EFFECTIVE STRUCTURE THEOREMS FOR SYMPLECTIC SPACES VIA HEIGHT

LENNY FUKSHANSKY

ABSTRACT. Given a $2k$-dimensional symplectic space (Z, F) in N variables, $1 < 2k \leq N$, over a global field K, we prove the existence of a symplectic basis for (Z, F) of bounded height. This can be viewed as a version of Siegel's lemma for a symplectic space. As corollaries of our main result, we prove the existence of a small-height decomposition of (Z, F) into hyperbolic planes, as well as the existence of two generating flags of totally isotropic subspaces. These present analogues of known results for quadratic spaces. A distinctive feature of our argument is that it works simultaneously for essentially any field with a product formula, algebraically closed or not. In fact, we prove an even more general version of these statements, where canonical height is replaced with twisted height. All bounds on height are explicit.

1. INTRODUCTION

Throughout this paper, we let K be either a number field, a function field (i.e. a finite algebraic extension of the field of rational functions in one variable over an arbitrary field), or the algebraic closure of one or the other. Let

$$F(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_{ij} X_i Y_j \tag{1}$$

be an alternating bilinear form in $N \geq 2$ variables with coefficients in K. We will also write $F = (f_{ij})_{1 \leq i,j \leq N}$ for the anti-symmetric $N \times N$ coefficient matrix of F, i.e. $f_{ij} = -f_{ji}$ for all $1 \leq i, j \leq N$. In particular, $f_{ii} = 0$ for all $1 \leq i \leq N$, and the associated quadratic form $F(\boldsymbol{X}) = F(\boldsymbol{X}, \boldsymbol{X})$ is identically zero on K^N.

Let Z be a $2k$-dimensional subspace of K^N, $1 \leq k \leq N/2$, and let us write (Z, F) for the symplectic space defined on Z by F. We will assume that (Z, F) is regular, meaning that for every $\boldsymbol{0} \neq \boldsymbol{x} \in Z$ there exists $\boldsymbol{y} \in Z$ such that $F(\boldsymbol{x}, \boldsymbol{y}) \neq 0$. Then (Z, F) has a symplectic basis (see for instance [7]), that is a basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k$ for Z over K such that

$$F(\boldsymbol{x}_i, \boldsymbol{x}_j) = F(\boldsymbol{y}_i, \boldsymbol{y}_j) = F(\boldsymbol{x}_i, \boldsymbol{y}_j) = 0 \ \forall \ 1 \leq i \neq j \leq k, \ F(\boldsymbol{x}_i, \boldsymbol{y}_i) = 1 \ \forall \ 1 \leq i \leq k. \tag{2}$$

A subspace V of Z is called totally isotropic if $F(\boldsymbol{x}, \boldsymbol{y}) = 0$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$, and a maximal totally isotropic subspace of (Z, F) is called a Lagrangian. All Lagrangians of (Z, F) have the same dimension; it is an easy consequence of (2) that this dimension is k. Indeed, it is easy to see that $V_1 = \mathrm{span}_K\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\}$ and $V_2 = \mathrm{span}_K\{\boldsymbol{y}_1, \dots, \boldsymbol{y}_k\}$ are Lagrangians in (Z, F). Moreover, (Z, F) is a

1991 *Mathematics Subject Classification.* Primary 11E12, 11G50, 11H55, 11D09.
Key words and phrases. quadratic and bilinear forms, symplectic spaces, heights.

hyperbolic space over these Lagrangians, meaning that

$$Z = \mathbb{H}_1 \perp \cdots \perp \mathbb{H}_k, \tag{3}$$

where for each $1 \le i \le k$, $\mathbb{H}_i = \mathrm{span}_K\{\boldsymbol{x}_i, \boldsymbol{y}_i\}$ is a hyperbolic plane, $\perp$ stands for orthogonal direct sum, and orthogonality throughout this paper is always meant with respect to F. This means that once we know how to find a symplectic basis for (Z, F), we immediately obtain two Lagrangians as well as an orthogonal decomposition of (Z, F) into hyperbolic planes. However the classical result about the existence of a basis satisfying (2) is ineffective, i.e. it provides no information as to how does one find such a basis.

The main goal of this paper is to prove an effective version of the existence theorem for a symplectic basis, and derive from it effective statements about existence of Lagrangians and a hyperbolic decomposition for a regular symplectic space. We use the approach of height functions, which will be formally introduced in section 2. We will define a height function H on the points of a projective space over K, and in particular will talk about height of vectors and subspaces of K^N to mean height of the corresponding projective points; specifically, subspaces of K^N will be viewed as points on a corresponding Grassmanian. We will also give a slightly different definition for the height $\mathcal{H}$ of our alternating bilinear form F. Loosely speaking, height measures the arithmetic complexity of objects in question, meaning that the smaller is the height of a projective point the less "arithmetically complicated" this point is. In particular, height satisfies the crucial finiteness property: any set of projective algebraic points of bounded height and degree is always finite (this will be rigorously discussed in section 2, especially see (11)). Therefore, proving the existence of a point or subspace of bounded height over K that satisfies some arithmetic conditions may provide a search bound for points satisfying such conditions. Hence our goal will be to prove effective theorems for symplectic spaces in the sense of providing bounds on height. We can now state our main result.

Theorem 1.1. *Let (Z, F) be a regular $2k$-dimensional symplectic space in N variables over K, where $1 \le k < 2k \le N$. Then there exists a symplectic basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k, \boldsymbol{y}_1, \dots, \boldsymbol{y}_k$ for Z satisfying (2) such that*

$$\prod_{i=1}^{2k} H(\boldsymbol{x}_i)H(\boldsymbol{y}_i) \le (C_K(N, 2k)H(Z))^{a_k}\, \mathcal{H}(F)^{b_k}, \tag{4}$$

where $C_K(N, 2k)$ is a field constant defined in section 2 below,

$$a_k = \begin{cases} \frac{k^2+4k}{4} & \text{if } 2|k \\ \frac{k^2+4k-1}{4} & \text{if } 2 \nmid k, \end{cases}$$

and

$$b_k = \begin{cases} \frac{2k^3+9k^2-14k}{12} & \text{if } 2|k \\ \frac{2k^3+9k^2-14k+3}{12} & \text{if } 2 \nmid k. \end{cases}$$

An immediate corollary of Theorem 1.1 is an effective version of Witt decomposition for (Z, F), which in the symplectic case is just a decomposition into hyperbolic planes.

