BOUNDED HEIGHT IN PENCILS OF SUBGROUPS OF FINITE RANK.

F. AMOROSO, D. MASSER AND U. ZANNIER

Abstract. In a recent paper we proved some new bounded height results for equations involving varying integer exponents. Here we make a start on the problem of generalizing to rational exponents, which corresponds to the step from groups that are finitely generated to groups of finite rank. We discover two unexpected obstacles. The first is that bounded height may genuinely fail in the neighbourhood of certain exponents. The second concerns vanishing subsums, which seem to be much harder to deal with than in classical situations like S-unit equations. But for certain simple and natural equations we are able to clarify the first obstacle and eliminate the second. The proofs are partly based on our earlier work but there are also new considerations about successive minima over function fields.

1. Introduction

In 1997 Beukers [4] proved a new kind of result about the equation

$$t^l + (1-t)^l = 1. \tag{1.1}$$

Namely that if $l \geq 2$ is an integer, then the absolute height (see below) of any solution t_0 to (1.1) is bounded above independently of l. More explicitly, his Lemma 3.5 (p. 100) implies that

$$h(t_0) \leq \log 216 \tag{1.2}$$

for the standard logarithmic Weil height. He was motivated by certain finiteness problems (see below) and irreducibility. Even before that, notably in Silverman's paper [19] from 1983, the importance of Bounded Height for more general problems has been recognized and developed, as in the work of Habegger (see for example [11]); see below for more on this.

In our own paper [2] of 2017 we obtained a broad extension of (1.2) to systems of equations with any number of exponents and any number of terms. For example we proved the existence of similar height bounds when

$$t^l + (1-t)^m = 1 \qquad (l \geq 1, m \geq 1, l+m \geq 3) \tag{1.3}$$

or

$$t^l + (1-t)^l + (1+t)^l = 1 \qquad (l \geq 1) \tag{1.4}$$

(which were made explicit by Denz [9]). The general theme of the present paper is to allow exponents that are not only integral but also rational.

To see why this is natural let us first state the main result of [2].

Let $\mathcal{C}$ be a projective smooth curve defined over $\overline{\mathbb{Q}}$, with function field denoted $\mathbb{F} = \overline{\mathbb{Q}}(\mathcal{C})$ and choose in the usual way a height function h on $\mathcal{C}(\overline{\mathbb{Q}})$.

Date: 31.8.2023.

In [2] we said that a subgroup Γ of $\mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ is constant-free if its image Γ' by any surjective homomorphism $\mathbb{G}_{\mathrm{m}}^r \to \mathbb{G}_{\mathrm{m}}$ satisfies $\Gamma' \cap \overline{\mathbb{Q}}^* = \Gamma'_{\mathrm{tors}}$ (which is the assumption of Theorem [6, Theorem 1']). But we may note that if the homomorphism is not surjective then this trivially holds for Γ' and so the surjectivity here is irrelevant. That implies for example that any homomorphic image of something constant-free is also constant-free.

Theorem 1.1 ([2], Theorem 1.3, p. 2601)**.** *Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ be a finitely generated constant-free subgroup and let V be a subvariety of $\mathbb{G}_{\mathrm{m}}^r$ defined over $\mathbb{F}$. Then the height of the points $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that for some $\boldsymbol{\gamma} \in \Gamma \backslash V$ the specialization $\boldsymbol{\gamma}_P$ is defined and lies in the specialization V_P, is effectively bounded above in terms of $\mathcal{C}, \Gamma$, and V.*

Here it is clear that the excluded points $\boldsymbol{\gamma}$ in V are necessary, for otherwise we could deduce that $\boldsymbol{\gamma}_P$ lies in V_P for all P.

Thus with $\mathcal{C}$ as affine $\mathbb{A}^1$ and $\mathbb{F}$ as $\overline{\mathbb{Q}}(t)$ we may take $r = 2$, Γ generated by $(t, 1-t)$, and V defined by $x_1 + x_2 = 1$ to get (1.1). And Γ generated by $(t, 1), (1, t-1)$ leads to (1.3). Similarly $r = 3$, $(t, 1-t, 1+t)$ and $x_1+x_2+x_3 = 1$ give (1.4).

Our rational exponents will then correspond to the division group Γ^{div} of Γ consisting of all γ for which there exists a positive integer n with $\gamma' = \gamma^n$ in Γ.

Now several classical diophantine results for finitely generated groups naturally extend without significant change to groups of finite rank such as Γ^{div}. This holds for example with the finiteness of solutions to S-unit equations when formulated in terms of groups. However the analogues for abelian varieties instead of $\mathbb{G}_{\mathrm{m}}^r$ do not go so simply; for example just for Γ the result [10] of Faltings implies the Mordell Conjecture, and the extension to Γ^{div} follows from the earlier work [12] of Hindry using non-trivial Kummer Theory.

Here at least we can easily define specializations $\boldsymbol{\gamma}_P$ of $\boldsymbol{\gamma}$ as follows. We choose n as above minimal; and then $\boldsymbol{\gamma}_P$ is any solution of $\boldsymbol{\gamma}_P^n = \boldsymbol{\gamma}'_P$. So there is an indeterminacy involving roots of unity ω with order dividing n.

For example with Γ generated by t in $\overline{\mathbb{Q}}(t)$ the specializations of $t^{1/n}$ are $\omega t_0^{1/n}$ for any fixed choice of $t_0^{1/n}$ and all such ω. But also if ω_1 has order exactly n then $(\omega_1 t)^n$ is in Γ and so the specializations of $\omega_1 t$ are all ωt_0. By making n large we get new results without really leaving Γ, for example height boundedness for $\omega_1 t^l + \omega_2(1-t)^l = 1$ independent of the roots of unity ω_1, ω_2.

But now the literal extension of our Theorem 1.1 to Γ^{div} is false, even in the very special case $\mathcal{C} = \mathbb{A}^1, \mathbb{F} = \overline{\mathbb{Q}}(t), r = 1$ and Γ generated by t. Here the determinations for any $\lambda \in \mathbb{Q}$ of t_0^λ correspond to n as the denominator of λ. Thus with $V = \{2\}$ and the example

$$t_0^\lambda = 2$$

we have

$$h(t_0) = \frac{\log 2}{|\lambda|}$$

which is not bounded above as $\lambda \to 0$.

Actually if λ is small then t^λ is in some geometric sense close to $\eta = 1$, and now

$$\gamma_P \eta_P^{-1} \eta \in V. \tag{1.5}$$

For the V treated in this paper we will see that this is an important obstruction to bounded height.

Similarly for the rational analogue of (1.1) there is an obstruction near $\lambda = 0$, although this is not quite so obvious (see Lemma 6.1). Nevertheless it is relatively easy to extend Beukers's result to rational exponents, where we define $t_0^\lambda, (1-t_0)^\lambda$ as above with $\boldsymbol{\gamma} = (t, 1-t)$. More precisely, if $\lambda = m/n$ with coprime $n \geq 1$ and m, we say that t_0 is a solution of the equation $t^\lambda + (1-t)^\lambda = 1$ if there exist u, v with

$$u^n = t_0^m, \quad v^n = (1-t_0)^m, \quad u + v = 1$$

(compare also (B.2) and the equation immediately preceding it).

Theorem 1.2. *For any $\varepsilon \in (0,1)$ there is an effective C, depending only on ε, such that for any rational λ with $|\lambda| > \varepsilon$ and $\lambda \neq 1$ the solutions $t_0 \in \overline{\mathbb{Q}}$ of $t^\lambda + (1-t)^\lambda = 1$ satisfy $h(t_0) \leq C$.*

More precisely, for any positive rational $\lambda \neq 1$ the solutions of the equation above satisfy

$$h(t_0) \leq 100 \max(1, \lambda^{-1}).$$

One cannot expect bounded height when $\lambda > 0$ is small, as a consequence of a well-known result of Zhang-Zagier (see Lemma 6.1).

In Appendix A we prove a more general result (Theorem A.1) which deals with the equation $\alpha t^\lambda + \beta(1-t)^\lambda = 1$ with α, β fixed non-zero algebraic numbers.

For the rational analogue of (1.3) we have

Theorem 1.3. *For any $\varepsilon \in (0,1)$ there is an effective C, depending only on ε, such that for any rationals λ, μ with $\max\{|\lambda|, |\mu|\} > \varepsilon$ and $(\lambda, \mu) \neq (1,1)$ the solutions $t_0 \in \overline{\mathbb{Q}}$ of $t^\lambda + (1-t)^\mu = 1$ satisfy $h(t_0) \leq C$.*

Again we cannot expect bounded height when $\lambda > 0$ and $\mu > 0$ are small (also see Lemma 6.1).

And for the rational analogue of (1.4) we have

Theorem 1.4. *For any $\varepsilon \in (0,1)$ there is an effective C, depending only on ε, such that for any rational λ with $|\lambda| > \varepsilon$ the solutions $t_0 \in \overline{\mathbb{Q}}$ of $t^\lambda + (1-t)^\lambda + (1+t)^\lambda = 1$ satisfy $h(t_0) \leq C$.*

Once more one cannot expect bounded height when $\lambda > 0$ is small, this time by an explicit lower bound of Warin (see Lemma 6.2).

These results will be deduced in section 6 from our main result, which we now state.

Given $\mathbf{f} = (f_1, \ldots, f_r)$ in $\mathbb{F}^r$, we use the affine geometric height

$$h_{\text{geo}}^{\mathbb{A}}(\mathbf{f}) = -\sum_P \min\{0, \text{ord}_P(f_1), \ldots, \text{ord}_P(f_r)\} \geq 0 \tag{1.6}$$

in $\mathbb{Z}$, and extend in the usual way to $\overline{\mathbb{F}}^r$. It vanishes precisely on $\overline{\mathbb{Q}}^r$. Given $\mathbf{f}, \mathbf{f}_1, \mathbf{f}_2 \in \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{F}})$ we use the obvious notations $\mathbf{f}^{-1}$ and $\mathbf{f}_1\mathbf{f}_2$ for the group operations.

For a single $f \in \mathbb{G}_{\mathrm{m}}(\overline{\mathbb{F}})$ we have $h^{\mathbb{A}}_{\mathrm{geo}}(f^{-1}) = h^{\mathbb{A}}_{\mathrm{geo}}(f)$ by the product formula $\sum_P \mathrm{ord}_P(f) = 0$ and it follows for $\mathbf{g} = (g_1, \ldots, g_r)$ that

$$\mathrm{dist}(\mathbf{f}, \mathbf{g}) = h^{\mathbb{A}}_{\mathrm{geo}}(f_1 g_1^{-1}) + \cdots + h^{\mathbb{A}}_{\mathrm{geo}}(f_r g_r^{-1}) \tag{1.7}$$

defines a distance on $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{F}})/\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$. Note that for $\mathbf{f} \in \mathbb{F}^r$ we have $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{f}) \leq \sum_{i=1}^r h^{\mathbb{A}}_{\mathrm{geo}}(f_i) \leq r h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{f})$. Thus $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{f}\mathbf{g}^{-1}) \leq \mathrm{dist}(\mathbf{f}, \mathbf{g}) \leq r h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{f}\mathbf{g}^{-1})$.

Our main theorem below concerns for technical reasons varieties V defined over $\overline{\mathbb{Q}}$. And for reasons connected with vanishing subsums, which are well-known to cause problems for example in the study of S-unit equations, we shall restrict further to linear hypersurfaces V. For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$ and $\mathbf{x} = (x_1, \ldots, x_r)$ we denote the scalar product $\alpha_1 x_1 + \cdots + \alpha_r x_r$ by $\boldsymbol{\alpha}.\mathbf{x}$, and we shall normalize V so that it is defined by

$$\boldsymbol{\alpha}.\mathbf{x} = 1$$

for $\boldsymbol{\alpha}$ in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$. Thus a vanishing subsum gives rise in particular to a vanishing subsum on the left-hand side.

Now for $\boldsymbol{\gamma}, \boldsymbol{\eta} \in \Gamma^{\mathrm{div}}$ and $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that $\boldsymbol{\gamma}_P, \boldsymbol{\eta}_P$ are defined, we have an analogue

$$\boldsymbol{\gamma}_P \boldsymbol{\eta}_P^{-1} \boldsymbol{\eta} \in V \tag{1.8}$$

of (1.5). Notice that we may specialize (1.8) to get $\boldsymbol{\gamma}_P \in V_P = V$ so that it is contained in the set we are interested in.

Theorem 1.5. *Let $\Gamma \subset \mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ be a constant-free finitely generated subgroup. Let $\boldsymbol{\alpha}$ be in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ and $\varepsilon \in (0,1)$. Then there is an effective $B_\varepsilon = B(\mathcal{C}, \Gamma, \boldsymbol{\alpha}, \varepsilon)$, an effective $C = C(\mathcal{C}, \Gamma)$ and an effective finite subset $H = H(\mathcal{C}, \Gamma)$ of Γ^{div} with the following property. The height of points $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that for some $\boldsymbol{\gamma} \in \Gamma^{\mathrm{div}}$ the value $\boldsymbol{\gamma}_P$ is defined and*

$$\boldsymbol{\alpha}.\boldsymbol{\gamma}_P = 1 \tag{1.9}$$

with non-vanishing subsums, is bounded above by B_ε, except possibly for the pairs $(P, \boldsymbol{\gamma})$ with $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\gamma}) \leq C$ and

$$\boldsymbol{\alpha}.(\boldsymbol{\gamma}_P \boldsymbol{\eta}_P^{-1} \boldsymbol{\eta}) = 1 \tag{1.10}$$

for some $\boldsymbol{\eta}$ in H with $\mathrm{dist}(\boldsymbol{\eta}, \gamma) < \varepsilon$.

Remark 1.6. *The presence of the non-specialized $\boldsymbol{\eta}$ in (1.10) gives a good possibility to go further by "descent". Indeed this is how we shall complete the proof of Theorem 1.4 in Appendix B.*

Note that the assumption on vanishing subsums is necessary, as the following example shows. Consider the equation

$$\frac{1}{2} t^{\lambda+\mu} - t^\lambda + \frac{1}{2} t^\mu = 1 \tag{1.11}$$

with Γ generated by $(t, t, 1), (t, 1, t)$. It has a solution $t_0 = 2^n$ corresponding to $\lambda = n$ and $\mu = 1/n$ so $\boldsymbol{\gamma} = (t^{n+1/n}, t^n, t^{1/n})$ in Γ^{div} and $h(t_0) \to \infty$ as

$n \to \infty$. If Theorem 1.5 were applicable, we would deduce $h^{\mathbb{A}}_{\rm geo}(\boldsymbol{\gamma}) \le C$; but already $h^{\mathbb{A}}_{\rm geo}(\boldsymbol{\gamma}) \ge h^{\mathbb{A}}_{\rm geo}(t^n) = n$. This contradiction comes from the vanishing subsum

$$\frac{1}{2}t_0^{n+1/n} - t_0^n = \frac{1}{2}t_0^n(t_0^{1/n} - 2) = 0.$$

It means that any extension of Theorem 1.5 to general V will not be straightforward.

On the other hand in Γ instead of $\Gamma^{\rm div}$ this kind of obstacle disappears, thanks to Theorem 1.1. For example with the integral analogue

$$\frac{1}{2}t^{l+m} - t^l + \frac{1}{2}t^m = 1$$

of (1.11), then because any solution t_0 satisfies $(t_0^l + 1)(t_0^m - 2) = 0$ we still have vanishing subsums but this does not invalidate bounded height $h(t_0) \le \log 2$. This suggests that the problem of vanishing subsums may be different from that of the S-unit situation, which shows the same phenomenon for $\Gamma^{\rm div}$ as for Γ.

In Theorems 1.2, 1.3, 1.4 (and even in the example (1.5) above) the set H of Theorem 1.5 consists only of the identity in $\mathbf{G}_{\rm m}^r$ (besides what may be called the functional solutions). But this does not always happen. An example is $x + 2y = 1$ with Γ generated by $(t, 1), (1, 1-t)$ as in Theorem 1.3. For $\lambda = 1$ and $\mu \ne 1$ there is a solution

$$(x, y) = (t_0, (1 - t_0)^\mu)$$

with $t_0 = 1 - 2^{-1/(\mu-1)}$. Here $h(t_0) \to \infty$ as $\mu \to 1$. Thus H must contain $(t, 1-t)$, because such (λ, μ) near $(1, 1)$ must be avoided (but not $(1, 1)$ itself).

