EFFECTIVE RESULTS FOR POINTS ON CERTAIN SUBVARIETIES OF TORI

ATTILA BÉRCZES, JAN-HENDRIK EVERTSE, KÁLMÁN GYŐRY, AND CORENTIN PONTREAU

ABSTRACT. Thanks to work of Laurent, Poonen and Rémond, who proved the conjecture of Lang-Bogomolov for tori in a more precise form, it is possible to give an accurate description of the set of $\overline{\mathbb{Q}}$-rational points of a given subvariety $\mathcal{X}$ of a linear torus defined over $\overline{\mathbb{Q}}$, that with respect to the height are "very close" to a given multiplicative group of finite rank.

In the present paper, we obtain, for certain special classes of varieties $\mathcal{X}$, effective versions of the results mentioned above, giving explicit upper bounds for the heights and degrees of the points under consideration. The main feature of our results is that the points we consider do not have to lie in a prescribed number field. Our main tools are Baker-type logarithmic forms estimates and Bogomolov-type estimates for the number of points on the variety $\mathcal{X}$ with very small height.

1. INTRODUCTION.

Choose an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. Recall that the group of $\overline{\mathbb{Q}}$-rational points of the N-dimensional torus is

$$\mathbb{G}_m^N(\overline{\mathbb{Q}}) = (\overline{\mathbb{Q}}^*)^N = \{\mathbf{x} = (x_1, \ldots, x_N) : x_i \in \overline{\mathbb{Q}}^* \text{ for } i = 1, \ldots, N\}$$

with coordinatewise multiplication, i.e., if $\mathbf{x} = (x_1, \ldots, x_N)$, $\mathbf{y} = (y_1, \ldots, y_N)$ then $\mathbf{xy} = (x_1y_1, \ldots, x_Ny_N)$. Denote by $h(x)$ the absolute logarithmic Weil

---

2000 Mathematics Subject Classification: 11D61.

Keywords and Phrases: Subvarieties of tori, Lang-Bogomolov conjecture, effective results, linear forms in logarithms.

The research was supported in part by the Hungarian Academy of Sciences (A.B.,K.G.), and 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.).

height of $x \in \overline{\mathbb{Q}}$. Define the height and degree of $\mathbf{x} = (x_1, \dots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ by $h(\mathbf{x}) := \sum_{i=1}^N h(x_i)$, and $[\mathbb{Q}(x_1, \dots, x_N) : \mathbb{Q}]$, respectively. Let $\mathcal{X}$ be an algebraic subvariety of $(\overline{\mathbb{Q}}^*)^N$ (i.e., the set of common zeros in $(\overline{\mathbb{Q}}^*)^N$ of a set of polynomials in $\overline{\mathbb{Q}}[X_1, \dots, X_N]$), and Γ a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$. We want to study the intersection of $\mathcal{X}$ with any of the sets

$$\overline{\Gamma} := \Big\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \exists m \in \mathbb{Z}_{>0} \text{ with } \mathbf{x}^m \in \Gamma \Big\} \quad \text{(the division group of } \Gamma\text{)},$$

$$\overline{\Gamma}_\varepsilon := \Big\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \exists \mathbf{y}, \mathbf{z} \in (\overline{\mathbb{Q}}^*)^N \text{ with } \mathbf{x} = \mathbf{y}\mathbf{z},\ \mathbf{y} \in \overline{\Gamma},\ h(\mathbf{z}) < \varepsilon \Big\},$$

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

where $\varepsilon > 0$.

Recall that by an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ we mean an algebraic subvariety that is a subgroup of $(\overline{\mathbb{Q}}^*)^N$, and by a translate of an algebraic subgroup a coset $\mathbf{x}\mathcal{H} = \{\mathbf{x} \cdot \mathbf{y} : \mathbf{y} \in \mathcal{H}\}$, where $\mathcal{H}$ is an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ and $\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N$.

It follows from work of Poonen [12] that there is $\varepsilon > 0$ depending only on N and the degree of $\mathcal{X}$, such that $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ is contained in a finite union of translates

$$\mathbf{x}_1\mathcal{H}_1 \cup \dots \cup \mathbf{x}_T\mathcal{H}_T \tag{1.1}$$

where $\mathbf{x}_i \in \overline{\Gamma}_\varepsilon$, $\mathcal{H}_i$ is an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$ and $\mathbf{x}_i\mathcal{H}_i \subset \mathcal{X}$ for $i = 1, \dots, T$. This encompasses earlier work of Liardet [9] and Laurent [8] (who considered $\mathcal{X} \cap \overline{\Gamma}$) and Zhang [17] (who considered $\mathcal{X} \cap \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : h(\mathbf{x}) < \varepsilon\}$).

Bombieri and Zannier [3] and Schmidt [15] proved precise quantitative versions for Zhang's result with an explicit positive value for ε and an explicit upper bound for the number T of translates, both depending only on N and the degree of $\mathcal{X}$ and their result was further improved by various authors. Later, Rémond [13] proved a quantitative version of Poonen's result with an explicit positive value for ε depending on N and the degree of $\mathcal{X}$ and an explicit upper bound for T depending only on N, the degree of $\mathcal{X}$ and the rank of Γ.

Define $\mathcal{X}^{\rm exc}$ to be the set of $\mathbf{x} \in \mathcal{X}$ with the property that there exists an algebraic subgroup $\mathcal{H}$ of $(\overline{\mathbb{Q}}^*)^N$ of dimension > 0 such that $\mathbf{x}\mathcal{H} \subset \mathcal{X}$, and let $\mathcal{X}^0 := \mathcal{X} \setminus \mathcal{X}^{\rm exc}$. The second author stated in the survey paper [7] that there exists $\varepsilon > 0$ depending on N, $\mathcal{X}$ and Γ such that $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$ is finite. This was proved in a more general form by Rémond [13]. In the case that $\mathcal{X}$ is a curve, Rémond gave, for some explicit value of ε depending on N, the rank of Γ and the height and degree of $\mathcal{X}$, an explicit upper bound for the cardinality of $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$; his result was recently improved by the fourth author [11] for curves in $(\overline{\mathbb{Q}}^*)^2$. For higher dimensional varieties, such a quantitative version has as yet not been established.

The purpose of the present paper is to derive, for certain special classes of varieties $\mathcal{X}$, effective versions of the results mentioned above. As for the intersection $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$, this means that we give an explicit value for ε and effectively computable upper bounds for the heights and degrees of the points $\mathbf{x}_1, \dots, \mathbf{x}_T$ in (1.1). As for $\mathcal{X}^0 \cap C(\overline{\Gamma}, \varepsilon)$, this means that we give an explicit value for ε and effectively computable upper bounds for the heights and degrees of the points in this intersection. We mention that to obtain fully effective results it is necessary to give effective upper bounds for the degrees as well since the points we are considering do not have their coordinates in a prescribed algebraic number field.

The classes of varieties we consider are such that they allow an application of logarithmic forms estimates. Two cases are worked out in detail. Firstly, we consider curves $\mathcal{C} : f(x, y) = 0$ in $(\overline{\mathbb{Q}}^*)^2$ where $f \in \overline{\mathbb{Q}}[X, Y]$ is not a binomial. Here we generalize a result of Bombieri and Gubler [4, p. 147, Theorem 5.4.5] and the first three authors [1, Theorems 2.1, 2.3 and 2.5] by giving explicit bounds for the heights of the points $\mathbf{x}$ contained both in $\mathcal{C}$ and in $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$, respectively. Our proofs are based on a new diophantine approximation theorem obtained in [1] (cf. Lemma 4.1 in Section 4 below). Secondly, we consider varieties in $(\overline{\mathbb{Q}}^*)^N$ given by equations $f_1(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0$ where each polynomial f_i is a binomial or trinomial. Here we apply effective results on linear equations $ax + by = 1$ established in [1].

In our proofs, the logarithmic forms estimates provide effective upper bounds for the heights; to obtain effective upper bounds for the degrees we need estimates for the number of points of small height in a variety. From these two basic cases one may deduce effective results for other classes of varieties; at the end of Section 2 we mention some possibilities. An important ingredient of our arguments (see Section 7 below) is an effective result of the following shape. Let $\mathbf{x}_0 \in (\overline{\mathbb{Q}}^*)^N$, and $\mathcal{H}$ a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$. If $\mathbf{x}_0\mathcal{H} \cap \Gamma$ or $\mathbf{x}_0\mathcal{H} \cap \overline{\Gamma}_\varepsilon$ is non-empty, then it contains a point with height and degree below some effectively computable bounds.

Our theorems are stated in Section 2. In Section 3 we introduce the necessary notation, in Section 4 we have collected our auxiliary results, and in the remaining sections we give the proofs.

2. Results

In the statements of our results the following notation is used.

Let K be an algebraic number field. The ring of integers of K is denoted by $\mathcal{O}_K$ and the set of places of K is denoted by M_K.

For every place $v \in M_K$ we choose an absolute value $|\cdot|_v$ in such a way that for $x \in \mathbb{Q}$ we have

$$|x|_v = |x|^{[K_v:\mathbb{R}]/[K:\mathbb{Q}]} \text{ if } v \text{ is infinite}, \;\; |x|_v = |x|_p^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]} \text{ if } v \text{ is finite},$$

where p is the prime below v. The absolute values $|\cdot|_v$ ($v \in M_K$) satisfy the Product formula $\prod_{v\in M_K} |x|_v = 1$ for $x \in K^*$.

For any finite set of places S of K, containing all infinite places, we define the ring of S-integers and group of S-units by

$$\begin{aligned} \mathcal{O}_S &= \{x \in K : |x|_v \le 1 \text{ for } v \in M_K \setminus S\}, \\ \mathcal{O}_S^* &= \{x \in K : |x|_v = 1 \text{ for } v \in M_K \setminus S\}, \end{aligned}$$

respectively.

The (absolute logarithmic) height of $x \in \overline{\mathbb{Q}}$ is defined by picking any number field K such that $x \in K$ and putting

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

This does not depend on the choice of K.

We define the height of $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ by

$$h(\mathbf{x}) := \sum_{i=1}^{n} h(x_i).$$

For a number field L and for $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ we define the extension $L(\mathbf{x}) = L(x_1, \ldots, x_N)$.

We write $\log^* x := \max(1, \log x)$ for $x > 0$ and $\log^* 0 := 1$.

If G is a finitely generated abelian group, we say that $\xi_1, \ldots, \xi_r$ generate G modulo G_{tors} if $\xi_1, \ldots, \xi_r \in G$ and if the reductions modulo G_{tors} of these elements generate G/G_{tors}. We call $\{\xi_1, \ldots, \xi_r\}$ a basis of G modulo G_{tors} if $\xi_1, \ldots, \xi_r \in G$ and the reductions of $\xi_1, \ldots, \xi_r$ modulo G_{tors} forms a basis of G/G_{tors}.

Let Γ be a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$, where $N \geq 2$. Further, let $\overline{\Gamma}$, $\overline{\Gamma}_\varepsilon$ and $C(\overline{\Gamma}, \varepsilon)$ be defined as in Section 1. Choose a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ of Γ modulo Γ_{tors} and put

$$h_0 := \max\big(1, h(\mathbf{w}_1), \ldots, h(\mathbf{w}_r)\big).$$

Denote by K the smallest number field such that $\Gamma \subset (K^*)^N$, and put $d := [K : \mathbb{Q}]$. Let S be the minimal finite set of places of K containing all the infinite places of K and having the property that $\Gamma \subset (\mathcal{O}_S^*)^N$ and denote by s the cardinality of S. Define

$$N(v) := 2 \text{ if } v \text{ is infinite}, \quad N(v) := \#\mathcal{O}_K/\mathfrak{p}_v \text{ if } v \text{ is finite}, \tag{2.2}$$

where $\mathfrak{p}_v$ is the prime ideal of $\mathcal{O}_K$ corresponding to v, and

$$\mathbf{N} := \max_{v \in S} N(v). \tag{2.3}$$

For the moment we assume that $N = 2$ and consider curves in $(\overline{\mathbb{Q}}^*)^2$. Thus, Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^2$; $\mathbf{w}_1, \ldots, \mathbf{w}_r$,$h_0$,$K, d, S, s, \mathbf{N}$ will have the same meaning as above. Let $f(X, Y) \in \overline{\mathbb{Q}}[X, Y]$ be an absolutely irreducible polynomial which is not of the shape $aX^mY^n - b$ or

$aX^m - bY^n$ for some $a, b \in \overline{\mathbb{Q}}$, $m, n \in \mathbb{Z}_{\geq 0}$. Let L be the field extension of K generated by the coefficients of f. Put

$$\delta := \deg_s f, \quad H := \max(1, h(f)),$$
$$C_1 := \left(e^{13}\delta^7 d^3 r\right)^{r+3} s \cdot \frac{\mathbf{N}^{2\delta^2}}{\log \mathbf{N}} h_0^r \cdot \log^* \left(\max(\delta d s \mathbf{N}, \delta h_0)\right).$$

Let $\mathcal{C} \subset (\overline{\mathbb{Q}}^*)^2$ be the curve defined by $f(x, y) = 0$. By our assumptions on f, $\mathcal{C}$ is not a translate of a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^2$.

Theorem 2.1. *For every point* $\mathbf{x} = (x, y) \in \mathcal{C} \cap \Gamma$ *we have*

$$h(\mathbf{x}) = h(x) + h(y) \leq C_1 H.$$

Notice that in this bound there is no dependence on the field L other than what is implicit from H.

The following results are obtained by combining the above theorem with estimates for the number of points of small height on a curve in $(\overline{\mathbb{Q}}^*)^2$. The notation will be the same as above.

