Radboud University

Faculty of Science

Institute for Mathematics, Astrophysics and Particle Physics

Zero estimates for polynomials inspired by Hilbert's seventh problem

Author:
Janet Flikkema
s4457242

Supervisor:
Prof. Dr. Wadim Zudilin

Second corrector:
Dr. Wieb Bosma

Bachelor Thesis Mathematics

December 20, 2018

Radboud University

Abstract

This thesis sketches the proof of the Gelfond-Schneider theorem, which is the solution to Hilbert's seventh problem. An important part of the proof is a so-called zero estimate for a two-variable complex polynomial. We will zoom in on this part and also derive zero estimates in different settings. In particular, we find that nonzero complex polynomials in ℓ variables cannot vanish on finite subsets of certain subgroups of $\mathbb{C}_+ \times \mathbb{C}_\times^{\ell-1}$. The work in this thesis is inspired by material in the course notes of Analytic Number Theory, taught to master's students on Radboud University.

Content

1 Introduction

In 1900, David Hilbert published a list of twenty-three open mathematical problems of which the seventh concerned the transcendence of particular numbers. The problem was a generalization of a conjecture already formulated by Leonhard Euler in his *Introductio in analysin infinitorum* in 1728. The latter had not yet been proven when Hilbert published his problems and is as follows:

If α is a positive rational number unequal to 1 and β is an irrational real algebraic number, then α^β is irrational.

Recall that an algebraic number is the root of a nonzero rational polynomial and a number that is not algebraic is called transcendental. Euler's conjecture implied in particular that $2^{\sqrt{2}}$ is irrational. In 1873, the transcendence of e was proven by Hermite, followed by the transcendence of π which Lindemann established nine years later by generalizing Hermite's methods. Hilbert wondered whether $2^{\sqrt{2}}$ and $e^\pi = (-1)^{-i}$ were also transcendental; these numbers served as an inspiration for his seventh problem:

Is α^β always transcendental, or at least irrational, for algebraic α, β, where $\alpha \neq 0, 1$ and β is irrational?

Hilbert believed that the Riemann hypothesis (his eighth problem, unsolved) and Fermat's last theorem (solved in 1994) would be solved long before his seventh problem. However, the solution to the seventh problem was found simultaneously and independently by Aleksandr Gelfond and Theodor Schneider in 1934. They proved that if $\alpha_1, \alpha_2, \beta_1, \beta_2$ are nonzero algebraic numbers such that $\ln \alpha_1$ and $\ln \alpha_2$ are linearly independent over the rationals, then

$$\beta_1 \ln \alpha_1 + \beta_2 \ln \alpha_2 \neq 0,$$

meaning that $\ln \alpha_1$ and $\ln \alpha_2$ are linearly independent over the field of algebraic numbers. This result forms the solution to Hilbert's seventh problem, as we will see in Chapter 2.

Naturally, the question arose if the above result also holds in the more general case, where we have a linear combination of an arbitrary number of logarithms of algebraic numbers. This was solved in 1966 by Alan Baker, who proved that if $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ are nonzero algebraic numbers such that $\ln \alpha_1, \ldots, \ln \alpha_n$ are linearly independent over the rationals, then

$$\beta_1 \ln \alpha_1 + \ldots + \beta_n \ln \alpha_n \neq 0,$$

meaning that $\ln \alpha_1, \ldots, \ln \alpha_n$ are linearly independent over the field of algebraic numbers. This has led to the theory of *Linear Forms in Logarithms*, which has many applications in number theory, for example solving Diophantine equations [1, 2].

This thesis elaborates on an important part of the proof of the Gelfond-Schneider theorem, which is given in Chapter 2. In the specific part of the proof, it is shown that a nonzero complex polynomial P in two variables cannot vanish on (i.e. cannot be zero in all points of) a particular set. As an illustration, this is comparable to the known property that a nonzero polynomial in one variable cannot have more distinct zeros than its degree. Generalizing the result used in the Gelfond-Schneider proof, we find that P cannot vanish on a finite subset of a group $\mathcal{A} \subset \mathbb{C}^2$ that satisfies certain conditions. Results of this type are called "zero estimates". In Chapter 3, we give zero estimates in different settings and describe the structure of the corresponding groups $\mathcal{A}$.

In the two variable case, where P is a nonzero two-variable complex polynomial, we find the following results. First of all, if $\alpha_1, ..., \alpha_m \in \mathbb{C} \setminus \{0\}$ are multiplicatively independent (this means that no product of integral powers of these numbers is equal to 1) and $\beta_1, \ldots, \beta_m \in \mathbb{C}$ are linearly independent over the rationals, then P cannot vanish on the set

$$\mathcal{A} = \{(r_1\beta_1 + \ldots + r_m\beta_m,\ \alpha_1^{r_1} \ldots \alpha_m^{r_m}) \mid (r_1, \ldots, r_m) \in \mathbb{Z}^m\}.$$

This $\mathcal{A}$ is a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times$ which is isomorphic to $\mathbb{Z}_+^m$. The dimension m can be chosen arbitrarily. Fixing the degrees of P, we establish the required number of elements in the subset of $\mathcal{A}$ on which P cannot vanish. Moreover, if $\mathcal{A} \subset \mathbb{C} \times \mathbb{R}_{>0}$, this required number of elements is smaller than (or at most the same as) in the general case where $\mathcal{A} \subset \mathbb{C}^2$. This is a consequence of the non-vanishing of the so-called *generalized Vandermonde determinant*, which has a "stronger" version in the positive real case than in the complex case. The proofs of these results are given in Sections 3.1 and 3.2.

Moreover, Section 3.3 describes a zero estimate for complex polynomials in more than two variables. In the case where P is a nonzero polynomial in $\ell > 2$ variables, we find that if $\beta, \alpha, \tilde{\alpha} \in \mathbb{C} \setminus \{0\}$ with α and $\tilde{\alpha}$ multiplicatively independent, then P cannot vanish on the set

$$\mathcal{A} = \{(r\beta, \alpha^r, \tilde{\alpha}^r) \mid r \in \mathbb{Z}\}.$$

This $\mathcal{A}$ is a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times^{\ell-1}$ isomorphic to $\mathbb{Z}_+$. Again, fixing the degrees of P, we establish the required number of elements in the subset of $\mathcal{A}$ on which P cannot vanish.