In [2] we derived Theorem 1.1 from [2, Proposition 6.1, p. 2639] which deals with a variety defined by a single linear equation and with a rank one subgroup. Since every variety is the intersection of hypersurfaces, we first reduced to a single hypersurface. By means of a morphism we have then supposed that this hypersurface was defined by a linear equation, and finally we reduced to rank one subgroups. Here too we could have made a similar deduction, following the first two steps, but at the price of a too technical statement, as the above example might suggest.

When the height of the relevant $\boldsymbol{\gamma}$ is large, the proof of Theorem 1.5 follows closely the reduction to rank one subgroups mentioned before. For the reader's convenience, we shall give all details in section 3. It is here that we need the assumption on non-vanishing subsums.

For $\boldsymbol{\gamma}$ of small height the proof is entirely new. It is postponed to section 5. We approximate $\boldsymbol{\gamma}$ with an element $\mathbf{v} \in \Gamma^{\rm div}$ of bounded "denominator". We then consider the successive minima of the geometric height on the vector space $\mathbf{v}^\perp$ of $\mathbf{g} \in \overline{\mathbb{F}}^r$ such that $\mathbf{g}\mathbf{v} \in V$. If their heights are of the same magnitude, we simply use the so-called height machine (see Lemma 2.1 below). Otherwise, there is a bounded $\boldsymbol{\eta}$ in $\Gamma^{\rm div}$ which is close to $\boldsymbol{\gamma}$ (with respect to the distance (1.7)) and such that $\boldsymbol{\gamma}_P \boldsymbol{\eta}_P^{-1} \boldsymbol{\eta} \in V$ as in (1.8). The relevant tools needed for this construction are developed in section 4, devoted to a close analysis on the behaviour of these successive minima.

In section 6 we deduce from Theorem 1.5 the first assertion of Theorem 1.2 on Beukers's equation with rational exponents, as well as Theorem 1.3. We also deduce Theorem 1.4 on Denz's equation with rational exponents, except for possible pathological solutions near $\lambda = 1$ which cannot be excluded from our main theorem. Appendix A is devoted to prove the more precise statement of Theorem 1.2 and more generally to study the solutions of $\alpha t^\lambda + \beta(1-t)^\lambda = 1$ with α, β fixed non-zero algebraic numbers. Finally in Appendix B we eliminate these pathological solutions of Denz's equation. As in Remark 1.6 this requires a sort of "descent procedure" and then some arguments involving more delicate properties of heights.

In connexion with the broader importance of Bounded Height mentioned earlier, we should point out that Theorem 1.1 for $r = 1$ and $V = \{1\}$ implies a 1999 result [6, Theorem 1'] of Bombieri, Masser and Zannier on a single multiplicative relation. In the same paper they prove some new finiteness results (Theorem 2) when there are two multiplicative relations. Already Theorem 4.1 (p. 101) of Beukers [4] implies something of the same type for (1.1), namely that if t_0 is a solution of

$$t^l + (1-t)^l = 1, \quad t^m + (1-t)^m = 1 \tag{1.12}$$

for different $l \geq 2, m \geq 2$, then t_0 must be $0, 1$ or one of the two primitive sixth roots of unity. It would be interesting to have a finiteness result in the context of Theorem 1.5 or even just for

$$t^\lambda + (1-t)^\lambda = 1, \quad t^\mu + (1-t)^\mu = 1. \tag{1.13}$$

2. Notation, auxiliary results and reduction.

Recall the geometric height $h^{\mathbb{A}}_{\text{geo}}$ from the introduction and the two arithmetic heights h (the height on the curve $\mathcal{C}$ and the standard affine Weil height on $\overline{\mathbb{Q}}$). We need moreover the affine height on $\overline{\mathbb{Q}}^r$, which we denote with $h^{\mathbb{A}}$ to emphasis his affine nature.

We shall use a functorial relation between them, which is an easy consequence of "Weil's Height Machine" (compare for instance [2], Lemma 3.3, p. 2610:

Lemma 2.1. *Let* $\mathbf{f} = (f_1, \ldots, f_r) \in \overline{\mathbb{F}}^r$ *and* $P \in \mathcal{C}(\overline{\mathbb{Q}})$, *not a pole of* $f_1, \ldots, f_r$. *Then*

$$h^{\mathbb{A}}(\mathbf{f}_P) = h^{\mathbb{A}}_{\text{geo}}(\mathbf{f})h(P) + O(1 + h(P)^{1/2})$$

where the implied constant may depend on $\mathbf{f}$ *but not on* P.

The next result is an estimate familiar in problems of multiplicative dependence, but we could not find a precise reference. We noted that $h^{\mathbb{A}}_{\text{geo}}(\mathbf{fg}) \leq h^{\mathbb{A}}_{\text{geo}}(\mathbf{f}) + h^{\mathbb{A}}_{\text{geo}}(\mathbf{g})$ on $\mathbb{G}^r_{\text{m}}(\mathbb{F})$; but we have only $h^{\mathbb{A}}_{\text{geo}}(\mathbf{f}^{-1}) \leq r h^{\mathbb{A}}_{\text{geo}}(\mathbf{f})$. It follows easily that

$$h^{\mathbb{A}}_{\text{geo}}(\mathbf{f}_1^{b_1} \cdots \mathbf{f}_m^{b_m}) \leq r(|b_1| h^{\mathbb{A}}_{\text{geo}}(\mathbf{f}_1) + \cdots + |b_m| h^{\mathbb{A}}_{\text{geo}}(\mathbf{f}_m)) \tag{2.1}$$

for $b_1, \ldots, b_m$ in $\mathbb{Z}$; and it extends at once to $b_1, \ldots, b_m$ in $\mathbb{Q}$ by raising to a suitable power.

Given $\mathbf{f}_1, \ldots, \mathbf{f}_m$ in $\mathbb{G}^r_{\text{m}}(\mathbb{F})$ we define their relation group modulo constants as the set of $(a_1, \ldots, a_m)$ in $\mathbb{Z}^m$ such that $\mathbf{f}_1^{a_1} \cdots \mathbf{f}_m^{a_m}$ is in $\mathbb{G}^r_{\text{m}}(\overline{\mathbb{Q}})$.

From now on we assume that the value groups of the various ord_P in (1.6) are $\mathbb{Z}$.

Lemma 2.2. *Suppose $m \geq 2$ and $\mathbf{f}_1, \ldots, \mathbf{f}_m$ in $\mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ have relation group modulo constants of rank 1. Then there is a generator $(a_1, \ldots, a_m)$ whose non-zero coordinates satisfy*

$$|a_i| \leq (rm)^{m-1} \prod_{j \neq i} h_{\mathrm{geo}}^{\mathbb{A}}(\mathbf{f}_j). \tag{2.2}$$

Proof. We can more or less imitate the arguments of Lemma 7.19 (p. 222) of [20] in the situation when $\mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ is replaced by $\mathbb{G}_{\mathrm{m}}(K)$ for a number field K. But we use also ideas from [7] and the "zero height group" (p. 454).

If some $\mathbf{f}_i$, say $\mathbf{f}_m$, is already in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ then the others cannot be, and then $(0, \ldots, 0, 1)$ is the required generator because $1 \leq h_{\mathrm{geo}}^{\mathbb{A}}(\mathbf{f}_1) \cdots h_{\mathrm{geo}}^{\mathbb{A}}(\mathbf{f}_{m-1})$. So we can assume that none of the $\mathbf{f}_i$ are in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$.

Next by induction on m we can suppose that there is a generator $(a_1, \ldots, a_m)$ with all coordinates non-zero.

Now $\boldsymbol{\theta} = \mathbf{f}_1^{a_1} \cdots \mathbf{f}_m^{a_m}$ is in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$. In particular $a_m \neq 0$ and we can suppose $a_m \geq 1$. We use Minkowski's First Theorem to find $q, p_1, \ldots, p_m$ in $\mathbb{Z}$, not all zero, such that

$$\left| q\frac{a_i}{a_m} - p_i \right| < c_i \quad (i = 1, \ldots, m-1), \qquad |q| \leq c_m$$

provided the positive real numbers $c_1, \ldots, c_{m-1}, c_m$ satisfy

$$c_i \leq 1 \quad (i = 1, \ldots, m-1), \qquad c_1 \cdots c_{m-1} c_m = 1. \tag{2.3}$$

Then $r_i = qa_i - p_i a_m$ has $|r_i| < a_m c_i$ $(i = 1, \ldots, m-1)$. Now

$$\boldsymbol{\theta}^q = \mathbf{f}_1^{a_1 q} \cdots \mathbf{f}_m^{a_m q} = \mathbf{f}_1^{r_1 + p_1 a_m} \cdots \mathbf{f}_{m-1}^{r_{m-1} + p_{m-1} a_m} \mathbf{f}_m^{a_m q}$$

so that $\boldsymbol{\beta} = \mathbf{f}_1^{p_1} \cdots \mathbf{f}_{m-1}^{p_{m-1}} \mathbf{f}_m^q$ satisfies $\boldsymbol{\beta}^{a_m} = \boldsymbol{\theta}^q \mathbf{f}_1^{-r_1} \cdots \mathbf{f}_{m-1}^{-r_{m-1}}$. Thus by (2.1)

$$a_m h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\beta}) \leq r(|r_1| h_1 + \cdots + |r_{m-1}| h_{m-1}) < r a_m (c_1 h_1 + \cdots + c_{m-1} h_{m-1})$$

where $h_i = h_{\mathrm{geo}}^{\mathbb{A}}(\mathbf{f}_i) \geq 1$ by our assumption about proper subsets.

We choose

$$c_i = \frac{1}{r(m-1)h_i} \quad (i = 1, \ldots, m-1)$$

and then $c_m = (r(m-1))^{m-1} h_1 \cdots h_{m-1}$ in accordance with (2.3). We get $h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\beta}) < 1$ so $\boldsymbol{\beta}$ lies in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$.

Because of the generator property, there is an integer $k \neq 0$ such that

$$(p_1, \ldots, p_{m-1}, q) = k(a_1, \ldots, a_{m-1}, a_m).$$

We deduce

$$|a_m| \leq |k a_m| = |q| \leq c_m = (r(m-1))^{m-1} h_1 \cdots h_{m-1},$$

in accordance with (2.2) for $i = m$. The other inequalities follow by symmetry. $\square$

We now prove a Northcott Finiteness Property for $h_{\mathrm{geo}}^{\mathbb{A}}$ on Γ (of course it does not hold on the full $\mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$, due to $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ and zero height). It may be useful in other contexts to use a property weaker than constant-free (which however is not homomorphism stable).

Lemma 2.3. *Suppose* Γ *in* $\mathbb{G}_{\mathrm{m}}^r(\mathbb{F})$ *is finitely generated with*

$$\Gamma \cap \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}}) = \Gamma_{\mathrm{tor}}. \tag{2.4}$$

(a) For any $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_n$ *in* Γ *which are multiplicatively independent there is a constant* $c = c(\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_n) > 0$ *such that*

$$h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\gamma}_1^{a_1} \cdots \boldsymbol{\gamma}_n^{a_n}) \geq c \max\{|a_1|, \ldots, |a_n|\} \tag{2.5}$$

for any $(a_1, \ldots, a_n)$ *in* $\mathbb{Z}^n$.

(b) For any real N *there are at most finitely many* $\boldsymbol{\gamma}$ *in* Γ *with* $h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\gamma}) \leq N$.

Proof. We remark straightaway that because of (2.4) multiplicative dependence in Γ is the same as multiplicative dependence modulo $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$.

For (a) we can assume that $a_1, \ldots, a_n$ are not all zero; and by induction that they are all not zero. We apply Lemma 2.2 to $\boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_n$ with $\boldsymbol{\gamma}_0 = \boldsymbol{\gamma}_1^{a_1} \cdots \boldsymbol{\gamma}_n^{a_n}$. No n of these are dependent; for example if $\boldsymbol{\gamma}_0^{b_0}\boldsymbol{\gamma}_1^{b_1} \cdots \boldsymbol{\gamma}_{n-1}^{b_{n-1}}$ is in $\mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$ then our hypothesis (2.4) would show that it is torsion. Then eliminating $\boldsymbol{\gamma}_0$ would give $a_n b_0 = 0$. Here $b_0 = 0$ is impossible and $a_n = 0$ has been excluded. Clearly $(-1, a_1, \ldots, a_n)$ is a generator in the sense of Lemma 2.2, and we deduce that

$$|a_i| \leq (rn)^{n+1} h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\gamma}_0) \prod_{j \neq i} h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\gamma}_j) \quad (i = 1, \ldots, n).$$

This is a more explicit version of (2.5).

For (b) we simply apply (a) to basis elements of $\Gamma/\Gamma_{\mathrm{tors}}$, so that $\boldsymbol{\gamma} = \boldsymbol{\mu}\boldsymbol{\gamma}_1^{a_1} \cdots \boldsymbol{\gamma}_n^{a_n}$ for torsion $\boldsymbol{\mu}$. We deduce that the exponents are bounded in terms of N (and Γ), and the result follows. □

3. Proof of Theorem 1.5. I. Large $\boldsymbol{\gamma}$.

In this section we prove that there exists $C = C(\mathcal{C}, \Gamma, \boldsymbol{\lambda})$ such that the height of points $P \in \mathcal{C}(\overline{\mathbb{Q}})$ such that, for some $\boldsymbol{\gamma} \in \Gamma^{\mathrm{div}}$ with

$$h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\gamma}) > C \tag{3.1}$$

the value $\boldsymbol{\gamma}_P$ is defined and satisfies (1.9) with non-vanishing subsums is bounded above. Thus there is no need for H and $\boldsymbol{\eta}$ as in (1.10). The proof follows closely the pattern of the deduction of [2, Proposition 6.1, p. 2639] from [2, Theorem 1.5, p. 2603] (see [2, p. 2639–2640]). Note that in the present situation the full property of constant-free is not needed, only (2.4) suffices.

For the reader's convenience, we recall the statement of [2, Theorem 1.5], in a non-homogeneous version which is more convenient for us.

Theorem 3.1. *Let* $r \geq 1$ *and* $f_1, \ldots, f_r \in \mathbb{F}$ *be non-zero rational functions such that* f_i *is non-constant for some* i. *Then there exists a positive real number* C_1 *depending only on* $f_1, \ldots, f_r$, *having the following properties. Let* $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r) \in \mathbb{G}_{\mathrm{m}}^r(\overline{\mathbb{Q}})$. *Consider, for a natural number* n, *a solution* $P \in \mathcal{C}(\overline{\mathbb{Q}})$ *of the equation*

$$\alpha_1 f_1(P)^n + \cdots + \alpha_r f_r(P)^n = 1. \tag{3.2}$$

Then, if $n \geq C_1$ and if there are no vanishing subsums, we have

$$h(P) \leq \frac{r+1}{n} h^{\mathbb{A}}(\alpha_1, \dots, \alpha_r) + C_1. \tag{3.3}$$

To deduce Theorem 3.1 from [2, Theorem 1.5, p. 2603], we replace therein r by $r+1$, Γ by $\Gamma \times \{1\}$ and we take $\alpha_{r+1} = -1$, $f_{r+1} = 1$.

Write $\Gamma = \Gamma_{\text{tors}} \oplus \Gamma_{\text{free}}$ and $\mathbf{1} = (1, \dots, 1)$. Let $\mathbf{f} \in \Gamma_{\text{free}}$ be constant. Then $\mathbf{f} \in \Gamma \cap \mathbb{G}_{\text{m}}^r(\overline{\mathbb{Q}}) = \Gamma_{\text{tors}}$ by (2.4). Since $\Gamma_{\text{free}} \cap \Gamma_{\text{tors}} = \{\mathbf{1}\}$ we get $\mathbf{f} = \mathbf{1}$. Thus

$$\mathbf{f} \in \Gamma_{\text{free}} \backslash \{\mathbf{1}\} \Longrightarrow f_i \notin \overline{\mathbb{Q}} \text{ for some } i. \tag{3.4}$$

Since Γ_{free} is finitely generated and torsion free, it is freely generated by, say, $\mathbf{a}_1, \dots, \mathbf{a}_\kappa$.

