Shukhrat Alladustov
sh.alladustov@gmail.com

A STRONG BOUND FOR THE NUMBER OF SOLUTIONS OF THUE EQUATIONS

Master's thesis, July 2015

Supervisor: Prof. Yuri Bilu

Leiden University

University of Bordeaux

Contents

Chapter 1

Introduction

A Thue equation is a Diophantine equation of the form

$$F(x, y) = m \tag{1.0.1}$$

where $F(x, y) \in \mathbb{Z}[x, y]$ is an irreducible homogeneous polynomial in two variables of degree $r \geqslant 3$ and $m \in \mathbb{Z}$ is an integer number. It is named after Axel Thue, who in 1909 has shown that equations of this kind has only finitely many integer solutions. We will discuss the proof later on. Other approaches to the problem were performed also by Th.Skolem and A.Baker. Th.Skolem used p-adic power series, under weak restrictions on F and A.Baker used lower bounds for linear forms in logarithms.

Later, there have been put another question: *Does there exist an upper bound for the number of solutions not depending on the polynomial F but only on the degree r and integer m ?* The first results were obtained by C.L.Siegel for a specific case when $r = 3$ and $F(x, y) = ax^3 - by^3 \in \mathbb{Z}$.

But, the question was answered by J.H.Evertse ([2]) in the general case. Evertse's method was based on the ideas of Siegel and Baker to reduce the problem to the specific cases which are more comfortable to work with. Firstly, he considered the equation of the form $|ax^r - by^r| = c$ where he obtained $2r^{w(c)} + 4$ as an upper bound for the number of solutions with $w(c)$ is an integer number depending on $c \in \mathbb{N}$. Also, the equation $|F(x, y)| = 1$ with F is a cubic polynomial with positive discriminant was investigated. As a result in this case, it was obtained that the equation has at most twelve solutions. Finally, some approaches in equations in number fields and techniques of Pade approximation were used to obtain the following upper bound for the number of primitive solutions of the equation (1.0.1)

$$7^{15\left(\binom{r}{3}+1\right)^2} + 6 \times 7^{2\binom{r}{3}(t+1)}$$

where t is the number of prime divisors of m.

In our work we consider a particular type of the Thue equations

$$|F(x, y)| = 1 \tag{1.0.2}$$

under the same conditions for $F(x, y) \in \mathbb{Z}[x, y]$ as above and we deal with the problem of finding an upper bound for the number of integer solutions, we describe the method of E.Bombieri and W.M.Schmidt.

It is clear that each integer solution of $(1.0.2)$, if exists, consist of coprime numbers (x, y). Also, if (x, y) is a solution then $(-x, -y)$ is so too. Therefore for simplicity we consider only one of these solutions, more precisely throughout the thesis we only work with solutions such that $y \geqslant 0$ without mentioning.

As a main result we derive the following relation for the number of solution of $(1.0.2)$.

Theorem 1.0.1. *Let $F(x, y) \in \mathbb{Z}[x, y]$ be an irreducible homogeneous polynomial of degree $r \geqslant 3$. Then for $r > c$, the number of integer solutions of the equation*

$$|F(x, y)| = 1$$

which solutions (x, y) and $(-x, -y)$ are regarded the same, does not exceed $215r$, where $c > 0$ is an absolute constant.

Firstly, we classify the polynomials satisfying the above conditions according to the number of solutions of $(1.0.2)$. Then it allows us to consider one representative from each class instead of working with all polynomials. We also show that as such a representative we can take the polynomials whose leading coefficient is 1, for this purpose we construct the so-called auxiliary forms.

Also, we derive a decomposition of $\mathbb{Z}^2$ into finitely many smaller sets $\mathbb{Z}^2 = \cup_{i=0}^{n} \mathbb{Z}_i$, using this decomposition we can only deal with polynomials whose discriminants are big enough. Here, a key point is the fact that the number of solutions of $(1.0.2)$ does not exceed the sum of the numbers of solutions in each $\mathbb{Z}_i$, $i = 0, ..., n$.

The proof of the main theorem is based on two steps, firstly we deal with finding a bound for the number of solutions (x, y) whose heights $H(x, y) = \max(|x|, |y|)$ is not less than some fixed number $M > 0$ and then for the solutions (x, y) with $y < M$, which we call large and small solutions respectively. Obviously, large and small solutions cover all the solutions.

In the first case, we need some additional theorems. We state *the Lewis and Mahler estimation* and *the Thue-Siegel principle* in specific cases ([5]) with proofs. Also, we develop the so-called *Strong gap principle.* Using these approaches we conclude that $3r$ is an upper bound provided r sufficiently large.

In the second case, we use an auxiliary polynomial and some tricks to show that the bound is $212r$.

Actually, this result can be applied to obtain a bound for the number of solutions of $(1.0.1)$ too, more precisely if the number of solutions of (1.0.2) does not exceed N_r then $N_r r^t$ is an upper bound for the number of solutions of $(1.0.1)$, where t is the number of prime factors of integer m.

Chapter 2

Additional theorems and facts

2.1 The Thue theorem about finiteness of solutions

As we mentioned before, we first show that (1.0.1) has only finitely many solutions under fixed coefficients. For this we need the following theorem due to Thue.

Theorem 2.1.1 (Thue)**.** *Let* $\alpha \in \mathbb{C}$ *be an algebraic integer number of degree* $r \geqslant 2$ *, then for any* $\varepsilon > 0$ *and* $c > 0$ *the inequality*

$$\left|\alpha - \frac{p}{q}\right| < \frac{c}{q^{\frac{r}{2}+1+\varepsilon}}$$

can have only finitely many solutions with $p \in \mathbb{Z}$ *and* $q \in \mathbb{N}$ *.*

This result will be directly used in the proof of the following theorem.

Theorem 2.1.2. *Let*

$$F(x, y) = a_0 x^r + a_1 x^{r-1} y + ... + a_r y^r \in \mathbb{Z}[x, y], \ r \geqslant 3$$

be an irreducible polynomial with $a_0 > 0$ *and let* $m \in \mathbb{Z}$ *be any integer, then the number of integer solutions of the equation*

$$F(x, y) = m \tag{2.1.1}$$

is finite.

Proof. We consider the following two cases separately:

1) $m = 0$ then the equation has no integer solutions except trivial one, because F is an irreducible homogeneous polynomial of degree $r \geqslant 3$.

2) Suppose $m \neq 0$. We consider the polynomial $f(z) = F(z, 1)$ then we have that $f(z) = a_0 z^r + a_1 z^{n-1} + ... + a_{r-1} z + a_r \in \mathbb{Z}[z]$ is also an irreducible polynomial in one variable, let $\alpha_1, ..., \alpha_n$ be its roots. Then we have the factorization

$$F(x, y) = y^r f\left(\frac{x}{y}\right) = a_0 \left(x - \alpha_1 y\right) ... \left(x - \alpha_r y\right)$$

So, solutions of (2.1.1) satisfy the relation

$$|x - \alpha_1 y| ... |x - \alpha_r y| = \frac{|m|}{a_0} \tag{2.1.2}$$

Product of r numbers equals to $\frac{|m|}{a_0}$ therefore there exists k, $1 \leqslant k \leqslant r$ with

$$|x - \alpha_k y| \leqslant \sqrt[r]{\frac{|m|}{a_0}} \tag{2.1.3}$$

Take $\gamma > 0$ such that $\gamma < \min |\alpha_i - \alpha_j|, i \neq j$, it is possible since $\alpha_1, ..., \alpha_r$ are the different roots of the irreducible polynomial. The inequality (2.1.3) then implies

$$|x - \alpha_i y| = |(\alpha_k - \alpha_i)y + (x - \alpha_k y)| \geqslant |\alpha_k - \alpha_i||y| - |x - \alpha_k y| \geqslant \gamma|y| - \sqrt[r]{\frac{|m|}{a_0}} \tag{2.1.4}$$

for all $i = 1, ..., r,\ i \neq k$.

It is clear that if y is bounded then the number of solutions is finite. Therefore, throwing away finitely many solutions, if needed, one can assume that

$$|y| > \frac{2\sqrt[r]{\frac{|m|}{a_0}}}{\gamma}$$

Then from (2.1.4) we obtain that

$$|x - \alpha_i y| > \frac{1}{2}\gamma|y|, i = 1, ..., r, i \neq k.$$

Thus, for the product of $r - 1$ components the following relation holds

$$\prod_{i \neq k} |x - \alpha_i y| > \left(\frac{1}{2}\gamma|y|\right)^{r-1}$$

Now, applying this fact in (2.1.2) we obtain

$$|x - \alpha_k y| < \frac{c}{|y|^{r-1}},\ c = \frac{|m|}{a_0\left(\frac{1}{2}\gamma\right)^{r-1}}$$

or equivalently, for almost all the solutions (x, y) we have that

$$\left|\alpha_k - \frac{x}{y}\right| < \frac{c}{|y|^r} \tag{2.1.5}$$

But, by the Thue theorem the last inequality has only finitely many solutions $x, y \in \mathbb{Z}$, $y \neq 0$ for any integer $n \geqslant 2$ and $c > 0$. We have that all, but finitely many integer solutions of (2.1.1) satisfy the relation (2.1.5), which implies that the number of solutions cannot be infinite.

Corollary 2.1.1. *Let F be as in the definition of the last theorem. Then the equation*

$$|F(x, y)| = 1 \tag{2.1.6}$$

has only finitely many solutions.

