EFFECTIVE RESULTS FOR UNIT EQUATIONS OVER FINITELY GENERATED DOMAINS

JAN-HENDRIK EVERTSE AND KÁLMÁN GYŐRY

ABSTRACT. Let $A \supset \mathbb{Z}$ be a commutative domain which is finitely generated over $\mathbb{Z}$ as a $\mathbb{Z}$-algebra and let a, b, c be non-zero elements of A. Extending earlier work of Siegel [25, 1921], Mahler [18, 1933] and Parry [20, 1950], Lang [13, 1960] proved that the equation (*) $a\varepsilon + b\eta = c$ in $\varepsilon, \eta \in A^*$ has only finitely many solutions. Using Baker's theory of logarithmic forms, Győry [6, 1974], [7, 1979] proved that the solutions of (*) can be determined effectively if A is contained in an algebraic number field. In this paper we prove, in a precise quantitative form, an effective finiteness result for equations (*) over an arbitrary domain A of characteristic 0 which is finitely generated over $\mathbb{Z}$. Our main tools are already existing effective finiteness results for (*) over number fields and function fields, an effective specialization argument of Győry [8, 1983], [9, 1984], and effective results of Seidenberg [24, 1974] and Aschenbrenner [1, 2004] on linear equations over polynomial rings. We prove also an effective result for the exponential equation $a\gamma_1^{v_1} \cdots \gamma_s^{v_s} + b\gamma_1^{w_1} \cdots \gamma_s^{w_s} = c$ in integers $v_1, \ldots, w_s$, where a, b, c and $\gamma_1, \ldots, \gamma_s$ are non-zero elements of A.

1. INTRODUCTION

Let $A \supset \mathbb{Z}$ be a commutative domain which is finitely generated over $\mathbb{Z}$ as a $\mathbb{Z}$-algebra. As usual, we denote by A^* the unit group of A. We consider equations

$$a\varepsilon + b\eta = c \quad \text{in } \varepsilon, \eta \in A^* \tag{1.1}$$

where a, b, c are non-zero elements of A. Such equations, usually called *unit equations*, have a great number of applications. For instance, the ring

2010 Mathematics Subject Classification: Primary 11D61; Secondary: 11J86.

Keywords and Phrases: Unit equations, finitely generated domains, effective finiteness results.

July 29, 2011.

of S-integers in an algebraic number field is finitely generated over $\mathbb{Z}$, so the S-unit equation in two unknowns is a special case of (1.1). In this paper, we consider equations (1.1) in the general case, where A may contain transcendental elements, too.

Siegel [25, 1921] proved that (1.1) has only finitely many solutions in the case that A is the ring of integers of a number field, and Mahler [18, 1933] did this in the case that $A = \mathbb{Z}[1/p_1 \cdots p_t]$ for certain primes $p_1, \ldots, p_t$. For S-unit equations over number fields, the finiteness of the number of solutions of (1.1) follows from work of Parry [20, 1950]. Finally, Lang [13, 1960] proved for arbitrary finitely generated domains A that (1.1) has only finitely many solutions. The proofs of all these results are ineffective.

Baker [2, 1968] and Coates [5, 1968/69] implicitly proved effective finiteness results for certain special (S-)unit equations. Later, Győry [6, 1974], [7, 1979] showed, in the case that A is the ring of S-integers in a number field, that the solutions of (1.1) can be determined effectively in principle. His proof is based on estimates for linear forms in ordinary and p-adic logarithms of algebraic numbers. In his papers [8, 1983] and [9, 1984], Győry introduced an effective specialization argument, and he used this to establish effective finiteness results for decomposable form equations and discriminant equations over a wide class of finitely generated domains A containing both algebraic and transcendental elements, of which the elements have some "good" effective representations. His results contain as a special case an effective finiteness result for equations (1.1) over these domains. Győry's method of proof could not be extended to arbitrary finitely generated domains A.

It is the purpose of this paper to prove an effective finiteness result for (1.1) over arbitrary finitely generated domains A. In fact, we give a quantitative statement, with effective upper bounds for the "sizes" of the solutions ε, η. The main new ingredient of our proof is an effective result by Aschenbrenner [1, 2004] on systems of linear equations over polynomial rings over $\mathbb{Z}$.

We introduce the notation used in our theorems. Let again $A \supset \mathbb{Z}$ be a commutative domain which is finitely generated over $\mathbb{Z}$, say $A = \mathbb{Z}[z_1, \ldots, z_r]$. Let I be the ideal of polynomials $f \in \mathbb{Z}[X_1, \ldots, X_r]$ such

that $f(z_1,\dots,z_r)=0$. Then I is finitely generated, hence

$$A \cong \mathbb{Z}[X_1,\dots,X_r]/I, \quad I=(f_1,\dots,f_m) \tag{1.2}$$

for some finite set of polynomials $f_1,\dots,f_m \in \mathbb{Z}[X_1,\dots,X_r]$. We observe here that given $f_1,\dots,f_m$, it can be checked effectively whether A is a domain containing $\mathbb{Z}$. Indeed, this holds if and only if I is a prime ideal of $\mathbb{Z}[X_1,\dots,X_r]$ with $I\cap\mathbb{Z}=(0)$, and the latter can be checked effectively for instance using Aschenbrenner [1, Prop. 4.10, Cor. 3.5].

Denote by K the quotient field of A. For $\alpha\in A$, we call f a *representative* for α, or say that f represents α if $f\in\mathbb{Z}[X_1,\dots,X_r]$ and $\alpha=f(z_1,\dots,z_r)$. Further, for $\alpha\in K$, we call (f,g) a *pair of representatives* for α or say that (f,g) represents α if $f,g\in\mathbb{Z}[X_1,\dots,X_r]$, $g\notin I$ and $\alpha=f(z_1,\dots,z_r)/g(z_1,\dots,z_r)$. We say that $\alpha\in A$ (resp. $\alpha\in K$) is given if a representative (resp. pair of representatives) for α is given.

To do explicit computations in A and K, one needs an *ideal membership algorithm* for $\mathbb{Z}[X_1,\dots,X_r]$, that is an algorithm which decides for any given polynomial and ideal of $\mathbb{Z}[X_1,\dots,X_r]$ whether the polynomial belongs to the ideal. In the literature there are various such algorithms; we mention only the algorithm of Simmons [26, 1970], and the more precise algorithm of Aschenbrenner [1, 2004] which plays an important role in our paper; see Lemma 2.5 below for a statement of his result. One can perform arithmetic operations on A and K by using representatives. Further, one can decide effectively whether two polynomials f_1,f_2 represent the same element of A, i.e., $f_1-f_2\in I$, or whether two pairs of polynomials $(f_1,g_1),(f_2,g_2)$ represent the same element of K, i.e., $f_1g_2-f_2g_1\in I$, by using one of the ideal membership algorithms mentioned above.

The *degree* $\deg f$ of a polynomial $f\in\mathbb{Z}[X_1,\dots,X_r]$ is by definition its total degree. By the *logarithmic height* $h(f)$ of f we mean the logarithm of the maximum of the absolute values of its coefficients. The *size* of f is defined by $s(f):=\max(1,\deg f,h(f))$. Clearly, there are only finitely many polynomials in $\mathbb{Z}[X_1,\dots,X_r]$ of size below a given bound, and these can be determined effectively.

Theorem 1.1. *Assume that $r\geq 1$. Let $\widetilde{a},\widetilde{b},\widetilde{c}$ be representatives for a,b,c, respectively. Assume that $f_1,\dots,f_m$ and $\widetilde{a},\widetilde{b},\widetilde{c}$ all have degree at most d and logarithmic height at most h, where $d\geq 1$, $h\geq 1$. Then for each solution (ε,η) of (1.1), there are representatives $\widetilde{\varepsilon},\widetilde{\varepsilon}',\widetilde{\eta},\widetilde{\eta}'$ of $\varepsilon,\varepsilon^{-1},\eta,\eta^{-1}$,*

respectively, such that

$$s(\widetilde{\varepsilon}),\ s(\widetilde{\varepsilon}'),\ s(\widetilde{\eta}),\ s(\widetilde{\eta}') \leq \exp\Big((2d)^{c_1^r}(h+1)\Big),$$

where c_1 is an effectively computable absolute constant > 1.

By a theorem of Roquette [22, 1958], the unit group of a domain finitely generated over $\mathbb{Z}$ is finitely generated. In the case that $A = O_S$ is the ring of S-integers of a number field it is possible to determine effectively a system of generators for A^*, and this was used by Győry in his effective finiteness proof for (1.1) with $A = O_S$. However, no general algorithm is known to determine a system of generators for the unit group of an arbitrary finitely generated domain A. In our proof of Theorem 1.1, we do not need any information on the generators of A^*.

By combining Theorem 1.1 with an ideal membership algorithm for $\mathbb{Z}[X_1, \ldots, X_r]$, one easily deduces the following:

Corollary 1.2. *Given $f_1, \ldots, f_m, a, b, c$, the solutions of* (1.1) *can be determined effectively.*

Proof. Clearly, ε, η is a solution of (1.1) if and only if there are polynomials $\widetilde{\varepsilon}, \widetilde{\varepsilon}', \widetilde{\eta}, \widetilde{\eta}' \in \mathbb{Z}[X_1, \ldots, X_r]$ such that $\widetilde{\varepsilon}, \widetilde{\eta}$ represent ε, η, and

$$\widetilde{a} \cdot \widetilde{\varepsilon} + \widetilde{b} \cdot \widetilde{\eta} - \widetilde{c},\ \ \widetilde{\varepsilon} \cdot \widetilde{\varepsilon}' - 1,\ \ \widetilde{\eta} \cdot \widetilde{\eta}' - 1 \in I. \tag{1.3}$$

Thus, we obtain all solutions of (1.1) by checking, for each quadruple of polynomials $\widetilde{\varepsilon}, \widetilde{\varepsilon}', \widetilde{\eta}, \widetilde{\eta}' \in \mathbb{Z}[X_1, \ldots, X_r]$ of size at most $\exp\Big((2d)^{c_1^r}(h+1)\Big)$ whether it satisfies (1.3). Further, using the ideal membership algorithm, it can be checked effectively whether two different pairs $(\widetilde{\varepsilon}, \widetilde{\eta})$ represent the same solution of (1.1). Thus, we can make a list of representatives, one for each solution of (1.1). □

Let $\gamma_1, \ldots, \gamma_s$ be multiplicatively independent elements of K^* (the multiplicative independence of $\gamma_1, \ldots, \gamma_s$ can be checked effectively for instance using Lemma 7.2 below). Let again a, b, c be non-zero elements of A and consider the equation

$$a\gamma_1^{v_1} \cdots \gamma_s^{v_s} + b\gamma_1^{w_1} \cdots \gamma_s^{w_s} = c \ \text{ in } v_1, \ldots, v_s,\ w_1, \ldots, w_s \in \mathbb{Z}. \tag{1.4}$$

Theorem 1.3. *Let $\widetilde{a}, \widetilde{b}, \widetilde{c}$ be representatives for a, b, c and for $i = 1, \ldots, s$, let (g_{i1}, g_{i2}) be a pair of representatives for γ_i. Suppose that $f_1, \ldots, f_m$,*

$\widetilde{a}, \widetilde{b}, \widetilde{c}$, and g_{i1}, g_{i2} $(i = 1, \dots, s)$ all have degree at most d and logarithmic height at most h, where $d \geq 1$, $h \geq 1$. Then for each solution $(v_1, \dots, w_s)$ of (1.4) we have

$$\max\big(|v_1|, \dots, |v_s|, |w_1|, \dots, |w_s|\big) \leq \exp\Big((2d)^{c_2^{r+s}}(h+1)\Big),$$

where c_2 is an effectively computable absolute constant > 1.

An immediate consequence of Theorem 1.3 is that for given $f_1, \dots, f_m$, a, b, c and $\gamma_1, \dots, \gamma_s$, the solutions of (1.4) can be determined effectively.

Since every domain finitely generated over $\mathbb{Z}$ has a finitely generated unit group, equation (1.1) maybe viewed as a special case of (1.4). But since no general effective algorithm is known to find a finite system of generators for the unit group of a finitely generated domain, we cannot deduce an effective result for (1.1) from Theorem 1.3. In fact, we argue reversely, and prove Theorem 1.3 by combining Theorem 1.1 with an effective result on Diophantine equations of the type $\gamma_1^{v_1} \cdots \gamma_s^{v_s} = \gamma_0$ in integers $v_1, \dots, v_s$, where $\gamma_1, \dots, \gamma_s, \gamma_0 \in K^*$ (see Corollary 7.3 below).

The idea of the proof of Theorem 1.1 is roughly as follows. We first estimate the degrees of the representatives of ε, η using Mason's effective result [19, 1983] on two term S-unit equations over function fields. Next, we apply many different specialization maps $A \to \overline{\mathbb{Q}}$ to (1.1) and obtain in this manner a large system of S-unit equations over number fields. By applying an existing effective finiteness result for such S-unit equations (e.g., Győry and Yu [10, 2006]) we collect enough information to retrieve an effective upper bound for the heights of the representatives of ε, η. In our proof, we apply the specialization maps on a domain $B \supset A$ of a special type which can be dealt with more easily. In the construction of B, we use an effective result of Seidenberg [24, 1974] on systems of linear equations over polynomial rings over arbitrary fields. To be able to go back to equation (1.1) over A, we need an effective procedure to decide whether a given element of B belongs to A^*. For this decision procedure, we apply an effective result of Aschenbrenner [1, 2004] on systems of linear equations over polynomial rings over $\mathbb{Z}$.

The above approach was already followed by Győry [8, 1983], [9, 1984]. However, in these papers the domains A are represented over $\mathbb{Z}$ in a different way. Hence, to select those solutions from B of the equations under

consideration which belong to A, certain restrictions on the domains A had to be imposed.

In a forthcoming paper, we will give some applications of our above theorems and our method of proof to other classes of Diophantine equations over finitely generated domains.

2. Effective linear algebra over polynomial rings

We have collected some effective results for systems of linear equations to be solved in polynomials with coefficients in a field, or with coefficients in $\mathbb{Z}$.

Here and in the remainder of this paper, we write

$$\log^* x := \max(1, \log x) \text{ for } x > 0, \ \log^* 0 := 1.$$

We use notation $O(\cdot)$ as an abbreviation for $c\times$ the expression between the parentheses, where c is an effectively computable absolute constant. At each occurrence of $O(\cdot)$, the value of c may be different.

Given a commutative domain R, we denote by $R^{m,n}$ the R-module of $m \times n$-matrices with entries in R and by R^n the R-module of n-dimensional column vectors with entries in R. Further, $\mathrm{GL}_n(R)$ denotes the group of matrices in $R^{n,n}$ with determinant in the unit group R^*. The degree of a polynomial $f \in R[X_1, \ldots, X_N]$, that is, its total degree, is denoted by $\deg f$.

From matrices A, B with the same number of rows, we form a matrix $[A, B]$ by placing the columns of B after those of A. Likewise, from two matrices A, B with the same number of columns we form $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$ by placing the rows of B below those of A.

