A NICE DIOPHANTINE EQUATION

MARIA CHIARA BRAMBILLA

1. Introduction

One of the most interesting and fascinating subjects in Number Theory is the study of diophantine equations. A **diophantine equation** is a polynomial equation with coefficient in $\mathbb{Z}$. Solving a diophantine equation means either to find all the integer solutions, or to prove that there are no integer solutions. The name of these equations comes from Diophantus of Alexandria (probably third century A.D.), whose main work is *"Arithmetica"*, a collection of arithmetical problems.

In general to solve a diophantine equation can be very difficult and involve deep mathematical results. For example, the famous Fermat's Last Theorem states that the diophantine equation $x^n + y^n = z^n$ has no nonzero integer solutions for $n \geq 3$. Proving this statement is extraordinarily difficult and it has been completed only in recent times. It is remarkable that Fermat wrote this theorem as a marginal note in his copy of *"Arithmetica"*, without giving a proof.

In this paper we will solve a diophantine equation of degree 2 (quadratic) in two variables by means of elementary techniques. Our **problem** is to find all the integer solutions (x, y) of the following equation

$$x^2 - Nxy + y^2 = 1, \tag{1.1}$$

where N is a fixed integer number.

We can suppose $N \geq 0$ since if $(\widetilde{x}, \widetilde{y})$ is a solution of (1.1) for $N = \widetilde{N}$ then $(\widetilde{x}, -\widetilde{y}), (-\widetilde{x}, \widetilde{y})$ are solutions of (1.1) for $N = -\widetilde{N}$.

Let us consider first the special cases $N = 0, 1, 2$. The cases $N = 0, 1$ are easy to solve because the number of solutions is finite.

• If $N = 0$, the equation becomes $x^2 + y^2 = 1$, and the set of real solutions is represented by a circle with centre $(0, 0)$ and radius 1 in the plane with coordinates x, y. It is clear that the only points with integer coordinates are $(1, 0), (0, 1), (-1, 0), (0, -1)$. There are no other solutions because if $|x| > 1$ there are no $y \in \mathbb{R}$ which satisfy (1.1).

• If $N = 1$, the set of real solutions is represented by an ellipse. From Figure 1 we can easily see that there exist exactly six points with integer

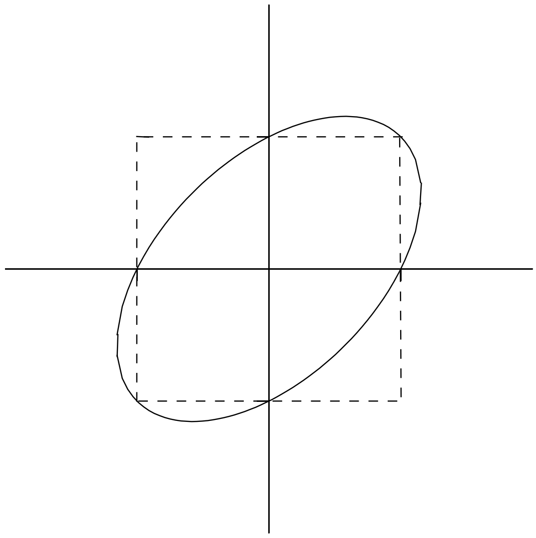

FIGURE 1. Case N=1

coordinates: $(1, 0), (1, 1), (0, 1)$ and $(-1, 0), (-1, -1), (0, -1)$. Indeed if $|x| > 1$ there are no solutions $y \in \mathbb{Z}$.

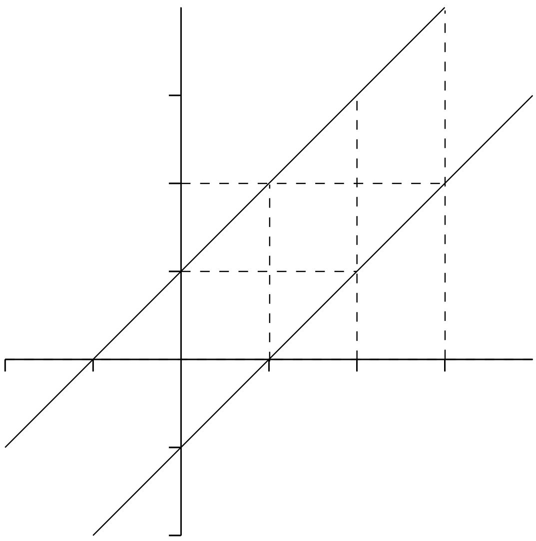

FIGURE 2. Case N=2

• If $N = 2$, our equation is $(x - y)^2 = 1$ i.e. $(x - y - 1)(x - y + 1) = 0$ and the set of real solutions is represented by two straight lines (see Figure 2). In this case, the solutions of (1.1) are all the integer points on the lines $y = x - 1$ and $y = x + 1$ and they are infinitely many. It is immediate to give the complete list of solutions, that is: $(x, y) = (n, n-1)$ or $(x, y) = (n, n+1)$ for all $n \in \mathbb{Z}$. It is clear that there are no other integer solutions.

