The cakravāla method

S. Raghavan

The object of this article spurred by a book review in *Current Science* (1996, 70, 753–754) is to try to place in proper perspective the well-known work of Indian mathematicians especially their kuttaka, bhāvanā and cakravāla methods evolved in connection with the solution in integers of certain indeterminate equations of degree one or two. In this effort, we base ourselves completely on André Weil's masterly, unbiased and incisive analysis of this topic in his beautiful book *Number Theory – An Approach Through History – From Hammurapi to Legendre*[1]. Any help needed to understand related results concerning continued fractions can be readily secured from the book *An Introduction to the Theory of Numbers*[2] (especially chapter 7), although we endeavour to skirt a reference thereto by the reader by providing as self-contained an account as possible of facts related to simple continued fractions (in the sequel).

Solving indeterminate equations of the first degree, say $ax + by + cz + \cdots = m$ in the variables $x, y, z, \ldots$ with integral coefficients $a, b, c, \ldots$ for integer values of the variables has been the key to cracking puzzles or finding integral solutions of simultaneous simple linear congruences, e.g. $x \equiv k(\text{mod } r)$, $y \equiv l(\text{mod } s)$ requiring $x - k$, $y - l$ to be divisible by integers r, s respectively for given integers k, l. A single linear congruence $px \equiv m(\text{mod } q)$ in the variable x for the modulus q is equivalent to a linear equation $px - qy = m$ in two variables x, y. Euclid's algorithm for finding the greatest common divisor of the integers p and q provides a method of solving the last-mentioned equation; expanding the rational number p/q (for $q \neq 0$) as a simple continued fraction gives an alternative (but equivalent) approach to the same problem. A clear description of the general solution of this linear equation in x, y can be found in the Sanskrit text *Aryabhatiya* of the 5th or 6th century A.D. In subsequent Sanskrit treatises, this method came to be known as the 'kuttaka' (='pulveriser') and is indeed a kind of forerunner of Fermat's powerful principle of 'infinite descent'. It is also the 'first ever explicit description' of the general solution from anywhere, not taking into account China. Indian astronomy at that time was under the influence of Greek sources and yet perhaps one cannot with certainty attribute 'kuttaka' to Greek mathematics. In utter disregard or (possible) ignorance of this Indian dimension to the 'kuttaka' and of the connection with the seventh book of Euclid's *Elementa*, Bachet inserted a strong claim to the method as his own, in the second edition of his book on *Problèmes plaisants et délectables*.

Let N be a natural number which is not the square of an integer; the simplest example is $N = 2$. In view of the connection with finding close approximations to the irrational number $\sqrt{N}$ by rational numbers, indeterminate equations of degree 2 like $x^2 - Ny^2 = \pm m$ for given integers m must have been indeed investigated by the Greek mathematicians. Special cases of the 'composition formulae'

$$(x^2 - Ny^2)(z^2 - Nt^2) = (xz \pm Nyt)^2 - N(xt \pm yz)^2, \qquad (1)$$

e.g. for $N = 3$, $z = 5$, $t = 3$ may have been applied by Archimedes for finding rational approximations to $\sqrt{3}$. However, the composition formula (1) occurs *explicitly* in the work of Brahmagupta (in the 7th century) while seeking the solution in integers x, y of equations of the form $x^2 - Ny^2 = \pm m$ for any fixed natural number m. For $m = 1$, the name Pell's equation has come to stay for the diophantine equation

$$x^2 - Ny^2 = 1. \qquad (2)$$

Such equations 'do occur in Diophantus..., but it is a rational solution that is asked for, even when accidentally a solution in integers is obtained...'. It may be reasonable to suppose that Archimedes was interested in equations of this type or, to be very optimistic, that he had even found a general method of solving equations like (2).