The logarithmic height $h(S)$ of a finite set $S = \{a_1, \ldots, a_t\} \subset \mathbb{Z}$ is defined by $h(S) := \log \max(|a_1|, \ldots, |a_t|)$. The logarithmic height $h(U)$ of a matrix with entries in $\mathbb{Z}$ is defined by the logarithmic height of the set of entries of U. The logarithmic height $h(f)$ of a polynomial with coefficients in $\mathbb{Z}$ is the logarithmic height of the set of coefficients of f.

Lemma 2.1. *Let $U \in \mathbb{Z}^{m,n}$. Then the $\mathbb{Q}$-vector space of $\mathbf{y} \in \mathbb{Q}^n$ with $U\mathbf{y} = \mathbf{0}$ is generated by vectors in $\mathbb{Z}^n$ of logarithmic height at most $mh(U) + \frac{1}{2}m \log m$.*

Proof. Without loss of generality we may assume that U has rank m, and moreover, that the matrix B consisting of the first m columns of U is invertible. Let $\Delta := \det B$. By multiplying with ΔB^{-1}, we can rewrite $U\mathbf{y} = \mathbf{0}$ as $[\Delta I_m, C]\mathbf{y} = \mathbf{0}$, where I_m is the $m \times m$-unit matrix, and C consists of $m \times m$-subdeterminants of U. The solution space of this system is generated by the columns of $\left[\begin{smallmatrix} -C \\ \Delta I_{n-m} \end{smallmatrix}\right]$. An application of Hadamard's inequality gives the upper bound from the lemma for the logarithmic heights of these columns. □

Proposition 2.2. *Let F be a field, $N \geq 1$, and $R := F[X_1, \ldots, X_N]$. Further, let A be an $m \times n$-matrix and $\mathbf{b}$ and m-dimensional column vector, both consisting of polynomials from R of degree $\leq d$ where $d \geq 1$.*

(i) The R-module of $\mathbf{x} \in R^n$ with $A\mathbf{x} = \mathbf{0}$ is generated by vectors $\mathbf{x}$ whose coordinates are polynomials of degree at most $(2md)^{2^N}$.

(ii) Suppose that $A\mathbf{x} = \mathbf{b}$ is solvable in $\mathbf{x} \in R^n$. Then it has a solution $\mathbf{x}$ whose coordinates are polynomials of degree at most $(2md)^{2^N}$.

Proof. See Aschenbrenner [1, Thms. 3.2, 3.4]. Results of this type were obtained earlier, but not with a completely correct proof, by Hermann [12, 1926] and Seidenberg [24, 1974]. □

Corollary 2.3. *Let $R := \mathbb{Q}[X_1, \ldots, X_N]$. Further, Let A be an $m \times n$-matrix of polynomials in $\mathbb{Z}[X_1, \ldots, X_N]$ of degrees at most d and logarithmic heights at most h where $d \geq 1$, $h \geq 1$. Then the R-module of $\mathbf{x} \in R^n$ with $A\mathbf{x} = \mathbf{0}$ is generated by vectors $\mathbf{x}$, consisting of polynomials in $\mathbb{Z}[X_1, \ldots, X_N]$ of degree at most $(2md)^{2^N}$ and height at most $(2md)^{6^N}(h+1)$.*

Proof. By Proposition 2.2 (i) we have to study $A\mathbf{x} = \mathbf{0}$, restricted to vectors $\mathbf{x} \in R^n$ consisting of polynomials of degree at most $(2d)^{2^N}$. The set of these $\mathbf{x}$ is a finite dimensional $\mathbb{Q}$-vector space, and we have to prove that it is generated by vectors whose coordinates are polynomials in $\mathbb{Z}[X_1, \ldots, X_N]$ of logarithmic height at most $(2md)^{6^N}(h+1)$.

If $\mathbf{x}$ consists of polynomials of degree at most $(2md)^{2^N}$, then $A\mathbf{x}$ consists of m polynomials with coefficients in $\mathbb{Q}$ of degrees at most $(2md)^{2^N} + d$, all whose coefficients have to be set to 0. This leads to a system of linear equations $U\mathbf{y} = \mathbf{0}$, where $\mathbf{y}$ consists of the coefficients of the polynomials in $\mathbf{x}$ and U consists of integers of logarithmic heights at most h. Notice

that the number m^* of rows of U is m times the number of monomials in N variables of degree at most $(2md)^{2^N}+d$, that is

$$m^* \le m\binom{(2md)^{2^N}+d+N}{N}.$$

By Lemma 2.1 the solution space of $U\mathbf{y}=\mathbf{0}$ is generated by integer vectors of logarithmic height at most

$$m^*h + \tfrac{1}{2}m^*\log m^* \le (2md)^{6^N}(h+1).$$

This completes the proof of our corollary. □

Lemma 2.4. *Let $U \in \mathbb{Z}^{m,n}$, $\mathbf{b}\in\mathbb{Z}^m$ be such that $U\mathbf{y}=\mathbf{b}$ is solvable in $\mathbb{Z}^n$. Then it has a solution $\mathbf{y}\in\mathbb{Z}^n$ with $h(\mathbf{y}) \le mh([U,\mathbf{b}]) + \frac{1}{2}m\log m$.*

Proof. Assume without loss of generality that U and $[U,\mathbf{b}]$ have rank m. By a result of Borosh, Flahive, Rubin and Treybig [4, 1989], $U\mathbf{y}=\mathbf{b}$ has a solution $\mathbf{y}\in\mathbb{Z}^n$ such that the absolute values of the entries of $\mathbf{y}$ are bounded above by the maximum of the absolute values of the $m\times m$-subdeterminants of $[U,\mathbf{b}]$. The upper bound for $h(\mathbf{y})$ as in the lemma easily follows from Hadamard's inequality. □

Proposition 2.5. *Let $N\ge 1$ and let $f_1,\dots,f_m,b\in\mathbb{Z}[X_1,\dots,X_N]$ be polynomials of degrees at most d and logarithmic heights at most h where $d\ge 1$, $h\ge 1$, such that*

$$f_1x_1+\cdots+f_mx_m = b \tag{2.1}$$

is solvable in $x_1,\dots,x_m\in\mathbb{Z}[X_1,\dots,x_N]$. Then (2.1) has a solution in polynomials $x_1,\dots,x_m\in\mathbb{Z}[X_1,\dots,X_N]$ with

$$\deg x_i \le (2d)^{\exp O(N\log^* N)}(h+1),\quad h(x_i)\le (2d)^{\exp O(N\log^* N)}(h+1)^{N+1} \tag{2.2}$$

for $i=1,\dots,m$.

Proof. Aschenbrenner's main theorem [1, Theorem A] states that Eq. (2.1) has a solution $x_1,\dots,x_m\in\mathbb{Z}[X_1,\dots,X_N]$ with $\deg x_i\le d_0$ for $i=1,\dots,m$, where

$$d_0 = (2d)^{\exp O(N\log^* N)}(h+1).$$

So it remains to show the existence of a solution with small logarithmic height.

Let us restrict to solutions $(x_1, \dots, x_m)$ of (2.1) of degree $\leq d_0$, and denote by $\mathbf{y}$ the vector of coefficients of the polynomials $x_1, \dots, x_m$. Then (2.1) translates into a system of linear equations $U\mathbf{y} = \mathbf{b}$ which is solvable over $\mathbb{Z}$. Here, the number of equations, i.e., number of rows of U, is equal to $m^* := \binom{d_0+d+N}{N}$. Further, $h(U, \mathbf{b}) \leq h$. By Lemma 2.4, $U\mathbf{y} = \mathbf{b}$ has a solution $\mathbf{y}$ with coordinates in $\mathbb{Z}$ of height at most

$$m^* h + \tfrac{1}{2} m^* \log m^* \leq (2d)^{\exp O(N \log^* N)} (h+1)^{N+1}.$$

It follows that (2.1) has a solution $x_1, \dots, x_m \in \mathbb{Z}[X_1, \dots, X_N]$ satisfying (2.2). □

Remarks. 1. Aschenbrenner gives in [1] an example which shows that the upper bound for the degrees of the x_i cannot depend on d and N only.

2. The above lemma gives an effective criterion for ideal membership in $\mathbb{Z}[X_1, \dots, X_N]$. Let $b \in \mathbb{Z}[X_1, \dots, X_N]$ be given. Further, suppose that an ideal I of $\mathbb{Z}[X_1, \dots, X_N]$ is given by a finite set of generators $f_1, \dots, f_m$. By the above lemma, if $b \in I$ then there are $x_1, \dots, x_m \in \mathbb{Z}[X_1, \dots, X_N]$ with upper bounds for the degrees and heights as in (2.2) such that $b = \sum_{i=1}^m x_i f_i$. It requires only a finite computation to check whether such x_i exist.

3. A reduction

We reduce the general unit equation (1.1) to a unit equation over a domain B of a special type which can be dealt with more easily.

Let again $A = \mathbb{Z}[z_1, \dots, z_r] \supset \mathbb{Z}$ be a commutative domain finitely generated over $\mathbb{Z}$ and denote by K the quotient field of A. We assume that $r > 0$. We have

$$A \cong \mathbb{Z}[X_1, \dots, X_r]/I \tag{3.1}$$

where I is the ideal of polynomials $f \in \mathbb{Z}[X_1, \dots, X_r]$ such that $f(z_1, \dots, z_r) = 0$. The ideal I is finitely generated. Let $d \geq 1$, $h \geq 1$ and assume that

$$I = (f_1, \dots, f_m) \text{ with } \deg f_i \leq d, \quad h(f_i) \leq h \ (i = 1, \dots, m). \tag{3.2}$$

Suppose that K has transcendence degree $q \geq 0$. In case that $q > 0$, we assume without loss of generality that $z_1, \dots, z_q$ form a transcendence basis of $K/\mathbb{Q}$. We write $t := r - q$ and rename $z_{q+1}, \dots, z_r$ as $y_1, \dots, y_t$,

respectively. In case that $t = 0$ we have $A = \mathbb{Z}[z_1, \ldots, z_q]$, $A^* = \{\pm 1\}$ and Theorem 1.1 is trivial. So we assume henceforth that $t > 0$.

Define

$$A_0 := \mathbb{Z}[z_1, \ldots, z_q], \quad K_0 := \mathbb{Q}(z_1, \ldots, z_q) \text{ if } q > 0,$$

$$A_0 := \mathbb{Z}, \quad K_0 := \mathbb{Q} \text{ if } q = 0.$$

Then

$$A = A_0[y_1, \ldots, y_t], \quad K = K_0(y_1, \ldots, y_t).$$

Clearly, K is a finite extension of K_0, so in particular an algebraic number field if $q = 0$. Using standard algebra techniques, one can show that there exist $y \in A$, $f \in A_0$ such that $K = K_0(y)$, y is integral over A_0, and

$$A \subseteq B := A_0[f^{-1}, y], \qquad a, b, c \in B^*.$$

If $\varepsilon, \eta \in A^*$ is a solution to (1.1), then $\varepsilon_1 := a\varepsilon/c$, $\eta_1 := b\eta/c$ satisfy

$$\varepsilon_1 + \eta_1 = 1, \quad \varepsilon_1, \eta_1 \in B^*. \tag{3.3}$$

At the end of this section, we formulate Proposition 3.8 which gives an effective result for equations of the type (3.3). More precisely, we introduce an other type of degree and height $\overline{\deg}\,(\alpha)$ and $\overline{h}(\alpha)$ for elements α of B, and give effective upper bounds for the $\overline{\deg}$ and $\overline{h}$ of ε_1, η_1. Subsequently we deduce Theorem 1.1.

The deduction of Theorem 1.1 is based on some auxiliary results which are proved first. We start with an explicit construction of y, f, with effective upper bounds in terms of r, d, h and a, b, c for the degrees and logarithmic heights of f and of the coefficients in A_0 of the monic minimal polynomial of y over A_0. Here we follow more or less Seidenberg [24, 1974]. Second, for a given solution ε, η of (1.1), we derive effective upper bounds for the degrees and logarithmic heights of representatives for ε, ε^{-1}, η, η^{-1} in terms of $\overline{\deg}\,(\varepsilon_1)$, $\overline{h}(\varepsilon_1)$, $\overline{\deg}\,(\eta_1)$, $\overline{h}(\eta_1)$. Here we use Proposition 2.5 (Aschenbrenner's result).

We introduce some further notation. First let $q > 0$. Then since $z_1, \ldots, z_q$ are algebraically independent, we may view them as independent variables, and for $\alpha \in A_0$, we denote by $\deg \alpha$, $h(\alpha)$ the total degree and logarithmic height of α, viewed as polynomial in $z_1, \ldots, z_q$. In case that $q = 0$, we have $A_0 = \mathbb{Z}$, and we agree that $\deg \alpha = 0$, $h(\alpha) = \log |\alpha|$ for $\alpha \in A_0$. We frequently use the following estimate, valid for all $q \geq 0$:

Lemma 3.1. *Let $g_1, \dots, g_n \in A_0$ and $g = g_1 \cdots g_n$. Then*

$$|h(g) - \sum_{i=1}^{n} h(g_i)| \le q \deg g.$$

Proof. See Bombieri and Gubler [3, Lemma 1.6.11, pp. 27]. □

We write $\mathbf{Y} = (X_{q+1}, \dots, X_r)$ and $K_0(\mathbf{Y}) := K_0(X_{q+1}, \dots, X_r)$, etc. Given $f \in \mathbb{Q}(X_1, \dots, X_r)$ we denote by f^* the rational function of $K_0(\mathbf{Y})$ obtained by substituting z_i for X_i for $i = 1, \dots, q$ (and $f^* = f$ if $q = 0$). We view elements $f^* \in A_0[\mathbf{Y}]$ as polynomials in $\mathbf{Y}$ with coefficients in A_0. We denote by $\deg_{\mathbf{Y}} f^*$ the (total) degree of $f^* \in K_0[\mathbf{Y}]$ with respect to $\mathbf{Y}$. We recall that $\deg g$ is defined for elements of A_0 and is taken with respect to $z_1, \dots, z_q$. With this notation, we can rewrite (3.1), (3.2) as

$$\text{(3.4)} \qquad \begin{cases} A \cong A_0[\mathbf{Y}]/(f_1^*, \dots, f_m^*), \\ \deg_{\mathbf{Y}} f_i^* \le d \text{ for } i = 1, \dots, m, \\ \text{the coefficients of } f_1^*, \dots, f_m^* \text{ in } A_0 \text{ have degrees at most } d \\ \text{and logarithmic heights at most } h. \end{cases}$$

Put $D := [K : K_0]$ and denote by $\sigma_1, \dots, \sigma_D$ the K_0- isomorphic embeddings of K in an algebraic closure $\overline{K_0}$ of K_0.

Lemma 3.2. *(i) We have $D \le d^t$.*
(ii) There exist integers $a_1, \dots, a_t$ with $|a_i| \le D^2$ for $i = 1, \dots, t$ such that for $w := a_1 y_1 + \cdots + a_t y_t$ we have $K = K_0(w)$.

Proof. (i) The set

$$\mathcal{W} := \{\mathbf{y} \in \overline{K_0}^t : \ f_1^*(\mathbf{y}) = \cdots = f_m^*(\mathbf{y}) = 0\}$$

consists precisely of the images of $(y_1, \dots, y_t)$ under $\sigma_1, \dots, \sigma_D$. So we have to prove that $\mathcal{W}$ has cardinality at most d^t.

