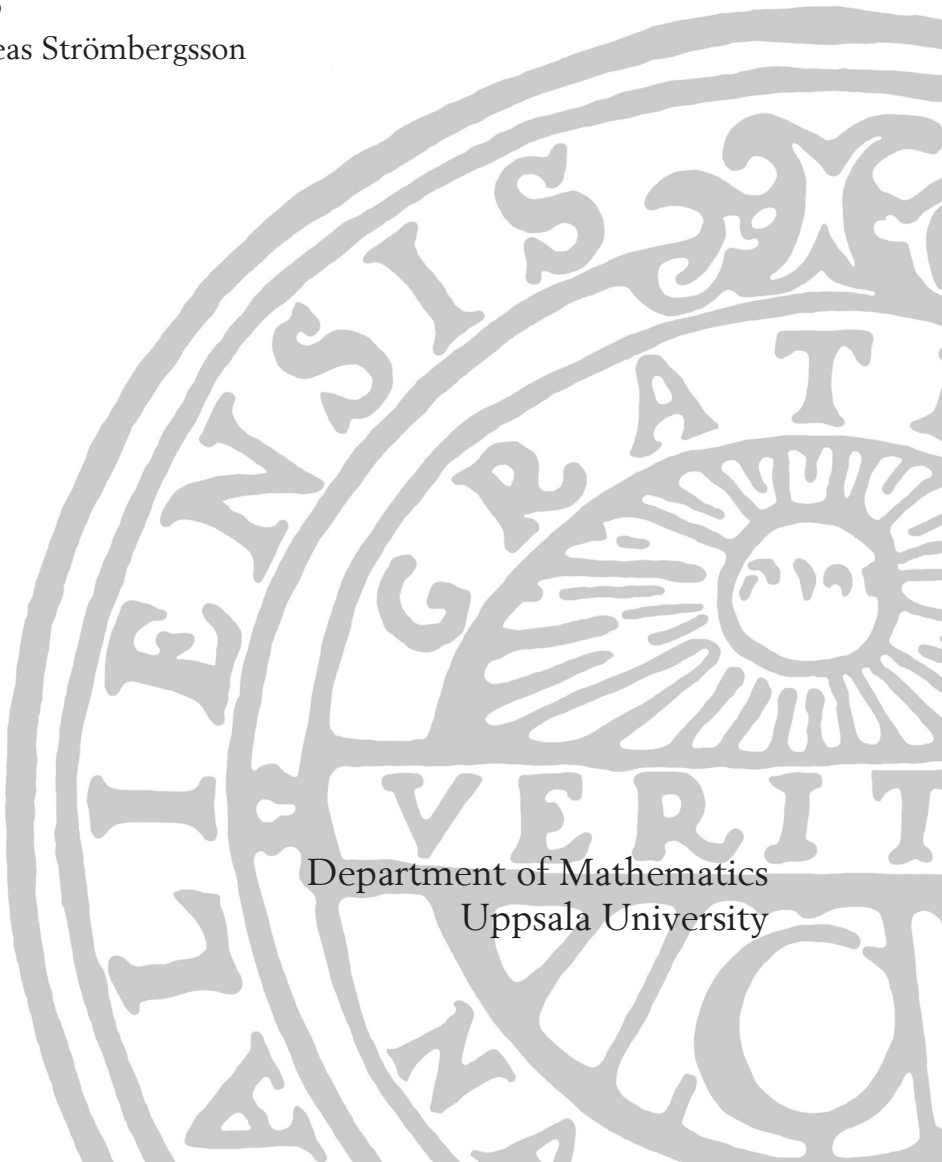
UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2008:12

The Thue–Siegel–Roth Theorem

Daniel Ishak

Examensarbete i matematik, 15 hp
Handledare och examinator: Andreas Strömbergsson
Juni 2008



Department of Mathematics
Uppsala University

THE THUE–SIEGEL–ROTH THEOREM

DANIEL ISHAK

Abstract. In this paper we will give a proof of the Thue-Siegel-Roth Theorem, which states that for any algebraic number α and any $\epsilon > 0$ there exists only a finite number of pairs of coprime integers p, q such that $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\epsilon}}$. We will follow the proof as it is presented Leveque's book, [8, ch 4]. This proof also deals with the more general case when $\frac{p}{q}$ is allowed to be an algebraic number in some fixed number field.

Contents

1. Background

Some properties of rational numbers make them easier to work with than irrational numbers. Because of this mathematicians have tried to approximate irrational numbers by rational numbers. Let n be a given positive integer. It can easily be shown by looking at the Farey sequence of order n (see [1]), $\mathfrak{F}_n$, that for every irrational number ξ there exists an irreducible fraction, $\frac{p}{q}$, such that

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q(n+1)},\ 0 < q \leq n.$$

And thus, by looking at Farey sequences of higher orders, this implies that there are infinitely many irreducible fractions $\frac{p}{q}$ satisfying the inequality

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Theorem 1.1 (Liouville's theorem)**.** *For any algebraic number α of degree $n > 1$, there exists a constant $c = c(\alpha)$ such that*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^n}$$

for all rationals p/q, with $q > 0$.

Proof. First note that this is trivial for $\left|\alpha - \frac{p}{q}\right| > 1$ since we can let $c < 1$.

Let $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ be the minimal polynomial having α as a root. It is assumed that $f(z) \in \mathbb{Z}[z]$ is irreducible over $\mathbb{Z}$, and thus also over $\mathbb{Q}$ by Gauss's lemma. So $f(\frac{p}{q}) \neq 0$ and

$$\left|q^n f(\frac{p}{q})\right| \geq 1. \tag{1}$$

Using the mean value theorem we get that $|f(\frac{p}{q})| = |(\alpha - \frac{p}{q})f'(\xi)|$ for some ξ between α and p/q. Using (1) with the fact that there exists a constant c such that $|f'(\xi)| < \frac{1}{c}$ we conclude that

$$\left|\alpha - \frac{p}{q}\right| = \frac{|f(\frac{p}{q})|}{|f'(\xi)|} > \frac{c}{q^n}.$$

□

Assume $\alpha < \frac{p}{q}$. In the proof above letting $\frac{1}{c} > \sup_{z\in(\alpha,\alpha+1)} |f'(z)|$ is enough. The same is true when $\alpha > \frac{p}{q}$, but clearly with $z \in (\alpha - 1, \alpha)$. Combining them we get $z \in (\alpha - 1, \alpha + 1)$. To summarize this: In Theorem 1.1 we can let
$c = \min\left\{\frac{1}{2}, \left(\sup_{z\in(\alpha-1,\alpha+1)} |f'(z)|\right)^{-1}\right\}$.

In 1909 the Axel Thue (see [4]) showed that given an algebraic number α of degree n greater than one, the equation

$$0 < \left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\kappa},$$

where $\kappa > \frac{n}{2} + 1$, has only a finite number of solutions in terms of the coprime integers p and q. Carl Ludwig Siegel further relaxed the conditions in 1921 ([5]) to $\kappa > \min_{\substack{1 \leq s \leq n-1 \\ s \in \mathbb{Z}}} (s + n/(s+1))$ and in particular $\kappa > 2\sqrt{n}$. This was, independently, further developed to $\kappa > \sqrt{2n}$ by Freeman Dyson (see [6]) and Alexander Gelfond in 1947.

It was finally shown in 1955 by Klaus Roth (see [7]) that given $\epsilon > 0$, the inequality has only a finite number of solutions where $\kappa = 2 + \epsilon$.

Theorem 1.2 (Thue-Siegel-Roth's Theorem)**.** *Let α be an algebraic number. For every $\epsilon > 0$, there exists a $c = c(\alpha, \epsilon)$ such that*

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^{2+\epsilon}},$$

for all irreducible rationals p/q.

We say that α is approximable to order 2 and to no higher. In this paper we will give a proof of this theorem, but generalized to the case of algebraic number in place of rationals (see Theorem 7.1). We will, as mentioned earlier, very closely follow the proof given by LeVeque in [8, ch 4]. Along the way we will point out a couple of mistakes in the proof in [8], which we have unfortunately not been able to work around (although we certainly believe that this is possible). We stress that these problems do *not* occur in the special case of working over the rationals, i.e. they do not pertain to the case of Theorem 1.2 above.

One interesting fact connecting Roth's theorem with (simple) continued fractions follows from the next theorem (cf. [2, Theorem 184]).

Theorem 1.3. *Let x be irrational. If*

$$\left| \frac{p}{q} - x \right| < \frac{1}{2q^2},$$

then p/q is a convergent.

The above theorem thus reduces the Thue-Siegel-Roth Theorem to checking the convergents of the algebraic number α. Although this seems very helpful, not enough is know about the continued fractions of an algebraic number in order to use this theorem in solving the Thue-Siegel-Roth Theorem. Another theorem involving simple continued fractions follows as.