Let $\boldsymbol{\gamma} \in \Gamma^{\text{div}}$ of sufficiently large height such that the value $\boldsymbol{\gamma}_P$ is defined and satisfies (1.9). We write

$$\boldsymbol{\gamma} = \boldsymbol{\omega} \mathbf{a}_1^{\lambda_1} \cdots \mathbf{a}_\kappa^{\lambda_\kappa}.$$

for some (uncontrolled) torsion $\boldsymbol{\omega}$ and $\lambda_1, \dots, \lambda_\kappa \in \mathbb{Q}$. Generally let $r_1, \dots, r_\kappa \in \mathbb{Q}$. Since $h^{\mathbb{A}}(\mathbf{x}) \leq \sum h^{\mathbb{A}}(x_i)$ and $h^{\mathbb{A}}(x^\nu) = |\nu| h^{\mathbb{A}}(x)$, applying Lemma 2.1 with $r = 1$ and $f_1 = a_{ji}$ we get

$$h^{\mathbb{A}}((\mathbf{a}_1^{r_1} \cdots \mathbf{a}_\kappa^{r_\kappa})_P) \leq \sum_{i,j} |r_j| h\big((a_{ji})_P\big) \leq c_1 (\max |r_j|) h(P) \tag{3.5}$$

(assuming $h(P) \geq 1$ as we may), where c_1 depends only on $\mathbf{a}_1, \dots, \mathbf{a}_\kappa$. We set

$$Q := [4(r+1)c_1] + 1. \tag{3.6}$$

Let $A := \max_j |\lambda_j|$. By Dirichlet's Theorem on simultaneous approximation (or Minkowski as in the proof of Lemma 2.2), there exists a positive integer $q \leq Q^\kappa$ and integers p_j such that

$$\left| q \frac{\lambda_j}{A} - p_j \right| < \frac{1}{Q}$$

for $j = 1, \dots, \kappa$. Let $n = [A/q] + 1$. We set $r_j = \lambda_j - n p_j$ and

$$\boldsymbol{\rho} = \prod_{j=1}^{\kappa} \mathbf{a}_j^{r_j} \in \Gamma^{\text{div}}, \qquad \mathbf{f} = \prod_{j=1}^{\kappa} \mathbf{a}_j^{p_j} \in \Gamma_{\text{free}}.$$

Let $\boldsymbol{\alpha}' = \boldsymbol{\alpha} \boldsymbol{\omega} \boldsymbol{\rho}_P$. Since $\boldsymbol{\gamma} = \boldsymbol{\omega} \boldsymbol{\rho} \mathbf{f}^n$, equation (1.9) takes the shape (3.2), with α replaced by $\boldsymbol{\alpha}'$ therein.

By definition of A there is an index j_0 such that $\lambda_{j_0} = \pm A$ and thus $p_{j_0} = \pm q \neq 0$. Since $\mathbf{a}_1, \dots, \mathbf{a}_\kappa$ is a basis of Γ_{free} we have $\mathbf{f} \neq \mathbf{1}$. By (3.4), f_i is non-constant for some i. This is one of the conditions we need to apply Theorem 3.1. We still need to show that $n \geq C_1$ where $C_1 = C_1(\mathbf{f})$ is the constant appearing in Theorem 3.1.

We have

$$|p_j| \leq \left| q \frac{\lambda_j}{A} \right| + \frac{1}{Q} \leq q + \frac{1}{Q} \leq 2Q^\kappa. \tag{3.7}$$

Thus, $f_1, \dots, f_r$ belong to a finite set, depending only on Γ. Moreover

$$|r_j| = \left| \frac{A}{q}\left(q\frac{\lambda_j}{A} - p_j\right) - \left(n - \frac{A}{q}\right)p_j \right| \leq \frac{n}{Q} + 2Q^\kappa. \tag{3.8}$$

So by (3.5) we have

(3.9)

$$h^{\mathbb{A}}(\boldsymbol{\alpha}') = h^{\mathbb{A}}(\boldsymbol{\alpha}\boldsymbol{\omega}\boldsymbol{\rho}_P) \leq h^{\mathbb{A}}(\boldsymbol{\alpha}) + h^{\mathbb{A}}(\boldsymbol{\rho}_P) \leq h^{\mathbb{A}}(\boldsymbol{\alpha}) + c_1\left(\frac{n}{Q} + 2Q^\kappa\right)h(P).$$

Let $C_1 = C_1(\mathbf{f})$ be the constant appearing in Theorem 3.1. This constant depends on the rational functions f_i, which by (3.6) and (3.7) belong to a finite set, depending in turn only on Γ. Thus $C_1 \leq c_2$ where now c_2 depends only on Γ. We choose

$$C = C(\mathcal{C}, \Gamma) = r\kappa Q^\kappa \max(c_2, 8(r+1)c_1 Q^\kappa) \sum_j h^{\mathbb{A}}_{\text{geo}}(\mathbf{a}_j).$$

By assumption (3.1) and (2.1)

$$C \leq h^{\mathbb{A}}_{\text{geo}}(\boldsymbol{\gamma}) = h^{\mathbb{A}}_{\text{geo}}(\mathbf{a}_1^{\lambda_1} \cdots \mathbf{a}_\kappa^{\lambda_\kappa}) \leq r\kappa \max_j |\lambda_j| \sum_j h^{\mathbb{A}}_{\text{geo}}(\mathbf{a}_j).$$

Thus $A = \max_j |\lambda_j| \geq Q^\kappa \max(c_2, 8(r+1)c_1 Q^\kappa)$. Since $q \leq Q^\kappa$, we then have

$$n = [A/q] + 1 \geq A/Q^\kappa \geq \max(c_2, 8(r+1)c_1 Q^\kappa). \tag{3.10}$$

In particular, $n \geq c_2 \geq C_1$. Now all the hypotheses of Theorem 3.1 are fulfilled. By (3.3) of this theorem and by the inequalities (3.9), $Q \geq 4(r+1)c_1$ (compare (3.6) above) and $n \geq 8(r+1)c_1 Q^\kappa$ (compare (3.10) above),

$$\begin{aligned} h(P) &\leq \frac{r+1}{n} h^{\mathbb{A}}(\boldsymbol{\alpha}') + C_1 \\ &\leq \frac{r+1}{n} h^{\mathbb{A}}(\boldsymbol{\alpha}) + (r+1)c_1\left(\frac{1}{Q} + \frac{2Q^\kappa}{n}\right)h(P) + c_2 \\ &\leq \frac{r+1}{n} h^{\mathbb{A}}(\boldsymbol{\alpha}) + \frac{h(P)}{2} + c_2. \end{aligned}$$

Thus $h(P) \leq 2(r+1)h^{\mathbb{A}}(\boldsymbol{\alpha}) + 2c_2$. This concludes the proof of Theorem 1.5 when the height of $\boldsymbol{\gamma}$ is large.

4. Functional results.

In order to prove Theorem 1.5 for the elements $\boldsymbol{\gamma} \in \Gamma^{\text{div}}$ of small geometric height, we need first some auxiliary results.

Recall that $\mathbb{F} = \overline{\mathbb{Q}}(\mathcal{C})$ is the function field of a curve $\mathcal{C}$. The following observation is the analogue in function fields of [1, Lemma 3.2, p. 606] and of a result of Rémond (see [15, Lemma 2.1, p. 3]). In it we shall have to leave the safety of $\mathbb{G}_{\text{m}}^r(\mathbb{F})$ for its affine closure $\mathbb{F}^r$ (for us an unfamilar step). Here products (but not inverses) are still well-defined.

Lemma 4.1. *Let Γ be a finitely generated constant-free subgroup of $\mathbb{G}_{\text{m}}^r(\mathbb{F})$. Then there are effective $C_2 = C_2(\mathcal{C}, \Gamma) > 0$ and an effective integer $n =$*

$n(\mathcal{C},\Gamma) \geq 1$ *with the following property. Suppose* $\mathbf{w} \in \mathbb{F}^r$ *and* $\mathbf{v} \in \Gamma^{\mathrm{div}}$ *are such that* $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{w}\mathbf{v}^{-1}) < C_2^{-1}$. *Then*

$$\mathbf{w}^n \in \overline{\mathbb{Q}}^r\Gamma. \tag{4.1}$$

Proof. Suppose first that $\mathbf{w}$ is in $\mathbb{G}^r_{\mathrm{m}}(\mathbb{F})$. With $\mathbf{a}_1,\ldots,\mathbf{a}_\kappa$ as basis elements for the free part of Γ, there are integers $q \geq 1, b_1,\ldots b_\kappa$ with $\mathbf{v}^q = \mathbf{a}_1^{b_1}\cdots\mathbf{a}_\kappa^{b_\kappa}$. Write $\mathbf{u} = \mathbf{w}\mathbf{v}^{-1}$ and $\mathbf{t} = \mathbf{u}^q \in \mathbb{G}^r_{\mathrm{m}}(\mathbb{F})$, so that

$$\mathbf{t}\mathbf{w}^{-q}\mathbf{a}_1^{b_1}\cdots\mathbf{a}_\kappa^{b_\kappa} = 1. \tag{4.2}$$

Thus the relation group modulo constants of $\mathbf{t},\mathbf{w},\mathbf{a}_1,\ldots,\mathbf{a}_\kappa$ has rank $\rho \geq 1$.

If $\rho = 1$ then (4.2) must correspond to a generator because there is an exponent 1. As $q \geq 1$ we get from Lemma 2.2

$$q \leq (r(\kappa+2))^{\kappa+1}h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{t})h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_1)\cdots h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_\kappa). \tag{4.3}$$

As $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{t}) = qh^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{u})$ we deduce $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{u}) \geq C_2^{-1}$ for

$$C_2 = (r(\kappa+2))^{\kappa+1}h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_1)\cdots h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_\kappa).$$

This contradicts a hypothesis; and it follows that $\rho \geq 2$.

So we can eliminate $\mathbf{t}$ from the relations to see that $\mathbf{w},\mathbf{a}_1,\ldots,\mathbf{a}_\kappa$ are multiplicatively dependent modulo constants. So the rank of this new relation group modulo constants is at least 1. But it cannot be bigger, otherwise the further elimination of $\mathbf{w}$ would show that $\mathbf{a}_1,\ldots,\mathbf{a}_\kappa$ are multiplicatively dependent modulo constants. However we have already noted that this is equivalent to plain multiplicative dependence in the constant-free Γ.

So the new rank is 1. A generator $(n,c_1,\ldots,c_\kappa)$ must have $n \neq 0$, and Lemma 2.2 shows that

$$|n| \leq (r(\kappa+1))^{\kappa}h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_1)\cdots h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_\kappa).$$

We can suppose $n \geq 1$, and this gives $\mathbf{w}^n \in \mathbb{G}^r_{\mathrm{m}}(\overline{\mathbb{Q}})\Gamma$ slightly stronger than required.

We now proceed by induction on r. If $\mathbf{w} = (w_1,\ldots,w_r)$ is not in $\mathbb{G}^r_{\mathrm{m}}(\mathbb{F})$ then some coordinates are zero. Say for simplicity that $w_r = 0$ but no others. We project to $\mathbb{G}^{r-1}_{\mathrm{m}}$ to get $\mathbf{w}' = (w_1,\ldots,w_{r-1})$ and also Γ' in $\mathbb{G}^{r-1}_{\mathrm{m}}(\mathbb{F})$ and $\mathbf{v}'$ in Γ'^{div}. Note that Γ' is also constant-free. Then $h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{w}'\mathbf{v}'^{-1}) = h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{w}\mathbf{v}^{-1})$ and so if this is small as above, then $\mathbf{w}'^n \in \mathbb{G}^{r-1}_{\mathrm{m}}(\overline{\mathbb{Q}})\Gamma'$ for some $n \geq 1$ bounded as above. Thus $\mathbf{w}'^n = (\theta_1,\ldots,\theta_{r-1})\boldsymbol{\gamma}'$ for some $\theta_1,\ldots,\theta_{r-1}$ in $\overline{\mathbb{Q}}^*$ and $\gamma' = (\gamma_1,\ldots,\gamma_{r-1})$ in Γ'. There is γ_r in $\mathbb{F}$ with $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_{r-1},\gamma_r)$ in Γ, and now finally $\mathbf{w}^n = (\theta_1,\ldots,\theta_{r-1},0)\boldsymbol{\gamma}$ is in $\overline{\mathbb{Q}}^r\Gamma$ as required. □

We define a projective geometric height h_{geo} on $\mathbb{F}^{r+1}\setminus\{0\}$ by

$$h_{\mathrm{geo}}(\mathbf{f}) = -\sum_P \min\{\mathrm{ord}_P(f_1),\ldots,\mathrm{ord}_P(f_r),\mathrm{ord}_P(f_{r+1})\} \geq 0 \tag{4.4}$$

for $\mathbf{f} = (f_1,\ldots,f_r,f_{r+1})$; and then extend to $\overline{\mathbb{F}}^{r+1}\setminus\{0\}$. Note that for $\mathbf{f} \in \overline{\mathbb{F}}^{r+1}$ with $f_{r+1} = 1$ we have

$$h_{\mathrm{geo}}(\mathbf{f}) = h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{f}) = h^{\mathbb{A}}_{\mathrm{geo}}(f_1,\ldots,f_r).$$

Thus for $\mathbf{x}$, $\mathbf{y} \in \mathbb{G}_{\mathrm{m}}^{r+1} \cap \{x_{r+1} = 1\}$ we also have, by definition (1.7) and by the remark that follows it,

$$\mathrm{dist}(\mathbf{x}, \mathbf{y}) \le r h_{\mathrm{geo}}^{\mathbb{A}}(\mathbf{x}\mathbf{y}^{-1}) = r h_{\mathrm{geo}}(\mathbf{x}\mathbf{y}^{-1}). \tag{4.5}$$

Given a non-zero vector $\mathbf{x} \in \overline{\mathbb{F}}^{r+1}$ we consider the r-dimensional $\overline{\mathbb{F}}$-vector space $\mathbf{x}^{\perp}$ in $\overline{\mathbb{F}}^{r+1}$ of vectors $\mathbf{g}$ orthogonal to $\mathbf{x}$, that is, $\mathbf{g}.\mathbf{x} = 0$.

Lemma 4.2. *Let $\Gamma_1 \subset \{x_{r+1} = 1\}$ be a finitely generated constant-free subgroup of $\mathbb{G}_{\mathrm{m}}^{r+1}(\mathbb{F})$. Let $\mathbf{v} \in (\overline{\mathbb{Q}}^*)^{r+1}\Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}$, $\mathbf{g} \in \mathbf{v}^{\perp}$ with $\mathbf{g} \in \overline{\mathbb{F}}^{r}$ non-zero, and consider the vector space $U \subseteq \overline{\mathbb{F}}^{r+1}$ spanned over $\overline{\mathbb{F}}$ by the conjugates $(\mathbf{g}\mathbf{v})^{\sigma}$ for $\sigma \in \mathrm{Gal}(\overline{\mathbb{F}}/\mathbb{F})$. Then $U\mathbf{v}^{-1}$ is a subspace of $\mathbf{v}^{\perp}$ with a basis of vectors of height $h_{\mathrm{geo}}(\mathbf{g})$.*

Moreover, if $h_{\mathrm{geo}}(\mathbf{g}) < C_3^{-1}$ for some $C_3 = C_3(\mathcal{C}, \Gamma_1) \ge 1$, there exist $n = n(\mathcal{C}, \Gamma_1) \in \mathbb{N}$ and $\boldsymbol{\eta} \in \Gamma_1^{1/n} \cap \{x_{r+1} = 1\}$ with

$$\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \le C_3 h_{\mathrm{geo}}(\mathbf{g})$$

such that $U\boldsymbol{\eta}^{-1}$ is defined[1] *over $\overline{\mathbb{Q}}$.*

Proof. Write for simplicity $\mathbf{u} = \mathbf{g}\mathbf{v}$. We prove the first assertion. We have

$$\mathbf{u}.\mathbf{1} = \mathbf{u}.(\mathbf{v}^{-1}\mathbf{v}) = (\mathbf{u}\mathbf{v}^{-1}).\mathbf{v} = \mathbf{g}.\mathbf{v} = 0$$

with $\mathbf{u} = (1, \ldots, 1)$. Thus also

$$0 = (\mathbf{u}.\mathbf{1})^{\sigma} = \mathbf{u}^{\sigma}.\mathbf{1} = \mathbf{u}^{\sigma}.(\mathbf{v}^{-1}\mathbf{v}) = (\mathbf{u}^{\sigma}\mathbf{v}^{-1}).\mathbf{v}$$

so $\mathbf{u}^{\sigma}\mathbf{v}^{-1} \in \mathbf{v}^{\perp}$.

Also there is a positive integer q with $\mathbf{v}^q \in (\overline{\mathbb{Q}}^*)^{r+1}\Gamma$. Thus $\mathbf{v}^{\sigma} = \boldsymbol{\zeta}_{\sigma}\mathbf{v}$ for torsion $\boldsymbol{\zeta}_{\sigma}$. Therefore

$$\mathbf{u}^{\sigma}\mathbf{v}^{-1} = (\mathbf{g}\mathbf{v})^{\sigma}\mathbf{v}^{-1} = \mathbf{g}^{\sigma}(\mathbf{v}^{\sigma}\mathbf{v}^{-1}) = \mathbf{g}^{\sigma}\boldsymbol{\zeta}_{\sigma}$$

has height $h_{\mathrm{geo}}(\mathbf{g})$; and a basis of $U\mathbf{v}^{-1}$ can be selected from these as asserted.