In fact, this follows from a repeated application of Bézout's Theorem. Given $g_1, \dots, g_k \in K_0[\mathbf{Y}]$, we denote by $\mathcal{V}(g_1, \dots, g_k)$ the common set of zeros of $g_1, \dots, g_k$ in $\overline{K_0}^t$. Let $g_1 := f_1^*$. Then by the version of Bézout's Theorem in Hartshorne [11, p. 53, Thm. 7.7], the irreducible components of $\mathcal{V}(g_1)$ have dimension $t - 1$, and the sum of their degrees is at most $\deg_{\mathbf{Y}} g_1 \le d$. Take a $\overline{K_0}$-linear combination g_2 of $f_1^*, \dots, f_m^*$ not vanishing identically on any of the irreducible components of $\mathcal{V}(g_1)$. For any of

these components, say $\mathcal{V}$, the intersection of $\mathcal{V}$ and $\mathcal{V}(g_2)$ is a union of irreducible components, each of dimension $t-2$, whose degrees have sum at most $\deg_{\mathbf{Y}} g_2 \cdot \deg \mathcal{V} \le d \deg \mathcal{V}$. It follows that the irreducible components of $\mathcal{V}(g_1, g_2)$ have dimension $t-2$ and that the sum of their degrees is at most d^2. Continuing like this, we see that there are linear combinations $g_1, \dots, g_t$ of $f_1^*, \dots, f_m^*$ such that for $i = 1, \dots, t$, the irreducible components of $\mathcal{V}(g_1, \dots, g_i)$ have dimension $d-i$ and the sum of their degrees is at most d^i. For $i = t$ it follows that $\mathcal{V}(g_1, \dots, g_t)$ is a set of at most d^t points. Since $\mathcal{W} \subseteq \mathcal{V}(g_1, \dots, g_t)$ this proves (i).

(ii) Let $a_1, \dots, a_t$ be integers. Then $w := \sum_{i=1}^t a_i y_i$ generates K over K_0 if and only if $\sum_{j=1}^t a_j \sigma_i(y_j)$ $(i = 1, \dots, D)$ are distinct. There are integers a_i with $|a_i| \le D^2$ for which this holds. □

Lemma 3.3. *There are $\mathcal{G}_0, \dots, \mathcal{G}_D \in A_0$ such that*

$$\sum_{i=0}^{D} \mathcal{G}_i w^{D-i} = 0, \quad \mathcal{G}_0 \mathcal{G}_D \neq 0, \tag{3.5}$$

$$\deg \mathcal{G}_i \le (2d)^{\exp O(r)}, \quad h(\mathcal{G}_i) \le (2d)^{\exp O(r)}(h+1) \quad (i = 0, \dots, D). \tag{3.6}$$

Proof. In what follows we write $\mathbf{Y} = (X_{q+1}, \dots, X_r)$ and $\mathbf{Y}^{\mathbf{u}} := X_{q+1}^{u_1} \cdots X_{q+t}^{u_t}$, $|\mathbf{u}| := u_1 + \cdots + u_t$ for tuples of non-negative integers $\mathbf{u} = (u_1, \dots, u_t)$. Further, we define $W := \sum_{j=1}^t a_j X_{q+j}$.

$\mathcal{G}_0, \dots, \mathcal{G}_D$ as in (3.5) clearly exist since w has degree D over K_0. By (3.4), there are $g_1^*, \dots, g_m^* \in A_0[\mathbf{Y}]$ such that

$$\sum_{i=0}^{D} \mathcal{G}_i W^{D-i} = \sum_{j=1}^{m} g_j^* f_j^*. \tag{3.7}$$

By Proposition 2.2 (ii), applied with the field $F = K_0$, there are polynomials $g_j^* \in K_0[\mathbf{Y}]$ (so with coefficients being rational functions in $\mathbf{z}$) satisfying (3.7) of degree at most $(2\max(d, D))^{2^t} \le (2d^t)^{2^t} =: d_0$ in $\mathbf{Y}$. By multiplying $\mathcal{G}_0, \dots, \mathcal{G}_D$ with an appropriate non-zero factor from A_0 we may assume that the g_j^* are polynomials in $A_0[\mathbf{Y}]$ of degree at most d_0 in $\mathbf{Y}$. By considering (3.7) with such polynomials g_j^*, we obtain

$$\sum_{i=0}^{D} \mathcal{G}_i W^{D-i} = \sum_{j=1}^{m} \Big(\sum_{|\mathbf{u}| \le d_0} g_{j,\mathbf{u}} \mathbf{Y}^{\mathbf{u}} \Big) \cdot \Big(\sum_{|\mathbf{v}| \le d} f_{j,\mathbf{v}} \mathbf{Y}^{\mathbf{v}} \Big), \tag{3.8}$$

where $g_{j,\mathbf{u}} \in A_0$ and $f_j^* = \sum_{|\mathbf{v}|\le d} f_{j,\mathbf{v}}\mathbf{Y}^{\mathbf{v}}$ with $f_{j,\mathbf{v}} \in A_0$. We view $\mathcal{G}_0, \dots, \mathcal{G}_D$ and the polynomials $g_{j,\mathbf{u}}$ as the unknowns of (3.8). Then (3.8) has solutions with $\mathcal{G}_0\mathcal{G}_D \neq 0$.

We may view (3.8) as a system of linear equations $\mathcal{A}\mathbf{x} = \mathbf{0}$ over K_0, where $\mathbf{x}$ consists of $\mathcal{G}_i$ $(i = 0, \dots, D)$ and $g_{j,\mathbf{u}}$ $(j = 1, \dots, m,\ |\mathbf{u}| \le d_0)$. By Lemma 3.2 and an elementary estimate, the polynomial $W^{D-i} = (\sum_{k=1}^{t} a_k X_{q+k})^{D-i}$ has logarithmic height at most $O(D\log(2D^2t)) \le (2d)^{O(t)}$. By combining this with (3.4), it follows that the entries of the matrix $\mathcal{A}$ are elements of A_0 of degrees at most d and logarithmic heights at most $h_0 := \max((2d)^{O(t)}, h)$. Further, the number of rows of $\mathcal{A}$ is at most the number of monomials in $\mathbf{Y}$ of degree at most $d_0 + d$ which is bounded above by $m_0 := \binom{d_0+d+t}{t}$. So by Corollary 2.3, the solution module of (3.8) is generated by vectors $\mathbf{x} = (\mathcal{G}_0, \dots, \mathcal{G}_D, \{g_{i,\mathbf{u}}\})$, consisting of elements from A_0 of degree and height at most

$$\big(2m_0d\big)^{2^q} \le (2d)^{\exp O(r)}, \quad \big(2m_0d\big)^{6^q}(h_0+1) \le (2d)^{\exp O(r)}(h+1),$$

respectively.

At least one of these vectors $\mathbf{x}$ must have $\mathcal{G}_0\mathcal{G}_D \neq 0$ since otherwise (3.8) would have no solution with $\mathcal{G}_0\mathcal{G}_D \neq 0$, contradicting (3.5). Thus, there exists a solution $\mathbf{x}$ whose components $\mathcal{G}_0, \dots, \mathcal{G}_D$ satisfy both (3.5), (3.6). This proves our lemma. □

It will be more convenient to work with

$$y := \mathcal{G}_0 w = \mathcal{G}_0 \cdot (a_1y_1 + \cdots + a_ty_t).$$

In the case $D = 1$ we set $y := 1$. The following properties of y follow at once from Lemmas 3.1–3.3.

Corollary 3.4. *We have $K = K_0(y)$, $y \in A$, y is integral over A_0, and y has minimal polynomial $\mathcal{F}(X) = X^D + \mathcal{F}_1X^{D-1} + \cdots + \mathcal{F}_D$ over K_0 with*

$$\mathcal{F}_i \in A_0, \quad \deg \mathcal{F}_i \le (2d)^{\exp O(r)},\ h(\mathcal{F}_i) \le (2d)^{\exp O(r)}(h+1)$$

for $i = 1, \dots, D$.

Recall that $A_0 = \mathbb{Z}$ if $q = 0$ and $\mathbb{Z}[z_1, \dots, z_q]$ if $q > 0$, where in the latter case, $z_1, \dots, z_q$ are algebraically independent. Hence A_0 is a unique factorization domain, and so the gcd of a finite set of elements of A_0 is well-defined and up to sign uniquely determined. With every element $\alpha \in K$ we

can associate an up to sign unique tuple $P_{\alpha,0},\dots,P_{\alpha,D-1},Q_\alpha$ of elements of A_0 such that

$$\alpha = Q_\alpha^{-1}\sum_{j=0}^{D-1} P_{\alpha,j}y^j \quad \text{with } Q_\alpha \neq 0,\ \gcd(P_{\alpha,0},\dots,P_{\alpha,D-1},Q_\alpha)=1. \tag{3.9}$$

Put

$$\begin{cases} \overline{\deg}\,\alpha := \max(\deg P_{\alpha,0},\dots,\deg P_{\alpha,D-1},\deg Q_\alpha), \\ \overline{h}(\alpha) := \max\big(h(P_{\alpha,0}),\dots,h(P_{\alpha,D-1}),h(Q_\alpha)\big) \end{cases}. \tag{3.10}$$

Then for $q=0$ we have $\overline{\deg}\,\alpha = 0$, $\overline{h}(\alpha) = \log\max\big(|P_{\alpha,0}|,\dots,|P_{\alpha,D-1}|,|Q_\alpha|\big)$.

Lemma 3.5. *Let $\alpha \in K^*$ and let (a,b) be a pair of representatives for α, with $a,b \in \mathbb{Z}[X_1,\dots,X_r]$, $b \notin I$. Put $d^* := \max(d,\deg a,\deg b)$, $h^* := \max(h,h(a),h(b))$. Then*

$$\overline{\deg}\,\alpha \le (2d^*)^{\exp O(r)},\quad \overline{h}(\alpha) \le (2d^*)^{\exp O(r)}(h^*+1). \tag{3.11}$$

Proof. Consider the linear equation

$$Q\cdot\alpha = \sum_{j=0}^{D-1} P_j y^j \tag{3.12}$$

in unknowns $P_0,\dots,P_{D-1},Q \in A_0$. This equation has a solution with $Q \neq 0$, since $\alpha \in K = K_0(y)$ and y has degree D over K_0. Write again $\mathbf{Y} = (X_{q+1},\dots,X_r)$ and put $Y := \mathcal{G}_0\cdot(\sum_{j=1}^t a_j X_{q+j})$. Let a^*, $b^* \in A_0[\mathbf{Y}]$ be obtained from a,b by substituting z_i for X_i for $i=1,\dots,q$ ($a^*=a$, $b^*=b$ if $q=0$). By (3.4), there are $g_j^* \in A_0[\mathbf{Y}]$ such that

$$Q\cdot a^* - b^*\sum_{j=0}^{D-1} P_j Y^j = \sum_{j=1}^m g_j^* f_j^*. \tag{3.13}$$

By Proposition 2.2 (ii) this identity holds with polynomials $g_j^* \in A_0[\mathbf{Y}]$ of degree in $\mathbf{Y}$ at most $(2\max(d^*,D))^{2^t} \le (2d^*)^{t2^t}$, where possibly we have to multiply $(P_0,\dots,P_{D-1},Q)$ with a non-zero element from A_0. Now completely similarly as in the proof of Lemma 3.3, one can rewrite (3.13) as a system of linear equations over K_0 and then apply Corollary 2.3. It follows that (3.12) is satisfied by $P_0,\dots,P_{D-1},Q \in A_0$ with $Q\neq 0$ and

$$\begin{aligned} &\deg P_i,\ \deg Q \le (2d^*)^{\exp O(r)}, \\ &h(P_i),\ h(Q) \le (2d^*)^{\exp O(r)}(h^*+1)\quad (i=0,\dots,D-1). \end{aligned}$$

By dividing $P_0, \dots, P_{D-1}, Q$ by their gcd and using Lemma 3.1 we obtain $P_{\alpha,0}, \dots, P_{D-1,\alpha}, Q_\alpha \in A_0$ satisfying both (3.9) and

$$\begin{aligned} &\deg P_{i,\alpha},\ \deg Q_\alpha \le (2d^*)^{\exp O(r)}, \\ &h(P_{i,\alpha}),\ h(Q_\alpha) \le (2d^*)^{\exp O(r)}(h^*+1) \quad (i = 0, \dots, D-1). \end{aligned}$$

□

Lemma 3.6. *Let $\alpha_1, \dots, \alpha_n \in K^*$. For $i = 1, \dots, n$, let (a_i, b_i) be a pair of representatives for α_i, with $a_i, b_i \in \mathbb{Z}[X_1, \dots, X_r]$, $b_i \notin I$. Put*

$$\begin{aligned} d^{**} &:= \max(d, \deg a_1, \deg b_1, \dots, \deg a_n, \deg b_n), \\ h^{**} &:= \max\big(h, h(a_1), h(b_1), \dots, h(a_n), h(b_n)\big). \end{aligned}$$

Then there is a non-zero $f \in A_0$ such that

(3.14) $A \subseteq A_0[y, f^{-1}],\ \alpha_1, \dots, \alpha_n \in A_0[y, f^{-1}]^*$,

(3.15) $\deg f \le (n+1)(2d^{**})^{\exp O(r)},\ h(f) \le (n+1)(2d^{**})^{\exp O(r)}(h^{**}+1)$.

Proof. Take

$$f := \prod_{i=1}^{t} Q_{y_i} \cdot \prod_{j=1}^{n} \big(Q_{\alpha_i} Q_{\alpha_i^{-1}}\big).$$

Since in general, $Q_\beta \beta \in A_0[y]$ for $\beta \in K^*$, we have $f\beta \in A_0[y]$ for $\beta = y_1, \dots, y_t, \alpha_1, \alpha_1^{-1}, \dots, \alpha_n, \alpha_n^{-1}$. This implies (3.14). The inequalities (3.15) follow at once from Lemmas 3.5 and 3.1. □

Lemma 3.7. *Let $\lambda \in K^*$ and let ε be a non-zero element of A. Let (a, b) with $a, b \in \mathbb{Z}[X_1, \dots, X_r]$ be a pair of representatives for λ. Put*

$$\begin{aligned} d_0 &:= \max(\deg f_1, \dots, \deg f_m, \deg a, \deg b, \overline{\deg}\, \lambda\varepsilon), \\ h_0 &:= \max\big(h(f_1), \dots, h(f_m), h(a), h(b), \overline{h}(\lambda\varepsilon)\big). \end{aligned}$$

Then ε has a representative $\widetilde{\varepsilon} \in \mathbb{Z}[X_1, \dots, X_r]$ such that

$$\deg \widetilde{\varepsilon} \le (2d_0)^{\exp O(r \log^* r)}(h_0+1), \quad h(\widetilde{\varepsilon}) \le (2d_0)^{\exp O(r \log^* r)}(h_0+1)^{r+1}.$$

If moreover $\varepsilon \in A^*$, then ε^{-1} has a representative $\widetilde{\varepsilon}' \in \mathbb{Z}[X_1, \dots, X_r]$ with

$$\deg \widetilde{\varepsilon}' \le (2d_0)^{\exp O(r \log^* r)}(h_0+1), \quad h(\widetilde{\varepsilon}') \le (2d_0)^{\exp O(r \log^* r)}(h_0+1)^{r+1}.$$

Proof. In case that $q > 0$, we identify z_i with X_i and view elements of A_0 as polynomials in $\mathbb{Z}[X_1, \ldots, X_q]$. Put $Y := \mathcal{G}_0 \cdot (\sum_{i=1}^{t} a_i X_{q+i})$. We have

$$\lambda\varepsilon = Q^{-1} \sum_{i=0}^{D-1} P_i y^i \tag{3.16}$$

with $P_0, \ldots, P_{D-1}, Q \in A_0$ and $\gcd(P_0, \ldots, P_{D-1}, Q) = 1$. According to (3.16), $\widetilde{\varepsilon} \in \mathbb{Z}[X_1, \ldots, X_r]$ is a representative for ε if and only if there are $g_1, \ldots, g_m \in \mathbb{Z}[X_1, \ldots, X_r]$ such that

$$\widetilde{\varepsilon} \cdot (Q \cdot a) + \sum_{i=1}^{m} g_i f_i = b \sum_{i=0}^{D-1} P_i Y^i. \tag{3.17}$$

We may view (3.17) as an inhomogeneous linear equation in the unknowns $\widetilde{\varepsilon}, g_1, \ldots, g_m$. Notice that by Lemmas 3.2–3.5 the degrees and logarithmic heights of Qa and $b\sum_{i=0}^{D-1} P_i Y^i$ are all bounded above by $(2d_0)^{\exp O(r)}$, $(2d_0)^{\exp O(r)}(h_0 + 1)$, respectively. Now Proposition 2.5 implies that (3.17) has a solution with upper bounds for $\deg \widetilde{\varepsilon}$, $h(\widetilde{\varepsilon})$ as stated in the lemma.

