EFFECTIVE RESULTS FOR LINEAR EQUATIONS IN TWO UNKNOWNS FROM A MULTIPLICATIVE DIVISION GROUP

ATTILA BÉRCZES, JAN-HENDRIK EVERTSE, AND KÁLMÁN GYŐRY

Abstract. We prove a new effective result for equations $a_1x_1+a_2x_2=1$ in $(x_1,x_2)\in\Gamma$, where a_1,a_2 are non-zero algebraic numbers and Γ is a finitely generated multiplicative subgroup of $(\overline{\mathbb{Q}}^*)^2$. In our result, we give an explicit upper bound in terms of a_1,a_2 and Γ for the heights of the solutions (x_1,x_2). More generally, we prove effective results for solutions (x_1,x_2) in the division group of Γ and for solutions 'close' to this division group. Here, the solutions do not lie anymore in a given number field. Therefore, to achieve effectiveness we give for each solution (x_1,x_2), explicit upper bounds both for its height and for the degree of the number field it generates.

1. Introduction

In the literature there are various effective results on S-unit equations in two unknowns. In our paper we work out effective results in a quantitative form for the more general equation

$$a_1x_1+a_2x_2=1 \qquad \text{in } (x_1,x_2)\in\Gamma, \tag{1.1}$$

where $a_1,a_2\in\overline{\mathbb{Q}}^*$ and Γ is an arbitrary finitely generated subgroup of positive rank of the multiplicative group $(\overline{\mathbb{Q}}^*)^2=\overline{\mathbb{Q}}^*\times\overline{\mathbb{Q}}^*$ endowed with coordinatewise multiplication (see Theorems 2.1 and 2.2 in Section 2). Such more general results can be used to improve upon existing effective bounds on the solutions of discriminant equations and certain decomposable form equations. These will be worked out in a forthcoming work.

In fact, in the present paper we prove even more general effective results for equations of the shape (1.1) with solutions (x_1,x_2) from a larger group, nemely the division group $\overline{\Gamma}=\{(x_1,x_2)\in(\overline{\mathbb{Q}}^*)^2\,|\,\exists k\in\mathbb{Z}_{>0}:\ (x_1^k,x_2^k)\in\Gamma\}$, and even with solutions (x_1,x_2) 'very close' to $\overline{\Gamma}$. To our knowledge, these are the first effective results of this kind. Our results give an effective upper bound for both the height of a solution (x_1,x_2) and the degree of the field $\mathbb{Q}(x_1,x_2)$; see Theorems 2.3 and 2.5 and Corollary 2.4 in Section 2.

2000 Mathematics Subject Classification: Primary 11D61, 11J68; Secondary 11D57.

Keywords and Phrases: effective results, generalized unit equations, approximation of algebraic numbers by elements of a finitely generated multiplicative group.

The research was supported in part by grants T67580 (A.B., K.G.) and T48791 (A.B.) of the Hungarian National Foundation for Scientific Research, the János Bolyai Research Scholarship (A.B.) and the Eötvös Scholarship (A.B.).

In the proofs of these Theorems we utilize Theorem 2.1 (on (1.1) with solutions from Γ), as well as a result of Beukers and Zagier [2], which asserts that (1.1) has at most two solutions $(x_1, x_2) \in (\overline{\mathbb{Q}}^*)^2$ with very small height.

The hard core of the proofs of our results mentioned above is a new effective lower bound for $|1 - \alpha\xi|_v$, where α is a fixed element from a given algebraic number field K, v is a place of K, and the unknown ξ is taken from a given finitely generated subgroup of K^* (see Theorem 4.1 in Section 4). This result is proved using linear forms in logarithms estimates. Our Theorem 4.1 has a consequence (cf. Theorem 4.2 in Section 4) which is of a similar flavour as earlier results by Bombieri [3], Bombieri and Cohen [4], [5], and Bugeaud [7] (see also Bombieri and Gubler [6, Chap. 5.4]) but it gives in many cases a better estimate. Consequently, Theorem 4.1 leads to an explicit upper bound for the heights of the solutions of (1.1) which is in many cases sharper than what is obtainable from the work of Bombieri et al.

In Section 2 we state our results concerning (1.1), in Section 3 we introduce some notation, in Section 4 we state our results concerning $|1 - \alpha\xi|_v$, and in Sections 5, 6 we prove our Theorems.

2. Results

To state our results we need the following notation. If G is a finitely generated abelian group, then $\{\xi_1, \dots, \xi_r\}$ is called a system of generators of G/G_{tors} if $\xi_1, \dots, \xi_r \in G$, $\xi_1, \dots, \xi_r \notin G_{\mathrm{tors}}$, and the reductions of $\xi_1, \dots, \xi_r$ modulo G_{tors} generate G/G_{tors}. Such a system is called a basis of G/G_{tors} if its reduction modulo G_{tors} forms a basis of G/G_{tors}.

We fix an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ and assume that all algebraic number fields occurring henceforth are contained in $\overline{\mathbb{Q}}$. We denote by $(\overline{\mathbb{Q}}^*)^2$ the group $\{(x_1, x_2) \,|\, x_1, x_2 \in \overline{\mathbb{Q}}^*\}$ with coordinatewise multiplication $(x_1, x_2)(y_1, y_2) = (x_1y_1, x_2y_2)$. Further, we denote by $h(x)$ the absolute logarithmic height of $x \in \overline{\mathbb{Q}}^*$ and define the height of $\mathbf{x} = (x_1, x_2) \in (\overline{\mathbb{Q}}^*)^2$ by $h(\mathbf{x}) = h(x_1, x_2) := h(x_1) + h(x_2)$. Notice that this definition differs from the usual one for points in $(\overline{\mathbb{Q}}^*)^2$.

The ring of integers of an algebraic number field K is denoted by $\mathcal{O}_K$ and the set of places of K by M_K. For $v \in M_K$ we define

$$N(v) := 2 \text{ if } v \text{ is infinite}, \quad N(v) := N(\wp_v) \text{ if } v \text{ is finite}, \tag{2.1}$$

where $\wp_v$ is the prime ideal of $\mathcal{O}_K$ corresponding to v and $N(\wp_v) = \#\big(\mathcal{O}_K/\wp_v\big)$ is the norm of $\wp_v$.

Finally, we define $\log^* a := \max(1, \log a)$ for $a > 0$ and $\log^* 0 := 1$.

We consider again the equation

$$a_1x_1 + a_2x_2 = 1 \quad \text{in } (x_1, x_2) \in \Gamma \tag{1.1}$$

where $a_1, a_2 \in \overline{\mathbb{Q}}^*$ and where Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^2$ of rank > 0. Let $\mathbf{w}_1 = (\xi_1, \eta_1), \ldots, \mathbf{w}_r = (\xi_r, \eta_r)$ be a system of generators of $\Gamma/\Gamma_{\text{tors}}$ which is not necessarily a basis. Notice that every element of Γ can be expressed as $\zeta\mathbf{w}_1^{x_1} \cdots \mathbf{w}_r^{x_r}$ where $x_1, \ldots, x_r \in \mathbb{Z}$ and $\zeta \in \Gamma_{\text{tors}}$, i.e., the coordinates of ζ are roots of unity.

