is purely periodic with length 3. One interesting, and later helpful, result is that if x is a non-square integer then the continued fraction expansion of $\sqrt{x}$ is purely periodic. Now that we have sufficient background on continued fractions, we are ready to take a look at Pell's equation.

Chapter 3: Pell's Equation for d = 2

Lets take a look at $x^2 - 2y^2 = \pm 1$. Notice for any d, $x = 1, y = 0$ will always be a solution to $x^2 - dy^2 = 1$; and it is referred to as the trivial solution denoted by x_0, y_0. Further, we can create sequences $\{x_n\}$ and $\{y_n\}$ where $\{x_n\}$ is strictly increasing and $\{y_n\}$ is ordered in correspondence with its x_n value. This notation will be helpful later on. It takes little time to find the first few solutions.

x_n	y_n	$x_n^2 - 2y_n^2$
1	0	1
1	1	–1
3	2	1
7	5	–1
17	12	1
41	29	–1
99	70	1

You might have noticed the pattern above that if x_n, y_n is a solution then the next solution appears to be $x_{n+1} = x_n + 2y_n$ and $y_{n+1} = x_n + y_n$. Whether it is the next solution is something we will discuss later. We can, however, show that it is a solution for if $x_n^2 - 2y_n^2 = \pm 1$ then $(x_n + 2y_n)^2 - 2(x_n + y_n)^2 = x_n + 4x_n y_n + 4y_n^2 - 2x_n - 4x_n y_n - 2y_n^2$

$$= -x_n^2 + 2y_n^2 = -(x_n^2 - 2y_n^2) = -(\pm 1) = \mp 1$$

One direct result of this is that there are infinitely many solutions for the $d = 2$ case. If we suppose there exists finitely many solutions, then there exists $x_k = \max\{x_n\}$. Then from the previous work $x_b = x_k + 2y_k, y_b = x_k + y_k$ is also a solution. But $x_b = x_k + 2y_k > x_k$ since $2y_k > 0$ which contradicts the fact that x_k is the largest solution. There is a connection between the solutions for $d = 2$ and the continued fraction expansion of $\sqrt{2}$ but first we need to calculate its convergents.

Recall that $\sqrt{2} = [1;\overline{2}]$. So,

$$C_1 = 1 + \frac{1}{2} = \frac{3}{2} \quad \Rightarrow p_1 = 3 \text{ and } q_1 = 2$$

$$C_2 = 1 + \frac{1}{2 + \frac{1}{2}} = \frac{7}{5} \quad \Rightarrow p_2 = 7 \text{ and } q_2 = 5$$

$$C_3 = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}} = \frac{17}{12} \quad \Rightarrow p_3 = 17 \text{ and } q_3 = 12$$

$$C_4 = ... = \frac{41}{29} \quad \Rightarrow \; p_4 = 41 \text{ and } q_4 = 29$$

$$C_5 = ... = \frac{99}{70} \quad \Rightarrow \; p_5 = 99 \text{ and } q_5 = 70$$

Have you noticed yet that our (p_n, q_n) values so far are precisely the (x_n, y_n) solutions we found earlier for $x^2 - 2y^2 = \pm 1$? Let's see if this approach works for $d = 3$.

Chapter 4: Pell's Equation for d = 3

Here we find that the first solutions to $x^2 - 3y^2 = \pm 1$ are

x_n	y_n	$x_n^2 - 3y_n^2$
1	0	1
2	1	1
7	4	1
26	15	1

So far there does not seem to be a solution to $x^2 - 3y^2 = -1$.

Lemma 1.1: There exists no solutions to $x^2 - dy^2 = -1$ when $d \equiv 3 (\text{mod}\, 4)$.

Proof: First we need to prove that if x is an integer then $x^2 \equiv 0 \pmod 4$

or $x^2 \equiv 1 \pmod 4$.

Case 1: x is even.

Then $x = 2k$ for some integer k and $x^2 = (2k)^2 = 4k^2 \equiv 0 \pmod 4$.

Case 2: x is odd.

Then $x = 2k+1$ for some integer k and

$x^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 4(k^2+k)+1 \equiv 1 \pmod 4$.

Now we can continue with the rest of the proof. Let $d \equiv 3 \pmod 4$. Then from what we just did, $dy^2 \equiv 0 \pmod 4$ or $dy^2 \equiv 3 \pmod 4$.

Case 1: $x^2 \equiv 0 \pmod 4$

Then $x^2 - dy^2 \equiv 0 - 0 \neq -1 \pmod 4$

or $x^2 - dy^2 \equiv 0 - 3 \neq -1 \pmod 4$

Case 2: $x^2 \equiv 1 \pmod 4$

Then $x^2 - dy^2 \equiv 1 - 0 \neq -1 \pmod 4$

or $x^2 - dy^2 \equiv 1 - 3 \neq -1 \pmod 4$ which concludes our proof.

Now let us see if there is a relation to the continued fraction expansion of $\sqrt{3}$ like there was in the $d = 2$ case. The continued fraction expansion of $\sqrt{3}$ is $[1;\overline{1,2}]$ and so the convergents are

$$C_1 = 1 + \frac{1}{1} = 2 = \frac{2}{1} \text{ so } p_1 = 2 \text{ and } q_1 = 1$$

$$C_2 = 1 + \frac{1}{1+\frac{1}{2}} = 1 + \frac{2}{3} = \frac{5}{3} \text{ so } p_2 = 5 \text{ and } q_2 = 3$$

$$C_3 = 1 + \frac{1}{1+\frac{1}{2+\frac{1}{1}}} = 1 + \frac{3}{4} = \frac{7}{4} \text{ so } p_3 = 7 \text{ and } q_3 = 4$$

$$C_4 = \text{.....} = \frac{19}{11} \text{ so } p_4 = 19 \text{ and } q_4 = 11$$

Notice that p_1, q_1 and p_3, q_3 are solutions however p_2, q_2 and p_4, q_4 are not. So it seems as though there is some relationship between the continued fraction expansion of $\sqrt{3}$ and solutions to the equation $x^2 - 3y^2 = \pm 1$, but apparently it is not as complete as in the case of $d = 2$.

Chapter 5: Pell's Equation for d = 4

When d = 4, we have the trivial solution 1,0 to $x^2 - 4y^2 = \pm 1$ but it seems as though we can not find any others. Actually, it is relatively easy to see if d is any perfect square then there exists no non-trivial solutions to $x^2 - dy^2 = \pm 1$. If you recall from the introduction, Fermat's challenge contained the condition that d was not a perfect square.

Lemma 1.2: If d is a perfect square, then there is no non-trivial solution to $x^2 - dy^2 = \pm 1$.

Proof: Let x, y be a solution to $x^2 - dy^2 = \pm 1$ where d is a perfect square.

Then if d is a perfect square, there exists an integer k such that $k^2 = d$ and

$$\pm 1 = x^2 - dy^2 = x^2 - (ky)^2$$

So x^2 and $(ky)^2$ are perfect squares that differ by 1. The only perfect squares, however, that differ by 1 are 1 and 0. Thus $x = 1$ and $y = 0$ is the only solution, which is trivial.

Chapter 6: All Solutions To $x^2 - dy^2 = 1$ Are Among Our Convergents

We will now focus on a general solution for the original equation proposed by Fermat and then discuss the case when $x^2 - dy^2 = -1$ later. Further, we will also restrict d to be

any positive integer that is not a perfect square since we just proved this case to be trivial. We have observed that there seems to be a relationship between the solutions and the convergents of the continued fraction expansion of $\sqrt{d}$. This chapter focuses on proving our next Theorem, Theorem 2, that all of our solutions must be among our convergents. This will ultimately lead to our general solution. First, however, we need to prove some Lemmas.