In the book entitled *Algebra with Arithmetic and Mensuration, from the Sanscrit of Brahmegupta and Bhάscara* by H. T. Colebrooke, one finds an entire section dealing with Brahmagupta's investigations (in the seventh century) on Vargaprakrti – the solution of equations $Ny^2 + m = x^2$ in integers x, y with N as above and m, a non-zero integer; N is called 'gunaka' (or 'prakrti') and m the ksepa (=additive). If the triple $(x, y; m)$ stands for a solution in *integers* x, y of the equation $x^2 - Ny^2 = m$, keeping N fixed all the while, the composition formula (1) is given by Brahmagupta in the form

$$(x, y; m) \cdot (z, t; n) \to (xz \pm Nyt, xt \pm yz; mn).$$

These laws acquired, in the post-Brahmagupta manuscripts in India, the name bhāvanā ('production') rules; in modern parlance, this is just the 'multiplicativity of the norm'. Brahmagupta shows, how composition of $(x, y; m)$ with a triple $(p, q; 1)$ gives a triple $(x', y'; m)$ for the same additive m. Under composition with itself, any triple $(x,y;m)$ yields a solution in rational numbers x'/m, y'/m of the equation $(x'/m)^2 - N(y'/m)^2 = 1$ and indeed a triple $(x'/m, y'/m; 1)$ if x'/m and y'/m are actually integers. Applying then his bhāvanā, Brahmagupta solves equation (2) for several cases of N including ones like $N = 92$ or $N = 83$. For $m = -1$ or ± 2, composition of a triple $(p, q; m)$ with itself leads to $(p^2 + Nq^2, 2pq; 1)$ or $((p^2 + Nq^2)/2, pq; 1)$ respectively.

Despite the remarkable results of Brahmagupta, the general solution of (2) is still not at hand. Actually, the cakravāla ('cyclic' method) for getting the general solution of equation (2) is to be found much later, around the twelfth century, in the work of Bhāskara; nearly the same description of the cakravāla is provided in a commentary of the eleventh century by 'an otherwise unknown author' Jayadeva, leaving one to guess who was the true inventor of the cakravāla. Following Weil, we can see how the brilliant cakravāla arises in a natural manner from the work of earlier Indian mathematicians.

Starting from a triple (p_0, q_0, m_0) with 'small' m_0, the idea is to use the bhāvanā to get a triple $(p_1, q_1; m_1)$ with m_1 also 'small' and eventually hope to hit upon a triple $(u, v; 1)$ giving a non-trivial solution of equation (2), of course. First we can assume, without loss of generality that the greatest common divisor d of p_0 and q_0 in the initial triple is already equal to 1, since if $d > 1$, we could start instead from $(p_0/d, q_0/d; m_0/d^2)$. Since p_0 and q_0 have greatest common divisor 1, so have q_0 and m_0 clearly. Then the kuttaka readily enables us to find an integer x_0 such that m_0 divides $p_0 + q_0x_0$ (i.e. solve the arithmetical congruence $q_0x_0 \equiv -p_0 \pmod{m_0}$). If, in addition, the chosen x_0 is fixed in its residue class modulo m_0 so as to satisfy the inequalities $x_0 < \sqrt{N} < x_0 + |m_0|$, then we see that $\sqrt{N} + x_0$ cannot be negative *provided that* $|m_0| < 2\sqrt{N}$ (since '$\sqrt{N} + x_0 < 0$' would imply that $2\sqrt{N} < \sqrt{N} + x_0 + |m_0| < |m_0|$) and consequently for $|m_0| < 2\sqrt{N}$,

$$0 < N - x_0^2 = (\sqrt{N} - x_0)(\sqrt{N} + x_0) < |m_0|(\sqrt{N} + \sqrt{N}) = 2|m_0|\sqrt{N}. \quad (3)$$

Composition of $(p_0, q_0; m_0)$ with $(x_0, 1; x_0^2 - N)$ gives rise to the triple $(p_1, q_1; m_1)$ where

$$p_1 := (p_0x_0 + Nq_0)/(m_0,$$
$$q_1 := (p_0 + q_0x_0)/m_0,$$
$$m_1 := (x_0^2 - N)/m_0 \quad (4)$$

are clearly integers. (In fact, by composition, $(m_0p_1)^2 - N(m_0q_1)^2 = m_0(x_0^2 - N) = m_0^2 m_1$ and $q_0^2(x_0^2 - N) = q_0^2x_0^2 - p_0^2 + m_0 = m_0((q_0x_0 - p_0)\ (q_0x_0 + p_0)/m_0 + 1)$ is divisible by m_0, i.e. $x_0^2 - N$ is divisible by m_0 in view of the greatest common divisor of m_0 and q_0^2 being 1 by our assumption above. Thus m_1 is an integer while the congruence condition on x_0 implies that q_1 is an integer as well and so are $p_1^2 = Nq_1^2 + m_1$ and p_1 too as a consequence!) Moreover, $|m_0m_1| = N - x_0^2 < 2|m_0|\sqrt{N}$, by (3) and therefore we have

$$|m_1| < 2\sqrt{N}. \quad (5)$$

Starting with the triple $(p_0\ q_0; m_0)$, the passage to $(p_1, q_1; m_1)$ as above gives an inductive construction of the triples $(p_i, q_i; m_i)$ as follows.

