MATHEMATICS OF COMPUTATION
Volume 71, Number 239, Pages 1281–1286
S 0025-5718(01)01381-3
Article electronically published on September 17, 2001

THUE'S THEOREM AND THE DIOPHANTINE EQUATION $x^2 - Dy^2 = \pm N$

KEITH MATTHEWS

Abstract. A constructive version of a theorem of Thue is used to provide representations of certain integers as $x^2 - Dy^2$, where $D = 2, 3, 5, 6, 7$.

1. Introduction

The idea of using Euclid's algorithm to construct solutions of $p = x^2 + y^2$ goes back to Serret [9] and Hermite [5]. (Also see Wagon [12] and Brillhart [1].) The method easily extends to $p = x^2 + ny^2$, $n = 2, 3, 5$. (See Wilker [13] for $n = 5$.)

Cornacchia [2, pp. 61–66] generalised the method to $N = x^2 + ny^2$, $n \geq 1$ and discussed the case $n < 0$ [2, pp. 66–70]. (Also see Nitaj [8] and Hardy, Muskat and Williams [3], [4], Muskat [6], Williams [14], [15].)

It is not so well known that the Serret–Hermite method can be used to find explicit solutions of $x^2 - Dy^2 = N$ when $D > 1$ is small. Nagell [7, pp. 210–212] used a nonconstructive form of a theorem of Thue [10, p. 587] to deal with $D = 2$ and 3, while a variant of Thue's theorem was also used in Uspensky and Heaslet [11, pp. 352–368] for $D = 2, 3, 5$.

In this paper we show how to obtain explicit representations of certain integers in the form $x^2 - Dy^2$ for small $D > 1$, using a constructive version of Thue's theorem based on Euclid's algorithm. Amongst other things, if $u^2 \equiv D \pmod{N}$, $D \not\equiv 1 \pmod{N}$ is soluble and $\gcd(D, N) = 1, N$ odd, we show how to find the following representations:

$N = 8k \pm 1$	$N = x^2 - 2y^2$ $-N = x^2 - 2y^2$
$N = 12k + 1$	$N = x^2 - 3y^2$
$N = 12k - 1$	$-N = x^2 - 3y^2$
$N = 5k + 1$	$N = x^2 - 5y^2$
$N = 5k - 1$	$-N = x^2 - 5y^2$
$N = 24k + 1$ or $24k - 5$	$N = x^2 - 6y^2$
$N = 24k - 1$ or $24k + 5$	$-N = x^2 - 6y^2$
$N = 28k + 1, 28k + 9$ or $28k + 25$	$N = x^2 - 7y^2$
$N = 28k - 1, 28k - 9$ or $28k - 25$	$-N = x^2 - 7y^2$

Received by the editor May 5, 2000 and, in revised form, September 4, 2000.
2000 *Mathematics Subject Classification.* Primary 11D09.

2. Euclid's algorithm and Thue's theorem

Euclid's algorithm. Let a and b be natural numbers, $a > b$, where b does not divide a. Let $r_0 = a$, $r_1 = b$ and for $1 \le k \le n$, $r_{k-1} = r_k q_k + r_{k+1}$, where $0 < r_{k+1} < r_k$ and $r_n = 0$. Define sequences $s_0, s_1, \dots, s_{n+1}$ and $t_0, t_1, \dots, t_{n+1}$ by

$$s_0 = 1, s_1 = 0, t_0 = 0, t_1 = 1, t_{k-1} = t_k q_k + t_{k+1}, s_{k-1} = s_k q_k + s_{k+1},$$

for $1 \le k \le n$. Then the following are easily proved by induction:

(i) $s_k = (-1)^k |s_k|$, $t_k = (-1)^{k+1}|t_k|$;
(ii) $0 = |s_1| < |s_2| < \cdots < |s_{n+1}|$;
(iii) $1 = |t_1| < |t_2| < \cdots < |t_{n+1}|$;
(iv) $a = |t_k| r_{k-1} + |t_{k-1}| r_k$ for $1 \le k \le n+1$;
(v) $r_k = s_k a + t_k b$ for $1 \le k \le n+1$.