Theorem 2.2. *Let*

$$\varepsilon := \left(2^{48}\delta(\log \delta)^5\right)^{-1}. \tag{2.4}$$

Then for every $\mathbf{x} \in \mathcal{C} \cap \overline{\Gamma}_\varepsilon$ *we have*

$$h(\mathbf{x}) \leq r h_0 \delta C_1 + C_1 H, \quad [L(\mathbf{x}) : L] \leq 2^{50}\delta(\log \delta)^6.$$

Theorem 2.3. *Let*

$$\varepsilon := \left(2^{50}\delta(\log \delta)^5\right)^{-1} \cdot \left(r h_0 \delta C_1 + C_1 H\right)^{-1}. \tag{2.5}$$

Then for every $\mathbf{x} \in \mathcal{C} \cap C(\overline{\Gamma}, \varepsilon)$ *we have*

$$h(\mathbf{x}) \leq 2 r h_0 \delta C_1 + 2 C_1 H, \quad [L(\mathbf{x}) : L] \leq 2^{50}\delta(\log \delta)^6.$$

Remark. In the special case when f is linear, (i.e., $\mathcal{C}$ is a line), our above theorems have been proved in [1] with larger ε's and sharper upper bounds.

Now we turn our attention to varieties of arbitrary dimension N. Let

$$\mathcal{X} := \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \ f_i(\mathbf{x}) = 0, \ i = 1, \ldots, m\}$$

be a subvariety of $(\overline{\mathbb{Q}}^*)^N$, where $f_1, \dots, f_m$ are non-constant polynomials in $\overline{\mathbb{Q}}[X_1, \dots, X_N]$ each consisting of 2 or 3 monomials. Put

$$\delta := \max(\deg f_1, \dots, \deg f_m), \quad H := \max(1, h(f_1), \dots, h(f_m)).$$

Further, let L be the smallest number field containing K and the coefficients of the polynomials f_i $(i = 1, \dots, m)$. Again, Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$, $K, d, S, s, h_0, \mathbf{N}$ have the same meaning as before. The stabilizer of $\mathcal{X}$ is given by

$$\mathrm{Stab}(\mathcal{X}) = \left\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \mid \mathbf{x}\mathcal{X} \subseteq \mathcal{X}\right\},$$

where $\mathbf{x}\mathcal{X} = \{\mathbf{xy} \ : \ \mathbf{y} \in \mathcal{X}\}$. $\mathrm{Stab}(\mathcal{X})$ is clearly an algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$, and it can be computed effectively in terms of the defining polynomials $f_1, \dots, f_m$ of $\mathcal{X}$.

Put

$$C^* := \left(e^{11} d^3 r\right)^{r+3} (\delta h_0)^r s \cdot \frac{\mathbf{N}}{\log \mathbf{N}} \cdot \log^* \max(ds\mathbf{N}, \delta h_0), \tag{2.6}$$

and

$$\begin{cases} C_2 := C^* N (2\delta)^{N-1}. \\ C_3 := C^* \cdot 2mh_0 \left(r4^{r+1} \cdot d(\log 3d)^3 \cdot m\delta h_0\right)^r. \end{cases} \tag{2.7}$$

Theorem 2.4. *Let $\mathcal{X}$ satisfy the conditions listed above, and put $\mathcal{H} :=$ $\mathrm{Stab}(\mathcal{X})$.*

(i) *Suppose that $\mathcal{H}$ is finite. Then for every $\mathbf{x} \in \mathcal{X} \cap \Gamma$ we have*

$$h(\mathbf{x}) \leq C_2 H.$$

(ii) *Suppose that $\mathcal{H}$ is not finite. Then $\mathcal{X} \cap \Gamma$ is contained in some finite union of translates*

$$\mathbf{x}_1 \mathcal{H} \cup \dots \cup \mathbf{x}_T \mathcal{H},$$

with

$$\mathbf{x}_i \mathcal{H} \subset \mathcal{X}, \quad \mathbf{x}_i \in \Gamma, \quad h(\mathbf{x}_i) \leq C_3 H \text{ for } i = 1, \dots, T. \tag{2.8}$$

Our results for $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ and $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ are as follows.

Theorem 2.5. *Put*

$$\varepsilon := \frac{0.03}{4\delta}. \tag{2.9}$$

(i) *Assume that* $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ *is finite. Then for every* $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$ *we have*

$$h(\mathbf{x}) < rh_0\delta C_2 + C_2H, \quad [L(\mathbf{x}) : L] \le 2^{m+N}\delta^N. \tag{2.10}$$

(ii) *Assume that* $\mathcal{H}$ *is not finite. Then* $\mathcal{X} \cap \overline{\Gamma}_\varepsilon$ *is contained in a finite union of translates*

$$\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H},$$

where for $i = 1, \ldots, T$, *we have* $\mathbf{x}_i \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$, $\mathbf{x}_i\mathcal{H} \subset \mathcal{X}$, *and where* $h(\mathbf{x}_i)$ *and* $[L(\mathbf{x}_i) : L]$ *are bounded above by effectively computable numbers depending only on* $\Gamma, f_1, \ldots, f_m$.

Remark. It is possible in principle to give explicit expressions for the effectively computable numbers in part (ii) of Theorem 2.5, but these are rather complicated.

Theorem 2.6. *Let*

$$\varepsilon := \frac{0.03}{4\delta(C_2\delta rh_0 + 2C_2H)}. \tag{2.11}$$

Assume that $\mathrm{Stab}(\mathcal{X})$ *is finite. Then for every* $\mathbf{x} \in \mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ *we have*

$$h(\mathbf{x}) \le 2rh_0\delta C_2 + 2C_2H, \quad [L(\mathbf{x}) : L] \le 2^{m+N}\delta^N.$$

Remark. If $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ is not finite, then in general $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ need not be contained in a finite union of translates $\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H}$. Indeed, suppose that $\dim \mathcal{X} > \dim \mathcal{H}$, and that $\mathcal{H} \cap \Gamma$ contains points of infinite order. Pick any $\mathbf{x}_0 \in \mathcal{X}$. Choose a point $\mathbf{u} \in \mathcal{H} \cap \Gamma$ of infinite order. Thus $h(\mathbf{u}) > 0$. Then for any sufficiently large integer n,

$$h(\mathbf{x}_0) \le \varepsilon(1 + nh(\mathbf{u}) - h(\mathbf{x}_0)) \le \varepsilon(1 + h(\mathbf{x}_0\mathbf{u}^n)).$$

Hence $\mathbf{x} := \mathbf{x}_0\mathbf{u}^n \in \mathbf{x}_0\mathcal{H} \cap C(\overline{\Gamma}, \varepsilon)$. That is, every translate $\mathbf{x}_0\mathcal{H}$ with $\mathbf{x}_0 \in \mathcal{X}$ contains elements from $C(\overline{\Gamma}, \varepsilon)$. If $\mathcal{X} \cap C(\overline{\Gamma}, \varepsilon)$ were contained in a finite union of translates $\cup_{i=1}^t \mathbf{x}_i\mathcal{H}$, then so were $\mathcal{X}$, which is impossible.

Possible extensions. We discuss some other cases, where one may get effective results similar to those discussed above.

1. First let $\mathcal{C}$ be an irreducible curve in $(\overline{\mathbb{Q}}^*)^N$ where $N \geq 2$. Assume that $\mathcal{C}$ is not contained in a translate $\mathbf{x}\mathcal{H}$ where $\mathcal{H}$ is a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$. Viewing the variables $X_1, \dots, X_N$ as functions on $\mathcal{C}$, at least one of them, X_1 say, is transcendental over $\overline{\mathbb{Q}}$, while the others are algebraically dependent on X_1. Hence there are polynomials $f_2, \dots, f_n \in \overline{\mathbb{Q}}(X, Y)$, which can be determined effectively from the data describing $\mathcal{C}$, such that for each point $(x_1, \dots, x_N) \in \mathcal{C}$ we have $f_i(x_1, x_i) = 0$ for $i = 2, \dots, N$. None of the polynomials $f_2, \dots, f_N$ can be a binomial since otherwise $\mathcal{C}$ would be contained in a translate of an algebraic group. Let $(x_1, \dots, x_N)$ be in the intersection of $\mathcal{C}$ with Γ, $\overline{\Gamma}_\varepsilon$ or $C(\overline{\Gamma}, \varepsilon)$. Then we obtain upper bounds for the heights and degrees of $x_1, \dots, x_N$ by applying Theorems 2.1, 2.2, 2.3 to $f_i(x_1, x_i) = 0$ $(i = 2, \dots, N)$.

2. Recall that a homomorphism of algebraic groups from $(\overline{\mathbb{Q}}^*)^N$ to $(\overline{\mathbb{Q}}^*)^M$ is given by

$$(x_1, \dots, x_N) \mapsto \Big(\prod_{j=1}^N x_j^{a_{1j}}, \dots, \prod_{j=1}^N x_j^{a_{Mj}}\Big)$$

where the exponents a_{ij} are integers. Now our Theorems 2.4, 2.5, 2.6 can be extended to varieties $\mathcal{X} = \bigcap_{i=1}^m \varphi_i^{-1}(\mathcal{C}_i)$, where for $i = 1, \dots, m$, $\mathcal{C}_i$ is a curve in $(\overline{\mathbb{Q}}^*)^2$ and φ_i a homomorphism of algebraic groups from $(\overline{\mathbb{Q}}^*)^N$ to $(\overline{\mathbb{Q}}^*)^2$.

We define the *rank* of a polynomial $f = \sum_{\mathbf{i} \in I} a(\mathbf{i}) X_1^{i_1} \cdots X_N^{i_N}$ (where $\mathbf{i} = (i_1, \dots, i_N)$, I is a finite set, and $a(\mathbf{i}) \in \overline{\mathbb{Q}}^*$ for $\mathbf{i} \in I$) to be the rank of the $\mathbb{Z}$-module generated by $\mathbf{i} - \mathbf{j}$ for all $\mathbf{i}, \mathbf{j} \in I$. Then a variety $\mathcal{X}$ as above can be given by polynomial equations $f_1(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0$ where $f_1, \dots, f_m$ are polynomials in $\overline{\mathbb{Q}}[X_1, \dots, X_N]$ of rank ≤ 2.

3. Heights

By the Product formula we have for any number field K and any $x \in K^*$ that

$$h(x) = \sum_{v \in M_K} \max(0, \log |x|_v) = \frac{1}{2} \sum_{v \in M_K} |\log |x|_v|. \tag{3.1}$$

Recall that we have defined

$$h(\mathbf{x}) := \sum_{i=1}^{n} h(x_i)$$

for $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$. Further, for $\xi \in \mathbb{Q}$ we define $\mathbf{x}^\xi := (x_1^\xi, \ldots, x_N^\xi)$. The point $\mathbf{x}^\xi$ is determined only up to multiplication with $(\overline{\mathbb{Q}}^*_{\rm tors})^N$ where $\overline{\mathbb{Q}}^*_{\rm tors} = \{\boldsymbol{\rho} \in \overline{\mathbb{Q}}^* \,:\, \exists m \in \mathbb{Z}_{>0} \text{ with } \boldsymbol{\rho}^m = 1\}$. But $h(\mathbf{x}^\xi)$ is well defined. It now follows easily that

$$h(\mathbf{x}\mathbf{y}) \le h(\mathbf{x}) + h(\mathbf{y}),\ h(\mathbf{x}^\xi) = |\xi| h(\mathbf{x}) \text{ for } \mathbf{x}, \mathbf{y} \in (\overline{\mathbb{Q}}^*)^N,\ \xi \in \mathbb{Q},$$

and $h(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in (\overline{\mathbb{Q}}^*_{\rm tors})^N$.

We define several heights for polynomials. Let f be a non-zero polynomial with coefficients in $\overline{\mathbb{Q}}$, and let $a_1, \ldots, a_R$ be its non-zero coefficients. Choose a number field K such that $a_1, \ldots, a_R \in K$. Recall that for every infinite place v of K there is an embedding $\sigma_v : K \hookrightarrow \mathbb{C}$ such that $|\cdot|_v = |\sigma_v(\cdot)|^{\varepsilon_v}$, where $\varepsilon_v := [K_v : \mathbb{R}]/[K : \mathbb{Q}]$. For $v \in M_K$ we put $\|f\|_v := \max_{1 \le i \le R} |a_i|_v$. Further, for every infinite place v of K and every $l \ge 1$ we put $\|f\|_{v,l} := \Big(\sum_{i=1}^{R} |\sigma_v(a_i)|^l\Big)^{\varepsilon_v/l}$. We have already defined

$$h(f) := \sum_{v \in M_K} \log \|f\|_v.$$

In addition, we define the heights

$$h_l(f) := \sum_{v|\infty} \log \|f\|_{v,l} + \sum_{v\nmid\infty} \log \|f\|_v \text{ for } l \ge 1,$$

and the Gauss-Mahler height

$$h_{GM}(f) := \sum_{v|\infty} \varepsilon_v \log M(f^{\sigma_v}) + \sum_{v\nmid\infty} \log \|f\|_v,$$

where f^σ is the polynomial obtained by applying σ to the coefficients of f and $M(\cdot)$ denotes the Mahler measure of a polynomial with complex coefficients. None of these heights depends on the choice of K. We have

$$h_{GM}(f) \le h_1(f), \quad h(f) \le h_1(f) \le h(f) + \log R, \tag{3.2}$$

where R is the number of non-zero coefficients of f. Further, for any non-zero polynomial $P \in \overline{\mathbb{Q}}[X]$ and any root ζ of P we have

$$h(\zeta) \leq h_{GM}(P) \leq h_1(P). \tag{3.3}$$

We use also exponential heights $H(x) = \exp(h(x))$ for $x \in \overline{\mathbb{Q}}$, and likewise $H(f), H_l(f), H_{GM}(f)$ for polynomials f with coefficients in $\overline{\mathbb{Q}}$.

4. Main tools

In this section we have collected the tools needed in the sequel.

We start with some results from [1] that have been derived from lower bounds for linear forms in logarithms. Let 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 $t > 0$. We fix a set of (not necessarily multiplicatively independent) generators $\{\xi_1, \ldots, \xi_r\}$ of G modulo G_{tors} and put

$$Q := \prod_{i=1}^{r} \max\big(1, h(\xi_i)\big). \tag{4.1}$$

Let $N(v)$ $(v \in M_K)$ be given by (2.2), i.e., $N(v) := 2$ if v is infinite and $N(v) := \#\mathcal{O}_K/\mathfrak{p}_v$ if v is finite, where $\mathfrak{p}_v$ is the prime ideal of $\mathcal{O}_K$ corresponding to v.

