1-1-2006

Siegel's Lemma with Additional Conditions

Lenny Fukshansky
Claremont McKenna College

SIEGEL'S LEMMA WITH ADDITIONAL CONDITIONS

LENNY FUKSHANSKY

ABSTRACT. Let K be a number field, and let W be a subspace of K^N, $N \geq 1$. Let $V_1, ..., V_M$ be subspaces of K^N of dimension less than dimension of W. We prove the existence of a point of small height in $W \setminus \bigcup_{i=1}^{M} V_i$, providing an explicit upper bound on the height of such a point in terms of heights of W and $V_1, ..., V_M$. Our main tool is a counting estimate we prove for the number of points of a subspace of K^N inside of an adelic cube. As corollaries to our main result we derive an explicit bound on the height of a non-vanishing point for a decomposable form and an effective subspace extension lemma.

1. Introduction and notation

The name Siegel's Lemma is usually used to denote results about small-height solutions of a system of linear equations. Such a result in a simple form was first proved by Thue in 1909 ([10], pp. 288-289) using the Dirichlet's box principle. Siegel ([9], Bd. I, p. 213, Hilfssatz) was the first to formally state this principle in the classical case.

Notice that a small-height solution to a system of linear equations is a point of small height in the nullspace of the matrix of this linear system. Thus this principle can be viewed as a statement about points of small height in a given vector space. We write H and $\mathcal{H}$ for appropriately selected height functions, which we will precisely define below. The following modern formulation of this result follows from a celebrated theorem of Bombieri and Vaaler, [2].

Theorem 1.1 ([2])**.** *Let K be a number field of degree d and discriminant $\mathcal{D}_K$, and let $N \geq 1$ be an integer. Let W be a non-zero subspace of K^N of dimension $w \leq N$. There exists a non-zero point $\boldsymbol{x} \in W$ such that*

$$H(\boldsymbol{x}) \leq \left\{ N |\mathcal{D}_K|^{1/d} \right\}^{1/2} \mathcal{H}(W)^{1/w}. \tag{1}$$

The exponent on $\mathcal{H}(W)$ in the upper bound of Theorem 1.1 is best possible, however the constant is not. The best possible constant for Siegel's Lemma was recently obtained by Vaaler in [13]. The actual Bombieri - Vaaler theorem is more general: it produces a full basis of small height for W. Results of this sort were originally treated as important technical lemmas used in transcendental number theory and Diophantine approximations for the purpose of constructing a certain auxiliary polynomial (see [2] and [1] for more information). Nowadays they have evolved as important results in their own right.

In this paper we consider a generalization of this problem. Let K be a number field, and let W be a subspace of K^N, $N \geq 2$. Let $V_1, ..., V_M$ be subspaces of K^N of

1991 *Mathematics Subject Classification.* Primary 11D04, 11H06; Secondary 11H46.
Key words and phrases. lattices, linear forms, diophantine approximation, height.

dimension less than dimension of W. We want to prove the existence of a non-zero point of small height in $W \setminus \bigcup_{i=1}^{M} V_i$ providing an explicit upper bound on the height of such a point. More precisely, our main result reads as follows.

Theorem 1.2. *Let K be a number field of degree d with discriminant $\mathcal{D}_K$. Let $N \geq 2$ be an integer, $l = \left[\frac{N}{2}\right]$, and let W be a subspace of K^N of dimension w, $1 \leq w \leq N$. Let $1 \leq s < w$ be an integer, and let $V_1, ..., V_M$ be nonzero subspaces of K^N with $\max_{1 \leq i \leq M}\{\dim_K(V_i)\} \leq s$. There exists a point $\boldsymbol{x} \in W \setminus \bigcup_{i=1}^{M} V_i$ such that*

$$H(\boldsymbol{x}) \leq \mathcal{C}_{K,N}(w,s)\mathcal{H}(W)^d \left\{ \left(\sum_{i=1}^{M} \frac{1}{\mathcal{H}(V_i)^d} \right)^{\frac{1}{(w-s)d}} + M^{\frac{1}{(w-s)d+1}} \right\}, \tag{2}$$

where

$$\mathcal{C}_{K,N}(w,s) = 2^{w(d+3)} |\mathcal{D}_K|^{\frac{w}{2}} \left((wd)^w \binom{Nd}{ld}^{\frac{1}{2d}} \right)^{\frac{1}{w-s}}. \tag{3}$$

The dependence on $\mathcal{H}(W)$ in the upper bound of Theorem 1.2 is sharp at least in the case $K = \mathbb{Q}$. Let $M = 1$, and take V_1 to be a subspace of W of dimension $w-1$ generated by the vectors corresponding to the first $w-1$ successive minima of W with respect to an adelic unit cube. Then the smallest vector in $W \setminus V_1$ will be the one corresponding to the w-th successive minimum, and its height can be as large as a constant multiple of $\mathcal{H}(W)$: this is a consequence of the adelic version of Minkowski's successive minima theorem and the Bombieri - Vaaler version of Siegel's lemma (see [2]).

We separately discuss a special case of our main result, which can be thought of as an inverse of Siegel's Lemma. Suppose that $W = K^N$, and let $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X})$ be M linear forms in N variables with coefficients in K. Then we can prove the existence of a point $\boldsymbol{x}$ in K^N of relatively small height such that $L_i(\boldsymbol{x}) \neq 0$ for every $i = 1, ..., M$ (i.e. $\boldsymbol{x}$ is outside of the union of nullspaces of linear forms). This discussion generalizes some results presented in the companion paper [5] in the case $K = \mathbb{Q}$ to any number field. In particular, Theorem 1.2 can be viewed as a generalization of Theorem 5.1 of [5]. Although we employ similar principles in the proof, the techniques and ideas of [5] are more elementary and combinatorial in nature.