Now suppose that $\varepsilon \in A^*$. Again by (3.16), $\widetilde{\varepsilon}' \in \mathbb{Z}[X_1, \ldots, X_r]$ is a representative for ε^{-1} if and only if there are $g_1', \ldots, g_m' \in \mathbb{Z}[X_1, \ldots, X_r]$ such that

$$\widetilde{\varepsilon}' \cdot b \sum_{i=0}^{D-1} P_i Y^i + \sum_{i=1}^{m} g_i' f_i = Q \cdot a.$$

Similarly as above, this equation has a solution with upper bounds for $\deg \widetilde{\varepsilon}'$, $h(\widetilde{\varepsilon}')$ as stated in the lemma. □

Recall that we have defined $A_0 = \mathbb{Z}[z_1, \ldots, z_q]$, $K_0 = \mathbb{Q}(z_1, \ldots, z_q)$ if $q > 0$ and $A_0 = \mathbb{Z}$, $K_0 = \mathbb{Q}$ if $q = 0$, and that in the case $q = 0$, degrees and $\overline{\deg}$-s are always zero. Theorem 1.1 can be deduced from the following Proposition, which makes sense also if $q = 0$. The proof of this Proposition is given in Sections 4–6.

Proposition 3.8. *Let $f \in A_0$ with $f \neq 0$, and let*

$$\mathcal{F} = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X] \quad (D \geq 1)$$

be the minimal polynomial of y over K_0. Let $d_1 \geq 1$, $h_1 \geq 1$ and suppose

$$\max(\deg f, \deg \mathcal{F}_1, \ldots, \deg \mathcal{F}_D) \leq d_1, \quad \max(h(f), h(\mathcal{F}_1), \ldots, h(\mathcal{F}_D)) \leq h_1.$$

Define the domain $B := A_0[y, f^{-1}]$. *Then for each pair* (ε_1, η_1) *with*

$$\varepsilon_1 + \eta_1 = 1, \quad \varepsilon_1, \eta_1 \in B^* \tag{3.18}$$

we have

$$\overline{\deg}\,\varepsilon_1, \overline{\deg}\,\eta_1 \le 4qD^2 \cdot d_1, \tag{3.19}$$

$$\overline{h}(\varepsilon_1), \overline{h}(\eta_1) \le \exp O\Big(2D(q+d_1)\log^*\{2D(q+d_1)\} + Dh_1\Big). \tag{3.20}$$

Proof of Theorem 1.1. Let $a, b, c \in A$ be the coefficients of (1.3), and $\widetilde{a}, \widetilde{b}, \widetilde{c}$ the representatives for a, b, c from the statement of Theorem 1.1. By Lemma 3.6, there exists non-zero $f \in A_0$ such that that $A \subseteq B := A_0[y, f^{-1}]$, $a, b, c \in B^*$, and moreover, $\deg f \le (2d)^{\exp O(r)}$ and $h(f) \le (2d)^{\exp O(r)}(h+1)$. By Corollary 3.4 we have the same type of upper bounds for the degrees and logarithmic heights of $\mathcal{F}_1, \dots, \mathcal{F}_D$. So in Proposition 3.8 we may take $d_1 = (2d)^{\exp O(r)}$, $h_1 = (2d)^{\exp O(r)}(h+1)$. Finally, by Lemma 3.2 we have $D \le d^t$.

Let (ε, η) be a solution of (1.1) and put $\varepsilon_1 := a\varepsilon/c$, $\eta_1 := b\eta/c$. By Proposition 3.8 we have

$$\overline{\deg}\,\varepsilon_1 \le 4qd^{2t}(2d)^{\exp O(r)} \le (2d)^{\exp O(r)}, \quad \overline{h}(\varepsilon_1) \le \exp\Big((2d)^{\exp O(r)}(h+1)\Big).$$

We apply Lemma 3.7 with $\lambda = a/c$. Notice that λ is represented by $(\widetilde{a}, \widetilde{c})$. By assumption, $\widetilde{a}$ and $\widetilde{c}$ have degrees at most d and logarithmic heights at most h. Letting $\widetilde{a}, \widetilde{c}$ play the role of a, b in Lemma 3.7, we see that in that lemma we may take $h_0 = \exp\Big((2d)^{\exp O(r)}(h+1)\Big)$ and $d_0 = (2d)^{\exp O(r)}$. It follows that $\varepsilon, \varepsilon^{-1}$ have representatives $\widetilde{\varepsilon}$, $\widetilde{\varepsilon}' \in \mathbb{Z}[X_1, \dots, X_r]$ such that

$$\deg\widetilde{\varepsilon},\ \deg\widetilde{\varepsilon}',\ h(\widetilde{\varepsilon}),\ h(\widetilde{\varepsilon}') \le \exp\Big((2d)^{\exp O(r)}(h+1)\Big).$$

We observe here that the upper bound for $\overline{h}(\varepsilon_1)$ dominates by far the other terms in our estimation. In the same manner one can derive similar upper bounds for the degrees and logarithmic heights of representatives for η and η^{-1}. This completes the proof of Theorem 1.1. □

Proposition 3.8 is proved in Sections 4–6. In Section 4 we deduce the degree bound (3.19). Here, our main tool is Mason's effective result on S-unit equations over function fields [19, 1983]. In Section 5 we work out a more precise version of an effective specialization argument of Győry [8,

1983], [9, 1984]. In Section 6 we prove (3.20) by combining the specialization argument from Section 5 with a recent effective result for S-unit equations over number fields, due to Győry an Yu [10, 2006].

4. Bounding the degree

We start with recalling some results on function fields in one variable. Let $\mathbf{k}$ be an algebraically closed field of characteristic 0 and let z be transcendental over $\mathbf{k}$. Let K be a finite extension of $\mathbf{k}(z)$. Denote by $g_{K/\mathbf{k}}$ the genus of K, and by M_K the collection of valuations of $K/\mathbf{k}$, i.e, the valuations of K with value group $\mathbb{Z}$ which are trivial on $\mathbf{k}$. Recall that these valuations satisfy the sum formula

$$\sum_{v\in M_K} v(x) = 0 \quad \text{for } x \in K^*.$$

As usual, for a finite subset S of M_K the group of S-units of K is given by

$$O_S^* = \{x \in K^* : v(x) = 0 \text{ for } v \in M_K \setminus S\}.$$

The (homogeneous) height of $\mathbf{x} = (x_1, \dots, x_n) \in K^n$ relative to $K/\mathbf{k}$ is defined by

$$H_K(\mathbf{x}) = H_K(x_1, \dots, x_n) := -\sum_{v\in M_K} \min(v(x_1), \dots, v(x_n)).$$

By the sum formula,

$$H_K(\alpha\mathbf{x}) = H_K(\mathbf{x}) \quad \text{for } \alpha \in K^*. \tag{4.1}$$

The height of $x \in K$ relative to $K/\mathbf{k}$ is defined by

$$H_K(x) := H_K(1, x) = -\sum_{v\in M_K} \min(0, v(x)).$$

If L is a finite extension of K, we have

$$H_L(x_1, \dots, x_n) = [L : K]H_K(x_1, \dots, x_n) \quad \text{for } (x_1, \dots, x_n) \in K^n. \tag{4.2}$$

By $\deg f$ we denote the total degree of $f \in \mathbf{k}[z]$. Then for $f_1, \dots, f_n \in \mathbf{k}[z]$ with $\gcd(f_1, \dots, f_n) = 1$ we have

$$H_{\mathbf{k}(z)}(f_1, \dots, f_n) = \max(\deg f_1, \dots, \deg f_n). \tag{4.3}$$

Lemma 4.1. *Let $y_1, \dots, y_m \in K$ and suppose that*

$$X^m + f_1 X^{m-1} + \cdots + f_m = (X - y_1) \cdots (X - y_m)$$

for certain $f_1, \dots, f_m \in \mathbf{k}[z]$. Then

$$[K : \mathbf{k}(z)] \max(\deg f_1, \dots, \deg f_m) = \sum_{i=1}^{m} H_K(y_i).$$

Proof. By Gauss' Lemma we have for $v \in M_K$,

$$\min(v(f_1), \dots, v(f_m)) = \sum_{i=1}^{m} \min(0, v(y_i)).$$

Now take the sum over $v \in M_K$ and apply (4.2), (4.3). □

Lemma 4.2. *Let K be the splitting field over $\mathbf{k}(z)$ of $F := X^m + f_1 X^{m-1} + \cdots + f_m$, where $f_1, \dots, f_m \in \mathbf{k}[z]$. Then*

$$g_{K/\mathbf{k}} \le (d-1)m \cdot \max_{1 \le i \le m} \deg f_i,$$

where $d := [K : \mathbf{k}(z)]$.

Proof. This is Lemma H of Schmidt [23, 1978]. □

In what follows, the cardinality of a set S is denoted by $|S|$.

Proposition 4.3. *Let K be a finite extension of $\mathbf{k}(z)$ and S be a finite subset of M_K. Then for every solution of*

$$x + y = 1 \quad \text{in } x, y \in O_S^* \setminus \mathbf{k}^* \tag{4.4}$$

we have $\max(H_K(x), H_K(y)) \le |S| + 2g_{K/\mathbf{k}} - 2$.

Proof. See Mason [19, 1983]. □

We keep the notation from Proposition 3.8. We may assume that $q > 0$ because the case $q = 0$ is trivial. Let as before $K_0 = \mathbb{Q}(z_1, \dots, z_q)$, $K = K_0(y)$, $A_0 = \mathbb{Z}[z_1, \dots, z_q]$, $B = \mathbb{Z}[z_1, \dots, z_q, f^{-1}, y]$.

Fix $i \in \{1, \dots, q\}$. Let $\mathbf{k}_i := \mathbb{Q}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_q)$ and $\overline{\mathbf{k}_i}$ its algebraic closure. Thus, the domain A_0 is contained in $\overline{\mathbf{k}_i}[z_i]$. Let $y^{(1)} = y, \dots, y^{(D)}$ denote the conjugates of y over K_0. Let M_i denote the splitting field of the polynomial $X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$ over $\overline{\mathbf{k}_i}(z_i)$, i.e.

$$M_i := \overline{\mathbf{k}_i}(z_i, y^{(1)}, \dots, y^{(D)}).$$

The subring

$$B_i := \overline{\mathbf{k}_i}[z_i, f^{-1}, y^{(1)}, \dots, y^{(D)}]$$

of M_i contains $B = \mathbb{Z}[z_1, \dots, z_q, f^{-1}, y]$ as a subring. Put $\Delta_i := [M_i : \overline{\mathbf{k}_i}(z_i)]$.

We apply Lemmas 4.1, 4.2 and Proposition 4.3 with $z_i, \mathbf{k}_i, M_i$ instead of $z, \mathbf{k}, K$. Denote by g_{M_i} the genus of $M_i/\overline{\mathbf{k}_i}$. The height H_{M_i} is taken with respect to $M_i/\overline{\mathbf{k}_i}$. For $g \in A_0$, we denote by $\deg_{z_i} g$ the degree of g in the variable z_i.

Lemma 4.4. *Let $\alpha \in K$ and denote by $\alpha^{(1)}, \dots, \alpha^{(D)}$ the conjugates of α over K_0. Then*

$$\overline{\deg}\,\alpha \le qD \cdot d_1 + \sum_{i=1}^{q} \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}).$$

Proof. We have

$$\alpha = Q^{-1} \sum_{j=0}^{D-1} P_j y^j$$

for certain $P_0, \dots, P_{D-1}, Q \in A_0$ with $\gcd(Q, P_0, \dots, P_{D-1}) = 1$. Clearly,

$$\overline{\deg}\,\alpha \le \sum_{i=1}^{q} \mu_i, \quad \text{where } \mu_i := \max(\deg_{z_i} Q, \deg_{z_i} P_0, \dots, \deg_{z_i} P_{D-1}). \tag{4.5}$$

Below, we estimate $\mu_1, \dots, \mu_q$ from above. We fix $i \in \{1, \dots, q\}$ and use the notation introduced above.

Obviously,

$$\alpha^{(k)} = Q^{-1} \sum_{j=0}^{D-1} P_j \cdot (y^{(k)})^j \text{ for } k = 1, \dots, D.$$

Let Ω be the $D \times D$-matrix with rows

$$(1, \dots, 1),\ (y^{(1)}, \dots, y^{(D)}), \dots, \big((y^{(1)})^{D-1}, \dots, (y^{(D)})^{D-1}\big).$$

By Cramer's rule, $P_j/Q = \delta_j/\delta$, where $\delta = \det \Omega$, and δ_j is the determinant of the matrix obtained by replacing the j-th row of Ω by $(\alpha^{(1)}, \dots, \alpha^{(D)})$.