2. Looking for solutions

Let us now consider the case $N \geq 3$. In this case the set of real solutions of equation (1.1) is represented by an hyperbola. Since the hyperbola is not bounded, it is not immediate to see which points have integer coordinates. A first remark is that the hyperbola is symmetric with respect to the lines $y = x$ and $y = -x$. It follows that if $(\widetilde{x}, \widetilde{y})$ is a solution of (1.1), then $(\widetilde{y}, \widetilde{x}), (-\widetilde{x}, -\widetilde{y}), (-\widetilde{y}, -\widetilde{x})$ are solutions too.

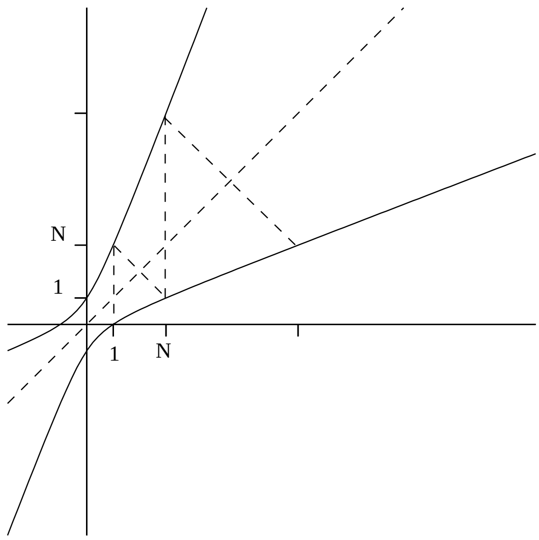

Figure 3. Case $N \geq 3$

Now we will use this symmetry in order to construct a list of solutions. If $x = 1$, equation (1.1) has two solutions $y = 0, N$. We represent the situation in Figure 3: starting from the point $(1, 0)$ we find the second intersection of $x = 1$ with the hyperbola, that is $(1, N)$. By symmetry with respect to $y = x$ we get the new solution $(N, 1)$ and we search for the second intersection of $x = N$ with the hyperbola.

We can repeat this construction starting from any solution (x_0, y_0) of (1.1). Indeed the intersection of $x = x_0$ and (1.1) gives the equation

$$y^2 - Nx_0y + x_0^2 - 1 = 0$$

whose two solutions y_0 and y_1 satisfy $y_0 + y_1 = Nx_0$. Therefore if $x_0, y_0 \in \mathbb{Z}$ also $y_1 \in \mathbb{Z}$. This means that any integer solution (x_0, y_0) gives a sequence of integer solutions in this way:

$$(x_0, y_0) \to (x_0, Nx_0 - y_0) \to (Nx_0 - y_0, x_0) \to$$

$$\to (Nx_0 - y_0, N(Nx_0 - y_0) - x_0) \to \dots$$

In particular starting from the solution $(1,0)$ we obtain infinitely many solutions. In order to write sintetically these solutions let us define recursively the sequence $\{a_k\}$ as follows.

$$\begin{cases} a_0 = 0, \\ a_1 = 1, \\ a_{k+1} = Na_k - a_{k-1}, \end{cases}$$

Then the solutions we obtain starting from $(0,1)$ are

$$(x,y) = (a_k, a_{k+1}) \quad \text{and} \quad (x,y) = (a_{k+1}, a_k).$$

Clearly $(-a_k, -a_{k+1})$ and $(-a_{k+1}, -a_k)$ are solutions too.

3. An interesting sequence

Let us study the sequence $\{a_k\}$. We want to obtain an explicit form starting from the recursive definition.