Proof. The proof follows from considering the two equations

$$F(x, y) = \pm 1.$$

Each of these equations is a particular case of the Theorem 2.1.2 with $m = 1$ and $m = -1$, therefore the number of solutions of 2.1.6 is also finite. □

2.2 Discriminant of the polynomial

Let $F(x,y) = a_0x^r + a_1x^{r-1}y + ... + a_ry^r$ be an irreducible homogeneous polynomial with integer coefficients, of degree $r \geqslant 3$. Suppose $\alpha_1, ..., \alpha_r$ be the roots of the polynomial

$$f(x) = F(x,1) = a_0x^r + a_1x^{r-1} + ... + a_r$$

then we have

$$F(x,y) = a_0(x - \alpha_1 y)...(x - \alpha_r y)$$

Now, we define the discriminant $D(F)$ of the polynomial F using the determinant of the Vandermonde matrix

$$V = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & ... & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & ... & \alpha_2^{n-1} \\ ... & ... & ... & ... & ... \\ 1 & \alpha_r & \alpha_r^2 & ... & \alpha_r^{n-1} \end{pmatrix}$$

as $D(F) = a_0^{2r-2}(\det V)^2$, more precisely

$$D(F) = a_0^{2r-2} \prod_{i<j} (\alpha_i - \alpha_j)^2 \tag{2.2.1}$$

Let us see what happens if we change (x,y) by $A(x,y)$, where A is a matrix with non-zero determinant of the form $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z})$ and $A(x,y) = (ax+by, cx+dy)$. Consider the polynomial $F_A(x,y) = F(ax+by, cx+dy)$. Then

$$F(ax+by, cx+dy) = b_0x^r + b_1x^{r-1}y + ... + b_ry^r$$

for some integers $b_0, ..., b_r$.

Evaluating F_A at $(1,0)$ we can see that

$$b_0 = F(a,c) \tag{2.2.2}$$

Consider the equation

$$F_A(x,1) = a_0(ax+b)^r + a_1(ax+b)^{r-1}(cx+d) + ... + a_r(cx+d)^r = 0$$

We note that we can assume $cx+d \neq 0$, otherwise the equation has a solution only if $ax+b=0$, but this contradicts to the assumption that $det(A) \neq 0$. So, dividing both sides of the equality by non-zero term $(cx+d)^r$ we obtain

$$a_0\left(\frac{ax+b}{cx+d}\right)^r + a_1\left(\frac{ax+b}{cx+d}\right)^{r-1} + ... + a_r = F\left(\left(\frac{ax+b}{cx+d}\right), 1\right) = 0 \tag{2.2.3}$$

According to our assumption, the equation $F(z,1) = 0$ has solutions $\alpha_1, ..., \alpha_r$, then solving the equations

$$\frac{ax+b}{cx+d} = \alpha_i$$

for $i = 1, ..., r$ we obtain that the roots of (2.2.3) are

$$\beta_1 = \frac{\alpha_1 - b}{a - \alpha_1 c}, ..., \beta_r = \frac{\alpha_r - b}{a - \alpha_r c}$$

Then for the difference of these roots we have

$$\beta_i - \beta_j = \frac{\alpha_i - b}{a - \alpha_i c} - \frac{\alpha_j - b}{a - \alpha_j c} = det(A)\frac{\alpha_j - \alpha_i}{(a - \alpha_i c)(a - \alpha_j c)}$$

and we obtain the following relation

$$\prod_{i<j}(\beta_i - \beta_j)^2 = det(A)^{r(r-1)}\prod_{i<j}\left(\frac{\alpha_j - \alpha_i}{(a - \alpha_i c)(a - \alpha_j c)}\right)^2 = det(A)^{r(r-1)}\frac{\prod_{i<j}(\alpha_j - \alpha_i)^2}{F(a,c)^{r(r-1)}}$$

Taking into account the fact (2.2.2) we get

$$D(F_A) = b_0^{2r-2}\prod_{i<j}(\beta_i - \beta_j)^2 = b_0^{2r-2}\cdot a_0^{2r-2}\cdot \det(A)^{2r-2)}\frac{\prod_{i<j}(\alpha_j - \alpha_i)^2}{F(a,c)^{r(r-1)}} = \det(A)^{r(r-1)}D(F)$$

Thus, the change of variables $(x, y) \to A(x, y)$ gives the following relation between the discriminants of the polynomials F_A and F

$$D(F_A) = \det(A)^{r(r-1)}D(F) \qquad (2.2.4)$$

2.3 Simplification of the form F and normalization

2.3.1 Decomposition for $\mathbb{Z}^2$

For simplicity we need to consider the forms satisfying certain properties. More precisely, we want to work with polynomials which have quite big discriminants.

Lemma 2.3.1. *For any prime p we have the following decomposition of $\mathbb{Z}^2$,*

$$\mathbb{Z}^2 = \bigcup_{j=0}^{p} A_j\mathbb{Z}^2$$

where $A_0 = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$, $A_j = \begin{pmatrix} 0 & -1 \\ p & j \end{pmatrix}$, $j = 1, ..., p$

Proof. Indeed, take any pair $(x, y) \in \mathbb{Z}^2$. If $x \equiv 0(\mod p)$, then $x = p \cdot x'$, some $x' \in \mathbb{Z}$, then

$$A_0\begin{pmatrix} x' \\ y \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y \end{pmatrix} = (px', y) = (x, y)$$

Now let $x \not\equiv 0(\mod p)$. We again consider two cases:

1) $y \equiv 0(\mod p)$, then $y = py'$ for some $y' \in \mathbb{Z}$ then taking $(y' + x, -x) \in \mathbb{Z}^2$ we obtain

$$A_p(y' + x, -x) = \begin{pmatrix} 0 & -1 \\ p & p \end{pmatrix}\begin{pmatrix} y' + x \\ -x \end{pmatrix} = (x, p(x + y') - px) = (x, y)$$

2) $y \not\equiv 0(\mod p)$. Then we want to find $(x', y') \in \mathbb{Z}^2$ and $j, 1 \leqslant j \leqslant p-1$ such that

$$A_j(x', y') = \begin{pmatrix} 0 & -1 \\ p & j \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = (-y', px' + jy') = (x, y)$$

Take $y' = -x$, then it is enough to show that the equation $px' - xj = y$ has a solution $x', j \in \mathbb{Z}$, $1 \leqslant j \leqslant p-1$. It is clear that $-xj = y(\mod p)$ has such a solution $1 \leqslant j \leqslant p-1$ since $x, y \neq 0(\mod p)$, i.e., $y + xj = 0(\mod p)$. Choose $x' = \frac{y+jx}{p}$, then $A_j(x', y') = (x, y)$.

Thus, for any pair $(x, y) \in \mathbb{Z}^2$ there exist j, $0 \leqslant j \leqslant p$ and $(x', y') \in \mathbb{Z}^2$ such that $(x, y) = A_j(x', y')$ i.e., the decomposition above holds.

□

This decomposition allows us to restrict ourselves with finding an upper bound for the number of solutions in each $A_i\mathbb{Z}^2$. Then, the number of all solutions of (1.0.2) in $\mathbb{Z}^2$ does not exceed the sum of number of solutions in all these sets.

We note that if $(x, y) = A_j(x', y')$ is a solution of (1.0.2) then

$$|F(x, y)| = |F(A_j(x', y'))| = |F_{A_j}(x', y')| = 1$$

and from (2.2.4) we have that $|D(F_{A_j})| = p^{r(r-1)}|D(F)|$, since $det(A_j) = p$ for any $0 \leqslant j \leqslant p$. Now denote by $N_r(p)$ the number of solutions of (1.0.2) for the polynomials with

$$|D(F)| \geqslant p^{r(r-1)}$$

Then the number of solutions in each set $A_j\mathbb{Z}^2$ does not exceed $N_r(p)$, therefore for N_r, the number of all the solutions in $\mathbb{Z}^2$ we have

$$N_r \leqslant (p+1)N_r(p)$$

We note that p is any prime number that we have not put any additional conditions yet, it will be fixed at the end of our discussions.

2.3.2 Classification of polynomials

Let F and G be two irreducible homogeneous polynomials in two variables of the same degree. Then these polynomials are called to be equivalent, if the equations

$$|F(x, y)| = 1 \text{ and } |G(x, y)| = 1$$

have the same number of solutions, and for equivalent polynomials F and G we use the notation $F \sim G$. The first easy example of equivalent polynomials is F and $-F$, where F satisfies the conditions above. Also, we note that if $A \in SL_2(\mathbb{Z})$, then $F_A \sim F$.

By this definition of equivalence the set of irreducible homogeneous polynomials of degree $r \geqslant 3$ is divided into several classes. Using, this classification we can replace polynomials with equivalent forms which is easier to work with.

Now suppose that (1.0.2) has a solution (x_0, y_0) for some fixed F. Then, there exists $A \in Sl_2(\mathbb{Z})$, such that $A^{-1}(x_0, y_0) = (1, 0)$. Therefore, $(1, 0)$ is a solution of of the equation $|F_A(x, y)| = 1$. We have that $F_A(1, 0) = b_0$, where b_0 is the leading coefficient of the polynomial F_A, hence $b_0 = \pm 1$. As F, F_A and $-F_A$ are equivalent, therefore considering an equivalent form F_A we can restrict ourselves to forms which are normalized, that is with the leading coefficient equal to 1. Therefore, without loss of generality we can assume that the polynomial F is normalized if needed.

2.4 Mahler height and the Mahler inequality

Let $F(x,y) \in \mathbb{Z}[x,y]$ be a polynomial as in (1.0.2), denote $f(x) = F(x,1)$, then it is irreducible polynomial in one variable and let $\alpha_1, ..., \alpha_r$ be its roots.

We define the Mahler height of the polynomial F as

$$M(F) = |a_0| \prod_{i=1}^{r} \max(1, |\alpha_i|)$$

with a_0 is the leading coefficient of F.

2.4.1 Hadamar's inequality

Is the Mahler height and discriminant of the polynomial F related to each other, if so, what is the relation? To answer this question we need the following theorem by Hadamar.