This paper is structured as follows. In section 2 we present a technical lemma on the problem of counting integer lattice points in a closed cube in $\mathbb{R}^N$. In section 3 we use this counting mechanism to prove Theorem 1.2. In section 4 we discuss some interesting corollaries of this result.

We start with some notation. let K be a number field of degree d over $\mathbb{Q}$, O_K its ring of integers, $\mathcal{D}_K$ its discriminant, and $M(K)$ its set of places. For each place $v \in M(K)$ we write K_v for the completion of K at v and let $d_v = [K_v : \mathbb{Q}_v]$ be the local degree of K at v, so that for each $u \in M(\mathbb{Q})$

$$\sum_{v \in M(K), v|u} d_v = d. \tag{4}$$

For each place $v \in M(K)$ we define the absolute value $\| \ \|_v$ to be the unique absolute value on K_v that extends either the usual absolute value on $\mathbb{R}$ or $\mathbb{C}$ if $v|\infty$, or the

usual p-adic absolute value on $\mathbb{Q}_p$ if $v|p$, where p is a prime. We also define the second absolute value $|\ |_v$ for each place v by $|a|_v = \|a\|_v^{d_v/d}$ for all $a \in K$. Then for each non-zero $a \in K$ the *product formula* reads

$$\prod_{v \in M(K)} |a|_v = 1. \tag{5}$$

For each finite place $v \in M(K)$, $v \nmid \infty$, we define the *local ring of v-adic integers* $O_v = \{x \in K : |x|_v \leq 1\}$, whose unique maximal ideal is $P_v = \{x \in K : |x|_v < 1\}$. Then $O_K = \bigcap_{v \nmid \infty} O_v$.

We extend absolute values to vectors by defining the local heights. For each $v \in M(K)$ define a local height H_v on K_v^N by

$$H_v(\boldsymbol{x}) = \max_{1 \leq i \leq N} |x_i|_v, \tag{6}$$

for each $\boldsymbol{x} \in K_v^N$. Also, for each $v|\infty$ we define another local height

$$\mathcal{H}_v(\boldsymbol{x}) = \left(\sum_{i=1}^{N} \|x_i\|_v^2 \right)^{d_v/2d} \tag{7}$$

Then we can define two slightly different global height functions on K^N:

$$H(\boldsymbol{x}) = \prod_{v \in M(K)} H_v(\boldsymbol{x}), \quad \mathcal{H}(\boldsymbol{x}) = \prod_{v \nmid \infty} H_v(\boldsymbol{x}) \times \prod_{v | \infty} \mathcal{H}_v(\boldsymbol{x}), \tag{8}$$

for each $\boldsymbol{x} \in K^N$. It is easy to see that

$$H(\boldsymbol{x}) \leq \mathcal{H}(\boldsymbol{x}) \leq \sqrt{N} H(\boldsymbol{x}). \tag{9}$$

All our inequalities will use height H for vectors, however we use $\mathcal{H}$ to define the conventional Schmidt height on subspaces in the manner described below. This choice of heights coincides with [2].

We extend both heights H and $\mathcal{H}$ to polynomials by viewing them as height functions of the coefficient vector of a given polynomial. We also define a height function on subspaces of K^N. Let $V \subseteq K^N$ be a subspace of dimension J, $1 \leq J \leq N$. Choose a basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_J$ for V, and write $X = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_J)$ for the corresponding $N \times J$ basis matrix. Then

$$V = \{X\boldsymbol{t} : \boldsymbol{t} \in K^J\}.$$

On the other hand, there exists an $(N-J) \times N$ matrix A with entries in K such that

$$V = \{\boldsymbol{x} \in K^N : A\boldsymbol{x} = 0\}.$$

Let $\mathcal{I}$ be the collection of all subsets I of $\{1, ..., N\}$ of cardinality J. For each $I \in \mathcal{I}$ let I' be its complement, i.e. $I' = \{1, ..., N\} \setminus I$, and let $\mathcal{I}' = \{I' : I \in \mathcal{I}\}$. Then

$$|\mathcal{I}| = \binom{N}{J} = \binom{N}{N-J} = |\mathcal{I}'|.$$

For each $I \in \mathcal{I}$, write X_I for the $J \times J$ submatrix of X consisting of all those rows of X which are indexed by I, and ${}_{I'}A$ for the $(N-J) \times (N-J)$ submatrix of A consisting of all those columns of A which are indexed by I'. By the duality principle of Brill-Gordan [6] (also see Theorem 1 on p. 294 of [7]), there exists a non-zero constant $\gamma \in K$ such that

$$\det(X_I) = (-1)^{\varepsilon(I')} \gamma \det({}_{I'}A), \tag{10}$$

where $\varepsilon(I') = \sum_{i \in I'} i$. Define the vectors of *Grassmann coordinates* of X and A respectively to be

$$Gr(X) = (\det(X_I))_{I \in \mathcal{I}} \in K^{|I|}, \quad Gr(A) = (\det({}_{I'}A))_{I' \in \mathcal{I}'} \in K^{|I'|},$$

and so by (10) and (5)

