TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 371, Number 1, 1 January 2019, Pages 357–386
http://dx.doi.org/10.1090/tran/7238
Article electronically published on April 25, 2018

DISTRIBUTION OF INTEGRAL DIVISION POINTS ON THE ALGEBRAIC TORUS

PHILIPP HABEGGER AND SU-ION IH

Abstract. Let K be a number field with algebraic closure $\overline{K}$, and let S be a finite set of places of K containing all the infinite ones. Let Γ_0 be a finitely generated subgroup of $\mathbb{G}_{\mathrm{m}}(\overline{K})$, and let $\Gamma \subset \mathbb{G}_{\mathrm{m}}(\overline{K})$ be the division group attached to Γ_0. Here is an illustration of what we will prove in this article. Fix a proper closed subinterval I of $[0,\infty)$ and a nonzero effective divisor D on $\mathbb{G}_{\mathrm{m}}$ which is not the translate of any torsion divisor on the algebraic torus $\mathbb{G}_{\mathrm{m}}$ by any point of Γ with height belonging to I.

Then we prove a statement which easily implies that the set of "integral division points on $\mathbb{G}_{\mathrm{m}}$ with height near Γ", i.e., the set of points of Γ with (standard absolute logarithmic Weil) height in J which are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D, is finite for some fixed subinterval J of $[0,\infty)$ properly containing I. We propose a conjecture on the nondensity of integral division points on semi-abelian varieties with prescribed height values, which generalizes some previously known conjectures as well as this finiteness result for $\mathbb{G}_{\mathrm{m}}$. Finally, we also propose an analogous version for a dynamical system on $\mathbb{P}^1$.

1. Introduction

Let K be a number field with algebraic closure $\overline{K}$. Let $\overline{K}_v$ be an algebraic closure of a completion K_v of K with respect to a place v of K. We assume $K \subset \overline{K}_v$ for convenience. A ring homomorphism from a field extension of K is called a K-*embedding* if its restriction to K is the identity. Suppose $\alpha, \gamma \in \mathbb{G}_{\mathrm{m}}(\overline{K}) = \overline{K}^{\times}$, where for simplicity $\mathbb{G}_{\mathrm{m}}$ denotes the algebraic torus over K. The Galois conjugates of α over K determine a divisor (α) on $\mathbb{G}_{\mathrm{m}}$. Let S be a finite set of places of K including all the infinite ones. We say that γ is S-*integral* on $\mathbb{G}_{\mathrm{m}}$ relative to the divisor (α) if, for all places $v \notin S$ of K and all K-embeddings $\sigma : K(\gamma) \to \overline{K}_v$ and $\tau : K(\alpha) \to \overline{K}_v$, we have

$$|\sigma(\gamma)|_v = 1 \quad \text{and} \quad |\sigma(\gamma) - \tau(\alpha)|_v = \max\{1, |\tau(\alpha)|_v\}. \tag{1}$$

Received by the editors August 22, 2015, and, in revised form, October 7, 2016, and February 20, 2017.

2010 *Mathematics Subject Classification.* Primary 11G50, 11J61, 11J71, 11J86, 11L15, 14G25, 14G40, 20G30, 37P05, 37P35.

Key words and phrases. Arithmetical dynamical system, canonical height, division group, division point, Erdős–Turán theorem, integral point, Koksma's inequality, linear forms in logarithms, logarithmic equidistribution, multiplicative group, Weyl sums.

The work of the first author was partially supported by the National Science Foundation, grant number DMS-1128155.

The work of the second author was partially supported by the Simons Foundation, grant number 267613.

The opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of those foundations.

Under the assumption of the first condition, the second one is equivalent to asking for $|\sigma(\gamma) - \tau(\alpha)|_v \geq 1$ if $|\tau(\alpha)|_v \leq 1$. In most of the problems concerning integral points we are interested in, we will be allowed to extend the ground field without loss of generality. If K is assumed to contain α, we can suppress τ in all the places above and get a simpler expression. We will keep all this notation throughout unless otherwise stated and refer the reader to section 2 for additional details.

The above definition of an S-integral point on $\mathbb{G}_{\mathrm{m}}$ relative to (α) extends to an arbitrary effective divisor D on $\mathbb{G}_{\mathrm{m}}$ in a natural way. The point γ is said to be *S-integral* on $\mathbb{G}_{\mathrm{m}}$ relative to D if it is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α) for all α in the support of D. Further down, in section 6.5, we will define S-integral points on the projective line $\mathbb{P}^1$ relative to an effective divisor on $\mathbb{P}^1$; see [10, pp. 2011–2012] for a unifying intrinsic definition of integral points on $\mathbb{P}^1$ and on $\mathbb{G}_{\mathrm{m}}$. We retain this geometric notation throughout the first section.

Let Γ_0 be a finitely generated subgroup of $\mathbb{G}_{\mathrm{m}}(\overline{K})$, and let $\Gamma \subset \mathbb{G}_{\mathrm{m}}(\overline{K})$ be the *division group* (over, above, or) attached to Γ_0 in $\mathbb{G}_{\mathrm{m}}(\overline{K})$, i.e., $\Gamma := \{\xi \in \mathbb{G}_{\mathrm{m}}(\overline{K}) : \xi^n \in \Gamma_0 \text{ for some integer } n \geq 1\}$.

Suppose that D is a nonzero effective divisor on $\mathbb{G}_{\mathrm{m}}$. We call D a *torsion divisor* if its support consists of roots of unity and the *translate* αD of D by an element α of $\mathbb{G}_{\mathrm{m}}(\overline{K})$ is the divisor obtained from D by multiplying each irreducible component of D by α.

Let h denote the standard absolute logarithmic Weil height on $\mathbb{G}_{\mathrm{m}}(\overline{K})$; see [6, Section 1.5], [13, Chapter B], and section 2 for further details.

The main result of this paper is the following theorem.

Theorem 1.1. (*Keep all the notation K, S, Γ, and h as above.*) *Suppose that D is a nonzero effective divisor on $\mathbb{G}_{\mathrm{m}}$.*

(i) *Let $(I_n)_{n\geq 1}$ be a decreasing sequence of closed subintervals of $[0,\infty)$. If*

$$\{\gamma \in \Gamma : \gamma \text{ is } S\text{-integral on } \mathbb{G}_{\mathrm{m}} \text{ relative to } D \text{ and } h(\gamma) \in I_n\}$$

is infinite for all integers $n \geq 1$, then D is the translate of a torsion divisor on $\mathbb{G}_{\mathrm{m}}$ by an element of

$$\{\alpha \in \Gamma : \alpha \text{ is } S\text{-integral on } \mathbb{G}_{\mathrm{m}} \text{ relative to } D \text{ and } h(\alpha) \in I\},$$

where $I = \bigcap_{n\geq 1} I_n$.

(ii) *If $(\gamma_n)_{n\geq 1}$ is a sequence of pairwise distinct elements of Γ that are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D, then D is the translate of a torsion divisor on $\mathbb{G}_{\mathrm{m}}$ by some $\alpha \in \Gamma$ that is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D with $\lim_{n\to\infty} h(\gamma_n \alpha^{-1}) = 0$.*

Let us spell out the detailed meaning of the conclusion of part (i). It means that the divisor D on $\mathbb{G}_{\mathrm{m}}$ is of the form $\sum_{j=1}^{r} a_j(\alpha\zeta_j)$ for some integers $r \geq 1, a_1 \geq 1, \ldots, a_r \geq 1$ and some roots of unity $\zeta_1, \ldots, \zeta_r$, where $\alpha \in \Gamma$ is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D and $h(\alpha) \in I$.

We thank Shou-Wu Zhang for informing us of the fact that Theorem 1.1(i) is equivalent to its apparently weaker version with the stronger assumption that $I_1 = I_2 = \cdots = I$. Indeed, the conclusion of the apparently weaker version of Theorem 1.1(i) implies that D must be of the form $\sum_{j=1}^{r} a_j(\alpha\zeta_j)$ as above with $h(\alpha) \in I_n$ for all integers $n \geq 1$. Then it follows that $h(\alpha) \in I$, as desired.

We note that Theorem 1.1 refines two statements in the literature. If $\Gamma_0 = \{1\}$ or $I_1 = I_2 = \cdots = \{0\}$, we recover [4, Theorem 0.1] by M. Baker, Ih, and Rumely that

is confined to the roots of unity. Taking $I_1 = I_2 = \cdots = [0, \infty)$ yields Grant and Ih's [10, Theorem 2.5]. Note, however, that this reference excludes the interesting case $D = (\alpha)$ with $\alpha \in \Gamma$. This case leads to new difficulties, such as the logarithmic singularity to be mentioned in section 3 in connection with equidistribution, but also to interesting features, as laid out in Theorem 1.3 below. See Theorem 6.3 below for a strengthening of Theorem 1.1(i).

In the formulation of Theorem 1.1, (ii) implies (i), and this is how we will proceed to prove Theorem 1.1(i) in this article; cf. see also Theorem 1.3(ii) and Theorem 1.6 with its preceding and subsequent comments. Indeed, it will be an easy exercise to see that (i) $\Rightarrow$ (ii) is also true, however.

Example 1.2. (We thank David Grant for discussing the following example that concerns Theorem 1.1(i).) Let $K = \mathbb{Q}$, $S = \{2, \infty\}$, and let μ_∞ be the set of all roots of unity in $\overline{\mathbb{Q}}$. Take Γ_0 to be the cyclic subgroup of $\mathbb{G}_{\mathrm{m}}(\overline{\mathbb{Q}})$ generated by 2. So we have $\Gamma = \{2^a\zeta \in \mathbb{G}_{\mathrm{m}}(\overline{\mathbb{Q}}) : a \in \mathbb{Q} \text{ and } \zeta \in \mu_\infty\}$. Let $D = (2)$, the translate of the torsion divisor (1) by $2 \in \Gamma$.

There are infinitely many S-integral points on $\mathbb{G}_{\mathrm{m}}$ relative to D with height in $I = \{\log 2\}$. Indeed, let ζ_N be a primitive root of unity of order N, where N is a positive integer divisible by at least two distinct prime numbers. Then all the points $2\zeta_N$ are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D, and it is obvious that they belong to Γ with height equal to $\log 2$. It is also possible to show that the points $\gamma_n = 2^{(n+1)/n} \in \Gamma$ are all S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D with height in $I_n = [\log 2, (1 + 1/n) \log 2]$ for all $n \geq 1$. Notice that $\bigcap_{n\geq 1} I_n = I$ and that all the points γ_n have distinct heights in contrast to $2\zeta_N$.

We will prove several results on the way to proving Theorem 1.1.

First, we will deal with the important special case $D = (\alpha)$ where α may possibly lie in Γ.

Theorem 1.3. (*Keep all the notation K, S, Γ, and h as above.*) *Let $\alpha \in \mathbb{G}_{\mathrm{m}}(\overline{K})$. Suppose $(\gamma_n)_{n\geq 1}$ is a sequence of pairwise distinct elements of Γ that are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α). Then the following statements hold true.*

(i) *There exists $B \in \mathbb{R}$ with $h(\gamma_n) \leq B$ for all n.*

(ii) *More precisely, we have*

$$\lim_{n\to\infty} h(\gamma_n\alpha^{-1}) = 0 \quad \textit{and} \quad \lim_{n\to\infty} h(\gamma_n) = h(\alpha).$$

The latter equality of part (ii) is an immediate consequence of the former one through an easy application of the triangle inequality of the Weil height with respect to multiplication, i.e., $|h(x) - h(y)| \leq h(xy^{-1}) \leq h(x) + h(y)$ for all $x, y \in \mathbb{G}_{\mathrm{m}}(\overline{K})$.

As a first byproduct of Theorem 1.3 we have the following corollary.

Corollary 1.4. (*Keep the notation as in Theorem* 1.3.) *If $\alpha \in \mathbb{G}_{\mathrm{m}}(\overline{K}) - \mu_\infty$, then*

$$\{\gamma \in \Gamma : \gamma \textit{ is } S\textit{-integral on } \mathbb{G}_{\mathrm{m}} \textit{ relative to } (\alpha) \textit{ and } h(\gamma) \leq \epsilon\} \tag{2}$$

is finite for some real number $\epsilon > 0$.

Its proof enables us to take ϵ to be any positive real number strictly less than $h(\alpha)$. We note that α may belong to Γ here.

We provide a second corollary which is a byproduct of Theorem 1.3. We see from its second part that any α in Theorem 1.3 must actually lie in Γ.

Corollary 1.5. (*Keep the notation as in Theorem* 1.3.)

(i) (The "two-point case") *If* $\alpha, \beta \in \mathbb{G}_{\mathrm{m}}(\overline{K})$ *satisfy* $\alpha\beta^{-1} \notin \mu_\infty$*, then the set of elements in* Γ *that are* S*-integral on* $\mathbb{G}_{\mathrm{m}}$ *relative to* $(\alpha)+(\beta)$ *is finite.*

(ii) (The "one-point case") *If* $\alpha \in \mathbb{G}_{\mathrm{m}}(\overline{K}) - \Gamma$*, then the set of elements in* Γ *that are* S*-integral on* $\mathbb{G}_{\mathrm{m}}$ *relative to* (α) *is finite.*

Notice that this corollary is equivalent to the full strength of [10, Theorem 2.5], as was observed in [10, Example 1.2].

We take two steps to complete the proof of Theorem 1.1. We will first prove the following weaker version whose proof relies heavily on Theorem 1.3.

Theorem 1.6. (*Keep the notation as in Theorem* 1.1.) *If* $(\gamma_n)_{n\geq 1}$ *is a sequence of pairwise distinct elements of* Γ *that are* S*-integral on* $\mathbb{G}_{\mathrm{m}}$ *relative to* D*, then* D *is the translate of a torsion divisor on* $\mathbb{G}_{\mathrm{m}}$ *by some* $\alpha \in \Gamma$ *with* $\lim_{n\to\infty} h(\gamma_n\alpha^{-1}) = 0$.

In order to conclude Theorem 1.1(ii), say, it suffices to strengthen the conclusion by replacing "by some $\alpha \in \Gamma$" with "by some $\alpha \in \Gamma$ that is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D".