Theorem 2.4.1. *Let*

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix}$$

be a matrix with complex entries. Then we have the following inequality

$$(\det A)^2 \leqslant \prod_{i=1}^{n} (|a_{i1}|^2 + |a_{i2}|^2 ... + |a_{in}|^2)$$

Now, we apply this theorem to the Vandermonde matrix,

$$V = \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & ... & \alpha_1^{r-1} \\ 1 & \alpha_2 & \alpha_2^2 & ... & \alpha_2^{r-1} \\ ... & ... & ... & ... & ... \\ 1 & \alpha_r & \alpha_r^2 & ... & \alpha_r^{r-1} \end{pmatrix}$$

Then we obtain that

$$(\det V)^2 \leqslant \prod_{i=1}^{r} (1 + |\alpha_i|^2 ... + |\alpha_i|^{2r-2}) \leqslant \prod_{i=1}^{r} r(\max\{1, |\alpha_i|\})^{2r-2} = r^r \left(\prod_{i=1}^{r} \max\{1, |\alpha_i|\} \right)^{2r-2}$$

That is,

$$|a_0|^{2r-2} \cdot (\det V)^2 \leqslant |a_0|^{2r-2} \cdot r^r \left(\prod_{i=1}^{r} \max\{1, |\alpha_i|\} \right)^{2r-2}$$

So, from the definitions of $D(F)$ and $M(F)$ we obtain the following relation between the Mahler height and the discriminant of the polynomial F

$$|D(F)| \leqslant r^r M(F)^{2r-2} \tag{2.4.1}$$

2.5 The Lewis and Mahler estimation

We recall that in this and all the next sections we assume that F is irreducible homogeneous polynomial in two variables with integer coefficients, of degree $r \geqslant 3$.

Let x and y be two coprime integer numbers, then we define the height of the point (x, y) by

$$H(x,y) = \max(|x|, |y|)$$

Then we have the following estimation by Lewis and Mahler.

Lemma 2.5.1. *For any pair of coprime integers* (x, y)*, with* $y \neq 0$ *we have that*

$$\min_{\alpha} \min \left(1, \left|\frac{x}{y} - \alpha\right|\right) \leqslant \frac{(2r^{1/2}M(F))^r |F(x,y)|}{H(x,y)^r}$$

where α *runs through all the roots of the polynomial* $f(x) = F(x, 1)$.

Proof. Let

$$g(x) = \frac{1}{x-\alpha} f(x)$$

where $\alpha = \alpha_i$ for some $i = 1, ..., r$ and $f(x) = a_0(x - \alpha_1)...(x - \alpha_r)$. Then g is a polynomial of degree $r - 1$.

We use the same notation for the discriminant of f as for F, i.e., $D(f) = D(F)$. Then we have

$$D(f) = a_0^{2r-2} \prod_{i<j} (\alpha_i - \alpha_j)^2 = f'(\alpha)^2 D(g)$$

Also, from (2.4.1) we have that $|D(F)| \leqslant r^r M(F)^{2r-2}$.

Then using the last fact for $G(x,y) = \frac{1}{x-\alpha y} F(x,y)$ taking into account the fact that $M(G) \leqslant M(F)$ we obtain

$$|f'(\alpha)| = \frac{|D(F)|^{\frac{1}{2}}}{|D(G)|^{\frac{1}{2}}} \geqslant \frac{|D(F)|^{\frac{1}{2}}}{r^{\frac{r-1}{2}} M(F)^{r-2}}$$

As $D(F)$ is a non-zero integer we have $|D(F)| \geqslant 1$, then

$$|f'(\alpha)| \geqslant \frac{1}{r^{\frac{r-1}{2}} M(F)^{r-2}} \tag{2.5.1}$$

Let (x, y) be a fixed solution, then without loss of generality one may assume that

$$\left|\frac{x}{y} - \alpha_r\right| = \min_{\alpha \in \{\alpha_1, \ldots, \alpha_r\}} \left|\frac{x}{y} - \alpha\right| = \delta$$

Also, reordering the roots if necessary we suppose that

$$|\alpha_r - \alpha_i| \leqslant 2\delta, \text{ if } i = 1, .., N$$

$$|\alpha_r - \alpha_i| > 2\delta, \text{ if } i = N+1, .., r-1$$

for some $1 \leqslant N \leqslant r - 1$. Then

$$\prod_{i=1}^{N} \left|\frac{x}{y} - \alpha_1\right| > \delta^N = 2^{-N}(2\delta)^N \geqslant 2^{-N} \prod_{i=1}^{N} |\alpha_r - \alpha_i| \tag{2.5.2}$$

For, $i = N+1, ..., r-1$ we have that

$$|\alpha_r - \alpha_i| > 2\delta = 2\left|\frac{x}{y} - \alpha_r\right|$$

Therefore

$$\left|\frac{x}{y} - \alpha_i\right| = \left|\frac{x}{y} - \alpha_r + \alpha_r - \alpha_i\right| \geqslant |\alpha_r - \alpha_i| - \left|\frac{x}{y} - \alpha_r\right| \geqslant \frac{1}{2}|\alpha_r - \alpha_i|$$

Then taking the product we obtain

$$\prod_{i=N+1}^{r-1}\left|\frac{x}{y} - \alpha_i\right| > 2^{-(r-N-1)}\prod_{i=N+1}^{r-1}|\alpha_r - \alpha_i| \tag{2.5.3}$$

We have that

$$|F(x,y)| = |a_0|\cdot|y|^r\left|\frac{x}{y} - \alpha_1\right|...\left|\frac{x}{y} - \alpha_r\right| = |a_0|\cdot|y|^r\left(\prod_{i=1}^{N}\left|\frac{x}{y} - \alpha_i\right|\right)\cdot\left(\prod_{i=N+1}^{r-1}\left|\frac{x}{y} - \alpha_i\right|\right)\left|\frac{x}{y} - \alpha_r\right|$$

Then (2.5.2) and (2.5.3) imply that

$$|F(x,y)| \geqslant 2^{-r+1}|y|^r\cdot\left(|a_0|\prod_{i=1}^{r-1}|\alpha_r - \alpha_i|\right)\cdot\left|\frac{x}{y} - \alpha_r\right|$$

using the fact

$$|f'(\alpha_r)| = |a_0|\prod_{i=1}^{r-1}|\alpha_r - \alpha_i|$$

we obtain

$$|F(x,y)| \geqslant 2^{-r+1}|y|^r\cdot|f'(\alpha_r)|\cdot\left|\frac{x}{y} - \alpha_r\right| \tag{2.5.4}$$

Now we consider two cases on $H(x,y)$.
First case, $H(x,y) = |y|$. Then from the relation (2.5.4) we have that

$$|F(x,y)| \geqslant 2^{-r+1}H(x,y)^r\cdot|f'(\alpha_r)|\cdot\left|\frac{x}{y} - \alpha_r\right|$$

Taking into account the relations (2.5.1) and the last inequality we obtain

$$\left|\frac{x}{y} - \alpha_r\right| \leqslant \frac{|F(x,y)|}{2^{-r+1}H(x,y)^r\cdot|f'(\alpha_r)|} \leqslant \frac{2^{r-1}\cdot r^{\frac{r-1}{2}}M(F)^{r-2}|F(x,y)|}{H(x,y)^r} \leqslant \frac{(2r^{\frac{1}{2}}M(F))^r|F(x,y)|}{H(x,y)^r}$$

Therefore

$$\min_{\alpha}\left(1, \left|\frac{x}{y} - \alpha\right|\right) \leqslant \frac{(2r^{\frac{1}{2}}M(F))^r|F(x,y)|}{H(x,y)^r}$$

Second case, $H(x,y) = |x|$. We have that

$$|F(x,y)| = |a_0||x - \alpha_1 y|...|x - \alpha_r y| = |x|^r\cdot|a_0|\cdot|\alpha_1|...|\alpha_r|\left|\frac{1}{\alpha_1} - \frac{y}{x}\right|...\left|\frac{1}{\alpha_r} - \frac{y}{x}\right|$$

Again without loss of generality one may assume

$$\min_{\alpha}\left|\frac{1}{\alpha}-\frac{y}{x}\right|=\left|\frac{1}{\alpha_r}-\frac{y}{x}\right|$$

where minimum is taken through all the roots of f.

Denote, $\varphi(x)=F(1,x)$. Then $\frac{1}{\alpha_1},...,\frac{1}{\alpha_r}$ are its roots, repeating the same procedure as we followed for f in this case we obtain that

$$|F(x,y)|\geqslant 2^{-r+1}|x|^r\cdot|\varphi'(\frac{1}{\alpha_r})|\cdot\left|\frac{y}{x}-\frac{1}{\alpha_r}\right|$$

Again using (2.5.1) for φ we have

$$\left|\frac{y}{x}-\frac{1}{\alpha_r}\right|\leqslant\frac{(2r^{\frac{1}{2}})^{r-1}M(F))^{r-2}|F(x,y)|}{|x|^r}=\frac{(2r^{\frac{1}{2}})^{r-1}M(F))^{r-2}|F(x,y)|}{H(x,y)^r}\tag{2.5.5}$$

Now, denote

$$\sigma=\max(1,|\alpha_1|,...,|\alpha_r|)$$

If we have

$$\left|\frac{1}{\alpha_r}-\frac{y}{x}\right|\geqslant\frac{1}{2\sigma}$$

then applying this directly to (2.5.5) we get

$$\frac{(2r^{\frac{1}{2}})^{r-1}M(F)^{r-2}|F(x,y)|}{H(x,y)^r}\geqslant\frac{1}{2\sigma}\geqslant\frac{1}{2\sigma}\min_{\alpha}\left(1,\left|\frac{x}{y}-\alpha\right|\right)\tag{2.5.6}$$

