Solving the Pell equation

H. W. Lenstra, Jr.

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 (1610–1685) 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's book [25] is the best guide; see also [5 (Chap. XII), 13, 26]. 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 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* [6, Abschn. 2, Cap. 7]. Modern textbooks usually give a formulation in terms of continued fractions, which is also due to Euler (see for example [19, Chap. 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 [25, Chap. II, § XIII]), and the first to publish such a proof was Lagrange (1736–1813) in 1768 [15].

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 $\mathbf{Z}[\sqrt{d}]$ of norm 1; here the norm $\mathbf{Z}[\sqrt{d}]^* \to \mathbf{Z}^* = \{\pm 1\}$ between unit groups multiplies each unit by its conjugate, and the units ± 1 of $\mathbf{Z}[\sqrt{d}]$ 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 nth 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; for the ring $\mathbf{Z}[\sqrt{d}]$, 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

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

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.

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 Wolfenbüttel library, and published by him in 1773 (see [17; 11, pp. 528–534]). It is now generally credited to Archimedes (see [7, 25]). 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, taken from [1], is shown in Figure 1. 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, 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 [5, 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 [8]). The first to solve the cattle problem in a satisfactory way was A. Amthor in 1880 (see [14]). 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* [18]. 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 demonstrates that the continued fraction method is not the last word on the Pell equation.

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 $\mathbf{Z}[\sqrt{d}]^* \to \mathbf{Z}^*$. 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, ... 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 [16, eq. (11.4)]); so 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 (see [20]), which depends on the fast Fourier transform, reduces this to $R_d \cdot (1 + \log d)^{c_3}$ for suitable c_3. 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 [12]. 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 $D = \{d_0 d_1^{2n} : n \in \mathbf{Z}, n \geq m\}$ 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 $\lim_{x\to\infty} \#\{d \in D : d \leq x\}/x = 1$, 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 results above, 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.

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 nth 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 (see [14]), 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

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

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 [23]). 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 in Figure 2: for the first time in history, *all* infinitely many solutions to the cattle problem are 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.) Archimedes had an interest in the representation of large numbers (see [11, pp. 215–259]), and there is little doubt that the solution in Figure 2 would have pleased and satisfied him.

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.

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 $\mathbf{Z}[\sqrt{d}]$. Then $x_1+y_1\sqrt{d}$ may still be a proper power in the field $\mathbf{Q}(\sqrt{d})$ of fractions of $\mathbf{Z}[\sqrt{d}]$. For example, the least d with $y_1>6$ is

$d = 13$, for which one has $x_1 = 649$, $y_1 = 180$, and

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

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

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

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 nth root in $\mathbf{Q}(\sqrt{d})$, then $n = 1$, 2, 3, or 6, the case $n = 2$ being possible only for $d \equiv 1$, 2, or 5 mod 8, and the cases $n = 3$ and 6 only for $d \equiv 5$ mod 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 $\mathbf{Q}(\sqrt{d})$, that is, expressions of the form

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

where t is a nonnegative integer, a_i, b_i, n_i are integers, $n_i \neq 0$, and for each i at least one of a_i and b_i is nonzero. We define the *length* of such an expression to be

$$\sum_{i=1}^{t}\big(\log|n_i| + \log(|a_i| + |b_i|\sqrt{d})\big).$$

This is roughly proportional to the amount of bits needed to specify the numbers a_i, b_i, and n_i. Each power product represents a nonzero element of $\mathbf{Q}(\sqrt{d})$, and that element can be expressed uniquely as $(a + b\sqrt{d})/c$, with a, b, $c \in \mathbf{Z}$, $\gcd(a, b, c) = 1$, $c > 0$. However, the number of bits of a, b, c will typically grow linearly with the exponents $|n_i|$ themselves rather than with their logarithms. So one avoids using the latter representation and works directly with the power products instead.

Several fundamental issues are raised by the representation of elements as power products. For example, can we recognize whether two power products represent the same element of $\mathbf{Q}(\sqrt{d})$, by means of a polynomial time algorithm? Here 'polynomial time' means, as before, that the run time is bounded by a polynomial function of the length of the input, which in this case equals the sum of the lengths of the two given power products. Similarly, can we decide in polynomial time whether a given power product represents an element of the form $a + b\sqrt{d}$ with a, $b \in \mathbf{Z}$, that is, an element of $\mathbf{Z}[\sqrt{d}]$? If it does, can we decide whether one has $a^2 - db^2 = 1$ and a, $b > 0$, so that we have a solution to Pell's equation, and can we compute the residue classes of a and b modulo a given positive integer m, all in polynomial time?