Theorem 1 (Thue). *Let a and b be integers, $a > b > 1$ with $\gcd(a,b) = 1$. Then the congruence $bx \equiv y \pmod{a}$ has a solution in nonzero integers x and y satisfying $|x| < \sqrt{a}, |y| \le \sqrt{a}$.*

Proof. As $r_n = \gcd(a,b) = 1$ and $a > \sqrt{a} > 1$ and the remainders $r_0, \dots, r_n$ in Euclid's algorithm decrease strictly to 1, there is a unique index k such that $r_{k-1} > \sqrt{a} \ge r_k$. Then the equation $a = |t_k| r_{k-1} + |t_{k-1}| r_k$ gives $a \ge |t_k| r_{k-1} > |t_k|\sqrt{a}$. Hence $|t_k| < \sqrt{a}$.

Finally, $r_k = s_k a + t_k b$, so $bt_k \equiv r_k \pmod{a}$ and we can take $x = t_k, y = r_k$. □

3. The equation $x^2 - Dy^2 = \kappa N$ with small κ

Let $N \ge 1$ be an odd integer, $D > 1$ and not a perfect square. Then a necessary condition for solvability of the equation $x^2 - Dy^2 = \kappa N$ with $\gcd(x,y) = 1$ is that the congruence $u^2 \equiv D \pmod{N}$ shall be soluble. From now on we assume this, together with $\gcd(D,N) = 1$ and $1 < u < N$. Then the Jacobi symbol $\left(\frac{D}{N}\right) = 1$. We note that if N is prime, then $\left(\frac{D}{N}\right) = 1$ also implies that $u^2 \equiv D \pmod{N}$ is soluble.

If we take $a = N$ and $b = u$ in Euclid's algorithm, the integers $r_k^2 - Dt_k^2$ decrease strictly for $k = 0, \dots, n$, from a^2 to $1 - Dt_n^2$ and are always multiples of N. For

$$r_k^2 - Dt_k^2 \equiv t_k^2 u^2 - Dt_k^2 \equiv t_k^2(u^2 - D) \equiv 0 \pmod{N}.$$

If k is chosen so that $r_{k-1} > \sqrt{N} > r_k$, as in the proof of Thue's theorem, then as

$$N = r_{k-1}|t_k| + r_k|t_{k-1}| > r_{k-1}|t_k|, \tag{1}$$

we have $|t_k| < \sqrt{N}$ and

$$-DN < r_k^2 - Dt_k^2 < N. \tag{2}$$

Hence $r_k^2 - Dt_k^2 = -lN$, $-1 < l < D$. In fact $1 \le l < D$. Hence

$$-DN < r_k^2 - Dt_k^2 \le -N. \tag{3}$$

Also $r_k^2 + lN = Dt_k^2$ and hence $Dt_k^2 > lN$. Hence

$$|t_k| > \sqrt{\frac{lN}{D}}. \tag{4}$$

From equation (1), $N > r_{k-1}|t_k|$ and hence inequality (4) implies

$$r_{k-1} < \sqrt{\frac{DN}{l}}. \tag{5}$$

4. The equation $x^2 - 2y^2 = \pm N$

The assumption $\left(\frac{2}{N}\right) = 1$ is equivalent to $N \equiv \pm 1 \pmod 8$. Also $1 \le l < 2$, so $l = 1$ and (3) gives $r_k^2 - 2t_k^2 = -N$. Hence from equation (5) with $D = 2$, $r_{k-1} < \sqrt{2N}$ and

$$-N = r_k^2 - 2t_k^2 < r_{k-1}^2 - 2t_{k-1}^2 < r_{k-1}^2 < 2N.$$

Hence $r_{k-1}^2 - 2t_{k-1}^2 = N$.

Example. Let $N = 10000000033$, a prime of the form $8n+1$. Then $u = 87196273$ gives $k = 10$, $r_{10} = 29015, t_{10} = -73627$, $r_9 = 118239, t_9 = 44612$ and $r_{10}^2 - 2t_{10}^2 = -N$, $r_9^2 - 2t_9^2 = N$.