Otherwise if

$$\left|\frac{1}{\alpha_r}-\frac{y}{x}\right|\leqslant\frac{1}{2\sigma}$$

then according to the definition of σ

$$\left|\frac{1}{\alpha_r}\right|\geqslant\frac{1}{\sigma}$$

therefore we have that

$$\left|\frac{y}{x}\right|=\left|\frac{1}{\alpha_r}+\left(\frac{y}{x}-\frac{1}{\alpha_r}\right)\right|\geqslant\left|\frac{1}{\alpha_r}\right|-\left|\frac{y}{x}-\frac{1}{\alpha_r}\right|\geqslant\frac{1}{2\sigma}$$

Then

$$\left|\frac{y}{x}-\frac{1}{\alpha_r}\right|=\left|\frac{y}{x}\cdot\frac{1}{\alpha_r}\left(\frac{x}{y}-\alpha_r\right)\right|\geqslant\frac{1}{2\sigma}\cdot\frac{1}{\sigma}\left|\frac{x}{y}-\alpha_r\right|=\frac{1}{2\sigma^2}\left|\frac{x}{y}-\alpha_r\right|$$

Therefore from (2.5.5) we obtain

$$\frac{(2r^{\frac{1}{2}})^{r-1}M(F)^{r-2}|F(x,y)|}{H(x,y)^r}\geqslant\frac{1}{2\sigma^2}\left|\frac{x}{y}-\alpha_r\right|\geqslant\frac{1}{2\sigma^2}\min_{\alpha}\left(1,\left|\frac{x}{y}-\alpha\right|\right)\tag{2.5.7}$$

From the relations (2.5.6) and (2.5.7), taking into account the fact that $\sigma\geqslant 1$ we can conclude that

$$\min_{\alpha}\left(1, \left|\frac{x}{y}-\alpha\right|\right) \leqslant 2\sigma^2 \frac{(2r^{\frac{1}{2}})^{r-1}M(F)^{r-2}|F(x,y)|}{H(x,y)^r}$$

Also, we note that

$$\sigma = \max(1, |\alpha_1|, ..., |\alpha_r|) \leqslant M(F)$$

Therefore we obtain

$$\min_{\alpha}\left(1, \left|\frac{x}{y}-\alpha\right|\right) \leqslant \frac{(2r^{\frac{1}{2}}M(F))^r|F(x,y)|}{H(x,y)^r}$$

Thus, in both cases the desired result is proven.

□

2.6 The Thue-Siegel principle

2.6.1 Some definitions and facts on number fields

Let K be a number field with $[K:\mathbb{Q}] = r \geqslant 3$.

Denote by M_K the set of places on K, then for any $\alpha \in K^\times$ we have the following product formula

$$\prod_{\nu \in M_K} |\alpha|_\nu^{d_\nu} = 1$$

where $d_\nu = [K_\nu : \mathbb{Q}_\nu]$ with $\mathbb{Q}_\nu$ and K_ν completions of $\mathbb{Q}$ and K respectively, with respect to an absolute value $\nu \in M_K$.

Definition 2.6.1. *The height of a number* $\alpha \in K$ *is*

$$H(\alpha) = \left[\prod_{\nu \in M_K} \max\{1, |\alpha|_\nu\}^{d_\nu}\right]^{\frac{1}{[K:\mathbb{Q}]}}$$

and logarithmic height is

$$h(\alpha) = \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} d_\nu \log^+ |\alpha|_\nu$$

where $\log^+ x = \max\{0, \log x\}$ *assuming that* $\log^+ 0 := 0$ *and* $d_\nu = [K_\nu : Q_\nu]$ *as above.*

We note that in the general case we can define the height of any algebraic number $\alpha \in \overline{\mathbb{Q}}$ in the same way, with K is any number field containing α, then this definition is well defined and does not depend on the choice of K.

Definition 2.6.2. *For a vector* $\underline{\alpha} = (\alpha_1, ..., \alpha_n) \in \mathbb{A}^n(K)$ *the affine height* $h_{\mathbb{A}}$ *is*

$$h_{\mathbb{A}}(\underline{\alpha}) = \frac{1}{[K:\mathbb{Q}]} \sum_{\nu \in M_K} d_\nu \log^+ |\underline{\alpha}|_\nu$$

where $|\underline{\alpha}|_\nu := \max\{1, |\alpha_0|_\nu, ..., |\alpha_n|_\nu\}$.

We also define the logarithmic height $h(P)$ of a polynomial $P(x,y) \in K[x,y]$ as the affine height of the vector consisting of the coefficients of P.

Let $\alpha \in K$ and suppose that $|\alpha|_\nu \leqslant 1$ for some place $\nu \in M_K$. We say that a rational number $\beta \in \mathbb{Q}$ approximates α if $|\alpha - \beta|_\nu < 1$.

2.6.2 Admissible triple

To describe the Thue-Siegel principle corresponding to our case, first we give a definition of admissible triple. Also, from now on we always assume that the polynomial F is normalized according to the discussions in the Section 2.3.2.

Definition 2.6.3. *Let $t > 0$ be any positive number and let β_1, β_2 be two rational numbers that approximate $\alpha \in K$. We say that (A_1, A_2, τ) is an admissible triple for $(\alpha, \beta_1, \beta_2, t, \delta)$, where $\delta > 0$ is an arbitrary number, if for all positive numbers d_1, d_2 with $d_2 \leqslant \delta d_1$ there exists a polynomial $P(x_1, x_2) \in \mathbb{Q}[x_1, x_2]$ that satisfies the following three conditions*

(i)

$$deg_{x_1} P \leqslant d_1, deg_{x_2} P \leqslant d_2$$

where $deg_{x_i} P$ is the highest degree of x_i in the polynomial P, $i = 1, 2$ and

$$\frac{\partial^{i_1+i_2}}{\partial x_1^{i_1} \partial x_2^{i_2}} P(\alpha, \alpha) = 0$$

for all (i_1, i_2) with

$$\frac{i_1}{d_1} + \frac{i_2}{d_2} < t$$

(ii) there exists (j_1, j_2) with

$$\frac{\partial^{j_1+j_2}}{\partial x_1^{j_1} \partial x_2^{j_2}} P(\beta_1, \beta_2) \neq 0$$

and

$$\frac{j_1}{d_1} + \frac{j_2}{d_2} < \tau$$

(iii) the following relation is satisfied

$$h(P) \leqslant A_1 d_1 + A_2 d_2 + o(d_1) + o(d_2)$$

when d_1, d_2 are big enough with $d_2 \leqslant \delta d_1$.

Theorem 2.6.1 (Thue-Siegel principle). *Let $0 < t < \sqrt{\frac{2}{r}}$ and let $\alpha, \beta_1, \beta_2, \delta$ be the same as in the definition above, and let (A_1, A_2, τ) be an admissible triple for this data. Suppose also $0 < \tau < t$, and*

$$|\alpha - \beta_1| < t - \tau \text{ and } |\alpha - \beta_2| < \frac{1}{2}(t - \tau)^2$$

If

$$|\alpha - \beta_1| < (3e^{A_1} h(\beta_1))^{-\lambda}$$

then we have that

$$|\alpha - \beta_2| > (3e^{A_2} h(\beta_2))^{-\lambda}$$

or

$$\log 3 + A_2 + h(\beta_2) < \delta^{-1}(\log 3 + A_1 + h(\beta_1))$$

where $\lambda = \frac{2}{t-\tau}$.

Proof. As (A_1, A_2, τ) is an admissible triple for $(\alpha, \beta_1, \beta_2, t, \delta)$, from the definition it follows that there exists a polynomial $P(x_1, x_2) \in \mathbb{Q}[x_1, x_2]$ satisfying the properties $(i), (ii)$ and (iii). Accordingly, there exist j_1, j_2, $\frac{j_1}{d_1} + \frac{j_2}{d_2} < \tau$ such that

$$\frac{1}{j_1!j_1!}\frac{\partial^{j_1+j_2}}{\partial x_1^{j_1}\partial x_2^{j_2}}P(\beta_1, \beta_2) \neq 0$$

we denote this non-zero value by γ.

Then from the product formula we have that

$$\sum_{\nu\in M_K} d_\nu \log|\gamma|_\nu = \sum_{\nu|\infty} d_\nu \log|\gamma|_\nu + \sum_{\nu\nmid\infty} d_\nu \log|\gamma|_\nu = 0$$

Using the Taylor expansion around a point (α, α) we have the following relation for the value of γ

$$\gamma = \sum_{i_1,i_2} \frac{1}{i_1!i_2!j_1!j_2!}\frac{\partial^{i_1+i_2+j_1+j_2}}{\partial x_1^{i_1+j_1}\partial x_2^{i_2+j_2}}P(\alpha,\alpha)(\beta_1-\alpha)^{i_1}(\beta_2-\alpha)^{i_2}$$

According to the choice of $P(x_1, x_2)$ we have that

$$\frac{\partial^{j_1+j_2+i_1+i_2}}{\partial x_1^{i_1+j_1}\partial x_2^{i_2+j_2}}P(\alpha,\alpha) = 0, \text{ for } \frac{j_1+i_1}{d_1} + \frac{j_2+i_2}{d_2} < t$$

also we have that

$$\frac{j_1}{d_1} + \frac{j_2}{d_2} < \tau$$

therefore it holds for all (i_1, i_2) with

$$\frac{i_1}{d_1} + \frac{i_2}{d_2} < t - \tau$$

Now, we estimate the values of $\log|\gamma|_\nu$ from above for all $\nu \in M_K$.

For this, we consider the following cases separately: 1) $\nu = \infty$, i.e., absolute value that is restricted from the usual archimedian absolute value on $\mathbb{C}$, 2) ν archimedian except the first case 3) ν is non-archimedian.