All questions just raised have affirmative answers, even in the context of general algebraic number fields. Algorithms proving this were exhibited recently by Guoqiang Ge (see [9, 10]). In particular, one can efficiently decide whether a given power product represents a solution to Pell's equation, and if it does, one can efficiently compute any desired number of 'least significant' decimal digits of that solution; taking the logarithm of the power product, one can do the same for the *leading* digits, and for the *number* of decimal digits, except possibly in the probably very rare cases that a or b is excessively close to a power of 10. There is *no* known polynomial time algorithm for deciding whether a given power product represents the *fundamental* solution to Pell's equation.

Infrastructure

Suppose now that, given d, we are not asking for the fundamental solution $x_1 + y_1\sqrt{d}$ to Pell's equation, but for a power product in $\mathbf{Q}(\sqrt{d})$ that represents it. The following theorem summarizes essentially all that is rigorously known about the smallest length of such a power product and about algorithms for finding one.

Theorem. *There are positive real numbers c_6 and c_7 with the following properties.*

(a) *For each positive integer d that is not a square there exists a power product that represents the fundamental solution to Pell's equation and that has length at most $c_6 \cdot (\log d)^2$.*

(b) *The problem of computing a power product representing the fundamental solution to Pell's equation is 'polynomial time equivalent' to the problem of computing an integer $\tilde{R}_d$ with $|R_d - \tilde{R}_d| < 1$.*

(c) *There is an algorithm that given d computes a power product representing the fundamental solution to Pell's equation in time at most $R_d^{1/2} \cdot (1 + \log d)^{c_7}$.*

Part (a) of the theorem, which is taken from [3], implies that the question we are asking does admit a brief answer, so that there is no obvious obstruction to the existence of a polynomial time algorithm for *finding* such an answer.

Part (b), which is not formulated too rigorously, asserts the existence of two polynomial time algorithms. The first takes as input a power product $\prod_i (a_i + b_i\sqrt{d})^{n_i}$ representing the fundamental solution to the Pell equation and gives as output an integer approximation to the regulator. There is no surprise here, one just uses the formula $R_d = \sum_i n_i \log |a_i + b_i\sqrt{d}|$ and applies a polynomial time algorithm for approximating logarithms (see [2]). The second algorithm takes as input the number d as well as an integer approximation $\tilde{R}_d$ to R_d, and it computes a power product representing the fundamental solution to Pell's equation. Since the algorithm runs in polynomial time, the length of the output is polynomially bounded, and this is in fact the way part (a) of the theorem is proved.

The key notion underlying the second algorithm is that of 'infrastructure', a word coined by Shanks (see [22]) to describe a certain multiplicative structure that he detected within the period of the continued fraction expansion of $\sqrt{d}$. It was subsequently shown (see [16]) that this period can be 'embedded' in a circle group of 'circumference' R_d, the embedding preserving the cyclical structure. In the modern terminology of Arakelov theory, one may describe that circle group as the kernel of the natural map $\mathrm{Pic}^0\,\overline{\mathbf{Z}[\sqrt{d}]} \to \mathrm{Pic}\,\mathbf{Z}[\sqrt{d}]$ from the group of 'metrized line bundles of degree 0' on the 'arithmetic curve' corresponding to $\mathbf{Z}[\sqrt{d}]$ to the usual class group of invertible ideals. By means of Gauss's reduced binary quadratic forms one can do explicit computations in $\mathrm{Pic}^0\,\overline{\mathbf{Z}[\sqrt{d}]}$ and in its 'circle' subgroup. For a fuller explanation of these notions and their algorithmic use we refer to the literature [3, 16, 21, 22, 27].

The equivalence stated in part (b) of the theorem has an interesting feature that is not commonly encountered in the context of equivalences. Namely, one may achieve an improvement by going 'back-and-forth'. Thus, starting from a power product representing the fundamental solution, one can first use it to compute $\tilde{R}_d$, and next use $\tilde{R}_d$ to find a *second* power product, possibly of smaller length than the initial one. And conversely, starting from any rough approximation to R_d one can compute a power product and use it to compute R_d to any desired accuracy.