Lastly, let

$$c(r,d) := 20(16ed)^{3(r+2)}\left(\frac{r}{e}\right)^r.$$

Lemma 4.1. *Let* $\alpha \in K^*$ *with* $\max(h(\alpha), 1) \leq H$, $v \in M_K$, *and* $0 < \kappa \leq 1$. *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 $h(\xi) < C_4(\kappa) \cdot H$, *where*

$$\begin{aligned} C_4(\kappa) \;&:=\; (c(r,d)/\kappa)\frac{N(v)}{\log N(v)} Q \cdot \\ &\qquad \cdot \max\{\log(c(r,d)N(v)/\kappa), \log^* Q\}. \end{aligned}$$

Proof. This is [1, Theorem 4.2], with instead of $c(r,d)$ a constant c depending also on the rank t of G. However, using $t \leq r$ an easy computation proves the estimate of our lemma. □

We keep the notation from above. In addition, let S be a finite set of places of K containing all infinite places such that $G \subset \mathcal{O}_S^*$. Put $s := \#S$ and define $\mathbf{N}$ by (2.3), that is $\mathbf{N} := \max_{v \in S} N(v)$. Consider the equation

$$\alpha x + \beta y = 1 \quad \text{in } x \in G,\ y \in \mathcal{O}_S^*, \tag{4.3}$$

where $\alpha, \beta \in K^*$ with $\max(h(\alpha), h(\beta), 1) \leq H$.

Lemma 4.2. *For every solution $x \in G, y \in \mathcal{O}_S^*$ of (4.3) we have*

$$\max\left(h(x), h(y)\right) < C_5 H, \tag{4.4}$$

where

$$C_5 := c(r,d) \cdot \frac{s\mathbf{N}}{\log \mathbf{N}} Q \cdot \max\{\log(c(r,d)s\mathbf{N}), \log^* Q\}.$$

Proof. This is [1, Theorem 2.2], again with a constant c depending on the rank t of G which we bounded above using $t \leq r$. □

Below we have collected some results on heights of algebraic points.

Lemma 4.3. *Suppose that α is a non-zero algebraic number of degree d, which is not a root of unity. Then*

$$h(\alpha) \geq c(d)^{-1}$$

where

$$c(1) = \frac{1}{\log 2}, \quad c(d) = \frac{d(\log 3d)^3}{2} \quad \text{if } d \geq 2.$$

Proof. This is the main result of Voutier [16]. □

Lemma 4.4. *(i) Let $\alpha, \beta \in \overline{\mathbb{Q}}^*$. Then there are at most two points $\mathbf{x} = (x,y) \in (\overline{\mathbb{Q}}^*)^2$ such that*

$$\alpha x + \beta y = 1, \quad h(\mathbf{x}) \leq 0.03.$$

(ii) Let $f(X,Y) \in \overline{\mathbb{Q}}[X,Y]$ be an irreducible polynomial which is not a binomial. Then the number of points $\mathbf{x} = (x,y) \in (\overline{\mathbb{Q}}^*)^2$ with

$$f(x,y) = 0, \quad h(\mathbf{x}) \leq \left(2^{47} \deg_s f (\log \deg_s f)^5\right)^{-1}$$

is at most

$$2^{50} \deg_s f (\log \deg_s f)^6.$$

Proof. (i) Beukers and Zagier [2, Corollary 2.4] proved that if there are three points $(x_1, y_1), (x_2, y_2), (x_3, y_3) \in (\overline{\mathbb{Q}}^*)^2$ satisfying $\alpha x_i + \beta y_i = 1$ for $i = 1, 2, 3$, then $\sum_{i=1}^3 h(x_i, y_i) \geq \log \rho$, where ρ denotes the real root of $\rho^{-6} + \frac{1}{2}\rho^{-2} = 1$ which is larger than 1. We have $\log \rho > 0.09$.

(ii) This is proved by the fourth author in [10, Proposition 5.1] (see also [11, Proposition 3.3]). □

Our last height result is an effective version of a special case of Bézout's Theorem.

Lemma 4.5. *Let $f, g \in \overline{\mathbb{Q}}[X, Y]$ be two coprime polynomials. Then for every common zero $\mathbf{x} = (x, y)$ of f and g we have*

$$h(\mathbf{x}) \leq \deg_s g \cdot h_{GM}(f) + \deg_s f \cdot h_1(g).$$

Proof. See [11, Lemma 3.7]. □

5. Proof of Theorem 2.1

We denote the partial degrees of f with respect to X, Y by δ_X, δ_Y, respectively, and put $\delta := \deg_s f = \delta_X + \delta_Y$. From our assumptions it follows that f is irreducible over $\overline{\mathbb{Q}}$, that f has at least three non-zero terms, and hence that $\delta_X \geq 1$, $\delta_Y \geq 1$.

We assume that one of the coefficients of f is 1 which is no loss of generality since the height of a polynomial is invariant under multiplication by a scalar.

Recall that we allow that f has its coefficients in $\overline{\mathbb{Q}}$; this will be needed in the proofs of Theorems 2.2, 2.3. But in fact there is no loss of generality to assume that $f \in K[X, Y]$. To see this, suppose that $f \notin K[X, Y]$. Then there is $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K)$ such that the polynomial f^σ obtained by applying σ to the coefficients of f is distinct from f. Since one of the coefficients of f is 1, f^σ cannot be proportional to f, and since f is irreducible over $\overline{\mathbb{Q}}$, f^σ has to be coprime to f. Now if $\mathbf{x} \in \Gamma$ is a zero of f then it is also a zero of f^σ. Thus, by Lemma 4.5, (3.2), noting that $\deg_s f = \deg_s f^\sigma = \delta$, it follows that

$$h(\mathbf{x}) \leq \delta(h_{GM}(f^\sigma) + h_1(f)) \leq 2\delta(H + 2\log \delta)$$

and this is much sharper than the bound from Theorem 2.1.

Write

$$f(X,Y)=\sum_{(i,j)\in\mathcal{F}} a_{ij}X^iY^j \text{ with } a_{ij}\in K^* \text{ for } (i,j)\in\mathcal{F}, \tag{5.1}$$

where $\mathcal{F}$ is a subset of $\{0,\dots,\delta_X\}\times\{0,\dots,\delta_Y\}$. Thus,

$$\#\mathcal{F}\le(\delta_X+1)(\delta_Y+1)\le\delta^2.$$

The height $H(f)$ remains unaltered under multiplication by a_{ij}^{-1} for any $(i,j)\in\mathcal{F}$, so we have for any place $v\in M_K$ and any two pairs $(i,j),(k,l)\in\mathcal{F}$,

$$|a_{pq}/a_{ij}|_v\le\max_{k,l}|a_{kl}/a_{ij}|_v\le H(f)$$

and by interchanging the role of a_{pq},a_{ij},

$$H(f)^{-1}\le|a_{pq}/a_{ij}^{-1}|_v\le H(f). \tag{5.2}$$

Put $s:=\#S$. Take a point $\mathbf{x}=(x,y)\in\mathcal{C}\cap\Gamma$ with

$$H(\mathbf{x})\ge(\delta^2H(f))^{24s\delta^4}. \tag{5.3}$$

Notice that the logarithm of the right-hand side is much smaller than the upper bound C_1H from our Theorem. By the product formula we have

$$\begin{aligned} H(\mathbf{x})^2=(H(x)H(y))^2 &= \prod_{v\in S}\max(|x|_v,|x|_v^{-1})\max(|y|_v,|y|_v^{-1}) \\ &\le \prod_{v\in S}\max(|x|_v,|x|_v^{-1},|y|_v,|y|_v^{-1})^2. \end{aligned}$$

Thus, there exists $v\in S$ such that

$$\max(|x|_v,|x|_v^{-1},|y|_v,|y|_v^{-1})\ge H(\mathbf{x})^{1/s}\ge(\delta^2H(f))^{24\delta^4}.$$

Replacing x by $x^{\pm1},y^{\pm1}$ and correspondingly f by a polynomial $\tilde{f}$ with $\tilde{f}(x^{\pm1},y^{\pm1})=0$ (which has the same partial degrees and height as f), we see that there is no loss of generality to assume that $\min(|x|_v,|y|_v)\ge1$ and moreover,

$$\max(|x|_v,|y|_v)\ge H(\mathbf{x})^{1/s}\ge(\delta^2H(f))^{24\delta^4}. \tag{5.4}$$

Now let us order the pairs in $\mathcal{F}$ according to

$$|x^py^q|_v\ge|x^ry^s|_v\ge\cdots.$$

Recall that f is not a binomial. Hence $\mathcal{F}$ contains pairs other than (p,q), (r,s). Further, $\delta_X, \delta_Y \geq 1$ so $\mathcal{F}$ contains pairs (i,j) with $i > 0$ and pairs with $j > 0$. Using also $\min(|x|_v, |y|_v) \geq 1$, it follows that $|x^p y^q|_v \geq \max(|x|_v, |y|_v)$. Now (5.4) gives

$$|x^p y^q|_v \geq H(\mathbf{x})^{\frac{1}{s}} \geq (\delta^2 H(f))^{24\delta^4}. \tag{5.5}$$

We compare $|x^p y^q|_v$, $|x^r y^s|_v$. Using that $f(x,y) = 0$ and also (5.1), (5.2), and the fact that $\#\mathcal{F} \leq \delta^2$, we obtain

$$|x^p y^q|_v \leq \delta^2 \max_{(i,j)\in\mathcal{F}} |a_{ij}|_v |a_{pq}|_v^{-1} |x^i y^j|_v \leq \delta^2 H(f) |x^r y^s|_v.$$

Hence

$$1 \leq \left|x^{p-r} y^{q-s}\right|_v \leq \delta^2 H(f). \tag{5.6}$$

We claim that (p,q) and (r,s) are linearly independent. Indeed assume there exists $u \in \mathbb{Q} \setminus \{1\}$ such that $(up, uq) = (r,s)$. We deduce from (5.6)

$$|x^p y^q|_v^{|1-u|} \leq \delta^2 H(f).$$

We note that from $p, q \leq \delta - 1$ it follows $|u-1| \geq \frac{1}{\delta-1}$, thus

$$|x^p y^q|_v \leq (\delta^2 H(f))^{\delta-1}$$

which contradicts (5.5).

Hence for all $(i,j) \in \mathcal{F}$ there are $A_{ij}, B_{ij} \in \mathbb{Q}$ with

$$i = A_{ij} p + B_{ij} r, \quad j = A_{ij} q + B_{ij} s.$$

Let $(i,j) \in \mathcal{F}$. Then using

$$x^i y^j = (x^p y^q)^{A_{ij}+B_{ij}} (x^{r-p} y^{s-q})^{B_{ij}} \tag{5.7}$$

and (5.6), we get

$$\begin{aligned} |x^p y^q|_v &\geq |x^i y^j|_v = |x^p y^q|_v^{A_{ij}+B_{ij}} |x^{r-p} y^{s-q}|_v^{B_{ij}} \\ &\geq |x^p y^q|_v^{A_{ij}+B_{ij}} \cdot (\delta^2 H(f))^{-|B_{ij}|}. \end{aligned}$$

Put $D = |ps - qr|$. Then $D, D \cdot A_{ij} = is - jr$ and $D \cdot B_{ij} = pj - qi \in \mathbb{Z}$ and moreover, $|D| \leq (\delta-1)^2$, $|DA_{ij}| \leq (\delta-1)^2$, $|DB_{ij}| \leq (\delta-1)^2$. Therefore,

$$|x^p y^q|_v^{D - D(A_{ij}+B_{ij})} \geq (\delta^2 H(f))^{-(\delta-1)^2}.$$

Since $|x^py^q|_v > (\delta^2 H(f))^{(\delta-1)^2}$ (by (5.5)) the integer $D - D(A_{ij} + B_{ij})$ is non-negative, in other words $A_{ij}+B_{ij} = 1$ or $A_{ij}+B_{ij} \le 1-\frac{1}{D}$. Now define $\mathcal{I}$ to be the set of $(i,j) \in \mathcal{F}$ such that $A_{ij} + B_{ij} = 1$. The set $\mathcal{I}$ contains at least the pairs (p,q) and (r,s). Choose a D-th root $z^{1/D}$ of $z := x^{r-p}y^{s-q}$. Then by (5.7) we have

$$\begin{aligned} &0 = f(x,y) = x^py^qR(z^{1/D}) + Q(x,y) \\ &\text{with } R(Z) := \sum_{(i,j)\in\mathcal{I}} a_{ij}Z^{DB_{ij}},\ Q(X,Y) := \sum_{(i,j)\in\mathcal{F}\setminus\mathcal{I}} a_{ij}X^iY^j. \end{aligned} \tag{5.8}$$

Let $m := -\min\{DB_{ij} : (i,j) \in \mathcal{I}\}$ and put $R^*(Z) := Z^mR(Z)$. Thus $R^*(Z)$ is a polynomial with $R^*(0) \ne 0$. Since $\mathcal{I}$ contains at least two pairs, the polynomial R^* is non-constant. Choose an extension of $|\cdot|_v$ to $\overline{\mathbb{Q}}$. We proceed to estimate from above $|R^*(z^{1/D})|_v$.

Let $(i,j) \in \mathcal{F}\setminus\mathcal{I}$. Then by (5.7), $A_{ij} + B_{ij} \le 1 - \frac{1}{D}$, $|B_{ij}| \le (\delta-1)^2/D$, (5.6) we have

$$\begin{aligned} |x^iy^j|_v &= |x^py^q|_v^{A_{ij}+B_{ij}} \cdot \left|x^{r-p}y^{q-s}\right|_v^{B_{ij}} \\ &\le |x^py^q|_v^{1-\frac{1}{D}} \cdot (\delta^2H(f))^{(\delta-1)^2/D}. \end{aligned}$$

Hence

$$|Q(x,y)|_v \le |x^py^q|_v^{1-1/D} \cdot (\delta^2H(f))^{1+(\delta-1)^2/D}.$$

Using this estimate together with (5.6), (5.5), we obtain

$$\begin{aligned} |R^*(z^{1/D})|_v &= |z|_v^{m/D}|R(z^{1/D})|_v = |z|_v^{m/D}|Q(x,y)|_v \\ &\le (\delta^2H(f))^{\delta^2/D}|x^py^q|_v^{-1/D}(\delta^2H(f))^{1+(\delta-1)^2/D} \\ &\le (\delta^2H(f))^{(3\delta^2)/D}H(\mathbf{x})^{-1/sD}. \end{aligned}$$

It is useful to observe here that in the above argument the D-th root $z^{1/D}$ was chosen arbitrarily. Thus, we have

$$|\prod_{\rho} R^*(\rho z^{1/D})|_v \le (\delta^2H(f))^{3\delta^2}H(\mathbf{x})^{-1/s} \tag{5.9}$$

where the product is taken over all D-th roots of unity.