Theorem 1.4. *If x is an irrational with a periodic continued fraction, then x is approximable to order two and no higher order.*

Proof. Let $x = [a_0, a_1, \dots]$. The quotients of the continued fraction are bounded, so there exists an M such that

$$0 < a_k < M,$$

for all $k \in \mathbb{N}$. We now define some variables.

$$\frac{p_n}{q_n} = [a_0, \dots, a_n],$$
$$a_n' = [a_n, a_{n+1}, \dots]$$

and

$$q_n^{'} = a_n^{'} q_{n-1} + q_{n-2}.$$

Since

$$a_n^{'} = a_n + \frac{1}{a_{n+1}^{'}}$$

for $n > 0$, we have that $a_n < a_n^{'} < a_n + 1$. So

$$q_{n+1}^{'} = a_{n+1}^{'} q_n + q_{n-1} < (a_{n+1} + 1)q_n + q_{n-1} < (M+2)q_n,$$

and since $q_n = a_n q_{n-1} + q_{m-2}$ (see [2, ch 10.2]), we get that

$$q_{n+1} = a_{n+1} q_n + q_{m-1} < a_{n+1}^{'} q_n + q_{m-1} = q_{n+1}^{'} < (M+2)q_n,$$

and similarly $q_n < (M+2)q_{n-1}$. Suppose that $q_{n-1} < q \leq q_n$. Then $q_n < (M+2)q$ and by [2, Theorem 181]

$$\left| \frac{p}{q} - x \right| \geq \left| \frac{p_n}{q_n} - x \right| = \frac{1}{q_n q_{n+1}^{'}} > \frac{1}{(M+2)q_n^2} > \frac{1}{(M+1)^3 q_{n-1}^2} > \frac{K}{q^2},$$

where $K = (M+2)^{-3}$ and the first equality follows from [2, Theorem 163]. □

Note that the proof holds for a stronger theorem, viz.

Theorem 1.5. *If x is an irrational with bounded continued fraction, then x is approximable to order two and no higher order.*

The main idea of the proof for the generalized version of Theorem 1.2 is to assume that there are infinitely many solutions, and by choosing m solutions $\zeta_1, \ldots, \zeta_m$, such that m satisfies an inequality $G(m) < 2 + \epsilon$, we can construct a polynomial $Q(z_1, \cdots, z_m)$ with special properties. These m solutions were chosen such that when considering the number $Q(\zeta_1, \ldots, \zeta_m)$, we can deduce that $G(m) > 2 + \epsilon$, which contradicts our previous inequality. The major part of this paper will be dedicated to finding this special polynomial $Q(z_1, \ldots, z_m)$ and describing its properties.

2. Polynomials

Let $P(z) = a_0 + a_1 z + \cdots + a_n z^n$ be a polynomial with complex coefficients. We start off with some definitions.

Definition 2.1. Let P be the polynomial above. Then $||P|| = \max\{|a_0|, |a_1|, \ldots, |a_n|\}$. Furthermore, if α is algebraic over $\mathbb{Q}$ with its minimal polynomial $f(z)$ over $\mathbb{Q}$, we define the *height* $H(\alpha) = ||f||$.

Theorem 2.1. *Let*

$$L(z) = l\prod_{k=1}^{h}(z - \lambda_k)$$

where $l, \lambda_k \in \mathbb{C}$ *for* $k = 1, \ldots, h$. *Then*

$$|l|\prod_{k=1}^{h}(1 + |\lambda_k|) \leq 6^h||L||. \tag{2}$$

Proof. It is easily seen that the constant l can be excluded. We now arrange the complex numbers $\lambda_1, \ldots, \lambda_h$ in such a way that $|\lambda_i| \leq 2$ for $i = 1, \ldots, t$ and $|\lambda_i| > 2$ for $i > t$.

For each $k = t + 1, ..., h$ we use a trick which involves dividing by $|z_0 - \lambda_k|$, where z_0 is a $(t + 1)$th root of unity. So

$$\frac{1 + |\lambda_k|}{|z_0 - \lambda_k|} \leq \frac{1 + |\lambda_k|}{|\lambda_k| - |z_0|} = \frac{1 + |\lambda_k|}{|\lambda_k| - 1} = 1 + \frac{2}{|\lambda_k| - 1} < 1 + \frac{2}{2 - 1} = 3,$$

and thus

$$\prod_{k=t+1}^{h}(1 + |\lambda_k|) < 3^{h-t}\left|\prod_{k=t+1}^{h}(z_0 - \lambda_k)\right|. \tag{3}$$

For the remaining $k = 1, 2, \ldots, t$ we get

$$\prod_{k=1}^{t}(1 + |\lambda_k|) \leq (1 + 2)^t = 3^t.$$

For now assume $|f(z_0)| \geq 1$, where $f(z) = \prod_{k=1}^{t}(z - \lambda_k)$ and z_0 is the previous $(t+1)$th root of unity. Using this assumption we obtain

$$\prod_{k=1}^{t}(1 + |\lambda_k|) \leq 3^t \leq 3^t\left|\prod_{k=1}^{t}(z_0 - \lambda_k)\right|,$$

and this together with (3) gives

$$\begin{aligned}\prod_{k=1}^{h}(1 + |\lambda_k|) &= \left(\prod_{k=1}^{t}(1 + |\lambda_k|)\right)\left(\prod_{k=t+1}^{h}(1 + |\lambda_k|)\right) \\ &< 3^h\left|\prod_{k=1}^{h}(z_0 - \lambda_k)\right| \leq 3^h||L||(|z_0|^h + \cdots + 1) \\ &= 3^h||L||(h + 1) \leq 6^h||L||.\end{aligned}$$

The only thing remaining is showing that such a z_0 in fact does exist. First let ϵ be a $(t+1)$th root of unity. Our objective is to show that the sum $\sum_{v=0}^{t} |f(\epsilon^v)| \geq t+1$. Set

$$f(z) = \sum_{r=0}^{t} \mu_r z^r, \qquad \mu_t = 1.$$

Then

$$\sum_{v=0}^{t} \epsilon^v f(\epsilon^v) = \sum_{v=0}^{t} \left(\epsilon^v \sum_{r=0}^{t} \mu_r \epsilon^{vr} \right) = \sum_{r=0}^{t} \left(\mu_r \sum_{v=0}^{t} \epsilon^{v(r+1)} \right).$$

In order to evaluate this we focus on the sum $\sum_{v=0}^{t} \epsilon^{v(r+1)}$. If $(t+1)|(r+1)$ this sum clearly equals $t+1$. Since $0 \leq r \leq t$ this is only true for $r = t$. If $(t+1) \nmid (r+1)$, this is just a case of the geometric sum

$$\sum_{v=0}^{t} z^v = \frac{z^{n+1} - 1}{z - 1},$$

and so the sum is zero. Thus

$$\sum_{v=0}^{t} \epsilon^v f(\epsilon^v) = \mu_t(t+1) = t+1.$$

Hence

$$\sum_{v=0}^{t} |f(\epsilon^v)| = \sum_{v=0}^{t} |\epsilon^v f(\epsilon^v)| \geq \left| \sum_{v=0}^{t} \epsilon^v f(\epsilon^v) \right| = t+1$$

as desired. Hence there exists some root of unity $z_0 \in \{1, \epsilon, \epsilon^2, \ldots, \epsilon^t\}$ such that $|f(z_0)| \geq 1$, and this concludes the proof. □

Some more theorems which we will need later will now follow.

Theorem 2.2. *Let $f(z)$ and $g(z)$ be complex polynomials of degree n and m respectively. Suppose the coefficient of z^m in $g(z)$ has absolute value greater than or equal to 1. Then*

$$||f|| \leq 6^{m+n} ||fg||.$$

This is a simple consequence of Theorem 2.1. See [8, Theorem 4-3].

Theorem 2.3. *Let $f(z)$ be a polynomial of degree n, having real coefficients. Then*

$$||f||^m \leq (mn+1)||f^m||.$$

The following is just a sketch of the proof. See [8, Theorem 4-4] for the full proof.

Proof. Let $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ and let $||f|| = a$. Firstly, we note that the theorem is true if either $a = |a_0|$ or $a = |a_n|$. Secondly, it can be shown that we can with no loss in generality assume that $a = |a_t|$ where $n/2 \leq t < n$. Now let

$$g(z, \theta) = f(z) - ae^{i\theta} z^n,$$

and let $\alpha = \alpha(\theta)$ be the numerically largest of the zeros of $g(z, \theta)$. Given any θ,

$$f(\alpha) = ae^{i\theta} \alpha^n$$

and thus

$$|f^m(\alpha)| = \left| ae^{i\theta} \alpha^n \right|^m.$$

If $|\alpha| \geq 1$ we will have the inequality

$$(1 + |\alpha| + \cdots + |\alpha|^{mn}) \leq (mn + 1)|\alpha|^{mn}.$$

Assuming $|\alpha| \geq 1$,

$||f||^m|\alpha|^{mn} = a^m|\alpha|^{mn} = |f^m(\alpha)| \leq ||f^m||(1 + |\alpha| + \cdots + |\alpha|^{mn}) \leq (mn+1)\,||f^m||\,|\alpha|^{mn}$,

and so $||f||^m \leq (mn+1)\,||f^m||$. The problem has now reduced to show that there exists a θ such that $|\alpha(\theta)| \geq 1$. □

We now turn to two definitions.

Definition 2.2. Let α is an algebraic number of degree n, with the corresponding minimal polynomial $p(z)$. Then the roots, $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$, of $p(z)$ are called the conjugates of α.

Definition 2.3. Let $P(z)$ be a polynomial with algebraic coefficients $c_1, c_1, \ldots, c_r$. Also let C_i equal the maximum of the absolute values of the conjugates of c_i. Then we define $|\overline{P}|$ as $|\overline{P}| = \max\{C_1, C_2, \ldots, C_r\}$. If α is algebraic we also define $|\overline{\alpha}|$ as the maximum of the absolute values of of the conjugates of α.

