AN ABSOLUTE SIEGEL'S LEMMA

Damien Roy and Jeffrey Lin Thunder

Dedicated to Wolfgang M. Schmidt on the occasion of his sixtieth birthday.

Abstract. We establish a version of Siegel's Lemma over the algebraic closure $\overline{k}$ of a global field k and apply it to show the existence of subspaces of $(\overline{k})^n$ of relatively small height between given subspaces of $(\overline{k})^n$.

Introduction

Let K be a finite algebraic extension of the field $\mathbb{Q}$ of rational numbers and let n be a positive integer. A celebrated result of Bombieri and Vaaler (Theorem 9 of [2]) shows that, given a system S of homogeneous linear equations in n variables with coefficients in K, there exists a basis of solutions of S in K^n whose product of the heights is bounded above by the product of the height of S and a constant $C(K, m, n)$ which depends only on the field K, the dimension m of the solution space and the integer n. In this statement, all heights are absolute multiplicative Weil's heights; their precise definitions are recalled in the text. The above result shows the existence of solutions of S of relatively small height in terms of the height of S. This type of result plays a central role in Diophantine approximation and transcendental number theory (for example, see [12]). It appeared first in the work of Thue [15]. However, it is usually referred to as Siegel's Lemma since Siegel was the first to state such a result formally [13].

Let $\overline{\mathbb{Q}}$ be a fixed algebraic closure of $\mathbb{Q}$. By duality (formula (4) in §1 of [11]), the above mentioned result can be translated in the following terms, where H stands for the height:

Let $V \subseteq (\overline{\mathbb{Q}})^n$ be a subspace of dimension $m > 0$ and let $K \subset \overline{\mathbb{Q}}$ be a field of definition of V of finite degree over $\mathbb{Q}$. Then there exists a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ of V which is contained

First author partially supported by NSERC

in K^n and satisfies

$$H(\mathbf{x}_1)\cdots H(\mathbf{x}_m) \le C(K,m,n)H(V). \tag{1}$$

For fixed integers m and n with $0 < m \le n$ and for subfields K of $\overline{\mathbb{Q}}$ of fixed degree d over $\mathbb{Q}$, the constant $C(K,m,n)$ grows like a power of the absolute value of the discriminant of K and thus can be arbitrarily large. In a preceding paper [10], the authors constructed examples showing that the optimal value of $C(K,m,n)$ for which the above statements hold is bounded below by a function $C_1(K,m,n)$ with similar behavior. In other words, the dependence of $C(K,m,n)$ on the field K is essential as long as one considers bases of V contained in K^n. On the other hand, $C(K,m,n)$ increases when K is replaced by a larger field. Thus, the above result does not provide more information about bases of V with coordinates in a larger field.

In this paper we show that, given a subspace $V \subseteq (\overline{\mathbb{Q}})^n$ of positive dimension m and a real number $\epsilon > 0$, there exists a basis $\{\mathbf{x}_1,\dots,\mathbf{x}_m\}$ of V (which depends on the choice of ϵ) such that (1) holds with $C(K,m,n)$ replaced by $2^{m(m-1)/2}+\epsilon$. This is our main result. We also show that, for any given integers m, n with $0 < m < n$, there exists a subspace $V \subseteq (\overline{\mathbb{Q}})^n$ of dimension m defined over $\mathbb{Q}$ such that the product in the left hand side of (1) is bounded below by $(m+1)^{-1/2}(\sqrt{2})^m H(V)$ for any basis $\{\mathbf{x}_1,\dots,\mathbf{x}_m\}$ of V in $(\overline{\mathbb{Q}})^n$. This shows that the best constant in this context is at least exponential in m.

As an application, we consider the problem of finding a complete flag of subspaces of $(\overline{\mathbb{Q}})^n$ of small height between two given subspaces W and V with $W \subset V$. This type of question was first considered by Schmidt in the case where $W = \{\mathbf{0}\}$ or $V = (\overline{\mathbb{Q}})^n$ (Theorem 2 of [11]). The general case was studied by Burger and Vaaler (Theorem 1 of [3]). However, their results assume that W and V are defined over a common number field $K \subset \overline{\mathbb{Q}}$ and they consider only intermediate subspaces defined over K. Their bounds for the heights of these subspaces therefore involve constants depending on the field K. Using our result and looking at all subspaces between W and V, we eliminate this dependence on K.

Let k be either $\mathbb{Q}$ or a field $k_0(T)$ of rational functions in one variable over a fixed constant field k_0, and let $\overline{k}$ be an algebraic closure of k. Then the preceding statements hold with $\mathbb{Q}$ replaced by k and $\overline{\mathbb{Q}}$ replaced by $\overline{k}$. Only the constants are different. The statements corresponding to the version of Siegel's Lemma of Bombieri and Vaaler above, as well as the fact that the constant involved is essentially optimal, have been extended to function fields by J. L. Thunder (Corollary 2 and Theorem 2 of [17]). Here we consider both the number field and the function field cases simultaneously.

The proof of our main result uses the notion of twisted height. We parametrize $V \subset (\overline{k})^n$ via a linear map $\phi\colon (\overline{k})^m \to (\overline{k})^n$ whose image is V. The new height H_ϕ on $(\overline{k})^m$ given by $H_\phi(\mathbf{y}) = H(\phi(\mathbf{y}))$ for each $\mathbf{y} \in (\overline{k})^m$ is an example of twisted height. This construction (see §3 of [14]) allows one to work within $(\overline{k})^m$ instead of V. We also define the successive minima associated to a twisted height and give an upper bound for their product. The main result then follows by applying this bound to the height H_ϕ.

Associated to a twisted height is a dual twisted height. There is a simple relationship between the twisted height of a subspace of $(\overline{k})^n$ and the dual twisted height of its orthogonal space with respect to the usual bilinear form of $(\overline{k})^n$. Using this relationship we establish, by induction on n, an upper bound for the product of the successive minima associated to a twisted height. However, the case $n = 2$ has to be done separately. To treat it, we extend to the symmetric algebra over $(\overline{k})^2$ the twisted height given on $(\overline{k})^2$. We then construct, for a given integer $r > 0$, a basis of the space of homogeneous elements of degree r such that the elements of this basis have small height and are defined over K. This uses an adelic version of the theorem of Minkowski on successive minima of a convex body (Theorem 3 of [2] or Theorem 5 of [7]). We factor the elements of this basis as products of homogeneous elements of degree 1, granted that the symmetric algebra over $(\overline{k})^2$ is isomorphic to the polynomial algebra $\overline{k}[X, Y]$. This gives elements of $(\overline{k})^2$ of relatively small height and the required estimates follow.

This paper is organized as follows. In section 1 we define the twisted heights. In section 2 we state our version of Siegel's Lemma. Sections 3 and 4 are devoted to preliminaries for the proof, mainly properties of the twisted heights. In section 5 we present the proof for the case of dimension 2. In section 6 we study the successive minima associated to a twisted height, and in section 7 we relate them to the successive minima of the dual twisted height. The proof of our main result is given in section 8 together with its application to the existence of flags of subspaces of small height. In section 9 we study a canonical family of subspaces and bound their heights. Finally, section 10 is devoted to a lower bound for the optimal constant.

1. Twisted heights

Throughout this paper, k denotes either the field $\mathbb{Q}$ of rational numbers or a field $k_0(T)$ of rational functions in one indeterminate over a field k_0 (finite or not). We fix an algebraic closure $\overline{k}$ of k and denote by $K \subset \overline{k}$ an extension of k of finite degree. We also use the letters m and n to denote positive integers with $m \le n$.

We denote by $M(K)$ the set of all places v of K, with the condition that v is trivial on k_0 if $k = k_0(T)$. For each $v \in M(K)$ we denote by K_v the completion of K at v. We write $\mathfrak{O}_v$ for the valuation ring of v in K_v and $\mathfrak{M}_v$ for the maximal ideal of $\mathfrak{O}_v$. We also normalize the absolute value $|\ |_v$ on K_v in the following way. For $k = \mathbb{Q}$, we require that:

(i) if $\mathfrak{M}_v = 0$ (ie. if v is archimedean), then $|\ |_v$ extends the usual absolute value on $\mathbb{Q}$;

(ii) if $\mathfrak{M}_v$ contains a prime number $p \in \mathbb{Z}$, then $|p|_v = 1/p$.

For $k = k_0(T)$, we fix a real number $c > 0$ and require that:

(i) if $1/T \in \mathfrak{M}_v$, then $|T|_v = e^c$;

(ii) if $\mathfrak{M}_v$ contains an irreducible polynomial $P(T) \in k_0[T]$, then $|P(T)|_v = e^{-cf}$, where f is the degree of $P(T)$.

In all cases, if $w \in M(k)$ is the place of k which lies below $v \in M(K)$, then the absolute value $|\ |_v$ on K_v extends $|\ |_w$ on k_w. We write $n_v = [K_v : k_w]$ for the local degree of K at v. The product formula then reads

$$\prod_{v \in M(K)} |x|_v^{n_v} = 1$$

for all non-zero $x \in K$. We denote by $K_{\mathbb{A}}$ the adele ring of K and view K as a subfield of $K_{\mathbb{A}}$ via the diagonal embedding.

For each $v \in M(K)$ we define a norm on $(K_v)^n$ by

$$\|(x_1, \ldots, x_n)\|_v = \begin{cases} \left(|x_1|_v^2 + \cdots + |x_n|_v^2\right)^{1/2} & \text{if } v \text{ is archimedean,} \\ \max\{|x_1|_v, \ldots, |x_n|_v\} & \text{otherwise.} \end{cases}$$

Let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$ with local components $A_v \in \mathrm{GL}_n(K_v)$. We define a *height* on K^n by

$$H_A(\mathbf{x}) = \Big(\prod_{v \in M(K)} \|A_v(\mathbf{x})\|_v^{n_v} \Big)^{1/[K:k]}$$

for all $\mathbf{x} \in K^n$. If $E \subset \overline{k}$ is an extension of K of finite degree, we view $K_{\mathbb{A}}$ as a subring of $E_{\mathbb{A}}$ and also denote by A the element of $\mathrm{GL}_n(E_{\mathbb{A}})$ which coincides with A on $(K_{\mathbb{A}})^n$. Then the corresponding height on E^n extends the above one on K^n. Whence H_A provides a height on $(\overline{k})^n$.

Let r be a positive integer. Given a commutative ring R we write $\bigwedge^r(R^n)$ for the r-th exterior power of R^n and $S^r(R^n)$ for its r-th symmetric power. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the canonical basis of R^n. If $r \le n$, we put $M = \binom{n}{r}$ and identify $\bigwedge^r(R^n)$ with R^M via the R-module homomorphism which sends the products $\mathbf{e}_{i_1} \wedge \cdots \wedge \mathbf{e}_{i_r}$ with $1 \le i_1 < \cdots < i_r \le n$ to the canonical basis of R^M, according to the lexicographical order. This gives a norm on $\bigwedge^r((K_v)^n)$ for all $v \in M(K)$. It also identifies $\bigwedge^r A$ with an element of $\mathrm{GL}_M(K_{\mathbb{A}})$, and so we get a height $H_{\wedge^r A}$ on $\bigwedge^r((\overline{k})^n)$.

