NORMS EXTREMAL WITH RESPECT TO THE MAHLER MEASURE

PAUL FILI AND ZACHARY MINER

Abstract. In this paper, we introduce and study several norms which are constructed in order to satisfy an extremal property with respect to the Mahler measure. These norms are a natural generalization of the metric Mahler measure introduced by Dubickas and Smyth. We show that bounding these norms on a certain subspace implies Lehmer's conjecture and in at least one case that the converse is true as well. We evaluate these norms on a class of algebraic numbers that include Pisot and Salem numbers, and for surds. We prove that the infimum in the construction is achieved in a certain finite dimensional space for all algebraic numbers in one case, and for surds in general, a finiteness result analogous to that of Samuels and Jankauskas for the t-metric Mahler measures.

1. Introduction

1.1. **Background.** Let K be a number field with set of places M_K. For each $v \in M_K$ lying over a rational prime p, let $\|\cdot\|_v$ be the absolute value on K extending the usual p-adic absolute value on $\mathbb{Q}$ if v is finite or the usual archimedean absolute value if v is infinite. Then for $\alpha \in K^\times$, the absolute logarithmic Weil height h is given by

$$h(\alpha) = \sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \log^+ \|\alpha\|_v$$

where $\log^+ t = \max\{\log t, 0\}$. As the right hand side above does not depend on the choice of field K containing α, h is a well-defined function mapping $\overline{\mathbb{Q}}^\times \to [0, \infty)$ which vanishes precisely on the roots of unity $\operatorname{Tor}(\overline{\mathbb{Q}}^\times)$. Related to the Weil height is the logarithmic *Mahler measure*, given by

$$m(\alpha) = (\deg \alpha) \cdot h(\alpha),$$

where $\deg \alpha = [\mathbb{Q}(\alpha) : \mathbb{Q}]$. Perhaps the most important open question regarding the Mahler measure is *Lehmer's conjecture* that there exists an absolute constant c such that

$$m(\alpha) \geq c > 0 \quad \text{for all} \quad \alpha \in \overline{\mathbb{Q}}^\times \setminus \operatorname{Tor}(\overline{\mathbb{Q}}^\times). \tag{1.1}$$

The question of the existence of algebraic numbers with small Mahler measure was first posed in 1933 by D.H. Lehmer [8]. The current best known lower bound, due to Dobrowolski [4], is of the form

$$m(\alpha) \gg \left(\frac{\log \log \deg \alpha}{\log \deg \alpha} \right)^3 \quad \text{for all} \quad \alpha \in \overline{\mathbb{Q}}^\times \setminus \operatorname{Tor}(\overline{\mathbb{Q}}^\times)$$

where the implied constant is absolute.

Date: May 4, 2011.
2000 *Mathematics Subject Classification.* 11R04.
Key words and phrases. Weil height, Mahler measure, Lehmer's problem.

The Weil height h naturally satisfies the conditions of being a metric on the space

$$\mathcal{G} = \overline{\mathbb{Q}}^{\times} / \operatorname{Tor}(\overline{\mathbb{Q}}^{\times})$$

of algebraic numbers modulo torsion, and in fact, viewing $\mathcal{G}$ as a vector space over $\mathbb{Q}$ written multiplicatively (see the paper of Allcock and Vaaler [1]), it is easy to see that h is a vector space norm. The study of the Mahler measure on the vector space of algebraic numbers modulo torsion presents several difficulties absent for the Weil height, first of which is that while m also vanishes precisely on $\operatorname{Tor}(\overline{\mathbb{Q}}^{\times})$, unlike h, it is not well-defined on the quotient space modulo torsion. To get around that difficulty, Dubickas and Smyth [6] first introduced the *metric Mahler measure*, which gave a well-defined metric on $\mathcal{G}$ satisfying the additional property of being the largest metric which descends from a function bounded above by the Mahler measure on $\overline{\mathbb{Q}}^{\times}$. Later, the first author and Samuels [10, 12] defined the *ultrametric Mahler measure* which satisfies the strong triangle inequality and gives a projective height on $\mathcal{G}$. It is easy to see that the metric and ultrametric Mahler measures each induce the discrete topology on $\mathcal{G}$ if and only if Lehmer's conjecture is true.

In this paper we will introduce vector space norms on $\mathcal{G}$ which satisfy an analogous extremal property with respect to the Mahler measure as the metric Mahler measure does. Before presenting our constructions, let us fix our notation. We denote the L^p Weil heights for $1 \leq p < \infty$ by

$$h_p(\alpha) = \left(\sum_{v \in M_K} \frac{[K_v : \mathbb{Q}_v]}{[K : \mathbb{Q}]} \cdot |\log \|\alpha\|_v|^p \right)^{1/p} \quad \text{for} \quad \alpha \in K^{\times},$$

noting that the classical Weil height satisfies $2h = h_1$ (see [1]) and is well-defined, independent of choice of K. For $p = \infty$, we let

$$h_\infty(\alpha) = \sup_{v \in M_K} |\log \|\alpha\|_v| \quad \text{for} \quad \alpha \in K^{\times},$$

noting that this height serves as a generalization of the (logarithmic) house of an algebraic integer (see Section 1.3 below for more details). Analogously, we have the L^p Mahler measure defined on $\overline{\mathbb{Q}}$ to be $m_p(\alpha) = (\deg \alpha) \cdot h_p(\alpha)$, where $m_1 = 2m$ is twice the usual Mahler measure.

For an algebraic number $\alpha \in \overline{\mathbb{Q}}^{\times}$, we let $\boldsymbol{\alpha} \in \mathcal{G}$ denote its equivalence class modulo torsion. Let $d : \mathcal{G} \to \mathbb{N}$ be given by

$$d(\boldsymbol{\alpha}) = \min_{\zeta \in \operatorname{Tor}(\overline{\mathbb{Q}}^{\times})} \deg \zeta\alpha,$$

where $\zeta\alpha$ ranges over all representatives of the equivalence class $\boldsymbol{\alpha}$. The *minimal logarithmic Mahler measure* is defined to be the function $m : \mathcal{G} \to [0, \infty)$ given by[1]

$$m(\boldsymbol{\alpha}) = d(\boldsymbol{\alpha}) h(\boldsymbol{\alpha}).$$

(Recall that h_p is constant on cosets modulo torsion, so that $h_p(\boldsymbol{\alpha}) = h_p(\alpha)$ for all $\alpha \in \overline{\mathbb{Q}}^{\times}$.) More generally, we define the *minimal logarithmic L^p Mahler measure*

$$m_p(\boldsymbol{\alpha}) = d(\boldsymbol{\alpha}) h_p(\boldsymbol{\alpha}).$$

[1] The usual Mahler measure is not defined on $\mathcal{G}$, thus our use of m for the minimal Mahler measure should result in no confusion.

This function is called minimal because it yields, for any element $\boldsymbol{\alpha} \in \mathcal{G}$, the minimal logarithmic L^p Mahler measure amongst all of the representatives in $\overline{\mathbb{Q}}^\times$ of our $\boldsymbol{\alpha} \in \mathcal{G}$, that is,

$$m_p(\boldsymbol{\alpha}) = \min_{\zeta \in \mathrm{Tor}(\overline{\mathbb{Q}}^\times)} m_p(\alpha\zeta).$$

Let us now recall the construction of the metric Mahler measure $\widehat{m} : \mathcal{G} \to [0, \infty)$ of Dubickas and Smyth [6]. This construction may be applied to any height function as in [7] and will in general produce metrics defined on $\mathcal{G}$ modulo its zero set. The (logarithmic) metric Mahler measure is defined by

$$\widehat{m}(\boldsymbol{\alpha}) = \inf_{\alpha = \alpha_1 \cdots \alpha_n} \sum_{i=1}^{n} m(\alpha_i),$$

where the infimum is taken over all possible ways of writing any representative of $\boldsymbol{\alpha}$ as a product of other algebraic numbers. This construction is extremal in the sense that any other function $g : \mathcal{G} \to [0, \infty)$ satisfying

(1) $g(\boldsymbol{\alpha}) \leq m(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$, and
(2) $g(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) \leq g(\boldsymbol{\alpha}) + g(\boldsymbol{\beta})$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{G}$, (*Triangle Inequality*)

is then smaller than $\widehat{m}$, that is, $g(\boldsymbol{\alpha}) \leq \widehat{m}(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$. Equivalently, lifting to $\overline{\mathbb{Q}}^\times$ in the natural way, it is easy to see that $\widehat{m}$ satisfies the same extremal property with respect to the logarithmic Mahler measure. This extremal property is characteristic of the metric construction for height functions [6, 7, 10].

1.2. **Main results.** The space $\mathcal{G}$ has a vector space structure over $\mathbb{Q}$ (written multiplicatively), so we might ask if there exists a vector space norm satisfying the same extremal property with respect to the Mahler measure. We define the *extremal norm* $\widetilde{m}_p$ associated to m_p to be:

$$\widetilde{m}_p(\boldsymbol{\alpha}) = \inf_{\boldsymbol{\alpha} = \boldsymbol{\alpha}_1^{r_1} \cdots \boldsymbol{\alpha}_n^{r_n}} \sum_{i=1}^{n} |r_i| m_p(\boldsymbol{\alpha}_i),$$

where the infimum is taken over all ways of writing $\boldsymbol{\alpha}$ as a linear combination of vectors $\boldsymbol{\alpha}_i \in \mathcal{G}$ with $r_i \in \mathbb{Q}$. (Observe that $m_p(\boldsymbol{\alpha}^r) \neq |r| m_p(\boldsymbol{\alpha})$ for general $\boldsymbol{\alpha} \in \mathcal{G}$ and $r \in \mathbb{Q}^\times$, so that in general the metric construction $\widehat{m}_p$ and the norm construction $\widetilde{m}_p$ will not agree.) We prove that $\widetilde{m}_p$ is a well-defined vector space norm on $\mathcal{G}$ which is extremal amongst all seminorms with respect to the Mahler measure, in the sense that if $g : \mathcal{G} \to [0, \infty)$ is a function satisfying

(1) $g(\boldsymbol{\alpha}) \leq m_p(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$,
(2) $g(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) \leq g(\boldsymbol{\alpha}) + g(\boldsymbol{\beta})$, and
(3) $g(\boldsymbol{\alpha}^r) = |r| g(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}, r \in \mathbb{Q}$,

then $g \leq \widetilde{m}_p$, that is, $g(\boldsymbol{\alpha}) \leq \widetilde{m}_p(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$.

Our main result is a finiteness theorem for the extremal norm $\widetilde{m}_1$ analogous to the main result of [12] for the infimum of the metric Mahler measure. Let K be a number field and let

$$V_K = \{\boldsymbol{\alpha}^r : r \in \mathbb{Q} \text{ and } \boldsymbol{\alpha} \in K^\times / \mathrm{Tor}(K^\times)\}$$

be the vector subspace inside $\mathcal{G}$ spanned by elements of $K^\times/\operatorname{Tor}(K^\times)$. Notice that $\boldsymbol{\alpha} \in V_K$ if and only if for any coset representative $\alpha \in \overline{\mathbb{Q}}^\times$ we have $\alpha^n \in K^\times$ for some $n \in \mathbb{N}$. Let M_K be the set of places of K, and let $S \subset M_K$ be a finite set of places of K, including all archimedean places. Then for any field extension L/K, define

$$V_{L,S} = \{\boldsymbol{\alpha} \in V_L : \|\boldsymbol{\alpha}\|_w = 1 \text{ for } w \mid v \in M_K \setminus S\}.$$

Observe that by Dirichlet's S-unit theorem, $V_{L,S}$ is a finite dimensional vector space. Our main result is the following:

Theorem 1. *Let $\boldsymbol{\alpha} \in V_K$, where K is Galois. Then there exists a finite set of rational primes S, containing the archimedean place, such that*

$$\widetilde{m}_1(\boldsymbol{\alpha}) = \sum_{F \subseteq K} [F : \mathbb{Q}] \cdot h_1(\boldsymbol{\alpha}_F)$$

where $\boldsymbol{\alpha}_F \in \overline{V_{F,S}}$, $\boldsymbol{\alpha} = \prod_{F \subseteq K} \boldsymbol{\alpha}_F$. Furthermore, for each pair of fields $E \subset F \subseteq K$,

$$h_1(\boldsymbol{\alpha}_F) = \inf_{\boldsymbol{\beta} \in V_{E,S}} h_1(\boldsymbol{\alpha}_F/\boldsymbol{\beta}).$$

We may interpret the last sentence of Theorem 1 as saying the norm of each $\boldsymbol{\alpha}_F$ is equal to the quotient norm of $\boldsymbol{\alpha}_F$ with respect to any subfield.