We emphasize that our proof of Theorem 1.1 does not rely on any results in [10]. Indeed, we proceed by a rather different independent method. Our proof strategy is closer to the work [4] of M. Baker, Ih, and Rumely. We will deduce a result, Proposition 3.1, of equidistribution type for elements in Γ which, in combination with the product formula, will eventually yield Theorem 1.3, its corollaries, and the other finiteness theorems. A large portion of this paper is dedicated to the proof of this proposition.

The results in [4] deal with the division group of rank 0, i.e., the group of all roots of unity. Working with division groups of positive rank introduces additional difficulties. A new feature in the present work is the use of an explicit version of Weyl's equidistribution criterion due to Erdős and Turán. In order to apply the Erdős–Turán theorem we will need to bound certain Weyl sums from above. As in [4] we need manageable lower bounds for linear forms in logarithms in the absolute value, for which we need a variant of the result originally obtained by A. Baker. But in contrast to [4], we also require an analogous result due to Yu in the nonarchimedean case. We also use a result of Schlickewei and Schmidt coming from the geometry of numbers that allows us to find a tuple of independent elements of Γ_0 that behave nicely with respect to the height.

Because it fits in well with the overall theme of this paper, we include in the final section a conjecture generalizing Theorem 1.1(i) to an arbitrary semi-abelian variety and an analogous conjecture on a dynamical system.

Finally, we hope that a suitably adapted method will give a corresponding result for elliptic curves.

2. Basic common notation

For the reader's convenience we collect some notation in this section.

The set of positive integers is denoted by $\mathbb{N} = \{1, 2, 3, \dots\}$. By $\varphi(\cdot)$ we mean Euler's totient function. The largest integer not exceeding a real number x is denoted by $\lfloor x \rfloor$. The cardinality of a set M is denoted by $\#M$. The group of units of a ring R is denoted by $R^\times$.

Through this article h denotes the *standard* (*or classical*) *absolute logarithmic Weil height* whose basic properties are found in, e.g., [6, Section 1.5] and [13,

Chapter B]. Let K be a number field, and let $\overline{K}$ be a fixed algebraic closure of K. Similar notation applies to any field in what follows. The group of all roots of unity in $\overline{K}^\times$ is denoted by μ_∞.

A *place* or *prime* v of K is an absolute value $|\cdot|_v$ that extends either the usual complex absolute value on $\mathbb{Q}$ or a p-adic absolute value on $\mathbb{Q}$ for some prime number p with $|p|_v = p^{-1}$. In the former case v is said to be *infinite* and we write $v \mid \infty$, while in the remaining cases v is said to be *finite* and we write $v \nmid \infty$. We denote by M_K the set of all places of K. If L is a finite extension of K and $w \in M_L$ restricts to $v \in M_K$, then we say that w *lies over* or *above* v and write $w \mid v$.

A completion of K with respect to v is denoted by K_v. By abuse of notation we sometimes write K_w instead of $K_{w|_K}$ for any $w \in M_L$, where L is a finite extension of K. The absolute value $|\cdot|_v$ extends uniquely to K_v. We will fix an algebraic closure $\overline{K}_v$ of K_v. For simplicity we suppose $K \subset K_v \subset \overline{K}_v$ and extend $|\cdot|_v$ further to $\overline{K}_v$. If L is any field extension of K, then

$$\mathrm{Hom}_K(L, \overline{K}_v) \;=\; \big\{\sigma : \sigma \text{ is a field embedding } L \to \overline{K}_v \text{ such that } \sigma|_K = \mathrm{id}_K\big\},$$

is the set of all K-*embeddings* of L into $\overline{K}_v$.

We often use Vinogradov's notation $f \ll g$ (resp. $f \gg g$) for real-valued functions f and g, which means that there exists a constant $c > 0$ such that $f \le cg$ (resp. $f \ge cg$) on the specified domain. We make an effort to indicate the dependency of the constant c on various involved data in case of any danger of confusion.

3. Logarithmic equidistribution in division groups

Proposition 3.1 below can be regarded as a result of *logarithmic equidistribution type* for elements in a division group in $\mathbb{G}_{\mathrm{m}}(\overline{K}) = \overline{K}^\times$ at a given place of K. It generalizes the result stated in [4, Section 1] that holds when the division group is the group of roots of unity. Our statements are valid for the finite and infinite places of K.

Let K be a number field, and let Γ_0 be a finitely generated subgroup of $\overline{K}^\times$. For the proof of Theorem 1.1, which is the main goal of this article, we may assume without loss of generality that $\Gamma_0 \subset K^\times$, hence we will do so below for convenience. We write $\Gamma \subset \overline{K}^\times$ for the division group attached to Γ_0.

Let v be a place of K. Let $\gamma \in \Gamma$ and $\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)$. As some positive integer power of γ lies in K, we see that $|\sigma(\gamma)|_v$ does not depend on σ. Thus for convenience we sometimes abuse notation and simply write $|\gamma|_v$ for $|\sigma(\gamma)|_v$ or the other way around when doing so is more suggestive.

The main purpose of the next few sections is to prove the following proposition.

Proposition 3.1. (*Keep the notation K, v, and Γ as above.*) *Let $\alpha \in K^\times$. If $\gamma \in \Gamma$ satisfies $\gamma \neq \alpha$, then*

$$\left| \frac{1}{D} \left(\sum_{\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)} \log |\sigma(\gamma) - \alpha|_v \right) - \log \max\{|\gamma|_v, |\alpha|_v\} \right| \ll \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}, \tag{3}$$

where $D = [K(\gamma) : K]$ and the implied constant depends on K, v, Γ, and α, but not on γ.

We state the proposition in the special case $\alpha = 1$ separately.

Proposition 3.2. (*Keep the notation K, v, and Γ as above.*) *If $\gamma \in \Gamma$ satisfies $\gamma \neq 1$, then*

$$\left| \frac{1}{D} \left(\sum_{\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)} \log|\sigma(\gamma) - 1|_v \right) - \log \max\{|\gamma|_v, 1\} \right| \ll \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}, \tag{4}$$

where $D = [K(\gamma) : K]$ and the implied constant depends on K, v, and Γ, but not on γ.

The seemingly weaker Proposition 3.2 easily implies Proposition 3.1.

Proof that Proposition 3.2 *implies Proposition* 3.1. The subgroup of $K^\times$ generated by Γ_0 and $\alpha \in K^\times$ is finitely generated. Let Ω be its attached division group in $\overline{K}^\times$. We will apply Proposition 3.2 to Ω. Indeed, if $\gamma \in \Gamma$ with $\gamma \neq \alpha$, then $\gamma\alpha^{-1} \in \Omega$ is not 1. We set $D = [K(\gamma\alpha^{-1}) : K] = [K(\gamma) : K]$ and observe that

$$\frac{1}{D} \sum_{\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)} \log|\sigma(\gamma) - \alpha|_v \;=\; \log|\alpha|_v + \frac{1}{D} \sum_{\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)} \log|\sigma(\gamma\alpha^{-1}) - 1|_v$$

equals $\log|\alpha|_v + \log\max\{|\gamma\alpha^{-1}|_v, 1\} = \log\max\{|\gamma|_v, |\alpha|_v\}$ up to an error term as in (3) with γ replaced by $\gamma\alpha^{-1}$. But $h(\gamma\alpha^{-1}) \leq h(\gamma) + h(\alpha^{-1})$ by basic height properties, and we deduce the conclusion of Proposition 3.1. □

Before we start the proof of Proposition 3.2, we mention a corollary for points of small height that is in the spirit of Bilu's equidistribution theorem [5]. In the archimedean case the corollary does not seem to follow from Bilu's theorem due to a logarithmic singularity of the test function. Our corollary does not contradict Autissier's example [2] either, since the members of the sequence below are required to belong to the group Γ. More precisely, he considered roots of the irreducible polynomials $(X^n - 1)(X - 2) + 3$ $(n \geq 1)$ in his counterexample to a *logarithmic equidistribution property* there. But these roots do not belong to a finite rank group in contrast to our assumption below.

Corollary 3.3. (*Keep the notation K, v, and Γ as above.*) *Let $\alpha \in K^\times$. Suppose that $(\gamma_n)_{n \geq 1}$ is a nonrepeating sequence of elements of Γ different from α such that $\lim_{n\to\infty} h(\gamma_n) = 0$. Then*

$$\lim_{n\to\infty} \frac{1}{[K(\gamma_n) : K]} \sum_{\sigma \in \mathrm{Hom}_K(K(\gamma_n), \overline{K}_v)} \log|\sigma(\gamma_n) - \alpha|_v \;=\; \log\max\{1, |\alpha|_v\}.$$

Proof. For all $n \geq 1$, the comment stated above Proposition 3.1 easily implies

$$[K(\gamma_n) : K] \cdot \log|\gamma_n|_v \;\leq\; [K(\gamma_n) : \mathbb{Q}] \cdot h(\gamma_n)$$

and thus that $\log|\gamma_n|_v \leq [K : \mathbb{Q}] \cdot h(\gamma_n)$. In conjunction with this result, the same argument applied to γ_n^{-1} yields

$$\left|\log|\gamma_n|_v\right| \;\leq\; [K : \mathbb{Q}] \cdot h(\gamma_n).$$

Since $h(\gamma_n) \to 0$ as $n \to \infty$ by hypothesis, this implies $|\gamma_n|_v \to 1$ as $n \to \infty$. Since the sequence of heights $h(\gamma_n)$ is bounded, Northcott's theorem implies that

$[K(\gamma_n) : K] \to \infty$ as $n \to \infty$. So the right-hand side of (3) applied to γ_n tends to 0 as n goes to ∞. Therefore the desired convergence immediately falls out. □

To prove Proposition 3.2, we will separate the nonarchimedean and archimedean cases. The former case is handled in section 4, and the latter case is proved in section 5.

A common tool in both cases is the following result due to Schlickewei and Schmidt that is related to the geometry of numbers. We will not need the full strength of their result in relation to the implied constant, however. In fact the statement below, which suffices for our application, has a rather short proof.

Lemma 3.4. (*Keep the notation Γ_0 and h as above.*) *There exist some nonnegative integer r and some $\mathbb{Z}$-linearly independent elements $\gamma_1, \dots, \gamma_r \in \Gamma_0$, whose residues generate Γ_0 modulo its torsion subgroup, such that*

$$h(\gamma_1^{a_1} \cdots \gamma_r^{a_r}) \gg \max\{|a_1|, \dots, |a_r|\}$$

for all $(a_1, \dots, a_r) \in \mathbb{Z}^r$. The implied constant depends on Γ_0, but not on $(a_1, \dots, a_r)$.

Proof. A version with an explicit value for the implied constant follows from [16, Theorem 1.1]. In this reference the lower case h is the exponential height $\exp \circ h$. See also the work of Schlickewei and Schmidt [17]. □

4. Proof of Proposition 3.2 in the nonarchimedean case

In this section we will prove Proposition 3.2 in the nonarchimedean case. Let K be a number field, let v be a finite place of K, and let p be its residue characteristic. Let Γ_0 be a finitely generated subgroup of $K^\times$, and let $\Gamma \subset \overline{K}^\times$ be its division group. In particular, K, Γ_0, and Γ are as in Proposition 3.2.

In this section, unless otherwise stated, the constants implicit in $\ll$ or $\gg$ may depend on K, Γ_0, and v. But they will not depend on $\gamma \in \Gamma$, an element which will appear below.

4.1. Preparatory statements.

We start with a simple reduction step that Proposition 3.2 holds if $|\gamma|_v \neq 1$. This will not be used until section 4.2, however.

Lemma 4.1. (*Keep all the notation Γ and v above.*) *Suppose that $\gamma \in \Gamma$. The left-hand side of inequality* (4) *of Proposition* 3.2 *vanishes if $|\gamma|_v \neq 1$.*

Proof. Recall that $|\sigma(\gamma)|_v$ does not depend on $\sigma \in \mathrm{Hom}_K(K(\gamma), \overline{K}_v)$. Since $|\gamma|_v \neq 1$ by hypothesis, each summand on the left-hand side of inequality (4) equals $\log \max\{|\gamma|_v, 1\}$ by the ultrametric triangle inequality. Therefore the difference in question is zero. □

Let us fix elements $\gamma_1, \dots, \gamma_r$ as in Lemma 3.4 applied to Γ_0. To avoid trivialities, we will assume that $r \geq 1$, which can be achieved if we enlarge Γ_0 by simply adding, e.g., 2 to Γ_0, if needed. Just for caution, notice that we are allowed to enlarge Γ_0 in Proposition 3.2 without doing any harm.

For any $\gamma \in \Gamma$, there is a least integer $d(\gamma) \geq 1$ such that there exist $a_1, \dots, a_r \in \mathbb{Z}$ and a $\xi(\gamma) \in \mu_\infty$, which is uniquely determined by γ, with

$$\gamma^{d(\gamma)} = \xi(\gamma) \cdot \gamma_1^{a_1} \cdots \gamma_r^{a_r}. \tag{5}$$

Lemma 4.2. (*Keep all the notation K, Γ, and r as above.*) *For any $\gamma \in \Gamma$, we have*

$$d(\gamma) \ll [K(\gamma) : K]^{2r},$$

where the implied constant may depend on K and $\gamma_1, \dots, \gamma_r$, but not on γ.

Proof. Of course, we may assume without loss of generality that $d(\gamma) \geq 3$. Let us write γ as in (5). By Dirichlet's theorem [18, Theorem II.1B] applied to $N := d(\gamma) - 1 \geq d(\gamma)/2 > 1$, there are integers $q, p_1, \dots, p_r$ such that

$$1 \leq q < d(\gamma) \quad \text{and} \quad \left| q \cdot \frac{a_n}{d(\gamma)} - p_n \right| < \frac{1}{N^{1/r}} \leq \left(\frac{d(\gamma)}{2} \right)^{-1/r}$$

for all $1 \leq n \leq r$. We set $\delta = \gamma^q \gamma_1^{-p_1} \cdots \gamma_r^{-p_r}$ and use standard height properties to bound

$$h(\delta) \leq \sum_{n=1}^{r} \left| q \cdot \frac{a_n}{d(\gamma)} - p_n \right| h(\gamma_n) \ll d(\gamma)^{-1/r},$$

where the implied constant depends on the heights $h(\gamma_n)$ and r, but not on γ or δ.