The algorithm referred to in part (c) is the fastest rigorously proven algorithm for computing a power product as desired. Its run time is roughly the square root of the run time of the continued fraction algorithm. It again makes use of the infrastructure just discussed, combining it with a search technique that is known as the 'baby step-giant step' method. The power product coming out of the algorithm may not have a very small length, but one can easily do something about this by using the algorithms of part (b). Our estimates for R_d show that the run time is at most $d^{1/4} \cdot (1 + \log d)^{c_8}$ for some c_8; here the exponent $1/4$ can be improved to $1/5$ if one is willing to assume certain generalized Riemann hypotheses (see [21]). According to [4], part (c) is valid with $c_7 = 1 + \epsilon$ for all $\epsilon > 0$ and all d exceeding a bound depending on ϵ.

Mathematically the infrastructure methods have great interest. Algorithmically one conjectures that something faster is available. But as we shall see, the final victory may belong to the infrastructure.

Smooth numbers

The algorithms for solving Pell's equation that we saw so far have an exponential run time as a function of $\log d$. One prefers to have an algorithm whose run time is polynomial in $\log d$. The method that we shall now discuss is believed to have a run time that is 'halfway' between exponential and polynomial. Like many subexponential algorithms in number theory, it makes use of *smooth numbers*, that is, nonzero integers that up to sign are built up from small prime factors. Smooth numbers have been used with great success in the design of algorithms for factoring integers and for computing discrete logarithms in multiplicative groups of rings. Here we shall see how they can be used for the solution of Pell's equation as well.

Instead of giving a formal description, we illustrate the algorithm on the case $d = 4\,729494 = 2 \cdot 3 \cdot 7 \cdot 11 \cdot 29 \cdot 353$ derived from the cattle problem. The computation is less laborious and more entertaining than the expansion of $\sqrt{d}$ in a continued fraction performed by Amthor. We shall explain the method on an intuitive level only; readers desirous to see its formal justification should acquaint themselves with the basic theorems of algebraic number theory.

The smooth numbers that the algorithm operates with are not ordinary integers, but elements of the ring $\mathbf{Z}[\sqrt{d}]$, with d as just chosen. There is a natural way of extending the notion of smoothness to such numbers. Namely, for $\alpha = a + b\sqrt{d} \in \mathbf{Q}(\sqrt{d})$, with a, $b \in \mathbf{Q}$, write $\alpha' = a - b\sqrt{d}$. Then $\alpha \mapsto \alpha'$ yields an automorphism of the field $\mathbf{Q}(\alpha)$ and the ring $\mathbf{Z}[\alpha]$, and the *norm* map $N\colon \mathbf{Q}(\sqrt{d}) \to \mathbf{Q}$ defined by $N(\alpha) = \alpha\alpha' = a^2 - db^2$ respects multiplication. It is now natural to expect that an element α of $\mathbf{Z}[\sqrt{d}]$ is smooth if and only if α' is smooth; so one may as well pass to their product $N(\alpha)$, which is an ordinary integer, and *define* α to be smooth if $|N(\alpha)|$ is built up from prime numbers that lie below a certain bound. The size of this bound depends on the circumstances; in the present computation we choose it empirically.

The first step in the algorithm is to find a good supply of smooth numbers $a+b\sqrt{d}$ in $\mathbf{Z}[\sqrt{d}]$, or, equivalently, pairs of integers a, b for which a^2-db^2 is smooth. One does this by trying $b=1, 2, 3, \ldots$ in succession, and trying integers a in the neighborhood of $b\sqrt{d}$; then $|a^2-db^2|$ is fairly small, which increases its chance to be smooth. For example, with $b=1$ one finds for a near $b\sqrt{d} \doteq 2174.74$ the following smooth values of a^2-d:

$$\begin{aligned}
2156^2-d&=-2\cdot 7\cdot 11\cdot 17\cdot 31, & 2178^2-d&=2\cdot 3\cdot 5\cdot 11\cdot 43,\\
2162^2-d&=-2\cdot 5^3\cdot 13\cdot 17, & 2184^2-d&=2\cdot 3\cdot 7\cdot 31^2,\\
2175^2-d&=3\cdot 13\cdot 29, & 2187^2-d&=3\cdot 5^2\cdot 23\cdot 31.
\end{aligned}$$