Define $K := \mathbb{Q}(\Gamma)$, i.e. the field generated by Γ over $\mathbb{Q}$. We do not require that $a_1, a_2 \in K$. Let S be the smallest set of places of K containing all infinite places such that $\mathbf{w}_1, \ldots, \mathbf{w}_r \in (\mathcal{O}_S^*)^2$, where $\mathcal{O}_S^*$ denotes the group of S-units in K. Put

$$\begin{aligned} &Q_\Gamma := h(\mathbf{w}_1) \cdots h(\mathbf{w}_r), \\ &d := [K : \mathbb{Q}], \quad s := \#S, \quad N := \max_{v \in S} N(v). \end{aligned}$$

Denote by t the maximum of the rank of the subgroup of $\overline{\mathbb{Q}}^*$ generated by $\xi_1, \ldots, \xi_r$ and the rank of the subgroup generated by $\eta_1, \ldots, \eta_r$. In view of $\operatorname{rank}\Gamma > 0$ we have $t > 0$. We define

$$\begin{aligned} &c_1(r, d, t) := 3(16ed)^{3(t+2)} \left(d(\log 3d)^3\right)^{r-t} (t/e)^t, \\ &A := 26c_1(r, d, t)s\frac{N}{\log N} Q_\Gamma \max\{\log(c_1(r, d, t)sN), \log^* Q_\Gamma\}, \qquad (2.2) \\ &H := \max(h(a_1), h(a_2), 1). \end{aligned}$$

Then our first result reads as follows:

Theorem 2.1. *For every solution $(x_1, x_2) \in \Gamma$ of (1.1) we have*

$$h(x_1, x_2) < AH. \tag{2.3}$$

We shall deduce Theorem 2.1 from Theorem 2.2 below. Let G be a finitely generated multiplicative subgroup of $\overline{\mathbb{Q}}^*$ of rank $t > 0$, and $\xi_1, \ldots, \xi_r$ a system of generators of G/G_{tors}. Let K be a number field containing G, and S a finite set of places of K containing the infinite places such that $G \subseteq \mathcal{O}_S^*$. We consider the equation

$$a_1x_1 + a_2x_2 = 1 \qquad \text{in } x_1 \in G,\ x_2 \in \mathcal{O}_S^*, \tag{2.4}$$

where $a_1, a_2 \in \overline{\mathbb{Q}}^*$. Let $d := [K : \mathbb{Q}]$. Let s be the cardinality of S, $N := \max_{v \in S} N(v)$ and put

$$Q_G := h(\xi_1) \cdots h(\xi_r).$$

Theorem 2.2. *Under the above assumptions and notation, every solution of (2.4) satisfies*

$$h(x_1) < c_1(r, d, t)s\frac{N}{\log N} Q_G H \log^*\left(\frac{Nh(x_1)}{H}\right) \tag{2.5}$$

and

$$\max(h(x_1), h(x_2)) < 6.5c_1(r,d,t)s\frac{N}{\log N}Q_G H\cdot \cdot \max\{\log(c_1(r,d,t)sN), \log^* Q_G\}, \tag{2.6}$$

where as before $H := \max(h(a_1), h(a_2), 1)$ *and* $c_1(r,d,t)$ *is the constant defined in (2.2).*

If in particular $r = t$ *and* $\{\xi_1, \dots, \xi_t\}$ *is a basis of* G/G_{tors}*, then, in (2.5) and (2.6) we can replace* $c_1(r,d,t)$ *by* $c_1(d,t) = 73(16ed)^{3t+5}$.

An important special case of equation (2.4) is when $G = \mathcal{O}_S^*$. Then (2.4) is called an S-unit equation. The first explicit upper bound for the height of the solutions of S-unit equations was given by Győry [12] by means of the theory of logarithmic forms. This bound was later improved by several authors. In this special case we have $t = s - 1$ and we may choose a basis $\{\xi_1, \dots, \xi_{s-1}\}$ for $\mathcal{O}_S^*/(\mathcal{O}_S^*)_{\rm tors}$ such that

$$h(\xi_1)\cdots h(\xi_{s-1}) \le \frac{((s-1)!)^2}{2^{s-2}d^{s-1}}R_S, \tag{2.7}$$

where R_S denotes the S-regulator in K (see e.g. Bugeaud and Győry [8]). The best known bounds for the solutions of S-unit equations are due to Bugeaud [7] and Győry and Yu [13]. As an immediate consequence, one can derive from our Theorem 2.2 and (2.7) an explicit bound for the solutions of S-unit equations which is comparable with the best known ones.

We now consider equations such as (1.1) but with solutions (x_1, x_2) from a larger set. We keep the notation introduced before Theorem 2.1.

The division group of Γ is given by

$$\overline{\Gamma} := \left\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists k \in \mathbb{Z}_{>0} \ \text{ with } \ \mathbf{x}^k \in \Gamma\right\}.$$

For any $\varepsilon > 0$ define the "cylinder" and "truncated cone" around $\overline{\Gamma}$ by

$$\overline{\Gamma}_\varepsilon := \left\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists \mathbf{y}, \mathbf{z} \text{ with } \mathbf{x} = \mathbf{y}\mathbf{z},\ \mathbf{y} \in \overline{\Gamma},\ \mathbf{z} \in (\overline{\mathbb{Q}}^*)^2,\ \ h(\mathbf{z}) < \varepsilon\right\} \tag{2.8}$$

and

$$C(\overline{\Gamma}, \varepsilon) := \Big\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^2 \mid \exists \mathbf{y}, \mathbf{z} \ \text{with } \mathbf{x} = \mathbf{y}\mathbf{z},\ \mathbf{y} \in \overline{\Gamma}, \mathbf{z} \in (\overline{\mathbb{Q}}^*)^2,\ h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))\Big\}, \tag{2.9}$$

respectively. The set $\overline{\Gamma}_\varepsilon$ was introduced by Poonen [15] and the set $C(\overline{\Gamma}, \varepsilon)$ by the second author [10] (both in a much more general context).

We emphasize that points from $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$ do not have their coordinates in a prescribed number field. So for effective results on Diophantine equations with solutions from $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$, we need an effective upper bound not only for the height of each solution, but also for the degree of the field which it generates. We fix $a_1, a_2 \in \overline{\mathbb{Q}}^*$ and define

$$K := \mathbb{Q}(\Gamma), \quad K_0 := \mathbb{Q}(a_1, a_2, \Gamma).$$

The quantities d, s, N, H and Q_Γ will have the same meaning as in Theorem 2.1 and A will be the constant defined in (2.2). Further, we put

$$h_0 := \max\{h(\xi_1), \dots, h_(\xi_r), h(\eta_1), \dots, h(\eta_r)\},$$

where $\mathbf{w}_i = (\xi_i, \eta_i)$ for $i = 1, \dots, r$ is the chosen system of generators for $\Gamma/\Gamma_{\text{tors}}$.

Consider now the equation

$$a_1 x_1 + a_2 x_2 = 1 \qquad \text{in } (x_1, x_2) \in \overline{\Gamma}_\varepsilon. \tag{2.10}$$

Theorem 2.3. *Suppose that (x_1, x_2) is a solution of (2.10) and that*

$$\varepsilon < 0.0225. \tag{2.11}$$

Then we have

$$h(x_1, x_2) \leq A h(a_1, a_2) + 3 r h_0 A \tag{2.12}$$

and

$$[K_0(x_1, x_2) : K_0] \leq 2. \tag{2.13}$$

The following consequence is immediate:

Corollary 2.4. *With the notation and assumptions from above, let (x_1, x_2) be a solution of*

$$a_1 x_1 + a_2 x_2 = 1 \qquad \text{in } (x_1, x_2) \in \overline{\Gamma}. \tag{2.14}$$

Then $h(x_1, x_2) \leq A h(a_1, a_2) + 3 r h_0 A$ and $[K_0(x_1, x_2) : K_0] \leq 2$.