Lemma 2.1: Let C_k denote the k^{th} convergent for the continued fraction expansion of $\sqrt{d} = [a_0; a_1, a_2, ...]$. If we define p_k and q_k to be

$$p_0 = a_0 \qquad q_0 = 1$$
$$p_1 = a_1 a_0 + 1 \qquad q_1 = a_1$$
$$p_k = a_k p_{k-1} + p_{k-2} \qquad q_k = a_k q_{k-1} + q_{k-2}$$

Then $C_k = \frac{p_k}{q_k}$

Proof:

This will be a proof by induction. For the base cases we have

$$\frac{p_0}{q_0} = \frac{a_0}{1} = a_0 = C_0 \quad , \quad \frac{p_1}{q_1} = \frac{a_1 a_0 + 1}{a_1} = a_0 + \frac{1}{a_1} = C_1 \quad ,$$

$$\frac{p_2}{q_2} = \frac{a_2 p_1 + p_0}{a_2 q_1 + q_0} = \frac{a_2(a_1 a_0 + 1) + a_0}{a_2 a_1 + 1} = \frac{a_2 a_1 a_0 + a_2 + a_0}{a_2 a_1 + 1}$$

$$= \frac{a_0(a_1 a_2 + 1) + a_2}{a_1 a_2 + 1} = a_0 + \frac{a_2}{a_1 a_2 + 1} = a_0 + \frac{1}{\frac{a_1 a_2 + 1}{a_2}}$$

$$= a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = C_2 .$$

Now for the inductive step, let it be true for some $k > 2$. Then

$[a_0; a_1, a_2, ..., a_k] = C_k = \dfrac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}}$. Further, the conditions that are true for any continued fraction must hold for the k^{th} convergent of the continued fraction, so

$[a_0; a_1, a_2, ..., a_{k-1}, a_k + \dfrac{1}{a_{k+1}}, a_{k+2}, ...]$ which means

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

$$= \frac{a_k p_{k-1} + \frac{p_{k-1}}{a_{k+1}} + p_{k-2}}{a_k q_{k-1} + \frac{q_{k-1}}{a_{k+1}} + q_{k-2}} = \frac{a_{k+1}}{a_{k+1}} \frac{a_k p_{k-1} + \frac{p_{k-1}}{a_{k+1}} + p_{k-2}}{a_k q_{k-1} + \frac{q_{k-1}}{a_{k+1}} + q_{k-2}}$$

$$= \frac{a_{k+1}(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{a_{k+1}(a_k q_{k-1} + q_{k-2}) + q_{k-1}} = \frac{a_{k+1} p_k + p_{k-1}}{a_{k+1} q_k + q_{k-1}}.$$

Further, $C_k' = [a_0; a_1, a_2, ..., a_{k-1}, a_k + \dfrac{1}{a_{k+1}}] = [a_0; a_1, a_2, ..., a_k, a_{k+1}] = C_{k+1}$ so

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

Lemma 2.2: If we define p_k and q_k as we did in Lemma 2.1, then

$p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}$ for all $k \in \mathbb{N}$.

Proof: (by induction)

For the base case,

$$p_1 q_0 - p_0 q_1 = (a_1 a_0 + 1) - a_0 a = p_1 - a_0 a_1 = 1 = (-1)^{1-1}$$

Now assume it holds for some $k \in \mathbb{N}$. Then,

$$p_{k+1}q_k - p_k q_{k+1} = (a_{k+1}p_k + p_{k-1})q_k - (a_{k+1}q_k + q_{k-1})p_k$$

$$= a_{k+1}p_k q_k + p_{k-1}q_k - a_{k+1}q_k p_k - q_{k-1}p_k$$

$$= p_{k-1}q_k - p_k q_{k-1}$$

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

$$= -(-1)^{k-1} = (-1)^k = (-1)^{(k+1)-1}$$

Therefore the Lemma holds.

One direct result of this lemma is that in defining p_k and q_k this way, the $\gcd(p_k, q_k) = 1$ so p_k and q_k are precisely the same ones we were discussing before.

Corollary 2.2: If p_k and q_k are defined as in Lemma 2.2, then $\gcd(p_k, q_k) = 1$.

Proof: First recall that $ax + by = 1$ has an integer solution if and only if $\gcd(p_k, q_k) = 1$.

Suppose k is odd. Then

$$p_{k+1}q_k - p_k q_{k+1} = 1$$

$$\Rightarrow p_{k+1}q_k + p_k(-q_{k+1}) = 1$$

$$\Rightarrow \gcd(p_k, q_k) = 1.$$

Suppose k is even. Then

$$p_{k+1}q_k - p_k q_{k+1} = -1$$

$$\Rightarrow p_k q_{k+1} - p_{k+1}q_k = 1$$

$$\Rightarrow p_k q_{k+1} + (-p_{k+1})q_k = 1$$

$$\Rightarrow \gcd(p_k, q_k) = 1.$$

Lemma 2.3 Let $\frac{p_k}{q_k}$ be the k^{th} convergent of the continued fraction expansion of the irrational number x. If a and b are integers with $1 \le b < q_{k+1}$, then $|q_k x - p_k| \le |bx - a|$.

Comment: One interesting byproduct of this lemma is that, given the same conditions, if $|q_k x - p_k| \le |bx - a|$ then $|q_{k+1} x - p_{k+1}| \le |bx - a|$ (something we did not cover but involves the fact that every convergent is closer to x than the last). Which means $\left| x - \frac{p_{k+1}}{q_{k+1}} \right| \le \left| x - \frac{a}{b} \right|$. This tells us that no rational number whose denominator is less than q_{k+1}, will be closer to the irrational number x than the convergent C_{k+1}. This, then, is a great method for finding rational approximations to irrational numbers. Merely calculate the convergents until you get to an acceptably sized denominator and, from this lemma, you will have the best approximation to x.

Proof: Let $n \in \mathbb{N}$ and let $a, b \in \mathbb{Z}$ such that $1 \le b < q_{n+1}$. First, we want to construct a system of equations

$$\begin{aligned} p_n s + p_{n+1} r &= a \\ q_n s + q_{n+1} r &= b \end{aligned} \quad \text{where } r, s \in \mathbb{Z}.$$

$p_n s = a - p_{n+1} r \Rightarrow s = \frac{a - p_{n+1} r}{p_n}$. So,

$$b = \frac{q_n a - q_n p_{n+1} r + p_n q_{n+1} r}{p_n}$$

$$= \frac{q_n a + r(p_n q_{n+1} - q_n p_{n+1})}{p_n}$$

$$= \frac{q_n a + r(-1)^{n+1}}{p_n} \quad \text{(by Lemma 2.2)}$$

$$\Rightarrow bp_n = q_n a + r(-1)^{n+1}$$

$$\Rightarrow bp_n - q_n a = r(-1)^{n+1}$$

$$\Rightarrow (-1)^{n+1}(bp_n - q_n a) = r\,.$$

Similarly, $(-1)^{n+1}(aq_{n+1} - p_{n+1}b) = s\,.$

Notice that s and r must be integers.

Claim 1: $s \neq 0$

Suppose $s = 0$. Then $aq_{n+1} = p_{n+1}b \Rightarrow \frac{a}{b} = \frac{p_{n+1}}{q_{n+1}}$ and since $\gcd(p_{n+1}, q_{n+1}) = 1$, q_{n+1} must divide b. Then $q_{n+1} \leq b$ which contradicts our hypothesis. So $s \neq 0$.

Case 1: $r = 0$

Then,

$$a = p_n s \text{ and } b = q_n s\,.$$

$$\Rightarrow |bx - a| = |q_n sx - p_n s|$$

$$= |s||q_n x - p_n|$$

$$\geq |q_n x - p_n| \qquad (\text{since } s \in \square)$$

Case 2: $r \neq 0$

Claim 2: r and s have opposite signs.

case 2.1: $r < 0$ (which forces $r \leq -1$ since $r \in \square$)

First consider $b = q_n s + q_{n+1} r$ which is equivalent to $q_n s = b - q_{n+1} r$. Since $1 \leq b < q_{n+1}$ and $r \leq -1$, we know that $q_n s > 0$ and so $s > 0$ (since $q_n > 0$). So r and s have opposite signs.

case 2.2: $r > 0$ (which forces $r \geq 1$ since $r \in \square$)

This case has the same strategy as the last. Consider $q_n s = b - q_{n+1} r$. $1 \leq b < q_{n+1}$ and $r \geq 1$, so $q_n s < 0$ which implies that $s < 0$ (since $q_n > 0$). Therefore r and s have opposite signs.

