

# INTEGRAL POINTS ON GENERIC FIBERS

ARNAUD BODIN

ABSTRACT. Let  $P(x, y)$  be a rational polynomial. If the curve  $(P(x, y) = k)$ ,  $k \in \mathbb{Q}$ , is irreducible and admits an infinite number of points whose coordinates are integers, Siegel's theorem implies that the curve is rational. We deal with the case where  $k$  is a generic value and prove, in the spirit of the Abhyankar-Moh-Suzuki theorem, that there exists an algebraic automorphism sending  $P(x, y)$  to the polynomial  $x$  or to  $x^2 - \ell y^2$ ,  $\ell \in \mathbb{N}$ . Moreover for such curves we give a sharp bound for the number of integral points  $(x, y)$  with  $x$  and  $y$  bounded.

## 1. INTRODUCTION

Let  $P \in \mathbb{Q}[x, y]$  be a non-constant polynomial and  $\mathcal{C} = (P(x, y) = 0) \subset \mathbb{C}^2$  be the corresponding algebraic curve. An old and famous result is the following, [12]:

**Theorem** (Siegel's theorem). *Suppose that  $\mathcal{C}$  is irreducible. If the number of integral points  $\mathcal{C} \cap \mathbb{Z}^2$  is infinite then  $\mathcal{C}$  is a rational curve.*

Our main goal is to prove a stronger version of Siegel's theorem for curve defined by an equation  $\mathcal{C} = (P(x, y) = k)$  where  $k$  is a "generic" value. It is known that there exists a finite set  $\mathcal{B}$  such that the topology of the complex plane curve  $(P(x, y) = k) \subset \mathbb{C}^2$  is independent of  $k \in \mathbb{C} \setminus \mathcal{B}$ . We say that  $k \in \mathbb{C} \setminus \mathcal{B}$  is a *generic value*.

**Theorem 1.** *Let  $P \in \mathbb{Q}[x, y]$  and let  $k \in \mathbb{Q} \setminus \mathcal{B}$  be a generic value. Suppose that the algebraic curve  $\mathcal{C} = (P(x, y) = k)$  is irreducible. If  $\mathcal{C}$  contains an infinite number of integral points  $(m, n) \in \mathbb{Z}^2$  then there exists an algebraic automorphism  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  such that*

$$P \circ \Phi(x, y) = x \quad \text{or} \quad P \circ \Phi(x, y) = \alpha(x^2 - \ell y^2) + \beta,$$

where  $\ell \in \mathbb{N}^*$  is a non-square and  $\alpha \in \mathbb{Q}^*$ ,  $\beta \in \mathbb{Q}$ .

We recall that an algebraic automorphism  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  is a map  $\Phi : \mathbb{Q}^2 \rightarrow \mathbb{Q}^2$  defined by a pair of polynomials  $\Phi(X, Y) = (\phi_1(X, Y), \phi_2(X, Y))$ ,  $\phi_1(X, Y), \phi_2(X, Y) \in \mathbb{Q}[X, Y]$ ; moreover it is invertible in the sense

---

*Date:* September 2, 2009.

that there exists a pair of polynomials  $(\psi_1(X, Y), \psi_2(X, Y))$  such that  $\Phi(\psi_1(X, Y), \psi_2(X, Y)) = (X, Y)$ . In particular the curve  $\mathcal{C} = (P(x, y) = k)$  is diffeomorphic to a line  $(x = 0)$ , in which case the set  $\mathcal{B}$  is empty or to a hyperbola  $(x^2 - dy^2 = 1)$  in which case  $\mathcal{B}$  is a singleton.

Theorem 1 can be seen as an arithmetic version of the Abhyankar-Moh-Suzuki theorem [2] and in fact we use this result. It can also be seen as a strong version, for generic values, of a result of Nguyen Van Chau [8] concerning possible counter-examples to the Jacobian conjecture.

Let us give a first example for which the theorem applies: let  $P(x, y) = x - y^d$ . The curve  $\mathcal{C} = (x - y^d = 0)$  has infinitely many integral points of type  $(n^d, n)$ ,  $n \in \mathbb{Z}$ . And for the algebraic automorphism  $\Phi(x, y) = (x + y^d, y)$  we have  $P \circ \Phi(x, y) = x$ . In particular the curve  $\mathcal{C}$  is sent (by  $\Phi^{-1}$ ) to the line  $(x = 0)$ . As pointed out by Kevin Buzzard the second case corresponds to Pell's equation and a second example for which the theorem applies is  $(x^2 - 2y^2 = 1)$ : it admits an infinite number of integral points. Of course a kind of reciprocal of Theorem 1 is true. Let  $Q_1(x, y) = x$ , (resp.  $Q_2(x, y) = x^2 - \ell y^2$ ) and  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  whose inverse  $\Phi^{-1}$  has integral coefficients. If we set  $P_1 = Q_1 \circ \Phi$  (resp.  $P_2 = Q_2 \circ \Phi$ ) then the curve  $(P_1 = 0)$  (resp.  $(P_2 = 1)$ ) has infinitely many integral points.

For non-generic values the result is not true, for example let  $P(x, y) = x^2 - y^3$  and  $\mathcal{C} = (x^2 - y^3 = 0)$ . The integral points  $(n^3, n^2)$ ,  $n \in \mathbb{Z}$  belong to  $\mathcal{C}$ , but as  $\mathcal{C}$  is singular it cannot be algebraically equivalent to a line not to a hyperbola.

