One diophantine equation

Benjamin M.M. de Weger
Econometric Institute, Erasmus University Rotterdam,
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.
(e-mail: dweger@wis.few.eur.nl)

December 1, 1995

Abstract

We solve a problem posed recently by J.H.E. Cohn, in proving that $x = y = 1$ is the only solution in nonnegative integers to the diophantine equation $x^2 - 3y^4 = -2$.

1 Introduction

Recently, J.H.E. Cohn [C2] asked to prove that the only solution in nonnegative integers to the diophantine equation

$$x^2 - 3y^4 = -2 \tag{1}$$

is given by $x = y = 1$. In this note we provide a proof. As Cohn showed, this implies that the diophantine equation $x^2 - kxy^2 + y^4 = -2$ (equation 4 in [C2]) has only two solutions in nonnegative integers, namely $x = 1$, $y = 1$, $k = 4$, and $x = 3$, $y = 1$, $k = 4$.

In 1966, Cohn [C1] also asked to prove that the only solutions in nonnegative integers to the diophantine equation

$$2x^2 = 3y^4 - 1 \tag{2}$$

are given by $x = y = 1$ and $x = 11, y = 3$. This was done already in 1967, by R. Bumby [B], and again in 1995 by N. Tzanakis [T].

The two problems of solving equations (1) and (2) look like each other, and indeed have a relation, although somewhat superficial, as we shall see below.

Unfortunately our proof is far from elementary, as Cohn expected, since it is based on the Gelfond-Baker theory of linear forms in logarithms, and computer calculations. However, these methods do allow a 'proof by routine', as this note intends to illustrate. Indeed, the method we use below to solve equation (1) can be used to give a third solution (after those by Bumby and Tzanakis) of equation (2). We did work out such a proof, for both equations (1) and (2), but we give details here only for equation (1). We note that the method of elliptic logarithms used by Tzanakis [T] for solving equation (2) should also work for solving equation (1), and it would not surprise us at all when Bumby's more elementary but rather complicated method [B] for solving equation (2) would also work for equation (1).

As a side remark we note that equation (1) is equivalent to the problem of determining the perfect squares in the binary recurrence sequence $1, 3, 11, 41, 153, \ldots$, given by $a_0 = 1$, $a_1 = 3$, and $a_{n+1} = 4a_n - a_{n-1}$ for all $n \geq 1$. Similarly, equation (2) is equivalent to the problem of determining the perfect squares in the binary recurrence sequence $1, 9, 89, 891, 8721, \ldots$, given by $a_0 = 1$, $a_1 = 9$, and $a_{n+1} = 10a_n - a_{n-1}$ for all $n \geq 1$. Cohn's elementary methods, given in [C1] and several other papers by him, for determining squares in recurrence sequences satisfying relations of the type $a_{n+1} = ka_n - a_{n-1}$ or $a_{n+1} = ka_n + a_{n-1}$, seem to work only when k is odd.

2 Deriving Thue equations

First note that both equations (1) and (2) can be brought together in the equation

$$U^2 + 2V^2 = 3W^4. \tag{3}$$

Namely, equation (1) is equation (3) with $U = x$, $V = 1$, $W = y$, and equation (2) is equation (3) with $U = 1$, $V = x$, $W = y$.

We factor both sides of equation (3) over the ring $\mathbb{Z}(\sqrt{-2})$. Let $\alpha \in \mathbb{Z}(\sqrt{-2})$ be the fourth-power-free part of $U + V\sqrt{-2}$, and write

$$U + V\sqrt{-2} = \alpha(E + F\sqrt{-2})^4 \tag{4}$$

for $E, F \in \mathbb{Z}$. Then (3), (4) and the definition of α imply that $N(\alpha)$ equals 3 times a fourth power. Let's use this in finding the primes in $\mathbb{Z}(\sqrt{-2})$ that may divide α.

The primes lying above 3 are $\pi = 1 + \sqrt{-2}$ and its complex conjugate $\overline{\pi}$. Put $k = \mathrm{ord}_\pi(\alpha)$, $\ell = \mathrm{ord}_{\overline{\pi}}(\alpha)$. Then by the definition of α we have $k, \ell \in \{0, 1, 2, 3\}$, but on the other hand $k + \ell = \mathrm{ord}_3(N(\alpha)) \equiv 1 \pmod 4$, so we find that $(k, \ell) = (1, 0)$, $(0, 1)$, $(3, 2)$, or $(2, 3)$. Without loss of generality we may assume that $k > \ell$ (otherwise we replace (4) by its complex conjugate, and change some signs that are irrelevant anyway). This implies that the contribution of π and $\overline{\pi}$ to α is either $\pi = 1 + \sqrt{-2}$ or $\pi^3\overline{\pi}^2 = 3^2(1 + \sqrt{-2})$.