For $p > 1$ we are able to show that $\widetilde{m}_p$ attains its infimum on roots of rationals by computing $\widetilde{m}_p(\boldsymbol{\alpha}) = h_p(\boldsymbol{\alpha})$ directly for such numbers in Proposition 3.7. (See Samuels and Jankauskas [13] for analogous results on the p-metric construction.[2])

In order to prove our main result on the infimum of $\widetilde{m}_1$, we prove in Section 4.1 several results related to heights of algebraic numbers modulo multiplicative group actions very much related to the results of A.C. de la Maza and E. Friedman [11], which we interpret as results about quotient norms. In particular, for L/K and $\boldsymbol{\alpha} \in V_{L,S}$, we use essentially the same proof used in [11] to show in Theorem 4, that the infimum $\inf_{\boldsymbol{\beta} \in V_{K,S}} h(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})$ is attained in the closure $\overline{V_{K,S}}$. In Theorem 5, we show that under certain extra conditions we may find the infimum $\inf_{\boldsymbol{\beta} \in V_{K,S}} h(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})$ within $V_{K,S}$, extending a result of [11]. In our final result on quotient norms, Theorem 6, we show that for $\boldsymbol{\alpha} \in V_{L,S}$ we can find an element $\boldsymbol{\eta} \in \overline{V_{K,S}}$ of minimal height which satisfies both

(1) $h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) = \inf_{\boldsymbol{\beta} \in V_{K,S}} h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})$, and
(2) $h_1(\boldsymbol{\eta}) + [L : K] h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) = \inf_{\boldsymbol{\beta} \in V_{K,S}} \big(h_1(\boldsymbol{\beta}) + [L : K] h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})\big)$.

We then construct an S-unit projection which allows us to reduce to finite dimensions, which we believe is new and of interest in itself as it is a nonincreasing map with respect to the height.

[2]The p-metric construction studied in [13] is very different from our metric construction $f \mapsto \widehat{f}$. The p-metric Mahler measure $\mathcal{M}_p$ is the infimum over all representations $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_n$ of the ℓ^p norm of the vector $(m(\boldsymbol{\alpha}_1), \ldots, m(\boldsymbol{\alpha}_n))$. One notable difference between the two constructions is that $\widehat{m}_p$ satisfies the triangle inequality $\widehat{m}_p(\boldsymbol{\alpha}\boldsymbol{\beta}) \le \widehat{m}_p(\boldsymbol{\alpha}) + \widehat{m}_p(\boldsymbol{\beta})$, while $\mathcal{M}_p$ satisfies a p-metric triangle inequality $\mathcal{M}_p(\boldsymbol{\alpha}\boldsymbol{\beta})^p \le \mathcal{M}_p(\boldsymbol{\alpha})^p + \mathcal{M}_p(\boldsymbol{\beta})^p$. However, when $p = 1$, they are essentially the same: $\widehat{m}_1(\boldsymbol{\alpha}) = \mathcal{M}_1(\boldsymbol{\alpha})$.

1.3. **Applications to Lehmer's problem.** Given that the norms $\widetilde{m}_p$ are extremal with respect to the Mahler measure, it is natural to ask what applications these norms have to the Lehmer problem. Define $\mathcal{A} \subset \mathcal{G}$ to be the set of $\mathbf{1} \neq \boldsymbol{\alpha} \in \mathcal{G}$ which have a representative α satisfying the following properties:

(1) α is an algebraic unit.
(2) $[\mathbb{Q}(\alpha^n) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ for all $n \in \mathbb{N}$.
(3) For any proper subfield F of $K = \mathbb{Q}(\alpha)$, $\mathrm{Norm}_F^K(\alpha) \in \mathrm{Tor}(F^\times)$.

The conditions of the set $\mathcal{A}$ are exactly, in the terminology of [9], that $\boldsymbol{\alpha}$ be a unit, representable[3], and projection irreducible, respectively. Then in [9, Theorem 4] it is proven that for any $1 \leq p \leq \infty$ there exists a constant c_p such that

$$m_p(\alpha) = (\deg \alpha) \cdot h_p(\alpha) \geq c_p > 0 \quad \text{for all} \quad \alpha \in \overline{\mathbb{Q}}^\times \setminus \mathrm{Tor}(\overline{\mathbb{Q}}^\times) \tag{1.2}$$

if and only if

$$m_p(\boldsymbol{\alpha}) = d(\boldsymbol{\alpha}) \cdot h_p(\boldsymbol{\alpha}) \geq c_p > 0 \quad \text{for all} \quad \boldsymbol{\alpha} \in \mathcal{A}. \tag{1.3}$$

We note that equation (1.2) is equivalent to the Lehmer conjecture for $p = 1$ and the Schinzel-Zassenhaus conjecture for $p = \infty$ [9, Proposition 4.1]. Therefore we formulate the following conjecture:

Conjecture 1. *For each $1 \leq p \leq \infty$, there exists a constant c_p such that*

$$\widetilde{m}_p(\boldsymbol{\alpha}) \geq c_p > 0 \quad \textit{for all} \quad \boldsymbol{\alpha} \in \mathcal{A}. \tag{1.4}$$

Theorem 2. *If Conjecture 1 is true, then* (1.2) *holds.*

In particular, for $p = 1$ (1.4) implies that Lehmer's conjecture is true, and for $p = \infty$ equation (1.4) implies that the Schinzel-Zassenhaus conjecture is true.

For $p \neq 2$, we are unable to prove the converse to Theorem 2. We nevertheless expect the result is true and make the following conjecture:

Conjecture 2. *If* (1.2) *holds, then Conjecture 1 is true.*

However, when $p = 2$ we are able to prove that:

Theorem 3. *There exists a constant c_2 such that*

$$m_2(\alpha) = (\deg \alpha) \cdot h_2(\alpha) \geq c_2 > 0 \quad \textit{for all} \quad \alpha \in \overline{\mathbb{Q}}^\times \setminus \mathrm{Tor}(\overline{\mathbb{Q}}^\times)$$

if and only if

$$\widetilde{m}_2(\boldsymbol{\alpha}) \geq c_2 > 0 \quad \textit{for all} \quad \boldsymbol{\alpha} \in \mathcal{A}.$$

Proof. In [9], we construct a norm $\|\cdot\|_{m,2}$, and prove in [9, Theorem 4] that bounding $\|\cdot\|_{m,2}$ away from zero on $\mathcal{A}$ is equivalent to bounding m_2 away from zero on $\overline{\mathbb{Q}}^\times \setminus \mathrm{Tor}(\overline{\mathbb{Q}}^\times)$. Further, in [9, Theorem 6 et seq.] we prove that $\|\boldsymbol{\alpha}\|_{m,2} \leq m_2(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$. It follows by the extremal property for $\widetilde{m}_2$ that

$$\|\boldsymbol{\alpha}\|_{m,2} \leq \widetilde{m}_2(\boldsymbol{\alpha}) \leq m_2(\boldsymbol{\alpha})$$

for all $\boldsymbol{\alpha} \in \mathcal{G}$, and the claim now follows. □

[3]Such numbers are called *Lehmer irreducible* in earlier drafts.

The format of this paper is as follows. In Section 2 we prove basic results about degree functions on $\mathcal{G}$ and projections onto subspaces. In Section 3 we construct the extremal norms $\widetilde{m}_p$ arrived at by the infimum process, prove a very useful alternative formulation of $\widetilde{m}_p$ in Proposition 3.6, and examine explicit classes of algebraic numbers for which we can compute the value of the norms (for example, on surds and in the $p = 1$ case on Salem and Pisot numbers). Lastly in Section 4 we study $\widetilde{m}_1$ in particular and prove our main result that for any given class of an algebraic number the infimum in the construction of $\widetilde{m}_1$ is attained in a finite dimensional vector space.

The authors would like to acknowledge Jeffrey Vaaler for many helpful conversations in general, and specifically for his contributions to Lemma 4.3, as well as Clayton Petsche and Felipe Voloch for helpful remarks regarding this same lemma. We also thank the referee of this paper for many helpful suggestions.

2. Preliminary Lemmas

2.1. **Subspaces associated to number fields.** We will now prove some lemmas regarding the relationship between certain subspaces determined by number fields. Let $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and let us define

$$\mathcal{K} = \{K/\mathbb{Q} : [K : \mathbb{Q}] < \infty\} \quad \text{and} \quad \mathcal{K}^G = \{K \in \mathcal{K} : \sigma K = K \ \forall \sigma \in G\}.$$

Let us briefly recall the combinatorial properties of the sets $\mathcal{K}$ and $\mathcal{K}^G$ partially ordered by inclusion. Recall that $\mathcal{K}$ and $\mathcal{K}^G$ are *lattices*, that is, partially ordered sets for which any two elements have a unique greatest lower bound, called the *meet*, and a least upper bound, called the *join*. Specifically, for any two fields K, L, the meet $K \wedge L$ is given by $K \cap L$ and the join $K \vee L$ is given by KL. If K, L are Galois then both the meet (the intersection) and the join (the compositum) are Galois as well, thus $\mathcal{K}^G$ is also a lattice. Both lattices have a minimal element, namely $\mathbb{Q}$, and are *locally finite*, that is, between any two fixed elements we have a finite number of intermediate elements.

For each $K \in \mathcal{K}$, let

$$V_K = \{\boldsymbol{\alpha}^r : r \in \mathbb{Q} \text{ and } \boldsymbol{\alpha} \in K^\times / \operatorname{Tor}(K^\times)\}.$$

Then V_K is the subspace of $\mathcal{G}$ spanned by elements of $K^\times/\operatorname{Tor}(K^\times)$. We call a subspace of the form V_K for $K \in \mathcal{K}$ a *distinguished* subspace. Suppose we fix an algebraic number $\boldsymbol{\alpha} \in \mathcal{G}$. Then the set

$$\{K \in \mathcal{K} : \boldsymbol{\alpha} \in V_K\}$$

forms a sublattice of $\mathcal{K}$, and by the finiteness properties of $\mathcal{K}$ this set must contain a unique minimal element.

Definition 2.1. For any $\boldsymbol{\alpha} \in \mathcal{G}$, the *minimal field* is defined to be the minimal element of the set $\{K \in \mathcal{K} : \boldsymbol{\alpha} \in V_K\}$. We denote the minimal field of $\boldsymbol{\alpha}$ by $K_{\boldsymbol{\alpha}}$.

Note that the action of $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $\mathcal{G}$ is well-defined (see [1]).

Lemma 2.2. *For any $\boldsymbol{\alpha} \in \mathcal{G}$, we have* $\mathrm{Stab}_G(\boldsymbol{\alpha}) = \mathrm{Gal}(\overline{\mathbb{Q}}/K_{\boldsymbol{\alpha}}) \leq G$.

Notation 2.3. By $\mathrm{Stab}_G(\boldsymbol{\alpha})$ we mean the $\sigma \in G$ such that $\sigma\boldsymbol{\alpha} = \boldsymbol{\alpha}$. As this tacit identification is convenient we shall use it throughout with no further comment.

Proof. Clearly $\mathrm{Gal}(\overline{\mathbb{Q}}/K_{\boldsymbol{\alpha}}) \le \mathrm{Stab}_G(\boldsymbol{\alpha})$, as $\alpha^\ell \in K_{\boldsymbol{\alpha}}$ for some $\ell \in \mathbb{N}$ by definition of $V_{K_{\boldsymbol{\alpha}}}$. To see the reverse containment, observe that $K_{\boldsymbol{\alpha}} = \mathbb{Q}(\alpha^\ell)$ for some $\ell \in \mathbb{N}$. Now, for $\sigma \in \mathrm{Stab}_G(\boldsymbol{\alpha})$, we have $\sigma\alpha = \zeta\alpha$ for $\zeta \in \mathrm{Tor}(\mathbb{Q}^\times)$. Then if $\sigma(\alpha^\ell) \ne \alpha^\ell$, there would exist an $m \in \mathbb{N}$ such that $\sigma(\alpha^{\ell m}) = \alpha^{\ell m}$. Thus, σ is contained in a proper supergroup of $\mathrm{Gal}(\overline{\mathbb{Q}}/K_{\boldsymbol{\alpha}})$, so there would be a proper subfield of $K_{\boldsymbol{\alpha}}$ containing $\alpha^{\ell m}$, contradicting the definition of $K_{\boldsymbol{\alpha}}$. □

2.2. **Representability.** Observe that the action of the absolute Galois group $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is well-defined on the vector space of algebraic numbers modulo torsion $\mathcal{G}$, and in fact it is easy to see that each Galois automorphism gives rise to a distinct isometry of $\mathcal{G}$ in the h_p norm (see [9, §2.1] for more details). Let us denote the image of the class $\boldsymbol{\alpha}$ under $\sigma \in G$ by $\sigma\boldsymbol{\alpha}$. In order to associate a notion of degree to a subspace in a meaningful fashion so that we can define our norms associated to the Mahler Measure we define the function $\delta : \mathcal{G} \to \mathbb{N}$ by

$$\delta(\boldsymbol{\alpha}) = \#\{\sigma\boldsymbol{\alpha} : \sigma \in G\} = [G : \mathrm{Stab}_G(\boldsymbol{\alpha})] = [K_{\boldsymbol{\alpha}} : \mathbb{Q}] \tag{2.1}$$

to be the size of the orbit of $\boldsymbol{\alpha}$ under the Galois action, with the last equality above following from Lemma 2.2.

