A VARIATION ON SIEGEL'S LEMMA

JAN HENDRIK EVERTSE

Appendix to the paper:
Quantitative Diophantine approximations on projective varieties
by Roberto G. Ferretti

1. Introduction

In many Diophantine approximation proofs, a major step is to construct a polynomial, a global section of a given line bundle, or some other type of auxiliary function with certain prescribed properties. In general this can be translated into the problem to find a non-zero n-dimensional vector of small height with coordinates in some algebraic number field K lying in some prescribed linear subspace of K^n. There are various results implying the existence of such a vector, see for instance Bombieri and Vaaler [1, Thm. 9]. These results are extensions of the so-called Siegel's Lemma, which states that a given system of m homogeneous linear equations with integer coefficients in $n > m$ unknowns has a non-zero solution in integers of small absolute value. Siegel was the first to state this formally ([11, Band I, p. 213]), but it was already implicitly proved by Thue ([12, pp. 288-289]).

In this note we will deduce the version of Siegel's lemma used by Ferretti in [7, Section 6]. Roughly speaking, the problem encountered by Ferretti is the following. Denote by O_K the ring of integers of K and define the size of $x \in O_K$ to be the maximum of the absolute values of the conjugates of x. Let I be a non-zero ideal of the polynomial ring $K[X_0, \dots, X_N]$ and let $\{f_{i1}, \dots, f_{i,n_i}\} \subset K[X_0, \dots, X_N]$ $(i = 1, \dots, s)$ be given sets of polynomials. Find numbers $x_{ij} \in O_K$ of small size, not all equal to 0, such that

$$\sum_{i=1}^{n_1} x_{1j} f_{1j} \equiv \cdots \equiv \sum_{i=1}^{n_s} x_{sj} f_{sj} \pmod{I}.$$

This can be translated into the following problem. Suppose we are given a linear subspace W of K^h and linearly independent sets of vectors $\{\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i}\}$ $(i =$

$1, \dots, s$) in the quotient space K^h/W. Show that there are numbers $x_{ij} \in O_K$ of small size, not all equal to 0, such that $\sum_{j=1}^{n_1} x_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s} x_{sj}\mathbf{b}_{sj}$.

We show that under some natural hypotheses there exist such numbers x_{ij} with sizes below some explicit bound depending on K, $n = \dim K^h/W$, the height of W and the norms of the vectors $\mathbf{b}_{ij}$ (cf. Theorem 2.2). It is essential for Ferretti's purposes, that in the special case of our result needed by him, our bound has a polynomial dependence on n. The precise statement of our result is given in the next section.

Our main tool is the result of Bombieri and Vaaler mentioned above. Our upper bound will have a dependence on the number field K. We will also prove an "absolute" result in which the upper bound for the sizes of the numbers x_{ij} is independent of K but in which the numbers x_{ij} may lie in some unspecified algebraic extension of K. To deduce the absolute result we replace the Bombieri-Vaaler theorem by a result of Zhang [15, Thm. 5.2] (see also Roy and Thunder [9, Thm. 2.2], [10, Thm. 1] for a weaker result).

We mention that our proof is not completely straightforward. By a more obvious application of the result of Bombieri and Vaaler we would have obtained a "basis-independent" result, giving upper bounds for the sizes of the coordinates of the vectors $\sum_{j=1}^{n_i} x_{ij}\mathbf{b}_{ij}$, rather than for the numbers x_{ij} themselves. Then subsequently we could have deduced upper bounds for the sizes of the numbers x_{ij} by invoking Cramer's rule, but due to the various determinant estimates the resulting bounds would have had a dependence on n of the order $n!$. This would have been useless for Ferretti's application mentioned above, which required upper bounds for the sizes of the x_{ij} depending at most polynomially on n. Therefore we had to use a more subtle argument which avoids the use of Cramer's rule.

2. The main result

2.1. We introduce some notation. The transpose of a matrix A is denoted by A^t. Given any ring R, we denote by R^n the module of n-dimensional column vectors with coordinates in R. Let k, n be integers with $1 \leqslant k \leqslant n$ and put $T := \binom{n}{k}$. Denote by $I_1, \dots, I_T$ the subsets of $\{1, \dots, n\}$ of cardinality k, in some given order. Then we define the exterior product of $\mathbf{a}_1 = (a_{11}, \dots, a_{1n})^t, \dots, \mathbf{a}_k = (a_{k1}, \dots, a_{kn})^t \in R^n$ by

$$\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k := (A_1, \dots, A_T)^t,$$

where A_l is defined such that if $I_l = \{i_1, \ldots, i_k\}$ with $i_1 < i_2 < \cdots < i_k$ then $A_l = \det \left(a_{p,i_q}\right)_{p,q=1,\ldots,k}$. Thus, if $\mathbf{b}_i = \sum_{j=1}^k \xi_{ij}\mathbf{a}_j$ for $i = 1, \ldots, k$ with $\xi_{ij} \in R$, then

$$\mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_k = \det \left(\xi_{ij}\right)_{i,j=1,\ldots,k} \cdot \mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k \,. \tag{2.1}$$

Let K be an algebraic number field. Denote by O_K the ring of integers, by Δ_K the discriminant, and by M_K the set of places of K. We have $M_K = M_K^\infty \cup M_K^0$ where M_K^∞ is the set of infinite places and M_K^0 the set of finite places of K. For $v \in M_K$ we denote by K_v the completion of K at v. The infinite places are divided into real places (i.e., with $K_v = \mathbf{R}$) and complex places (with $K_v = \mathbf{C}$).

Put $d := [K : \mathbf{Q}]$ and $d_v := [K_v : \mathbf{Q}_p]$ for $v \in M_K$, where p is the place of $\mathbf{Q}$ lying below v and $\mathbf{Q}_p$ is the completion of $\mathbf{Q}$ at p. In particular, $d_v = 1$ if v is a real place while $d_v = 2$ if v is a complex place. Denote by r_1 the number of real places and by r_2 the number of complex places of K; then $r_1 + 2r_2 = \sum_{v\in M_K^\infty} d_v = d$.

