Exponential-polynomial equations

Zoeteman, M.

M. Zoeteman

Exponential-polynomial equations

Master thesis

July 30, 2019

Thesis supervisor: Dr. J.-H. Evertse

Leiden University
Mathematical Institute

Contents

1 Introduction

Let $\alpha_1, ..., \alpha_n \in \mathbb{C}$ be non-zero multiplicatively independent numbers. We write $\alpha = (\alpha_1, ..., \alpha_n)$, and for $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{Z}^n$ we use the notation

$$\alpha^{\mathbf{x}} = \alpha_1^{x_1} \cdots \alpha_n^{x_n}.$$

We consider the Diophantine equation

$$\alpha^{\mathbf{x}} = f(\mathbf{x}) \text{ in } \boldsymbol{x} \in \mathbb{Z}^n, \tag{1.1}$$

where $f(\mathbf{x}) = f(x_1, ..., x_n) \in \mathbb{C}[x_1, ..., x_n]$ is a polynomial of total degree δ. By a special case of a theorem of Laurent [6, 7, 8] the above equation has only finitely many solutions. Explicit upper bounds for the number of solutions have been studied as well. Let K be the field $K := \mathbb{Q}(\alpha_1, ..., \alpha_n)$ and write $\Delta := \binom{n+\delta}{n}$ and $B := \Delta + 1$. In 2000, Schlickewei and Schmidt [15] proved the following upper bound in the case where K is a number field.

Theorem 1.1. *If K is a number field of degree d, then* (1.1) *has at most $d^{6B^2} 2^{35B^3}$ solutions.*

This is in fact a special case of what Schlickewei and Schmidt proved, namely Theorem 2.12 in §2, which gives an upper bound for the number of solutions to so called exponential-polynomial equations, of which (1.1) is a special case. Then, in 2009, Schmidt [19] generalised this result as follows.

Theorem 1.2. *If the set of all roots of unity in K generates a number field of finite degree d_0, then* (1.1) *has at most $d_0^{6B^2} \exp\left(B^{9B}\right)$ solutions.*

Notice that the previous two results only give an upper bound under certain assumptions on the α_i, and that the upper bounds depend on the α_i. Building further on Schmidt's ideas, Corvaja, Schmidt and Zannier [2] were in 2010 the first to prove an upper bound that only depends on the number of variables n and the total degree δ of f. Their result is as follows.

Theorem 1.3. *Equation* (1.1) *has at most $\exp\left(B^{9B}\right)$ solutions.*

In this thesis we prove the following sharpening of the upper bound in Theorem 1.3.

Theorem 1.4. *Equation* (1.1) *has at most $(8B)^{9B^6}$ solutions.*

Writing the upper bounds as

$$(8B)^{9B^6} = \exp\left(9B^6 \log(8B)\right) \text{ and } \exp\left(B^{9B}\right) = \exp\left(\exp\left(9B \log B\right)\right),$$

we see that our gain is in loosing one exponential. In [2], it is remarked that Theorem 1.3 is easily generalised to the same upper bound for the number of rational solutions to (1.1), if we fix the values of $\log \alpha_1, ..., \log \alpha_n$ (so that we can define $\alpha_i^{x_i} := \mathrm{e}^{x_i \log \alpha_i}$ for $x_i \in \mathbb{Q}$). In the following corollary, we generalise Theorem 1.4 in this way. The proof is given in §4.6.

Corollary 1.5. *If we fix values for $\log \alpha_1, ..., \log \alpha_n$, then the equation*

$$\alpha_1^{x_1} \cdots \alpha_n^{x_n} = f(x_1, ..., x_n) \text{ in } \mathbf{x} \in \mathbb{Q}^n \tag{1.2}$$

has at most $(8B)^{9B^6}$ solutions.

1.1 Lower bounds

One could wonder whether the upper bound in Theorem 1.4 is close to being sharp as $n \to \infty$. Obviously there is no general lower bound for the number of solutions to (1.1). For example, if the α_i are algebraically independent and $f \in \mathbb{Q}[x_1, ..., x_n]$ and $f(0, ..., 0) \neq 1$, then (1.1) has no solutions at all. The question remains whether there are particular classes of equations of the form (1.1) for which the number of solutions is large when $n \to \infty$, and in particular whether this number can get close to the upper bound from Theorem 1.4. Therefore we consider classes of equations

$$\alpha_1^{x_1} \cdots \alpha_n^{x_n} = f(x_1, ..., x_n)$$

for which the number of solutions $g(n)$ satisfies $g(n) \to \infty$ as $n \to \infty$. We construct two of such examples.

Example 1.6. For any multiplicatively independent complex numbers $\alpha_1, ..., \alpha_n$, let f be the linear polynomial $f(x_1, ..., x_n) := \alpha_1 x_1 + ... + \alpha_n x_n$. Let $\mathbf{e_i} \in \mathbb{Z}^n$ denote the vector with i-th coordinate equal to 1 and the other coordinates equal to 0. Then $\mathbf{e_1}, ..., \mathbf{e_n}$ are n solutions to (1.1) for our choice of f. For this f we have $\delta = 1$, so $B = \binom{n+1}{n} + 1 = n + 2$, and (1.1) has at least $n \geq B - 2$ solutions.

Notation. For functions $g, h : \mathbb{Z}_{>0} \to \mathbb{R}_{>0}$ we use the notation $f(n) \sim g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$.

In the following example we apply *Stirling's estimate* $n! \sim \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$ in order to study the growth of the number of solutions as a function of n.

Example 1.7. Let again $\alpha_1, ..., \alpha_n$ be any multiplicatively independent complex numbers. For a vector $\mathbf{y} = (y_1, ..., y_n) \in \{0, 1\}^n$, we write $|\mathbf{y}| := y_1 + ... + y_n$. Now let f be the polynomial

$$f(x_1, ..., x_n) := \sum_{\mathbf{y}\in\{0,1\}^n} (-1)^{|\mathbf{y}|} \alpha_1^{y_1} \cdots \alpha_n^{y_n} (1 - y_1 - x_1) \cdots (1 - y_n - x_n),$$

then we have $f(\mathbf{x}) = \alpha^{\mathbf{x}}$ for each $\mathbf{x} \in \{0, 1\}^n$. Since f has total degree n, we have $\Delta = \binom{2n}{n}$, so by Stirling's estimate

$$B \sim \Delta \sim \frac{(2n)!}{(n!)^2} \sim \frac{4^n}{\sqrt{\pi n}}.$$

Thus, for n large enough we have $B \leq 4^n$, and the equation (1.1) has at least $2^n \geq \sqrt{B}$ solutions. This is a smaller number of solutions than in Example 1.6, but in this new example the total degree δ of f satsfies $\delta \to \infty$ as $n \to \infty$, while the polynomial in Example 1.6 is linear for any n.

There is still a large gap between the lower bounds $B-2$ and $\sqrt{B}$ that we found in our examples, and the upper bound $(8B)^{9B^6}$ from Theorem 1.4. To the best of the author's knowledge, such lower bounds have not been studied yet in the literature, so it is an open challenge to "close the gap", i.e. to construct equations of the form (1.1) with more solutions and/or to improve the upper bound from Theorem 1.4.

In the following example we fix $n = 1$, but the degree δ of f tends to infinity.

Example 1.8. Consider the equation $\alpha^x = f(x)$ in $x \in \mathbb{Z}$, where $\alpha \in \mathbb{C}$ is not a root of unity and where

$$f(x) = \sum_{i=1}^{k} \left(\prod_{\substack{j=1 \\ j\neq i}}^{k} \frac{x - j}{i - j} \right) \alpha^i,$$

with k a positive integer. Notice that for all integers $1 \leq x \leq k$ we have $f(x) = \alpha^x$. In this example we have $B = \binom{1+k}{1} + 1 = k + 2$, and our equation $\alpha^x = f(x)$ has at least $k = B - 2$ solutions.

1.2 Outline of the thesis

We now provide a sketch of how we improve the upper bound from Theorem 1.3 to the one from Theorem 1.4. Let K be a field of characteristic 0, let $n \geq 2$ be an integer and let $\Gamma \subset (K^*)^n$ be a subgroup of rank r. For coefficients $a_1, ..., a_n \in K^*$, we consider the equation

$$a_1 x_1 + ... + a_n x_n = 1 \text{ in } (x_1, ..., x_n) \in \Gamma. \tag{1.3}$$

A solution to this equation is called *non-degenerate* if no proper subsum of the left-hand side vanishes. Evertse, Schlickewei and Schmidt [3] proved in 2002 that the number of non-degenerate solutions to (3.4) is at most $\exp\left((6n)^{3n}(r+1)\right)$. In 2009, Amoroso and Viada [1] improved this upper bound to $(8n)^{4n^4(n+r+1)}$. In §4 we prove Theorem 1.4. In this proof we follow the arguments given in the proofs of Theorem 1.2 and Theorem 1.3. At the point where these proofs apply the upper bound of Evertse, Schlickewei and Schmidt, we apply the new upper bound from Amoroso and Viada. Using this new upper bound in the further calculations, we will arrive at our result as stated in Theorem 1.4.

In §2 we study some general theory of exponential-polynomial equations and their relation to linear recurrence sequences. In §3, we discuss some basic terminology related to linear equations, the results of Evertse, Schlickewei and Schmidt, and of Amoroso and Viada, and (very briefly) some of the techniques behind the proofs of these results.

Notation. Throughout this thesis, we use the notation $\mathbb{N} := \{1, 2, 3, ...\}$, and by $\overline{\mathbb{Q}}$ we denote an algebraic closure of $\mathbb{Q}$. For a set S and $n \in \mathbb{N}$, we use the notation $S^n = S \times ... \times S$ for the n-fold cartesian product. For a multiplicative group G and $n \in \mathbb{N}$, we use the notation $G^{[n]} := \{g^n : g \in G\}$.

2 Linear recurrence sequences and exponential-polynomial equations

In this section we treat some standard theory of linear recurrence sequences and exponential-polynomial equations, as well as the current best known upper bounds for the number of solutions of such equations. This section is for a large part based on Section 10.11 from [5]. For an extensive treatment of the theory of linear recurrence sequences and exponential-polynomial equations, we refer to the chapter on linear recurrence sequences by Schmidt in [11].

2.1 Linear recurrence sequences

We start with treating some standard theory of linear recurrence sequences. Let K be an algebraically closed field of characteristic 0.

Definition 2.1. A *linear recurrence sequence* in K is a sequence $L = (a_n)_{n=0}^{\infty} \subset K$ given by a linear recurrence

$$a_n = c_1 a_{n-1} + ... + c_k a_{n-k} \text{ for } n \geq k, \tag{2.1}$$

where $c_1, ..., c_k \in K$ and $c_k \neq 0$, and by initial values $a_0, ..., a_{k-1} \in K$, not all zero.

The most famous example of a linear recurrence sequence is the Fibonacci sequence $F = (F_n)_{n=0}^{\infty} = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, ...)$ given by the recurrence $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$ and initial values $F_0 = 0$ and $F_1 = 1$. Of course this sequence can also be given by the recurrence $F_n = 2F_{n-2} + F_{n-3}$ and by initial values $F_0 = 0, F_1 = F_2 = 1$. Therefore we need the following definition.

Definition 2.2. Let $L = (a_n)_{n=0}^{\infty}$ be a linear recurrence sequence as in Definition 2.1. The minimal $k \in \mathbb{N}$ for which an expression such as (2.1) exists, is called *the order of* L.

Lemma 2.3. *Let $L = (a_n)_{n=0}^{\infty}$ be a linear recurrence sequence as in Definition 2.1, where k is the order of L. Then the coefficients $c_1, ..., c_k$ for which (2.1) holds are unique.*

Proof. Suppose we have for all $n \geq k$ that

$$c_1 a_{n-1} + ... + c_k a_{n-k} = a_n = d_1 a_{n-1} + ... + d_k a_{n-k},$$

with $d_1, ..., d_k, c_1, ..., c_k \in \mathbb{C}$, with c_k, d_k non-zero, and assume that $(c_1, ..., c_k) \neq (d_1, ..., d_k)$. Let i be the smallest $1 \leq j \leq k$ such that $c_j \neq d_j$. Combining the two expressions for a_n yields

$$a_{n-i} = \frac{1}{d_i - c_i} \left((c_{i+1} - d_{i+1}) a_{n-i-1} + ... + (c_k - d_k a_{n-k}) \right),$$

contradicting the minimality of k. □

Definition 2.4. Let $L = (a_n)_{n=0}^{\infty}$ be a linear recurrence sequence as in Definition 2.1, where k is the order of L. Then we associate to L its *companion polynomial* $f_L \in \mathbb{C}[X]$ given by $f_L(X) := X^k - c_1 X^{k-1} - ... - c_{k-1} X - c_k$.

Note that the companion polynomial is well-defined, because of Lemma 2.3. For example, the Fibonacci sequence F has companion polynomial

$$f_F = X^2 - X - 1 = (X - \varphi)(X - \psi),$$

where $\varphi := \frac{1+\sqrt{5}}{2}$ is the golden ratio and $\psi := \frac{1-\sqrt{5}}{2}$ is its conjugate in the field $\mathbb{Q}\left(\sqrt{5}\right)$. It is well known that the Fibonacci numbers can be calculated directly without using recursion, via the expression

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}} \text{ for all } n \geq 0,$$

This turns out to be a general phenomenon, as is indicated by Theorem 2.5, which states that there is a one to one correspondence between linear recurrence sequences and exponential-polynomial expressions in one variable. Let L be a linear recurrence sequence as in Definition 2.1 of order k. Since K is algebraically closed, we can factorise the companion polynomial f_L as $f_L(X) = (X - \alpha_1)^{e_1} \cdots (X - \alpha_m)^{e_m}$, with $\alpha_1, ..., \alpha_m \in K$ pairwise different and $e_1, ..., e_m \in \mathbb{N}$.