Notice that the constant term of R^* is a coefficient of f, say a_{i_0,j_0}. By dividing f by a_{i_0,j_0} as we may since it does not affect the above estimates, we get that the constant term of R^* is 1. Thus we have

$$R^*(Z) = \prod_{\zeta}(1 - \zeta^{-1}Z)$$

where the product is taken over all zeros of R^*. So

$$\prod_{\rho} R^*(\rho z^{1/D}) = \prod_{\zeta}(1 - \zeta^{-D}z).$$

Choose some ζ for which $|1 - \zeta^{-D}z|_v$ is minimal. Using (5.9), (5.5), and also that R^* has degree at most $2\delta^2$ and that $H(z) \le H(\mathbf{x})^{\delta}$ we arrive at

$$\begin{aligned}|1 - \zeta^{-D}z|_v &\le \{(\delta^2 H(f))^{3\delta^2} H(\mathbf{x})^{-1/s}\}^{1/\deg R^*} \\ &\le \left(H(\mathbf{x})^{-2/3s}\right)^{1/2\delta^2} \le H(z)^{-1/3s\delta^3}.\end{aligned}$$

The number ζ^{-D} may lie outside K. Let $K' = K(\zeta^D)$. Then $[K' : K] \le 2\delta^2$ and there is a place v' of K' lying above v such that $|\gamma|_{v'} = |\gamma|_v^{[K'_{v'}:K_v]/[K':K]}$ for $\gamma \in K'$ where $|\cdot|_{v'}$ is normalized with respect to K'. Thus we finally obtain

$$\log|1 - \zeta^{-D}z|_{v'} \le -\frac{1}{6s\delta^5} \cdot h(z). \tag{5.10}$$

Now we apply Lemma 4.1 to (5.10) with $\alpha = \zeta^{-D}$, $\kappa = (6s\delta^5)^{-1}$, K' instead of K, v' instead of v and we take for G the group $\{x^{r-p}y^{s-q} : (x, y) \in \Gamma\}$. Notice that by (3.3), (3.2),

$$\begin{aligned}&h(\zeta^D) \le Dh_1(R^*) \le \delta^2 h_1(f) \le \delta^2(H + 2\log\delta), \\ &[K' : \mathbb{Q}] \le 2\delta^2 d,\ N(v') \le N(v)^{2\delta^2}.\end{aligned}$$

So in the bound $C_4(\kappa)H$ from Lemma 4.1 we have to replace H by $\delta^2(H + 2\log\delta)$, κ by $(6\delta^5 s)^{-1}$, d by $2\delta^2 d$ and $N(v)$ by $N(v') \le N(v)^{2\delta^2} \le \mathbf{N}^{2\delta^2}$. Further, if $\{\mathbf{w}_i = (w_{1i}, w_{2i}) : i = 1, \dots, r\}$ is a basis of Γ modulo Γ_{tors}, the group G is generated modulo G_{tors} by the numbers $\xi_i := w_{1i}^{r-p} w_{2i}^{s-q}$ $(i = 1, \dots, r)$ and so for the quantity Q defined by (4.1) we have

$$Q = \prod_{i=1}^{r} \max(1, h(\xi_i)) \le (\delta h_0)^r.$$

A straightforward computation shows that with these replacements for H, κ, $N(v)$ and the upper bound for Q, the constant $c(r, d)$ becomes $c' :=$

$20(32e\delta^2 d)^{3r+6}(32^3e^2r)^r$, and $C_4(\kappa)$ can be estimated from above by

$$c'\cdot 6\delta^5 s\cdot\frac{\mathbf{N}^{2\delta^2}}{2\delta^2\log\mathbf{N}}\cdot(\delta h_0)^r\cdot$$
$$\cdot\max\Big(\log(c'\mathbf{N}^{2\delta^2}\cdot 6\delta^5 s),\log^*\big((\delta h_0)^r\big)\Big).$$

Using that the maximum is at most $100r\delta^2\log^*\big(\max(\delta ds\mathbf{N},\delta h_0)\big)$, we obtain for $C_4(\kappa)$ the upper bound

$$C:=e^{36}(e^{13}r)^r\delta^{7r+17}d^{3r+6}sh_0^r\cdot\frac{\mathbf{N}^{2\delta^2}}{\log\mathbf{N}}\cdot r^2\log^*\big(\max(\delta ds\mathbf{N},\delta h_0)\big).$$

Thus, if $z\neq\zeta^D$ we get

$$h(z)<C\max(1,h(\zeta^{-D}))\leq C\delta^2(H+2\log\delta),$$

while if $z=\zeta^D$ we get $h(z)\leq\delta^2(H+2\log\delta)$ which is much smaller.

We proved that $\mathbf{x}=(x,y)$ verifies an equation $x^{r-p}y^{s-q}=\mu$ for some $\mu\in K$ with

$$h(\mu)\leq C\cdot\delta^2(H+2\log\delta).$$

Since f is irreducible and not a binomial, we can apply Lemma 4.5 and obtain, using $h_{GM}(X^rY^s-\mu X^pY^q)=h(\mu)$, $h_1(f)\leq H+2\log\delta$, the upper bound

$$\begin{aligned}h(\mathbf{x})&\leq\delta(h_1(f)+h(\mu))\leq\delta(\delta^2C+1)\cdot(H+2\log\delta)\\&\leq 3\delta^4CH\leq C_1H.\end{aligned}$$

Our Theorem follows. □

6. Proofs of Theorems 2.2 and 2.3

Theorems 2.2 and 2.3 are proved in the same manner. We prove only Theorem 2.3 and then indicate the necessary modifications to obtain a proof of Theorem 2.2.

Proof of Theorem 2.3. Let $\mathbf{x}\in\mathcal{C}\cap C(\overline{\Gamma},\varepsilon)$ with the value of ε given by (2.5). We may write $\mathbf{x}=\mathbf{yz}$ with $\mathbf{y}\in\overline{\Gamma}$ and $\mathbf{z}\in(\overline{\mathbb{Q}}^*)^2$ with $h(\mathbf{z})<\varepsilon(1+h(\mathbf{y}))$. We may further split up $\mathbf{y}$ as

$$\mathbf{y}=\mathbf{vw}\quad\text{with }\mathbf{v}\in\Gamma,\mathbf{w}=\prod_{i=1}^r\mathbf{w}_i^{\gamma_i},\tag{6.1}$$

where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$. Here $\mathbf{w}$ is determined only up to a root of unity but this will not cause problems.

Define now a new polynomial $f^*(\mathbf{V}) := f(\mathbf{wz}\cdot\mathbf{V})$. Notice that $f^*(\mathbf{v}) = 0$. First observe that $\deg_s f^* = \deg_s f$ which we write again as δ. Further, $h(f^*) \leq h(f) + \delta h(\mathbf{wz}) \leq h(f) + \delta(h(\mathbf{w}) + h(\mathbf{z}))$. By applying Theorem 2.1 to f^* we obtain

$$\begin{aligned} h(\mathbf{v}) &\leq C_1 H + C_1\delta(h(\mathbf{w}) + h(\mathbf{z})) \\ &\leq C_1 H + C_1\delta \cdot \Big(\varepsilon\big(1 + h(\mathbf{vw})\big) + h(\mathbf{w})\Big) \\ &\leq C_1\delta\varepsilon h(\mathbf{v}) + C_1\delta(\varepsilon + (1+\varepsilon)h(\mathbf{w})) + C_1 H. \end{aligned} \tag{6.2}$$

Here it is essential that the bound of Theorem 2.1 does not depend on the field generated by the coefficients of f^*. Further,

$$\begin{aligned} h(\mathbf{x}) &\leq h(\mathbf{vw}) + \varepsilon \cdot (1 + h(\mathbf{vw})) \\ &\leq \varepsilon + (1+\varepsilon)\cdot(h(\mathbf{v}) + h(\mathbf{w})) \\ &\leq \varepsilon + (1+\varepsilon)h(\mathbf{w}) + (1+\varepsilon)h(\mathbf{v}). \end{aligned} \tag{6.3}$$

By our choice of ε we have $(1+\varepsilon)(1 - C_1\varepsilon\delta)^{-1} \leq 2$. Further,

$$h(\mathbf{w}) \leq \sum_{i=1}^{r} |\gamma_i| \cdot h(\mathbf{w}_i) \leq \frac{1}{2} r h_0.$$

By inserting this bound as well as the upper bound for $h(\mathbf{v})$ resulting from (6.2) into (6.3), we obtain

$$\begin{aligned} h(\mathbf{x}) &\leq \Big(\varepsilon + (1+\varepsilon)h(\mathbf{w})\Big)\cdot\Big(1 + 2C_1\delta\Big) + 2C_1 H \\ &\leq \Big(\varepsilon + (1+\varepsilon)rh_0/2\Big)\cdot\Big(1 + 2C_1\delta\Big) + 2C_1 H \\ &\leq 2rh_0\delta C_1 + 2C_1 H. \end{aligned} \tag{6.4}$$

This is the upper bound for $h(\mathbf{x})$ in Theorem 2.3.

We now estimate from above $[L(\mathbf{x}) : L]$ where L is the number field generated by Γ and the coefficients of f. This degree is equal to the number of distinct points among $\sigma(\mathbf{x})$ where $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. So we have to estimate from above the latter. $\mathbf{y}$, $\mathbf{v}$, $\mathbf{w}$ will be as above.

Pick $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Define $g(\mathbf{X}) := f(\mathbf{x}\cdot\mathbf{X})$. Notice that $\deg_s g = \deg_s f = \delta$. Since some integer power of $\mathbf{y}$ belongs to $\Gamma \subseteq L^2$ and σ is a

L-isomorphism, we infer that $\sigma(\mathbf{y})\mathbf{y}^{-1}$ is a root of unity. It follows that

$$h(\sigma(\mathbf{x})\mathbf{x}^{-1}) = h(\sigma(\mathbf{z})\mathbf{z}^{-1}) \leq 2h(\mathbf{z}).$$

The point $\sigma(\mathbf{x})\mathbf{x}^{-1}$ belongs to the curve defined by g. So, under the assumption

$$2h(\mathbf{z}) \leq \Big(2^{47}\delta(\log\delta)^5\Big)^{-1} \tag{6.5}$$

we deduce from Lemma 4.4,(ii) that the number of distinct points $\sigma(\mathbf{x})$ is at most

$$2^{50}\delta^2(\log\delta)^6.$$

and this is precisely the upper bound from Theorem 2.3.

It remains to prove (6.5). We have $h(\mathbf{z}) \leq \varepsilon \cdot \Big(1 + h(\mathbf{w}) + h(\mathbf{v})\Big)$ so as in (6.2) we obtain

$$h(\mathbf{z}) \leq \varepsilon \cdot \Big(1 + h(\mathbf{w}) + C_1 H + C_1\delta \cdot h(\mathbf{w}) + h(\mathbf{z})\Big)$$

implying

$$\Big(1 - \varepsilon C_1\delta\Big)h(\mathbf{z}) \leq \varepsilon \cdot \Big(\Big(1 + C_1\delta\Big) \cdot h(\mathbf{w}) + 1 + C_1 H\Big).$$

Then inserting $h(\mathbf{w}) \leq \frac{1}{2}rh_0$ and using (2.5) we get

$$h(\mathbf{z}) \leq \varepsilon \cdot \Big(C_1\delta r h_0 + 2C_1 H\Big). \tag{6.6}$$

Now our choice of ε in (2.5) implies indeed (6.5). □

Proof of Theorem 2.2. The proof is very similar to that of Theorem 2.3. We indicate only the necessary changes.

So let $\mathbf{x} \in \mathcal{C}(\overline{\mathbb{Q}}) \cap \overline{\Gamma}_\varepsilon$ with ε given by (2.4). Then $\mathbf{x} = \mathbf{y}\mathbf{z}$ with $\mathbf{y} \in \overline{\Gamma}_\varepsilon$ and $h(\mathbf{z}) < \varepsilon$. Write again $\mathbf{y} = \mathbf{v}\mathbf{w}$ with $\mathbf{v} \in \Gamma$ and $\mathbf{w} = \prod_{i=1}^r \mathbf{w}_i^{\gamma_i}$ with $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$.

Now using $h(\mathbf{z}) < \varepsilon$ we obtain instead of (6.2),

$$h(\mathbf{v}) \leq C_1\delta(h(\mathbf{w}) + \varepsilon) + C_1 H.$$

Further,

$$h(\mathbf{x}) \leq h(\mathbf{v}) + h(\mathbf{w}) + h(\mathbf{z}) \leq (1 + \delta C_1)h(\mathbf{w}) + \varepsilon + C_1 H$$

and by inserting $h(\mathbf{w}) \leq \frac{1}{2}rh_0$, we obtain

$$h(\mathbf{x}) \leq rh_0\delta C_1 + C_1 H$$

which is the bound from Theorem 2.2.