First we can search for solutions of the relation $a_{k+1} = Na_k - a_{k-1}$ with the special form $a_k = x^k$, for some $x \in \mathbb{R}$. We obtain the equation $x^2 - Nx + 1 = 0$, which has two solutions $\alpha = \frac{N+\sqrt{N^2-4}}{2}$ and $\beta = \frac{N-\sqrt{N^2-4}}{2}$. Hence for any $u, v \in \mathbb{R}$, we have that $a_k = u\alpha^k + v\beta^k$ satisfies the recurrence relation. By imposing the initial condition $a_0 = 0, a_1 = 1$ we obtain $u = \frac{1}{\sqrt{N^2-4}}$ and $v = -\frac{1}{\sqrt{N^2-4}}$ Thus the explicit form of an element of the sequence is the following:

$$a_k = \frac{\left(\frac{N+\sqrt{N^2-4}}{2}\right)^k - \left(\frac{N-\sqrt{N^2-4}}{2}\right)^k}{\sqrt{N^2-4}}.$$

Notice that if $N = 3$ we have $\{a_k\} = \{0, 1, 3, 8, 21, 55, \dots\}$. All the terms of this sequence are Fibonacci numbers. Recall that the well known **Fibonacci sequence** is defined by

$$\begin{cases} f_0 = 0, \\ f_1 = 1, \\ f_{k+1} = f_k + f_{k-1}. \end{cases}$$

If $N = 3$ the family $\{a_k\}$ is exactly the even subsequence of the Fibonacci sequence, that is $a_k = f_{2k}$. Let us prove this fact by means of the **induction principle**:

(i) First step: $k = 0$, then $a_0 = 0$, $f_0 = 0$.

(ii) Induction step: suppose that $a_h = f_{2h}$ for all $h \le k$. Let us show that $a_{k+1} = f_{2k+2}$. Indeed

$$a_{k+1} = 3a_k - a_{k-1} = 3f_{2k} - f_{2k-2} = 3f_{2k} - (f_{2k} - f_{2k-1}) =$$

$$= 2f_{2k} + (f_{2k+1} - f_{2k}) = f_{2k} + f_{2k+1} = f_{2k+2}.$$

The name of Fibonacci numbers comes from Leonardo Pisano, son (*filius*) of Bonacci, whose book *"Liber abaci"* (1202) was fundamental to introduce Arabic and Indian arithmetic into Europe. Fibonacci numbers satisfy amazing properties and appear in mathematics and in nature in unexpected ways. For instance, the ratio of consecutive Fibonacci numbers, that is $\frac{f_{k+1}}{f_k}$, converges to the **golden ratio** $\frac{1+\sqrt{5}}{2}$.

In general, if $N \neq 3$, our sequence $\{a_k\}$ is not constituted by Fibonacci numbers, nevertheless it satisfies several good properties analogously to Fibonacci sequence. Let us see some of these properties.

• The sequence of ratios is decreasing

$$\frac{a_k}{a_{k-1}} \geq \frac{a_{k+1}}{a_k}$$

and converges to

$$\lim_{k\to\infty} \left(\frac{a_{k+1}}{a_k}\right) = \frac{N+\sqrt{N^2-4}}{2}.$$

Notice that if $N = 3$ this limit is $\frac{3+\sqrt{5}}{2} = (\frac{1+\sqrt{5}}{2})^2$, that is the square of the golden ratio.

• For all $k \geq 1$

$$a_k^2 - a_{k+1}a_{k-1} = 1$$

from which it follows that $GCD(a_k, a_{k-1}) = 1$, that is a_k and a_{k-1} are relatively prime.

• For any fixed $N \geq 3$,

$$a_{k-1}^2 + a_k^2 - Na_{k-1}a_k = 1 \quad \text{for all} \quad k \geq 1 \tag{3.1}$$

The previous properties can be easily proved by induction. Let us see for example the proof of equality (3.1).

(i) If $k = 1$, since $a_0 = 0, a_1 = 1$, obviously the equality $a_0^2 + a_1^2 - Na_0a_1 = 1$ is true.

(ii) Let us suppose that (3.1) holds for all $k \leq n$ and let us prove that $a_n^2 + a_{n+1}^2 - Na_na_{n+1} = 1$. By substituing $a_{n+1} = Na_n - a_{n-1}$ and by induction hypothesis we obtain:

$$a_n^2 + (Na_n - a_{n-1})^2 - Na_n(Na_n - a_{n-1}) = a_n^2 + a_{n-1}^2 - Na_na_{n-1} = 1.$$

In other words we have verified that (a_{k+1}, a_k) is a solution of equation (1.1), as we already knew from the previous geometric construction.

4. Uniqueness

Our next goal is to prove that the solutions we have found are the only possible solutions. Because of the symmetry it is enough to prove that any solution (x, y) with $x \geq y$ belongs to the family $\{(a_{k+1}, a_k)\}$.