For $b=2$, 3, 4, one finds, restricting to values of a that are coprime to b:

$$\begin{aligned}
4329^2-2^2d&=-3\cdot 5\cdot 17^2\cdot 41, & 4399^2-2^2d&=5^2\cdot 13\cdot 31\cdot 43,\\
4341^2-2^2d&=-3\cdot 5\cdot 17^3, & 6514^2-3^2d&=-2\cdot 5^3\cdot 13\cdot 41,\\
4351^2-2^2d&=5^2\cdot 23^2, & 6524^2-3^2d&=-2\cdot 5\cdot 7\cdot 41,\\
4363^2-2^2d&=13^2\cdot 17\cdot 41, & 6538^2-3^2d&=2\cdot 7\cdot 13\cdot 23\cdot 43,\\
4389^2-2^2d&=3\cdot 5\cdot 7\cdot 11\cdot 13\cdot 23, & 8699^2-4^2d&=17\cdot 41.
\end{aligned}$$

The prime numbers occurring in these sixteen factorizations are the small prime factors 2, 3, 7, 11, 29 of d, as well as the prime numbers $p\le 43$ with $\left(\frac{d}{p}\right)=1$. It is only the latter primes that matter, and there are seven of them: 5, 13, 17, 23, 31, 41, and 43. It is important that the number of smooth expressions a^2-db^2 exceeds the number of those primes, which is indeed the case: $16>7$. If one uses only the prime numbers up to 31 and the eight factorizations that do not contain 41 or 43, there is still a good margin: $8>5$. Thus, one decides to work with the 'smoothness bound' 31.

The next step is to write down the prime *ideal* factorizations of the eight numbers $(a+b\sqrt{d})/(a-b\sqrt{d})$. Consider, for example, the case $a=2162$, $b=1$. Since 2162^2-d contains a factor 13, the element $2162+\sqrt{d}$ has a prime ideal factor of norm 13, and from $2162\equiv 4 \bmod 13$ one sees that this is the prime ideal $\mathbf{p}_{13}=(13, 4+\sqrt{d})$; it is the kernel of the ring homomorphism $\mathbf{Z}[\sqrt{d}]\to\mathbf{Z}/13\mathbf{Z}$ sending $\sqrt{d}$ to $-4 \pmod{13}$. The conjugate prime ideal $\mathbf{q}_{13}=(13,4-\sqrt{d})$ then occurs in $2162-\sqrt{d}$. Likewise, $2162+\sqrt{d}$ is divisible by the cube of the prime ideal $\mathbf{p}_5=(5,2+\sqrt{d})$ and by $\mathbf{p}_{17}=(17,3+\sqrt{d})$, and $2162-\sqrt{d}$ by $\mathbf{q}_5^3\mathbf{q}_{17}$, where $\mathbf{q}_5=(5,2-\sqrt{d})$ and $\mathbf{q}_{17}=(17,3-\sqrt{d})$. Finally, $2162+\sqrt{d}$ has the prime ideal factor $(2,\sqrt{d})$, but since 2 divides d, this prime ideal equals its own conjugate, so it cancels when one divides $2162+\sqrt{d}$ by its conjugate. Altogether one finds the prime ideal factorization

$$\big((2162+\sqrt{d})/(2162-\sqrt{d})\big) = (\mathbf{p}_5/\mathbf{q}_5)^3 \cdot (\mathbf{p}_{13}/\mathbf{q}_{13}) \cdot (\mathbf{p}_{17}/\mathbf{q}_{17}).$$

As a second example, consider $a = 4351$, $b = 2$. We have $4351^2 - 2^2 d = 5^2 \cdot 23^2$, and from $4351/2 \equiv -2 \bmod 5$ one sees that $4351 + 2\sqrt{d}$ belongs to $\mathbf{q}_5$ rather than $\mathbf{p}_5$. Similarly, $4351/2 \equiv 2 \bmod 23$ implies that it belongs to $\mathbf{p}_{23} = (23, 2+\sqrt{d})$. Writing $\mathbf{q}_{23} = (23, 2-\sqrt{d})$, one obtains

$$\big((4351+2\sqrt{d})/(4351-2\sqrt{d})\big) = (\mathbf{p}_5/\mathbf{q}_5)^{-2} \cdot (\mathbf{p}_{23}/\mathbf{q}_{23})^2.$$

Doing this for all eight pairs a, b, one arrives at the table in Figure 3. The first row lists the prime numbers p we are using. The first column lists the eight expressions $\alpha = a + b\sqrt{d}$. In the αth row and the pth column, one finds the exponent of $\mathbf{p}_p/\mathbf{q}_p$ in the prime ideal factorization of α/α'; here $\mathbf{p}_p$, $\mathbf{q}_p$ are as above, with $\mathbf{p}_{31} = (31, 14+\sqrt{d})$ and $\mathbf{q}_{31} = (31, 14-\sqrt{d})$. Thus, each α gives rise to an 'exponent vector' that belongs to $\mathbf{Z}^5$.