Finally, the last chapter shortly discusses the found results and open problems.

2 Hilbert's seventh problem

The seventh problem of Hilbert is one of the twenty-three open mathematical problems posed by David Hilbert in 1900. The question asked by Hilbert is the following:

2.1. Hilbert's seventh problem. *Is α^β always transcendental or at least irrational for algebraic α, β, where $\alpha \neq 0, 1$ and β is irrational?*

Let us recall shortly what it means for a number to be algebraic and transcendental.

Definition 2.2. A number $\alpha \in \mathbb{C}$ is called *algebraic* if it is a root of a non-zero polynomial in $\mathbb{Q}[x]$, i.e. if there is a $P \in \mathbb{Q}[x]$ with $P(\alpha) = 0$. A complex number that is not algebraic is called *transcendental*.

The solution to Hilbert's seventh problem was found and proven independently by Aleksandr Gelfond and Theodor Schneider in 1934. This resulted in the Gelfond–Schneider theorem, which is given below in two equivalent versions:

Theorem 2.3. (Gelfond–Schneider)
(a) *Let α, β be algebraic, with $\alpha \neq 0, 1$ and β irrational. Then α^β is transcendental.*
(b) *Let $\alpha_1, \alpha_2, \beta_1, \beta_2$ be nonzero algebraic numbers such that $\ln \alpha_1$ and $\ln \alpha_2$ are linearly independent over $\mathbb{Q}$. Then*

$$\beta_1 \ln \alpha_1 + \beta_2 \ln \alpha_2 \neq 0.$$

Proof of the equivalence:
To show the equivalence of the two statements, first suppose that (a) holds and (b) does not hold, so there exist nonzero algebraic numbers $\alpha_1, \alpha_2, \beta_1, \beta_2$ with $\ln \alpha_1$ and $\ln \alpha_2$ linearly independent and $\beta_1 \ln \alpha_1 + \beta_2 \ln \alpha_2 = 0$. Note that the linear independence of $\ln \alpha_1$ and $\ln \alpha_2$ implies that $-\beta_1/\beta_2$ is irrational and $\alpha_1 \neq 1$. It follows that $\alpha_1^{-\beta_1/\beta_2} = \alpha_2$ is algebraic, contradicting (a). So (a) implies (b). Conversely, suppose (b) holds and (a) does not hold, so there exist algebraic α, β with $\alpha \neq 0, 1$ and β irrational, such that α^β is algebraic. The irrationality of β implies that $\ln \alpha$ and $\ln \alpha^\beta$ are linearly independent. We have

$$\beta \ln \alpha - \ln \alpha^\beta = 0,$$

contradicting (b). Thus (b) implies (a). □

Remark. The use of the property $\ln \alpha^\beta = \beta \ln \alpha$ in the proof is justified, because we fix a branch of the complex logarithm. This means that if we write each $z \in \mathbb{C}$ as $z = |z| e^{i\theta}$, then we fix the ln function only on (for example) the domain $\theta \in (-\pi, \pi]$. Then we can "define" $\ln \alpha^\beta = \beta \ln \alpha$ for all $\alpha, \beta \in \mathbb{C}$.

Inspired by the equivalence of the two Theorems 2.3(a) and 2.3(b), we write $\alpha = \alpha_1$ and $\alpha^\beta = \alpha_2$. Furthermore, we define the following matrix:

$$\mathcal{M} = (a_{r,s}^{u,v}), \quad a_{r,s}^{u,v} = (r + s\beta)^u (\alpha_1^r \alpha_2^s)^v,$$

where $0 \le r, s < 2N$, $0 \le u < K = \lfloor N \ln N \rfloor$, $0 \le v < L = \left\lfloor \frac{N}{\ln N} \right\rfloor$. The rows of $\mathcal{M}$ are indexed by all pairs (r, s), which are $4N^2$ rows in total. The columns of $\mathcal{M}$ are indexed by all pairs (u, v), which are $KL \sim N^2$ columns in total. The proof of the Gelfond–Schneider theorem is now based on the following properties of the matrix $\mathcal{M}$:

Lemma 2.4. *The rank of $\mathcal{M}$ is equal to KL (so $\mathcal{M}$ has full rank).*

As a result of Lemma 2.4, we can select KL linearly independent rows of the matrix $\mathcal{M}$. Denote the collection of these rows by $\mathcal{L}$ and define

$$\Delta = \det\left(a_{r,s}^{u,v}\right)_{(r,s)\in\mathcal{L}}^{0\le u<K,\ 0\le v<L} \neq 0.$$

We have the following estimates for $\ln|\Delta|$:

Lemma 2.5. *For all sufficiently large N, the estimate $\ln|\Delta| \le -N^4$ holds.*

Lemma 2.6. *If α, β, and α^β are algebraic, then $\ln|\Delta| \ge -\frac{1}{2}N^4$ for all sufficiently large N.*

Proof of Theorem 2.3: It follows directly from Lemma 2.5 and Lemma 2.6 that α^β is transcendental: suppose α, β, and α^β are all algebraic, then the two estimates contradict each other. This proves Theorem 2.3. □

It remains to show Lemmas 2.4, 2.5 and 2.6. Lemma 2.5 can be proven analytically, using so-called *interpolation determinants*, whereas Lemma 2.6 follows from Liouville's inequality. For complete proofs of more general versions of these two lemmas, see [3] (Lemmas 6 and 7). A detailed proof of Lemma 2.4 is given below.