Theorem 2.4. *Let $f_1(z), f_2(z), \ldots, f_t(z)$ be polynomials with algebraic coefficients. Then*

$$\left|\overline{\prod_{v=1}^{t} f_v}\right| \leq \prod_{v=1}^{t}(1 + \deg f_v) \prod_{v=1}^{t} \left|\overline{f_v}\right|.$$

Proof. Suppose, without loss in generality, that $\deg f_1 \geq \deg f_2 \geq \cdots \geq \deg f_t$.

Each coefficient of $f_1 f_2$ will be the sum of products of a coefficient of f_1 and a coefficient of f_2. Suppose $\alpha\beta$ is such a product. Then each conjugate of $\alpha\beta$ could be written as the product of a conjugate of α and a conjugate of β, thus $|\overline{\alpha\beta}| \leq |\overline{\alpha}|\,\left|\overline{\beta}\right|$. The same argument is used to show that $\left|\overline{\alpha + \beta}\right| \leq |\overline{\alpha}| + |\overline{\beta}|$. Since the number of sums of products will be at most $\deg f_2 + 1$ we get that

$$|\overline{f_1 f_2}| \leq (\deg f_2 + 1)|\overline{f_1}||\overline{f_2}|.$$

The argument is repeated and the result follows. □

Theorem 2.5. *Let p and r be positive integers with $1 \leq r < p$. Suppose that $F(z_1, \ldots, z_p)$, $G(z_1, \ldots, z_r)$ and $H(z_{r+1}, \ldots, z_p)$ are polynomials with coefficients in an algebraic number field K, those of F being integers. Also suppose that*

$$F(z_1, \ldots, z_p) = G(z_1, \ldots, z_r)H(z_{r+1}, \ldots, z_p).$$

Then if γ is any coefficient in F, there is a factorization $\gamma = \alpha\beta$ in K such that the coefficients in αH and βG are integers in K.

This is not hard to show. See [8, Theorem 4-6] for the proof.

3. Generalized Wronskians

Now we turn our attention to Wronskians. The results from this section will be needed in the next one.

Throughout this section K will be an algebraic number field and $f_0, \dots, f_{l-1} \in K[z_1, \dots, z_p]$.

Definition 3.1. $f_0, \dots, f_{l-1}$ are said to be linearly independent if

$$k_0 f_1 + \cdots k_{l-1} f_{l-1} = 0,\ k_i \in K \text{ for each } i = 0, \dots, l-1,$$

has only the trivial solution; $k_0, \dots, k_{l-1}$ are all zeros in K. If $f_0, \dots, f_{l-1}$ are not linearly independent they are said to be linearly dependent.

If $p = 1$, then we define the Wronskian as

$$W(z) = \begin{vmatrix} \frac{1}{0!} f_0(z) & \frac{1}{0!} f_1(z) & \cdots & \frac{1}{0!} f_{l-1}(z) \\ \frac{1}{1!} f_0'(z) & \frac{1}{1!} f_1'(z) & \cdots & \frac{1}{l!} f_{l-1}'(z) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(l-1)!} f_0^{(l-1)}(z) & \frac{1}{(l-1)!} f_1^{(l-1)}(z) & \cdots & \frac{1}{(l-1)!} f_{l-1}^{(l-1)}(z) \end{vmatrix}$$

$$= \det\left(\frac{1}{\mu!} \frac{d^\mu}{dz^\mu} f_v(z) \right),\ \mu, v = 0, 1, \dots, l-1.$$

The difference between this definition of the Wronskian and the usual one is the precense of the factors $\frac{1}{\mu!}$, $\mu = 0, \dots, l-1$. We define the Wronskian in this way because it will be convenient for us.

The Generalized Wronskian deals with all the cases $p \geq 1$. Let $\Delta_0, \dots, \Delta_{l-1}$ be differential operators of the form

$$\Delta_\mu = \frac{1}{j_1! \cdots j_p!} \left(\frac{\partial}{\partial z_1} \right)^{j_1} \cdots \left(\frac{\partial}{\partial z_p} \right)^{j_p}$$

such that the order, $j_1 + \cdots j_p$, och Δ_μ does not exceed μ for any $\mu = 0, \dots, l-1$. We now define the generalized Wronskian as

$$G(Z) = \begin{vmatrix} \Delta_0 f_0(z) & \Delta_0 f_1(z) & \cdots & \Delta_0 f_{l-1}(z) \\ \Delta_1 f_0(z) & \Delta_1 f_1(z) & \cdots & \Delta_1 f_{l-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{l-1} f_0(z) & \Delta_{l-1} f_1(z) & \cdots & \Delta_{l-1} f_{l-1}(z) \end{vmatrix}.$$

Since there are several different Δ_μ's, we see that there is not just one generalized Wronskian. The number of different Δ_μ's are $(p+1)^\mu$.

We now turn to two theorems which will be needed later. The proofs are rather long and will not be in this paper. For the full proofs see [8, Chapter 4-3].

Theorem 3.1. *(a) If $f_0, \dots, f_{l-1}$ are l polynomials over K in the single variable z, whose Wronskian $W(z)$ vanishes identically, then they are dependent.*

(b) If $f_0, \dots, f_{l-1}$ are l polynomials over K in the variables $z_1, \dots, z_p$, for which every generalized Wronskian $G_l(z_1, \dots, z_p)$ vanishes identically, then they are dependent.

Theorem 3.2. *Let $R(z_1, \ldots, z_p)$ be a polynomial in $p \geq 2$ variables, with integral coefficients in K such that*

$$0 < |\overline{R}| \leq B.$$

Let R be of degree at most r_j in z_j, for $j = 1, \ldots, p$. Then there is an l in $\mathbb{Z}$ with

$$1 \leq l \leq r_p + 1,$$

there is an integer β in K, and there are differential operators $\Delta_0, \ldots, \Delta_{l-1}$ on the variables $z_1, \ldots, z_{p-1}$ of orders at most $0, \ldots, l-1$ respectively, such that if

$$F(z_1, \ldots, z_p) = \beta \det \left(\Delta_\mu \frac{1}{v!} \left(\frac{\partial}{\partial z_p} \right)^v R \right), \ \mu, v = 0, \ldots, l-1,$$

then (a) F has integral coefficients in K and is not identically zero; (b) a decomposition

$$F(z_1, \ldots, z_p) = U(z_1, \ldots, z_{p-1}) V(z_p)$$

holds, where U and V have integral coefficients in K, U is of degree at most lr_j in z_j for $j = 1, \ldots, p-1$, and V is of degree at most lr_p in z_p; (c) the following bound holds:

$$|\overline{F}| \leq \{(r_1 + 1) \ldots (r_p + 1)\}^{2l} 2^{2(r_1 + \cdots + r_p)l} l!^2 B^{2l}.$$

4. The index

In this section $P(z_1, \ldots, z_p)$ is a non-zero polynomial. Also, $\overline{\alpha}_p = (\alpha_1, \ldots, \alpha_p)$ and $\overline{r}_p = (r_1, \ldots, r_p)$ are lists of complex numbers and positive numbers, respectively.

Now, expand the polynomial $P(\alpha_1 + y_1, \ldots, \alpha_p + y_p)$ in $y_1, \ldots, y_p$, i.e.,

$$P(\alpha_1 + y_1, \ldots, \alpha_p + y_p) = \sum_{j_1=0}^{\infty} \cdots \sum_{j_p=0}^{\infty} c(j_1, \ldots, j_p) y_1^{j_1} \cdots y_p^{j_p}.$$

Definition 4.1. The index θ of P at the point $\overline{\alpha}_p$* relative to $\overline{r}_p$ is

$$\theta = \min \left(\frac{j_1}{r_1} + \cdots + \frac{j_p}{r_p} \right),$$

where the minimum is taken over the set of p-tuples $(j_1, \ldots, j_p)$ of non-negative integers for which $c(j_1, \ldots, j_p) \neq 0$.

Theorem 4.1. *Let $P(z_1, \ldots, z_p)$ and $Q(z_1, \ldots, z_p)$ be non-zero polynomials. If the indices are considered at the same point $\overline{\alpha}_p$ relative to the same list of numbers $\overline{r}_p$, then the following relations hold:*

$$\textit{index } (P + Q) \geq \min(\textit{index } P, \textit{index } Q),$$

$$\textit{index } PQ = \textit{index } P + \textit{index } Q.$$

Proof. This is easy to verify. □

Let K be an algebraic number field, $B \geq 1$ and consider the set $\mathcal{R}_m = \mathcal{R}_m(B; \overline{r}_m)$ of polynomials $R \in K[z_1, \ldots, z_m]$ with the properties:

(i) R is of degree at most r_j in z_j, for $j = 1, \ldots, m$,
(ii) and $|\overline{R}| \leq B$.

Let $\zeta_1, \ldots, \zeta_m$ be algebraic numbers in $\mathbb{C}$ of heights $q_1, \ldots, q_m$ respectively. Let $\theta(R)$ denote the index of R at the point $\overline{\zeta}_m = (\zeta_1, \ldots, \zeta_m)$ relative to $\overline{r}_m$. We now go on and make a definition which will be useful.

Definition 4.2. If we iterate over all $R \in \mathcal{R}_m$ and all lists $\overline{\zeta}_m$ with elements of heights $q_1, \ldots, q_m$†, then Θ_m is defined as

$$\Theta_m(B; \overline{q}_m; \overline{r}_m) = \sup\{\theta(R)\}.$$

Our aim is to find an upper bound for $\Theta_m(B; \overline{q}_m; \overline{r}_m)$. We will use induction on m and start off with the theorem which will later be used in the case $m = 1$.