Since $q < d(\gamma)$, the element δ cannot be a root of unity. So we may apply a weak version of Dobrowolski's theorem [9] to deduce

$$h(\delta) \gg [K(\delta) : K]^{-2} \geq [K(\gamma) : K]^{-2}.$$

Comparing the upper and lower bounds for $h(\delta)$ completes the proof. □

Next we bound the order $\operatorname{ord} \xi(\gamma)$ of the root of unity $\xi(\gamma)$ appearing in (5).

Lemma 4.3. (*Keep all the notation K, Γ, and $\xi(\gamma)$ as above.*) *For any $\gamma \in \Gamma$, we have*

$$\operatorname{ord} \xi(\gamma) \ll [K(\gamma) : K]^2,$$

where the implied constant may depend on K, but not on γ.

Proof. Clearly, $\xi(\gamma) \in K(\gamma)$ by (5). Thus it follows that

$$[K(\gamma) : K] \geq [K(\xi(\gamma)) : K] \geq [\mathbb{Q}(\xi(\gamma)) : \mathbb{Q}]/[K : \mathbb{Q}].$$

Now $[\mathbb{Q}(\xi(\gamma)) : \mathbb{Q}]$ is Euler's totient function evaluated at $\operatorname{ord} \xi(\gamma)$. The crude lower bound $[\mathbb{Q}(\xi(\gamma)) : \mathbb{Q}] \gg \big(\operatorname{ord} \xi(\gamma)\big)^{1/2}$, with an absolute implied constant, follows from [11, Theorem 327] and establishes the lemma. □

We follow a line of argumentation similar to the one of [4] and set

$$M(\gamma) := \inf\{|\zeta\gamma - 1|_v : \zeta \in \mu_\infty \text{ satisfies } \zeta\gamma \neq 1\}$$

for any $\gamma \in \overline{K}_v$. As we will see in the lemma below,

$$N(\gamma, \epsilon) := \#\{\zeta \in \mu_\infty : |\zeta\gamma - 1|_v \leq 1 - \epsilon\} \tag{6}$$

is finite if $\epsilon > 0$. We now use a Tate–Voloch style argument to bound $N(\gamma, \epsilon)$ from above.

Lemma 4.4. (*Keep all the notation K, v, and p as above.*) *If $\epsilon \in (0, 1)$ and $\gamma \in \overline{K}_v$, then*

$$N(\gamma, \epsilon) \leq p(\log p)\epsilon^{-1}.$$

Proof. We may assume without loss of generality that the set appearing in (6) contains at least one element, say ζ_0. By the ultrametric triangle inequality, we find $|\gamma^{-1}|_v = |\zeta_0|_v = 1$. Thus $|\zeta_0 - \gamma^{-1}|_v \leq 1 - \epsilon$. A further application of the ultrametric triangle inequality shows $|\zeta\zeta_0^{-1} - 1|_v \leq 1 - \epsilon$ for any other ζ in the mentioned set. Hence without loss of generality we may assume that $\gamma = 1$.

If $|\zeta - 1|_v \leq 1 - \epsilon$, then ζ has order p^n for some $n \geq 0$ (cf. [15, Lemma I.10.1]) and the fact that $\zeta - 1$ must be a unit if ζ has order divisible by at least two distinct prime numbers. If $n \geq 1$, then this reference and our normalization of the v-adic absolute value yield

$$1 - \epsilon \;\geq\; |\zeta - 1|_v \;=\; p^{-1/((p-1)p^{n-1})} \;\geq\; p^{-1/p^{n-1}}. \tag{7}$$

Taking the logarithm leads to $-\log(1-\epsilon) \leq (\log p)/p^{n-1}$. We use $-\log(1-\epsilon) \geq \epsilon$ to deduce $p^{n-1} \leq (\log p)\epsilon^{-1}$. Let n_0 be the largest integer with $p^{n_0} \leq p(\log p)\epsilon^{-1}$. Then $n_0 \geq 0$ and $n \leq n_0$, so $\zeta^{p^{n_0}} = 1$. The lemma follows since this leaves us with at most $p(\log p)\epsilon^{-1}$ possibilities for ζ. □

The next lemma provides a lower bound for $M(\gamma)$. Its proof relies on Yu's p-adic analog of A. Baker's estimate [22].

Lemma 4.5. (*Keep all the notation K and Γ as above.*) *If $\gamma \in \Gamma$, then*

$$\log M(\gamma) \;\gg\; -\log([K(\gamma) : K] \cdot \max\{2, h(\gamma)\}),$$

where the implied constant does not depend on γ.

Proof. We may assume without loss of generality that $M(\gamma) < p^{-1/(p-1)}$, since the implied constant is allowed to depend on p. Let $\epsilon > 0$, and suppose ζ is a root of unity such that $\zeta\gamma \neq 1$,

$$|\zeta\gamma - 1|_v \;<\; p^{-1/(p-1)} \quad \text{and} \quad |\zeta\gamma - 1|_v \;<\; M(\gamma) + \epsilon. \tag{8}$$

Later on, we will take ϵ toward zero, and we remark that the implied constants to appear in this proof are independent of γ and ϵ. We note that $|\gamma|_v = 1$ by the ultrametric triangle inequality.

The equality in (7) and the upper bound in the first inequality in (8) would be inconsistent with each other if $\zeta\gamma \neq 1$ were a root of unity. So $\zeta\gamma$ and hence γ are not roots of unity.

Let $e \in \mathbb{N}$ be the ramification index of K at v. We fix k to be the least integer with $k > e/(p-1)$. Thus $k \ll 1$. Using the notation in (5), write

$$\gamma^{d(\gamma)\cdot \operatorname{ord} \xi(\gamma)} \;=\; \gamma_1^{a_1 \cdot \operatorname{ord} \xi(\gamma)} \cdots \gamma_r^{a_r \cdot \operatorname{ord} \xi(\gamma)}. \tag{9}$$

This is an element of K since $\gamma_1, \ldots, \gamma_r \in K^\times$. However, $\gamma^{d(\gamma)\cdot \operatorname{ord} \xi(\gamma)}$ is also a unit in $\mathcal{O}_v$, the ring of integers of K_v. It represents a unit in $\mathcal{O}_v/\mathcal{M}_v^k$ where $\mathcal{M}_v$ is the maximal ideal of $\mathcal{O}_v$. Observe that $\mathcal{O}_v/\mathcal{M}_v^k$ is a finite ring which depends only on K and v; thus

$$g \;:=\; \#(\mathcal{O}_v/\mathcal{M}_v^k)^\times \;\ll\; 1.$$

By basic group theory γ^d represents 1 in the mentioned quotient ring, where

$$d \;:=\; g \cdot d(\gamma) \cdot \operatorname{ord} \xi(\gamma).$$

Any $\alpha \in \mathcal{M}_v^k$ satisfies $|\alpha|_v \leq p^{-k/e} < p^{-1/(p-1)}$ by our choice of k, so

$$|\gamma^d - 1|_v \;<\; p^{-1/(p-1)}. \tag{10}$$

Now $|(\zeta\gamma)^d - 1|_v \leq |\zeta\gamma - 1|_v$ by the ultrametric triangle inequality, so $|(\zeta\gamma)^d - 1|_v < p^{-1/(p-1)}$ because of (8). Combining this with (10) using the ultrametric triangle inequality gives

$$|1-\zeta^d|_v \;=\; |(1-\zeta^d)\gamma^d|_v \;=\; |(\gamma^d-1)-((\zeta\gamma)^d-1)|_v \;<\; p^{-1/(p-1)}.$$

As we have already observed, 1 is the only root of unity that lies v-adically this close to 1. Therefore $\zeta^d = 1$ and

$$|\gamma_1^{g\cdot a_1\cdot \operatorname{ord}\xi(\gamma)}\cdots\gamma_r^{g\cdot a_r\cdot\operatorname{ord}\xi(\gamma)}-1|_v = |\gamma^d-1|_v = |(\zeta\gamma)^d-1|_v \leq |\zeta\gamma-1|_v < M(\gamma)+\epsilon$$

by (8).

This chain of inequalities can lead to a nontrivial lower bound for $M(\gamma)+\epsilon$ only if $\gamma^d \neq 1$. This is the case, as γ has infinite order. So we can apply Yu's result [22, Theorem 1′] to bound the absolute value on the left-hand side from below and obtain

$$\log(M(\gamma)+\epsilon) \;\gg\; -\log\bigl(\operatorname{ord}\xi(\gamma)\cdot\max\{2,|a_1|,\dots,|a_r|\}\bigr).$$

Next we use Lemma 3.4 and (9) to estimate

$$d(\gamma)\cdot\operatorname{ord}\xi(\gamma)\cdot h(\gamma) \;=\; h(\gamma^{d(\gamma)\cdot\operatorname{ord}\xi(\gamma)}) \;\gg\; \operatorname{ord}\xi(\gamma)\cdot\max\{|a_1|,\dots,|a_r|\}.$$

Therefore we get

$$\log(M(\gamma)+\epsilon) \;\gg\; -\log\bigl(d(\gamma)\cdot\operatorname{ord}\xi(\gamma)\cdot\max\{2,h(\gamma)\}\bigr).$$

We insert the bounds from Lemmas 4.2 and 4.3 into this inequality and let $\epsilon\to 0^+$. Then the desired bound is immediate. □

4.2. **Finishing the proof of Proposition 3.2 in the nonarchimedean case.** We are ready to bound the absolute value in (4) from above to finish the proof of Proposition 3.2 in the nonarchimedean case.

By Lemma 4.1, we may assume that $|\gamma|_v = 1$. We note that $\sigma\in\operatorname{Hom}_K(K(\gamma),\overline{K}_v)$ implies

$$\phi(\sigma) \;:=\; \frac{\sigma(\gamma)}{\gamma} \;\in\; \mu_\infty,$$

since some power of γ lies in $\Gamma_0\subset K$. Motivated by (4), we write

$$S \;:=\; \frac{1}{D}\sum_{\sigma\in\operatorname{Hom}_K(K(\gamma),\overline{K}_v)}\log|\sigma(\gamma)-1|_v \;=\; \frac{1}{D}\sum_{\sigma\in\operatorname{Hom}_K(K(\gamma),\overline{K}_v)}\log|\phi(\sigma)\cdot\gamma-1|_v,$$

where we recall $D=[K(\gamma):K]$. Observe that $\gamma\neq 1$ by hypothesis, and hence that the logarithms are well-defined in the means above. We have $|\gamma|_v = 1$, so $|\phi(\sigma)\cdot\gamma-1|_v\leq 1$ by the ultrametric triangle inequality and thus

$$S \;\leq\; 0. \tag{11}$$

It remains to bound S from below. Let $\epsilon\in(0,1/2]$ be a parameter which we will fix properly further down. The various implied constants to appear below will

be independent of γ and ϵ. We have

$$S \;=\; \frac{1}{D}\left(\sum_{\substack{\sigma\in \mathrm{Hom}_K(K(\gamma),\overline{K}_v)\\ |\phi(\sigma)\cdot\gamma-1|_v\le 1-\epsilon}} \log|\phi(\sigma)\cdot\gamma-1|_v + \sum_{\substack{\sigma\in \mathrm{Hom}_K(K(\gamma),\overline{K}_v)\\ |\phi(\sigma)\cdot\gamma-1|_v> 1-\epsilon}} \log|\phi(\sigma)\cdot\gamma-1|_v\right)$$

(12)

$$\geq\; \frac{1}{D}\left(N(\gamma,\epsilon)\cdot\log M(\gamma)+D\cdot\log(1-\epsilon)\right),$$

where we recall that $M(\cdot)$ and $N(\cdot,\cdot)$ were defined near (6). Lemma 4.5 applied to γ yields

$$\log M(\gamma) \;\gg\; -\log(D\cdot\max\{2,h(\gamma)\}).$$

Next we multiply this with $N(\gamma,\epsilon)/D$ to get

$$\frac{N(\gamma,\epsilon)\cdot\log M(\gamma)}{D} \;\gg\; -N(\gamma,\epsilon)\cdot\frac{\log(D\cdot\max\{2,h(\gamma)\})}{D}. \tag{13}$$

We insert (13) into (12) to deduce

$$S \;\geq\; -c\cdot N(\gamma,\epsilon)\cdot\frac{\log(D\cdot\max\{2,h(\gamma)\})}{D}+\log(1-\epsilon),$$

where $c>0$ is a constant independent of γ and ϵ.

Finally, $\log(1-\epsilon)\geq -2\epsilon$ for $\epsilon\in(0,1/2]$ and $N(\gamma,\epsilon)\ll\epsilon^{-1}$ by Lemma 4.4. So it follows that

$$S \;\gg\; -\epsilon^{-1}\cdot\frac{\log(D\cdot\max\{2,h(\gamma)\})}{D}-\epsilon.$$

In the nonarchimedean case, the proposition follows from this estimate and (11) with an ample margin on choosing $\epsilon := 1/(2D^{1/2})$. □

5. Proof of Proposition 3.2 in the archimedean case

We keep K and Γ from Proposition 3.2 as well as some common notation from section 2. We also recall that Γ_0 is a finitely generated subgroup of $K^\times$ whose division group is Γ.

5.1. **Preparatory statements.** To handle the quantities appearing in (4) in the archimedean case, we need to have full control over the Galois conjugates of elements of Γ over a suitable ground field. The following lemma will provide such a ground field.

Lemma 5.1. (*Keep the notation K and Γ as above.*) *The group $\Gamma\cap K(\mu_\infty)^\times/\mu_\infty$ is torsion free and finitely generated. In particular, there exists a number field $F\subset\overline{K}$ such that*

$$K \;\subset\; F \;\subset\; K(\mu_\infty) \quad \textit{and} \quad \Gamma\cap K(\mu_\infty)^\times \;\subset\; \mu_\infty F^\times.$$

Proof. The existence of F follows from the first statement. To prove the first statement, we note that $\Gamma' := \Gamma\cap K(\mu_\infty)^\times/\mu_\infty$ is torsion free. So it suffices to show that Γ' is finitely generated. We observe that $\dim_{\mathbb{Q}}(\Gamma'\otimes_{\mathbb{Z}}\mathbb{Q})<\infty$, since $\dim_{\mathbb{Q}}(\Gamma\otimes_{\mathbb{Z}}\mathbb{Q})<\infty$. So Γ' has finite rank.