Proof of Lemma 2.4: Suppose that the statement is not true, so that $\operatorname{rank}(\mathcal{M}) < KL$. This means that there exist KL constants $C_{u,v} \in \mathbb{C}$ (not all equal to zero) such that the sum of all columns (u, v) of $\mathcal{M}$, each of them multiplied by $C_{u,v}$, is zero:

$$\sum_{u,v} C_{u,v}(r + s\beta)^u(\alpha_1^r\alpha_2^s)^v = 0 \text{ for all } 0 \le r, s < 2N.$$

In other words, there exists a polynomial $P \in \mathbb{C}[x, y]$,

$$P(x, y) = \sum_{u=0}^{K-1} \sum_{v=0}^{L-1} C_{u,v} x^u y^v \not\equiv 0$$

such that $P(r + s\beta, \alpha_1^r\alpha_2^s) = 0$ for all $0 \le r, s < 2N$. We want to show that this is not possible. In order to do this, we use the following lemma:

Lemma 2.7. *Let $P \in \mathbb{C}[x, y]$, $P \not\equiv 0$, $\deg_x(P) < K$ and $\deg_y(P) < L$. Assume that the set*

$$\{\alpha_1^r \alpha_2^s \mid 0 \le r < R_1,\ 0 \le s < S_1\},\ \text{where } \alpha_1, \alpha_2 \in \mathbb{C} \setminus \{0\};\ r, s, R_1, S_1 \text{ are integers,} \tag{2.1}$$

has at least L distinct elements. Furthermore, assume that the set

$$\{r + s\beta \mid 0 \le r < R_2,\ 0 \le s < S_2\},\ \text{where } \beta \in \mathbb{C};\ r, s, R_2, S_2 \text{ are integers,} \tag{2.2}$$

has more than $(K-1)L$ distinct elements. Then at least one of the numbers

$$P(r + s\beta, \alpha_1^r \alpha_2^s),\ \text{where } 0 \le r < R_1 + R_2 - 1 \text{ and } 0 \le s < S_1 + S_2 - 1, \tag{2.3}$$

is nonzero.

Before we prove Lemma 2.7, let us show how to apply the result to our case. Note that at least one of the numbers $\alpha_1 = \alpha$ and $\alpha_2 = \alpha^\beta$ is not a root of unity, because if they both are, say $\alpha_1 = e^{2\pi i k_1/m}$ and $\alpha_2 = e^{2\pi i k_2/n}$, then

$$\frac{m \ln \alpha}{n\beta \ln \alpha} = \frac{2\pi i k_1}{2\pi i k_2},$$

making β a rational number. This contradicts the assumptions in Theorem 2.3. So we know that α_j is not a root of unity for some $j \in \{1, 2\}$. This means that all numbers α_j^k, where $0 \le k < N$, are distinct so that the set (2.1) (with $R_1 = S_1 = N$) contains at least $N \ge L = \lfloor \frac{N}{\ln N} \rfloor$ distinct numbers. Furthermore, because β is irrational, all numbers in the set (2.2) (with $R_2 = S_2 = N$) are distinct, which are $N^2 \ge KL > (K-1)L$ elements in total. Now the application of Lemma 2.7 tells us that at least one of the numbers $P(r + s\beta, \alpha_1^r \alpha_2^s)$, where $0 \le r, s < 2N$, is nonzero. This proves Lemma 2.4. $\square$

In order to prove Lemma 2.7, we will need the following statement about the so-called *generalized Vandermonde determinant*:

Lemma 2.8. *Let $k_1, k_2, \ldots, k_n$ be integers, $0 \le k_1 < k_2 < \ldots < k_n < L$, and let $\mathcal{E} \subset \mathbb{C} \setminus \{0\}$ be a certain numerical set which contains at least L distinct elements. Then there exist n numbers $a_1, \ldots, a_n \in \mathcal{E}$ such that the square matrix $(a_j^{k_i})_{1 \le i,j \le n}$ is nondegenerate.*

Proof: Take L distinct numbers $b_1, \ldots, b_L \in \mathcal{E}$ and consider $\det(b_j^{k-1})_{1 \le k,j \le L}$. This determinant is known as the *Vandermonde determinant* and its value is

$$\det(b_j^{k-1})_{1 \le k,j \le L} = \prod_{1 \le i < j \le L} (b_j - b_i) \ne 0.$$

The determinant is nonzero because $b_1, \ldots, b_L$ are all distinct. It follows that the rows of the matrix $(b_j^{k-1})_{1 \le k,j \le L}$ are linearly independent; in particular the rows with indices

$k_1+1,\ldots,k_n+1$. Consequently, the matrix $(b_j^{k_i})_{1\le i\le n;\ 1\le j\le L}$ has full rank and therefore we can select n linearly independent columns of this matrix. This gives us n numbers $a_1,\ldots,a_n \in \{b_j \mid 1\le j\le L\} \subset \mathcal{E}$ with $\det(a_j^{k_i})_{1\le i,j\le n} \ne 0$. $\square$

Proof of Lemma 2.7: We prove the lemma by contradiction. Assume that the statement is not true, so that all the numbers in (2.3) are zero. We can rewrite the polynomial $P(x,y)$ as:

$$P(x,y) = \sum_{i=1}^{n} Q_i(x) y^{k_i}, \text{ where } Q_i(x) \not\equiv 0 \text{ for } 1\le i\le n;\ 0\le k_1<k_2<\ldots<k_n<L.$$

Define the set

$$\mathcal{E} = \{\alpha_1^r\alpha_2^s \mid 0\le r<R_1, 0\le s<S_1\} \subset \mathbb{C}\setminus\{0\},$$

with at least L distinct elements by the hypothesis. Apply Lemma 2.8 to obtain an n-element subset $\mathcal{L} = \{(r,s)\} \subset \mathcal{E}$ such that

$$B = \det\big((\alpha_1^r\alpha_2^s)^{k_i}\big)_{1\le i\le n;\ (r,s)\in\mathcal{L}} \ne 0.$$

Consider the polynomials

$$P_{r,s}(x,y) = P(x+r+s\beta, \alpha_1^r\alpha_2^s y) = \sum_{i=1}^{n} Q_i(x+r+s\beta)(\alpha_1^r\alpha_2^s)^{k_i} y^{k_i}, \text{ where } (r,s)\in\mathcal{L}.$$