We now prove the second assertion. We assume $h_{\mathrm{geo}}(\mathbf{g}) < C_3^{-1}$ for C_3 to be determined shortly. Let d be the $\overline{\mathbb{F}}$-dimension of U. Note that $d \ge 1$ since $\mathbf{g}$ is non-zero and $d \le r$ since $U\mathbf{v}^{-1}$ is a subspace of $\mathbf{v}^{\perp}$. It is easy to see that this is the $\mathbb{F}$-dimension of the space spanned over $\mathbb{F}$ by the coordinates of $\mathbf{u}$ and that U has a $\overline{\mathbb{F}}$-basis in $\mathbb{F}^{r+1}$, consisting of say $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(d)}$.

Given a multi-index $\mathbf{i} = (i_1, \ldots, i_d)$ with $1 \le i_1 < \cdots < i_d \le r+1$, let

$$w_{\mathbf{i}} = \begin{vmatrix} u_{i_1}^{(1)} & \cdots & u_{i_d}^{(1)} \\ \vdots & \vdots & \vdots \\ u_{i_1}^{(d)} & \cdots & u_{i_d}^{(d)} \end{vmatrix} \in \mathbb{F}$$

formed with the coordinates of this basis (of course we are speaking of grassmannians and wedge products). Let $\mathbf{w} = (w_{\mathbf{i}})_{\mathbf{i}} \in \mathbb{F}^R$ with $R = \binom{r+1}{d}$. Note that $\mathbf{w} \neq 0$, since $d = \dim U$. Without loss of generality we can assume $\mathbf{w}_R \neq 0$ (in a suitable ordering). We consider the height h_{geo} on a subspace

[1] That is, it has a basis of vectors in $\overline{\mathbb{Q}}^{r}$, or equivalently a set of defining linear equations over $\overline{\mathbb{Q}}$.

of $\overline{\mathbb{F}}^{r+1}$, namely the projective, geometric height of the exterior product of any basis of the subspace, see definition in [17, Chapter 1, §8, p. 28] for more details. Then for every $\mathbf{x} \in \mathbb{G}_{\mathrm{m}}^{r+1}(\overline{\mathbb{F}})$

$$h_{\mathrm{geo}}(U\mathbf{x}) = h_{\mathrm{geo}}(\mathbf{w}\Phi(\mathbf{x})), \tag{4.6}$$

where $\Phi = \Phi_d$ is the homomorphism from $\mathbb{G}_{\mathrm{m}}^{r+1}$ to $\mathbb{G}_{\mathrm{m}}^{R}$ defined by

$$\Phi(x_1, \ldots, x_{r+1}) = (x_{i_1} \cdots x_{i_d})_{\mathbf{i}}.$$

In particular, $h_{\mathrm{geo}}(U) = h_{\mathrm{geo}}(\mathbf{w})$ and $h_{\mathrm{geo}}(U\mathbf{v}^{-1}) = h_{\mathrm{geo}}(\mathbf{w}\mathbf{v}'^{-1})$ with $\mathbf{v}' = \Phi(\mathbf{v})$. Moreover, since $U\mathbf{v}^{-1}$ has a basis of vectors of height $h_{\mathrm{geo}}(\mathbf{g})$ by the first assertion of this lemma,

$$h_{\mathrm{geo}}(\mathbf{w}\mathbf{v}'^{-1}) = h_{\mathrm{geo}}(U\mathbf{v}^{-1}) \leq d h_{\mathrm{geo}}(\mathbf{g}) \leq r h_{\mathrm{geo}}(\mathbf{g}). \tag{4.7}$$

We pause to note that

$$h_{\mathrm{geo}}(\mathbf{x}) \leq r h_{\mathrm{geo}}(\Phi(\mathbf{x})) \tag{4.8}$$

for $\mathbf{x} = (x_1, \ldots, x_{r+1})$ in $\mathbb{G}_{\mathrm{m}}^{r+1}(\overline{\mathbb{F}})$; for example

$$\begin{aligned} h_{\mathrm{geo}}^{\mathbb{A}}(x_1/x_2) &= h_{\mathrm{geo}}^{\mathbb{A}}((x_1x_3\cdots x_{d+1})/(x_2x_3\cdots x_{d+1})) \\ &= h_{\mathrm{geo}}(x_1x_3\cdots x_{d+1}, x_2x_3\cdots x_{d+1}) \leq H \end{aligned}$$

for $H = h_{\mathrm{geo}}(\Phi(\mathbf{x}))$. Similarly for any x_i/x_j. Thus

$$\begin{aligned} h_{\mathrm{geo}}(\mathbf{x}) &= h_{\mathrm{geo}}^{\mathbb{A}}(x_1/x_{r+1}, \ldots, x_r/x_{r+1}) \\ &\leq h_{\mathrm{geo}}^{\mathbb{A}}(x_1/x_{r+1}) + \cdots + h_{\mathrm{geo}}^{\mathbb{A}}(x_r/x_{r+1}) \leq rH, \end{aligned}$$

giving (4.8).

We want to apply Lemma 4.1 to a suitable subgroup of $\mathbb{G}_{\mathrm{m}}^{R}$. By assumption, there exists $\boldsymbol{\alpha} \in (\overline{\mathbb{Q}}^*)^{r+1}$ such that $\boldsymbol{\alpha}^{-1}\mathbf{v} \in \Gamma^{\mathrm{div}}$. Thus $\Phi(\boldsymbol{\alpha}^{-1})\mathbf{v}' = \Phi(\boldsymbol{\alpha}^{-1}\mathbf{v}) \in \Phi(\Gamma_1^{\mathrm{div}}) = \Phi(\Gamma_1)^{\mathrm{div}}$. We de-homogenize taking the map ψ from $\mathbb{F}^R\backslash\{x_R \neq 0\}$ to $\mathbb{F}^R$ defined by $\psi(y_1, \ldots, y_R) = (y_1/y_R, \ldots, y_{R-1}/y_R, 1)$. We let $\tilde{\Gamma} = \tilde{\Gamma}_d = (\psi \circ \Phi)(\Gamma_1)$, also finitely generated and constant-free. We also define

$$\tilde{\mathbf{w}} = \psi(\Phi(\boldsymbol{\alpha})\mathbf{w}) \in \mathbb{F}^R, \qquad \text{and} \quad \tilde{\mathbf{v}} = \psi(\Phi(\boldsymbol{\alpha}^{-1})\mathbf{v}') \in \tilde{\Gamma}^{\mathrm{div}}.$$

(remember that $\mathbf{w}_R \neq 0$). By assumption $h_{\mathrm{geo}}(\mathbf{g}) < C_3^{-1}$. If $C_3 \geq rC_2(\mathcal{C}, \tilde{\Gamma})$, then by (4.7),

$$h_{\mathrm{geo}}^{\mathbb{A}}(\tilde{\mathbf{w}}\tilde{\mathbf{v}}^{-1}) = h_{\mathrm{geo}}(\Phi(\boldsymbol{\alpha})\mathbf{w} \cdot \Phi(\boldsymbol{\alpha}^{-1})\mathbf{v}') = h_{\mathrm{geo}}(\mathbf{w}\mathbf{v}'^{-1}) < C_2(\mathcal{C}, \tilde{\Gamma})^{-1}.$$

By Lemma 4.1, $\tilde{\mathbf{w}}^n \in \overline{\mathbb{Q}}^R\tilde{\Gamma}$ for some integer $n = n(\mathcal{C}, \tilde{\Gamma}) \geq 1$. Since $\Phi(\boldsymbol{\alpha}) \in (\overline{\mathbb{Q}}^*)^R$, this gives, after (slowly) re-homogenizing,

$$\mathbf{w}^n \in \mathbb{F}^*\overline{\mathbb{Q}}^R\Phi(\Gamma_1).$$

Since $\Gamma_1 \subset \{x_{r+1} = 1\}$ there is $\boldsymbol{\eta} \in \Gamma_1^{1/n} \cap \{x_{r+1} = 1\}$ such that

$$(\mathbf{w}\Phi(\boldsymbol{\eta}^{-1}))^n \in \mathbb{F}^*\overline{\mathbb{Q}}^R. \tag{4.9}$$

Taking powers, we can replace n by $n(\mathcal{C}, \Gamma_1) = \mathrm{lcm}_{1\leq d\leq r}\, n(\mathcal{C}, \tilde{\Gamma}_d)$; and similarly we can assume $C_3 \geq r \max_{1\leq d\leq r} C_2(\mathcal{C}, \tilde{\Gamma}_d)$. By (4.6) and (4.9),

$$h_{\mathrm{geo}}(U\boldsymbol{\eta}^{-1}) = h_{\mathrm{geo}}(\mathbf{w}\Phi(\boldsymbol{\eta}^{-1})) = 0$$

which shows that $U\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$. Moreover, again by (4.9) and by (4.7),

$$h_{\mathrm{geo}}(\Phi(\boldsymbol{\eta})\mathbf{v}'^{-1}) = h_{\mathrm{geo}}(\mathbf{w}\mathbf{v}'^{-1}) \leq r h_{\mathrm{geo}}(\mathbf{g}).$$

Hence, by (4.5) and (4.8) we get

$$\begin{aligned}\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) &\leq r h_{\mathrm{geo}}(\boldsymbol{\eta}\mathbf{v}^{-1}) \leq r^2 h_{\mathrm{geo}}(\Phi(\boldsymbol{\eta}\mathbf{v}^{-1})) = r^2 h_{\mathrm{geo}}(\Phi(\boldsymbol{\eta})\mathbf{v}'^{-1}) \\ &\leq r^3 h_{\mathrm{geo}}(\mathbf{g}).\end{aligned}$$

□

Let $\mathbf{v} \in \overline{\mathbb{F}}^{r+1} \cap \{x_{r+1} = 1\}$. We consider the successive infima

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$$

of h_{geo} on $\mathbf{v}^{\perp}$. The next proposition contains some useful information on these infima, when $\mathbf{v} \in (\overline{\mathbb{Q}}^*)^{r+1}\Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}$ for some finitely generated constant-free subgroup Γ_1 of $\mathbb{G}_{\mathrm{m}}^{r+1}(\mathbb{F}) \cap \{x_{r+1} = 1\}$. To simplify the statement and the proof, we assume that the μ_i are attained. Given a k-vector space V and $v_1, \ldots, v_l \in V$ we define as usual $\mathrm{Span}_k(v_1, \ldots, v_l)$ as the k-vector subspace spanned by $v_1, \ldots, v_l$.

Lemma 4.3. *Let $\Gamma_1 \subset \{x_{r+1} = 1\}$ be a finitely generated constant-free subgroup of $\mathbb{G}_{\mathrm{m}}^{r+1}(\mathbb{F})$. Let*

$$\mathbf{v} \in (\overline{\mathbb{Q}}^*)^{r+1}\Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}.$$

Denote by $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ the successive infima of h_{geo} on $\mathbf{v}^{\perp}$. Then there exists a basis $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(r)}$ of $\mathbf{v}^{\perp}$ with $h_{\mathrm{geo}}(\mathbf{g}^{(i)}) = \mu_i$ and integers $0 = d_0 < d_1 < \cdots < d_k = r$ such that

$$\begin{aligned}(4.10)\quad \mu_1 = \cdots = \mu_{d_1} \leq \mu_{d_1+1} = \cdots = \mu_{d_2} \leq \cdots \\ \leq \mu_{d_{k-1}+1} = \cdots = \mu_{d_k} = \mu_r\end{aligned}$$

with the following properties. Let $C_3 = C_3(\mathcal{C}, \Gamma_1)$ and $n = n(\mathcal{C}, \Gamma_1) \in \mathbb{N}$ as in Lemma 4.2. If for some j with $1 \leq j \leq k$ we have

$$\mu_{d_{j-1}+1} < (2nC_3)^{-1},$$

then there exists $\boldsymbol{\eta} \in \Gamma_1^{1/n}$ with $\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \leq C_3\mu_{d_{j-1}+1}$ such that

$$\mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d_j)})\mathbf{v}\boldsymbol{\eta}^{-1}$$

is defined over $\overline{\mathbb{Q}}$.

Proof. We first construct a basis $\mathbf{g}^{(i)}$ of $\mathbf{v}^{\perp}$ and integers d_j satisfying (4.10), such that the following additional assertion is satisfied.

(1) For $j = 1, \ldots, k$, let $U^{(j)}$ be the vector space spanned by $(\mathbf{g}^{(d_l)}\mathbf{v})^{\sigma}$ for $l = 1, \ldots, j$ and $\sigma \in \mathrm{Gal}(\overline{\mathbb{F}}/\mathbb{F})$. Then $\mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d_j)}) = U^{(j)}\mathbf{v}^{-1}$.

Let $\mathbf{g}^{(1)} \in \mathbf{v}^{\perp}$ realize the first minimum and let $U^{(1)}$ be the vector space spanned by $(\mathbf{g}^{(1)}\mathbf{v})^{\sigma}$. We denote by d_1 its dimension. By the first assertion of Lemma 4.2, $U^{(1)}\mathbf{v}^{-1}$ is a subspace of $\mathbf{v}^{\perp}$ and there exists a $\overline{\mathbb{F}}$-basis $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d_1)}$ of $U^{(1)}\mathbf{v}^{-1}$ consisting of vectors of height μ_1. Thus $\mu_1 = \cdots = \mu_{d_1}$.

If $d_1 = r$ we have finished with (1). Otherwise, let $\mathbf{g}^{(d_1+1)}$ realize μ_{d_1+1} and let U be the vector space spanned by the $(\mathbf{g}^{(d_1+1)}\mathbf{v})^{\sigma}$ $(\sigma \in \mathrm{Gal}(\overline{\mathbb{F}}/\mathbb{F}))$.

By the first assertion of Lemma 4.2, $U\mathbf{v}^{-1}$ is a subspace of $\mathbf{v}^\perp$ and there exists a $\overline{\mathbb{F}}$-basis of $U\mathbf{v}^{-1}$ consisting of vectors of height μ_{d_1+1}. Let $U^{(2)} = U^{(1)} + U$ and $d_2 = \dim(U^{(2)})$. Thus we can complete $\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(d_1)}$ to a basis of $U^{(2)}\mathbf{v}^{-1}$ with vectors $\mathbf{g}^{(d_1+1)}, \dots, \mathbf{g}^{(d_2)}$ of height μ_{d_1+1}. Hence $\mu_{d_1+1} = \dots = \mu_{d_2}$.

Continuing in this way we define integers $0 = d_0 < d_1 < \dots < d_k = r$ and a basis $\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(r)}$ satisfying (1).

We now prove

(2) If $\mu_1 < C_3^{-1}$, there exists $\boldsymbol{\eta} \in \Gamma_1^{1/n}$ with $\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \le C_3\mu_1$ such that $\mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(d_1)})\mathbf{v}\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$.

(3) Let us assume $\mu_{d_{j-1}+1} < (2nC_3)^{-1}$ for some j with $2 \le j \le k$. Let $\boldsymbol{\eta}$ as in (2) [2]. Then $\mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(d_j)})\mathbf{v}\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$.

which clearly imply our claim.

To prove (2) we simply use (1) with $j = 1$ and the second part of Lemma 4.2. By this lemma there exists $\boldsymbol{\eta} \in \Gamma_1^{1/n}$ with

$$\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \le C_3 h_{\mathrm{geo}}(\mathbf{g}^{(1)}) = C_3\mu_1$$

such that $U^{(1)}\boldsymbol{\eta}^{-1} = \mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(d_1)})\mathbf{v}\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$.

