Algorithmic Number Theory
MSRI Publications
Volume **44**, 2008

Solving the Pell equation

HENDRIK W. LENSTRA, JR.

ABSTRACT. We illustrate recent developments in computational number theory by studying their implications for solving the Pell equation. We shall see that, if the solutions to the Pell equation are properly represented, the traditional continued fraction method for solving the equation can be significantly accelerated. The most promising method depends on the use of smooth numbers. As with many algorithms depending on smooth numbers, its run time can presently only conjecturally be established; giving a rigorous analysis is one of the many open problems surrounding the Pell equation.

1. Pell's equation

The *Pell equation* is the equation

$$x^2 = dy^2 + 1,$$

to be solved in positive integers x, y for a given nonzero integer d. For example, for $d = 5$ one can take $x = 9$, $y = 4$. We shall always assume that d is positive but not a square, since otherwise there are clearly no solutions.

The English mathematician John Pell (1611–1685) has nothing to do with the equation. Euler (1707–1783) mistakenly attributed to Pell a solution method that had in fact been found by another English mathematician, William Brouncker (1620–1684), in response to a challenge by Fermat (1601–1665); but attempts to change the terminology introduced by Euler have always proved futile.

Pell's equation has an extraordinarily rich history, to which Weil [1984] is the best guide; see also [Dickson 1920, Chapter XII; Konen 1901; Whitford 1912]. Brouncker's method is in substance identical to a method that was known to Indian mathematicians at least six centuries earlier. As we shall see, the equation

This paper appeared in slightly different form in *Notices Amer. Math. Soc.* **49** (2002), 182–192, with the permission of MSRI and the editors of the present volume.

also occurred in Greek mathematics, but no convincing evidence that the Greeks could solve the equation has ever emerged.

A particularly lucid exposition of the "Indian" or "English" method of solving the Pell equation is found in Euler's *Algebra* [Euler 1770, Abschnitt 2, Capitel 7]. Modern textbooks usually give a formulation in terms of continued fractions, which is also due to Euler (see for example [Niven et al. 1991, Chapter 7]). Euler, as well as his Indian and English predecessors, appears to take it for granted that the method always produces a solution. That is true, but it is not obvious — all that is obvious is that *if* there is a solution, the method will find one. Fermat was probably in possession of a proof that there is a solution for every d (see [Weil 1984, Chapter II, § XIII]), and the first to publish such a proof was Lagrange (1736–1813) [1773].

One may rewrite Pell's equation as

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

so that finding a solution comes down to finding a nontrivial unit of the ring $\mathbb{Z}[\sqrt{d}]$ of norm 1; here the norm $\mathbb{Z}[\sqrt{d}]^* \to \mathbb{Z}^* = \{\pm 1\}$ between unit groups multiplies each unit by its conjugate, and the units ± 1 of $\mathbb{Z}[\sqrt{d}]$ are considered trivial. This reformulation implies that once one knows a solution to Pell's equation, one can find infinitely many. More precisely, if the solutions are ordered by magnitude, then the n-th solution x_n, y_n can be expressed in terms of the first one, x_1, y_1, by

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n.$$

Accordingly, the first solution x_1, y_1 is called the *fundamental solution* to the Pell equation, and *solving* the Pell equation means finding x_1, y_1 for given d. By abuse of language, we shall also refer to $x + y\sqrt{d}$ instead of the pair x, y as a solution to Pell's equation and call $x_1 + y_1\sqrt{d}$ the fundamental solution.

One may view the solvability of Pell's equation as a special case of *Dirichlet's unit theorem* from algebraic number theory, which describes the structure of the group of units of a general ring of algebraic integers [Stevenhagen 2008a]; for the ring $\mathbb{Z}[\sqrt{d}]$, it is the product of $\{\pm 1\}$ and an infinite cyclic group.