Next, let π be a prime lying above a rational prime $p \neq 3$ that ramifies or splits in $\mathbb{Z}(\sqrt{-2})$. Then $N\pi = p$, and on the one hand $\mathrm{ord}_\pi(\alpha) \in \{0, 1, 2, 3\}$ by the definition of α, but on the other hand $\mathrm{ord}_\pi(\alpha) = \mathrm{ord}_p(N(\alpha))$, so is divisible by 4. Hence $\pi \not| \, \alpha$.

Finally, let p be a rational prime that remains prime in $\mathbb{Z}(\sqrt{-2})$. Then $N(p) = p^2$, and on the one hand $\mathrm{ord}_p(\alpha) \in \{0, 1, 2, 3\}$ by the definition of α, but on the other hand $2\mathrm{ord}_p(\alpha) = \mathrm{ord}_p(N(\alpha))$ is divisible by 4. Hence $p \not| \, \alpha$, or $\mathrm{ord}_p(\alpha) = 2$.

It follows that there is a $g \in \mathbb{Z}$, squarefree, and having as possible prime divisors only 3 and the primes that remain prime in $\mathbb{Z}(\sqrt{-2})$, such that $\alpha = \pm g^2(1 + \sqrt{-2})$. Now we expand the right-hand side of (4) and take real and imaginary parts, and also we take the norm in (4). Thus we obtain the following parametrization for the solutions of (3):

$$\left\{ \begin{array}{rcl} U &=& \pm g^2(E^4 - 8E^3F - 12E^2F^2 + 16EF^3 + 4F^4) \\ V &=& \pm g^2(E^4 + 4E^3F - 12E^2F^2 - 8EF^3 + 4F^4) \\ W &=& g(E^2 + 2F^2) \end{array} \right. . \tag{5}$$

Now we return to Cohn's original problems. In the case of equation (1) we have $V = 1$. It follows at once from the second equation of (5) that $g^2 = 1$, and looking at this equation modulo 4 we see that $\pm = +$. Hence the second equation of (5) now becomes the quartic Thue equation

$$E^4 + 4E^3F - 12E^2F^2 - 8EF^3 + 4F^4 = 1. \tag{6}$$

And in the case of equation (2) we have $U = 1$, and in a similar way the first equation of (5) now becomes the quartic Thue equation

$$E^4 - 8E^3F - 12E^2F^2 + 16EF^3 + 4F^4 = 1. \tag{7}$$

Below we will show that equation (6) admits only the solutions $(E, F) = (1, 0), (-1, 0)$. This provides a complete answer to Cohn's question on equation (1). We also solved equation (7) by the same method as we use below, and thus showed that it admits only the solutions $(E, F) = (1, 0), (-1, 0), (1, 1), (-1, -1)$. This provides a new proof of the result of Bumby [B] and Tzanakis [T] on equation (2). We will give details of our proofs only for equation (6). The reader is invited to fill in the details for equation (7) himself.

We could make it easy for ourselves, and use the software package KANT, which contains a program that solves Thue equations in an automated way. This program works essentially in the same way as our arguments below, and on referring to the program we could at this point end the paper in a few lines. We chose not to do so for two reasons. The first reason is didactical: we want to give the reader insight in what is going on in our proof, and we feel that a program that works as a black box does not give enough of such insight. The second reason is more mathematical: we will see below that in our specific situation we can make improvements on the general method, so that we obtain smaller upper bounds and a faster reduction procedure. However, we encourage those readers who are only interested in the truth of results and not in their proofs to stop reading here and invoke the KANT program. Note that KANT is available by anonymous ftp from `ftp.math.tu-berlin.de`.

3 A quartic field

So from now on we concentrate on equation (6). We need some information on the quartic field associated to the binary form in equation (6). We computed the data below by Pari-1.39.03 on a personal computer.