Corollary 1.2. *Let the notation be as in Theorem 1.1, then there exists a decomposition (3) for (Z, F) with*

$$\prod_{i=1}^{2k} H(\mathbb{H}_i) \leq (C_K(N, 2k)H(Z))^{a_k}\, \mathcal{H}(F)^{b_k}. \tag{5}$$

Proof. For each $1 \leq i \leq k$, take $\mathbb{H}_i = \operatorname{span}_K\{\boldsymbol{x}_i, \boldsymbol{y}_i\}$, then by Lemma 2.3 below $H(\mathbb{H}_i) \leq H(\boldsymbol{x}_i)H(\boldsymbol{y}_i)$, and the statement of the corollary follows from (4). □

We can now also establish the existence of flags of totally isotropic subspaces of bounded height, whose union generates Z.

Corollary 1.3. *Let the notation be as in Theorem 1.1. For each $1 \leq n \leq k$, there exist totally isotropic subspaces V_n and W_n of (Z, F) such that $\dim_K V_n = \dim_K W_n = n$, $V_n \cap W_n = \{\mathbf{0}\}$,*

$$V_1 \subset V_2 \subset \cdots \subset V_k,\ W_1 \subset W_2 \subset \cdots \subset W_k, \tag{6}$$

and

$$H(V_n)H(W_n) \leq \left(C_K(N, 2k)^{a_k} H(Z)^{a_k} \mathcal{H}(F)^{b_k}\right)^{\frac{n}{k}}. \tag{7}$$

In particular, (Z, F) is generated by the two small-height Lagrangians V_k and W_k, i.e. $Z = \operatorname{span}_K\{V_k, W_k\}$.

Proof. With notation of Theorem 1.1, assume without loss of generality that the symplectic basis vectors are ordered in such a way that

$$H(\boldsymbol{x}_1)H(\boldsymbol{y}_1) \leq H(\boldsymbol{x}_2)H(\boldsymbol{y}_2) \leq \ldots H(\boldsymbol{x}_k)H(\boldsymbol{y}_k).$$

Then let $V_n = \operatorname{span}_K\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ and $W_n = \operatorname{span}_K\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$, for each $1 \leq n \leq k$, and notice that by Lemma 2.3,

$$H(V_n)H(W_n) \leq H(\boldsymbol{x}_1) \ldots H(\boldsymbol{x}_n)H(\boldsymbol{y}_1) \ldots H(\boldsymbol{y}_n).$$

The statement of the corollary now follows from (4). □

These results should be viewed as symplectic space analogues of Siegel's lemma and effective decomposition theorems for quadratic spaces. The name Siegel's lemma usually refers to results about the existence of a basis of small height for a vector space over a global field (see [1], [12], and [5]). In case of a quadratic space (i.e. when F is a symmetric bilinear form), a version of Siegel's lemma with additional conditions, asserting the existence of an orthogonal basis of small height, has been proved in [4] over a number field and in [3] over $\overline{\mathbb{Q}}$. Theorem 1.1 is precisely a symplectic space analogue of these theorems.

There has been a large number of results on small-height zeros of quadratic forms, starting with a classical theorem of Cassels [2]. One of the directions generalizing Cassels' theorem produced results on small-height linear subspaces of a quadratic space on which the quadratic form vanishes identically (see [8], [9], [13], and [14]). Corollary 1.3 should be viewed as an analogue of these results for a symplectic space. Finally, the structural results for a quadratic space, such as effective Witt decomposition, have been proved in [4] and [3]; Corollary 1.2 serves as a symplectic space analogue of this.

In the case of a quadratic space, such problems were usually treated separately by different methods over a number field, function field, or algebraic closures. The distinctive feature of the symplectic situation is that, because it is much more

linear, we are able to treat these problems at once over any global field with a product formula for which a Siegel's lemma type result exists - this is due to the purely combinatorial nature of our argument. Moreover, we prove our main result in terms of more general twisted heights (see Theorem 4.2), from which Theorem 1.1 follows immediately.

This paper is structured as follows. In section 2 we set the notation and define the height functions, and present a few technical lemmas on properties of heights. In section 3 we prove a combinatorial lemma (Lemma 3.1), which we later use to obtain Theorem 1.1. In section 4 we derive Theorem 1.1 by means of proving the more general Theorem 4.2 with the use of Siegel's lemma, stated as Theorem 4.1, and Lemma 3.1.

2. Notation and heights

We start with some notation, following [5]. Throughout this paper, K will either be a number field (finite extension of $\mathbb{Q}$), a function field, or algebraic closure of one or the other; in fact, for the rest of this section, unless explicitly specified otherwise, we will assume that K is either a number field or a function field, and will write $\overline{K}$ for its algebraic closure. By a function field we will always mean a finite algebraic extension of the field $\mathfrak{K} = \mathfrak{K}_0(t)$ of rational functions in one variable over a field $\mathfrak{K}_0$, where $\mathfrak{K}_0$ can be any field. When K is a number field, clearly $K \subset \overline{K} = \overline{\mathbb{Q}}$; when K is a function field, $K \subset \overline{K} = \overline{\mathfrak{K}}$, the algebraic closure of $\mathfrak{K}$. In the number field case, we write $d = [K : \mathbb{Q}]$ for the global degree of K over $\mathbb{Q}$; in the function field case, the global degree is $d = [K : \mathfrak{K}]$, and we also define the effective degree of K over $\mathfrak{K}$ to be