Claim 3: $(q_n x - p_n)$ and $(q_{n+1} x - p_{n+1})$ have opposite signs.

Recall that the sequence of convergents $\{C_k\}$ is alternating, so if $\frac{p_n}{q_n} < x$ then $\frac{p_{n+1}}{q_{n+1}} > x$,

or if $\frac{p_n}{q_n} > x$ then $\frac{p_{n+1}}{q_{n+1}} < x$.

case 3.1: $\frac{p_n}{q_n} < x$

If $\frac{p_n}{q_n} < x$ then $p_n < q_n x \Rightarrow 0 < q_n x - p_n$, similarly, $\frac{p_{n+1}}{q_{n+1}} > x \Rightarrow q_{n+1} x - p_{n+1} < 0$, and so $(q_n x - p_n)$ and $(q_{n+1} x - p_{n+1})$ have opposite signs.

case 3.2: $\frac{p_n}{q_n} > x$

If $\frac{p_n}{q_n} > x$ then $p_n > q_n x \Rightarrow 0 > q_n x - p_n$, similarly, $\frac{p_{n+1}}{q_{n+1}} < x \Rightarrow q_{n+1} x - p_{n+1} > 0$, and so $(q_n x - p_n)$ and $(q_{n+1} x - p_{n+1})$ have opposite signs.

The purpose of this is to establish the fact that $(q_n x - p_n)$ and $(q_{n+1} x - p_{n+1})$ have opposite signs and s and r also have opposite signs. Then $s(q_n x - p_n)$ and $r(q_{n+1} x - p_{n+1})$ must have the same sign. So,

$$|bx - a| = |(q_n s + q_{n+1} r)x - (p_n s + p_{n+1} r)|$$

$$= |s(q_n x - p_n) + r(q_{n+1} x - p_{n+1})|$$

$$= |s(q_n x - p_n)| + |r(q_{n+1} x - p_{n+1})|$$

(since the two terms have the same sign)

$$= |s||(q_n x - p_n)| + |r||(q_{n+1} x - p_{n+1})|$$

$$\geq |s||(q_n x - p_n)|$$

$$\geq |(q_n x - p_n)| \text{ . (since } s \in \mathbb{Z})$$

Therefore $|q_k x - p_k| \leq |bx - a|$ which concludes our proof.

Lemma 2.4: Let x be an irrational number. If the rational number $\frac{a}{b}$, where $b \geq 1$ and $\gcd(a,b) = 1$, satisfies $\left| x - \frac{a}{b} \right| < \frac{1}{2b^2}$ then $\frac{a}{b}$ is one of the convergents in the continued fraction expansion of x.

Proof: Proof by contradiction. Let $\frac{p_n}{q_n}$ denote the convergents of the continued fraction expansion of x. Suppose the conditions hold but $\frac{a}{b}$ is not a convergent for the expansion of x. Then, since $\{q_n\}$ is an increasing sequence, there exists $k \in \mathbb{Z}$ such that

$$q_k < b < q_{k+1} \text{ (1).}$$

Then by Lemma 2.3,

$$|q_k x - p_k| \leq |bx - a|$$

$$\Rightarrow q_k \left| x - \frac{p_k}{q_k} \right| \le b \left| x - \frac{a}{b} \right| < b \frac{1}{2b^2} = \frac{1}{2b} \qquad \text{(by hypothesis)}$$

$$\Rightarrow \left| x - \frac{p_k}{q_k} \right| < \frac{1}{2bq_k} \quad (2)$$

Further, since $\frac{a}{b} \neq \frac{p_k}{q_k}$ we know that $bp_k \neq aq_k$ which implies

$$|bp_k - aq_k| \geq 1 \text{ (3).}$$

So,

$$\frac{1}{bq_k} \le \left| \frac{bp_k - aq_k}{bq_k} \right| = \left| \frac{p_k}{q_k} - \frac{a}{b} \right| \le \left| \frac{p_k}{q_k} - x \right| + \left| x - \frac{a}{b} \right| \quad \text{(triangle inequality and (3))}$$

$$< \frac{1}{2bq_k} + \frac{1}{2b^2} \text{ by (2) and hypothesis}$$

$$\text{So } \frac{1}{bq_k} < \frac{1}{2bq_k} + \frac{1}{2b^2} \Rightarrow \frac{2b^2 q_k}{bq_k} < \frac{2b^2 q_k}{2bq_k} + \frac{2b^2 q_k}{2b^2}$$

$$\Rightarrow 2b < b + q_k$$

$$\Rightarrow b < q_k$$

Which contradicts (1), therefore $\frac{a}{b} = \frac{p_n}{q_n}$ for some $n \in \square$. Further, since $\gcd(a,b) = 1$ and $\gcd(p_n, q_n) = 1$, we know that $a = p_n$ and $b = q_n$ which completes our proof. Now we are finally ready to prove Theorem 2.

Theorem 2: If x, y is a solution to $x^2 - dy^2 = 1$ then there exists a natural number k such that $x = p_k$ and $y = q_k$ where $C_k = \frac{p_k}{q_k}$ is the k^{th} convergent for the continued fraction expansion of $\sqrt{d}$.

Proof: Let $x, y \in \mathbb{Z}$ such that $x^2 - dy^2 = 1$. Then

$$(x - \sqrt{d}y)(x + \sqrt{d}y) = 1 \quad \text{(which means } x > \sqrt{d}y\text{)}$$

$$\Rightarrow x - \sqrt{d}y = \frac{1}{x + \sqrt{d}y}$$

$$\Rightarrow \frac{x}{y} - \sqrt{d} = \frac{1}{y(x + \sqrt{d}y)}$$

$$< \frac{\sqrt{d}}{y(\sqrt{d}y + \sqrt{d}y)} \quad \text{(since } \sqrt{d} > 1 \text{ and } x > \sqrt{d}y\text{)}$$

$$= \frac{\sqrt{d}}{2y^2\sqrt{d}}$$

$$= \frac{1}{2y^2}$$

Then by Lemma 2.4, there exists $k \in \mathbb{Z}$ such that $x = p_k$ and $y = q_k$ where $\frac{p_k}{q_k} = C_k$ for the continued fraction expansion of $\sqrt{d}$.

Chapter 7: General Solutions To $x^2 - dy^2 = 1$

Now we know all the solutions to $x^2 - dy^2 = 1$ can be found in the convergents of $\sqrt{d}$.

Our next goal is to find a way to figure out *which* convergents are going to be solutions.

To do this, we must first start with defining some things.

Lemma 3.1: Let $\sqrt{d} = [a_0; a_1, a_2, ...]$. Define

$$s_0 = 0,\ t_0 = 1,\ r_0 = \sqrt{d}$$

$$s_{k+1} = a_k t_k - s_k,\quad t_{k+1} = \frac{d - s_{k+1}^2}{t_k}, \text{ and } r_{k+1} = \frac{1}{r_k - a_k}$$

Then $\forall k \in \mathbb{N}$,

(a) s_k and t_k are integers with $t_k \neq 0$.

(b) t_k divides $d - s_k^2$

(c) $\dfrac{s_k + \sqrt{d}}{t_k} = r_k$

Proof: (by induction)

Base: (a) s_0 and t_0 are clearly integers by definition

(b) $t_0 = 1$ divides $d - s_k^2 = d$ (since 1 divides any integer d)

(c) $\dfrac{s_o + \sqrt{d}}{t_o} = \dfrac{0 + \sqrt{d}}{1} = \sqrt{d} = r_0$

Induction: Assume (a), (b), (c) hold for some $k \in \mathbb{N}$.

(a)

s_k, t_k, a_k are integers so $s_{k+1} = a_k t_k - s_k$ will also be an integer. Further

$$t_{k+1} = \frac{d - s_{k+1}^2}{t_k} = \frac{d - (a_k^2 t_k^2 - 2 s_k a_k t_k + s_k^2)}{t_k}$$

$$= \frac{d - s_k^2}{t_k} + 2 s_k a_k - a_k^2 t_k$$

Which is an integer since t_k divides $d - s_k^2$. Also $t_{k+1} \neq 0$ because this would imply that $d = s_{k+1}^2$ which means d is a perfect square.