As an example, consider $d = 14$. One has

$$\sqrt{14} = 3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3 + \sqrt{14}}}}},$$

so the continued fraction expansion of $3+\sqrt{14}$ is purely periodic with period length 4. Truncating the expansion at the end of the first period, one finds that the fraction

$$3+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{\cfrac{1}{1}}}}=\frac{15}{4}$$

is a fair approximation to $\sqrt{14}$. The numerator and denominator of this fraction yield the fundamental solution $x_1 = 15$, $y_1 = 4$; indeed one has $15^2 = 14 \cdot 4^2 + 1$. Furthermore, one computes $(15+4\sqrt{14})^2 = 449 + 120\sqrt{14}$, so $x_2 = 449$, $y_2 = 120$; and so on. One finds:

n	x_n	y_n
1	15	4
2	449	120
3	13455	3596
4	403201	107760
5	12082575	3229204
6	362074049	96768360

The shape of the table reflects the exponential growth of x_n and y_n with n.

For general d, the continued fraction expansion of $[\sqrt{d}]+\sqrt{d}$ is again purely periodic, and the period displays a symmetry similar to the one visible for $d = 14$. If the period length is even, one proceeds as above; if the period length is odd, one truncates at the end of the *second* period [Buhler and Wagon 2008].

2. The cattle problem

An interesting example of the Pell equation, both from a computational and from a historical perspective, is furnished by the *cattle problem* of Archimedes (287–212 B.C.). A manuscript containing this problem was discovered by Lessing (1729–1781) in the Wolffenbüttel library, and published by him in 1773 (see [Lessing 1773; Heiberg 1913, pp. 528–534]). It is now generally credited to Archimedes [Fraser 1972; Weil 1984]. In twenty-two Greek elegiac distichs, the problem asks for the number of white, black, dappled, and brown bulls and cows belonging to the Sun god, subject to several arithmetical restrictions. A version in English heroic couplets, published in [Archimedes 1999], is shown on page 4. In modern mathematical notation the problem is no less elegant. Writing x, y, z, t for the numbers of white, black, dappled, and brown bulls,

PROBLEM

that Archimedes conceived in verse
and posed to the specialists at Alexandria
in a letter to Eratosthenes of Cyrene.

The Sun god's cattle, friend, apply thy care
to count their number, hast thou wisdom's share.
They grazed of old on the Thrinacian floor
of Sic'ly's island, herded into four,
colour by colour: one herd white as cream,
the next in coats glowing with ebon gleam,
brown-skinned the third, and stained with spots the last.
Each herd saw bulls in power unsurpassed,
in ratios these: count half the ebon-hued,
add one third more, then all the brown include;
thus, friend, canst thou the white bulls' number tell.
The ebon did the brown exceed as well,
now by a fourth and fifth part of the stained.
To know the spotted — all bulls that remained —
reckon again the brown bulls, and unite
these with a sixth and seventh of the white.
Among the cows, the tale of silver-haired
was, when with bulls and cows of black compared,
exactly one in three plus one in four.
The black cows counted one in four once more,
plus now a fifth, of the bespeckled breed
when, bulls withal, they wandered out to feed.
The speckled cows tallied a fifth and sixth
of all the brown-haired, males and females mixed.
Lastly, the brown cows numbered half a third
and one in seven of the silver herd.
Tell'st thou unfailingly how many head
the Sun possessed, o friend, both bulls well-fed
and cows of ev'ry colour — no-one will
deny that thou hast numbers' art and skill,
though not yet dost thou rank among the wise.
But come! also the foll'wing recognise.
Whene'er the Sun god's white bulls joined the black,
their multitude would gather in a pack
of equal length and breadth, and squarely throng
Thrinacia's territory broad and long.
But when the brown bulls mingled with the flecked,
in rows growing from one would they collect,
forming a perfect triangle, with ne'er
a diff'rent-coloured bull, and none to spare.
Friend, canst thou analyse this in thy mind,
and of these masses all the measures find,
go forth in glory! be assured all deem
thy wisdom in this discipline supreme!

respectively, one reads in lines 8–16 the restrictions

$$\begin{aligned} x &= (\tfrac{1}{2}+\tfrac{1}{3})y+t, \\ y &= (\tfrac{1}{4}+\tfrac{1}{5})z+t, \\ z &= (\tfrac{1}{6}+\tfrac{1}{7})x+t. \end{aligned}$$