$$\mathfrak{m}(K, \mathfrak{K}) = \frac{[K : \mathfrak{K}]}{[K_0 : \mathfrak{K}_0]},$$

where K_0 is the algebraic closure of $\mathfrak{K}_0$ in K. If K is a number field, we let $\mathcal{D}_K$ be its discriminant; if K is a function field, we will also write $g(K)$ for the genus of K, as defined by the Riemann-Roch theorem (see [12] for details). We can now define the field constant $C_K(N, L)$, which appears in our upper bounds:

$$C_K(N,L) = \begin{cases} N^{\frac{L}{2}} |\mathcal{D}_K|^{\frac{L}{2d}} & \text{if } K \text{ is a number field} \\ \exp\left(\frac{g(K)-1+\mathfrak{m}(K,\mathfrak{K})}{\mathfrak{m}(K,\mathfrak{K})}\right) & \text{if } K \text{ is a function field} \\ 3^{\frac{L(L-1)}{2}} & \text{if } K = \overline{\mathbb{Q}} \\ 2 & \text{if } K = \overline{\mathfrak{K}}, \end{cases}$$

Next we discuss absolute values on K. Let $M(K)$ be the set of all places of K when K is a number field or $\overline{\mathbb{Q}}$, and the set of all places of K which are trivial over the field of constants when K is a function field or its algebraic closure. For each place $v \in M(K)$ we write K_v for the completion of K at v and let d_v be the local degree of K at v, which is $[K_v : \mathbb{Q}_v]$ in the number field case, and $[K_v : \mathfrak{K}_v]$ in the function field case.

If K is a number field, then 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 rational prime.

If K is a function field, then all absolute values on K are non-archimedean. For each $v \in M(K)$, let $\mathfrak{O}_v$ be the valuation ring of v in K_v and $\mathfrak{M}_v$ the unique maximal ideal in $\mathfrak{O}_v$. We choose the unique corresponding absolute value $|\ |_v$ such that:

(i) if $1/t \in \mathfrak{M}_v$, then $|t|_v = e$,
(ii) if an irreducible polynomial $p(t) \in \mathfrak{M}_v$, then $|p(t)|_v = e^{-\deg(p)}$.

In both cases, for each non-zero $a \in K$ the *product formula* reads

$$\prod_{v \in M(K)} |a|_v^{d_v} = 1. \tag{8}$$

We can now define local norms on vectors. For each $v \in M(K)$ define a local norm $\| \ \|_v$ on K_v^N by

$$\|\boldsymbol{x}\|_v = \begin{cases} \max_{1 \le i \le N} |x_i|_v & \text{if } v \nmid \infty \\ \left(\sum_{i=1}^N |x_i|_v^2\right)^{1/2} & \text{if } v|\infty \end{cases}$$

for each $\boldsymbol{x} \in K_v^N$. We define the following global height function on K^N:

$$H(\boldsymbol{x}) = \left(\prod_{v \in M(K)} \|\boldsymbol{x}\|_v^{d_v} \right)^{1/d}, \tag{9}$$

for each $\boldsymbol{x} \in K^N$. More generally, let us define the *twisted height* on K^N as introduced by J. L. Thunder. We write $K_{\mathbb{A}}$ for the ring of adeles of K, and view K as a subfield of $K_{\mathbb{A}}$ under the diagonal embedding (see [16] for details). Let $A \in GL_N(K_{\mathbb{A}})$ with local components $A_v \in GL_N(K_v)$. The corresponding twisted height on K^N is defined by

$$H_A(\boldsymbol{x}) = \left(\prod_{v \in M(K)} \|A_v \boldsymbol{x}\|_v^{d_v} \right)^{1/d}, \tag{10}$$

for all $\boldsymbol{x} \in K^N$. Given any finite extension E/K, $K_{\mathbb{A}}$ can be viewed as a subring of $E_{\mathbb{A}}$, and let us also write A for the element of $GL_N(E_{\mathbb{A}})$ which coincides with A on $K_{\mathbb{A}}^N$. The corresponding twisted height on E^N extends the one on K^N, hence H_A is a height on $\overline{K}$. Notice also that the usual height H as defined above is simply H_I, where I is the identity element of $GL_N(K_{\mathbb{A}})$ all of whose local components are given by $N \times N$ identity matrices. Due to the normalizing exponent $1/d$, our height functions are absolute, i.e. for points over $\overline{\mathbb{Q}}$ or $\overline{\mathfrak{K}}$, respectively, their value does not depend on the field of definition. This means that if $\boldsymbol{x}$ is in $\overline{\mathbb{Q}}^N$ or $\overline{\mathfrak{K}}^N$, then for every $A \in GL_N(K_{\mathbb{A}})$, $H_A(\boldsymbol{x})$ can be evaluated over any number field or function field, respectively, containing the coordinates of $\boldsymbol{x}$, and so H_A provides a height on $\overline{K}^N$.

A fundamental property of heights (in case K is a number field or a function field with a finite field of constants $\mathfrak{K}_0$), sometimes referred to as the Northcott property, is that for every $A \in GL_N(K_{\mathbb{A}})$,

$$\left| \left\{ [\boldsymbol{\alpha}] \in \mathbb{P}^{N-1}(\overline{K}) \ : \ \deg([\boldsymbol{\alpha}]) \le B, H_A(\boldsymbol{\alpha}) \le C \right\} \right| < \infty, \tag{11}$$

where $\mathbb{P}^{N-1}(\overline{K})$ is $(N-1)$-dimensional projective space over $\overline{K}$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is in $\overline{K}^N$, and so $[\boldsymbol{\alpha}]$ is the corresponding projective point, B, C are positive real numbers, and $\deg(\boldsymbol{\alpha}) = [\mathbb{Q}(\alpha_1, \ldots, \alpha_N) : \mathbb{Q}]$ if K is a number field, or $[\mathfrak{K}(\alpha_1, \ldots, \alpha_N) : \mathfrak{K}]$ if K is a function field, i.e. it is the algebraic degree of $\boldsymbol{\alpha}$ over the ground field

over which K is defined. We define $\deg([\boldsymbol{\alpha}])$, the degree of the projective point represented by $\boldsymbol{\alpha}$, by

$$\deg([\boldsymbol{\alpha}]) = \min\{\deg(\boldsymbol{\alpha}') \ : \ \boldsymbol{\alpha}' \in \overline{K}^N, \ [\boldsymbol{\alpha}'] = [\boldsymbol{\alpha}]\}.$$