1) $\nu = \infty$. Then from the Taylor expansion above we obtain that

$$\log|\gamma|_\nu \leqslant \max_{i_1,i_2} \log\left|\frac{1}{i_1!i_2!j_1!j_2!}\frac{\partial^{i_1+i_2+j_1+j_2}}{\partial x_1^{i_1+j_1}\partial x_2^{i_2+j_2}}P(\alpha,\alpha)(\beta_1-\alpha)^{i_1}(\beta_2-\alpha)^{i_2}\right|_\nu + o(d_1) + o(d_2)$$

Let $\deg_{x_1} P = d_1$ and $\deg_{x_2} P = d_2$, then non-zero coefficients in the expansion of γ are of the form

$$\frac{k_1!k_2!}{j_1!j_2!i_1!i_2!(k_1-i_1-j_1)!(k_2-i_2-j_2)!}$$

for some k_1, k_2 such that $i_1 + j_1 \leqslant k_1 \leqslant d_1$ and $i_2 + j_2 \leqslant k_2 \leqslant d_2$.

Now, we want to find a bound for the logarithmic absolute values of these coefficients, the following lemma describes this relation.

Lemma 2.6.1. *The following inequality holds*

$$\max_{k_1\leqslant d_1,k_2\leqslant d_2} \log \frac{k_1!k_2!}{i_1!i_2!j_1!j_2!(k_1-i_1-j_1)!(k_2-i_2-j_2)!} \leqslant d_1 F\left(\frac{i_1}{d_1},\frac{j_1}{d_1}\right) + d_2 F\left(\frac{i_2}{d_2},\frac{j_1}{d_2}\right)$$

where

$$F(u,v) = u\log\frac{1}{u} + v\log\frac{1}{v} + (1-u-v)\log\frac{1}{1-u-v} \quad \textit{if } u \leqslant \frac{1-v}{2}$$

$$F(u,v) = u\log\frac{1}{u} + (1-v)\log\frac{2}{1-v} \quad \textit{if } u > \frac{1-v}{2}$$

Applying this lemma we obtain

$$\log|\gamma|_\nu \leqslant \log|P|_\nu + d_1\log^+|\beta_1|_\nu + d_2\log^+|\beta_2|_\nu$$

$$+\max_{x_1,x_2}\left\{d_1F\left(x_1,\frac{j_1}{d_1}\right) + d_2F\left(x_2,\frac{j_2}{d_2}\right) + d_1x_1\log|\alpha-\beta_1|_\nu + d_2x_1\log|\alpha-\beta_2|_\nu\right\}$$

where maximum is taken over $x_1+x_2 \geqslant t-\tau, 0\leqslant x_1 \leqslant 1-\frac{j_1}{d_1}, 0\leqslant x_2\leqslant 1-\frac{j_2}{d_2}$.

Differentiating and taking into account the facts that

$$|\alpha-\beta_1|_\nu < t-\tau \text{ and } |\alpha-\beta_2|_\nu < \frac{1}{2}(t-\tau)^2$$

and after some calculations we get that the maximum is reached for $x_1+x_2=t-\tau$, then $x_1\leqslant t-\tau$ and $x_2\leqslant t-\tau$, then

$$F\left(x_1,\frac{j_1}{d_1}\right) \leqslant F\left(t-\tau,\frac{j_1}{d_1}\right) \leqslant \log 3$$

and

$$F\left(x_2,\frac{j_2}{d_1}\right) \leqslant \log 3$$

Also, considering the maximum over $x_1+x_2=t-\tau$ we obtain

$$\max\{d_1x_1\log|\alpha-\beta_1|_\nu + d_2x_2\log|\alpha-\beta_2|_\nu\} = -(t-\tau)\min\left\{d_1\log\frac{1}{|\alpha-\beta_1|_\nu} + d_2\log\frac{1}{|\alpha-\beta_2|_\nu}\right\}$$

So, we have that

$$\log|\gamma|_\nu \leqslant \log|P|_\nu + d_1\log^+|\beta_1|_\nu + d_2\log^+|\beta_2|_\nu + d_1\log 3 + d_2\log 3$$

$$-(t-\tau)\min\left\{d_1\log\frac{1}{|\alpha-\beta_1|_\nu} + d_2\log\frac{1}{|\alpha-\beta_2|_\nu}\right\}$$

2) ν archimedian except the first case. Then, using the fact that

$$\max_{k\leqslant d}\log\binom{k}{i} \leqslant d\log 3$$

we obtain

$$\log|\gamma|_\nu = \log\left|\frac{1}{j_1!j_1!}\frac{\partial^{j_1+j_2}}{\partial^{j_1}\partial^{j_2}}P(\beta_1,\beta_2)\right|_\nu \leqslant \log|P|_\nu + d_1\log^+|\beta_1|_\nu + d_2\log^+|\beta_2|_\nu$$

$$+d_1\log 3 + d_2\log 3 + o(d_1) + o(d_2)$$

3) ν is non-archimedian. Then we have that $|n|_\nu \leqslant 1$ for all $n\in\mathbb{Z}$, therefore

$$\log|\gamma|_\nu = \log\left|\frac{1}{j_1!j_1!}\frac{\partial^{j_1+j_2}}{\partial^{j_1}\partial^{j_2}}P(\beta_1,\beta_2)\right|_\nu \leqslant \log|P|_\nu + d_1\log^+|\beta_1|_\nu + d_2\log^+|\beta_2|_\nu$$

Then 1), 2) and 3) imply that

$$\sum_{\nu\in K} d_v \log|\gamma|_\nu = \sum_{\nu=\infty} d_v \log|\gamma|_\nu + \sum_{\nu\neq\infty,\nu|\infty} d_v \log|\gamma|_\nu + \sum_{\nu\nmid\infty} d_v \log|\gamma|_\nu$$

$$\leqslant \sum_{\nu\in K} d_v \log|P|_\nu + d_1 \sum_{\nu\in K} d_v \log^+|\beta_1|_\nu + d_2 \sum_{\nu\in K} d_v \log^+|\beta_2|_\nu + d_1\log 3 + d_2 \log 3$$

$$-(t-\tau)\min\left\{d_1 \log\frac{1}{|\alpha-\beta_1|}, d_2\log\frac{1}{|\alpha-\beta_2|}\right\} + o(d_1) + o(d_2)$$

$$= h(P) + d_1 h(\beta_1) + d_2 h(\beta_2) + d_1 \log 3 + d_2 \log 3$$

$$-(t-\tau)\min\left\{d_1 \log\frac{1}{|\alpha-\beta_1|}, d_2\log\frac{1}{|\alpha-\beta_2|}\right\} + o(d_1) + o(d_2)$$

On the other hand, for $\gamma \in \mathbb{Q}$ from product formula we have

$$\sum_{\nu\in K} d_v \log|\gamma|_\nu = 0$$

According to the choice of A_1 and A_2, for $d_2 \leqslant \delta d_1$ we have

$$h(P) \leqslant A_1 d_1 + A_2 d_2 + o(d_1) + o(d_2).$$

Now, choose

$$d_1 = \left[\frac{D}{\log 3 + A_1 + h(\beta_1)}\right]$$

and

$$d_2 = \left[\frac{D}{\log 3 + A_2 + h(\beta_2)}\right]$$

where $D \in \mathbb{R}_{>0}$ large enough. If $d_2 > \delta d_1$ then

$$\log 3 + A_2 + h(\beta_2) \leqslant \delta^{-1}(\log 3 + A_1 + h(\beta_1))$$

Otherwise, if $d_2 \leqslant \delta d_1$ then,

$$h(P) \leqslant A_1 d_1 + A_2 d_2 + o(d_1) + o(d_2)$$

for D big enough, applying this we obtain

$$0 \leqslant d_1(\log 3 + A_1 + h(\beta_1)) + d_2(\log 3 + A_2 + h(\beta_2)) - (t-\tau)\min\left\{d_1 \log\frac{1}{|\alpha-\beta_1|}, d_2\log\frac{1}{|\alpha-\beta_2|}\right\}$$

Then, the condition

$$|\alpha-\beta_1| < (3e^{A_1}h(\beta_1))^{-\lambda}$$

and

$$(t-\tau)\min\left\{d_1 \log\frac{1}{|\alpha-\beta_1|}, d_2\log\frac{1}{|\alpha-\beta_2|}\right\} \leqslant d_1(\log 3 + A_1 + h(\beta_1)) + d_2(\log 3 + A_2 + h(\beta_2))$$

imply

$$|\alpha-\beta_2| > (3e^{A_2}h(\beta_2))^{-\lambda}$$

□

Example of an admissible triple

To apply the Thue-Siegel principle we need some particular admissible triple for $(\alpha, \beta_1, \beta_2, t, \delta)$. For this purpose we describe two facts below from the discussions in [4].

Lemma 2.6.2. *Let* $\sqrt{\frac{2}{r+1}} < t < \sqrt{\frac{2}{r}} < 1$ *and let* $\alpha \in \bar{\mathbb{Q}}$ *be an algebraic number of degree* $r \geqslant 3$ *over* $\mathbb{Q}$, *if* (A_1, A_2, τ) *is admissible then* (A_1, A_2, τ_0) *with*

$$\tau_0 = \sqrt{2 - rt^2 + (r-1)\delta}$$

is also admissible.

Lemma 2.6.3. *We have that*

$$A_1 = A_2 = \frac{rt^2}{2 - rt^2}\left(h(\alpha) + \frac{1}{2}\right)$$

with $\tau_0 = \sqrt{2 - rt^2 + (r-1)\delta}$ *is an admissible triple for* $(\alpha, \beta_1, \beta_2, t, \delta)$.

Corollary 2.6.1. *Let* $t < \sqrt{\frac{2}{r}}, \sqrt{2 - rt^2} < \tau < t$ *and* $A_1 = \frac{t^2}{2-rt^2}\left(\log M(F) + \frac{r}{2}\right)$. *Also suppose that* $\lambda = \frac{2}{t-\tau} < r$.