It is convenient to take $q_0 = 1$ and $P_0 := [\sqrt{N}]$, the largest integer not exceeding the (positive) square root $[\sqrt{N}]$ of N, so that $0 < \sqrt{N} - p_0 < 1$. The congruence condition on x_0 now looks simpler, viz. $x_0 \equiv -p_0 \pmod{m_0}$ with $m_0 := p_0^2 - N < 0$, i.e. x_0 is any integer such that $x_0 + p_0$ is divisible by m_0 (but subject to the additional conditions $x_0 < \sqrt{N} < x_0 + |m_0|$, of course!). Due to the special choice of q_0, the kuttaka does not need to be invoked here (for solving a congruence for x_0) but also at every subsequent step under the induction, as nicely emphasized by Weil. We reproduce his comments in this regard verbatim[1]. 'Strangely enough, this does not seem to have been noticed by any of our Indian authors (nor even by their later commentators, down to the sixteenth century); they make no mention of it, and invariably refer to the kuttaka for the choice of x, even though their abundant numerical evidence could easily have convinced them that this was unnecessary.'

Let us assume the triples $(p_j, q_j; m_j)$, x_j for $0 \le j \le i$ constructed inductively with $|m_j| < 2\sqrt{N}$, $x_j \equiv -x_{j-1} \pmod{m_j}$ i.e. $x_j + x_{j-1}$ divisible by m_j, $x_j < \sqrt{N} < x_j + |m_j|$, $m_{-1} := 1$, $x_{-1} := p_0$ and $x_{j-1}^2 - N = m_{j-1}m_j$ for $j \ge 0$. Then composition of $(p_i, q_i; m_i)$ with the triple $(x_i, 1; x_i^2 - N)$ leads to the definition of $p_{i+1}, q_{i+1}, m_{i+1}$, viz. $p_{i+1} := (p_ix_i + Nq_i)/m_i$, $q_{i+1} := (q_ix_i + p_i)/m_i$ and $m_{i+1} := (x_i^2 - N)/m_i$. The congruence condition on x_i above coupled with the relation $-q_ix_{i-1} + p_i = -x_{i-1}\ (q_{i-1}x_{i-1} + p_{i-1})/m_{i-1} + (p_{i-1}x_{i-1} + Nq_{i-1})/m_{i-1} = q_{i-1}\ (N - x_{i-1}^2)/m_{i-1} = -q_{i-1}m_i$ ensures that q_{i+1} is an integer. Since m_{i+1} is an integer by the same congruence condition on x_i, p_{i+1} is an integer too. The same kind of argument applied to derive the bound (5) leads to $|m_{i+1}| < 2\sqrt{N}$ and further ensures that $(-1)^i m_{i-1}$, $\sqrt{N} \pm x_i$ are all positive (and so $N - x^2 > 0$ for all i). We have thus, on hand, an infinite sequence of integers (the 'additives') $m_0, m_1, m_2, \ldots$ bounded by $2\sqrt{N}$ in absolute value. Hence infinitely many among them must coincide by Dirichlet's box principle. But, as we will see presently, much more is true, namely, (i) there exists an integer s such that $m_{j+s} = m_j$ for $j \ge 1$ and (ii) $m_j = 1$ for an integer j, leading to a solution of equation (2).

In other words, the m_i repeat themselves in a periodical fashion (actually, corresponding to the periodicity of the infinite simple continued fraction expansion for the quadratic irrationality $\sqrt{N}$).