The third step in the algorithm is finding linear relations with integer coefficients between the eight exponent vectors. The set of such relations forms a free abelian group of rank 3, which is 8 minus the rank of the 8×5 matrix formed by the eight vectors. A set of three independent generators for the relation group is given by the last three columns of Figure 3; in general, one can find such a set by applying techniques of linear algebra over $\mathbf{Z}$.

In the final step of the algorithm one inspects the relations one by one. Consider for example the first relation. It expresses that the sum of the exponent vectors corresponding to $2156+\sqrt{d}$ and $2162+\sqrt{d}$ equals the sum of the exponent vectors for $2187+\sqrt{d}$ and $4389+2\sqrt{d}$. In other words, if we put

$$\alpha = \frac{(2156+\sqrt{d})\cdot(2162+\sqrt{d})}{(2187+\sqrt{d})\cdot(4389+2\sqrt{d})},$$

then the element $\epsilon = \alpha/\alpha'$ has all exponents in its prime ideal factorization equal to 0. This is the same as saying that ϵ is a unit $x + y\sqrt{d}$ of the ring $\mathbf{Z}[\sqrt{d}]$; also, the norm $\epsilon\epsilon' = x^2 - dy^2$ of this unit equals $N(\alpha)/N(\alpha') = 1$, so we obtain an integral solution to Pell's equation $x^2 - dy^2 = 1$, except that it is uncertain whether x and y are positive. We can write $\epsilon = \alpha/\alpha' = \alpha^2/N(\alpha)$, where the prime factorization of $N(\alpha)$ is available from the factorizations of $a^2 - db^2$ that we started with; one finds in this manner the following two power product representations of ϵ:

$$\begin{aligned}\epsilon &= \frac{(2156+\sqrt{d})\cdot(2162+\sqrt{d})\cdot(2187-\sqrt{d})\cdot(4389-2\sqrt{d})}{(2156-\sqrt{d})\cdot(2162-\sqrt{d})\cdot(2187+\sqrt{d})\cdot(4389+2\sqrt{d})} \\ &= \frac{3^2\cdot 23^2\cdot(2156+\sqrt{d})^2\cdot(2162+\sqrt{d})^2}{2^2\cdot 17^2\cdot(2187+\sqrt{d})^2\cdot(4389+2\sqrt{d})^2}.\end{aligned}$$

In the second representation, ϵ is 'visibly' a square, or, equivalently, $N(\alpha)$ is a square; this is a bad sign, since it is certain to happen when $\epsilon = 1$, in which case one has $\alpha \in \mathbf{Q}$, $N(\alpha) = \alpha^2$, $x = 1$, and $y = 0$. That is indeed what occurs here. (Likewise, it would have been a bad sign if ϵ were visibly $-d$ times a square; this is certain to happen if $\epsilon = -1$.) In the present case, the numbers are small enough that one can directly verify that $\epsilon = 1$. For larger power products, one can decide whether ϵ equals ± 1 by computing $\log|\epsilon|$ to a suitable precision and proving that the logarithm of a positive unit of $\mathbf{Z}[\sqrt{d}]$ cannot be close to 0 without being equal to 0.

Thus, the first relation disappointingly gives rise to a trivial solution to the Pell equation. The reader may check that the unit

$$\frac{29^2\cdot(4351+2\sqrt{d})^2\cdot(4389+2\sqrt{d})^4}{5^4\cdot 7^2\cdot 11^2\cdot 23^4\cdot(2175+\sqrt{d})^4}$$

obtained from the second relation is also equal to 1. The third relation yields the unit

$$\eta = \frac{2^4\cdot 5^{14}\cdot(2175+\sqrt{d})^{18}\cdot(2184+\sqrt{d})^{10}\cdot(2187+\sqrt{d})^{20}\cdot(4341+2\sqrt{d})^6}{3^{27}\cdot 7^5\cdot 29^9\cdot 31^{20}\cdot(2162+\sqrt{d})^{18}\cdot(4351+2\sqrt{d})^{10}}.$$