Let α *be an algebraic number of degree* r, *and*

$$\left|\alpha - \frac{x}{y}\right| < (4e^{A_1}H(x,y))^{-\lambda} \text{ and } \left|\alpha - \frac{x'}{y'}\right| < (4e^{A_1}H(x',y'))^{-\lambda}$$

then

$$\log(4e^{A_1}) + \log H(x', y') \leqslant \delta^{-1}\{\log(4e^{A_1}) + \log H(x,y)\}$$

where

$$\delta = \frac{rt^2 + \tau^2 - 2}{r - 1}$$

Proof. For the chosen value for A_1 we already have that

$$\left|\alpha - \frac{x}{y}\right| < (4e^{A_1}H(x,y))^{-\lambda} < (t - \tau)$$

and

$$\left|\alpha - \frac{x'}{y'}\right| < (4e^{A_1}H(x',y'))^{-\lambda} < \frac{1}{2}(t-\tau)^2$$

Then, if $|\alpha| \leqslant 1$, the result immediately follows from the Thue-Siegel principle above. Otherwise, if $|\alpha| > 1$ then

$$\left|\alpha - \frac{x}{y}\right| < (4e^{A_1}H(x,y))^{-\lambda}$$

and the fact that x/y approximates α strongly (i.e., the difference is small enough) imply

$$\left|\alpha^{-1} - \frac{y}{x}\right| < |\alpha|^{-1}\left|\frac{y}{x}\right|(4e^{A_1}H(x,y))^{-\lambda} < (3e^{A_1}H(x,y))^{-\lambda}$$

then again Thue-Siegel principle for $\alpha^{-1}, y/x, y'/x'$ gives the desired result, since Mahler heights of $F(y,x)$ and $F(x,y)$ are the same.

$\square$

Chapter 3

An upper bound for the number of solutions

3.1 Large solutions

We first consider the solutions of the equation

$$|F(x,y)| = 1 \tag{3.1.1}$$

which are 'large', i.e., solutions (x,y) such that $H(x,y) \geqslant M$ for some $M > 0$ big enough, which will be fixed later.

3.1.1 Strong gap principle

Now, classify the roots of (3.1.1) dividing them into r classes, we call that two solutions (x,y) and (x',y') belong to the same class if

$$\min_{\alpha} \left| \alpha - \frac{x}{y} \right| = \left| \alpha_0 - \frac{x}{y} \right| \text{ and } \min_{\alpha} \left| \alpha - \frac{x'}{y'} \right| = \left| \alpha_0 - \frac{x'}{y'} \right|$$

for some α_0 , a root of the polynomial f , where α runs through all the roots.

Now, take any such a class, then we can numerate the elements as $(x_1,y_1),(x_2,y_2),...$ and reordering if necessary suppose that

$$H(x_1,y_1) \leqslant H(x_2,y_2) \leqslant ....$$

Then for two consequent solutions (x_n,y_n) and (x_{n+1},y_{n+1}) from one class we have the following relation

$$\frac{1}{y_n y_{n+1}} \leqslant \left| \frac{x_n}{y_n} - \frac{x_{n+1}}{y_{n+1}} \right| = \left| \frac{x_n}{y_n} - \alpha_0 + \alpha_0 - \frac{x_{n+1}}{y_{n+1}} \right| \leqslant \left| \frac{x_n}{y_n} - \alpha_0 \right| + \left| \alpha_0 - \frac{x_{n+1}}{y_{n+1}} \right|$$

Then Lewis and Mahler estimation stated before implies

$$\left| \frac{x_n}{y_n} - \alpha_0 \right| + \left| \frac{x_{n+1}}{y_{n+1}} - \alpha_0 \right| \leqslant \frac{C|F(x_n,y_n)|}{H(x_n,y_n)^r} + \frac{C|F(x_{n+1},y_{n+1})|}{H(x_{n+1},y_{n+1})^r}$$

where $C = (2r^{\frac{1}{2}}M(F))^r$.

According to our assumption $H(x_{n+1}, y_{n+1}) \geqslant H(x_n, y_n)$ and $|F(x_{n+1}, y_{n+1})| = |F(x_n, y_n)| = 1$ as (x_n, y_n) and (x_{n+1}, y_{n+1}) are solutions, therefore

$$\frac{1}{y_n y_{n+1}} \leqslant \left|\frac{x_n}{y_n} - \frac{x_{n+1}}{y_{n+1}}\right| \leqslant \frac{2C}{H(x_n, y_n)^r} \tag{3.1.2}$$

Theorem 3.1.1 (Strong gap principle)**.** *Suppose that* $H(x_1, y_1) \geqslant C^{\frac{1}{r}}$. *Then for each* $n = 1, 2, ...$ *we have*

$$H(x_n, y_n) \geqslant \left\{(2C)^{-\frac{1}{r-2}} H(x_1, y_1)\right\}^{(r-1)^{n-1}}$$

Proof. Firstly, taking into account the relation (3.1.2) we have

$$H(x_{i+1}, y_{i+1}) \geqslant y_{i+1} = \frac{y_{i+1} y_i}{y_i} \geqslant H(x_i, y_i)^r/(2Cy_i) \geqslant H(x_i, y_i)^{r-1}/2C$$

for all $i = 1, 2, \ldots$. Applying this fact several times we obtain

$$H(x_n, y_n) \geqslant \frac{H(x_{n-1}, y_{n-1})^{r-1}}{2C} \geqslant \frac{H(x_{n-2}, y_{n-2})^{(r-1)^2}}{(2C)^{1+(r-1)}} \geqslant \ldots \geqslant \frac{H(x_1, y_1)^{(r-1)^{n-1}}}{(2C)^{1+(r-1)+(r-1)^2+\ldots(r-1)^{n-2}}}$$

We have that

$$(2C)^{1+(r-1)+(r-1)^2+\ldots(r-1)^{n-2}} = (2C)^{\frac{(r-1)^{n-1}-1}{r-2}} \leqslant \left((2C)^{\frac{1}{r-2}}\right)^{(r-1)^{n-1}}$$

Thus,

$$H(x_n, y_n) \geqslant \frac{H(x_1, y_1)^{(r-1)^{n-1}}}{(2C)^{1+(r-1)+(r-1)^2+\ldots(r-1)^{n-2}}} \geqslant \left\{(2C)^{-\frac{1}{r-2}} H(x_1, y_1)\right\}^{(r-1)^{n-1}}$$

Now, we combine the strong gap principle, the corollary of the Thue-Siegel principle and the Lewis and Mahler estimation. For this purpose we choose M such that $\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots$ approximate α_0 'good' enough, α_0 is some root of the polynomial f.

From the Lewis and Mahler estimation we have that

$$\left|\alpha_0 - \frac{x_i}{y_i}\right| \leqslant \frac{C}{H(x_i, y_i)^r}$$

for every i. So, if we assume that

$$M \geqslant C^{\frac{1}{r-\lambda}} (4e^{A_1})^{\frac{\lambda}{r-\lambda}}$$

then

$$\frac{C}{H(x_i, y_i)^r} \leqslant \frac{1}{(4e^{A_1} H(x_i, y_i))^{\lambda}}$$

Then the corollary of the Thue-Siegel principle implies

$$\log(4e^{A_1}) + \log H(x_n, y_n) < \delta^{-1}(\log(4e^{A_1}) + \log H(x_1, y_1))$$

as

$$H(x_n, y_n) \geqslant \left\{(2C)^{-\frac{1}{r-2}} H(x_1, y_1)\right\}^{(r-1)^{n-1}}$$

we have that

$$\log H(x_n,y_n) \geqslant (r-1)^{n-1}\log\left((2C)^{-\frac{1}{r-2}}H(x_1,y_1)\right) = (r-1)^{n-1}\left(\log H(x_1,y_1) - \frac{1}{r-2}\log(2C)\right)$$

Therefore,

$$\log(4e^{A_1}) + (r-1)^{n-1}\left(\log H(x_1,y_1) - \frac{1}{r-2}\log(2C)\right) < \delta^{-1}(\log(4e^{A_1}) + \log H(x_1,y_1))$$

or

$$(r-1)^{n-1} \leqslant \delta^{-1}\frac{\log M + \log(4e^{A_1})}{\log M - (r-2)^{-1}\log(2C)}$$

when $\log M > (r-2)^{-1}\log(2C)$, which holds in our case according to requirements for M and the choice of A_1. If we choose

$$M = (2C)^{\frac{1}{r-\lambda}}(4e^{A_1})^{\frac{\lambda}{r-\lambda}} \tag{3.1.3}$$

then we get

$$(r-1)^{n-1} \leqslant \delta^{-1}\frac{(r-\lambda)^{-1}\log(2C) + r(r-\lambda)^{-1}\log(4e^{A_1})}{(\lambda-2)(r-2)^{-1}(r-\lambda)^{-1}\log(2C) + \lambda(r-\lambda)^{-1}\log(4e^{A_1})}$$

$$\leqslant \delta^{-1}\frac{\log(2C) + r\log(4e^{A_1})}{(\lambda-2)(r-2)^{-1}\log(2C) + \lambda\log(4e^{A_1})} = \delta^{-1}\frac{(r-2)(\log(2C) + r\log(4e^{A_1}))}{(\lambda-2)(\log(2C) + \frac{r-2}{\lambda-2}\lambda\log(4e^{A_1}))} \leqslant \delta^{-1}\frac{r-2}{\lambda-2}$$

since according to our assumption $\lambda \leqslant r$. Then

$$(n-1)\log(r-1) \leqslant \log(\delta^{-1}\frac{r-2}{\lambda-2})$$

$$n \leqslant 1 + \frac{\log(\delta^{-1}(\lambda-2)^{-1}) + \log(r-2)}{\log(r-1)} < 2 + \frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)}$$

So, the number of coprime solutions (x,y), provided (x,y) and $(-x,-y)$ are counted as one, of the equation

$$|F(x,y)| = 1$$

such that $H(x,y) \geqslant M$, where M is the same in (3.1.3) does not exceed

$$\left(2 + \left[\frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)}\right]\right) r$$

since we divided the solutions into r different classes and in each such a class there is no more than $2 + \left[\frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)}\right]$ elements.