Let θ be a root of $t^4 - 6t^2 + 6$. Then $\psi = -4 - 4\theta + \theta^2 + \theta^3$ is a root of $t^4 + 4t^3 - 12t^2 - 8t + 4$. The quartic field $\mathbb{K} = \mathbb{Q}(\theta)$ is totally real, has discriminant $13824 = 2^9\, 3^3$, integral basis $\{1, \theta, \theta^2, \theta^3\}$, trivial class group, Galois group $\mathbb{D}_4$, and a system of fundamental units is

$$\epsilon = 1 + \theta, \quad \sigma(\epsilon) = 1 - \theta, \quad \eta = 5 - 2\theta - 2\theta^2,$$

all of norm equal to 1. The symbol σ denotes an automorphism of $\mathbb{K}$ sending θ to $-\theta$. We denote conjugates by lower indices, numbered as follows:

$$\theta_1 = \sqrt{3 + \sqrt{3}}, \quad \theta_2 = \sqrt{3 - \sqrt{3}}, \quad \theta_3 = -\theta_1, \quad \theta_4 = -\theta_2.$$

4 Deriving linear forms in logarithms

Equation (6) leads to

$$E - F\psi = \pm\epsilon^a\sigma(\epsilon)^b\eta^c \tag{8}$$

for some $a, b, c \in \mathbb{Z}$, and $N(E - F\psi) = 1$. Since the differences of the four numbers $E - F\psi_j$ are of the size of $|F|$, these four numbers are far apart when $|F|$ is large. Since their product equals 1, it must be the case that three of them are also of the size of $|F|$, and one of them is extremely small, of the size of $|F|^{-3}$. Let $i \in \{1, 2, 3, 4\}$, depending on E, F, thus not known in advance, be such that $|E - F\psi_i| = \min_{j=1,2,3,4} |E - F\psi_j|$. Then this is the extremely small one.

We consider three conjugates of equation (8), with indices i, j, k, where j, k are taken as follows:

$$\begin{array}{lll} j = 2, k = 4 & \text{if} & i = 1 \text{ or } i = 3, \\ j = 1, k = 3 & \text{if} & i = 2 \text{ or } i = 4. \end{array}$$

Although we could have made other choices, we will see below that this particular choice turns out to be convenient.

Siegel's identity now reads

$$(\psi_i - \psi_j)(E - F\psi_k) + (\psi_j - \psi_k)(E - F\psi_i) + (\psi_k - \psi_i)(E - F\psi_j) = 0,$$

and together with (8) this leads to

$$\frac{\psi_i - \psi_j}{\psi_i - \psi_k}\left(\frac{\epsilon_k}{\epsilon_j}\right)^a\left(\frac{\sigma(\epsilon)_k}{\sigma(\epsilon)_j}\right)^b\left(\frac{\eta_k}{\eta_j}\right)^c - 1 = \frac{\psi_j - \psi_k}{\psi_k - \psi_i}\,\frac{E - F\psi_i}{E - F\psi_j}. \tag{9}$$

In this formula the right-hand side is very small, in fact of the size of $|F|^{-4}$, by the definition of i.

Put

$$\Lambda_i = \log\left|\frac{\psi_i - \psi_j}{\psi_i - \psi_k}\right| + a\log\left|\frac{\epsilon_k}{\epsilon_j}\right| + b\log\left|\frac{\sigma(\epsilon)_k}{\sigma(\epsilon)_j}\right| + c\log\left|\frac{\eta_k}{\eta_j}\right|,$$

so that the left hand side of (9) equals $e^{\pm\Lambda_i} - 1$. Then also $|\Lambda_i|$ is extremely small, of the size of $|F|^{-4}$.

Put $A = \max\{|a|, |b|, |c|\}$. Equation (8) suggests that $|F|$ is of the size of $\exp(A)$, so that in fact $|\Lambda_i|$ is exponentially small in terms of A. Indeed, following the details of [TW], one finds, under the condition $|F| \geq 3$, that

$$|\Lambda_i| < 28692.1\exp(-3.01818A). \tag{10}$$

Heuristically speaking, this is extraordinary for a linear form with coefficients of the size of A, which generically will be only polynomially small in terms of A. We will exploit this heuristic argument below.

5 A large upper bound

In general one now has to proceed with a four-term linear form Λ_i. The number of terms in the linear form is very important for the efficience of the method. The upper bound to be derived for A grow more than exponentially in terms of this number, and also the complexity of the reduction procedure depends very much on it. Although with the present state of the art a four-term linear form is very well doable in general, it's to the author's taste a matter of elegance to use a method as efficient as possible, by using properties special to the particular case, taking the shortcuts this suggests, rather than to follow blindly the longer route set out by the general theory.