We can now extend our notation to define Schmidt twisted height on matrices, which is the same as height function on subspaces of $\overline{K}^N$. Let $A \in GL_N(K_{\mathbb{A}})$, $\boldsymbol{e}_1, \dots, \boldsymbol{e}_N$ be the standard basis for K^N, and $1 \le J \le N$. Then J-th exterior component $\bigwedge^J K^N$ can be identified with the vector space $K^{\binom{N}{J}}$ via the cannonical isomorphism that sends the wedge products $\boldsymbol{e}_{i_1} \wedge \cdots \wedge \boldsymbol{e}_{i_J}$, $1 \le i_1 < \cdots < i_J \le N$, to the standard basis elements of $K^{\binom{N}{J}}$ in lexicographic order. This also identifies $\bigwedge^J A$ with an element of $GL_{\binom{N}{J}}(K_{\mathbb{A}})$, and so we can talk about the height $H_{\bigwedge^J A}$ on $\bigwedge^J \overline{K}^N$. Let X be an $N \times J$ matrix of rank J whose column vectors are $\boldsymbol{x}_1, ..., \boldsymbol{x}_J \in K^N$, then we define

$$H_A(X) = H_{\bigwedge^J A}(\boldsymbol{x}_1 \wedge \ ... \ \wedge \boldsymbol{x}_J).$$

In the same manner, we define the height of a $J \times N$ matrix to be the height of the wedge product of its row vectors instead of column vectors. Now let $V \subseteq \overline{K}^N$ be a subspace of dimension J, $1 \le J \le N$, defined over K. Choose a basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_J$ for V over K, and write $X = (\boldsymbol{x}_1 \dots \boldsymbol{x}_J)$ for the corresponding $N \times J$ basis matrix. Define the height of V by $H_A(V) = H_A(X)$. This definition is legitimate, since it does not depend on the choice of the basis for V: let $\boldsymbol{y}_1, ..., \boldsymbol{y}_J$ be another basis for V over K and $Y = (\boldsymbol{y}_1 \dots \boldsymbol{y}_J)$ the corresponding $N \times J$ basis matrix, then there exists $W \in GL_J(K)$ such that $Y = XW$, and so

$$\boldsymbol{y}_1 \wedge \ ... \ \wedge \boldsymbol{y}_J = (\det W) \ \boldsymbol{x}_1 \wedge \ ... \ \wedge \boldsymbol{x}_J,$$

hence, by the product formula

$$H_A(Y) = H_{\bigwedge^J A}(\boldsymbol{y}_1 \wedge \ ... \ \wedge \boldsymbol{y}_J) = H_{\bigwedge^J A}(\boldsymbol{x}_1 \wedge \ ... \ \wedge \boldsymbol{x}_J) = H_A(X).$$

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

$$V = \left\{ \boldsymbol{x} \in \overline{K}^N : B\boldsymbol{x} = 0 \right\}. \tag{12}$$

An important duality principle relates heights of V and B. For $A \in GL_N(K_{\mathbb{A}})$ with local components $A_v \in GL_N(K_v)$ for every $v \in M(K)$, let $A^* \in GL_N(K_{\mathbb{A}})$ be given by the local components $(A_v^t)^{-1} \in GL_N(K_v)$ for every $v \in M(K)$. We also define

$$|\det A|_{\mathbb{A}} = \left(\prod_{v \in M(K)} |\det A_v|_v^{d_v} \right)^{1/d}.$$

The following is Theorem 1.1 of [5] (see also Duality Theorem in section 2 of [11]).

Lemma 2.1. *For any subspace $V \subseteq \overline{K}^N$ and $A \in GL_N(K_{\mathbb{A}})$, we have*

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

where B is as in (12).

In particular, this implies that $H(V) = H(B)$ if B is as in (12), since clearly for the identity $I \in GL_N(K_{\mathbb{A}})$, $I^* = I$ and $|\det I|_{\mathbb{A}} = 1$.

We also define height of our bilinear form F in the following conventional way: let $\mathcal{H}(F)$ be the usual height H of the anti-symmetric matrix $(f_{ij})_{1\leq i,j\leq N}$, viewed as a vector in K^{N^2}. Notice that it is different from the height on matrices defined above, which is why we denote it by $\mathcal{H}$ instead of H.

Finally, we define certain dilation constants for an element $A \in GL_N(K_{\mathbb{A}})$ that will appear in our bounds (see Lemmas 3.1, 3.2, and Proposition 4.1 of [5]; see also [6]). Roughly speaking, as we will see in Lemma 2.2 below, these constants indicate by how much does a given automorphism A of $K_{\mathbb{A}}^N$ "'distort"' the corresponding twisted height H_A as compared to H, the canonical height. Let $A_v = (a_{ij}^v)_{1\leq i,j\leq N} \in GL_N(K_v)$ be local components of A for each $v \in M(K)$, and let us write $A_v^{-1} = (b_{ij}^v)_{1\leq i,j\leq N}$. Then for all but finitely many places $v \in M(K)$ the corresponding map A_v is an isometry; in fact, let $M_A(K) \subset M(K)$ be the finite (possibly empty) subset of places v at which A_v is *not* an isometry. For each $v \notin M_A(K)$, define $\mathcal{C}_1^v(A) = \mathcal{C}_2^v(A) = 1$, and for each $v \in M_A(K)$, let

$$\mathcal{C}_1^v(A) = \left(\sum_{l=1}^N \sum_{m=1}^N |b_{lm}^v|_v\right)^{-1},\ \mathcal{C}_2^v(A) = \sum_{i=1}^N \sum_{j=1}^N |a_{ij}^v|_v. \tag{13}$$