$$\mathcal{H}(Gr(X)) = \mathcal{H}(Gr(A)).$$

Define the height of V denoted by $\mathcal{H}(V)$ to be this common value. This definition is legitimate, since it does not depend on the choice of the basis for V. In particular, notice that if

$$L(X_1, ..., X_N) = \sum_{i=1}^{N} q_i X_i \in K[X_1, ..., X_N]$$

is a linear form with a non-zero coefficient vector $\boldsymbol{q} \in K^N$, and $V = \{\boldsymbol{x} \in K^N : L(\boldsymbol{x}) = 0\}$ is an $(N-1)$-dimensional subspace of K^N, then

$$\mathcal{H}(V) = \mathcal{H}(L) = \mathcal{H}(\boldsymbol{q}). \tag{11}$$

The method of proof of Theorem 1.2 is the following. For a positive $R \geq 1$ we estimate cardinalities of sets

$$S_R(W) = \{\boldsymbol{x} \in W \cap O_K^N : \max_{v|\infty} H_v(\boldsymbol{x})^{d/d_v} \leq R\},$$

and $S_R(V_i) = S_R(W) \cap V_i$ for each $1 \leq i \leq M$. In other words, we count the number of points in sections of the adelic cube with "sidelength" R by W and by each V_i. Then we find R large enough so that $|S_R(W)|$ is greater than $\sum_{i=1}^{M} |S_R(V_i)|$. A related estimate for the number of points of bounded height in a subspace of K^N is provided by J. Thunder in [12]. Thunder's estimate, however, is asymptotic with an implicit constant in the error term. This is not suitable for our purposes, since we need explicit upper and lower bounds. Our estimates are different from Thunder's also in the way that we are considering points inside of an adelic cube, which is a smaller set than the one considered in [12]. We formulate our counting estimate precisely in Lemma 3.2 at the end of section 3. We are now ready to proceed. Results of this paper also appear as a part of [4].

2. Lattice points in cubes

In this section we state some bounds on the number of points of a lattice in $\mathbb{R}^N$ inside of a closed cube. These will later be used to prove our main result.

For the rest of this paper, let $R \geq 1$, and define

$$C_R^N = \{\boldsymbol{x} \in \mathbb{R}^N : \max_{1 \leq i \leq N} |x_i| \leq R\},$$

to be a cube in $\mathbb{R}^N$ centered at the origin with sidelength $2R$. Given a lattice Λ in $\mathbb{R}^N$ of rank N and determinant Δ, we want to estimate the quantity $|\Lambda \cap C_R^N|$. First suppose that $\mathrm{rk}(\Lambda) = N$. Then there exists an uppertriangular, nonsigular $N \times N$ matrix $A = (a_{mn})$ with positive real entries such that $\Lambda = \{A\boldsymbol{\xi} : \boldsymbol{\xi} \in \mathbb{Z}^N\}$. Then by Corollary 3.3 of [5], we have:

$$\prod_{m=1}^{N} \left[\frac{2R}{a_{mm}}\right] \leq |\Lambda \cap (C_R^N + \boldsymbol{z})| \leq \prod_{m=1}^{N} \left(\left[\frac{2R}{a_{mm}}\right] + 1\right), \tag{12}$$

for each point $\boldsymbol{z}$ in $\mathbb{R}^N$. Notice that if $2R \geq \max_{1\leq m\leq N} a_{mm}$, then the lower bound of (12) is greater or equal than $\prod_{m=1}^{N}\left(\frac{2R}{a_{mm}}-1\right)$.

If the matrix A as above with fixed determinant Δ is such that all diagonal entries $a_{mm} \geq c$ for some positive constant c, then the right hand side of (12) takes its maximum value and the left hand side takes its minimum value when $a_{mm} = c$ for $N-1$ distinct values of m. This leads to the following lemma.

Lemma 2.1. *Let Λ be a lattice of full rank in $\mathbb{R}^N$ of determinant Δ such that there exists a positive constant c and an uppertriangular basis matrix $A = (a_{mn})_{1\leq m,n\leq N}$ of Λ with diagonal entries $a_{mm} \geq c$ for all $1 \leq m \leq N$ (in particular, this is true with $c = 1$ if $\Lambda \subseteq \mathbb{Z}^N$). Assume that $2R \geq \max\left\{\frac{\Delta}{c^{N-1}}, c\right\}$. Then for each point $\boldsymbol{z}$ in $\mathbb{R}^N$ we have*

$$\begin{aligned}\left(\frac{2Rc^{N-1}}{\Delta}-1\right)\left(\frac{2R}{c}-1\right)^{N-1} &\leq |\Lambda\cap(C_R^N+\boldsymbol{z})| \\ &\leq \left(\frac{2Rc^{N-1}}{\Delta}+1\right)\left(\frac{2R}{c}+1\right)^{N-1}. \qquad (13)\end{aligned}$$

Notice that the assumption on R is not needed for the upper bound of (13). Moreover, this upper bound is sharp: consider the lattice $\Lambda = \Delta\mathbb{Z}\times\mathbb{Z}^{N-1}$ for a fixed Δ.

3. Proof of Theorem 1.2

In fact, we prove a slightly sharper bound that reads as follows.