Remark. We can express r_{k-1} and t_{k-1} in terms of r_k and t_k. The method is useful later for delineating cases when $D = 5, 6, 7$:

Using the identities

$$(6) \qquad (r_k r_{k-1} - Dt_k t_{k-1})^2 - D(t_k r_{k-1} - t_{k-1} r_k)^2 = (r_k^2 - Dt_k^2)(r_{k-1}^2 - Dt_{k-1}^2)$$

and

$$(7) \qquad (-1)^k N = r_k t_{k-1} - r_{k-1} t_k,$$

we deduce that

$$(8) \qquad r_k r_{k-1} - Dt_k t_{k-1} = \epsilon N,$$

where $\epsilon = \pm 1$.

From equation (8), we see that $\epsilon = 1$, as $t_k t_{k-1} < 0$. Hence

$$(9) \qquad r_k r_{k-1} + DT_k T_{k-1} = N,$$

where $T_k = |t_k|$. Then solving equations (7) and (9) with $D = 2$ for r_{k-1} and T_{k-1} yields

$$r_{k-1} = -r_k + 2T_k, \quad T_{k-1} = T_k - r_k,$$

5. The equation $x^2 - 3y^2 = \pm N$

The assumption $\left(\frac{3}{N}\right) = 1$ is equivalent to $N \equiv \pm 1 \pmod{12}$. From equation (3), we have $-3N < r_k^2 - 3t_k^2 \le -N$. Hence $r_k^2 - 3t_k^2 = -2N$ or $-N$.

Case 1. Assume $N \equiv 1 \pmod{12}$. Then $r_k^2 - 3t_k^2 = -N$ would imply the contradiction $r_k^2 \equiv -1 \pmod 3$.

Hence $r_k^2 - 3t_k^2 = -2N$ and inequality (5) implies $r_{k-1} < \sqrt{\frac{3N}{2}}$. Hence

$$-2N = r_k^2 - 3t_k^2 < r_{k-1}^2 - 3t_{k-1}^2 < r_{k-1}^2 < \frac{3N}{2}.$$

Consequently $r_{k-1}^2 - 3t_{k-1}^2 = N$.

We find $2r_{k-1} = -r_k + 3T_k$ and $2T_{k-1} = -r_k + T_k$.

Case 2. Assume $N \equiv -1 \pmod{12}$. Then $r_k^2 - 3t_k^2 = -2N$ would imply the contradiction $r_k^2 \equiv 2 \pmod 3$. Hence $r_k^2 - 3t_k^2 = -N$ and inequality (5) implies $r_{k-1} < \sqrt{3N}$. Hence

$$-N = r_k^2 - 3t_k^2 < r_{k-1}^2 - 3t_{k-1}^2 < r_{k-1}^2 < 3N.$$

Consequently $r_{k-1}^2 - 3t_{k-1}^2 = N$ or $2N$. However $r_{k-1}^2 - 3t_{k-1}^2 = N$ implies the contradiction $r_{k-1}^2 \equiv -1 \pmod 3$. Hence $r_{k-1}^2 - 3t_{k-1}^2 = 2N$.

We find $r_{k-1} = -r_k + 3T_k$ and $T_{k-1} = -r_k + T_k$.

6. The equation $x^2 - 5y^2 = \pm N$

The assumption $\left(\frac{5}{N}\right) = 1$ is equivalent to $N \equiv \pm 1 \pmod 5$. Then from equation (3), we have $-5N < r_k^2 - 5t_k^2 \le -N$. Hence $r_k^2 - 5t_k^2 = -4N, -3N, -2N$ or $-N$.

We cannot have $r_k^2 - 5t_k^2 = -3N$ as then $\left(\frac{5}{3}\right) = 1$. Neither can we have $r_k^2 - 5t_k^2 = -2N$, as N is odd.

Case 1. Assume $N \equiv 1 \pmod 5$. Then $r_k^2 - 5t_k^2 = -N$ would imply the contradiction $r_k^2 \equiv -1 \pmod 5$. Hence $r_k^2 - 5t_k^2 = -4N$. Then r_k and t_k are both odd. Also inequality (5) implies $r_{k-1} < \sqrt{\frac{5N}{4}}$. Hence $-N \le r_{k-1}^2 - 5t_{k-1}^2 \le N$.