Theorem 4.2.

$$\Theta_1(B; \overline{q}_1; \overline{r}_1) \leq \frac{3N(N+1)}{\log q_1} + \frac{N \log B}{r_1 \log q_1},$$‡

where $N = [K : \mathbb{Q}]$.

To prove this we first need to state a theorem which we will need.

*Note that we refer to $\overline{\alpha}_p$ as a point here.

†I.e., $H(\zeta_1) = q_1, \ldots, H(\zeta_m) = q_m$.

‡If we only consider the index of a polynomial at the point $\overline{\alpha}_p$ of real numbers, then the term $\frac{3N(N+1)}{\log q_1}$ could be neglected. This was what Roth did in his original proof. See [7].

Theorem 4.3 (Symmetric Function Theorem)**.** *If P is a polynomial in the roots of an equation $f(x) = 0$ of degree n, and if P is symmetric in $n-1$ of the roots, then P is equal to a polynomial, with integral coefficients, in the remaining root and the coefficients of $f(x)$ and P.*

Proof of Theorem 4.2. Let $\chi(z)$ be the defining polynomial of ζ_1, with degree h. We then have that

$$||\chi|| = H(\zeta_1) = q_1.$$

Assume $K = \mathbb{Q}(\gamma)$ and that $\gamma_1 = \gamma, \ldots, \gamma_k$ are the conjugates of γ. Thus K is a vector space over $\mathbb{Q}$ with the base $\{1, \gamma, \ldots, \gamma^{n-1}\}$ and every coefficient, c, of R is in $span\{1, \gamma, \ldots, \gamma^{n-1}\}$. Let $\overline{\gamma} = [1, \gamma, \ldots, \gamma^{n-1}]$ be $1 \times (n-1)$ matrix. Now, the coefficients of $R \in \mathcal{R}_1$ are in $\mathcal{O}_K$ and any coefficient of R can be written as a linear combination of the base vectors, say $c_i = \delta_0 + \delta_1\gamma^1 + \cdots + \delta_n\gamma^n = \overline{\gamma}\mathbf{X}^T$, where $\mathbf{X} \in \mathbb{Q}^n$. By letting $d_i = \overline{\gamma_i}\mathbf{X}^T$, where γ_i is a conjugate of γ, we get another coefficient. If we replace the coefficient c_i in R by d_i, where γ_j is a conjugate of γ, we get a new polynomial R_2. Doing this with all the coefficients, and not just one by one, and multiply the N polynomials we get R^*, where $\deg R^* = Nr_1$. By the Symmetric Function Theorem, R^* has coefficients in $\mathbb{Z}$. By Theorem 2.4 and the fact that $R \in \mathcal{R}_1$

$$||R^*|| \leq (1 + r_1)^N B^N. \tag{4}$$

We also have that

$$R(\zeta_1 + y_1) = y_1^{r_1\theta} P(y_1),$$

where P is a polynomial, and so by the definition of the index and substituting $y_1 = z_1 - \zeta_1$, $(z_1 - \zeta_1)^{r_1\theta} | R(z_1)$. Since $R^* \in \mathbb{Z}[z]$, and from the clear fact that χ is the minimal polynomial of all the conjugates of γ, $R^*(z_1)$ is divisible by $\chi^{r_1\theta}$ and therefore $hr_1\theta \leq Nr_1$. Now let $Q = R^*/\chi^{r_1\theta}$ of degree $Nr_1 - hr_1\theta$. Q is a polynomial with coefficients in $\mathbb{Z}$, of which the coefficients of $z^{Nr_1 - hr_1\theta}$ clearly has an absolute value of greater than or equal to 1. By applying Theorem 2.2,

$$||\chi^{r_1\theta}|| \leq 6^{Nr_1 - hr_1\theta + hr_1\theta} ||Q\chi^{r_1\theta}|| = 6^{Nr_1} ||R^*||. \tag{5}$$

Now it follows that

$$\begin{aligned} q_1^{r_1\theta} = ||\chi||^{r_1\theta} &\leq (hr_1\theta + 1)||\chi^{r_1\theta}|| \\ &\leq (hr_1\theta + 1)6^{Nr_1}||R^*|| \\ &\leq (Nr_1 + 1)^{N+1} 6^{Nr_1} B^N \\ &< 2^{Nr_1(N+1)} 6^{Nr_1} B^N \\ &< 12^{N(N+1)r_1} B^N, \end{aligned}$$

where we made use of (5) and Theorem 2.3 in the second inequality, and finally (4) in the third inequality. Hence

$$\theta < \frac{N(N+1)\log 12}{\log q_1} + \frac{N \log B}{r_1 \log q_1}$$

and since $\log 12 < 3$, the theorem follows. □

Now we proceed and find a recurrence relation between Θ_m and Θ_{m-1}.

Theorem 4.4. *Let $p \geq 2$ be a positive integer, let $r_1, \dots, r_p$ be positive integers such that*

$$r_p > 10\delta^{-1}, \frac{r_{j-1}}{r_j} > \delta^{-1}, \text{ for } j = 2, \dots, p, \tag{6}$$

where $0 < \delta < 1$, and let $q_1, \dots, q_p$ be positive integers. Then

$$\Theta_p(B; \overline{q}_p; \overline{r}_p) \leq 2 \max(\Phi + \Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}}), \tag{7}$$

where the maximum is taken over integers l satisfying

$$1 \leq l \leq r_p + 1,$$

and where

$$\Phi = \Theta_1(M; q_p; lr_p) + \Theta_{p-1}(M; \overline{q}_{p-1}; \overline{lr}_{p-1})$$

and

$$M = (r_1 + 1)^{2pl} 2^{2r_1 pl} l!^2 B^{2l}.$$

Proof. If $R \in \mathcal{R}_p$ and $\zeta_1, \dots, \zeta_p$ are algebraic numbers of heights $q_1, \dots, q_p$ respectively, then we need to show that the index θ of R does not exceed the right hand side of (7).

Since R satisfies the hypothesis of Theorem 3.2, there are numbers l and β and a polynomial $F \in K[z_1, \dots, z_p]$ having the properties listed there. By (6), $r_1 > r_2 > \cdots > r_p$ and Theorem 3.2,

$$|\overline{F}| < \{(r_1 + 1) \dots (r_p + 1)\}^{2l} 2^{2(r_1 + \cdots + r_p)l} l!^2 B^{2l} < (r_1 + 1)^{2pl} 2^{2pr_1 l} l!^2 B^{2l} = M.$$

F can be factorized into

$$F(z_1, \dots, z_p) = U(z_1, \dots, z_{p-1}) V(z_p),$$

by Theorem 3.2. Clearly $|\overline{U}| < M$, $|\overline{V}| < M$[§] and by the same theorem z_j has degree at most lr_j for $j = 1, \dots, p$. Therefore $U \in \mathcal{R}_{p-1}(M; \overline{lr}_{p-1})$ and $V \in \mathcal{R}_1(M; lr_p)$. This means that the index of U at $(\zeta_1, \dots, \zeta_{p-1})$ relative to $lr_1, \dots, lr_{p-1}$ is at most $\Theta_{p-1}(M; \overline{q}_{p-1}; \overline{lr}_{p-1})$, and thus by definition the index of U at $(\zeta_1, \dots, \zeta_{p-1})$ relative to $r_1, \dots, r_{p-1}$ is at most $l\Theta_{p-1}(M; \overline{q}_{p-1}; \overline{lr}_{p-1})$. Similarly the index of V at ζ_p relative to r_p is at most $l\Theta(M; q_p; lr_p)$. Thus

$$index\ F = index\ (UV) = index\ U + index\ V \leq l\Phi,$$

and we now need to find a relation between θ and the index of F.

Consider any differential operator of the form

$$\Delta = \frac{1}{i_1! \cdots i_{p-1}!} \left(\frac{\partial}{\partial z_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial z_{p-1}}\right)^{i_{p-1}},$$

[§]This is not always true, but true when $K = \mathbb{Q}$. The coefficients of $F(z_1, \dots, z_p) = U(z_1, \dots, z_{p-1})V(z_p)$ are the product of a coefficient in U and a coefficients in V. However, for many K the group of units of $\mathcal{O}_K$, which we denote $U(\mathcal{O}_K)$, contains α's with $|\overline{\alpha}|$ arbitrarily large; note that for such α we have that $\beta = \alpha^{-1}$ *also* belongs to $\mathcal{O}_K$ (in fact to $U(\mathcal{O}_K)$), and $|\overline{\alpha\beta}| = |\overline{1}| = 1$ whereas both $|\overline{\alpha}|$ and $|\overline{\beta}|$ are large. This can happen for instance when K is a real quadratic number field.

of order $w \leq l-1$. If the polynomial

$$\Delta \frac{1}{v!}\left(\frac{\partial}{\partial z_p}\right)^v R(z_1,\dots,z_p)$$

is not equal to zero, its index at $(\zeta_1,\dots,\zeta_{p-1})$ relative to $r_1,\dots,r_{p-1}$ is at least

$$\theta - \frac{i_1}{r_1} - \dots - \frac{i_{p-1}}{r_{p-1}} - \frac{v}{r_p} \geq \theta - \frac{w}{r_{p-1}} - \frac{v}{r_p}.$$