Then define

$$\mathcal{C}_1(A) = \prod_{v\in M(K)} (\mathcal{C}_1^v)^{d_v/d},\ \mathcal{C}_2(A) = \prod_{v\in M(K)} (\mathcal{C}_2^v)^{d_v/d}, \tag{14}$$

both of which are products of only a finite number of non-trivial terms. With this notation, it will also be convenient to define

$$\mathfrak{C}(A) = \frac{\mathcal{C}_2(A)}{\mathcal{C}_1(A)} = \prod_{v\in M_A(K)} \left(\sum_{i,j,l,m=1}^N |a_{ij}^v b_{lm}^v|_v\right)^{d_v/d}, \tag{15}$$

and

$$\mathfrak{C}'(A) = \frac{\mathfrak{C}(A)|\det A|_{\mathbb{A}}^{1/2}}{\mathcal{C}_1(A)^2}. \tag{16}$$

Clearly, in the case when $A = I$ is the identity element of $GL_N(K_{\mathbb{A}})$, $\mathfrak{C}'(A) = \mathfrak{C}(A) = \mathcal{C}_1(A) = \mathcal{C}_2(A) = 1$. Another important observation is that, since for every $v \in M(K)$, $(A_v^t)^{-1} = (A_v^{-1})^t$, therefore

$$\begin{aligned} &\mathcal{C}_1^v(A^*)^{-1} = \mathcal{C}_2^v(A),\ \mathcal{C}_2^v(A^*) = \mathcal{C}_1^v(A)^{-1}, \\ &\mathcal{C}_1(A^*)^{-1} = \mathcal{C}_2(A),\ \mathcal{C}_2(A^*) = \mathcal{C}_1(A)^{-1}, \\ &\mathfrak{C}(A^*) = \mathfrak{C}(A). \end{aligned} \tag{17}$$

Next we present some technical lemmas that we use later in our main proof, detailing the key properties of height functions. The first one shows that the canonical height H and the twisted height H_A are comparable for each $A \in GL_N(K_{\mathbb{A}})$ with the comparison constants being precisely the dilation constants $\mathcal{C}_1(A), \mathcal{C}_2(A)$ defined above. This is Proposition 4.1 of [5].

Lemma 2.2. *Let $A \in GL_N(K_{\mathbb{A}})$. Then*

$$\mathcal{C}_1(A)H(\boldsymbol{x}) \le H_A(\boldsymbol{x}) \le \mathcal{C}_2(A)H(\boldsymbol{x}), \tag{18}$$

for all $\boldsymbol{x} \in \overline{K}^N$, where $\mathcal{C}_1(A)$ and $\mathcal{C}_2(A)$ are as in (14) above.

Remark 2.1. A simple consequence of Lemma 2.2 and (17) which will be useful to us is that for all $\boldsymbol{x} \in \overline{K}^N$,

$$H_{A^*}(\boldsymbol{x}) \le \mathcal{C}_1(A)^{-2} H_A(\boldsymbol{x}). \tag{19}$$

The next lemma is a consequence of Laplace's expansion, and can be found as Lemma 4.7 of [5] (also see pp. 15-16 of [1]).

Lemma 2.3. *Let X be a $N \times J$ matrix over $\overline{K}$ with column vectors $\boldsymbol{x}_1, ..., \boldsymbol{x}_J$, and let $A \in GL_N(K_{\mathbb{A}})$. Then*

$$H_A(X) = H_{\bigwedge^J A}(\boldsymbol{x}_1 \wedge \boldsymbol{x}_1 \ ... \ \wedge \boldsymbol{x}_J) \le \prod_{i=1}^{J} H_A(\boldsymbol{x}_i). \tag{20}$$

More generally, if the $N \times J$ matrix X can be partitioned into blocks as $X = (X_1 \ X_2)$, then

$$H_A(X) \le H_A(X_1)H_A(X_2). \tag{21}$$

The following well known fact is an immediate corollary of Theorem 1 of [10] adapted over $\overline{K}$ and extended to twisted height.

Lemma 2.4. *Let U_1 and U_2 be subspaces of $\overline{K}^N$, and let $A \in GL_N(K_{\mathbb{A}})$. Then*

$$H_A(U_1 \cap U_2) \le H_A(U_1)H_A(U_2).$$

The next one is a generalization of Lemma 2.3 of [4] over $\overline{K}$ and with the twisted height H_A replacing canonical height H. We present the proof here for the purposes of self-containment.

Lemma 2.5. *Let X be a $N \times J$ matrix over $\overline{K}$ with column vectors $\boldsymbol{x}_1, ..., \boldsymbol{x}_J$, $A \in GL_N(K_{\mathbb{A}})$, and let F be a bilinear form in N variables, as above (we also write F for its $N \times N$ coefficient matrix). Then*

$$H_A(FX) \le \mathfrak{C}(A)^J \mathcal{H}(F)^J \prod_{i=1}^{J} H_A(\boldsymbol{x}_i), \tag{22}$$

where $\mathfrak{C}(A)$ is as in (15). In particular, this implies that

$$H(FX) \le \mathcal{H}(F)^J \prod_{i=1}^{J} H(\boldsymbol{x}_i). \tag{23}$$

Proof. By Lemmas 2.3 and 2.2,

$$H_A(FX) = H_{\bigwedge^J A}(\boldsymbol{x}_1^t F \wedge \ ... \ \wedge \boldsymbol{x}_J^t F) \le \prod_{i=1}^{J} H_A(\boldsymbol{x}_i^t F) \le \mathcal{C}_2(A)^J \prod_{i=1}^{J} H(\boldsymbol{x}_i^t F). \tag{24}$$

For each $1 \le i \le J$,

$$\boldsymbol{x}_i^t F = \left(\sum_{j=1}^{N} f_{j1} x_{ij}, ..., \sum_{j=1}^{N} f_{jN} x_{ij} \right).$$