Let us now prove (3) when $j = 2$. Let, as in proof of (1), U be the vector space spanned by $(\mathbf{g}^{(d_1+1)}\mathbf{v})^\sigma$ so that $U^{(2)} = U^{(1)} + U$ by (1) with $i = 1$. Let us assume $\mu_{d_1+1} < (2nC_3)^{-1}$. Thus $\mu_1 \le \mu_{d_1+1} < C_3^{-1}$ and, by (2) and (1) with $i = 1$, there exists $\boldsymbol{\eta} \in \Gamma_1^{1/n}$ with $\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \le C_3\mu_1$ such that $U^{(1)}\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$. By Lemma 4.2 there exists $\boldsymbol{\eta}' \in \Gamma_1^{1/n}$ with $\mathrm{dist}(\boldsymbol{\eta}', \mathbf{v}) \le C_3 h_{\mathrm{geo}}(\mathbf{g}^{(d_1+1)}) = C_3\mu_{d_1+1}$ such that $U\boldsymbol{\eta}'^{-1}$ is defined over $\overline{\mathbb{Q}}$. The triangular inequality gives

$$\mathrm{dist}(\boldsymbol{\eta}', \boldsymbol{\eta}) \le \mathrm{dist}(\boldsymbol{\eta}', \mathbf{v}) + \mathrm{dist}(\mathbf{v}, \boldsymbol{\eta}) \le C_3\mu_{d_1+1} + C_3\mu_1 \le 2C_3\mu_{d_1+1} < \frac{1}{n}.$$

On the other hand, by the remark which follows (1.7),

$$\mathrm{dist}(\boldsymbol{\eta}', \boldsymbol{\eta}) \ge h_{\mathrm{geo}}^{\mathbb{A}}(\boldsymbol{\eta}'\boldsymbol{\eta}^{-1}) \in \frac{1}{n}\mathbb{Z},$$

since $(\boldsymbol{\eta}')^n$, $\boldsymbol{\eta}^n \in \Gamma_1 \subset \mathbb{G}_{\mathrm{m}}^{r+1}(\mathbb{F})$. Thus $\mathrm{dist}(\boldsymbol{\eta}', \boldsymbol{\eta}) = 0$ and $\boldsymbol{\eta}'\boldsymbol{\eta}^{-1} \in \overline{\mathbb{Q}}^{r+1}$. Therefore also $U\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$. Since $U^{(1)}\boldsymbol{\eta}^{-1}$ is again defined over $\overline{\mathbb{Q}}$ and $U^{(2)} = U^{(1)} + U$, we deduce that $U^{(2)}\boldsymbol{\eta}^{-1} = \mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(d_2)})\mathbf{v}\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$. Thus does $j = 2$. The proof of (3) when $j > 2$ follows the same lines. □

We also need the following lemma.

Lemma 4.4. *Let U be a $\overline{\mathbb{F}}$-vector subspace of $\overline{\mathbb{F}}^{r+1}$ with basis $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(d)}$ on which $x_1 + \dots + x_{r+1} = 0$. Let $\Gamma_1 \subset \{x_{r+1} = 1\}$ be a finitely generated subgroup of $\mathbb{G}_{\mathrm{m}}^{r+1}(\mathbb{F})$. We assume $U\boldsymbol{\psi}^{-1}$ defined over $\overline{\mathbb{Q}}$ for some $\boldsymbol{\psi} \in \Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}$ so that the specialization U_P is defined as $U\boldsymbol{\psi}^{-1}\boldsymbol{\psi}_P$.*

[2] By assumption $\mu_1 \le \mu_{d_{j-1}+1} < C_3^{-1}$, which is the condition we need to apply (2).

Then for all $P \in \mathcal{C}$ with at most finitely many exceptions, $U_P(\overline{\mathbb{Q}})$ has a $\overline{\mathbb{Q}}$-basis $\mathbf{u}_P^{(1)}, \ldots, \mathbf{u}_P^{(d)}$.

Further, if $\boldsymbol{\phi} \in \overline{\mathbb{Q}}^{r+1}\Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}$ is such that $\boldsymbol{\phi}_P \in U_P$, then $(\boldsymbol{\phi}_P \boldsymbol{\psi}_P^{-1}).\boldsymbol{\psi} = 0$.

Proof. The first part is standard reduction theory. As for the second part, the space $U\boldsymbol{\psi}^{-1}$ is defined by equations of the form $\mathbf{a}.\mathbf{x} = 0$ for $\mathbf{a} \in \overline{\mathbb{Q}}^{r+1}$. So U is defined by the $(\mathbf{a}\boldsymbol{\psi}^{-1}).\mathbf{x} = 0$. Thus U_P is defined by the $(\mathbf{a}\boldsymbol{\psi}_P^{-1}).\mathbf{x} = 0$. Therefore

$$0 = (\mathbf{a}\boldsymbol{\psi}_P^{-1}).\boldsymbol{\phi}_P = \mathbf{a}.(\boldsymbol{\psi}_P^{-1}\boldsymbol{\phi}_P)$$

showing that $\boldsymbol{\psi}_P^{-1}\boldsymbol{\phi}_P \in U\boldsymbol{\psi}^{-1}$ so $\mathbf{u} = \boldsymbol{\psi}_P^{-1}\boldsymbol{\phi}_P\boldsymbol{\psi} \in U$. Now

$$0 = \mathbf{u}.\mathbf{1} = (\boldsymbol{\psi}_P^{-1}\boldsymbol{\phi}_P\boldsymbol{\psi}).\mathbf{1} = (\boldsymbol{\phi}_P\boldsymbol{\psi}_P^{-1}).\boldsymbol{\psi}$$

as required. □

5. Proof of Theorem 1.5. II. Small $\boldsymbol{\gamma}$.

In this section we prove Theorem 1.5, assuming further that $\boldsymbol{\gamma} \in \Gamma^{\mathrm{div}}$ satisfies $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\gamma}) \leq C$ for some $C = C(\mathcal{C}, \Gamma)$. We shall introduce along the proof constants $c_1, \ldots, c_5 \geq 1$ depending only on $\mathcal{C}$, Γ, $\boldsymbol{\alpha}$ and ε (but not on $\boldsymbol{\gamma}$).

Instead of proving that $\boldsymbol{\eta}$ lies in the set H, we will show that it lies in a subgroup $\Gamma^{1/n}$ of Γ^{div} consisting of all $\boldsymbol{\zeta}$ with $\boldsymbol{\zeta}^n$ in Γ. Here the positive integer $n = n(\mathcal{C}, \Gamma)$ will be effective. As $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\gamma}) \leq C$ and $\mathrm{dist}(\boldsymbol{\eta}, \gamma) < \varepsilon < 1$ we deduce $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\eta}^n) \leq n(C+1)$. So by Lemma 2.3 (b) there are at most finitely possibilities for $\boldsymbol{\eta}^n$ and therefore also $\boldsymbol{\eta}$ itself, thus yielding our H.

Let $P \in \mathcal{C}(\overline{\mathbb{Q}})$, such that the value $\boldsymbol{\gamma}_P$ is defined and lies in V. We homogenize replacing Γ with $\Gamma_1 = \Gamma \times \{1\}$ and $\boldsymbol{\gamma}$ by $(\boldsymbol{\gamma}, 1) \in \Gamma^{\mathrm{div}} \cap \{x_{r+1} = 1\}$. Then, after changing the signs of the coefficients α_i in (1.9), this equation reads

$$\boldsymbol{\alpha}.\boldsymbol{\gamma}_P = \alpha_1(\gamma_1)_P + \cdots + \alpha_r(\gamma_r)_P + \alpha_{r+1}(\gamma_{r+1})_P = 0. \tag{5.1}$$

with $\alpha_{r+1} = 1$ so $\boldsymbol{\alpha}$ gets replaced by $(\boldsymbol{\alpha}, 1)$. Given $\mathbf{a} \in \Gamma_1$ and $\lambda \in \mathbb{Q}$, we choose from now on $\mathbf{a}^\lambda \in \Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}$.

We write $\Gamma_1 = \Gamma_{\mathrm{tors}} \oplus \Gamma_{\mathrm{free}}$ where Γ_{free} is freely generated by, say, $\mathbf{a}_1, \ldots, \mathbf{a}_\kappa \in \mathbb{G}_{\mathrm{m}}^{r+1} \cap \{x_{r+1} = 1\}$. We have

$$c_1^{-1} \max |r_j| \leq h^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{a}_1^{r_1} \cdots \mathbf{a}_\kappa^{r_\kappa}) \leq c_1 \max |r_j| \tag{5.2}$$

where the second inequality is trivial and the first is Lemma 2.3. Since $h^{\mathbb{A}}(\mathbf{x}) \leq \sum h^{\mathbb{A}}(x_i)$ and $h^{\mathbb{A}}(x^r) = |r| h^{\mathbb{A}}(x)$, applying Lemma 2.1 with $r = 1$ and $f_1 = a_{ji}$ we get

$$h^{\mathbb{A}}((\mathbf{a}_1^{r_1} \cdots \mathbf{a}_\kappa^{r_\kappa})_P) \leq \sum_{i,j} |r_j| h\big((a_{ji})_P\big) \leq c_2(\max |r_j|) h(P) \tag{5.3}$$

(assuming $h(P) \geq 1$ as we may). We write

$$\boldsymbol{\gamma} = \boldsymbol{\omega}\mathbf{a}_1^{\lambda_1} \cdots \mathbf{a}_\kappa^{\lambda_\kappa}.$$

where $\boldsymbol{\omega} \in (\overline{\mathbb{Q}}^*_{\mathrm{tors}})^{r+1} \cap \{x_{r+1} = 1\}$ and $\lambda_1, \ldots, \lambda_\kappa \in \mathbb{Q}$. Since $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\gamma}) \leq C$, by the lower bound in (5.2) we have

$$|\lambda_j| \leq c_1 C, \qquad j = 1, \ldots, \kappa. \tag{5.4}$$

Let $q \in \mathbb{N}$ and $p_1, \ldots, p_\kappa \in \mathbb{Z}$ defined by

$$q := 1 + [(4nc_2C_3 + rc_1)/\varepsilon], \qquad p_j := [q\lambda_j] \tag{5.5}$$

where $C_3 = C_3(\mathcal{C}, \Gamma_1)$ and $n = n(\mathcal{C}, \Gamma_1)$ are as in Lemma 4.2. Thus

$$\left|\lambda_j - \frac{p_j}{q}\right| < \frac{1}{q} \quad \text{for } i = 1, \ldots, \kappa. \tag{5.6}$$

We consider the vector

$$\mathbf{v} = \boldsymbol{\alpha}\mathbf{a}_1^{p_1/q} \cdots \mathbf{a}_\kappa^{p_\kappa/q} \in (\overline{\mathbb{Q}}^*)^{r+1}\Gamma_1^{\mathrm{div}} \cap \{x_{r+1} = 1\}.$$

By (5.5) and (5.4) we have

$$\begin{aligned} q &\leq 1 + (4nc_2C_3 + rc_1)/\varepsilon, \\ |p_j| &\leq q|\lambda_j| + 1 \leq (1 + (4nc_2C_3 + rc_1)/\varepsilon)c_1C + 1, \qquad j = 1, \ldots, \kappa. \end{aligned}$$

Thus $\mathbf{v}$ belongs to a finite set depending only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$. By (4.5), by the upper bound in (5.2) and by (5.6),

$$\mathrm{dist}(\mathbf{v}, \boldsymbol{\gamma}) \leq rh^{\mathbb{A}}_{\mathrm{geo}}(\mathbf{v}\boldsymbol{\gamma}^{-1}) = rh^{\mathbb{A}}_{\mathrm{geo}}\big(\mathbf{a}_1^{p_1/q-\lambda_1} \cdots \mathbf{a}_\kappa^{p_\kappa/q-\lambda_\kappa}\big) \leq rc_1/q. \tag{5.7}$$

We denote by $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_r$ the successive infima of h_{geo} on $\mathbf{v}^\perp$. By [16, Theorem 2.2], $\mu_1 + \cdots + \mu_r \leq h_{\mathrm{geo}}(\mathbf{v}^\perp)$. On the other hand a standard argument shows that we have indeed equality. Moreover, by well-known facts on the height of subspaces (see [17, §8, p. 28]), $h_{\mathrm{geo}}(\mathbf{v}^\perp) = h_{\mathrm{geo}}(\mathbf{v})$. Thus

$$\mu_1 + \mu_2 + \cdots + \mu_r = h_{\mathrm{geo}}(\mathbf{v}). \tag{5.8}$$

To simplify the proof, we assume from now on that the μ_i are attained. Then Lemma 4.3 gives us integers $0 = d_0 < d_1 < \cdots < d_k = r$ and a $\overline{\mathbb{F}}$-basis $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(r)}$ of $\mathbf{v}^\perp$ such that $h_{\mathrm{geo}}(\mathbf{g}^{(i)}) = \mu_i$. The $\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(r)}$ belong to a finite set depending only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$. They are linearly independent over $\overline{\mathbb{F}}$. It is easy to see that outside a finite set of P in $\mathcal{C}(\overline{\mathbb{Q}})$ (depending only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$), their specializations $\mathbf{g}^{(1)}_P, \ldots, \mathbf{g}^{(r)}_P$ remain linearly independent over $\overline{\mathbb{Q}}$ (for example consider maximal minors of the matrix of coordinates). Similarly $\mathbf{v}_P \neq 0$, so $\mathbf{v}_P^\perp$ remains of dimension r over $\overline{\mathbb{Q}}$. So $\mathbf{g}^{(1)}_P, \ldots, \mathbf{g}^{(r)}_P$ are a $\overline{\mathbb{Q}}$-basis for $\mathbf{v}_P^\perp$. Let

$$\boldsymbol{\beta} = \boldsymbol{\omega}\big(\mathbf{a}_1^{\lambda_1-p_1/q} \cdots \mathbf{a}_\kappa^{\lambda_\kappa-p_\kappa/q}\big)_P \in \overline{\mathbb{Q}}^{r+1} \cap \{x_{r+1} = 1\}.$$

By the upper bound in (5.3) and by (5.6),

$$h^{\mathbb{A}}(\boldsymbol{\beta}) \leq c_2 \max_j \left|\lambda_j - \frac{p_j}{q}\right| h(P) \leq \frac{c_2}{q} h(P). \tag{5.9}$$

Note that

$$\boldsymbol{\beta}\mathbf{v}_P = \boldsymbol{\alpha}\boldsymbol{\gamma}_P, \tag{5.10}$$

hence $\boldsymbol{\beta}.\mathbf{v}_P = \boldsymbol{\alpha}.\boldsymbol{\gamma}_P = 0$ by (5.1). Thus $\boldsymbol{\beta}$ is a (non-zero) vector in $\mathbf{v}_P^{\perp}$. We get a basis of $\mathbf{v}_P^{\perp}$ from $\boldsymbol{\beta}, \mathbf{g}_P^{(1)}, \ldots, \mathbf{g}_P^{(r)}$ by discarding a vector $\mathbf{g}_P^{(i)}$ for some $i \in \{1, \ldots, r\}$. More precisely, let[3]

$$j' = \text{largest integer } 0 \le j \le k \text{ s.t. } \boldsymbol{\beta} \notin \mathrm{Span}_{\overline{\mathbb{Q}}}(\mathbf{g}_P^{(1)}, \ldots, \mathbf{g}_P^{(d_j)}). \tag{5.11}$$

Note that j' is well-defined and $< k$ since $\boldsymbol{\beta}$ is a non-zero vector in $\mathbf{v}_P^{\perp}$. By (5.11) we get a basis from $\boldsymbol{\beta}, \mathbf{g}_P^{(1)}, \ldots, \mathbf{g}_P^{(r)}$ by discarding a vector $\mathbf{g}_P^{(i')}$ for some i' with

$$d_{j'} + 1 \le i' \le r. \tag{5.12}$$

We now denote by h and h_2 the projective height on $\overline{\mathbb{Q}}^{r+1}$ defined on choosing respectively the L_∞ and the L_2-norm at the infinite places. We also denote by the same letter h_2 the height on $\overline{\mathbb{Q}}$-subspaces (see [17, §8, p. 28]). Note that $h(\mathbf{x}) \le h_2(\mathbf{x}) \le h(\mathbf{x}) + \frac{1}{2}\log(r+1)$ for $\mathbf{x} \in \overline{\mathbb{Q}}^{r+1}$ and $h(\mathbf{x}) = h^{\mathbb{A}}(\mathbf{x})$ if moreover $x_{r+1} = 1$. By well-known facts (see again [17, §8, p. 28])

$$\begin{aligned} h(\mathbf{v}_P) \le h_2(\mathbf{v}_P) &= h_2(\mathrm{Span}_{\overline{\mathbb{Q}}}(\mathbf{v}_P)) \\ &= h_2(\mathbf{v}_P^{\perp}) \\ &\le h_2(\boldsymbol{\beta}) + \sum_{j \ne i'} h_2(\mathbf{g}_P^{(j)}) \\ &\le h^{\mathbb{A}}(\boldsymbol{\beta}) + \sum_{j \ne i'} h(\mathbf{g}_P^{(j)}) + c_3. \end{aligned} \tag{5.13}$$

By the projective version of Lemma 2.1 and taking into account that $\mathbf{g}^{(j)}$ belongs to a finite set depending only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$, we find

$$\begin{aligned} h(\mathbf{g}_P^{(i)}) &\le h_{\mathrm{geo}}(\mathbf{g}^{(i)})h(P) + c_4(1 + h(P)^{1/2}) \\ &= \mu_i h(P) + c_4(1 + h(P)^{1/2}) \end{aligned}$$

for $i = 1, \ldots, r$. Thus, by (5.8),

$$\sum_{i \ne i'} h(\mathbf{g}_P^{(i)}) \le (h_{\mathrm{geo}}(\mathbf{v}) - \mu_{i'})h(P) + rc_4(1 + h(P)^{1/2}).$$

Inserting (5.9) and this last inequality in (5.13) we get

$$h(\mathbf{v}_P) \le \left(\frac{c_2}{q} + h_{\mathrm{geo}}(\mathbf{v}) - \mu_{i'}\right)h(P) + rc_4(1 + h(P)^{1/2}).$$

On the other hand, again by the projective version of Lemma 2.1 and since $\mathbf{v}$ belongs to a finite set depending only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$,

$$h(\mathbf{v}_P) \ge h_{\mathrm{geo}}(\mathbf{v})h(P) - c_5(1 + h(P)^{1/2}).$$

Comparing the upper bound and the lower bound for $h^{\mathbb{A}}(\mathbf{v}_P)$ we get:

$$\left(\mu_{i'} - \frac{c_2}{q}\right) h(P) \le (rc_4 + c_5)(1 + h(P)^{1/2}).$$

[3]With the convention $\mathrm{Span}_{\overline{\mathbb{Q}}}(\mathbf{g}_P^{(1)}, \ldots, \mathbf{g}_P^{(d_0)}) = \{0\}$.