By our assumption,

$$P_{r,s}(r'+s'\beta, \alpha_1^{r'}\alpha_2^{s'}) = P((r+r')+(s+s')\beta, \alpha_1^{r+r'}\alpha_2^{s+s'}) = 0$$
$$\text{for all } (r,s)\in\mathcal{L},\ 0\le r'<R_2 \text{ and } 0\le s'<S_2.$$

Finally, define

$$\begin{aligned}\Gamma \quad &= \big(Q_i(x+r+s\beta)(\alpha_1^r\alpha_2^s)^{k_i}\big)_{1\le i\le n;\ (r,s)\in\mathcal{L}};\\ \Delta(x) &= \det\Gamma,\end{aligned}$$

where the rows of the matrix Γ are indexed by i, $1\le i\le n$, and the columns by $(r,s)\in\mathcal{L}$. Write each Q_i as $Q_i(x) = b_i x^{m_i} + \mathcal{O}(x^{m_i-1})$, where $b_i\ne 0$ for all $1\le i\le n$. Then we can rewrite the determinant $\Delta(x)$ as follows:

$$\begin{aligned}\Delta(x) =\ & \det\big(b_i(\alpha_1^r\alpha_2^s)^{k_i}x^{m_i} + \mathcal{O}(x^{m_i-1})\big)_{1\le i\le n;\ (r,s)\in\mathcal{L}}\\ =\ & \det\left(\Big[\big(b_i(\alpha_1^r\alpha_2^s)^{k_i}\big)_{1\le i\le n;\ (r,s)\in\mathcal{L}} + \big(\mathcal{O}(x^{-1})\big)_{n\times n}\Big]\operatorname{diag}(x^{m_1},\ldots,x^{m_n})\right),\\ & \text{where } \big(\mathcal{O}(x^{-1})\big)_{n\times n} \text{ is an } n\times n \text{ matrix with all entries of order } \mathcal{O}(x^{-1}).\end{aligned}$$

Consequently, we find that $\Delta(x) = Ax^{m_1+\ldots+m_n} + \mathcal{O}(x^{m_1+\ldots+m_n-1})$, where

$$A = \det\big(b_i(\alpha_1^r\alpha_2^s)^{k_i}\big)_{1\le i\le n;\ (r,s)\in\mathcal{L}} = B \cdot b_1 \ldots b_n \neq 0.$$

Thus, $\Delta(x) \neq 0$ and $\deg \Delta(x) = m_1 + \ldots + m_n$. Consider now the system of n linear equations

$$\sum_{i=1}^{n} Q_i(x + r + s\beta)(\alpha_1^r\alpha_2^s)^{k_i} y^{k_i} = P_{r,s}(x, y), \text{ where } (r, s) \in \mathcal{L},$$

in n unknowns y^{k_i}. The corresponding matrix of the system is $\Gamma^\top$, so that the determinant of the system is exactly $\Delta(x)$. To solve this system, we use Cramer's rule:

$$\Delta(x) \cdot y^{k_i} = \Delta_i \text{ for all } 1 \le i \le n,$$

where Δ_i is the determinant of the matrix obtained by replacing the i-th column of $\Gamma^\top$ with the column vector $\big(P_{r,s}(x, y)\big)_{(r,s)\in\mathcal{L}}$. Substituting $x = r' + s'\beta$, $y = \alpha_1^{r'}\alpha_2^{s'}$, where $r' = 1, \ldots, R_2 - 1$ and $s' = 0, 1, \ldots, S_2 - 1$, in this column vector makes it zero so that the determinants Δ_i are zero as well:

$$\Delta(r' + s'\beta) \cdot (\alpha_1^{r'}\alpha_2^{s'})^{k_i} = 0 \quad \text{for all } 1 \le i \le n;\ 0 \le r' < R_2;\ 0 \le s' < S_2.$$

This result implies that $\Delta(r'+s'\beta) = 0$ for all $0 \le r' < R_2$ and $0 \le s' < S_2$. By the hypothesis there are more than $(K - 1)L$ distinct numbers in the set (2.2), so that the number of zeros of the polynomial $\Delta(x)$ is greater than $(K - 1)L$, hence $\deg \Delta(x) > (K - 1)L$. On the other hand,

$$\deg \Delta(x) = m_1 + \ldots + m_n \le (K - 1)n \le (K - 1)L,$$

leading to a contradiction. In conclusion, at least one of the numbers in (2.3) is nonzero. $\Box$

3 Zero estimates

By the fundamental theorem of algebra, any nonzero single-variable complex polynomial of degree n has at most n distinct complex zeros. This means that any such polynomial cannot vanish on any subset of $\mathbb{C}$ with more than n distinct elements. For polynomials in two variables, there is no similar result. For example, consider the polynomial $P \in \mathbb{C}[x, y]$ given by $P = x - y$; then the set of all points $(\alpha, \alpha) \in \mathbb{C}^2$ forms an infinite set of zeros of P. However, inspired by Lemma 2.7 and its proof, we will see that any nonzero $P \in \mathbb{C}[x, y]$ cannot vanish on subsets $\mathcal{A} \subset \mathbb{C}^2$ that have a certain group structure. Moreover, we find that if we fix the degrees of P, then P cannot vanish on particular finite subsets of these groups $\mathcal{A}$. Results of this type are called "zero estimates". In this chapter, we give zero estimates in different settings, specifically for complex polynomials in two variables. First, two analogous statements are given for the "two-variable complex" and "two-variable positive real" cases. The last section of this chapter discusses zero estimates for complex polynomials in more than two variables.