Finally we consider the equation

$$a_1 x_1 + a_2 x_2 = 1 \qquad \text{in } (x_1, x_2) \in C(\overline{\Gamma}, \varepsilon). \tag{2.15}$$

Theorem 2.5. *Suppose that (x_1, x_2) is a solution of (2.15) and that*

$$\varepsilon < \frac{0.09}{8 A h(a_1, a_2) + 20 r h_0 A}. \tag{2.16}$$

Then we have

$$h(x_1, x_2) \leq 3 A h(a_1, a_2) + 5 r h_0 A \tag{2.17}$$

and

$$[K_0(x_1, x_2) : K_0] \leq 2. \tag{2.18}$$

The paper [1] gives explicit upper bounds for the number of solutions of (2.14), while from [11] one can deduce explicit upper bounds for the number of solutions for multivariate generalizations of (2.14), (2.10), (2.15). The sets $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ have been defined in the much more general context of semi-abelian varieties (see [15], [18]), and in [16], [17], Rémond proved quantitative analogues of the work of [11] for subvarieties of abelian varieties and subvarieties of tori. We mention that the results of [1], [11] and [15]–[18] are ineffective, in that they do not provide an algorithm to determine the solutions.

In a forthcoming work, to be written with Pontreau, we extend our effective results concerning (2.10) and (2.15) to equations $f(x_1, x_2) = 0$ in (x_1, x_2) from $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$, where $f \in \overline{\mathbb{Q}}[X_1, X_2]$ is an arbitrary polynomial. Further, we apply the results from the present paper to obtain effective and quantitative results for points in algebraic subvarieties of $\mathbb{G}_m^n$ from a restricted class.

3. Notation

In this section we collect the notation used in our paper. Let K be an algebraic number field of degree d. Denote by $\mathcal{O}_K$ its ring of integers and by M_K its set of places. For $v \in M_K$, we define an absolute value $|\cdot|_v$ as follows. If v is infinite and corresponds to $\sigma : K \to \mathbb{C}$, then we put $|x|_v = |\sigma(x)|^{d_v/d}$ for $x \in K$, where $d_v = 1$ or 2 according as $\sigma(K)$ is contained in $\mathbb{R}$ or not; if v is a finite place corresponding to a prime ideal $\wp$ of $\mathcal{O}_K$, then we put $|x|_v = N(\wp)^{-\operatorname{ord}_\wp x/d}$ for $x \in K \setminus \{0\}$, and $|0|_v = 0$. Here $N(\wp)$ denotes the norm of $\wp$, and $\operatorname{ord}_\wp x$ the exponent of $\wp$ in the prime ideal factorization of the principal fractional ideal (x). The absolute logarithmic height $h(x)$ of $x \in K$ is defined by

$$h(x) = \sum_{v \in M_K} \max(0, \log |x|_v). \tag{3.1}$$

More generally, if $x \in \overline{\mathbb{Q}}$ then choose an algebraic number field K such that $x \in K$ and define $h(x)$ by (3.1). This definition does not depend on the choice of K. Notice that $h(x) = 0$ if and only if $x \in \overline{\mathbb{Q}}^*_{\mathrm{tors}}$, where $\overline{\mathbb{Q}}^*_{\mathrm{tors}}$ is the group of roots of unity in $\overline{\mathbb{Q}}^*$.

Let S denote a finite subset of M_K containing all infinite places. Then $x \in K$ is called an S-integer if $|x|_v \le 1$ for all $v \in M_K \setminus S$. The S-integers form a ring in K, denoted by $\mathcal{O}_S$. Its unit group, denoted by $\mathcal{O}_S^*$, is called the group of S-units. It follows from (3.1) and the product formula that

$$h(x) = \frac{1}{2} \sum_{v \in S} |\log |x|_v| \quad \text{if } x \in \mathcal{O}_S^*. \tag{3.2}$$

For $\mathbf{x} = (x_1, x_2) \in (\overline{\mathbb{Q}}^*)^2$ define

$$h(\mathbf{x}) := h(x_1) + h(x_2).$$

Notice that for $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in (\overline{\mathbb{Q}}^*)^2$

$$h(\mathbf{xy}) \leq h(\mathbf{x}) + h(\mathbf{y}),$$
$$h(\mathbf{x}) = 0 \iff \mathbf{x} \in (\overline{\mathbb{Q}}^*_{\text{tors}})^2 \ ,$$
$$h(\mathbf{x}^a) = |a| h(\mathbf{x}) \ \text{ for } a \in \mathbb{Q},$$

where for $a \in \mathbb{Q}$ we define $\mathbf{x}^a := (x_1^a, x_2^a)$. The point $\mathbf{x}^a$ is determined only up to multiplication with elements from $(\overline{\mathbb{Q}}^*_{\text{tors}})^2$, but $h(\mathbf{x}^a)$ is well defined.

4. Diophantine approximation by elements from a finitely generated multiplicative group

Let again K be an algebraic number field of degree d, M_K the set of places on K, and G a finitely generated multiplicative subgroup of K^* of rank $t > 0$. Further, let $\{\xi_1, \ldots, \xi_r\}$ be a system of (not necessarily multiplicatively independent) generators of G such that $\xi_1, \ldots, \xi_r$ are not roots of unity. Put

$$Q_G := h(\xi_1) \cdots h(\xi_r).$$

Further, for any $v \in M_K$ let $N(v)$ be as in (2.1).

Theorem 2.2 and then subsequently Theorem 2.1 will be deduced from the following theorem.

Theorem 4.1. *Let $\alpha \in K^*$ with $\max(h(\alpha), 1) \leq H$ and let $v \in M_K$. Then for every $\xi \in G$ for which $\alpha\xi \neq 1$, we have*

$$\log |1 - \alpha\xi|_v > -c_2(r, d, t) \frac{N(v)}{\log N(v)} Q_G H \log^* \left(\frac{N(v) h(\xi)}{H} \right), \tag{4.1}$$

where

$$c_2(r, d, t) = (16ed)^{3(t+2)} \left(d (\log 3d)^3 \right)^{r-t} (t/e)^t.$$

If in particular $r = t$ and $\{\xi_1, \ldots, \xi_t\}$ is a basis of G/G_{tors}, then (4.1) holds with $c_2(d, t) = 36(16ed)^{3t+5} (\log^* d)^2$ instead of $c_2(r, d, t)$.

It should be observed that $c_2(d, t)$ does not contain a t^t factor.

The following theorem is in fact an immediate consequence of Theorem 4.1.

Theorem 4.2. *Let $\alpha \in K^*$ with $\max(h(\alpha), 1) \leq H$, let $v \in M_K$, and let $0 < \kappa \leq 1$. Then for every $\xi \in G$ with $\alpha\xi \neq 1$ and*

$$\log |1 - \alpha\xi|_v < -\kappa h(\xi) \tag{4.2}$$

we have

$$h(\xi) < (c_2(r, d, t)/\kappa) \frac{N(v)}{\log N(v)} Q_G H \log^* \left(\frac{N(v) h(\xi)}{H} \right) \tag{4.3}$$

and

$$h(\xi) < 6.4(c_2(r,d,t)/\kappa)\frac{N(v)}{\log N(v)}Q_G H\cdot \cdot \max\Big(\log\big((c_2(r,d,t)/\kappa)N(v)\big), \log^* Q_G\Big) \tag{4.4}$$

with the constant $c_2(r,d,t)$ *specified in Theorem 4.1.*

If in particular $r=t$ *and* $\{\xi_1,\dots,\xi_t\}$ *is a basis of* G/G_{tors}, *(4.3) and (4.4) hold with* $c_2(d,t)$ *instead of* $c_2(r,d,t)$.

We note that when applying Theorem 4.2 to equation (2.4), inequality (4.3) yields better bounds in Theorem 2.2 than (4.4).