(b)

$$t_{k+1} = \frac{d - s_{k+1}^2}{t_k} \Rightarrow t_{k+1} t_k = d - s_{k+1}^2$$

and since t_k is an integer this means that t_{k+1} must divide $d - s_{k+1}^2$.

(c)

$$r_{k+1} = \frac{1}{r_k - a_k} = \frac{1}{\frac{s_k + \sqrt{d}}{t_k} - a_k}$$

$$= \frac{t_k}{s_k + \sqrt{d} - a_k t_k}$$

$$= \frac{t_k}{\sqrt{d} - s_{k+1}}$$

$$= \frac{t_k}{d - s_{k+1}^2} (\sqrt{d} + s_{k+1})$$

$$= \frac{1}{t_{k+1}} (\sqrt{d} + s_{k+1})$$

$$= \frac{\sqrt{d} + s_{k+1}}{t_{k+1}}$$

Therefore (a), (b), and (c) hold.

Lemma 3.2: If $[a_0; a_1, a_2, ...]$ is the continued fraction expansion of $\sqrt{d}$ then

$\sqrt{d} = \frac{r_{k+1}p_k + p_{k-1}}{r_{k+1}q_k + q_{k-1}}$ where $\frac{p_k}{q_k}$ is the k^{th} convergent of $[a_0; a_1, a_2, ...]$ and r_{k+1} is defined as in Lemma 3.1.

Proof:

Let $k \in \mathbb{N}$. Then consider $[a_0; a_1, a_2, ..., a_k, r_{k+1}]$. Well,

$$[a_0; a_1, a_2, ..., a_k, r_{k+1}] = [a_0; a_1, a_2, ..., a_k + \frac{1}{r_{k+1}}]$$

$$= [a_0; a_1, a_2, ..., a_k + \frac{1}{\frac{1}{r_k - a_k}}]$$

$$= [a_0; a_1, a_2, ..., a_k + r_k - a_k]$$

$$= [a_0; a_1, a_2, ..., r_k]$$

Which holds for any $k \in \mathbb{N}$. Therefore $[a_0; a_1, a_2, ..., a_k, r_{k+1}] = [r_0] = r_0 = \sqrt{d}$.

Further, by lemma 2.1, $[a_0; a_1, a_2, ..., a_k, r_{k+1}] = \frac{r_{k+1}p_k + p_{k-1}}{r_{k+1}q_k + q_{k-1}}$. Therefore,

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

Theorem 3: If $\frac{p_k}{q_k}$ denotes the convergents of the continued fraction expansion of $\sqrt{d}$ then $p_k^2 - dq_k^2 = (-1)^{k+1}t_{k+1}$ where $t_{k+1} > 0$.

Proof:

$$\sqrt{d} = \frac{r_{k+1}p_k + p_{k-1}}{r_{k+1}q_k + q_{k-1}} \quad \text{(by Lemma 3.2)}$$

$$= \frac{\left(\frac{s_{k+1} + \sqrt{d}}{t_{k+1}}\right)p_k + p_{k-1}}{\left(\frac{s_{k+1} + \sqrt{d}}{t_{k+1}}\right)q_k + q_{k-1}}$$

$$= \frac{\left(s_{k+1} + \sqrt{d}\right)p_k + t_{k+1}p_{k-1}}{\left(s_{k+1} + \sqrt{d}\right)q_k + t_{k+1}q_{k-1}}$$

$$\Rightarrow \sqrt{d}\left[\left(s_{k+1} + \sqrt{d}\right)q_k + t_{k+1}q_{k-1}\right] = \left(s_{k+1} + \sqrt{d}\right)p_k + t_{k+1}p_{k-1}$$

$$\Rightarrow \sqrt{d}\left(q_k s_{k+1} + t_{k+1}q_{k-1} - p_k\right) = p_k s_{k+1} + t_{k+1}p_{k-1} - dq_k$$

The right hand side of the above expression must be rational, while the left hand side must be irrational unless they both equal zero. This then implies that

$$\left(q_k s_{k+1} + t_{k+1}q_{k-1} - p_k\right) = 0 \text{ and } p_k s_{k+1} + t_{k+1}p_{k-1} - dq_k = 0$$

$$\Rightarrow q_k s_{k+1} + t_{k+1}q_{k-1} = p_k \text{ and } p_k s_{k+1} + t_{k+1}p_{k-1} = dq_k$$

So,

$$p_k^2 - dq_k^2 = p_k\left(q_k s_{k+1} + t_{k+1}q_{k-1}\right) - \left(p_k s_{k+1} + t_{k+1}p_{k-1}\right)q_k$$

$$= p_k t_{k+1}q_{k-1} - t_{k+1}p_{k-1}q_k$$

$$= \left(p_k q_{k-1} - p_{k-1}q_k\right)t_{k+1}$$

$$= (-1)^{k-1}t_{k+1} \text{ (by lemma 2.2)}$$

$$= (-1)^{k+1}t_{k+1}$$

Finally, we want to prove that $t_k > 0$ for all $k \in \mathbb{N}$. We know $t_0, t_1 > 0$ so lets assume it holds for some $k \in \mathbb{N}$. We know our convergents are an alternating sequence with limit $\sqrt{d}$, so we have two cases.

Case 1: $C_{k-1} < \sqrt{d} < C_k$

Then $p_{k-1}^2 - dq_{k-1}^2 < 0$ and $p_k^2 - dq_k^2 > 0$. Now consider

$$\frac{p_k^2 - dq_k^2}{p_{k-1}^2 - dq_{k-1}^2} = \frac{(-1)^{k+1} t_{k+1}}{(-1)^k t_k} = -\frac{t_{k+1}}{t_k}.$$

Since $\frac{p_k^2 - dq_k^2}{p_{k-1}^2 - dq_{k-1}^2} < 0$, $\frac{t_{k+1}}{t_k} > 0$. This implies that $t_{k+1} > 0$ since $t_k > 0$.

Case 2: $C_k < \sqrt{d} < C_{k-1}$

Then $p_{k-1}^2 - dq_{k-1}^2 > 0$ and $p_k^2 - dq_k^2 < 0$. Then consider

$$\frac{p_k^2 - dq_k^2}{p_{k-1}^2 - dq_{k-1}^2} = \frac{(-1)^{k+1} t_{k+1}}{(-1)^k t_k} = -\frac{t_{k+1}}{t_k}.$$

Since $\frac{p_k^2 - dq_k^2}{p_{k-1}^2 - dq_{k-1}^2} < 0$, $\frac{t_{k+1}}{t_k} > 0$. This implies that $t_{k+1} > 0$ since $t_k > 0$. Our proof of Theorem 3 is now complete.

We know that all of our solutions can be found among our convergents of the continued fraction expansion of $\sqrt{d}$, and now we know for all of our convergents $C_k = \frac{p_k}{q_k}$, $p_k^2 - dq_k^2 = (-1)^{k+1} t_{k+1}$. So if we can figure out when $t_{k+1} = 1$, we will be able to find all of our solutions. We will now discuss a result that allows us to do this.

Theorem 4: If n is the length of the period of the expansion of $\sqrt{d}$, then $t_k = 1$ if and only if n divides k.

Recall from our discussion on continued fractions that if d is an integer that is not a perfect square then the continued fraction expansion of $\sqrt{d}$ is purely periodic. Therefore, we can let $\sqrt{d} = [a_o; \overline{a_1, a_2, ..., a_n}]$. We will start by proving that if n divides k then $t_k = 1$.