Next, for the numbers x', y', z', t' of cows of the same respective colors, the poet requires in lines 17–26

$$\begin{aligned} x' &= (\tfrac{1}{3}+\tfrac{1}{4})(y+y'), & z' &= (\tfrac{1}{5}+\tfrac{1}{6})(t+t'), \\ y' &= (\tfrac{1}{4}+\tfrac{1}{5})(z+z'), & t' &= (\tfrac{1}{6}+\tfrac{1}{7})(x+x'). \end{aligned}$$

Whoever can solve the problem thus far is called merely competent by Archimedes; to win the prize for supreme wisdom, one should also meet the conditions formulated in lines 33–40 that $x+y$ be a *square* and that $z+t$ be a *triangular number*.

The first part of the problem is just linear algebra, and there is indeed a solution in *positive* integers. The general solution to the first three equations is given by $(x, y, z, t) = m\cdot(2226,\ 1602,\ 1580,\ 891)$, where m is a positive integer. The next four equations turn out to be solvable if and only if m is divisible by 4657; with $m = 4657\cdot k$ one has

$$(x', y', z', t') = k\cdot(7206360,\ 4893246,\ 3515820,\ 5439213).$$

The true challenge is now to choose k such that $x+y = 4657\cdot 3828\cdot k$ is a square and $z+t = 4657\cdot 2471\cdot k$ is a triangular number. From the prime factorization $4657\cdot 3828 = 2^2\cdot 3\cdot 11\cdot 29\cdot 4657$ one sees that the first condition is equivalent to $k = al^2$, where $a = 3\cdot 11\cdot 29\cdot 4657$ and l is an integer. Since $z+t$ is a triangular number if and only if $8(z+t)+1$ is a square, we are led to the equation $h^2 = 8(z+t)+1 = 8\cdot 4657\cdot 2471\cdot al^2+1$, which is the Pell equation $h^2 = dl^2+1$ for

$$d = 2\cdot 3\cdot 7\cdot 11\cdot 29\cdot 353\cdot(2\cdot 4657)^2 = 410\,286423\,278424.$$

Thus, by Lagrange's theorem, the cattle problem admits infinitely many solutions.

In 1867 the otherwise unknown German mathematician C. F. Meyer set out to solve the equation by the continued fraction method [Dickson 1920, p. 344]. After 240 steps in the continued fraction expansion for $\sqrt{d}$ he had still not detected the period, and he gave up. He may have been a little impatient; it was later discovered that the period length equals 203254; see [Grosjean and De Meyer 1991]. The first to solve the cattle problem in a satisfactory way was A. Amthor in 1880 (see [Krumbiegel and Amthor 1880]). Amthor did *not* directly apply the continued fraction method; what he did do we shall discuss below. Nor did he spell out the decimal digits of the fundamental solution to the Pell equation or the corresponding solution of the cattle problem. He did show that, in the smallest solution to the cattle problem, the total number of cattle is given by a number of 206545 digits; of the four leading digits 7766 that he gave, the fourth was wrong, due to the use of insufficiently precise logarithms. The full number occupies forty-seven pages of computer printout, reproduced in reduced size on twelve pages of the *Journal of Recreational Mathematics* [Nelson 1980/81]. In abbreviated form, it reads

$$77602714\ldots237983357\ldots55081800,$$

each of the six dots representing 34420 omitted digits.

Several nineteenth century German scholars were worried that so many bulls and cows might not fit on the island of Sicily, contradicting lines 3 and 4 of the poem; but, as Lessing remarked, the Sun god, to whom the cattle belonged, will have coped with it.

The story of the cattle problem shows that the continued fraction method is not the last word on the Pell equation.

3. Efficiency

We are interested in the *efficiency* of solution methods for the Pell equation. Thus, how much time does a given algorithm for solving the Pell equation take? Here *time* is to be measured in a realistic way, which reflects, for example, that large positive integers are more time-consuming to operate with than small ones; technically, one counts *bit operations*. The input to the algorithm is d, and the running time estimates are accordingly expressed as functions of d. If one supposes that d is specified in binary or in decimal, then the *length of the input* is approximately proportional to $\log d$. An algorithm is said to run in *polynomial time* if there is a positive real number c_0 such that for all d the running time is at most $(1+\log d)^{c_0}$, in other words, if the time that it takes the algorithm to *solve* the Pell equation is not much greater than the time required to *write down* the equation.