As an application, we draw the following corollary:

**Corollary 2.** *Suppose that the hypotheses of Theorem 1 are true. For infinitely many values  $k \in \mathbb{Q}$ , the curve  $\mathcal{C}_k = (P(x, y) - k = 0)$  contains infinitely many integral points.*

Let us sketch the proof of Theorem 1, we start with a curve  $\mathcal{C}$  having infinitely many integral points. By Siegel's theorem  $\mathcal{C}$  is a rational curve. In fact original Siegel's theorem says more: the curve has one or two places at infinity. Firstly suppose that  $\mathcal{C}$  has only one place at infinity; as  $\mathcal{C}$  is defined by  $(P(x, y) = k)$  for a generic  $k$ ,  $\mathcal{C}$  is a smooth curve. Using the Abhyankar-Moh-Suzuki theorem [2], the curve can be sent to a line by an algebraic automorphism. Next if  $\mathcal{C}$  has two places at infinity the curve can be sent to a hyperbola by a result of Neumann [7]. Then we are able to explicitly write down equations for the polynomials.

We will apply Theorem 1 to obtain new bounds for the number of integral points on algebraic curves. Let  $\mathcal{C} = (P(x, y) = 0)$  be an algebraic curve, and let  $d = \deg P$ . Let

$$N(\mathcal{C}, B) = \# \{ (x, y) \in \mathcal{C} \cap \mathbb{Z}^2 \mid |x| \leq B \text{ and } |y| \leq B \}.$$

An upper bound for  $N(\mathcal{C}, B)$  is given by Bombieri and Pila in [4]; we shall use results of Heath-Brown [6] that have been made explicit by Walkowiak [13]:

**Theorem.** *For all irreducible curves  $\mathcal{C}$  of degree  $d$  and all  $B > 0$ :*

- (1) (Heath-Brown)  $N(\mathcal{C}, B) \leq C_{d,\varepsilon} B^{\frac{1}{d} + \varepsilon}$  for some constant  $C_{d,\varepsilon}$ ,
- (2) (Walkowiak)  $N(\mathcal{C}, B) \leq 2^{48} d^8 \ln(B)^5 B^{\frac{1}{d}}$ .

The term  $B^{\frac{1}{d}}$  in the theorem above is sharp but the term  $\ln(B)^5$  (corresponding to  $B^\varepsilon$ ) and especially the constant  $2^{48} d^8$  are probably far from being best possible.

For curves  $\mathcal{C}$  as in Theorem 1 we will give sharp bounds for  $N(\mathcal{C}, B)$ . First of all if  $\mathcal{C} = (P = k)$  and the polynomial  $P$  is algebraically equivalent to  $x^2 - \ell y^2$  (i.e. there exists  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  such that  $P \circ \Phi(x, y) = x^2 - \ell y^2$ ) it is known [11, p. 135] that there exists  $C > 0$  such that

$$N(\mathcal{C}, B) \leq C \cdot \ln(B).$$

Of course it implies  $N(\mathcal{C}, B) \leq B^{\frac{1}{d}}$  for sufficiently large  $B$  and we shall omit this case.

So if  $P$  is not algebraically equivalent to  $x^2 - \ell y^2$  then, as a corollary of Theorem 1, such a curve  $\mathcal{C}$  admits a parametrization by polynomials: let  $(p(t), q(t))$  be a parametrization of  $\mathcal{C}$  with rational coefficients. Moreover  $\deg P$  is equal to  $\deg p$  or  $\deg q$  (see [9, Lemma 2.1]). We suppose  $\deg P = \deg p$  and write

$$p(t) = \frac{1}{b} (a_d t^d + a_{d-1} t^{d-1} + \cdots + a_0)$$

where  $a_0, \dots, a_d, b \in \mathbb{Z}$ ,  $b > 0$  and  $\gcd(a_0, \dots, a_d, b) = 1$ .

**Theorem 3.** *Let  $P \in \mathbb{Q}[x, y]$  be a polynomial of degree  $d$ , let  $k \in \mathbb{Q} \setminus \mathcal{B}$  be a generic value. Suppose that  $\mathcal{C} = (P(x, y) = k)$  is an irreducible algebraic curve with infinitely many integral points. Then for all sufficiently large  $B$ , the number  $N(\mathcal{C}, B)$  of integral points on  $\mathcal{C}$  bounded by  $B$  verifies:*

$$N(\mathcal{C}, B) \leq 2 |a_d|^{1 - \frac{1}{d}} b^{\frac{1}{d}} B^{\frac{1}{d}} + 2,$$

where  $a_d$  and  $b$  are defined above.

*Acknowledgments:* This work has been done during a visit at the University of Zaragoza; I wish to thank the people of the Department of Mathematics and especially Enrique Artal Bartolo for hospitality. I also thank Pierre Dèbes for his support and the referees for useful comments.

## 2. PARAMETRIZATION

**2.1. Topology of polynomials.** By a result of Thom, for a polynomial  $P \in \mathbb{C}[x, y]$  seen as a map  $P : \mathbb{C}^2 \rightarrow \mathbb{C}$  there exists a finite set  $\mathcal{B} \subset \mathbb{C}$  such that

$$P : P^{-1}(\mathbb{C} \setminus \mathcal{B}) \longrightarrow \mathbb{C} \setminus \mathcal{B}$$

is a topological locally trivial fibration. A value  $k \notin \mathcal{B}$  is a *generic value*.

For example we have the following characterization of the generic values: the Euler characteristic of the complex plane curve  $(P(x, y) = k) \subset \mathbb{C}^2$  is independent of  $k \notin \mathcal{B}$  and jumps if and only if  $k \in \mathcal{B}$ .