Similarly, we put $N = \binom{r+n-1}{n-1}$ and identify $S^r(R^n)$ with R^N via the R-module homomorphism which sends the products $\mathbf{e}_{i_1} \cdots \mathbf{e}_{i_r}$ with $1 \le i_1 \le \cdots \le i_r \le n$ to the canonical basis of R^N, according to the lexicographical order. This provides a norm on $S^r((K_v)^n)$ and a height $H_{S^r A}$ on $S^r((\overline{k})^n)$ for each $v \in M(K)$ and each integer $r > 0$.

We define the height $H_A(V)$ of a one-dimensional subspace $V \subseteq (\overline{k})^n$ as $H_A(\mathbf{x})$, where $\mathbf{x}$ is any non-zero element of V. The product formula shows that this is independent of the choice of $\mathbf{x}$. In general, if $V \subseteq (\overline{k})^n$ is a subspace of dimension $m \ge 1$, then $\bigwedge^m V$ is a one-dimensional subspace of $\bigwedge^m((\overline{k})^n)$ and we put

$$H_A(V) = H_{\wedge^r A}(\textstyle\bigwedge^m V).$$

If $V = \{\mathbf{0}\}$, we simply put $H_A(V) = 1$. In particular, we have $H_A((\overline{k})^n) = |\det A|_{\mathbb{A}}$, where

$$|\det A|_{\mathbb{A}} = \Big(\prod_{v \in M(K)} |\det A_v|_v^{n_v} \Big)^{1/[K:k]}.$$

The canonical bilinear form of R^n is the map $b\colon R^n \times R^n \to R$ given by

$$b(\mathbf{x}, \mathbf{y}) = x_1 y_1 + \cdots + x_n y_n$$

for all $\mathbf{x} = (x_1, \dots, x_n)$ and all $\mathbf{y} = (y_1, \dots, y_n)$ in R^n. If $\phi\colon R^m \to R^n$ is an R-module homomorphism, we denote its transpose by ${}^t\phi\colon R^n \to R^m$. Given $A \in \mathrm{GL}_n(K_{\mathbb{A}})$ as above, we define

$$A^* = ({}^t A)^{-1} \in \mathrm{GL}_n(K_{\mathbb{A}}).$$

Given a subspace $V \subseteq (\overline{k})^n$, we also denote by $V^\perp$ the subspace of $(\overline{k})^n$ orthogonal to V with respect to the canonical bilinear form of $(\overline{k})^n$. The following duality theorem plays an important role in the present study (see §2 of [16]).

Theorem 1.1. *For any subspace $V \subseteq (\overline{k})^n$, we have*

$$H_{A^*}(V^\perp) = |\det A|_{\mathbb{A}}^{-1} H_A(V).$$

Finally, for each positive integer r we denote by $R[X_1, \dots, X_n]_r$ the homogeneous part of degree r of the polynomial algebra $R[X_1, \dots, X_n]$. We identify the algebra $R[X_1, \dots, X_n]$ with the symmetric algebra $S(R^n)$ of R^n via the R-algebra isomorphism which sends X_i to $\mathbf{e}_i$ for $i = 1, \dots, n$. By restriction this identifies $R[X_1, \dots, X_n]_r$ with $S^r(R^n)$, and thus with R^N. This gives a norm on $K_v[X_1, \dots, X_n]_r$ for each $v \in M(K)$. We also let A act on $K_{\mathbb{A}}[X_1, \dots, X_n]_r$ by the formula $AP = P \circ A$, and denote by $H_A(P)$ the corresponding height of a polynomial $P \in \overline{k}[X_1, \dots, X_n]_r$. If α is the element of $S^r((\overline{k})^n)$ associated to P via the isomorphism $\overline{k}[X_1, \dots, X_n]_r \simeq S^r((\overline{k})^n)$, we have $H_A(P) = H_{S^r({}^tA)}(\alpha)$.

When A is the identity map of $(K_{\mathbb{A}})^n$ the above heights coincide with the ones defined in [2] and [11] and are simply denoted by H. The duality theorem gives $H(V^\perp) = H(V)$ (see §V.2 of [2] or §1 of [11]).

2. The main result

Let the notations be as in §1. To present in one statement the results for number fields and for function fields, we introduce two more notations:
(i) if $k = \mathbb{Q}$ we put

$$\delta = 1 \qquad \text{and} \qquad C(K) = |\mathrm{Disc}(K)|^{1/(2[K:k])},$$

where $\mathrm{Disc}(K)$ denotes the discriminant of K;
(ii) if $k = k_0(T)$ we put

$$\delta = 0 \qquad \text{and} \qquad C(K) = \exp\left\{ \frac{c}{\mathfrak{m}(K,k)} \big(g(K) - 1 + \mathfrak{m}(K,k) \big) \right\},$$

where c is the positive constant chosen in §1, $g(K)$ is the genus of K and $\mathfrak{m}(K,k)$ is the effective degree of the extension $k \subseteq K$ as defined in [1]. In this situation

$$\mathfrak{m}(K,k) = [K : k]/[K_0 : k_0],$$

where K_0 is the algebraic closure of k_0 in K.

The height of a system of linear equations in n unknowns with coefficients in $\overline{k}$ is by definition the height $H(W)$ of the rowspace W of its coefficient matrix. By duality (Theorem 1.1) this is also the height $H(V)$ of the solution space $V = W^{\perp}$ of the system. This allows one to state Siegel's Lemma only in terms of subspaces of $(\overline{k})^n$. Specifically, we have

Theorem 2.1. *Let $V \subseteq (\overline{k})^n$ be a subspace of dimension $m > 0$ defined over K. Then there exists a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V in K^n such that*

$$H(\mathbf{x}_1) \cdots H(\mathbf{x}_m) \leq n^{\delta m/2} C(K)^m H(V).$$

When $k = k_0(T)$ this is Corollary 2 of Theorem 1 in [17]. When $k = \mathbb{Q}$ the result follows from Theorem 9 of [2]. If we disregard constants depending only on m and n, both results are even equivalent: the main difference between them comes from the definition of the height of an element $\mathbf{x}$ of $(\overline{k})^n$. In [2] the maximum norm is used for each local factor and, if we denote the resulting height by $H_{\max}$, both heights are related by

$$H_{\max}(\mathbf{x}) \leq H(\mathbf{x}) \leq \sqrt{n} H_{\max}(\mathbf{x}).$$

The upper bound in Theorem 2.1 shows a dependency on the field K. This cannot be avoided as long as one considers only bases of V contained in K^n. More precisely, it follows from Theorem 2 of [10] and from Theorem 2 of [17] that, for any choice of integers m, n with $2 \leq m < n$, any integer $d \geq 2$, and any separable extension K of k with $K \subset \overline{k}$ and $[K : k] = d$, there exists an m-dimensional subspace V of $(\overline{k})^n$ defined over K such that

$$H(\mathbf{x}_1) \cdots H(\mathbf{x}_m) \geq C_1 C(K)^{t/(d-1)} H(V)$$

for any basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V contained in K^n. Here t denotes the integral part of $(m-1)/2$ and C_1 is a constant depending only on m and d.

We will show that this dependency on the field of definition of V disappears by looking at all bases of V in $(\overline{k})^n$.

Theorem 2.2. *Let V be as in Theorem 2.1 and let $\epsilon > 0$. There exists a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V (which depends on the choice of ϵ) such that*

$$H(\mathbf{x}_1) \cdots H(\mathbf{x}_m) \leq (2^{\delta m(m-1)/2} + \epsilon) H(V).$$

In §5 we prove an effective version of Theorem 2.2 for the case $m = 2$. The general case is proven in §8 as a consequence of a more general result where H is replaced by a twisted height. Note that, for any basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V, we have

$$H(\mathbf{x}_1) \cdots H(\mathbf{x}_m) \geq H(V)$$

(see Lemma 4.7 below). Thus, when $k = k_0(T)$ the constant in the inequality of Theorem 2.2 is essentially best possible.

Remark. Assume $k = \mathbb{Q}$ and let the notations be as in Theorem 2.1. Let $\mathfrak{O}_K$ be the ring of integers of K and, following §V.3 of [2], define the inhomogeneous height of an element $\mathbf{x} = (x_1, \ldots, x_n)$ of $(\overline{\mathbb{Q}})^n$ by

$$H_{\mathrm{inh}}(\mathbf{x}) = H_{\max}(1, x_1, \ldots, x_n).$$

Then Corollary 11 of [2] shows that there exists a basis $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ of V which is contained in $(\mathfrak{O}_K)^n$ and satisfies

$$H_{\mathrm{inh}}(\mathbf{y}_1) \cdots H_{\mathrm{inh}}(\mathbf{y}_m) \leq C(K)^m H(V).$$

Similarly, let $\mathfrak{O}$ be the ring of all algebraic integers in $\overline{\mathbb{Q}}$. We can show that, given $\epsilon > 0$, there exists a basis $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ of V which is contained in $\mathfrak{O}^n$ and satisfies

$$H_{\mathrm{inh}}(\mathbf{y}_1) \cdots H_{\mathrm{inh}}(\mathbf{y}_m) \leq (2^{m(m-1)/2} + \epsilon) H(V).$$

This follows from Theorem 2.2 using the fact that, for each non-zero $\mathbf{x} \in (\overline{\mathbb{Q}})^n$, there exists a non-zero $a \in \overline{\mathbb{Q}}$ such that $a\mathbf{x} \in \mathfrak{O}^n$ and $H_{\mathrm{inh}}(a\mathbf{x}) = H_{\max}(\mathbf{x})$. We omit the proof.

3. Preliminaries: local results

By a *local field* we mean a field which is complete with respect to a discrete or archimedean absolute value (such as K_v for any $v \in M(K)$). Given such a field L, we define a norm on L^n as in §1 by

$$\|(x_1, \ldots, x_n)\| = \begin{cases} \left(|x_1|^2 + \cdots + |x_n|^2\right)^{1/2} & \text{if } L \text{ is archimedean,} \\ \max\{|x_1|, \ldots, |x_n|\} & \text{otherwise.} \end{cases}$$

We also put a norm on each graded part of the algebras $\bigwedge(L^n)$, $S(L^n)$ and $L[X_1, \ldots, X_n]$ by identifying them with some power L^N of L as in §1.

If F is a finite algebraic extension of L, we view F as a local field with respect to the unique absolute value of F which extends the given absolute value of L. Then the corresponding norm on F^n extends the one on L^n. We say that a linear map $\phi\colon L^m \to L^n$ is *norm preserving* if

$$\|\phi(\mathbf{x})\| = \|\mathbf{x}\|$$

for all $\mathbf{x} \in L^m$. If L is non-archimedean this is equivalent to the condition

$$\phi^{-1}((\mathfrak{O}_L)^n) = (\mathfrak{O}_L)^m,$$

where $\mathfrak{O}_L = \{a \in L : |a| \leq 1\}$ denotes the ring of integers of L. If L is archimedean and if we identify it with $\mathbb{R}$ or $\mathbb{C}$ (see Theorem 4 in Chapter 2 of [1]), then the norms on L^m and L^n are associated to the usual Hermitian scalar products of $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively, (both denoted by $(\ ,\ )$) and the condition that ϕ is norm preserving becomes

$$(\phi(\mathbf{x}), \phi(\mathbf{y})) = (\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in L^m$. We establish four lemmas.