Now notice that $h(\alpha)=h(\zeta\alpha)$ for all algebraic numbers α and all roots of unity ζ. Hence $h:\Gamma\to[0,\infty)$ factors through a well-defined function $\Gamma'\to[0,\infty)$ which we also call h. If $\gamma,\gamma'\in\Gamma'$, then the triangle inequality $h(\gamma\gamma')\leq h(\gamma)+h(\gamma')$ as well as the homogeneity $h(\gamma^k)=|k|h(\gamma)$ for any $k\in\mathbb{Z}$ holds. Furthermore,

by Kronecker's theorem, $h(\gamma) = 0$ if and only if γ is the identity element of Γ'. Therefore, h determines an *abelian group norm* on Γ' in the sense of [23, p. 172].

The field $K(\mu_\infty)$, being an abelian extension of K, cannot contain elements of arbitrarily small positive height by Amoroso and Zannier's result in [1]. Thus the height induces a discrete abelian group norm on the countable group Γ'. Therefore, [23, p. 173, Theorem] implies that Γ' is a free abelian group. Being of finite rank, it must then be finitely generated. □

It is obvious that F is not uniquely determined. For the remainder of this section we fix such an F as in Lemma 5.1. The following lemma, reminiscent of classical Kummer theory, describes the Galois conjugates of elements in Γ and will later lead to an exponential sum.

Lemma 5.2. (*Keep the notation K and Γ as above.*) *Let F be a field as in Lemma 5.1. Let $\gamma \in \Gamma$, and let d be the least positive integer such that $\gamma^d \in K(\mu_\infty)$. Suppose that $\zeta \in \mu_\infty$ has order $N \in \mathbb{N}$ and satisfies $\gamma^d \in \zeta F^\times$. Then the following assertions hold.*

(i) *The extension $F(\mu_\infty, \gamma)/F(\mu_\infty)$ is cyclic of degree d.*

(ii) *The extension $F(\gamma^d)/F$ is Galois and the extension $F(\gamma)/F(\gamma^d)$ has degree d.*

(iii) *If $\sigma_0 : F(\gamma) \to \mathbb{C}$ is an arbitrary field embedding, then there is a subgroup $H \subset (\mathbb{Z}/N\mathbb{Z})^\times$ of index at most $[F : \mathbb{Q}]$ and an integer a_0 coprime to N such that*

$$\left\{ \frac{\sigma(\gamma)}{\sigma_0(\gamma)} \in \mathbb{C} : \sigma \text{ is a field embedding } F(\gamma) \to \mathbb{C} \text{ and } \sigma|_F = \sigma_0|_F \right\} = \left\{ e^{2\pi i \left(a_0 \frac{a-1}{dN} + \frac{k}{d}\right)} \in \mathbb{C} : a \in H \text{ and } k \text{ is an integer with } 0 \le k < d \right\}. \tag{14}$$

Moreover,

$$dN^{1/2} \ll [F(\gamma) : F] = d \cdot \#H \le d \cdot \varphi(N), \tag{15}$$

where the implied constant depends only on $[F : \mathbb{Q}]$ and we recall that $\varphi(\cdot)$ denotes Euler's totient function.

Observe that the factor $e^{2\pi i a_0 \frac{a-1}{dN}}$ appearing in (14) is only well-defined modulo a dth root of unity if $a \in (\mathbb{Z}/N\mathbb{Z})^\times$. But the set in the right-hand side of (14) makes perfect sense as k runs over $\{0, 1, \dots, d-1\}$.

Proof. By basic field theory, $K(\mu_\infty, \gamma)/K(\mu_\infty)$ is cyclic of degree d. Part (i) follows from $K(\mu_\infty) = F(\mu_\infty)$, a consequence of Lemma 5.1.

The extension $F(\gamma^d)/F$ is Galois since $F(\gamma^d) = F(\zeta)$. It remains to prove the second statement in part (ii). Let us consider the map

$$\mathrm{Hom}_{F(\mu_\infty)}(F(\mu_\infty, \gamma), \overline{K}) \to \mathrm{Hom}_{F(\mu_\infty) \cap F(\gamma)}(F(\gamma), \overline{K})$$

determined by restricting the field embeddings. It is injective, so comparing the cardinalities of the left- and right-hand sides yields

$$d = [F(\mu_\infty, \gamma) : F(\mu_\infty)] \le [F(\gamma) : F(\mu_\infty) \cap F(\gamma)],$$

where we used part (i). Now $F(\gamma^d) \subset F(\mu_\infty) \cap F(\gamma)$ and so $[F(\gamma) : F(\gamma^d)] \ge d$. But the reversed inequality clearly holds, and this concludes the proof of part (ii).

To prove part (iii), we may assume without loss of generality that $F(\gamma) \subset \mathbb{C}$ and that σ_0 is the inclusion. We begin by determining the various quantities in the

assertion. We note that $F(\gamma^d) = F(\zeta)$ since $\gamma^d \in \zeta F^\times$. There is $a_0 \in \mathbb{Z}$ coprime to N such that $\zeta = e^{2\pi i a_0/N}$. If $\tau \in \mathrm{Gal}(F(\gamma^d)/F)$, then

$$\frac{\tau(\gamma^d)}{\gamma^d} \;=\; \frac{\tau(\zeta)}{\zeta} \;=\; \zeta^{\psi(\tau)-1} \;=\; e^{2\pi i a_0 \frac{\psi(\tau)-1}{N}}, \tag{16}$$

where $\psi : \mathrm{Gal}(F(\gamma^d)/F) \to (\mathbb{Z}/N\mathbb{Z})^\times$ is an injective group homomorphism. Let us write H for the image of ψ. We note that $\#H = [F(\zeta) : F]$ and that the index of H in $(\mathbb{Z}/N\mathbb{Z})^\times$ is $\varphi(N)/[F(\zeta) : F] = [\mathbb{Q}(\zeta) : \mathbb{Q}]/[F(\zeta) : F] \le [F : \mathbb{Q}]$. Thus H satisfies the index bound claimed in part (iii).

We are interested in the set of values $\sigma(\gamma)/\gamma$ as σ runs over all the field embeddings $F(\gamma) \to \mathbb{C}$ that are the identity on F. In (16) above we described the set of their dth powers. Since $[F(\gamma) : F(\gamma^d)] = d$ by part (ii), we obtain the desired set by taking all the possible dth roots of elements in (16). This implies (14).

The final claim on $[F(\gamma) : F]$ follows from $[F(\gamma) : F(\gamma^d)] = d$, $[F(\gamma^d) : F] = \#H$, and $N^{1/2} \ll \varphi(N) \le N$, where the implicit constant is absolute. □

Below and in similar situations we will sometimes identify elements in $\mathbb{Z}/N\mathbb{Z}$ with their representatives respectively in $\{0, 1, \dots, N-1\}$.

Let $N \in \mathbb{N}$. For any subgroup $H \subset (\mathbb{Z}/N\mathbb{Z})^\times$ and any $m \in \mathbb{Z}$, we set

$$S(H, m) \;=\; \sum_{a \in H} e^{2\pi i \frac{ma}{N}}.$$

If m is coprime to N, then $S((\mathbb{Z}/N\mathbb{Z})^\times, m)$ is a Ramanujan sum and its modulus is known to be at most 1. In the next two lemmas we evaluate and bound similar Weyl sums attached to any subgroups $H \subset (\mathbb{Z}/N\mathbb{Z})^\times$.

Lemma 5.3. *Suppose that we are given $d, N \in \mathbb{N}$, a subgroup $H \subset (\mathbb{Z}/N\mathbb{Z})^\times$, and an integer a_0 coprime to N. Let $m \in \mathbb{Z}$, $g = \gcd(N, m/d)$, $m' = m/(dg)$, and $N' = N/g$. And let H' be the image of H under the natural group homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times$. Then we have*

$$\left| \sum_{\substack{a \in H \\ 0 \le k < d}} e^{2\pi i m \left(a_0 \frac{a-1}{dN} + \frac{k}{d}\right)} \right| \;=\; \begin{cases} 0 & \text{if } d \nmid m \quad \text{and} \\ d \cdot \frac{\#H}{\#H'} \cdot |S(H', a_0 m')| & \text{if } d \mid m. \end{cases}$$

Proof. Let us abbreviate $\xi = e^{2\pi i a_0/(dN)}$. The sum inside the absolute value is equal to

$$\sum_{a \in H} \xi^{m(a-1)} \sum_{k=0}^{d-1} e^{2\pi i \frac{mk}{d}}.$$

The second sum vanishes if $d \nmid m$ and the lemma is true in this case. Thus let us suppose that $d \mid m$, in which case the left-hand side of the assertion is equal to

$$d \left| \sum_{a \in H} \xi^{m(a-1)} \right| \;=\; d \left| \sum_{a \in H} \xi^{ma} \right| \;=\; d \left| \sum_{a \in H} e^{2\pi i a_0 \frac{m'a}{N'}} \right|$$

with m', N' as in the statement of the lemma. The sum on the right-hand side can be expressed in terms of the sum over the elements of $H' \subset (\mathbb{Z}/N'\mathbb{Z})^\times$ as described in the assertion, since the kernel of the surjective group homomorphism $H \to H'$ has $\#H/\#H'$ elements. This completes the proof. □

We now turn to the problem of bounding $S(H,m)$ from above for m coprime to N. The following clean bound and its proof were kindly provided by Bob Vaughan.

Lemma 5.4. *Let $N \in \mathbb{N}$ and let H be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$. If $m \in \mathbb{Z}$ is coprime to N, then*

$$|S(H,m)| \ \leq \ N^{1/2}.$$

Proof. Since H is a group, we have $S(H,m) = \sum_{a\in H} e^{2\pi i mab/N}$ for all $b \in H$. We may thus take the sum over H to obtain

$$\#H \cdot |S(H,m)|^2 \ = \ \sum_{b\in H} \left| \sum_{a\in H} e^{2\pi i \frac{mab}{N}} \right|^2 \ \leq \ \sum_{b=0}^{N-1} \left| \sum_{a\in H} e^{2\pi i \frac{mab}{N}} \right|^2 .$$

We expand the square on the right-hand side and switch the order of summation to get

$$\#H \cdot |S(H,m)|^2 \ \leq \ \sum_{a\in H} \sum_{a'\in H} \sum_{b=0}^{N-1} e^{2\pi i \frac{mab}{N}} e^{-2\pi i \frac{ma'b}{N}}.$$

Observe that $b \mapsto e^{2\pi i mab/N}$ determines a character of $\mathbb{Z}/N\mathbb{Z}$ for any $a \in H$. The different a yield different characters since m and N are coprime. So the triple sum equals $N \cdot \#H$ by the orthogonality of characters. The lemma follows on canceling $\#H$ and taking the square root of the resulting inequality. □

We use this last estimate to bound from above the Weyl sum arising from Lemma 5.3.

Lemma 5.5. *Let $B, d, m, N \in \mathbb{N}$ with $d \mid m$, and let $a_0 \in \mathbb{Z}$ be coprime to N. If $H \subset (\mathbb{Z}/N\mathbb{Z})^\times$ is a subgroup of index at most B, then*

$$\frac{1}{dm \cdot \#H} \left| \sum_{\substack{a\in H \\ 0\leq k<d}} e^{2\pi i m\left(a_0 \frac{a-1}{dN} + \frac{k}{d}\right)} \right| \ \ll \ \frac{\log\log(3N)}{(dmN)^{1/2}}, \tag{17}$$

where the implied constant depends only on B.

Proof. We set

$$\widetilde{m} = \frac{m}{d} \in \mathbb{Z}, \quad g = \gcd(N, \widetilde{m}), \quad m' = \frac{\widetilde{m}}{g}, \quad \text{and } N' = \frac{N}{g}.$$

By Lemma 5.3 the left-hand side of (17) is equal to $|S(H', a_0 m')|/(m \cdot \#H')$, where H' denotes the image of H under the natural group homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times$. We observe $\#H' \geq \varphi(N')/B$ from the surjectivity of $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/N'\mathbb{Z})^\times$ and the hypothesis $[(\mathbb{Z}/N\mathbb{Z})^\times : H] \leq B$. Then Lemma 5.4 yields

$$\frac{1}{m \cdot \#H'} \cdot |S(H', a_0 m')| \ \leq \ \frac{{N'}^{1/2}}{m \cdot \#H'} \ \leq \ B \cdot \frac{{N'}^{1/2}}{m \cdot \varphi(N')}.$$

Contrary to the situation in Lemma 4.3, it does not suffice to bound $\varphi(N')$ from below by a multiple of ${N'}^{1/2}$. However, [11, Theorem 328] provides $\varphi(N') \gg$

$N'/\log\log(3N')$, where the implied constant is absolute. So we have

$$\begin{aligned}\frac{1}{m\cdot\#H'}\cdot|S(H',a_0m')| &\ll \frac{\log\log(3N')}{mN'^{1/2}} = \frac{g^{1/2}\log\log(3N')}{mN^{1/2}}\\ &\le \frac{g^{1/2}\log\log(3N)}{mN^{1/2}} \le \frac{\widetilde{m}^{1/2}\log\log(3N)}{mN^{1/2}},\end{aligned}$$

where we notice that the implied constant depends only on B and that the last inequality comes from the divisibility $g \mid \widetilde{m}$. The lemma then follows from the equality $\widetilde{m}^{1/2}/m = (dm)^{-1/2}$. □

For any $D \in \mathbb{N}$, the *discrepancy* of a D-tuple $(x_1,\dots,x_D)$ of elements of $[0,1)$ is defined to be

$$\mathcal{D}(x_1,\dots,x_D) := \sup_{I\subset[0,1)} \big|\#\{1\le j\le D : x_j\in I\} - D\cdot\mathrm{leng}(I)\big|,$$

where I runs over all open, closed, or half-open intervals of $[0,1)$ and $\mathrm{leng}(I)$ denotes its length.

