Summer 2023

Lattice Extensions and Zeros of Multilinear Polynomials

Maxwell Forst
Claremont Graduate University

Lattice Extensions and Zeros of Multilinear Polynomials

BY

Maxwell Forst

Claremont Graduate University
2023

Approval of the Dissertation Commitee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Maxwell Forst as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Mathematics.

Lenny Fukshansky, Chair
Claremont McKenna College
Professor of Mathematics

Michael Orrison
Harvey Mudd College
Professor of Mathematics

Allon Percus
Claremont Graduate University
Professor of Mathematics

Jeffrey D. Vaaler
University of Texas at Austin
Professor Emeritus

ABSTRACT

Lattice Extensions and Zeros of Multilinear Polynomials

by

Maxwell Forst

Claremont Graduate University: 2023

We treat several problems related to the existence of lattice extensions preserving certain geometric properties and small-height zeros of various multilinear polynomials. An extension of a Euclidean lattice L_1 is a lattice L_2 of higher rank containing L_1 so that the intersection of L_2 with the subspace spanned by L_1 is equal to L_1. Our first result provides a counting estimate on the number of ways a primitive collection of vectors in a lattice can be extended to a basis for this lattice. Next, we discuss the existence of lattice extensions with controlled determinant, successive minima and covering radius. In the two-dimensional case, we also present some observations about the deep holes of a lattice as elements of the quotient torus group. Looking for basis extensions additionally connects to a search for small-height zeros of multilinear polynomials, for which we obtain several results over arbitrary number fields. These include bounds for a system of polynomials under appropriate hypotheses, as well as for a single polynomial with some additional avoidance conditions. In addition to several height inequalities that we need for these bounds, we obtain a new absolute version of Siegel's lemma which is proved using only linear algebra tools.

ACKNOWLEDGEMENTS

I would like to thank my advisor Lenny Fukshansky for all of his time, patience and dedication during my time at the Claremont Graduate University and for helping me find my field of interest. Undeniably, I would not be where I am without his help.

I would also like to thank the many great professors in the Claremont colleges for there many great classes they offer. In particular I would like to thank Allon Percus and Marina Chugunova for some of the most challenging classes I have ever taken.

I would also like to thank my parents and my grandmother for their love and support over the years and for helping me chase my ambitions.

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS v

I. Introduction 1

1.1 Lattice extensions 2
1.2 Planar Lattices 10
1.3 Height Functions 12
1.4 A New Version of Siegel's Lemma 18
1.5 Multilinear Forms 22

II. On Lattice Extensions 26

2.1 Small Determinant Extensions 27
2.2 Counting Lattice extensions 30
2.3 Successive minima extensions 40

III. On Planar Lattices 45

3.1 Farey Fractions 46
3.2 Deep Holes of Planar Lattices 52
3.3 Covering radius of planar lattices 57

IV. On Heights and Siegel's Lemma 66

4.1 Additional Properties of Heights 67
4.2 Proof of Theorem 1.4.1 69
4.3 Proof of Theorem 1.4.2 78

V. On Multilinear Forms 82

5.1 Height Inequalities 83
5.2 Zeros of multiple polynomials 86

5.3 Zeros of one polynomial 89
5.4 Zeros of multilinear forms 92

Chapter I

Introduction

1.1 Lattice extensions

Let V be a real m-dimensional vector space and let $r > 0$, define $B_V(r)$ to be the open ball of radius r in V centered at $\mathbf{0}$. Write ω_m for the volume of the unit ball $B_V(1)$ so that $\text{vol}(B_V(r)) = \omega_m r^m$.

A lattice Λ is a subgroup of V that satisfies two geometric properties: Λ is **discrete** in that there exists $\epsilon > 0$ so that for any $\boldsymbol{x}, \boldsymbol{y} \in \Lambda, \boldsymbol{x} \neq \boldsymbol{y}$

$$(\boldsymbol{x} + B_V(\epsilon)) \cap (\boldsymbol{y} + B_V(\epsilon)) = \emptyset;$$

and Λ is **co-compact** in that there exists a compact set $C \subset V$ with $\text{vol}(C) > 0$ so that

$$\bigcup_{\boldsymbol{x} \in \Lambda} (\boldsymbol{x} + C) = V.$$

Lattices can be define in terms of a **basis** of linearly independent vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ where

$$\begin{aligned} \Lambda &= \text{span}_{\mathbb{Z}}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \\ &= \{a_1\boldsymbol{a}_1 + \ldots + a_m\boldsymbol{a}_m : a_1, \ldots, a_m \in \mathbb{Z}\}. \end{aligned}$$

Equivalently, Λ can be defined in terms of a basis matrix:

$$\Lambda = A\mathbb{Z}^m$$

where A is the full rank $n \times m$ matrix

$$A = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_m).$$

With this we can define the **determinant** of Λ as

$$\det \Lambda = \sqrt{|\det(A^T A)|}.$$

Two $n \times m$ basis matrices A, B define the same lattice if and only if $A = BU$ for some $U \in \mathrm{GL}_m(\mathbb{Z})$. For this reason the determinant of Λ does not depend on a particular choice of basis.