We also know that

$$\frac{w}{r_{p-1}} \leq \frac{l-1}{r_{p-1}} \leq \frac{r_p}{r_{p-1}} \leq \delta.$$

This means that the index must be at least

$$\max\left(0, \theta - \delta - \frac{v}{r_p}\right) \geq \max\left(0, \theta - \frac{v}{r_p}\right) - \delta. \tag{8}$$

We know, from Theorem 3.2, that

$$F(z_1,\dots,z_p) = \beta \det\left(\Delta_\mu \frac{1}{v!}\left(\frac{\partial}{\partial z_p}\right)^v R\right),\ \mu,v = 0,\dots,l-1,$$

and by expanding the determinant we end up with $l!$ terms of the form

$$\pm\beta(\Delta_{\mu_0}R)\left(\Delta_{\mu_1}\frac{1}{1!}\frac{\partial}{\partial z_p}R\right)\cdots\left(\Delta_{\mu_{l-1}}\frac{1}{(l-1)!}\left(\frac{\partial}{\partial z_p}\right)^{l-1}R\right),$$

where $\Delta_{\mu_0},\dots,\Delta_{\mu_{l-1}}$ are of orders at most $l-1$. By (8) the index of such a term is at most

$$\sum_{v=0}^{l-1}\left(0,\theta-\frac{v}{r_p}\right) - l\delta,$$

thus

$$l\Phi \geq index\ F \geq \sum_{v=0}^{l-1}\left(0,\theta-\frac{v}{r_p}\right) - l\delta.$$

We suppose that $\theta r_p > 10$, since otherwise (7) would be trivially true. This means that $\lfloor \theta r_p \rfloor^2 > 2\theta^2 r_p^2/3$. If $\theta r_p < l$, this means that

$$\begin{aligned}\sum_{v=0}^{l-1}\max\left(0,\theta-\frac{v}{r_p}\right) &= r_p^{-1}\sum_{v=0}^{\lfloor\theta r_p\rfloor}(\theta r_p - v)\\ &\geq \frac{1}{2}r_p^{-1}\lfloor\theta r_p\rfloor^2\\ &\geq \frac{1}{3}\theta^2 r_p,\end{aligned}$$

while

$$\sum_{v=0}^{l-1}\max\left(0,\theta-\frac{v}{r_p}\right) = \sum_{v=0}^{l-1}\left(\theta-\frac{v}{r_p}\right) = l\theta - \frac{l(l-1)}{2r_p} \geq \frac{1}{2}l\theta,$$

if $\theta r_p \geq l$. Thus

$$index\ F \geq \min(\frac{1}{2}l\theta, \frac{1}{3}\theta^2 r_p) - l\delta,$$

and so either $\theta < 2(\Phi + \delta)$, and we are done, or

$$\frac{1}{3}\theta^2 r_p \le l(\Phi + \delta) \le (r_p + 1)(\Phi + \delta).$$

Since $r_p + 1 < \frac{4}{3} r_p$,

$$\theta \le 2(\Phi + \delta)^{\frac{1}{2}} \le 2(\Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}}),$$

and we are done. □

Now we use what is known from Theorem 4.2 and Theorem 4.4 for the main theorem in this section.

Theorem 4.5. *Let m be a positive integer, and suppose that*

$$0 < \delta < \frac{1}{m 2^m (N+1)^2}. \tag{9}$$

Let $r_1, \dots, r_m$ be positive integers such that

$$r_m > 10\delta^{-1}, \frac{r_{j-1}}{r_j} > \delta^{-1}, \text{ for } j = 2, \dots, m. \tag{10}$$

Let $q_1, \dots, q_m$ be positive integers such that

$$\log q_1 > 2\delta^{-1} m(2m+1), \tag{11}$$

$$r_j \log q_j \ge r_1 \log q_1, \text{ for } j = 2, \dots, m, \tag{12}$$

$$\log q_1 > 3\delta^{-1} N(N+1). \tag{13}$$

Then

$$\Theta(q_1^{\delta r_1}; \overline{q}_m; \overline{r}_m) < 10^m \delta^{\left(\frac{1}{2}\right)^m}. \tag{14}$$

Proof. We will use induction on m. If $m = 1$, then by Theorem 4.2, (13) and (9)

$$\begin{aligned} \Theta_1(q_1^{\delta r_1}; \overline{q}_1; \overline{r}_1) &\le \frac{3N(N+1)}{\log q_1} + \frac{N \log q_1^{\delta r_1}}{r_1 \log q_1} = \frac{3N(N+1)}{\log q_1} + N\delta \\ &\le (N+1)\delta < (\delta/2)^{\frac{1}{2}} < 10\delta^{\frac{1}{2}}, \end{aligned} \tag{15}$$

which is the desired inequality.

Now suppose $p > 1$ is an integer and that the theorem holds for $m = p - 1$. We will now look at the case $m = p$ and we will need to show that

$$\Theta_p(q_1^{\delta r_1}; \overline{q}_p; \overline{r}_p) \le 2 \max(\Phi + \Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}}) < 10^p \delta^{\left(\frac{1}{2}\right)^p},$$

where Φ is as in Theorem 4.4. To do this we will look at $\Theta_1(M; q_p; l r_p)$ and $\Theta_{p-1}(M; \overline{q}_{p-1}; \overline{l r}_{p-1})$ separately. We start by estimating M.

$$M = (r_1+1)^{2pl} 2^{2r_1 pl} l!^2 B^{2l} \le \left\{(r_1+1)^{2p} 2^{2r_1 p} l^2 q_1^{2\delta r_1}\right\}^l.$$

Since $l \le r_p + 1 < r_1 + 1 \le 2^{r_1}$, it follows that

$$M < \left\{2^{r_1(4p+2)} q_1^{2\delta r_1}\right\}^l < \left\{e^{r_1(4p+2)} q_1^{2\delta r_1}\right\}^l.$$

By (14) we get that $4p + 2 < \delta p^{-1} \log q_1$, so

$$M < q_1^{\delta_1 l r_1},$$

where

$$\delta_1 = 2\delta(1+p^{-1}).$$

Thus

$$\Theta_1(M; q_p; lr_p) < \Theta_1(q_1^{\delta_1 lr_1}; q_p; lr_p) \tag{16}$$

and

$$\Theta_{p-1}(M; \overline{q}_{p-1}; \overline{lr}_{p-1}) < \Theta_{p-1}(q_1^{\delta_1 lr_1}; \overline{q}_{p-1}; \overline{lr}_{p-1}). \tag{17}$$

By (9) we get that

$$\delta_1 < \frac{1+p^{-1}}{p2^{p-1}(N+1)^2} < \frac{1}{(p-1)2^{p-1}(N+1)^2}. \tag{18}$$

By Theorem 4.2,

$$\begin{aligned}\Theta_1(q_1^{\delta_1 lr_1}; q_p; lr_p) &< \frac{3N(N+1)}{\log q_1} + \frac{N\log q_1^{\delta_1 lr_1}}{r_1 \log q_1} < \delta + \frac{N\delta_1 lr_1 \log q_1}{lr_p \log q_p} \\ &< \delta + N\delta_1 < (N+1)\delta_1 < \delta_1^{\frac{1}{2}},\end{aligned} \tag{19}$$

where we used (13) and the fact that $(N+1)\delta_1 < \delta_1^{\frac{1}{2}}$.

We now go on and estimate the right-hand side of (17). (10) − (13) remain true if we replace δ by δ_1. By the result obtained at (18) we see that (9) is also true. By the induction hypothesis

$$\Theta_{p-1}(q_1^{\delta_1 lr_1}; \overline{q}_{p-1}; \overline{lr}_{p-1}) < 10^{p-1}\delta_1^{\left(\frac{1}{2}\right)^{p-1}}. \tag{20}$$

Since $\delta_1 < 4\delta$, we can combine the results obtained at (19) and (20),

$$\Phi < 2\delta^{\frac{1}{2}} + 2\left(10^{p-1}\delta^{\left(\frac{1}{2}\right)^{p-1}}\right) < 3\left(10^{p-1}\delta^{\left(\frac{1}{2}\right)^{p-1}}\right).$$

Finally, (7) gives

$$\begin{aligned}\Theta_p(q_1^{\delta_1 r_1}; \overline{q}_p; \overline{r}_p) &< 2\left\{3\left(10^{p-1}\delta^{\left(\frac{1}{2}\right)^{p-1}}\right) + 3^{\frac{1}{2}}\left(10^{\frac{1}{2}(p-1)}\delta^{\left(\frac{1}{2}\right)^{p}}\right) + \delta^{\frac{1}{2}}\right\} \\ &< 2\left(\frac{3}{10} + \frac{3^{\frac{1}{2}}}{10^{\frac{3}{2}}} + \frac{1}{10^2}\right)10^p\delta^{\left(\frac{1}{2}\right)^{p}} < 10^p\delta^{\left(\frac{1}{2}\right)^{p}}.\end{aligned}$$

□

5. A combinatorical lemma

Definition 5.1. Let $r_1, \dots, r_m$ be positive integers, and $\lambda > 0$. Let $(j_1, \dots, j_m)$ be a list of integers such that

$$0 \le j_1 \le r_1, \dots, 0 \le j_m \le r_m \tag{21}$$

and

$$\frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \le \frac{1}{2}(m - \lambda). \tag{22}$$

If A is the set of all such lists of integers, then $A_m(\lambda)$ is defined as

$$A_m(\lambda) = |A|.$$

Theorem 5.1. *If $r_1, \dots, r_m$ and λ are as in the above definition, then*

$$A_m(\lambda) \le 2\sqrt{m}\lambda^{-1}(r_1 + 1) \cdots (r_m + 1). \tag{23}$$

Proof. We will prove this by induction on m. The theorem clearly holds for the case $m = 1$, since for $\lambda \le 1$

$$2\sqrt{m}\lambda^{-1}(r_1 + 1) \ge 2(r_1 + 1),$$

and $A_m(\lambda) = 0$ for $\lambda > 1$.