The main tool in the proofs of Theorems 4.1 and 4.2 is the theory of logarithmic forms, more precisely Theorem C in Section 5. Bombieri [3] and Bombieri and Cohen [4], [5] have developed another effective method in Diophantine approximation, based on an extended version of the Thue-Siegel principle, the Dyson lemma and some geometry of numbers. Bugeaud [7], following their approach and combining it with estimates for linear forms in two and three logarithms, obtained sharper results than Bombieri and Cohen. Bugeaud deduced an explicit upper bound for $h(\xi)$ from the inequality

$$\log|1-\alpha\xi|_v < -\kappa h(\alpha\xi). \tag{4.5}$$

It is easy to check that apart from the trivial case $\min(h(\xi), h(\alpha\xi)) \le h(\alpha)$ when $h(\xi) \le 2H$ follows, we have

$$h(\xi)/2 \le h(\alpha\xi) \le 2h(\xi).$$

Hence, if ξ and $\alpha\xi$ are not roots of unity, (4.5) and (4.2) can be deduced from each other with κ replaced by $\kappa/2$. It follows from Bugeaud's theorem that if (4.2) holds with $0<\kappa\le 1$, then

$$h(\xi) \le \begin{cases} 10T\max(H,T) & \text{if } v \text{ is infinite,} \\ 8c_3(d,\kappa)T\max(H,40T) & \text{if } v \text{ is finite,} \end{cases} \tag{4.6}$$

where

$$c_3(d,\kappa) = \begin{cases} 8\cdot 10^{19}\,(d^4(\log 3d)^7/\kappa)\log^*(2d/\kappa) & \text{if } v \text{ is infinite,} \\ 8\cdot 10^{6}\,(d^5/\kappa)\,(\log^*(2d/\kappa))^2 & \text{if } v \text{ is finite,} \end{cases}$$

and

$$T = (2rc_3(d,\kappa))^r N(v)(\log N(v))Q_G.$$

It is easily seen that the bound in (4.4) has a better dependence on each parameter than the bound in (4.6), except possibly Q_G and H. In fact, the bound in (4.6) is smaller than that in (4.4) precisely when both Q_G and $H\cdot\log Q_G/Q_G$ are large relative to $N(v)$, d, r, t and κ, and in that case, the bound (4.6) is at most a factor $\log Q_G$ better than (4.4).

Finally it should be observed, that in contrast with (4.6), our bound in (4.4) contains only the factor t^t, but not r^r. Furthermore, if in particular $r = t$ and $\{\xi_1, \ldots, \xi_t\}$ is a basis of G/G_{tors}, there is no factor t^t at all in our bound in (4.4). We note that in the general case, even the factor t^t has been removed from (4.4) by the second and third authors in a forthcoming work.

We should remark here, that we obtained Theorem 4.1 and its consequences by applying lower bounds for linear forms in logarithms in an arbitrary number of variables, whereas Bugeaud used lower bounds for linear forms in two or three logarithms. We used a lemma from the Geometry of Numbers to replace a dependence on the coefficients of the linear form in logarithms associated with (4.1) by one on the height of ξ.

5. Proofs of Theorems 4.1 and 4.2

We need several auxiliary results.

Keeping the notation of Section 4, let K be an algebraic number field of degree d and assume that it is embedded in $\mathbb{C}$. Let

$$\Lambda = \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1, \tag{5.1}$$

where $\alpha_1, \ldots, \alpha_n$ are n (≥ 2) non-zero elements of K, and $b_1, \ldots, b_n$ are rational integers, not all zero. Put

$$B^* = \max\{|b_1|, \ldots, |b_n|\}.$$

Let $A_1, \ldots, A_n$ be reals with

$$A_i \geq \max\{dh(\alpha_i), \pi\} \quad (i = 1, \ldots, n). \tag{5.2}$$

Theorem A. (Matveev [14]) *Let $n \geq 2$. Suppose that $\Lambda \neq 0$, $b_n = \pm 1$, and let B be a real number with*

$$B \geq \max\left\{B^*, 2e\max\left(\frac{n\pi}{\sqrt{2}}, A_1, \ldots, A_{n-1}\right)A_n\right\}. \tag{5.3}$$

Then we have

$$\log|\Lambda| > -c_1(n,d)A_1 \cdots A_n \log(B/(\sqrt{2}A_n)), \tag{5.4}$$

where

$$c_1(n,d) = \min\left\{1.451(30\sqrt{2})^{n+4}(n+1)^{5.5}, \pi 2^{6.5n+27}\right\} d^2\log(ed).$$

Proof. This is a consequence of Corollary 2.3 of Matveev [14]; see Proposition 4 in Győry and Yu [13]. □

Let B and B_n be real numbers satisfying

$$B \geq \max\{|b_1|, \ldots, |b_n|\}, \quad B \geq B_n \geq |b_n|. \tag{5.5}$$

Denote by $\wp$ a prime ideal of the ring of integers $\mathcal{O}_K$ and let $e_\wp$ and $f_\wp$ be the ramification index and the residue class degree of $\wp$, respectively. Thus $N(\wp) = p^{f_\wp}$, where p is the prime number below $\wp$.

Theorem B. (Yu [22]) *Let $n \geq 2$. Assume that $\operatorname{ord}_p b_n \leq \operatorname{ord}_p b_i$ for $i = 1, \ldots, n$, and set*

$$h'_i = \max\{h(\alpha_i), 1/(16e^2d^2)\}, \quad i = 1, \ldots, n.$$

If $\Lambda \neq 0$, then for any real δ with $0 < \delta \leq 1/2$ we have

$$\begin{aligned} \operatorname{ord}_\wp \Lambda \leq c_2(n,d) e_\wp^n \frac{N(\wp)}{(\log N(\wp))^2} \cdot \\ \cdot \max\left\{ h'_1 \cdots h'_n \log(M\delta^{-1}), \frac{\delta B}{B_n c_3(n,d)} \right\}, \end{aligned} \tag{5.6}$$

where

$$\begin{aligned} c_2(n,d) =& (16ed)^{2(n+1)} n^{3/2} \log(2nd) \log(2d), \\ c_3(n,d) =& (2d)^{2n+1} \log(2d) \log^3(3d), \end{aligned}$$

and

$$M = B_n c_4(n,d) N(\wp)^{n+1} h'_1 \ldots h'_{n-1},$$

with

$$c_4(n,d) = 2e^{(n+1)(6n+5)} d^{3n} \log(2d).$$

Proof. This is the second consequence of the Main Theorem in Yu [22]. □

The following theorem is a consequence of Theorems A and B.

Theorem C. *Let $n \geq 2$ and $v \in M_K$. Suppose that in (5.1) we have $\Lambda \neq 0$, $b_n = \pm 1$ and that $\alpha_1, \ldots, \alpha_{n-1}$ are not roots of unity. Let*

$$Q_\alpha := h(\alpha_1) \cdots h(\alpha_{n-1}), \quad H := \max(h(\alpha_n), 1).$$

If

$$B \geq \max(|b_1|, \ldots, |b_{n-1}|, 2e(3d)^{2n} Q_\alpha H), \tag{5.7}$$

then

$$\log |\Lambda|_v > -c_5(n,d) \frac{N(v)}{\log N(v)} Q_\alpha H \log^* \left(\frac{BN(v)}{H} \right), \tag{5.8}$$

where

$$c_5(n,d) = \lambda (16ed)^{3n+2} (\log^* d)^2,$$

where $\lambda = 1$ or 12 according as $n \geq 3$ or $n = 2$.

To deduce Theorem C from Theorems A and B, we need the following.

Lemma 5.1. (Voutier [21]). *Suppose that α is a non-zero algebraic number of degree d which is not a root of unity. Then*

$$dh(\alpha) \geq \begin{cases} \log 2 & \text{if } d = 1, \\ 2/(\log 3d)^3 & \text{if } d \geq 2. \end{cases} \tag{5.9}$$

Proof. For $d \geq 2$ this is due to Voutier [21]. He showed also that for $d \geq 2$ this lower bound may be replaced by $(1/4)(\log\log d/\log d)^3$. □

Proof of Theorem C. First assume that v is infinite. We apply Theorem A with $A_i = \max\{dh(\alpha_i), \pi\}$ for $i = 1, \ldots, n$. Then using (5.9), it is easy to see that

$$A_1 \cdots A_n \leq (2.52d)^{2n} Q_\alpha H.$$

Further, we have $\sqrt{2}A_n > H/N(v)$ and

$$2e \max\left\{\frac{n\pi}{\sqrt{2}}, A_1, \ldots, A_{n-1}\right\} A_n \leq 2e(3d)^{2n} Q_\alpha H.$$

Now (5.7) implies (5.3), and (5.8) follows from the inequality (5.4) of Theorem A.