3.1 Zero estimates in the two-variable complex case

Let us start this section by stating a more general version of Lemma 2.7:

Proposition 3.1. *Let $P \in \mathbb{C}[x, y]$, $P \not\equiv 0$, $\deg_x(P) < K$ and $\deg_y(P) < L$. Moreover, let $\mathcal{N}, \mathcal{N}' \subset \mathbb{Z}^m_{\geq 0}$ with $\#\mathcal{N} \geq L$ and $\#\mathcal{N}' > (K-1)L$. Assume that all elements in the set*

$$\{\alpha_1^{r_1} \ldots \alpha_m^{r_m} \mid (r_1, \ldots, r_m) \in \mathcal{N}\}, \text{ where } \alpha_1, \ldots, \alpha_m \in \mathbb{C} \setminus \{0\}, \tag{3.1}$$

are pairwise distinct. Furthermore, assume that all elements in the set

$$\{r_1\beta_1 + \ldots + r_m\beta_m \mid (r_1, \ldots, r_m) \in \mathcal{N}'\}, \text{ where } \beta_1, \ldots, \beta_m \in \mathbb{C}, \tag{3.2}$$

are pairwise distinct. Then at least one of the numbers

$$P(r_1\beta_1 + \ldots + r_m\beta_m,\ \alpha_1^{r_1} \ldots \alpha_m^{r_m}), \text{ where } (r_1, \ldots, r_m) \in \mathcal{N} + \mathcal{N}',$$

is nonzero.

Proof: The proof is completely analogous to the proof of Lemma 2.7: replace each $\alpha_1^r\alpha_2^s$ with $\alpha_1^{r_1} \ldots \alpha_m^{r_m}$, each $r + s\beta$ with $r_1\beta_1 + \ldots + r_m\beta_m$ and let the indices $(r_1, \ldots, r_m)$ be in the corresponding subsets of $\mathbb{Z}^m_{\geq 0}$, instead of including all $0 \leq r_i < R_i$ for certain integers R_i. □

Corollary. *Let $P \in \mathbb{C}[x, y]$, $P \not\equiv 0$. Then P cannot vanish on any subset $\mathcal{A} \subset \mathbb{C}^2$ of the form*

$$\mathcal{A} = \{(r_1\beta_1 + \ldots + r_m\beta_m,\ \alpha_1^{r_1} \ldots \alpha_m^{r_m}) \mid (r_1, \ldots, r_m) \in \mathbb{Z}^m\},$$

where $\beta_1, \ldots, \beta_m \in \mathbb{C}$ are linearly independent (over $\mathbb{Q}$) and $\alpha_1, \ldots, \alpha_m \in \mathbb{C} \setminus \{0\}$ are multiplicatively independent. Moreover, in these settings, $\mathcal{A}$ forms a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times$ with the group operation $(a_1, b_1) \cdot (a_2, b_2) = (a_1 + a_2, b_1b_2)$.

In the corollary above, we used the following definition:

Definition 3.2. Numbers $\alpha_1, \dots, \alpha_m \in \mathbb{C} \setminus \{0\}$ are called *multiplicatively independent* if $\alpha_1^{r_1} \dots \alpha_m^{r_m} \neq 1$ for all $(r_1, \dots, r_m) \in \mathbb{Z}^m$ with $r_i \neq 0$ for at least one $1 \leq i \leq m$.

Proof of the Corollary: The fact that P cannot vanish on the set $\mathcal{A}$ follows directly from applying Proposition 3.1. The linear and multiplicative independence of $\beta_1, \dots, \beta_m$ and $\alpha_1, \dots, \alpha_m$ respectively, ensures that the numbers in both (3.1) and (3.2) are pairwise distinct. Finally, it is straightforward to see that $\mathcal{A}$ is a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times$. In fact (by construction), $\mathcal{A}$ as a group is isomorphic to $\mathbb{Z}_+^m$, with an isomorphism $\mathbb{Z}_+^m \to \mathcal{A}$ given by $(r_1, \dots, r_m) \mapsto (r_1\beta_1 + \dots + r_m\beta_m,\ \alpha_1^{r_1} \dots \alpha_m^{r_m})$. $\square$

Note that for the subgroup $\mathcal{A} \subset \mathbb{C}_+ \times \mathbb{C}_\times$, such that P cannot vanish on $\mathcal{A}$, it is required that not only $\mathcal{A}$, but also the projections $\mathcal{A}_+ \subset \mathbb{C}_+$ and $\mathcal{A}_\times \subset \mathbb{C}_\times$ are isomorphic to $\mathbb{Z}_+^m$. For example, the subgroup $\mathcal{A}' = \{(r_1\beta_1 + \dots + r_m\beta_m,\ 1) \mid (r_1, \dots, r_m) \in \mathbb{Z}^m\} \subset \mathbb{C}_+ \times \mathbb{C}_\times$ is isomorphic to $\mathbb{Z}_+^m$, but this set does not satisfy the conditions of Proposition 3.1.

Remark. We have now described the structure of the set $\mathcal{A}$ such that any nonzero $P \in \mathbb{C}[x, y]$ cannot vanish on this set. Denote the isomorphism $\mathbb{Z}_+^m \to \mathcal{A}$ by ϕ. If we fix $\deg_x(P) < K$ and $\deg_y(P) < L$, then P cannot vanish on any subset of $\mathcal{A}$ that is of the form $\phi(\mathcal{N} + \mathcal{N}')$, where $\mathcal{N},\ \mathcal{N}' \subset \mathbb{Z}_{\geq 0}^m$, with $\#\mathcal{N} \geq L$ and $\#\mathcal{N}' > (K-1)L$. In particular, P cannot vanish on the "m-dimensional cube" $\phi(\{0, \dots, KL-1\}^m) \subset \mathcal{A}$.

3.2 Zero estimates in the two-variable positive real case

An important part of the proof of Lemma 2.7 (and therefore of Proposition 3.1) is the statement about the generalized Vandermonde determinant in Lemma 2.8. In the positive real case, there is a stronger version of this statement, which in turn implies a stronger version of Proposition 3.1. The version of Lemma 2.8 for positive real numbers is as follows (see also [4]):

Lemma 3.3. *Let $k_1, k_2, \dots, k_n$ be integers with $0 \leq k_1 < k_2 \dots < k_n$, and let $a_1 < a_2 < \dots < a_n$ be positive real numbers. Then the matrix $(a_j^{k_i})_{1 \leq i,j \leq n}$ has a positive determinant.*

Before we prove the lemma, let us show how it implies a stronger version of Proposition 3.1. For this, we need the following definition:

Definition 3.4. Let $P \in \mathbb{C}[x_1, \dots, x_n]$. For $1 \leq i_1 < \dots < i_k \leq n$, the *pseudo-degree* of P with respect to $x_{i_1}, \dots, x_{i_k}$, denoted $\operatorname{pdeg}_{x_{i_1}, \dots, x_{i_k}}(P)$, is defined as the number of non-zero terms of P considered as a polynomial in the variables $x_{i_1}, \dots, x_{i_k}$ with its coefficients in the set of polynomials in the remaining variables.