First let us show this fact by means of a **geometrical** argument. Suppose that (x_0, y_0) is any integer point on the lower branch of the hyperbola, that is with $x_0 \geq y_0$.

We know by symmetry that also (y_0, x_0) is an integer point. On the other hand, also the second intersection of $x = y_0$ with the hyperbola is an integer point with coordinates $(y_0, Ny_0 - x_0)$.

In other words we can apply backwards the same argument used in Section 2 and represented in Figure 3. Thus starting from any integer point we get a sequence of integer points

$$(x_0, y_0) \to (y_0, x_0) \to (y_0, Ny_0 - x_0) \to \dots$$

Notice that there is necessarily a point (x_1, y_1) in this sequence that lies on the lower branch of the hyperbola (which is the graph of an increasing function) between $(1, 0)$ and $(N, 1)$. Consequently, we have that $0 \leq y_1 \leq 1$, that is $(x_1, y_1) = (1, 0)$ or $(x_1, y_1) = (N, 1)$. This implies that the starting point (x_0, y_0) belongs to the family $\{(a_{k+1}, a_k)\}$ as we claimed.

We see now an **algebraic** proof of the uniqueness of the family $\{(a_{k+1}, a_k)\}$ of solutions. Suppose that (x, y) with $x \geq y$ is an integer solution of (1.1).

From the inequality $x^2 - Nxy + y^2 = 1 \geq 0$, since $x \geq y$ and $N \geq 3$, we obtain that $\frac{x}{y} > (\frac{N+\sqrt{N^2-4}}{2})$. Since $\frac{a_{k+1}}{a_k}$ is decreasing to $(\frac{N+\sqrt{N^2-4}}{2})$, it follows that there exists $k \geq 0$ such that

$$\frac{a_{k+1}}{a_k} \leq \frac{x}{y} < \frac{a_k}{a_{k-1}}.$$

Consider the following system

$$\begin{cases} x = na_k + ma_{k+1}, \\ y = na_{k-1} + ma_k. \end{cases}$$

This system has discriminant $\Delta = a_k^2 - a_{k+1}a_{k-1} = 1$, consequently it admits a solution (n, m), with $n, m \in \mathbb{Z}$. In particular, $n \geq 0$ because $\frac{x}{y} \geq \frac{a_{k+1}}{a_k}$, and $m > 0$ because $\frac{x}{y} < \frac{a_k}{a_{k-1}}$.

Then let us compute

$$x^2 - Nxy + y^2 = (na_k + ma_{k+1})^2 - N(na_k + ma_{k+1})(na_{k-1} + ma_k) + (na_{k-1} + ma_k)^2 =$$

$$n^2(a_k^2 + a_{k-1}^2 - Na_k a_{k-1}) + m^2(a_{k+1}^2 + a_k^2 - Na_{k+1}a_k) +$$

$$+ nm(2a_k a_{k+1} + 2a_{k-1}a_k - Na_k^2 - Na_{k+1}a_{k-1}) =$$

$$n^2 + m^2 + Nnm,$$

indeed

$$2a_k a_{k+1} + 2a_{k-1}a_k - Na_k^2 - Na_{k+1}a_{k-1} =$$

$$= 2a_k(Na_k - a_{k-1}) + 2a_{k-1}a_k - Na_k^2 - N(Na_k - a_{k-1})a_{k-1} =$$

$$= N(a_k^2 - Na_k a_{k-1} + a_{k-1}^2) = N.$$

Therefore, since

$$1 = x^2 - Nxy + y^2 = n^2 + m^2 + Nnm$$

and $n \geq 0, m > 0$, the only possible solution is $n = 0, m = 1$.

This implies that all the solutions are of the form (a_{k+1}, a_k) as claimed.

5. Conclusion

For any $N \in \mathbb{Z}$, all the solutions of the diophantine equation

$$x^2 - Nxy + y^2 = 1$$

are given by the family

$$\{(x, y) = (a_k, a_{k+1}), (a_{k+1}, a_k), (-a_k, -a_{k+1}), (-a_{k+1}, -a_k)\}$$

where the sequence a_k is defined as

$$\begin{cases} a_0 = 0, \\ a_1 = 1, \\ a_{k+1} = Na_k - a_{k-1}, \qquad \text{for } k \geq 1. \end{cases}$$

Dipartimento di Matematica e Applicazioni per l'Architettura, Università di Firenze, piazza Ghiberti, 27, 50122, Florence, Italy

E-mail address: brambilla@math.unifi.it