Theorem 3.1. *Let K be a number field of degree d with discriminant $\mathcal{D}_K$ and r_2 complex places. Let $N \geq 2$ be an integer, and let W be a subspace of K^N of dimension w, $1 \leq w \leq N$. Let $1 \leq s < w$ be an integer, and let $V_1, ..., V_M$ be nonzero subspaces of K^N of corresponding dimensions $l_1, ..., l_M \geq 1$ with $\max_{1\leq i\leq M}\{l_i\} \leq s$. Define*

$$R_1 = \left(\left(\mathcal{C}_K^1(w)\mathcal{H}(W)\right)^{\frac{1}{w-s}}+1\right)\left\{\left(\sum_{i=1}^{M}\frac{\mathcal{C}_{K,N}^2(l_i)}{\mathcal{H}(V_i)^d}\right)^{\frac{1}{(w-s)d}} + M^{\frac{1}{(w-s)d+1}}\right\}, \qquad (14)$$

where

$$\mathcal{C}_K^1(w) = 4^{\frac{w(2d-r_2)+1}{2d}}(wd)^w|\mathcal{D}_K|^{\frac{w}{2d}}, \qquad \mathcal{C}_{K,N}^2(l_i) = \frac{2^{l_i r_2}\binom{Nd}{l_i d}^{1/2}}{|\mathcal{D}_K|^{l_i/2}}, \qquad (15)$$

and

$$R_2 = 2^{\frac{w(d-2r_2)}{2}}wd|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d. \qquad (16)$$

There exists a point $\boldsymbol{x} \in W \setminus \bigcup_{i=1}^{M} V_i$ such that

$$H(\boldsymbol{x}) \leq \max\{R_1, R_2\}.$$

Proof. Let

$$\sigma_1, ..., \sigma_{r_1}, \tau_1, ..., \tau_{r_2}, ..., \tau_{2r_2}$$

be the embeddings of K into $\mathbb{C}$ with $\sigma_1, ..., \sigma_{r_1}$ being real embeddings and $\tau_i, \tau_{r_2+i} = \bar{\tau}_i$ for each $1 \leq i \leq r_2$ being the pairs of complex conjugate embeddings. For each $\alpha \in K$ and each complex embedding τ_i, write $\tau_{i1}(\alpha) = \Re(\tau_i(\alpha))$ and $\tau_{i2}(\alpha) =$

$\Im(\tau_i(\alpha))$, where $\Re$ and $\Im$ stand respectively for real and imaginary parts of a complex number. We will view $\tau_i(\alpha)$ as a pair $(\tau_{i1}(\alpha), \tau_{i2}(\alpha)) \in \mathbb{R}^2$. Then $d = r_1 + 2r_2$, and for each $N \geq 1$ we define an embedding

$$\sigma^N = (\sigma_1^N, ..., \sigma_{r_1}^N, \tau_1^N, ..., \tau_{r_2}^N) : K^N \longrightarrow K_\infty^N,$$

where

$$K_\infty = \prod_{v|\infty} K_v = \prod_{v|\infty} \mathbb{R}^{d_v} = \mathbb{R}^d,$$

since $\sum_{v|\infty} d_v = d$. Then $\sigma^N(O_K^N)$ can be viewed as a lattice of full rank in $\mathbb{R}^{Nd}$.

For $R \geq 1$ let C_R^{Nd} be the cube with sidelength $2R$ centered at the origin in $\mathbb{R}^{Nd}$, as above. Let V be a subspace of K^N of dimension l, $1 \leq l \leq N$. We want to estimate the number of lattice points in the slice of a cube by $\sigma^N(V)$. Let

$$\Lambda(V) = \sigma^N\left(V \cap O_K^N\right),$$

then, by Theorem 2 of [11], $\Lambda(V)$ is a lattice in $\mathbb{R}^{Nd}$ of rank ld, and

$$|\det(\Lambda(V))| = \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l \mathcal{H}(V)^d. \tag{17}$$

Notice that the exponent d on $\mathcal{H}(V)$ appears because our height is absolute unlike the one in Theorem 2 of [11]. Also, the constant 2^{-r_2} appears because we use a slightly different embedding into $\mathbb{R}^{Nd}$ than that in Theorem 2 of [11] (see Lemma 2 on p. 115 of [8]).

On the other hand, let $\boldsymbol{x}_1, ..., \boldsymbol{x}_{ld}$ be a basis for $\Lambda(V)$ as a lattice in $\mathbb{R}^{Nd}$, and write $X = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_{ld}) = (x_{ij})$ for the $Nd \times ld$ basis matrix. Then each row of X consists of blocks of all conjugates of l algebraic integers from O_K. If $I \subset \{1, ..., Nd\}$ with $|I| = ld$, then write X_I for the $ld \times ld$ submatrix of X whose rows are rows of X indexed by I. In other words, X_I is the I-th Grassmann component matrix of X. Then each row of X_I again consists of blocks of all conjugates of l algebraic integers from O_K.

Let $\{v_1, ..., v_{r_1}\} \subset M(K)$ be the places corresponding to the real embeddings $\sigma_1, ..., \sigma_{r_1}$, and let $\{u_1, ..., u_{r_2}\} \subset M(K)$ be the places corresponding to the complex embeddings $\tau_1, ..., \tau_{r_2}$. Let $\alpha \in O_K$, then $|\alpha|_v \leq 1$ for all $v \nmid \infty$, and so $|\alpha|_v \geq 1$ for at least one $v|\infty$, call this place v_*. If v_* is real, say $v_* = v_j$ for some $1 \leq j \leq r_1$, then $|\sigma_j(\alpha)| \geq 1$. If v_* is complex, say $v_* = u_j$ for some $1 \leq j \leq r_2$, then $\sqrt{\tau_{j1}(\alpha)^2 + \tau_{j2}(\alpha)^2} \geq 1$, hence $\max\{|\tau_{j1}(\alpha)|, |\tau_{j2}(\alpha)|\} \geq \frac{1}{\sqrt{2}}$. Therefore,

$$\max\{|\sigma_1(\alpha)|, ..., |\sigma_{r_1}(\alpha)|, |\tau_{11}(\alpha)|, |\tau_{12}(\alpha)|, ..., |\tau_{r_2 1}(\alpha)|, |\tau_{r_2 2}(\alpha)|\} \geq \frac{1}{\sqrt{2}},$$

in other words the maximum of the Euclidean absolute values of all conjugates of an algebraic integer is at least $\frac{1}{\sqrt{2}}$. Therefore the maximum of the Euclidean absolute values of the entries of every row of X_I is at least $\frac{1}{\sqrt{2}}$.