Theorem 2.5. *There are polynomials $g_1(X), ..., g_m(X)$ with $\deg g_i = e_i - 1$ for each i, such that for all $n \geq 0$ we have*

$$a_n = g_1(n)\alpha_1^n + ... + g_m(n)\alpha_m^n. \tag{2.2}$$

On the other hand, every sequence $(a_n)_{n=0}^{\infty}$ of the form (2.2) is a linear recurrence sequence.

Proof. The idea of the proof is to write the formal power series $\sum_{n=0}^{\infty} a_n X^n$ in two different ways, which gives relation (2.2) for a_n. The polynomial $f(X) := X^k f_L\left(\frac{1}{X}\right) = 1 - c_1 X - ... - c_k X^k$ factorises as

$$f(X) = (1 - \alpha_1 X)^{e_1} \cdots (1 - \alpha_m X)^{e_m}.$$

Define $g(X) := f(X) \sum_{n=0}^{\infty} a_n X^n$, and write $c_0 := -1$ and $c_i := 0$ for $i \geq k+1$. Then we have

$$g(X) = -\left(\sum_{i=0}^{\infty} c_i X^i\right)\left(\sum_{n=0}^{\infty} a_n X^n\right) = -\sum_{l=0}^{\infty}\left(\sum_{i+n=l} c_i a_n\right) X^l.$$

For $l \geq k$, the recurrence relation (2.1) gives that $\sum_{i+n=l} c_i a_n = 0$, so g is in fact a polynomial of degree at most $k-1$. Partial fraction decomposition gives that there are coefficients $c_{i,j} \in K$ for $1 \leq i \leq m$ and $1 \leq j \leq e_i$, with $c_{i,j} \neq 0$ if $j = e_i$, such that

$$\begin{aligned}
\sum_{n=0}^{\infty} a_n X^n &= \frac{g(X)}{f(X)} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{e_i} \frac{c_{i,j}}{(1-\alpha_i X)^j} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{e_i} c_{i,j} \sum_{n=0}^{\infty} \binom{n+j-1}{n-1} \alpha_i^n X^n = \sum_{n=0}^{\infty}\left(\sum_{i=1}^{m}\left(\sum_{j=1}^{e_i} c_{i,j}\binom{n+j-1}{n-1}\right)\alpha_i^n\right) X^n.
\end{aligned}$$

Comparing coefficients on both sides, we see that a_n satisfies a relation of the form (2.2).

Now assume $(a_n)_{n=0}^{\infty}$ is a sequence of the form (2.2). We show that a_n satisfies the relation (2.1). Writing $g_i(x) = \sum_{j=0}^{e_i-1} c_{i,j} x^j$, we have for all $n \geq k$ that

$$\begin{aligned}
\sum_{h=1}^{k} c_h a_{n-h} &= \sum_{h=1}^{k} c_h \sum_{i=1}^{m}\sum_{j=0}^{e_i-1} c_{i,j}(n-h)^j \alpha_i^{n-h} \\
&= \sum_{i=1}^{m}\sum_{j=0}^{e_i-1} c_{i,j} \sum_{h=1}^{k} c_h \alpha_i^{n-h}(n-h)^j.
\end{aligned} \tag{2.3}$$

Let $x\frac{d}{dx}$ be the operator that first takes the derivative of a function in x and then multiplies it with x. By induction on j one can show that for each $j \geq 0$ we have

$$\left(x\frac{d}{dx}\right)^j \left(x^{n-k} f_L(x)\right) = \sum_{h=0}^{k} c_h x^{n-h}(n-h)^j. \tag{2.4}$$

Because of the factorisation of f_L, we have $f_L^{(j)}(\alpha_i) = 0$ for each $0 \leq j \leq e_i - 1$, so evaluating (2.4) in α_i yields

$$\sum_{h=0}^{k} c_h \alpha_i^{n-h}(n-h)^j = 0.$$

Combining this with (2.3) provides us with

$$\sum_{h=1}^{k} c_h a_{n-h} = \sum_{i=1}^{m} \sum_{j=0}^{e_i-1} c_{i,j} n^j \alpha_i^n = a_n.$$

□

Definition 2.6. For $L = (a_n)_{n=0}^{\infty}$ a linear recurrence sequence, we call $N(L) := \#\{n \in \mathbb{Z}_{\geq 0} : a_n = 0\}$ the *zero multiplicity* of L (which is possibly infinite).

Because of Theorem 2.5, determining the zero multiplicity of linear recurrence sequences is equivalent to determining the number of solutions to exponential-polynomial equations

$$g_1(n)\alpha_1^n + ... + g_m(n)\alpha_m^n = 0 \text{ in } n \in \mathbb{Z}_{\geq 0}, \tag{2.5}$$

with $g_1, ..., g_m \in K[X]$ non-zero polynomials and $\alpha_1, ..., \alpha_m \in K^*$ pairwise different.

Definition 2.7. Let $L = (a_n)_{n=0}^{\infty}$ be a linear recurrence sequence with companion polynomial $f_L = X^k - c_1 X^{k-1} - ... - c_{k-1}X - c_k$, where f_L factorizes as $f_L(X) = (X-\alpha_1)^{e_1} \cdots (X-\alpha_m)^{e_m}$. Then L is called *non-degenerate* if $\frac{\alpha_i}{\alpha_j}$ is not a root of unity for all $i \neq j$, and L is called *degenerate* if it is not non-degenerate.

Degenerate linear recurrence sequences can have infinitely many zeroes. For example, the sequence $L = (a_n)_{n=0}^{\infty}$ with $a_0 = 2$, $a_1 = 0$ and $a_n = 4a_{n-2}$ for $n \geq 2$, satifies $a_n = 2^n + (-2)^n$ and has companion polynomial $f_L(X) = (X-2)(X+2)$, while $\frac{2}{-2} = -1$ is a root of unity. The following theorem states that this cannot happen for non-degenerate linear recurrence sequences.

Theorem 2.8. *If L is a non-degenerate linear recurrence sequence, then L has finite zero multiplicity.*

In 1935, Skolem [20] proved Theorem 2.8 in the case $K = \mathbb{Q}$, and in the same year Mahler [10] proved the case $K = \overline{\mathbb{Q}}$. Finally, in 1953 Lech [9] proved Theorem 2.8 for any field K of characteristic 0. Their proofs are based on p-adic analysis. After these results were proven, study has been made of explicit upper bounds for zero multiplicities of linear recurrence sequences. Schmidt [18] was the first to prove an upper bound for the zero multiplicity only depending on the order of the sequence.

Theorem 2.9. *If L is a non-degenerate linear recurrence sequence of order k, then*

$$N(L) \leq \exp(\exp(\exp(3k \log k))).$$

The current best known upper bound was obtained by Amoroso and Viada [1].

Theorem 2.10. *If L is a non-degenerate linear recurrence sequence of order k, then*

$$N(L) \leq \exp(\exp(70k)).$$

This main reason for this improvement is that Amoroso and Viada used their improved upper bound for the number of solutions to linear equations in a multiplicative group [1] (Theorem 3.6 in this thesis) instead of an earlier result from Evertse, Schlickewei and Schmidt [3] (Theorem 3.5 in this thesis).

2.2 General exponential-polynomial equations

Equation of the type (2.5), which correspond to zeroes of linear recurrence sequences, are exponential polynomial equations in one variable. Now we consider general exponential-polynomial equations. Let again K be an algebraically closed field of characteristic 0, and let $n \in \mathbb{N}$. Recall the notation $\alpha^{\mathbf{x}} = \alpha_1^{x_1} \cdots \alpha_n^{x_n}$, for $\alpha = (\alpha_1, ..., \alpha_n) \in (K^*)^n$ and $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{Z}^n$. Let $f_1, ..., f_m \in K[x_1, ..., x_n]$ be non-zero polynomials, and let $\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}} \in (K^*)^n$, and consider the equation

$$f_1(\mathbf{x})\alpha_{\mathbf{1}}{}^{\mathbf{x}} + ... + f_m(\mathbf{x})\alpha_{\mathbf{m}}{}^{\mathbf{x}} = 0 \text{ in } \mathbf{x} \in \mathbb{Z}^n. \tag{2.6}$$

A solution $\mathbf{x} \in \mathbb{Z}^n$ to (2.6) is called *non-degenerate* if there is no strict non-empty subset I of $\{1, ..., m\}$ such that

$$\sum_{i \in I} f_i(\mathbf{x})\alpha_{\mathbf{i}}{}^{\mathbf{x}} = 0.$$

Let $\Lambda(\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}}) \subset \mathbb{Z}^n$ be the subgroup

$$\Lambda(\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}}) := \{\mathbf{x} \in \mathbb{Z}^n : \alpha_{\mathbf{1}}{}^{\mathbf{x}} = ... = \alpha_{\mathbf{m}}{}^{\mathbf{x}}\}.$$

Laurent ([6], [7] [8]) proved the following finiteness result.

Theorem 2.11. *If* $\Lambda(\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}}) = \{\mathbf{0}\}$*, then* (2.6) *has finitely many non-degenerate solutions.*

As mentioned in the introduction, this implies that (1.1) has finitely many solutions, as we now show. Writing $\alpha_{\mathbf{1}} := \alpha$, $\alpha_{\mathbf{2}} := (1, ..., 1)$, $f_1 = 1$ and $f_2 = -f$, equation (1.1) is equivalent to

$$f_1(\mathbf{x})\alpha_{\mathbf{1}}{}^{\mathbf{x}} + f_2(\mathbf{x})\alpha_{\mathbf{2}}{}^{\mathbf{x}} = 0.$$

Since $\alpha_1, ..., \alpha_n$ are multiplicatively independent, we have $\Lambda(\alpha_{\mathbf{1}}, \alpha_{\mathbf{2}}) = \{\mathbf{0}\}$, so indeed (1.1) has finitely many solutions by Theorem 2.11.

Theorem 2.11 also implies Theorem 2.8. By Theorem 2.5 we have to show that (2.5) has finitely many solutions, where $g_1, ..., g_m \in K[x]$ are non-zero polynomials and where none of the quotients $\frac{\alpha_i}{\alpha_j}$ $(i \neq j)$ is a root of unity. We proceed by induction on m. If $m = 1$, then the result is obvious, since a non-zero polynomial in one variable has only finitely many zeroes. Now let $m \geq 2$ and suppose that equations of the form (2.5) in fewer than m variables have only finitely many solutions. Since none of the quotients $\frac{\alpha_i}{\alpha_j}$ $(i \neq j)$ is a root of unity, we have $\Lambda(\alpha_1, ..., \alpha_m) = \{\mathbf{0}\}$, so by the one variable case of Theorem 2.11, equation (2.5) has only finitely many non-degenerate solutions. Now assume $n \in \mathbb{Z}_{\geq 0}$ is a degenerate solution to (2.5), then there is a strict non-empty subset $I \subset \{1, ..., m\}$ with

$$\sum_{i \in I} g_i(n)\alpha_i^n = 0,$$

such that the above does not hold for any strict non-empty subset of I. But then n is a non-degenerate solution to the equation above, so by the induction hypothesis there are finitely many possibilities for n for each I. Since there are finitely many possibilities for I as well, there are finitely many degenerate solutions to (2.5), completing the inductive proof of Theorem 2.8.

It is still an open problem to find an upper bound for the number of non-degenerate solutions to (2.6) that only depends on the number of variables n, the total degrees of the polynomials f_i, and the number of polynomials m (and *not* on the values of the $\alpha_{\mathbf{i}}$), under the assumption that

$\Lambda(\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}}) = \{\mathbf{0}\}$. Such upper bounds are known for exponential-polynomial equations with only one polynomial involved, i.e. equations of the form (1.1), and for exponential-polynomial equations in one variable, i.e. equations of the form (2.5), as we discussed in §1 resp. §2.1. The following theorem of Schlickewei and Schmidt [15] from 2000 provides an upper bound for the number of non-degenerate solutions to (2.6) over number fields, with a dependence on the $\alpha_{\mathbf{i}}$.

Theorem 2.12. *Assume that* $\Lambda(\alpha_{\mathbf{1}}, ..., \alpha_{\mathbf{m}}) = \{\mathbf{0}\}$. *Further suppose that the coordinates of the vectors* $\alpha_{\mathbf{i}}$ *and the coefficients of the polynomials* f_i $(i = 1, ..., m)$ *are contained in an algebraic number field of degree* d. *For* $i = 1, ..., m$, *let* δ_i *be the total degree of* f_i *and let*

$$B := \max\left\{n, \sum_{i=1}^{m} \binom{n+\delta_i}{n}\right\}.$$

Then the number of non-degenerate solutions to (2.6) *is at most* $d^{6B^2}2^{35B^3}$.

Theorem 1.1 is a special case of this upper bound (in order to see this, note that the we may assume that the coeffients of f lie in K, as we show in §4.3.1).

3 Linear equations

3.1 Results on linear equations

Definition 3.1. A multiplicatively written abelian group Γ has rank $r \in \mathbb{Z}_{\geq 0}$ if it has a free subgroup Γ_0 of rank r such that for each $\gamma \in \Gamma$ there is an $n \in \mathbb{N}$ with $\gamma^n \in \Gamma_0$.