Note that for $P \in \mathbb{C}[x,y]$, $\mathrm{pdeg}_y(P) - 1$ is always bounded above by $\deg_y(P)$. Consequently, if $\mathrm{pdeg}_y(P) = n$ and $\deg_y(P) < L$, then we always have $n \leq L$. For sparse polynomials, which are polynomials with high degrees but only few terms, this pseudo-degree becomes considerably smaller than the degree. Consider also the example below.

Example 3.5. If $P = x^2y^2 + 2xy^2 + y^{999} \in \mathbb{C}[x,y]$, then $P = y^2x^2 + 2y^2x + y^{999} \in (\mathbb{C}[y])[x]$ so that $\mathrm{pdeg}_x(P) = 3 = \deg_x(P) + 1$. Also $P = y^{999} + (x^2 + 2x)y^2 \in (\mathbb{C}[x])[y]$, so that $\mathrm{pdeg}_y(P) = 2 \ll 1000 = \deg_y(P) + 1$. Moreover, $\mathrm{pdeg}_{x,y}(P) = 3$.

Now we can state a zero estimate for the case where we let all $\alpha_1, \ldots, \alpha_m$ be positive real numbers. This result is "stronger" than Proposition 3.1 in the sense that the required number of elements in the set $\mathcal{N}$ now depends on $\mathrm{pdeg}_y(P)$ instead of $\deg_y(P)$:

Proposition 3.6. *Let $P \in \mathbb{C}[x,y]$, $P \not\equiv 0$, $\deg_x(P) < K$ and $\mathrm{pdeg}_y(P) = n$. Moreover, let $\mathcal{N}, \mathcal{N}' \subset \mathbb{Z}^m_{\geq 0}$ with $\#\mathcal{N} \geq n$ and $\#\mathcal{N}' > (K-1)n$. Assume that all elements in the set*

$$\{\alpha_1^{r_1} \ldots \alpha_m^{r_m} \mid (r_1, \ldots, r_m) \in \mathcal{N}\}, \ \textit{where } \alpha_1, \ldots, \alpha_m \in \mathbb{R}_{>0},$$

are pairwise distinct. Furthermore, assume that all elements in the set

$$\{r_1\beta_1 + \ldots + r_m\beta_m \mid (r_1, \ldots, r_m) \in \mathcal{N}'\}, \ \textit{where } \beta_1, \ldots, \beta_m \in \mathbb{C},$$

are pairwise distinct. Then at least one of the numbers

$$P(r_1\beta_1 + \ldots + r_m\beta_m, \ \alpha_1^{r_1} \ldots \alpha_m^{r_m}), \ \textit{where } (r_1, \ldots, r_m) \in \mathcal{N} + \mathcal{N}',$$

is nonzero.

Proof: Again, this proof is completely analogous to the proof of Lemma 2.7 (Proposition 3.1), by applying Lemma 3.3 to ensure that the generalized Vandermonde determinant with the numbers $\alpha_1^{r_1} \ldots \alpha_m^{r_m}$, where $(r_1, \ldots, r_m) \in \mathcal{N}$, does not vanish. $\square$

Corollary. *Let $P \in \mathbb{C}[x,y]$, $P \not\equiv 0$. Then P cannot vanish on any subset $\mathcal{A} \subset \mathbb{C} \times \mathbb{R}_{>0}$ of the form*

$$\mathcal{A} = \{(r_1\beta_1 + \ldots + r_m\beta_m, \ \alpha_1^{r_1} \ldots \alpha_m^{r_m}) \mid (r_1, \ldots, r_m) \in \mathbb{Z}^m\},$$

where $\beta_1, \ldots, \beta_m \in \mathbb{C}$ are linearly independent (over $\mathbb{Q}$) and $\alpha_1, \ldots, \alpha_m \in \mathbb{R}_{>0}$ are multiplicatively independent. In these settings, the set $\mathcal{A}$ forms a subgroup of $\mathbb{C}_+ \times (\mathbb{R}_{>0})_\times$ which is isomorphic to $\mathbb{Z}^m_+$. $\square$

Remark. Let $\mathcal{A}$ be a set of the form that is described in the corollary above, and denote the isomorphism $\mathbb{Z}^m_+ \to \mathcal{A}$ by ϕ. We find that a nonzero polynomial $P \in \mathbb{C}[x,y]$, with $\deg_x(P) < K$ and $\mathrm{pdeg}_y(P) = n$, cannot vanish on the m-dimensional cube $\phi(\{0, \ldots, Kn-1\}^m) \subset \mathcal{A}$. Thus, particularly for sparse polynomials, this cube can be taken significantly smaller than in the $\mathcal{A} \subset \mathbb{C}^2$ case.

It remains to prove Lemma 3.3. There are several ways to prove the lemma, including combinatorially and algebraically. Here we give an analytic proof.

Proof of Lemma 3.3: By induction on n, it follows that a nonzero function

$$f(x) = \sum_{i=1}^{n} c_i x^{k_i},$$

with $c_i, k_i \in \mathbb{R}$ has at most $n-1$ positive zeros. For $n = 1$ we have $f(x) = c_1 x^{k_1}$ which has no positive zero. Consider now arbitrary n and suppose $f(x)$ has n positive zeros. Then also $x^{-k_1} f(x)$ has n positive zeros, and Rolle's theorem provides $n-1$ positive zeros for the derivative $g(x) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{-k_1} f(x))$. But we have

$$g(x) = \sum_{i=1}^{n-1} (k_{i+1} - k_1) c_{i+1} x^{k_{i+1} - k_1 - 1},$$

with at most $n-2$ positive zeros by the induction hypothesis, which leads to a contradiction. Therefore, $f(x)$ has at most $n-1$ positive zeros.