Next assume that v is finite. Keeping the notation of Theorem B and using again (5.9), we infer that

$$h_i' = h(\alpha_i) \text{ for } i = 1, \ldots, n-1 \qquad h_n' = h(\alpha_n).$$

Hence $h_n' = H$ if $h(\alpha_n) \geq 1$ and $H = 1$ otherwise. Choosing $\delta = h_1' \cdots h_n'/B$ and $B_n = 1$ in Theorem B, (5.7) implies that $\delta \leq \frac{1}{2}$. Using the fact that $|\Lambda|_v = N(\wp)^{-\operatorname{ord}_\wp \Lambda}$, after some computation (5.8) follows from (5.6) of Theorem B. □

Theorem 4.1 will be proved by combining Theorem C with the following result from the geometry of numbers. Let t be a positive integer. A convex distance function on $\mathbb{R}^t$ is a function $f := \mathbb{R}^t \to \mathbb{R}_{\geq 0}$ such that

$$\begin{aligned} &f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^t, \\ &f(\lambda \mathbf{x}) = |\lambda| f(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathbb{R}^t, \lambda \in \mathbb{R}, \\ &f(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}. \end{aligned}$$

Lemma 5.2. *Let f be a convex distance function on $\mathbb{R}^t$. Let $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_t\}$ be any basis of $\mathbb{Z}^t$ for which the product $f(\boldsymbol{a}_1) \cdots f(\boldsymbol{a}_t)$ is minimal. Let $\boldsymbol{x} \in \mathbb{Z}^t$ and suppose that $\boldsymbol{x} = b_1\boldsymbol{a}_1 + \cdots + b_t\boldsymbol{a}_t$ with $b_1, \ldots, b_t \in \mathbb{Z}$. Then*

$$\max(|b_1| f(\boldsymbol{a}_1), \ldots, |b_t| f(\boldsymbol{a}_t)) \leq c_6(t) f(\boldsymbol{x}), \tag{5.10}$$

where $c_6(t) = t^{2t}$.

Remark. Schlickewei [19] proved that there exists a basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_t\}$ of $\mathbb{Z}^t$ satisfying (5.10) with 4^t instead of $c_6(t)$, but it is not clear whether for this basis, the product $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is minimal. In our proof of Theorem 4.1, the minimality of $f(\mathbf{a}_1) \cdots f(\mathbf{a}_t)$ is crucial, while an improvement of $c_6(t)$ would have only little influence on the final result.

Proof. Let $C = \{\mathbf{x} \in \mathbb{R}^t : f(\mathbf{x}) \leq 1\}$. This is a compact, convex body which is symmetric around $\mathbf{0}$. Let $\lambda_1, \ldots, \lambda_t$ denote the successive minima of C

with respect to the lattice $\mathbb{Z}^t$. Since $\lambda_1 \le \cdots \le \lambda_t$, it follows from a result of Mahler (see e.g. Cassels [9], pp. 135-136, Lemma 8) that there exists a basis $\mathbf{y}_1, \dots, \mathbf{y}_t$ of $\mathbb{Z}^t$ such that $f(\mathbf{y}_i) \le \max(1, i/2)\lambda_i$ for $i = 1, \dots, t$. Together with Minkowski's theorem on successive minima, this gives

$$f(\mathbf{a}_1)\cdots f(\mathbf{a}_t) \le f(\mathbf{y}_1)\cdots f(\mathbf{y}_t) \le 2t! \cdot \mathrm{Vol}(C)^{-1}, \tag{5.11}$$

where $\mathrm{Vol}(C)$ denotes the volume of C.

By Jordan's theorem or John's Lemma (see e.g. Schmidt [20], pp. 87–89) there is a t-dimensional ellipsoid E in $\mathbb{R}^t$ such that $E \subseteq C \subseteq (\sqrt{t})E$. Further, there is a $t \times t$ real non-singular matrix A such that $E = \{\mathbf{x} \in \mathbb{R}^t : \|A\mathbf{x}\| \le 1\}$, where $\|\cdot\|$ denotes the Euclidean norm. Thus

$$\frac{1}{\sqrt{t}}\|A\mathbf{x}\| \le f(\mathbf{x}) \le \|A\mathbf{x}\| \quad \text{for} \ \ \mathbf{x} \in \mathbb{R}^t. \tag{5.12}$$

Consequently,

$$V(t)|\det(A)|^{-1} \le \mathrm{Vol}(C) \le t^{t/2}V(t)|\det(A)|^{-1}, \tag{5.13}$$

where $V(t)$ denotes the volume of the t-dimensional unit ball.

Now let $\mathbf{x} = b_1\mathbf{a}_1 + \cdots + b_t\mathbf{a}_t$ with $b_1, \dots, b_t \in \mathbb{Z}$. Then $A\mathbf{x} = b_1(A\mathbf{a}_1) + \cdots + b_t(A\mathbf{a}_t)$. Let B be the matrix with columns $A\mathbf{a}_1, \dots, A\mathbf{a}_t$. Since $|\det(\mathbf{a}_1, \dots, \mathbf{a}_t)| = 1$, we have $|\det(B)| = |\det(A)|$. By this fact, Cramer's rule and Hadamard's inequality, we have for $i = 1, \dots, t$,

$$\begin{aligned} |b_i| &= |\det(A\mathbf{a}_1, \dots, A\mathbf{a}_{i-1}, A\mathbf{x}, A\mathbf{a}_{i+1}, \dots A\mathbf{a}_t)| \Big/ |\det(B)| \\ &\le \|A\mathbf{a}_1\|\cdots\|A\mathbf{a}_{i-1}\| \cdot \|A\mathbf{x}\| \cdot \|A\mathbf{a}_{i+1}\|\cdots\|A\mathbf{a}_t\| \Big/ |\det(A)|. \end{aligned}$$

Together with (5.12), (5.11) and (5.13), this implies

$$\begin{aligned} |b_i|f(\mathbf{a}_i) &\le t^{(t-1)/2}\big(f(\mathbf{a}_1)\cdots f(\mathbf{a}_t)/|\det(A)|\big)f(\mathbf{x}) \\ &\le t^{(t-1)/2} \cdot 2t!V(t)^{-1}f(\mathbf{x}) \qquad\qquad \text{for} \ \ i = 1, \dots, t. \end{aligned}$$

By inserting $V(t) = \pi^{t/2}/(t/2)!$ if t is even and $V(t) = \pi^{(t-1)/2} \Big/ \big(\frac{1}{2} \cdot \frac{3}{2} \cdots \frac{t}{2}\big)$ if t is odd, we get the bound in (5.10). □

Lemma 5.3. *Let G be a finitely generated multiplicative subgroup of K^* of rank $t > 0$. Let $\delta_1, \dots, \delta_t \in G$ be multiplicatively independent such that $h(\delta_1) \le \cdots \le h(\delta_t)$. Then G/G_{tors} has a basis $\{\gamma_1, \dots, \gamma_t\}$ such that*

$$h(\gamma_i) \le \max(1, i/2)h(\delta_i) \quad \textit{for} \ \ i = 1, \dots, t. \tag{5.14}$$

Proof. Let $\{\rho_1, \dots, \rho_t\}$ be a basis for G/G_{tors}. Then we can write

$$\delta_i = \zeta_i\rho_1^{b_{i1}}\cdots\rho_t^{b_{it}}, \quad i = 1, \dots, t,$$

where $\zeta_i \in G_{\mathrm{tors}}$, and

$$\mathbf{b}_1 = (b_{11}, \dots, b_{1t}), \dots, \mathbf{b}_t = (b_{t1}, \dots, b_{tt})$$

are linearly independent vectors in $\mathbb{Z}^t$.