We now estimate from above $[L(\mathbf{x}) : L]$ and for this we have to estimate the number of distinct points among $\sigma(\mathbf{x})$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. As above we have

$$h(\sigma(\mathbf{x})\mathbf{x}^{-1}) = h(\sigma(\mathbf{z})\mathbf{z}^{-1}) < 2\varepsilon.$$

Thanks to our choice of ε in (2.4) we have (6.5), and our upper bound for $[L(\mathbf{x}) : L]$ follows in the same manner as above. □

7. Points in translates of algebraic groups

In the present section we prove effective results on the intersection of Γ or $\overline{\Gamma}_\varepsilon$ with a translate $\mathbf{x}_0\mathcal{H}$, where Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$, $\varepsilon > 0$, $\mathbf{x}_0 \in (\overline{\mathbb{Q}}^*)^N$ is fixed and $\mathcal{H}$ is a proper algebraic subgroup of $(\overline{\mathbb{Q}}^*)^N$. In fact we show that if $\mathbf{x}_0\mathcal{H}$ contains a point from Γ or $\overline{\Gamma}_\varepsilon$ then it contains such a point with height and degree below some effectively computable constants. Thus, it can be decided effectively whether or not $\mathbf{x}_0\mathcal{H}$ contains points from Γ or $\overline{\Gamma}_\varepsilon$.

For $\mathbf{x} = (x_1, \ldots, x_N) \in (\overline{\mathbb{Q}}^*)^N$ and an $N \times M$-matrix $A = \big(a_{ij}\big)_{1\le i\le M, 1\le j\le N}$, with $a_{ij} \in \mathbb{Z}$ we define $\mathbf{x}^A \in (\overline{\mathbb{Q}}^*)^M$ by

$$\mathbf{x}^A := (x_1^{a_{11}} \ldots x_N^{a_{N1}}, \ldots, x_1^{a_{1M}} \ldots x_N^{a_{NM}}).$$

Thus, $(\mathbf{x}^A)^B = \mathbf{x}^{AB}$ whenever the product of the matrices A, B is defined. It is well-known that for every $(N - M)$-dimensional algebraic subgroup $\mathcal{H}$ of $(\overline{\mathbb{Q}}^*)^N$ there is an integer $N \times M$-matrix A of rank M such that $\mathcal{H}$ is the set of points $\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N$ with $\mathbf{x}^A = \mathbf{1} = (1, \ldots, 1)$ (M times) (see for instance [4, Theorem 3.2.19]). Moreover, every translate of $\mathcal{H}$ can be described as the set of solutions of $\mathbf{x}^A = \mathbf{c}$ for some fixed $\mathbf{c} \in (\overline{\mathbb{Q}}^*)^M$. (See for instance again [4].)

As before, we choose a basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ of Γ modulo Γ_{tors}. Let K be the smallest number field such that $\Gamma \subset (K^*)^N$ and let S be the smallest set of places of K that contains all infinite places and such that $\Gamma \subset (\mathcal{O}_S^*)^N$. Put

$$h_0 := \max\{1, h(\mathbf{w}_1), \ldots, h(\mathbf{w}_r)\},\ d := [K : \mathbb{Q}],\ s := \#S.$$

Notice that by the product formula we have for $\mathbf{x} = (x_1, \ldots, x_N) \in \Gamma$,

$$h(\mathbf{x}) = \frac{1}{2} \sum_{v \in S} \sum_{i=1}^{N} |\log |x_i|_v|. \tag{7.1}$$

Let $A = \big(a_{ij}\big)_{1 \le i \le N,\, 1 \le j \le M}$ be an integer $N \times M$-matrix, where we do not require that A has rank M. Further, let $\mathbf{c}$ be a fixed point of $(\overline{\mathbb{Q}}^*)^M$, and δ, H reals such that

$$\max_{i,j} |a_{ij}| \le \delta, \quad \max(1, h(\mathbf{c})) \le H.$$

Let $c(d)$ be the constant from Lemma 4.3.

Our first result is as follows.

Proposition 7.1. *Assume that*

$$\mathbf{x}^A = \mathbf{c} \quad \text{in } \mathbf{x} \in \Gamma \tag{7.2}$$

is solvable. Then (7.2) has a solution $\mathbf{x}_0 \in \Gamma$ *such that*

$$h(\mathbf{x}_0) \le h_0 \cdot \big(r 4^r c(d) M \delta h_0\big)^r \cdot H.$$

In the proof we need some results on lattice points. We start with recalling a result of Schlickewei [14, Proposition 4.2].

Lemma 7.2. *Let Λ be a discrete subgroup of rank r in $\mathbb{R}^m$ and $\|\cdot\|$ a norm on $\mathbb{R}^m$. Then there exists a basis $\mathbf{a}_1, \ldots, \mathbf{a}_r$ of Λ such that for any $x_1, \ldots, x_r \in \mathbb{Z}$ we have*

$$\|x_1 \mathbf{a}_1 + \cdots + x_r \mathbf{a}_r\| \ge 4^{-r} \max\{|x_1| \|\mathbf{a}_1\|, \ldots, |x_r| \|\mathbf{a}_r\|\}. \tag{7.3}$$

Proof. Schlickewei proved this only for $\mathbb{Z}^r$ instead of arbitrary lattices Λ, but using a suitable linear transformation the above more general result follows in a straightforward way. □

In the sequel let $\|\cdot\|_l$ denote the usual l-norm defined by $\|\mathbf{x}\|_l = \big(\sum_i |x_i|^l\big)^{1/l}$ if $1 \le l < \infty$ and $\|\mathbf{x}\|_\infty = \max_i |x_i|$.

Lemma 7.3. *Let U be an $r \times k$ integer matrix of rank k and $\mathbf{m} \in \mathbb{Z}^k$. Further, let R, V be reals such that the coordinates of $\mathbf{m}$ have absolute values* *at most R and the entries of U have absolute values at most V. Suppose that the equation*

$$\mathbf{x}U = \mathbf{m} \quad \textit{in } \mathbf{x} \in \mathbb{Z}^r \tag{7.4}$$

has a solution. Then equation (7.4) has a solution $\mathbf{x}_0 \in \mathbb{Z}^r$ such that

$$\|\mathbf{x}_0\|_\infty \leq k^{k/2} V^{k-1} \max(V, R).$$

Proof. According to a result of Borosh, Flahive, Rubin and Treybig [5], (7.4) has a solution $\mathbf{x}_0$ with $\|\mathbf{x}_0\|_\infty \leq W$, where W is the maximum of the absolute values of the minors of the augmented matrix with U on the first r rows and $\mathbf{m}$ on the last row. Now our Lemma follows easily by applying Hadamard's inequality. □

Proof of Proposition 7.1. Put $s := \#S$. For any positive integer t, we denote by φ_t the group homomorphism from $(\mathcal{O}_S^*)^t$ to $\mathbb{R}^{st}$, given by

$$\varphi_t : \ \mathbf{x} \mapsto (\log |x_i|_v : v \in S, i = 1, \dots, t)\,,$$

where we have written $\mathbf{x} = (x_1, \dots, x_t)$. Further, denote by $\|\cdot\|$ the 1-norm on $\mathbb{R}^{Ns}$ and by $\|\cdot\|^*$ the 1-norm on $\mathbb{R}^{Ms}$.

The kernel of $\varphi := \varphi_N|_\Gamma$ is Γ_{tors}, and the image Λ of φ in $\mathbb{R}^{Ns}$ is a discrete subgroup of rank r. Equation (7.2) can be written in the form

$$\mathbf{y}B = \mathbf{b} \ \text{ in } \ \mathbf{y} \in \Lambda, \tag{7.5}$$

where $\mathbf{b} := \varphi_M(\mathbf{c})$ and

$$B := \begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix}$$

is an integer $Ns \times Ms$-matrix. Notice that $\varphi_M(\mathbf{w}^A) = \varphi_N(\mathbf{w})B$ for $\mathbf{w} \in (\mathcal{O}_S^*)^N$. By assumption, equation (7.5) is solvable, and in view of (7.1), we need to find a solution $\mathbf{y}_0$ of (7.5) such that $\|\mathbf{y}_0\|$ is at most two times the upper bound from Proposition 7.1.

Put $B(\Lambda) := \{\mathbf{y}B \ : \ \mathbf{y} \in \Lambda\}$. Clearly, $B(\Lambda)$ is a discrete subgroup in $\mathbb{R}^{Ms}$. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be the images of the chosen basis $\mathbf{w}_1, \dots, \mathbf{w}_r$ of Γ modulo Γ_{tors} under φ. Then $\mathbf{v}_1, \dots, \mathbf{v}_r$ form a basis of Λ,

$$\|\mathbf{v}_i\| \leq 2h_0 \text{ for } i = 1, \dots, r, \tag{7.6}$$

and $\mathbf{v}_1B,\dots,\mathbf{v}_rB$ form a system of generators for $B(\Lambda)$. Suppose that the rank of $B(\Lambda)$ is k. By Lemma 7.2 there exists a basis $\mathbf{a}_1,\dots,\mathbf{a}_k$ of $B(\Lambda)$, such that for every $\mathbf{x}=n_1\mathbf{a}_1+\cdots+n_k\mathbf{a}_k\in B(\Lambda)$ with $n_1,\dots,n_k\in\mathbb{Z}$ we have

$$\|\mathbf{x}\|^*\geq 4^{-k}\max(|n_1|\|\mathbf{a}_1\|^*,\dots,|n_k|\|\mathbf{a}_k\|^*). \tag{7.7}$$

Since $\mathbf{b}\in B(\Lambda)$, there exist $m_1,\dots,m_k\in\mathbb{Z}$, such that $\mathbf{b}=m_1\mathbf{a}_1+\cdots+m_k\mathbf{a}_k$. Using $\|\mathbf{b}\|^*=2h(\mathbf{c})\leq 2H$ (in view of (7.1)), $\|\mathbf{a}_i\|^*\geq 2c(d)^{-1}$ (by Lemma 4.3, (7.1) and the fact that $\mathbf{a}_i\in\varphi_M((\mathcal{O}_S^*)^M)$ and (7.7) we have

$$|m_i|\leq 4^k\frac{\|\mathbf{b}\|^*}{\|\mathbf{a}_i\|^*}\leq 4^kc(d)\cdot H\quad\text{for } i=1,\dots,k. \tag{7.8}$$

Further, since $\mathbf{v}_iB\in B(\Lambda)$ we can write $\mathbf{v}_iB=\sum_{j=1}^k u_{ij}\mathbf{a}_j$ for $i=1,\dots,r$. Using $\|\mathbf{v}_iB\|^*=2h(\mathbf{w}_i^A)\leq 2M\delta h_0$ and again $\|\mathbf{a}_j\|^*\geq 2c(d)^{-1}$, (7.7) we get

$$|u_{ij}|\leq 4^kc(d)M\delta h_0\quad\text{for } i=1,\dots,r,\ j=1,\dots,k. \tag{7.9}$$

Let $\mathbf{y}$ be a solution of (7.5). Then $\mathbf{y}\in\Lambda$ and so we have $\mathbf{y}=\sum_{i=1}^r\mu_i\mathbf{v}_i$ with $\mu_i\in\mathbb{Z}$ for $i=1,\dots,r$. Using that on the one hand $\mathbf{b}=m_1\mathbf{a}_1+\cdots+m_k\mathbf{a}_k$ and on the other hand

$$\mathbf{b}=\mathbf{y}B=\sum_{i=1}^r\mu_i(\mathbf{v}_iB)=\sum_{i=1}^r\mu_i\left(\sum_{j=1}^k u_{ij}\mathbf{a}_j\right)=\sum_{j=1}^k\left(\sum_{i=1}^r u_{ij}\mu_i\right)\mathbf{a}_j,$$

we obtain

$$\sum_{i=1}^r u_{ij}\mu_i=m_j\quad\text{for } j=1,\dots,k. \tag{7.10}$$

Further we have (7.9) and (7.8) to bound the coefficients and the right hand side of the system of linear equations (7.10). On applying Lemma 7.3 with $V=4^kc(d)M\delta h_0$, $R=4^kc(d)H$, we see that the system (7.10) has a solution $\mu\in\mathbb{Z}^r$ with

$$\sum_{i=1}^r|\mu_i|\leq rk^{k/2}V^{k-1}\max(V,R)\leq\big(r\cdot 4^rc(d)M\delta h_0\big)^r\cdot H.$$

Now in view of (7.6), the vector $\mathbf{y}_0=\sum_{i=1}^r\mu_i\mathbf{v}_i$ is a solution to (7.5) such that

$$\|\mathbf{y}_0\|\leq 2h_0\cdot\big(r4^rc(d)M\delta h_0\big)^r\cdot H$$

and this is indeed twice the bound of our Proposition. □

As before, Γ is a finitely generated subgroup of $(\overline{\mathbb{Q}}^*)^N$ of rank r, A an integer $N \times M$-matrix and $\mathbf{c}$ a point in $(\overline{\mathbb{Q}}^*)^M$. The set $\overline{\Gamma}_\varepsilon$ ($\varepsilon > 0$) is defined as in the Introduction. We assume that A has rank $N - P$.

Proposition 7.4. *Let $\varepsilon > 0$. There exist effectively computable constants C_6, C_7 depending only on $\Gamma, A, \mathbf{c}, \varepsilon$, such that if*

$$\mathbf{x}^A = \mathbf{c} \text{ in } \mathbf{x} \in \overline{\Gamma}_\varepsilon \tag{7.11}$$

is solvable, then there exists $\mathbf{x}_0 \in \overline{\Gamma}_\varepsilon$ with

$$\mathbf{x}_0^A = \mathbf{c}, \quad h(\mathbf{x}_0) \leq C_6, \quad [\mathbb{Q}(\mathbf{x}_0) : \mathbb{Q}] \leq C_7. \tag{7.12}$$

We deduce Proposition 7.4 from Proposition 7.5 below.