By the Cauchy-Binet formula,

$$\begin{aligned}\max_{|I|=ld} |\det(X_I)| &\leq |\det(\Lambda(V))| \\ &= \left(\sum_{|I|=ld} |\det(X_I)|^2\right)^{1/2} \\ &\leq \binom{Nd}{ld}^{1/2} \max_{|I|=ld} |\det(X_I)|. \qquad (18)\end{aligned}$$

Let $J \subset \{1, ..., Nd\}$ with $|J| = ld$ be such that $|\det(X_J)| = \max_{|I|=ld} |\det(X_I)|$, and let $\Omega(V)$ be the lattice of full rank in $\mathbb{R}^{ld}$ spanned over $\mathbb{Z}$ by the column vectors of X_J. By combining (17) and (18), we see that

$$\begin{aligned}\binom{Nd}{ld}^{-1/2} \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l \mathcal{H}(V)^d &= \binom{Nd}{ld}^{-1/2} |\det(\Lambda(V))| \\ &\leq \det(\Omega(V)) = |\det(X_J)| \\ &\leq |\det(\Lambda(V))| = \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l \mathcal{H}(V)^d. \qquad (19)\end{aligned}$$

For convenience, we denote $\det(\Omega(V))$ by $\Delta(V)$. By Corollary 1 on p. 13 of [3], we can select a basis for $\Omega(V)$ so that the basis matrix is upper triangular, all of its nonzero entries are positive, and the maximum entry of each row occurs on the diagonal. Each of these maximum values is at least $\frac{1}{\sqrt{2}}$, since each row still consists of blocks of all conjugates of l algebraic integers from O_K. Therefore the lattice $\Omega(V)$ satisfies the conditions of Lemma 2.1 with $c = \frac{1}{\sqrt{2}}$. Hence

$$|\Omega(V) \cap C_R^{ld}| \leq \left(\frac{2^{\frac{3}{2}}R}{2^{\frac{ld}{2}}\Delta(V)} + 1\right)(2^{\frac{3}{2}}R + 1)^{ld-1}. \qquad (20)$$

On the other hand, by Theorem 4.3 of [5] (in particular see equation (31) of [5]), we have

$$|\Lambda(V) \cap C_R^{Nd}| \geq |\Omega(V) \cap C_{\frac{R}{ld}}^{ld}|. \qquad (21)$$

Assume that $R \geq 2^{\frac{ld}{2}} ld\Delta(V)$. Then combining (21) with the lower bound of Lemma 2.1, we obtain

$$\begin{aligned}|\Lambda(V) \cap C_R^{Nd}| &\geq \left(\frac{2^{\frac{3}{2}}R}{2^{\frac{ld}{2}}ld\Delta(V)} - 1\right)\left(\frac{2^{\frac{3}{2}}R}{ld} - 1\right)^{ld-1} \\ &\geq \frac{1}{2^{\frac{ld}{2}}\Delta(V)}\left(\frac{R\left(2^{\frac{3}{2}} - 1\right)}{ld}\right)^{ld} \\ &> \frac{R^{ld}}{(ld)^{ld}\Delta(V)}, \qquad (22)\end{aligned}$$

since $2^{\frac{3}{2}} - 1 > \frac{3}{2} > 2^{\frac{1}{2}}$.

For future use, we also need to define a projection $\varphi_V : \Lambda(V) \longrightarrow \Omega(V)$, given by our construction. Namely, if $X\boldsymbol{y} \in \Lambda(V)$ for some $\boldsymbol{y} \in \mathbb{Z}^{Nd}$, then $\varphi_V(X\boldsymbol{y}) = X_J\boldsymbol{y}_J$,

where $\boldsymbol{y}_J \in \mathbb{Z}^{ld}$ is obtained from $\boldsymbol{y}$ by removing all the coordinates which are not indexed by J. It is quite easy to see that φ_V is a $\mathbb{Z}$-module isomorphism.

Now let W be a w-dimensional subspace of K^N, and let $V_1, ..., V_M$ be M proper subspaces of W of respective dimensions $1 \leq l_1, ..., l_M \leq s$. For $R \geq 1$, let

$$S_R(W) = \{\boldsymbol{x} \in W \cap O_K^N : \max_{v|\infty} H_v(\boldsymbol{x})^{d/d_v} \leq R\}, \tag{23}$$

and for each $1 \leq i \leq M$, let $S_R(V_i) = S_R(W) \cap V_i$. Define a counting function

$$f_W(R) = |S_R(W)| - \left|\bigcup_{i=1}^M S_R(V_i)\right| \geq |S_R(W)| - \sum_{i=1}^M |S_R(V_i)|,$$