Now consider the rows $(a_1^{k_i} \ a_2^{k_i} \ \dots \ a_n^{k_i})$ for $1 \leq i \leq n$. Suppose that the rows are linearly dependent. Then there are $c_1, c_2, \dots, c_n \in \mathbb{R}$ with at least one $c_i \neq 0$, and

$$\sum_{i=1}^{n} c_i (a_1^{k_i} \ a_2^{k_i} \ \dots \ a_n^{k_i}) = (0 \ 0 \ \dots \ 0).$$

In other words, the nonzero function

$$f(x) = \sum_{i=1}^{n} c_i x^{k_i}$$

has n positive zeros, namely the points $a_1, \dots, a_n$. This is in contradiction with the property stated above. Thus the rows $(a_1^{k_i} \ a_2^{k_i} \ \dots \ a_n^{k_i})$ are linearly independent, making the determinant of $(a_j^{k_i})_{1 \leq i,j \leq n}$ nonzero. Note that this is true in general when the k_i are distinct real numbers and the a_j are distinct positive real numbers.

The positivity of the determinant is proven as follows. Recall that the Vandermonde determinant is given by

$$\det(a_j^{k-1})_{1 \leq k,j \leq n} = \prod_{1 \leq i < j \leq n} (a_j - a_i) > 0.$$

The determinant is positive because $a_1 < a_2 < \dots < a_n$. We would like to know the determinant of $(a_j^{k_i})_{1 \leq i,j \leq n}$, where $0 \leq k_1 < k_2 < \dots < k_n$ are integers. Denote the determinant of $(a_j^{x_i})_{1\le i,j\le n}$ as $D(x_1, x_2, \ldots, x_n)$, this is a function that depends continuously on the parameters $x_i \in \mathbb{R}$. The Vandermonde determinant is in this way given by $D(0, 1, \ldots, n-1)$, and we would like to know $D(k_1, k_2, \ldots, k_n)$. Suppose $k_n > n-1$, otherwise we are done. Suppose also that $D(0, 1, \ldots, n-2, k_n) < 0$. Because D depends continuously on x_n, this would mean that there would be some $r_n \in \mathbb{R}$, $n-1 < r_n < k_n$ with $D(0, 1, \ldots, n-2, r_n) = 0$, which gives a contradiction because we already found this determinant to be nonzero. So $D(0, 1, \ldots, n-2, k_n) > 0$. We can do the same thing for $k_{n-1} > n-2$ and we find that $D(0, 1, \ldots, k_{n-1}, k_n) > 0$, etcetera. Inductively, this proves that $D(k_1, k_2, \ldots, k_n) = \det(a_j^{k_i})_{1\le i,j\le n} > 0$. $\square$

3.3 Zero estimates for polynomials in more than two variables

We start this section by giving a zero estimate for polynomials $P \in \mathbb{C}[x, y, z]$, which will be generalized to a zero estimate for polynomials $P \in \mathbb{C}[x_1 \ldots, x_\ell]$ with $\ell > 2$.

Proposition 3.7. *Let $P \in \mathbb{C}[x, y, z]$, $P \not\equiv 0$, $\deg_x(P) < K$, $\operatorname{pdeg}_{y,z}(P) = n$, so that*

$$P(x, y, z) = \sum_{i=1}^{n} Q_i(x) y^{k_i} z^{\tilde{k}_i}, \text{ where } Q_i(x) \not\equiv 0 \text{ for all } 1 \le i \le n; \ (k_i, \tilde{k}_i) \ne (k_j, \tilde{k}_j) \text{ for all } i \ne j.$$

Let $\beta \in \mathbb{C} \setminus \{0\}$ and let R be an integer with $R > (K-1)n$. Furthermore, assume that all n elements in the set

$$\{\alpha^{k_i} \tilde{\alpha}^{\tilde{k}_i} \mid 1 \le i \le n\}, \text{ where } \alpha, \tilde{\alpha} \in \mathbb{C} \setminus \{0\},$$

are pairwise distinct. Then at least one of the numbers

$$P(r\beta, \alpha^r, \tilde{\alpha}^r), \text{ where } 0 \le r < n + R - 1, \tag{3.3}$$

is nonzero.

Proof: The proof is very similar to the proof of Lemma 2.7. Consider the polynomials

$$\begin{aligned} P_r(x, y, z) &= P(x + r\beta,\ \alpha^r y,\ \tilde{\alpha}^r z) \\ &= \sum_{i=1}^{n} Q_i(x + r\beta)(\alpha^r)^{k_i} y^{k_i} (\tilde{\alpha}^r)^{\tilde{k}_i} z^{\tilde{k}_i} \\ &= \sum_{i=1}^{n} Q_i(x + r\beta)(\alpha^{k_i} \tilde{\alpha}^{\tilde{k}_i})^r y^{k_i} z^{\tilde{k}_i}, \quad 0 \le r < n. \end{aligned}$$

Suppose that all the numbers in (3.3) are zero. Then

$$\begin{aligned} P_r(r'\beta, \alpha^{r'}, \tilde{\alpha}^{r'}) = P((r + r')\beta, \alpha^{r+r'}, \tilde{\alpha}^{r+r'}) = 0 \\ \text{for all } 0 \le r < n,\ 0 \le r' < R. \end{aligned} \tag{3.4}$$

Finally, we define

$$\begin{aligned}\Gamma \quad &= (Q_i(x+r\beta)(\alpha^{k_i}\tilde{\alpha}^{\tilde{k_i}})^r)_{1\le i\le n;\ 0\le r<n};\\ \Delta(x) &= \det\Gamma.\end{aligned}$$