Recall that for the purposes of evaluating height we view the coefficient matrix $F = (f_{ij})_{1 \leq i,j \leq N}$ as a vector in K^{N^2}, and we write $\|F\|_v$ for the local norm of this vector at the place v. Then for each $v \nmid \infty$,

$$\|\boldsymbol{x}_i^t F\|_v \leq \|F\|_v \|\boldsymbol{x}_i\|_v, \tag{25}$$

and for $v|\infty$, by Cauchy-Schwarz inequality

$$\begin{aligned} \|\boldsymbol{x}_i^t F\|_v &= \left\{ \sum_{k=1}^{N} \left\| \sum_{j=1}^{N} f_{jk} x_{ij} \right\|_v^2 \right\}^{d_v/2d} \\ &\leq \left\{ \sum_{k=1}^{N} \left(\sum_{j=1}^{N} \|f_{jk}\|_v^2 \right) \left(\sum_{j=1}^{N} \|x_{ij}\|_v^2 \right) \right\}^{d_v/2d} = \|F\|_v \|\boldsymbol{x}_i\|_v. \end{aligned} \tag{26}$$

Therefore for each $1 \leq i \leq J$,

$$H(\boldsymbol{x}_i^t F) \leq H(\boldsymbol{x}_i)\mathcal{H}(F) \leq \mathcal{C}_1(A)^{-1} H_A(\boldsymbol{x}_i)\mathcal{H}(F), \tag{27}$$

where the last inequality follows by Lemma 2.2. Now the lemma follows by combining (24) with (27). □

Remark 2.2. Notice that Lemma 2.5 is true for any bilinear form F, symmetric, alternating, or none of the above - the proof carries over word for word. Moreover, F can just as well be any $N \times N$ matrix, viewed as a vector in K^{N^2} for the purposes of defining the height $\mathcal{H}(F)$.

We are now ready to proceed.

3. A combinatorial lemma

In this section we prove a certain graph-theoretic lemma, which we later use in the proof of our main result. We start with some notation. A graph G is connected if there is a path in G connecting every two of its vertices. On the other hand, we will call a pair of vertices connected if they are connected by a single edge, and disconnected otherwise. A graph in which every two vertices are connected is called complete. A complete subgraph on n vertices of a graph G will be called maximal if G does not contain a complete subgraph on any larger number of vertices. Two pairs of vertices in a graph G will be called disjoint if they do not have a vertex in common. We can now state the lemma.

Lemma 3.1. *Let G be a graph on $2k$ vertices, $k \geq 1$, such that a maximal complete subgraph of G has at most k vertices. Then there exist at least $\left[\frac{k+1}{2}\right]$ disjoint pairs of disconnected vertices. Moreover, this bound is sharp, meaning that there are such graphs in which any maximal (with respect to cardinality) set of disjoint pairs of disconnected vertices has cardinality precisely $\left[\frac{k+1}{2}\right]$.*

Proof. Let $v_1, \ldots, v_{2k}$ be the vertices of G. For each $1 \leq i \neq j \leq 2k$, define

$$\delta_{ij} = \delta_{ji} = \begin{cases} 1 & \text{if } v_i \text{ is connected to } v_j, \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$S_1 = \{1, \ldots, k+1\},$$

then there must exist $i_1 \neq j_1 \in S_1$ such that $\delta_{i_1 j_1} = 0$: if this was not true, then G would contain a complete subgraph on $k+1$ vertices $v_1, \dots, v_{k+1}$. Next, let

$$S_2 = (S_1 \setminus \{i_1, j_1\}) \cup \{k+2, k+3\}.$$

Since $|S_2| = k+1$, by the same reasoning, there must exist $i_2 \neq j_2 \in S_2$ such that $\delta_{i_2 j_2} = 0$, and next define

$$S_3 = (S_2 \setminus \{i_2, j_2\}) \cup \{k+4, k+5\}.$$

Continuing in this manner, in each set

$$S_n = (S_{n-1} \setminus \{i_{n-1}, j_{n-1}\}) \cup \{k+2n-2, k+2n-1\}, \tag{28}$$

we will find vertices v_{i_n}, v_{j_n} such that $\delta_{i_n j_n} = 0$. From (28), we see that $1 \leq n \leq M = \left[\frac{k+1}{2}\right]$, and so we get a collection of distinct vertices

$$\{v_{i_1}, v_{j_1}, \dots, v_{i_M}, v_{j_M}\} \subset \{v_1, \dots, v_{2k}\}, \tag{29}$$

which satisfy the condition

$$\delta_{i_n j_n} = 0, \quad \forall\ 1 \leq n \leq M = \left[\frac{k+1}{2}\right]. \tag{30}$$

This is precisely a collection of $\left[\frac{k+1}{2}\right]$ disjoint pairs of disconnected vertices in G.

Next we show that $\left[\frac{k+1}{2}\right]$ is sharp. Let G be a graph on vertices $v_1, \dots, v_{2k}$ as above so that $\delta_{ij} = 1$ for all $i \neq j$ such that $1 \leq i \leq k-1$ and $1 \leq j \leq 2k$, and $\delta_{ij} = 0$ for all $k \leq i \neq j \leq 2k$; in other words, each of the first $k-1$ vertices is connected to every other vertex in G, but no two vertices out of $v_k, \dots, v_{2k}$ are connected to each other. Clearly, any maximal complete subgraph of G will have k vertices; in fact, these will be precisely the $k+1$ subgraphs on the sets of vertices $\{v_1, \dots, v_{k-1}, v_j\}$ for each $k \leq j \leq 2k$. Then a maximal (with respect to cardinality) set of disjoint pairs of disconnected vertices is, for instance the set of pairs $v_k, v_{k+1}; \dots; v_{2k-2}, v_{2k-1}$, if k is even, and $v_k, v_{k+1}; \dots; v_{2k-1}, v_{2k}$, if k is odd. In both cases, the cardinality of such a set is $\left[\frac{k+1}{2}\right]$. This completes the proof. $\square$

4. Proof of Theorem 1.1

In this section we prove a more general version of Theorem 1.1, stated as Theorem 4.2 below, where the canonical height H is replaced with the twisted height H_A, as defined in section 2; since H is simply H_I with $I \in \mathrm{GL}_N(K_{\mathbb{A}})$ being the identity, Theorem 1.1 readily follows from Theorem 4.2. We start with a conventional twisted height version of Siegel's lemma.