Of course the image by  $P$  of a singular point of any curve  $(P(x, y) = k)$ ,  $k \in \mathbb{C}$ , is not a generic value, but for example  $P(x, y) = x(xy - 1)$  has no singular points while  $\mathcal{B} = \{0\}$ . Then if  $k \notin \mathcal{B}$  is a generic value the plane algebraic curve  $\mathcal{C} = (P(x, y) = k) \subset \mathbb{C}^2$  does not have singular points. Hence if  $\mathcal{C}$  is connected then  $\mathcal{C}$  is irreducible.

The connectedness of a generic fiber  $\mathcal{C}$  is equivalent to  $P(x, y)$  being non-composite [1]. We recall that  $P(x, y)$  is *composite* if there exist  $h \in \mathbb{C}[t]$ ,  $\deg h \geq 2$ , and  $Q \in \mathbb{C}[x, y]$  such that  $P(x, y) = h \circ Q(x, y)$ . By [3, Theorem 7] we even can choose  $h$  and  $Q$  with rational coefficients. Consequently it has been noticed by Janusz Gwozdziejewicz that the hypothesis “ $\mathcal{C}$  is irreducible” in Theorem 1 can be removed. In that case the conclusion becomes  $P \circ \Phi(x, y) = h(x)$  or  $P \circ \Phi(x, y) = h(x^2 - ly^2)$ , where  $h \in \mathbb{Q}[t]$  is a one-variable polynomial of positive degree.

**2.2. Algebraic automorphisms.** For  $K = \mathbb{Q}$  or  $K = \mathbb{C}$  an *algebraic automorphism*  $\Phi \in \text{Aut } \mathbb{A}_K^2$  is a polynomial map  $\Phi = (\phi_1, \phi_2) : K^2 \rightarrow K^2$  (where  $\phi_1(X, Y), \phi_2(X, Y) \in K[X, Y]$ ) which is invertible (that is to say there exists  $\psi_1(X, Y), \psi_2(X, Y) \in K[X, Y]$  such that  $\Phi(\psi_1(X, Y), \psi_2(X, Y)) = (X, Y)$ ). The polynomials  $P, Q \in K[x, y]$  are *algebraically equivalent* if there exists  $\Phi \in \text{Aut } \mathbb{A}_K^2$  such that  $Q = P \circ \Phi$ . And in fact such  $P$  and  $Q$  have the same topological and algebraic properties.

**2.3. Siegel’s theorem and the topology of  $\mathcal{C}$ .** Siegel’s theorem as stated in the introduction says that an irreducible curve having infinitely many integral points is rational, that is to say its genus is

zero. Another formulation is that the curve admits a parametrization by rational fractions. In fact Siegel's original theorem [12] gives more information than the rationality of the curve.

**Theorem 4** (Siegel's theorem). *Suppose that  $\mathcal{C}$  is an irreducible curve with infinitely many integral points; then  $\mathcal{C}$  is a rational curve with one or two places at infinity.*

As a corollary we get in our situation:

**Lemma 5.** *If  $k$  is a generic value and the curve  $\mathcal{C} = (P(x, y) = k)$  contains infinitely many integral points then  $\mathcal{C}$  is homeomorphic to  $\mathbb{C}$  or to  $\mathbb{C}^*$ .*

In the language of algebraic geometry we could also say that  $\mathcal{C}$  is isomorphic, as a complex algebraic variety, to  $\mathbb{A}_{\mathbb{C}}^1$  or to  $\mathbb{A}_{\mathbb{C}}^1 \setminus \{0\}$ .

*Proof.* As  $k$  is a generic value it implies that the curve  $\mathcal{C} = (P(x, y) = k)$  is smooth. If  $\mathcal{C}$  has one place at infinity then  $\mathcal{C}$  is homeomorphic to  $\mathbb{P}_{\mathbb{C}}^1 \setminus \{\infty\} \simeq \mathbb{C}$ . If  $\mathcal{C}$  has two places at infinity then  $\mathcal{C}$  is homeomorphic to  $\mathbb{P}_{\mathbb{C}}^1 \setminus \{0, \infty\} \simeq \mathbb{C}^*$ .  $\square$

In fact the homeomorphisms can be replaced by diffeomorphism and even algebraic isomorphism. In Lemma 7 we will treat the case  $\mathbb{C}$  and in Lemma 9 the case  $\mathbb{C}^*$ .

**2.4. Case of  $\mathcal{C}$  being homeomorphic to  $\mathbb{C}$ .** We recall the Abhyankar-Moh-Suzuki theorem [1, 2, 10]:

**Theorem 6.**

- (1) *Let  $t \mapsto (p(t), q(t))$  be an injective polynomial map from  $\mathbb{C}$  to  $\mathbb{C}^2$  such that the tangent vector  $(p'(t), q'(t))$  is never  $(0, 0)$ ; then  $\deg p$  divides  $\deg q$ , or  $\deg q$  divides  $\deg p$ .*
- (2) *Let  $\mathcal{C} = (P(x, y) = 0) \subset \mathbb{C}^2$  be an algebraic plane curve, non-singular and homeomorphic to  $\mathbb{C}$ ; then there exists an algebraic automorphism  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{C}}^2$  such that*

$$P \circ \Phi(x, y) = x.$$

The second statement is the usual form of the Abhyankar-Moh-Suzuki theorem, it is in fact a consequence of the first (see the proof below), for which a more general statement exists [2].

