Making A Brief Acquaintance with Pell's Equation

Leo Vaysman

December 9, 2024

1 Introduction

1.1 Motivation and Abstract

Pell's equation, also known as the Pell-Fermat Equation, is a Diophantine equation much like the equation in Fermat's Last Theorem. Pell's equation does not immediately yield solutions, but is simpler to work with than Fermat's Last Theorem. I chose to research Pell's equation because it has very clear and interesting ramifications – it yields rational approximations of the square root of any non-square integer. Additionally, it has interesting ramifications in the study of rings and fields, as well as in number theory beyond Diophantine equations. This paper aims to give a brief explanation of a method for determining solutions to Pell's equation, with certain algebraically lengthy proofs ommitted (such proofs can be found in *Solving the Pell Equation* by Jacobson and Williams)..

1.2 What is Pell's equation?

Pell's equation is a class of quadratic Diophantine equations of the form

$$x^2 - ny^2 = 1$$

Here, n is any positive non-square integer, and x and y are integers. Pell's equation has infinitely many solutions for x and y for every valid choice of n, these solutions being integral points on a hyperbola.

1.3 History

Pell's equation has been studied by mathematicians for several millennia. Archimedes likely knew of Pell's equation in some form, as there is a famous math puzzle known as the cattle problem attributed to him that essentially simplifies to Pell's equation (we'll see the cattle problem at the end of this paper). In the seventh century A.D., Brahmagupta found a method for determining solutions to the Pell equation (as we shall discuss, this boils down to finding a fundamental solution for every n), but unfortunately, as n grows, Brahmagupta's method is inconvenient (to use computer science language, it does not run in polynomial time with regards to the size of $\log(n)$).

2 Solving the Pell Equation: Fundamental Solutions

2.1 The modified Pell equation

Consider a more generalized form of the equation given in **1.2**. Let us look at the integer solutions of

$$x^2 + ny^2 = z^2$$

Without loss of generality, we can assume that $n < 0$ and n is not a perfect square, so clearly any solution to this is a solution to Pell's equation. In fact, this equation has at least one solution for every negative non-square n; we would like, for any n, to find a way of generating every solution.

We will now examine the solutions of the Diophantine equation

$$X^2 - nY^2 = 4\sigma, \sigma \in \{-1, 1\}$$

If we have some X, Y, σ satisfying this equation, it is clear that $X \equiv nY \pmod 2$.

When $X \equiv nY \equiv 0 \pmod 2$ and $\sigma = 1$, we have two subcases. Our first option is that $X \equiv Y \equiv 0 \pmod 2$. In this case, let $x = X/2$, $y = Y/2$. This yields a solution to Pell's equation. Alternatively, it is possible that $X \equiv 0 \pmod 2$, $Y \equiv 1 \pmod 2$. Clearly, by modular arithmetic, we have $n \equiv 0 \pmod 4$. Now let $x = X/2$, $y = Y$. This also yields a solution to Pell's equation, but now it is a solution for $n/4$ as our determiner rather than for n itself.

Given two solutions to the equation we are currently considering (call this the Modified Pell Equation), we can, given certain constraints, derive a third.

Theorem 1 *If (x_1, y_1, σ_1) and (x_2, y_2, σ_2) are solutions to the Modified Pell Equation, where $x_1 \neq \eta x_2, y_1 \neq -\eta y_2$ for $\eta \in \{-1, 1\}$, then we have a third solution to the Modified Pell Equation (x_3, y_3, σ_3) where $x_3 = \frac{x_1 x_2 + D y_1 y_2}{2}$, $y_3 = \frac{x_1 y_2 + x_2 y_1}{2}$, $\sigma_3 = \sigma_1 \sigma_2$.*

This is not difficult to prove algebraically, but for a proof for this (and for proceeding) theorems, see Jacobson and Williams, ch. 1.

For any solution to the Modified Pell Equation, the following lemmas hold:

Lemma 2 *If (x, y, σ) is a solution to the Modified Pell Equation, then $x + y\sqrt{n} > 2$ if and only if $x, y > 0$.*

Lemma 3 *If (x, y, σ) is a solution to the Modified Pell Equation, and $x, y > 0$, then $2yn \geq 8$.*

The first lemma can be shown quickly by checking parity; the second lemma can be shown simply by looking at all possible options for which $2yn$ would be less than 8.