If $\mu_{i'} \geq \frac{2c_2}{q}$ the height of P is uniformly bounded. Thus we can assume

$$\mu_{i'} < \frac{2c_2}{q}.$$

Since $d_{j'} + 1 \leq i'$ by (5.12) and since $q \geq 4nc_2C_3$ by (5.5), we have

$$\mu_{d_{j'}+1} < \frac{2c_2}{q} \leq (2nC_3)^{-1}. \tag{5.14}$$

Let $U = \mathrm{Span}_{\overline{\mathbb{F}}}(\mathbf{g}^{(1)}, \ldots, \mathbf{g}^{(d_{j'}+1)})\mathbf{v}$. Then U is a non-zero subspace on which $x_1 + \cdots + x_{r+1} = 0$. By Lemma 4.3 (with $j = j' + 1$), and by (5.14), there exists $\boldsymbol{\eta} \in \Gamma_1^{1/n} \cap \{x_{r+1} = 1\}$ with

$$\mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) \leq C_3\mu_{d_{j'}+1} \leq \frac{2c_2C_3}{q}$$

such that $U\boldsymbol{\eta}^{-1}$ is defined over $\overline{\mathbb{Q}}$. By the triangular inequality and by (5.7),

$$\mathrm{dist}(\boldsymbol{\eta}, \boldsymbol{\gamma}) \leq \mathrm{dist}(\boldsymbol{\eta}, \mathbf{v}) + \mathrm{dist}(\mathbf{v}, \boldsymbol{\gamma}) \leq \frac{2c_2C_3 + rc_1}{q}.$$

Since $q > (2c_2C_3 + rc_1)/\varepsilon$ (again by (5.5)), we get

$$\mathrm{dist}(\boldsymbol{\eta}, \boldsymbol{\gamma}) < \varepsilon.$$

By (5.10) and by definition (5.11) of j' we have

$$\boldsymbol{\alpha}\boldsymbol{\gamma}_P = \boldsymbol{\beta}\mathbf{v}_P \in \mathrm{Span}_{\overline{\mathbb{Q}}}(\mathbf{g}_P^{(1)}, \ldots, \mathbf{g}_P^{(d_{j'}+1)}) \cdot \mathbf{v}_P.$$

We shall apply Lemma 4.4 to U with $\boldsymbol{\psi} = \boldsymbol{\eta}$. Since all the quantities there now depend only on $\mathcal{C}$, Γ, ε and $\boldsymbol{\alpha}$, the space on the right-hand side of the displayed formula can assumed to be U_P. Thus $(\boldsymbol{\alpha}\boldsymbol{\gamma}_P\boldsymbol{\eta}_P^{-1}).\boldsymbol{\eta} = 0$ in $\overline{\mathbb{F}}^{r+1}$, which on descending to $\overline{\mathbb{F}}^r$ is what we want in Theorem 1.5. This concludes the proof for $\boldsymbol{\gamma}$ of small height.

6. Proofs of Theorems 1.2, 1.3, and 1.4.

As Theorem 1.2 is considerably sharpened in the Appendix A, we shall present here only a brief deduction of the first assertion of that theorem from Theorem 1.5. In fact we carry this out for the stronger Theorem 1.3.

As mentioned in the introduction, we take $\mathcal{C} = \mathbb{A}^1$ and $\mathbb{F} = \overline{\mathbb{Q}}(t)$, with Γ generated by $(t, 1), (1, 1 - t)$, and $\boldsymbol{\alpha} = (1, 1)$. Trivially, we do not have vanishing subsums. We obtain effective C and an effective finite subset H of Γ^{div} such that for $P = t_0$ and $\boldsymbol{\gamma} = (t^\lambda, (1 - t)^\mu)$, the height $h(t_0)$ is bounded by a function of ε unless $h^{\mathbb{A}}_{\mathrm{geo}}(\boldsymbol{\gamma}) \leq C$ and there is $\boldsymbol{\eta} \in H$ with

$$\mathrm{dist}(\boldsymbol{\eta}, \boldsymbol{\gamma}) < \varepsilon, \quad \boldsymbol{\alpha}.(\boldsymbol{\gamma}_P\boldsymbol{\eta}_P^{-1}\boldsymbol{\eta}) = 1.$$

We write $\boldsymbol{\eta} = (t^a, (1 - t)^b)$ for rational a, b and get the equations

$$\max\{|\lambda|, |\mu|\} \leq C, \quad \max\{|\lambda-a|, |\mu-b|\} < \varepsilon, \quad t_0^{\lambda-a}t^a+(1-t_0)^{\mu-b}(1-t)^b = 1$$

identically in t. It is easy to see that there are β_1, $\beta_2 \in \overline{\mathbb{Q}}^*$ with $\beta_1t^a + \beta_2(1 - t)^b = 1$ if and only if $(a, b) = (0, 0)$ and $\beta_1 + \beta_2 = 1$ or $(a, b) = (1, 1)$ and $\beta_1 = \beta_2 = 1$. The first case is ruled out because $\max\{|\lambda|, |\mu|\} > \varepsilon$ and in the second case t_0 or $1 - t_0$ is a root of unity because $(\lambda, \mu) \neq (1, 1)$.

This completes the proof of Theorem 1.3 and so of the first part of Theorem 1.2.

We next show that we do not have uniformly bounded height for the solutions of $t^\lambda + (1-t)^\lambda = 1$ when $\lambda \to 0$ and even the analogue for Theorem 1.3.

Lemma 6.1. *For rational $\lambda > 0, \mu > 0$ let $t_0 \in \overline{\mathbb{Q}}$ be a solution of $t^\lambda + (1-t)^\mu = 1$. If t_0^λ is not $0, 1$, or $\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$, then*

$$h(t_0) \geq \frac{1}{4\max\{\lambda, \mu\}} \log\left(\frac{1+\sqrt{5}}{2}\right) - \frac{1}{2}\log 2.$$

Proof. Put $\alpha = t_0^\lambda$. By assumption, $\alpha \neq 0, 1, \frac{1}{2} \pm \frac{i}{2}\sqrt{3}$. By a result of Zagier [22] (which makes explicit in this very special setting the toric case of Bogomolov's conjecture, now a theorem of Zhang [23]), $h(\alpha) + h(1-\alpha) \geq c$, where $c = \frac{1}{2}\log(\frac{1+\sqrt{5}}{2})$ (this constant is best possible). We have $h(\alpha) = \lambda h(t_0)$, $h(1-\alpha) = \mu h(1-t_0)$ and $h(1-t_0) \leq h(t_0) + \log 2$. Putting all together we get the desired result. □

Note that such t_0 really do exist. To see this we observe that for any positive integer n there is a polynomial $P_n(t)$ which for any s_1, s_2 with $s_1^n = t, s_2^n = 1-t$ is the product of $\omega_1 s_1 + \omega_2 s_2 - 1$ over all roots of unity ω_1, ω_2 of order dividing n. We take any n divisible by 6, and using the Puiseux expansions

$$s_1 = t^{1/n}, \quad s_2 = \zeta(t-1)^{1/n} = \zeta(t^{1/n} + \cdots) \qquad (\zeta = e^{\pi i/n})$$

at $t = \infty$ we check that P_n has degree $n^2/n = n$ (note that $\omega_1 + \zeta\omega_2 \neq 0$ because Euler's $\phi(2n) = 2\phi(n) > \phi(n)$). And using $s_1 = t^{1/n}, s_2 = 1 + \cdots$ at $t = 0$ we check that $\mathrm{ord}_{t=0} P_n = 1$; and similarly $\mathrm{ord}_{t=1} P_n = 1$.

Thus P_n has a zero $t_0 \neq 0, 1$, which is then a solution of $t^\lambda + (1-t)^\mu = 1$ for $\lambda = \mu = 1/n$. And t_0^λ cannot be 1 or a sixth root of unity $\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$ because $t_0 \neq 1$. So t_0 is as in Lemma 6.1. And if n varies we see from the lower bound there that we get infinitely many t_0 in this way.

It may be interesting to note that if we choose this n such that the lower bound is bigger than $\log 216$ in (1.2), then the resulting t_0 cannot satisfy $t^m + (1-t)^m = 1$ for any integer $m \geq 2$ (compare (1.12) and (1.13)).

Now we turn to Theorem 1.4. If we have vanishing subsums in $t_0^\lambda + (1-t_0)^\lambda + (1+t_0)^\lambda = 1$, we find easily $t_0 = \pm i$ which we can exclude. Then a similar argument leads to bounded $h(t_0)$ unless there is a rational a with

$$\lambda < C, \quad |\lambda - a| < \varepsilon, \quad t_0^{\lambda-a}t^a + (1-t_0)^{\lambda-a}(1-t)^a + (1+t_0)^{\lambda-a}(1+t)^a = 1$$

identically in t. The last now forces $a = 0, 1, 2$. As above $a = 0$ is ruled out. If $a = 2$ then we must have

$$t_0^{\lambda-2} = -1, \quad (1-t_0)^{\lambda-2} = \frac{1}{2}, \quad (1+t_0)^{\lambda-2} = \frac{1}{2}.$$

Here $\lambda = 2$ is impossible, and then we see again that t_0 is a root of unity. Finally $a = 1$ leads to the pair

$$t_0^{\lambda-1} - 2(1-t_0)^{\lambda-1} = -1, \quad t_0^{\lambda-1} + 2(1+t_0)^{\lambda-1} = 1 \tag{6.1}$$

with $|\lambda - 1| < \varepsilon$. We shall show in the Appendix B that these equations have no solutions. This completes the proof of Theorem 1.4.

We notice that again we do not have uniformly bounded height for the solutions of $t^\lambda + (1-t)^\lambda + (1+t)^\lambda = 1$ when $\lambda \to 0$.

Lemma 6.2. *For rational $\lambda > 0$ let $t_0 \in \overline{\mathbb{Q}}$ be a solution of $t^\lambda + (1-t)^\lambda + (1+t)^\lambda = 1$. If $t_0^\lambda, (1-t_0)^\lambda, (1+t_0)^\lambda$ are all not 0 or 1, then*

$$h(t_0) \geq \frac{1}{6\lambda} \log\left(\frac{1+\sqrt{5}}{2}\right) - \frac{2}{3}\log 2.$$

Proof. Warin obtained the following explicit version of another special case of Zhang's result. Namely his Theorem III.1 (p. 23) of [21] says that if α, β, γ are in $\overline{\mathbb{Q}}^*$ and all not 1 with $\alpha + \beta + \gamma = 1$ then

$$h(\alpha) + h(\beta) + h(\gamma) \geq c$$

for the same c as above (also here best possible). We take $\alpha = t_0^\lambda, \beta = (1-t_0)^\lambda, \gamma = (1+t_0)^\lambda$, none of which are $0, 1$, and argue as before. □

Also as before we can show that such t_0 exist, this time by considering the product Q_n of $\omega_1 s_1 + \omega_2 s_2 + \omega_3 s_3 - 1$, with s_1, s_2 as above and $s_3^n = 1 + t$. We can check, now for n prime to 6, that Q_n has degree n^2 (here we use $\omega_1 - \omega_2 + \omega_3 \neq 0$). And $Q_n(0) \neq 0$ (here $\omega_2 + \omega_3 - 1 \neq 0$) as well as $Q_n(1) \neq 0$ (here $\omega_1 + \omega_3 2^{1/n} - 1 \neq 0$ because $2^{1/n}$ has degree $n > \phi(n)$) and $Q_n(-1) \neq 0$ (here $-\omega_1 + \omega_2 2^{1/n} - 1 \neq 0$ similarly). Thus Q_n has a zero $t_0 \neq 0, 1, -1$. For $\lambda = 1/n$ it follows that $t_0^\lambda, (1-t_0)^\lambda, (1+t_0)^\lambda$ are not $0, 1$ and so t_0 is as in Remark 6.2. Again by varying n we get infinitely many t_0.

Appendix A. Beukers' equation with a rational exponent

In this appendix we prove the following theorem which generalizes Theorem 1.2 stated in the introduction.

Theorem A.1. *Let α, β be non-zero algebraic numbers and let $\lambda \in \mathbb{Q}$ be positive. Let $t_0 \in \overline{\mathbb{Q}} \backslash \{0, 1\}$. We fix determinations of t_0^λ and of $(1-t_0)^\lambda$. Let us assume*

$$\alpha t_0^\lambda + \beta(1-t_0)^\lambda = 1$$

and

$$(\alpha t_0^{\lambda-1}, \beta(1-t_0)^{\lambda-1}) \neq (1, 1).$$

Then

$$h(t_0) \leq 100 \max(1, \lambda^{-1}) + 121\lambda^{-1}(h(\alpha) + h(\beta)).$$

The special algebraic numbers t_0 such that $\alpha t_0^{\lambda-1} = \beta(1-t_0)^{\lambda-1} = 1$ trivially satisfy the equation $\alpha t^\lambda + \beta(1-t)^\lambda = 1$. They can be directly handled as we explain after the proof of the theorem. Thus Theorem 1.2 follows from Theorem A.1; see again the discussion after the proof for details.

Our theorem extends the following result of Beukers and Schlickewei which deals with even integers λ, and will be deduced from it.

Theorem A.2 ([5], Lemma 2.3)**.** *Let $a,\ b,\ A,\ B \in \overline{\mathbb{Q}}^*$ such that*

$$A + B = 1 \quad \textit{and} \quad aA^{2n} + bB^{2n} = 1$$

for some integer $n \in \mathbb{N}$. Then[4] $H(A, B) \le 6\sqrt{3} \cdot 2^{1/n} H(a, b)^{1/n}$.

Proof of Theorem A.1. Let $\lambda > 0$ be a rational parameter and let $t_0 \in \overline{\mathbb{Q}}$ which satisfies $\alpha t_0^\lambda + \beta(1 - t_0)^\lambda = 1$ and such that

$$(\alpha t_0^{\lambda-1}, \beta(1 - t_0)^{\lambda-1}) \neq (1, 1).$$

We shall prove:

$$h(t_0) \le \begin{cases} 8\lambda^{-1} + 10\lambda^{-1}(h(\alpha) + h(\beta)), & \text{if } \lambda \le 1/6; \\ 100\lambda^{-1} + 121\lambda^{-1}(h(\alpha) + h(\beta)), & \text{if } 1/6 \le \lambda \le 1; \\ 100 + 20(h(\alpha) + h(\beta)), & \text{if } 1 \le \lambda \le 6; \\ 8 + 9\lambda^{-1}(h(\alpha) + h(\beta)), & \text{if } \lambda \ge 6. \end{cases} \tag{A.1}$$

Theorem A.1 follows after a simple computation.

The strategy of the proof of (A.1) is the following. We distinguish three cases: for $\lambda \ge 6$ we apply Theorem A.2 choosing for n the integer part of $\lambda/2$. For $\lambda \in [1, 6)$ we simply use the relations between roots and coefficients. Finally, we reduce the case $\lambda \in (0, 1]$ to the previous ones by a duality argument involving exponent $1/\lambda$.