Let us retain the notation of Lemma 5.2; in particular σ_0 denotes an arbitrary field embedding of $F(\gamma)$ into $\mathbb{C}$. If $\sigma : F(\gamma)\to\mathbb{C}$ is a field embedding with $\sigma|_F = \sigma_0|_F$, then by (14) there is a rational number $x_\sigma\in[0,1)$ such that

$$\frac{\sigma(\gamma)}{\sigma_0(\gamma)} = e^{2\pi i x_\sigma}. \tag{18}$$

These x_σ are pairwise distinct and their number is equal to $[F(\gamma):F] = d\cdot\#H$. For convenience, we fix some order and use $x_1,\dots,x_D$ to denote the x_σ, where $D = [F(\gamma):F]$.

Lemma 5.6. (*Keep the notation D, F, etc. as above.*) *We have*

$$\frac{\mathcal{D}(x_1,\dots,x_D)}{D} \ll \frac{(\log\log(3D))^{2/3}}{D^{1/3}},$$

where the implied constant depends only on $[F:\mathbb{Q}]$.

Proof. In this proof the implied constants to appear will depend only on $[F:\mathbb{Q}]$. In a moment we will prove

$$\frac{\mathcal{D}(x_1,\dots,x_D)}{D} \ll \frac{(\log\log(3N))^{2/3}}{dN^{1/3}}, \tag{19}$$

where we recall that d and N come from Lemma 5.2. The right-hand side is at most $(\log\log(3dN))^{2/3}/(dN)^{1/3}$. This quantity is decreasing in sufficiently large dN. So the desired result follows from the bound

$$D = [F(\gamma):F] = d\cdot\#H \le d\cdot\varphi(N) \le dN.$$

Let $L\ge d$ be an integer that is to be determined below. We apply a theorem of Erdős and Turán [12, Theorem 5.5] to L and obtain

$$\begin{aligned}\frac{\mathcal{D}(x_1,\dots,x_D)}{D} &\le \frac{1}{L+1} + \frac{2(1+\pi^{-1})}{D}\sum_{m=1}^{L}\frac{1}{m}\left|\sum_{k=1}^{D} e^{2\pi i m x_k}\right|\\ &\le \frac{1}{L+1} + 2(1+\pi^{-1})\sum_{m=1}^{L}\frac{1}{dm\cdot\#H}\left|\sum_{\substack{a\in H\\ 0\le k<d}} e^{2\pi i m\left(a_0\frac{a-1}{dN}+\frac{k}{d}\right)}\right|,\end{aligned}$$

where we recall that H is as in Lemma 5.2, and we use $D = d \cdot \#H$. By Lemma 5.3 the sum on the right can only be nonzero if m is a multiple of d. In this case we write $m = d\widetilde{m}$ and rewrite the sum as a sum over $\widetilde{m}$. Next we use the bound in Lemma 5.5 with $B := [F : \mathbb{Q}]$ to bound the exponential sum and get

$$\frac{\mathcal{D}(x_1, \dots, x_D)}{D} \ll \frac{1}{L+1} + \frac{\log\log(3N)}{dN^{1/2}} \sum_{\widetilde{m}=1}^{\lfloor L/d \rfloor} \frac{1}{\widetilde{m}^{1/2}},$$

where the implied constant depends only on B. We can bound the sum as

$$\sum_{\widetilde{m}=1}^{\lfloor L/d \rfloor} \frac{1}{\widetilde{m}^{1/2}} \le 1 + \int_1^{\lfloor L/d \rfloor} \frac{dx}{x^{1/2}} \le 2 \left\lfloor \frac{L}{d} \right\rfloor^{1/2} \le 2 \left(\frac{L}{d} \right)^{1/2}.$$

We fix an absolute constant $c \geq 1$ such that

$$L = \left\lfloor c \cdot \frac{dN^{1/3}}{(\log\log(3N))^{2/3}} \right\rfloor \geq d.$$

This choice of L leads to (19) and thus to the lemma. □

We write $V(f)$ for the total variation over $[0, 1]$ of a real-valued function f whose domain contains $[0, 1]$.

Lemma 5.7. *Suppose that $\gamma \in \mathbb{C}$ and $\kappa \in (0, 1]$. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by*

$$f(x) := \log \max\{\kappa, |e^{2\pi i x}\gamma - 1|\}$$

for all $x \in \mathbb{R}$. Then f is continuous and satisfies $V(f) \leq 2\log(3/\kappa)$.

Proof. The continuity of f is obvious. To show the remaining assertion, we fix any $t \in [0, 1)$ such that $e^{2\pi i t}\gamma =: \gamma' \in [0, \infty)$ and introduce

$$\tilde{f}(x) := \log \max\{\kappa, |e^{2\pi i x}\gamma' - 1|\} = f(x + t)$$

for all $x \in \mathbb{R}$. By periodicity we have $V(\tilde{f}) = V(f)$. So it suffices to bound $V(\tilde{f})$ by the claimed value.

Observe that $\tilde{f}(x) = \tilde{f}(1 - x)$, since $\gamma' \in \mathbb{R}$. Moreover, $\tilde{f}(x)$ increases on $[0, 1/2]$ and decreases on $[1/2, 1]$, since $\gamma' \geq 0$. Therefore,

$$\begin{aligned} V(\tilde{f}) &= \tilde{f}(1/2) - \tilde{f}(0) + \tilde{f}(1/2) - \tilde{f}(1) \\ &= 2 \cdot \log \frac{\max\{\kappa, |\gamma' + 1|\}}{\max\{\kappa, |\gamma' - 1|\}} \\ &= 2 \cdot \log \frac{\gamma' + 1}{\max\{\kappa, |\gamma' - 1|\}}. \end{aligned}$$

We conclude the proof by checking the two cases $\gamma' < 2$ and $\gamma' \geq 2$. If $\gamma' < 2$, then we get $V(\tilde{f}) \leq 2\log((\gamma' + 1)/\kappa) < 2\log(3/\kappa)$. On the other hand, if $\gamma' \geq 2$, then we also get $V(\tilde{f}) \leq 2\log((\gamma' + 1)/(\gamma' - 1)) \leq 2\log 3 \leq 2\log(3/\kappa)$. □

The following estimate will be useful later; cf. [21, Exercise 1.1(b)].

Remark. Suppose $0 \leq \lambda < 1$ and $w \in \mathbb{C}$ such that $|w - 1| \leq \lambda$. Then the principal value $\log w$ of the logarithm satisfies $|\log w| \leq |w - 1|/(1 - \lambda)$.

Lemma 5.8. (*Keep all the notation K and Γ as in Proposition* 3.2 *and the notation F as in Lemma* 5.2.) *Let $\sigma_0 : F \to \mathbb{C}$ be an arbitrary field embedding. Let $\gamma \in \Gamma$ and $D = [F(\gamma) : F]$. If $\gamma \neq 1$, then*

$$\left| \frac{1}{D} \left(\sum_{\substack{\sigma : F(\gamma) \to \mathbb{C} \\ \sigma|_F = \sigma_0}} \log|\sigma(\gamma) - 1| \right) - \log \max\{|\sigma_0(\gamma)|, 1\} \right| \ll \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}},$$

where the implied constant depends on K, Γ, and F, but not on γ.

Even though γ may not lie in K, we recall that $|\sigma_0(\gamma)|$ is well-defined by the comment just before the statement of Proposition 3.1. In other words, it is independent of any choice of an extension of σ_0 to $F(\gamma)$. Also note that the result here is nothing but Proposition 3.2 in the archimedean case except for the fact that the ground field has changed from an arbitrary field K to a manageable field F. In section 5.2 below we will complete the proof of Proposition 3.2.

Proof of Lemma 5.8. In what follows, the implied constants to appear will be allowed to depend on K, Γ, and F, but not on γ.

Let S denote the sum

$$S := \frac{1}{D} \sum_{\substack{\sigma : F(\gamma) \to \mathbb{C} \\ \sigma|_F = \sigma_0}} \log|\sigma(\gamma) - 1|$$

from the statement of the lemma.

Let $x_1, \ldots, x_D \in [0, 1)$ be rational numbers as near (18). So

$$\sigma(\gamma) = \frac{\sigma(\gamma)}{\sigma_0(\gamma)} \sigma_0(\gamma) = e^{2\pi i x_j} \sigma_0(\gamma)$$

for all σ as in the definition of S, here j depends on σ. By abuse of notation we write γ for the complex number $\sigma_0(\gamma)$. In this notation we have $\sigma(\gamma) = e^{2\pi i x_j} \gamma$.

Fix an absolute constant $c > 0$ such that

$$\kappa := c \cdot \frac{(\log\log(3D))^{2/3}}{D^{1/3}} \in (0, 1/2]. \tag{20}$$

We split up S into the sum $S_1 + S_2$, where

$$S_1 := \frac{1}{D} \sum_{\substack{1 \leq j \leq D \\ |e^{2\pi i x_j} \gamma - 1| \leq \kappa}} \log|e^{2\pi i x_j} \gamma - 1|$$

and

$$S_2 := \frac{1}{D} \sum_{\substack{1 \leq j \leq D \\ |e^{2\pi i x_j} \gamma - 1| > \kappa}} \log|e^{2\pi i x_j} \gamma - 1|.$$

As $\kappa \leq 1$, we get $S_1 \leq 0$.

Let j correspond to a summand of S_1, and let $\theta \in (-1/2, 1/2]$ such that

$$e^{2\pi i \theta} |\gamma| = e^{2\pi i x_j} \gamma.$$

The remark before this lemma, applied with $\lambda := 1/2 \geq \kappa$ and $w := e^{2\pi i x_j}\gamma$, yields

$$|2\pi\theta| \ \leq \ |e^{2\pi i x_j}\gamma - 1|/(1-\lambda) \ \leq \ 2\kappa$$

by the choice of the j. Thus θ lies in an interval of length at most $2\kappa/\pi$. Now we may use the discrepancy bound in Lemma 5.6 together with the choice of (20) to deduce that the total number of summands in S_1 is $\ll \kappa D + (D \cdot \log\log(3D))^{2/3} \ll (D \cdot \log\log(3D))^{2/3}$. Thus it follows that

$$S_1 \ \gg \ \frac{(\log\log(3D))^{2/3}}{D^{1/3}} \cdot \min_{\substack{1\leq j\leq D \\ |e^{2\pi i x_j}\gamma - 1|\leq \kappa}} \log|e^{2\pi i x_j}\gamma - 1|. \tag{21}$$

We now use A. Baker's theory on linear forms in logarithms to bound this minimum from below. We refer to A. Baker and Wüstholz's explicit estimate in [3].

By Dirichlet's theorem on S-units, $\Gamma \cap F^\times$ is a finitely generated abelian group. We fix $\gamma_1, \ldots, \gamma_r$ as in Lemma 3.4 applied to $\Gamma \cap F^\times$. They generate $\Gamma \cap F^\times$ modulo its torsion subgroup. Notice that this choice does not depend on γ.

Let $d \geq 1$ be minimal with $\gamma^d \in K(\mu_\infty)$, as in Lemma 5.2. Then $\gamma^d \in \zeta F^\times$ for some root of unity ζ of order N by Lemma 5.1. Let t be the cardinality of the torsion subgroup of $\Gamma \cap F^\times$. We observe that $t \ll 1$ and that $(\gamma^d\zeta^{-1})^t$ lies in the group generated by $\gamma_1, \ldots, \gamma_r$, i.e., that

$$\gamma^{dt} \ = \ \zeta^t \cdot \gamma_1^{a_1} \cdots \gamma_r^{a_r} \tag{22}$$

for some integers $a_1, \ldots, a_r$.

We take $l_1, \ldots, l_r \in \mathbb{C}$ such that

$$e^{l_n} \ = \ \gamma_n \quad \text{and} \quad \mathrm{Im}(l_n) \ := \ \text{the imaginary part of } l_n \ \in \ (-\pi, \pi]$$

for each $1 \leq n \leq r$. So the l_n are the principal values of logarithms. By (22) there is an integer k such that

$$\gamma \ = \ e^{2\pi i \frac{k}{dtN} + \frac{a_1 l_1 + \cdots + a_r l_r}{dt}}.$$

Either from noticing that $|e^{2\pi i x_j}\gamma - 1| \leq \kappa \leq 1/2$ in S_1 and using the same remark before the current lemma or from the Taylor expansion of e^z at $z = 0$, we then see that

$$\left| 2\pi i \left(\frac{k}{dtN} + x_j \right) + \sum_{n=1}^{r} \frac{a_n}{dt} l_n \right| \ \ll \ |e^{2\pi i x_j}\gamma - 1|$$

after possibly adding an integer multiple of dtN to k to make sure that the left-hand side is close enough to 0. We multiply the expression inside the absolute value symbol by dtN to get the linear form

$$\Lambda \ := \ 2\pi i(k + x_j dtN) + \sum_{n=1}^{r} a_n l_n N$$

in logarithms. It is small in the sense that

$$|\Lambda| \ \ll \ dN \cdot |e^{2\pi i x_j}\gamma - 1| \ \leq \ \kappa dN, \tag{23}$$

where the first inequality comes from $t \ll 1$ and the second inequality comes from the choice of j.

If $\Lambda = 0$, then evaluating $e^{\Lambda/(dtN)}$ there shows that $1 = e^{2\pi i x_j}\gamma = \sigma(\gamma)$ for the field embedding σ attached to x_j. Hence $\gamma = 1$ which contradicts our hypothesis.

Therefore $\Lambda \neq 0$ and to find a lower bound for $|\Lambda|$, we first need to control the integer coefficients appearing in Λ. The imaginary part of Λ is

$$2\pi(k + x_j dtN) + \sum_{n=1}^{r} a_n \cdot \mathrm{Im}(l_n) \cdot N.$$

We have $|\mathrm{Im}(\Lambda)| \leq |\Lambda| \ll \kappa dN \leq dN$ and so

$$|k| \ \ll \ N \cdot \max\{d, |a_1|, \ldots, |a_r|\}, \tag{24}$$

since $t \ll 1, |\mathrm{Im}(l_n)| \leq \pi$ for all $0 \leq n \leq r$, and $|x_j| \leq 1$ for all $1 \leq j \leq D$. On the other hand, basic height properties, Lemma 3.4, and (22) yield

$$dt \cdot h(\gamma) \ = \ h(\gamma^{dt}) \ = \ h(\gamma_1^{a_1} \cdots \gamma_r^{a_r}) \ \gg \ \max\{|a_1|, \ldots, |a_r|\}.$$

In combination with (24) we obtain

$$\max\{k, |a_1|N, \ldots, |a_r|N\} \ \ll \ dN \cdot \max\{1, h(\gamma)\}.$$

By (14) we observe that $x_j dN$ is an integer in $[0, dN)$.