Observe that since taking roots or powers does not affect the $\mathbb{Q}$-vector space span, and in particular the minimal field $K_{\boldsymbol{\alpha}}$, the function δ is invariant under nonzero scaling in $\mathcal{G}$, that is, $\delta(\boldsymbol{\alpha}^r) = \delta(\boldsymbol{\alpha})$ for all $0 \ne r \in \mathbb{Q}$. In order to better understand the relationship between our elements in $\mathcal{G}$ and their representatives in $\overline{\mathbb{Q}}^\times$, we need to understand when an $\boldsymbol{\alpha} \in V_K$ has a representative $\alpha \in K^\times$ (or is merely a root of an element $\alpha^n \in K^\times$ for some $n > 1$). Naturally, the choice of coset representative modulo torsion affects this, and we would like to avoid such considerations. Therefore we define the function $d : \mathcal{G} \to \mathbb{N}$ by

$$d(\boldsymbol{\alpha}) = \min\{\deg \zeta\alpha : \alpha \in \overline{\mathbb{Q}}^\times,\ \zeta \in \mathrm{Tor}(\overline{\mathbb{Q}}^\times)\}. \tag{2.2}$$

In other words, for a given $\boldsymbol{\alpha} \in \mathcal{G}$, which is an equivalence class of an algebraic number modulo torsion, $d(\boldsymbol{\alpha})$ gives us the minimum degree amongst all of the coset representatives in $\overline{\mathbb{Q}}^\times$ modulo the torsion subgroup.

A number $\boldsymbol{\alpha} \in \mathcal{G}$ can then be represented by an algebraic number in $K_{\boldsymbol{\alpha}}^\times$ if and only if $d(\boldsymbol{\alpha}) = \delta(\boldsymbol{\alpha})$. We therefore make the following definition:

Definition 2.4. We define the set of *representable* elements of $\mathcal{G}$ to be the set

$$\mathcal{R} = \{\boldsymbol{\alpha} \in \mathcal{G} : \delta(\boldsymbol{\alpha}) = d(\boldsymbol{\alpha})\}. \tag{2.3}$$

The set $\mathcal{R}$ consists precisely of the $\boldsymbol{\alpha} \in \mathcal{G}$ that can be represented by some $\alpha \in \overline{\mathbb{Q}}^\times$ of degree equal to the degree of the minimal field $K_{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha}$.

We recall the terminology from [5] that a number $\alpha \in \overline{\mathbb{Q}}^\times$ is *torsion-free* if $\alpha/\sigma\alpha \notin \mathrm{Tor}(\overline{\mathbb{Q}}^\times)$ for all distinct Galois conjugates $\sigma\alpha$. Thus, torsion-free numbers give rise to distinct elements $\sigma\boldsymbol{\alpha} \in \mathcal{G}$ for each distinct Galois conjugate $\sigma\alpha$ of α in $\overline{\mathbb{Q}}$.

Lemma 2.5. *We have the following:*

1. *For each $\boldsymbol{\alpha} \in \mathcal{G}$, there is a unique minimal exponent $\ell(\boldsymbol{\alpha}) \in \mathbb{N}$ such that $\boldsymbol{\alpha}^{\ell(\boldsymbol{\alpha})} \in \mathcal{R}$.*
2. *For any $\alpha \in \overline{\mathbb{Q}}^\times$, we have $\delta(\boldsymbol{\alpha}) \mid \deg \alpha$.*

(3) *$\boldsymbol{\alpha} \in \mathcal{R}$ if and only if it has a representative in $\overline{\mathbb{Q}}^\times$ which is torsion-free.*

Proof. For $\boldsymbol{\alpha} \in \mathcal{G}$, choose a representative $\alpha \in \overline{\mathbb{Q}}^\times$ and let

$$\ell = \operatorname{lcm}\{\operatorname{ord}(\alpha/\sigma\alpha) : \sigma \in G \text{ and } \alpha/\sigma\alpha \in \operatorname{Tor}(\overline{\mathbb{Q}}^\times)\}$$

where $\operatorname{ord}(\zeta)$ denotes the order of an element $\zeta \in \operatorname{Tor}(\overline{\mathbb{Q}}^\times)$. Then observe that α^ℓ is torsion-free. Now if a number $\beta \in \overline{\mathbb{Q}}^\times$ is torsion-free, then each distinct conjugate $\sigma\beta$ determines a distinct element in $\mathcal{G}$, so we have

$$\deg \beta = [G : \operatorname{Stab}_G(\boldsymbol{\beta})] = [K_{\boldsymbol{\beta}} : \mathbb{Q}] = \delta(\boldsymbol{\beta}).$$

Thus $\deg \alpha^\ell = \delta(\boldsymbol{\alpha}^\ell)$. This proves existence in the first claim, and the existence of a minimum value follows since $\mathbb{N}$ is discrete. To prove the second claim, observe that $\mathbb{Q}(\alpha^\ell) \subset \mathbb{Q}(\alpha)$, so with the choice of ℓ as above, we have $\delta(\boldsymbol{\alpha}) = [\mathbb{Q}(\alpha^\ell) : \mathbb{Q}] \mid [\mathbb{Q}(\alpha) : \mathbb{Q}] = \deg \alpha$ for all $\alpha \in \overline{\mathbb{Q}}^\times$. The third now follows immediately. □

It is proven in [9] that in fact, the minimal value $\ell(\boldsymbol{\alpha})$ satisfies $d(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha})\delta(\boldsymbol{\alpha})$.

2.3. **Projections to distinguished subspaces.** Suppose $\boldsymbol{\beta} \in V_K$ for a number field K. In our proof of Theorem 1, it will be necessary to replace an arbitrary representation $\boldsymbol{\beta} = \boldsymbol{\beta}_1 \cdots \boldsymbol{\beta}_n$ with another representation $\boldsymbol{\beta} = \boldsymbol{\beta}'_1 \cdots \boldsymbol{\beta}'_n$ where each $\boldsymbol{\beta}'_i$ belongs to V_K and satisfies $\delta h_1(\boldsymbol{\beta}'_i) \le \delta h_1(\boldsymbol{\beta}_i)$. To this end, we define an operator that projects an element $\boldsymbol{\alpha} \in \mathcal{G}$ onto the subspace V_K. Let $H = \operatorname{Gal}(\overline{\mathbb{Q}}/K) \le G$, and let $\sigma_1, \ldots, \sigma_k$ be right coset representatives from $\operatorname{Stab}_H(\boldsymbol{\alpha}) \le H$, with $k = [H : \operatorname{Stab}_H(\boldsymbol{\alpha})]$. Define the map P_K on elements of $\mathcal{G}$ via

$$P_K(\boldsymbol{\alpha}) = \left(\prod_{i=1}^{k} \sigma_i(\boldsymbol{\alpha})\right)^{1/k}. \tag{2.4}$$

Lemma 2.6. *Let $\boldsymbol{\alpha} \in \mathcal{G}$. Then:*

(1) $P_K(\boldsymbol{\alpha}^r) = P_K(\boldsymbol{\alpha})^r$.
(2) $P_K(\boldsymbol{\alpha}) \in V_K$.
(3) $P_K(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$ *for all* $\boldsymbol{\alpha} \in V_K$.

Proof. (1) $P_K(\boldsymbol{\alpha}^r) = P_K(\boldsymbol{\alpha})^r$ follows from its definition in terms of the Galois action on $\mathcal{G}$.

(2) By scaling if necessary, we may assume $\boldsymbol{\alpha} \in \mathcal{R}$. Choose a torsion-free representative $\alpha \in \overline{\mathbb{Q}}^\times$. Then, the result will follow from $N_K^{K(\alpha)} = \prod_{i=1}^k \sigma_i(\alpha)$, since $N_K^{K(\alpha)}(\alpha) \in K^\times$. To see this, note that for α torsion-free, $\alpha/\sigma(\alpha)$ is never a nontrivial torsion element, so its orbit in $\overline{\mathbb{Q}}^\times$ and its $\mathcal{G}$ orbit coincide.

(3) Again, we may assume $\boldsymbol{\alpha} \in \mathcal{R}$, so that a torsion-free representative $\alpha \in K^\times$. Then $N_K^{K(\alpha)} = \alpha^k$, and $P_K(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$ follows. □

Proposition 2.7. *Let K be a number field. Then P_K is a projection onto V_K of norm 1 with respect to the L^p norms for $1 \le p \le \infty$.*

Proof. By Lemma 2.6, it follows that $P_K^2 = P_K$ is a projection onto V_K. Then for $1 \leq p \leq \infty$,

$$h_p(P_K\boldsymbol{\alpha}) = h_p\left(\sigma_1(\boldsymbol{\alpha})\cdots\sigma_k(\boldsymbol{\alpha})\right)^{1/k} \leq \frac{1}{k}\sum_{i=1}^{k} h_p(\sigma_i\boldsymbol{\alpha}) = \frac{1}{k}\sum_{i=1}^{k} h_p(\boldsymbol{\alpha}) = h_p(\boldsymbol{\alpha}),$$

since the Weil p-height is invariant under the Galois action. This proves P_K has operator norm $\|P_K\| \leq 1$, and since $V_{\mathbb{Q}}$ is fixed for every P_K, we get $\|P_K\| = 1$. □

As a corollary, if we let $\mathcal{G}_p$ denote the completion of $\mathcal{G}$ under the Weil p-norm h_p and extend P_K by continuity, we obtain:

Corollary 2.8. *The subspace $\overline{V_K} \subset \mathcal{G}_p$ is complemented in $\mathcal{G}_p$ for all $1 \leq p \leq \infty$.*

As $\mathcal{G}_2 = L^2(Y, \lambda)$ is the L^2 space for a certain measure space (Y, λ) constructed explicitly in [1], and thus a Hilbert space, more is in fact true:

Proposition 2.9. *For each $K \in \mathcal{K}$, P_K is the orthogonal projection onto the subspace $\overline{V_K} \subset \mathcal{G}_2$.*

Proof. Observe that P_K is idempotent and has operator norm $\|P_K\| = 1$ with respect to the L^2 norm, and any such projection in a Hilbert space is orthogonal (see [15, Theorem III.1.3]). □

We now explore the relationship between the Galois group and the projection operators P_K for $K \in \mathcal{K}$.

Lemma 2.10. *For any field $K \subseteq \overline{\mathbb{Q}}$ and $\sigma \in G$,*

$$\sigma P_K = P_{\sigma K}\,\sigma.$$

Equivalently, $P_K\,\sigma = \sigma P_{\sigma^{-1}K}$.

Proof. We prove the first form, the second obviously being equivalent. Let $H = \mathrm{Gal}(\overline{\mathbb{Q}}/K)$, and note that if $\tau \in H$, then $\sigma\tau\sigma^{-1} \in \mathrm{Gal}(\overline{\mathbb{Q}}/\sigma K)$. Then by the definition of P_K:

$$\begin{aligned}
\sigma P_K\boldsymbol{\alpha} &= \sigma\left(\sigma_1\boldsymbol{\alpha}\cdots\sigma_k\boldsymbol{\alpha}\right)^{1/k} \\
&= \left(\sigma\sigma_1\boldsymbol{\alpha}\cdots\sigma\sigma_k\boldsymbol{\alpha}\right)^{1/k} \\
&= \left(\sigma\sigma_1(\sigma^{-1}\sigma)\boldsymbol{\alpha}\cdots\sigma\sigma_k(\sigma^{-1}\sigma)\boldsymbol{\alpha}\right)^{1/k} \\
&= \left((\sigma\sigma_1\sigma^{-1})\sigma\boldsymbol{\alpha}\cdots(\sigma\sigma_k\sigma^{-1})\sigma\boldsymbol{\alpha}\right)^{1/k} \\
&= P_{\sigma K}(\sigma\boldsymbol{\alpha}).
\end{aligned}$$

□

We will be particularly interested in the case where the projections P_K, P_L commute with each other (and thus $P_K P_L$ is a projection to the intersection of their ranges). To that end, let us determine the intersection of two distinguished subspaces:

Lemma 2.11. *Let $K, L \subset \overline{\mathbb{Q}}$ be extensions of $\mathbb{Q}$ of arbitrary degree. Then the intersection $V_K \cap V_L = V_{K\cap L}$.*

Proof. Let $\alpha^m \in K$ and $\alpha^n \in L$ for some $m, n \in \mathbb{N}$. Then $\alpha^{mn} \in K \cap L$, so $\boldsymbol{\alpha} \in V_{K\cap L}$. The reverse inclusion is obvious. □

Lemma 2.12. *Suppose $K \in \mathcal{K}$ and $L \in \mathcal{K}^G$. Then P_K and P_L commute, that is,*

$$P_K P_L = P_{K\cap L} = P_L P_K.$$

In particular, the family of operators $\{P_K : K \in \mathcal{K}^G\}$ is commuting.