For $v \in M_K$ we choose the absolute value $|\cdot|_v$ on K_v representing v such that if v is infinite then $|\cdot|_v$ extends the standard absolute value, while if v is finite and lies above the prime number p, then $|\cdot|_v$ extends the standard p-adic absolute value, i.e. with $|p|_p = p^{-1}$. These absolute values satisfy the product formula $\prod_{v\in M_K} |x|_v^{d_v} = 1$ for $x \in K^*$. For $x \in K$ we have

$$\max_{v\in M_K^\infty} |x|_v = \max(|x^{(1)}|, \ldots, |x^{(d)}|)$$

where $x^{(1)}, \ldots, x^{(d)}$ are the conjugates of x.

We now define norms and heights. Put

$$\|\mathbf{x}\|_v := \Big(\sum_{i=1}^n |x_i|_v^2\Big)^{1/2} \quad \text{for } v \in M_K^\infty,\ \mathbf{x} \in K_v^n$$
$$\|\mathbf{x}\|_v := \max(|x_1|_v, \ldots, |x_n|_v) \quad \text{for } v \in M_K^0,\ \mathbf{x} \in K_v^n$$

where $\mathbf{x} = (x_1, \ldots, x_n)^t$. Then the absolute height of $\mathbf{x} \in K^n$ is given by

$$H(\mathbf{x}) := \prod_{v\in M_K} \|\mathbf{x}\|_v^{d_v/d}.$$

By the product formula we have $H(\lambda\mathbf{x}) = H(\mathbf{x})$ for $\lambda \in K^*$.

More generally, we define the height of a linear subspace V of K^n by $H(V) = 1$ if $V = (\mathbf{0})$ and

$$H(V) := H(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_k)$$

if $V \neq (\mathbf{0})$ where $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ is any basis of V. By (2.1) and the product formula, this is well-defined, i.e., independent of the choice of the basis.

An M_K*-constant* is a tuple of constants $C = \{C_v : v \in M_K\}$ with $C_v > 0$ for $v \in M_K$ and with $C_v = 1$ for all but finitely many v.

For a linear subspace V of K^n and a field extension L of K we denote by $V \otimes_K L$ the L-linear subspace of L^n generated by V. Given any finite extension L of K we define $O_L, M_L, M_L^\infty, M_L^0, |\cdot|_w, \|\cdot\|_w$ $(w \in M_L)$ completely similarly as for K.

Lastly, for $v \in M_K$ and for any proper linear subspace W of K^h, we denote by $\rho_{W,v}$ the canonical map from K_v^h to $K_v^h/(W \otimes_K K_v)$. Further, for $\mathbf{x} \in K_v^h/(W \otimes_K K_v)$ we put

$$\|\mathbf{x}\|_v^W := \inf\{\|\mathbf{x}^*\|_v : \ \mathbf{x}^* \in K_v^h, \ \rho_{W,v}(\mathbf{x}^*) = \mathbf{x}\}.$$

Then the precise statement of the result mentioned in the introduction reads as follows.

Theorem 2.2. *Let h be a positive integer, let W be a proper linear subspace of K^h and let $C = \{C_v : v \in M_K\}$ be an M_K-constant. Further, let $V_1, \dots, V_s$ $(s \geqslant 2)$ be linear subspaces of K^h/W such that*

$$\dim(V_1 + \cdots + V_s) \ =: n > 0, \tag{2.2}$$

$$\dim(V_1 \cap \cdots \cap V_s) \ =: m > 0 \tag{2.3}$$

and such that for $i = 1, \dots, s$, V_i has a basis $\{\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i}\}$ with

$$\|\mathbf{b}_{ij}\|_v^W \leqslant C_v \quad \text{for } j = 1, \dots, n_i, \ v \in M_K. \tag{2.4}$$

Lastly, let U be the inverse image of $V_1 + \cdots + V_s$ under the canonical map from K^h to K^h/W.

Then there are $x_{ij} \in O_K$ $(i = 1, \dots, s, \ j = 1, \dots, n_i)$, not all 0, such that

$$\sum_{j=1}^{n_1} x_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s} x_{sj}\mathbf{b}_{sj}, \tag{2.5}$$

$$\max_{v \in M_K^\infty} |x_{ij}|_v \leqslant \Big(\frac{2}{\pi}\Big)^{2r_2/d} |\Delta_K|^{1/d} \cdot \Big\{(ns)^{n/2}\Big(\prod_{v \in M_K} C_v^{d_v/d}\Big)^n \cdot \frac{H(W)}{H(U)}\Big\}^{(s-1)/m} \tag{2.6}$$

for $i = 1, \dots, s, \ j = 1, \dots, n_i$.

Moreover, there are a finite extension L of K and numbers $x_{ij} \in O_L$ ($i = 1, \dots, s$, $j = 1, \dots, n_i$), not all 0, satisfying (2.5) (viewed as indentities in $L^h/(W \otimes_K L)$) and

$$\max_{w \in M_L^\infty} |x_{ij}|_w \leqslant m^{1/2} \cdot \Big\{ (ns)^{n/2} \Big(\prod_{v \in M_K} C_v^{d_v/d} \Big)^n \cdot \frac{H(W)}{H(U)} \Big\}^{(s-1)/m} \tag{2.7}$$

$$\text{for } i = 1, \dots, s,\ j = 1, \dots, n_i.$$

Remark. This result is applied by Ferretti for n, m satisfying $n/m \leqslant 4/3$. In this case, the upper bounds in (2.6), (2.7) depend polynomially on n.