Let $S \subset M_K$ be minimal such that S contains all infinite places and $G \subseteq \mathcal{O}_S^*$. We define

$$f(\mathbf{x}) := \frac{1}{2} \sum_{v \in S} \left| x_1 \log |\rho_1|_v + \cdots + x_t \log |\rho_t|_v \right|,$$

where $\mathbf{x} = (x_1, \ldots, x_t) \in \mathbb{R}^t$. This is a convex distance function. Further, by (3.2) we have

$$f(\mathbf{b}_i) = h(\delta_i) \quad \text{for} \ \ i = 1, \ldots, t. \tag{5.15}$$

Using again Mahler's result mentioned above, we infer that there is a basis $\mathbf{a}_i = (a_{i1}, \ldots, a_{it})$ $(i = 1, \ldots, t)$ of $\mathbb{Z}^t$ for which

$$f(\mathbf{a}_i) \leq \max(1, i/2) f(\mathbf{b}_i) \quad \text{for} \ \ i = 1, \ldots, t. \tag{5.16}$$

Putting $\gamma_i = \rho_1^{a_{i1}} \cdots \rho_t^{a_{it}}$ for $i = 1, \ldots, t$, we infer that $\{\gamma_1, \ldots, \gamma_t\}$ is a basis for G/G_{tors} , which in view of (5.15), (5.16) satisfies (5.14). □

We first prove Theorem 4.1 and then Theorem 4.2.

Proof of Theorem 4.1. Since G has rank $t > 0$, there are t multiplicatively independent elements among the generators $\xi_1, \ldots, \xi_r$, say $\xi_1, \ldots, \xi_t$. Then by Lemma 5.1

$$h(\xi_1) \cdots h(\xi_t) \leq c_7(d)^{r-t} Q_G, \tag{5.17}$$

where $c_7(d) = \frac{d}{2}(\log 3d)^3$ if $d \geq 2$ and $c_7(d) = (\log 2)^{-1}$ if $d = 1$. Let $\delta_1, \ldots, \delta_t$ be multiplicatively independent elements of G such that $h(\delta_1) \cdots h(\delta_t)$ is minimal. Then

$$h(\delta_1) \cdots h(\delta_t) \leq h(\xi_1) \cdots h(\xi_t). \tag{5.18}$$

Further, by Lemma 5.3, G/G_{tors} has a basis $\{\gamma_1, \ldots, \gamma_t\}$ such that

$$h(\gamma_1) \cdots h(\gamma_t) \leq c_8(t) h(\delta_1) \cdots h(\delta_t), \tag{5.19}$$

with $c_8(t) := t!/2^{t-1}$. We may assume that $\{\gamma_1, \ldots, \gamma_t\}$ is such a basis for which $h(\gamma_1) \cdots h(\gamma_t)$ is minimal.

For $\xi \in G$, we can write

$$\xi = \zeta \gamma_1^{b_1} \cdots \gamma_t^{b_t}, \tag{5.20}$$

where $\zeta \in G_{\text{tors}}$ and $\mathbf{b} = (b_1, \ldots, b_t) \in \mathbb{Z}^t$. As in the proof of Lemma 5.3, consider the following convex distance function on $\mathbb{R}^t$:

$$f(\mathbf{x}) := \frac{1}{2} \sum_{v \in S} \left| x_1 \log |\gamma_1|_v + \cdots + x_t \log |\gamma_t|_v \right|,$$

where $\mathbf{x} = (x_1, \ldots, x_t) \in \mathbb{R}^t$ and S is the same as in the proof of Lemma 5.3. Then $f(\mathbf{b}) = h(\xi)$. Consider the standard basis $\mathbf{a}_1 = (1, 0, \ldots, 0)$, $\mathbf{a}_2 = (0, 1, 0, \ldots, 0)$, …, $\mathbf{a}_t = (0, \ldots, 0, 1)$ in $\mathbb{Z}^t$. Then

$$f(\mathbf{a}_i) = h(\gamma_i) \quad \text{for} \ \ i = 1, \ldots, t,$$

and $f(\mathbf{a}_1)\cdots f(\mathbf{a}_t)$ is minimal among the bases of $\mathbb{Z}^t$.

We can now apply Lemma 5.2 to this basis $\mathbf{a}_1,\dots,\mathbf{a}_t$, and infer that

$$|b_i|h(\gamma_i) = |b_i|f(\mathbf{a}_i) \le c_6(t)f(\mathbf{b}) = c_6(t)h(\xi), \quad i=1,\dots,t.$$

Together with Lemma 5.1 this gives

$$\max(|b_1|,\dots,|b_t|) \le c_6(t)c_7(d)h(\xi). \tag{5.21}$$

We apply now Theorem C with $v\in S$ and with

$$\Lambda = 1-\alpha\xi = 1-\alpha'\gamma_1^{b_1}\cdots\gamma_t^{b_t},$$

where $\alpha'=\zeta\alpha$. Let $Q_\gamma := h(\gamma_1)\cdots h(\gamma_t)$. First assume that

$$c_6(t)c_7(d)h(\xi) \ge 2e(3d)^{2(t+1)}Q_\gamma H. \tag{5.22}$$

Further suppose that

$$h(\xi) \ge \left(c_6(t)c_7(d)\right)^{1/2} H. \tag{5.23}$$

Then putting $B=c_6(t)c_7(d)h(\xi)$, it follows that

$$\log^*\left(\frac{BN(v)}{H}\right) \le 3\log\left(\frac{h(\xi)N(v)}{H}\right). \tag{5.24}$$

Together with (5.17), (5.18), (5.19) and (5.24), Theorem C gives (4.1) after some computation.

Consider now the case when at least one of (5.22) and (5.23) does not hold. We cover this remaining case by assuming that

$$h(\xi) < \frac{1}{2}c_2(r,d,t)Q_G H$$

with the $c_2(r,d,t)$ occurring in Theorem 4.1. By the product formula and Liouville's inequality we get

$$\begin{aligned}|1-\alpha\xi|_v &= \prod_{\substack{w\in M_K\\ w\neq v}} |1-\alpha\xi|_w^{-1} \ge \frac{1}{2}\prod_{\substack{w\in M_K\\ w\neq v}} \max(1,|\alpha\xi|_w)^{-1}\\ &\ge \frac{1}{2}\exp(-h(\alpha\xi)) \ge \frac{1}{2}\exp\left(-\left(H+\frac{1}{2}c_2(r,d,t)Q_G H\right)\right),\end{aligned}$$

whence (4.1) follows again.

Finally, assume that $r=t$ and that $\{\xi_1,\dots,\xi_t\}$ is a basis of $G/G_{\rm tors}$. We may assume without loss of generality that $Q_G = h(\xi_1)\cdots h(\xi_t)$ is minimal among all bases of $G/G_{\rm tors}$. Then in our above proof we can choose $\gamma_i=\xi_i$ for $i=1,\dots,t$ and we do not need $\delta_1,\dots,\delta_t$. This simplification in the proof gives (4.1) with $c_2(d,t)$ in place of $c_2(r,d,t)$. □

Proof of Theorem 4.2. Together with the estimate (4.1) of Theorem 4.1, (4.2) gives (4.3), and then (4.4) easily follows. □

6. Proofs of Theorems 2.1, 2.2, 2.3 and 2.5

Taking as a starting point Theorem 4.2, we first deduce Theorem 2.2, then Theorem 2.1, and from the latter Theorems 2.3 and 2.5.