Let K be a field of characteristic 0, let $n \geq 2$ be an integer and let $\Gamma \subset (K^*)^n$ be a subgroup of some finite rank r, where $(K^*)^n = K \times ... \times K$ is the n-fold cartesian product. Let $\lambda_1, ..., \lambda_n \in K^*$ be some coefficients, and consider the linear equation

$$\lambda_1 x_1 + ... + \lambda_n x_n = 0 \text{ in } (x_1, ..., x_n) \in \Gamma. \tag{3.1}$$

A solution $(x_1, ..., x_n)$ of (3.1) is called *non-degenerate* if there is no strict and non-empty subset I of $\{1, ..., n\}$ satisfying

$$\sum_{i \in I} \lambda_i x_i = 0. \tag{3.2}$$

Similarly, a solution to the equation

$$\lambda_1 x_1 + ... + \lambda_n x_n = 1 \text{ in } (x_1, ..., x_n) \in \Gamma \tag{3.3}$$

is called non-degenerate if (3.2) does not hold for any I. We call two solutions $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$ to (3.1) *proportional*, denoted as $(x_1, ..., x_n) \sim (y_1, ..., y_n)$, if there is a $\lambda \in K^*$ such that $(x_1, ..., x_n) = (\lambda y_1, ..., \lambda y_n)$. Equations of the type (3.3) and upper bounds for the number of non-degenerate solutions to such equations play a crucial role in the proof of Theorem 1.4. The following lemma, which is standard, allows us to switch between linear equations of the form (3.1) and (3.3).

Lemma 3.2. *Let S_0 denote the set of non-degenerate solutions to (3.1) in Γ, and let $S_0/\sim$ denote the corresponding collection of proportionality classes. Let $\Gamma' \subset (K^*)^{n-1}$ be the subgroup $\Gamma' := \left\{ \left(\frac{x_1}{x_n}, ..., \frac{x_{n-1}}{x_n} \right) : (x_1, ..., x_n) \in \Gamma \right\}$ and let S_1 be the set of non-degenerate solutions $(x_1, ..., x_{n-1}) \in \Gamma'$ to the equation*

$$\gamma_1 x_1 + ... + \gamma_{n-1} x_{n-1} = 1,$$

where $\gamma_i := -\frac{\lambda_i}{\lambda_n}$ for all $1 \leq i \leq n-1$. Then the sets $S_0/\sim$ and S_1 are in bijection.

Proof. Define the map $f : (S_0/\sim) \to S_1$ by

$$[(x_1, ..., x_n)] \mapsto \left(\frac{x_1}{x_n}, ..., \frac{x_{n-1}}{x_n} \right),$$

and the map $g : S_1 \to S_0/\sim$ by

$$(y_1, ..., y_{n-1}) \mapsto [(y_1, ..., y_{n-1}, 1)].$$

Writing out the definitions shows that these maps are well-defined and each other's inverses. □

Remark 3.3. The reason we consider only non-degenerate solutions to (3.3), is that we are interested in upper bounds for the number of solutions, while there may be infinitely many degenerate solutions. For example, assume we have an infinite multiplicative subgroup $G \subset K^*$ of finite rank, and let $\Gamma := G^n \subset (K^*)^n$. Suppose we have a solution $(x_1, ..., x_n) \in \Gamma$ to (3.3), satisfying $\lambda_1 x_1 + ... + \lambda_m x_m = 1$ and $\lambda_{m+1} x_{m+1} + ... + \lambda_n x_n = 0$, for some $1 \leq m < n$. Then for each $g \in G$, $(x_1, ..., x_m, g x_{m+1}, ..., g x_n) \in \Gamma$ is a solution to (3.3) as well, giving us infinitely many degenerate solutions to (3.3).

Evertse [4] proved in 1999 the following result for linear equations in roots of unity.

Theorem 3.4. *For any $\lambda_1, ..., \lambda_n \in \mathbb{C}^*$, equation (3.3) has at most $(n+1)^{3(n+1)^2}$ non-degenerate solutions $(x_1, ..., x_n)$ for which $x_1, ..., x_n \in \mathbb{C}$ are all roots of unity.*

Next, we consider more general results on linear equations. For coefficients $a_1, ..., a_n \in K^*$, we consider the equation

$$a_1 x_1 + ... + a_n x_n = 1 \text{ in } (x_1, ..., x_n) \in \Gamma. \tag{3.4}$$

Evertse, Schlickewei and Schmidt proved the following upper bound for the number of non-degenerate solutions to this equation, which remarkably only depends on the number of variables n and on the rank r of Γ, but not on the field K, the coefficients a_i or the group Γ itself.

Theorem 3.5. *Equation (3.4) has at most $\exp\left((6n)^{3n}(r+1)\right)$ non-degenerate solutions.*

In 2009, Amoroso and Viada [1] improved this upper bound as follows.

Theorem 3.6. *Equation (3.4) has at most $(8n)^{4n^4(n+r+1)}$ non-degenerate solutions.*

3.2 Some ideas and techniques behind the proofs of Theorems 3.5 and 3.6

We very briefly sketch some of the ideas and techniques behind the proofs of these results. Since the proofs have a considerable overlap, we treat both proofs at the same time, and point out where they differ. This subsection is based on the proof of Theorem 3.5 given in Section 6.3 of [5]. The first step in the proof is to show via a so called specialisation argument that it suffices to consider the case where K is an algebraic number field, i.e. a finite extension of $\mathbb{Q}$ (see §4.2 for another specialisation argument). In the proof a distinction is made between small and large solutions in terms of algebraic heights. In §3.2.1, we recall some algebraic number theory needed to define these heights. For a more extensive treatment of this material, we refer to [12]. In order to estimate the number of non-degenerate solutions to (3.4), a version of the so called Subspace Theorem is used, which we discuss in §3.2.2. Then we provide a sketch of the proof in §3.2.3.

3.2.1 Algebraic number theory and heights

Let $\mathcal{O}_K$ denote the *ring of integers* of a number field K, i.e. $\mathcal{O}_K$ consists of all elements $\alpha \in K$ for which there is a monic polynomial $f \in \mathbb{Z}[X]$ with $f(\alpha) = 0$. A *fractional $\mathcal{O}_K$-ideal* is a non-zero $\mathcal{O}_K$-submodule I of K, such that for some $x \in K^*$ we have $xI \subset \mathcal{O}_K$. Henceforth, we mean by a prime ideal of $\mathcal{O}_K$ a non-zero prime ideal. Fractional $\mathcal{O}_K$-ideals have unique prime ideal factorisation. That is, any fractional $\mathcal{O}_K$-ideal I can be factorised in a unique way as

$$I = \prod_{\mathfrak{p}} \mathfrak{p}^{\mathrm{ord}_{\mathfrak{p}}(I)},$$

where the product runs over all prime ideals $\mathfrak{p}$, and the exponents $\mathrm{ord}_{\mathfrak{p}}(I)$ are all integers, and are non-zero for at most finitely many $\mathfrak{p}$. For $x \in K^*$, let $\mathrm{ord}_{\mathfrak{p}}(x)$ be the order of the principal fractional ideal $x\mathcal{O}_K$ at $\mathfrak{p}$, and let $\mathrm{ord}_{\mathfrak{p}}(0) := \infty$. Notice that a fractional $\mathcal{O}_K$-ideal I is an integral $\mathcal{O}_K$-ideal (i.e., an ideal of the ring $\mathcal{O}_K$) if and only if $\mathrm{ord}_{\mathfrak{p}}(I) \geq 0$ for each prime $\mathfrak{p}$. Applying this to principal fractional ideals, we see that $\mathcal{O}_K$ can be described as

$$\mathcal{O}_K = \{x \in K : \mathrm{ord}_{\mathfrak{p}}(x) \geq 0 \text{ for each prime } \mathfrak{p}\}, \tag{3.5}$$

and that its unit group $\mathcal{O}_K^*$ equals

$$\mathcal{O}_K^* = \{x \in K : \mathrm{ord}_{\mathfrak{p}}(x) = 0 \text{ for each prime } \mathfrak{p}\},$$

Finally, a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ has *absolute norm* $N_K(\mathfrak{p}) := |\mathcal{O}_K/\mathfrak{p}|$.

For $\sigma : K \to \mathbb{R}$ a real embedding of K (i.e. a field homomorphism), we call the singleton set $\{\sigma\}$ a *real place* of K. For $\sigma : K \to \mathbb{C}$ a complex embedding (i.e. a field homomorphism with $\sigma(K) \not\subset \mathbb{R}$), we call the set $\{\sigma, \overline{\sigma}\}$ a *complex place* of K. An *infinite place* of K is a real or complex place of K. A *finite place* of K is a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$. Finally, a place of K is a finite or infinite place of K, and we let M_K denote the set of all places of K.

For each place $v \in M_K$, there is an absolute value $|\cdot|_v$ on K, defined as follows: for $x \in K$, let

$$|x|_v := \begin{cases} |\sigma(x)|, & \text{if } v = \{\sigma\} \text{ is a real place,} \\ |\sigma(x)|^2, & \text{if } v = \{\sigma, \overline{\sigma}\} \text{ is a complex place,} \\ N_K(\mathfrak{p})^{-\mathrm{ord}_{\mathfrak{p}}(x)}, & \text{if } v = \mathfrak{p} \text{ is a finite place,} \end{cases}$$

where $|\cdot|$ denotes the standard Euclidean absolute value on $\mathbb{C}$.

By (3.5), $\mathcal{O}_K$ consists of all elements $x \in K$ having $|x|_{\mathfrak{p}} \leq 1$ for *each* prime. Allowing finitely many exceptions brings us to the concept of *S-integers*. Let S be a finite set of places, containing all the infinite places. Then the ring of S-integers $\mathcal{O}_S$ is defined by

$$\mathcal{O}_S := \{x \in K : |x|_v \leq 1 \text{ for each } v \in M_K \setminus S\},$$

and its unit group $\mathcal{O}_S^*$ equals

$$\mathcal{O}_S^* := \{x \in K : |x|_v = 1 \text{ for each } v \in M_K \setminus S\}.$$

Now for $\mathbf{x} = (x_1, ..., x_n) \in K^n$, the *height* $H(\mathbf{x})$ is defined as

$$H(\mathbf{x}) := \prod_{v \in M_K} \max\left(1, |x_1|_v, ..., |x_n|_v\right)^{1/[K:\mathbb{Q}]}. \tag{3.6}$$

This definition turns out to be independent on the choice of K, so the height is well-defined. By large solutions of (3.4), we mean solutions $\mathbf{x}$ of this equation for which $H(\mathbf{x})$ is large.

3.2.2 The Subspace Theorem

A *linear form* in $\mathbb{C}[X_1, ..., X_n]$ is a homogeneous polynomial $L \in \mathbb{C}[X_1, ..., X_n]$ of degree 1, i.e. $L = \alpha_1 X_1 + ... + \alpha_n X_n$ with $\alpha_1, ..., \alpha_n \in \mathbb{C}$ not all zero. A collection of n linear forms $L_i = \alpha_{i,1} X_1 + ... + \alpha_{i,n} X_n$ $(1 \leq i \leq n)$ is called linearly dependent if there are $\lambda_1, ..., \lambda_n \in \mathbb{C}$, not all zero, such that $\lambda_1 L_1 + ... + \lambda_n L_n$ is zero in $\mathbb{C}[X_1, ..., X_n]$, and is called linearly independent otherwise. Further, for a vector $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{C}^n$, we denote by $||\mathbf{x}||$ the maximum norm of $\mathbf{x}$, i.e.

$$||\mathbf{x}|| := \max\{|x_1|, ..., |x_n|\}.$$

Finally, we denote by $\overline{\mathbb{Q}}$ the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. The following theorem is the original Subspace Theorem, proved by Schmidt in 1972.

Theorem 3.7. *Let $L_1, ..., L_n \in \mathbb{C}[X_1, ..., X_n]$ be n linearly independent linear forms, with coefficients in $\overline{\mathbb{Q}}$, and let $C > 0, \epsilon > 0$. Then the set of solutions of the inequality*

$$|L_1(\mathbf{x}) \cdots L_n(\mathbf{x})| \leq C||\mathbf{x}||^{-\epsilon} \text{ in } \mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}, \tag{3.7}$$

is contained in a union of finitely many proper linear subspaces of $\mathbb{Q}^n$.

Later, many variations and generalisations of Schmidt's Subspace Theorem were obtained. We state a particular generalisation of Theorem 3.7 that is convenient for our purposes. Let K be a number field of degree d. For $v \in M_K$, denote by K_v the completion of K w.r.t. the absolute value $|\cdot|_v$, and let $\overline{K_v}$ be an algebraic closure of K_v. It turns out that $|\cdot|_v$ can be extended in a unique way to an absolute value on $\overline{K_v}$, which we also denote by $|\cdot|_v$.

Theorem 3.8. *Let S be a finite set of places of K containing all the infinite places. For any $v \in S$, let $L_{1,v}, ..., L_{n,v} \in \overline{K}_v[X_1, ..., X_n]$ be n linearly independent forms $(n \geq 2)$, with coefficients that are algebraic over K, let C_v be a positive real constant, and let $c_{1,v}, ..., c_{n,v} \in \mathbb{R}$ be such that*

$$\sum_{v \in S} \sum_{i=1}^{n} c_{i,v} < 0.$$

Then the set of solutions to the system of inequalities

$$|L_{i,v}(\mathbf{x})|_v \leq C_v H(\mathbf{x})^{d \cdot c_{i,v}} \ (v \in S, i = 1, ..., n) \ \textit{in} \ \mathbf{x} \in \mathcal{O}_S^n,$$

is contained in a union of finitely many proper linear subspaces of K^n*, where* $H(\mathbf{x})$ *is the height as defined in* (3.6).