Proof. It suffices to prove $P_K(V_L) \subset V_L$, as this will imply that $P_K(V_L) \subset V_K \cap V_L = V_{K\cap L}$ by the above lemma, and thus that $P_K P_L$ is itself a projection onto $V_{K\cap L}$, implying $P_K P_L = P_{K\cap L}$, and since any orthogonal projection is equal to its adjoint, we find that $P_{K\cap L} = P_L P_K$ as well. To prove that $P_K(V_L) \subset V_L$, observe that for $\boldsymbol{\alpha} \in V_L$,

$$P_K(\boldsymbol{\alpha}) = (\sigma_1\boldsymbol{\alpha} \cdots \sigma_k\boldsymbol{\alpha})^{1/k}$$

where the σ_i are right coset representatives of $\mathrm{Stab}_H(\boldsymbol{\alpha})$ in $H = \mathrm{Gal}(\overline{\mathbb{Q}}/K)$. However, $\sigma(V_L) = V_L$ for $\sigma \in G$ since L is Galois, and thus, $P_K(\boldsymbol{\alpha}) \in V_L$ as well. But $P_K(\boldsymbol{\alpha}) \in V_K$ by construction and the proof is complete. □

From these facts, we derive the following useful lemma:

Lemma 2.13. *If $K \in \mathcal{K}^G$, then $\delta(P_K\boldsymbol{\alpha}) \le \delta(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$.*

Proof. Let $F = K_{\boldsymbol{\alpha}}$. Since $K \in \mathcal{K}^G$, we have by Lemma 2.12 that $P_K\boldsymbol{\alpha} = P_K(P_F\boldsymbol{\alpha}) = P_{K\cap F}\boldsymbol{\alpha}$. Thus, $P_K\boldsymbol{\alpha} \in V_{K\cap F}$, and so $\delta(P_K\boldsymbol{\alpha}) \le [K \cap F : \mathbb{Q}] \le [F : \mathbb{Q}] = \delta(\boldsymbol{\alpha})$. □

3. Extremal metric heights and norms

3.1. **Construction.** The aim of this section is to construct norms extremal with respect to the minimal Mahler measure. Let us begin by recalling the metric construction, as applied in [6]:

Definition 3.1. For $f : \mathcal{G} \to [0, \infty)$, the *metric height associated to f* is defined to be the function $\widehat{f} : \mathcal{G} \to [0, \infty)$ given by

$$\widehat{f}(\boldsymbol{\alpha}) = \inf_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_1\cdots\boldsymbol{\alpha}_n} \sum_{i=1}^{n} f(\boldsymbol{\alpha}_i),$$

where the infimum ranges over all possible factorizations $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_n$ in $\mathcal{G}$.

Proposition 3.2. *Suppose $f(\boldsymbol{\alpha}^{-1}) = f(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$. Then the function $\widehat{f}$ satisfies:*

1. *$\widehat{f}(\boldsymbol{\alpha}) \le f(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$.*
2. *$\widehat{f}(\boldsymbol{\alpha}^{-1}) = \widehat{f}(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$.*
3. *$\widehat{f}(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) \le \widehat{f}(\boldsymbol{\alpha}) + \widehat{f}(\boldsymbol{\beta})$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{G}$.*
4. *The zero set $Z(\widehat{f}) = \{\boldsymbol{\alpha} \in \mathcal{G} : \widehat{f}(\boldsymbol{\alpha}) = 0\}$ is a subgroup of $\mathcal{G}$, and $\widehat{f}$ is a metric on $\mathcal{G}/Z(\widehat{f})$.*

It is the largest function that does so, that is, for any other function g which satisfies the above conditions, we have $g(\boldsymbol{\alpha}) \le \widehat{f}(\boldsymbol{\alpha})$ for all $\boldsymbol{\alpha} \in \mathcal{G}$. In particular, if f already satisfies the triangle inequality, then $\widehat{f} = f$.

This last property of being the largest metric less than or equal to f we call the *extremal property*. The construction of metric heights only uses the group structure of $\mathcal{G}$, and ignores the vector space structure. If we wish to respect scaling in $\mathcal{G}$ as well, we arrive at the notion of a norm height:

Definition 3.3. Let $f : \mathcal{G} \to [0, \infty)$ be a given function. We define the *norm height associated to* f to be the function $\widetilde{f} : \mathcal{G} \to [0, \infty)$ given by

$$\widetilde{f}(\boldsymbol{\alpha}) = \inf_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_1^{r_1}\cdots\boldsymbol{\alpha}_n^{r_n}} \sum_{i=1}^{n} |r_i| \, f(\boldsymbol{\alpha}_i),$$

where the infimum ranges over all possible factorizations $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1^{r_1} \cdots \boldsymbol{\alpha}_n^{r_n}$ in $\mathcal{G}$ with $\boldsymbol{\alpha}_i \in \mathcal{G}$ and $r_i \in \mathbb{Q}$.

Proposition 3.4. *The norm height* $\widetilde{f}$ *satisfies the properties:*

(1) $\widetilde{f}(\boldsymbol{\alpha}) \leq f(\boldsymbol{\alpha})$ *for all* $\boldsymbol{\alpha} \in \mathcal{G}$.
(2) $\widetilde{f}(\boldsymbol{\alpha}^r) = |r|\widetilde{f}(\boldsymbol{\alpha})$ *for all* $r \in \mathbb{Q}$ *and* $\boldsymbol{\alpha} \in \mathcal{G}$.
(3) $\widetilde{f}(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) \leq \widetilde{f}(\boldsymbol{\alpha}) + \widetilde{f}(\boldsymbol{\beta})$ *for all* $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{G}$.
(4) *The zero set* $Z(\widetilde{f}) = \{\boldsymbol{\alpha} \in \mathcal{G} : \widetilde{f}(\boldsymbol{\alpha}) = 0\}$ *is a vector subspace of* $\mathcal{G}$.

Thus, $\widetilde{f}$ *is a seminorm on* $\mathcal{G}$*, and a norm on* $\mathcal{G}/Z(\widetilde{f})$*. It is the largest function on* $\mathcal{G}$ *that satisfies the above properties, that is, for any other function* g *which satisfies the above conditions, we have* $g(\boldsymbol{\alpha}) \leq \widetilde{f}(\boldsymbol{\alpha})$ *for all* $\boldsymbol{\alpha} \in \mathcal{G}$*. In particular, if* f *is already a seminorm on* $\mathcal{G}$*, then* $\widetilde{f} = f$*.*

The proof of Proposition 3.4 follows easily from the definitions. As above, we refer to the last part of the proposition as the *extremal property* of the norm height construction. Observe that if f satisfies the scaling property $f(\boldsymbol{\alpha}^r) = |r|f(\boldsymbol{\alpha})$, then $\widetilde{f} = \widehat{f}$ and the construction is the same.

Proposition 3.5. $\widetilde{m}_p$ *is a vector space norm on* $\mathcal{G}$.

Proof. It only remains to show that the $\widetilde{m}_p$ vanishes precisely on the zero subspace of the vector space $\mathcal{G}$, which is $\{\mathbf{1}\}$. Observe that $h_p \leq m_p$, and therefore, by the extremal property,

$$h_p(\boldsymbol{\alpha}) \leq \widetilde{m}_p(\boldsymbol{\alpha}) \quad \text{for all} \quad \boldsymbol{\alpha} \in \mathcal{G}.$$

In particular, we see that $\widetilde{m}_p(\boldsymbol{\alpha}) = 0$ if and only if $h_p(\boldsymbol{\alpha}) = 0$, which occurs precisely when $\boldsymbol{\alpha} = \mathbf{1}$. □

Note that as the degree function δ is invariant under nonzero scaling, we have that $\delta h_p(\boldsymbol{\alpha}^r) = |r|\delta h_p(\boldsymbol{\alpha})$ for rational r. Therefore, the metric construction $\delta h_p \mapsto \widehat{\delta h_p}$ results in a norm on $\mathcal{G}$. The following result shows that this norm is exactly $\widetilde{m}_p$, which is computationally very useful and will be tacitly used several times in our proofs below:

Proposition 3.6. *The norm* $\widetilde{m}_p$ *extremal with respect to the Mahler measure* m_p *is precisely* $\widehat{\delta h_p}$*, that is,* $\widetilde{m}_p(\boldsymbol{\alpha}) = \widehat{\delta h_p}(\boldsymbol{\alpha})$ *for all* $\boldsymbol{\alpha} \in \mathcal{G}$.

Proof. By Lemma 2.5, there is a unique minimal $\ell \in \mathbb{N}$ such that

$$d(\boldsymbol{\alpha}^\ell) = \delta(\boldsymbol{\alpha}).$$

Then it is easy to see that for $\boldsymbol{\alpha} \in \mathcal{G}$, the expression

$$|s/r| \, m_p(\boldsymbol{\alpha}^{r/s}) = |s/r| \, d(\boldsymbol{\alpha}^{r/s}) h_p(\boldsymbol{\alpha}^{r/s})$$

is minimized for $r/s = \ell$, and for that value,

$$|1/\ell|\, d(\boldsymbol{\alpha}^\ell) h_p(\boldsymbol{\alpha}^\ell) = \delta(\boldsymbol{\alpha}) h_p(\boldsymbol{\alpha}).$$

We may then conclude that

$$\widetilde{m}_p(\boldsymbol{\alpha}) = \inf_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_1^{r_1}\cdots\boldsymbol{\alpha}_n^{r_n}} \sum_{i=1}^n |r_i|\, m_p(\boldsymbol{\alpha}_i) = \inf_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_1\cdots\boldsymbol{\alpha}_n} \sum_{i=1}^n \delta(\boldsymbol{\alpha}_i) h_p(\boldsymbol{\alpha}_i) = \widehat{\delta h_p}(\boldsymbol{\alpha})$$

which is the desired result. □

3.2. **Explicit values.** We will now compute the values of the norms $\widetilde{m}_p$ on certain classes of algebraic numbers.

Recall that a *surd* is an algebraic number $\alpha \in \overline{\mathbb{Q}}^\times$ such that $\alpha^n \in \mathbb{Q}^\times$ for some $n \in \mathbb{N}$. Call $\boldsymbol{\alpha} \in \mathcal{G}$ a surd if one (and therefore all) coset representatives of $\boldsymbol{\alpha}$ are surds.

Proposition 3.7. *If $\boldsymbol{\alpha} \in \mathcal{G}$ is a surd, then $\widetilde{m}_p(\boldsymbol{\alpha}) = h_p(\boldsymbol{\alpha})$.*

Proof. Observe that $\delta(\boldsymbol{\alpha}) = 1$ for any surd. Since $h_p \leq m_p$ is a norm, we have by the extremal property of $\widetilde{m}_p$ that

$$h_p(\boldsymbol{\alpha}) \leq \widetilde{m}_p(\boldsymbol{\alpha}) \leq m_p(\boldsymbol{\alpha}) \quad \text{for all} \quad \boldsymbol{\alpha} \in \mathcal{G}.$$

But then

$$\widetilde{m}_p(\boldsymbol{\alpha}) = \widehat{\delta h_p}(\boldsymbol{\alpha}) \leq \delta(\boldsymbol{\alpha}) h_p(\boldsymbol{\alpha}) = h_p(\boldsymbol{\alpha}),$$

and therefore we have equality. □

For a comparison of computations of $\widetilde{m}_p$ on surds with the p-metric Mahler measures, see [6, 13].

We now consider a class of numbers analogous to the CPS numbers of [6].