so that if $f_W(R) > 0$ then there exists a point of height at most R in $W \cap O_K^N$ outside of $\bigcup_{i=1}^M V_i$. Thus we want to find the minimal possible R for which $f_W(R) > 0$. Notice that for each $\boldsymbol{x} \in K^N$,

$$\max_{v|\infty} H_v(\boldsymbol{x})^{d/d_v} = \max_{1 \leq j \leq N} \max\{|\sigma_1(x_j)|, ..., |\sigma_{r_1}(x_j)|, |\tau_1(x_j)|, ..., |\tau_{r_2}(x_j)|\},$$

hence $\sigma^N(S_R(W)) = \sigma^N(W \cap O_K^N) \cap C_R^{Nd}$, and so $|S_R(W)| = |\sigma^N(S_R(W))| = |\Lambda(W) \cap C_R^{Nd}|$, since σ^N is injective. Also, for each $1 \leq i \leq M$ the map $\varphi_{V_i} \circ \sigma^N$ is injective, and if for some $\boldsymbol{x} \in S_R(V_i)$, $\boldsymbol{y} = \varphi_{V_i} \circ \sigma^N(\boldsymbol{x})$, then

$$R \geq \max_{v|\infty} H_v(\boldsymbol{x})^{d/d_v} \geq \max_{1 \leq j \leq l_i d} |y_j|,$$

therefore $\boldsymbol{y} \in \Omega(V_i) \cap C_R^{l_i d}$. This means that for each $1 \leq i \leq M$, we have $|S_R(V_i)| \leq |\Omega(V_i) \cap C_R^{l_i d}|$. Hence we have proved that

$$f_W(R) \geq |\Lambda(W) \cap C_R^{Nd}| - \sum_{i=1}^M |\Omega(V_i) \cap C_R^{l_i d}|,$$

where the notation is as above. From here on assume that $R \geq 2^{\frac{wd}{2}} wd\Delta(W)$. Applying (20) and (22) we obtain

$$\begin{aligned}
f_W(R) &\geq \frac{R^{wd}}{(wd)^{wd}\Delta(W)} - \sum_{i=1}^M \left(\frac{R}{2^{\frac{l_i d - 3}{2}}\Delta(V_i)} + 1\right)(2^{\frac{3}{2}}R + 1)^{l_i d - 1} \\
&\geq \frac{R^{wd}}{(wd)^{wd}\Delta(W)} - (2^{\frac{3}{2}}R + 1)^{sd-1} \sum_{i=1}^M \left(\frac{R}{2^{\frac{d-3}{2}}\Delta(V_i)} + 1\right) \\
&\geq \frac{R^{wd}}{(wd)^{wd}\Delta(W)} - 4^{\left(s - \frac{1}{4}\right)d - \frac{1}{4}} \left(\sum_{i=1}^M \frac{1}{\Delta(V_i)}\right) R^{sd} - 4^{sd-1} M R^{sd-1} \\
&\geq \left(\frac{R^{sd-1}}{(wd)^{wd}\Delta(W)}\right) \times \\
&\times \left\{R^{(w-s)d+1} - (4wd)^{wd}\Delta(W)\left(\sum_{i=1}^M \frac{1}{\Delta(V_i)}\right) R - (4wd)^{wd}\Delta(W)M\right\}. \qquad (24)
\end{aligned}$$

Let $x = \sum_{i=1}^M \frac{1}{\Delta(V_i)}$, and let $\mathcal{A}_W = (4wd)^{wd}\Delta(W)$, and define

$$g_W(R) = R^{(w-s)d+1} - \mathcal{A}_W x R - \mathcal{A}_W M,$$

so that $f_W(R) \geq \frac{R^{sd-1}}{(wd)^{wd}\Delta(W)} g_W(R)$. Hence we want to determine a value of R for which $g_W(R) > 0$. Let $\mathcal{B}_W$ be a positive number to be specified later. Then

$$\begin{aligned}
& g_W\left(\mathcal{B}_W\left(M^{\frac{1}{(w-s)d+1}} + x^{\frac{1}{(w-s)d}}\right)\right) \\
&= \mathcal{B}_W^{(w-s)d+1}\left(M^{\frac{1}{(w-s)d+1}} + x^{\frac{1}{(w-s)d}}\right)^{(w-s)d+1} \\
&\quad - \mathcal{A}_W\mathcal{B}_W\left(M^{\frac{1}{(w-s)d+1}} + x^{\frac{1}{(w-s)d}}\right)x - \mathcal{A}_W M \\
&\geq (\mathcal{B}_W^{(w-s)d+1} - \mathcal{A}_W)M \\
&\quad + \mathcal{B}_W(\mathcal{B}_W^{(w-s)d} - \mathcal{A}_W)x^{1+\frac{1}{(w-s)d}} - \mathcal{A}_W\mathcal{B}_W M^{\frac{1}{(w-s)d+1}} \\
&\geq (\mathcal{B}_W^{(w-s)d+1} - \mathcal{A}_W(\mathcal{B}_W+1))M + \mathcal{B}_W(\mathcal{B}_W^{(w-s)d} - \mathcal{A}_W)x^{1+\frac{1}{(w-s)d}} \\
&> 0, \qquad (25)
\end{aligned}$$

for all M and x if $\mathcal{B}_W \geq 1$, and $\mathcal{B}_W^{(w-s)d} - 2\mathcal{A}_W > 0$, hence we can choose