3. An auxiliary result

3.1. We state an auxiliary result dealing with vectors in K^h (i.e., not in a quotient space) but with modified norms. From this result we will deduce Theorem 2.2. We keep the notation introduced before. In addition, an *M_K-matrix of order n* is a tuple of matrices $D = \{D_v : v \in M_K\}$ with $D_v \in GL_n(K_v)$ for $v \in M_K$ and with $|\det D_v|_v = 1$ for all but finitely many v.

Theorem 3.2. *Let n be a positive integer. Let $D = \{D_v : v \in M_K\}$ be an M_K-matrix of order n. Assume that K^n has a basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ with*

$$\|D_v \mathbf{b}_i\|_v \leqslant 1 \quad \text{for } i = 1, \dots, n,\ v \in M_K. \tag{3.1}$$

Further, let $V_1, \dots, V_s$ ($s \geqslant 2$) be linear subspaces of K^n such that

$$\dim(V_1 \cap \dots \cap V_s) =: m > 0 \tag{3.2}$$

and such that for $i = 1, \dots, s$, V_i has a basis $\{\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i}\}$ with

$$\|D_v \mathbf{b}_{ij}\|_v \leqslant 1 \quad \text{for } j = 1, \dots, n_i,\ v \in M_K. \tag{3.3}$$

Then there are $x_{ij} \in O_K$ $(i = 1, \dots, s,\ j = 1, \dots, n_i)$, *not all* 0, *such that*

$$\sum_{j=1}^{n_1} x_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s} x_{sj}\mathbf{b}_{sj}, \tag{3.4}$$

$$\max_{v \in M_K^\infty} |x_{ij}|_v \leqslant \Big(\frac{2}{\pi}\Big)^{2r_2/d} |\Delta_K|^{1/d} \cdot \Big\{ (ns)^{n/2} \prod_{v \in M_K} |\det D_v|_v^{-d_v/d} \Big\}^{(s-1)/m} \tag{3.5}$$

for $i = 1, \dots, s,\ j = 1, \dots, n_i$.

Moreover, there are a finite extension L *of* K *and numbers* $x_{ij} \in O_L$ $(i = 1, \dots, s,$ $j = 1, \dots, n_i)$, *not all* 0, *satisfying (3.4) and*

$$\max_{w \in M_L^\infty} |x_{ij}|_w \leqslant m^{1/2} \cdot \Big\{ (ns)^{n/2} \prod_{v \in M_K} |\det D_v|_v^{-d_v/d} \Big\}^{(s-1)/m} \tag{3.6}$$

for $i = 1, \dots, s,\ j = 1, \dots, n_i$.

Remark. (3.1) is a technical condition needed in the proof. In all applications we know of, this condition can be satisfied.

4. Preparations

4.1. Let K be a number field and $v \in M_K$. Let B be a $(n-m) \times n$-matrix with entries in K_v where $0 < m < n$ and let $\mathbf{b}_1, \dots, \mathbf{b}_{n-m}$ denote the rows of B. Put

$$H_v(B) := \|\mathbf{b}_1 \wedge \cdots \wedge \mathbf{b}_{n-m}\|_v,$$

where the exterior product is defined similarly as for column vectors. Then by (2.1) we have

$$H_v(CB) = |\det C|_v \cdot H_v(B) \quad \text{for } C \in GL_{n-m}(K_v). \tag{4.1}$$

Further, by applying Hadamard's inequality if $v \in M_K^\infty$ and the ultrametric inequality if $v \in M_K^0$ we obtain

$$H_v(B) \leqslant \|\mathbf{b}_1\|_v \cdots \|\mathbf{b}_{n-m}\|_v\,. \tag{4.2}$$

If B has its entries in K then we define the height of B by

$$H(B) := \prod_{v \in M_K} H_v(B)^{d_v/d},$$

where as before, $d_v = [K_v : \mathbf{Q}_p]$ and $d = [K : \mathbf{Q}]$. Thus $H(B) \geqslant 1$ if $\operatorname{rank} B = n - m$.

We recall some versions of Siegel's Lemma. Let again m, n be integers with $n > m > 0$ and let B be an $(n-m) \times n$-matrix with entries in K, satisfying

$$\operatorname{rank} B = n - m. \tag{4.3}$$

Consider the system of linear equations

$$B\mathbf{x} = \mathbf{0} \tag{4.4}$$

to be solved in either $\mathbf{x} \in K^n$ or $\mathbf{x} \in L^n$ where L is a finite extension of K.

Lemma 4.2. *Equation (4.4) has a non-zero solution* $\mathbf{x} = (x_1, \ldots, x_n)^t \in O_K^n$ *with*

$$|x_i|_v \leqslant \left(\frac{2}{\pi}\right)^{2r_2/d} |\Delta_K|^{1/d} \cdot H(B)^{1/m} \quad \text{for } i = 1, \ldots, n,\ v \in M_K^\infty. \tag{4.5}$$

Proof. For $\mathbf{x} = (x_1, \ldots, x_n)^t \in K^n$ we put

$$\|\mathbf{x}\|_{v,\infty} := \max(|x_1|_v, \ldots, |x_n|_v) \quad \text{for } v \in M_K^\infty,$$
$$H_\infty(\mathbf{x}) := \prod_{v \in M_K^\infty} \|\mathbf{x}\|_{v,\infty}^{d_v/d} \cdot \prod_{v \in M_K^0} \|\mathbf{x}\|_v^{d_v/d}.$$

By the version of Siegel's Lemma due to Bombieri and Vaaler [1, Theorem 9], there is a non-zero solution $\mathbf{y} \in K^n$ of (4.4) with

$$H_\infty(\mathbf{y}) \leqslant \left(\frac{2}{\pi}\right)^{r_2/d} |\Delta_K|^{1/2d} \cdot H(B)^{1/m}. \tag{4.6}$$

By [1, Theorem 3] with $L = 1$ (the one-dimensional version of the adèlic Minkowski's theorem) there is a non-zero $\lambda \in K$ with

$$|\lambda|_v \leqslant \left(\frac{2}{\pi}\right)^{r_2/d} |\Delta_K|^{1/2d} \cdot H_\infty(\mathbf{y}) \cdot \|\mathbf{y}\|_{v,\infty}^{-1} \quad \text{for } v \in M_K^\infty,$$
$$|\lambda|_v \leqslant \|\mathbf{y}\|_v^{-1} \quad \text{for } v \in M_K^0.$$

(Let $K_{\mathbf{A}}$ denote the ring of adèles of K and let $\mathcal{S}$ be the set of $\lambda \in K_{\mathbf{A}}$ satisfying these inequalities. It can be checked that $\mathcal{S}$ has Haar measure $V(\mathcal{S}) = 2^d$, and this guarantees the existence of a non-zero $\lambda \in \mathcal{S} \cap K$.)