Lemma 3.8. *Suppose that $\boldsymbol{\alpha} \in \mathcal{G}$ satisfies $\widehat{m}_p(\boldsymbol{\alpha}^n) = n\,\widehat{m}_p(\boldsymbol{\alpha})$ for all $n \in \mathbb{N}$. Then $\widetilde{m}_p(\boldsymbol{\alpha}) = \widehat{m}_p(\boldsymbol{\alpha})$.*

Proof. Using Theorem 1 we may choose a factorization $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_n$ so that

$$\widehat{\delta h_p}(\boldsymbol{\alpha}) = \sum_{i=1}^n \delta(\boldsymbol{\alpha}_i) h_p(\boldsymbol{\alpha}_i).$$

By Lemma 2.5 we have an exponent $\ell_i = \ell(\boldsymbol{\alpha}_i) \in \mathbb{N}$ such that $\boldsymbol{\alpha}_i^{\ell_i} \in \mathcal{R}$ for $1 \leq i \leq n$. Let $k = \operatorname{lcm}\{\ell_1, \ldots, \ell_n\}$. Then observe that

$$k \cdot \widehat{\delta h_p}(\boldsymbol{\alpha}) = \sum_{i=1}^n \delta(\boldsymbol{\alpha}_i^k) h_p(\boldsymbol{\alpha}_i^k) = \sum_{i=1}^n d(\boldsymbol{\alpha}_i^k) h_p(\boldsymbol{\alpha}_i^k) \geq \widehat{m}_p(\boldsymbol{\alpha}^k) = k\,\widehat{m}_p(\boldsymbol{\alpha}).$$

By the extremal property, $\widehat{\delta h_p}(\boldsymbol{\alpha}) \leq \widehat{m}_p(\boldsymbol{\alpha})$, so we must have equality, as claimed. □

Definition 3.9. Call $\boldsymbol{\tau} \in \mathcal{G}$ a *Pisot/Salem number* if it has a representative $\tau \in \overline{\mathbb{Q}}^\times$ that can be written as $\tau = \tau_1 \cdots \tau_k$ where each $\tau_i > 1$ is a Pisot number (that is, an algebraic integer with all of its conjugates strictly inside the unit circle) or a Salem number (an algebraic integer with all of its conjugates on or inside the unit circle, and at least one on the unit circle).

Proposition 3.10. *Every Pisot/Salem number $\boldsymbol{\tau}$ is representable, that is, $\boldsymbol{\tau} \in \mathcal{R}$.*

Proof. It is easy to see that for a Pisot/Salem number $\boldsymbol{\tau} \in \mathcal{G}$ and its given representative $\tau > 1$ that all other Galois conjugates $\tau' \neq \tau$ have $|\tau'| < |\tau|$. Therefore τ is torsion-free, since otherwise there would be a conjugate $\tau' = \zeta\tau$ for some $1 \neq \zeta \in \mathrm{Tor}(\overline{\mathbb{Q}}^{\times})$, having the same modulus as τ, a contradiction. It follows by Lemma 2.5 that $\boldsymbol{\tau} \in \mathcal{R}$. □

For a Pisot/Salem number $\boldsymbol{\tau} \in \mathcal{G}$, it is shown in [6, Theorem 1(c)] that

$$\widehat{m}_1(\boldsymbol{\tau}^n) = 2\log\lceil\tau^n\rceil \quad \text{for all} \quad n \in \mathbb{N}.$$

Thus, by Lemma 3.8 above, we have the following result:

Proposition 3.11. *Let $\boldsymbol{\tau} \in \mathcal{G}$ be a Pisot/Salem number with given representative $\tau \in \overline{\mathbb{Q}}^{\times}$. Then*

$$\widetilde{m}_1(\boldsymbol{\tau}) = \widehat{m}_1(\boldsymbol{\tau}) = 2\log\lceil\tau\rceil.$$

Since there exist Pisot and Salem numbers of arbitrarily large degree, and for a Pisot or Salem number $\tau > 1$ we have $h_1(\tau) = (2/\deg\tau)\log\lceil\tau\rceil$, we easily see that the norms h_1 and $\widetilde{m}_1$ are inequivalent.

4. The infimum in the $\widetilde{m}_1$ norm

4.1. **S-unit subspaces and quotient norms.** Let $K \in \mathcal{K}$ be a number field with places M_K. Let $S \subset M_K$ be a finite set of places of K, including all archimedean places. Then for any field L, let

$$V_{L,S} = \{\boldsymbol{\alpha} \in V_L : \|\boldsymbol{\alpha}\|_w = 1 \text{ for } w \mid v \in M_K \setminus S\}. \tag{4.1}$$

Then $V_{L,S}$ is the $\mathbb{Q}$-vector space span inside V_L of the S'-units of L, where S' is the set of places w of L such that $w \mid v \in S$. Since we always require that S include the archimedean places, $V_{L,S}$ will always include the vector space span of the units of L.

Dirichlet's S-unit theorem and, in particular, the non-vanishing of the S-regulator, imply the following result:

Proposition 4.1. *If $S \subset M_K$ as above, then the $\mathbb{Q}$-vector space $V_{K,S}$ and its completion $\overline{V_{K,S}}$ have finite dimension $\#S - 1$. For $L \neq K$, the space $V_{L,S}$ has dimension $\#S' - 1$ where S' is the set of places w of L such that $w \mid v \in S$.*

In what follows below, we will primarily require S to be a set of rational primes, including the infinite prime. Notice that under these definitions, if $K \subset L$, then $V_{K,S} \subset V_{L,S}$. One of the goals of this section will be to determine the properties of the quotient norm of $V_{L,S}/V_{K,S}$, in a manner inspired by the initial work of A.M. Bergé and J. Martinet [2, 3] and in particular the more recent work of A.C. de la Maza and E. Friedman [11].

The main result of this section is the following theorem, which is essentially an analogue for the norm $\widetilde{m}_1$ of the main result of [12] for the infimum of the metric Mahler measure:

Theorem 1. *Let $\boldsymbol{\alpha} \in V_K$, where K is Galois. Then there exists a finite set of rational primes S, containing the archimedean place, such that*

$$\widetilde{m}_1(\boldsymbol{\alpha}) = \sum_{F \subseteq K} [F : \mathbb{Q}] \cdot h_1(\boldsymbol{\alpha}_F)$$

where $\boldsymbol{\alpha}_F \in \overline{V_{F,S}}$, $\boldsymbol{\alpha} = \prod_{F \subseteq K} \boldsymbol{\alpha}_F$. *Furthermore, for each pair of fields* $E \subseteq F \subseteq K$,

$$h_1(\boldsymbol{\alpha}_F) = \inf_{\boldsymbol{\beta} \in V_{E,S}} h_1(\boldsymbol{\alpha}_F/\boldsymbol{\beta}).$$

In other words, the norm of each $\boldsymbol{\alpha}_F$ *is equal to the quotient norm of* $\boldsymbol{\alpha}_F$ *with respect to any subfield.*

In contrast to the main result of [12], we are unable to prove that this infimum is in fact attained in the vector space of classical algebraic numbers $\mathcal{G}$, rather than the completion. However, our result is strengthened by the fact that the S-unit spaces in which the infimum is attained are finite dimensional real vector spaces. Therefore, if we must pass to the completion, we know that the terms in the infimum are limits of the form $\lim_{n\to\infty} \boldsymbol{\alpha}^{r_n}$ where $\boldsymbol{\alpha} \in \mathcal{G}$ and r_n is a sequence of rational numbers tending to a real limit r as $n \to \infty$.

Before we can prove Theorem 1, we must first prove several quotient norm results very much related to the results of [11], and then we will construct an S-unit projection which will allow us to reduce to the specified situation. Let $S \subset M_K$ be a finite set of places to be specified later, and consider two number fields $K \subset L$. Again let $V_{K,S}$ denote the vector subspace of V_K spanned by the S-units of K and let $V_{L,S}$ denote the corresponding subspace of V_L. Now for each $v \in S$ let $d_v = [K_v : \mathbb{Q}_v]$ be the local degree. Rather than following the usual convention and considering the places of L which lie above the places S of K, we will consider the $[L : K]$ absolute values which restrict to each place v (that is, we will not consider equivalence on L nor weight such by local degrees). Thus we get $\#S \cdot [L : K]$ absolute values on L. Let us fix the $\boldsymbol{\alpha} \in V_{L,S} \setminus V_{K,S}$ for which we want to compute the quotient norm modulo $V_{K,S}$. For a given $v \in S$, order the $[L : K]$ absolute values on L which extend $\|\cdot\|_v$ so that

$$\|\boldsymbol{\alpha}\|_{v,1} \leq \|\boldsymbol{\alpha}\|_{v,2} \leq \cdots \leq \|\boldsymbol{\alpha}\|_{v,[L:K]}.$$

Now we associate to $\boldsymbol{\alpha}$ a vector $a \in \mathbb{R}^{S\times[L:K]}$ via

$$\begin{aligned} \varphi : V_{L,S} &\to \mathbb{R}^{S\times[L:K]} \\ \boldsymbol{\alpha} &\mapsto a = (d_v \log \|\boldsymbol{\alpha}\|_{v,i})_{v\in S,\ 1\leq i\leq[L:K]} \end{aligned}$$

Note that by the product formula and our normalization above, the sum of the components of a is zero. By the ordering above, we also have

$$a_{v,i} \leq a_{v,i+1}$$

for all $v \in S$ and $1 \leq i < [L : K]$. The goal of this section is to prove the following results which will be needed below:

Theorem 4 (de la Maza, Friedman 2008)**.** *For* $\boldsymbol{\alpha} \in V_{L,S}$ *and the vector* $a = \varphi(\boldsymbol{\alpha}) \in \mathbb{R}^{S\times[L:K]}$ *with indices ordered as above,*

$$\inf_{\boldsymbol{\beta}\in V_{K,S}} h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) = \frac{1}{[L:\mathbb{Q}]} \sum_{i=1}^{[L:K]} \left| \sum_{v\in S} a_{v,i} \right|.$$

Equivalently,

$$\|\boldsymbol{\alpha}\|_{V_{L,S}/V_{K,S}} = \sum_{i=1}^{[L:K]} \left| \sum_{v\in S} a_{v,i} \right|,$$

where we are loosely using the notation $V_{L,S}/V_{K,S}$ for the quotient space of the vector spaces $\overline{\varphi(V_{K,S})} \subset \overline{\varphi(V_{L,S})} \subset \mathbb{R}^{S\times[L:K]}$ endowed with the L^1 norm.

Remark 4.2. When the authors of [11] claimed the infimum appearing in Theorem 4 occurs in $V_{K,S}$, they use in eqn. (2.19) that for a dense open subset of $\overline{V_{L,S}} = V_{L,S}\otimes\mathbb{R}$ (the real span of the image of the logarithmic embedding) the components $a_{v,i}$ for a given v may be assumed distinct. However, this is not always true as the distinct places of L lying above v might not be $[L : K]$ in number (for example, if v is a finite place which ramifies or has inertia). Thus we might always have a certain number of equalities amongst the $\{a_{v,i} : 1 \le i \le [L : K]\}$ for a given v. However, if we make the very minor modification of working inside of $\mathbb{R}^{S\times[L:K]}$ rather than $\overline{V_{L,S}}$, then we have no such number theoretic restrictions and may assume $[L : K]$ distinct places for a given v. After adjusting for this, the remainder of the proof in [11] carries through to show the slightly weaker result that the infimum actually occurs in the completion $\overline{V_{K,S}}$.

We make a slight extension of another result of [11]:

Theorem 5. *Let $\boldsymbol{\alpha} \in V_K$ have nonzero support at only the infinite places and one finite place v of K. Let W denote the subspace of V_K spanned by the units of K. Then there exists $\boldsymbol{\beta} \in W$ such that*

$$h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1}) = \inf_{\boldsymbol{\gamma}\in W} h_1(\boldsymbol{\alpha}\boldsymbol{\gamma}^{-1}) = \frac{1}{[K:\mathbb{Q}]}\left(|d_v \log\|\boldsymbol{\alpha}\|_v| + \left|\sum_{w|\infty} d_w \log\|\boldsymbol{\alpha}\|_w\right|\right).$$

We conclude with a new theorem that will be used to describe the infimum of $\widetilde{m}_1$:

Theorem 6. *For a given $\boldsymbol{\alpha} \in V_{L,S}$, there exists $\boldsymbol{\eta} \in \overline{V_{K,S}}$ such that the following conditions all hold:*

(1) $h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) = \inf\limits_{\boldsymbol{\beta}\in V_{K,S}} h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})$, *and*

(2) $h_1(\boldsymbol{\eta}) + [L : K]h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) = \inf\limits_{\boldsymbol{\beta}\in V_{K,S}} \big(h_1(\boldsymbol{\beta}) + [L : K]h_1(\boldsymbol{\alpha}\boldsymbol{\beta}^{-1})\big)$.

We now provide the proofs for the above results.