Gauss' Lemma implies that $\gcd(P_0, \dots, P_{D-1}, Q) = 1$ in the ring in $\mathbf{k}_i[z_i]$. By (4.3) (with z_i in place of z) we have

$$\begin{aligned} \mu_i &= \max(\deg_{z_i} Q, \deg_{z_i} P_0, \dots, \deg_{z_i} P_{D-1}) \\ &= H_{\overline{\mathbf{k}}(z_i)}(Q, P_0, \dots, P_{D-1}). \end{aligned}$$

Using $[M_i : \overline{\mathbf{k}_i}(z_i)] = \Delta_i$, the identities (4.2), (4.1) (with z_i instead of z) and the fact that $(\delta, \delta_1, \dots, \delta_D)$ is a scalar multiple of $(Q, P_0, \dots, P_{D-1})$ we obtain

$$\Delta_i \mu_i = H_{M_i}(Q, P_0, \dots, P_{D-1}) = H_{M_i}(\delta, \delta_1, \dots, \delta_D). \tag{4.6}$$

We bound from above the right-hand side. A straightforward estimate yields that for every valuation v of $M_i/\overline{\mathbf{k}_i}$,

$$\begin{aligned} &-\min(v(\delta), v(\delta_1), \dots, v(\delta_D)) \\ &\qquad \le -D \sum_{j=1}^{D} \min(0, v(y^{(j)})) - \sum_{j=1}^{D} \min(0, v(\alpha^{(j)})). \end{aligned}$$

Then summation over v and an application of Lemma 4.1 lead to

$$\begin{aligned} H_{M_i}(\delta, \delta_1, \dots, \delta_D) &\le D \sum_{j=1}^{D} H_{M_i}(y^{(j)}) + \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}), \\ &\le D\Delta_i \max(\deg_{z_i} \mathcal{F}_1, \dots, \deg \mathcal{F}_D) + \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}) \\ &\le \Delta_i \cdot Dd_1 + \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}), \end{aligned}$$

and then a combination with (4.6) gives

$$\mu_i \le Dd_1 + \Delta_i^{-1} \sum_{j=1}^{D} H_{M_i}(\alpha^{(j)}).$$

Now these bounds for $i = 1, \dots, q$ together with (4.5) imply our Lemma. □

Proof of (3.19). We fix again $i \in \{1, \dots, q\}$ and use the notation introduced above. By Lemma 4.2, applied with $\mathbf{k}_i, z_i, M_i$ instead of $\mathbf{k}, z, K$ and with $F = \mathcal{F} = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$, we have

$$g_{M_i} \le (\Delta_i - 1) D \max_j \deg_{z_i}(\mathcal{F}_j) \le (\Delta_i - 1) \cdot Dd_1. \tag{4.7}$$

Let S denote the subset of valuations v of $M_i/\overline{\mathbf{k}_i}$ such that $v(z_i) < 0$ or $v(f) > 0$. Each valuation of $\overline{\mathbf{k}_i}(z_i)$ can be extended to at most $[M_i : \overline{\mathbf{k}_i}(z_i)] =$

Δ_i valuations of M_i. Hence M_i has at most Δ_i valuations v with $v(z_i) < 0$ and at most $\Delta_i \deg f$ valuations with $v(f) > 0$. Thus,

$$|S| \le \Delta_i + \Delta_i \deg_{z_i} f \le \Delta_i(1 + \deg f) \le \Delta_i(1 + d_1). \tag{4.8}$$

Every $\alpha \in M_i$ which is integral over $\overline{\mathbf{k}}_i[z_i, f^{-1}]$ belongs to O_S. The elements $y^{(1)}, \dots, y^{(D)}$ belong to M_i and are integral over $A_0 = \mathbb{Z}[z_1, \dots, z_q]$ so they certainly belong to O_S. As a consequence, the elements of B and their conjugates over $\mathbb{Q}(z_1, \dots, z_q)$ belong to O_S. In particular, if $\varepsilon_1, \eta_1 \in B^*$ and $\varepsilon_1 + \eta_1 = 1$, then

$$\varepsilon_1^{(j)} + \eta_1^{(j)} = 1,\ \varepsilon_1^{(j)}, \eta_1^{(j)} \in O_S^* \ \text{ for } j = 1, \dots, D. \tag{4.9}$$

We apply Proposition 4.3 and insert the upper bounds (4.7), (4.8). It follows that for $j = 1, \dots, D$ we have either $\varepsilon_1^{(j)} \in \overline{\mathbf{k}_i}$ or

$$H_{M_i}(\varepsilon_1^{(j)}) \le |S| + 2g_{M_i} - 2 \le 3\Delta_i \cdot Dd_1;$$

in fact the last upper bound is valid also if $\varepsilon_1^{(j)} \in \overline{\mathbf{k}_i}$. Together with Lemma 4.4 this gives

$$\overline{\deg}\, \varepsilon_1 \le qDd_1 + qD \cdot 3Dd_1 \le 4qD^2 d_1.$$

For $\overline{\deg}\, \eta_1$ we derive the same estimate. This proves (3.19). □

5. Specializations

In this section we prove some results about specialization homomorphisms from the domain B from Proposition 3.8 to $\overline{\mathbb{Q}}$. We start with some notation and some preparatory lemmas.

The set of places of $\mathbb{Q}$ is $M_{\mathbb{Q}} = \{\infty\} \cup \{\text{primes}\}$. By $|\cdot|_\infty$ we denote the ordinary absolute value on $\mathbb{Q}$ and by $|\cdot|_p$ (p prime) the p-adic absolute value, with $|p|_p = p^{-1}$. More generally, let L be an algebraic number field and denote by M_L its set of places. Given $v \in M_L$, we define the absolute value $|\cdot|_v$ in such a way that its restriction to $\mathbb{Q}$ is $|\cdot|_p$ if v lies above $p \in M_{\mathbb{Q}}$. These absolute values satisfy the product formula $\prod_{v \in M_L} |x|_v^{d_v} = 1$ for $x \in L^*$, where $d_v := [L_v : \mathbb{Q}_p]/[L : \mathbb{Q}]$.

The (absolute logarithmic) height of $\mathbf{x} = (x_1, \dots, x_m) \in L^m \setminus \{\mathbf{0}\}$ is defined by

$$h(\mathbf{x}) = h(x_1, \dots, x_m) = \log \prod_{v \in M_L} \big(\max(|x_1|_v, \dots, |x_m|_v)\big)^{d_v}.$$

By the product formula, $h(\alpha\mathbf{x}) = h(\mathbf{x})$ for $\alpha \in L^*$. Moreover, $h(\mathbf{x})$ depends only on $\mathbf{x}$ and not on the choice of the field L such that $\mathbf{x} \in L^m$. So it defines a height on $\overline{\mathbb{Q}}^m \setminus \{\mathbf{0}\}$. The (absolute logarithmic) height of $\alpha \in \overline{\mathbb{Q}}$ is defined by $h(\alpha) := h((1, \alpha))$. In case that $\alpha \in L$ we have

$$h(\alpha) = \log \prod_{v \in M_L} \max(1, |\alpha|_v^{d_v}).$$

For $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{Z}^m$ with $\gcd(a_1, \ldots, a_m) = 1$ we have

$$h(\mathbf{a}) = \log \max(|a_1|, \ldots, |a_m|). \tag{5.1}$$

It is easy to verify that for $a_1, \ldots, a_m, b_1, \ldots, b_m \in \overline{\mathbb{Q}}$,

$$h(a_1b_1 + \cdots + a_mb_m) \leq h(1, a_1, \ldots, a_m) + h(1, b_1, \ldots, b_m) + \log m. \tag{5.2}$$

Let G be a polynomial with coefficients in L. If $a_1, \ldots, a_r$ are the non-zero coefficients of G, we put $|G|_v := \max(|a_1|_v, \ldots, |a_r|_v)$ for $v \in M_L$. For a polynomial G with coefficients in $\mathbb{Z}$ we define $h(G) := \log |G|_\infty$.

We start with four auxiliary results that are used in the construction of our specializations.

Lemma 5.1. *Let $m \geq 1$, $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}}$ and suppose that $G(X) := \prod_{i=1}^m (X - \alpha_i) \in \mathbb{Z}[X]$. Then*

$$|h(G) - \sum_{i=1}^m h(\alpha_i)| \leq m.$$

Proof. See Bombieri and Gubler [3, Theorem 1.6.13, pp. 28]. □

Lemma 5.2. *Let $m \geq 1$, let $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{Q}}$ be distinct and suppose that $G(X) := \prod_{i=1}^m (X - \alpha_i) \in \mathbb{Z}[X]$. Let $q, p_0, \ldots, p_{m-1}$ be integers with*

$$\gcd(q, p_0, \ldots, p_{m-1}) = 1,$$

and put

$$\beta_i := \sum_{j=0}^{m-1} (p_j/q)\alpha_i^j \quad (i = 1, \ldots, m).$$

Then

$$\log \max(|q|, |p_0|, \ldots, |p_{m-1}|) \leq 2m^2 + (m-1)h(G) + \sum_{j=1}^m h(\beta_j).$$

Proof. For $m = 1$ the assertion is obvious, so we assume $m \geq 2$. Let Ω be the $m \times m$ matrix with rows $(\alpha_1^i, \ldots, \alpha_m^i)$ $(i = 0, \ldots, m-1)$. By Cramer's rule we have $p_i/q = \delta_i/\delta$ $(i = 0, \ldots, m-1)$, where $\delta = \det \Omega$ and δ_i is the determinant of the matrix, obtained by replacing the i-th row of Ω by $(\beta_1, \ldots, \beta_m)$. Put $\mu := \log \max(|q|, |p_0|, \ldots, |p_{m-1}|)$. Then by (5.1),

$$\mu = h(q, p_0, \ldots, p_{m-1}) = h(\delta, \delta_0, \ldots, \delta_{m-1}).$$

Let $L = \mathbb{Q}(\alpha_1, \ldots, \alpha_m)$. By Hadamard's inequality for the infinite places and the ultrametric inequality for the finite places, we get

$$\max(|\delta|_v, |\delta_1|_v, \ldots, |\delta_m|_v) \leq c_v \prod_{i=1}^{m} \max(1, |\alpha_i|_v)^{m-1} \max(1, |\beta_i|_v)$$

for $v \in M_L$, where $c_v = m^{m/2}$ if v is infinite and $c_v = 1$ if v is finite. By taking the product over $v \in M_L$ and then logarithms, it follows that

$$\mu \leq \tfrac{1}{2} m \log m + \sum_{i=1}^{m} \big((m-1)h(\alpha_i) + h(\beta_i)\big).$$

A combination with Lemma 5.1 implies our lemma. □

Lemma 5.3. *Let $g \in \mathbb{Z}[z_1, \ldots, z_q]$ be a non-zero polynomial of degree d and $\mathcal{N}$ a subset of $\mathbb{Z}$ of cardinality $> d$. Then*

$$|\{\mathbf{u} \in \mathcal{N}^q : \ g(\mathbf{u}) = 0\}| \leq d|\mathcal{N}|^{q-1}.$$

Proof. We proceed by induction on q. For $q = 1$ the assertion is clear. Let $q \geq 2$. Write $g = \sum_{i=0}^{d_0} g_i(z_1, \ldots, z_{q-1}) z_q^i$ with $g_i \in \mathbb{Z}[z_1, \ldots, z_{q-1}]$ and $g_{d_0} \neq 0$. Then $\deg g_{d_0} \leq d - d_0$. By the induction hypothesis, there are at most $(d-d_0)|\mathcal{N}|^{q-2} \cdot |\mathcal{N}|$ tuples $(u_1, \ldots, u_q) \in \mathcal{N}^q$ with $g_{d_0}(u_1, \ldots, u_{q-1}) = 0$. Further, there are at most $|\mathcal{N}|^{q-1} \cdot d_0$ tuples $\mathbf{u} \in \mathcal{N}^q$ with $g_{d_0}(u_1, \ldots, u_{q-1}) \neq 0$ and $g(u_1, \ldots, u_q) = 0$. Summing these two quantities implies that g has at most $d|\mathcal{N}|^{q-1}$ zeros in $\mathcal{N}^q$. □

Lemma 5.4. *Let $g_1, g_2 \in \mathbb{Z}[z_1, \ldots, z_q]$ be two non-zero polynomials of degrees D_1, D_2, respectively, and let N be an integer $\geq \max(D_1, D_2)$. Define*

$$\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q : \ |\mathbf{u}| \leq N, \, g_2(\mathbf{u}) \neq 0\}.$$

Then $\mathcal{S}$ is non-empty, and

$$|g_1|_p \leq (4N)^{qD_1(D_1+1)/2} \max\{|g_1(\mathbf{u})|_p : \ \mathbf{u} \in \mathcal{S}\} \tag{5.3}$$

for $p \in M_{\mathbb{Q}} = \{\infty\} \cup \{\text{primes}\}$.

Proof. Put $C_p := \max\{|g_1(\mathbf{u})|_p : \ \mathbf{u} \in \mathcal{S}\}$ for $p \in M_{\mathbb{Q}}$. We proceed by induction on q, starting with $q = 0$. In the case $q = 0$ we interpret g_1, g_2 as non-zero constants with $|g_1|_p = C_p$ for $p \in M_{\mathbb{Q}}$. Then the lemma is trivial. Let $q \geq 1$. Write

$$g_1 = \sum_{j=0}^{D_1'} g_{1j}(z_1, \ldots, z_{q-1}) z_q^j, \quad g_2 = \sum_{j=0}^{D_2'} g_{2j}(z_1, \ldots, z_{q-1}) z_q^j,$$

where $g_{1,D_1'}, g_{2,D_2'} \neq 0$. By the induction hypothesis, the set

$$\mathcal{S}' := \{\mathbf{u}' \in \mathbb{Z}^{q-1} : \ |\mathbf{u}'| \leq N, \ g_{2,D_2'}(\mathbf{u}') \neq 0\}$$

is non-empty and moreover,

$$\max_{0 \leq j \leq D_1'} |g_{1j}|_p \leq (4N)^{(q-1)D_1(D_1+1)/2} C_p' \ \text{ for } p \in M_{\mathbb{Q}} \tag{5.4}$$

where

$$C_p' := \max\{|g_{1j}(\mathbf{u}')|_p : \ \mathbf{u}' \in \mathcal{S}', \ j = 0, \ldots, D_1'\}.$$

We estimate C_p' from above in terms of C_p. Fix $\mathbf{u}' \in \mathcal{S}'$. There are at least $2N+1-D_2' \geq D_1'+1$ integers u_q with $|u_q| \leq N$ such that $g_2(\mathbf{u}', u_q) \neq 0$. Let $a_0, \ldots, a_{D_1'}$ be distinct integers from this set. By Lagrange's interpolation formula,

$$\begin{aligned} g_1(\mathbf{u}', X) &= \sum_{j=0}^{D_1'} g_{1j}(\mathbf{u}') X^j \\ &= \sum_{j=0}^{D_1'} g_1(\mathbf{u}', a_j) \prod_{\substack{i=0 \\ i \neq j}}^{D_1'} \frac{X - a_i}{a_j - a_i}. \end{aligned}$$

From this we deduce

$$\begin{aligned} \max_{0 \leq j \leq D_1'} |g_{1j}(\mathbf{u}')| &\leq C_\infty \sum_{j=0}^{D_1'} \prod_{\substack{i=0 \\ i \neq j}}^{D_1'} \frac{1 + |a_i|}{|a_j - a_i|} \\ &\leq C_\infty (D_1' + 1)(N+1)^{D_1'} \leq (4N)^{D_1'(D_1'+1)/2} C_\infty. \end{aligned}$$

Now let p be a prime and put $\Delta := \prod_{1 \leq i < j \leq D_1'} |a_j - a_i|$. Then

$$\max_{0 \leq j \leq D_1'} |g_{1j}(\mathbf{u}')|_p \leq C_p |\Delta|_p^{-1} \leq \Delta C_p \leq (4N)^{D_1'(D_1'+1)/2} C_p.$$

It follows that $C_p' \le (4N)^{D_1'(D_1'+1)/2} C_p$ for $p \in M_{\mathbb{Q}}$. A combination with (5.4) gives (5.3). $\square$

We now introduce our specializations $B \to \overline{\mathbb{Q}}$ and prove some properties. We assume $q > 0$ and apart from that keep the notation and assumptions from Proposition 3.8. In particular, $A_0 = \mathbb{Z}[z_1, \dots, z_q]$, $K_0 = \mathbb{Q}(z_1, \dots, z_q)$ and

$$K = \mathbb{Q}(z_1, \dots, z_q, y), \quad B = \mathbb{Z}[z_1, \dots, z_q, f^{-1}, y],$$

where f is a non-zero element of A_0, y is integral over A_0, and y has minimal polynomial

$$\mathcal{F} := X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in A_0[X]$$

over K_0. In the case $D = 1$, we take $y = 1$, $\mathcal{F} = X - 1$.