How fast is the continued fraction method? Can the Pell equation be solved in polynomial time? The central quantity that one needs to consider in order to answer such questions is the *regulator* R_d, which is defined by

$$R_d = \log(x_1 + y_1\sqrt{d}),$$

where $x_1 + y_1\sqrt{d}$ denotes, as before, the fundamental solution to Pell's equation. The regulator coincides with what in algebraic number theory would be called the regulator of the kernel of the norm map $\mathbb{Z}[\sqrt{d}]^* \to \mathbb{Z}^*$. From

$x_1 - y_1\sqrt{d} = 1/(x_1 + y_1\sqrt{d})$ one deduces that $0 < x_1 - y_1\sqrt{d} < 1/(2\sqrt{d})$, and combining this with $x_1 + y_1\sqrt{d} = \mathrm{e}^{R_d}$, one finds that

$$\frac{\mathrm{e}^{R_d}}{2} < x_1 < \frac{\mathrm{e}^{R_d}}{2} + \frac{1}{4\sqrt{d}}, \qquad \frac{\mathrm{e}^{R_d}}{2\sqrt{d}} - \frac{1}{4d} < y_1 < \frac{\mathrm{e}^{R_d}}{2\sqrt{d}}.$$

This shows that R_d is very close to $\log(2x_1)$ and to $\log(2y_1\sqrt{d})$. That is, if x_1 and y_1 are to be represented in binary or in decimal, then R_d is approximately proportional to the *length of the output* of any algorithm solving the Pell equation. Since the time required for spelling out the output is a lower bound for the total running time, we may conclude: *there exists c_1 such that any algorithm for solving the Pell equation takes time at least $c_1 R_d$*. Here c_1 denotes, just as do $c_2, c_3, \ldots$ below, a positive real number that does not depend on d.

The continued fraction method almost meets this lower bound. Let l be the period length of the continued fraction expansion of $[\sqrt{d}] + \sqrt{d}$ if that length is even and twice that length if it is odd. Then one has

$$\frac{\log 2}{2} \cdot l < R_d < \frac{\log(4d)}{2} \cdot l;$$

see [Lenstra 1982, (11.4)]. Thus R_d and l are approximately proportional. Using this, one estimates easily that the time taken by a straightforward implementation of the continued fraction method is at most $R_d^2 \cdot (1 + \log d)^{c_2}$ for suitable c_2; and a more refined implementation, which depends on the fast Fourier transform, reduces this to $R_d \cdot (1 + \log d)^{c_3}$ for suitable c_3; see [Schönhage 1971]. We conclude that the latter version of the continued fraction method is optimal, apart from a logarithmic factor.

In view of these results it is natural to ask how the regulator grows as a function of d. It turns out that it fluctuates wildly. One has

$$\log(2\sqrt{d}) < R_d < \sqrt{d} \cdot (\log(4d) + 2),$$

the lower bound because of the inequality $y_1 < \mathrm{e}^{R_d}/(2\sqrt{d})$ above and the upper bound by [Hua 1942]. The gap between the two bounds is very large, but it cannot be helped: if d ranges over numbers of the form $k^2 - 1$, for which one has $x_1 = k$ and $y_1 = 1$, then $R_d - \log(2\sqrt{d})$ tends to 0; and one can show that there exist an infinite set D of d's and a constant c_4 such that all $d \in D$ have $R_d = c_4\sqrt{d}$. In fact, if d_0, d_1 are integers greater than 1 and d_0 is not a square, then there exists a positive integer $m = m(d_0, d_1)$ such that $D = \{d_0 d_1^{2n} : n \in \mathbb{Z}, n \geq m\}$ has this property for some $c_4 = c_4(d_0, d_1)$.