We now apply the main theorem of [3] and obtain

$$\log|\Lambda| \ \gg \ -\log(dN \cdot \max\{2, h(\gamma)\}).$$

We take the logarithm of (23) and immediately see that

$$\log|e^{2\pi i x_j}\gamma - 1| \ \gg \ -\log(dN \cdot \max\{2, h(\gamma)\}).$$

The first bound in (15) yields $\log(dN) \ll \log(2D)$. Thus we have

$$\log|e^{2\pi i x_j}\gamma - 1| \ \gg \ -\log(D \cdot \max\{2, h(\gamma)\}).$$

Recall that this bound holds for all the j over which the minimum in (21) is taken. Since $S_1 \leq 0$, it then follows that

$$|S_1| \ \ll \ \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}. \tag{25}$$

So as to control S, we must also find a bound for S_2. Here we will again require the discrepancy bound proved in Lemma 5.6. We work with the function

$$f(x) \ := \ \log\max\{\kappa, |e^{2\pi i x}\gamma - 1|\}$$

from Lemma 5.7 defined on $[0, 1]$, where κ is as in (20).

Another way to write S_2 is

$$S_2 \ = \ \frac{1}{D}\left(\#\left\{1 \leq j \leq D : |e^{2\pi i x_j}\gamma - 1| \leq \kappa\right\} \cdot \log(\kappa^{-1}) + \sum_{j=1}^{D} f(x_j)\right).$$

The cardinality of the set in this expression is just the number of terms in S_1 and this number has been established as being $\ll (D \cdot \log\log(3D))^{2/3}$ above. The definition of κ in (20) yields

$$\left|S_2 - \frac{1}{D}\sum_{j=1}^{D} f(x_j)\right| \ \ll \ \frac{\log(2D) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}. \tag{26}$$

To proceed, we use Koksma's inequality [12, Theorem 5.4] and compare the sum $\frac{1}{D}\sum_{j=1}^{D} f(x_j)$ with the integral $\int_0^1 f(x)\, dx$. Now note that $f : [0, 1] \to \mathbb{R}$ is

a continuous function with $V(f) \leq 2\log(3/\kappa)$ by Lemma 5.7. Again we use the definition of κ to get the bound

$$V(f) \ll |\log \kappa| \ll \log(2D).$$

Koksma's inequality yields

$$\left| \frac{1}{D} \sum_{j=1}^{D} f(x_j) - \int_0^1 f(x)\, dx \right| \leq V(f) \cdot \frac{\mathcal{D}(x_1, \dots, x_D)}{D}.$$

We apply the usual discrepancy bound to get

$$\left| \frac{1}{D} \sum_{j=1}^{D} f(x_j) - \int_0^1 f(x)\, dx \right| \ll \frac{\log(2D) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}.$$

So

$$\left| S_2 - \int_0^1 f(x) dx \right| \ll \frac{\log(2D)(\log\log(3D))^{2/3}}{D^{1/3}} \tag{27}$$

by (26).

It remains to estimate the integral of f. Look at

$$\int_0^1 f(x)\, dx \geq \int_0^1 \log|e^{2\pi i x}\gamma - 1|\, dx = \log\max\{1, |\gamma|\}, \tag{28}$$

where the equality is nothing but Jensen's formula. To obtain an upper bound, we use $\kappa \leq 1$ to estimate

$$\begin{aligned} \int_0^1 f(x)\, dx &\leq \int_{|e^{2\pi i x}\gamma - 1| > \kappa} \log|e^{2\pi i x}\gamma - 1|\, dx \\ &= \int_0^1 \log|e^{2\pi i x}\gamma - 1|\, dx - \int_{|e^{2\pi i x}\gamma - 1| \leq \kappa} \log|e^{2\pi i x}\gamma - 1|\, dx \\ &= \log\max\{1, |\gamma|\} - \int_{|e^{2\pi i x}\gamma - 1| \leq \kappa} \log|e^{2\pi i x}\gamma - 1|\, dx, \end{aligned}$$

where Jensen's formula is used again. We use the remark just above (21) to see that the last integral is over an arc of length $\ll \kappa$. Using $-\int_0^\epsilon \log x\, dx = \epsilon(1 - \log\epsilon)$ and (28), we conclude

$$\left| \int_0^1 f(x)\, dx - \log\max\{1, |\gamma|\} \right| \ll \kappa(1 - \log\kappa) \ll \frac{\log(2D) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}.$$

We recall (27) and conclude

$$|S_2 - \log\max\{1, |\gamma|\}| \ll \frac{\log(2D) \cdot (\log\log(3D))^{2/3}}{D^{1/3}}. \tag{29}$$

Then the lemma is immediate from $S = S_1 + S_2$ and the bounds in (25) and (29). □

5.2. **Finishing the Proof of Proposition 3.2 in the archimedean case.** To complete the proof of the proposition in the archimedean case, we have only to use a standard fact about the behavior of absolute values under field extensions to reduce the case of an arbitrary ground field K to that of our preferable field F and whence to Lemma 5.8.

In our archimedean case, the sum inside the absolute value symbol in (4) equals

$$\frac{1}{D} \sum_{\substack{\sigma : K(\gamma) \to \mathbb{C} \\ \sigma|_K = \sigma_0}} \log |\sigma(\gamma) - 1|,$$

where σ_0 is a fixed field embedding of K into $\mathbb{C}$.

Recall that F comes from Lemma 5.1 and contains K. A basic fact about absolute values under field extensions implies

$$\sum_{\substack{\sigma : K(\gamma) \to \mathbb{C} \\ \sigma|_K = \sigma_0}} \log |\sigma(\gamma) - 1| \;=\; \frac{1}{[F(\gamma) : K(\gamma)]} \sum_{\substack{\sigma : F(\gamma) \to \mathbb{C} \\ \sigma|_K = \sigma_0}} \log |\sigma(\gamma) - 1|.$$

On dividing by $D = [K(\gamma) : K]$, we get the equality of means

$$\frac{1}{D} \sum_{\substack{\sigma : K(\gamma) \to \mathbb{C} \\ \sigma|_K = \sigma_0}} \log |\sigma(\gamma) - 1| \;=\; \frac{1}{[F : K]} \sum_{\substack{\tau : F \to \mathbb{C} \\ \tau|_K = \sigma_0}} \left(\frac{1}{[F(\gamma) : F]} \sum_{\substack{\sigma : F(\gamma) \to \mathbb{C} \\ \sigma|_F = \tau}} \log |\sigma(\gamma) - 1| \right).$$

Therefore it suffices to prove the proposition with K replaced by F. But we have already done so in Lemma 5.8. □

6. Proof of the theorems

The heart of our proof of Theorem 1.1 lies in the proof of Theorem 1.3.

6.1. **Proof of Theorem 1.3.** We will introduce a lemma which gives a height bound for integral points. This will readily imply Theorem 1.3.

Lemma 6.1. *Let K, S, Γ, and α be as in Theorem* 1.3. *Then there exists a constant $c > 0$ depending only on K, S, Γ, and α such that for all $\gamma \in \Gamma$ that are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α), we have*

$$h(\gamma) \;\le\; c \quad \textit{and} \quad h(\gamma\alpha^{-1}) \;\le\; c \cdot \frac{\log(2D) \cdot (\log\log(3D))^{2/3}}{D^{1/3}},$$

where $D = [K(\gamma) : K]$.

Proof. Without loss of generality, we may extend K and take all the primes lying over those of S, and assume that $\alpha \in \mathbb{G}_{\mathrm{m}}(K)$. The property of being S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α) becomes easier to satisfy when we enlarge S. Thus we may assume $|\alpha|_v = 1$ for all places $v \notin S$ of K.

Let γ be as in the hypothesis, and note that $\gamma \neq \alpha$. Let $L = K(\gamma)$. Then by the product formula we observe that

$$\begin{aligned} 0 &= \sum_{w \in M_L} [L_w : \mathbb{Q}_w] \cdot \log |\gamma - \alpha|_w \\ &= \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \sum_{\substack{w \in M_L \\ w|v}} [L_w : K_v] \cdot \log |\gamma - \alpha|_w \\ &= \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \sum_{\sigma \in \mathrm{Hom}_K(L, \overline{K}_v)} \log |\sigma(\gamma - \alpha)|_v, \end{aligned}$$

where the last equality follows from [6, Lemma 1.3.7]

Recall that we assume above that $|\alpha|_v = 1$ for all places $v \notin S$ of K. Since γ is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α), the terms coming from the $v \in M_K - S$ in the last sum above must vanish by (1). If we omit them, then we obtain

$$\sum_{v \in S} [K_v : \mathbb{Q}_v] \sum_{\sigma \in \mathrm{Hom}_K(L, \overline{K}_v)} \log |\sigma(\gamma) - \alpha|_v \;=\; 0.$$

We divide both sides by $D = [L : K]$ and apply Proposition 3.1 to each element of S. Thus there is a constant $c_1 > 0$, independent of γ, such that

$$\sum_{v \in S} [K_v : \mathbb{Q}_v] \left(\log \max\{|\gamma|_v, |\alpha|_v\} - c_1 \cdot \frac{\log(D \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}} \right) \leq 0;$$

we recall the convention introduced just before Proposition 3.1 on $|\gamma|_v$. If we rearrange this sum, then we get

(30)

$$\sum_{v \in S} [K_v : \mathbb{Q}_v] \log \max\{|\gamma|_v, |\alpha|_v\} \;\leq\; c_2 \cdot \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}},$$

where $c_2 = c_1 \cdot \sum_{v \in S} [K_v : \mathbb{Q}_v]$ is also independent of γ.

Recall that $|\gamma|_v = 1$ if $v \in M_K - S$ by (1) and $|\alpha|_v = 1$ for the same v. So we may replace "$v \in S$" in (30) by "$v \in M_K$" without changing the value of the sum. We divide the resulting quantity in the left-hand side by $[K : \mathbb{Q}]$. This obtained quotient equals $h(\gamma\alpha^{-1})$ by the product formula. Therefore (30) results in

$$h(\gamma\alpha^{-1}) \;\leq\; c_3 \cdot \frac{\log(D \cdot \max\{2, h(\gamma)\}) \cdot (\log\log(3D))^{2/3}}{D^{1/3}} \tag{31}$$

for some constant $c_3 > 0$, where we emphasize that c_3 does not depend on the individual γ. Recall that $h(\gamma) \leq h(\gamma\alpha^{-1}) + h(\alpha)$. As the right-hand side of (31) is asymptotically logarithmic in $h(\gamma)$, we conclude that $h(\gamma)$ is bounded.

The second inequality in the assertion follows immediately from the first one by inserting the height upper bound into (31). □

Proof of Theorem 1.3. Part (i) is nothing but the first assertion of Lemma 6.1. Below Theorem 1.3 in the Introduction, we already showed that the second convergence in part (ii) is an easy consequence of the first convergence. Thus only the first equality of convergence in part (ii) needs justification. Indeed, let $(\gamma_n)_{n \geq 1}$ be as in the statement of part (ii) of the theorem. Part (i) says that the heights of the γ_n are bounded, and so we have $[K(\gamma_n) : K] \to \infty$ as $n \to \infty$, by Northcott's

theorem. Then the right-hand side of the second assertion of Lemma 6.1 applied to $\gamma := \gamma_n$ must go to 0 as n goes to ∞. □

6.2. **Proof of Corollary 1.4.** This corollary is a simple application of Theorem 1.3. We observe that $h(\alpha) > 0$ by Kronecker's theorem applied to our hypothesis on α. An infinite sequence of elements in (2) would imply $h(\alpha) \leq \epsilon$ by Theorem 1.3(ii). So any ϵ in $(0, h(\alpha))$ will do. □

6.3. **Proof of Corollary 1.5.** As was mentioned in the Introduction, the following proof will be independent of any results in [10] and hence will give a new proof to Theorem 2.5 there.

We begin with part (ii). Thus assume that $\alpha \in \mathbb{G}_{\mathrm{m}}(\overline{K}) - \Gamma$. Suppose to a contradiction that $(\gamma_n)_{n\geq 1}$ is a sequence of pairwise distinct elements in Γ that are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α).

We will use the hypothesis that $\alpha \notin \Gamma$ to derive a contradiction. Let Ω be the division group attached to the subgroup Ω_0 of $\mathbb{G}_{\mathrm{m}}(\overline{K})$ that is generated by Γ_0 and α. The hypothesis that $\alpha \notin \Gamma$ implies that $\alpha \notin \mu_\infty$. So Ω_0 has rank $r \geq 1$.

We fix independent elements $\omega_1, \ldots, \omega_r$ as in Lemma 3.4 applied to Ω_0. Hence for each integer $n \geq 1$, there are integers $a_{n0}, a_{n1}, \ldots, a_{nr}$ such that

$$\gamma_n^{a_{n0}} \;=\; \omega_1^{a_{n1}} \cdots \omega_r^{a_{nr}}$$

and $a_{n0} \geq 1$. Similarly, write

$$\alpha^{b_0} \;=\; \omega_1^{b_1} \cdots \omega_r^{b_r}$$

for some integers $b_0, b_1, \ldots, b_r$ with $b_0 \geq 1$. Using basic height properties and Lemma 3.4, we find

$$h(\gamma_n \alpha^{-1}) \;\gg\; \sum_{j=1}^{r} \left| \frac{a_{nj}}{a_{n0}} - \frac{b_j}{b_0} \right|, \tag{32}$$

where the implied constant does not depend on n. By Theorem 1.3(ii) the height on the left-hand side tends to zero as n goes to ∞. This means that the γ_n converge to α in the height metric. And from (32) we conclude that the $a_n := (a_{n1}/a_{n0}, \ldots, a_{nr}/a_{n0})$ converge to $b := (b_1/b_0, \ldots, b_r/b_0)$ in the euclidean metric on $\mathbb{R}^r$.