Since this is not visibly a square, we can be certain that it is not 1. Since it is positive, it is not -1 either. So η is of the form $x + y\sqrt{d}$, where x, $y \in \mathbf{Z}$ satisfy $x^2 - dy^2 = 1$ and $y \neq 0$; thus, $|x|$, $|y|$ solve Pell's equation. From the power product, one computes the logarithm of the unit to be about 102.101583. This implies that $\eta > 1$, so that η is the largest of the four numbers η, $\eta' = 1/\eta$, $-\eta$, and $-\eta'$; in other words, $x + y\sqrt{d}$ is the largest of the four numbers $\pm x \pm y\sqrt{d}$, which is equivalent to x and y being *positive*. In general one can achieve this by first replacing η by $-\eta$ if η is negative, and next by η' if $\eta < 1$.

We conclude that the power product defining η does represent a solution to Pell's equation. The next question is whether it is the *fundamental* solution. In the present case we can easily confirm this, since from Amthor's computation we know that $R_d \doteq 102.101583$, and the logarithm of any *non*fundamental solution would be at least $2 \cdot R_d$. Therefore, η is equal to the solution u found by Amthor, and it is indeed fundamental. In particular, the numbers $\log \eta \doteq 102.101583$ and $\log u \doteq 102.101583$ are exactly equal, not just to a precision of six decimals.

The power product representation we found for η is a little more compact than the standard representation we gave for u. Indeed, its length, as defined earlier, is about 93.099810, as compared to $R_d \doteq 102.101583$ for u. The power product

$$\frac{(2175+\sqrt{d})^{18}(2184+\sqrt{d})^{10}(2187+\sqrt{d})^{20}(4341+2\sqrt{d})^{6}(2162-\sqrt{d})^{18}(4351-2\sqrt{d})^{10}}{(2175-\sqrt{d})^{18}(2184-\sqrt{d})^{10}(2187-\sqrt{d})^{20}(4341-2\sqrt{d})^{6}(2162+\sqrt{d})^{18}(4351+2\sqrt{d})^{10}},$$

which also represents u, has length about 125.337907.

Performance

The smooth numbers method for solving Pell's equation exemplified in the previous section can be extended to any value of d. There is unfortunately not much one can currently prove either about the run time or about the correctness of the method. Regarding the run time, however, one can make a reasonable conjecture.

For $x > \mathrm{e}$, write

$$L(x) = \exp\bigl(\sqrt{(\log x)\log\log x}\bigr).$$

The conjecture is that, for some positive real number c_9 and all $d > 2$, the smooth numbers method runs in time at most $L(d)^{c_9}$. This is, at a doubly logarithmic level, the exact average of $x^{c_9} = \exp(c_9 \log x)$ and $(\log x)^{c_9} = \exp(c_9 \log\log x)$; so conjecturally, the run time of the smooth numbers method is in a sense halfway between exponential time and polynomial time.

The main ingredient of the heuristic reasoning leading to the conjecture is the following proven theorem: for fixed positive real numbers c, c', and $x \to \infty$, the probability for a random positive integer $\leq x^{c'}$ (drawn from a uniform distribution) to have all its prime factors $\leq L(x)^c$ equals $1/L(x)^{c'/(2c)+o(1)}$. This theorem explains the importance of the function L in the analysis of algorithms depending on smooth numbers. Other ingredients of the heuristic run time analysis are the belief that the expressions $a^2 - db^2$ that one hopes to be smooth are so with the same probability as if they were random numbers, and the belief that the units produced by the algorithm have a substantial probability of being different from ± 1. These beliefs appear to be borne out in practice.

Probably one can take $c_9 = 3/\sqrt{8} + \epsilon$ in the conjecture just formulated, for any $\epsilon > 0$ and all d exceeding a bound depending on ϵ; one has $3/\sqrt{8} \doteq 1.06066$. One of the bottlenecks is the time spent on solving a large sparse linear system over $\mathbf{Z}$. If one is very optimistic about developing a better algorithm for doing this, it may be possible to achieve 1 instead of $3/\sqrt{8}$.