Write $\mathbf{x} = (x_1, \ldots, x_n)^t = \lambda \mathbf{y}$. Then $\mathbf{x}$ is a non-zero solution of (4.4). We have $\|\mathbf{x}\|_v \leqslant 1$ for $v \in M_K^0$, hence $\mathbf{x} \in O_K^n$. Further, $\max_i |x_i|_v = \|\mathbf{x}\|_{v,\infty} \leqslant (2/\pi)^{r_2/d} |\Delta_K|^{1/2d} H_\infty(\mathbf{y})$ for $v \in M_K^\infty$, which together with (4.6) implies (4.5). □

Lemma 4.3. *There is a finite extension L of K such that (4.4) has a non-zero solution $\mathbf{x} = (x_1, \dots, x_n)^t \in O_L^n$ with*

$$|x_i|_w \leqslant m^{1/2} \cdot H(B)^{1/m} \quad \textit{for } i = 1, \dots, n,\ w \in M_L^\infty. \tag{4.7}$$

Proof. For $\mathbf{x} \in K^n$, put $h(\mathbf{x}) := \log H(\mathbf{x})$. As is well-known, this height is absolute, i.e. independent of K, and invariant under scalar multiplication so that it gives rise to a height on $\mathbf{P}^{n-1}(\overline{\mathbf{Q}})$. Let $X \subset \mathbf{P}^{n-1}$ be the linear projective space given by (4.4). Denote by $h_F(X)$ the absolute Faltings height of X (cf. [8, p. 435, Definition 5.1]). A very special case of Zhang [15, Theorem 5.2] gives that for every $\varepsilon > 0$ there is a point $\mathbf{y} \in X(\overline{\mathbf{Q}})$ with

$$h(\mathbf{y}) \leqslant \frac{1+\varepsilon}{m} \cdot h_F(X). \tag{4.8}$$

For instance by [8, p. 437, Prop. 5.5] we have

$$h_F(X) = \log H(X) + \sigma_m \ \text{ with } \sigma_m := \frac{1}{2} \sum_{j=1}^{m-1} \sum_{k=1}^{j} \frac{1}{k}$$

where we have used X also to denote the linear subspace of K^n defined by (4.4). Lastly, by [1, p. 28] we have $H(X) = H(B)$. By combining these facts with (4.8) we obtain that for every $\varepsilon > 0$ there is a non-zero solution $\mathbf{y} \in \overline{\mathbf{Q}}^n$ of (4.4) such that

$$H(\mathbf{y}) \leqslant \Big\{ \exp(\sigma_m) \cdot H(B) \Big\}^{(1+\varepsilon)/m}. \tag{4.9}$$

We mention that Roy and Thunder [10, Theorem 1] proved a similar result with $m(m-1)/4$ instead of σ_m.

By e.g., [4, Lemma 6.3] there are a finite extension L of K and a non-zero $\lambda \in L$ such that $\mathbf{y} \in L^n$ and such that

$$|\lambda|_w \leqslant \Big(\frac{H(\mathbf{y})}{\|\mathbf{y}\|_w} \Big)^{1+\varepsilon} \text{ for } w \in M_L^\infty, \quad |\lambda|_w \leqslant \|\mathbf{y}\|_w^{-1} \text{ for } w \in M_L^0.$$

Let $\mathbf{x} = (x_1, \dots, x_n)^t = \lambda \mathbf{y}$. Then $\mathbf{x}$ is a non-zero solution of (4.4). Further, $\|\mathbf{x}\|_w \leqslant 1$ for $w \in M_L^0$ which implies $\mathbf{x} \in O_L^n$. Lastly, in view of (4.9) we have $\max_i |x_i|_w \leqslant \|\mathbf{x}\|_w \leqslant \Big\{ \exp(\sigma_m) \cdot H(B) \Big\}^{(1+\varepsilon)^2/m}$ for $w \in M_L^\infty$. Using that $\sigma_m < \frac{1}{2} m \log m$ and letting $\varepsilon \downarrow 0$ we obtain that there are a finite extension L of K and a non-zero solution $\mathbf{x} \in O_L^n$ of (4.4) satisfying (4.7). □

5. Proof of Theorem 3.2

5.1. We keep the notation and assumptions from Theorem 3.2. From elementary linear algebra we know that $n - \dim(V_1 \cap \cdots \cap V_s) \geqslant \sum_{i=1}^{s}(n - \dim V_i)$. We want to reduce this to the case that

$$n - \dim(V_1 \cap \cdots \cap V_s) = \sum_{i=1}^{s}(n - \dim V_i). \tag{5.1}$$

This is provided by the following lemma.