To allow for other applications (e.g., Lemma 7.2 below), we consider a more general situation than what is needed for the proof of Proposition 3.8. Let $d_1 \ge d_0 \ge 1$, $h_1 \ge h_0 \ge 1$ and assume that

$$\begin{cases} \max(\deg \mathcal{F}_1, \dots, \deg \mathcal{F}_D) \le d_0, \quad \max(d_0, \deg f) \le d_1, \\ \max\big(h(\mathcal{F}_1), \dots, h(\mathcal{F}_D)\big) \le h_0, \quad \max(h_0, h(f)) \le h_1. \end{cases} \tag{5.5}$$

Let $\mathbf{u} = (u_1, \dots, u_q) \in \mathbb{Z}^q$. Then the substitution $z_1 \mapsto u_1, \dots, z_q \mapsto u_q$ defines a ring homomorphism (specialization)

$$\varphi_{\mathbf{u}} : \ \alpha \mapsto \alpha(\mathbf{u}) : \ \{g_1/g_2 : \ g_1, g_2 \in A_0, \ g_2(\mathbf{u}) \ne 0\} \to \mathbb{Q}.$$

We want to extend this to a ring homomorphism from B to $\overline{\mathbb{Q}}$ and for this, we have to impose some restrictions on $\mathbf{u}$. Denote by $\Delta_{\mathcal{F}}$ the discriminant of $\mathcal{F}$ (with $\Delta_{\mathcal{F}} := 1$ if $D = \overline{\deg}\, \mathcal{F} = 1$), and let

$$\mathcal{H} := \Delta_{\mathcal{F}} \mathcal{F}_D \cdot f. \tag{5.6}$$

Then $\mathcal{H} \in A_0$. Using that $\Delta_{\mathcal{F}}$ is a polynomial of degree $2D-2$ with integer coefficients in $\mathcal{F}_1, \dots, \mathcal{F}_D$, it follows easily that

$$\deg \mathcal{H} \le (2D-1)d_0 + d_1 \le 2Dd_1. \tag{5.7}$$

Now assume that

$$\mathcal{H}(\mathbf{u}) \ne 0. \tag{5.8}$$

Then $f(\mathbf{u}) \ne 0$ and moreover, the polynomial

$$\mathcal{F}_{\mathbf{u}} := X^D + \mathcal{F}_1(\mathbf{u}) X^{D-1} + \cdots + \mathcal{F}_D(\mathbf{u})$$

has D distinct zeros which are all different from 0, say $y_1(\mathbf{u}), \dots, y_D(\mathbf{u})$. Thus, for $j = 1, \dots, D$ the assignment

$$z_1 \mapsto u_1, \dots, z_q \mapsto u_q, \quad y \mapsto y_j(\mathbf{u})$$

defines a ring homomorphism $\varphi_{\mathbf{u},j}$ from B to $\overline{\mathbb{Q}}$; in the case $D = 1$ it is just $\varphi_{\mathbf{u}}$. The image of $\alpha \in B$ under $\varphi_{\mathbf{u},j}$ is denoted by $\alpha_j(\mathbf{u})$. Recall that we may express elements α of B as

$$\alpha = \sum_{i=0}^{D-1} (P_i/Q) y^i \tag{5.9}$$
$$\text{with } P_0, \dots, P_{D-1}, Q \in A_0, \ \gcd(P_0, \dots, P_{D-1}, Q) = 1.$$

Since $\alpha \in B$, the denominator Q must divide a power of f, hence $Q(\mathbf{u}) \neq 0$. So we have

$$\alpha_j(\mathbf{u}) = \sum_{i=0}^{D-1} (P_i(\mathbf{u})/Q(\mathbf{u})) y_j(\mathbf{u})^i \quad (j = 1, \dots, D). \tag{5.10}$$

It is obvious that $\varphi_{\mathbf{u},j}$ is the identity on $B \cap \mathbb{Q}$. Thus, if $\alpha \in B \cap \overline{\mathbb{Q}}$, then $\varphi_{\mathbf{u},j}(\alpha)$ has the same minimal polynomial as α and so it is conjugate to α.

For $\mathbf{u} = (u_1, \dots, u_q) \in \mathbb{Z}^q$, we put $|\mathbf{u}| := \max(|u_1|, \dots, |u_q|)$. It is easy to verify that for any $g \in A_0$, $\mathbf{u} \in \mathbb{Z}^q$,

$$\log |g(\mathbf{u})| \leq q \log \deg g + h(g) + \deg g \log \max(1, |\mathbf{u}|). \tag{5.11}$$

In particular,

$$h(\mathcal{F}_{\mathbf{u}}) \leq q \log d_0 + h_0 + d_0 \log \max(1, |\mathbf{u}|) \tag{5.12}$$

and so by Lemma 5.2 (ii),

$$\sum_{j=1}^{D} h(y_j(\mathbf{u})) \leq D + q \log d_0 + h_0 + d_0 \log \max(1, |\mathbf{u}|). \tag{5.13}$$

Define the algebraic number fields $K_{\mathbf{u},j} := \mathbb{Q}(\mathbf{y}_j(\mathbf{u}))$ $(j = 1, \dots, D)$. Denote by Δ_L the discriminant of an algebraic number field L. We derive an upper bound for the discriminant $\Delta_{K_{\mathbf{u},j}}$ of $K_{\mathbf{u},j}$.

Lemma 5.5. *Let $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \neq 0$. Then for $j = 1, \dots, D$ we have $[K_{\mathbf{u},j} : \mathbb{Q}] \leq D$ and*

$$|\Delta_{K_{\mathbf{u},j}}| \leq D^{2D-1} \left(d_0^q \cdot e^{h_0} \max(1, |\mathbf{u}|)^{d_0} \right)^{2D-2}.$$

Proof. Let $j \in \{1, \ldots, D\}$. The estimate for the degree is obvious. To estimate the discriminant, let $\mathcal{P}_j$ be the monic minimal polynomial of $y_j(\mathbf{u})$. Then $\Delta_{K_{\mathbf{u},j}}$ divides the discriminant $\Delta_{\mathcal{P}_j}$ of $\mathcal{P}_j$. Using the expression of the discriminant of a monic polynomial as the product of the squares of the differences of its zeros, one easily shows that $\Delta_{\mathcal{P}_j}$ divides $\Delta_{\mathcal{F}_\mathbf{u}}$ in the ring of algebraic integers and so also in $\mathbb{Z}$. Therefore, $\Delta_{K_{\mathbf{u},j}}$ divides $\Delta_{\mathcal{F}_\mathbf{u}}$ in $\mathbb{Z}$.

It remains to estimate from above the discriminant of $\mathcal{F}_\mathbf{u}$. By, e.g., Lewis and Mahler [14, bottom of p. 335], we have

$$|\Delta_{\mathcal{F}_\mathbf{u}}| \le D^{2D-1}|\mathcal{F}_\mathbf{u}|^{2D-2},$$

where $|\mathcal{F}_\mathbf{u}|$ denotes the maximum of the absolute values of the coefficients of $\mathcal{F}_\mathbf{u}$. By (5.12), this is bounded above by $d_0^q e^{h_0} \max(1, |\mathbf{u}|)^{d_0}$, so

$$|\Delta_{\mathcal{F}_\mathbf{u}}| \le D^{2D-1}\big(d_0^q e^{h_0} \max(1, |\mathbf{u}|)^{d_0}\big)^{2D-2}.$$

This implies our lemma. □

We finish with two lemmas, which relate the height of $\alpha \in B$ to the heights of $\alpha_j(\mathbf{u})$ for $\mathbf{u} \in \mathbb{Z}^q$.

Lemma 5.6. *Let $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \ne 0$. Let $\alpha \in B$. Then for $j = 1, \ldots, D$,*

$$h(\alpha_j(\mathbf{u})) \le D^2 + q(D \log d_0 + \log \overline{\deg}\, \alpha) + Dh_0 + \overline{h}(\alpha) + \\ + (Dd_0 + \overline{\deg}\, \alpha) \log \max(1, |\mathbf{u}|).$$

Proof. Let $P_0, \ldots, P_{D-1}, Q$ as in (5.9) and write $\alpha_j(\mathbf{u})$ as in (5.10). By (5.2),

$$h(\alpha_j(\mathbf{u})) \le \log D + \\ + h\big(1, P_0(\mathbf{u})/Q(\mathbf{u}), \ldots, P_{D-1}(\mathbf{u})/Q(\mathbf{u})\big) + (D-1)h(y_j(\mathbf{u})). \tag{5.14}$$

From (5.11) we infer

$$\begin{aligned} &h(1, P_0(\mathbf{u})/Q(\mathbf{u}), \ldots, P_{D-1}(\mathbf{u})/Q(\mathbf{u})) \\ &\quad \le \log \max(|Q(\mathbf{u})|, |P_0(\mathbf{u})|, \ldots, |P_{D-1}(\mathbf{u})|) \\ &\quad \le q \log \overline{\deg}\, \alpha + \overline{h}(\alpha) + \overline{\deg}\, \alpha \cdot \log \max(1, |\mathbf{u}|). \end{aligned}$$

By combining (5.14) with this inequality and with (5.13), our lemma easily follows. □

Lemma 5.7. *Let $\alpha \in B$, $\alpha \neq 0$, and let N be an integer with*

$$N \geq \max\left(\overline{\deg}\,\alpha,\, 2Dd_0 + 2(q+1)(d_1+1)\right).$$

Then the set

$$\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q : \ |\mathbf{u}| \leq N,\ \mathcal{H}(\mathbf{u}) \neq 0\}$$

is non-empty, and

$$\overline{h}(\alpha) \leq 5N^4(h_1+1)^2 + 2D(h_1+1)H$$

where $H := \max\{h(\alpha_j(\mathbf{u})) : \mathbf{u} \in \mathcal{S},\ j = 1, \ldots, D\}$.

Proof. It follows from our assumption on N, (5.7), and Lemma 5.4 that $\mathcal{S}$ is non-empty. We proceed with estimating $\overline{h}(\alpha)$.

Let $P_0, \ldots, P_{D-1}, Q \in A_0$ be as in (5.9). We analyse Q more closely. Let

$$f = \pm p_1^{k_1} \cdots p_m^{k_m} g_1^{l_1} \cdots g_n^{l_n}$$

be the unique factorization of f in A_0, where $p_1, \ldots, p_m$ are distinct prime numbers, and $\pm g_1, \ldots, \pm g_n$ distinct irreducible elements of A_0 of positive degree. Notice that

$$m \leq h(f)/\log 2 \leq h_1/\log 2, \tag{5.15}$$

$$\sum_{i=1}^{n} l_i h(g_i) \leq qd_1 + h_1, \tag{5.16}$$

where the last inequality is a consequence of Lemma 5.1. Since $\alpha \in B$, the polynomial Q is also composed of $p_1, \ldots, p_m$, $g_1, \ldots, g_n$. Hence

$$Q = a\widetilde{Q} \ \text{ with } a = \pm p_1^{k_1'} \cdots p_m^{k_m'},\ \widetilde{Q} = g_1^{l_1'} \cdots g_n^{l_n'} \tag{5.17}$$

for certain non-negative integers $k_1', \ldots, l_n'$. Clearly,

$$l_1' + \cdots + l_n' \leq \deg Q \leq \overline{\deg}\,\alpha \leq N,$$

and by Lemma 3.1 and (5.16),

$$h(\widetilde{Q}) \leq q \deg Q + \sum_{i=1}^{n} l_i' h(g_i) \leq N(q + qd_1 + h_1) \leq N^2(h_1+1). \tag{5.18}$$

In view of (5.11), we have for $\mathbf{u} \in \mathcal{S}$,

$$\begin{aligned} \log|\widetilde{Q}(\mathbf{u})| &\leq q\log d_1 + h(\widetilde{Q}) + \deg Q \log N \\ &\leq \tfrac{3}{2}N\log N + N^2(h_1+1) \leq N^2(h_1+2). \end{aligned}$$

Hence

$$h(\widetilde{Q}(\mathbf{u})\alpha_j(\mathbf{u})) \le N^2(h_1+2)+H$$

for $\mathbf{u}\in\mathcal{S}$, $j=1,\dots,D$. Further, by (5.10), (5.16) we have

$$\widetilde{Q}(\mathbf{u})\alpha_j(\mathbf{u}) = \sum_{i=0}^{D-1}(P_i(\mathbf{u})/a)y_j(\mathbf{u})^i.$$

Put

$$\delta(\mathbf{u}) := \gcd(a, P_0(\mathbf{u}),\dots,P_{D-1}(\mathbf{u})).$$

Then by applying Lemma 5.2 and then (5.12) we obtain

$$\begin{aligned}(5.19)\quad &\log\left(\frac{\max(|a|,|P_0(\mathbf{u})|,\dots,|P_{D-1}(\mathbf{u})|}{\delta(\mathbf{u})}\right)\\ &\le 2D^2+(D-1)h(\mathcal{F}_{\mathbf{u}})+D\big(N^2(h_1+2)+H\big)\\ &\le 2D^2+(D-1)(q\log d_1+h_1+d_1\log N)+D\big(N^2(h_1+2)+H\big)\\ &\le N^3(h_1+2)+DH.\end{aligned}$$

Our assumption that $\gcd(Q,P_0,\dots,P_{D-1})=1$ implies that the gcd of a and the coefficients of $P_0,\dots,P_{D-1}$ is 1. Let $p\in\{p_1,\dots,p_m\}$ be one of the prime factors of a. There is $j\in\{0,\dots,D-1\}$ such that $|P_j|_p=1$. Our assumption on N and (5.7) imply that $N\ge\max(\deg\mathcal{H},\deg P_j)$. This means that Lemma 5.4 is applicable with $g_1=P_j$ and $g_2=\mathcal{H}$. It follows that

$$\max\{|P_j(\mathbf{u})|_p:\ \mathbf{u}\in\mathcal{S}\}\ge(4N)^{-qN(N+1)/2}.$$

That is, there is $\mathbf{u}_0\in\mathcal{S}$ with $|P_j(\mathbf{u}_0)|_p\ge(4N)^{-qN(N+1)/2}$. Hence

$$|\delta(\mathbf{u}_0)|_p\ge(4N)^{-qN(N+1)/2}.$$

Together with (5.19), this implies

$$\begin{aligned}\log|a|_p^{-1} &\le \log|a/\delta(\mathbf{u}_0)|+\log|\delta(\mathbf{u}_0)|_p^{-1}\\ &\le N^3(h_1+2)+DH+\tfrac12N^3\log 4N\le N^4(h_1+3)+DH.\end{aligned}$$

Combining this with the upper bound (5.15) for the number of prime factors of a, we obtain

$$(5.20)\qquad \log|a|\le 2N^4h_1(h_1+3)+2Dh_1\cdot H.$$

Together with (5.17), (5.18), this implies

$$\begin{aligned} h(Q) &\leq 2N^4h_1(h_1+3)+2Dh_1\cdot H+N^2(h_1+1) \\ &\leq 3N^4(h_1+1)^2+2Dh_1\cdot H. \end{aligned} \tag{5.21}$$

Further, the right-hand side of (5.20) is also an upper bound for $\log\delta(\mathbf{u})$, for $\mathbf{u}\in\mathcal{S}$. Combining this with (5.19) gives

$$\begin{aligned} &\log\max\{|P_j(\mathbf{u})|:\ \mathbf{u}\in\mathcal{S},\ j=0,\dots,D-1\} \\ &\quad\leq N^3(h_1+2)+DH+3N^4(h_1+1)^2+2Dh_1\cdot H \\ &\quad\leq 4N^4(h_1+1)^2+2D(h_1+1)\cdot H. \end{aligned}$$

Another application of Lemma 5.4 yields

$$\begin{aligned} h(P_j) &\leq \tfrac{1}{2}qN(N+1)\log 4N+4N^4(h_1+1)^2+2D(h_1+1)\cdot H \\ &\leq 5N^4(h_1+1)^2+2D(h_1+1)\cdot H \end{aligned}$$

for $j=0,\dots,D-1$. Together with (5.21) this gives the upper bound for $\overline{h}(\alpha)$ from our lemma. □

6. Completion of the proof of Proposition 3.8

It remains only to prove the height bound in (3.20). We use an effective result of Győry and Yu [10, 2006] on S-unit equations in number fields. To state this, we need some notation.