First case: $\lambda \ge 6$. Let n be the integer part of $\lambda/2$. By assumption $n \ge 3$. Let $a_0 = t_0^{\lambda-2n}$, $b_0 = (1 - t_0)^{\lambda-2n}$ and $a = \alpha a_0$, $b = \beta b_0$. Thus $a t_0^{2n} + b(1 - t_0)^{2n} = 1$. By Theorem A.2,

$$\begin{aligned} H(t_0, 1 - t_0) &\le 6\sqrt{3} \cdot 2^{1/n} H(a, b)^{1/n} \\ &\le 6\sqrt{3} \cdot 2^{1/n} H(\alpha : \beta : 1)^{1/n} H(a_0, b_0)^{1/n}. \end{aligned}$$

Since $\lambda - 2n \ge 0$,

$$H(a_0, b_0)^{1/n} = H(t_0, 1 - t_0)^{(\lambda-2n)/n}.$$

Since $\lambda < 2(n + 1)$ and $n \ge 3$ we have

$$1 - \frac{\lambda - 2n}{n} > 1 - \frac{2}{n} \ge \frac{1}{3}$$

and $2^{1/n} \le 2^{1/3}$. Thus

$$H(t_0, 1 - t_0) \le (6\sqrt{3} \cdot 2^{1/3})^3 H(\alpha : \beta : 1)^{3/n}.$$

We have $3 \log(6\sqrt{3} \cdot 2^{1/3}) \le 8$ and $n > \lambda/2 - 1 \ge \lambda/3$ (since $\lambda \ge 6$). Thus

$$h(t_0) \le h^{\mathbb{A}}(t_0, 1 - t_0) \le 8 + 9\lambda^{-1} h^{\mathbb{A}}(\alpha, \beta) \le 8 + 9\lambda^{-1}(h(\alpha) + h(\beta)).$$

Second case: $1 \le \lambda \le 6$. Note that we can find a rational p/q with $\gcd(p, q) = 1$ and $q = 1$ or $q = 2$ such that

$$\varepsilon := \left|\lambda - \frac{p}{q}\right| \le \frac{2}{5q^2}$$

[4]We denote by $H = \exp(h)$ the affine non-logarithmic Weil's height.

(indeed, if the fractional part $\{\lambda\}$ is ≤ 0.4 or ≥ 0.6, then the inequality holds with $q = 1$; otherwise $|\{\lambda\} - 1/2| \leq 0.1$ and it holds with $q = 2$). Since $\lambda \geq 1$ and $\varepsilon < 1$ we have $p \neq 0$. Let $a_0 = t_0^{\lambda - p/q}$, $b_0 = (1-t_0)^{\lambda - p/q}$ and $a = \alpha a_0$, $b = \beta b_0$. Thus

$$bat_0^{p/q} + b(1-t_0)^{p/q} = 1$$

which means that there exists $s_0 \in \overline{\mathbb{Q}}$ such that $s_0^q = t_0$ and $1 - as_0^p = b(1-s_0^q)^{p/q}$. Taking q powers, we see that s_0 is a root of the polynomial

$$f := (1-as^p)^q - b^q(1-s^q)^p \in \overline{\mathbb{Q}}[s].$$

Fact. $f \neq 0$.

Proof. We have

$$(1-as^p)^q = 1 - qas^p + \text{higher terms},$$
$$b^q(1-s^q)^p = b^q - pb^q s^q + \text{higher terms}.$$

Assume first $(p,q) \neq (1,1)$. Since $\gcd(p,q) = 1$ we have $p \neq q$. We also recall that p, $q \geq 1$. Comparing the two expansions above, we deduce that $b^q = 1$. Moreover, if $q < p$ then $b^q = 0$, which is clearly impossible. But $a \neq 0$ since $\alpha \neq 0$, $a \neq 0$. Thus $(p,q) = (1,1)$ and $f = (1-b) + (b-a)s$ which is $= 0$ if and only if $a = b = 1$, that is

$$\alpha t_0^{\lambda-1} = \beta(1-t_0)^{\lambda-1} = 1$$

which we have excluded in our assumption. □

We now recall a classical relation between the height of the roots of a polynomial and the height of its coefficients. Let $K := \mathbb{Q}(a,b)$. Given a place v of K we denote by $M_v(f)$ the Mahler measure of f^σ, if v is archimedean and corresponds to the immersion $\sigma\colon K \hookrightarrow \overline{\mathbb{Q}}$. If v is non-archimedean, we let $M_v(f)$ be the maximum of the v-adic absolute values of the coefficients of f. We then define the normalized height of f as

$$\hat{h}(f) = \frac{1}{[K:\mathbb{Q}]} \sum_v d_v \log M_v(f)$$

where v runs over the places of K and where d_v denote the local degree. It is well known that $\hat{h}(f) = \sum_s h(s)$ for s running over the roots of f, counted with multiplicities. Thus $h^{\mathbb{A}}(s_0) \leq \hat{h}(f)$ and we have to estimate this last quantity.

Given a non-archimedean place of K we have

$$M_v(f) \leq \max\{|a|_v, |b|_v, 1\}^q.$$

Let now v be an archimedean place, corresponding to the immersion $\sigma\colon K \hookrightarrow \overline{\mathbb{Q}}$. Since the Mahler measure of a polynomial is bounded by the maximum of the absolute value of the polynomial on the disk of radius 1, we have:

$$\begin{aligned} M_v(f) &\leq \max_{|s|=1} |(1-a^\sigma s^p)^q - (b^\sigma)^q(1-s^q)^p| \\ &\leq (1+|a|_v)^q + 2^p|b|_v^q \\ &\leq (2^q+2^p)\max\{|a|_v, |b|_v, 1\}^q. \end{aligned}$$

Thus

$$\hat{h}(f) \leq qh^{\mathbb{A}}(a,b) + \log(2^q + 2^p) \leq q(h^{\mathbb{A}}(\alpha,\beta) + h^{\mathbb{A}}(a_0,b_0) + \log(2 + 2^{p/q})).$$

By our choices of a_0, b_0 and ε,

$$\begin{aligned} h^{\mathbb{A}}(a_0,b_0) &\leq h(a_0) + h(b_0) = \varepsilon(h(t_0) + h(1-t_0)) \leq 2\varepsilon h(t_0) + \varepsilon \log 2 \\ &\leq \frac{4h(t_0)}{5q^2} + \frac{2\log 2}{5q^2} \end{aligned}$$

and $p/q \leq \lambda + \epsilon \leq 7$. Thus, taking into account $q \leq 2$,

$$\begin{aligned} h(t_0) &= qh(s_0) \leq q\hat{h}(f) \\ &\leq q^2(h^{\mathbb{A}}(\alpha,\beta) + h^{\mathbb{A}}(a_0,b_0) + \log(2 + 2^{p/q})) \\ &\leq \frac{4h(t_0)}{5} + \frac{2\log 2}{5} + 4(\log(2+2^7) + h^{\mathbb{A}}(\alpha,\beta)) \end{aligned}$$

and so

$$\begin{aligned} h(t_0) &\leq 2\log 2 + 5 \cdot 4(\log(2+2^7) + h^{\mathbb{A}}(\alpha,\beta)) \\ &\leq 100 + 20(h(\alpha) + h(\beta)). \end{aligned}$$

Third case: $\lambda \leq 1$. We reduce to the previous two cases. Note that $t_0' = \alpha t_0^{\lambda}$ is a solution of

$$\alpha^{-\lambda^{-1}}(t_0')^{\lambda^{-1}} + \beta^{-\lambda^{-1}}(1-t_0')^{\lambda^{-1}} = 1$$

and

$$h(t_0) = \lambda^{-1}h^{\mathbb{A}}(\alpha^{-1}t_0') \leq \lambda^{-1}h^{\mathbb{A}}(t_0') + \lambda^{-1}h(\alpha).$$

Since $\lambda^{-1} \geq 1$ the results of the previous cases apply. Suppose first $\lambda \leq 1/6$. Then $\lambda^{-1} \geq 6$ and

$$\begin{aligned} h(t_0) &\leq \lambda^{-1}h(t_0') + \lambda^{-1}h(\alpha) \\ &\leq \lambda^{-1}\Big(8 + 9\lambda h(\alpha^{-\lambda^{-1}}) + 9\lambda h(\beta^{-\lambda^{-1}})\Big) + \lambda^{-1}h(\alpha) \\ &\leq \lambda^{-1}\big(8 + 9h(\alpha) + 9h(\beta)\big) + \lambda^{-1}h(\alpha) \\ &\leq 8\lambda^{-1} + 10\lambda^{-1}(h(\alpha) + h(\beta)). \end{aligned}$$

Suppose now $1/6 \leq \lambda \leq 1$. Then $1 \leq \lambda^{-1} \leq 6$ and

$$\begin{aligned} h(t_0) &\leq \lambda^{-1}h(t_0') + \lambda^{-1}h(\alpha) \\ &\leq \lambda^{-1}\big(100 + 20h(\alpha^{-\lambda^{-1}}) + 20h(\beta^{-\lambda^{-1}})\big) + \lambda^{-1}h(\alpha) \\ &\leq 100\lambda^{-1} + (20\lambda^{-1} + 1)\lambda^{-1}(h(\alpha) + h(\beta)) \\ &\leq 100\lambda^{-1} + 121\lambda^{-1}(h(\alpha) + h(\beta)). \end{aligned}$$

□

We remark that the duality trick which we have used in the third case is only needed to get a better bound when λ is close to zero. We could indeed modify the proof in the second case in order to get a result, depending on a fixed $\varepsilon \in (0,1]$, which holds for $\varepsilon \leq \lambda \leq 1$.

As promised, we now study the special algebraic numbers t_0 such that

$$\alpha t_0^{\lambda-1} = \beta(1-t_0)^{\lambda-1} = 1. \tag{A.2}$$

Then t_0 trivially satisfies the equation $\alpha t^\lambda + \beta(1-t)^\lambda = 1$.

Let us first assume $\lambda = 1$. Then (A.2) is satisfied if and only if $\alpha = \beta = 1$ and in this case the equation $\alpha t^\lambda + \beta(1-t)^\lambda = 1$ becomes trivial. Thus we do not have bounded height.

Let now suppose $\lambda \neq 1$. Since

$$t_0 = \alpha^{-1/(\lambda-1)} \text{ and } 1 - t_0 = \beta^{-1/(\lambda-1)} \tag{A.3}$$

we still have bounded height, but the bound seems to go to infinity as $\lambda \to 1$, unless α, β are both roots of unity. However we have the additional equation

$$\alpha^{-1/(\lambda-1)} + \beta^{-1/(\lambda-1)} = 1 \tag{A.4}$$

and it may be seen in several way that if α, β are not both roots of unity, this equation determines at most finitely many values of $\lambda \neq 1$ in terms of α and β.

For example, this is one of the few effective instances of the Skolem-Mahler-Lech Theorem, which follows from linear forms in logarithms combining [13, Theorem 1, p. 65] with Kummer theory. Or we could simply apply Liardet's Theorem (as made effective for example by Bérczes, Evertse and Győry [3] Theorem 2.3); this possibility was mentioned in [6] pp. 1121-1122.

In any case we get an effective bound $h(t_0) \leq C(\alpha, \beta)$.

Examples of (A.4) are

$$\alpha = 2, \quad \beta^q = \frac{2^q}{2^q - 1}, \quad \lambda = 1 + \frac{1}{q}$$

for positive integer q with $t_0 = 2^{-q}$. Letting $q \to \infty$ shows that such a $C(\alpha, \beta)$ cannot be bounded as any function of $h(\alpha) + h(\beta)$.

In the special case $\alpha = \beta = 1$, the solutions t_0 of the equation (A.3) are roots of unity, and thus, by the previous remarks, any solution of $t^\lambda + (1-t)^\lambda = 1$ with $\lambda > 0$ and $\lambda \neq 1$ satisfies $h(t_0) \leq 100 \max(1, \lambda^{-1})$. This completes the proof of Theorem 1.2 stated in the introduction.

Appendix B. Exceptional solutions of the Denz equation.

We return to the system (6.1); this will finally settle the Denz equation as in Theorem 1.4.

Proposition B.1. *Suppose δ is rational with $0 < |\delta| \leq 10^{-330}$. Then there is no $t_0 \neq 0, 1, -1$ for which the determinations satisfy*

$$t_0^\delta - 2(1-t_0)^\delta = -1, \quad t_0^\delta + 2(1+t_0)^\delta = 1. \tag{B.1}$$

We will need the following "two-circle" results, in which the adjectives refer to their mode of intersection.

Lemma B.2. *For complex z, w with $|z| = |w| = 1$ we have*

$$|z^2 - z + 1| \leq 3|z + w - 1|.$$

Proof. With $z = e^{i\theta}, w = e^{i\phi}$ this is equivalent (after squaring) to $F \geq 0$ $(0 \leq \theta, \phi \leq 2\pi)$ with

$$F = F(\theta, \phi) = 12 - 7\cos\theta - 9\cos\phi + 9\cos(\theta - \phi) - \cos 2\theta.$$

We have

$$\frac{\partial F}{\partial \phi} = 9(\sin\phi + \sin(\theta - \phi))$$

which vanishes in the interior only for $\theta \equiv \pi + 2\phi$ modulo $2\pi\mathbb{Z}$. At these points $F = G(\cos\phi)$ with

$$G(x) = 2(2 + x)(1 - x)(1 - 2x)^2 \geq 0$$

for $-1 \leq x \leq 1$. And on the boundary $F(\theta, 0) = F(\theta, 2\pi) = H(\cos\theta)$ with

$$H(x) = 2(1 + x)(2 - x) \geq 0$$

and $F(0, \phi) = 4$. This completes the proof. □

The constant 3 here is best possible, as the example $z = -1, w = 1$ shows.

Lemma B.3. *For complex z, w with $|z| = |w| = 1$ we have*

$$|z + 1| \leq 2|z + 2w - 1|^{1/2}.$$

Proof. With $\epsilon = |z + 2w - 1|$ we have

$$\epsilon \geq |2w| - |z - 1| = 2 - |z - 1|$$

so $|z - 1| \geq 2 - \epsilon$. If $\epsilon \leq 2$ then drawing a picture shows that

$$|z + 1| \leq \sqrt{4 - (2 - \epsilon)^2} = \sqrt{4\epsilon - \epsilon^2} \leq 2\sqrt{\epsilon}.$$

And if $\epsilon > 2$ then even

$$|z + 1| \leq 2 < \sqrt{2}\sqrt{\epsilon}$$

which completes the proof. □

The exponent $1/2$ comes from tangency. The multiplying constant 2 cannot be replaced by anything smaller, as the example

$$z = -1 + \frac{4\epsilon - \epsilon^2}{2} + i\frac{2 - \epsilon}{2}\sqrt{4\epsilon - \epsilon^2}, \quad w = 1 - \frac{\epsilon}{2} - i\frac{1}{2}\sqrt{4\epsilon - \epsilon^2}$$

with

$$|z + 1| = \sqrt{4 - \epsilon}|z + 2w - 1|^{1/2}$$

and $\epsilon \to 0$ shows.

Proof of Proposition B.1. To keep better track of the determinations and their Galois conjugates, we write $\delta = r/q$ for integers $q > 0, r \neq 0$ and consider the set Y_{rq} in affine $\mathbf{A}^4$ defined by the equations

$$y_1^q = y^r, \quad y_2^{2q} = (1 - y)^{2r}, \quad y_3^{2q} = (1 + y)^{2r}$$

$$y_1 - 2y_2 = -1, \qquad y_1 + 2y_3 = 1 \tag{B.2}$$

(and $y \neq 0, 1, -1$). Note that this is a finite set in $\overline{\mathbb{Q}}^4$.

One checks that there is an involution on Y_{rq} defined by sending (y, y_1, y_2, y_3) to

$$\left(\frac{1}{y}, \frac{1}{y_1}, \frac{y_2}{y_1}, -\frac{y_3}{y_1}\right). \tag{B.3}$$

Now a solution of (B.1) leads to a point

$$(y, y_1, y_2, y_3) = (t_0, t_0^\delta, (1-t_0)^\delta, (1+t_0)^\delta)$$

on Y_{rq}. So it suffices to show that Y_{rq} is empty whenever

$$0 < |\delta| \leq \ 10^{-330}. \tag{B.4}$$

Any element σ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on points $(\eta, \eta_1, \eta_2, \eta_3)$ of Y_{rq}.