Proposition 7.5. *Let $\mathbf{c}_0 \in (\overline{\mathbb{Q}}^*)^N$, B an integer $P \times N$ matrix of rank P and $\varepsilon > 0$. There exist effectively computable constants C_8, C_9 depending only on $\Gamma, B, \mathbf{c}_0, \varepsilon$, such that if there is $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ with*

$$\mathbf{c}_0 \mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \tag{7.13}$$

then there exists $\mathbf{t}_0 \in (\overline{\mathbb{Q}}^*)^P$ such that

$$\mathbf{c}_0 \mathbf{t}_0^B \in \overline{\Gamma}_\varepsilon, \quad h(\mathbf{t}_0) \leq C_8, \quad [\mathbb{Q}(\mathbf{t}_0) : \mathbb{Q}] \leq C_9. \tag{7.14}$$

Proposition 7.5 $\Longrightarrow$ Proposition 7.4. Let $A, \mathbf{c}, \varepsilon$ be as in Proposition 7.4. Let $\mathbf{x} \in \overline{\Gamma}_\varepsilon$ with $\mathbf{x}^A = \mathbf{c}$. There are matrices $U_1 \in \mathrm{GL}_N(\mathbb{Z})$, $U_2 \in \mathrm{GL}_M(\mathbb{Z})$ such that

$$U_1 A U_2 = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where D is an invertible integer $(N-P) \times (N-P)$-matrix. Let $\mathbf{x}^* := \mathbf{x}^{U_1^{-1}}$. Write $\mathbf{x}^* = (\mathbf{s}, \mathbf{t})$ where $\mathbf{s} \in (\overline{\mathbb{Q}}^*)^{N-P}$, $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$. We can decompose $\mathbf{x}^*$ as $(\mathbf{s}, \mathbf{1}) \cdot (\mathbf{1}, \mathbf{t})$, where in the first component $\mathbf{1}$ stands for P ones and in the second component for $N - P$ ones. Notice that $\mathbf{s}^D = \mathbf{c}'$, where $\mathbf{c}'$ consists of the first $N - P$ coordinates of $\mathbf{c}^{U_2}$ and hence $\mathbf{s}^\Delta = \mathbf{c}'^{\Delta D^{-1}}$, where $\Delta = \det D$. This shows that $\mathbf{s}$ belongs to a finite, effectively determinable set depending only on $A, \mathbf{c}$.

Put $\mathbf{c}_0 := (\mathbf{s}, \mathbf{1})^{U_1^{-1}}$, and let B be the matrix consisting of the last P rows of U_1. Then B is a $P \times N$-matrix of rank P. Notice that $\mathbf{c}_0^A = \mathbf{c}$, $BA = 0$ and $\mathbf{c}_0\mathbf{t}^B = \mathbf{x} \in \overline{\Gamma}_\varepsilon$.

By Proposition 7.5, there is $\mathbf{t}_0 \in (\overline{\mathbb{Q}}^*)^P$ with $\mathbf{c}_0\mathbf{t}_0^B \in \overline{\Gamma}_\varepsilon$, $h(\mathbf{c}_0) \le C_8$, $[\mathbb{Q}(\mathbf{c}_0) : \mathbb{Q}] \le C_9$, where C_8, C_9 are effectively computable in terms of $B, \mathbf{c}_0, \Gamma, \varepsilon$.

Now put $\mathbf{x}_0 := \mathbf{c}_0\mathbf{t}_0^B$. Then $\mathbf{x}_0 \in \overline{\Gamma}_\varepsilon$, $\mathbf{x}_0^A = \mathbf{c}_0^A\mathbf{t}_0^{BA} = \mathbf{c}$ and $h(\mathbf{x}_0) \le C_6$, $[\mathbb{Q}(\mathbf{x}_0) : \mathbb{Q}] \le C_7$ with C_6, C_7 effectively computable in terms of $B, \mathbf{c}_0, \Gamma, \varepsilon$. Since $\mathbf{c}_0 = (\mathbf{s}, \mathbf{1})^{U_1^{-1}}$ belongs to to a finite set effectively computable in terms of $\mathbf{c}, A$ and since B is effectively computable in terms of A, we may choose C_6, C_7 to be effectively computable in terms of $A, \mathbf{c}, \Gamma, \varepsilon$. This proves Proposition 7.4. □

We proceed to prove Proposition 7.5. Let K' be the number field generated by the coordinates of $\mathbf{c}_0$ and by the coordinates of a system of generators for Γ.

Lemma 7.6. *Assume there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *with (7.13). Then there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *such that*

$$\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \quad \exists m \in \mathbb{Z}_{>0} \text{ with } \mathbf{t}^m \in (K'^*)^P. \tag{7.15}$$

Proof. First observe that if $\mathbf{u} \in \overline{\Gamma}_\varepsilon$ and $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/K')$, then $\sigma(\mathbf{u}) \in \overline{\Gamma}_\varepsilon$. Indeed, write $\mathbf{u} = \mathbf{u}_1\mathbf{u}_2$ with $\mathbf{u}_1 \in \overline{\Gamma}$, $h(\mathbf{u}_2) < \varepsilon$. There is $k \in \mathbb{Z}_{>0}$ such that $\sigma(\mathbf{u}_1)^k = \mathbf{u}_1^k \in \Gamma$, implying that $\sigma(\mathbf{u}_1) \in \overline{\Gamma}$. Further, $h(\sigma(\mathbf{u}_2)) < \varepsilon$. So $\sigma(\mathbf{u}) \in \overline{\Gamma}_\varepsilon$.

Now let $\sigma_1, \dots, \sigma_m$ be the distinct K'-isomorphic embeddings of $K'(\mathbf{t})$ into $\overline{\mathbb{Q}}$. Take

$$\mathbf{t}' := \left(\prod_{i=1}^m \sigma_i(\mathbf{t})\right)^{1/m}.$$

This is determined only up to a factor in $(\overline{\mathbb{Q}}^*_{\mathrm{tors}})^P$, but this is not causing any problem. Write $\mathbf{c}_0\mathbf{t}^B = \mathbf{u}_1\mathbf{u}_2$ with $\mathbf{u}_1 \in \overline{\Gamma}$, $h(\mathbf{u}_2) < \varepsilon$. Then

$$\mathbf{c}_0\mathbf{t}'^B = \left(\prod_{i=1}^m \sigma_i(\mathbf{u}_1)\right)^{1/m} \left(\prod_{i=1}^m \sigma_i(\mathbf{u}_2)\right)^{1/m},$$

which belongs to $\overline{\Gamma}_\varepsilon$. Clearly, $\mathbf{t}^m \in (K'^*)^P$. □

Let S' be the smallest set of places of K', containing all infinite places and such that $\mathbf{c}_0 \in (\mathcal{O}_{S'}^*)^N$, $\Gamma \subseteq (\mathcal{O}_{S'}^*)^N$. Put $s' := \#S'$.

Lemma 7.7. *Assume there exists* $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ *with (7.13). Then there exists* $\mathbf{t}$ *with*

$$\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon, \quad \mathbf{t} \in (\overline{\mathcal{O}_{S'}^*})^P, \tag{7.16}$$

where $\overline{\mathcal{O}_{S'}^*} = \{\mathbf{x} \in \overline{\mathbb{Q}}^* : \exists m \in \mathbb{Z}_{>0}$ *with* $\mathbf{x}^m \in \mathcal{O}_{S'}^*\}$.

Proof. Let $\mathbf{t} \in (\overline{\mathbb{Q}}^*)^P$ be as in (7.15), i.e., $\mathbf{t}^m \in (K'^*)^P$ for some $m \in \mathbb{Z}_{>0}$. Write

$$\mathbf{c}_0\mathbf{t}^B = \mathbf{y}\mathbf{z} \ \text{ with } \ \mathbf{y} \in \overline{\Gamma}, \quad h(\mathbf{z}) < \varepsilon. \tag{7.17}$$

Let $n \in \mathbb{Z}_{>0}$ be such that $\mathbf{y}^n \in \Gamma$ and let k be any positive multiple of $\operatorname{lcm}(m,n)$. Thus

$$\mathbf{z}^k = \mathbf{c}_0^k(\mathbf{t}^k)^B\mathbf{y}^{-k} \in (K'^*)^P. \tag{7.18}$$

Write $\mathbf{t} = (t_1, \ldots, t_P)$. By the Dirichlet-Chevalley-Weil S-unit theorem, there are $\varepsilon_1, \ldots, \varepsilon_P \in \mathcal{O}_{S'}^*$ such that

$$\left| \log |\varepsilon_i|_v - \log \left(\frac{|t_i^k|_v}{\left\{\prod_{w \in S} |t_i^k|_w\right\}^{1/s}} \right) \right| \le C \quad \text{for } \ i = 1, \ldots, P, \ v \in S',$$

where C is an effectively computable constant depending only on K', S', and independent of k. Now define

$$\begin{aligned} &\mathbf{t}' := (\varepsilon_1^{1/k}, \ldots, \varepsilon_P^{1/k}), \quad \mathbf{z}' := \mathbf{c}_0(\mathbf{t}'^B)\mathbf{y}^{-1}, \\ &\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N) := (\varepsilon_1, \ldots, \varepsilon_P)^B. \end{aligned} \tag{7.19}$$

(with a suitable choice of the k-th roots). Write $\mathbf{z} := (z_1, \ldots, z_P)$, $\mathbf{z}' := (z_1', \ldots, z_P')$, $\mathbf{v} := (v_1, \ldots, v_N) = \mathbf{t}^B$, $(\mathbf{c}_0\mathbf{y}^{-1})^k := (\alpha_1, \ldots, \alpha_P)$. Then since $\alpha_1, \ldots, \alpha_P \in \mathcal{O}_{S'}^*$ (by our choice of k and S') we have for $i = 1, \ldots, N$,

$v \in S'$,

$$\begin{aligned}
&\left|\log|{z'_i}^k|_v - \log\left(\frac{|z_i^k|_v}{\left\{\prod_{w\in S'}|z_i^k|_w\right\}^{1/s'}}\right)\right| \\
&\quad = \left|\log|\alpha_i\eta_i|_v - \log\left(\frac{|\alpha_i v_i^k|_v}{\left\{\prod_{w\in S'}|\alpha_i v_i^k|_w\right\}^{1/s'}}\right)\right| \\
&\quad = \left|\log|\eta_i|_v - \log\left(\frac{|v_i^k|_v}{\left\{\prod_{w\in S'}|v_i^k|_w\right\}^{1/s'}}\right)\right| \le C',
\end{aligned}$$

where C' is an effectively computable constant depending only on K', S' and B, but which is independent of k. Together with the product formula this implies

$$\left|\log|{z'_i}^k|_v - \log|z_i^k|_v - \frac{1}{s'}\sum_{w\notin S'}\log|z_i^k|_w\right| \le C'$$

for $v \in S'$, $i = 1, \dots, N$. Now we get

$$\begin{aligned}
h(\mathbf{z}'^k) &= \sum_{i=1}^N \sum_{v\in S'} \max\Big(0, \log|{z'_i}^k|_v\Big) \\
&\le \sum_{i=1}^N \sum_{v\in S'} \max\left(0, C' + \log|{z_i}^k|_v + \frac{1}{s'}\sum_{w\notin S'}\log|z_i^k|_w\right) \\
&\le Ns'C' + \sum_{i=1}^N \sum_{v\in S'} \max\big(0, \log|{z_i}^k|_v\big) + \sum_{i=1}^N \sum_{v\notin S'} \max\big(0, \log|{z_i}^k|_v\big) \\
&= Ns'C' + \sum_{i=1}^N h(z_i^k) = Ns'C' + h(\mathbf{z}^k).
\end{aligned}$$

Consequently,

$$h(\mathbf{z}') \le h(\mathbf{z}) + \frac{Ns'C'}{k}.$$

By assumption, $h(\mathbf{z}) < \varepsilon$. We had chosen k to be any positive multiple of $\mathrm{lcm}(m,n)$. By choosing k large enough, we can achieve that $h(\mathbf{z}') < \varepsilon$. Now from our choice of $\mathbf{t}'$ in (7.19) it follows that $\mathbf{t}' \in (\overline{\mathcal{O}_S^*})^P$ and $\mathbf{c}_0{\mathbf{t}'}^B = \mathbf{y}\mathbf{z}' \in \overline{\Gamma}_\varepsilon$. This proves Lemma 7.7. □

The proof of Proposition 7.5 rests upon linear programming.

Define the group

$$G := \left\{ \mathbf{y}\mathbf{t}^B \ : \ \mathbf{y} \in \overline{\Gamma}, \mathbf{t} \in (\overline{\mathcal{O}_{S'}^*})^P \right\}.$$

This is a group of finite rank q. Choose a maximal multiplicatively independent subset $\mathbf{t}_1, \dots, \mathbf{t}_s$ of $(\mathcal{O}_{S'}^*)^P$. Then $\mathbf{u}_1 := \mathbf{t}_1^B, \dots, \mathbf{u}_s := \mathbf{t}_s^B$ are multiplicatively independent since $\operatorname{rank} B = P$. Choose $\mathbf{u}_{s+1}, \dots, \mathbf{u}_q \in \Gamma$ such that $\{\mathbf{u}_1, \dots, \mathbf{u}_q\}$ form a maximal multiplicatively independent subset of G. After a suitable choice of roots of $\mathbf{u}_1, \dots, \mathbf{u}_q$, we may express G as

$$G = \left\{ \boldsymbol{\rho}\mathbf{u}_1^{\xi_1} \dots \mathbf{u}_q^{\xi_q} \ : \ \boldsymbol{\rho} \in (\overline{\mathbb{Q}}_{\text{tors}}^*)^N, \xi_1, \dots, \xi_q \in \mathbb{Q} \right\}.$$

We are searching for $\mathbf{t} \in (\overline{\mathcal{O}_{S'}}^*)^P$ such that

$$\mathbf{c}_0\mathbf{t}^B = \mathbf{y}\mathbf{z}, \quad \mathbf{y} \in \overline{\Gamma}, \quad h(\mathbf{z}) < \varepsilon.$$

For such $\mathbf{t}$ we have $\mathbf{z} = \mathbf{c}_0\mathbf{y}^{-1}\mathbf{t}^B \in \mathbf{c}_0 G$. So we are searching for $\mathbf{z} \in \mathbf{c}_0 G$ with $h(\mathbf{z}) < \varepsilon$.