The vectors $a_1, a_2, \ldots$ generate a vector subspace of $\mathbb{Q}^r$ that does not contain b; indeed, otherwise, we could write some positive integer power of α as a product of integer powers of some γ_n's, which contradicts the hypothesis that $\alpha \notin \Gamma$. Thus we can find some $(l_1, \ldots, l_r) \in \mathbb{Z}^r$ such that

$$\sum_{j=1}^{r} a_{nj} l_j \;=\; 0 \quad \text{for all } n \geq 1 \quad \text{and} \quad \sum_{j=1}^{r} b_j l_j \;\neq\; 0.$$

But this is impossible, since $a_n \to b$ as $n \to \infty$. This contradiction completes the proof of part (ii).

To prove part (i), we assume additionally that the γ_n are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (β), too. Then

$$h(\alpha\beta^{-1}) \;=\; h(\gamma_n^{-1}\alpha\gamma_n\beta^{-1}) \;\leq\; h(\gamma_n\alpha^{-1}) + h(\gamma_n\beta^{-1})$$

holds for all $n \geq 1$, from basic height properties. Thus part (ii) of Theorem 1.3 implies $h(\alpha\beta^{-1}) = 0$ on taking $n \to \infty$. Therefore $\alpha\beta^{-1}$ is a root of unity by Kronecker's theorem, which is a contradiction to the hypothesis. □

6.4. **Proof of Theorem 1.6.** We will use Theorem 1.3 and its Corollary 1.5. It follows from the assumption that $(\gamma_n)_{n\geq 1}$ is a sequence of pairwise distinct elements of Γ that are S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to (α) for any element α in the support of D. By Corollary 1.5(ii) all such α must lie in Γ. Part (i) of the same corollary implies that D is the translate of a torsion divisor on $\mathbb{G}_{\mathrm{m}}$ by some α in the support of D. Finally, the first equality in Theorem 1.3(ii) implies that $\lim_{n\to\infty} h(\gamma_n\alpha^{-1}) = 0$. □

6.5. **The completion of the proof of Theorem 1.1.** Let K and S be as in Theorem 1.1, and let $\alpha, \gamma \in \overline{K}$. We require the notion of an S-integral point on $\mathbb{P}^1$ relative to an effective divisor. We identify α, γ with $[1:\alpha], [1:\gamma] \in \mathbb{P}^1(\overline{K})$ respectively, and say that γ is *S-integral* on $\mathbb{P}^1$ relative to the divisor (α) on $\mathbb{P}^1$, if for all places $v \notin S$ of K and all K-embeddings $\sigma : K(\gamma) \to \overline{K}_v$ and $\tau : K(\alpha) \to \overline{K}_v$ we have

$$\begin{cases} |\sigma(\gamma) - \tau(\alpha)|_v \ \geq\ 1 & \text{if } |\tau(\alpha)|_v \leq 1, \\ |\sigma(\gamma)|_v \ \leq\ 1 & \text{if } |\tau(\alpha)|_v > 1. \end{cases} \tag{33}$$

We say that γ is S-integral on $\mathbb{P}^1$ relative to (∞), with $\infty = [0:1]$, if $|\sigma(\gamma)|_v \leq 1$ for all v and σ as above. By symmetry ∞ is said to be S-integral on $\mathbb{P}^1$ relative to (γ) in this case. If D is any effective divisor on $\mathbb{P}^1$, the set of S-integral points on $\mathbb{P}^1$ relative to D is defined as in the case of $\mathbb{G}_{\mathrm{m}}$. By employing notation of [10], we briefly let $\mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S})$ denote this set, where $\overline{\mathcal{O}}_{K,S}$ stands for the integral closure in $\overline{K}$ of the ring of S-integers of K. If E is an effective divisor on $\mathbb{G}_{\mathrm{m}}$, we also write $\mathbb{G}_{\mathrm{m},E}(\overline{\mathcal{O}}_{K,S})$ for the set of algebraic numbers that are S-integral points on $\mathbb{G}_{\mathrm{m}}$ relative to E that is equal to $\mathbb{P}^1_{E+(0)+(\infty)}(\overline{\mathcal{O}}_{K,S})$. Observe that $\mathbb{P}^1_E(\overline{\mathcal{O}}_{K,S}) \supset \mathbb{G}_{\mathrm{m},E}(\overline{\mathcal{O}}_{K,S})$ not only because of $\mathbb{P}^1 \supset \mathbb{G}_{\mathrm{m}}$.

We begin by proving the following proposition.

Proposition 6.2. (*Keep the notation as in Theorem* 1.1.) *Suppose that $D = (\alpha)$ for some $\alpha \in \Gamma$. If $\mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$ is infinite, then S must contain all the finite primes of K that divide α or α^{-1}.*

Proof. Without loss of generality, we may assume that $\alpha \in K^\times$ by enlarging K and replacing S with the set of all the primes lying over the elements of S if needed. This does not change the problem. Since $\mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$ is infinite by hypothesis, there are infinitely many elements $\gamma \in \Gamma$ such that $\alpha\gamma \in \mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$.

If γ is as above, we claim $\gamma \in \mathbb{P}^1_{(1)}(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$. Indeed, let v be a finite place of K not in S.

On the one hand, let us suppose $|\alpha|_v \leq 1$. Then for any K-embedding $\sigma : K(\gamma) \to \overline{K}_v$, we have $1 \leq |\sigma(\alpha\gamma) - \alpha|_v = |\alpha|_v \cdot |\sigma(\gamma) - 1|_v$ by (33) and thus

$$|\sigma(\gamma) - 1|_v \geq |\alpha|_v^{-1} \geq 1. \tag{34}$$

On the other hand, suppose $|\alpha|_v > 1$. Then by (33) we have $|\sigma(\gamma)|_v \leq |\alpha|_v^{-1} < 1$ for any σ as before. This implies that $|\sigma(\gamma) - 1|_v = 1 \geq 1$.

In both cases we concluded $|\sigma(\gamma) - 1|_v \geq 1$ for all σ as above and all $v \in M_K - S$. This settles our claim on γ above.

For a suitably large but finite set of places $S' \supset S$ of K we have $\mathbb{P}^1_{(1)}(\overline{\mathcal{O}}_{K,S}) \cap \Gamma \subset \mathbb{G}_{\mathrm{m},(1)}(\overline{\mathcal{O}}_{K,S'}) \cap \Gamma$. By Theorem 1.3(ii) applied to S', we know that $h(\gamma)$ should be arbitrarily close to 0 for all but finitely many γ. However, arguing by using

[6, Lemma 1.3.7], we find that (34) implies

$$h(\gamma) \;\geq\; \frac{1}{[K:\mathbb{Q}]}\log(|\alpha|_v^{-1}) \;>\; 0$$

if $|\alpha|_v < 1$ and $v \in M_K - S$. As the quantity in the middle is independent of γ, we arrive at a contradiction. Therefore, $|\alpha|_v \geq 1$ for all $v \in M_K - S$.

It remains only to show that S must contain all the finite primes of K that divide α^{-1} too. We will do a case-by-case study and deduce this statement from what we have already proved.

On the one hand, suppose that $|\alpha|_v > 1$ for some $v \in M_K - S$. Using (33) and the ultrametric triangle inequality, we find

$$\left|\sigma\left(\frac{1}{\alpha\gamma}\right) - \frac{1}{\alpha}\right|_v \;=\; \left|\sigma\left(\frac{1}{\alpha\gamma}\right)\right|_v \;\geq\; 1$$

for any K-embedding $\sigma : K(\gamma) \to \overline{K}_v$.

On the other hand, suppose that $|\alpha|_v = 1$. Then for any K-embedding $\sigma : K(\gamma) \to \overline{K}_v$, we have

$$\left|\sigma\left(\frac{1}{\alpha\gamma}\right) - \frac{1}{\alpha}\right|_v \;=\; \left|\frac{1}{\alpha}\cdot\left(\frac{1}{\sigma(\gamma)} - 1\right)\right|_v \;=\; \left|\frac{\sigma(\gamma)-1}{\sigma(\gamma)}\right|_v.$$

If $|\sigma(\gamma)|_v \neq 1$, then the rightmost quotient is clearly ≥ 1. If $|\sigma(\gamma)|_v = 1$, then the rightmost quotient is equal to $|\sigma(\gamma)-1|_v$, which is ≥ 1 by (34). Thus the quotient is always ≥ 1 and it follows that $(\alpha\gamma)^{-1} \in \mathbb{P}^1_{(\alpha^{-1})}(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$. Since there are infinitely many such elements $(\alpha\gamma)^{-1}$, we conclude, from what we proved above, that S must contain all the finite primes of K that divide α^{-1} too, as desired. □

The completion of the proof of Theorem 1.1. We begin with part (ii). From Theorem 1.6 we already know that $D = \alpha T$, the translate of some torsion divisor T of $\mathbb{G}_{\mathrm{m}}$ by some element $\alpha \in \Gamma$ with $\lim_{n\to\infty} h(\gamma_n\alpha^{-1}) = 0$. (Note that D is not assumed to be irreducible here.) Without loss of generality we may assume $\alpha \in K^\times$ and that 1 is in the support of T. By Proposition 6.2 above we know that S must contain all the finite primes of K that divide α or α^{-1}. Choose an element $\zeta \in \mu_\infty$ whose order is divisible by two distinct and sufficiently large prime numbers, depending only on the roots of unity appearing in T, so that it is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to T. Then it is easy to see that $\alpha\zeta$ lies in Γ and that it is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D. Therefore we have $D = \alpha\zeta \cdot \zeta^{-1}T$, the translate of the torsion divisor $\zeta^{-1}T$ by the element $\alpha\zeta \in \Gamma$ that is S-integral on $\mathbb{G}_{\mathrm{m}}$ relative to D with $\lim_{n\to\infty} h(\gamma_n(\alpha\zeta)^{-1}) = \lim_{n\to\infty} h(\gamma_n\alpha^{-1}) = 0$, as desired.

Now we prove part (i). Let $n \geq 1$ be an integer. There exist infinitely many γ as in the hypothesis of (i) with $h(\gamma) \in I_n$. Thus already $D = \alpha T$ is as above. We use the fact that $|h(\gamma) - h(\alpha)| \leq h(\gamma\alpha^{-1})$, observed below Theorem 1.3, to get $h(\alpha) \in I_n$ from the already proved part (ii), since I_n is closed. As this holds for all n, we find $h(\alpha) \in I$. □

Here we mention the following slightly strengthened version of the first part of the theorem we just proved. It is easy to see the analogous strengthenings of the second part and of Theorem 1.6 by replacing the first $\mathbb{G}_{\mathrm{m}}$ with $\mathbb{P}^1$ in the statements too, the detail of which we omit here for simplicity.

Theorem 6.3. (*Keep all the notation K, S, Γ, h, $(I_n)_{n\geq 1}$, and I as in Theorem* 1.1(i).) *Suppose that D is a nonzero effective divisor on $\mathbb{G}_{\mathrm{m}}$. If $\{\gamma \in \mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \Gamma : h(\gamma) \in I_n\}$ is infinite for all integers $n \geq 1$, then D is the translate of a torsion divisor on $\mathbb{G}_{\mathrm{m}}$ by an element of $\{\alpha \in \mathbb{G}_{\mathrm{m},D}(\overline{\mathcal{O}}_{K,S}) \cap \Gamma : h(\alpha) \in I\}$.*

Note that this is nothing but the replacement of $\mathbb{G}_{\mathrm{m}}$ by $\mathbb{P}^1$ in the hypothesis of Theorem 1.1(i). We indicate now how to prove Theorem 6.3. As usual, we will enlarge K freely. Theorem 1.6 is invariant under the enlargement of S. So its conclusion holds true when all the points in the support of D are in K and are units at the places of K outside S. So we may replace $\mathbb{G}_{\mathrm{m}}$ by $\mathbb{P}^1$ in the hypothesis of Theorem 1.6. As in the proof of Theorem 1.1 we write $D = \alpha T$. By Proposition 6.2 we have $|\alpha|_v = 1$ for all places v of K outside the original S. So α is S-integral on $\mathbb{P}^1$ relative to $(0) + (\infty)$. This remains true after multiplying α by a root of unity. As in the proof of Theorem 1.1(ii) we deduce $\alpha \in \mathbb{G}_{\mathrm{m},D}(\overline{\mathcal{O}}_{K,S}) \cap \Gamma$ and as in the proof of Theorem 1.1(i) we find $h(\alpha) \in I$.

Besides, similarly to a remark below Theorem 1.1, this theorem is also equivalent to its apparently weaker version with the stronger assumption that $I_1 = I_2 = \cdots = I$.

7. The generality and the case of dynamical systems

7.1. **In pursuit of a general phenomenon.** Instead of repeating the details here, we refer the reader to [10, pp. 2011–2012] for the general definitions of *integral points* on an arbitrary variety and of *division groups, torsion divisors*, and the *translate of a divisor* by a point on a semi-abelian variety.

According to Chambert-Loir [7], we have a notion of a (*logarithmic*) *canonical height* on a semi-abelian variety A defined over a number field K, which will simply be called a height and used in our conjecture below with the general notation $\widehat{h}$. This height is defined not only for points but also for closed subvarieties, more generally for nonzero effective cycles of pure dimension, and it depends on additional data such as the choice of an appropriate line bundle. We do not go into such details here, however, except to say that this can have negative values and yet that it agrees with the standard notion of a canonical height when restricted to the case where A is either an abelian variety or $\mathbb{G}_{\mathrm{m}}^g$, g, an integer ≥ 1. For convenience we use additive notation for the group law of A in general.

Definition. Let Γ be a division group in $A(\overline{K})$, and let D be an effective divisor on A. For any subinterval J of $[0, \infty)$, we define

$$A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,J} := \{P \in \Gamma : P \text{ is } S\text{-integral on } A \text{ relative to } D \text{ and } \widehat{h}(P) \in J\}.$$

With Theorem 1.1 in mind one may naively wonder whether there could be something true more generally. Below are two candidates on the nondensity of integral division points on semi-abelian varieties that one may, a bit boldly, suspect to be true.

Suppose that D is a nonzero effective divisor on A.

- **(A)** Let $(I_n)_{n\geq 1}$ be a decreasing sequence of closed subintervals of $[0, \infty)$ with intersection $I = \bigcap_{n\geq 1} I_n$ such that in case of $\dim A \geq 2$, we assume $I \cap \widehat{h}(\Gamma) \neq \emptyset$, where $\widehat{h}(\Gamma) := \{\widehat{h}(\gamma) \in [0, \infty) : \gamma \in \Gamma\}$. If $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I_n}$ is Zariski dense in A for all integers $n \geq 1$, then D is the translate of a torsion divisor on A by an element of $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I}$.