Lemma 3.1. *Let L be a local field and let $A \in \mathrm{GL}_n(L)$. There exist positive real numbers C_1, C_2 depending on A with the following property. For any finite algebraic extension F of L and any $\mathbf{x} \in F^n$, we have*

$$C_1\|\mathbf{x}\| \leq \|A_F(\mathbf{x})\| \leq C_2\|\mathbf{x}\|,$$

where A_F denotes the element of $\mathrm{GL}_n(F)$ which extends A.

Proof. Let (a_{ij}) be the matrix of A with respect to the canonical basis of L^n. For each finite algebraic extension F of L and each $\mathbf{x} = (x_1, \dots, x_n) \in F^n$, we get

$$\|A_F(\mathbf{x})\| \leq \sum_{i=1}^{n} \Big| \sum_{j=1}^{n} a_{ij} x_j \Big| \leq \Big(\sum_{i,j=1}^{n} |a_{ij}| \Big) \cdot \|\mathbf{x}\|.$$

This proves the existence of C_2. Similarly, there exists $C_3 > 0$ such that $\|A_F^{-1}(\mathbf{x})\| \leq C_3\|\mathbf{x}\|$ for all F and $\mathbf{x}$ as above. This implies $C_3^{-1}\|\mathbf{x}\| \leq \|A_F(\mathbf{x})\|$ for the same choices of F and $\mathbf{x}$. This completes the proof.

Lemma 3.2. *Let L be a local field and let $\phi\colon L^m \to L^n$ be an injective linear map.*

(i) *There exists $\lambda \in \mathrm{GL}_m(L)$ such that $\phi \circ \lambda$ is norm preserving.*

(ii) *If ϕ is norm preserving, then for all integers $r \geq 0$ and all finite algebraic extensions F of L, the maps $\bigwedge^r \phi$, $S^r\phi$ and the F-linear map $\phi_F\colon F^m \to F^n$ which extends ϕ are also norm preserving.*

Proof. (i) If L is non-archimedean, $\phi^{-1}((\mathfrak{O}_L)^n)$ is a free $\mathfrak{O}_L$-submodule of L^m of rank m. Thus, there exists $\lambda \in \mathrm{GL}_m(L)$ such that $\lambda((\mathfrak{O}_L)^m) = \phi^{-1}((\mathfrak{O}_L)^n)$. Then $\phi \circ \lambda$ is norm preserving since $(\mathfrak{O}_L)^m = (\phi \circ \lambda)^{-1}((\mathfrak{O}_L)^n)$. If L is archimedean, choose $\mathbf{u}_1, \dots, \mathbf{u}_m \in L^m$ such that $\{\phi(\mathbf{u}_1), \dots, \phi(\mathbf{u}_m)\}$ is an orthonormal basis of $\phi(L^m)$ and define λ as the element of $\mathrm{GL}_m(L)$ which sends the canonical basis of L^m to $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$. Then $\phi \circ \lambda$ maps the canonical basis of L^m to an orthonormal subset of L^n, and so is norm preserving.

(ii) Assume that ϕ is norm preserving. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ be the image under ϕ of the canonical basis of L^m. If L is non-archimedean this set is a basis of $\phi(L^m) \cap (\mathfrak{O}_L)^n$ over $\mathfrak{O}_L$, so it can be extended to a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of $(\mathfrak{O}_L)^n$ over $\mathfrak{O}_L$. If L is archimedean the set $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is an orthonormal basis of $\phi(L^m)$, so it can be extended to an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of L^n. In both cases the map $\psi \in \mathrm{GL}_n(L)$ which sends the canonical basis of L^n to $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an isometry and we have $\phi = \psi \circ \imath$, where $\imath\colon L^m \to L^n$ is given by

$$\imath(x_1, \dots, x_m) = (x_1, \dots, x_m, 0, \dots, 0) \tag{3.1}$$

for all $(x_1, \dots, x_m) \in L^m$. Since the maps $\bigwedge^r \imath$, $S^r\imath$ and $\imath_F$ all preserve the norm, we only have to show that the same is true of $\bigwedge^r \psi$, $S^r\psi$ and ψ_F.

In the non-archimedean case ψ maps $(\mathfrak{O}_L)^n$ onto itself. Therefore, $\bigwedge^r \psi$, $S^r\psi$ and ψ_F map, respectively, $\bigwedge^r((\mathfrak{O}_L)^n)$, $S^r((\mathfrak{O}_L)^n)$ and $(\mathfrak{O}_F)^n$ onto themselves, and so they preserve the norm.

In the archimedean case we may assume $\mathbb{R} \subseteq L \subseteq F = \mathbb{C}$. If $L = \mathbb{R}$, ψ is an isometry of $\mathbb{R}^n$. Then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we have

$$\|\psi_{\mathbb{C}}(\mathbf{x} + i\mathbf{y})\|^2 = \|\psi(\mathbf{x}) + i\psi(\mathbf{y})\|^2 = \|\psi(\mathbf{x})\|^2 + \|\psi(\mathbf{y})\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} + i\mathbf{y}\|^2.$$

Thus, $\psi_{\mathbb{C}}$ is an isometry of $\mathbb{C}^n$. If $L = \mathbb{C}$, then $\psi_{\mathbb{C}} = \psi$ is also an isometry of $\mathbb{C}^n$. Thus, in all cases $\psi_{\mathbb{C}}$ leaves invariant the Hermitian scalar product of $\mathbb{C}^n$: we have

$$(\psi_{\mathbb{C}}(\mathbf{x}), \psi_{\mathbb{C}}(\mathbf{y})) = (\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. The Hermitian scalar products induced on $\bigwedge^r(\mathbb{C}^n)$ and $S^r(\mathbb{C}^n)$ via the isomorphisms $\bigwedge^r(\mathbb{C}^n) \simeq \mathbb{C}^M$ and $S^r(\mathbb{C}^n) \simeq \mathbb{C}^N$ described in §1 are characterized by the conditions

$$(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_r, \mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_r) = \det((\mathbf{x}_i, \mathbf{y}_j))$$

and

$$(\mathbf{x}_1 \cdots \mathbf{x}_r, \mathbf{y}_1 \cdots \mathbf{y}_r) = \operatorname{perm}((\mathbf{x}_i, \mathbf{y}_j))$$

for all $\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_r \in \mathbb{C}^n$, where $\det(a_{ij})$ and $\operatorname{perm}(a_{ij})$ denote, respectively, the determinant and the permanent of a matrix (a_{ij}) (see §5.6 and §7.4 of [5]). Thus, these scalar products are left invariant by $\bigwedge^r(\psi_{\mathbb{C}})$ and $S^r(\psi_{\mathbb{C}})$. In particular, $\bigwedge^r \psi$ and $S^r\psi$ are isometries of $\bigwedge^r(L^n)$ and $S^r(L^n)$, respectively.

Lemma 3.3. *Let L be a local field, let*

$$\{\mathbf{0}\} \subset W_1 \subset W_2 \subset \cdots \subset W_n = L^n$$

be a nested sequence of subspaces of L^n with $\dim(W_i) = i$ for $i = 1, \ldots, n$, and let $a_1, \ldots, a_n \in L$ with

$$|a_1| \geq |a_2| \geq \cdots \geq |a_n| > 0.$$

Then there exists $\phi \in \mathrm{GL}_n(L)$ with $|\det\phi| = |a_1 a_2 \cdots a_n|$ such that, for each finite algebraic extension F of L, each $i = 1, \ldots, n$ and each element $\mathbf{x}$ in the subspace of F^n spanned by W_i, we have

$$\|\phi_F(\mathbf{x})\| \geq |a_i| \|\mathbf{x}\|,$$

where $\phi_F \in \mathrm{GL}_n(F)$ is the automorphism of F^n which extends ϕ.

Proof. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be the canonical basis of L^n and, for each $i = 1, \ldots, n$, let S_i be the subspace of L^n generated by $\mathbf{e}_1, \ldots, \mathbf{e}_i$. If L is non-archimedean, choose $\mathbf{w}_1, \ldots, \mathbf{w}_n \in L^n$ such that $\{\mathbf{w}_1, \ldots, \mathbf{w}_i\}$ is a basis of $W_i \cap (\mathfrak{O}_L)^n$ over $\mathfrak{O}_L$ for $i = 1, \ldots, n$. Otherwise, choose them such that $\{\mathbf{w}_1, \ldots, \mathbf{w}_i\}$ is an orthonormal basis of W_i for each i. Then the map $\psi \in \mathrm{GL}_n(L)$ determined by $\psi(\mathbf{w}_i) = \mathbf{e}_i$ for $i = 1, \ldots, n$ is an isometry of L^n with $\psi(W_i) = S_i$ for each i. Define $\phi = \rho \circ \psi$, where $\rho \in \mathrm{GL}_n(L)$ satisfies $\rho(\mathbf{e}_i) = a_i \mathbf{e}_i$ for $i = 1, \ldots, n$. Since $|\det\psi| = 1$, we have $|\det\phi| = |\det\rho| = |a_1 \cdots a_n|$.

Let F be a finite algebraic extension of L. For any $\mathbf{y} = (y_1, \ldots, y_i, 0, \ldots, 0)$ in the subspace of F^n generated by S_i we have

$$\|\rho_F(\mathbf{y})\| = \|(a_1 y_1, \ldots, a_i y_i, 0, \ldots, 0)\| \geq |a_i| \|\mathbf{y}\|.$$

Since ψ_F is an isometry of F^n (Lemma 3.2), this implies $\|\phi_F(\mathbf{x})\| \geq |a_i|\|\psi_F(\mathbf{x})\| = |a_i|\|\mathbf{x}\|$ for any $\mathbf{x}$ in the subspace of F^n spanned by W_i.

The last lemma compares the norm of a product of homogeneous elements of degree 1 to the product of their norms in the algebras $\bigwedge(L^n)$, $S(L^n)$ and $L[X_1, \ldots, X_n]$ for a given local field L.

Lemma 3.4. *Let r be a positive integer, let L be a local field, let $\mathbf{y}_1, \ldots, \mathbf{y}_r \in L^n$ and let $P_1, \ldots, P_r \in L[X_1, \ldots, X_n]_1$. Put $\theta = 1$ if L is archimedean and $\theta = 0$ otherwise. Then the products $\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_r \in \bigwedge^r(L^n)$, $\mathbf{y}_1 \cdots \mathbf{y}_r \in S^r(L^n)$ and $P_1 \cdots P_r \in L[X_1, \ldots, X_n]_r$ satisfy*

(i) $\|\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_r\| \leq \|\mathbf{y}_1\| \cdots \|\mathbf{y}_r\|$,
(ii) $n^{-\theta r/2}\|\mathbf{y}_1\| \cdots \|\mathbf{y}_r\| \leq \|\mathbf{y}_1 \cdots \mathbf{y}_r\| \leq n^{\theta r/2}\|\mathbf{y}_1\| \cdots \|\mathbf{y}_r\|$,
(iii) $n^{-\theta r/2}\|P_1\| \cdots \|P_r\| \leq \|P_1 \cdots P_r\| \leq n^{\theta r/2}\|P_1\| \cdots \|P_r\|$.