**Lemma 7.** *Let the curve  $\mathcal{C} = (P(x, y) = k)$  contain infinitely many integral points,  $k$  being a generic value. If  $\mathcal{C}$  is homeomorphic to  $\mathbb{C}$  then there exists  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  whose inverse has integral coefficients such that  $P \circ \Phi(x, y) = ax + b$ ,  $a, b \in \mathbb{Q}$ .*

*Proof.* The curve  $\mathcal{C}$  is a rational curve with one place at infinity and then admits a parametrization  $(p(t), q(t))$  by polynomials with rational coefficients. As  $k$  is a generic value, the curve  $\mathcal{C}$  is smooth and then  $(p'(t), q'(t))$  never vanishes. Hence the existence of  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{C}}^2$  comes from Theorem 6-(2). But here we ask the coefficients of  $\Phi$  to be rationals and those of  $\Phi^{-1}$  to be integers: we will apply Theorem 6-(1). Hence  $\deg p$  divides  $\deg q$  or  $\deg q$  divides  $\deg p$ . Suppose that  $\delta = \deg p > 0$  divides  $\deg q$  and write  $p(t) = a_{\delta}t^{\delta} + a_{\delta-1}t^{\delta-1} + \dots$  and  $q(t) = b_{\ell\delta}t^{\ell\delta} + b_{\ell\delta-1}t^{\ell\delta-1} + \dots$  with  $a_i, b_i \in \mathbb{Q}$  and  $\ell \geq 1$ . Write  $a_{\delta} = \frac{\alpha}{\beta}$  and  $b_{\ell\delta} = \frac{\alpha'}{\beta'}$ .

Set the algebraic automorphism of  $\text{Aut } \mathbb{A}_{\mathbb{Q}}^2$ ,

$$\Phi_1(x, y) = \left( x, \frac{1}{\alpha^{\ell}\beta'}y - \frac{\beta^{\ell}}{\alpha^{\ell}}\frac{\alpha'}{\beta'}x^{\ell} \right),$$

its inverse is

$$\Phi_1^{-1}(x, y) = \left( x, \alpha^{\ell}\beta'y + \alpha'\beta^{\ell}x^{\ell} \right),$$

whose coefficients are integers.

The composition with  $\Phi_1$  yields a parametrization of  $(P-k) \circ \Phi_1(x, y)$  given by  $(p'(t), q'(t)) = \Phi_1^{-1}(p(t), q(t))$  with  $q'(t) \in \mathbb{Q}[t]$  and  $\deg q' < \deg q$ . We repeat this process until one of  $p(t)$  or  $q(t)$  is a constant the other one being of degree 1 (this is possible because  $\mathcal{C}$  has no singular points). Then by the algebraic automorphism  $\Phi = \Phi_1 \circ \Phi_2 \circ \dots$ , whose inverse has integral coefficients, we get  $(P-k) \circ \Phi(x, y) = ax + b$ ,  $a, b \in \mathbb{Q}$ .  $\square$

As pointed out by the referee, a more direct proof of Lemma 7 can be given: by Hilbert's Theorem 90 there is no non-trivial  $\mathbb{Q}$ -form of the affine line so our curve isomorphic to  $\mathbb{A}_{\mathbb{C}}^1$  is isomorphic to  $\mathbb{A}_{\mathbb{Q}}^1$  over  $\mathbb{Q}$ .

**2.5. Case of  $\mathcal{C}$  being homeomorphic to  $\mathbb{C}^*$ .** We will need the classification over  $\mathbb{C}$  of polynomials with a generic fiber homeomorphic to  $\mathbb{C}^*$ , due to W. Neumann [7, §8].

**Theorem 8.** *If  $\mathcal{C} = (P(x, y) = k)$ ,  $k \neq 0$  a generic value, is homeomorphic to  $\mathbb{C}^*$  then there exists an algebraic automorphism  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{C}}^2$  such that*

$$P \circ \Phi(x, y) = x^p y^q + \beta$$

$$\text{or } P \circ \Phi(x, y) = x^p (yx^r + a_{r-1}x^{r-1} + \dots + a_0)^q + \beta,$$

with  $\beta \in \mathbb{C}$ ,  $p > 0$ ,  $q > 0$ ,  $\gcd(p, q) = 1$ ,  $r > 0$ ,  $a_0, \dots, a_{r-1} \in \mathbb{C}$  and  $a_0 \neq 0$ .

We will prove that only some special polynomials of the first type can have an infinite number of integral points.

The main result of this part is the following lemma.

**Lemma 9.** *If  $\mathcal{C} = (P(x, y) = k)$ ,  $k$  a generic value, is homeomorphic to  $\mathbb{C}^*$  and has an infinite number of integral points, then there exists  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  such that  $P \circ \Phi(x, y) = \alpha(x^2 - \ell y^2) + \beta$ ,  $\ell \in \mathbb{Z}$ ,  $\alpha \in \mathbb{Q}^*$ ,  $\beta \in \mathbb{Q}$ .*

*Proof.* By Theorem 8 we know that the polynomial  $P - \beta$  has exactly two absolute irreducible factors. For simplicity of the redaction we suppose in the following that  $\beta = 0$ . Notice that the curve  $(P = 0)$  is *not* the curve  $\mathcal{C}$ .

**First case :  $P$  is reducible in  $\mathbb{Q}[x, y]$ .**

Once again we will prove that from the automorphism  $\Phi$  of Theorem 8, that *a priori* has complex coefficients, we can construct an automorphism  $\Psi$  with rational coefficients and its inverse with integral coefficients. We write  $P = \alpha A^p B^q$  the decomposition into irreducible factors with  $A, B \in \mathbb{Q}[x, y]$ . Again for simplicity we suppose  $\alpha = 1$ . We will decompose the proof according to the cases of Theorem 8. In both cases we see that the curve  $(P = 0)$  has a non-singular irreducible component homeomorphic to  $\mathbb{C}$  (the one sent by  $\Phi^{-1}$  to  $(x = 0)$ ). This component homeomorphic to  $\mathbb{C}$  is either  $(A = 0)$  or  $(B = 0)$ ; say it is  $(A = 0)$ . Then, as  $A \in \mathbb{Q}[x, y]$ , by the version of Abhyankar-Moh-Suzuki theorem used as in Lemma 7 above, there exists  $\Psi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$ , whose inverse has integral coefficients, such that:  $A \circ \Psi(x, y) = ax + b$ , this implies :