Lemma 5.2. *There are integers $n_1' \geqslant n_1, \ldots, n_s' \geqslant n_s$ and vectors $\mathbf{b}_{ij} \in K^n$ for $i = 1, \ldots, s$, $j = n_i + 1, \ldots, n_i'$ such that the following conditions are satisfied:*
(i) for $i = 1, \ldots, s$ the vectors $\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i'}$ are linearly independent and if V_i' is the vector space generated by these vectors then $V_1' \cap \cdots \cap V_s' = V_1 \cap \cdots \cap V_s$;
(ii) $n - \dim(V_1' \cap \cdots \cap V_s') = \sum_{i=1}^{s}(n - \dim V_i')$;
(iii) $\|D_v \mathbf{b}_{ij}\|_v \leqslant 1$ for $i = 1, \ldots, s$, $j = 1, \ldots, n_i'$, $v \in M_K$;
(iv) If for some extension L of K we have $\sum_{j=1}^{n_1'} x_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s'} x_{sj}\mathbf{b}_{sj}$ with $x_{ij} \in L$, then $x_{ij} = 0$ for $i = 1, \ldots, s$, $j = n_i + 1, \ldots, n_i'$.

Proof. We choose $n_1' = n_1$ so that $V_1' = V_1$. Let $i \in \{2, \ldots, s\}$. Put $t_i := \dim((V_1 \cap \cdots \cap V_{i-1}) + V_i)$ and $n_i' = n_i + n - t_i$. We start with the basis $\{\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i}\}$ of V_i given by (3.3). We extend this to a basis $\{\mathbf{c}_1, \ldots, \mathbf{c}_{t_i - n_i}\} \cup \{\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i}\}$ of $(V_1 \cap \cdots \cap V_{i-1}) + V_i$. We extend this further to a basis $\{\mathbf{c}_1, \ldots, \mathbf{c}_{t_i - n_i}\} \cup \{\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i}\} \cup \{\mathbf{b}_{i,n_i+1}, \ldots, \mathbf{b}_{i,n_i'}\}$ of K^n where $\mathbf{b}_{ij}$ $(j = n_i + 1, \ldots, n_i')$ are chosen from the basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of K^n satisfying (3.1). Thus, $\{\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i'}\}$ is linearly independent and (iii) is satisfied. Let V_i' be the vector space generated by $\mathbf{b}_{i1}, \ldots, \mathbf{b}_{i,n_i'}$.

In order to prove (i) and (ii), we prove by induction on i that $V_1 \cap \cdots \cap V_i = V_1' \cap \cdots \cap V_i'$ and $n - \dim(V_1' \cap \cdots \cap V_i') = \sum_{j=1}^{i}(n - \dim V_j')$ for $i = 1, \ldots, s$. For $i = 1$ this is clear. Assume this has been proved for $i - 1$ in place of i, where $i \geqslant 2$. Thus $V_1' \cap \cdots \cap V_i' = (V_1 \cap \cdots \cap V_{i-1}) \cap V_i'$. Suppose $\mathbf{x} \in V_1' \cap \cdots \cap V_i'$. Then on the one hand, $\mathbf{x} \in V_1 \cap \cdots \cap V_{i-1}$, on the other hand $\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{y} \in V_i$ and $\mathbf{z}$ is a linear combination of the vectors $\mathbf{b}_{i,n_i+1}, \ldots, \mathbf{b}_{i,n_i'}$. But then $\mathbf{z} = \mathbf{x} - \mathbf{y}$ is also a linear combination of the vectors $\mathbf{c}_1, \dots, \mathbf{c}_{t_i-n_i}$, $\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i}$. Hence $\mathbf{z} = \mathbf{0}$, and therefore, $\mathbf{x} \in V_1 \cap \cdots \cap V_i$. It follows that $V_1' \cap \cdots \cap V_i' = V_1 \cap \cdots \cap V_i$. Further, noting that $\dim((V_1' \cap \cdots \cap V_{i-1}') + V_i') = \dim((V_1 \cap \cdots \cap V_{i-1}) + V_i') = n$, we obtain

$$\begin{aligned} n - \dim(V_1' \cap \cdots \cap V_i') &= n - \dim(V_1' \cap \cdots \cap V_{i-1}') - \dim V_i' + n \\ &= \sum_{j=1}^{i-1}(n - \dim V_j') + n - \dim V_i' = \sum_{j=1}^{i}(n - \dim V_j'). \end{aligned}$$

This completes the induction step, hence completes the proof of (i) and (ii).

Let L be an extension of K. For a linear subspace V of K^n, put $V^L := V \otimes_K L$. Let $\mathbf{x} = \sum_{j=1}^{n_1'} x_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s'} x_{sj}\mathbf{b}_{sj}$ with $x_{ij} \in L$. Then $\mathbf{x} \in V_1'^L \cap \cdots \cap V_s'^L$. By (i) we have $V_1'^L \cap \cdots \cap V_s'^L = V_1^L \cap \cdots \cap V_s^L$. Hence there are $y_{ij} \in L$ such that $\mathbf{x} = \sum_{j=1}^{n_1} y_{1j}\mathbf{b}_{1j} = \cdots = \sum_{j=1}^{n_s} y_{sj}\mathbf{b}_{sj}$. Since by (i) each set $\{\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i'}\}$ is linearly independent over L, this implies $x_{ij} = y_{ij}$ for $j = 1, \dots, n_i$ and $x_{ij} = 0$ for $j = n_i + 1, \dots, n_i'$. This proves (iv). □

5.3. Proof of Theorem 3.2.

According to Lemma 5.2, in order to prove Theorem 3.2 it suffices to prove this result for the sets $\{\mathbf{b}_{ij} : j = 1, \dots, n_i'\}$ in place of $\{\mathbf{b}_{ij} : j = 1, \dots, n_i\}$. Therefore, there is no loss of generality to assume (5.1) and we shall do so in the sequel.