Proof. Inequality (i) follows by induction on r from inequality (2.6) of [2] (this inequality of [2] is stated for the completions of a number field but its proof shows that it holds in general for any local field).

Put $Q = P_1 \cdots P_r \in L[X_1, \ldots, X_n]_r$. If L is non-archimedean, Gauss' lemma (Proposition 2.1 in Chapter 3 of [6]) gives

$$\|Q\| = \|P_1\| \cdots \|P_r\|.$$

If L is archimedean we may assume $L = \mathbb{C}$. In that case denote by $\mathcal{M}(P)$ the Mahler measure of a polynomial $P \in \mathbb{C}[X_1, \ldots, X_n]$ (see [9]). Denote by $\|P\|_1$ its length, i.e., the sum of the absolute values of its coefficients, and denote by $\|P\|_\infty$ the maximum of the absolute values of its coefficients. Inequality (3) of [9] shows that for any homogeneous linear form $P \in \mathbb{C}[X_1, \ldots, X_n]_1$ we have $\|P\|_\infty \leq \mathcal{M}(P)$; thus $\|P\| \leq \sqrt{n}\mathcal{M}(P)$. Since $\|Q\| \geq \mathcal{M}(Q)$ (Proposition 2.10 in Chapter 3 of [6]), the multiplicativity of the Mahler measure gives

$$\|Q\| \geq \mathcal{M}(Q) = \mathcal{M}(P_1) \cdots \mathcal{M}(P_r) \geq n^{-r/2}\|P_1\| \cdots \|P_r\|.$$

On the other hand, we find

$$\|Q\| \leq \|Q\|_1 \leq \|P_1\|_1 \cdots \|P_r\|_1 \leq n^{r/2}\|P_1\| \cdots \|P_r\|.$$

This proves (iii). Inequality (ii) follows from (iii) using the fact that the isomorphism of algebras $L[X_1, \ldots, X_n] \simeq S(L^n)$ described in §1 is an isometry at each degree.

4. Preliminaries: global results

We first observe that the twisted heights introduced in §1 are bounded above and below by non-zero constant multiples of the usual canonical height on $\overline{k}$; cf. Proposition 7.2 of [14] dealing with the projective heights associated to metrized line bundles.

Proposition 4.1. *Let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. There exist constants $C_1, C_2 > 0$, depending only on A, such that*

$$C_1 H(\mathbf{x}) \leq H_A(\mathbf{x}) \leq C_2 H(\mathbf{x})$$

for all $\mathbf{x} \in (\overline{k})^n$.

Proof. Let E be a finite algebraic extension of K in $\overline{k}$, let $v \in M(K)$ and let $w \in M(E)$ with $w|v$. Let A_w denote the element of $\mathrm{GL}_n(E_w)$ which extends A_v. Lemma 3.1 shows that there exist constants $C_{1v}, C_{2v} > 0$, independent of the choice of E and w, such that

$$C_{1v}\|\mathbf{x}\|_w \leq \|A_w(\mathbf{x})\|_w \leq C_{2v}\|\mathbf{x}\|_w$$

for each $\mathbf{x} \in E^n$ By lemma 3.2 (ii), we may take $C_{1v} = C_{2v} = 1$ for each non-archimedean $v \in M(K)$ such that $A_v \in \mathrm{GL}_n(\mathfrak{O}_v)$, since in that case A_w is an isometry of $(E_w)^n$. Now the products

$$C_i = \prod_{v \in M(K)} (C_{iv})^{n_v/[K:k]}$$

with $i = 1, 2$ involve only a finite number of non-trivial terms and, given this choice of C_1 and C_2, the inequality of the proposition holds for all $\mathbf{x} \in E^n$. Since C_1 and C_2 are independent of the choice of E, this inequality holds for all $\mathbf{x} \in (\overline{k})^n$.

The next proposition and its corollaries concern the behavior of the twisted heights under a linear map.

Proposition 4.2. *Let $\phi\colon (\overline{k})^m \to (\overline{k})^n$ be an injective linear map defined over K and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. Denote by the same letter ϕ the element of $\mathrm{Hom}_{K_{\mathbb{A}}}((K_{\mathbb{A}})^m, (K_{\mathbb{A}})^n)$ which coincides with ϕ on K^m. Then there exists $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $A_v \circ \phi_v \circ B_v^{-1}$ is norm preserving for all $v \in M(K)$. For such a choice of B we have*

$$H_A(\phi(\mathbf{x})) = H_B(\mathbf{x}) \tag{4.1}$$

for all $\mathbf{x} \in (\overline{k})^m$.

Proof. Let ψ be an automorphism of $(\overline{k})^n$ defined over K such that $\phi = \psi \circ \imath$, where $\imath$ is given by (3.1). Use the same letters ψ and $\imath$ to denote the elements of $\mathrm{GL}_n(K_{\mathbb{A}})$ and

$\mathrm{Hom}_{K_{\mathbb{A}}}((K_{\mathbb{A}})^m, (K_{\mathbb{A}})^n)$ which coincide with ψ on K^n and with $\imath$ on K^m, respectively. Since ψ_v is an isometry of $(K_v)^n$ for almost all $v \in M(K)$, the map $\phi_v = \psi_v \circ \imath_v$ is norm preserving for almost all $v \in M(K)$. On the other hand, for each $v \in M(K)$ the composite map $A_v \circ \phi_v \colon (K_v)^m \to (K_v)^n$ is injective. Thus, by Lemma 3.2 (i) there exists $C_v \in \mathrm{GL}_m(K_v)$ such that $A_v \circ \phi_v \circ C_v$ is norm preserving. Since A_v and ϕ_v are norm preserving for almost all $v \in M(K)$, the same is true of the family $C = (C_v)_{v\in M(K)}$, and thus $C \in \mathrm{GL}_m(K_{\mathbb{A}})$. This proves the existence of B: take $B = C^{-1}$.

Now let $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $A_v \circ \phi_v \circ B_v^{-1}$ is norm preserving for all $v \in M(K)$. Consider a finite algebraic extension E of K in $\overline{k}$ and choose $v \in M(K)$ and $w \in M(E)$ such that $w|v$. By Lemma 3.2 (ii) the linear map $A_w \circ \phi_w \circ B_w^{-1} \colon (E_w)^m \to (E_w)^n$ is also norm preserving. Thus, we have

$$\|A_w(\phi(\mathbf{x}))\|_w = \|B_w(\mathbf{x})\|_w$$

for all $\mathbf{x} \in E^m$. This being true for any $w \in M(E)$, we deduce that (4.1) holds for all $\mathbf{x} \in E^m$. The choice of E being arbitrary, this inequality also holds for all $\mathbf{x} \in (\overline{k})^m$.

Corollary 4.3. *Let the notations be as in Proposition 4.2 and choose $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $A_v \circ \phi_v \circ B_v^{-1}$ is norm preserving for all $v \in M(K)$. Then*

(i) *$H_A(\phi(W)) = H_B(W)$ for all subspaces $W \subseteq (\overline{k})^m$,*

(ii) *$H_A(P \circ {}^t\phi) = H_B(P)$ for all homogeneous polynomials $P \in \overline{k}[X_1, \dots, X_m]$.*

Moreover, if V is the image of ϕ, then $H_A(V) = |\det B|_{\mathbb{A}}$.

Proof. It suffices to show (i) for subspaces of $(\overline{k})^m$ of positive dimension and (ii) for homogeneous polynomials of positive degree.

Let $r \in \mathbb{Z}$ with $1 \le r \le m$. Lemma 3.2 (ii) shows that the map

$$\textstyle\bigwedge^r (A_v \circ \phi_v \circ B_v^{-1}) = (\bigwedge^r A)_v \circ (\bigwedge^r \phi)_v \circ (\bigwedge^r B)_v^{-1} \tag{4.2}$$

from $\bigwedge^r((K_v)^m)$ to $\bigwedge^r((K_v)^n)$ is norm preserving for all $v \in M(K)$. Thus, by Proposition 4.2 we have

$$H_{\wedge^r A}\big((\textstyle\bigwedge^r \phi)(\alpha)\big) = H_{\wedge^r B}(\alpha) \tag{4.3}$$

for all $\alpha \in \bigwedge^r((\overline{k})^m)$. If W is a subspace of $(\overline{k})^m$ of dimension r and if α is a generator of $\bigwedge^r W$, then $(\bigwedge^r \phi)(\alpha)$ is a generator of $\bigwedge^r(\phi(W))$ and (4.3) becomes equality (i) of the corollary. In particular, for $r = m$ we get

$$H_A(V) = H_A\big(\phi((\overline{k})^m)\big) = H_B((\overline{k})^m) = |\det B|_{\mathbb{A}}.$$

Similarly, Lemma 3.2 (ii) shows that the map

$$S^r(A_v \circ \phi_v \circ B_v^{-1}) = (S^r A)_v \circ (S^r \phi)_v \circ (S^r B)_v^{-1} \tag{4.4}$$

from $S^r((K_v)^m)$ to $S^r((K_v)^n)$ is norm preserving for all positive integers r and all $v \in M(K)$. Thus, by Proposition 4.2 we have

$$H_{S^r A}\big((S^r\phi)(\beta)\big) = H_{S^r B}(\beta) \tag{4.5}$$

for all integers $r > 0$ and all $\beta \in S^r((\overline{k})^m)$. If $P \in \overline{k}[X_1, \dots, X_m]$ is a homogeneous polynomial of degree $r > 0$ and if $\beta \in S^r((\overline{k})^m)$ is its image under the algebra isomorphism $\overline{k}[X_1, \dots, X_n] \simeq S((\overline{k})^m)$ described in §1, then $(S^r\phi)(\beta)$ is the image of $P \circ {}^t\phi$ under the same isomorphism and (4.5) becomes equality (ii) of the corollary.

Corollary 4.4. *Let the notations be as in Proposition 4.2. Then for the same choices of B we also have*

(i) $H_{\wedge^r A}\big((\bigwedge^r\phi)(W)\big) = H_{\wedge^r B}(W)$ *for all* $r = 1, \dots, m$ *and all subspaces* W *of* $\bigwedge^r((\overline{k})^m)$,

(ii) $H_{S^r A}\big((S^r\phi)(W)\big) = H_{S^r B}(W)$ *for all* $r \geq 1$ *and all subspaces* W *of* $S^r((\overline{k})^m)$.

Proof. This follows from Corollary 4.3 using the fact that the maps (4.2) (with $1 \leq r \leq m$) and (4.4) (with $r \geq 1$) are norm preserving for all $v \in M(K)$.

We will also need the following result.

Lemma 4.5. *Let r be a positive integer, let R be a commutative ring and let A be an endomorphism of R^m as R-module. Then*

$$\det(S^r A) = (\det A)^s, \tag{4.6}$$

where $s = \binom{r+m-1}{m}$.