$$\begin{aligned}
\mathcal{B}_W &= (2\mathcal{A}_W)^{\frac{1}{(w-s)d}} + 1 = \left(4^{wd+\frac{1}{2}}(wd)^{wd}\Delta(W)\right)^{\frac{1}{(w-s)d}} + 1 \\
&\leq \left(4^{\frac{w(2d-r_2)+1}{2}}(wd)^{wd}|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d\right)^{\frac{1}{(w-s)d}} + 1, \qquad (26)
\end{aligned}$$

where the last inequality follows by (19). Therefore, $f_W(R) > 0$ if R is such that

$$\begin{aligned}
R &\geq \left\{\left(4^{\frac{w(2d-r_2)+1}{2}}(wd)^{wd}|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d\right)^{\frac{1}{(w-s)d}} + 1\right\} \times \\
&\times \left\{\left(\sum_{i=1}^{M}\frac{1}{\Delta(V_i)}\right)^{\frac{1}{(w-s)d}} + M^{\frac{1}{(w-s)d+1}}\right\}. \qquad (27)
\end{aligned}$$

Estimating the latter from above using (19), we infer that $f_W(R) > 0$ if

$$\begin{aligned}
R &\geq \left\{\left(4^{\frac{w(2d-r_2)+1}{2}}(wd)^{wd}|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d\right)^{\frac{1}{(w-s)d}} + 1\right\} \times \\
&\times \left\{\left(\sum_{i=1}^{M}\frac{2^{l_i r_2}\binom{Nd}{l_i d}^{1/2}}{|\mathcal{D}_K|^{l_i/2}\mathcal{H}(V_i)^d}\right)^{\frac{1}{(w-s)d}} + M^{\frac{1}{(w-s)d+1}}\right\}. \qquad (28)
\end{aligned}$$

By our original assumption R must also be greater or equal than $2^{\frac{wd}{2}}wd\Delta(W)$. To accomplish this, by (19) we can take

$$R \geq 2^{\frac{w(d-2r_2)}{2}}wd|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d. \qquad (29)$$

Combining (28) with (29) completes the proof. □

Notice that the main part of this argument can be treated as a separate result on the number of points of a subspace of K^N in the adelic cube. Write $K_{\mathbb{A}}$ for the ring of the adeles of K. Define the N-dimensional adelic cube with "sidelength" R to be

$$C_{\mathbb{A}}^N(R) = \prod_{v\nmid\infty} O_v^N \times \prod_{v|\infty}\{\boldsymbol{x} \in K_v^N : H_v(\boldsymbol{x})^{d/d_v} \leq R\}, \qquad (30)$$

for $R \geq 1$. This is a basic example of a compact convex symmetric set in the adelic geometry of numbers (see [2] for details). K^N can be viewed as a lattice in $K_{\mathbb{A}}^N$ under the standard diagonal embedding. For a subspace W of K^N we also write W for its image under this embedding. Clearly $C_{\mathbb{A}}^N(R) \cap W$ is a finite set. In fact, it is precisely the set $S_R(W)$ as defined by (23). The following lemma follows from the argument in the proof of Theorem 3.1 above.

Lemma 3.2. *Let $W \subseteq K^N$ be a w-dimensional subspace, $1 \leq w \leq N$, and let $R \geq 1$. Then*

$$\left(\frac{2^{\frac{w(2r_2-d)+3}{2}}R}{wd|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d} - 1\right)\left(\frac{2^{\frac{3}{2}}R}{wd} - 1\right)^{wd-1} \leq |C_{\mathbb{A}}^N(R) \cap W|$$

$$\leq \left(\frac{\binom{Nd}{wd}^{\frac{1}{2}}2^{\frac{w(2r_2-d)+3}{2}}R}{|\mathcal{D}_K|^{\frac{w}{2}}\mathcal{H}(W)^d} + 1\right)(2^{\frac{3}{2}}R+1)^{wd-1}. \tag{31}$$

Lemma 3.2 presents the counting principle that is our main tool.

4. Corollaries

Notice that in case $K = \mathbb{Q}$ and $s = w - 1$ the bound of Theorem 1.2 becomes

$$(16w)^w \binom{N}{l}^{1/2} \mathcal{H}(W)\left\{\sum_{i=1}^{M}\frac{1}{\mathcal{H}(V_i)} + \sqrt{M}\right\}, \tag{32}$$

which is essentially (up to a constant) the bound of Theorem 5.1 in [5].

Here is another interesting observation that generalizes some ideas of [5]. Suppose that $W = K^N$ and $V_1, ..., V_M$ is a collection of nullspaces of linear forms $L_1, ..., L_M$ in N variables with coefficients in K (i.e. $w = N$ and $l_i = s = N - 1$ for each $1 \leq i \leq M$). Let

$$F(X_1, ..., X_N) = \prod_{i=1}^{M} L_i(X_1, ..., X_N).$$

Then F is a homogeneous polynomial of degree M in N variables with coefficients in K. Hence Theorem 3.1 produces a point $\boldsymbol{x} \in K^N$ of small height at which F does not vanish. In fact, a simple explicit bound on $H(\boldsymbol{x})$ that depends only on K, N, and M follows from Theorem 3.1 in this case:

$$H(\boldsymbol{x}) \leq 2^{N(d+3)+1}\left(Nd|\mathcal{D}_K|\right)^{\frac{N}{2}}\binom{Nd}{Nd-d}^{\frac{1}{2d}}M^{1/d}. \tag{33}$$

Notice that this is a certain inverse of Siegel's Lemma: we produce a point of small height outside of a collection of subspaces. This can also be viewed as an effective instance of the following more general non-effective simple lemma.