Let L be an algebraic number field of degree d_L. We denote by O_L, M_L, Δ_L, h_L, R_L the ring of integers, set of places, discriminant, class number and regulator of L. The norm of an ideal $\mathfrak{a}$ of O_L, i.e., $|O_L/\mathfrak{a}|$, is denoted by $N\mathfrak{a}$.

Further, let S be a finite set of places of L, containing all infinite places. Suppose S has cardinality s. Recall that the ring of S-integers O_S and the group of S-units O_S^* are given by

$$\begin{aligned} O_S &= \{x\in L:\ |x|_v\leq 1 \text{ for } v\in M_L\setminus S\}, \\ O_S^* &= \{x\in L:\ |x|_v=1 \text{ for } v\in M_L\setminus S\}. \end{aligned}$$

If case that S consists only of the infinite places of L, we put $P:=2$, $Q:=2$. If S contains also finite places, let $\mathfrak{p}_1,\dots,\mathfrak{p}_t$ denote the prime

ideals corresponding to the finite places of S, and put

$$P := \max\{N\mathfrak{p}_1, \dots, N\mathfrak{p}_t\}, \quad Q := N(\mathfrak{p}_1 \cdots \mathfrak{p}_t).$$

Further, let R_S denote the S-regulator associated with S. In case that S consists only of the infinite places of L it is equal to R_L, while otherwise

$$R_S = h_S R_L \prod_{i=1}^{t} \log N\mathfrak{p}_i,$$

where h_S is a divisor of h_L whose definition is not important here. By, e.g., formula (59) of [10] (which is an easy consequence of formula (2) of Louboutin [16, 2000]) we have

$$h_L R_L \le |\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1}.$$

By the inequality of the geometric and arithmetic mean, we have for $t > 0$,

$$\prod_{i=1}^{t} \log N\mathfrak{p}_i \le \Big(t^{-1} \log(N\mathfrak{p}_1 \cdots N\mathfrak{p}_t)\Big)^t \le (\log Q)^s$$

and hence,

$$R_S \le |\Delta_L|^{1/2} (\log^* |\Delta_L|)^{d_L - 1} \cdot (\log^* Q)^s. \tag{6.1}$$

This is clearly true also if $t = 0$.

Proposition 6.1. *Let ε, η such that*

$$\varepsilon + \eta = 1, \quad \varepsilon, \eta \in O_S^*. \tag{6.2}$$

Then

$$\max(h(\varepsilon), h(\eta)) \le c_1 P R_S \left(1 + \log^* R_S / \log P\right), \tag{6.3}$$

where

$$c_1 = \max(1, \pi/d_L) s^{2s+3.5} 2^{7s+27} (\log 2s) d_L^{2(s+1)} (\log^* 2d_L)^3.$$

Proof. This is Theorem 1 of Győry, Yu [10] with $\alpha = \beta = 1$. □

Proof of (3.20). As before, we use $O(\cdot)$ to denote a quantity which is $c\times$ the expression between the parentheses, where c is an effectively computable absolute constant which may be different at each occurrence of the O-symbol.

We first consider the case $q > 0$. Let ε_1, η_1 be a solution of (3.18). Pick $\mathbf{u} \in \mathbb{Z}^q$ with $\mathcal{H}(\mathbf{u}) \ne 0$, pick $j \in \{1, \dots, D\}$ and put $L := K_{\mathbf{u},j}$. Further, let the set of places S consist of all infinite places of L, and all finite places

of L lying above the rational prime divisors of $f(\mathbf{u})$. Note that $y_j(\mathbf{u})$ is an algebraic integer, and $f(\mathbf{u}) \in O_S^*$. Hence $\varphi_{\mathbf{u},j}(B) \subseteq O_S$ and $\varphi_{\mathbf{u},j}(B^*) \subseteq O_S^*$. So

$$\varepsilon_{1,j}(\mathbf{u}) + \eta_{1,j}(\mathbf{u}) = 1, \quad \varepsilon_{1,j}(\mathbf{u}),\, \eta_{1,j}(\mathbf{u}) \in O_S^*, \tag{6.4}$$

where $\varepsilon_{1,j}(\mathbf{u}), \eta_{1,j}(\mathbf{u})$ are the images of ε_1, η_1 under $\varphi_{\mathbf{u},j}$.

We estimate from above the upper bound (6.3) from Proposition 6.1. By assumption, f has degree at most d_1 and logarithmic height at most h_1, hence

$$|f(\mathbf{u})| \le d_1^q e^{h_1} \max(1, |\mathbf{u}|)^{d_1} =: R(\mathbf{u}). \tag{6.5}$$

Since the degree of L is $d_L \le D$, the cardinality s of S is at most $s \le D(1+\omega)$, where ω is the number of prime divisors of $f(\mathbf{u})$. Using the inequality from prime number theory, $\omega \le O(\log|f(\mathbf{u})|/\log\log|f(\mathbf{u})|)$, we obtain

$$s \le O\Big(\frac{D \log^* R(\mathbf{u})}{\log^* \log^* R(\mathbf{u})}\Big). \tag{6.6}$$

From this, one easily deduces that

$$c_1 \le \exp O(D \log^* R(\mathbf{u})). \tag{6.7}$$

Next, we estimate P and R_S. By (6.5), we have

$$P \le Q \le |f(\mathbf{u})|^D \le \exp O(D \log^* R(\mathbf{u})). \tag{6.8}$$

To estimate R_S, we use (6.1). By Lemma 5.5 (using $d_0 \le d_1$) we have

$$|\Delta_L| \le D^{2D-1}\big(d_1^q e^{h_1} \max(1, |\mathbf{u}|)^{d_1}\big)^{2D-2} \le \exp O(D \log^* R(\mathbf{u})),$$

and this easily implies

$$|\Delta_L|^{1/2} (\log^* \Delta_L)^{D-1} \le \exp O(D \log^* R(\mathbf{u})).$$

Together with the estimates (6.6),(6.8) for s and Q, this leads to

$$R_S \le \exp O\Big(D \log^* R(\mathbf{u}) + s \log^* \log^* Q\Big) \le \exp O(D \log^* R(\mathbf{u})). \tag{6.9}$$

Now by collecting (6.7)–(6.9), we infer that the right-hand side of (6.3) is bounded above by $\exp O(D \log^* R(\mathbf{u}))$. So applying Proposition 6.1 to (6.4) gives

$$h(\varepsilon_{1,j}(\mathbf{u})),\, h(\eta_{1,j}(\mathbf{u})) \le \exp O(D \log^* R(\mathbf{u})). \tag{6.10}$$

We apply Lemma 5.7 with $N := 4D^2(q+d_1+1)^2$. From the already established (3.19) it follows that $\overline{\deg}\,\varepsilon_1$, $\overline{\deg}\,\eta_1 \le N$. Further, since $d_1 \ge d_0$ we have $N \ge 2Dd_0+2(d_1+1)(q+1)$. So indeed, Lemma 5.7 is applicable with this value of N. It follows that the set $\mathcal{S} := \{\mathbf{u} \in \mathbb{Z}^q : |\mathbf{u}| \le N,\ \mathcal{H}(\mathbf{u}) \ne 0\}$ is not empty. Further, for $\mathbf{u} \in \mathcal{S}$, $j = 1, \dots, D$, we have

$$\begin{aligned} h(\varepsilon_{1,j}(\mathbf{u})) &\le \exp O(Dq \log d_1 + Dh_1 + Dd_1 \log^* N) \\ &\le \exp O(N^{1/2} \log^* N + Dh_1), \end{aligned}$$

and so by Lemma 5.7,

$$\overline{h}(\varepsilon_1) \le \exp O(N^{1/2} \log^* N + Dh_1).$$

For $\overline{h}(\eta_1)$ we obtain the same upper bound. This easily implies (3.20) in the case $q > 0$.

Now assume $q = 0$. In this case, $K_0 = \mathbb{Q}$, $A_0 = \mathbb{Z}$ and $B = \mathbb{Z}[f^{-1}, y]$ where y is an algebraic integer with minimal polynomial $\mathcal{F} = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D \in \mathbb{Z}[X]$ over $\mathbb{Q}$, and f is a non-zero rational integer. By assumption, $\log|f| \le h_1$, $\log|\mathcal{F}_i| \le h_1$ for $i = 1, \dots, D$. Denote by $y_1, \dots, y_D$ the conjugates of y, and let $L = \mathbb{Q}(y_j)$ for some j. By a similar argument as in the proof of Lemma 5.5, we have $|\Delta_L| \le D^{2D-1} e^{(2D-2)h_1}$. The isomorphism given by $y \mapsto y_j$ maps K to L and B to O_S, where S consists of the infinite places of L and of the prime ideals of O_L that divide f. The estimates (6.5)–(6.9) remain valid if we replace $R(\mathbf{u})$ by e^{h_1}. Hence for any solution ε_1, η_1 of (3.18),

$$h(\varepsilon_{1,j}),\ h(\eta_{1,j}) \le \exp O(Dh_1),$$

where $\varepsilon_{1,j}$,$\eta_{1,j}$ are the j-th conjugates of ε_1, η_1, respectively. Now an application of Lemma 5.2 with $g = \mathcal{F}$, $m = D$, $\beta_j = \varepsilon_{1,j}$ gives

$$\overline{h}(\varepsilon_1) \le \exp O(Dh_1).$$

Again we derive the same upper bound for $\overline{h}(\eta_1)$, and deduce (3.20). This completes the proof of Proposition 3.8. □

7. Proof of Theorem 1.3

We start with some results on multiplicative (in)dependence.

Lemma 7.1. *Let L be an algebraic number field of degree d, and $\gamma_0, \dots, \gamma_s$ non-zero elements of L such that $\gamma_0, \dots, \gamma_s$ are multiplicatively dependent,*

but any s elements among $\gamma_0, \dots, \gamma_s$ are multiplicatively independent. Then there are non-zero integers $k_0, \dots, k_s$ such that

$$\gamma_0^{k_0} \cdots \gamma_s^{k_s} = 1,$$

$$|k_i| \le 58(s!e^s/s^s)d^{s+1}(\log d)h(\gamma_0)\cdots h(\gamma_s)/h(\gamma_i) \quad \textit{for } i = 0, \dots, s.$$

Proof. This is Corollary 3.2 of Loher and Masser [15, 2004]. They attribute this result to Yu Kunrui. Another result of this type was obtained earlier by Loxton and van der Poorten [17, 1983]. □

We prove a generalization for arbitrary finitely generated domains. As before, let $A = \mathbb{Z}[z_1, \dots, z_r] \supseteq \mathbb{Z}$ be a domain, and suppose that the ideal I of polynomials $f \in \mathbb{Z}[X_1, \dots, X_r]$ with $f(z_1, \dots, z_r) = 0$ is generated by $f_1, \dots, f_m$. Let K be the quotient field of A. Let $\gamma_0, \dots, \gamma_s$ be non-zero elements of K, and for $i = 1, \dots, s$, let (g_{i1}, g_{i2}) be a pair of representatives for γ_i, i.e., elements of $\mathbb{Z}[X_1, \dots, X_r]$ such that

$$\gamma_i = \frac{g_{i1}(z_1, \dots, z_r)}{g_{i2}(z_1, \dots, z_r)}.$$

Lemma 7.2. *Assume that $\gamma_0, \dots, \gamma_s$ are multiplicatively dependent. Further, assume that $f_1, \dots, f_m$ and g_{i1}, g_{i2} $(i = 0, \dots, s)$ have degrees at most d and logarithmic heights at most h, where $d \ge 1$, $h \ge 1$. Then there are integers $k_0, \dots, k_s$, not all equal to 0, such that*

$$\gamma_0^{k_0} \cdots \gamma_s^{k_s} = 1, \tag{7.1}$$

$$|k_i| \le (2d)^{\exp O(r+s)}(h+1)^s \quad \textit{for } i = 0, \dots, s. \tag{7.2}$$

Proof. We assume without loss of generality that any s numbers among $\gamma_0, \dots, \gamma_s$ are multiplicatively independent (if this is not the case, take a minimal multiplicatively dependent subset of $\{\gamma_0, \dots, \gamma_s\}$ and proceed further with this subset). We first assume that $q > 0$. We use an argument of van der Poorten and Schlickewei [21, 1991]. We keep the notation and assumptions from Sections 3–5. In particular, we assume that $z_1, \dots, z_q$ is a transcendence basis of K, and rename $z_{q+1}, \dots, z_r$ as $y_1, \dots, y_t$, respectively. For brevity, we have included the case $t = 0$ as well in our proof. But it should be possible to prove in this case a sharper result by means of a more elementary method. In the case $t > 0$, y and $\mathcal{F} = X^D + \mathcal{F}_1 X^{D-1} + \cdots + \mathcal{F}_D$ will be as in Corollary 3.4. In the case $t = 0$ we take $m = 1$, $f_1 = 0$,

$d = h = 1$, $y = 1$, $\mathcal{F} = X - 1$, $D = 1$. We construct a specialization such that among the images of $\gamma_0, \ldots, \gamma_s$ no s elements are multiplicatively dependent, and then apply Lemma 7.1.