2.2 The fundamental solution of a Pell equation

Theorem 4 *If we have two solutions (x_1, y_1, σ_1) and (x_2, y_2, σ_2) to the Modified Pell Equation, with all x and y positive, then*

$$x_2 + y_2\sqrt{n} > x_1 + y_1\sqrt{n}$$

if and only if $x_2 > x_1$ and $y_2 \geq y_1$.

This allows us to have a concept of a certain solution being "smaller" than another solution. Given that smaller values for x and y biconditionally yield smaller values for $x + y\sqrt{n}$, we can define a smallest solution such that $x + y\sqrt{n}$ is minimal (though greater than 2). We call this solution the *fundamental solution*, and we let $\epsilon = \frac{x + y\sqrt{n}}{2}$.

Critically, we can now show that every solution of Pell's equation for a given n is derived from this value of ϵ.

Theorem 5 *If (x', y', σ') is any solution of the Modified Pell Equation for a given n, then*

$$\eta = (x' + y'\sqrt{n})/2 = \pm\epsilon^n$$

This gives us a good way to attack the problem of finding every solution of the Pell Equation.

3 Solving Pell's equation: continued fractions

3.1 What are continued fractions?

Given any real number r and any given sequence of integers $\{q_n\}$ known as *partial quotients*. Define $r_0 = r$, and recursively define $r_{j+1} = \frac{1}{r_j - q_j}$ for $0 \leq j \leq i$. Now, we can express r_0 as the continued

fraction

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots q_{i-1} + \cfrac{1}{q_i + \frac{1}{r_{i+1}}}}}}$$

Now, we denote this by $r_0 = \langle q_0, q_1 ... q_i, r_{i+1} \rangle$.

3.2 The convergents of a continued fraction

We define the *convergents* of a continued fraction as follows: Let $A_{-2} = B_{-1} = 0$, $A_{-1} = B_{-2} - 1$. Recursively define $A_{j+1} = q_{j+1}A_j + A_{j-1}$, $B_{j+1} = q_{j+1}B_j + B_{j-1}$. The following facts are easily shown through arithmetic manipulation:

$$A_j B_{j-1} - B_j A_{j-1} = (-1)^j - 1$$

$$\frac{A_i}{B_i} = \langle q_0, q_1 ... q_i \rangle$$

$$\frac{A_i}{A_{i-1}} = \langle q_i, q_{i-1} ... q_0 \rangle$$

$$\frac{B_i}{B_{i-1}} = \langle q_i, q_{i-1} ... q_1 \rangle$$

Call $\frac{A_i}{B_i}$ the *ith convergent* of a continued fraction. We can now define a continued fraction as either convergent or divergent, depending on – as intuition would have it – whether the sequence of convergents converges or diverges. Note that the convergents depend on r and on the sequence of qs initially chosen.

These convergents provide a very good rational approximation for the real number r, as is obvious from the definition of the series. However, what is not as obvious is the fact that if any rational fraction provides a "very good approximation" of a real number, it must be a convergent of the continued fraction of the number. Unfortunately such a proof is outside the scope of this paper, but is given in Jacobson and Williams. This yields the following result that makes the connection to Pell's equation clear:

Theorem 6 *If $x, y, n, z \in Z$, $x, y, > 0$, $\sqrt{n} \notin Q$, $|z| \leq \sqrt{n}$, and*

$$x^2 - ny^2 = z$$

then x/y is a convergent in the continued fraction expansion of $\sqrt{n}$.

3.3 Periodic and purely periodic continued fractions

We define a continued fraction as *periodic* if, at some point, the sequence eventually repeats; that is, if the partial denominators are comprised of a preperiod q_0 through q_m, and a repeating block (the period) q_{m+1} through q_k that then repeats infinitely (e.g. $q_{k+1} = q_{m+1}$, $q_{k+2} = q_{m+2} ...$)

Euler and Lagrange proved the following –

Theorem 7 *If x is a regular continued fraction that is periodic, x is a quadratic irrational number (Euler). The converse is also true (Lagrange).*

(For an elegant proof of this theorem, see Ben Lynn's notes in the bibliography.)

We turn our attention to the quadratic irrationals with *no* preperiod; that is, the ones whose continued fraction expansion is solely comprised of a repeating block. We call these *purely periodic*. For example, consider

$$\Phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\cdots}}} = \langle 1, 1, 1, 1... \rangle$$

Of course, the period need not be of length 1 as it is with Φ.

3.4 Using periodic continued fractions to find ϵ

As discussed in **2.2**, finding every solution to a Pell equation boils down to finding ϵ. Now, putting it all together, we can use continued fractions to find ϵ as follows.