Then as in the remark above, we can show

(i) if $r_{k-1}^2 - 5t_{k-1}^2 = -N$, then

$$4r_{k-1} = -3r_k + 5T_k, \quad 4T_{k-1} = -r_k + 3T_k$$

and hence $r_k \equiv -T_k \pmod 4$;

(ii) if $r_{k-1}^2 - 5t_{k-1}^2 = N$, then

$$4r_{k-1} = -r_k + 5T_k, \quad 4T_{k-1} = -r_k + T_k$$

and hence $r_k \equiv T_k \pmod 4$.

Case 2. Assume $N \equiv -1 \pmod 5$. Then $r_k^2 - 5t_k^2 = -4N$ would imply the contradiction $r_k^2 \equiv 4 \pmod 5$. Hence $r_k^2 - 5t_k^2 = -N$. Then not both r_k and t_k are odd. Also inequality (5) implies $r_{k-1} < \sqrt{5N}$ and we deduce that $-N < r_{k-1}^2 - 5t_{k-1}^2 \le 4N$. Consequently $r_{k-1}^2 - 5t_{k-1}^2 = N$ or $4N$.

Then as in the remark above, we can show

(i) if $r_{k-1}^2 - 5t_{k-1}^2 = N$, then

$$r_{k-1} = -2r_k + 5T_k, \quad T_{k-1} = -r_k + 2T_k$$

and hence $r_{k-1} \equiv -2r_k \pmod 5$;

(ii) if $r_{k-1}^2 - 5t_{k-1}^2 = 4N$, then

$$r_{k-1} = -r_k + 5T_k, \quad T_{k-1} = -r_k + T_k$$

and hence $r_{k-1} \equiv -r_k \pmod 5$.

Here is a complete classification of the possible cases:

1. $N = 5k + 1$. Then $r_k^2 - 5t_k^2 = -4N$, while r_k and t_k are odd.
   (i) $r_k \equiv -T_k \pmod 4$. Then $r_{k-1}^2 - 5t_{k-1}^2 = -N$.
   (ii) $r_k \equiv T_k \pmod 4$. Then $r_{k-1}^2 - 5t_{k-1}^2 = N$.
2. $N = 5k - 1$. Then $r_k^2 - 5t_k^2 = -N$, while r_k and t_k are not both odd.
   (i) $r_{k-1} \equiv -2r_k \pmod 5$. Then $r_{k-1}^2 - 5t_{k-1}^2 = N$.
   (ii) $r_{k-1} \equiv -r_k \pmod 5$. Then $r_{k-1}^2 - 5t_{k-1}^2 = 4N$.

7. The equation $x^2 - 6y^2 = \pm N$

The assumption $\left(\frac{6}{N}\right) = 1$ is equivalent to $N \equiv \pm 1 \pmod{24}$ or $N \equiv \pm 5 \pmod{24}$. Then from equation (3), we have $-6N < r_k^2 - 6t_k^2 \le -N$. Hence $r_k^2 - 6t_k^2 = -5N, -4N, -3N, -2N$ or $-N$. Only $-4N$ is ruled out immediately and the other possibilities can occur.

As with the case $D = 5$, there is a complete classification of the possible cases:

1. $N = 24k - 1$ or $24k + 5$.
   (i) $r_k \equiv 0 \pmod 3$. Then $r_k^2 - 6t_k^2 = -3N$, $r_{k-1}^2 - 6t_{k-1}^2 = -N$.

(ii) $r_k \not\equiv 0 \pmod 3$. Then $r_k^2 - 6t_k^2 = -N$.
(a) $r_{k-1} \equiv 0 \pmod 2$. Then $r_{k-1}^2 - 6t_{k-1}^2 = 2N$.
(b) $r_{k-1} \equiv 1 \pmod 2$. Then $r_{k-1}^2 - 6t_{k-1}^2 = 5N$.