Lemma 4.1. *Let K be a number field of degree d, and let F be a polynomial in $N \geq 2$ variables of degree $M \geq 1$ with coefficients in K. There exists a constant $\mathcal{C}_K(N)$ and $\boldsymbol{x} \in O_K^N$ such that $F(\boldsymbol{x}) \neq 0$, and*

$$H(\boldsymbol{x}) \leq \mathcal{C}_K(N)M^{1/d}. \tag{34}$$

Proof. Let

$$S_M(K) = \left\{ x \in K : |x|_v \leq 1 \ \forall \ v \nmid \infty, \quad |x|_v^{d/d_v} \leq \mathcal{C}(K) M^{1/d} \ \forall \ v | \infty \right\},$$

where $\mathcal{C}(K)$ is a positive field constant to be specified later. By [8] (Theorem 0, p. 102) there exist constants $\mathcal{A}(K)$ and $\mathcal{B}(K)$ such that

$$\mathcal{A}(K)\mathcal{C}(K)^d M \leq |S_M(K)| \leq \mathcal{B}(K)\mathcal{C}(K)^d M. \tag{35}$$

Let

$$\mathcal{C}(K) = \left(\frac{2}{\mathcal{A}(K)} \right)^{1/d}, \tag{36}$$

so that $|S_M(K)| \geq 2M \geq M+1$. It is a well-known fact (see for instance Lemma 1 on p. 261 of [3], also Lemma 2.1 of [5]) that a non-zero polynomial of degree M in N variables cannot vanish on the whole set S^N if S is a set of cardinality larger than M. Hence there must exist $\boldsymbol{x} \in S_M(K)^N$ such that $F(\boldsymbol{x}) \neq 0$, and so

$$H(\boldsymbol{x}) \leq \prod_{v|\infty} \left(\mathcal{C}(K) M^{1/d} \right)^{d_v/d} = \mathcal{C}(K) M^{1/d}. \tag{37}$$

This completes the proof. □

Notice that the upper bound in (34) has the correct order of magnitude in the following sense. It is conceptual for the cardinality of the set $S_M(K)$ in the proof of Lemma 4.1 to be at least $M+1$, since there are polynomials of degree M that vanish on a set S^N if $|S| \leq M$: let $S = \{\alpha_1, ..., \alpha_M\} \subset \mathbb{Z}$, and let

$$F(X_1, ..., X_N) = \sum_{i=1}^{N} \prod_{j=1}^{M} (X_i - \alpha_j).$$

Another interesting immediate corollary of Theorem 1.2 in the case $M = 1$ is the following subspace extension lemma.

Corollary 4.2. *Let K be a number field as in Theorem 3.1. Let $N \geq 2$ be an integer, and let W be a subspace of K^N of dimension w, $1 < w \leq N$. Let $V \subseteq W$ be a proper subspace of W of dimension $(w-1) \geq 1$. There exists a point $\boldsymbol{x} \in O_K^N$ such that $W = \mathrm{span}_K\{V, \boldsymbol{x}\}$, and*

$$H(\boldsymbol{x}) \leq \mathcal{C}_{K,N}(w, w-1) \mathcal{H}(W)^d \left(1 + \frac{1}{\mathcal{H}(V)} \right), \tag{38}$$

where the constant $\mathcal{C}_{K,N}(w, w-1)$ is as in (3).

Aknowledgements. I want to thank Professor Jeffrey D. Vaaler for his valuable advice and numerous useful conversations on the subject of this paper. I would also like to thank Professor Preda Mihailescu and the referee for their helpful comments.

References

[1] E. Bombieri and P. B. Cohen. Siegel's lemma, Pade approximations and Jacobians. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25(1-2):155–178, 1998.

[2] E. Bombieri and J. D. Vaaler. On Siegel's lemma. *Invent. Math.*, 73(1):11–32, 1983.

[3] J. W. S. Cassels. *An Introduction to the Geometry of Numbers*. Springer-Verlag, 1959.

[4] L. Fukshansky. *Algebraic points of small height with additional arithmetic conditions*. PhD thesis, University of Texas at Austin, 2004.

[5] L. Fukshansky. Integral points of small height outside of a hypersurface. *Monatsh. Math.*, 147(1):25–41, 2006.

[6] P. Gordan. Uber den grossten gemeinsamen factor. *Math. Ann.*, 7:443–448, 1873.

[7] W. V. D. Hodge and D. Pedoe. *Methods of Algebraic Geometry, Volume 1*. Cambridge Univ. Press, 1947.

[8] S. Lang. *Algebraic Number Theory*. Addison-Wesley, 1970.

[9] C. L. Siegel. Uber einige Anwendungen diophantischer Approximationen. *Abh. der Preuss. Akad. der Wissenschaften Phys.-math Kl.*, Nr. 1:209–266, 1929.

[10] A. Thue. Uber Annaherungswerte algebraischer Zahlen. *J. Reine Angew. Math.*, 135:284–305, 1909.

[11] J. L. Thunder. An asymptotic estimate for heights of algebraic subspaces. *Trans. Amer. Math. Soc.*, 331:395–424, 1992.

[12] J. L. Thunder. The number of solutions of bounded height to a system of linear equations. *J. Number Theory*, 43:228–250, 1993.

[13] J. D. Vaaler. The best constant in Siegel's lemma. *Monatsh. Math.*, 140(1):71–89, 2003.

Department of Mathematics, 3368 TAMU, Texas A&M University, College Station, Texas 77843-3368

E-mail address: lenny@math.tamu.edu