The Search For An Exhaustive Solution to Pell's Equation

Derek Smith

Abstract

A Pell Equation with variables x and y in parameter D is given by the following expression:

$$x^2 = 1 + Dy^2$$

The search for a solution to such an equation involves finding a method that generates all non-trivial solutions $(x, y) \neq (1, 0)$. Through the years it has been noted that such an equation has no solutions if D is a perfect square and an infinitude of solutions otherwise. The first known mention of these equations appears in Ancient Greece. This paper traces the search for an exhaustive solution beginning with Brahmagupta's method of composition, picking up again with Fermat's independent work. Fermat's subsequent challenge was finally put to rest by Lagrange in his *Additions* to Euler's *Elements of Algebra*. Lagrange proved that Euler's method produces all solutions given a least positive initial solution. This paper aims to provide an introduction to Pell equations by placing selected results in a historical context.

1. Introduction

A Pell Equation with variables x and y in parameter D is given by the following expression:

$$x^2 = 1 + Dy^2 \tag{1}$$

By fundamental solution to this equation, we mean the two smallest, positive integers $(x, y) \neq (1, 0)$ satisfying the Pell equation. We talk only of positive solutions to dispose of the trivial solutions $(-x, -y), (-x, y), (x, -y)$ given positive solution (x, y). There are also restrictions on the parameter D. It must be a positive integer but not a perfect square. Without these restrictions this equation has only the trivial solution $(1, 0)$. We begin with a quick proof of the second of these facts.

Proposition 1. If $D = k^2$ for some integer k, then the equation $x^2 - Dy^2 = 1$ has only the trivial solution $(1, 0)$.

Proof. Assume, to the contrary, that $x^2 - Dy^2 = 1$ has some other least positive solution, $(a, b) \neq (1, 0)$. Upon substitution, $x^2 - Dy^2 = (a + kb)(a - kb) = 1$, where a, b, k are all integers. But the product of two integers cannot equal one unless they both are one. We already know we are not in the case $(a, b) = (1, 0)$, so the assumption that the equation has some other solution must be incorrect. $\diamondsuit$

A similar proof could be constructed for the case $D = -k$. It is helpful to see that a solution (a, b) to (1) provides a good rational approximation to $\sqrt{D}$. Dividing both sides of (1) by y^2 and taking the square root gives $a/b = \sqrt{1/y^2 + D}$. For large values of y, the approximation becomes better. This concept leads to a technique for producing a the fundamental solution. For example

$$x^2 = 1 + 2y^2$$

has solutions:

$$(a_1, b_1) = (17, 12) \Rightarrow a/b = 1.416667$$
$$(a_2, b_2) = (577, 408) \Rightarrow a/b = 1.414216$$

The Pell Equation has a long and broken history. It was first inadvertantly studied by Diophantus and Archimedes. Diophantus solved equations of this form in specific cases. The solution to Archimedes' Cattle Problem hinges on a Pell equation, although it is not known whether he intended for this. In addition to his many other significant contributions, the Indian mathematician Brahmagutpa provided the first known general solution method, though not exhaustive. Fermat brought these equations into modern Western mathematics with a single example as one of his famous challenge problems. Euler mistakenly named the equations after Pell. Lagrange set the solution to this class of equations in stone by further proving the continued fraction method given by Euler.

2. Early Contributions

As stated earlier, part of Archimedes Cattle Problem can be formulated as a Pell equation. This problem is purported "to be one proposed by Archimedes, in a letter to Eratosthenes, to the mathematicians of Alexandria." [Dickson 342] This information was put forth in a manuscript published in 1773 by Gotthold Lessing. The first part of the problem consists of a large system of linear equations resulting from conditions on the relative number of cattle of various colors. "G. H. F. Nesselmann argued that the final part of the epigram leading to conditions [of square and triangular numbers] was a later addition." [Dickson 344] Of course this is the part of the problem that is most troublesome. It leads to the following Pell equation:

$$x^2 = 1 + 4,729,494y^2$$

It is not known whether Archimedes possessed the capabilities to solve such an equation. Brahmagutpa was the first mathematician to put forth a method for solving exactly these types of equations. Not only did he develop methods for generating a single solution, but he realized that a single solution could be modified to produce a large number of solutions.