Theorem 3.8 is equivalent to the p-adic Subspace Theorem as proved by Schlickewei [16, Theorem 2.1]. In the particular case that S consists only of the infinite places it follows from work of Schmidt [17]. Theorem 3.7 and Theorem 3.8 are both ineffective, in the sense that the proofs do not reveal in which particular subspaces the solutions lie, and do not give an upper bound for how many such subspaces there are. In the proofs of Theorem 3.5 and 3.6, a quantitative version of Theorem 3.8 is used, which gives an upper bound for the number of subspaces independent of the cardinality of S.

3.2.3 Sketch of the proofs

The proofs of Theorems 3.5 and 3.6 consist of the following steps.

(i): As mentioned earlier, the theorem is reduced to the case where K is a number field. Furthermore, it is shown that we may suppose that Γ is finitely generated. Choose a finite set T of generators for Γ. Let $S \subset M_K$ consist of all infinite places and the primes that occur in the factorisations of the coordinates of the elements of T. Then S is finite, and $\Gamma \subset (\mathcal{O}_S^*)^n$, which enables us to apply a quantitative version of Theorem 3.8 to the solutions of (3.4).

(ii): A quantitative version of Theorem 3.8 is applied to show that the solutions of (3.4) with *large* height are contained in a bounded number of subspaces of K^n.

(iii): One finds an upper bound for the number of subspaces of K^n in which the solutions to (3.4) with *small* height lie. Schmidt was the first to find such an upper bound, and this upper bound is applied in the proof of Theorem 3.5. Using techniques from commutative algebra, Amoroso and Viada improved this upper bound, and obtained Theorem 3.6 as a corollary. We remark that in the paper of Evertse, Schlickewei and Schmidt [3], the upper bound for the number of small solutions is larger than the upper bound for number of large solutions.

(iv) One applies induction on n to bound the number of non-degenerate solutions to (3.4). By step (ii) and (iii), it suffices to find a convenient upper bound for the number of non-degenerate solutions to (3.4) lying in an $(n-1)$-dimensional subspace of K^n. Combining the equation for such a subspace with (3.4) then yields an equation in fewer than n variables, which enables us to apply the induction hypothesis.

For more details on the proof, we refer to Section 6.3 from [5].

4 Proof of Theorem 1.4

4.1 Two results from the paper of Corvaja, Schmidt and Zannier

The main new ingredients in the paper of Corvaja, Schmidt and Zannier [2] compared to the paper of Schmidt [19], are the following two results. Let K be any number field of degree D.

Theorem 4.1. *Let $u \in \overline{\mathbb{Q}}^*$, and assume there is a positive integer n with $u^n \in K^*$. Let q be the smallest such n, and let ζ be a primitive q-th root of unity. Then we have*

$$[K(u) : K] \geq \frac{\varphi(q)}{[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]}.$$

Theorem 4.2. *Let G be a multiplicative subgroup of $\overline{\mathbb{Q}}^*$ generated by n elements.*
Define $H := G \cap K^*$, and assume that the quotient group G/H is finite. Let $V \subset \overline{\mathbb{Q}}$ be a K-vector space of finite dimension r. Then V has non-empty intersection with at most $g(r, D)^n$ cosets of H in G, where we may take

$$g(r, D) = g_1(r, D) = 2(r+1)^{12(r+1)^2} D^4,$$

or

$$g(r, D) = g_2(r, D) = 52320(r+1)^{6.3(r+1)^2} D^{1+3\left(\log^+ \log^+ D\right)^{-1}},$$

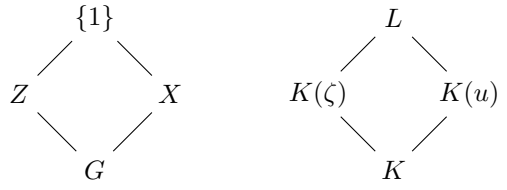
where $\log^+ x = \max\{1, \log x\}$.

Remark 4.3. In [2] this lemma is proved with $g_2(r, D) = c(r+1)^{12(r+1)^2} D^{1+\left(\log^+ \log^+ D\right)^{-1}}$ instead, for some unknown absolute constant $c > 0$. We follow the same proof, but make a little extra effort with technicalities to obtain our bound for $g_2(r, D)$. However, for our proof of Theorem 1.4, the upper bound $g_1(r, D)$ suffices.

We prove these theorems following the paper of Corvaja, Schmidt and Zannier [2].

4.1.1 Proof of Theorem 4.1

Let ζ be a primitive q-th root of unity, and let $L := K(u, \zeta)$. As L is the decomposition field of the polynomial $X^q - u^q$ over K, the extension L/K is Galois. Writing $G := \mathrm{Gal}(L/K)$, $X := \mathrm{Gal}(L/K(u))$ and $Z := \mathrm{Gal}(L/K(\zeta))$, we have the following Galois correspondence,



. (4.1)

Using multiplicativity of field degrees in towers, we see that

$$[K(u) : K] = \frac{|Z|}{|X|}[K(\zeta) : K]. \tag{4.2}$$

In the Lemmas 4.4, 4.5 and 4.7 we estimate the quantities $|X|, |Z|$ and $[K(\zeta) : K]$, and then we combine these estimates with (4.2) to prove Theorem 4.1. Crucial in the proofs of these lemmas are the maps $\alpha : G \to (\mathbb{Z}/q\mathbb{Z})^*$ and $\beta : G \to \mathbb{Z}/q\mathbb{Z}$, which we define as follows. For $g \in G$, let $\alpha(g) \in (\mathbb{Z}/q\mathbb{Z})^*$ and $\beta(g) \in \mathbb{Z}/q\mathbb{Z}$ be such that $g(\zeta) = \zeta^{\alpha(g)}$ and $g(u) = u\zeta^{\beta(g)}$. Notice that any $g \in G$ is completely determined by $\alpha(g)$ and $\beta(g)$. Further, for $g, h \in G$ we have

$$\zeta^{\alpha(gh)} = gh(\zeta) = g\left(\zeta^{\alpha(h)}\right) = \zeta^{\alpha(g)\alpha(h)},$$
$$\zeta^{\beta(gh)} u = gh(u) = g\left(\zeta^{\beta(h)} u\right) = \zeta^{\alpha(g)\beta(h)+\beta(g)} u,$$

hence

$$\alpha(gh) = \alpha(g)\alpha(h) \text{ and } \beta(gh) = \alpha(g)\beta(h) + \beta(g). \tag{4.3}$$

By our Galois correspondence (4.1) we have

$$Z = \{g \in G : \alpha(g) = 1\} \text{ and } X = \{g \in G : \beta(g) = 0\}. \tag{4.4}$$

From (4.3) and (4.4) we conclude that α and the restriction of β to Z are homomorphisms.

Lemma 4.4. *We have $|Z| = \frac{q}{m}$ for some divisor m of q.*

Proof. By (4.4) and by the Galois correspondence (4.1), $\beta|_Z$ has kernel $Z \cap X = \{1\}$, hence $\beta|_Z$ is an injective homomorphism from Z into $\mathbb{Z}/(q\mathbb{Z})$. Thus, for some divisor m of q we have

$$Z \cong \beta(Z) = \{mx : x \in \mathbb{Z}/(q\mathbb{Z})\}, \tag{4.5}$$

and therefore $|Z| = \frac{q}{m}$. □

Write $q = m'q'$, with m' composed of prime factors dividing m and with q' coprime to m. For any $x \in \mathbb{Z}/q\mathbb{Z}$, write $\overline{x} := x \mod m$ for its reduction in $\mathbb{Z}/m\mathbb{Z}$.

Lemma 4.5. *The order of X is at most $\frac{\varphi(q')m'}{m}$.*

Proof. The proof consists of the following five observations.

(i): *Via α, we can interpret G/Z as a subgroup of $(\mathbb{Z}/q\mathbb{Z})^*$.*
Indeed, this follows by the isomorphism theorem, as α has kernel Z by (4.4). This also implies that G/Z is commutative and that we can interpret $\alpha(X)$ as a subgroup of G/Z.

(ii): *The elements $\beta(g)$, $g \in G$ are coprime in $\mathbb{Z}/q\mathbb{Z}$, i.e. there is no divisor $d > 1$ of q such that $\beta(G) \subset d(\mathbb{Z}/q\mathbb{Z})$.*
Suppose the contrary, then for such a d we have for all $g \in G$ that

$$g\left(u^{q/d}\right) = \left(\zeta^{\frac{\beta(g)}{d}}\right)^q u^{\frac{q}{d}} = u^{\frac{q}{d}}.$$

Then we have $u^{\frac{q}{d}} \in K$, contradicting the minimality of q.

(iii): *For $g \in G$, the values $\overline{\alpha(g)}$ and $\overline{\beta(g)}$ only depend on the class of g in G/Z. Hence we can define $\overline{\alpha_\tau} := \overline{\alpha(g)}$ and $\overline{\beta_\tau} := \overline{\beta(g)}$ if $\tau \in G/Z$ is the class of g.*
Namely, for all $z \in Z$ we have $\alpha(z) = 1$, so $\overline{\alpha(z)} = 1$, and $\overline{\beta(z)} = 0$ by (4.5). Combining these facts with (4.3) we see that for all $g \in G$ and $z \in Z$ we have the desired congruences

$$\alpha(gz) \equiv \alpha(g)\alpha(z) \equiv \alpha(g) \equiv \alpha(z)\alpha(g) \equiv \alpha(zg) \mod m,$$

$$\beta(gz) \equiv \beta(g) + \alpha(g)\beta(z) \equiv \beta(g) \equiv \beta(z) + \alpha(z)\beta(g) \equiv \beta(zg) \mod m.$$

Definition 4.6. Let $T \subset G/Z$ be the set of all $\tau \in G/Z$ with $\overline{\beta_\tau} = 0$.

(iv): *We have* $|X| \leq |T|$.
Recalling from observation (i) that via α we can interpret G/Z as a subgroup of $(\mathbb{Z}/q\mathbb{Z})^*$, we can consider $\alpha(X)$ as a subgroup of G/Z as well. For $g \in X$ we have $\beta(g) = 0$ by (4.4), and thus $\overline{\beta(g)} = 0$. Thus for $\tau \in \alpha(X) \subset G/Z$ we have $\overline{\beta_\tau} = 0$, hence $\tau \in T$, which shows that $\alpha(X) \subset T$. By (4.4) and the Galois correspondence (4.1) we have $\ker(\alpha|_X) = Z \cap X = \{1\}$, so α is injective on X, and $|X| = |\alpha(X)| \leq |T|$.

(v): *We have* $|T| \leq \frac{\varphi(q')m'}{m}$.
Again we identify G/Z with a subgroup of $(\mathbb{Z}/q\mathbb{Z})^*$. This identification shows that G/Z is abelian, hence for all $\sigma, \tau \in G/Z$ we have by (4.3),

$$\overline{\beta_\sigma} + \overline{\alpha_\sigma}\overline{\beta_\tau} = \overline{\beta_{\sigma\tau}} = \overline{\beta_{\tau\sigma}} = \overline{\beta_\tau} + \overline{\alpha_\tau}\overline{\beta_\sigma},$$

which yields

$$\overline{\beta_\sigma}(\overline{\alpha_\tau} - 1) = \overline{\beta_\tau}(\overline{\alpha_\sigma} - 1).$$

Therefore, we have for all $\sigma \in G/Z$ and $\tau \in T$ that $\overline{\beta_\sigma}(\overline{\alpha_\tau} - 1) = 0$ in $\mathbb{Z}/m\mathbb{Z}$. Fixing $\tau \in T$ and letting σ run through G/Z, it follows by observation (ii) that $\overline{\alpha_\tau} = 1$. Using this and the inclusion $T \subset G/Z \subset (\mathbb{Z}/q\mathbb{Z})^*$, we get an inclusion

$$T \subset \{x \in \mathbb{Z}/q\mathbb{Z} : (x, q) = 1 \text{ and } x \equiv 1 \mod m\} = \{x \in \mathbb{Z}/q\mathbb{Z} : (x, q') = 1 \text{ and } x \equiv 1 \mod m\}, \tag{4.6}$$

since $x \equiv 1 \mod m$ implies that $(x, m') = 1$. The equation $(x, q') = 1$ has $\varphi(q')$ solutions in $\mathbb{Z}/q'\mathbb{Z}$, and the equation $x \equiv 1 \mod m$ has $\frac{m'}{m}$ solutions in $\mathbb{Z}/m'\mathbb{Z}$. Since $q = q'm'$ and $(q', m') = 1$, it follows by the Chinese remainder theorem and (4.6) that $|T| \leq \varphi(q')\frac{m'}{m}$.

Combining observations (iv) and (v) completes the proof of Lemma 4.5. □

Lemma 4.7. *We have* $[K(\zeta) : K] = \frac{\varphi(q)}{[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]}$.