$$P \circ \Psi(x, y) = (ax + b)^p Q(x, y)^q.$$

**Sub-case  $P \circ \Phi(x, y) = x^p y^q$ .**

Then  $(Q(x, y) = 0)$  is non-singular, homeomorphic to  $\mathbb{C}$  and the intersection multiplicity with the line  $(ax + b = 0)$  is 1. Then if  $(p(t), q(t))$  is a polynomial parametrization of  $(Q(x, y) = 0)$  we have  $\deg p = 1$ . Then as in the proof of Lemma 7 by algebraic automorphisms of type  $(x, y) \mapsto (\alpha x, \beta y - \gamma x^\ell)$  whose inverse have integral coefficients, we can suppose that  $q(t)$  is a constant. Notice that such automorphisms preserve vertical lines.

Then  $Q(x, y)$  becomes  $cy + d$ ,  $c, d \in \mathbb{Q}$ , while  $ax + b$  remains unchanged. Then we have found  $\Psi' \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  whose inverse has integral coefficients such that:

$$P \circ \Psi'(x, y) = (ax + b)^p (cy + d)^q.$$

**Sub-case  $P \circ \Phi(x, y) = x^p (yx^r + a_{r-1}x^{r-1} + \dots + a_0)^q$ .**

$P \circ \Phi(x, y) = x^p(yx^r + a_{r-1}x^{r-1} + \dots + a_0)^q$  is algebraically equivalent to  $P \circ \Psi(x, y) = (ax + b)^p Q(x, y)^q$  by the algebraic automorphism  $\Phi \circ \Psi^{-1}$ . Moreover  $\Phi \circ \Psi^{-1}$  should send  $x$  to  $ax + b$ . Then  $\Phi \circ \Psi^{-1}$  is the composition of algebraic automorphisms of type  $(x, y) \mapsto (ax + b, y)$  and  $(x, y) \mapsto (\alpha x, \beta y - \gamma x^\ell)$ . This implies that the degree in the variable  $y$  remains unchanged. Then  $\deg_y Q(x, y) = \deg_y(yx^r + a_{r-1}x^{r-1} + \dots + a_0) = 1$ . Then  $Q(x, y) = q_1(x)y + q_2(x)$ . Due to the asymptotic branches we have  $q_1(x, y) = (ax + b)^r$ . And by algebraic automorphisms whose inverse have integral coefficients of type  $(x, y) \mapsto (\alpha x, \beta y - \gamma x^\ell)$  we can suppose  $\deg q_2 < r$ . Then we have found  $\Psi' \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  with an inverse having integral coefficients such that:

$$P \circ \Psi'(x, y) = (ax + b)^p (y(ax + b)^r + b_{r-1}x^{r-1} + \dots + b_0)^q,$$

$b_0, \dots, b_{r-1} \in \mathbb{Q}$ ,  $b_0 \neq 0$ .

**Conclusion for both sub-cases.**

Now the curve  $(P \circ \Psi'(x, y) = k)$  has a finite number of integral points since the branches at infinity are asymptotic to horizontal or vertical lines (with equation  $(ax + b = 0)$ ,  $(cy + d = 0)$  in the first case and  $(ax + b = 0)$ ,  $(y = 0)$  in the second case ; they correspond to the two points at infinity  $(0 : 1 : 0)$  and  $(1 : 0 : 0)$ ). Now as  $\Psi'^{-1}$  has integral coefficients, an integral point  $(m, n) \in (P(x, y) = k) \cap \mathbb{Z}^2$  is sent to an integral point  $\Psi'^{-1}(m, n) \in (P \circ \Psi'(x, y) = k) \cap \mathbb{Z}^2$  it implies that  $\mathcal{C} = (P(x, y) = k)$  also have a finite number of integral points.

**Second case :  $P$  is irreducible in  $\mathbb{Q}[x, y]$ .**

We still apply Theorem 8. Then by Lemma 10 below it implies that there exist  $C, D \in \mathbb{Q}[x, y]$ ,  $\ell \in \mathbb{Z}$  such that  $P = C^2 - \ell D^2$ . Then  $P = (C - \sqrt{\ell}D)(C + \sqrt{\ell}D)$  is the decomposition into irreducible factors.

**Sub-case  $P \circ \Phi(x, y) = x^p y^q$ .**

Then by Lemma 10 we know that  $p = 1$ ,  $q = 1$ . And equivalently there exists  $\Phi' \in \text{Aut } \mathbb{A}_{\mathbb{C}}^2$  such that  $P \circ \Phi'(x, y) = (x - \sqrt{\ell}y)(x + \sqrt{\ell}y)$ . Then  $P \circ \Phi'(x, y) = (C^2 - \ell D^2) \circ \Phi'(x, y) = (x - \sqrt{\ell}y)(x + \sqrt{\ell}y)$ . We may suppose that  $(C - \sqrt{\ell}D) \circ \Phi'(x, y) = (x - \sqrt{\ell}y)$  and  $(C + \sqrt{\ell}D) \circ \Phi'(x, y) = (x + \sqrt{\ell}y)$ , by addition and subtraction we get  $C \circ \Phi'(x, y) = x$  and  $D \circ \Phi'(x, y) = y$ . Then  $(CD) \circ \Phi'(x, y) = xy$ , with  $C, D \in \mathbb{Q}[x, y]$ . As in the first case above we are now enable to find  $\Psi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  such that  $C \circ \Psi(x, y) = x$ ,  $D \circ \Psi(x, y) = y$  and  $CD \circ \Psi(x, y) = xy$ . Now  $P \circ \Psi(x, y) = (C^2 - \ell D^2) \circ \Psi(x, y) = x^2 - \ell y^2$ .