This method is called composition. It produces a new solution from two known solutions. He obviously realized that this leads to an infinite number of solutions. His method is a generalization of the following idea.

Proposition 2. If (a, b) and (c, d) are solutions to the Pell equation $x^2 - Dy^2 = 1$, then $(ac + Dbd, ad + bc)$ and $(ac - Dbd, ad - bc)$ are also solutions to the same equation.

Proof. Brahmagupta's method of composition relies on the following identities:

$$\begin{aligned}(a^2 - Db^2)(c^2 - Dd^2) &= a^2c^2 - Da^2d^2 - Db^2c^2 + D^2b^2d^2 \\ &= a^2c^2 + D^2b^2d^2 - D(a^2d^2 + b^2c^2) \\ &= a^2c^2 + 2Dabcd + D^2b^2d^2 - D(a^2d^2 + 2abcd + b^2c^2) \\ &= (ac + Dbd)^2 - D(ad + bc)^2\end{aligned}$$

$$\begin{aligned}(a^2 - Db^2)(c^2 - Dd^2) &= a^2c^2 - Da^2d^2 - Db^2c^2 + D^2b^2d^2 \\ &= a^2c^2 + D^2b^2d^2 - D(a^2d^2 + b^2c^2) \\ &= a^2c^2 - 2Dabcd + D^2b^2d^2 - D(a^2d^2 - 2abcd + b^2c^2) \\ &= (ac - Dbd)^2 - D(ad - bc)^2\end{aligned}$$

If (a,b) and (c,d) are solutions to $x^2 - Dy^2 = 1$, then

$$\begin{aligned}a^2 - Db^2 &= 1 \\ c^2 - Dd^2 &= 1\end{aligned}$$

By substitution into the left hand sides of the previous two identities, we arrive at

$$\begin{aligned}(ac + Dbd)^2 - D(ad + bc)^2 &= 1 \\ (ac - Dbd)^2 - D(ad - bc)^2 &= 1\end{aligned}$$

Therefore $(ac + Dbd, ad + bc)$ and $(ac - Dbd, ad - bc)$ both sovle the Pell equation in parameter D. ♢

This proof was paraphrased from Robertson and Edmund.

3. European Contributions

"The European interest began in 1657 when Fermat issued a challenge to the mathematicians of Europe and England." [Edmund & Robertson] This challenge included a Pell equation. After spending many years reproducing the years of earlier Indian mathematicians, Euler finally developed a method to generate a single solution to the Pell Equation. This method is described in Section 6 without proof.

A method was also given in Eulers *Elements of Algebra* to generate what was believed to be all solutions to the Pell equation, given the fundamental solution.

Proposition 3. P(n): The n-th solution (x_n, y_n) can be expressed in terms of the first one, (x_1, y_1), by $x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$. [Lenstra 1]

Proof. This proof will be by induction. Show that P(n) holds for all $n \in N$.

P(1) is trivial. Show that P(k) $\Rightarrow$ P($k+1$).

It is known that (x_1, y_1) is a solution by assumption and (x_k, y_k) is a solution by the induction hypothesis. It remains to show that (x_{k+1}, y_{k+1}) is a solution, where

$$\begin{aligned}
x_{k+1} + y_{k+1}\sqrt{D} &= (x_1 + y_1\sqrt{D})^{k+1} \\
&= (x_1 + y_1\sqrt{D})^k(x_1 + y_1\sqrt{D}) \\
&= (x_k + y_k\sqrt{D})(x_1 + y_1\sqrt{D}) \\
&= x_1x_k + y_1y_kD + (y_1x_k + x_1y_k)\sqrt{D}
\end{aligned}$$