Proof. The extension $\mathbb{Q}(\zeta)/(K \cap \mathbb{Q}(\zeta))$ is Galois. Taking the compositum with K, we see that the extension $K(\zeta)/K$ is Galois as well with degree

$$[K(\zeta) : K] = [\mathbb{Q}(\zeta) : K \cap \mathbb{Q}(\zeta)] = \frac{[\mathbb{Q}(\zeta) : \mathbb{Q}]}{[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]} = \frac{\varphi(q)}{[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]}.$$

□

Since $q = q'm'$ and q' and m' are coprime, we have $\varphi(q) = \varphi(q')\varphi(m')$. Combining this observation with the previous three lemmas and equation (4.2), we get

$$[K(u) : K] \geq \frac{q/m}{\varphi(q')m'/m} \frac{\varphi(q)}{[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]} = \frac{q\varphi(m')}{m'[K \cap \mathbb{Q}(\zeta) : \mathbb{Q}]}. \tag{4.7}$$

By the expression $\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right)$ and because m' divides q, we have $\frac{\varphi(m')}{m'} \geq \frac{\varphi(q)}{q}$. Applying this to (4.7) completes the proof of Theorem 4.1. □

4.1.2 Proof of Theorem 4.2

Preparatory work

Lemma 4.8. *For all $n \geq 2$ we have*

$$\varphi(n) \geq \frac{n}{5 \log^+ \log^+ n}.$$

Proof. For $2 \leq n \leq 15$ we have $\log^+ \log^+ n = 1$ and the claim can be checked by inspection. For $n \geq 16$, we have $\log^+ \log^+ n = \log \log n$. Letting γ denote the Euler–Mascheroni constant, we use the estimate [14]

$$\varphi(n) > \frac{n}{\mathrm{e}^{-\gamma} \log \log n + \frac{3}{\log \log n}}.$$

We conclude the proof by noting that for all $n \geq 16$ we have

$$\mathrm{e}^{-\gamma} \log \log n + \frac{3}{\log \log n} \leq 5 \log \log n.$$

□

Lemma 4.9. *For all $x, y \in \mathbb{R}_{>0}$ we have*

$$\log^+ \log^+(xy) \leq 2 \log^+ \log^+(x) \cdot \log^+ \log^+(y). \tag{4.8}$$

Proof. If $x < 16$ and $y < 16$, then

$$\log^+ \log^+(xy) < \log^+ \log^+(256) < 2.$$

If $x \geq 16$ and $y \geq 16$, then (4.8) is equivalent to the inequality

$$\log \log(xy) \leq 2 \log \log(x) \cdot \log \log(y),$$

which holds. Finally, if $x \geq 16$ and $y < 16$, then

$$\begin{aligned} \log^+ \log^+(xy) &< \log^+ \log^+(16x) \\ &\leq \log^+ \log^+(16) \log^+ \log^+(x) \leq 2 \log^+ \log^+(y) \log^+ \log^+(x). \end{aligned}$$

□

Let K be any subfield of $\mathbb{C}$.

Definition 4.10. A complex number ξ is called a radical of K if there is a positive integer n such that $\xi^n \in K^*$.

Definition 4.11. For $\xi_1, ..., \xi_n \in \mathbb{C}^*$, let $K(\xi_1 : ... : \xi_n)$ denote the extension of K generated by all the quotients $\frac{\xi_i}{\xi_j}$, where $1 \leq i, j \leq n$ with $\xi_j \neq 1$.

Lemma 4.12. *Let $n \geq 2$ be an integer, and suppose we have coefficients $\lambda_1, ..., \lambda_n \in K^*$. Assume $(\xi_1, ..., \xi_n) \in \mathbb{C}^n$ is a non-degenerate solution to the equation*

$$\lambda_1 \xi_1 + ... + \lambda_n \xi_n = 0, \tag{4.9}$$

with the ξ_i radicals of K. Then we have

$$[K(\xi_1 : ... : \xi_n) : K] \leq n^{3n^2}.$$

In the case that K is a number field of degree D, there is $q \in \mathbb{N}$ with

$$\left(\frac{\xi_i}{\xi_j}\right)^q \in K^* \text{ for all } 1 \leq i, j \leq n, \tag{4.10}$$

where q can both be bounded by

$$q \leq Q_1(n, D) = n^{6n^2} D^2$$

and by

$$q \leq Q_2(n, D) = 80 n^{3n^2} \log^+(n) D \log^+ \log^+(D).$$

Remark 4.13. In [2] this lemma is proved with $Q_2(n, D) = cn^{6n^2} D \log^+ \log^+ D$ instead, for some unknown absolute constant $c > 0$. We follow the same proof, but make a little extra effort with technicalities to obtain our bound for $Q_2(n, D)$.

Proof. In order to prove the first statement, we show that $K(\xi_1 : ... : \xi_n)$ has at most n^{3n^2} embeddings in $\mathbb{C}$ that are the identity on K. Let σ be such an embedding. For any i there is an $n_i \in \mathbb{N}$ with $\xi_i^{n_i} \in K^*$ and thus $\sigma(\xi_i)^{n_i} = \xi_i^{n_i}$. Therefore, $\sigma(\xi_i) = \xi_i \zeta_{\sigma,i}$ where $\zeta_{\sigma,i}$ is an n_i-th root of unity. Now by (4.9) we have

$$\lambda_1 \xi_1 \zeta_{\sigma,1} + ... + \lambda_n \xi_n \zeta_{\sigma,n} = 0.$$

Considering the above as an equation in the variable $(\zeta_{\sigma,1}, ..., \zeta_{\sigma,n})$, it follows from combining Theorem 3.4 and Lemma 3.2 that there are at most n^{3n^2} solutions $(\zeta_{\sigma,1}, ..., \zeta_{\sigma,n})$ modulo proportionality. Since

$$\sigma\left(\frac{\xi_i}{\xi_j}\right) = \frac{\xi_i}{\xi_j} \frac{\zeta_{\sigma,i}}{\zeta_{\sigma,j}},$$

σ is determined by the quotients $\frac{\zeta_{\sigma,i}}{\zeta_{\sigma,j}}$ and thus by the proportionality class of $(\zeta_{\sigma,1}, ..., \zeta_{\sigma,n})$. We conclude that there are at most n^{3n^2} possible values for σ, which concludes the proof of the first statement.

Now assume K is a number field of degree D. Consider the multiplicative subgroup G of $\mathbb{C}^*$ generated by the elements $\frac{\xi_i}{\xi_j}$ $(1 \leq i, j \leq n)$, and let H be the subgroup $H := G \cap K^*$. Now G/H is generated by the classes of the elements $\frac{\xi_i}{\xi_j}$, and since the ξ_i are radicals of K, the element $\frac{\xi_i}{\xi_j}$ have finite order as elements of G/H. Hence, G/H is a finite abelian group, so by the fundamental theorem for abelian groups we have a group isomorphism

$$G/H \cong \mathbb{Z}/n_1\mathbb{Z} \oplus ... \oplus \mathbb{Z}/n_t\mathbb{Z}$$

for some natural numbers $n_1, ..., n_t$ with $n_1 | n_2 | ... | n_{t-1} | n_t$. Thus, letting $q := n_t$, we can pick an element $u \in G$, such that its class $\overline{u}$ in G/H has order q and note that each element of G/H has order dividing q. In particular, this q satisfies (4.10). Now we can bound $[K(u) : K]$ from below by Theorem 4.1 and from above by the first statement of the current lemma, and combine these bounds to bound q as desired. To be precise, we have

$$\frac{\varphi(q)}{D} \leq [K(u) : K] \leq n^{3n^2}, \tag{4.11}$$

which yields $\varphi(q) \leq Dn^{3n^2}$. Without loss of generality we may assume that $q \geq 3$ (otherwise q certainly satisfies the upper bounds we wish to prove), and therefore $\varphi(q) > \sqrt{q}$, which yields $q \leq D^2 n^{6n^2}$. Combining (4.11) with Lemma 4.8 we get

$$q \leq 5Dn^{3n^2} \log^+ \log^+ q. \tag{4.12}$$

Combining the bounds $\varphi(q) > \sqrt{q}$ and (4.11) with Lemma 4.9 we get

$$\log^+ \log^+ q \leq 2 \log^+ \log^+ \varphi(q) \leq 4 \log^+ \log^+(D) \cdot \log^+ \log^+ \left(n^{3n^2}\right). \tag{4.13}$$

For $n \geq 3$ we have

$$\log^+ \log^+ \left(n^{3n^2}\right) = \log^+ \left(3n^2 \log^+(n)\right) \leq \log^+ \left(n^4\right) \leq 4 \log^+(n), \tag{4.14}$$

while one can check that $\log^+ \log^+ \left(n^{3n^2}\right) \leq 4 \log^+(n)$ also holds for $n = 2$. Combining the estimates (4.12), (4.13) and (4.14) yields the second bound for q. □

Having established this lemma, we start with the proof of Theorem 4.2.

Introducing a lattice

The first key idea of the proof is to introduce a lattice as follows. Let $\xi_1, ..., \xi_n \in G$ be a set of generators of G. For $\mathbf{m} = (m_1, ..., m_n) \in \mathbb{Z}^n$, we write

$$\xi^{\mathbf{m}} = \xi_1^{m_1} \cdots \xi_n^{m_n}.$$

Composing the surjective group homomorphism $\mathbb{Z}^n \to G,\ \mathbf{m} \mapsto \xi^{\mathbf{m}}$ with the projection map $\pi : G \to G/H$, we obtain a surjective homomorphism $\varphi : \mathbb{Z}^n \to G/H,\ \mathbf{m} \mapsto [\xi^{\mathbf{m}}]$. Now

$$L := \ker(\varphi) = \{\mathbf{m} \in \mathbb{Z}^n : \xi^{\mathbf{m}} \in K^*\}$$

is a subgroup of $\mathbb{Z}^n$, and by the isomorphism theorem φ yields an isomorphism

$$\overline{\varphi} : \mathbb{Z}^n / L \xrightarrow{\sim} G/H \tag{4.15}$$

Since G/H is finite, there is for each $1 \leq i \leq n$ a positive integer m_i such that $\xi_i^{m_i} \in H$. Thus the n vectors $(m_1, 0, ..., 0), ..., (0, ..., 0, m_n) \in \mathbb{Z}^n$ all lie in L, so L is a lattice of full rank in $\mathbb{Z}^n$.

Reduction to representatives

The next step in the proof is the observation that whether a residue class of G/H has non-empty intersection with V, only depends on one representative of the class lying in V or not. Let R be a set of representatives for the classes of $\mathbb{Z}^n/L$, then by the isomorphism (4.15) the set $\{\xi^{\mathbf{m}} : \mathbf{m} \in R\}$ is a set of representatives of the classes from G/H. For any $\mathbf{m} \in R$ and $\mathbf{m}' \in L$, we have $\xi^{\mathbf{m}} \in V$ if and only if $\xi^{\mathbf{m}} \cdot \xi^{\mathbf{m}'} \in V$, because $\xi^{\mathbf{m}'}$ lies in K^* and V is a K-vector space. Thus, the number of cosets of H in G that have non-empty intersection with V equals $|M|$, where

$$M := \{\mathbf{m} \in R : \xi^{\mathbf{m}} \in V\}. \tag{4.16}$$

Applying induction and constructing a special basis of V

We prove Theorem 4.2 by induction on the dimension r of V. Firstly, consider the case $r = 1$, then $V = \lambda \cdot K$ for some $\lambda \in \overline{\mathbb{Q}}^*$. Assume we have vectors $\mathbf{m}, \mathbf{m}' \in M$. Then there are $\alpha, \alpha' \in K^*$ such that $\xi^{\mathbf{m}} = \lambda\alpha$ and $\xi^{\mathbf{m}'} = \lambda\alpha'$. This shows that $\xi^{\mathbf{m}-\mathbf{m}'} = \frac{\alpha}{\alpha'}$ lies in H, hence $\mathbf{m} - \mathbf{m}' \in L$. Since $\mathbf{m}$ and $\mathbf{m}'$ lie in R this implies that $\mathbf{m} = \mathbf{m}'$, so $|M| \leq 1 \leq \min\{g_1(1, D), g_2(1, D)\}$, which completes the base case of our induction.

Now let $r \geq 2$ be given and assume that for any vector space V of dimension $r' < r$ we have $|M| \leq \min\{g_1(r', D), g_2(r', D)\}$. We may assume without loss of generality that V is generated by the set $\{\xi^{\mathbf{m}} : \mathbf{m} \in M\}$. Namely, assume this is not the case, then the subspace $V' \subset V$ generated by $\{\xi^{\mathbf{m}} : \mathbf{m} \in M\}$ has some dimension $r' < r$, and we have

$$M = \{\mathbf{m} \in R : \xi^{\mathbf{m}} \in V'\}.$$

Now the induction hypothesis yields $|M| \leq \min\{g_1(r', D)^n, g_2(r', D)^n\} \leq \min\{g_1(r, D)^n, g_2(r, D)^n\}$. In order to construct a convenient basis for V, we use this observation combined with some more lattice theory.

Definition 4.14. For $h \in \mathbb{N}$ and for vectors $\mathbf{a_1}, ..., \mathbf{a_h} \in M$, let $q(\mathbf{a_1}, ..., \mathbf{a_h})$ be the smallest $q \in \mathbb{N}$ such that for all $1 \leq i, j \leq h$ we have $q(\mathbf{a_i} - \mathbf{a_j}) \in L$.

Remark 4.15. As $\mathbb{Z}^n/L$ is finite, there is a $q \in \mathbb{N}$ satisfying $q\mathbf{m} \in L$ for each $\mathbf{m} \in \mathbb{Z}^n$. Therefore, $q(\mathbf{a_1}, ..., \mathbf{a_h})$ is a well-defined number.

Remark 4.16. For any $1 \leq i, j \leq h$, we have that $q(\mathbf{a_1}, ..., \mathbf{a_h}) \cdot (\mathbf{a_i} - \mathbf{a_j}) \in L$, so $q(\mathbf{a_1}, ..., \mathbf{a_h})$ is divisble by the order $q(\mathbf{a_i}, \mathbf{a_j})$ of $\mathbf{a_i} - \mathbf{a_j}$ in $\mathbb{Z}^n/L$.