We note that in our specific situations the linear forms Λ_i can be simplified considerably. Namely, due to our special choice of j, k, we have $\sigma(\epsilon)_j = \epsilon_k$ and $\sigma(\epsilon)_k = \epsilon_j$, so that

$$\frac{\sigma(\epsilon)_k}{\sigma(\epsilon)_j} = \left(\frac{\epsilon_k}{\epsilon_j}\right)^{-1}.$$

It also turns out that

$$-\left(\frac{\psi_1-\psi_2}{\psi_1-\psi_4}\right)^2 = -\left(\frac{\psi_3-\psi_2}{\psi_3-\psi_4}\right)^2 = \frac{\epsilon_2}{\epsilon_4}, \quad -\left(\frac{\psi_2-\psi_1}{\psi_2-\psi_3}\right)^2 = -\left(\frac{\psi_4-\psi_1}{\psi_4-\psi_3}\right)^2 = \frac{\epsilon_1}{\epsilon_3}.$$

(We note that in solving equation (7) by the same method, we find exactly the same phenomena occurring. It's not clear to what extent this is coincidental.) It means that

$$\Lambda_i = \left(a - b - \frac{1}{2}\right)\log\left|\frac{\epsilon_k}{\epsilon_j}\right| + c\log\left|\frac{\eta_k}{\eta_j}\right|.$$

Not only do these linear forms have only two terms, also the number of linear forms itself has been halved, as we observe that $\Lambda_1 = \Lambda_3$ and $\Lambda_2 = \Lambda_4$.

Put $d = 2a - 2b - 1 \in \mathbb{Z}$. It now is useful to consider the linear forms

$$2\Lambda_i = d\log\left|\frac{\epsilon_k}{\epsilon_j}\right| + 2c\log\left|\frac{\eta_k}{\eta_j}\right|. \tag{11}$$

To these forms we apply a result from the Gelfond-Baker (transcendence) theory of linear form in logarithms of algebraic numbers, stating that such a linear form cannot be too small. Put $B = \max\{|d|, |2c|\}$, then $B \leq 4A + 1$. From the result of Baker and Wüstholz [BW] we find, under the condition $B \geq 3$, that

$$|2\Lambda_i| > \exp\left(-1.24112 \times 10^{12}\log B\right) \tag{12}$$

(note that the algebraic numbers inside the logarithms lie in the field $\mathbb{K}$ itself). Combining (10), (11) and (12) we immediately obtain the absolute upper bound

$$A < 1.29867 \times 10^{13}. \tag{13}$$

6 Reduction of the upper bound

An efficient search below the upper bound (13) goes as follows, following again [TW]. The basic idea is to use the heuristic argument mentioned above, that the linear forms Λ_i should behave generically, i.e. should be not essentially smaller than Dirichlet's diophantine approximation theorem suggests, i.e. polynomial in the maximum of the coefficients, of degree one less than the number of terms of the linear form. As we now have an upper bound on these coefficients, namely (13), we have a good idea how small the linear forms can be, namely of the size of A^{-1}, which is at worst $\approx 10^{-13}$. If we can check that indeed for all coefficients below these bounds the linear forms are larger than $\approx 10^{-13}$, we can use (10) to obtain a reduced upper bound for A.

In practice it works as follows. Fix a linear form Λ_i (we only have to look at $i = 1$ and $i = 2$). Define $\lambda \in \mathbb{Z}$ by

$$\begin{pmatrix} 1 & 0 \\ [10^{30} \log |\epsilon_k/\epsilon_j|] & [10^{30} \log |\eta_k/\eta_j|] \end{pmatrix} \begin{pmatrix} d \\ 2c \end{pmatrix} = \begin{pmatrix} d \\ \lambda \end{pmatrix},$$

where $[\cdot]$ denotes rounding to an integer. This formula says that the point $(d, \lambda)^\top$ is in the lattice spanned by the columns of the above matrix, and because d is odd this point is not the origin. Here the constant 10^{30} is chosen conveniently, i.e. somewhat larger than the square of the upper bound (13) for A.

By a variant of the euclidean algorithm we can find the nonzero lattice point closest to the origin, and this gives us a lower bound for the length of the vector $(d, \lambda)^\top$. For details, see the Appendix at the end of this note. These lower bounds turned out to be $> 6.09448 \times 10^{14}$. Hence

$$\left(6.09448 \times 10^{14}\right)^2 \leq d^2 + \lambda^2 < (4A+1)^2 + \lambda^2,$$

and now the upper bound (13) for A yields

$$|\lambda| > 6.09309 \times 10^{14}.$$

By the definition of λ and (11) we see that

$$|\lambda - 2 \times 10^{30} \Lambda_i| \leq 6A + 1 < 7.79203 \times 10^{13},$$

hence

$$|\Lambda_i| > 2.65694 \times 10^{-16},$$

as expected (if one accepts that 10^{-16} is of the same size as 10^{-13}). This combined with (10) immediately gives the reduced upper bound $A \leq 15$.