Note that the determinant $B = \det((\alpha^{k_i}\tilde{\alpha}^{\tilde{k_i}})^r)_{0\le i\le n;\ 0\le r<n}$ is nonzero, as this is just the Vandermonde determinant with all entries distinct. Calculating the determinant $\Delta(x)$ in the same way as in Lemma 2.7, we find that $\Delta(x) \neq 0$ and $\deg\Delta(x) \le (K-1)n$. Now consider the system of n linear equations

$$\sum_{i=1}^{n} Q_i(x+r\beta)(\alpha^{k_i}\tilde{\alpha}^{\tilde{k_i}})^r y^{k_i} z^{\tilde{k_i}} = P_r(x,y,z), \text{ where } 0\le r<n,$$

in n unknowns $y^{k_i}z^{\tilde{k_i}}$. The determinant of the matrix corresponding to this system is exactly $\Delta(x)$. Solving the system using (3.4) and Cramer's rule, we obtain that $\Delta(r'\beta) = 0$ for all $0 \le r' < R$. But $\beta \in \mathbb{C}\setminus\{0\}$, so $\Delta(x)$ has at least $R > (K-1)n$ distinct zeros, which is in contradiction with the degree of $\Delta(x)$. $\square$

Corollary. *Let $P \in \mathbb{C}[x,y,z]$, $P \not\equiv 0$. Then P cannot vanish on any subset $\mathcal{A} \subset \mathbb{C}^3$ of the form*

$$\mathcal{A} = \{(r\beta, \alpha^r, \tilde{\alpha}^r) \mid r \in \mathbb{Z}\},$$

where $\beta, \alpha, \tilde{\alpha} \in \mathbb{C}\setminus\{0\}$, and the numbers $\alpha, \tilde{\alpha}$ are multiplicatively independent. Under these conditions, $\mathcal{A}$ forms a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times^2$ isomorphic to $\mathbb{Z}_+$. $\square$

Remark. Let $\mathcal{A}$ be a group as described in the corollary above and denote the isomorphism $\mathbb{Z}_+ \to \mathcal{A}$ by ϕ. Let $P \in \mathbb{C}[x,y,z]$ be nonzero with $\deg_x(P) < K$, $\operatorname{pdeg}_{y,z}(P) = n$; then P cannot vanish on the subset $\phi(\{0,\dots,Kn-1\}) \subset \mathcal{A}$.

Generalization to the ℓ-variable case:

As stated before, Proposition 3.7 can be generalized to a completely analogous statement for nonzero polynomials $P \in \mathbb{C}[x_1,\dots,x_\ell]$ with $\ell > 2$. For this we need to simply let x_1 take the role of x in the three-variable case, and let $x_2,\dots,x_\ell$ take the role of y,z. We obtain that any nonzero $P \in \mathbb{C}[x_1,\dots,x_\ell]$ cannot vanish on any subset $\mathcal{A} \subset \mathbb{C}^\ell$ of the form

$$\mathcal{A} = \{(r\beta, \alpha_1^r, \dots, \alpha_{\ell-1}^r) \mid r \in \mathbb{Z}\},$$

where $\beta, \alpha_1, \dots, \alpha_{\ell-1} \in \mathbb{C}\setminus\{0\}$, and the numbers $\alpha_1,\dots,\alpha_{\ell-1}$ are multiplicatively independent. Moreover, $\mathcal{A}$ is a subgroup of $\mathbb{C}_+ \times \mathbb{C}_\times^{\ell-1}$ isomorphic to $\mathbb{Z}_+$. If we consider the isomorphism $\phi : \mathbb{Z}_+ \to \mathcal{A}$ and fix $\deg_{x_1}(P) < K$, $\operatorname{pdeg}_{x_2,\dots,x_\ell}(P) = n$, then P cannot vanish on the subset $\phi(\{0,\dots,Kn-1\}) \subset \mathcal{A}$.

4 Discussion

In the two-variable case, described in Sections 3.1 and 3.2, we have found that the number of elements in the subset of $\mathcal{A}$ on which $P \in \mathbb{C}[x, y]$ cannot vanish, is dependent on $\deg_y(P)$ if $\mathcal{A} \subset \mathbb{C}^2$ and on $\operatorname{pdeg}_y(P)$ if $\mathcal{A} \subset \mathbb{C} \times \mathbb{R}_{>0}$. This result could be useful in particular when considering sparse polynomials, i.e. polynomials that have high degrees but only few terms.

Furthermore, in the ℓ-variable case with $\ell > 2$, we found that the subgroup $\mathcal{A}$ of $\mathbb{C}_+ \times \mathbb{C}_\times^{\ell-1}$, on which P cannot vanish, is isomorphic to $\mathbb{Z}_+$. The question remains if we can find some $\mathcal{A}$ isomorphic to $\mathbb{Z}_+^m$, such as in the two-variable case, on which P does not vanish. In particular, in three variables and with $m = 2$, we are looking for a group $\mathcal{A}$ that is of the form $\mathcal{A} = \{(r\beta_1 + s\beta_2,\ \alpha_1^r\alpha_2^s,\ \tilde{\alpha_1}^r\tilde{\alpha_2}^s) \mid (r, s) \in \mathbb{Z}^2\}$. If we wish to approach this in the same way as in Proposition 3.7, we need the determinant $\det\big((\alpha_1^r\alpha_2^s)^{k_i}(\tilde{\alpha_1}^r\tilde{\alpha_2}^s)^{\tilde{k_i}}\big)_{0\leq i\leq n;\ (r,s)\in\mathcal{L}}$ to be nonzero for some n-element subset $\mathcal{L} \subset \mathbb{Z}^2$. However, solving (estimating) this determinant is much more complicated than the generalized Vandermonde determinant.

Another remaining question in the general ℓ-variable case is if there exist subgroups $\mathcal{A} \subset \mathbb{C}_+^s \times \mathbb{C}_\times^{\ell-s}$, where $0 \leq s \leq \ell$, on which P cannot vanish.

References

[1] Y. Bugeaud: *Linear Forms in Logarithms*, European Mathematical Society, 2018.

[2] A. Baker: *Transcendental Number Theory*, Cambridge University Press, 1975.

[3] M. Laurent, M. Mignotte & Y. Nesterenko: *Formes linéaires en deux logarithmes et déterminants d'interpolation*, Journal of Number Theory **55**, 285–321, 1995.

[4] S. Fischler, J. Sprang & W. Zudilin: *Many odd zeta values are irrational*, preprint arXiv:1803.08905 [math.NT], 2018.