2. $N = 24k + 1$ or $24k - 5$:
(i) $r_k \equiv 0 \pmod 2$. Then $r_k^2 - 6t_k^2 = -2N$, $r_{k-1}^2 - 6t_{k-1}^2 = N$.
(ii) $r_k \equiv 1 \pmod 2$. Then $r_k^2 - 6t_k^2 = -5N$.
(a) $r_k \equiv T_k \pmod 5$. Then $r_{k-1}^2 - 6t_{k-1}^2 = N$.
(b) $r_k \equiv -T_k \pmod 5$. Then
$$r_{k-1}^2 - 6t_{k-1}^2 = -2N,\ r_{k-2}^2 - 6t_{k-2}^2 = N.$$

8. The equation $x^2 - 7y^2 = \pm N$

The assumption $\left(\frac{7}{N}\right) = 1$ is equivalent to $N \equiv 1, 3, 9, 19, 25, 27 \pmod{28}$.
As with the case $D = 6$, there is a complete classification of the possible cases:

1. $N = 28k + 1, 28k + 9$, or $28k + 25$.
(i) $r_k \equiv T_k \pmod 2$. Then $r_k^2 - 7t_k^2 = -6N$.
(a) $r_k \equiv -T_k \pmod 6$. Then $r_{k-1}^2 - 7t_{k-1}^2 = -3N$.
(1) $r_{k-1} \equiv -T_{k-1} \pmod 3$. Then $r_{k-2}^2 - 7t_{k-2}^2 = N$.
(2) $r_{k-1} \equiv T_{k-1} \pmod 3$. Then $r_{k-2}^2 - 7t_{k-2}^2 = 2N$.
(b) $r_k \equiv T_k \pmod 6$. Then $r_{k-1}^2 - 7t_{k-1}^2 = N$.
(ii) $r_k \not\equiv T_k \pmod 2$. Then $r_k^2 - 7t_k^2 = -3N$.
(a) $r_k \equiv -T_k \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = N$.
(b) $r_k \equiv T_k \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = 2N$.

2. $N = 28k + 3, 28k + 19$, or $28k + 27$.
(i) $r_k \equiv T_k \pmod 2$. Then $r_k^2 - 7t_k^2 = -2N$.
(a) $r_{k-1} \equiv -T_{k-1} \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = -N$.
(b) $r_{k-1} \equiv T_{k-1} \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = 3N$.
(ii) $r_k \not\equiv T_k \pmod 2$. Then $r_k^2 - 7t_k^2 = -N$.
(a) $r_{k-1} \equiv -T_{k-1} \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = 3N$.
(b) $r_{k-1} \equiv T_{k-1} \pmod 3$. Then $r_{k-1}^2 - 7t_{k-1}^2 = 6N$.

In cases 1(a)(2) and 2(i), the equations $r_{k-2}^2 - 7t_{k-2}^2 = 2N$ and $r_k^2 - 7t_k^2 = -2N$ give rise to equations $x^2 - 7y^2 = N, -N$, respectively, if we write $x + y\sqrt{7} = (r_{k-2} + t_{k-2}\sqrt{7})/(3 + \sqrt{7})$ and $(r_k + t_k\sqrt{7})/(3 + \sqrt{7})$, respectively. For if $x + y\sqrt{7} = (r + t\sqrt{7})/(3 + \sqrt{7})$, where r and t are odd, then $x = \frac{3r-7t}{2}$ and $y = \frac{3t-r}{2}$ are integers and $x^2 - 7y^2 = (r^2 - 7t^2)/2$.

We note that 1(a)(2) cannot occur unless $N \equiv 0 \pmod 3$ for we have

$$r_{k-1} = \frac{-5r_k + 7T_k}{6}, \qquad T_{k-1} = \frac{-r_k + 5T_k}{6} \tag{10}$$

$$r_{k-2} = \frac{-r_{k-1} + 7T_{k-1}}{3}, \qquad T_{k-2} = \frac{-r_{k-1} + T_{k-1}}{3}. \tag{11}$$

Then (10) implies $r_{k-1} + T_{k-1} = -r_k + 2T_k \equiv -r_k - T_k \equiv 0 \pmod 3$. Also (11) implies $r_{k-1} \equiv T_{k-1} \pmod 3$. Hence 3 divides r_{k-1} and T_{k-1}, and the equation $r_{k-1}^2 - 7t_{k-1}^2 = -3N$ then implies 3 divides N.