As this fact true for all t and λ satisfying the conditions

$$t < \sqrt{\frac{2}{r}},\ \sqrt{2-rt^2} < \tau < t \text{ and } \lambda < r$$

Now, we choose $t = \sqrt{\frac{2}{r+a^2}}, \tau = bt$ with $0 < a < b < 1$ then they satisfy the above conditions. For this data we have that

$$\delta^{-1} = \frac{r-1}{rt^2+\tau^2-2} = \frac{r-1}{\frac{2r}{r+a^2}+b^2r^2-2} < \frac{r^2}{2(b^2-a^2)}$$

and

$$\lambda = \frac{2}{(1-b)t} > \frac{\sqrt{2r}}{1-b}$$

Then we obtain

$$\frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)} < \frac{\log(\frac{r^2}{2(b^2-a^2)}(\lambda-2)^{-1})}{\log(r-1)} < 2$$

that is

$$\left[\frac{\log(\delta^{-1}(\lambda-2)^{-1})}{\log(r-1)}\right] \leqslant 1$$

for r big enough, therefore the upper bound is $3r$ in this case. □

3.2 Small solutions

Now, we deal with 'small' solutions, i.e., the solutions (x,y) with $y \leqslant M$, where M is the same as in the last section. We remind that we classified the polynomials in the part 2.3.2. and in this section we always assume that the polynomial F has the smallest Mahler height in the class that it belongs.

The idea we are going to follow is changing the polynomial F with another polynomial which has the same number of solutions as F which is easier to work with. We note that if $A \in SL_2(\mathbb{Z})$, then F_A and F are equivalent, that is the equations

$$|F(x,y)| = 1 \text{ and } |G(x,y)| = 1 \tag{3.2.1}$$

have the same number of solutions, where $G(x,y) = F_A(x,y) = F(ax+by, cx+dy)$ with $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. Therefore for simplicity we can change F with F_A, $A \in SL_2(\mathbb{Z})$ if needed. And we also assume that the leading coefficient of F is 1.

Auxiliary polynomial

As before, let $\alpha_1, ..., \alpha_r$ be the roots of the polynomial $f(x) = F(x,1)$ then

$$F(x,y) = (x-\alpha_1 y)...(x-\alpha_r y)$$

We use the notations $L_i(x,y) := x - \alpha_i y, i = 1,..,r$, then $F(x,y) = L_1(x,y) \cdot ... \cdot L_r(x,y)$ Then, if (x_0, y_0) is an integer solution of the equation

$$|F(x,y)| = 1$$

then $L_i(x_0,y_0) \neq 0$, for all $i = 1,...,r$. As $\gcd(x_0,y_0) = 1$, there exists a pair $(x_0', y_0') \in \mathbb{Z}^2$ with $x_0y_0' - y_0x_0' = 1$, i.e., $\begin{pmatrix} x_0 & y_0 \\ x_0' & y_0' \end{pmatrix} \in SL_2(\mathbb{Z})$ that is (x_0,y_0) and (x_0',y_0') is a basis for $\mathbb{Z}^2$. Then for any $(x,y) \in \mathbb{Z}^2$ we have a decomposition $(x,y) = a(x_0,y_0) + b(x_0',y_0')$, for some $a, b \in \mathbb{Z}$. In fact,

$$x_0y - y_0x = x_0(ay_0 + by_0') - y_0(ax_0 + bx_0') = b(x_0y_0' - y_0x_0') = b$$

Then,

$$(x, y) = a(x_0, y_0) + (x_0 y - y_0 x)(x_0', y_0')$$

Therefore for all solutions (x, y) and for all $i = 1, ..., r$ we have that

$$\frac{L_i(x, y)}{L_i(x_0, y_0)} = a + (x_0 y - y_0 x)\frac{L_i(x_0', y_0')}{L_i(x_0, y_0)} = a - (x_0 y - y_0 x)\beta_i$$

where $\beta_i = -L_i(x_0', y_0')/L_i(x_0, y_0)$. Then,

$$\frac{L_i(x, y)}{L_i(x_0, y_0)} - \frac{L_j(x, y)}{L_j(x_0, y_0)} = (x_0 y - y_0 x)(\beta_i - \beta_j) \tag{3.2.2}$$

Now, let (x_0, y_0) be a fixed solution, then define an auxiliary polynomial G, as

$$G(v, w) = (v - \beta_1 w)...(v - \beta_2 w) = \left(v + \frac{L_1(x_0', y_0')}{L_1(x_0, y_0)}w\right)...\left(v + \frac{L_r(x_0', y_0')}{L_r(x_0, y_0)}w\right)$$

We have that $|L_1(x_0, y_0)| \cdot ... \cdot |L_r(x_0, y_0)| = 1$ therefore

$$|G(v, w)| = \prod_{i=1}^{r} |L_i(x_0, y_0)v + L_i(x_0', y_0')w| = \prod_{i=1}^{r} |(x_0 v + x_0' w) - (y_0 v + y_0' w)\alpha_i| =$$

$$= |F(x_0 v + x_0' w, y_0 v + y_0' w)| = |F_X(v, w)|$$

where $X = \begin{pmatrix} x_0 & x_0' \\ y_0 & y_0' \end{pmatrix} \in SL_2(\mathbb{Z})$. Therefore G is equivalent to F, we have that $|G(u, v)| = 1$ implies $G(u, v) = \pm 1$, we choose a sign in such a way that $G(1, 0) = 1$, i.e., leading coefficient equal to 1. As G is well defined for all the solutions (x_0, y_0), we can consider the specific case when $(x_0, y_0) = (1, 0)$ which is clearly a solution. In this case we have that $L_i(x_0, y_0) = L_i(1, 0) = 1$ for all $1 \leqslant i \leqslant r$ and (3.2.2) can be rewritten as

$$\frac{1}{L_i(x, y)} - \frac{1}{L_j(x, y)} = y(\beta_i - \beta_j) \tag{3.2.3}$$

for all solutions (x, y). If for the product of r positive numbers we have $a_1 \cdot .... \cdot a_r = 1$ then there is at least one with $a_i \geqslant 1$. Using this fact for $|L_1(x, y)| \cdot ... \cdot |L_r(x, y)| = 1$ we can conclude that $\exists i, 1 \leqslant i \leqslant r$ such that $|L_i(x, y)| \geqslant 1$. Then (3.2.3) implies that

$$\frac{1}{|L_j(x, y)|} = \left|\frac{1}{L_j(x, y)} + y(\beta_i - \beta_j)\right| \geqslant |y||\beta_i - \beta_j| - 1 \tag{3.2.4}$$

for all $1 \leqslant j \leqslant r$. For any $x, y \in \mathbb{R}$ and $\alpha \in \mathbb{C}$ we have the equivalence

$$|x - \alpha y| \geqslant 1 \Leftrightarrow |\overline{x - \alpha y}| \geqslant 1$$

therefore (3.2.4) is equivalent to

$$\frac{1}{|L_j(x, y)|} \geqslant |y||\beta_i - \overline{\beta_j}| - 1 \tag{3.2.5}$$

From (3.2.4) and (3.2.5) we obtain the following

$$\frac{2}{|L_j(x, y)|} \geqslant |y||\beta_i - \overline{\beta_j}| - 1 + |y||\beta_i - \beta_j| - 1 \geqslant |y||2\beta_i - (\beta_j + \overline{\beta_j})| - 2 = |y||2\beta_i - 2Re(\beta_j)| - 2$$

that is we have
$$\frac{1}{|L_j(x,y)|} \geqslant |y||\beta_i - Re(\beta_j)| - 1$$
For all $j \neq i$ there exists an integer m, depending on β_j, that is on the solution (x,y) such that $|m - Re(\beta_j)| \leqslant \frac{1}{2}$, then

$$\frac{1}{|L_j(x,y)|} \geqslant |y||\beta_i - m + (m - Re(\beta_j))| - 1 \geqslant |y|\left(|\beta_i - m| - \frac{1}{2}\right) - 1 \qquad (3.2.6)$$

As we mentioned before, we only deal with solutions $(x,y), y > 0$. Now, we again classify the solutions (x,y) of $(1.0.2)$, with $1 \leqslant y \leqslant M$. We note that the set of all the solution (x,y) with $H(x,y) \leqslant M$ is covered in this way. Denote X_i the set of solutions (x,y) such that $1 \leqslant y \leqslant M$ and $|L_i(x,y)| \leqslant \frac{1}{2y}$. Among the elements of X_i we have the following relations.