**Sub-case  $P \circ \Phi(x, y) = x^p(yx^r + a_{r-1}x^{r-1} + \dots + a_0)^q$ .**



Again  $p = 1$ ,  $q = 1$ , and we may suppose that  $(C - \sqrt{\ell}D) \circ \Phi(x, y) = x$ . We denote  $Q = y \circ \Phi^{-1}(x, y)$  i.e.  $Q \circ \Phi(x, y) = y$ . Then

$$\begin{aligned} (C - \sqrt{\ell}D)(C + \sqrt{\ell}D) &= P \\ &= x(yx^r + a_{r-1}x^{r-1} + \cdots + a_0) \circ \Phi^{-1}(x, y) \\ &= (C - \sqrt{\ell}D)(Q(C - \sqrt{\ell}D)^r + \cdots) \end{aligned}$$

Then

$$(C + \sqrt{\ell}D) = (Q(C - \sqrt{\ell}D)^r + \cdots).$$

But as  $C, D \in \mathbb{Q}[x, y]$  we have  $d = \deg(C + \sqrt{\ell}D) = \deg(C - \sqrt{\ell}D)$  and we get  $d = \deg Q + rd$ . As  $r \geq 1$  we get  $\deg Q = 0$  which is in contradiction with the definition of  $Q$ . Then this sub-case does not occur.  $\square$

It remains to prove the following technical lemma used in the second case above.

**Lemma 10.** *Let  $P \in \mathbb{Q}[x, y]$  such that  $P = \alpha A^p B^q$ , with  $\gcd(p, q) = 1$  and with  $\alpha \in \mathbb{Q}^*$ ,  $A, B \in \overline{\mathbb{Q}}[x, y]$  monic and irreducible (that is to say  $P$  admits exactly two absolute irreducible factors). Then either  $A, B \in \mathbb{Q}[x, y]$  or  $p = 1$ ,  $q = 1$  and there exist  $C, D \in \mathbb{Q}[x, y]$ ,  $\ell \in \mathbb{Z}$  non-square such that  $P = \alpha(C^2 - \ell D^2)$ .*

The following proof is due to Pierre Dèbes.

*Proof.* Let  $a_{i,j} \in \overline{\mathbb{Q}}$  be the coefficients of  $A$ . Let  $n$  be the degree of the finite extension  $\mathbb{Q}((a_{i,j}))/\mathbb{Q}$ . Then there exist exactly  $n$  distinct conjugates of  $A$ . But for all  $\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ ,  $\sigma(A) \in \{A, B\}$ . Then  $A$  has at most two distinct algebraic conjugates. Thus  $n = 1$  or  $n = 2$ . If  $A \notin \mathbb{Q}[x, y]$  then there exists  $a_{i_0, j_0} \notin \mathbb{Q}$  and a  $\sigma_0 \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$  such that  $\sigma_0(A) = B$ . Then  $n = 2$  so that the extension  $\mathbb{Q}((a_{i,j}))/\mathbb{Q}$  is quadratic. Moreover  $p = q = 1$ . This implies the existence of a non-square integer  $\ell$  such that  $A, B \in \mathbb{Q}(\sqrt{\ell})[x, y]$ . Now if we write  $A = C + \sqrt{\ell}D$ ,  $C, D \in \mathbb{Q}[x, y]$  then its algebraic conjugate is  $B = C - \sqrt{\ell}D$ .  $\square$

Lemma 7 and Lemma 9 imply Theorem 1 of the introduction. We now prove Corollary 2.

*Proof.* Suppose that the polynomial  $P(x, y)$  is algebraically equivalent to the polynomial  $x$ . We have constructed an automorphism  $\Phi \in \text{Aut } \mathbb{A}_{\mathbb{Q}}^2$  such that:  $\Phi(0, 0) = (0, 0)$ ,  $\Phi^{-1}$  has integral coefficients and  $P \circ \Phi(x, y) = ax + b$ ,  $a, b \in \mathbb{Q}$ . Let  $(m_j, n_j)_{j \in \mathbb{N}} \in \mathbb{Z}^2$  integral points of  $C = (P(x, y) = k)$ . Then  $(m'_j, n'_j) = \Phi^{-1}(m_j, n_j) \in \mathbb{Z}^2$  are solutions of

$ax + b = k$  (so that  $m'_j$  is independent of  $j$ ). Then for each  $i \in \mathbb{N} \setminus \{0\}$ , and all  $j \in \mathbb{N}$ ,  $(i \cdot m'_j, i \cdot n'_j)$  is a solution of  $ax + b = k_i$ , for some  $k_i \in \mathbb{Q}$ . As  $\Phi$  has coefficients in  $\mathbb{Q}$  and  $\Phi(0, 0) = (0, 0)$  there exists an arithmetical sequence  $p_i \in \mathbb{N}$ , such that  $\Phi(p_i \cdot m'_j, p_i \cdot n'_j) \in \mathbb{Z}^2$ . By construction for each  $p_i$  and all  $j \in \mathbb{N}$ ,  $P(\Phi(p_i \cdot m'_j, p_i \cdot n'_j)) = k_{p_i}$ .