Which leads to

$$\begin{aligned}
x_{k+1} &= x_1x_k + y_1y_kD \\
y_{k+1} &= y_1x_k + x_1y_k
\end{aligned}$$

Show $x_{k+1}^2 - y_{k+1}^2D = 1$.

$$\begin{aligned}
x_{k+1}^2 - y_{k+1}^2D &= (x_1x_k + y_1y_kD)^2 - (y_1x_k + x_1y_k)^2D \\
&= x_1^2x_k^2 + 2y_1y_kx_1x_kD + y_1^2y_k^2D^2 - (y_1^2x_k^2 + 2y_1y_kx_1x_k + y_k^2x_1^2)D \\
&= x_1^2x_k^2 + y_1^2y_k^2D^2 - (y_1^2x_k^2 + y_k^2x_1^2)D \\
&= x_1^2x_k^2 - (y_1^2x_k^2)D - (y_k^2x_1^2)D + y_1^2y_k^2D^2 \\
&= (x_1^1 - y_1^2D)(x_k^2 - y_k^2D) \\
&= (1)(1) = 1
\end{aligned}$$

Py the principle of mathematical induction, P(n) holds for all $n \in N$. $\diamondsuit$

4. Lagrange's Complete Solution

It now seemed that this very long search was over, and in fact it was. With the method of Proposition 3, it seemed that every solution to the Pell Equation could be produced given some non-trivial initial solution. This fact was not verified until Lagrange in his *Additions* to Euler's *Elements of Algebra*.

Next we present Lagrange's work as taken from *Elementary Number Theory*. The following Lemma is from page 347. The proof has been edited for flow.

Lemma. If ξ is a real number, and $\tau > 1$ is any integer, then two integers r and s can be found satisiying the inequalities

$$|s\xi - r| < \frac{1}{\tau}, \qquad 0 < s \leq r$$

Proof. For each $x = 0, 1, 2, \ldots, \tau$ determine y so that $0 \leq x_i\xi - y_i < 1$. We therefore get $\tau + 1$ distinct differences contained in the interval between 0 and 1. If we divide this interval into τ intervals of size $\frac{1}{\tau}$ we can find two differences $x_i\xi - y_i$ and $x_j\xi - y_j$ contained in the same interval by the pigeonhole principle. Choosing $x_j > x_i$ and letting

$$x_j - x_i = s, \qquad y_j - y_i = r,$$

Since both x_j and x_i are two distinct integers less than τ, their difference satisfies $0 < s \leq \tau$. Upon further computation, we find

$$\begin{aligned}(x_j\xi - y_j) - (x_i\xi - y_i) &= (x_j - x_i)\xi - (y_j - y_i) \\ &= s\xi - r\end{aligned}$$

Therefore $|s\xi - r| < \frac{1}{\tau}$, since it is the difference of two numbers contained in an interval of length $\frac{1}{\tau}$. ♢

We will now use this lemma to show that the Pell equation has a non-trivial solution for non-square parameter, D. The following proof is fairly technical and as such remains only edited slightly.

Proposition 4. The Pell Equation has a non-trivial integer solution for every $D > 1$ not a square. [Upensky & Heaslet 348].

Proof. Let D be a positive integer that is not a square. By the previous lemma, two integers r, s can be found for some integer $\tau > 1$ so that

$$|r - s\sqrt{D}| < \frac{1}{\tau}, \qquad 0 < s \leq \tau$$

Then also

$$|r + s\sqrt{D}| < \frac{1}{\tau} + 2\tau\sqrt{D}$$

and by the product of the previous two

$$|r^2 - Ds^2| < \frac{1}{\tau^2} + 2\sqrt{D} < 1 + 2\sqrt{D}$$

That is, integers r, s making $r - s\sqrt{D}$ as small as we please numerically, can be found in an infinite number satisfying this inequality string. Let $[1 + 2\sqrt{D}] = g$, the greatest integer; then the number of integers excluding zero in the interval between $-1 - 2\sqrt{D}$ and $1 + 2\sqrt{D}$ will be $2g$. Take $n = 2g^4$ pairs (r_i, s_i) satisfying the given inequalities so that