We define the quantity $T := \max\{q(\mathbf{u}, \mathbf{v}) : \mathbf{u}, \mathbf{v} \in M\}$. Since no two elements of M are congruent modulo L, we have $q(\mathbf{u}, \mathbf{v}) = 1$ if and only if $\mathbf{u} = \mathbf{v}$. Choose $\mathbf{a_1}, \mathbf{a_2} \in M$ with $q(\mathbf{a_1}, \mathbf{a_2}) = T$. We may clearly suppose that $|M| > 1$, and thus $T > 1$ and $\mathbf{a_1} \neq \mathbf{a_2}$. Since $\mathbf{a_1}$ and $\mathbf{a_2}$ represent different classes of $\mathbb{Z}^n/L$, we have $\mathbf{a_1} - \mathbf{a_2} \notin L$ and thus $\xi^{\mathbf{a_1}-\mathbf{a_2}} \notin K$. Thus, $\xi^{\mathbf{a_1}}$ and $\xi^{\mathbf{a_2}}$ are linearly independent over K, while we assumed that V is generated by the set $\{\xi^{\mathbf{m}} : \mathbf{m} \in M\}$, so we can extend $\{\xi^{\mathbf{a_1}}, \xi^{\mathbf{a_2}}\}$ to a basis $(\xi^{\mathbf{a_1}}, ..., \xi^{\mathbf{a_r}})$ of V for some $\mathbf{a_1}, ..., \mathbf{a_r} \in M$.

A case distinction

By the previous paragraph, for each $\mathbf{m} \in M$, there are unique coefficients $\lambda_{1,\mathbf{m}}, ..., \lambda_{r,\mathbf{m}} \in K$ such that

$$\xi^{\mathbf{m}} = \lambda_{1,\mathbf{m}}\xi^{\mathbf{a_1}} + ... + \lambda_{r,\mathbf{m}}\xi^{\mathbf{a_r}}. \tag{4.17}$$

Viewing the above as an equation of the form (4.9) in $r + 1$ variables, we make a distinction between the degenerate and the non-degenerate solutions - the former can be estimated with the induction hypothesis, while the latter can be estimated with Lemma 4.12. More precisely, writing

$$M_i := \{\mathbf{m} \in M : \lambda_{i,\mathbf{m}} = 0 \text{ in } (4.17)\}$$

for $1 \leq i \leq r$, and

$$M' := \{\mathbf{m} \in M : \lambda_{i,\mathbf{m}} \neq 0 \text{ in } (4.17) \text{ for each } i\},$$

we can estimate $|M|$ by

$$|M| \leq |M'| + \sum_{i=1}^{r} |M_i|. \tag{4.18}$$

Fix $1 \leq i \leq r$, and let $\mathbf{m} \in M_i$. Then the K-vectorspace V_i generated by the vectors $\{\xi^{\mathbf{a_l}} : 1 \leq l \leq r, l \neq i\}$, has dimension $r-1$, and

$$M_i \subset \{\mathbf{m} \in R : \xi^{\mathbf{m}} \in V_i\}.$$

Thus, by the induction hypothesis we obtain

$$|M_i| \leq \min\{g_1(r-1, D), g_2(r-1, D)\}.$$

This provides us with

$$\sum_{i=1}^{r} |M_i| \leq r \min\{g_1(r-1, D), g_2(r-1, D)\} \leq \frac{1}{2} \min\{g_1(r, D), g_2(r, D)\}. \tag{4.19}$$

For any $\mathbf{m} \in M'$, the second part of Lemma 4.12 shows that

$$p(\mathbf{m}) := q(\mathbf{m}, \mathbf{a_1}, ..., \mathbf{a_h}) \leq \min\{Q_1(r+1, D), Q_2(r+1, D)\} =: S. \tag{4.20}$$

By Remark 4.16, it follows that $l := \frac{p(\mathbf{m})}{q(\mathbf{a_1}, \mathbf{a_2})}$ is a positive integer that is at most $\frac{S}{T}$. Hence there are at most $\frac{S}{T}$ possible values for l and thus for $p(\mathbf{m})$. To deduce from this an estimate for the number of possible values for $\mathbf{m} \in M'$, we use the following definition and lemma.

Definition 4.17. For $n \in \mathbb{N}$, let $\tau(n)$ be the number of positive divisors of n. For any real $x \geq 1$, let

$$h(x) := \max_{\substack{n \in \mathbb{N} \\ n \leq x}} \tau(n).$$

Lemma 4.18. *We have*

$$h(x) \leq x^{\frac{1.07}{\log^+ \log^+ x}}.$$

Proof. For those x with $\log^+ \log^+ x = 1$ the result is obviously true. For larger x, the result follows from the bound

$$\tau(n) \leq 2^{1.538 \frac{\log n}{\log \log n}},$$

which is established in [13]. □

By Remark 4.16, $q(\mathbf{m}, \mathbf{a_1})$ is a divisor of $p(\mathbf{m})$. Thus, if the value of $p(\mathbf{m})$ is given, there are by (4.20) at most $h(S)$ possible values of $q(\mathbf{m}, \mathbf{a_1})$. We conclude that there are at most $\frac{S}{T} \cdot h(S)$ possible values of $q(\mathbf{m}, \mathbf{a_1})$ in total. Finally, we estimate the number of choices for $\mathbf{m} \in M'$ when $q(\mathbf{m}, \mathbf{a_1})$ has a fixed value, say $q(\mathbf{m}, \mathbf{a_1}) = t$. Now because of the definitions, we have $t \leq T$ and

$$t(\mathbf{m} - \mathbf{a_1}) \in L.$$

Letting $\varphi_t : \mathbb{Z}^n/L \to \mathbb{Z}^n/L$ be given by $x \mapsto t \cdot x$, this means that $\mathbf{m} - \mathbf{a_1} \in \ker(\varphi_t)$, and we estimate $\#\ker(\varphi_t)$ with the following lemma.

Lemma 4.19. *Let $\mathbb{Z}/k\mathbb{Z}$ be the cyclic group of order $k \in \mathbb{N}$ and let $\psi_t : \mathbb{Z}/k\mathbb{Z} \to \mathbb{Z}/k\mathbb{Z}$ be the homomorphism given by $\psi_t(x) = t \cdot x$. Then the kernel of ψ_t has at most t elements.*

Proof. Writing $t = t_1 t_2$ with $t_1 = \gcd(t, k)$ and $\gcd(t_2, k) = 1$, the congruence $tx \equiv 0 \mod k$ is equivalent to x being a multiple of $\frac{k}{t_1}$. The number of such $1 \leq x \leq k$ is at most $\frac{k}{k/t_1} = t_1 \leq t$. $\square$

Now since $\mathbb{Z}^n/L$ is a product of at most n cyclic groups it follows that $\#\ker(\varphi_t) \leq t^n \leq T^n$, so there are at most T^n possibilities for $\mathbf{m}$ given that $q(\mathbf{m}, \mathbf{a_1}) = t$. Putting everything together we arrive at

$$|M'| \leq \frac{S}{T} \cdot h(S) \cdot T^n \leq h(S) S^n, \tag{4.21}$$

as $T = q(\mathbf{a_1}, \mathbf{a_2})$ divides $p(\mathbf{m})$ for any $\mathbf{m} \in M'$, while we have the estimate (4.20).

Explicit estimates

Using (4.21) and that $h(S) \leq S \leq Q_1(r+1, D)$, we get

$$|M'| \leq S^{n+1} \leq S^{2n} = \left(\frac{1}{2} g_1(r, D)\right)^n \leq \frac{1}{2} \left(g_1(r, D)\right)^n. \tag{4.22}$$

By Lemma 4.18, we can bound $h(S)$ by

$$\begin{aligned} h(S) &\leq Q_2(r+1, D)^{\frac{1.07}{\log^+ \log^+ Q_2(r+1, D)}} \\ &\leq \left(80(r+1)^{3(r+1)^2} \log^+(r+1)\right)^{1.07} \left(D \log^+ \log^+(D)\right)^{\frac{1.07}{\log^+ \log^+ D}}. \end{aligned}$$

Since $r \geq 2$ we have $\log^+(r+1) \leq (r+1)^{0.01(r+1)^2}$. Writing $D_0 := D^{1/\log^+ \log^+ D}$ and using that

$$(\log^+ \log^+ D)^{\frac{1.07}{\log^+ \log^+ D}} \leq 3,$$

we arrive at

$$h(S) \leq 80^{1.07} (r+1)^{3.23(r+1)^2} \cdot 3 D_0^{1.07} \leq 327 (r+1)^{3.3(r+1)^2} D_0^{1.07} \tag{4.23}$$

Combining (4.23) with (4.21) and with the estimate

$$S \leq 80(r+1)^{3(r+1)^2} \log^+(r) D \log^+ \log^+(D) \leq 80(r+1)^{3.01(r+1)^2} D \log^+ \log^+(D),$$

we get

$$\begin{aligned} |M'| &\leq 327(r+1)^{3.23(r+1)^2} D_0^{1.07} \left(80(r+1)^{3.01(r+1)^2} D \log^+ \log^+(D)\right)^n \\ &\leq \frac{1}{2} \left(52320(r+1)^{6.3(r+1)^2} D D_0^3\right)^n, \end{aligned} \tag{4.24}$$

where the last inequality is obtained using

$$D_0^{1.07} \left(\log^+ \log^+(D)\right)^n \leq D_0^{3n}.$$

This inequality holds, since for $D \leq 15$ we have $\log^+ \log^+ D = 1$ and then it is obvious, and for $D \geq 16$ the inequality is equivalent to

$$n \log\log\log D \leq (3n - 1.07) \frac{\log(D)}{\log\log D}.$$

Combining (4.18), (4.19) (4.22) and (4.24) completes the induction step, and thus the proof of Theorem 4.2. $\square$

4.2 A specialisation argument

Let us recall the notation from §1. Let $\alpha_1, ..., \alpha_n \in \mathbb{C}$ be non-zero multiplicatively independent numbers, and let $f(\mathbf{x}) = f(x_1, ..., x_n) \in \mathbb{C}[x_1, ..., x_n]$ be a polynomial of total degree δ. We write $\alpha = (\alpha_1, ..., \alpha_n)$, and for $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{Z}^n$ we use the notation

$$\alpha^{\mathbf{x}} = \alpha_1^{x_1} \cdots \alpha_n^{x_n}.$$

We consider the Diophantine equation

$$\alpha^{\mathbf{x}} = f(\mathbf{x}) \text{ in } \boldsymbol{x} \in \mathbb{Z}^n, \tag{4.25}$$

Let K be the field $K := \mathbb{Q}(\alpha_1, ..., \alpha_n)$ and write $\Delta := \binom{n+\delta}{n}$ and $B := \Delta + 1$. Our aim is to show that (4.25) has at most $(8B)^{9B^6}$ solutions.

Via a so called specialisation argument, we reduce our proof to the case where the α_i and the coefficients of f lie in an algebraic number field. This slightly eases the arguments from Schmidt's paper [19]. Let U be the set containing $\alpha_1, ..., \alpha_n$, the non-zero coefficients of f, and the multiplicative inverses of these elements. The following lemma is a direct consequence of [21, Lemma 2].

Lemma 4.20. *There is a ring homomorphism $\varphi : \mathbb{Z}[U] \to \overline{\mathbb{Q}}$, such that $\varphi(\alpha_1), ..., \varphi(\alpha_n)$ are multiplicatively independent.*

Assume we have proven that (4.25) has at most $(8B)^{9B^6}$ solutions in the case that the α_i and the coefficients of f lie in a number field. Now consider (4.25) in the general case. Let $\varphi(f)$ denote the polynomial obtained by applying φ to the coefficients of f. Then each solution $\mathbf{x} \in \mathbb{Z}^n$ of (4.25) is also a solution to the equation

$$\varphi(\alpha_1)^{x_1} \cdots \varphi(\alpha_n)^{x_n} = \varphi(f)(x_1, ..., x_n). \tag{4.26}$$

Thus we can injectively map the set of solutions of (4.25) to the set of solutions of (4.26). But (4.26) is an equation of the type (4.25) with the numbers $\varphi(\alpha_i)$ and the coefficients of $\varphi(f)$ lying in an algebraic number field. So by assumption, (4.26) has at most $(8B)^{9B^6}$ solutions, so the same holds for (4.25).

4.3 Schmidt's paper

So far we have established the two results from the paper of Corvaja, Schmidt and Zannier [2], and we have seen that we may suppose without loss of generality that $\alpha_1, ..., \alpha_n$ and the coefficients of f lie in an algebraic number field. In particular, $K = \mathbb{Q}(\alpha_1, ..., \alpha_n)$ is an algebraic number field. The arguments in this section are based on Schmidt's earlier paper [19].

4.3.1 Some reductions

We first show that we may assume that all coefficients of f lie in K. Let W be the K-vector space generated by the coefficients of f. For each $\mathbf{x} \in \mathbb{Z}^n$, the left-hand side $\alpha^{\mathbf{x}}$ in (4.25) lies in K^*, while the right-hand side $f(\mathbf{x})$ lies in W. Thus, if $W \cap K = \{0\}$, then (4.25) has no solutions at all and we are done. So we may assume that $K \subset W$, and W has a basis of the form

$1, \lambda_1, ..., \lambda_t$ for some $\lambda_i \in \mathbb{C}$. Writing the coefficients of f as linear combinations in this basis and grouping terms we can write f as

$$f(\mathbf{x}) = f_0(\mathbf{x}) + \lambda_1 f_1(\mathbf{x}) + ... + \lambda_t f_t(\mathbf{x})$$

with $f_0(\mathbf{x}), ..., f_t(\mathbf{x}) \in K[\mathbf{x}]$. Now for any solution $\mathbf{x} \in \mathbb{Z}^n$ to (4.25), we have the following K-linear equality,

$$(f_0(\mathbf{x}) - \alpha^{\mathbf{x}}) \cdot 1 + f_1(\mathbf{x}) \cdot \lambda_1 + ... + f_t(\mathbf{x}) \cdot \lambda_t = 0,$$

and by linear independence it follows that $\mathbf{x}$ is a solution to $\alpha^{\mathbf{x}} = f_0(\mathbf{x})$. Thus, (4.25) is equivalent to an equation of the same form with $f(\mathbf{x}) \in K[\mathbf{x}]$.