Now suppose $m > 1$. We can clearly assume that $\lambda > 2\sqrt{m}$. The aim is to first find a recurrence relation between A_m and A_{m-1}.

Fix j_m to any integer satisfying (21). If we look for the number of lists $(j_1, \dots, j_{m-1})$, together with j_m, satisfying (21) and (22), then if we iterate over all the j_m, we get a sum equal to $A_m(\lambda)$. Thus, by letting

$$m - \lambda - \frac{2j_m}{r_m} = (m - 1) - \lambda',$$

where $\lambda' = \lambda'(j_m)$, we get that

$$A_m(\lambda) = \sum_{j_m=0}^{r_m} A_{m-1}(\lambda'(j_m)).$$

Now we take the induction step. Assume (23) is true for $m = p - 1$. Then, since $r_j + 1 \ge 1$ for all $j = 1, \dots, p - 1$,

$$A_m(\lambda) \le 2\sqrt{m-1}(r_1 + 1) \cdots (r_{p-1} + 1) \sum_{j_m=0}^{r_m} \left(\lambda - 1 + \frac{2j_m}{r_m}\right)^{-1},$$

and so we have reduced the problem to showing that

$$\sum_{j_m=0}^{r_m} \left(\lambda - 1 + \frac{2j_m}{r_m}\right)^{-1} \le \lambda^{-1}(m-1)^{-\frac{1}{2}}\sqrt{m}(r + 1),$$

for all positive integers r and m.

We now consider the two cases; r is even and r is odd. If r is even, we put $j = \frac{1}{2}r + k$ and obtain the sum

$$\begin{aligned}
\sum_{k=-\frac{1}{2}r}^{\frac{1}{2}r} \left(\lambda + \frac{2k}{r}\right)^{-1} &= \lambda^{-1} + \sum_{k=1}^{\frac{1}{2}r} \left\{ \left(\lambda + \frac{2k}{r}\right)^{-1} + \left(\lambda - \frac{2k}{r}\right)^{-1} \right\} \\
&= \lambda^{-1} + \sum_{k=1}^{\frac{1}{2}r} 2\lambda \left(\lambda^2 - \frac{4k^2}{r^2}\right)^{-1} \\
&\leq \lambda^{-1} + 2\lambda \sum_{k=1}^{\frac{1}{2}r} (\lambda^2 - 1)^{-1} \\
&= \lambda^{-1} + 2\lambda^{-1} \sum_{k=1}^{\frac{1}{2}r} (1 - \lambda^{-2})^{-1} \\
&= \lambda^{-1}(r+1)(1-\lambda^{-2})^{-1} \\
&< \lambda^{-1}(r+1)(1 - m^{-1}/4)^{-1} \\
&< \lambda^{-1}(r+1)(1-m^{-1})^{-\frac{1}{2}} = \lambda^{-1}(r+1)m^{\frac{1}{2}}(m-1)^{-\frac{1}{2}},
\end{aligned}$$

which is the desired inequality.

Now suppose r is odd. If we put $j = (r-1)/2 + k$ we obtain

$$\begin{aligned}
&\sum_{k=-\frac{1}{2}(r-1)}^{\frac{1}{2}(r+1)} \left(\lambda + \frac{2k-1}{r}\right)^{-1} \\
&\qquad = \sum_{k=1}^{\frac{1}{2}(r+1)} \left\{ \left(\lambda + \frac{2k-1}{r}\right)^{-1} + \left(\lambda - \frac{2k-1}{r}\right)^{-1} \right\} \\
&\qquad = 2\lambda \sum_{k=1}^{\frac{1}{2}(r+1)} \left(\lambda^2 - \frac{(2k-1)^2}{r^2}\right)^{-1} \\
&\qquad \leq \lambda(\lambda^2 - 1)^{-1}(r+1).
\end{aligned}$$

□

6. The approximation plynomial

In this section we will prove the theorem which is the only one referenced to in the proof of the Thue-Siegel-Roth theorem.

Let α be an algebraic integer of degree $n \geq 2$ over K. Let $\omega_1, \ldots, \omega_N$ be an integral basis for K, and put

$$|\overline{\alpha}| = b_1 \geq 1, \ \max(|\overline{\omega_1}|, \ldots, |\overline{\omega_N}|) = b_2.$$

That $\omega_1, \ldots, \omega_N$ is an integral basis for K mean that every element $\mathcal{O}_K$ can be written uniquely as a linear combination of $\{\omega_1, \ldots, \omega_N\}$ with coefficients in $\mathbb{Z}$.

We will later choose the variables m, δ, $q_1, \zeta_1, \ldots, q_m, \zeta_m$, $r_1, \ldots, r_m$, in the given order just specified, such that they satisfy the following conditions:

$$0 < \delta < \frac{1}{m2^m(N+1)^2}, \tag{24}$$

$$10^m \delta^{(\frac{1}{2})^m} + 2(1+3\delta)n\sqrt{m} < \frac{m}{2}, \tag{25}$$

$$r_m > 10\delta^{-1}, \frac{r_{j-1}}{r_j} > \delta^{-1}, \text{ for } j = 2, \ldots, m, \tag{26}$$

$$\delta^2 \log q_1 > 2m + 1 + m\log(b_1 + 1) + 4b_2 N, \tag{27}$$

$$r_j \log q_j \geq r_1 \log q_1, \text{ for } j = 2, \ldots, m, \tag{28}$$

$$\log q_1 > 3\delta^{-1} N(N+1). \tag{29}$$

We also define some new variables which will make the calculations less messy.

$$\lambda = 4(1+3\delta)n\sqrt{m} \tag{30}$$

$$\mu = \frac{1}{2}(m - \lambda), \tag{31}$$

$$\eta = 10^m \delta^{(\frac{1}{2})^m}, \tag{32}$$

$$B_1 = \lfloor q_1^{\delta r_1} \rfloor. \tag{33}$$

Theorem 6.1. *Suppose that the condtions* $(24) - (29)$ *are satisfied, and suppose that* $\zeta_1, \ldots, \zeta_m$ *are algebraic numbers of heights* $q_1, \ldots, q_m$*, respectively. Then there exists a polynomial* $Q \in K[z_1, \ldots, z_m]$ *with integral coefficients in* K *and of degree at most* r_j *in* z_j*, for* $j = 1, \ldots, m$*, such that*

(i) *the index of* Q *at the point* $(\alpha, \ldots, \alpha)$ *relative to* $r_1, \ldots, r_m$ *is at least* $\mu - \eta$*;*
(ii) $Q(\zeta_1, \ldots, \zeta_m) \neq 0$*;*
(iii) *for all derivatives*

$$Q_{i_1 \cdots i_m}(z_1, \ldots, z_m) = \frac{1}{i_1! \cdots i_m!} \left(\frac{\partial}{\partial z_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{i_m} Q,$$

where $i_1, \ldots, i_m$ *are non-negative integers, the inequality*

$$|Q_{i_1 \cdots i_m}(z_1, \cdots, z_m)| < B_1^{1+3\delta}(1 + |z_1|)^{r_1} \cdots (1 + |z_m|)^{r_m}$$

holds, and the corresponding inequality also holds if the coefficients in Q are replaced by their respective field conjugates.

Proof. Let C be the set of integers of K of the form

$$c_1\omega_1 + \cdots + c_N\omega_N,$$

where $c_1, \ldots, c_N$ range over all non-negative integers not exceeding B_1. If we put

$$(1 + r_1) \cdots (1 + r_m) = r,$$

then there are $|C|^r = (1 + B_1)^{Nr}$ polynomials

$$P(z_1, \ldots, z_m) = \sum_{s_1=0}^{r_1} \cdots \sum_{s_m=0}^{r_m} \gamma(s_1, \ldots, s_m) z_1^{s_1} \cdots z_m^{s_m}$$

whose coefficients $\gamma(s_1, \ldots, s_m) \in C$. If we put

$$\begin{aligned} &P_{i_1 \cdots i_m}(z_1, \ldots, z_m) \\ &\qquad = \frac{1}{i_1! \cdots i_m!} \left(\frac{\partial}{\partial z_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{i_m} P(z_1, \ldots, z_m) \\ &\qquad = \sum_{s_1=0}^{r_1} \cdots \sum_{s_m=0}^{r_m} \gamma(s_1, \ldots, s_m) \binom{r_1}{j_1} \cdots \binom{r_m}{j_m} z_1^{s_1 - j_1} \cdots z_m^{s_m - j_m}, \end{aligned}$$

then

$$|\overline{P_{j_1 \cdots j_m}}| \leq 2^{r_1 + \cdots + r_m} b_2 B_1 N \leq b_2 N 2^{m r_1} B_1 < b_2 N B_1^{1+\delta},$$

since

$$|\overline{\gamma(s_1, \ldots, s_m)}| \leq b_2 B_1 N, \tag{34}$$

$$\begin{aligned} m r_1 \log 2 < r_1[m(1 + \log \sqrt{b_1 + 1})] &< \frac{1}{2} r_1 [2m + 1 + m \log(b_1 + 1) + 4 b_2 N] \\ &< \frac{1}{2} \delta^2 r_1 \log q_1, \end{aligned}$$

$$\binom{r_k}{j_k} < 2^{r_k}$$

and $q_1^{\frac{1}{2} r_1} < B_1$. Also, since, by (27),

$$r \leq 2^{r_1 + \cdots + r_m} \leq 2^{m r_1} \leq (b_1 + 1)^{m r_1} < B_1^{\delta},$$

we obtain the bound

$$\begin{aligned} |\overline{P_{j_1 \cdots j_m}(\alpha, \ldots, \alpha)}| &\leq b_2 N B_1^{1+\delta} r b_1^{r_1 + \cdots + r_m} \\ &\leq b_2 N B_1^{1+3\delta}. \end{aligned} \tag{35}$$