Proof of Theorem 4. Notice that

$$\sum_{v\in S} a_{v,1} \le \sum_{v\in S} a_{v,2} \le \cdots \le \sum_{v\in S} a_{v,[L:K]}.$$

Let k be an index such that

$$\sum_{v\in S} a_{v,k} \le 0 \le \sum_{v\in S} a_{v,k+1}$$

where we let $k = 0$ or $k = [L : K]$ if $\sum_{v\in S} a_{v,1} \ge 0$ or $\sum_{v\in S} a_{v,[L:K]} \le 0$, respectively. We will assume for the moment that $1 \le k < [L : K]$ and defer the proof for the extreme cases for the moment. Let X denote the set of $x \in \overline{\varphi(V_{K,S})} \subset \mathbb{R}^{S\times[L:K]}$ which satisfy the conditions:

$$a_{v,k} \le x_v \le a_{v,k+1} \quad \text{for all} \quad v \in S$$

and

$$\sum_{v\in S} x_v = 0,$$

where we use x_v to denote the common value of $x_{v,i}$, which must be equal for all i since x arises from $V_{K,S}$. It is easy to see that X is nonempty as it contains, for example,

$$x_v = a_{v,k} + \frac{-s_k}{s_{k+1} - s_k}(a_{v,k+1} - a_{v,k})$$

where $s_i = \sum_{v\in S} a_{v,i}$. Notice that

$$\begin{aligned}\|a - x\|_1 = \sum_{i=1}^{[L:K]} \sum_{v\in S} |a_{v,i} - x_v| &= \sum_{v\in S}\left(\sum_{i=k+1}^{[L:K]} (a_{v,i} - x_v) - \sum_{i=1}^{k}(a_{v,i} - x_v)\right) \\ &= \sum_{i=1}^{[L:K]}\left|\sum_{v\in S} a_{v,i}\right| - ([L:K] - 2k)\sum_{v\in S} x_v = \sum_{i=1}^{[L:K]}\left|\sum_{v\in S} a_{v,i}\right|.\end{aligned}$$

Since $x = \varphi(\boldsymbol{\eta})$ for some $\boldsymbol{\eta} \in \overline{V_{K,S}}$ and $[L:\mathbb{Q}]\, h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) = \|a - x\|_1$, the result will be proven if we can show that the above value is minimal for the function $F_a : \mathbb{R}^S \to \mathbb{R}$ given by $y \mapsto \|a - y\|_1$ where we again view y as a vector in $\mathbb{R}^{S\times[L:K]}$ via $y_{v,i} = y_v$.

The function F_a is clearly convex. As remarked earlier, we must work inside of $\mathbb{R}^{S\times[L:K]}$ in order to be assured of our vectors having distinct components. Observe that for any $\epsilon > 0$ we may find an $a' \in \mathbb{R}^{S\times[L:K]}$ such that its components $a'_{v,i}$ for a given v are distinct and further that $\|a - a'\|_1 < \epsilon$. In order that the above computation remains unchanged for $\|a' - x\|_1$, we will construct a' from a by adding sufficiently small $\epsilon_i < 0$ to each component $a_{v,i}$ for $1 \le i \le k$, and adding sufficiently small $\epsilon_i > 0$ for $k + 1 \le i \le [L:K]$ in such a way that $\sum_i \epsilon_i = 0$. We will determine the minimum of $F_{a'}$ and this will in turn tell us the minimum of F_a. Let Y denote the subset of $\mathbb{R}^S$ defined by

$$a'_{v,k} < y_v < a'_{v,k+1} \quad \text{for all} \quad v \in S,$$

and $\sum_{v\in S} y_v = 0$. Observe that by our choice of a', we have that $X \subset Y$. For our vector x from above, observe that by the triangle inequality, we have

$$|\|x - a\|_1 - \|x - a'\|_1| \le \|a - a'\|_1 < \epsilon,$$

and hence

$$|F_a(x) - F_{a'}(x)| \le \|a - a'\|_1 < \epsilon.$$

Thus, by allowing ϵ to approach 0 we see that in showing $F_a(x)$ is the minimum value on X, it suffices to show $F_{a'}(x)$ is the minimum value on Y. Notice that Y is an open set of $\mathbb{R}^S$ and that our vector x lies in Y so it is nonempty. Notice further that by construction of a' the above computation at x works out still to give the same value for $F_{a'}$ at any $y \in Y$. Therefore, as $F_{a'}$ is a convex function of $\mathbb{R}^S$ which is constant on the open set Y, we conclude that $F_{a'}$ is minimal on Y, as any convex function which is constant on an open set attains its minimum on that set, which completes the proof for all $1 \le k < [L:K]$.

For the remaining cases where $k = 0$ or $k = [L:K]$ we make some trivial modifications to our set X. For the case $k = 0$, we let $X \subset \overline{\varphi(V_{K,S})} \subset \mathbb{R}^{[L:K]\times S}$ be given by

$$x_v < a_{v,1} \quad \text{for all} \quad v \in S$$

and

$$\sum_{v\in S} x_v = 0,$$

where we again use x_v to denote the common value of $x_{v,i}$. Now we demonstrate that X is nonempty by constructing

$$x_v = a_{v,1} - \frac{s_1}{\#S}.$$

where $s_i = \sum_{v\in S} a_{v,i}$. In the case $k = [L : K]$ likewise we take $\sum_{v\in S} x_v = 0$ and

$$x_v > a_{v,[L:K]} \quad \text{for all} \quad v \in S,$$

to define our set X and observe that we have a point given by

$$x_v = a_{v,[L:K]} - \frac{s_{[L:K]}}{\#S}$$

(noting that $s_{[L:K]} \leq 0$ in this case). The remainder of the proof continues exactly as above. □

Proof of Theorem 5. This is in essence an application of the above theorem with W substituted as the subspace; the primary difference is that we wish to show that in this instance, the infimum claimed is in fact attained in W, rather than $\overline{W}$. Suppose without loss of generality that $d_v \log \|\boldsymbol{\alpha}\|_v < 0$ so that $a_v < 0$ (for otherwise we may replace $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha}^{-1}$ and the height is unaffected). Then

$$s = \sum_{w|\infty} d_w \log \|\boldsymbol{\alpha}\|_w = \sum_{w|\infty} a_w > 0.$$

Let $X \subset \overline{\varphi(W)} \subset \mathbb{R}^S$ (where $S = \{w \in M_K : w \mid \infty\} \cup \{v\}$) be the set of x satisfying

$$x_v = 0, \quad x_w < a_w, \quad \text{for all} \quad w \mid \infty,$$

and

$$\sum_{w|\infty} x_w = 0.$$

The set X is nonempty as it contains

$$x_w = a_w - s/n \quad \text{for all} \quad w \mid \infty,$$

where $n = \#\{w \in M_K : w \mid \infty\}$. But then

$$\|a - x\|_1 = |a_v| + \sum_{w|\infty} |a_w - x_w| = |a_v| + \sum_{w|\infty} (a_w - x_w) = |a_v| + \left|\sum_{w|\infty} a_w\right|,$$

and the claim will follow if we can show that this value is minimal, since $[K : \mathbb{Q}]h_1(\gamma) = \|\varphi(\gamma)\|_1$ for $\gamma \in V_K$. But X is nonempty and is open as a subspace of the hyperplane $\overline{\varphi(W)}$, where the convex function $F : \overline{\varphi(W)} \to \mathbb{R}$ given by $y \mapsto \|a - y\|_1$ is constant, therefore, it is the minimum of this function. Since we have an open subset of $\overline{\varphi(W)}$ clearly we have a $\boldsymbol{\beta} \in W$ such that $y = \varphi(\boldsymbol{\beta}) \in X$ and the proof is complete. □

Proof of Theorem 6. With notation as in Theorem 4, recall that we defined X to be the set of all vectors $x \in \overline{\varphi(V_{K,S})} \subset \mathbb{R}^{S\times[L:K]}$ satisfying:

$$a_{v,k} \leq x_v \leq a_{v,k+1} \quad \text{for all} \quad v \in S$$

and

$$\sum_{v\in S} x_v = 0,$$

where we use x_v to denote the common value of $x_{v,i}$, which must be equal for all i since x arises from $V_{K,S}$. It was shown in Theorem 4 that for $\boldsymbol{\eta} \in \varphi^{-1}(X) \subset \overline{V_{K,S}}$, we have the first condition that $h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1})$ is minimized. Our goal will be to show that if we choose $\boldsymbol{\eta} \in \varphi^{-1}(X)$ of minimal height, then the remaining two conditions will be satisfied. Let us determine then what the minimal height of $x = \varphi(\boldsymbol{\eta}) \in \mathbb{R}^S$ can be. For a real number t we will denote $t^+ = \max\{t, 0\}$ and $t^- = \max\{-t, 0\}$, so that $t = t^+ - t^-$ and $|t| = t^+ + t^-$. Assume for the moment that $1 \leq k < [L : K]$ and let

$$\epsilon_v = a_{v,k}^- - x_v^- \quad \text{and} \quad \epsilon_v' = x_v^+ - a_{v,k}^+. \tag{4.2}$$

Note that $\epsilon_v, \epsilon_v' \geq 0$ since $a_{v,k} \leq x_v$. It follows that we may write

$$x_v = a_{v,k} + \epsilon_v + \epsilon_v', \tag{4.3}$$

and

$$|x_v| = |a_{v,k}| - \epsilon_v + \epsilon_v'. \tag{4.4}$$

To minimize $\|x\|_1$ we want to let $\sum_v \epsilon_v$ be as large as possible, and it is easy to see that we must have $0 \leq \epsilon_v \leq \min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\}$. Observe that $\min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\} = a_{v,k}^- - a_{v,k+1}^-$, for suppose the minimum is $a_{v,k}^-$. Then $a_{v,k+1}^- = 0$, and

$$\min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\} = a_{v,k}^- = a_{v,k}^- - a_{v,k+1}^-.$$

Now, suppose $\min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\} = a_{v,k+1} - a_{v,k}$. Then we must have $a_{v,k} \leq a_{v,k+1} \leq 0$, and so

$$\min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\} = a_{v,k+1} - a_{v,k} = a_{v,k}^- - a_{v,k+1}^-.$$

Thus in general $\min\{a_{v,k}^-, a_{v,k+1} - a_{v,k}\} = a_{v,k}^- - a_{v,k+1}^-$. Define

$$C = \sum_{v\in S} \left(a_{v,k}^- - a_{v,k+1}^-\right)$$

to be the largest possible value for $\sum_v \epsilon_v$. Our proof will break into two cases. First, assume that $C \geq -\sum_v a_{v,k}$, and note that this condition is equivalent to

$$\sum_v a_{v,k}^+ \geq \sum_v a_{v,k+1}^-. \tag{4.5}$$

Recall that $\sum_v a_{v,k} \leq 0$ and observe that this is equivalent to

$$\sum_v a_{v,k}^+ \leq \sum_v a_{v,k}^-. \tag{4.6}$$

Then by equations (4.5) and (4.6) we may subtract from $a_{v,k}^-$ a real number b_v satisfying

$$b_v \geq a_{v,k+1}^- \quad \text{and} \quad \sum_v b_v = \sum_v a_{v,k}^+, \tag{4.7}$$

and since $a_{v,k} \leq a_{v,k+1}$ implies that $a_{v,k+1}^- \leq a_{v,k}^-$, the value b_v may further be chosen to satisfy $a_{v,k}^- - b_v \geq 0$ for each v. Thus, when $C \geq -\sum_v a_{v,k}$, we define ϵ_v to be $a_{v,k}^- - b_v$ and ϵ_v' to be 0. It then follows that $x_v = a_{v,k} + \epsilon_v \in [a_{v,k}, a_{v,k+1}]$ for each v and $\sum_v x_v = \sum_v a_{v,k}^+ - \sum_v b_v = 0$, giving us

$$\|x\|_1 = \sum_v |a_{v,k}| - \sum_v \epsilon_v = \sum_v 2a_{v,k}^+,$$

and $\|x\|_1$ is minimal since $\sum_v \epsilon_v$ is maximized. (Our choices of ϵ_v and ϵ_v' agree with our previous definitions (4.2), by observing that $x_v = a_{v,k} + \epsilon_v = a_{v,k}^+ - b_v$. Thus, if $a_{v,k} = -a_{v,k}^-$, then $x_v = -b_v \leq 0$, so that $\epsilon_v = a_{v,k}^- - x_v^- = x_v - a_{v,k}$, and from (4.3) we get $\epsilon_v' = 0$. While if $a_{v,k} = a_{v,k}^+$, we have $0 \leq b_v \leq a_{v,k}^- = 0$, so that $x_v = a_{v,k}^+ \geq 0$, implying $\epsilon_v' = 0$, and $\epsilon_v = x_v - a_v = a_{v,k}^+ - a_{v,k}^+ = 0 = a_{v,k}^- - x_v^-$.)