Proof:

Let n divide k. Recall, from Lemma 3.2, that $[a_o; a_1, a_2, ...] = [a_o; a_1, a_2, ..., r_k]$ for any k. Therefore, $[a_o; \overline{a_1, a_2, ..., a_n}] = [a_0; r_1]$ and $[a_o; \overline{a_1, a_2, ..., a_n}] = [a_o; a_1, a_2, ..., a_n, r_{n+1}]$ and since our continued fraction is purely periodic, $[a_o; \overline{a_1, a_2, ..., a_n}] = [a_o; a_1, a_2, ..., a_n, r_1]$. So $r_{n+1} = r_1$. This would hold true for any multiple of n as well, which means $r_{mn+1} = r_1$ where $m \in \square$. So,

$$\frac{s_{mn+1} + \sqrt{d}}{t_{mn+1}} = \frac{s_1 + \sqrt{d}}{t_1} \text{ (by lemma 3.1)}$$

$\Rightarrow s_{mn+1} = s_1$ and $t_{mn+1} = t_1$ since all the s_i and t_i terms are integers and $\sqrt{d}$ is irrational. Further,

$$t_1 = \frac{d - s_1^2}{t_0} \quad \text{(by definition)}$$

$$= d - s_1^2 \quad (\text{since } t_0 = 1)$$

$$= d - s_{mn+1}^2 \quad (\text{since } s_{mn+1} = s_1)$$

$$= t_{mn+1}t_{mn} \quad \text{(by definition } \; t_{mn+1} = \frac{d - s_{mn+1}^2}{t_{mn}}\text{)}$$

$$= t_1 t_{mn} \quad \text{(since } \; t_{mn+1} = t_1\text{)}$$

Therefore, $t_{mn} = 1$, for all $m \in \square$. Well, n divides k, so $t_k = 1$.

Now suppose $t_k = 1$. We want to show that n must divide k.

$$r_k = \frac{s_k + \sqrt{d}}{t_k} \quad \text{(by lemma 3.1)}$$

$$= s_k + \sqrt{d} \quad \text{(since } \; t_k = 1\text{)} \quad (1)$$

$$\Rightarrow \lfloor r_k \rfloor = s_k + \left\lfloor \sqrt{d} \right\rfloor \quad \text{(since } \; s_k \text{ is an integer)}$$

$$\Rightarrow \lfloor r_k \rfloor = s_k + a_0$$

$$\text{(since } \; a_0 = \left\lfloor \sqrt{d} \right\rfloor \text{ from our discussion on continued fractions)} \quad (2)$$

$$r_{k+1} = \frac{1}{r_k - a_k} \quad \text{(by definition)}$$

$$\Rightarrow r_{k+1} = \frac{1}{r_k - \lfloor r_k \rfloor}$$

(from our discussion on continued fractions)

$$\Rightarrow r_k = \frac{1}{r_{k+1}} + \lfloor r_k \rfloor \quad (3)$$

Further,

$$\sqrt{d} = [a_o; r_1] = \frac{1}{r_1} + a_0 \quad (4).$$

And, $$\sqrt{d} = r_k - s_k \quad \text{(by (1))}$$

$$= \frac{1}{r_{k+1}} + \lfloor r_k \rfloor - s_k \quad \text{(by (3))}$$

$$= \frac{1}{r_{k+1}} + (s_k + a_0) - s_k \quad \text{(by (2))}$$

$$= \frac{1}{r_{k+1}} + a_0 .$$

Therefore,

$$a_0 + \frac{1}{r_1} = a_0 + \frac{1}{r_{k+1}} \quad \text{(by (4))}$$

$$\Rightarrow r_1 = r_{k+1}$$

And therefore the period length, n, of our continued fraction must divide k.

Now we have all the means necessary to define a general solution.

Theorem 5: Let $\frac{p_k}{q_k}$ be the convergents of the continued fraction expansion of $\sqrt{d}$ and let n be the length of the period of the expansion. If n is even, then all positive solutions of $x^2 - dy^2 = 1$ are given by $x = p_{kn-1}$ and $y = q_{kn-1}$ where $k \in \mathbb{N}$. If n is odd, then all positive solutions of $x^2 - dy^2 = 1$ are given by $x = p_{2kn-1}$ and $y = q_{2kn-1}$ where $k \in \mathbb{N}$.

Proof:

Case 1: n is even.

Let r, s be a solution to $x^2 - dy^2 = 1$. Then by Theorem 2, there exists a natural number z such that $x = p_z$ and $y = q_z$. Further, by Theorem 3 we know that $p_z^2 - dq_z^2 = (-1)^{z+1} t_{z+1}$, i.e., $p_{z-1}^2 - dq_{z-1}^2 = (-1)^z t_z = 1$. So $t_z = 1$ and $(-1)^z = 1$. For this to be the case, by Theorem 4, z must be a multiple of n. This ensures that $t_z = 1$ and, since n is even, that $(-1)^z = 1$.

Therefore $r = p_{kn-1}$ and $s = q_{kn-1}$ for some $k \in \mathbb{N}$.

For the other direction, if we let $r = p_{kn-1}$ and $s = q_{kn-1}$. Then $r^2 - ds^2 = (-1)^{kn} t_{kn}$.

Further, since n divides kn, $t_{kn} = 1$ and since n is even, $(-1)^{kn} = 1$. Therefore, $r^2 - ds^2 = 1$.

Case 2: n is odd.

Through the same argument $t_z = 1$ and $(-1)^z = 1$. For this to be the case, by Theorem 4, z must be a multiple of n, and z must be even in order for $(-1)^z = 1$. Therefore $r = p_{2kn-1}$ and $s = q_{2kn-1}$ for some $k \in \mathbb{N}$.

For the other direction, if we let $r = p_{2kn-1}$ and $s = q_{2kn-1}$. Then $r^2 - ds^2 = (-1)^{2kn} t_{2kn}$.

Further, since n divides $2kn$, $t_{2kn} = 1$ and since $2kn$ is even, $(-1)^{2kn} = 1$. Therefore, $r^2 - ds^2 = 1$ and this completes our proof.

Now we have a method to find all of our solutions. Lets look at an example of when $d = 8$. The continued fraction expansion for $\sqrt{8}$ is $[2;\overline{1,4}]$. By Theorem 5, since the period length of $[2;\overline{1,4}]$ is 2, which is even, all of our solutions are $x = p_{2k-1}$ and $y = q_{2k-1}$. So lets check the first few.

When $k=1$, we have p_1 and q_1 as a solution since,

$$\frac{p_1}{q_1}=C_1=2+\frac{1}{1}=\frac{3}{1} \text{ and } 3^2-8(1)^2=1.$$

When $k=2$ we have,

$$\frac{p_3}{q_3}=C_3=2+\frac{1}{1+\frac{1}{4+\frac{1}{1}}}=\frac{17}{6} \text{ and } 17^2-8(6)^2=1.$$

When $k=3$,

$$\frac{p_5}{q_5}=C_5=2+\frac{1}{1+\frac{1}{4+\frac{1}{1+\frac{1}{4+\frac{1}{1}}}}}=\frac{99}{35} \text{ and } 99^2-8(35)^2=1.$$

As you can see, calculating the convergents can take a lot of work. Imagine if we wanted to calculate the first 20 convergents for $d=8$. Luckily, there is another way to find all of our solutions that only requires us to calculate one convergent. As we will see from the next two theorems, if x_1, y_1 is our first non-trivial solution, which is typically referred to as the fundamental solution, then $x_k+y_k\sqrt{d}=(x_1+y_1\sqrt{d})^k$.

Theorem 6: Let x_1, y_1 be the fundamental solution of $x^2-dy^2=1$. Then every pair of integers x_k, y_k defined by the condition $x_k+y_k\sqrt{d}=(x_1+y_1\sqrt{d})^k$, where $k\in\square$, is also a positive solution.

Proof: Let $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^k$. We first want to show that $x_k - y_k\sqrt{d} = (x_1 - y_1\sqrt{d})^k$ when $k > 1$. Well,

$$x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^k = (x_1 + y_1\sqrt{d})^{k-1}(x_1 + y_1\sqrt{d})$$

$$= (x_{k-1} + y_{k-1}\sqrt{d})(x_1 + y_1\sqrt{d})$$

$$= (x_1x_{k-1} + y_1y_{k-1}d) + (x_1y_{k-1} + x_{k-1}y_1)\sqrt{d}$$

$$\Rightarrow x_k = x_1x_{k-1} + y_1y_{k-1}d \text{ and } y_k = x_1y_{k-1} + x_{k-1}y_1.$$

So,

$$x_k - y_k\sqrt{d} = x_1x_{k-1} + y_1y_{k-1}d - (x_1y_{k-1} + x_{k-1}y_1)\sqrt{d}$$

$$= (x_{k-1} - y_{k-1}\sqrt{d})(x_1 - y_1\sqrt{d})$$

$$= (x_1 - y_1\sqrt{d})^{k-1}(x_1 - y_1\sqrt{d})$$

$$= (x_1 - y_1\sqrt{d})^k$$

Then the proof follows rather easily since,

$$x_k^2 - dy_k^2 = (x_k - y_k\sqrt{d})(x_k + y_k\sqrt{d})$$

$$= (x_1 - y_1\sqrt{d})^k(x_1 + y_1\sqrt{d})^k$$

$$= [(x_1 - y_1\sqrt{d})(x_1 + y_1\sqrt{d})]^k$$

$$= (x_1^2 - dy_1^2)^k = (1)^k = 1$$

Theorem 7: If x_1, y_1 is the fundamental solution then all solutions x_k, y_k must be of the form $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^k$.