(B) If $(\gamma_n)_{n\geq 1}$ is a generic sequence (meaning no infinite subsequence is contained in any proper Zariski-closed subset of A) of elements of Γ that are S-integral on A relative to D, then D is the translate of a torsion divisor T on A by some $\alpha \in \Gamma$ that is S-integral on A relative to D with $\limsup_{n\to\infty} \widehat{h}(\overline{-\gamma_n+\alpha+T}) \leq 0$, where $\overline{-\gamma_n+\alpha+T}$ is the divisor on the completion of A (used in [7]) obtained by taking the Zariski closures of the irreducible components of the divisor $-\gamma_n+\alpha+T$ without changing multiplicities.

Remarks.

(1) Suppose that the part "by an element of $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I}$" of **(A)** is replaced with "by an element of Γ with height in I", cf. Theorem 1.6. Then note that to establish this resulting weakened statement, it suffices to do so after enlarging S or extending K (and taking all the places lying above those of S), if desired. In particular, as noted above, the veracity of this weakened statement is independent of the choice of completion $\overline{A}$ of A or the choice of integral model for $\overline{A}$. However, the veracity of **(A)** itself (keeping the part "by an element of $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I}$") is independent of the choice of completion $\overline{A}$ of A or the choice of integral model for $\overline{A}$ up to an enlargement of S in general. A similar remark applies to **(B)** too.

In addition, let us explain the reason for the hypothesis that $I \cap \widehat{h}(\Gamma) \neq \emptyset$ in **(A)**. It could happen that $I_n \cap \widehat{h}(\Gamma) \neq \emptyset$ for all $n \geq 1$, while $I \cap \widehat{h}(\Gamma) = \emptyset$ (e.g., the obviously uninteresting case that $I := \{\lambda\}$, $\lambda \in [0,\infty) - \widehat{h}(\Gamma)$), in which case there exists a priori no element of $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I}$ that could be used to translate a torsion divisor and get D.

(2) Exactly as in [10, p. 2012, Remark (ii)], the veracity of the above weakened statement will be unaffected if we replace $A_D(\overline{\mathcal{O}}_{K,S})_{\Gamma,I_n} = \overline{A}_{\overline{\mathrm{Supp}D}\cup\partial A}(\overline{\mathcal{O}}_{K,S}) \cap \{P \in \Gamma : \widehat{h}(P) \in I_n\}$ by $\overline{A}_{\overline{\mathrm{Supp}D}}(\overline{\mathcal{O}}_{K,S}) \cap \{P \in \Gamma : \widehat{h}(P) \in I_n\}$ in its statement. However, the same unaffected veracity is valid for **(A)** (as it is) up to an enlargement of S in general. A similar remark applies to **(B)** too.

(3) It is possible to formulate **(B)** too in the fashion of **(A)** with $(I_n)_{n\geq 1}$ involved, but we prefer the version of **(B)**.

(4) For example, if A is most interestingly an abelian variety or $\mathbb{G}_{\mathrm{m}}^g$ (g, an integer ≥ 1) as a first step, then the inequality in **(B)** can equivalently be replaced by an equality with the replacement of lim sup by lim. We do not go into any further detail here. Another option would be to use the nonnegative height of subvarieties of semi-abelian varieties introduced by David and Philippon [8].

7.2. **The relationship with previous conjectures.** First, when Γ_0 is the trivial subgroup of $A(\overline{K})$, note that Γ is equal to $A(\overline{K})_{\mathrm{tor}} := \{\text{all torsion points of } A(\overline{K})\}$. Alternatively, in **(A)** we can set $I = I_1 = I_2 = \cdots = \{0\}$, in which case the points under consideration should be torsion. Hence in either of these special cases **(A)** (not only recovers, but also) refines [4, Conjecture 3.2].

Second, slightly more generally, we set $I_n = [0, 1/n]$ for any $n \geq 1$, in which case $I = \{0\}$. Then we have the following.

Conjecture 7.1 (= A special case of **(A)**)**.** (*Keep all the notation K, S, A, Γ, and $\widehat{h}$ as above.*) *Suppose that D is a nonzero effective divisor on A. If*

$$\{P \in \Gamma : P \text{ is } S\text{-integral on } A \text{ relative to } D \text{ and } \widehat{h}(P) \leq \epsilon\}$$

is Zariski dense in A for all real numbers $\epsilon > 0$, then D is the translate of a torsion divisor on A by an element of $A_D(\overline{\mathcal{O}}_{K,S}) \cap A(\overline{K})_{\mathrm{tor}}$.

It should be mentioned that Conjecture 7.1 has been formulated earlier by Grant and Ih, and independently also by Sookdeo and Tucker.

Third, if we set $I = I_1 = I_2 = \cdots = [0, \infty)$ in **(A)**, then we can immediately (not only recover, but also) refine [10, Conjecture 1.1].

Finally, we assume that $\dim A = 1$. Keep the hypotheses of Conjecture 7.1 except to suppose, now in contrast, that D is torsion. Then **(A)** suggests a new phenomenon complementing Conjecture 7.1 in the case of $\dim A = 1$, i.e., that the set

$$\{P \in \Gamma : P \text{ is } S\text{-integral on } A \text{ relative to } D \text{ and } \widehat{h}(P) \geq \epsilon\}$$

is finite for any real number $\epsilon > 0$. (Indeed, it suffices to set $I = I_1 = I_2 = \cdots = [\epsilon, \infty)$ in **(A)**.) Thanks to Theorem 1.1(i), this is now a theorem in the case of $A = \mathbb{G}_{\mathrm{m}}$, while the case of elliptic curves remains open.

7.3. **The case of dynamical systems.** As well as a number field K and a finite set S of places of K including all the infinite ones, we fix the following notation: a K-morphism $\varphi : \mathbb{P}^1 \to \mathbb{P}^1$ of finite degree ≥ 2, a Call–Silverman canonical height $\widehat{h}$ associated to φ, and a point Q_0 of $\mathbb{P}^1(\overline{K})$. Then we recall the exact notation and definitions of $[\varphi]$, Γ_0 attached to Q_0, and Γ from [10, p. 2032] without repeating them here. Here is a conjecture analogous to Theorem 1.1(i) for dynamical systems.

Conjecture 7.2. (*Keep the notation as above.*) *Let D be a nonzero effective divisor on $\mathbb{P}^1$. Suppose that $(I_n)_{n \geq 1}$ is a decreasing sequence of closed subintervals of $[0, \infty)$ with intersection I. Consider the following three statements.*

(i) *For any integer $n \geq 1$,*

$$\#\big(\mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \{P \in \Gamma : \widehat{h}(P) \in I_n\}\big) \;=\; \infty.$$

(ii) *We have*

$$\#\big(\mathbb{P}^1_D(\overline{\mathcal{O}}_{K,S}) \cap \{P \in \Gamma : \widehat{h}(P) \in I\}\big) \;=\; \infty.$$

(iii) *All the irreducible components of D belong to Γ and have height values in $I \cup \{0\}$ at most one of which may be different from 0, and furthermore the irreducible components, if any, of D that are not exceptional for any $\phi \in [\varphi]$ have the same height value in I.*

Then the implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) *hold.*

See [19, p. 807 (and 798)] (also [20, pp. 16–17]) for the term "exceptional" in detail. In the special case that $I_n = [0, 1/n]$ for all $n \geq 1$, (the contrapositive of) the implication (i) $\Rightarrow$ (iii) gives rise to the following conjecture that has been formulated earlier by Grant and Ih, and independently also by Sookdeo and Tucker.

Besides, similar to a remark below Theorem 1.1, the conjecture (i) $\Rightarrow$ (iii) is equivalent to the apparently weaker conjecture (ii) $\Rightarrow$ (iii).

Conjecture 7.3 (= A special case of (i) $\Rightarrow$ (iii) of Conjecture 7.2)**.** (*Keep the notation as above.*) *If $Q \in \mathbb{P}^1(\overline{K}) - \mathbb{P}^1(\overline{K})_{\varphi\text{-preper}}$, then*

$$\#\big(\mathbb{P}^1_{(Q)}(\overline{\mathcal{O}}_{K,S}) \cap \{P \in \Gamma : \widehat{h}(P) \leq \epsilon\}\big) \;<\; \infty,$$

i.e., there are only finitely many points in Γ with height $\leq \epsilon$ which are S-integral on $\mathbb{P}^1$ relative to (Q), for some fixed real number $\epsilon > 0$.

Here $\mathbb{P}^1(\overline{K})_{\varphi\text{-preper}}$ is the set of all φ-preperiodic points of $\mathbb{P}^1(\overline{K})$, i.e., of those points whose forward orbits under φ are finite sets. Note that the implication of (i) $\Rightarrow$ (iii) of Conjecture 7.2 actually enables us to take ϵ to be any positive real number strictly less than $\widehat{h}(Q)$. All the exact same comments (including the possible extensions and restrictions of the definitions of Γ_0 and Γ) as in [10, Section 4.2] apply to this context, but we do not repeat them here. One thing to add though is that, e.g., in case of using the ones in [10, pp. 2032–2033, (i)–(ii)], we replace the part "any $\phi \in [\varphi]$" by "$\phi_n \circ \cdots \circ \phi_1$ with any $\phi_1, \ldots, \phi_n \in [\varphi]$" in Conjecture 7.2(iii) if $[\varphi]$ need not be closed under composition. Finally, see [14] for an application of Theorem 1.1 in relation to arithmetical dynamical systems.

Acknowledgments

The first author thanks the Institute for Advanced Study, and the second author thanks the Korea Institute for Advanced Study and Princeton University for their hospitality during their stays at the respective institutions. Both authors thank David Grant, Vijay Sookdeo, Tom Tucker, Bob Vaughan, and Shou-Wu Zhang for useful discussions too. They also thank the referee for various inquiries and useful suggestions for clarification of parts of the article.

References

[1] F. Amoroso and U. Zannier, *A relative Dobrowolski lower bound over abelian extensions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **29** (2000), no. 3, 711–727. MR1817715

[2] P. Autissier, *Sur une question d'équirépartition de nombres algébriques* (French, with English and French summaries), C. R. Math. Acad. Sci. Paris **342** (2006), no. 9, 639–641, DOI 10.1016/j.crma.2006.02.021. MR2225867

[3] A. Baker and G. Wüstholz, *Logarithmic forms and group varieties*, J. Reine Angew. Math. **442** (1993), 19–62, DOI 10.1515/crll.1993.442.19. MR1234835

[4] M. Baker, S. Ih, and R. Rumely, *A finiteness property of torsion points*, Algebra Number Theory **2** (2008), no. 2, 217–248, DOI 10.2140/ant.2008.2.217. MR2377370

[5] Y. Bilu, *Limit distribution of small points on algebraic tori*, Duke Math. J. **89** (1997), no. 3, 465–476, DOI 10.1215/S0012-7094-97-08921-3. MR1470340

[6] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006. MR2216774

[7] A. Chambert-Loir, *Points de petite hauteur sur les variétés semi-abéliennes* (French, with English and French summaries), Ann. Sci. École Norm. Sup. (4) **33** (2000), no. 6, 789–821, DOI 10.1016/S0012-9593(00)01053-3. MR1832991

[8] S. David and P. Philippon, *Sous-variétés de torsion des variétés semi-abéliennes* (French, with English and French summaries), C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), no. 8, 587–592, DOI 10.1016/S0764-4442(00)01634-7. MR1799094

[9] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*, Acta Arith. **34** (1979), no. 4, 391–401. MR543210

[10] D. Grant and S. Ih, *Integral division points on curves*, Compos. Math. **149** (2013), no. 12, 2011–2035, DOI 10.1112/S0010437X13007318. MR3143704

[11] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6th ed., Oxford University Press, Oxford, 2008. Revised by D. R. Heath-Brown and J. H. Silverman; With a foreword by Andrew Wiles. MR2445243

[12] G. Harman, *Metric number theory*, London Mathematical Society Monographs. New Series, vol. 18, The Clarendon Press, Oxford University Press, New York, 1998. MR1672558

[13] M. Hindry and J. H. Silverman, *Diophantine geometry: An introduction*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000. MR1745599

[14] S. Ih, *Integral points on the Chebyshev dynamical systems*, J. Korean Math. Soc. **52** (2015), no. 5, 955–964, DOI 10.4134/JKMS.2015.52.5.955. MR3393112

[15] J. Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher; With a foreword by G. Harder. MR1697859

[16] H. P. Schlickewei, *Lower bounds for heights on finitely generated groups*, Monatsh. Math. **123** (1997), no. 2, 171–178, DOI 10.1007/BF01305970. MR1430503

[17] H. P. Schlickewei and W. M. Schmidt, *On polynomial-exponential equations*, Math. Ann. **296** (1993), no. 2, 339–361, DOI 10.1007/BF01445109. MR1219906

[18] W. M. Schmidt, *Diophantine approximations and Diophantine equations*, Lecture Notes in Mathematics, vol. 1467, Springer-Verlag, Berlin, 1991. MR1176315

[19] J. H. Silverman, *Integer points, Diophantine approximation, and iteration of rational maps*, Duke Math. J. **71** (1993), no. 3, 793–829, DOI 10.1215/S0012-7094-93-07129-3. MR1240603

[20] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics, vol. 241, Springer, New York, 2007. MR2316407

[21] M. Waldschmidt, *Diophantine approximation on linear algebraic groups*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 326, Springer-Verlag, Berlin, 2000. Transcendence properties of the exponential function in several variables. MR1756786

[22] K. R. Yu, *Linear forms in p-adic logarithms. III*, Compositio Math. **91** (1994), no. 3, 241–276. MR1273651

[23] F. Zorzitto, *Discretely normed abelian groups*, Aequationes Math. **29** (1985), no. 2-3, 172–174, DOI 10.1007/BF02189825. MR819306

Departement Mathematik und Informatik, Fachbereich Mathematik, University of Basel, Spiegelgasse 1, 4051 Basel, Switzerland

Email address: philipp.habegger@unibas.ch

Department of Mathematics, University of Colorado at Boulder, Boulder, Colorado 80309-0395

Email address: ih@math.colorado.edu