Let $V \geq 2d$ be a positive integer. Later we shall make our choice of V more precise. Let

$$\mathcal{V} := \{\mathbf{v} = (v_0, \ldots, v_s) \in \mathbb{Z}^{s+1} \setminus \{\mathbf{0}\} : \quad |v_i| \leq V \text{ for } i = 0, \ldots, s, \text{and with } v_i = 0 \text{ for some } i\}. \tag{7.3}$$

Then

$$\gamma_{\mathbf{v}} := \Big(\prod_{i=0}^{s} \gamma_i^{v_i}\Big) - 1 \quad (\mathbf{v} \in \mathcal{V})$$

are non-zero elements of K. It is not difficult to show that for $\mathbf{v} \in \mathcal{V}$, $\gamma_{\mathbf{v}}$ has a pair of representatives $(g_{1,\mathbf{v}}, g_{2,\mathbf{v}})$ such that

$$\deg g_{1,\mathbf{v}},\ \deg g_{2,\mathbf{v}} \leq sdV.$$

In the case $t > 0$, there exists by Lemma 3.6 a non-zero $f \in A_0$ such that

$$A \subseteq B := A_0[y, f^{-1}], \quad \gamma_{\mathbf{v}} \in B^* \text{ for } \mathbf{v} \in \mathcal{V}$$

and

$$\deg f \leq V^{s+1}(2sdV)^{\exp O(r)} \leq V^{\exp O(r+s)}.$$

In the case $t = 0$ this holds true as well, with $y = 1$ and $f = \prod_{\mathbf{v}\in\mathcal{V}}((g_{1,\mathbf{v}} \cdot g_{2,\mathbf{v}})$. We apply the theory on specializations explained in Section 5 with this f. We put $\mathcal{H} := \Delta_{\mathcal{F}}\mathcal{F}_D f$, where $\Delta_{\mathcal{F}}$ is the discriminant of $\mathcal{F}$. Using Corollary 3.4 and inserting the bound $D \leq d^t$ from Lemma 3.2 we get for $t > 0$,

$$\begin{cases} d_0 := \max(\deg f_1, \ldots, \deg f_m, \deg \mathcal{F}_1, \ldots, \deg \mathcal{F}_D) \leq (2d)^{\exp O(r)}, \\ h_0 := \max\big(h(f_1), \ldots, h(f_m), h(\mathcal{F}_1), \ldots, h(\mathcal{F}_D)\big) \leq (2d)^{\exp O(r)}(h+1)\,; \end{cases} \tag{7.4}$$

with the provision $\deg 0 = h(0) = -\infty$ this is true also if $t = 0$. Combining this with Lemma 3.5, we obtain

$$\deg \mathcal{H} \leq (2D-1)d_0 + \deg f \leq V^{\exp O(r+s)}.$$

By Lemma 5.3 there exists $\mathbf{u} \in \mathbb{Z}^q$ with

$$\mathcal{H}(\mathbf{u}) \neq 0, \quad |\mathbf{u}| \leq V^{\exp O(r+s)}. \tag{7.5}$$

We proceed further with this $\mathbf{u}$.

As we have seen before, $\gamma_{\mathbf{v}} \in B^*$ for $\mathbf{v} \in \mathcal{V}$. By our choice of $\mathbf{u}$, there are D distinct specialization maps $\varphi_{\mathbf{u},j}$ $(j = 1, \dots, D)$ from B to $\overline{\mathbb{Q}}$. We fix one of these specializations, say $\varphi_{\mathbf{u}}$. Given $\alpha \in B$, we write $\alpha(\mathbf{u})$ for $\varphi_{\mathbf{u}}(\alpha)$. As the elements $\gamma_{\mathbf{v}}$ are all units in B, their images under $\varphi_{\mathbf{u}}$ are non-zero. So we have

$$\prod_{i=0}^{s} \gamma_i(\mathbf{u})^{v_i} \neq 1 \quad \text{for } \mathbf{v} \in \mathcal{V}, \tag{7.6}$$

where $\mathcal{V}$ is defined by (7.3).

We use Lemma 5.6 to estimate the heights $h(\gamma_i(\mathbf{u}))$ for $i = 0, \dots, s$. Recall that by Lemma 3.5 we have

$$\overline{\deg}\, \gamma_i \leq (2d)^{\exp O(r)}, \quad \overline{h}(\gamma_i) \leq (2d)^{\exp O(r)}(h+1)$$

for $i = 0, \dots, s$. By inserting these bounds, together with the bound $D \leq d^t$ from Lemma 3.2, those for d_0, h_0 from (7.4) and that for $\mathbf{u}$ from (7.5) into the bound from Lemma 5.6, we obtain for $i = 0, \dots, s$,

$$\begin{aligned} h(\gamma_i(\mathbf{u})) &\leq (2d)^{\exp O(r)}(1 + h + \log \max(1, |\mathbf{u}|)) \\ &\leq (2d)^{\exp O(r+s)}(1 + h + \log V). \end{aligned} \tag{7.7}$$

Assume that among $\gamma_0(\mathbf{u}), \dots, \gamma_s(\mathbf{u})$ there are s numbers which are multiplicatively dependent. By Lemma 7.1 there are integers $k_0, \dots, k_s$, at least one of which is non-zero and at least one of which is 0, such that

$$\begin{aligned} &\prod_{i=0}^{s} \gamma_i(\mathbf{u})^{k_i} = 0, \\ &|k_i| \leq (2d)^{\exp O(r+s)}(1 + h + \log V)^{s-1} \quad \text{for } i = 0, \dots, s. \end{aligned}$$

Now for

$$V = (2d)^{\exp O(r+s)}(h+1)^{s-1} \tag{7.8}$$

(with a sufficiently large constant in the O-symbol), the upper bound for the numbers $|k_i|$ is smaller than V. But this would imply that $\prod_{i=0}^{s} \gamma_i(\mathbf{u})^{v_i} = 1$ for some $\mathbf{v} \in \mathcal{V}$, contrary to (7.6). Thus we conclude that with the choice (7.8) for V, there exists $\mathbf{u} \in \mathbb{Z}^q$ with (7.5), such that any s numbers among $\gamma_0(\mathbf{u}), \dots, \gamma_s(\mathbf{u})$ are multiplicatively independent. Of course, the numbers $\gamma_0(\mathbf{u}), \dots, \gamma_s(\mathbf{u})$ are multiplicatively dependent, since they are the images

under $\varphi_{\mathbf{u}}$ of $\gamma_0, \ldots, \gamma_s$ which are multiplicatively dependent. Substituting (7.8) into (7.7) we obtain

$$h(\gamma_i(\mathbf{u})) \leq (2d)^{\exp O(r+s)}(h+1) \quad \text{for } i = 0, \ldots, s. \tag{7.9}$$

Now Lemma 7.1 implies that there are non-zero integers $k_0, \ldots, k_s$ such that

$$\prod_{i=0}^{s} \gamma_i(\mathbf{u})^{k_i} = 1, \tag{7.10}$$

$$|k_i| \leq (2d)^{\exp O(r+s)}(h+1)^s \quad \text{for } i = 0, \ldots, s. \tag{7.11}$$

Our assumption on $\gamma_0, \ldots, \gamma_s$ implies that there are non-zero integers $l_0, \ldots, l_s$ such that $\prod_{i=0}^{s} \gamma_i^{l_i} = 1$. Hence $\prod_{i=0}^{s} \gamma_i(\mathbf{u})^{l_i} = 1$. Together with (7.10) this implies

$$\prod_{i=1}^{s} \gamma_i(\mathbf{u})^{l_0 k_i - l_i k_0} = 1.$$

But $\gamma_1(\mathbf{u}), \ldots, \gamma_s(\mathbf{u})$ are multiplicatively independent, hence $l_0 k_i - l_i k_0 = 0$ for $i = 1, \ldots, s$. That is,

$$l_0(k_0, \ldots, k_s) = k_0(l_0, \ldots, l_s).$$

It follows that

$$\prod_{i=0}^{s} \gamma_i^{k_i} = \rho$$

for some root of unity ρ. But $\varphi_{\mathbf{u}}(\rho) = 1$ and it is conjugate to ρ. Hence $\rho = 1$. So in fact we have $\prod_{i=0}^{s} \gamma_i^{k_i} = 1$ with non-zero integers k_i satisfying (7.11). This proves our Lemma, but under the assumption $q > 0$. If $q = 0$ then a much simpler argument, without specializations, gives $h(\gamma_i) \leq (2d)^{\exp O(r+s)}(h+1)$ for $i = 0, \ldots, s$ instead of (7.9). Then the proof is finished in the same way as in the case $q > 0$. □

Corollary 7.3. *Let $\gamma_0, \gamma_1, \ldots, \gamma_s \in K^*$, and suppose that $\gamma_1, \ldots, \gamma_s$ are multiplicatively independent and*

$$\gamma_0 = \gamma_1^{k_1} \cdots \gamma_s^{k_s}$$

for certain integers $k_1, \ldots, k_s$. Then

$$|k_i| \leq (2d)^{\exp O(r+s)}(h+1)^s \quad \textit{for } i = 1, \ldots, s.$$

Proof. By Lemma 7.2, and by the multiplicative independence of $\gamma_1, \dots, \gamma_s$, there are integers $l_0, \dots, l_m$ such that

$$\prod_{i=0}^{m} \gamma_i^{l_i} = 1,$$

$$l_0 \neq 0, \quad |l_i| \leq (2d)^{\exp O(r+s)}(h+1)^s \text{ for } i = 0, \dots, s.$$

Now clearly, we have also

$$\prod_{i=1}^{s} \gamma_i^{l_0 k_i - l_i} = 1,$$

hence $l_0 k_i - l_i = 0$ for $i = 1, \dots, s$. It follows that $|k_i| = |l_i/l_0| \leq (2d)^{\exp O(r+s)}(h+1)^s$ for $i = 1, \dots, s$. This implies our Corollary. □

Proof of Theorem 1.3. We keep the notation and assumptions from the statement of Theorem 1.3. Define the domain

$$\widetilde{A} := A[\gamma_1, \gamma_1^{-1}, \dots, \gamma_s, \gamma_s^{-1}].$$

Then

$$\widetilde{A} \cong \mathbb{Z}[X_1, \dots, X_r, X_{r+1}, \dots, X_{r+2s}]/\widetilde{I}$$

with

$$\widetilde{I} = \Big(f_1, \dots, f_m, g_{12}X_{r+1} - g_{11}, g_{11}X_{r+2} - g_{12}, \dots \\ \dots, g_{s2}X_{r+2s-1} - g_{s1}, g_{s1}X_{r+2s} - g_{s2}\Big).$$

Let $(v_1, \dots, w_s)$ be a solution of (1.4), and put $\varepsilon := \prod_{i=1}^{s} \gamma_i^{v_i}$, $\eta := \prod_{i=1}^{s} \gamma_i^{w_i}$. Then

$$a\varepsilon + b\eta = c, \quad \varepsilon, \eta \in \widetilde{A}^*.$$

By Theorem 1.1, ε has a representative $\widetilde{\varepsilon} \in \mathbb{Z}[X_1, \dots, X_{r+2s}]$ of degree and logarithmic height both bounded above by

$$\exp\Big((2d)^{\exp O(r+s)}(h+1)\Big).$$

Now Corollary 7.3 implies

$$|v_i| \leq \exp\Big((2d)^{\exp O(r+s)}(h+1)\Big) \text{ for } i = 1, \dots, s.$$

For $|w_i|$ $(i = 1, \dots, s)$ we derive a similar upper bound. This completes the proof of Theorem 1.3. □

References

[1] M. Aschenbrenner, *Ideal membership in polynomial rings over the integers*, J. Amer. Math. Soc. **17** (2004), 407–442.

[2] A. Baker, *Contributions to the theory of Diophantine equations*, Philos. Trans. Roy. Soc. London, Ser. A **263**, 173–208.

[3] E. Bombieri, W. Gubler, *Heights in Diophantine Geometry*, Cambridge University Press, 2006.

[4] I. Borosh, M. Flahive, D. Rubin, B. Treybig, *A sharp bound for solutions of linear Diophantine equations*, Proc. Amer. Math. Soc. **105** (1989), 844–846.

[5] J. Coates, *An effective p-adic analogue of a theorem of Thue*, Acta Arith. **15** (1968/69), 279–305.

[6] K. Győry, *Sur les polynômes à coefficients entiers et de discriminant donné* II, Publ. Math. Debrecen **21** (1974), 125–144.

[7] K. Győry, *On the number of solutions of linear equations in units of an algebraic number field*, Comment. Math. Helv. **54** (1979), 583–600.

[8] K. Győry, *Bounds for the solutions of norm form, discriminant form and index form equations in finitely generated domains*, Acta Math. Hung. **42** (1983), 45–80.

[9] K. Győry, *Effective finiteness theorems for polynomials with given discriminant and integral elements with given discriminant over finitely generated domains*, J. reine angew. Math. **346**, 54–100.

[10] K. Győry, Kunrui Yu, *Bounds for the solutions of S-unit equations and decomposable form equations*, Acta Arith. **123** (2006), 9–41.

[11] R. Hartshorne, *Algebraic Geometry*, Springer Verlag, 1977.

[12] G. Hermann, *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann. **95** (1926), 736–788.

[13] S. Lang, *Integral points on curves*, Inst. Hautes Études Sci. Publ. Math. **6** (1960), 27–43.

[14] D.J. Lewis, K. Mahler, *On the representation of integers by binary forms*, Acta Arith. **6** (1961), 333–363.

[15] T. Loher, D. Masser, *Uniformly counting points of bounded height*, Acta Arith. **111** (2004), 277–297.

[16] S. Louboutin, *Explicit bounds for residues of dedekind zeta functions, values of L-functions at $s = 1$, and relative class numbers*, J. Number Theory **85** (2000), 263–282.

[17] J.H. Loxton, A.J. van der Poorten, *Multiplicative dependence in number fields*, Acta Arith. **42** (1983), 291–302.

[18] K. Mahler, *Zur Approximation algebraischer Zahlen, I. (Über den größten Primteiler binärer Formen)*, Math. Ann. **107** (1933), 691–730.

[19] R.C. Mason, *The hyperelliptic equation over function fields*, Math. Proc. Camb. Philos. Soc. **93** (1983), 219–230.

[20] C.J. Parry, *The $\mathfrak{p}$-adic generalisation of the Thue-Siegel theorem*, Acta Math. **83** (1950), 1–100.

[21] A.J. van der Poorten, H.P. Schlickewei, *Additive relations in fields*, J. Austral. Math. Soc. (Ser. A) **51** (1991), 154–170.

[22] P. Roquette, *Einheiten und Divisorenklassen in endlich erzeugbaren Körpern*, Jber. Deutsch. Math. Verein **60** (1958), 1–21.

[23] W.M. Schmidt, *Thue's equation over function fields*, J. Austral. Math. Soc. Ser. A **25** (1978), 385–422.

[24] A. Seidenberg, *Constructions in algebra*, Trans. Amer. Math. Soc. **197** (1974), 273–313.

[25] C.L. Siegel, *Approximation algebraischer Zahlen*, Math. Zeitschrift **10** (1921), 173–213.

[26] H. Simmons, *The solution of a decision problem for several classes of rings*, Pacific J. Math. **34** (1970), 547–557.

J.-H. Evertse
Universiteit Leiden, Mathematisch Instituut,
Postbus 9512, 2300 RA Leiden, The Netherlands

E-mail address: evertse@math.leidenuniv.nl

K. Győry
Institute of Mathematics, University of Debrecen
Number Theory Research Group, Hungarian Academy of Sciences and University of Debrecen
H-4010 Debrecen, P.O. Box 12, Hungary

E-mail address: gyory@science.unideb.hu