Proof. Let (a_{ij}) be the matrix of A with respect to the canonical basis $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ of R^m. Since both sides of (4.6) are polynomials in $a_{11}, \dots, a_{mm}$ with coefficients in $\mathbb{Z}$ (independent of the choice of R), it suffices to prove this equality when $R = \mathbb{C}$. We may also restrict ourselves to proving it for all A in a Zariski dense subset of $\mathrm{GL}_m(\mathbb{C})$.

Consider the basis of $S^r(\mathbb{C}^m)$ consisting of all products $\mathbf{e}_{i_1} \cdots \mathbf{e}_{i_r}$ with $1 \leq i_1 \leq \cdots \leq i_r \leq m$ in the lexicographical order. If (a_{ij}) is triangular, one verifies that the matrix of $S^r A$ in the above basis of $S^r(\mathbb{C}^m)$ is also triangular and that (4.6) holds. By the multiplicativity of the determinant, this implies that (4.6) holds for all $A \in \mathrm{GL}_m(\mathbb{C})$ whose matrix is the product LU of a lower triangular matrix L by an upper triangular matrix U. Since those products form a non-empty Zariski open subset of $\mathrm{GL}_m(\mathbb{C})$, inequality (4.6) holds for all endomorphisms A of $\mathbb{C}^m$ and the lemma is proved.

From the above lemma we deduce

Proposition 4.6. *Let r be a positive integer and let V be a subspace of $(\overline{k})^n$ of dimension m. Then*

$$H(S^rV) = H(V)^s,$$

where $s = \binom{r+m-1}{m}$.

Proof. Let $\phi\colon (\overline{k})^m \to (\overline{k})^n$ be an injective linear map whose image is V and let K be a field of definition of ϕ which contains k and has finite degree over k. By Corollaries 4.3 and 4.4, there exists $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $H(V) = |\det B|_{\mathbb{A}}$ and

$$H\big((S^r\phi)(W)\big) = H_{S^rB}(W)$$

for all subspaces W of $S^r((\overline{k})^m)$. Applying this with $W = S^r((\overline{k})^m)$ and using Lemma 4.5, we get

$$H(S^rV) = H_{S^rB}\big(S^r((\overline{k})^m)\big) = |\det(S^rB)|_{\mathbb{A}} = |\det B|_{\mathbb{A}}^s = H(V)^s,$$

with s as in Lemma 4.5.

To conclude this section we mention global analogs of the inequalities of Lemma 3.4. Since their proofs are quite similar, we only present the first one.

Lemma 4.7. *Let V be a subspace of $(\overline{k})^n$ of dimension m and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. Then for any basis $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ of V we have*

$$H_A(V) \le H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_m).$$

Proof. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ be a basis of V and let E be the finite extension of K generated by the coordinates of $\mathbf{x}_1, \dots, \mathbf{x}_m$. View A as an element of $\mathrm{GL}_n(E_{\mathbb{A}})$. By inequality (i) of Lemma 3.4 we have

$$\|\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_m\|_w \le \|\mathbf{y}_1\|_w \cdots \|\mathbf{y}_m\|_w$$

for all $w \in M(E)$ and all $\mathbf{y}_1, \ldots, \mathbf{y}_m \in (E_w)^n$; thus,

$$\begin{aligned} H_A(V) &= H_{\wedge^m A}(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_m) \\ &= \prod_{w \in M(E)} \|A_w(\mathbf{x}_1) \wedge \cdots \wedge A_w(\mathbf{x}_m)\|^{n_w/[E:k]} \\ &\leq \prod_{i=1}^{m} \prod_{w \in M(E)} \|A_w(\mathbf{x}_i)\|^{n_w/[E:k]} \\ &= \prod_{i=1}^{m} H_A(\mathbf{x}_i). \end{aligned}$$

Lemma 4.8. *Let r be a positive integer, let $\mathbf{x}_1, \ldots, \mathbf{x}_r \in (\overline{k})^n$ and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. The product $\beta = \mathbf{x}_1 \cdots \mathbf{x}_r \in S^r((\overline{k})^n)$ satisfies*

$$n^{-\delta r/2} H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_r) \leq H_{S^r A}(\beta) \leq n^{\delta r/2} H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_r).$$

Lemma 4.9. *Let r be a positive integer, let $P_1, \ldots, P_r \in \overline{k}[X_1, \ldots, X_n]_1$ be homogeneous linear forms and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. The product $Q = P_1 \cdots P_r \in \overline{k}[X_1, \ldots, X_n]_r$ satisfies*

$$n^{-\delta r/2} H_A(P_1) \cdots H_A(P_r) \leq H_A(Q) \leq n^{\delta r/2} H_A(P_1) \cdots H_A(P_r).$$

5. The case of dimension 2

In this section we establish the following effective version of Theorem 2.2 for a two-dimensional space.

Theorem 5.1. *Assume $n \geq 2$. Let V be a subspace of $(\overline{k})^n$ of dimension 2 defined over K and let r be a positive integer. Then there exists a basis $\{\mathbf{x}_1, \mathbf{x}_2\}$ of V with $[K(\mathbf{x}_i) : K] \leq r$ for $i = 1, 2$ such that*

$$H(\mathbf{x}_1) H(\mathbf{x}_2) \leq 2^{\delta} (r+1)^{\delta/r} C(K)^{2/r} H(V).$$

In the above statement $K(\mathbf{x}_i)$ stands for the extension of K generated by the coordinates of $\mathbf{x}_i$. The proof requires intermediate results.

Lemma 5.2. *Let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. There exists a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of K^n such that*

$$H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_n) \le n^{\delta n/2} C(K)^n |\det A|_{\mathbb{A}}.$$

Proof. If $k = k_0(T)$ this is precisely Theorem 1 of [17]. Assume $k = \mathbb{Q}$. The result then follows from Theorem 3 of [2] which provides an adelic version of Minkowski's second convex body theorem (see also Theorem 5 of [7]). To be precise, consider the set $\mathcal{S} = A^{-1}(\mathcal{C}^\circ)$ where $\mathcal{C}^\circ$ denotes the interior of

$$\mathcal{C} = \{\mathbf{x} \in (K_{\mathbb{A}})^n \ ; \ \|\mathbf{x}_v\|_v \le 1 \ \text{ for all } v \in M(K)\}.$$

As in [2], choose the Haar measure on $K_{\mathbb{A}}$ such that the quotient $K_{\mathbb{A}}/K$ has volume 1, and put the product measure on $(K_{\mathbb{A}})^n$. Then the volume of $\mathcal{C}^\circ$ is

$$V(n)^{r_1} V(2n)^{r_2} 2^{nr_2} |\mathrm{Disc}(K)|^{-n/2},$$

where r_1 and r_2 denote, respectively, the number of real and complex places of K, and $V(m)$ denotes the volume of the Euclidean unit ball of $\mathbb{R}^m$ for a given positive integer m. Since

$$\mathrm{vol}(\mathcal{S}) = |\det A|_{\mathbb{A}}^{-d} \mathrm{vol}(\mathcal{C}^\circ),$$

where $d = r_1 + 2r_2$ is the degree of K over $\mathbb{Q}$ (Proposition 3 in Chapter IV of [18]), and since

$$V(m) \ge \frac{1}{2\sqrt{m}} \Big(\frac{2\pi e}{m}\Big)^{m/2} \ge \Big(\frac{4}{m}\Big)^{m/2}$$

for each $m \ge 1$, we get

$$\mathrm{vol}(\mathcal{S}) \ge \Big(\frac{4}{n}\Big)^{dn/2} C(K)^{-dn} |\det A|_{\mathbb{A}}^{-d}.$$

Let $\lambda_1, \dots, \lambda_n$ be the successive minima of $\mathcal{S}$ with respect to the lattice K^n in $(K_{\mathbb{A}})^n$, and let $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be a corresponding basis of K^n. Then $H_A(\mathbf{x}_i) \le \lambda_i$ for $i = 1, \dots, n$. On the other hand, by Theorem 3 of [2] the successive minima of $\mathcal{S}$ satisfy

$$(\lambda_1 \cdots \lambda_n)^d \mathrm{vol}(\mathcal{S}) \le 2^{nd}.$$

The conclusion follows.

Proposition 5.3. *Let $A \in \mathrm{GL}_2(K_{\mathbb{A}})$ and let r be a positive integer. There exists a basis $\{\mathbf{y}_1, \mathbf{y}_2\}$ of $(\overline{k})^2$ with $[K(\mathbf{y}_i) : K] \le r$ for $i = 1, 2$ such that*

$$H_A(\mathbf{y}_1)H_A(\mathbf{y}_2) \le 2^{\delta}(r+1)^{\delta/r}C(K)^{2/r}|\det A|_{\mathbb{A}}.$$

Proof. Identify $S^r((K_{\mathbb{A}})^2)$ with $(K_{\mathbb{A}})^{r+1}$ via the isomorphism described in §1. View S^rA as an element of $\mathrm{GL}_{r+1}(K_{\mathbb{A}})$. By Lemma 5.2 there is a basis $\{\alpha_1, \dots, \alpha_{r+1}\}$ of $S^r(K^2)$ such that

$$H_{S^rA}(\alpha_1)\cdots H_{S^rA}(\alpha_{r+1}) \le (r+1)^{\delta(r+1)/2}C(K)^{r+1}|\det(S^rA)|_{\mathbb{A}}.$$

On the other hand, since $S(K^2)$ and $K[X, Y]$ are isomorphic as graded algebras and since each element of $K[X, Y]_r$ factors in $\overline{k}[X, Y]$ as a product of r linear forms, each element α of $S^r(K^2)$ also factors in $S((\overline{k})^2)$ as a product of r elements of $S^1((\overline{k})^2) = (\overline{k})^2$, each of them having coordinates in an extension of K of relative degree $\le r$. Thus, for $i = 1, \dots, r+1$ we can write

$$\alpha_i = \mathbf{x}_{i1}\cdots\mathbf{x}_{ir}, \tag{5.1}$$

where $\mathbf{x}_{i1}, \dots, \mathbf{x}_{ir} \in (\overline{k})^2$ have degree $\le r$ over K. For such a decomposition, Lemma 4.8 gives

$$H_A(\mathbf{x}_{i1})\cdots H_A(\mathbf{x}_{ir}) \le 2^{\delta r/2}H_{S^rA}(\alpha_i). \tag{5.2}$$

Lemma 4.5 also gives

$$|\det(S^rA)|_{\mathbb{A}} = |\det A|_{\mathbb{A}}^{r(r+1)/2}.$$

Thus, we get

$$\prod_{i=1}^{r+1}\prod_{j=1}^{r} H_A(\mathbf{x}_{ij}) \le \Big(2^{\delta}(r+1)^{\delta/r}C(K)^{2/r}|\det A|_{\mathbb{A}}\Big)^{r(r+1)/2}. \tag{5.3}$$

Among the points $\mathbf{x}_{ij}$ choose one for which $H_A(\mathbf{x}_{ij})$ is minimal; call it $\mathbf{y}_1$. For each $s = 1, \dots, r$, the set of elements α of $S^r((\overline{k})^2)$ which are divisible by $\mathbf{y}_1^s$ is a subspace of $S^r((\overline{k})^2)$ of dimension $r+1-s$. Therefore, at most $r+1-s$ elements among $\alpha_1, \dots, \alpha_{r+1}$ are divisible by $\mathbf{y}_1^s$. This implies that $\mathbf{x}_{ij}$ is a multiple of $\mathbf{y}_1$ for at most $r(r+1)/2$ pairs (i, j). Among the remaining points $\mathbf{x}_{ij}$ choose again one for which $H_A(\mathbf{x}_{ij})$ is minimal and call it

$\mathbf{y}_2$. Since $H_A(\mathbf{y}_2) \le H_A(\mathbf{x}_{ij})$ for at least $r(r+1)/2$ pairs (i,j) and since $H_A(\mathbf{y}_1) \le H_A(\mathbf{x}_{ij})$ for the remaining ones, the left hand side of (5.3) is at least $(H_A(\mathbf{y}_1)H_A(\mathbf{y}_2))^{r(r+1)/2}$. The conclusion follows.