Let B_i be the $n \times n_i$-matrix with columns $\mathbf{b}_{i1}, \dots, \mathbf{b}_{i,n_i}$, respectively and let $\mathbf{x}_i = (x_{i1}, \dots, x_{i,n_i})^t$ for $i = 1, \dots, s$. Then we may rewrite (3.4) as $B_1\mathbf{x}_1 = \cdots = B_s\mathbf{x}_s$ or as

$$\begin{pmatrix} B_1 & -B_2 & 0 & \cdots & 0 \\ B_1 & 0 & -B_3 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ B_1 & 0 & 0 & \cdots & -B_s \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_s \end{pmatrix} = \mathbf{0}. \tag{5.2}$$

We denote the matrix by B and the vector by $\mathbf{x}$, so that we have to solve $B\mathbf{x} = \mathbf{0}$. Note that B is an $n(s-1) \times (n_1 + \cdots + n_s)$-matrix. Since the solution space of (5.2) has dimension $\dim(V_1 \cap \cdots \cap V_s) = m$, the rank of B is $n_1 + \cdots + n_s - m$. Our assumption (5.1) says that $n - m = \sum_{j=1}^{s}(n - n_j)$, which implies $n_1 + \cdots + n_s - m = n(s-1)$. Therefore, B satisfies (4.3) with $n_1 + \cdots + n_s$ in place of n. Hence Lemma 4.2 and Lemma 4.3 are applicable. Recall that if we write $\mathbf{x} =$

$(x_{11},\dots,x_{1,n_1},\dots,x_{s1},\dots,x_{s,n_s})^t$, then $\mathbf{x}$ is a solution of (5.2) if and only if the numbers x_{ij} satisfy (3.4). Thus, by applying Lemma 4.2 to (5.2) we obtain that there are numbers $x_{ij}\in O_K$, not all 0 satisfying (3.4) and

$$(5.3)\qquad |x_{ij}|_v \leqslant \Big(\frac{2}{\pi}\Big)^{2r_2/d}|\Delta_K|^{1/d}\cdot H(B)^{1/m} \quad \text{for } i=1,\dots,s,\ j=1,\dots,n_i,\ v\in M_K^\infty.$$

Moreover, by applying Lemma 4.3 to (5.2) we obtain that there are a finite extension L of K, and numbers $x_{ij}\in O_L$, not all 0, satisfying (3.4) and

$$(5.4)\qquad |x_{ij}|_w \leqslant m^{1/2}\cdot H(B)^{1/m} \quad \text{for } i=1,\dots,s,\ j=1,\dots,n_i,\ w\in M_L^\infty.$$

It remains to estimate from above the height $H(B)$. Let $v\in M_K$. We express the matrix B in (5.2) as a product

$$\begin{pmatrix} D_v^{-1} & & & 0\\ & D_v^{-1} & & \\ & & \ddots & \\ 0 & & & D_v^{-1}\end{pmatrix}\cdot\begin{pmatrix} D_vB_1 & -D_vB_2 & 0 & \cdots & 0\\ D_vB_1 & 0 & -D_vB_3 & \cdots & 0\\ \vdots & \vdots & & \ddots & \vdots\\ D_vB_1 & 0 & 0 & \cdots & -D_vB_s\end{pmatrix},$$

where the left matrix has $s-1$ blocks D_v^{-1} on the diagonal and is zero at the other places. We denote the left matrix by E_v and the right matrix by F_v. Then $\det E_v=(\det D_v)^{1-s}$. By (3.3), the entries of F_v all have v-adic absolute value $\leqslant 1$. So by (4.2), $H_v(F_v)\leqslant (n_1+\cdots+n_s)^{n(s-1)/2}\leqslant (ns)^{n(s-1)/2}$ if $v\in M_K^\infty$ and $H_v(F_v)\leqslant 1$ if $v\in M_K^0$. Now (4.1) implies $H_v(B)=|\det E_v|_v\cdot H_v(F_v)\leqslant (ns)^{n(s-1)/2}|\det D_v|_v^{1-s}$ if $v\in M_K^\infty$, $H_v(B)\leqslant |\det D_v|_v^{1-s}$ if $v\in M_K^0$. On raising these inequalities to the power d_v/d and taking the product over $v\in M_K$ we obtain

$$H(B)\leqslant (ns)^{n(s-1)/2}\Big(\prod_{v\in M_K}|\det D_v|_v^{d_v/d}\Big)^{1-s}.$$

By inserting this into (5.3), (5.4), respectively we obtain (3.5) and (3.6). This proves Theorem 3.2. □

6. Proof of Theorem 2.2

6.1. We recall some facts about orthonormal sets of vectors. Let $v \in M_K$. We call a set of vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ in K_v^n *orthonormal* if for every $\mathbf{y} = (y_1, \dots, y_k)^t \in K_v^k$ we have

$$\|\sum_{i=1}^{k} y_i \mathbf{e}_i\|_v = \|\mathbf{y}\|_v = \begin{cases} \Big(\sum_{i=1}^{k} |y_i|_v^2\Big)^{1/2} & \text{if } v \in M_K^\infty, \\ \max(|y_1|_v, \dots, |y_k|_v) & \text{if } v \in M_K^0. \end{cases} \tag{6.1}$$

For $v \in M_K^\infty$ this coincides with the usual notion of orthonormality of a set of vectors in $\mathbf{R}^n$ or $\mathbf{C}^n$, while for $v \in M_K^0$ this is inspired by Weil [14, p. 26]. Obviously, orthonormal sets of vectors are linearly independent. An orthonormal basis of a subspace of K_v^n is a basis which is an orthonormal set of vectors.

Most of the material in this section can be deduced from the theory of orthogonal projections in K_v^n developed by Vaaler [13] and Burger and Vaaler [3]. Instead of using their results, we have given direct proofs since this turned out to be more convenient.

Lemma 6.2. *Let $\mathbf{a}_1, \dots, \mathbf{a}_k$ be linearly independent vectors in K_v^n. Then there is an orthonormal set of vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ in K_v^n such that*

$$\mathbf{a}_i = \sum_{j=1}^{i} \gamma_{ij} \mathbf{e}_j \quad \textit{for } i = 1, \dots, k,$$

with $\gamma_{ij} \in K_v$ for $i = 1, \dots, k$, $j = 1, \dots, i$ and $\gamma_{ii} \neq 0$ for $i = 1, \dots, k$.