Example. $N = 57$. The congruence $u^2 \equiv 7 \pmod{57}$ has solutions $u \equiv \pm 8, \pm 11 \pmod{57}$. Then $u = 8$ gives $k = 2, r_1 = 8, t_1 = 1, r_2 = 1, t_2 = -7$, $r_k^2 - 7t_k^2 = -6N$ and $r_{k-1}^2 - 7t_{k-1}^2 = N$, while $u = 11$ gives $k = 2, r_1 = 11, t_1 = 1, r_2 = 2, t_2 = -5$ and $r_k^2 - 7t_k^2 = -3N$ and $r_{k-1}^2 - 7t_{k-1}^2 = 2N$.

Acknowledgments

The author is grateful to Christina Miller for noticing the decreasing property of the integers $r_k^2 - Dt_k^2$.

The author is also grateful to Dr. Terence Jackson for his comments on an earlier draft of the paper.

The calculations were carried out with the author's number theory calculator program CALC and a UNIX **bc** program `thue`, both available at `http://www.maths.uq.edu.au/~krm/`.

References

1. J. Brillhart, *Note on representing a prime as a sum of two squares*, Math. Comp. **26** (1972) 1011–1013. MR **47:**3297
2. G. Cornacchia, *Su di un metodo per la risoluzione in numeri interi dell' equazione* $\sum_{h=0}^{n} C_h x^{n-h} = P$, Giornale di Matematiche di Battaglini **46** (1908) 33–90.
3. K. Hardy, J.B. Muskat, K.S. Williams, *A deterministic algorithm for solving* $n = fu^2 + gv^2$ *in coprime integers* u *and* v, Math. Comp. **55** (1990) 327–343. MR **91d:**11164
4. K. Hardy, J.B. Muskat, K.S. Williams, *Solving* $n = au^2 + buv + cv^2$ *using the Euclidean algorithm*, Utilitas Math. **38** (1990) 225–236. MR **92c:**11038
5. C. Hermite, *Note au sujet de l'article précedent*, J. Math. Pures Appl., **13** (1848) 15.
6. J.B. Muskat, *A refinement of the Hardy–Muskat–Williams algorithm for solving* $n = fu^2 + gv^2$, Utilitas Math. **41** (1992) 109–117. MR **93h:**11030
7. T. Nagell, *Introduction to Number Theory*, Chelsea Publishing Company, NY 1981. MR **30:**4714
8. A. Nitaj, *L'algorithme de Cornacchia*, Expositiones Mathematicae **13** (1995) 358–365. MR **97a:**11044
9. J.A. Serret, *Sur un théorème rélatif aux nombres entières*, J. Math. Pures Appl. **13** (1848) 12–14.
10. A. Thue, *Et par antydninger til en taltheorisk methode*, Selected Mathematical Papers of Axel Thue, Universitetsforlaget, Oslo 1977. MR **57:**46
11. J.V. Uspensky and M.A. Heaslet, *Elementary Number Theory*, McGraw–Hill, NY 1939. MR **1:**38d
12. S. Wagon, *The Euclidean algorithm strikes again*, Amer. Math. Monthly **97** (1990) 125–129. MR **91b:**11039
13. P. Wilker, *An efficient algorithmic solution of the diophantine equation* $u^2 + 5v^2 = m$, Math. Comp. **35** (1980) 1347–1352. MR **81m:**10021
14. K.S. Williams, *On finding the solutions of* $n = au^2 + buv + cv^2$ *in integers* u *and* v, Utilitas Math. **46** (1994) 3–19. MR **95g:**11019
15. K.S. Williams, *Some refinements of an algorithm of Brillhart*, Number Theory (Halifax) CMS Conference Proc. **15** (1994) 409–416. MR **96f:**11169

Department of Mathematics, University of Queensland, Brisbane, Australia, 4072

E-mail address: `krm@maths.uq.edu.au`