Proof: This will be a proof by contradiction. If we suppose not, then there exists $a, b \in \mathbb{Z}$ such that $a^2 - db^2 = 1$ and $(a + b\sqrt{d}) \neq (x_1 + y_1\sqrt{d})^k$ for all $k \in \mathbb{Z}$. Then there must exist $n \in \mathbb{Z}$ such that

$$(x_1 + y_1\sqrt{d})^n < a + b\sqrt{d} < (x_1 + y_1\sqrt{d})^{n+1}$$

$$\Rightarrow x_n + y_n\sqrt{d} < a + b\sqrt{d} < (x_n + y_n\sqrt{d})^n (x_1 + y_1\sqrt{d})$$

$$\Rightarrow (x_n - y_n\sqrt{d})(x_n + y_n\sqrt{d}) < (x_n - y_n\sqrt{d})(a + b\sqrt{d}) < (x_n - y_n\sqrt{d})(x_n + y_n\sqrt{d})^n (x_1 + y_1\sqrt{d})$$

since $(x_n - y_n\sqrt{d})$ must be positive.

$$\Rightarrow x_n^2 - dy_n^2 < x_n a + x_n b\sqrt{d} - y_n a\sqrt{d} - y_n bd < (x_n^2 - dy_n^2)(x_1 + y_1\sqrt{d})$$

$$\Rightarrow 1 < (x_n a - y_n bd) + (x_n b - y_n a)\sqrt{d} < (x_1 + y_1\sqrt{d})$$

Let $w = x_n a - y_n bd$ and $z = x_n b - y_n a$. Then $1 < w + z\sqrt{d} < (x_1 + y_1\sqrt{d})$. Further,

$$\begin{aligned} w^2 - dz^2 &= (x_n a - y_n bd)^2 - d(x_n b - y_n a)^2 \\ &= x_n^2 a^2 - 2x_n y_n abd + y_n^2 b^2 d^2 - x_n^2 b^2 d + 2x_n y_n abd - y_n^2 a^2 d \\ &= (x_n^2 - dy_n^2)(a^2 - db^2) = 1. \end{aligned}$$

So we have that w, z is a solution and $1 < w + z\sqrt{d} < (x_1 + y_1\sqrt{d})$, which means $w - z\sqrt{d} \in (0,1)$. Therefore $w > 0$ and

$$w + z\sqrt{d} - (w - z\sqrt{d}) > 1 - 1$$

$$\Rightarrow 2z\sqrt{d} > 0$$

$$\Rightarrow z > 0$$

So we have a positive solution w, z such that $w + z\sqrt{d} < x_1 + y_1\sqrt{d}$ and so x_1, y_1 is not the fundamental solution, which is a contradiction. Therefore all solutions must be of the form $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^k$.

Example: Let's take a look at the case of $d = 7$. By theorem 5, $x_1 = p_3$ and $y_1 = q_3$ since $\sqrt{7} = [2;\overline{1,1,1,4}]$. Well,

$$\frac{p_3}{q_3} = C_3 = 2 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}} = 2 + \frac{2}{3} = \frac{8}{3}$$

So $x_1 = 8$ and $y_1 = 3$. We can check our work to see that indeed $8^2 - 7(3)^2 = 1$. Then by theorems 6 and 7, the next solution x_2, y_2 can be found by the relation

$$\begin{aligned} x_2 + y_2\sqrt{d} &= (x_1 + y_1\sqrt{d})^2 \\ &= (8 + 3\sqrt{7})^2 \\ &= 127 + 48\sqrt{7} \end{aligned}$$

So $x_2 = 127$ and $y_2 = 48$. We can check to make sure that indeed $127^2 - 7(48)^2 = 16129 - 16128 = 1$. Now lets find the next solution.

$$\begin{aligned} x_3 + y_3\sqrt{d} &= (x_1 + y_1\sqrt{d})^3 \\ &= (x_1 + y_1\sqrt{d})^2 (x_1 + y_1\sqrt{d}) \\ &= (x_2 + y_2\sqrt{d})(x_1 + y_1\sqrt{d}) \\ &= (127 + 48\sqrt{7})(8 + 3\sqrt{7}) \\ &= 2024 + 765\sqrt{7} \end{aligned}$$

So $x_3 = 2024$ and $y_3 = 765$. Again, lets check and see that

$2024^2 - 7(765)^2 = 4096576 - 4096575 = 1$.

As you can see, this method seems to be much easier than calculating all of them using convergents. What if, however, it is difficult to calculate even the fundamental solution. For example, in the case when $d = 1000099$, the continued fraction expansion of $\sqrt{d}$ has period length 2174, which means you would need to calculate C_{2173} in order to find the fundamental solution. Thankfully, there are other methods for finding the fundamental solution.

Chapter 8: Finding the Fundamental Solution

First, when presented with a specific value for d you want to check if $d = a^2 - 1$ for some $a \in \mathbb{Z}$ because then $x_1 = a$ and $y_1 = 1$. It is relatively obvious to see why, but I will give the proof just in case.

Corollary 7.1: If $d = a^2 - 1$ for some $a \in \mathbb{Z}$ then $x_1 = a$ and $y_1 = 1$.

Proof: Let $d = a^2 - 1$. Then $a^2 - d1^2 = a^2 - (a^2 - 1) = 1$. Further, suppose there exists $b \in \mathbb{Z}$ such that $1 < b < a$ and $b^2 - dc^2 = 1$ for some $c \in \mathbb{Z}$. Since $b < a$ then $c < 1$, which is a contradiction since $c \in \mathbb{Z}$.

So for example, when $d = 63$ we know $x_1 = 8$ and $y_1 = 1$. Another way to find the fundamental solution is by using the **Chakravala Method.** The method was developed in India around 628 A.D. by the mathematician Brahmagupta who was trying to find the

first solution to $x^2 - 61y^2 = 1$, which would not be solved in the western world until 1767 by Langrange who had to calculate the 21st convergent. His work was based off of a Lemma done around the same time by a mathematician named Bhaskara which stated that if $Ny^2 + k = x^2$ (or $x^2 - Ny^2 = k$) then $N\left(\frac{my+x}{k}\right)^2 + \frac{m^2-N}{k} = \left(\frac{mx+Ny}{k}\right)^2$.

Chakravala Method: The best way to illustrate the method is to give a specific example. So let's look at the case $x^2 - 31y^2 = 1$.