The smooth numbers method needs to be supplemented with an additional technique if one wishes to be reasonably confident that the unit it produces is the fundamental solution to Pell's equation. We forgo a discussion of this technique, since there is no satisfactory method for testing whether it achieves its purpose. More precisely, there is currently no known way of verifying in subexponential time that a solution to the Pell equation that is given by means of a power product is the fundamental one. The most promising technique for doing this employs the *analytic class number formula*, but its effectiveness depends on the truth of the *generalized Riemann hypothesis*. The latter hypothesis, abbreviated 'GRH', asserts that there does not exist an algebraic number field whose associated zeta function has a complex zero with real part greater than $\frac{1}{2}$. The GRH can also be used to corroborate the heuristic run time analysis, albeit in a probabilistic setting. This leads to the following theorem.

Theorem. *There is a probabilistic algorithm that for some positive real number c_{10} has the following properties.*

(a) *Given any positive integer d that is not a square, the algorithm computes a positive integer R that differs by less than* 1 *from some positive integer multiple $m \cdot R_d$ of R_d.*

(b) *If the GRH is true, then* (a) *is valid with $m = 1$.*

(c) *If the GRH is true, then for each $d > 2$ the expected run time of the algorithm is at most $L(d)^{c_{10}}$.*

The algorithm referred to in the theorem is *probabilistic* in the sense that it employs a random number generator; every time the random number generator is called, it draws, in unit time, a random bit from the uniform distribution, independently of previously drawn bits. The run time and the output of a probabilistic algorithm depend not only on the input, but also on the random bits that are drawn; so given the input, they may be viewed as random variables. In the current case, the expectation of the run time for fixed d is considered in part (c) of the theorem, and (a) and (b) describe what we know about the output. In particular, the algorithm always terminates, and if GRH is true, then it is guaranteed to compute an integer approximation to the regulator.

The theorem just stated represents the efforts of several people, an up-to-date list of references being given by Ulrich Vollmer [24]. According to a recent unpublished result of Ulrich Vollmer, one may take $c_{10} = 3/\sqrt{8} + \epsilon$ for any $\epsilon > 0$ and all d exceeding a bound depending on ϵ; this improves the value $\sqrt{2} + \epsilon$ found in [24].

The last word on algorithms for solving Pell's equation has not been spoken yet. Very recently, Sean Hallgren exhibited a *quantum algorithm* that computes, in polynomial time, a power product representing the fundamental solution. His algorithm depends on infrastructure, but not on smooth numbers. For practical purposes, the smooth numbers method will remain preferable until quantum computers become available.

Acknowledgements

This paper was written while I held the 2000–2001 HP-MSRI Visiting Research Professorship. I thank Sean Hallgren, Mike Jacobson, Jr., and Ulrich Vollmer for answering my questions, John Voight for writing a first draft, and Bart de Smit for numerical assistance. A special word of thanks is due to Hugh Williams, whose version [27] of the same story contains many details omitted in mine.

References

1. Archimedes, *The cattle problem*, in English verse by S. J. P. Hillion & H. W. Lenstra jr., Mercator, Santpoort, 1999.
2. R. P. Brent, *Fast multiple-precision evaluation of elementary functions*, J. Assoc. Comput. Mach. **23** (1976), 242–251.
3. J. Buchmann, C. Thiel, H. C. Williams, *Short representation of quadratic integers*, Computational algebra and number theory (Sydney, 1992), Math. Appl. **325**, Kluwer, Dordrecht, 159–185.
4. J. Buchmann, U. Vollmer, in preparation.
5. L. E. Dickson, *History of the theory of numbers*, vol. II, *Diophantine analysis*, Carnegie Institution of Washington, Washington, 1920.
6. L. Euler, *Vollständige Anleitung zur Algebra*, Zweyter Theil, Kays. Acad. der Wissenschaften, St. Petersburg, 1770; *Opera mathematica*, ser. I, vol. 1, B. G. Teubner, Leipzig, 1911. English translation: *Elements of algebra*, Springer, New York, 1984.
7. P. M. Fraser, *Ptolemaic Alexandria*, Oxford University Press, Oxford, 1972.
8. C. C. Grosjean, H. E. de Meyer, *A new contribution to the mathematical study of the cattle-problem of Archimedes*, Th. M. Rassias, N. J. Teaneck (eds), Constantin

Carathéodory: an international tribute, Vols. I–II, World Sci. Publishing, 1991, 404–453.