Proof. For $v \in M_K^\infty$ this is simply the Gram-Schmidt orthogonalization procedure, while for $v \in M_K^0$ this is a consequence of [14, p. 26, Prop. 3]. □

Lemma 6.3. *Let $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ be an orthonormal set of vectors in K_v^n. Then*

$$\|\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_k\|_v = 1 . \tag{6.2}$$

Proof. For $v \in M_K^\infty$ this follows from a well-known fact for orthonormal sets of vectors in $\mathbf{R}^n$ or $\mathbf{C}^n$. Assume $v \in M_K^0$. Let $O_v = \{x \in K_v : |x|_v \leqslant 1\}$, $M_v = \{x \in K_v : |x|_v < 1\}$, $k_v = O_v/M_v$ denote the ring of v-adic integers, the maximal ideal of O_v and the residue field of v, respectively. (6.1) implies that $\mathbf{e}_i \in O_v^n$ for $i = 1, \dots, n$. Denote by $\mathbf{e}_i^*$ the reduction of $\mathbf{e}_i$ modulo M_v. Assume that (6.2) is incorrect, i.e., $\|\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_k\|_v < 1$. Then $\mathbf{e}_1^* \wedge \cdots \wedge \mathbf{e}_k^* = \mathbf{0}$, which implies that $\mathbf{e}_1^*, \dots, \mathbf{e}_k^*$ are linearly dependent in k_v^n. Hence there are $y_i^* \in k_v$, not all 0, such that $\sum_{i=1}^k y_i^* \mathbf{e}_i^* = \mathbf{0}$. By lifting this to O_v, we see that there are $y_i \in O_v$ with $\max(|y_1|_v, \dots, |y_k|_v) = 1$ such that $\| \sum_{i=1}^k y_i \mathbf{e}_i\|_v < 1$. But this contradicts (6.1). □

6.4. Proof of Theorem 2.2.

We keep the notation and assumptions from Theorem 2.2. We assume that for $v \in M_K^0$, C_v belongs to the value group $G_v = \{|x|_v : x \in K_v^*\}$. This is no loss of generality. For suppose that for some $v \in M_K^0$, $C_v \notin G_v$ and let C_v' be the largest number in G_v which is smaller than C_v. Then if we replace C_v by C_v', condition (2.4) is unaltered while the right-hand sides of (2.6), (2.7) decrease.

Let $r := \dim W$. Then $\dim U = r + n$. Choose a basis $\{\mathbf{a}_1, \dots, \mathbf{a}_{r+n}\}$ of U such that $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ is a basis of W. Let $v \in M_K$. Put $W_v := W \otimes_K K_v$, $U_v := U \otimes_K K_v$. According to Lemma 6.2, U_v has an orthonormal basis $\{\mathbf{e}_1, \dots, \mathbf{e}_{r+n}\}$ such that

$$\mathbf{a}_i = \sum_{j=1}^{i} \gamma_{ij} \mathbf{e}_j \quad \text{for } i = 1, \dots, r+n, \tag{6.3}$$

with $\gamma_{ij} \in K_v$ for $i = 1, \dots, r+n$, $j = 1, \dots, i$ and $\gamma_{ii} \neq 0$ for $i = 1, \dots, r+n$. Since $\mathbf{a}_1, \dots, \mathbf{a}_r$ are linear combinations of $\mathbf{e}_1, \dots, \mathbf{e}_r$ and vice-versa, $\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ is an orthonormal basis of W_v.

Let $\mathbf{x} \in V_1 + \cdots + V_s$. Choose any $\mathbf{x}^* \in U$ mapping to $\mathbf{x}$ under the canonical map from K^h to K^h/W. Write $\mathbf{x}^* = \sum_{i=1}^{r+n} x_i \mathbf{a}_i$ with $x_i \in K$. Then the vector

$$\varphi(\mathbf{x}) := (x_{r+1}, \dots, x_{r+n})^t \in K^n$$

is independent of the choice of $\mathbf{x}^*$. Notice that φ is a linear isomorphism from $V_1 + \cdots + V_s$ to K^n. We may express $\mathbf{x}^*$ otherwise as $\mathbf{x}^* = \sum_{i=1}^{r+n} y_i \mathbf{e}_i$ with $y_i \in K_v$. Then

$$\psi_v(\mathbf{x}) := (y_{r+1}, \dots, y_{r+n})^t \in K_v^n$$

is also independent of the choice of $\mathbf{x}^*$. Clearly, $\sum_{i=r+1}^{r+n} y_i \mathbf{e}_i$ maps to $\mathbf{x}$ under the canonical map from K_v^h to K_v^h/W_v. Further, from (6.1) it is clear that $\|\mathbf{x}^*\|_v \geqslant$

$\| \sum_{i=r+1}^{r+n} y_i \mathbf{e}_i \|_v = \|\psi_v(\mathbf{x})\|_v$. Therefore,

$$\|\mathbf{x}\|_v^W = \|\psi_v(\mathbf{x})\|_v. \tag{6.4}$$

Moreover, from (6.3) it follows that

$$\psi_v(\mathbf{x}) = E_v\varphi(\mathbf{x}) \quad \text{with } E_v = \begin{pmatrix} \gamma_{r+1,r+1} & \cdots & \cdots & \gamma_{r+n,r+1} \\ & \gamma_{r+2,r+2} & \cdots & \vdots \\ & & \ddots & \vdots \\ 0 & & & \gamma_{r+n,r+n} \end{pmatrix}, \tag{6.5}$$

where the elements of E_v below the diagonal are zero. By our assumption on C_v, there is an $\alpha_v \in K_v^*$ with $|\alpha_v|_v = C_v$. Now define the matrix $D_v := \alpha_v^{-1}E_v$. Then from (6.4) and (6.5) it follows that for $\mathbf{x} \in V_1 + \cdots + V_s$,

$$\|\mathbf{x}\|_v^W \leqslant C_v \Longleftrightarrow \|D_v\varphi(\mathbf{x})\|_v \leqslant 1. \tag{6.6}$$