It is believed that for most d the upper bound is closer to the truth. More precisely, a folklore conjecture asserts that there is a set D of nonsquare positive integers that has density 1 in the sense that $\lim_{x\to\infty} \#\{d \in D : d \le x\}/x = 1$, and that satisfies

$$\lim_{d\in D} \frac{\log R_d}{\log \sqrt{d}} = 1.$$

This conjecture, however, is wide open. The same is true for the much weaker conjecture that $\limsup_d (\log R_d)/\log\sqrt{d}$, with d ranging over the *squarefree* integers > 1, is *positive*.

If the folklore conjecture is true, then for most d the factor R_d entering the running time is about $\sqrt{d}$, which is an exponential function of the length $\log d$ of the input.

Combining the preceding results, one concludes that the continued fraction method takes time at most $\sqrt{d}\cdot(1+\log d)^{c_5}$; that conjecturally it is exponentially slow for *most* values of d; and that *any* method for solving the Pell equation that spells out x_1 and y_1 in full is exponentially slow for *infinitely many* d and will therefore fail to run in polynomial time.

If we want to improve upon the continued fraction method, then we need a way of representing x_1 and y_1 that is more compact than the decimal or binary notation.

4. Amthor's solution

Amthor's solution to the cattle problem depended on the observation that the number $d = 410\,286423\,278424$ can be written as $(2\cdot 4657)^2 \cdot d'$, where $d' = 4\,729494$ is squarefree. Hence, if x, y solves the Pell equation for d, then x, $2\cdot 4657\cdot y$ solves the Pell equation for d' and will therefore for some n be equal to the n-th solution x'_n, y'_n (say) of that equation:

$$x + 2\cdot 4657\cdot y\cdot\sqrt{d'} = (x'_1 + y'_1\sqrt{d'})^n.$$

This reduces the cattle problem to two easier problems: first, solving the Pell equation for d'; and second, finding the least value of n for which y'_n is divisible by $2\cdot 4657$.

Since d' is much smaller than d, Amthor could use the continued fraction algorithm for d'. In a computation that could be summarized in three pages, as in [Krumbiegel and Amthor 1880], he found the period length to be 92 and $x'_1 + y'_1\sqrt{d'}$ to be given by

$$\begin{aligned} u = {} & 109\,931986\,732829\,734979\,866232\,821433\,543901\,088049 \\ & + 50549\,485234\,315033\,074477\,819735\,540408\,986340\cdot\sqrt{4\,729494}. \end{aligned}$$

In order to save space, one can write

$$u = \bigl(300\,426607\,914281\,713365\cdot\sqrt{609} + 84\,129507\,677858\,393258\cdot\sqrt{7766}\bigr)^2.$$

$$w = 300\,426607\,914281\,713365 \cdot \sqrt{609} + 84\,129507\,677858\,393258 \cdot \sqrt{7766}$$

$$k_j = (w^{4658 \cdot j} - w^{-4658 \cdot j})^2 / 368\,238304 \qquad (j = 1, 2, 3, \ldots)$$

j-th solution	*bulls*	*cows*	*all cattle*
white	$10\,366482 \cdot k_j$	$7\,206360 \cdot k_j$	$17\,572842 \cdot k_j$
black	$7\,460514 \cdot k_j$	$4\,893246 \cdot k_j$	$12\,353760 \cdot k_j$
dappled	$7\,358060 \cdot k_j$	$3\,515820 \cdot k_j$	$10\,873880 \cdot k_j$
brown	$4\,149387 \cdot k_j$	$5\,439213 \cdot k_j$	$9\,588600 \cdot k_j$
all colors	$29\,334443 \cdot k_j$	$21\,054639 \cdot k_j$	$50\,389082 \cdot k_j$

All solutions to the cattle problem of Archimedes.

This is derived from the identity $x + y\sqrt{d} = \left(\sqrt{(x-1)/2} + \sqrt{(x+1)/2}\right)^2$, which holds whenever $x^2 = dy^2 + 1$. The regulator is found to be $R_{d'} \doteq 102.101583$.

In order to determine the least feasible value for n, Amthor developed a little theory, which one would nowadays cast in the language of finite fields and rings. Using that $p = 4657$ is a prime number for which the Legendre symbol $\left(\frac{d'}{p}\right)$ equals -1, he deduced from his theory that the least value for n divides $p+1 = 4658$; had he been a little more careful, he would have found that it must divide $(p+1)/2 = 2329 = 17 \cdot 137$ (see [Vardi 1998]). In any case, trying a few divisors, one discovers that the least value for n is actually *equal* to 2329. One has $R_d = 2329 \cdot R_{d'} \doteq 237794.586710$.