If  $P$  is algebraically equivalent to  $\alpha(x^2 - \ell y^2) + \beta$  then a similar proof holds as integral solutions  $(m'_j, n'_j)$  to  $\alpha(x^2 - \ell y^2) + \beta = k$  give rises for each  $i \in \mathbb{N} \setminus \{0\}$  to integral solutions  $(i \cdot m'_j, i \cdot n'_j)$  to  $\alpha(x^2 - \ell y^2) + \beta = k_i$  for some  $k_i \in \mathbb{Q}$ .  $\square$

### 3. NUMBER OF INTEGRAL POINTS ON POLYNOMIAL CURVES

We want to estimate the number of integral points on a “polynomial curve” parametrized by polynomial equations  $(p(t), q(t))$ . An integral point of our generic –hence non-singular– curve  $\mathcal{C}$  corresponds to a rational parameter  $t \in \mathbb{Q}$ . It remains to count the number of parameters  $t$  that yield to an integral value such that  $|p(t)| \leq B$  and  $|q(t)| \leq B$ .

**Lemma 11.** *Let  $p(t) = a_d t^d + a_{d-1} t^{d-1} + \dots + a_0 \in \mathbb{Q}[t]$ ,  $a_d > 0$ . Let  $\sigma = -\frac{a_{d-1}}{da_d}$ . There exists  $B_0 > 0$  such that for all  $B \geq B_0$ , if we set*

$$t_+ = \left(\frac{B}{a_d}\right)^{\frac{1}{d}} + \sigma + \frac{1}{2}, \quad t_- = -\left(\frac{B}{a_d}\right)^{\frac{1}{d}} + \sigma - \frac{1}{2},$$

then

$$\text{for all } t \geq t_+, \quad |p(t)| > B \quad \text{and} \quad \text{for all } t \leq t_-, \quad |p(t)| > B.$$

A similar result holds if  $a_d < 0$ .

*Proof.* Set  $\varepsilon = \frac{1}{2}$  and write  $t = s + \sigma + \varepsilon$ ; we look at the asymptotic behavior for  $p(t)$  when  $t$  (and  $s$ ) is large.

$$\begin{aligned} p(t) &= p(s + \sigma + \varepsilon) \\ &= a_d(s + \sigma + \varepsilon)^d + a_{d-1}(s + \sigma + \varepsilon)^{d-1} + \dots \\ &= a_d s^d + (da_d(\sigma + \varepsilon) + a_{d-1})s^{d-1} + o(s^{d-1}) \\ &= a_d s^d + da_d \varepsilon s^{d-1} + o(s^{d-1}). \end{aligned}$$

For  $s = \left(\frac{B}{a_d}\right)^{\frac{1}{d}}$  then  $a_d s^d = B$  we have

$$\begin{aligned} p(t_+) &= p(s + \sigma + \varepsilon) \\ &= B \cdot \left(1 + \varepsilon d \frac{1}{s} + o\left(\frac{1}{s}\right)\right). \end{aligned}$$

Then for all sufficiently large  $B$  (such that  $s > 0$  is large enough) we have  $p(t_+) \geq B \left(1 + \frac{\varepsilon d}{2s}\right)$  then  $p(t_+) > B$ . Now the function  $t \mapsto p(t)$  is an increasing function for sufficiently large  $t$ . Then for all sufficiently large  $B$ : if  $t \geq t_+$  then  $p(t) \geq p(t_+) > B$ .

Now

$$\begin{aligned} p(t_-) &= p(-s + \sigma - \varepsilon) \\ &= (-1)^d B \cdot \left(1 + \varepsilon d \frac{1}{s} + o\left(\frac{1}{s}\right)\right). \end{aligned}$$

Then for all sufficiently large  $B$ ,  $|p(t_-)| > B$ . And again if  $t < t_-$  then  $|p(t)| \geq |p(t_-)| > B$ .  $\square$

For a polynomial  $p(t) \in \mathbb{Q}[t]$  in one variable we define:

$$M(p, B) = \{t \in \mathbb{Q} \mid p(t) \in \mathbb{Z} \text{ and } |p(t)| \leq B\}.$$

**Lemma 12.** *Let  $p(t) = \frac{1}{b}(a_d t^d + \dots + a_0) \in \mathbb{Q}[t]$ ,  $a_0, \dots, a_d, b \in \mathbb{Z}$ ,  $\gcd(a_0, \dots, a_d, b) = 1$  and  $a_d > 0$ ,  $b > 0$ . There exists  $B_0 > 0$  such that for all  $B \geq B_0$  we have*

$$M(p, B) \leq 2a_d^{1-\frac{1}{d}} b^{\frac{1}{d}} B^{\frac{1}{d}} + 2.$$

*Proof.* If  $t = \frac{\alpha}{\beta} \in \mathbb{Q}$  with  $\gcd(\alpha, \beta) = 1$  and  $p(\frac{\alpha}{\beta}) = k \in \mathbb{Z}$  then it is well-known that  $\beta$  divides  $a_d$ . Then such  $t$  belongs to  $\frac{1}{a_d}\mathbb{Z}$ . Let  $B_0$  be as in Lemma 11. Again by Lemma 11 if  $t > 0$  and  $|p(t)| \leq B$  then  $t < t_+ = \left(\frac{B}{a_d/b}\right)^{\frac{1}{d}} + \sigma + \frac{1}{2}$ . If  $t < 0$  and  $|p(t)| \leq B$  then  $|t| < |t_-| = -t_- = \left(\frac{B}{a_d/b}\right)^{\frac{1}{d}} - \sigma + \frac{1}{2}$ . Now the cardinal of  $\frac{1}{a_d}\mathbb{Z} \cap [t_-, t_+]$  is less than  $a_d|t_+| + a_d|t_-| + 1 = 2a_d \left(\frac{B}{a_d/b}\right)^{\frac{1}{d}} + \sigma - \sigma + 2 = 2a_d \left(\frac{bB}{a_d}\right)^{\frac{1}{d}} + 2$ .  $\square$

Of course if  $p(t)$  is a monic polynomial with integral coefficients, i.e.  $b = 1$ ,  $a_d = 1$ , then  $M(p, B) \leq 2B^{\frac{1}{d}} + 2$ . For example if  $p(t) = t^d$  then  $M(p, B) = 2B^{\frac{1}{d}} + 1$ . The following example shows that the bound of Lemma 12 is the best one (at least for  $a_d = 1$ ).