Notation. Let S denote the set of all embeddings (i.e. injective homomorphisms) of K into $\mathbb{C}$. For $\xi \in K$ and $\sigma \in S$, write $\xi^{(\sigma)} := \sigma(\xi)$, and for a vector $\xi \in K^n$, write $\xi^{(\sigma)} = (\sigma(\xi_1), ..., \sigma(\xi_n))$. Further, let V be the $\mathbb{Q}$-vector space generated by the coefficients of f.

We may assume that K is not a number field of degree at most Δ, for otherwise, Theorem 1.1 already provides us with a convenient upper bound for the number of solutions to (4.25).

4.3.2 Reducing our main equation to a determinant equation

Since we have $\#\{(i_1, ..., i_n) \in (\mathbb{Z}_{\geq 0})^n : i_1 + ... + i_n \leq \delta\} = \Delta$, we see that f has at most Δ non-zero coefficients, hence V has a $\mathbb{Q}$-basis $b_1, ..., b_r$ with $r \leq \Delta$. Now let M be the matrix with rows

$$\left(b_1^{(\sigma)} \quad ... \quad b_r^{(\sigma)}\right), \sigma \in S,$$

with the rows in some arbitrary order.

Lemma 4.21. *The matrix M has rank r.*

Proof. We have to show that the vectors $\mathbf{b_i} := \left(b_i^{(\sigma)} : \sigma \in S\right)$, for $i = 1, ..., r$, are linearly independent over $\mathbb{C}$. Assume this is not the case, then these vectors span a $\mathbb{C}$-vector space of some dimension $m < r$. By symmetry we may suppose that $\mathbf{b_1}, ..., \mathbf{b_m}$ are linearly independent over $\mathbb{C}$. Now there are unique $c_1, ..., c_m \in \mathbb{C}$ such that

$$\mathbf{b_{m+1}} = c_1 \mathbf{b_1} + ... + c_m \mathbf{b_m}.$$

This means that $(c_1, ..., c_m)$ is the unique solutions in $\mathbb{C}^m$ to the system of equations

$$\sigma(b_{m+1}) = c_1 \sigma(b_1) + ... + c_m \sigma(b_m), \quad \text{for } \sigma \in S. \tag{4.27}$$

Now the compositum N of the fields $\sigma(K)$ ($\sigma \in S$) is a finite normal extension of K, so the elements of S are embeddings of K into N. Since all coefficients $\sigma(b_i)$ ($\sigma \in S, 1 \leq i \leq m+1$) lie in N, it follows by Cramer's rule that $c_1, ..., c_m \in N$. Applying $\tau \in \mathrm{Gal}(N/\mathbb{Q})$ to both sides of (4.27) yields

$$\tau(\sigma(b_{m+1})) = \tau(c_1)\tau(\sigma(b_1)) + ... + \tau(c_m)\tau(\sigma(b_m)), \quad \text{for } \sigma \in S, \tau \in \mathrm{Gal}(N/\mathbb{Q}).$$

When τ runs through $\mathrm{Gal}(N/\mathbb{Q})$ and σ runs through S, the composition $\tau \circ \sigma$ runs through S as well. Thus, for all $\tau \in \mathrm{Gal}(N/\mathbb{Q})$ and $\sigma \in S$ we have

$$\sigma(b_{m+1}) = \tau(c_1)\sigma(b_1) + ... + \tau(c_m)\sigma(b_m).$$

Hence, for each $\tau \in \mathrm{Gal}(N/\mathbb{Q})$, $(\tau(c_1), ..., \tau(c_m)) \in \mathbb{C}^m$ is a solution to the system (4.27). Since $(c_1, ..., c_m)$ is the unique solution in $\mathbb{C}^m$ to this system, we get for each $1 \leq i \leq m$ that $\tau(c_i) = c_i$ for all $\tau \in \mathrm{Gal}(N/\mathbb{Q})$, hence $c_i \in \mathbb{Q}$. Now (4.27) with σ the identity, gives a non-trivial $\mathbb{Q}$-linear relation between $b_1, ..., b_{m+1}$, contradicting that $b_1, ..., b_r$ are linearly independent over $\mathbb{Q}$. $\square$

Let $\mathbf{x} \in \mathbb{Z}^n$ be a solution of (4.25), then there are $\mu_1, ..., \mu_r \in \mathbb{Q}$ satisfying

$$\alpha^{\mathbf{x}} = \mu_1 b_1 + ... + \mu_r b_r.$$

Applying every $\sigma \in S$ on both sides we obtain

$$\left(\alpha^{(\sigma)}\right)^{\mathbf{x}} = \mu_1 b_1^{(\sigma)} + ... + \mu_r b_r^{(\sigma)} \text{ for all } \sigma \in S.$$

It follows that the vectors $\left(\left(\alpha^{(\sigma)}\right)^{\mathbf{x}} : \sigma \in S\right)$, $\left(b_1^{(\sigma)} : \sigma \in S\right)$,..., $\left(b_r^{(\sigma)} : \sigma \in S\right)$ satisfy a non-trivial linear relation over $\mathbb{C}$. Hence, for any $\sigma_0, ..., \sigma_r \in S$ we have a vanishing determinant

$$\begin{vmatrix} \left(\alpha^{(\sigma_0)}\right)^{\mathbf{x}} & b_1^{(\sigma_0)} & \cdots & b_r^{(\sigma_0)} \\ \vdots & \vdots & & \vdots \\ \left(\alpha^{(\sigma_r)}\right)^{\mathbf{x}} & b_1^{(\sigma_r)} & \cdots & b_r^{(\sigma_r)} \end{vmatrix} = 0. \tag{4.28}$$

Notation. For $\eta_1, ..., \eta_r \in S$, let $A^{(\eta_1,...,\eta_r)}$ be the determinant of the $(r \times r)$-matrix with $\left(b_i^{(\eta_1)} \quad ... \quad b_i^{(\eta_r)}\right)^T$ as its i-th column.

Expanding the determinant in (4.28) along the first column we can now write this equation as

$$\left(\alpha^{(\sigma_0)}\right)^{\mathbf{x}} A^{(\sigma_1,...,\sigma_r)} - \left(\alpha^{(\sigma_1)}\right)^{\mathbf{x}} A^{(\sigma_0,\sigma_2...,\sigma_r)} + ... + (-1)^r \left(\alpha^{(\sigma_r)}\right)^{\mathbf{x}} A^{(\sigma_0,\sigma_1...,\sigma_{r-1})} = 0. \tag{4.29}$$

By Lemma 4.21, M has a $(r \times r)$-submatrix with non-zero determinant. That is, for some $\tau_1, ..., \tau_r \in S$ we have $A^{(\tau_1,...,\tau_r)} \neq 0$. We fix such $\tau_1, ..., \tau_r$.

Notation. We write $A^{(1,...,r)} := A^{(\tau_1,...,\tau_r)}$, and for any $\sigma \in S$, we let $A^{(i+1,...,r,\sigma,1,...,i-1)} := A^{(\tau_{i+1},...,\tau_r,\sigma,\tau_1,...,\tau_{i-1})}$.
For $\xi \in K$, write $\xi^{(i)} := \xi^{(\tau_i)}$, and for $\xi \in K^n$, we write $\xi^{(i)} := \xi^{(\tau_i)}$.

Now in (4.29), we let $\sigma = \sigma_0$ run through S, while we fix $\sigma_1 = \tau_1$,...,$\sigma_r = \tau_r$, to obtain that for any $\sigma \in S$ we have

$$\left(\alpha^{(\sigma)}\right)^{\mathbf{x}} A^{(1,...,r)} - \left(\alpha^{(1)}\right)^{\mathbf{x}} A^{(\sigma,2...,r)} + ... + (-1)^r \left(\alpha^{(r)}\right)^{\mathbf{x}} A^{(\sigma,1,...,r-1)} = 0. \tag{4.30}$$

4.3.3 A case distinction

Notation. For $\sigma \in S$, and $1 \leq k \leq r$, we let $G(\sigma, k)$ be the subgroup of $\mathbb{Z}^n$ defined by

$$G(\sigma, k) := \left\{\mathbf{y} \in \mathbb{Z}^n : \left(\alpha^{(\sigma)}\right)^{\mathbf{y}} = \left(\alpha^{(k)}\right)^{\mathbf{y}}\right\}.$$

As $\mathbb{Z}^n$ is free of rank n, each group $G(\sigma, k)$ has rank at most n. We consider the assumption that

$$\textit{there is } \sigma \in S \textit{ such that for all } 1 \leq k \leq r \textit{ we have rank } G(\sigma, k) < n. \tag{4.31}$$

In our proof of Theorem 1.4 we make a distinction between the cases where (4.31) does or does not hold.

4.4 Application of Theorem 4.2 to equation (4.25)

In this subsection we prove Theorem 1.4 in the case where (4.31) does not hold. The arguments in this subsection are based on the paper of Corvaja, Schmidt and Zannier [2].

Let G be the multiplicative subgroup of $\overline{\mathbb{Q}}^*$ generated by $\alpha_1, ..., \alpha_n$, and for $m \in \mathbb{N}$, let $G^{[m]} := \{\alpha^m : \alpha \in G\}$.

Lemma 4.22. *Assume that* (4.31) *does not hold. Then there there are a number field L of degree at most r, and a number $m \in \mathbb{N}$, such that L is contained in K and $G^{[m]}$ is contained in L.*

Proof. Let d be the degree of K over $\mathbb{Q}$. Denote the elements of S by $\xi \mapsto \xi^{(i)}$, for $i = 1, ..., d$, where $\xi^{(i)} = \xi^{(\tau_i)}$ for $i = 1, ..., r$. For $i = 1, ..., d$, write $G(i, k) = G(\sigma, k)$ if σ is the embedding of K given by $\sigma : \xi \mapsto \xi^{(i)}$. By assumption, for each $1 \leq i \leq d$, there is $k \in \{1, ..., r\}$ such that $G(i, k)$ has rank n. By the pigeonhole principle, there are $1 \leq k \leq r$ and a subset $I \subset \{1, ..., d\}$ with $|I| \geq \frac{d}{r}$, such that for all $i \in I$ we have rank $G(i, k) = n$. But then the intersection $\bigcap_{i \in I} G(i, k)$ still has rank n, hence there is an $m \in \mathbb{N}$ such that for all $\mathbf{v} \in \mathbb{Z}^n$ we have $m\mathbf{v} \in \bigcap_{i \in I} G(i, k)$. Thus, for all $i, j \in I$ and all $\mathbf{v} \in \mathbb{Z}^n$ we have

$$((\alpha^{\mathbf{v}})^m)^{(i)} = ((\alpha^{\mathbf{v}})^m)^{(k)} = ((\alpha^{\mathbf{v}})^m)^{(j)}$$

Letting L be the number field generated over $\mathbb{Q}$ by $G^{[m]}$, this means that all embeddings $\xi \mapsto \xi^{(i)}, i \in I$, act the same on L. Thus, some embedding of L has at least $|I| \geq \frac{d}{r}$ extensions to K, hence $[K : L] \geq \frac{d}{r}$. This yields

$$[L : \mathbb{Q}] = \frac{[K : \mathbb{Q}]}{[K : L]} \leq \frac{d}{d/r} = r,$$

while the required property $G^{[m]} \subset L$ is clear from the definition of L. □

The following lemma establishes a stronger upper bound than the one in Theorem 1.4 in the case that (4.31) not hold.

Lemma 4.23. *If* (4.31) *does not hold, then equation* (4.25) *has at most 2^{54B^4} solutions.*

Proof. Let L and $m \in \mathbb{N}$ be as in Lemma 4.22, and let $H := G \cap L^*$. Since $G^{[m]} \subset L^*$, we see that H has finite index in G. Let W denote the L-vector space spanned by the coefficients of f. Since f has at most Δ non-zero coefficients, W has dimension at most $\Delta = B - 1$, and L has degree at most $r \leq B$. Thus, by Theorem 4.2, W has non-empty intersection with at most $g_1(B - 1, B)^n$ cosets of H. Now let $\mathbf{x} \in \mathbb{Z}^n$ be a solution to (4.25). Then we have $\alpha^{\mathbf{x}} \in W$, so $\alpha^{\mathbf{x}}$ lies in a union of at most $g_1(B - 1, B)^n$ cosets of H. Since $B \geq 3$, we have the estimate

$$g_1(B - 1, B) = 2 \cdot B^{12B^2} B^4 \leq 2^{13B^3}.$$

Now let us estimate the number of possibilities for $\mathbf{x}$ given that $\alpha^{\mathbf{x}}$ lies in a fixed coset of H, say $\alpha^{\mathbf{x}} \in \alpha^{\mathbf{z}} H$ for some $\mathbf{z} \in \mathbb{Z}^n$. Since H is of finite index in the free group G of rank n, H is also free of rank n. Hence there are $\mathbf{h_1}, ..., \mathbf{h_n} \in \mathbb{Z}^n$ such that $\beta_1 := \alpha^{\mathbf{h_1}}, ..., \beta_n := \alpha^{\mathbf{h_n}}$ form a basis of H. Therefore, there are $y_1, ..., y_n \in \mathbb{Z}$ such that

$$\mathbf{x} = \mathbf{z} + y_1\mathbf{h_1} + ... + y_n\mathbf{h_n}. \tag{4.32}$$

Since $\mathbf{x}$ is a solution to (4.25), it follows that

$$\beta_1^{y_1} \cdots \beta_n^{y_n} = \alpha^{-\mathbf{z}} f(\mathbf{z} + y_1\mathbf{h_1} + ... + y_n\mathbf{h_n}) = g(y_1, ..., y_n), \tag{4.33}$$

where $g \in K[x_1, ..., x_n]$ is a polynomial that only depends on the coset $\alpha^{\mathbf{z}}H$. Now (4.33) is an equation of the form (1.1) in the vector $\mathbf{y} = (y_1, ..., y_n)$, but with $\beta_1, ..., \beta_n \in L$, while L is a number field of degree at most $r \leq B$. Hence by Theorem 1.1, there are at most $B^{6B^2}2^{35B^3}$ possibilities for $\mathbf{y}$ and thus for $\mathbf{x}$, by (4.32). We claim that

$$B^{6B^2}2^{35B^3} \leq 2^{41B^3}.$$

Indeed, for $B = 3$ the inequality holds, while the right-hand side grows faster in B than the left-hand side, which can be seen by taking logarithms of both sides. Multiplying this amount with the number of classes in which $\alpha^{\mathbf{x}}$ can lie, we conclude that the number of solutions to (4.25) is at most

$$2^{13nB^3} \cdot 2^{41B^3} \leq 2^{54B^4}.$$

□

4.5 Proof of Theorem 1.4 by induction

The arguments in this subsection are based on Schmidt's paper [19]. Since $n \leq B$, the following lemma implies Theorem 1.4.