First we need to find any positive integer solution to $x^2 - 31y^2 = k$, so we will take $x = 6$, $y = 1$, and $k = 5$. Then by Bhaskara's Lemma we have,

$31\left(\frac{m+6}{5}\right)^2 + \frac{m^2-31}{5} = \left(\frac{6m+31}{5}\right)^2$. Now, we want $\frac{m+6}{5}$ to be an integer so we set $m = 5t - 1$. Next take t so that $\left|m^2 - 31\right|$ is minimized, which occurs when $t = 1$. Then this makes $m = 4$ and now we have $31\left(\frac{4+6}{5}\right)^2 + \frac{(4)^2-31}{5} = \left(\frac{6(4)+31}{5}\right)^2$.

$$\Rightarrow 31(2)^2 + (-3) = (11)^2$$

$$\Rightarrow (11)^2 - 31(2)^2 = -3$$

Since we did not get a solution we must repeat the process with $x = 11$, $y = 2$, and $k = -3$. Then we have, $31\left(\frac{2m+11}{-3}\right)^2 + \frac{m^2-31}{-3} = \left(\frac{11m+62}{-3}\right)^2$. We want $m = 3t - 1$ so that $\frac{2m+11}{-3}$ is an integer. We then choose $t = 2$ to minimize $\left|m^2 - 31\right|$ which makes $m = 5$. Then we have $31\left(\frac{2(5)+11}{-3}\right)^2 + \frac{(5)^2-31}{-3} = \left(\frac{11(5)+62}{-3}\right)^2$

$$\Rightarrow 31(-7)^2 + 2 = (-39)^2$$

$$\Rightarrow (39)^2 - 31(7)^2 = 2$$

Again, we did not get a solution so we must repeat the process. We now have $x = 39$, $y = 7$, and $k = 2$. Then, $31\left(\frac{7m+39}{2}\right)^2 + \frac{m^2-31}{2} = \left(\frac{39m+217}{2}\right)^2$. Let $m = 2t+1$, and choose $t = 2$ to minimize $|m^2 - 31|$. Then we have,

$31\left(\frac{7(5)+39}{2}\right)^2 + \frac{(5)^2-31}{2} = \left(\frac{39(5)+217}{2}\right)^2 \Rightarrow (206)^2 - 31(37)^2 = -3$. If we follow the same method again for $x = 206$, $y = 37$, and $k = -3$ we come to the result that $(657)^2 - 31(118)^2 = 5$. One more time gives us $(1520)^2 - 31(273)^2 = 1$ which is our solution.

We now have two general solutions to the equation $x^2 - dy^2 = 1$ and some different methods for making the process a bit less strenuous. We will now focus on the case when $x^2 - dy^2 = -1$.

Chapter 9: General Solutions To $x^2 - dy^2 = -1$

To find a general solution for $x^2 - dy^2 = -1$, we will use some of the same techniques as before. Notice first that lemmas 2.1- 2.4 are still true for this case so we can jump to the proof that all solutions to $x^2 - dy^2 = -1$ are found among the convergents for the continued fraction expansion of $\sqrt{d}$.

Theorem 8: If x, y is a solution to $x^2 - dy^2 = -1$ then there exists a natural number k such that $x = p_k$ and $y = q_k$ where $C_k = \frac{p_k}{q_k}$ is the k^{th} convergent for the continued fraction expansion of $\sqrt{d}$.

Proof: Let x, y be a solution to the equation. Then,

$$(x - y\sqrt{d})(x + y\sqrt{d}) = -1$$

$$\Rightarrow (x - y\sqrt{d}) = \frac{-1}{(x + y\sqrt{d})}$$

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

Now $x + y\sqrt{d} > 0$, so $x - y\sqrt{d} < 0 \Rightarrow \frac{x}{y} - \sqrt{d} < 0$. Further $\frac{-1}{y(x + y\sqrt{d})} < 0$ since $y(x + y\sqrt{d}) > 0$. So we then have,

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

$$< \frac{1}{y^2 + y^2\sqrt{d})} \quad (\text{since } x > y)$$

$$< \frac{1}{2y^2} \quad (\text{since } \sqrt{d} > 0)$$

Then by Lemma 2.4, there exists a $k \in \mathbb{N}$ such that $x = p_k$ and $y = q_k$ where $\frac{p_k}{q_k} = C_k$ for the continued fraction expansion of $\sqrt{d}$. So again, we can find all of our solutions from our convergents. Further, we can use Theorems 3 and 4 to find a general solution.

Theorem 9: Let $\frac{p_k}{q_k}$ be the convergents of the continued fraction expansion of $\sqrt{d}$ and let n be the length of the period of the expansion. If n is even, then there are no positive integer solutions to $x^2 - dy^2 = -1$. If n is odd, then all positive solutions of $x^2 - dy^2 = -1$ are given by $x = p_{(2k-1)n-1}$ and $y = q_{(2k-1)n-1}$ where $k \in \mathbb{Z}$.

Proof: The proof of this follows directly from previous theorems.

Let $s, r \in \mathbb{Z}$ such that $s^2 - dr^2 = -1$. By Theorem 8, there must exist $z \in \mathbb{Z}$ such that $s = p_z$ and $r = q_z$. Further, by Theorem 3, $p_z^2 - dq_z^2 = (-1)^{z+1} t_{z+1}$ which implies that $(-1)^{z+1} = 1$ and $t_{z+1} = 1$.

Case 1: n is even

Since $(-1)^{z+1} t_{z+1} = -1$, $t_{z+1} = 1$ and $(-1)^{z+1} = -1$. Then by Theorem 4, n must divide $z+1$, which means that $z+1$ must be even. Then $(-1)^{z+1} = 1$ which is a contradiction. So there can not be any solutions when n is even. Therefore there are no solutions since we have established that all solutions must be found among our convergents.

Case 2: n is odd

Since $(-1)^{z+1} t_{z+1} = -1$, this implies that $t_{z+1} = 1$ and $(-1)^{z+1} = -1$. Then by Theorem 4, n must divide $z+1$, further, since $(-1)^{z+1} = -1$, $z+1$ must be odd. So there exists $s \in \mathbb{Z}$ such that $ns = z+1$. Further, n is odd and $z+1$ is odd, so s must be odd also. Then there must exist $k \in \mathbb{Z}$ such that $s = 2k-1$. Then $(2k-1)n - 1 = z$.

For the other direction. Let $(2k-1)n-1=z$ for some $k\in\mathbb{Z}$. Then $2k-1$ divides $z+1$. By Theorem 4, this implies $t_{z+1}=1$. Further, since $2k-1$ is odd and n is odd, $(-1)^{z+1}=-1$. Then by theorem 3, $p_z^2-dq_z^2=(-1)^{z+1}t_{z+1}=-1$. Therefore solutions exist only when our period length is odd. If this is the case, then all solutions are of the form $x=p_{(2k-1)n-1}$ and $y=q_{(2k-1)n-1}$.

Example: Take a look at the case when $d=29$. $\sqrt{29}=[5;\overline{2,1,1,2,10}]$ so the continued fraction expansion of $\sqrt{29}$ has period length of 5. Therefore there are solutions to $x^2-29y^2=-1$. The first one occurring at $x=p_{(2-1)5-1}=p_4$ and $y=q_{(2-1)5-1}=q_4$. So let us calculate them.

$$\frac{p_4}{q_4}=C_4=5+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2}}}}=\frac{70}{13}$$

And if we check our answer, indeed $70^2-29(13)^2=-1$.

We can also find another general solution which is similar to the result from Theorems 6 and 7. By Theorems 10 and 11, we will see that if there exists a fundamental solution x_1,y_1 (which means the period length must be odd), then all solutions are of the form $x_k+y_k\sqrt{d}=(x_1+y_1\sqrt{d})^{2k-1}$ where $k\in\mathbb{Z}$. Further, for any k, x_k, y_k will be a solution.

Theorem 10: Let the period length of the continued fraction expansion of $\sqrt{d}$ be odd. Let x_1,y_1 be the fundamental solution of $x^2-dy^2=-1$. Then every pair of integers x_k,

y_k defined by the condition, $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^{2k-1}$ where $k \in \mathbb{N}$, is also a positive solution.

Proof: Let $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^{2k-1}$ for some $k \in \mathbb{N}$. Then, as we established earlier, $x_k - y_k\sqrt{d} = (x_1 - y_1\sqrt{d})^{2k-1}$. So,

$$\begin{aligned} x_k^2 - dy_k^2 &= (x_k + y_k\sqrt{d})(x_k - y_k\sqrt{d}) \\ &= (x_1 + y_1\sqrt{d})^{2k-1}(x_1 - y_1\sqrt{d})^{2k-1} \\ &= [(x_1 + y_1\sqrt{d})(x_1 - y_1\sqrt{d})]^{2k-1} \\ &= (x_1^2 - dy_1^2)^{2k-1} \\ &= (-1)^{2k-1} \\ &= -1 \end{aligned}$$

Theorem 11: If x_1, y_1 is the fundamental solution to $x^2 - dy^2 = -1$ then all solutions x_k, y_k must be of the form $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^{2k-1}$.