We give an expression for the height of an element $\mathbf{z} \in \mathbf{c}_0 G$. Such an element can be expressed as

$$\mathbf{z} = \mathbf{c}_0\boldsymbol{\rho}\mathbf{u}_1^{\xi_1} \dots \mathbf{u}_q^{\xi_q} \ \text{ with } \ \boldsymbol{\rho} \in (\overline{\mathbb{Q}}_{\text{tors}}^*)^N, \xi_1, \dots, \xi_q \in \mathbb{Q}.$$

Write $\boldsymbol{\xi} := (\xi_1, \dots, \xi_q)$. Let k be a positive integer such that $\boldsymbol{\rho}^k = \mathbf{1}$, $k\boldsymbol{\xi} \in \mathbb{Z}^q$. Further, write $\mathbf{u}_i = (u_{i1}, \dots, u_{iN})$ $(i = 1, \dots, r)$, $\mathbf{c}_0 = (c_{01}, \dots, c_{0N})$. Then

$$\begin{aligned} h(\mathbf{z}) &= \frac{1}{k}h(\mathbf{z}^k) = \frac{1}{k}\sum_{i=1}^N \sum_{v \in S'} \max\left(0, k\log|c_{0i}|_v + \sum_{j=1}^q k\xi_j \log|u_{ij}|_v\right) \\ &= \sum_{i=1}^N \sum_{v \in S'} \max\left(0, \log|c_{0i}|_v + \sum_{j=1}^q \xi_j \log|u_{ij}|_v\right) \qquad (7.20) \\ &= \frac{1}{2}\sum_{i=1}^N \sum_{v \in S'} \left|\log|c_{0i}|_v + \sum_{j=1}^q \xi_j \log|u_{ij}|_v\right| =: f(\boldsymbol{\xi}), \end{aligned}$$

where we have used $\sum_{v \in S'} \log|c_{0i}|_v = 0$, $\sum_{v \in S'} \log|u_{ij}|_v = 0$ for all i, j. The function f can be extended to $\mathbb{R}^q$. We prove some properties of this function.

Lemma 7.8. (i) *For every $R \geq 0$, the set $\{\boldsymbol{\xi} \in \mathbb{R}^q : f(\boldsymbol{\xi}) \leq R\}$ is compact with respect to the topology in $\mathbb{R}^q$.*
(ii) *There is an effectively computable constant $C > 0$ such that $|f(\boldsymbol{\xi}_1) - f(\boldsymbol{\xi}_2)| \leq C\|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|_\infty$ for all $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathbb{R}^q$.*

Proof. (i). We can express $f(\boldsymbol{\xi})$ as $\|\alpha(\boldsymbol{\xi})\|$ where $\|\cdot\|$ is a norm on $\mathbb{R}^{Ns}$ and α an injective affine map from $\mathbb{R}^q$ to $\mathbb{R}^{Ns}$. So our set under consideration is homeomorphic to a closed subset of a compact set, hence compact.

(ii). Obvious. □

Lemma 7.9. *The function f assumes a minimum on $\mathbb{R}^r$ and it is possible to determine effectively*

$$\varepsilon_0 := \min\{f(\boldsymbol{\xi}) \ : \ \boldsymbol{\xi} \in \mathbb{R}^q\}$$

and $\boldsymbol{\xi}_0$ with $f(\boldsymbol{\xi}_0) = \varepsilon_0$.

Proof. It clearly suffices to prove that f assumes a minimum on

$$D := \{\boldsymbol{\xi} \in \mathbb{R}^q \ : \ f(\boldsymbol{\xi}) \leq f(\mathbf{0})\}$$

and to determine the minimum of f on D and a point in D where this minimum is assumed. By Lemma 7.8,(i) the set D is compact, so f does indeed assume its minimum on D.

We can rewrite f as

$$f(\boldsymbol{\xi}) = \max\left(L_1(\boldsymbol{\xi}) + \beta_1, \ldots, L_A(\boldsymbol{\xi}) + \beta_A\right),$$

where $L_1, \ldots, L_A$ are linear forms with real coefficients and $\beta_1, \ldots, \beta_A \in \mathbb{R}$. For $i = 1, \ldots, A$, let

$$D_i := \{\boldsymbol{\xi} \in D \ : \ L_i(\boldsymbol{\xi}) + \beta_i \geq L_j(\boldsymbol{\xi}) + \beta_j \ \text{ for } \ j = 1, \ldots, A, j \neq i\}.$$

The set D_i is a closed subset of D, hence compact. Thus D_i is a compact polytope. Notice that $f(\boldsymbol{\xi}) = L_i(\boldsymbol{\xi}) + \beta_i$ for $\boldsymbol{\xi} \in D_i$. From the theory of linear programming it follows that f assumes its minimum on D_i in a vertex of D_i. The vertices of D_i can be determined effectively. So we can effectively determine $\varepsilon_i := \min\{L_i(\boldsymbol{\xi}) + \beta_i \ : \ \boldsymbol{\xi} \in \mathcal{D}_i\}$ and $\boldsymbol{\xi}_i \in D_i$ with $f(\boldsymbol{\xi}_{\mathbf{i}}) = \varepsilon_i$.

Now $\varepsilon_0 = \min(\varepsilon_1, \ldots, \varepsilon_A)$, and $f(\boldsymbol{\xi}_0) = \varepsilon_0$, where $\boldsymbol{\xi}_0$ is the point $\boldsymbol{\xi}_i$ among $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_A$ such that $\varepsilon_i = \varepsilon_0$. □

Proof of Proposition 7.5. Assume that there exists $\mathbf{t} \in (\overline{\mathcal{O}_{S'}^*})^P$ such that $\mathbf{c}_0\mathbf{t}^B \in \overline{\Gamma}_\varepsilon$. Write $\mathbf{c}_0\mathbf{t}^B = \mathbf{y}\mathbf{z}$ with $\mathbf{y} \in \overline{\Gamma}$, $h(\mathbf{z}) < \varepsilon$. Then $\mathbf{z} = \mathbf{c}_0\mathbf{y}^{-1}\mathbf{t}^B = \mathbf{c}_0\boldsymbol{\rho}\mathbf{u}_1^{\xi_1}\dots\mathbf{u}_q^{\xi_q}$ with $\boldsymbol{\rho} \in (\overline{\mathbb{Q}}_{\mathrm{tors}}^*)^N$ and $\xi_1,\dots,\xi_q \in \mathbb{Q}$. So $h(\mathbf{z}) = f(\boldsymbol{\xi}) < \varepsilon$ and therefore, $\varepsilon > \varepsilon_0$. Let C be the constant from Lemma 7.8,(ii) and define the integer k by

$$k := \left[\frac{2C}{\varepsilon - \varepsilon_0}\right] + 1. \tag{7.21}$$

Let $\boldsymbol{\xi}_0$ be as in Lemma 7.9 and write $\boldsymbol{\xi}_0 = (\xi_{01},\dots,\xi_{0q})$. Define integers $n_1,\dots,n_q$ by

$$|k\xi_{0i} - n_i| < 1 \quad (i = 1,\dots,q)$$

and let

$$\mathbf{t}_0 = \mathbf{t}_1^{n_1/k}\dots\mathbf{t}_s^{n_s/k}, \quad \mathbf{z}_0 = \mathbf{c}_0\mathbf{u}_1^{n_1/k}\dots\mathbf{u}_q^{n_q/k}.$$

By Lemma 7.8,(ii) and (7.21),

$$\begin{aligned} h(\mathbf{z}_0) = f\left(\frac{n_1}{k},\dots,\frac{n_q}{k}\right) &\le f(\boldsymbol{\xi}_0) + C\max_{1\le i\le q}\left|\xi_{0i} - \frac{n_i}{k}\right| \\ &\le \varepsilon_0 + \frac{C}{k} < \varepsilon_0 + \frac{C(\varepsilon-\varepsilon_0)}{2C} < \varepsilon. \end{aligned}$$

Further,

$$\begin{aligned} \mathbf{c}_0\mathbf{t}_0^B &= \mathbf{u}_{s+1}^{-n_{s+1}/k}\dots\mathbf{u}_q^{-n_q/k}\mathbf{z}_0 \in \overline{\Gamma}_\varepsilon, \\ h(\mathbf{t}_0) &\le \sum_{i=1}^{s}\left|\frac{n_i}{k}\right| h(\mathbf{t}_i) \le \sum_{i=1}^{q}(|\xi_{0i}| + \frac{1}{k})h(\mathbf{t}_i) \le C_8 \end{aligned}$$

and $\mathbf{t}_0^k \in (K'^*)^P$, implying $[\mathbb{Q}(\mathbf{t}_0):\mathbb{Q}] \le C_9$. The quantities C, ε_0, as well as $\mathbf{t}_1,\dots,\mathbf{t}_s$ are effectively computable in terms of $\Gamma, B, \mathbf{c}_0$, while k is effectively computable in terms of these parameters and also ε. Hence the constants C_8, C_9 are indeed effectively computable in terms of $\Gamma, B, \mathbf{c}_0, \varepsilon$, but they have been defined only for $\varepsilon > \varepsilon_0$. For completeness, we define $C_8 := 1$, $C_9 := 1$ if $\varepsilon \le \varepsilon_0$. Then clearly, Proposition 7.5 holds with these C_8, C_9. □

8. Proof of Theorem 2.4

We write $\mathbf{x}^{\mathbf{a}} := x_1^{a_1}\cdots x_N^{a_N}$ for $\mathbf{x} = (x_1,\dots,x_N) \in (\overline{\mathbb{Q}}^*)^N$, $\mathbf{a} = (a_1,\dots,a_N) \in \mathbb{Z}^N$.

By assumption

$$\mathcal{X} = \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ : \ f_1(\mathbf{x}) = 0, \dots, f_m(\mathbf{x}) = 0 \right\},$$

where each polynomial f_i belongs to $\overline{\mathbb{Q}}[X_1, \dots, X_N]$ and has at most three non-zero terms. Further, $\deg f_i \le \delta$ and $\max(1, h(f_i)) \le H$ for $i = 1, \dots, m$. Without loss of generality we assume that f_i $(i = 1, \dots, n)$ are trinomials and f_i $(i = n+1, \dots, m)$ are binomials, where $0 \le n \le m$. Thus, by dividing each f_i by one of its terms we see that $\mathcal{X}$ is given by equations

$$\alpha_{i1}\mathbf{x}^{\mathbf{a}_{i1}} + \alpha_{i2}\mathbf{x}^{\mathbf{a}_{i2}} = 1 \ \ (i = 1, \dots, n), \quad \alpha_{i1}\mathbf{x}^{\mathbf{a}_{i1}} = 1 \ \ (i = n+1, \dots, m), \tag{8.1}$$

where $\alpha_{ij} \in \overline{\mathbb{Q}}^*$, $\mathbf{a}_{ij} \in \mathbb{Z}^N$ for $(i,j) \in I := \{(1,1), \dots, (m,1), (1,2), \dots, (n,2)\}$. We should observe here that since each polynomial f_i has total degree at most δ, we have estimates for the maximum norm and the sum norm,

$$\|\mathbf{a}_{ij}\|_\infty \le \delta, \quad \|\mathbf{a}_{ij}\|_1 \le 2\delta \quad \text{for } (i,j) \in I. \tag{8.2}$$

Clearly the stabilizer of $\mathcal{X}$ is given by

$$\begin{aligned} \mathcal{H} := \mathrm{Stab}(\mathcal{X}) &= \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ \mid \mathbf{x}^{\mathbf{a}_{ij}} = 1 \text{ for } i = 1, \dots, m, \ j = 1, 2 \right\} \\ &= \left\{ \mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \ \mid \mathbf{x}^{A} = \mathbf{1} \right\}, \end{aligned} \tag{8.3}$$

where A is the $N \times (2m - n)$ matrix with columns $\mathbf{a}_{ij}$, $((i,j) \in I)$.

Let $i \in \{1, \dots, n\}$. Let $\mathbf{x} \in \mathcal{X} \cap \Gamma$. Denote by G be the subgroup of K^* generated by $\xi_1 := \mathbf{w}_1^{a_{i1}}, \dots, \xi_r := \mathbf{w}_r^{a_{i1}}$. Then for the quantity Q defined by (4.1), we have

$$Q := \prod_{j=1}^{r} \max\left(1, h(\mathbf{w}_1^{a_{ij}})\right) \le (\delta h_0)^r.$$

We apply Lemma 4.2 to the equation $\alpha_{i1}x + \alpha_{i2}y = 1$ in $x \in G$, $y \in \mathcal{O}_S^*$. Notice that $\max(1, h(\alpha_{i1}), h(\alpha_{i2})) \le H$. Replacing Q by $(\delta h_0)^r$ in the expression for C_5, we obtain a constant bounded above by C^*. In fact, this can be shown by a straightforward computation, using that the term with the maximum in C_5 is bounded above by $46r^2 \log^* \max(ds\mathbf{N}, \delta h_0)$. It follows that

$$h(\mathbf{x}^{\mathbf{a}_{i,j}}) < C^* H \text{ for } i = 1, \dots, n, \ j = 1, 2. \tag{8.4}$$

We clearly also have $h(\mathbf{x}^{\mathbf{a}_{i1}}) = h(\alpha_{i1}^{-1}) \leq H$ for $i = n+1, \ldots, m$. So we have (8.4) for $(i,j) \in I$. This implies

$$\mathbf{x}^A = \mathbf{c}, \tag{8.5}$$

where A is the $N \times (2m-n)$-matrix from above and where $\mathbf{c} \in (K^*)^{2m-n}$ with

$$h(\mathbf{c}) \leq (2m-n)C^*H \leq 2mC^*H. \tag{8.6}$$

Further, the entries of A have absolute values at most δ, and of each column of A the sum of its absolute values is at most 2δ.

We first assume that the stabilizer $\mathcal{H}$ is finite. Then A has rank N. Suppose for convenience that the first N columns, $\mathbf{a}_1, \ldots, \mathbf{a}_N$, say, of A form an invertible matrix D, with determinant Δ. Let $\mathbf{c}'$ consist of the first N coordinates of $\mathbf{c}$. Then $\mathbf{x}^\Delta = \mathbf{c}'^{\Delta D^{-1}}$. By Hadamard's inequality and (8.2), the entries of ΔD^{-1} have absolute value at most

$$\max_{1 \leq i \leq N} \prod_{j \neq i} \|\mathbf{a}_j\|_2 \leq (2\delta)^{N-1}. \tag{8.7}$$

Further, $h(\mathbf{c}') \leq NC^*H$. So $h(\mathbf{x}) \leq N(2\delta)^{N-1}C^*H = C_2H$. This proves part (i).