Proof of Theorem 2.2. First suppose that $a_1, a_2 \in K$. Let (x_1, x_2) be a solution of (2.4). Then (2.4) gives

$$h(x_1) \leq 3H + h(x_2) + \log 2. \tag{6.1}$$

First assume that $h(x_2) < 4 \cdot 10^2 sH$. Then (6.1) gives $h(x_1) \leq 404sH$, whence $h(x_1)N/H \leq 404sN$. Using now the fact that the function $X/\log X$ is monotone increasing for $X > e$, (2.5) and (2.6) easily follow.

Now assume that

$$h(x_2) \geq 4 \cdot 10^2 sH. \tag{6.2}$$

Choose $v \in S$ for which $|x_2|_v$ is minimal. Then we infer from (2.4) that

$$\log |1 - a_1 x_1|_v = \log |a_2 x_2|_v \leq -\frac{1}{s} h(x_2) + H. \tag{6.3}$$

Further, it follows from (6.1) and (6.2) that $h(x_1) \leq 1.01h(x_2)$. Hence we get from (6.2) and (6.3) that

$$\log |1 - a_1 x_1|_v < -\kappa h(x_1)$$

with the choice $\kappa = 1/(2.02s)$. By applying the estimate (4.3) of Theorem 4.2 we deduce (2.5) and subsequently we get for $h(x_1)$ the upper bound in (2.6) with 6.5 replaced by 6.4. Finally, it follows from (2.4) that $h(x_2) \leq 3H + h(x_1) + \log 2$, so we obtain (2.6) for $h(x_2)$ as well.

Now suppose that $(a_1, a_2) \notin (K^*)^2$. Then we choose a nontrivial embedding σ of the extension K_0/K into $\mathbb{C}$, where $K_0 = K(a_1, a_2)$. Then equation (2.4) leads to

$$\sigma(a_1)x_1 + \sigma(a_2)x_2 = 1. \tag{6.4}$$

Now expressing x_1 and x_2 by Cramer's rule from the system consisting of (2.4) and (6.4) we get an estimate for $h(x_1)$ and $h(x_2)$ which is much sharper than (2.5) and (2.6). □

Proof of Theorem 2.1. Suppose that $\xi_1, \dots, \xi_r$ generate a multiplicative subgroup, say G, of $\overline{\mathbb{Q}}^*$ of rank $t > 0$. Clearly G is contained in K^*. We may assume that $\xi_1, \dots, \xi_{r'}$ are not roots of unity. Then $t \leq r' \leq r$ and $\xi_1, \dots, \xi_{r'}$ is a system of generators of G/G_{tors}. By the assumption made on $\mathbf{w}_1, \dots, \mathbf{w}_r$, $\eta_{r'+1}, \dots, \eta_r$ are not roots of unity. Put

$$Q_G := h(\xi_1) \cdots h(\xi_{r'}).$$

Using Lemma 5.1 we infer that

$$Q_G \leq c_7(d)^{r-r'} Q_\Gamma, \tag{6.5}$$

where $c_7(d) = (1/2)d(\log 3d)^3$ if $d \geq 2$ and $c_7(d) = (\log 2)^{-1}$ if $d = 1$.

Let (x_1, x_2) be a solution of (1.1). Then $x_1 \in G$ and $x_2 \in \mathcal{O}_S^*$. We can now apply Theorem 2.2 to this solution and we obtain (2.6) with r replaced by r'. Using $r' \leq r$, (6.5) and

$$h(x_1, x_2) \leq h(x_1) + h(x_2),$$

(2.3) easily follows from (2.6). $\square$

In the proofs of Theorems 2.3 and 2.5 we need the following lemma.

Lemma 6.1. (Beukers and Zagier). *Let $(b_1, b_2) \in (\overline{\mathbb{Q}}^*)^2$, and let (x_{i1}, x_{i2}) $(i = 1, 2, 3)$ be points in $(\overline{\mathbb{Q}}^*)^2$ with $b_1 x_{i1} + b_2 x_{i2} = 1$ for $i = 1, 2, 3$. Then we have*

$$\sum_{i=1}^{3} h(x_{i1}, x_{i2}) \geq 0.09. \tag{6.6}$$

Proof. By Corollary 2.4 in [2] we have $\sum_{i=1}^{3} h(x_{i1}, x_{i2}) \geq \log \rho$, where ρ denotes the real root of $\rho^{-6} + \frac{1}{2}\rho^{-2} = 1$ which is larger than 1. We have $\log \rho \geq 0.09$. $\square$

The proofs of Theorems 2.3 and 2.5 are very similar. We work out the proof of Theorem 2.5 in detail, and then indicate which changes have to be made to obtain Theorem 2.3.

Proof of Theorem 2.5. Fix a solution (x_1, x_2) of equation (2.15). Since $(x_1, x_2) \in C(\overline{\Gamma}, \varepsilon)$ we can write

$$\begin{aligned} &(x_1, x_2) = (y_1, y_2)(z_1, z_2) \quad \text{with} \\ &(y_1, y_2) \in \overline{\Gamma}, \quad h(z_1, z_2) < \varepsilon(1 + h(y_1, y_2)). \end{aligned} \tag{6.7}$$

Further, we can write

$$\begin{aligned} &(y_1, y_2) = (y_1', y_2')(w_1, w_2) \quad \text{with} \\ &(y_1', y_2') \in \Gamma, \\ &(w_1, w_2) = \prod_{i=1}^{r} (\xi_i, \eta_i)^{c_i} \text{ with } c_i \in \mathbb{Q},\ |c_i| \leq \frac{1}{2} \quad (i = 1, \ldots, r). \end{aligned} \tag{6.8}$$

(Note that w_1, w_2 are defined up to roots of unity.) Thus we have

$$h(w_1, w_2) \leq \sum_{i=1}^{r} |c_i| h(\xi_i, \eta_i) \leq r h_0. \tag{6.9}$$

Write

$$(a_1', a_2') := (a_1, a_2)(w_1, w_2)(z_1, z_2). \tag{6.10}$$

Then by (6.9), (6.7),

$$h(a_1', a_2') \leq h(a_1, a_2) + r h_0 + \varepsilon(1 + h(y_1, y_2))$$

which leads to

$$h(a_1', a_2') \le h(a_1, a_2) + rh_0 + \varepsilon(1 + h(y_1', y_2') + rh_0). \tag{6.11}$$

Further, equation (2.15) can be written in the form

$$a_1'y_1' + a_2'y_2' = 1 \qquad \text{in } (y_1', y_2') \in \Gamma. \tag{6.12}$$

Using Theorem 2.1 we get

$$h(y_1', y_2') \le A \max\{h(a_1', a_2'), 1\} \tag{6.13}$$

where A is the constant defined in (2.2). Notice that this constant does not depend on the field generated by a_1', a_2'. Further, using (6.11) we get

$$h(y_1', y_2') \le Ah(a_1, a_2) + rh_0A + \varepsilon A + \varepsilon Ah(y_1', y_2') + rh_0\varepsilon A.$$

Since in view of (2.16) we have $\varepsilon < \frac{1}{2A}$ we obtain

$$h(y_1', y_2') \le 2Ah(a_1, a_2) + (1 + 2Arh_0 + rh_0). \tag{6.14}$$

Now by

$$h(y_1, y_2) \le h(y_1', y_2') + h(w_1, w_2) \le 2Ah(a_1, a_2) + (1 + 2Arh_0 + 2rh_0) \tag{6.15}$$

and

$$h(x_1, x_2) \le h(y_1, y_2) + \varepsilon(1 + h(y_1, y_2)) \le (\varepsilon + 1)h(y_1, y_2) + \varepsilon$$

we get

$$h(x_1, x_2) \le 3Ah(a_1, a_2) + 5Arh_0 \tag{6.16}$$

which proves assertion (2.17) of our Theorem 2.5.