Let ϑ be a primitive element of L, so that $L = \mathbb{Q}(\vartheta)$. Order the conjugates of ϑ so that $\vartheta_1, \ldots, \vartheta_{\rho_1}$ are real and $\overline{\vartheta_{\rho_1 + v}} = \vartheta_{\rho_1 + \rho_2 + v}$ are complex conjugates for $v = 1, \ldots, \rho_2$, so that $\rho_1 + 2\rho_2 = [\mathbb{Q}(\vartheta) : \mathbb{Q}] = nN$. Let ξ be fixed equal to one of the numbers $P_{j_1 \cdots j_m}(\alpha, \ldots, \alpha)$, where $j_1, \ldots, j_m$ satisfy the inequalities

$$0 \leq j_1 \leq r_1, \ \ldots, 0 \leq j_m \leq r_m, \qquad \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \leq \mu. \tag{36}$$

Then ξ can be written as a polynomial in ϑ, with rational coefficients, with the field conjugates $\xi^{(v)}$, for $v = 1, \dots, nN$, i.e., if

$$\xi = a_0 + a_1\vartheta^1 + \cdots + a_{n-1}\vartheta^{nN-1},$$

where $a_i \in \mathbb{Q}$, for $i = 1, \dots, nN$, then

$$\xi^{(v)} = a_0 + a_1\vartheta_v^1 + \cdots + a_{n-1}\vartheta_v^{nN-1},$$

is a field conjugate, where ϑ_v is one of the conjugates of ϑ. We can define nN real numbers $\xi_1, \dots, \xi_{nN}$ by the equations

$$\begin{aligned} \xi_v &= \xi^{(v)}, \qquad \text{for } v = 1, \dots, \rho_1, \\ \xi_v + i\xi_{v+\rho_2} &= \xi^{(v)}, \qquad \text{for } v = \rho_1 + 1, \dots, \rho_1 + \rho_2. \end{aligned}$$

If we fix the coefficients $\gamma(s_1, \dots, s_m)$, then we can arrange the ξ_v's in a fixed order and each of these numbers can be viewed as coordinates of a point. Doing this for all $j_1, \dots, j_m$ satisfying (36), we get, by Theorem 5.1,

$$M \le 2nN\sqrt{m}\lambda^{-1}r$$

coordinates. Furthermore, from (35) we see that each of the coordinates have absolute values smaller than $\lfloor b_2 N B_1^{1+3\delta} \rfloor + 1 = t$. Thus all points for various $\gamma(s_1, \dots, s_m) \in C$ lie in a cube of edge $2t$ in M-dimensional space. We can divide the cube into $(3t)^M$ subcubes of edge $\frac{2}{3}$. If

$$|C|^r = (1 + B_1)^{Nr} > (3t)^M, \tag{37}$$

then there exists more points than subcubes and thus the points corresponding to two different polynomials $P^*(z_1, \dots, z_m)$ and $P^{**}(z_1, \dots, z_m)$ lie in the same subcube. If we put

$$\overline{P} = P^* - P^{**},$$

then the point $\overline{P}_{j_1 \cdots j_m}(\alpha, \dots, \alpha)$ is in one of the 2^M subcubes closest to the origin, thus

$$|\overline{\overline{P}_{j_1 \cdots j_m}(\alpha, \dots, \alpha)}| \le \sqrt{2} \times \frac{2}{3} < 1,$$

for all $j_1, \dots, j_m$ satisfying (36). But, since $\overline{P}_{j_1 \cdots j_m}(\alpha, \dots, \alpha)$ is an algebraic integer, this can only be true if it equals to zero. Since this is true for all $j_1, \dots, j_m$ satisfying (36), the index of $\overline{P}$ at $(\alpha, \dots, \alpha)$ relative to $r_1, \dots, r_m$ must be greater than μ. The coefficients of $\overline{P}$ are the differences of two elements of C, and thus it is not hard to see that the inequality (34) holds for them.

We now verify that (37) indeed holds, so that the above conclusions are valid. Notice that by (27)

$$q_1^{\delta r_1} > 4b_2 N,$$

so

$$\begin{aligned} B_1 &> 4b_2N, \\ B_1^{Nr} &> (4b_2NB_1)^{\frac{1}{2}Nr}, \\ B_1^{Nr} &> (3b_2NB_1^{1+3\delta}+3)^{\frac{1}{2}Nr(1+3\delta)^{-1}} \\ (1+B_1)^{Nr} &> (3t)^M, \end{aligned}$$

where the third inequality follows from the fact that $B_1 > e^{15} > 3$.

Now, $\overline{P} \in \mathcal{R}_m(q_1^{\delta r_1}; \overline{r}_m)$[¶], its index at $(\zeta_1, \ldots, \zeta_m)$ relative to $r_1, \ldots, r_m$ must be less than η, by Theorem 4.5. Hence there exists a differential operator

$$\Delta_k = \frac{1}{k_1! \ldots k_m!} \left(\frac{\partial}{\partial z_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{k_m}$$

with

$$Q(z_1, \ldots, z_m) = \Delta_k \overline{P},$$

so that if

$$\frac{k_1}{r_1} + \cdots \frac{k_m}{r_m} < \eta,$$

then

$$Q(\zeta_1, \ldots, \zeta_m) \neq 0.$$

The index of Q at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least $\mu - \eta$. Notice that by (25) $\mu > \eta$. Thus (i) and (ii) are satisfied.

From (34) and the inequality $r < B_1^{\delta}$,

$$|\overline{Q}| \leq 2^{r_1+\cdots+r_m} b_2 N B_1 < 2^{mr_1} b_2 N B_1 < b_2 N B_1^{1+\delta},$$

and hence

$$|\overline{Q_{i_1 \cdots i_m}}| < 2^{r_1+\cdots+r_m} b_2 N B_1^{1+\delta} < b_2 N B_1^{1+2\delta}.$$

Finally,

$$\begin{aligned} |Q_{i_1 \cdots i_m}(z_1, \ldots, z_m)| &< b_2 N B_1^{1+2\delta} \prod_{v=1}^{m} (1 + |z_v| + \cdots + |z_v|^{r_v}) \\ &< b_2 N B_1^{1+2\delta} \prod_{v=1}^{m} (1 + |z_v|)^{r_v} \\ &< B_1^{1+3\delta} \prod_{v=1}^{m} (1 + |z_v|)^{r_v}, \end{aligned}$$

where the last inequality follows since $b_2 N < B_1^{\delta}$ by (27). The same inequality holds if the coefficients are replaced by their respective field conjugates, and thus we are done. □

[¶]This is the second error made by LeVeque. However, this holds when $K = \mathbb{Q}$ since then $|\overline{(\overline{P})}| = ||(\overline{P})|| \leq ||P|| < B_1 < q_1^{\delta r_1}$.

7. The Thue-Siegel-Roth theorem

Theorem 7.1 (Generalization of the Thue-Siegel-Roth theorem)**.** *Let K be an algebraic number field of degree N, i.e., $[K:\mathbb{Q}]=N$, and let α be algebraic of degree $n\geq 2$ over K. Then for each $\kappa>2$, the inequality*

$$|\alpha-\zeta|<\frac{1}{[H(\zeta)]^{\kappa}} \tag{38}$$

has only finitely many solutions for ζ in K.

Proof. This will be proved by assuming the theorem is false and finally arrive at a contradiction.

Let the (monic) defining polynomial of α be $p(z)=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0$. Further, let a equal the least common multiple of the denominators of the rational coefficients of $p(z)$. Then $q(z)=a^np\left(\frac{z}{a}\right)$ has integral coefficients, is monic and irreducible. $a\alpha$ is a root of the equation and thus $a\alpha$ is an algebraic integer. Suppose ζ is a solution of (38). Then

$$|a\alpha-a\zeta|<\frac{\alpha}{[H(\zeta)]^{\kappa}}\leq\frac{a^{\kappa N+1}}{[H(a\zeta)]^{\kappa}},$$

where the last inequality follows from the fact that ζ, and viz. $a\zeta$, can be at most of degree N, thus $H(a\zeta)\leq a^NH(\zeta)$. Hence, for arbitrary $\epsilon>0$, and for all solutions ζ with $H(\zeta)$ sufficiently large,

$$|a\alpha-a\zeta|<\frac{1}{[H(a\zeta)]^{\kappa-\epsilon}},$$

and ϵ can be chosen so small that $\kappa-\epsilon>2$. Thus we can assume that α is an algebraic integer. Note that we can ignore the $\zeta's$ for which $H(\zeta)$ is not sufficiently large since K is an algebraic number field, and thus a finite extension, which means that these cases are not infinitely many. Note that we only need to prove the theorem for primitive elements ζ in K. This is true because the number of subfields of an algebraic number field is finite and each element of K is a primitive element in such a subfield, thus the proof can be repeated.