Before moving on to indicate proofs for assertions (i) and (ii) above, it will be quite in order to quote some interesting observations by Weil (see ref. 1 pp. 23, 24, 94–97, 230–232) on the construction of $(p_i, q_i; m_i)$. 'The Indian prescription' for the choice of x_i within its congruence class modulo m_i is not quite the one described above, since their rule is 'to make $N - x_i^2$ "small" (i.e. in actual practice, as small as possible), but as the context shows, in absolute value' or in other words, to replace x_i by $y_i := x_i + |m_i|$ if $y_i^2 - N$ turns out to be less than $N - x_i^2$. 'It can be shown that this has merely the effect of abbreviating the procedure somewhat when that is the case,' but though 'numerically useful', can 'make the theoretical discussion much more cumbersome.' Moreover, the above rigorous treatment for constructing (p_i, q_i, m_i) 'may have been known to the Indians only experimentally; there is nothing to indicate whether they had proofs for them, or even for part of them'. 'In order to carry out the cakravāla', a 'starting point' which 'invariably they choose' is the triple $(p_0, 1; m_0)$ 'for which p_0^2 is the closest square to N, above or below.' Finally we are told to iterate the process only till we find an "additive" m with one of the values ±1, ±2, ±4 and then to make use of the bhāvanā, i.e. Brahmagupta's procedure for that case. Actually this is no more than a shortcut, since it can be shown that the cakravāla applied in a straightforward manner, would inevitably lead to a triple $(p, q; 1)$ as desired; while this shortcut is quite effective from the point of view of the numerical solution, it destroys the 'cyclic' character of the method, which otherwise would appear from the fact that the additives ··· would repeat themselves periodically corresponding to the periodicity of the continued fraction of $\sqrt{N}$.

'For the Indians, of course, the effectiveness of "cakravāla" could be no more than an experimental fact, based on their treatment of a great many special cases, some of them of considerable complexity and involving (to their delight, no doubt) quite large numbers. Fermat was the first to perceive the need for a general proof and Lagrange the first to publish one. Nevertheless, to have developed the cakravāla and to have applied it successfully to such difficult cases as $N = 61$ or $N = 67$ had been no mean achievement.'

We now go on to the promised proofs for assertions (i) and (ii) above. The triples $(p_i, q_i; m_i)$ and x_i for $i \geq 0$ as constructed above enable us to obtain an infinite simple continued fraction for $\sqrt{N}$. Let $\xi_0 := \sqrt{N} + p_0$, $a_0 := 2p_0$, $a_i := (-1)^i(x_{i-1} + x_{i-2})/m_{i-1}$ for $i \geq 1$ and η_j be defined by $\sqrt{N} = x_j + \eta_j$ for $j \geq 0$. Then, in view of the relation $x_{j-1}^2 - N = m_{j-1}m_j$ for $j \geq 0$, we can derive the following:

$$\xi_0 = a_0 + \frac{1}{\xi_1} \text{ with}$$

$$\xi_1 := \frac{1}{\sqrt{N} - p_0} = \frac{\sqrt{N} + p_0}{-m_0}$$

$$= \frac{x_0 + \eta_0 + p_0}{-m_0} = \frac{x_0 + x_{-1}}{-m_0} + \frac{\eta_0}{-m_0}$$

$$\xi_1 = a_1 + \frac{1}{\xi_2} \text{ with } \xi_2 = -\frac{m_0}{\sqrt{N} - x_0}$$

$$= -\frac{m_0(\sqrt{N} + x_0)}{-m_0 m_1} = \frac{x_1 + x_0 + \eta_1}{m_1}$$

$$\xi_2 = a_2 + \frac{1}{\xi_3},\ \xi_3 = \frac{m_1}{\sqrt{N} - x_1} = \frac{\sqrt{N} + x_1}{-m_2}$$

$$= \frac{x_2 + \eta_2 + x_1}{-m_2} = \frac{x_1 + x_2}{-m_2} + \frac{\eta_2}{-m_2}$$

$$\ldots$$

$$\xi_i = \frac{\sqrt{N} + x_{i-2}}{(-1)^i m_{i-1}} = a_i + \frac{1}{\xi_{i+1}} \ldots\ldots$$

$$\ldots \tag{7}$$

Using terminology from the theory of continued fractions[2], we see that $\xi_0 = \sqrt{N} + [\sqrt{N}]$ has the infinite continued fraction expansion

$$\xi_0 = \langle a_0, a_1, a_2, \cdots \rangle,$$

with the natural numbers $a_0, a_1, a_2, \cdots$ occurring as partial quotients.