Now for the second case, assume that $C < -\sum_v a_{v,k}$. Again, in order to minimize $\|x\|_1$ we want to let $\sum_v \epsilon_v$ be as large as possible; which by construction is equal to C. But we require

$$\sum_v (\epsilon_v + \epsilon_v') = -\sum_v a_{v,k} \quad (\geq 0)$$

in order to have $\sum_v x_v = 0$, so this implies that we will need ϵ_v', precisely such that

$$\sum_v \epsilon_v' = -\sum_v a_{v,k} - \sum_v \epsilon_v = -\sum_v a_{v,k} - C.$$

Then clearly

$$\begin{aligned}\|x\|_1 = \sum_v |a_{v,k}| - \sum_v \epsilon_v + \sum_v \epsilon_v' &= \sum_v |a_{v,k}| - \sum_v a_{v,k} - 2C \\ &= \sum_v 2a_{v,k}^- - 2\sum_v (a_{v,k}^- - a_{v,k+1}^-) = \sum_v 2a_{v,k+1}^-.\end{aligned}$$

So, by (4.5) we may express the minimal height of x in both cases as

$$\|x\|_1 = \max\left\{\sum_v 2a_{v,k}^+, \sum_v 2a_{v,k+1}^-\right\}.$$

Using such a minimal $\boldsymbol{\eta} = \varphi^{-1}(x) \in \overline{V_{K,S}}$ we see that the first two claims are satisfied. It remains to show that the third claim is true, specifically, that

$$h_1(\boldsymbol{\eta}) + [L:K]h_1(\boldsymbol{\alpha}\boldsymbol{\eta}^{-1}) \leq [L:K]h_1(\boldsymbol{\alpha}).$$

Translated into the appropriate L^1-norms, this claim is equivalent to:

$$\|x\|_{L^1(\mathbb{R}^S)} + \|a - x\|_{L^1(\mathbb{R}^{[L:K]\times S})} \leq \|a\|_{L^1(\mathbb{R}^{[L:K]\times S})}.$$

Where in the term $\|a-x\|_{L^1(\mathbb{R}^{[L:K]\times S})}$ we view x as a vector in $\mathbb{R}^{[L:K]\times S}$ via $x_{v,i}=x_v$ for all i. Writing this expression out, we have

$$\sum_v |x_v| + \sum_{i=1}^{[L:K]} \left| \sum_v a_{v,i} \right| \leq \sum_{i=1}^{[L:K]} \sum_v |a_{v,i}|,$$

equivalently, rearranging these terms,

$$2\max\left\{ \sum_v a_{v,k}^+, \sum_v a_{v,k+1}^- \right\} \leq 2\sum_v \left(\sum_{i=1}^{k} a_{v,i}^+ + \sum_{i=k+1}^{[L:K]} a_{v,i}^- \right), \tag{4.8}$$

which is clearly true and completes the proof for the cases $1 \leq k < [L:K]$. For the remaining cases, observe that for $k=0$ we have $x_v < a_{v,1}$ and thus it is easy to see that our minimal height is

$$\sum_v |x_v| = \sum_v 2a_{v,1}^-$$

and since the right hand side of (4.8) holds for $k=0$, the inequality still holds. The $k=[L:K]$ case is similar, as $a_{v,[L:K]} < x_v$ implies our minimal height is

$$\sum_v |x_v| = \sum_v 2a_{v,[L:K]}^+. \qquad \square$$

4.2. **S-unit projections and proof of Theorem 1.** Let K be a finite Galois extension of $\mathbb{Q}$ with set of places M_K. We normalize our absolute values by letting $\|\cdot\|_v$ be the absolute value which extends $|\cdot|_p$ for the rational prime p such that $v|p$, and let $|\cdot|_v = \|\cdot\|_v^{[K_v:\mathbb{Q}_v]/[K:\mathbb{Q}]}$. Denote by S a finite set of places to be fixed later which includes all of the archimedean places. Let O_K be the ring of algebraic integers of K and let U_S be the group of S-units of K. Since S is finite and contains the archimedean places, we know by Dirichlet's S-unit theorem that U_S is a free abelian group of finite rank $s=\#S-1$. Recall that the class group is the group of nonzero fractional ideals of K modulo principal ideals. It is well-known that for number fields, the class group of a number field has a finite order, and we will denote the order of the class group of K by h. It follows immediately that if for some finite place $v \in M_K$ the ideal

$$\mathcal{P}_v = \{\alpha \in K : \|\alpha\|_v < 1\} \subset O_K$$

is not principal, then

$$\mathcal{P}_v^h = (\alpha) \subset O_K \tag{4.9}$$

is a principal ideal of O_K, since the class of $\mathcal{P}_v^h$ is trivial in the class group.

The goal of this section is to construct a projection $P_S : V_K \to V_{K,S}$, with $V_{K,S}$ defined by equation (4.1), which will be instrumental in the proof of the main theorem. Let S consist of the following places of K:

1. The archimedean places of K.
2. The support of $\boldsymbol{\alpha}$ (all places where $\boldsymbol{\alpha}$ has nontrivial valuation).
3. The Galois conjugates of the above places under the natural action $\|\cdot\|_{\sigma v} = \|\sigma^{-1}(\cdot)\|_v$.

It is clear that S is finite. We now proceed to associate a generator to each place outside of S:

Lemma 4.3. *Let $K/\mathbb{Q}$ be galois. For any $v \in M_K \setminus S$, we can find $\boldsymbol{\alpha}_v \in V_K$ such that*

1. $\|\boldsymbol{\alpha}_v\|_v < 1$,
2. $\|\boldsymbol{\alpha}_v\|_w = 1$ *for all* $w \in M_K \setminus S$ *with* $w \neq v$, *and*
3. $\|\boldsymbol{\alpha}_v\|_w \geq 1$ *for all* $w \in S$.
4. $h_1(\boldsymbol{\alpha}_v) = \inf_{\boldsymbol{\beta} \in V_{K,S}} h_1(\boldsymbol{\alpha}_v/\boldsymbol{\beta})$.

Proof. If $\mathcal{P}_v = \{\alpha \in K : \|\alpha\|_v < 1\} \subset O_K$ is a principal ideal, then let β be a generator. Otherwise, let $\mathcal{P}_v^h = (\alpha)$ as in (4.9) and let $\boldsymbol{\beta} = \boldsymbol{\alpha}^{1/h} \in V_K$. Clearly, $\boldsymbol{\beta}$ has a nontrivial finite valuation only at v of $\|\boldsymbol{\beta}\|_v = p^{-1/e}$, where e is the ramification index of $v \mid p$. By Theorem 5 above, we can find $\boldsymbol{\eta} \in V_{K,S}$ such that

$$\begin{aligned} h_1(\boldsymbol{\beta\eta}) &= \sum_{w \in M_K} |\log|\beta\eta|_w| \\ &= \sum_{w \in M_K \setminus S} |\log|\beta|_w| + \sum_{w \in S} |\log|\beta\eta|_w| \\ &= |\log|\beta|_v| + \left| \sum_{w \in S} \log|\beta\eta|_w \right| \\ &= |\log|\beta|_v| + \left| \sum_{w \in S} \log|\beta|_w \right|, \end{aligned}$$

where the last equality follows from the product formula for η. That we have equality in the third step above implies that either $\log|\beta\eta|_w \geq 0$ for all $w \in S$ or $\log|\beta\eta|_w \leq 0$ for all $w \in S$. By our choice of β we have $\log|\beta|_v < 0$, and hence, by the product formula, all of the S valuations of $\beta\eta$ must be nonnegative. We therefore can choose $\boldsymbol{\alpha}_v = \boldsymbol{\beta\eta}$ and we are done. □

Let $v \in M_K$ and suppose $v \mid p$ for the rational prime p. Let $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be the absolute Galois group, and let

$$H = \mathrm{Stab}_G(v)$$

be the decomposition group associated to the finite place v. Let $\boldsymbol{\alpha} \in V_K$ and take $\{\sigma_1, \ldots, \sigma_k\}$ to be a set of right coset representatives for $\mathrm{Stab}_H(\boldsymbol{\alpha})$ in H (where $k = [H : \mathrm{Stab}_H(\boldsymbol{\alpha})]$) and then define $P_H : V_K \to V_K$ to be

$$P_H\boldsymbol{\alpha} = (\sigma_1(\boldsymbol{\alpha}) \cdots \sigma_k(\boldsymbol{\alpha}))^{1/k}.$$

Then by Proposition 2.7, P_H is a projection to $V_F \subseteq V_K$ for $F \subseteq K$ the fixed field of H, of operator norm 1 with respect to the Weil p-height h_p for $1 \leq p \leq \infty$. We will now construct a system of $\boldsymbol{\alpha}_v$ for each place $v \in M_K \setminus S$.

Lemma 4.4. *There exists a set $\{\boldsymbol{\alpha}_v \in V_K : v \in M_K \setminus S\}$ such that each $\boldsymbol{\alpha}_v$ satisfies the conditions of Lemma 4.3 above with the following additional property: for any $w \in S$ and $\sigma \in G$, if $\|\alpha_v\|_w \neq \|\alpha_v\|_{\sigma w}$ then $\sigma v \neq v$.*

Proof. For each rational prime p which has a place in $M_K \setminus S$ lying above it, pick one particular place $v \mid p$ lying above it. Choose $\boldsymbol{\alpha}'_v$ to be the number constructed by Lemma 4.3 above, and let $\boldsymbol{\alpha}_v = P_H \boldsymbol{\alpha}'_v$ where $H = \mathrm{Stab}_G(v)$ is the stabilizer of the place v in the absolute Galois group as above. Notice that by the fact that P_H has norm 1 and by minimality modulo $V_{K,S}$ of $\boldsymbol{\alpha}'_v$ in Lemma 4.3, $h_1(P_H \boldsymbol{\alpha}'_v) = h_1(\boldsymbol{\alpha}'_v)$. Since S is closed under the Galois action, and H fixes the place v, $\boldsymbol{\alpha}_v$ still satisfies the criteria of Lemma 4.3. For any other place $w \mid p$ lying above the same rational prime p, observe that there exists $\sigma \in G$ with $\sigma v = w$. Define $\boldsymbol{\alpha}_w = \sigma^{-1}(\boldsymbol{\alpha}_v)$, and repeat this construction for every rational prime p whose extensions to K lie in $M_K \setminus S$. This gives us the entire set of $\boldsymbol{\alpha}_v$ whose existence we need to establish, and since the Galois action permutes the places v lying over p, the $\boldsymbol{\alpha}_v$ thus constructed all meet the conditions of Lemma 4.3.

It now remains to see that this set has the additional property claimed. This is guaranteed by the "averaging" over H done by P_H in constructing the original $\boldsymbol{\alpha}_v$ whose orbit we took in the above construction. Observe that if $\sigma \in G$ fixes the v-adic valuation of $\boldsymbol{\alpha}_v$, then $\sigma \in H$. Let $F \subseteq K$ denote the fixed field of H and view P_H as the projection to V_F. Then $\boldsymbol{\alpha}_v \in V_F$ is some power of an element of $F^\times / \mathrm{Tor}(F^\times)$, so by linearity, we have $\sigma \boldsymbol{\alpha}_v = \boldsymbol{\alpha}_v$. Thus we see that for such $\sigma \in G$, $\|\sigma \boldsymbol{\alpha}_v\|_w = \|\boldsymbol{\alpha}_v\|_{\sigma^{-1} w}$ unless $\sigma v \neq v$, in which case we have the desired conclusion. ☐

Corollary 4.5. *For $v \mid p$ and $\boldsymbol{\alpha}_v$ in the set as constructed in Lemma 4.4, $\delta(\boldsymbol{\alpha}_v)$ is precisely the number of places of K which lie over p.*

Proof. As seen in the proof, if $\sigma(\boldsymbol{\alpha}_v) \neq \boldsymbol{\alpha}_v$, then $\sigma v \neq v$. While, if $\sigma(\boldsymbol{\alpha}_v) = \boldsymbol{\alpha}_v$, then $1 > \|\boldsymbol{\alpha}_v\|_v = \|\sigma(\boldsymbol{\alpha}_v)\|_v = \|\boldsymbol{\alpha}_v\|_{\sigma^{-1} v}$, which gives $\sigma v = v$. ☐

We are now ready to construct the projection $P_S : V_K \to V_{K,S}$ which is fundamental to the proof of Theorem 1.