Case 0. *For some σ we have*

$$|\sigma(\eta)| \geq \frac{1}{2}, \quad |1-\sigma(\eta)| \geq \frac{1}{2}, \quad |1+\sigma(\eta)| \geq \frac{1}{2}, \quad |\sigma(\eta)| \leq 2.$$

We note that for any z with $1/2 \leq |z| \leq 3$ we have

$$\left||z|^\delta - 1\right| = \left|\sum_{n=1}^\infty \frac{\delta^n (\log|z|)^n}{n!}\right| \leq |\delta| \sum_{n=1}^\infty \frac{|\log|z||^n}{n!} = |\delta|(\exp(|\log|z||)-1) \leq 2|\delta|.$$

Thus with $z = \sigma(\eta), 1-\sigma(\eta), 1+\sigma(\eta)$ and corresponding $s_1 = \sigma(\eta_1), s_2 = \sigma(\eta_2), s_3 = \sigma(\eta_3)$, taking conjugates in (B.2) gives

$$s_1 - 2s_2 = -1, \qquad s_1 + 2s_3 = 1$$

and

$$||s_1|-1| \leq 2|\delta|, \quad ||s_2|-1| \leq 2|\delta|, \quad ||s_3|-1| \leq 2|\delta|$$

(note that $s_1^q = \sigma(\eta)^r$ so $|s_1| = |\sigma(\eta)|^\delta$, and similarly for $|s_2|, |s_3|$). As

$$\left|\frac{w}{|w|} - w\right| = ||w|-1| \qquad (w \neq 0)$$

we get for $s_1' = s_1/|s_1|, s_2' = s_2/|s_2|, s_3' = s_3/|s_3|$ on the unit circle the inequalities

$$|s_1' - 2s_2' + 1| \leq 6|\delta|, \qquad |s_1' + 2s_3' - 1| \leq 6|\delta|$$

(note that $s_1' - 2s_2' + 1 = (s_1' - s_1) - 2(s_2' - s_2)$ and so on). By Lemma B.3 with $z = -s_1', w = s_2'$ we get

$$|s_1' - 1| \leq 2(6|\delta|)^{1/2};$$

and then with $z = s_1', w = s_3'$

$$|s_1' + 1| \leq 2(6|\delta|)^{1/2}.$$

These contradict each other provided $|\delta| < 1/24$, certainly guaranteed by (B.4).

Thus for every σ there are four possibilities coming from the failure of Case 0, and we consider each in turn.

`Case 1a`. *For some σ we have*

$$|1-\sigma(\eta)| < \frac{1}{2}.$$

We write $\sigma(\eta) = 1-u$ so that $|u| < 1/2$. Now

$$\sigma(\eta_1)^q = \sigma(\eta_1^q) = \sigma(\eta^r) = \sigma(\eta)^r = (1-u)^r$$

so that $\sigma(\eta_1)$ is a determination of $(1-u)^\delta$. There is a canonical such determination $b_1 = 1-\delta v$ for

$$v = u - \frac{\delta-1}{2}u^2 + \frac{\delta-1}{2}\frac{\delta-2}{3}u^3 - + \cdots ,$$

and because

$$\left|\frac{\delta-k}{k+1}\right| \le 1 \quad (k=1,2,\ldots)$$

we deduce

$$|b_1 - 1| \le |\delta|\frac{|u|}{1-|u|} \le |\delta|. \tag{B.5}$$

Thus $\sigma(\eta_1) = b_1\zeta_1$ for a root of unity ζ_1.

(Almost) similarly $\sigma(\eta_3)^{2q} = (2-u)^{2r}$ and so there is a root of unity ζ such that $\zeta\sigma(\eta_3)$ is a determination of $(2-u)^\delta$, and there is a canonical b_3 with

$$|b_3 - 2^\delta| \le 2^\delta|\delta| \tag{B.6}$$

(with of course the natural determination of 2^δ). So $\sigma(\eta_3) = b_3\zeta_3$ for a root of unity ζ_3.

The second equation in (B.2) now leads to

$$b_1\zeta_1 + 2b_3\zeta_3 = 1, \tag{B.7}$$

so that by (B.5) and (B.6) we get first $|\zeta_1 + 2^{1+\delta}\zeta_3 - 1| \le (2^{1+\delta}+1)|\delta|$ and then

$$|\zeta_1 + 2\zeta_3 - 1| \le |2^\delta - 1| + (2^{1+\delta}+1)|\delta| \le 6|\delta|. \tag{B.8}$$

Now Lemma B.3 gives

$$|\zeta_1 + 1| \le 2(6|\delta|)^{1/2}. \tag{B.9}$$

The first equation in (B.2) gives $2\sigma(\eta_2) = 1 + b_1\zeta_1 = (b_1-1)\zeta_1 + (\zeta_1+1)$ so we deduce

$$|\sigma(\eta_2)| \le \frac{1}{2}(|\delta| + 2(6|\delta|)^{1/2}) \le 5|\delta|^{1/2} < 1 \tag{B.10}$$

by (B.5) and (B.9). Similarly $\sigma(\eta_1) + 1 = 1 + b_1\zeta_1$ leads to

$$|\sigma(\eta_1) + 1| \le 10|\delta|^{1/2}. \tag{B.11}$$

We may say that these $\sigma(\eta_1)$ cluster near -1. Also

$$|\sigma(\eta_1) - 1| \le 2 + 10|\delta|^{1/2} < 3. \tag{B.12}$$

Now $\eta_2^{2q} = (1-\eta)^{2r}$ gives $|\sigma(\eta_2)| = |1-\sigma(\eta)|^\delta$. Thus

$$\delta = \frac{\log|\sigma(\eta_2)|}{\log|1-\sigma(\eta)|} > 0 \tag{B.13}$$

by(B.10) and the basic assumption in this Case 1a.

We leave this case unresolved for the moment, and proceed straight to

`Case 1b`. *For some σ we have*

$$|1+\sigma(\eta)| < \frac{1}{2}.$$

Here the same arguments with $-b_1\zeta_1 + 2b_2\zeta_2 = 1$ instead of (B.7) lead to

$$|\sigma(\eta_1) - 1| \le 10|\delta|^{1/2} \tag{B.14}$$

in place of (B.11), so clustering at 1, but also

$$|\sigma(\eta_1) + 1| < 3 \tag{B.15}$$

as in (B.12), as well as

$$\delta > 0 \tag{B.16}$$

as in (B.13). Again we jump to the next (somewhat critical)

`Case 2a`. *For some σ we have*

$$|\sigma(\eta)| < \frac{1}{2}.$$

Now subtracting the two equations in (B.2) and using $\sigma(\eta_2) = b_2\zeta_2$ with $|b_2 - 1| \le |\delta|$ and $\sigma(\eta_3) = b_3\zeta_3$ with $|b_3 - 1| \le |\delta|$ leads to

$$|\zeta_2 + \zeta_3 - 1| \le 2|\delta|$$

as in (B.8). By Lemma B.2 we deduce $|\zeta_2^2 - \zeta_2 + 1| \le 6|\delta|$ and then using $|\sigma(\eta_2) - \zeta_2| \le |\delta|$ also $|\sigma(\eta_2)^2 - \sigma(\eta_2) + 1| \le 12|\delta|$ and, multiplying by 4,

$$|\sigma(\eta_1)^2 + 3| \le 48|\delta|. \tag{B.17}$$

Here we may say that these $\sigma(\eta_1)$ cluster near $\pm\sqrt{-3}$.

In particular

$$|\sigma(\eta_1)|^2 \ge 3 - 48|\delta| > 1 \tag{B.18}$$

and also

$$|\sigma(\eta_1)^{-2} + 3| < 4. \tag{B.19}$$

Now we get

$$\delta = \frac{\log|\sigma(\eta_1)|}{\log|\sigma(\eta)|} < 0. \tag{B.20}$$

Next we claim that η_1 is a unit. Write $p = -r > 0$ so that $\delta = -p/q$. Pick any τ with $\tau^p = \eta_1^{-1}$. Then $\tau^{pq} = \eta_1^{-q} = \eta^p$ so $\eta = \zeta\tau^q$ for a root of unity ζ. Now $2\eta_2 = \eta_1 + 1$ so

$$2^{2q}(1-\zeta\tau^q)^{-2p} = 2^{2q}(1-\eta)^{-2p} = (2\eta_2)^{2q} = (\eta_1+1)^{2q} = (\tau^{-p}+1)^{2q},$$

or as a polynomial in τ

$$(1-\zeta\tau^q)^{2p}(1+\tau^p)^{2q} = 2^{2q}\tau^{2pq}.$$

Comparing highest and lowest coefficients we see that τ is a unit and therefore also $\eta_1 = \tau^{-p}$ as claimed (here it is crucial that $\delta < 0$ and it is highly unlikely that it works when $\delta > 0$).

Before proceeding further we jump to the final

Case 2b. *For some σ we have*

$$|\sigma(\eta)| > 2.$$

This is $|\sigma(\eta^{-1})| < 1/2$, and according to (B.3) the point $(\eta^{-1}, \eta_1^{-1}, \eta_2\eta_1^{-1}, -\eta_3\eta_1^{-1})$ lies in Y_{rq}. By Case 2a we get

$$|\sigma(\eta_1^{-1})^2 + 3| \le 48|\delta| \tag{B.21}$$

in place of (B.17), so clustering near $\pm 1/\sqrt{-3}$, as well as

$$|\sigma(\eta_1)|^{-2} \ge 3 - 48|\delta| > 1$$

as in (B.18), and also

$$|\sigma(\eta_1)^2 + 3| < 4 \tag{B.22}$$

as in (B.19). Then

$$\delta < 0 \tag{B.23}$$

as in (B.20). Also η_1^{-1}, so η_1 too, is a unit.

Now let us sum up. As mentioned, we can ignore Case 0, so that every σ falls into the other cases.

Suppose first that $\delta < 0$.

Then Cases 1a and 1b are impossible by (B.13) and (B.16). We consider the number

$$\eta' = (\eta_1^2 + 3)(\eta_1^{-2} + 3)$$

which is an algebraic integer.

In Case 2a we use (B.17) and (B.19) to see that

$$|\sigma(\eta')| < 192|\delta|.$$

And in Case 2b we use (B.21) and (B.22) to get the same inequality.

Thus we see that η' has norm of absolute value at most $(192|\delta|)^{d'}$ for its degree d'. So as soon as $|\delta| < 1/192$ we deduce $\eta' = 0$.

But this would imply that $\eta_1 = i\sqrt{3}$ (say), so $\eta_2 = (1 + \eta_1)/2$ and $\eta_3 = (1 - \eta_1)/2$ are roots of unity so also $1 - \eta, 1 + \eta$ which is impossible as $\eta \neq 0$; and a similar argument works with say $\eta_1 = 1/(i\sqrt{3})$ using $\eta^{-1} \neq 0$. This settles things when $\delta < 0$.

It remains to deal with $\delta > 0$.

Then Cases 2a and 2b are impossible by (B.20) and (B.23). But now we no longer know that η_1 is a unit (and probably it need not be), so we cannot use this method.

Instead we use Theorem A.1 on the equation $\eta_1 - 2\eta_2 = -1$. We note that $\eta_1^q = \eta^r$ so η_1 is a determination of η^δ. Also $\eta_2^{2q} = (1-\eta)^{2r}$ so there is a root of unity ζ such that η_2/ζ is a determination of $(1-\eta)^\delta$. Now

$$\alpha\eta_1 + \beta(\eta_2/\zeta) = 1$$

for $\alpha = -1, \beta = 2\zeta$. It follows from Theorem A.1 with $\lambda = \delta > 0$ that

$$h(\eta) \leq 100\delta^{-1} + 121\delta^{-1}\log 2$$

(as long as we don't have $\alpha\eta^{\delta-1} = 1$, in which case $h(\eta) = 0$ anyway). Therefore

$$h(\eta_1) = |\delta| h(\eta) \leq 100 + 121\log 2.$$

Now we consider $\eta' = \eta_1^2 - 1$, so that

$$h(\eta') \leq 200 + 243\log 2. \tag{B.24}$$

In Case 1a we use (B.11) and (B.12) to get

$$|\sigma(\eta')| < 30|\delta|^{1/2}.$$

And in Case 1b the same using (B.14) and (B.15).

If $\eta' \neq 0$ the Product Formula gives $1 < (30|\delta|^{1/2})^{d'} e^{d' h(\eta')}$ which by (B.24) is at most $(30|\delta|^{1/2} e^{200} 2^{243})^{d'}$. So as soon as

$$|\delta| \leq \frac{1}{30^2 e^{400} 2^{486}} = 10^{-322.972613\ldots}$$

(accounting for (B.4) above) we conclude $\eta' = 0$, now easily seen to be impossible. This settles things when $\delta > 0$, thereby completing the proof of the Proposition.

$\square$

References

[1] F. Amoroso, "On a conjecture of G. Rémond", *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, **15** (2016) 599–608.

[2] F. Amoroso, D. Masser and U. Zannier, "Bounded Height Problems and Silverman Specialization Theorem" , *Duke Math J.*, **166**, no. 13 (2017), 2599–2642.

[3] A. Bérczes, J.-H. Evertse and K. Györy, "Effective results for linear equations in two unknowns from a multiplicative division group", Acta Arithmetica **136** (2009), 331–349.

[4] F. Beukers, "On a sequence of polynomials." In *Algorithms for algebra. J. Pure Appl. Algebra* **117-118** (1997), 97–103.

[5] F. Beukers and F. Schlickewei, "The equation $x+y=1$ in finitely generated groups". *Acta Arithmetica* **78** (1996), no. 2, 189–199.

[6] E. Bombieri, D. Masser, and U. Zannier, "Intersecting a curve with algebraic subgroups of multiplicative groups", *Internat. Math. Res. Notices* **1999**, no. 20, 1119–1140.

[7] E. Bombieri, D. Masser, and U. Zannier, "Finiteness results for multiplicatively dependent points on complex curves". *Michigan Math. J.* **51** (2003), 451–466.

[8] W. D. Brownawell and D. W. Masser, "Vanishing sums in function fields". *Math. Proc. Cambridge Philos. Soc.* **100** (1986), no. 3, 427–434.

[9] A. Denz, "Bounding the height of certain algebraic numbers", Master Thesis, University of Basel 2016.

[10] G. Faltings, "The general case of S. Lang's conjecture". *Barsotti Symposium in Algebraic Geometry (Abano Terme, 1991)*, 175–182, Perspect. Math., **15**, Academic Press, San Diego, CA, 1994.

[11] P. Habegger, "Intersecting subvarieties of $\mathbb{G}_{\mathrm{m}}^n$ with algebraic subgroups." *Math. Ann.* **342** no. 2, 449–466 (2008).

[12] M. Hindry, "Autour d'une conjecture de S. Lang". *Invent. Math.*, **94**, 575-603 (1988).

[13] M. Mignotte, T. N. Shorey and R. Tijdeman, "The distance between terms of an algebraic recurrence sequence". *J. Reine Angew. Math.* **349**, 63–76 (1984).

[14] P. Philippon, "Sur des hauteurs alternatives. III." *J. Math. Pures Appl.(9)* **74** no. 4, 345–365 (1995).

[15] L. Pottmeyer, "Fields Generated by Finite Rank Subgroups of Tori and Elliptic Curves". *Int. J. Number Theory* **17**, no. 5, 1079–1089 (2021)

[16] D. Roy and J. Thunder, "An absolute Siegel's lemma." *J. Reine Angew. Math.* **476** (1996), 1–26. Addendum and erratum: *J. Reine Angew. Math.* **508**, 47–51 (1999).

[17] W. M. Schmidt, "Diophantine approximations and Diophantine equations". LNM 1467, Springer-Verlag, Berlin, 1991.

[18] W. M. Schmidt. "Heights of points on subvarieties of $\mathbb{G}_{\mathrm{m}}^n$". In"Number Theory 93-94", S. David editor, London Math. Soc. Ser., volume **235**, Cambridge University Press, 1996.

[19] J.H. Silverman, "Heights and the specialization map for families of abelian varieties", J. reine angew. Math. **342**, 197–211 (1983).

[20] M. Waldschmidt, "Diophantine approximation on linear algebraic groups. Transcendence properties of the exponential function in several variables." Grundlehren der mathematischen Wissenschaften, 326. Springer-Verlag, Berlin, 2000.

[21] O. Warin, "On $x + y + z + w = 1$ and heights". Master Thesis, University of Basel 2012.

[22] D. Zagier, "Algebraic numbers close to both 0 and 1", *Math. Computation* **61**, 485–491 (1993).

[23] S. Zhang, "Positive line bundles on arithmetic surfaces", *Annals of Math.* **136**, 569–587 (1992).