Proof of Theorem 5.1. Let $\phi\colon (\overline{k})^2 \to (\overline{k})^n$ be a linear map defined over K whose image is V. By proposition 4.2 and Corollary 4.3, there exists $B \in \mathrm{GL}_2(K_{\mathbb{A}})$ such that $|\det B|_{\mathbb{A}} = H(V)$ and $H(\phi(\mathbf{y})) = H_B(\mathbf{y})$ for all $\mathbf{y} \in (\overline{k})^2$. Applying Proposition 5.3 with $A = B$, we find that there exists a basis $\{\mathbf{y}_1, \mathbf{y}_2\}$ of $(\overline{k})^2$ with $[K(\mathbf{y}_i) : K] \le r$ for $i = 1, 2$, such that

$$H_B(\mathbf{y}_1)H_B(\mathbf{y}_2) \le 2^\delta (r+1)^{\delta/r} C(K)^{2/r} H(V).$$

The image of $\{\mathbf{y}_1, \mathbf{y}_2\}$ under ϕ is a basis $\{\mathbf{x}_1, \mathbf{x}_2\}$ of V with $[K(\mathbf{x}_i) : K] = [k(\mathbf{y}_i) : K] \le r$ and $H(\mathbf{x}_i) = H_B(\mathbf{y}_i)$ for $i = 1, 2$. This basis of V thus verifies all the conditions of the theorem.

Remark. A more direct approach to Theorem 5.1 is to consider $S^r(V)$ as a subspace of $S^r((\overline{k})^n)$ of dimension $r+1$ defined over K. By Theorem 2.1, it admits a basis $\{\alpha_1, \dots, \alpha_{r+1}\}$ contained in $S^r(K^n)$ such that

$$H(\alpha_1)\cdots H(\alpha_{r+1}) \le \binom{r+n-1}{n-1}^{\delta(r+1)/2} C(K)^{r+1} H(S^r(V)).$$

Since there is an isomorphism of algebras $S(V) \simeq \overline{k}[X, Y]$ defined over K, we also have decompositions like (5.1) with $\mathbf{x}_{i1}, \dots, \mathbf{x}_{ir} \in S^1(V) = V$ of degree $\le r$ over K. But now (5.2) is replaced by

$$H(\mathbf{x}_{i1})\cdots H(\mathbf{x}_{ir}) \le n^{\delta r/2} H(\alpha_i).$$

Finally, Proposition 4.6 gives $H(S^r(V)) = H(V)^{r(r+1)/2}$. The rest of the argument is the same but the final constant is larger.

6. Successive minima

Definition 6.1. Given $A \in \mathrm{GL}_n(K_{\mathbb{A}})$ and an integer i with $1 \le i \le n$, we define the *i-th absolute minimum* $\mu_i(A)$ of A as the infimum of the set of all real numbers $\mu > 0$ for which there exist i linearly independent elements $\mathbf{x}_1, \dots, \mathbf{x}_i$ of $(\overline{k})^n$ with $H_A(\mathbf{x}_j) < \mu$ for $j = 1, \dots, i$.

By construction these numbers satisfy

$$0 \le \mu_1(A) \le \mu_2(A) \le \cdots \le \mu_n(A).$$

The purpose of this section is to show that their product $\mu_1(A)\cdots\mu_n(A)$ is bounded above and below by non-zero constant multiples of $|\det A|_{\mathbb{A}}$, with constants depending only on n. The first step is to establish an upper bound for the first minimum of A.

Proposition 6.2. *Assume $n \geq 2$ and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. Then*

(i) $\mu_1(A^*)^{n-1} \leq 2^{\delta(n-1)(n-2)/2}|\det A|_{\mathbb{A}}^{-1}\mu_1(A)$,

(ii) $\mu_1(A) \leq 2^{\delta(n-1)/2}|\det A|_{\mathbb{A}}^{1/n}$.

Proof. We proceed by induction on n. When $n = 2$ the duality theorem (Theorem 1.1) gives

$$H_{A^*}(V^{\perp}) = |\det A|_{\mathbb{A}}^{-1}H_A(V) \tag{6.1}$$

for all one-dimensional subspaces V of $(\overline{k})^2$, thus $\mu_1(A^*) = |\det A|_{\mathbb{A}}^{-1}\mu_1(A)$ and (i) is verified. On the other hand, when $n = 2$ inequality (ii) follows from Proposition 5.3.

Assume $n > 2$. Choose a real number $\epsilon > 0$ and a one-dimensional subspace $V \subseteq (\overline{k})^n$ with

$$H_A(V) \leq \mu_1(A) + \epsilon.$$

By duality we still have (6.1). Let $\phi\colon (\overline{k})^{n-1} \to (\overline{k})^n$ be an injective linear map whose image is $V^{\perp}$ and let E be a finite extension of K in $\overline{k}$ such that ϕ is defined over E. Proposition 4.2 and Corollary 4.3 show that there exists $B \in \mathrm{GL}_{n-1}(E_{\mathbb{A}})$ which satisfies

$$|\det B|_{\mathbb{A}} = H_{A^*}(V^{\perp})$$

and $H_B(\mathbf{y}) = H_{A^*}(\phi(\mathbf{y}))$ for all $\mathbf{y} \in (\overline{k})^{n-1}$. This implies

$$\mu_1(A^*) \leq \mu_1(B).$$

By induction on n, we may assume

$$\mu_1(B) \leq 2^{\delta(n-2)/2}|\det B|_{\mathbb{A}}^{1/(n-1)}.$$

Therefore, we get

$$\mu_1(A^*)^{n-1} \leq 2^{\delta(n-1)(n-2)/2}|\det A|_{\mathbb{A}}^{-1}\big(\mu_1(A) + \epsilon\big).$$

The choice of $\epsilon > 0$ being arbitrary, this proves (i).

If we apply inequality (i) to A^* we get

$$\mu_1(A)^{n-1} \leq 2^{\delta(n-1)(n-2)/2}|\det A|_{\mathbb{A}}\mu_1(A^*),$$

since $\det(A^*) = (\det A)^{-1}$. The inequality (ii) follows by combining this with (i), granted that $n > 2$.

Theorem 6.3. *Assume $n \geq 2$ and let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. Then*

$$|\det A|_{\mathbb{A}} \leq \prod_{i=1}^{n} \mu_i(A) \leq 2^{\delta n(n-1)/2} |\det A|_{\mathbb{A}}.$$

Proof. The lower bound follows from Lemma 4.7. It shows in particular that $\mu_1(A) > 0$. To prove the upper bound, choose $\epsilon > 0$ such that $\mu_i(A) + \epsilon < \mu_{i+1}(A)$ for each $i = 1, \ldots, n-1$ with $\mu_i(A) < \mu_{i+1}(A)$. There exist a finite extension E of K in $\overline{k}$ and a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of E^n such that

$$H_A(\mathbf{x}_i) \leq \mu_i(A) + \epsilon$$

for $i = 1, \ldots, n$. For the same integers i, let V_i be the subspace of $(\overline{k})^n$ spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_i$.

Fix a place $u_0 \in M(k)$ and let $b \in k$ with $|b|_{u_0} > 1$. By adjoining to E an N-th root of b for a sufficiently large integer $N \geq 1$, we may assume without loss of generality that there exist an element a of E and a place v_0 of E above u_0 such that

$$1 < |a|_{v_0} < 1 + \epsilon.$$

Put $t = [E_{v_0} : k_{u_0}]/[E : k]$. Since $0 < t < 1$, there exist integral powers $a_1, \ldots, a_n$ of a which satisfy $|a_1|_{v_0} \geq \cdots \geq |a_n|_{v_0}$ and

$$1 \leq |a_i|_{v_0}^t \mu_i(A) \leq 1 + \epsilon$$

for $i = 1, \ldots, n$. For the same values of i, let W_i be the subspace of $(E_{v_0})^n$ generated by $A_{v_0}(\mathbf{x}_1), \ldots, A_{v_0}(\mathbf{x}_i)$. Take $\phi \in \mathrm{GL}_n(E_{v_0})$ as in Lemma 3.3 according to this choice of subspaces W_i and numbers $a_i \in E$. Define $B \in \mathrm{GL}_n(E_{\mathbb{A}})$ by

$$B_v = \begin{cases} A_v & \text{if } v \in M(E) \text{ and } v \neq v_0, \\ \phi \circ A_{v_0} & \text{if } v = v_0. \end{cases}$$

By virtue of this choice of ϕ, we get

$$|\det B|_{\mathbb{A}} = |\det A|_{\mathbb{A}} \, |\det \phi|_{v_0}^t \leq (1+\epsilon)^n |\det A|_{\mathbb{A}} \prod_{i=1}^{n} \mu_i(A)^{-1}.$$

We claim that $\mu_1(B) \geq 1$.

To verify this, take any non-zero $\mathbf{x} \in (\overline{k})^n$. Choose a finite algebraic extension F of E such that $\mathbf{x} \in F^n$ and let i be the smallest integer such that $\mathbf{x} \in V_i$. For all $w \in M(F)$ with $w|v_0$, we have

$$\|B_w(\mathbf{x})\|_w \geq |a_i|_{v_0} \|A_w(\mathbf{x})\|_w$$

by virtue of the choice of ϕ, since $A_w(\mathbf{x})$ belongs to the subspace of $(F_w)^n$ spanned by W_i. For any other $w \in M(F)$ we have $\|B_w(\mathbf{x})\|_w = \|A_w(\mathbf{x})\|_w$, since $B_w = A_w$. This implies

$$\begin{aligned} H_B(\mathbf{x}) &\geq \Big(\prod_{w|v_0} |a_i|_{v_0}^{[F_w : k_{u_0}]/[F:k]} \Big) H_A(\mathbf{x}) \\ &\geq |a_i|_{v_0}^t H_A(\mathbf{x}) \\ &\geq \mu_i(A)^{-1} H_A(\mathbf{x}). \end{aligned}$$

To conclude from this that $H_B(\mathbf{x}) \geq 1$, it remains only to show $H_A(\mathbf{x}) \geq \mu_i(A)$. Let j be the smallest integer such that $\mu_j(A) = \mu_i(A)$. It is equivalent to show $H_A(\mathbf{x}) \geq \mu_j(A)$. If $j = 1$, this inequality follows from the fact that $\mathbf{x}$ is $\neq 0$. If $j \geq 2$, then $\mathbf{x} \notin V_{j-1}$ and the j vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}, \mathbf{x}$ are linearly independent over $\overline{k}$. Thus, one of them has height at least $\mu_j(A)$. Since $\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}$ all have height no greater than $\mu_{j-1}(A) + \epsilon$ and since $\mu_{j-1}(A) + \epsilon < \mu_j(A)$, we must have $H_A(\mathbf{x}) \geq \mu_j(A)$. Hence $H_B(\mathbf{x}) \geq 1$. The choice of $\mathbf{x} \in (\overline{k})^n$ with $\mathbf{x} \neq 0$ being arbitrary, this proves $\mu_1(B) \geq 1$.