Proposition 4.6. *There exists a linear projection $P_S : V_K \to V_{K,S}$ which satisfies the following properties:*

1. $h_1(P_S \boldsymbol{\alpha}) \leq h_1(\boldsymbol{\alpha})$, *so $\|P_S\| = 1$ with respect to the Weil height norm, and*
2. $\delta(P_S \boldsymbol{\alpha}) \leq \delta(\boldsymbol{\alpha})$, *and thus $\|P_S\| = 1$ with respect to the Mahler norm.*

Proof. For our given S, let $\{\boldsymbol{\alpha}_v : v \in M_K \setminus S\}$ be the set constructed by Lemma 4.4. For each $v \in M_K \setminus S$, define the map $n_v : V_K \to \mathbb{Q}$ via the requirement that

$$\left\| \boldsymbol{\beta} \boldsymbol{\alpha}_v^{-n_v(\boldsymbol{\beta})} \right\|_v = 1 \quad \text{for all} \quad \boldsymbol{\beta} \in V_K.$$

It is easy to see that such a value for n_v must exist and be unique, since the v-adic valuations are discrete.[4] Further, observe that

$$n_v(\boldsymbol{\beta}\boldsymbol{\gamma}) = n_v(\boldsymbol{\beta}) + n_v(\boldsymbol{\gamma}) \quad \text{for all} \quad \boldsymbol{\beta}, \boldsymbol{\gamma} \in V_K.$$

[4] The reader will note that by our choice of $\boldsymbol{\alpha}_v$, the function $n_v(\cdot)$ is essentially the linear extension of $\mathrm{ord}_v(\cdot)$ from $K^\times / \mathrm{Tor}(K^\times)$ to V_K.

Define the map

$$\begin{aligned} P_S : V_K &\to V_{K,S} \\ \boldsymbol{\alpha} &\mapsto \boldsymbol{\alpha} \cdot \prod_{v \in M_K \setminus S} \boldsymbol{\alpha}_v{}^{-n_v(\boldsymbol{\alpha})}. \end{aligned}$$

That this is well-defined follows from the fact that $n_v(\boldsymbol{\alpha}) = 0$ for all but finitely many v and from the fact that by our choice of $\boldsymbol{\alpha}_v$ and $n_v(\boldsymbol{\alpha})$, $P_S\boldsymbol{\alpha}$ has support only in S and thus belongs to the $\mathbb{Q}$-vector space span of the S-units $V_{K,S}$. It follows easily from the definition that $P_S(\boldsymbol{\beta}^r\boldsymbol{\gamma}^s) = P_S(\boldsymbol{\beta})^r P_S(\boldsymbol{\gamma})^s$, hence P_S is linear.

We will now prove that P_S satisfies the first desired property. Fix our $\boldsymbol{\alpha} \in V_K$ and let $\boldsymbol{\beta} = P_S\boldsymbol{\alpha} \in V_{K,S}$. Let T denote the Galois orbit of $\operatorname{supp}(\boldsymbol{\alpha}) \setminus S$ inside M_K. The claim is then that

$$h_1(\boldsymbol{\beta}) \leq h_1\left(\boldsymbol{\beta} \prod_{v \in T} \boldsymbol{\alpha}_v^{n_v}\right) = h_1(\boldsymbol{\alpha}),$$

where we will suppress the argument in the exponents $n_v = n_v(\boldsymbol{\alpha})$. Denote $S' = M_K \setminus S$. Then

$$h_1(\boldsymbol{\beta}) = \sum_{w \in S} |\log|\boldsymbol{\beta}|_w| + \sum_{w \in S'} |\log|\boldsymbol{\beta}|_w| \,. \tag{4.10}$$

Now, $\sum_{w \in S'} |\log|\boldsymbol{\beta}|_w| = 0$, since $\boldsymbol{\beta} \in V_{K,S}$. We apply the triangle inequality to the remaining term:

$$\sum_{w \in S} |\log|\boldsymbol{\beta}|_w| \leq \sum_{w \in S} \left|\log|\boldsymbol{\beta}|_w + \sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right| + \sum_{w \in S} \left|\sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right|. \tag{4.11}$$

Observe that by our choice of $\boldsymbol{\alpha}_v$ for $v \in T$ in the lemmas above, we have $|\boldsymbol{\alpha}_v|_w \geq 1$ for all $w \in S$ and $|\boldsymbol{\alpha}_v|_w \leq 1$ for all $w \in S'$. Thus,

$$\sum_{w \in S} \left|\sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right| \leq \sum_{w \in S} \sum_{v \in T} |n_v| \log|\boldsymbol{\alpha}_v|_w = \sum_{w \in S'} \sum_{v \in T} |n_v|(-\log|\boldsymbol{\alpha}_v|_w),$$

where the last equality follows from the product formula. But $|\boldsymbol{\alpha}_v|_w = 1$ for all $w \in S' \setminus \{v\}$ and $|\boldsymbol{\alpha}_v|_v < 1$, so in fact,

$$\sum_{w \in S} \left|\sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right| \leq \sum_{w \in S'} \left|\sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right|.$$

On observing that $|\boldsymbol{\beta}|_w = 1$ for all $w \in S'$, we may write this same expression as:

$$\sum_{w \in S} \left|\sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right| \leq \sum_{w \in S'} \left|\log|\boldsymbol{\beta}|_w + \sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right|. \tag{4.12}$$

Combining equations (4.10), (4.11), and (4.12), we find that

$$h_1(\boldsymbol{\beta}) \leq \sum_{w \in M_K} \left|\log|\boldsymbol{\beta}|_w + \sum_{v \in T} n_v \log|\boldsymbol{\alpha}_v|_w\right| = h_1\left(\boldsymbol{\beta} \prod_{v \in T} \boldsymbol{\alpha}_v^{n_v}\right),$$

which is the desired result.

It now remains to prove the second claim, namely that

$$\delta(\boldsymbol{\beta}) \leq \delta\biggl(\boldsymbol{\beta}\prod_{v\in T}\boldsymbol{\alpha}_v^{n_v}\biggr) = \delta(\boldsymbol{\alpha}).$$

Suppose for some $\sigma \in G$ that $\boldsymbol{\beta} \neq \sigma\boldsymbol{\beta}$ but $\sigma(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$. Then for some $w \in S$, $\|\boldsymbol{\beta}\|_w \neq \|\boldsymbol{\beta}\|_{\sigma w}$, and so we must have

$$\biggl\|\prod_{v\in T}\boldsymbol{\alpha}_v^{n_v}\biggr\|_w \neq \biggl\|\prod_{v\in T}\boldsymbol{\alpha}_v^{n_v}\biggr\|_{\sigma w}.$$

It follows then by Lemma 4.4 that for some $v \in T$ we must have $\sigma v \neq v$ and $n_v \neq n_{\sigma v}$, else the w-adic valuation would not differ. But then it is easy to see that

$$\biggl\|\prod_{u\in T}\boldsymbol{\alpha}_u^{n_u}\biggr\|_v = \|\boldsymbol{\alpha}_v\|_v^{n_v} = p^{-n_v/e} \neq p^{-n_{\sigma v}/e} = \|\boldsymbol{\alpha}_{\sigma v}\|_{\sigma v}^{n_{\sigma v}} = \biggl\|\prod_{u\in T}\boldsymbol{\alpha}_u^{n_u}\biggr\|_{\sigma v},$$

where e is the ramification index of $v \mid p$. Thus any contribution the $\prod_{u\in T}\boldsymbol{\alpha}_u^{n_u}$ term might have towards decreasing the orbit of $\boldsymbol{\alpha} = \boldsymbol{\beta}\prod_{v\in T}\boldsymbol{\alpha}_v^{n_v}$ by equating two w-adic valuations of $\boldsymbol{\alpha}$ for $w \in S$ will nevertheless result in distinct v-adic valuations for some $v \in T$ and thus the new orbit will be at least as large, proving the claim. □

We are now ready to prove Theorem 1.

Proof of Theorem 1. Let $\boldsymbol{\alpha} \in V_K$, where K is the Galois closure of the minimal field of $\boldsymbol{\alpha}$. Let $P_K : \mathcal{G} \to V_K$ be the projection to V_K, S the set constructed above for K so that in fact $\boldsymbol{\alpha} \in V_{K,S}$, and $P_S : V_K \to V_{K,S}$ the projection defined in Proposition 4.6, where $V_{K,S}$ is the $\mathbb{Q}$-vector space span of the S-units in $K^\times$ modulo torsion. Notice that in fact, for some set $S' \subset M_{\mathbb{Q}}$, we have

$$\bigcup_{v\in S}\{w \in M_{\overline{\mathbb{Q}}} : w \mid v \in M_K\} = \bigcup_{p\in S'}\{w \in M_{\overline{\mathbb{Q}}} : w \mid p \in M_{\mathbb{Q}}\}$$

by the requirement that S be closed under the Galois action. S, as a set of places on K, meets the criteria set forth in the theorem statement. Let $P = P_S P_K : \mathcal{G} \to V_{K,S}$. By Lemma 2.13 and Propositions 2.7 and 4.6, we have that $\delta h_1(P\boldsymbol{\beta}) \leq \delta h_1(\boldsymbol{\beta})$ for all $\boldsymbol{\beta} \in \mathcal{G}$. Since P is linear and $\boldsymbol{\alpha} \in V_{K,S}$, note that $\boldsymbol{\alpha} = P\boldsymbol{\alpha}$, so if we have a factorization of $\boldsymbol{\alpha}$ into $\boldsymbol{\alpha}_i \in \mathcal{G}$ for $i = 1, \ldots, n$, then

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_n \implies \boldsymbol{\alpha} = (P\boldsymbol{\alpha}_1)\cdots(P\boldsymbol{\alpha}_n),$$

and $P\boldsymbol{\alpha}_i \in V_{K,S}$ for all $i = 1, \ldots, n$. Then by our established inequalities for P_K andP_S with respect to δh_1,

$$\sum_{i=1}^n \delta h_1(P\boldsymbol{\alpha}_i) \leq \sum_{i=1}^n \delta h_1(\boldsymbol{\alpha}_i).$$

Hence we may take the infimum within $V_{K,S}$. Associate to each term in the infimum its minimal subspace $V_{F,S} \subseteq V_{K,S}$ containing it for $F \subset K$. If we have more than one term for any given minimal subspace $V_{F,S}$, notice that the δ values are equal and we can combine any such terms by the triangle inequality for h_1. Thus, the first part of the claim is proven. The remaining criterion easily follows from observing that the choice of $\boldsymbol{\alpha}_F$ can be made in accord with Theorem 6. □

References

[1] D. Allcock, J.D. Vaaler. *A Banach Space determined by the Weil Height.* Acta Arith. 136 (2009), no. 3, 279–298.
[2] A.-M. Bergé, J. Martinet. *Minorations de hauteurs et petits régulateurs relatifs.* (French) Séminaire de Théorie des Nombres de Bordeaux (Talence, 1987–1988), Exp. no. 11.
[3] A.-M. Bergé, J. Martinet. *Notions relatives de régulateurs et de hauteurs.* (French) Acta Arith. 54 (1989), no. 2, 155–170.
[4] E. Dobrowolski. *On a question of Lehmer and the number of irreducible factors of a polynomial.* Acta Arith. 34 (1979), no. 4, 391–401.
[5] A. Dubickas. *Two exercises concerning the degree of the product of algebraic numbers.* Publ. Inst. Math. (Beograd) (N.S.) 77(91) (2005), 67–70.
[6] A. Dubickas, C.J. Smyth. *On the metric Mahler measure.* J. Number Theory 86 (2001), no. 2, 368–387.
[7] A. Dubickas, C.J. Smyth. *On metric heights.* Periodica Mathematica Hungarica Vol. 46 (2), 2003, 135–155.
[8] D.H. Lehmer. *Factorization of certain cyclotomic functions.* Ann. of Math. (2) 34 (1933), no. 3, 461–479.
[9] P. Fili, Z. Miner. *Orthogonal decomposition of the space of algebraic numbers and Lehmer's problem*, Preprint available at `arXiv:0911.1975`.
[10] P. Fili, C.L. Samuels. *On the non-Archimedean metric Mahler measure.* J. Number Theory 129 (2009), no. 7, 1698–1708.
[11] A.C. de la Maza, E. Friedman. *Heights of algebraic numbers modulo multiplicative group actions.* J. Number Theory 128 (2008), no. 8, 2199–2213.
[12] C.L. Samuels. *The infimum in the metric Mahler measure.* Canad. Math. Bull., to appear.
[13] C.L. Samuels, J. Jankauskas. *The t-metric Mahler measures of surds of rational numbers.* Acta Math. Hungar., to appear.
[14] J.D. Vaaler. *Notes on Dobrowolski's Theorem and Lehmer's Conjecture.* Course notes, 2007.
[15] K. Yosida. *Functional analysis* (Sixth edition). *Grundlehren der Mathematischen Wissenschaften*, 123. Springer-Verlag, Berlin-New York, 1980. xii+501 pp.

Department of Mathematics, University of Texas at Austin, TX 78712
E-mail address: `pfili@math.utexas.edu`

Department of Mathematics, University of Texas at Austin, TX 78712
E-mail address: `zminer@math.utexas.edu`