Theorem 4.1. *Let K be either a number field, function field, or the algebraic closure of one or the other, and let $Z \subseteq K^N$ be an L-dimensional subspace, $1 \leq L < N$. Then for each $A \in \mathrm{GL}_N(K_{\mathbb{A}})$, there exists a basis $\boldsymbol{z}_1, ..., \boldsymbol{z}_L$ for Z over K such that*

$$\prod_{i=1}^{L} H_A(\boldsymbol{z}_i) \leq C_K(N, L) H_A(Z), \tag{31}$$

where all the notation is as in section 2.

Proof. When K is a number field, this is the Bombieri-Vaaler version of Siegel's lemma [1] with canonical height replaced by twisted height (see [15]); when K is a function field, this is proved in [12]; when K is the algebraic closure of a number

field or a function field, this follows from the Roy-Thunder twisted height version of absolute Siegel's lemma (see Theorem 8.1 of [5]). □

Remark 4.1. The constant $C_K(N,L)$ in Theorem 4.1 can be replaced by a slightly sharper one, leading to a slightly better constant in Theorem 1.1 (see [15], [6]); however, this would make the inequalities harder to read, and some of the constants that would appear in the upper bound would not be easily computable, for instance the generalized Hermite's constant.

Theorem 4.2. *Let (Z,F) be a regular $2k$-dimensional symplectic space in N variables over K, where $1 \leq k < 2k \leq N$. Then for each $A \in \mathrm{GL}_N(K_{\mathbb{A}})$, there exists a symplectic basis $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k,\boldsymbol{y}_1,\dots,\boldsymbol{y}_k$ for Z satisfying (2) such that*

$$\prod_{i=1}^{2k} H_A(\boldsymbol{x}_i)H_A(\boldsymbol{y}_i) \leq \left(C_K(N,2k)H_A(Z)\right)^{a_k} \left(\mathfrak{C}'(A)\mathcal{H}(F)\right)^{b_k}, \tag{32}$$

where $\mathfrak{C}'(A)$ is as in (16), and the rest of notation is as in the statement of Theorem 1.1. In particular, if $A = I$ is the identity element of $GL_N(K_{\mathbb{A}})$, $\mathfrak{C}'(A) = 1$.

Proof. Fix $A \in \mathrm{GL}_N(K_{\mathbb{A}})$, and let $\boldsymbol{z}_1,\dots,\boldsymbol{z}_{2k}$ be the basis for Z guaranteed by Theorem 4.1. We argue by induction on k. If $k = 1$, then $F(\boldsymbol{z}_1,\boldsymbol{z}_2) \neq 0$, since otherwise (Z,F) would be singular. Let $\boldsymbol{x}_1 = \frac{1}{F(\boldsymbol{z}_1,\boldsymbol{z}_2)}\boldsymbol{z}_1$, $\boldsymbol{y}_1 = \boldsymbol{z}_2$, then $F(\boldsymbol{x}_1,\boldsymbol{y}_1) = 1$, and $H_A(\boldsymbol{x}_1) = H_A(\boldsymbol{z}_1)$. The result follows from (31).

Now assume $k > 1$. We construct a graph $G(Z)$ on $2k$ vertices in the following way: for each $1 \leq i \leq 2k$, a vertex v_i will correspond to the vector $\boldsymbol{z}_i$, and two vertices v_i and v_j will be connected if and only if $F(\boldsymbol{z}_i,\boldsymbol{z}_j) = 0$. Since a Lagrangian of (Z,F) has dimension k, the corresponding graph $G(Z)$ satisfies the condition of Lemma 3.1, which implies that there exists a collection of distinct vectors

$$\{\boldsymbol{z}_{i_1},\boldsymbol{z}_{j_1},\dots,\boldsymbol{z}_{i_M},\boldsymbol{z}_{j_M}\} \subset \{\boldsymbol{z}_1,\dots,\boldsymbol{z}_{2k}\}, \tag{33}$$

where $M = \left[\frac{k+1}{2}\right]$, which satisfy the condition

$$F(\boldsymbol{z}_{i_n},\boldsymbol{z}_{j_n}) \neq 0, \quad \forall\ 1 \leq n \leq M = \left[\frac{k+1}{2}\right]. \tag{34}$$

We can assume without loss of generality that the ordering in (33) satisfies the condition

$$H_A(\boldsymbol{z}_{i_1})H_A(\boldsymbol{z}_{j_1}) \leq \cdots \leq H_A(\boldsymbol{z}_{i_M})H_A(\boldsymbol{z}_{j_M}). \tag{35}$$

Then, combining (35) and (31) we have:

(36)

$$\left(H_A(\boldsymbol{z}_{i_1})H_A(\boldsymbol{z}_{j_1})\right)^M \leq \prod_{n=1}^{M} H_A(\boldsymbol{z}_{i_n})H_A(\boldsymbol{z}_{j_n}) \leq \prod_{m=1}^{2k} H_A(\boldsymbol{z}_m) \leq C_K(N,2k)H_A(Z).$$