$$|r_1^2 - Ds_1^2| > |r_2^2 - Ds_2^2| > \ldots > |r_n^2 - Ds_n^2|$$

The n differences

$$r_1^2 - Ds_1^2, r_2^2 - Ds_2^2, \ldots, r_n^2 - Ds_n^2$$

are all integers contained between $-1-2\sqrt{D}$ and $1+2\sqrt{D}$ since r, s, D are all integers. Let these integers be $L_1, L_2, \ldots, L_n$ and let M_i denote the number of times L_i occurs in this series of integers. Then

$$M_1 + M_2 + \cdots + M_n = n = 2g^4$$

and the greatest of the numbers $M_1, M_2, \cdots, M_n$ is necessarily $\geq g^3$ by the pigeonhole principle with $2g^4$ pigeons and $2g$ holes. That is to say, for some $k = \pm 1, \pm 2, \ldots, \pm g$, the equation

$$r^2 - Ds^2 = k$$

is satisfied by at least $g^3 > g^2 \geq k^2$ pairs of integers r, s.

We call two pairs x_1, y_1 and x_2, y_2 congruent mod k if and only if

$$x_2 \equiv x_1, \qquad y_2 \equiv y_1 \ (mod\ k)$$

Then the number of incongruent pairs is k^2, and among the k^2+1 pairs at least two pairs are congruent. Since the equation $r^2 - Ds^2 = k$ is satisfied by more than k^2 pairs, at least two of them, r_1, s_1 and r_2, s_2 will be congruent mod k, so that

$$r_2^2 - Ds_2^2 = r_1^2 - Ds_1^2 = k$$

$$r_2 \equiv r_1,\ \ s_2 \equiv s_1 \ (mod\ k)$$

Moreover, we can suppose that

$$|r - 2 - s_2\sqrt{D}| < |r_1 - s_1\sqrt{D}|$$

Consider the quotient

$$\frac{r_1 - s_1\sqrt{D}}{r - 2 - s_2\sqrt{D}} = \frac{r_1r_2 - Ds_1s_2 + (r_1s_2 - r_2s_1)\sqrt{D}}{k}$$

By virtue of the congruences $r_2 \equiv r_1$, $s_2 \equiv s_1$ (mod k),

$$r_1r_2 - Ds_1s_2 \equiv r_1^2 - Ds_1^2 \equiv 0, \qquad r_1s_2 - r_2s_1 \equiv 0 \ (mod\ k)$$

so that

$$\frac{r_1r_2 - Ds_1s_2}{k} = a, \qquad \frac{r_1s_2 - r_2s_1}{k} = b$$

are integers and

$$(r_1 - s_1\sqrt{D}) = (r_2 - s_2\sqrt{D})(a + b\sqrt{D})$$

implying

$$(r_1^2 - s_1^2D) = (r_2^2 - Ds_2^2)(a^2 + Db^2)$$

and cancelling $r_2^2 - Ds_2^2 = r_1^2 - s_1^2 D = k \neq 0$, gives

$$a^2 - Db^2 = 1$$

We also have $|a + b\sqrt{D}| > 1$ implying that $b \neq 0$. $\diamondsuit$

The existence of a non-trivial solution, and consequently an infinite number of solutions, to the Pell equation has now been proved. It is only left to confirm that the method of Proposition 3 generates all positive solutions to a given equation. The proof of this fact follows closely to Lagrange's work but has been modified to aid in flow.

Proposition 5. The method given in Proposition 3 provides a complete solution to the Pell Equation in the non-square parameter D.