It is good to realise again what really is going on here, as this reduction of the upper bound from 1.29867×10^{13} to only 15 is dramatic and seems mysterious. In fact, we have shown above that the existence of a solution of (10) with $16 \leq A < 1.29867 \times 10^{13}$ implies the existence of a nonzero lattice point in the given lattice with a distance to the origin less than or equal to 6.09448×10^{14}. By inspection, after having performed the euclidean algorithm, we know that such lattice points do not exist, hence that solutions in the given range do not exist.

7 Conclusion

It remains to find the solutions with $A \leq 15$, and those with $B \leq 2$ or $|F| \leq 2$. This can be done easily as follows.

When $B \leq 2$ we do not immediately have a bound on A. But there are only the possibilities ± 1 for d, and $0, \pm 1$ for c. In these few cases we compute Λ_i, and always find $|\Lambda_i| > 0.138259$. Then (10) yields $A \leq 4$ or $|F| \leq 2$.

The solutions with $|F| \leq 2$ can be found directly from equation (6) itself. We only found $(E, F) = \pm(1, 0)$.

To find the solutions with $A \leq 15$, we checked inequality (10) for the triples (a, b, c) with $6 \leq A \leq 15$, and for the triples (a, b, c) with $A \leq 5$ we wrote out both sides of (8) in terms of the integral basis of $\mathbb{K}$, and solved the linear system of four equations in two unknowns that resulted from comparing coefficients. We found only the solution $(a, b, c) = (0, 0, 0)$, leading to the solutions $(E, F) = \pm(1, 0)$, already found above. Hence this concludes our proof.

Appendix

In this appendix we give some numerical data. All computations were perfomed on a Pentium personal computer, and took in total less than one second CPU-time.

In the case $i = 1$ the lattice is defined by the numbers

$$\begin{aligned} [10^{30} \log |\epsilon_4/\epsilon_2|] &= -2\,82547\,30309\,52546\,31946\,91734\,96988 \\ [10^{30} \log |\eta_4/\eta_2|] &= 3\,10199\,26241\,35462\,31106\,09658\,79829 \end{aligned}$$

and a reduced basis of the lattice, computed by a variant of the euclidean algorithm, is

$$\left\{ \begin{pmatrix} 34617\,77635\,07452 \\ 1\,43902\,88628\,32293 \end{pmatrix}, \begin{pmatrix} -1\,77977\,14126\,72849 \\ 1\,56234\,81605\,96936 \end{pmatrix} \right\}.$$

The first vector in this basis points to the nonzero lattice point nearest to the origin. The length of this vector is $> 1.48007 \times 10^{15}$.

In the case $i = 2$ the lattice is defined by the numbers

$$\begin{aligned} [10^{30} \log |\epsilon_3/\epsilon_1|] &= -99386\,38120\,13933\,68031\,84316\,58312 \\ [10^{30} \log |\eta_3/\eta_1|] &= -4\,35285\,45224\,79817\,28086\,80421\,78698 \end{aligned}$$

and a reduced basis of the lattice, computed by a variant of the euclidean algorithm, is

$$\left\{ \begin{pmatrix} 59868\,73258\,57508 \\ 11402\,02692\,54778 \end{pmatrix}, \begin{pmatrix} 6\,11197\,23975\,82793 \\ -6\,10663\,63460\,97968 \end{pmatrix} \right\}.$$

The first vector in this basis points to the nonzero lattice point nearest to the origin. The length of this vector is $> 6.09448 \times 10^{14}$.

References

[BW] A. Baker and G. Wüstholz, "Logarithmic forms and group varieties", *J. reine angew. Math.* **442** [1993], 19–62.

[B] R. Bumby, "The diophantine equation $3x^4 - 2y^2 = 1$", *Math. Scand.* **21** [1967], 144–148.

[C1] J.H.E. Cohn, "Eight diophantine equations", *Proc. London Math. Soc. (3)* **16** [1966], 153–166.

[C2] J.H.E. Cohn, "Twelve diophantine equations", *Arch. Math.* **65** [1995], 130–133.

[T] N. Tzanakis, "Solving elliptic diophantine equations by estimating linear forms in elliptic logarithms. The case of quartic equations", *Acta Arith.* [1995], to appear.

[TW] N. Tzanakis and B.M.M. de Weger, "On the practical solution of the Thue equation", *J. Number Th.* **31** [1989], 99–132.