By Proposition 6.2 we have $\mu_1(B) \leq 2^{\delta(n-1)/2} |\det B|_{\mathbb{A}}^{1/n}$. Using the lower bound for $\mu_1(B)$ and the upper bound for $|\det B|_{\mathbb{A}}$ established above, this gives

$$1 \leq (1+\epsilon)^n 2^{\delta n(n-1)/2} |\det A|_{\mathbb{A}} \prod_{i=1}^{n} \mu_i(A)^{-1}.$$

The conclusion follows by letting ϵ tend to 0.

7. Duality

Proposition 6.2 shows a relation between the first absolute minimum of an element A of $\mathrm{GL}_n(K_{\mathbb{A}})$ and the first absolute minimum of A^*. In this section, we show more generally that all the absolute minima of A are closely related to those of A^*.

Theorem 7.1. *Let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$. For $i = 1, \ldots, n$ we have*

$$1 \leq \mu_i(A)\mu_{n+1-i}(A^*) \leq 2^{\delta n(n-1)}.$$

These inequalities are analogous to results of Mahler on polar lattices ([8], see also §VIII.5 of [4]). We first establish a lemma.

Lemma 7.2. *Let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$ and let b be the canonical bilinear form of $(\overline{k})^n$ (see §1). For each $\mathbf{u} \in (\overline{k})^n$ and each $\mathbf{v} \in (\overline{k})^n$ with $b(\mathbf{u}, \mathbf{v}) \neq 0$ we have $H_A(\mathbf{u})H_{A^*}(\mathbf{v}) \geq 1$.*

Proof. Let $\mathbf{u}, \mathbf{v} \in (\overline{k})^n$ with $b(\mathbf{u}, \mathbf{v}) \neq 0$ and let E be a finite extension of K in $\overline{k}$ such that $\mathbf{u}, \mathbf{v} \in E^n$. For each $w \in M(E)$, we denote by b_w the canonical bilinear form of $(E_w)^n$. It satisfies

$$|b_w(\mathbf{x}, \mathbf{y})|_w \leq \|\mathbf{x}\|_w \|\mathbf{y}\|_w \quad \text{and} \quad b_w(A_w(\mathbf{x}), A_w^*(\mathbf{y})) = b_w(\mathbf{x}, \mathbf{y})$$

for each $(\mathbf{x}, \mathbf{y}) \in (E_w)^n \times (E_w)^n$. Since $b(\mathbf{u}, \mathbf{v})$ is a non-zero element of E, the product formula together with the above relations give

$$\begin{aligned}
1 &= \prod_{w \in M(E)} \Big(|b_w(\mathbf{u}, \mathbf{v})|_w \Big)^{n_w/[E:k]} \\
&= \prod_{w \in M(E)} \Big(|b_w(A_w(\mathbf{u}), A_w^*(\mathbf{v}))|_w \Big)^{n_w/[E:k]} \\
&\leq \prod_{w \in M(E)} \Big(\|A_w(\mathbf{u})\|_w \|A_w^*(\mathbf{v})\|_w \Big)^{n_w/[E:k]} \\
&= H_A(\mathbf{u}) H_{A^*}(\mathbf{v}).
\end{aligned}$$

Proof of Theorem 7.1. To prove the lower bound, choose $\epsilon > 0$ arbitrarily small and, according to this choice of ϵ, choose bases $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of $(\overline{k})^n$ such that

$$H_A(\mathbf{u}_j) \leq \mu_j(A) + \epsilon \quad \text{and} \quad H_{A^*}(\mathbf{v}_j) \leq \mu_j(A^*) + \epsilon \tag{7.1}$$

for $j = 1, \dots, n$. Fix also an integer i with $1 \leq i \leq n$. Since the canonical bilinear form b of $(\overline{k})^n$ is non-degenerate, there exist integers r, s with $1 \leq r \leq i$ and $1 \leq s \leq n + 1 - i$ such that $b(\mathbf{u}_r, \mathbf{v}_s) \neq 0$. By virtue of Lemma 7.2 and inequalities (7.1), this implies

$$1 \leq H_A(\mathbf{u}_r) H_{A^*}(\mathbf{v}_s) \leq \big(\mu_i(A) + \epsilon\big)\big(\mu_{n+1-i}(A^*) + \epsilon\big).$$

Hence $\mu_i(A)\mu_{n+1-i}(A^*) \geq 1$ for $i = 1, \dots, n$.

By virtue of Theorem 6.3, we have

$$\begin{aligned}
\prod_{i=1}^{n} \big(\mu_i(A)\mu_{n+1-i}(A^*)\big) &= \Big(\prod_{i=1}^{n} \mu_i(A)\Big)\Big(\prod_{i=1}^{n} \mu_i(A^*)\Big) \\
&\leq 2^{\delta n(n-1)} |\det A|_{\mathbb{A}} |\det A^*|_{\mathbb{A}} \\
&\leq 2^{\delta n(n-1)},
\end{aligned}$$

using $\det(A^*) = (\det A)^{-1}$. Since all products $\mu_i(A)\mu_{n+1-i}(A^*)$ which enter in the left hand side of this inequality are at least 1, these products must also be no greater than $2^{\delta n(n-1)}$.

8. Small flags

We first prove the following generalization of Theorem 2.2.

Theorem 8.1. *Let V be a subspace of $(\overline{k})^n$ of dimension m, let $A \in \mathrm{GL}_n(K_{\mathbb{A}})$ and let $\epsilon > 0$. Then there exists a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V such that*

$$H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_m) \leq (2^{\delta m(m-1)/2} + \epsilon) H_A(V).$$

Note that Lemma 4.7 gives $H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_m) \geq H_A(V)$ for any basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V.

Proof. Without loss of generality, we may assume that V is defined over K. Then there exists a linear map $\phi\colon (\overline{k})^m \to (\overline{k})^n$ which is defined over K and whose image is V. According to this choice of ϕ, Proposition 4.2 and Corollary 4.3 show that there exists $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $|\det B|_{\mathbb{A}} = H_A(V)$ and $H_A(\phi(\mathbf{y})) = H_B(\mathbf{y})$ for all $\mathbf{y} \in (\overline{k})^m$. Fix $\eta > 0$ and choose a basis $\{\mathbf{y}_1, \ldots, \mathbf{y}_m\}$ of $(\overline{k})^m$ which satisfies

$$H_B(\mathbf{y}_i) \leq (1 + \eta)\mu_i(B)$$

for $i = 1, \ldots, m$. Its image under ϕ is a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ of V. Applying Theorem 6.3 to get an upper bound for $\mu_1(B) \cdots \mu_m(B)$, we find

$$\begin{aligned} H_A(\mathbf{x}_1) \cdots H_A(\mathbf{x}_m) &= H_B(\mathbf{y}_1) \cdots H_B(\mathbf{y}_m) \\ &\leq (1+\eta)^m \mu_1(B) \cdots \mu_m(B) \\ &\leq (1+\eta)^m 2^{\delta m(m-1)/2} |\det B|_{\mathbb{A}} \\ &= (1+\eta)^m 2^{\delta m(m-1)/2} H_A(V). \end{aligned}$$

For sufficiently small η, this basis of V thus satisfies the inequality of the theorem.

Theorem 2.2 follows from the above result by taking A to be the identity element of $\mathrm{GL}_n(K_{\mathbb{A}})$. Another consequence of Theorem 8.1 is the existence of a complete flag of subspaces of V with relatively small height.

Corollary 8.2. *Let the notations be as in Theorem 8.1. There exists a nested sequence of subspaces of V*

$$\{\mathbf{0}\} = V_0 \subset V_1 \subset \cdots \subset V_m = V$$

with $\dim(V_i) = i$ *and*

$$H_A(V_i) \le (2^{\delta i(m-1)/2} + \epsilon) H_A(V)^{i/m}$$

for $i = 0, \dots, m$.

Proof. Let $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ be a basis of V satisfying the inequality of Theorem 8.1. Without loss of generality, we may assume $H_A(\mathbf{x}_1) \le \cdots \le H_A(\mathbf{x}_m)$. For $i = 1, \dots, m$, define V_i as the subspace of V spanned by $\mathbf{x}_1, \dots, \mathbf{x}_i$, and put $V_0 = \{\mathbf{0}\}$. Using Lemma 4.7, we find

$$H_A(V_i) \le \prod_{j=1}^{i} H_A(\mathbf{x}_j) \le \Big(\prod_{j=1}^{m} H_A(\mathbf{x}_j) \Big)^{i/m} \le (2^{\delta i(m-1)/2} + \epsilon) H_A(V)^{i/m}$$

for $i = 1, \dots, m$. Since $H_A(V_0) = 1$, this inequality also holds for $i = 0$.

Finally, we prove the following absolute version over $\overline{k}$ of Theorem 1 of [3].

Theorem 8.3. *Let W and V be distinct subspaces of $(\overline{k})^n$ with $W \subset V$ and let $\epsilon > 0$. Put $l = \dim W$ and $m = \dim V$. Then there exists a nested sequence of subspaces of $(\overline{k})^n$*

$$W = V_l \subset V_{l+1} \subset \cdots \subset V_m = V$$

such that $\dim(V_i) = i$ *and*

$$H(V_i) \le (2^{\delta(m-i)(m-l-1)/2} + \epsilon) H(W)^{(m-i)/(m-l)} H(V)^{(i-l)/(m-l)}$$

for $i = l, \dots, m$.