Lemma 3.2.1. *Let $(x_1,y_1),(x_2,y_2) \in X_i$ be two different elements and suppose $y_1 \leqslant y_2$. Then*
$$y_2/y_1 \geqslant \frac{2}{7}\max(1, |\beta_i - m|) \qquad (3.2.7)$$

Proof. As $(x_1,y_1),(x_2,y_2) \in X_i$ are different solutions we have that $(x_1y_2 - x_2y_1) \neq 0$, also it is an integer number, so $|x_1y_2 - x_2y_1| \geqslant 1$. On the other hand

$$|x_1y_2 - x_2y_1| = \left|\det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\right| = \left|\det\begin{pmatrix} x_1 - \alpha_i y_1 & y_1 \\ x_2 - \alpha_i y_2 & y_2 \end{pmatrix}\right| = |y_2(x_1 - \alpha_i y_1) - y_1(x_2 - \alpha_i y_2)|$$

$$\leqslant y_1|L_i(x_2,y_2)| + y_2|L_i(x_1,y_1)| \leqslant \frac{y_1}{2y_2} + y_2|L_i(x_1,x_2)| \leqslant \frac{1}{2} + y_2|L_i(x_1,y_1)|$$

Therefore $y_2|L_i(x_1,y_1)| \geqslant \frac{1}{2}$. Then this relation together with $(3.2.6)$ imply that

$$y_2 \geqslant \frac{1}{2|L_i(x_1,y_1)|} \geqslant \frac{y_1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2}$$

That is

$$\frac{y_2}{y_1} \geqslant \frac{1}{2y_1|L_i(x_1,y_1)|} \geqslant \frac{1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2y_1} \geqslant \frac{1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2}$$

or

$$\frac{y_2}{y_1} \geqslant \max\left(1, \frac{1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2}\right)$$

Also, it can be easily checked that $\max(1, \frac{1}{2}z - \frac{3}{4}) \geqslant \frac{2}{7}\max(1,z)$ for any real number $z \geqslant 0$. Thus,

$$\frac{y_2}{y_1} \geqslant \frac{1}{2|L_i(x_1,y_1)|} \geqslant \frac{1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2y_1} \geqslant \frac{1}{2}\left(|\beta_i - m| - \frac{1}{2}\right) - \frac{1}{2} \geqslant \frac{2}{7}\max(1, |\beta_i - m|)$$

□

According to this fact, if a solution (x,y) with $y>0$ belongs to a class X_i, for some $1\leqslant i\leqslant r$ i.e., $|L_i(x,y)|\leqslant \frac{1}{2y}$, then for other elements $(x',y')\in X_i$, with $y\leqslant y'$ we have $|y'/y|\geqslant \frac{2}{7}\max(1,|\beta_i-m|)$. Now we want to investigate the case when (x,y) is a solution with $1\leqslant y\leqslant M$ not belonging to X_i, i.e., $|L_i(x,y)|>\frac{1}{2y}$. In this case, from (3.2.6) we obtain that

$$2y>\frac{1}{|L_j(x,y)|}\geqslant y\left(|\beta_i-m|-\frac{1}{2}\right)-1$$

Then dividing both sides by positive integer y we obtain

$$\frac{7}{2}\geqslant 2+\frac{1}{y}+\frac{1}{2}>|\beta_i-m| \tag{3.2.8}$$

We consider the set X, which consist of all solutions of the equation $|F(x,y)|=1$, with $1\leqslant y\leqslant M$, but the elements with the largest y from each class X_i, $i=1,...,r$ is excluded if X_i is not empty, i.e., at most r elements removed.

At the beginning we took a prime p without any conditions on it. Now, we consider a prime number p, such that $p>\left(\frac{7}{2}\right)^2$. Then we have the following theorem.

Theorem 3.2.1. *For any $\varepsilon>0$ and $r>r_1(p,\varepsilon)$ the cardinality of the set X satisfies the inequality*

$$|X|<\frac{r(1+\varepsilon)}{1-(2\log\left(\frac{7}{2}\right))/\log p}$$

for some $r_1(p,\varepsilon)>0$, depending on prime number p and ε.

Proof. Take some X_i and order the elements $(x_1,y_1),...,(x_v,y_v)$, such that $y_1\leqslant ...\leqslant y_v$. Then according to the definition, the solution (x_v,y_v) is not included in X. Then for these elements from (3.2.7) we have that

$$y_{k+1}/y_k\geqslant\frac{2}{7}\max(1,|\beta_i(x_k,y_k)-m(x_k,y_k)|) \tag{3.2.9}$$

for $k=1,...,v-1$. Therefore,

$$\prod_{x\in X_i\cap X}\frac{2}{7}\max(1,|\beta_i(x_k,y_k)-m(x_k,y_k)|)\leqslant\frac{y_v}{y_{v-1}}\cdot\frac{y_{v-1}}{y_{v-2}}\cdot...\cdot\frac{y_2}{y_1}=\frac{y_v}{y_1}\leqslant y_v\leqslant M$$

and for other elements (x,y) in X but not in X_i, from (3.2.8) we have that

$$\frac{2}{7}\max(1,|\beta_i(x_k,y_k)-m(x_k,y_k)|)\leqslant 1$$

Therefore

$$\prod_{x\in X}\frac{2}{7}\max(1,|\beta_i(x_k,y_k)-m(x_k,y_k)|)\leqslant M \tag{3.2.10}$$

We had that $G(v,w)=\prod_{i=1}^{r}(v-\beta_i w)$ is equivalent to F. Also, we should note that G is equivalent to $\hat{G}=\prod_{i=1}^{r}(v-(\beta_i-m)w)$, therefore $F\sim\hat{G}$. We assumed that F has the smallest height in its equivalence class, therefore $M(\hat{G})\geqslant M(F)$.

Hence

$$\prod_{i=1}^{r}\max(1,|\beta_i(x,y)-m(x,y)|)=M(\hat{G})\geqslant M(F)$$

Using the fact (3.2.10) for all $i = 1, ...r$ and taking product of all of them we obtain

$$\left(\left(\frac{2}{7}\right)^r M(F)\right)^{|X|} \leqslant M^r$$

Since $p > \left(\frac{7}{2}\right)^2$ we have that $M(F) > \left(\frac{7}{2}\right)^r$ for some $r > r_0(p)$ big enough and then

$$|X| \leqslant \frac{r \log M}{\log M(F) - r \log\left(\frac{7}{2}\right)}$$

According to the choice of A_1, t, τ and λ we have that

$$A_1 = \frac{1}{a^2}\left(\log M(F) + \frac{1}{2}r\right)$$

Also, we chose M as

$$M = (2C)^{\frac{1}{r-\lambda}}(4e^{A_1})^{\frac{\lambda}{r-\lambda}}$$

therefore,

$$\log M = \frac{r}{r-\lambda}\left(\log M(F) + \log(2r^{\frac{1}{2}}) + \frac{\log 2}{r}\right) + \frac{\lambda}{r-\lambda}\left(\log 4 + \frac{1}{a^2}\left(\log M(F) + \frac{r}{2}\right)\right)$$

According to the choice of t and τ we have $t = \sqrt{\frac{2}{r+a^2}}$, $\tau = bt$ with $0 < a < b < 1$ therefore $\lambda = \frac{2}{(1-b)t}$. Using these fact we obtain

$$\log M = (1 + O(r^{-\frac{1}{2}}))\log M(F) + O(r^{\frac{1}{2}})$$

Taking into account the fact that $\log M(F) > r\log\left(\frac{7}{2}\right)$ we obtain

$$\log M < \left(1 + \frac{\varepsilon}{2}\right)\log M(F)$$

for $r > r_1(\varepsilon)$ sufficiently large satisfying the condition $r > r_0(p)$ too. Then

$$|X| < r\left(1 + \frac{\varepsilon}{2}\right)\frac{\log M(F)}{\log M(F) - r\log\left(\frac{7}{2}\right)} = r\left(1 + \frac{\varepsilon}{2}\right)\frac{1}{1 - r\log\left(\frac{7}{2}\right)/\log M(F)} \qquad (3.2.11)$$

Note that we are considering the polynomial with $D(F) \geqslant p^{r(r-1)}$ also from the relation

$$D(F) \leqslant r^r M(F)^{2r-2}$$

we have that

$$M(F) \geqslant p^{r/2} r^{-r/(2r-2)}$$

Then applying this to (3.2.11) we obtain

$$|X| < r\left(1 + \frac{\varepsilon}{2}\right)\frac{1}{1 - r\log\left(\frac{7}{2}\right)/\log M(F)} < r\left(1 + \frac{\varepsilon}{2}\right)\frac{1}{1 - \log\left(\frac{7}{2}\right)/(\frac{1}{2}\log p - ((\log r)/(2r-2))}$$

$$< r(1+\varepsilon)\frac{1}{1 - \log\left(\frac{7}{2}\right)/(\frac{1}{2}\log p)}$$

Now we sum up and make the things more concrete choosing a prime p. In fact, by considering the solutions (x, y) with $H(x, y) \geqslant M$ and the solutions with $y \leqslant M$ (which contains all the solutions with $H(x, y) \leqslant M$) of the equation $|F(x, y)| = 1$ we cover all the solutions. Therefore it is enough to obtain upper bounds for the number of the solutions in each of these sets. We have shown that in the first case the bound is $3r$. In the second case, the set X contains all the solutions (x, y) with $y \leqslant M$, except at most r numbers which has highest heights from each set X_i, $i = 1, ..., r$ and the solution $(1, 0)$. Therefore we have an upper bound

$$|X| + 4r + 1$$

If $p > \left(\frac{7}{2}\right)^2$ and r is large enough, then we obtained that

$$|X| < r(1+\varepsilon)\frac{1}{1 - \log\left(\frac{7}{2}\right)/(\frac{1}{2}\log p)}$$

Now choose $p = 19$, then we have that $|X| < 211r$.

In conclusion, for $r > c$, where c big enough, which also satisfies all the conditions we required for r in the last sections, we have that $N_r < 215r$, that is the number of integer solutions of the equation (1.0.2) is bounded by $215r$ from above.

□

Bibliography

[1] E.Bombieri, W.M.Schmidt: On Thue's equation. Invent. math. **88**, 69-81 (1983)

[2] Evertse, J.H.: Upper bounds for the number of solutions of Diophantine equations. Math. Centrum. Amsterdam, 1-127 (1983)

[3] Yuri Bilu: Lecture notes on p-adic numbers and Diophantine equations

[4] E.Bombieri, J.Mueller: On effective measures of irrationality for $\sqrt[r]{\frac{a}{b}}$ and related numbers. J.Reine Angew. Math. **342**, 173-196 (1983)

[5] Lewis, D., Mahler, K.: Representation of integers by binary forms. Acta Arith. **6**, 333-363 (1961)

[6] Evertse, J-H.: Lecture notes on Diophantine Approximations delivered in the Spring term in Leiden University, 2014.

[7] Shidlovskiy A.B.: Diophantine approximations and transcendental numbers / Diofantovy priblizheniya i transtsendentnye chisla (Russian), (1982)