Proof: (By contradiction). Suppose not. Then there exists $a, b \in \mathbb{N}$ such that $a^2 - db^2 = -1$ and $a + b\sqrt{d} \neq (x_1 + y_1\sqrt{d})^{2k-1}$, for all $k \in \mathbb{N}$. Then there exists $n \in \mathbb{N}$ such that

$$(x_1 + y_1\sqrt{d})^{2n-1} < a + b\sqrt{d} < (x_1 + y_1\sqrt{d})^{2n+1}$$

$$\Rightarrow x_n + y_n\sqrt{d} < a + b\sqrt{d} < (x_n + y_n\sqrt{d})(x_1 + y_1\sqrt{d})^2$$

$$\Rightarrow (x_n - y_n\sqrt{d})(x_n + y_n\sqrt{d}) > (x_n - y_n\sqrt{d})(a + b\sqrt{d}) > (x_n - y_n\sqrt{d})(x_n + y_n\sqrt{d})(x_1 + y_1\sqrt{d})^2$$

since $x_n - y_n\sqrt{d} < 0$.

$$\Rightarrow -(x_1 - y_1\sqrt{d}) < (x_1 - y_1\sqrt{d})(x_n - y_n\sqrt{d})(a + b\sqrt{d}) < (x_1 + y_1\sqrt{d})$$

Further,

$$(x_1 - y_1\sqrt{d})(x_n - y_n\sqrt{d})(a + b\sqrt{d}) = (x_1x_na - x_1y_nbd - y_1x_nbd + y_1y_nad)$$

$$+(x_1x_nb - x_1y_na - y_1x_na + y_1y_nbd)\sqrt{d}$$

Let $r = (x_1x_na - x_1y_nbd - y_1x_nbd + y_1y_nad)$ and $s = (x_1x_nb - x_1y_na - y_1x_na + y_1y_nbd)$. Then,

$$r^2 - ds^2 = x_1^2x_n^2a^2 + x_1^2y_n^2b^2d^2 + y_1^2x_n^2b^2d^2 + y_1^2y_n^2a^2d^2$$

$$-x_1^2x_n^2b^2d - x_1^2y_n^2a^2d - y_1^2x_n^2a^2d - y_1^2y_n^2b^2d^3$$

$$\Rightarrow r^2 - ds^2 = (x_1^2 - dy_1^2)(x_n^2 - dy_n^2)(a^2 - db^2)$$

$$= (-1)^3$$

$$= -1$$

Now we have a solution r, s where $-(x_1 - y_1\sqrt{d}) < r + s\sqrt{d} < x_1 + y_1\sqrt{d}$.

Further, since $x_1 - y_1\sqrt{d} < 0 \Rightarrow -(x_1 - y_1\sqrt{d}) > 0$. Then, $r + s\sqrt{d} > 0$. This leads to the fact then that $r - s\sqrt{d} < 0 \Rightarrow r < s\sqrt{d}$. Therefore,

$$0 < r + s\sqrt{d} < s\sqrt{d} + s\sqrt{d} = 2s\sqrt{d} \Rightarrow s > 0.$$

Case 1: $r > 0$.

Then we have a contradiction, since r, s is a positive solution and $r + s\sqrt{d} < x_1 + y_1\sqrt{d}$ means x_1, y_1 is not the fundamental solution.

Case 2: $r < 0$

Then consider $-r, s$ which would be a positive solution. Further,

$-(x_1 - y_1\sqrt{d}) < r + s\sqrt{d} \Rightarrow -r - s\sqrt{d} < x_1 - y_1\sqrt{d}$. Since $-r - s\sqrt{d}$ and $x_1 - y_1\sqrt{d}$ are both negative, $\left|-r - s\sqrt{d}\right| > \left|x_1 - y_1\sqrt{d}\right|$. Further,

$$\left|(-r - s\sqrt{d})(-r + s\sqrt{d})\right| = \left|(x_1 - y_1\sqrt{d})(x_1 + y_1\sqrt{d})\right|$$

$$\Rightarrow \left|-r - s\sqrt{d}\right|(-r + s\sqrt{d}) = \left|x_1 - y_1\sqrt{d}\right|(x_1 + y_1\sqrt{d})$$

And since $\left|-r - s\sqrt{d}\right| > \left|x_1 - y_1\sqrt{d}\right|$, this implies $-r + s\sqrt{d} < x_1 + y_1\sqrt{d}$.

Then we have a positive solution $-r, s$ such that $-r + s\sqrt{d} < x_1 + y_1\sqrt{d}$ which implies that x_1, y_1 is not the fundamental solution which is a contradiction. Therefore all solutions must be of the form $x_k + y_k\sqrt{d} = (x_1 + y_1\sqrt{d})^{2k-1}$. Now we also have two ways of calculating the solutions to $x^2 - dy^2 = -1$. Next we will take a look at some problems involving Pell's equation.

Chapter 10: Problems Involving Pell's Equation

A.) Determine the smallest value of n such that $1+5+9+13+.....$ is a perfect square. (A number of the form $1+5+9+13+.....$ is often referred to as a hexagonal number.)

Let $n \in \square$. Then $1+5+9+13+..... = \sum_0^n 4i + 1$

$$= 4\sum_0^n i \; + \; \sum_0^n 1$$

$$= 4\frac{n(n+1)}{2} + n + 1$$

$$= 2n^2 + 3n + 1$$

So we want to find the smallest value of n such that $2n^2 + 3n + 1 = k^2$ for some $k \in \mathbb{Z}$. Well, $2n^2 + 3n + 1 = k^2 \Rightarrow (2n+1)(n+1) = k^2$ and since $2n+1$ and $n+1$ are relatively prime, there must exist $r, s \in \mathbb{Z}$ such that $2n+1 = r^2$ and $n+1 = s^2$. Further,

$$n + 1 = s^2 \Rightarrow n = s^2 - 1$$

Then,

$$r^2 = 2(s^2 - 1) + 1$$

$$\Rightarrow r^2 - 2s^2 = -1.$$

From our discussion on Pell's equation we know the smallest positive solution to this problem. $1,1$ will not work in this context, so the next solution is $7,5$. Therefore, $n+1 = s^2 \Rightarrow n = 24$. We can check to make sure this works.

$$\sum_{0}^{24} 4i + 1 = 4\frac{(24)(25)}{2} + 24 + 1$$

$$= 1225$$

$$= 35^2$$

So 1225 is the first hexagonal number that is a perfect square (besides 1 of course).

B.) Prove there exists infinitely many numbers that are square and triangular.

Let x be square and triangular. Then there exists $k \in \mathbb{N}$ such that $x = k^2$ and there exists $n \in \mathbb{N}$ such that $x = \frac{n(n+1)}{2}$. Therefore,

$$k^2 = \frac{n(n+1)}{2}$$

$$\Rightarrow 2k^2 = n^2 + n$$

$$\Rightarrow 8k^2 = 4n^2 + 4n$$

$$\Rightarrow 8k^2 + 1 = 4n^2 + 4n + 1$$

$$\Rightarrow 8k^2 + 1 = (2n+1)^2$$

$$\Rightarrow (2n+1)^2 - 8k^2 = 1$$

This is Pell's equation for the $d = 8$ case and where our x_n values are odd, but all solutions to $x^2 - 8y^2 = 1$ must have odd values for x. Therefore, there are infinitely many numbers that are square and triangular.

Conclusion

We now know numerous ways to solve the equations $x^2 - dy^2 = \pm 1$ if solutions exist and how to determine whether or not there are solutions. We have also seen that this knowledge is useful. Cases of Pell's Equation can arise in unexpected places like the problems in the last section. We also learned that we can use our solutions to make very accurate rational approximations to irrational numbers. There are, however, many more questions left unanswered and topics left uncovered. We could continue our discussion into the more general case of $x^2 - dy^2 = N$ where $N \in \mathbb{N}$ (Theorem 3 seems like it might

be a good starting place), or we could look at why Pell's Equation is significant in the theory of quadratic fields. I hope the reader found this paper to be useful and interesting. Hopefully, this paper will spark interest in the reader to pursue the knowledge left uncovered.