Proof. Let $\phi\colon (\overline{k})^m \to (\overline{k})^n$ be a linear map whose image is V and let K be a finite extension of k such that ϕ is defined over K. By Corollary 4.3, there exists $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $H(\phi(S)) = H_B(S)$ for all subspaces S of $(\overline{k})^m$. Put $S = \phi^{-1}(W)$. We get

$$H_B(S) = H(W) \quad \text{and} \quad |\det B|_{\mathbb{A}} = H_B((\overline{k})^m) = H(V).$$

Since $\dim(S^\perp) = m-l$, Corollary 8.2 shows that there exists a nested sequence of subspaces of $S^\perp$

$$\{\mathbf{0}\} = T_0 \subset T_1 \subset \cdots \subset T_{m-l} = S^\perp$$

with $\dim(T_i) = i$ and

$$H_{B^*}(T_i) \le (2^{\delta i(m-l-1)/2} + \epsilon) H_{B^*}(S^\perp)^{i/(m-l)}$$

for $i = 0, \ldots, m-l$. Put

$$V_i = \phi\big((T_{m-i})^\perp\big)$$

for $i = l, \ldots, m$. Then these subspaces form a nested sequence

$$W = V_l \subset V_{l+1} \subset \cdots \subset V_m = V$$

with $\dim(V_i) = i$. Using Theorem 1.1 and the preceding relations, we get

$$\begin{aligned}
H(V_i) &= H_B\big((T_{m-i})^\perp\big) \\
&= H_{B^*}(T_{m-i})\,|\det B|_{\mathbb{A}} \\
&\le (2^{\delta(m-i)(m-l-1)/2} + \epsilon) H_{B^*}(S^\perp)^{(m-i)/(m-l)} |\det B|_{\mathbb{A}} \\
&= (2^{\delta(m-i)(m-l-1)/2} + \epsilon) H_B(S)^{(m-i)/(m-l)} \big(|\det B|_{\mathbb{A}}\big)^{1-(m-i)/(m-l)} \\
&= (2^{\delta(m-i)(m-l-1)/2} + \epsilon) H(W)^{(m-i)/(m-l)} H(V)^{(i-l)/(m-l)}.
\end{aligned}$$

9. A canonical filtration

In the preceding section we showed the existence of a complete flag of subspaces of relatively small height for a given subspace V of $(\overline{k})^n$. In this section, we consider a filtration of V by subspaces that are, in a sense, canonically associated to it.

Definition 9.1. Let V be a subspace of $(\overline{k})^n$ of dimension m defined over K.
(i) For each real number $\mu > 0$, we define $V(\mu)$ as the subspace of V generated by the set of all elements $\mathbf{x}$ of V with $H(\mathbf{x}) < \mu$.
(ii) For each integer $i = 1, \ldots, m$, we define the *i-th absolute minimum* $\mu_i(V)$ of V as the infimum of the set of all real numbers $\mu > 0$ for which there exist i linearly independent elements $\mathbf{x}$ of V with $H(\mathbf{x}) < \mu$.

Equivalently, $\mu_i(V)$ is the infimum of the set of all real numbers $\mu > 0$ for which $\dim(V(\mu)) \le i$. From the above definitions we have

$$1 \le \mu_1(V) \le \mu_2(V) \le \cdots \le \mu_m(V)$$

and also

$$V(\mu_1) \subseteq V(\mu_2)$$

whenever $\mu_1 \le \mu_2$.

Proposition 9.2. *Let V be a subspace of $(\overline{k})^n$ of dimension m defined over K. Then*

$$\mu_1(V)\cdots\mu_m(V) \le 2^{\delta m(m-1)/2} H(V). \tag{9.1}$$

Moreover, for each $\mu > 0$ the subspace $V(\mu)$ of V is defined over a purely inseparable extension of K and satisfies

$$H(V(\mu)) \le 2^{\delta m_\mu (m-1)/2} H(V)^{m_\mu/m}, \tag{9.2}$$

where $m_\mu = \dim(V(\mu))$.

Proof. Let $\phi\colon (\overline{k})^m \to (\overline{k})^n$ be a linear map which is defined over K and whose image is V. As in Proposition 4.2 and Corollary 4.3, choose $B \in \mathrm{GL}_m(K_{\mathbb{A}})$ such that $|\det B|_{\mathbb{A}} = H(V)$ and $H(\phi(\mathbf{y})) = H_B(\mathbf{y})$ for all $\mathbf{y} \in (\overline{k})^m$. On the basis of Definitions 6.1 and 9.1, this implies $\mu_i(V) = \mu_i(B)$ for $i = 1, \dots, m$. Therefore, the upper bound (9.1) follows from Theorem 6.3 applied to B.

Let $\mu > 0$ be fixed. Each automorphism of $\overline{k}$ over K maps $V(\mu)$ into itself since it permutes the set of all $\mathbf{x} \in V$ with $H(\mathbf{x}) < \mu$. Thus, $V(\mu)$ is defined over a purely inseparable extension of K. Let $i = \dim(V(\mu))$. If $i = 0$, both sides of (9.2) are equal to 1. Assume $i > 0$. Then we have $\mu_i(V) < \mu$. Choose $\epsilon > 0$ such that $(1+\epsilon)\mu_i(V) < \mu$ and, according to that choice of ϵ, choose linearly independent elements $\mathbf{x}_1, \dots, \mathbf{x}_i$ of V with

$$H(\mathbf{x}_j) \le (1+\epsilon)\mu_j(V)$$

for $j = 1, \dots, i$. By construction, these elements belong to $V(\mu)$. Thus, they constitute a basis of $V(\mu)$. Using Lemma 4.7 we deduce

$$H(V(\mu)) \le \prod_{j=1}^{i} H(\mathbf{x}_j) \le (1+\epsilon)^i \prod_{j=1}^{i} \mu_j(V) \le (1+\epsilon)^i \Big(\prod_{j=1}^{m} \mu_j(V) \Big)^{i/m}.$$

Thus, by (9.1)

$$H(V(\mu)) \le (1+\epsilon)^i 2^{\delta i(m-1)/2} H(V)^{i/m}.$$

Inequality (9.2) follows from this, granted that ϵ can be chosen arbitrarily small.

10. A lower bound for the constant

As mentioned above, the constant appearing in the inequality of Theorem 2.2 is optimal when $k = k_0(T)$. When $k = \mathbb{Q}$ we do not know what the optimal constant is. The following result shows that it must grow at least exponentially with m.

Proposition 10.1. *Assume $m < n$. There is a subspace V of $(\overline{\mathbb{Q}})^n$, defined over $\mathbb{Q}$ of dimension m, which satisfies*

$$H(\mathbf{x}_1)\cdots H(\mathbf{x}_m) \geq (m+1)^{-1/2}2^{m/2}H(V)$$

for all bases $\{\mathbf{x}_1,\ldots,\mathbf{x}_m\}$ of V over $\overline{\mathbb{Q}}$.

The proof requires two lemmas.

Lemma 10.2. *Let d be a positive integer and let c be a positive real number. For any choice of positive real numbers $a_1,\ldots,a_d$ with product $a_1\cdots a_d \geq c$, we have*

$$(1+a_1)\cdots(1+a_d) \geq (1+c^{1/d})^d.$$

Proof. Let $f(x) = \log(1+e^x)$. This is a convex function on $\mathbb{R}$ since its second derivative $f''(x) = e^x(1+e^x)^{-2}$ is positive for all $x \in \mathbb{R}$. Let $b_i = \log a_i$ for $i = 1,\ldots,d$ and let $b = (b_1+\cdots+b_d)/d$. Since f is convex, we get

$$(f(b_1)+\cdots+f(b_d))/d \geq f(b).$$

The conclusion follows by taking the exponential of both sides of this inequality, granted that $b \geq (\log c)/d$.

Lemma 10.3. *Any point of $(\overline{\mathbb{Q}})^n$ with more than one non-zero coordinate has height at least $\sqrt{2}$.*

Proof. It suffices to show $H(1,x) \geq \sqrt{2}$ for all non-zero $x \in \overline{\mathbb{Q}}$. If x is not an algebraic integer, we have $H(1,x) \geq a \geq 2$, where a denotes the leading coefficient of the irreducible polynomial of x over $\mathbb{Z}$ normalized so that $a > 0$. Therefore, we may assume that x is an algebraic integer. Let d be its degree over $\mathbb{Q}$ and let $x^{(1)},\ldots,x^{(d)}$ be the conjugates of x in $\mathbb{C}$. Since the norm of x is a non-zero integer, we have $|x^{(1)}|\cdots|x^{(d)}| \geq 1$, and by Lemma 10.2 we find

$$H(1,x) = \prod_{i=1}^{d}(1+|x^{(i)}|^2)^{1/(2d)} \geq \sqrt{2}.$$

Proof of Proposition 10.1. Let V be the subspace of $(\overline{\mathbb{Q}})^n$ defined by the equations

$$x_1+x_2+\cdots+x_{m+1} = 0 \quad \text{and} \quad x_{m+2} = \cdots = x_n = 0$$

Then the height of V is $\sqrt{m+1}$ and V contains no point with only one non-zero coordinate. The conclusion follows from Lemma 10.3.

References

1. E. Artin, *Algebraic Numbers and Algebraic Functions*, Gordon and Breach, New York, 1967.
2. E. Bombieri and J. Vaaler, *On Siegel's Lemma*, Invent. Math. **73** (1983), 11–32.
3. E. B. Burger and J. D. Vaaler, *On the Decomposition of Vectors over Number Fields*, J. reine angew. Math. **435** (1993), 197–219.
4. J. W. S. Cassels, *An Introduction to the Geometry of Numbers*, Springer-Verlag, Berlin, 1959.
5. W. H. Greub, *Multilinear algebra*, Springer-Verlag, New York, 1967.
6. S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, New York, 1983.
7. R. B. Macfeat, *Geometry of Numbers in Adele Spaces*, Dissertationes Mathematicae (Rozprawy Matematyczne) LXXXVIII (1971), 54 pp.
8. K. Mahler, *Ein Übertragungsprinzip für konvexe Körper*, Časopis Pěst. Mat. Fys. **68** (1939), 93–102.
9. ———, *On some Inequalities for Polynomials in several Variables*, J. London Math. Soc. **37** (1962), 341–344.
10. D. Roy and J. Thunder, *A Note on Siegel's Lemma over Number Fields*, Monatsh. Math. (to appear).
11. W. M. Schmidt, *On Heights of algebraic Subspaces and Diophantine Approximations*, Annals of Math. **85** (1967), 430–472.
12. ———, *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Math., vol. 1467, Springer-Verlag, New York, 1991.
13. C. L. Siegel, *Über einige Anwendungen diophantischer Approximationen*, Abh. der Preuß. Akad. der Wissenschaften, Phys.-math. Kl., Nr.1 (1929) (=Ges. Abh., I, 209–266).
14. J. H. Silverman, *The Theory of Height Functions*, Chapter VI in: Arithmetic Geometry (G. Cornell and J.H. Silverman, eds.), Springer Verlag, New York, 1986.
15. A. Thue, *Über Annäherungswerte algebraischer Zahlen*, J. reine angew. Math. **135** (1909), 284-305.
16. J. L. Thunder, *Asymptotic Estimates for Rational Points of Bounded Height on Flag Varieties*, Comp. Math. **88** (1993), 155–186.
17. J. Thunder, *Siegel's Lemma for Function Fields*, Michigan Math. J. **42** (1995), 147–162.
18. A. Weil, *Basic Number Theory*, Springer verlag, New York, 1974.

Dépt. de Math., Univ. d'Ottawa, Ottawa, Ontario, Canada K1N 6N5
E-mail address: droy@acadvm1.uottawa.ca

Dept. of Math., Northern Illinois Univ., DeKalb, IL 60115
E-mail address: jthunder@math.niu.edu