There exists, as we shall see now, a minimal positive integer r such that $a_{i+r} = a_i$ for all i from a certain index l, say. Now $\xi_i = (\sqrt{N} + x_{i-2})/(-1)^i m_{i-1}$ as derived inductively in (7) and coupled with the bound $|m_i| < 2\sqrt{N}$ for all $i \geq 0$, the relation $x_i^2 = N + m_i m_{i+1}$ leads to the bound $|x_i|^2 < 5N$ for all $i \geq 0$. The number of distinct pairs $(x_{i-1}, (-1)^{i-1}m_i)$ of integers for $i = 1,2,3,\ldots$ subject to such fixed bounds (depending only on the given non-square natural number N) can only be finite. Thus there exist indices l and $k > l$ such that

$$(x_{l-1}, (-1)^{l-1}m_l) = (x_{k-1}, (-1)^{k-1}m_k), \text{ i.e.}$$

$$\xi_l = \langle a_l, a_{l+1}, \ldots, a_{k-1}, \xi_k \rangle$$

$$= \xi_k = \langle a_k, a_{k+1}, \ldots \rangle, \tag{8}$$

and therefore

$$\xi_0 = \langle a_0, a_1, \ldots, a_{l-1}, a_l, a_{l+1}, \ldots, a_{k-1}, \xi_k \rangle$$

$$= \langle a_0, a_1, \ldots, a_{l-1}, a_l, a_{l+1}, \ldots a_{k-1}, a_l, a_{l+1}, \ldots, a_{k-1} \ldots \rangle$$

$$= \langle a_0, a_1, \ldots, a_{l-1}, \overline{a_l, a_{l+1}, \ldots, a_{k-1}} \rangle$$

is a periodic continued fraction with period $r := k - l$, proving that $a_{i+r} = a_i$ for all $i \geq l$. Clearly r can be chosen to be minimal.

In the case of $\xi_0 = \sqrt{N} + [\sqrt{N}]$, we can even show that $l = 0$, i.e.

$$\xi_0 = \langle \overline{a_0, a_1, a_2, \ldots, a_{r-1}} \rangle \tag{9}$$

and so is 'purely periodic'. For this purpose, we note first that while $\xi_0 = p_0 + \sqrt{N} > 1$, its 'conjugate' $\xi_0' := p_0 - \sqrt{N}$ satisfies the inequalities $-1 < \xi_0' < 0$. Denoting for $\xi_i = (\sqrt{N} + x_{i-2})/(-1)^i m_{i-1}$ its 'conjugate' $(-\sqrt{N} + x_{i-2})/(-1)^i m_{i-1}$ by ξ_i', we have

$$\xi_i = a_i + 1/\xi_{i+1},\quad \xi_i' = a_i + 1/\xi_{i+1}'. \tag{10}$$

Now $a_i \geq 1$ for every i, $(-1)^i m_{i-1}$, $\sqrt{N} - x_{i-2}$ are all positive (by construction) and so, $\xi_i' < 0$, leading to $1/\xi_{i+1}' = \xi_i' - a_i < -1$, i.e. $-1 < \xi_{i+1}' < 0$ for every $i \geq -1$. But then, using (10), we have $0 < (-1/\xi_{i+1}') - a_i < 1$ so that a_i is just the largest integer $[-1/\xi_{i+1}']$ not exceeding the positive real number $-1/\xi_{i+1}'$. Since, for $k > l$, $\xi_k = \xi_l$ by (8), the 'conjugates' ξ_k', ξ_l' coincide and so $a_{k-1} = a_{l-1}$. Thus from $\xi_k = \xi_l$ we conclude that $\xi_{l-1} = a_{l-1} + 1/\xi_l = a_{k-1} + 1/\xi_k = \xi_{k-1}$. Iteration yields $\xi_r = \xi_{k-l} = \xi_0$, proving (9).