First, note (as discussed) that if the Modified Pell Equation is solvable for $X \equiv nY \equiv 1 \pmod 2$, then $n \equiv 1 \pmod 4$. For some n, define s and q as:

$$s = \begin{cases} 2 & \text{if } 4\text{—}n \text{ or } n \equiv 1 \pmod 4 \\ 1 & \text{otherwise} \end{cases}$$

$$q = \begin{cases} 0 & \text{if } n \not\equiv 1 \pmod 4 \\ 1 & \text{otherwise} \end{cases}$$

Now the following algorithm allows us to determine ϵ. Let $P_0 = q$, $Q_0 = s$, $q_0 = \lfloor (\sqrt{n} + q)/s \rfloor$, $B_{-1} = 0$, $B_0 = 1$, $G_{-1} = s$ and $G_0 = sq_0 - q$. We define the recurrences

$$P_{i+1} = q_i Q_i - P_i$$

$$Q_{i+1} = \frac{n - P_{i+1}^2}{Q_i}$$

$$q_{i+1} = \lfloor \frac{P_{i+1} + \sqrt{n}}{Q_i} \rfloor$$

$$G_{i+1} = q_{i+1} G_i + G_{i-1}$$

$$B_{i+1} = q_{i+1} B_i + B_{i-1}$$

for successive i until we find the least positive p for which $Q_p = s$. Then,

$$\epsilon = \frac{G_{p-1} + \sqrt{n} B_{p-1}}{s}$$

and $\sigma = (-1)^p$.

If one is wondering why this holds, a full algebraic proof is offered in Jacobson and Williams. However, to summarize roughly, it involves proving facts regarding the length of the period of the continued fraction expression of $\delta = \frac{q + \sqrt{n}}{s}$, and proving that various expressions including x_1, y_1, n, σ_1, and ϵ are all within small distances of ϵ (and thus belong to the continued fraction expression).

3.5 Pell's theorem for approximation of square roots

Essentially, the method described previously boils down to nothing more than finding the continued fraction representation of $\sqrt{n}$, and then using the "best" convergent (e.g. the one with yielding $s = 1$) to derive the fundamental solution (which all the following solutions can be determined by). Intuitively, this makes sense, given the fact that convergents are "good approximations" of real numbers, and the fact that we can rearrange a given Pell equation as follows:

$$x^2 - ny^2 = 1$$

$$x^2 = 1 + ny^2$$
$$x = \sqrt{1 + ny^2}$$

And now, given sufficiently large x and y:

$$x \approx \sqrt{ny^2}$$
$$x \approx y\sqrt{n}$$
$$x/y \approx \sqrt{n}$$

This is no coincidence; in fact, the Pell equation is as famous as it is in part because it allows for these approximations. The Pythagoreans, the Baudhayana Śulbasûtra (an ancient Sanskrit mathematical text containing a separately-derived assertion of the Pythagorean theorem), and writings of Archimedes all draw a connection between solutions to the Pell equation and approximations of the square roots of small numbers (generally 2 and 3).

4 Further results and curiosities involving the Pell equation

The Pell equation is closely related to modern algebra, specifically to the study of quadratic number fields. Quadratic number fields bear the concept of a norm, defined as $N(x + y\sqrt{k}) = x^2 - ky^2$. Thus, we can see that finding the solutions to the Pell equation $x^2 - ny^2 = 1$ is equivalent to finding all the elements of the ring $Q[\sqrt{n}]$ with norm of 1.

The Pell equation also appears in an important number-theoretic result, Størmer's Theorem, which says that, for any finite set P of prime numbers, there are only a finite number of pairs of consecutive integers whose prime factors are all in P. Though this result might not appear at first to have any relation to the Pell equation, not only is the Pell equation instrumental in its proof, but considering simultaneous solutions to Pell equations yields a method for which one can determine all these consecutive pairs for any set P.

These are just two among many of the curious applications of this relatively simple Diophantine equation.

References

B. Lynn. "Periodic Continued Fractions". Stanford University Cryptography.
Accessed from https://crypto.stanford.edu/pbc/notes/contfrac/periodic.html.
D. H. Lehmer. "On a problem of Størmer," Illinois Journal of Mathematics, Illinois J. Math. 8(1), 57-79, (March 1964).
H. W. Lenstra. "Solving the Pell Equation". Algorithmic Number Theory, Cambridge University Press (2008).
M. J. Jacobson and H. C. Williams. "Solving the Pell Equation". CMS Books in Mathematics (2009).