*Example 13.* Let  $p(t) = t^d - 1$  where  $d$  is an even number. There exist infinitely many integers  $B$  such that:

$$M(p, B) > 2B^{\frac{1}{d}} + 1.$$

In fact for  $k$  any positive integer, set  $B_k = p(k) = k^d - 1$ . Then as  $d$  is even for all  $t \in [-k, k]$  we have  $t^d - 1 \leq k^d - 1 = B_k$  then  $M(p, B_k) = 2k + 1 > 2(k^d - 1)^{\frac{1}{d}} + 1 = 2B_k^{\frac{1}{d}} + 1$ .

We apply these computations to the situation of our curves.

Let  $P(x, y) \in \mathbb{Q}[x, y]$  be irreducible, let  $\mathcal{C} = (P(x, y) = 0)$ . Then  $\mathcal{C}$  is a *polynomial curve* if it admits a polynomial parametrization  $(p(t), q(t))$ ,  $p(t), q(t) \in \mathbb{Q}[t]$ . Equivalently  $\mathcal{C}$  is a rational curve with one place at infinity. Moreover  $\deg P = \max(\deg p, \deg q)$ . We will suppose  $\deg P = \deg p$  and we write  $p(t) = \frac{1}{b}(a_d t^d + \cdots + a_0)$  as before.

**Lemma 14.** *Let  $\mathcal{C}$  be a polynomial curve. Suppose  $\deg P = d = \deg p$ ,  $p(t) = \frac{1}{b}(a_d t^d + \cdots + a_0)$ . Then there exists  $B_0 > 0$  such that for all  $B \geq B_0$ :*

$$N(\mathcal{C}, B) \leq 2a_d^{1-\frac{1}{d}} b^{\frac{1}{d}} B^{\frac{1}{d}} + \frac{(d-1)(d-2)}{2} + 2.$$

The term  $\frac{(d-1)(d-2)}{2}$  comes from the number of singular points; for non-singular curves we get the bound of Theorem 3. Moreover by inspection of the proofs, for all  $\varepsilon > 0$  we can change the bound to be  $N(\mathcal{C}, B) \leq 2a_d^{1-\frac{1}{d}} b^{\frac{1}{d}} B^{\frac{1}{d}} + \frac{(d-1)(d-2)}{2} + 1 + \varepsilon$ .

*Proof.* An algebraic curve of degree  $d$  must have less than  $\frac{(d-1)(d-2)}{2}$  singular points, see [5, p.117]. The other integral points  $(p(t), q(t))$  of  $\mathcal{C}$  correspond to rational parameters  $t$ , see (see [9] or [5, p.160]). Now we apply Lemma 12.  $\square$

Example 13 gives a curve parametrized by  $(p(t), t)$  that proves that the bound of Lemma 14 and Theorem 3 is asymptotically sharp.

## REFERENCES

- [1] E. Artal-Bartolo, *Une démonstration géométrique du théorème d'Abhyankar-Moh*. J. Reine Angew. Math. 464 (1995), 97–108.
- [2] S.S. Abhyankar, T.T. Moh, *Embeddings of the line in the plane*. J. Reine Angew. Math. 276 (1975), 148–166.
- [3] M. Ayad, *Sur les polynômes  $f(X, Y)$  tels que  $K[f]$  est intégralement fermé dans  $K[X, Y]$* . Acta Arith. 105 (2002), 9–28.
- [4] E. Bombieri, J. Pila, *The number of integral points on arcs and ovals*. Duke Math. J. 59 (1989), 337–357.
- [5] W. Fulton, *Algebraic curves*. Addison-Wesley, reprint of 1969.
- [6] D.R. Heath-Brown, *The density of rational points on curves and surfaces*. Ann. of Math. 155 (2002), 553–595.
- [7] W.D. Neumann, *Complex algebraic plane curves via their links at infinity*. Invent. Math. 98 (1989), 445–489.
- [8] Nguyen Van Chau, *Integer points on a curve and the plane Jacobian problem*. Ann. Polon. Math. 88 (2006), 53–58.
- [9] D. Poulakis, E. Voskos, *On the practical solution of genus zero Diophantine equations*. J. Symbolic Comput. 30 (2000), 573–582.
- [10] L. Rudolph, *Embeddings of the line in the plane*. J. Reine Angew. Math. 337 (1982), 113–118.

- [11] J.-P. Serre, *Lectures on the Mordell-Weil theorem*. Aspects of Mathematics, Vieweg 1997.
- [12] C. L. Siegel, *Über einige Anwendungen diophantischer Approximation*, Abh. Preuss. Akad. Wiss. Phys.-Math. Kl. 1929 ; Gesammelte Abhandlungen, Vol. I, Springer, Berlin, 1966, 209–266.
- [13] Y. Walkowiak, *Théorème d'irréductibilité de Hilbert effectif*. Acta Arith. 116 (2005), 343–362.

LABORATOIRE PAUL PAINLEVÉ, MATHÉMATIQUES, UNIVERSITÉ DE LILLE 1,  
59655 VILLENEUVE D'ASCQ, FRANCE.

*E-mail address:* `Arnaud.Bodin@math.univ-lille1.fr`