Let $\boldsymbol{x}_1 = \frac{1}{F(\boldsymbol{z}_{i_1},\boldsymbol{z}_{j_1})}\boldsymbol{z}_{i_1}$, $\boldsymbol{y}_1 = \boldsymbol{z}_{j_1}$, then $F(\boldsymbol{x}_1,\boldsymbol{y}_1) = 1$ and

$$H_A(\boldsymbol{x}_1)H_A(\boldsymbol{y}_1) \leq \left(C_K(N,2k)H_A(Z)\right)^{1/M}, \tag{37}$$

where $M = \left[\frac{k+1}{2}\right]$. Let

$$Z_1 = \mathrm{span}_K\{\boldsymbol{x}_1,\boldsymbol{y}_1\}^{\perp_F} \cap Z = \left\{\boldsymbol{z} \in \overline{K}^N : (\boldsymbol{x}_1\ \boldsymbol{y}_1)^t F\boldsymbol{z} = \boldsymbol{0}\right\} \cap Z,$$

then combining Lemmas 2.1, 2.4, and 2.5 with (17), (19), and (37), we obtain:

$$\begin{aligned} H_A(Z_1) &\leq |\det A|_{\mathbb{A}} H_{A^*}\left((\boldsymbol{x}_1\ \boldsymbol{y}_1)^t F\right) H_A(Z) \\ &\leq |\det A|_{\mathbb{A}} \mathfrak{C}(A^*)^2 H_{A^*}(\boldsymbol{x}_1) H_{A^*}(\boldsymbol{y}_1) \mathcal{H}(F)^2 H_A(Z) \\ &\leq C_K(N,2k)^{\frac{1}{M}} \left(\frac{\mathfrak{C}(A)|\det A|_{\mathbb{A}}^{1/2}}{\mathcal{C}_1(A)^2}\right)^2 H_A(Z)^{\frac{M+1}{M}} \mathcal{H}(F)^2. \end{aligned} \tag{38}$$

Moreover, notice that $\dim_K Z_1 = 2(k-1)$ and Z_1 is non-singular, since Z and $\mathrm{span}_K\{\boldsymbol{x}_1, \boldsymbol{y}_1\}$ are non-singular. By induction hypothesis, there exists a symplectic basis $\boldsymbol{x}_2, \dots, \boldsymbol{x}_k, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k$ for Z_1 so that

$$F(\boldsymbol{x}_i, \boldsymbol{x}_j) = F(\boldsymbol{y}_i, \boldsymbol{y}_j) = F(\boldsymbol{x}_i, \boldsymbol{y}_j) = 0\ \forall\ 2 \leq i \neq j \leq k,\ F(\boldsymbol{x}_i, \boldsymbol{y}_i) = 1\ \forall\ 2 \leq i \leq k,$$

and

$$\prod_{i=2}^{k} H_A(\boldsymbol{x}_i) H_A(\boldsymbol{y}_i) \leq (C_K(N, 2(k-1)) H_A(Z_1))^{a_{k-1}} \left(\mathfrak{C}'(A)\mathcal{H}(F)\right)^{b_{k-1}}. \tag{39}$$

Combining (37), (38), and (39), and using the fact that $C_K(N, L_1) \leq C_K(N, L_2)$ whenever $L_1 \leq L_2$, we obtain:

$$\prod_{i=1}^{k} H_A(\boldsymbol{x}_i) H_A(\boldsymbol{y}_i) \leq (C_K(N,2k) H_A(Z))^{\frac{(M+1)a_{k-1}+1}{M}} \left(\mathfrak{C}'(A)\mathcal{H}(F)\right)^{b_{k-1}+2a_{k-1}}. \tag{40}$$

The result now follows by a routine calculation. □

Remark 4.2. Clearly, versions of Corollary 1.2 and Corollary 1.3 with the twisted height H_A instead of the canonical height H follow immediately from Theorem 4.2.

References

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

[2] J. W. S. Cassels. Bounds for the least solutions of homogeneous quadratic equations. *Proc. Cambridge Philos. Soc.*, 51:262–264, 1955.

[3] L. Fukshansky. Small zeros of quadratic forms over $\overline{\mathbf{Q}}$. *to appear in Int. J. Number Theory, arxiv:math.NT/0512132.*

[4] L. Fukshansky. On effective Witt decomposition and Cartan-Dieudonné theorem. *Canad. J. Math.*, 59(6):1284–1300, 2007.

[5] D. Roy and J. L. Thunder. An absolute Siegel's lemma. *J. Reine Angew. Math.*, 476:1–26, 1996.

[6] D. Roy and J. L. Thunder. Addendum and erratum to: An absolute Siegel's lemma [J. Reine Angew. Math. 476 (1996), 1–26; MR1401695 (97h:11075)]. *J. Reine Angew. Math.*, 508:47–51, 1999.

[7] W. Scharlau. *Quadratic and Hermitian Forms.* Springer-Verlag, 1985.

[8] H. P. Schlickewei. Kleine nullstellen homogener quadratischer gleichungen. *Monatsh. Math.*, 100(1):35–45, 1985.

[9] H. P. Schlickewei and W. M. Schmidt. Quadratic geometry of numbers. *Trans. Amer. Math. Soc.*, 301(2):679–690, 1987.

[10] T. Struppeck and J. D. Vaaler. Inequalities for heights of algebraic subspaces and the Thue-Siegel principle. *Analytic number theory (Allerton Park, IL, 1989), Progr. Math.*, 85:493–528, 1990.

[11] J. L. Thunder. Asymptotic estimates for rational points of bounded height on flag varieties. *Comp. Math.*, 88(2):155–186, 1993.

[12] J. L. Thunder. Siegel's lemma for function fields. *Michigan Math. J.*, 42(1):147–162, 1995.

[13] J. D. Vaaler. Small zeros of quadratic forms over number fields. *Trans. Amer. Math. Soc.*, 302(1):281–296, 1987.

[14] J. D. Vaaler. Small zeros of quadratic forms over number fields, II. *Trans. Amer. Math. Soc.*, 313(2):671–686, 1989.
[15] J. D. Vaaler. The best constant in Siegel's lemma. *Monatsh. Math.*, 140(1):71–89, 2003.
[16] A. Weil. *Basic Number Theory*. Springer-Verlag, 1973.

Department of Mathematics, 850 Columbia Avenue, Claremont McKenna College, Claremont, CA 91711
E-mail address: lenny@cmc.edu