We now assume that $\mathcal{H}$ is infinite. Notice that we have to consider finitely many systems (8.5) as $\mathbf{c}$ runs through a finite set. If such a system has a solution $\mathbf{x}$ with $\mathbf{x} \in \mathcal{X}$, then each element of the translate $\mathbf{x}\mathcal{H}$ is also a solution of this system. On the other hand $\mathbf{x}\mathcal{H} \subset \mathcal{X}$. Thus we have proved that $\mathcal{X} \cap \Gamma$ is contained in some finite union of translates

$$\mathbf{x}_1\mathcal{H} \cup \cdots \cup \mathbf{x}_T\mathcal{H}.$$

with $\mathbf{x}_i\mathcal{H} \subset \mathcal{X}$ for $i = 1, \ldots, T$.

Fix any of these translates, which means that we have fixed one of the systems from (8.5). By assumption this system has a solution in $\mathbf{x} \in \Gamma$. Now by Proposition 7.1 (with $M = 2m-n \leq 2m$) and (8.6), this fixed system of type (8.5) has a solution $\mathbf{x} \in \Gamma$ such that

$$h(\mathbf{x}) \leq h_0\big(2r4^r c(d) m\delta h_0\big)^r \cdot 2mC^*H \leq C_3H.$$

□

9. Proof of Theorems 2.5 and 2.6

Proof of Theorem 2.6. Let $\mathbf{x} \in \mathcal{X}(\overline{\mathbb{Q}}) \cap C(\overline{\Gamma}, \varepsilon)$, with the value of ε given in (2.11).

As before, we write $\mathbf{x} = \mathbf{yz}$ with $\mathbf{y} \in \overline{\Gamma}$ and $\mathbf{z} \in (\overline{\mathbb{Q}}^*)^2$ with $h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))$ and we may further split up $\mathbf{y}$ as $\mathbf{y} = \mathbf{vw}$ with $\mathbf{v} \in \Gamma$, $\mathbf{w} = \prod_{i=1}^r \mathbf{w}_i^{\gamma_i}$, where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$. Define new polynomials $f_i^*(\mathbf{V}) := f_i(\mathbf{wz} \cdot \mathbf{V})$ $(i = 1, \dots, m)$ and let $\mathcal{X}^*$ be the variety given by $f_i^* = 0$ for $i = 1, \dots, m$, i.e., $\mathcal{X}^* := (\mathbf{wz})^{-1}\mathcal{X}$. Then $\mathbf{v} \in \mathcal{X}^* \cap \Gamma$. Notice that $\deg f_i^* \leq \delta$, and $\max(1, h(f_i^*)) \leq H + \delta h(\mathbf{wz}) \leq H + \delta(h(\mathbf{w}) + h(\mathbf{z}))$ for $i = 1, \dots, m$.

We observe that $\mathcal{X}$ and $\mathcal{X}^*$ have the same stabilizer $\mathcal{H}$, and this stabilizer is assumed to be finite.

We obtain the upper bound for $h(\mathbf{x})$ by applying Theorem 2.4 to $\mathcal{X}^*$ and then following the proof of Theorem 2.3, replacing everywhere C_1 by C_2.

Now we estimate $[L(\mathbf{x}) : L]$. To this end, it suffices to estimate the number of distinct points among $\sigma(\mathbf{x})$, $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$.

Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Write again $\mathbf{x} = \mathbf{yz}$ such that $\mathbf{y} \in \overline{\Gamma}$, $h(\mathbf{z}) < \varepsilon(1 + h(\mathbf{y}))$. Put $\mathbf{u}_\sigma := \sigma(\mathbf{x})\mathbf{x}^{-1}$. Since $\sigma(\mathbf{y})\mathbf{y}^{-1}$ is a torsion point, we have

$$h(\mathbf{u}_\sigma) = h(\sigma(\mathbf{x})\mathbf{x}^{-1}) = h(\sigma(\mathbf{z})\mathbf{z}^{-1}) \leq 2h(\mathbf{z}).$$

Completely similarly as (6.6) we obtain

$$h(\mathbf{z}) \leq \varepsilon \cdot \big(C_2 \delta r h_0 + 2C_2 H\big). \tag{9.1}$$

Hence

$$h(\mathbf{u}_\sigma) \leq 2\varepsilon \cdot \big(C_2 \delta r h_0 + 2C_2 H\big) =: \eta.$$

We assume again that f_i is a trinomial for $i = 1, \dots, n$ and a binomial for $i = n+1, \dots, m$. Then (8.1) holds for certain integer vectors $\mathbf{a}_{ij}$ and we obtain

$$\tilde{\alpha}_{i1}\mathbf{u}_\sigma^{\mathbf{a}_{i1}} + \tilde{\alpha}_{i2}\mathbf{u}_\sigma^{\mathbf{a}_{i2}} = 1 \text{ for } i = 1, \dots, n,$$
$$\tilde{\alpha}_{i1}\mathbf{u}_\sigma^{\mathbf{a}_{i1}} = 1 \text{ for } i = n+1, \dots, m.$$

Let $i \in \{1, \dots, n\}$. By our choice of ε in (2.11) we have

$$h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}}) \leq 2\delta\eta = 0.03.$$

Thus by Lemma 4.4 (i) we see that there are at most 2 possibilities for each pair $(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}})$. These facts imply that $\mathbf{u}_\sigma^A = \mathbf{c}_\sigma$ where $\mathbf{c}_\sigma$ runs through a set of cardinality at most 2^m if σ runs through $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Fix $\mathbf{c}_0$ and then σ_0 with $\mathbf{u}_{\sigma_0}^A = \mathbf{c}_0$. Then for every $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$ with $\mathbf{u}_\sigma^A = \mathbf{c}_0$ we have $\left(\frac{\mathbf{u}_\sigma}{\mathbf{u}_{\sigma_0}}\right)^A = \mathbf{1}$. This shows that for fixed $\mathbf{c}_0$ we have at most $t := \#\mathcal{H}$ possibilities for $\mathbf{u}_\sigma$, where $\mathcal{H} := \mathrm{Stab}(\mathcal{X}) = \left\{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N \mid \mathbf{x}^A = \mathbf{1}\right\}$. Hence for $\mathbf{u}_\sigma$ we have altogether at most $2^m t$ possibilities, implying $[L(\mathbf{x}) : L] \leq 2^m t$.

It remains to estimate $t = \#\mathcal{H}$. By assumption, $\mathcal{H}$ is finite hence is zero-dimensional, therefore the matrix A has rank N. Suppose for instance that the first N columns of A form an invertible matrix D. Then $\mathcal{H}$ is contained in $\mathcal{H}' = \{\mathbf{x} \in (\overline{\mathbb{Q}}^*)^N : \mathbf{x}^D = \mathbf{1}\}$. There are matrices $U_1 \in \mathrm{GL}_N(\mathbb{Z})$, $U_2 \in \mathrm{GL}_M(\mathbb{Z})$ such that $U_1 D U_2 = D_0$ is a diagonal matrix with positive integers $d_1, \ldots, d_N$ on the diagonal. Now $\mathbf{x} \mapsto \mathbf{x}^{U_1^{-1}}$ maps $\mathcal{H}'$ isomorphically to the algebraic group given by $x_1^{d_1} = 1, \ldots, x_N^{d_N} = 1$ and the latter clearly has cardinality $d_1 \cdots d_N$.

By an estimate similar to (8.7), using (8.2), we have $d_1 \cdots d_N = |\det D| \leq (2\delta)^N$. Hence $t \leq (2\delta)^N$. This leads to $[L(\mathbf{x}) : L] \leq 2^m t \leq 2^{m+N}\delta^N$.

□

Proof of Theorem 2.5. First suppose that $\mathrm{Stab}(\mathcal{X})$ is finite. Let $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$. We write $\mathbf{x} = \mathbf{y}\mathbf{z}$ with $\mathbf{y} \in \overline{\Gamma}$, $h(\mathbf{z}) < \varepsilon$ and then as usual $\mathbf{y} = \mathbf{v}\mathbf{w}$ with $\mathbf{v} \in \Gamma$ and $\mathbf{w} = \prod_{i=1}^r \mathbf{w}_i^{\gamma_i}$, where $\gamma_i \in \mathbb{Q}$, $|\gamma_i| \leq \frac{1}{2}$. Like in the proof of Theorem 2.6, we define the polynomials $f_i^*(\mathbf{V}) = f_i(\mathbf{w}\mathbf{z} \cdot \mathbf{V})$ $(i = 1, \ldots, m)$ and let $\mathcal{X}^*$ be the variety given by $f_i^* = 0$ $(i = 1, \ldots, m)$. Then again, $\mathbf{v} \in \mathcal{X}^* \cap \Gamma$. Recall that $\deg f_i^* \leq \delta$, and that

$$\max(1, h(f_i^*)) \leq H + \delta h(\mathbf{w}\mathbf{z}) \leq H + \delta(h(\mathbf{w}) + h(\mathbf{z})) \leq H + \delta(\frac{rh_0}{2} + \varepsilon)$$

for $i = 1, \ldots, m$. Now applying part (i) of Theorem 2.4 to $\mathcal{X}^* = (\mathbf{w}\mathbf{z})^{-1}\mathcal{X}$, we obtain

$$h(\mathbf{v}) \leq C_2(H + \delta\frac{rh_0}{2} + \varepsilon)$$

and together with $h(\mathbf{x}) \leq h(\mathbf{v}) + h(\mathbf{w}) + h(\mathbf{z}) \leq h(\mathbf{v}) + \frac{rh_0}{2} + \varepsilon$ this leads to the upper bound for $h(\mathbf{x})$ in (2.10).

As for the estimation of $[L(\mathbf{x}) : L]$, instead of (9.1) we have $h(\mathbf{z}) < \varepsilon$, then our assumption $\varepsilon = \frac{0.03}{4\delta}$ leads to the same conclusion $h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}}) \leq 0.03$ for $i = 1, \dots, n$, and the proof is concluded in the same way as that of Theorem 2.6.

We now assume that $\mathcal{H} := \mathrm{Stab}(\mathcal{X})$ is infinite. We define $\mathbf{z}, \mathbf{v}, \mathbf{w}$ as above and keep the notation from the proof of Theorem 2.4. Thus we obtain, completely similarly as in (8.4),

$$h(\mathbf{v}^{\mathbf{a}_{i,j}}) < C^*(H + \delta\frac{rh_0}{2} + \delta\varepsilon) \text{ for } (i,j) \in I,$$

and together with $h(\mathbf{w}) \leq \frac{rh_0}{2}$, $h(\mathbf{z}) < \varepsilon$, this leads to

$$h(\mathbf{x}^{\mathbf{a}_{i,j}}) < C^*(H + \delta rh_0) \text{ for } (i,j) \in I.$$

Then, similarly as (8.6) we obtain,

$$\mathbf{x}^A = \mathbf{c} \text{ with } h(\mathbf{c}) \leq 2mC^*(H + \delta rh_0). \tag{9.2}$$

Let $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/L)$, and put $\mathbf{u}_\sigma := \sigma(x) \cdot \mathbf{x}^{-1}$, $\mathbf{c}_\sigma := \sigma(\mathbf{c}) \cdot \mathbf{c}^{-1}$. Thus, $\mathbf{u}_\sigma^A = \mathbf{c}_\sigma$. Following the argument in the proof of Theorem 2.6, using our choice $\varepsilon = \frac{0.03}{4\delta}$ for ε, we infer again that $h(\mathbf{u}_\sigma^{\mathbf{a}_{i1}}, \mathbf{u}_\sigma^{\mathbf{a}_{i2}}) \leq 0.03$ for $i = 1, \dots, n$, and subsequently, that $\mathbf{c}_\sigma$ runs through a set of cardinality at most 2^m if σ runs through $\mathrm{Gal}(\overline{\mathbb{Q}}/L)$. This implies that we have at most 2^m possibilities for $\sigma(\mathbf{c})$. Hence

$$[L(\mathbf{c}) : L] \leq 2^m. \tag{9.3}$$

Now from (9.2), (9.3) we infer that for every $\mathbf{x} \in \mathcal{X} \cap \overline{\Gamma}_\varepsilon$ there is $\mathbf{c}$ from a finite, effectively determinable set depending only on Γ and $f_1, \dots, , f_m$, such that $\mathbf{x}^A = \mathbf{c}$. We conclude by applying Proposition 7.4 to each of the equations $\mathbf{x}^A = \mathbf{c}$. $\square$

References

[1] A. Bérczes, J.-H. Evertse and K. Györy, Effective results for linear equations in two unknowns from a multiplicative division group, *Acta Arith.*, to appear.

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

[3] E. Bombieri and U. Zannier, Algebraic points on subvarieties of $\mathbb{G}_m^n$, *Internat. Math. Res. Notices*, **1995:7**, 333-347.

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

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

[6] S. David and P. Philippon, Minorations des hauteurs normalisées des sous-variétés de variétés abeliennes. II, *Comment. Math. Helv.*, **77** (2002), 639–700.

[7] 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.

[8] M. Laurent, Equations diophantiennes exponentielles, *Invent. Math.*, **78** (1984), 299–327.

[9] P. Liardet, Sur une conjecture de Serge Lang, *Astérisque*, **24-25**, Soc. Math. France (1975), 187–210.

[10] C. Pontreau, Petits points d'une surface, *Canad Journ. of Math.*, to appear.

[11] C. Pontreau, A Mordell-Lang plus Bogolomov type result for curves in $\mathbb{G}_{\mathbf{m}}^2$, *Monatshefte für Math.*, to appear.

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

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

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

[15] W. M. Schmidt, Heights of points on subvarieties of $\mathbb{G}_m^n$, in *Number Theory* (Paris, 1993-1994), *London Math. Soc. Lecture Note Ser.*, **235**, Cambridge Univ. Press, 1996, 157–187.

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

[17] S. Zhang, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc.*, **8** (1995), 187–221.

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

C. Pontreau
Laboratoire de mathématiques Nicolas Oresme
CNRS UMR 6139
Université de Caen
14032 Caen cedex, France
E-mail address: pontreau@math.unicaen.fr