The conclusion is that the fundamental solution to the Pell equation for d itself is given by $x_1 + y_1\sqrt{d} = u^{2329}$, with u as just defined. Amthor failed to put everything together, but I did this for the convenience of the reader: for the first time in history, *all* infinitely many solutions to the cattle problem displayed in a handy little table! It does, naturally, not contain the full decimal expansion of any of the numbers asked for, but what it does contain should be considered more enlightening. For example, it enables the reader not only to verify easily that the total number of cattle in the smallest solution has 206545 decimal digits and equals 77602714 . . . 55081800, but also to discover that the number of dappled bulls in the 1494 195300th solution equals 111111 . . . 000000, a number of 308 619694 367813 digits. (Finding the middle digits is probably much harder.) There is no doubt that Archimedes, who wrote a lengthy epistle about the representation of large numbers to King Gelon (see [Dijksterhuis 1956] or [Heiberg 1913, pp. 215–259]), would have been pleased and satisfied by the solution as expressed in the table.

5. Power products

Suppose one wishes to solve the Pell equation $x^2 = dy^2 + 1$ for a given value of d. From Amthor's approach to the cattle problem we learn that for two reasons it may be wise to find the smallest divisor d' of d for which d/d' is a square: it saves time when performing the continued fraction algorithm, and it saves both time and space when expressing the final answer. There is no known algorithm for finding d' from d that is essentially faster than factoring d. In addition, if we want to determine *which* power of the fundamental solution for d' yields the fundamental solution for d — that is, the number n from the previous section — we also need to know the prime factorization of $\sqrt{d/d'}$, as well as the prime factorization of $p - \left(\frac{d'}{p}\right)$ for each prime p dividing $\sqrt{d/d'}$. Thus, if one wants to solve the Pell equation, one may as well start by factoring d. Known factoring algorithms may not be very fast for large d, but for most values of d they are still expected to be orders of magnitudes faster than any known method for solving the Pell equation [Stevenhagen 2008b].

Let it now be assumed that d is *squarefree*, and write $x_1 + y_1\sqrt{d}$ for the fundamental solution of the Pell equation, which is a unit of $\mathbb{Z}[\sqrt{d}]$. Then $x_1 + y_1\sqrt{d}$ may still be a proper power in the field $\mathbb{Q}(\sqrt{d})$ of fractions of $\mathbb{Z}[\sqrt{d}]$. For example, the least d with $y_1 > 6$ is $d = 13$, for which one has $x_1 = 649$, $y_1 = 180$, and

$$649 + 180\sqrt{13} = \left(\frac{3+\sqrt{13}}{2}\right)^6.$$

Also in the case $d = 109$, which Fermat posed as a challenge problem in 1657, the fundamental solution is a sixth power:

$$158\,070671\,986249 + 15\,140424\,455100\sqrt{109} = \left(\frac{261+25\sqrt{109}}{2}\right)^6.$$

However, this is as far as it goes: it is an elementary exercise in algebraic number theory to show that if n is a positive integer for which $x_1 + y_1\sqrt{d}$ has an n-th root in $\mathbb{Q}(\sqrt{d})$, then $n = 1, 2, 3,$ or 6, the case $n = 2$ being possible only for $d \equiv 1, 2,$ or $5 \bmod 8$, and the cases $n = 3$ and 6 only for $d \equiv 5 \bmod 8$. Thus, for large squarefree d one cannot expect to save much space by writing $x_1 + y_1\sqrt{d}$ as a power. This is also true when one allows the root to lie in a composite of quadratic fields, as we did for the cattle problem.

Let d again be an arbitrary positive integer that is not a square. Instead of powers, we consider *power products* in $\mathbb{Q}(\sqrt{d})$, that is, expressions of the form

$$\prod_{i=1}^{t}(a_i + b_i\sqrt{d})^{n_i}$$