9. G. Ge, *Algorithms related to multiplicative representations*, Ph. D. thesis, University of California, Berkeley, 1993.
10. G. Ge, *Recognizing units in number fields*, Math. Comp. **63** (1994), 377–387.
11. J. L. Heiberg, *Archimedis opera omnia cum commentariis Eutocii*, vol. II, B. G. Teubner, Leipzig, 1913; reprint, Stuttgart, 1972.
12. L.-K. Hua, *On the least solution to Pell's equation*, Bull. Amer. Math. Soc. **48** (1942), 731–735; *Selected papers*, Springer, New York, 1983, 119–123.
13. H. Konen, *Geschichte der Gleichung* $t^2 - Du^2 = 1$, S. Hirzel, Leipzig, 1901.
14. B. Krumbiegel, A. Amthor, *Das Problema Bovinum des Archimedes*, Historisch-literarische Abteilung der Zeitschrift für Mathematik und Physik **25** (1880), 121–136, 153–171.
15. J.-L. de la Grange, *Solution d'un problème d'arithmétique*, Mélanges de philosophie et de mathématique de la Société Royale de Turin **4** (1766–1769), 44–97 (this paper was written and submitted for publication in 1768, and it appeared in 1773; see [25, pp. 314–315]); *Œuvres*, vol. I, Gauthier-Villars, Paris, 1867, 669–731.
16. H. W. Lenstra, Jr., *On the calculation of regulators and class numbers of quadratic fields*, J. Armitage (ed.), Journées Arithmétiques 1980, London Math. Soc. Lecture Note Ser. **56**, Cambridge University Press, Cambridge, 1982, 123–150.
17. G. E. Lessing, *Zur Griechischen Anthologie*, Zur Geschichte und Litteratur, Zweyter Beytrag, Fürstl. Waysenhaus-Buchhandlung, Braunschweig, 1773, nr. XIII, 419–446.
18. H. L. Nelson, *A solution to Archimedes' cattle problem*, J. Recreational Math. **13** (3) (1980–81), 162–176.
19. I. Niven, H. S. Zuckerman, H. L. Montgomery, *An introduction to the theory of numbers*, 5th ed., John Wiley & Sons, New York, 1991.
20. A. Schönhage, *Schnelle Berechnung von Kettenbruchentwicklungen*, Acta Inform. **1** (1971), 139–144.
21. R. J. Schoof, *Quadratic fields and factorization*, H. W. Lenstra, Jr., R. Tijdeman (eds), Computational methods in number theory, Part II, Math. Centre Tracts **155**, Math. Centrum, Amsterdam, 1982, 235–286.
22. D. Shanks, *The infrastructure of a real quadratic field and its applications*, Proceedings of the Number Theory Conference (Univ. Colorado, Boulder, Colo., 1972), Univ.

Colorado, Boulder, 217–224.

23. I. Vardi, *Archimedes' cattle problem*, Amer. Math. Monthly **105** (1998), 305–319.
24. U. Vollmer, *Asymptotically fast discrete logarithms in quadratic number fields*, W. Bosma (ed.), Algorithmic number theory (ANTS-IV), Lecture Notes in Comput. Sci. **1838**, Springer, Berlin, 2000, 581–594.
25. A. Weil, *Number theory, an approach through history*, Birkhäuser, Boston, 1984.
26. E. E. Whitford, *The Pell equation*, E. E. Whitford, New York, 1912.
27. H. C. Williams, *Solving the Pell equation*, Proc. Millennial Conference on Number Theory, A. K. Peters, to appear.

Department of Mathematics # 3840, University of California,
Berkeley, CA 94720–3840, U. S. A. hwl@math.berkeley.edu

Mathematisch Instituut, Universiteit Leiden,
Postbus 9512, 2300 RA Leiden, The Netherlands hwl@math.leidenuniv.nl

Figure 1.

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!

Figure 2.

All solutions to the cattle problem of Archimedes

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

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

jth 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	$29\,334443 \cdot k_j$	$21\,054639 \cdot k_j$	$50\,389082 \cdot k_j$

Figure 3.

	5	13	17	23	31			
$2156+\sqrt{d}$	0	0	-1	0	-1	1	0	0
$2162+\sqrt{d}$	3	1	1	0	0	1	0	-9
$2175+\sqrt{d}$	0	1	0	0	0	0	-2	9
$2184+\sqrt{d}$	0	0	0	0	2	0	0	5
$2187+\sqrt{d}$	2	0	0	1	-1	-1	0	10
$4341+2\sqrt{d}$	-1	0	3	0	0	0	0	3
$4351+2\sqrt{d}$	-2	0	0	2	0	0	1	-5
$4389+2\sqrt{d}$	1	1	0	-1	0	-1	2	0