Now we have to prove the explicit upper bound on $[K_0(x_1, x_2) : K_0]$, where (x_1, x_2) is any solution of (2.15) and K_0 is the field generated by Γ, a_1, a_2. Let us fix such a solution. Choose (y_1, y_2), (z_1, z_2) as in (6.7) and then (y_1', y_2'), (w_1, w_2) as in (6.8). Finally, define (a_1', a_2') by (6.10). Define the field $L := K_0(a_1', a_2')$. We first prove that $[L : K_0] \le 2$.

Assume that this is false, that is, $[L : K_0] \ge 3$. Then there are at least 3 distinct embeddings of L to $\mathbb{C}$ which leave fixed the field K_0, call them $\sigma_1, \sigma_2, \sigma_3$. We consider again equation (6.12). Since $(y_1', y_2') \in \Gamma \subset (K_0^*)^2$ we have

$$\sigma_i(a_1')y_1' + \sigma_i(a_2')y_2' = 1 \quad \text{for } i = 1, 2, 3.$$

This means that the equation

$$(a_1'y_1')X + (a_2'y_2')Y = 1 \qquad \text{in } (X, Y) \in (\overline{\mathbb{Q}}^*)^2$$

has at least 3 distinct solutions, namely $\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right)$ $(i = 1, 2, 3)$. Now using Lemma 6.1 we know that

$$\sum_{i=1}^{3} h\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right) \ge 0.09. \tag{6.17}$$

On the other hand by (6.10) we have for any embedding $\sigma : L \to \mathbb{C}$,

$$\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) = \left(\frac{\sigma(a_1)}{a_1}, \frac{\sigma(a_2)}{a_2}\right)\left(\frac{\sigma(w_1)}{w_1}, \frac{\sigma(w_2)}{w_2}\right)\left(\frac{\sigma(z_1)}{z_1}, \frac{\sigma(z_2)}{z_2}\right).$$

However, $a_1, a_2 \in K_0$. Further, $(w_1, w_2) \in \overline{\Gamma}$, hence there exists a positive integer m such that $(w_1, w_2)^m \in \Gamma$. This means that $\left(\frac{\sigma(w_1)}{w_1}\right)^m = 1$ and $\left(\frac{\sigma(w_2)}{w_2}\right)^m = 1$. Thus we see that there exist roots of unity ζ_1, ζ_2 such that $\sigma(w_1) = \zeta_1 w_1$ and $\sigma(w_2) = \zeta_2 w_2$. So

$$\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) = (\zeta_1, \zeta_2)\left(\frac{\sigma(z_1)}{z_1}, \frac{\sigma(z_2)}{z_2}\right)$$

and together with $(x_1, x_2) \in C(\overline{\Gamma}, \varepsilon)$ and (6.15),

$$\begin{aligned} h\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) &\le 2h(z_1, z_2) \le 2\varepsilon(1 + h(y_1, y_2))) \\ &\le 2\varepsilon(2Ah(a_1, a_2) + (1 + 2Arh_0 + 2rh_0)) \\ &\le 2\varepsilon(2Ah(a_1, a_2) + 5Arh_0). \end{aligned} \tag{6.18}$$

This shows that

$$\sum_{i=1}^{3} h\left(\frac{\sigma_i(a_1')}{a_1'}, \frac{\sigma_i(a_2')}{a_2'}\right) < 4\varepsilon(2Ah(a_1, a_2) + 5Arh_0) < 0.09, \tag{6.19}$$

and this contradicts (6.17). Thus, we have proved that $[L : K_0] \le 2$.

In view of $y_1', y_2' \in K_0$ this shows that $[K_0(a_1'y_1', a_2'y_2') : K_0] \le 2$, consequently, $[K_0(a_1x_1, a_2x_2) : K_0] \le 2$ and finally, using that $a_1, a_2 \in K_0$ we get $[K_0(x_1, x_2) : K_0] \le 2$.

□

Proof of Theorem 2.3. The proof of Theorem 2.3 is completely similar to the proof of Theorem 2.5. The only difference is that the estimate (6.7) for $h(z_1, z_2)$ has to be replaced by $h(z_1, z_2) < \varepsilon$. This slightly modifies the estimates in the proof of Theorem 2.5 and instead of (6.14) we get

$$h(y_1', y_2') \le Ah(a_1, a_2) + A(\varepsilon + rh_0).$$

This in turn (instead of (6.16)) leads to the estimate

$$h(x_1, x_2) \le Ah(a_1, a_2) + 3Arh_0,$$

and this proves the assertion (2.12). In order to prove (2.13) we proceed in precisely the same way as we did for proving (2.18) in Theorem 2.5. The only difference is that instead of (6.18) we have

$$h\left(\frac{\sigma(a_1')}{a_1'}, \frac{\sigma(a_2')}{a_2'}\right) \le 2h(z_1, z_2) \le 2\varepsilon$$

which using now (2.11) leads to the same contradiction (6.19). Thus, (2.13) follows. □

References

[1] F. Beukers and H.P. Schlickewei, The equation $x + y = 1$ in finitely generated groups, *Acta Arith.*, **78** (1996), 189–199.

[2] F. Beukers and D. Zagier, Lower bounds of heights of points on hypersurfaces, *Acta Arith.*, **79** (1997), 103–111.

[3] E. Bombieri, Effective Diophantine approximation on $\mathbf{G}_m$, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, **20** (1993), 61–89.

[4] E. Bombieri and P. B. Cohen, Effective Diophantine approximation on $\mathbb{G}_M$. II, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, **24** (1997), 205–225.

[5] E. Bombieri and P. B. Cohen, An elementary approach to effective Diophantine approximation on $\mathbb{G}_m$, in: *Number theory and algebraic geometry*, Cambridge Univ. Press, Cambridge, 2003, pp. 41–62.

[6] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, Cambridge University Press, Cambridge, 2006.

[7] Y. Bugeaud, Bornes effectives pour les solutions des équations en S-unités et des équations de Thue-Mahler, *J. Number Theory*, **71** (1998), 227–244.

[8] Y. Bugeaud and K. Győry, Bounds for the solutions of unit equations, *Acta Arith.*, **74** (1996), 67–80.

[9] J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1959.

[10] J.-H. Evertse, Points on subvarieties of tori, in: *A panorama of number theory or the view from Baker's garden (Zürich, 1999)*, Cambridge Univ. Press, 2002, pp. 214–230.

[11] J.-H. Evertse, H.P. Schlickewei, W.M. Schmidt, Linear equations in variables which lie in a multiplicative group, *Ann. Math.*, **155** (2002), 807–836.

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

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

[14] E. M. Matveev, An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II, *Izv. Ross. Akad. Nauk Ser. Mat.*, **64** (2000), 125–180.

[15] B. Poonen, Mordell-Lang plus Bogomolov, *Invent. Math.*, **137** (1999), 413–425.

[16] G. Rémond, Décompte dans une conjecture de Lang, *Invent. Math.*, **142** (2000), 513–545.

[17] G. Rémond, Sur les sous-variétés des tores, *Compos. Math.*, **134** (2002), 337–366.

[18] G. Rémond, Approximation diophantienne sur les variétés semi-abeliennes, *Ann. Sci. École Norm. Sup.*, **36** (2003), 191–212.

[19] H. P. Schlickewei, Lower bounds for heights on finitely generated groups, *Monatsh. Math.*, **123** (1997), 171–178.

[20] W. M. Schmidt, *Diophantine approximation*, vol. 785 of *Lecture Notes in Mathematics*, Springer, Berlin, 1980.

[21] P. Voutier, An effective lower bound for the height of algebraic numbers, *Acta Arith.*, **74** (1996), 81–95.

[22] K. Yu, p-adic logarithmic forms and group varieties. III, *Forum Math.*, **19** (2007), 187–280.

A. Bérczes
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: berczesa@math.klte.hu

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@math.klte.hu