Choose m to be a rational integer so that $m>4n\sqrt{m}$ and

$$\frac{2m}{m-4n\sqrt{m}}<\kappa. \tag{39}$$

Note that the by first inequality we get that the right-hand side of (39) is positive and $\frac{2m}{m-4n\sqrt{m}}\longrightarrow 2$ as $m\longrightarrow\infty$. So such an m exists since κ is strictly greater than 2. Further, for sufficiently small δ we have

$$m-4(1+3\delta)n\sqrt{m}-2\eta>0,$$

where η was defined in (32). This inequality is the same as the one in (25).We choose δ to satisfy this, (24) and the inequality

$$\frac{2m(1+\delta)+2\delta N(2+5\delta)}{m-4(1+3\delta)n\sqrt{m}-2\eta}<\kappa \tag{40}$$

which is possible because of (39). By using (30) and (31) we can write this inequality as

$$\frac{m(1+\delta)+\delta N(2+5\delta)}{\mu-\eta} < \kappa. \tag{41}$$

We now choose a primitive solution ζ_1 of (38) such that $q_1 = H(\zeta_1)$ satisfies (27) and (29). We then choose further primitive solutions $\zeta_2, \ldots, \zeta_m$ of heights $q_2, \ldots, q_m$, respectively, such that for $j = 2, \ldots, m$,

$$\frac{\log q_j}{\log q_{j-1}} > \frac{2}{\delta}. \tag{42}$$

We now let r_1 te be any rational integer such that

$$r_1 > \frac{10 \log q_m}{\delta \log q_1}, \tag{43}$$

and define r_j, for $j = 2, \ldots, m$, by

$$\frac{r_1 \log q_1}{\log q_j} \leq r_j < \frac{r_1 \log q_1}{\log q_j} + 1. \tag{44}$$

This satisfies (28). Notice that this gives us

$$\frac{r_j \log q_j}{r_1 \log q_1} < 1 + \frac{\log q_j}{r_1 \log q_1} < 1 + \frac{\log q_m}{r_1 \log q_1} < 1 + \frac{\delta}{10}, \tag{45}$$

where (42) is used for the second inequality and (43) for the third. The conditions (26) are satisfied since

$$r_m \geq \frac{r_1 \log q_1}{\log q_m} > 10\delta^{-1},$$

by (44) and (43), and

$$\begin{aligned}
\frac{r_{j-1}}{r_j} &> \left(\frac{\log q_j}{r_1 \log q_1 + \log q_j}\right)\left(\frac{r_1 \log q_1}{\log q_{j-1}}\right) = \frac{\log q_j}{\log q_{j-1}}\left(\frac{r_1 \log q_1}{r_1 \log q_1 + \log q_j}\right) \\
&= \frac{\log q_j}{\log q_{j-1}}\left(1 + \frac{\log q_j}{r_1 \log q_1}\right)^{-1} > \frac{\log q_j}{\log q_{j-1}}\left(1 + \frac{\delta^{-1}}{10}\right)^{-1} \\
&> \frac{2}{\delta}\left(1 + \frac{\delta^{-1}}{10}\right)^{-1} > \frac{2}{\delta} > \delta^{-1}.
\end{aligned}$$

Let $Q(z_1, \ldots, z_m)$ be the polynomial as in Theorem 6.1. Let $\zeta_1, \ldots, \zeta_m \in K$ be zeros of irreducible polynomials of degree N with relatively prime coefficients in $\mathbb{Z}$, and the coefficients of z^N being $k_1, \ldots, k_m$, respectively. Then the number

$$\phi = Q(\zeta_1, \ldots, \zeta_m)$$

is in K. If the field conjugates of ζ_i are $\zeta_i^{'}, \zeta_i^{''}, \ldots$, for $i = 1, \ldots, m$, then $\mathrm{No}_{K/\mathbb{Q}}(\phi)$ is a sum of products of powers of the $\zeta_i^{(j)}$ with integral coefficients from K. In each such product, the factor $\zeta_i^{(j)}$ occurs to the power r_i at most. It can be shown that the product of k_i and any set of distinct conjugates of $\zeta_i^{(j)}$ is an algebraic integer. For each i, the field conjugates of ζ_i are distinct, because ζ_i is a primitive element of K. It follows that

$k_1^{r_1}\cdots k_m^{r_m}\mathrm{No}_{K/\mathbb{Q}}(\phi)$ is an algebraic integer, and since it is also rational it is a rational integer, hence

$$|k_1^{r_1}\cdots k_m^{r_m}\mathrm{No}_{K/\mathbb{Q}}(\phi)|\geq 1. \tag{46}$$

From (i) in Theorem 6.1 we have that the terms in

$$Q(\zeta_1,\dots,\zeta_m)=\sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}Q_{i_1\cdots i_m}(\alpha,\dots,\alpha)(\zeta_1-\alpha)^{i_1}\cdots(\zeta_m-\alpha)^{i_m}$$

are equal to zero whenever

$$\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}<\mu-\eta.$$

For the non-zero terms we have

$$\begin{aligned}|(\zeta_1-\alpha)^{i_1}\cdots(\zeta_m-\alpha)^{i_m}|&<(q_1^{i_1}\cdots q_m^{i_m})^{-\kappa}\\&=\left[q_1^{i_1/r_1}(q_2^{r_2/r_1})^{i_2/r_2}\cdots(q_m^{r_m/r_1})^{i_m/r_m}\right]^{-r_1\kappa}\\&\leq(q_1^{i_1/r_1}\cdots q_m^{i_m/r_m})^{-r_1\kappa}\\&<q_1^{-r_1\kappa(\mu-\eta)},\end{aligned}$$

where the first inequality follows from our assumption of $\zeta_1,\dots,\zeta_m$ as solutions to (38), and the third inequality follows from (28). By (iii) in Theorem 6.1

$$\begin{aligned}|\phi|&<(r_1+1)\cdots(r_m+1)B_1^{1+3\delta}(1+b_1)^{mr_1}q_1^{-r_1\kappa(\mu-\eta)}\\&<B_1^{1+5\delta}q_1^{-r_1\kappa(\mu-\eta)},\end{aligned}$$

and by using it once again together with Theorem 2.1 we get

$$\begin{aligned}|k_1^{r_1}\cdots k_m^{r_m}\mathrm{No}_{K/\mathbb{Q}}(\phi)|<k_1^{r_1}\cdots k_m^{r_m}|\phi||\phi'|\cdots|\phi^{(N)}|&<B_1^{1+5\delta}q_1^{-r_1\kappa(\mu-\eta)}B_1^{(N-1)(1+5\delta)}\\&\times\prod_{i=1}^{m}\left\{k_i\prod_{j=1}^{N}(1+|\zeta_i^{(j)}|)\right\}^{r_i}\\&<B_1^{N(1+5\delta)}q_1^{-r_1\kappa(\mu-\eta)}\prod_{i=1}^{m}(6^Nq_i)^{r_i}.\end{aligned}$$

In the proof of Theorem 6.1 it was shown that

$$2^{r_1+\cdots+r_m}<B_1^{\delta},$$

so $6^{N(r_1+\cdots+r_m)}<q_1^{\delta Nr_1}$ and by combining all terms we get

$$\begin{aligned}|k_1^{r_1}\cdots k_m^{r_m}\mathrm{No}_{K/\mathbb{Q}}(\phi)|&<q_1^{\delta Nr_1(1+5\delta)+\delta Nr_1+mr_1-r_1\kappa(\mu-\eta)}\\&<q_1^{\delta Nr_1(2+5\delta)+mr_1(1+\delta)-r_1\kappa(\mu-\eta)}.\end{aligned}$$

This together with (46) implies that

$$\delta N(2+5\delta)+m(1+\delta)>\kappa(\mu-\eta),$$

or

$$\kappa<\frac{\delta N(2+5\delta)+m(1+\delta)}{\mu-\eta},$$

which contradicts (41). The proof is thus completed. □

Acknowledgements

My deepest gratitude to my supervisor Andreas Strömbergsson for his help and comments during the process of this undergraduate thesis.

References

[1] E. Landau, "Vorlesungen über Zahlentheorie Aus der elementaren Zahlentheorie", Chelsea Publishing Company, 1946.
[2] G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers", Oxford University Press, 1979.
[3] K. Ireland and M. Rosen, "A Classical Introduction to Modern Number Theory", Springer, 1998.
[4] A. Thue, "J. Reine Angew. Math.", 135 (1909), 284-305.
[5] C. L. Siegel, Approximation algebraiseher Zahlen, Mathematische Zeitschrift, **10** (1921), 173-213.
[6] F. J. Dyson, The approximation to algebraic numbers by rationals, Acta Mathematica, **79** (1947), 225-240.
[7] K. F. Roth, Rational approximations to algebraic numbers, Mathematika, **2** (1955), 1-20.
[8] W. J. LeVeque, "Topics In Number Theory, Vol. I & II", Addison-Wesley Publishing Company, Inc, 1956.