From (6.3), (2.1), Lemma 6.3 we obtain,

$$\begin{aligned} \|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{r+n}\|_v &= |\gamma_{11} \cdots \gamma_{r+n,r+n}|_v \cdot \|\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{r+n}\|_v = |\gamma_{11} \cdots \gamma_{r+n,r+n}|_v, \\ \|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r\|_v &= |\gamma_{11} \cdots \gamma_{rr}|_v \cdot \|\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_r\|_v = |\gamma_{11} \cdots \gamma_{rr}|_v\,. \end{aligned}$$

Together with (6.5) this implies

$$|\det D_v|_v = |\alpha_v^{-n}\gamma_{r+1,r+1} \cdots \gamma_{r+n,r+n}|_v = C_v^{-n}\frac{\|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{r+n}\|_v}{\|\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r\|_v}. \tag{6.7}$$

We have a matrix D_v for every $v \in M_K$. The quantities in the right-hand side of (6.7) are equal to 1 for all but finitely many v. Therefore, $|\det D_v|_v = 1$ for all but finitely many v. That is, $D := \{D_v : v \in M_K\}$ is an M_K-matrix of order n. By (6.7) we have

$$\begin{aligned} \prod_{v\in M_K} |\det D_v|_v^{d_v/d} &= \Big(\prod_{v\in M_K} C_v^{d_v/d}\Big)^{-n}\frac{H(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_{r+n})}{H(\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r)} \\ &= \Big(\prod_{v\in M_K} C_v^{d_v/d}\Big)^{-n} \cdot H(U) \cdot H(W)^{-1}. \end{aligned} \tag{6.8}$$

From the bases of $V_1, \ldots, V_s$ with (2.4) we select a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of $V_1 + \cdots + V_s$. Now we apply Theorem 3.2 with the M_K-matrix D constructed above, with the vectors $\varphi(\mathbf{b}_i)$, $\varphi(\mathbf{b}_{ij})$ in place of $\mathbf{b}_i$, $\mathbf{b}_{ij}$ and with the spaces $\varphi(V_i)$ in place of V_i. Then the assumptions (2.2)-(2.4) of Theorem 2.2 in conjunction with (6.6) and the fact that φ is a linear isomorphism from $V_1 + \cdots + V_s$ to K^n, imply that the conditions (3.1)-(3.3) of Theorem 3.2 are satisfied. It follows that there are $x_{ij} \in O_K$, not all 0, satisfying (3.4) (with $\varphi(\mathbf{b}_{ij})$ instead of $\mathbf{b}_{ij}$) and (3.5). Since φ is an isomorphism, these x_{ij} satisfy (2.5), and by substituting (6.8) into (3.5) it follows that they also satisfy (2.6). Furthermore, there are a finite extension L of K and numbers $x_{ij} \in O_L$, not all 0, satisfying (3.4) (with again $\varphi(\mathbf{b}_{ij})$ instead of $\mathbf{b}_{ij}$) and (3.6), and similarly as above it follows that these numbers satisfy (2.5) and (2.7). This completes the proof of Theorem 2.2. $\square$

References

[1] E. Bombieri, J.D. Vaaler, On Siegel's Lemma, *Invent. math.*, **73** (1983) 11–32.

[2] J.-B. Bost, H. Gillet, C. Soulé, Height of Projective Varieties and Positive Green Forms, *J. Amer. Math. Soc*, **7** (1994) 903–1027.

[3] E.B. Burger, J.D. Vaaler, On the decomposition of vectors over number fields, *J. reine angew. Math.*, **435** (1993) 197–219.

[4] J.-H. Evertse, H. P. Schlickewei, A Quantitative Version of the Absolute Subspace Theorem, *J. reine u. angew. Math*, to appear.

[5] G. Faltings, Diophantine Approximations on Abelian Varieties, *Ann. of Math.*, **133** (1991) 549–576.

[6] G. Faltings, G. Wüstholz, Diophantine Approximations on Projective Spaces, *Invent. Math.*, **116** (1994) 109–138.

[7] R. G. Ferretti, Quantitative Diophantine approximations on projective varieties, in preparation.

[8] W. Gubler, Höhentheorie, *Math. Ann.*, **298** (1994) 427–455.

[9] D. Roy, J.L. Thunder, An absolute Siegel's Lemma, *J. reine angew. Math.*, **476** (1996) 1–26.

[10] D. Roy, J.L. Thunder, Addendum and Erratum to "An absolute Siegel's Lemma", *J. reine angew. Math.*, **508** (1999) 47–51.

[11] C.L. Siegel, Über einige Anwendungen Diophantischer Approximationen, *Abh. der Preuß. Akad. der Wissenschaften Phys.-math. Kl.*, **1** (1929) 209–266 (=*Ges. Abh. I*).

[12] A. Thue, Über Annäherungswerte algebraischer Zahlen, *J. reine angew. Math.*, **135** (1909) 284–305.

[13] J.D. Vaaler, Small zeros of quadratic forms over number fields, *Trans. AMS*, **302** (1987) 281–296.

[14] A. Weil, *Basic Number Theory*, Grundl. math. Wiss. **144**, Springer Verlag, Berlin 1973.

[15] S. Zhang, Positive line bundles on arithmetic varieties, *J. Amer. Math. Soc. (1)*, **8** (1995) 187–221.

Universiteit Leiden, Mathematisch Instituut, Postbus 9512, 2300 RA Leiden, The Netherlands

E-mail address: evertse@math.leidenuniv.nl