Lemma 4.24. *Equation* (4.25) *has at most* $(8B)^{9nB^5}$ *solutions.*

Remark 4.25. This upper bound is improved in comparison to the one derived in the paper of Corvaja, Schmidt and Zannier [2], due to the fact that we apply Theorem 3.6 instead of Theorem 3.5 in the proof.

Proof. The proof is by induction on n. By Lemma 4.23, we may assume in both the base case and the induction step that (4.31) holds. Indeed, we have $2^{54B^4} \leq (8B)^{9B^5}$, since the inequality holds for $B = 3$ and the right-hand side grows faster in B than the left-hand side. Throughout the proof, we fix σ as in (4.31). Let $\mathbf{x} \in \mathbb{Z}^n$ be a solution to (4.25), then $\mathbf{x}$ is a solution to (4.30) for our fixed σ. Since $A^{(1,...,r)} \neq 0$, we can find a non-empty subset $\mathcal{K} \subset \{1, ..., r\}$, such that

$$A^{(1,...,r)} \left(\alpha^{(\sigma)}\right)^{\mathbf{x}} + \sum_{k \in \mathcal{K}} (-1)^k A^{(1,...,k-1,\sigma,k+1,...,r)} \left(\alpha^{(k)}\right)^{\mathbf{x}} = 0, \tag{4.34}$$

and such that the above does not hold for any strict subset of $\mathcal{K}$. The number of possibilities for $\mathcal{K}$ is at most 2^r. For given $\mathcal{K}$, we can interpret the vector $w(\mathbf{x}) := \left(\left(\alpha^{(\sigma)}\right)^{\mathbf{x}}, \left(\alpha^{(k)}\right)^{\mathbf{x}} : k \in \mathcal{K}\right)$ as a non-degenerate solution to (4.34). Since $1 + |\mathcal{K}| \leq 1 + r \leq 1 + \Delta = B$, (4.34) is an equation in at most B variables, and our solutions $w(\mathbf{x})$ lie in the subgroup $\langle (w(\mathbf{e_1}), ..., w(\mathbf{e_n}) \rangle$ of $(\mathbb{C}^*)^n$, which has rank at most n. Thus, by Lemma 3.2 and by Theorem 3.6, we see that for given $\mathcal{K}$ the solutions $w(\mathbf{x})$ fall into at most $(8B)^{4B^4(B+n)}$ proportionality classes. Multiplying this with the number of possibilities for $\mathcal{K}$, we see that the solutions $\mathbf{x}$ of (4.25) are distributed over at most

$$(8B)^{4B^4(B+n)} \cdot 2^r \leq (8B)^{8B^5} \cdot 2^B \leq (8B)^{9B^5}$$

proportionality classes. The next step is to investigate how many solutions each class contains. Let $\mathbf{x}, \mathbf{x}' \in \mathbb{Z}^n$ be solutions to (4.25) such that $w(\mathbf{x})$ and $w(\mathbf{x}')$ are proportional, then for any $k \in \mathcal{K}$ we have

$$\left(\alpha^{(\sigma)}\right)^{\mathbf{x}-\mathbf{x}'} = \left(\alpha^{(k)}\right)^{\mathbf{x}-\mathbf{x}'}.$$

By definition this means that $\mathbf{x} - \mathbf{x}' \in G(\sigma, k)$. Letting $s := \operatorname{rank} G(\sigma, k)$, we get in the case $s = 0$ that $\mathbf{x} = \mathbf{x}'$, hence the number of solutions to (4.25) is at most $(8B)^{9B^5}$. In particular, for $n = 1$ we have $s = 0$ by (4.31), so this concludes the base case. Now assume that $n \geq 2$ and suppose that we have proved the lemma for all $n' < n$. We can assume that $0 < s < n$, and $G(\sigma, k)$ has some basis $\mathbf{v_1}, ..., \mathbf{v_s}$. There are unique $a_1, ..., a_s \in \mathbb{Z}$ such that

$$\mathbf{x} - \mathbf{x}' = a_1\mathbf{v_1} + ... + a_s\mathbf{v_s}, \tag{4.35}$$

hence for $\beta_i := \alpha_{\mathbf{i}}^{\mathbf{v_i}} \in K$ we have by (4.25),

$$\beta_1^{a_1} \cdots \beta_s^{a_s} = \alpha^{-\mathbf{x}'} f(\mathbf{x}). \tag{4.36}$$

Notice that $g(a_1, ..., a_s) := \alpha^{-\mathbf{x}'} f(\mathbf{x})$ is a polynomial in $s < n$ variables, depending on $\mathbf{x}', \mathbf{v_1}, ..., \mathbf{v_s}$, so by the induction hypothesis applied to (4.36), there are at most $(8B)^{9(n-1)B^5}$ possibilities for $(a_1, ..., a_s)$. Thus, by (4.35) this is an upper bound for the number of solutions in each class. Multiplying with the number of classes, and using $n \leq B$, we find that the number of solutions to (4.25) is bounded above by

$$(8B)^{9B^5} \cdot (8B)^{9(n-1)B^5} \leq (8B)^{9nB^5},$$

which concludes the induction step. The lemma now follows by induction. □

Concluding remarks

Let us compare the proofs of Theorems 1.2, 1.3 and 1.4. In the case that (4.31) holds, the papers of Schmidt [19], and of Corvaja, Schmidt and Zannier [2] use the same arguments. The improvement in the latter paper comes from the case that (4.31) does not hold, due to the application of Theorem 4.2 instead of other arguments of Schmidt that we did not discuss in this thesis. In the case that (4.31) does not hold, we obtained the same upper bound as Corvaja, Schmidt and Zannier, namely 2^{54B^4}.

In our proof of Lemma 4.24, we followed the same arguments as in the two papers for the case that (4.31) holds. However, we applied Theorem 3.6 instead of Theorem 3.5 in our induction, and this is the reason we obtain a sharper upper bound than Corvaja, Schmidt and Zannier. Finally we remark that the proofs of Theorem 1.2, Theorem 1.3 and Theorem 1.4 all make use of Theorem 1.1 in the case that (4.31) does not hold, by reducing equation (1.1) to an equation of the same form, but with the α_i lying in the number field L.

4.6 Proof of Corollary 1.5

We show that the equation

$$\alpha_1^{x_1} \cdots \alpha_n^{x_n} = f(x_1, ..., x_n) \text{ in } \mathbf{x} = (x_1, ..., x_n) \in \mathbb{Q}^n \tag{4.37}$$

also has at most $(8B)^{9B^6}$ solutions, if we fix values for $\log \alpha_1, ..., \log \alpha_n$. For $N \in \mathbb{N}$, let $S(N)$ be the set of solutions $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{Q}^n$ to (4.37), with the property that $N!x_i \in \mathbb{Z}$ for each $1 \leq i \leq n$. Now fix $N \in \mathbb{N}$. For $1 \leq i \leq n$, define $\beta_i := \alpha_i^{\frac{1}{N!}} = \exp\left(\frac{1}{N!} \log \alpha_i\right)$, and note that $\beta_1, ..., \beta_n$ are multiplicatively independent. Further define the polynomial $g(x_1, ..., x_n) := f\left(\frac{x_1}{N!}, ..., \frac{x_n}{N!}\right)$ and note that f and g have the same total degree. Now for any $\mathbf{x} = (x_1, ..., x_n) \in S(N)$, note that $\mathbf{y} = (y_1, ..., y_n)$, given by $y_i := N!x_i$, is an integer solution to the equation

$$\beta_1^{z_1} \cdots \beta_n^{z_n} = g(z_1, ..., z_n) \text{ in } (z_1, ..., z_n) \in \mathbb{Z}^n.$$

Now it follows by Theorem 1.4 that $|S(N)| \leq (8B)^{9B^6}$. Since we have a chain of inclusions $S(1) \subset S(2) \subset S(3) \subset ...$, and the cardinalities of the sets $S(N)$ are uniformly bounded from above, there is $N_0 \in \mathbb{N}$ such that $S(N) = S(N_0)$ for each $N \geq N_0$. Hence $S(N_0)$ is the complete set of rational solutions to (4.37), so (4.37) has at most $(8B)^{9B^6}$ solutions.

References

[1] F. Amoroso and E. Viada, "Small points on subvarieties of a torus," *Duke Math. J.*, vol. 150, no. 3, pp. 407–442, 2009.

[2] P. Corvaja, W. M. Schmidt, and U. Zannier, "The Diophantine equation $\alpha_1^{x_1} \cdots \alpha_n^{x_n} = f(x_1, \ldots, x_n)$. II," *Trans. Amer. Math. Soc.*, vol. 362, no. 4, pp. 2115–2123, 2010.

[3] J.-H. Evertse, H. P. Schlickewei, and W. M. Schmidt, "Linear equations in variables which lie in a multiplicative group," *Ann. of Math. (2)*, vol. 155, no. 3, pp. 807–836, 2002.

[4] J.-H. Evertse, "The number of solutions of linear equations in roots of unity," *Acta Arith.*, vol. 89, no. 1, pp. 45–51, 1999.

[5] J.-H. Evertse and K. Győry, *Unit equations in Diophantine number theory*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2015, vol. 146.

[6] M. Laurent, "Équations diophantiennes exponentielles," *Invent. Math.*, vol. 78, no. 2, pp. 299–327, 1984.

[7] ——, "Équations exponentielles polynômes et suites récurrentes linéaires," *Astérisque*, no. 147-148, pp. 121–139, 343–344, 1985, journées arithmétiques de Besançon.

[8] ——, "Équations exponentielles-polynômes et suites récurrentes linéaires. II," *J. Number Theory*, vol. 31, no. 1, pp. 24–53, 1989.

[9] C. Lech, "A note on recurring series," *Ark. Mat.*, vol. 2, pp. 417–421, 1953.

[10] K. Mahler, "Eine arithmetische Eigenschaft der Taylor-Koeffizienten rationaler Funktionen," *Proc. Kon. Ned. Akad. Wetensch.*, vol. 38, pp. 50–60–61, 1935.

[11] D. Masser, Y. V. Nesterenko, H. P. Schlickewei, W. M. Schmidt, and M. Waldschmidt, *Diophantine approximation*, ser. Lecture Notes in Mathematics. Springer-Verlag, Berlin; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 2003, vol. 1819, lectures from the C.I.M.E. Summer School held in Cetraro, June 28–July 6, 2000, Edited by F. Amoroso and U. Zannier, Fondazione CIME/CIME Foundation Subseries.

[12] J. Neukirch, *Algebraic number theory*, ser. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, vol. 322, translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[13] J.-L. Nicolas and G. Robin, "Majorations explicites pour le nombre de diviseurs de N," *Canad. Math. Bull.*, vol. 26, no. 4, pp. 485–492, 1983.

[14] J. B. Rosser and L. Schoenfeld, "Approximate formulas for some functions of prime numbers," *Illinois J. Math.*, vol. 6, pp. 64–94, 1962.

[15] H. P. Schlickewei and W. M. Schmidt, "The number of solutions of polynomial-exponential equations," *Compositio Math.*, vol. 120, no. 2, pp. 193–225, 2000.

[16] H. P. Schlickewei, "The $\mathfrak{p}$-adic Thue-Siegel-Roth-Schmidt theorem," *Arch. Math.*, vol. 29, no. 3, pp. 267–270, 1977.

[17] W. M. Schmidt, "Simultaneous approximation to algebraic numbers by elements of a number field," *Monatsh. Math.*, vol. 79, pp. 55–66, 1975.

[18] ——, "The zero multiplicity of linear recurrence sequences," *Acta Math.*, vol. 182, no. 2, pp. 243–282, 1999.

[19] ——, "The Diophantine equation $\alpha_1^{x_1} \cdots \alpha_n^{x_n} = f(x_1, \ldots, x_n)$," in *Analytic number theory. Essays in honour of Klaus Roth.* Cambridge Univ. Press, Cambridge, 2009, pp. 414–420.

[20] T. Skolem, "Ein Verfahren zur Behandlung gewisser exponentialer Gleichungen," *Oslo Vid. akad. Skrifter*, vol. 6, pp. 1–61, 1933.

[21] A. J. van der Poorten and H. P. Schlickewei, "Additive relations in fields," *J. Austral. Math. Soc. Ser. A*, vol. 51, no. 1, pp. 154–170, 1991.