Now for any $j \geq 1$, ξ_{jr} has the same continued fraction expansion

$$\langle a_0, a_1, a_2, \ldots, a_{r-1}, a_0, a_1, a_2, \ldots, a_{r-1}, \ldots \rangle$$
$$= \langle \overline{a_0, a_1, a_2, \ldots, a_{r-1}} \rangle$$

as ξ_0. Hence

$$\frac{\sqrt{N} + x_{jr-2}}{(-1)^{jr} m_{jr-1}} = \xi_{jr} = \xi_0 = \sqrt{N} + p_0,$$

$\sqrt{N}(1 - (-1)^{jr}m_{jr-1}) = (-1)^{jr}m_{jr-1}p_0 - x_{jr-2}$. But if $c + d\sqrt{N} = 0$ for integers c, d, then necessarily $c = d = 0$. Hence, in particular,

$$(-1)^{jr}m_{jr-1} = 1.$$

Whenever jr is even (e.g. $j = 2$ for odd r and $j = 1$ for even integers r), we have $m_{jr-1} = 1$ and the triple $(p_{jr-1}; q_{jr-1}; 1)$ gives a solution of the diophantine equation (2) proving assertion (ii). Assertion (i) also follows from above easily.

An example of how unsparing a critical analysis can tend to be (when intended or called for) may be found in Weil's remarks on Euler's contribution to the topic of 'Pell's equation' and the related continued fraction algorithm, on pages 232–233 (ref. 1): 'while Euler drew attention' to the periodicity and 'palindromic' property of the partial quotients a_i in the 'continued fractions for square roots $\sqrt{N}$, as well as to their use in solving Pell's equation, there is no sign that he (Euler) ever sought to back up his findings by anything more than experimental evidence. He (Euler) did mention that the values obtained by his process for the integers B_i, A_i, m_i are necessarily bounded...; from this he (Euler) could at least have derived the conclusion that the sequence (m_i) is periodic from a certain point onwards, but he failed to mention this, or did not bother to do so'. When in his later years (after Lagrange gave a 'definitive treatment' of the subject, based on the continued fraction algorithm...), 'Euler came back to the topic of Pell's equation, he added nothing of substance to what by that time was already public knowledge on that subject'.

Fermat must have been in the dark about the contribution of the Indian mathematicians to the solution of (2) and also possibly about Archimedes' *Problema bovinum.* He offered (in 1657) the problem of solving equation (2) (in integers, of course!) as a challenge to the English mathematicians and all others. In a personal letter to Huygens, a few months later, commenting on the solution by Wallis and Brouncker, he observed that 'the English had failed to give a general proof'; such a ('general') proof, according to Fermat, could only be 'obtained by descent'. But perhaps 'Fermat's method of solution (for (2)) did not greatly differ from the one he got from Wallis and Brouncker' and 'he had been able to extract from it a formal proof of the fact that it always leads to a solution'. The method 'which Wallis credits to Brouncker' is 'equivalent to the Indian cakravāla method as well as to the modern treatments based on continued fraction'.

1. Weil, André, *Number Theory – An Approach Through History – From Hammurapi to Legendre*, Birkhäuser, Boston, 1983.
2. Niven, I. and Zuckerman, H. S., *An Introduction to the Theory of Numbers*, Wiley Eastern, New Delhi, 1976.

S. Raghavan is with the SPIC Mathematical Institute, East Coast Chambers, IV Floor, 92, G. N. Chetty Road, T. Nagar, Madras 600 017, India.

Addendum

Fifty years of the exact solution of the two-dimensional Ising model by Onsager

Somendra M. Bhattacharjee and Avinash Khare

[*Curr. Sci.*, 1995, **69**, 816–821]

We would like to make the following corrections/additions which we learnt after the paper was written.

W. Lenz was working in Rostock University (not Rostalk as mentioned in the text).

E. Ising was a teacher (and later became the headmaster) of a Jewish school in Germany from 1934 to November 1938. When his school was damaged, he managed to leave Germany. It was as late as 1949 that he realized that his name had become famous.

It seems that Ising also agrees that the model should be named the Lenz–Ising model. The name 'Ising model' became popular following the title of Peierls' paper[12].

It appears that Heisenberg in his 1928 paper[8] also thought that at least eight nearest neighbours are needed for a phase transition. We are not sure whether this comment by Heisenberg refers to the Ising model or to the new model he proposed in that paper.

We thank Prof. Sigismund Kobe of University of Dresden for several clarification.

Edited and published by Prof. P. Balaram and Prof. S. Ramaseshan, Current Science Association, Bangalore 560 080.
Typeset by WINTECS TYPESETTERS (Ph: 3327311), Bangalore 560 021, Printed at Printek Printers, Bangalore (Ph: 3357763).