Proof. Let (A, B) be the fundamental, non-trivial solution to the Pell Equation in paramter D. If (a, b) is any other solution distinct from the fundamental, then $b > B$ and

$$a + b\sqrt{D} > A + B\sqrt{D}$$

Assume, to the contrary, that (a, b) is a non-trivial solution not given by the method outlined in Proposition 3. In the series of powers

$$A + B\sqrt{D}, (A + B\sqrt{D})^2, (A + B\sqrt{D})^3, \ \ldots$$

there are two consecutive terms such that

$$(A + B\sqrt{D})^n \leq a + b\sqrt{D} < (A + B\sqrt{D})^{n+1}$$

Multiplying through by $(A - B\sqrt{D})^n$ gives

$$1 \leq (a + b\sqrt{D})(A - B\sqrt{D})^n < A + B\sqrt{D}$$

Note that $(a + b\sqrt{D})(A - B\sqrt{D})^n$ can be reduced to the form $p + q\sqrt{D}$ where p, q are integers and $p^2 - Dq^2 = 1$, giving

$$1 \leq p + q\sqrt{D} < A + B\sqrt{D}$$

and

$$0 < A - B\sqrt{D} < p - q\sqrt{D} \leq 1$$

These inequalities imply that $p > 0$ and $0 \leq q < B$. But q can not be positive, otherwise there would be a solution in positive integers smallter and (A, B). Therefore $p = 1$, $q = 0$ and

$$a + b\sqrt{D} = (A + B\sqrt{D})^n$$

for some positive n. $\diamondsuit$

In addition to proving this Theorem, Lagrange also succeeded in proving the correctness of Eulers method of generating the fundamental solution, described next.

6. Final Comments

No mention of how to find initial solution. Here we outline the general solution method with an example.

Solve $x^2 = 1 + 17y^2$.

We begin with the continued fraction expansion of $\sqrt{17}$.

$$\begin{aligned}
x = \sqrt{17} &= 4 + \frac{1}{x} \\
\Rightarrow \frac{1}{x} &= \sqrt{17} - 4 \\
\Rightarrow x &= \frac{1}{\sqrt{17}-4} \cdot \frac{\sqrt{17}+4}{\sqrt{17}+4} = \sqrt{17} + 4
\end{aligned}$$

Substituting yields:

$$\begin{aligned}
\sqrt{17} = 4 + \frac{1}{x} &= 4 + \frac{1}{4+\sqrt{17}} \\
&= 4 + \frac{1}{8 + \frac{1}{8+\ldots}} \\
&= [4; 8, 8, 8, \ldots]
\end{aligned}$$

Truncating this pattern at the first repetition, we get

$$\frac{x}{y} = 4 + \frac{1}{8} = \frac{33}{8}$$

Therefore, the pair $(33, 8)$ is the least positive solution to the given Pell equation.

$$33^2 - 17 \cdot 8^2 = 1$$

To create another solution, solve the following:

$$\begin{aligned}
(x + y\sqrt{17}) &= (33 + 8\sqrt{17})^2 \\
&= 33^2 + 33 \cdot 16\sqrt{17} + 8^2 \cdot 17 \\
&= 2177 + 528\sqrt{17}
\end{aligned}$$

Therefore, the pair $(2177, 528)$ is the next solution to the given Pell equation.

This is an interesting topic that leads into a number of algebraic and number theoretical topics. Though we can no longer continue this 2000 year search there are still some paths that we have not been down yet, such as the one that leads to a polynomial time algorithm to generate these solutions or a proof that none exists.

References

[1] Uspensky, J. V. and Heaslet, M. A. (1939) *Elementary Number Theory.* Chelsea Publishing Company, New York.

[2] Dickson, Leonard (1971) *History of the Theory of Numbers, Volume II.* McGraw-Hill Book Company, New York and London.

[3] O'Connor, John and Robertson, Edmund (2002) *The MacTutor History of Mathematics archive.* http://www-gap.dcs.st-and.ac.uk/ history/HistTopics/Pell.html, Accessed Decemeber 1, 2002.

[4] Lenstra, H. W. (February 2002) *Solving the Pell Equation.* notices of the AMS, Volume 49, Number 2.