We also define two equivalence relations on lattices. Two lattices L and L' in $\mathbb{R}^n$ are **isometric** if there exists an $n \times n$ orthogonal matrix U so that $L = UL'$. Likewise, two lattices are said to be **similar** if there exists orthogonal matrix U and non-zero scalar s so that $L = sUL'$.

We will now recall several definitions from the geometry of numbers (see [37] for a more detailed exposition). A given set $F \subset \mathrm{span}_{\mathbb{R}}(\Lambda)$ is said to be a **fundamental domain** of Λ if

$$\mathrm{span}_{\mathbb{R}}(\Lambda) = \bigcup_{\boldsymbol{x} \in \Lambda} \boldsymbol{x} + F,$$

and $(\boldsymbol{x} + F) \cap (\boldsymbol{y} + F) = \emptyset$ for all $\boldsymbol{x}, \boldsymbol{y} \in \Lambda, \boldsymbol{x} \neq \boldsymbol{y}$. If F is a fundamental domain of Λ, then the volume of F in $\mathrm{span}_{\mathbb{R}}(\Lambda)$ equals $\det(\Lambda)$. Thus $\det(\Lambda)$ is also referred to as the **co-volume** of Λ in $\mathrm{span}_{\mathbb{R}}(\Lambda)$. Often the most useful fundamental domains are fundamental parallelepipeds. If $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ is a basis for Λ where $m = \operatorname{rank} \Lambda$, we can define a **fundamental parallelepiped** as

$$\mathcal{P} = \left\{ \sum_{i=1}^{m} a_i \boldsymbol{x}_i : a_i \in [0, 1) \ \forall i \right\}.$$

Let Λ be a lattice of rank m, then the **successive minima** of Λ are ordered positive real numbers

$$0 \leq \lambda_1(\Lambda) \leq \ldots \leq \lambda_m(\Lambda),$$

such that

$$\lambda_i(\Lambda) := \inf\{r \in \mathbb{R} : \dim_{\mathbb{R}} \operatorname{span}_{\mathbb{R}}(\Lambda \cap B_V(r)) \geq i\} \quad 1 \leq i \leq m.$$

A vector $\boldsymbol{x} \in \Lambda$ is said to correspond to the i-th successive minima if $\|\boldsymbol{x}\| = \lambda_i(\Lambda)$. In this section, $\|\cdot\|$ refers to the standard Euclidean norm on $\mathbb{R}^n$ and $|\cdot|$ refers to the sup-norm (maximum of the absolute values of the coordinates). The product of the successive minima is bounded by Minkowski's successive minima theorem:

$$\frac{2^m}{m!\det(\Lambda)} \leq \omega_m \prod_{i=1}^{m} \lambda_i(\Lambda) \leq \frac{2^n}{\det(\Lambda)}.$$

The **packing radius** of Λ in a vector space V is defined as

$$\sup\left\{r \in \mathbb{R}_{\geq 0} : \boldsymbol{x} + B_V(r) \bigcap \boldsymbol{y} + B_V(r) = \emptyset\right\},$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \Lambda, \boldsymbol{x} \neq \boldsymbol{y}$. It is easy to see that the packing radius of Λ is equal to $\frac{1}{2}\lambda_1(\Lambda)$.

The **covering radius**, also called the **inhomogeneous minimum**, of a full rank lattice Λ in a vector space V is defined as

$$\mu(\Lambda) = \inf\left\{r \in \mathbb{R} : \bigcup_{\boldsymbol{x} \in \Lambda} (\boldsymbol{x} + B_V(r)) = V\right\}.$$

There is a classical inequality of Jarnik that asserts

$$\mu(\Lambda) \leq \frac{1}{2}\sum_{i=1}^{m} \lambda_i(\Lambda).$$

Now, let L_1, L_2 be lattices in $\mathbb{R}^n$ of rank m_1, m_2 respectively so that $L_1 \subset L_2$ and

$1 \leq m_1 < m_2 \leq n$. L_2 is said to be an **extension** of L_1 if

$$L_2 \cap \operatorname{span}_{\mathbb{R}} L_1 = L_1.$$

This is equivalent to saying that the quotient

$$L_2/L_1 \cong \mathbb{Z}^{m_2-m_1},$$

i.e. L_2/L_1 is torsion free. Alternatively, L_2 is an extension of L_1 if there exists a basis of L_2, $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m_1}, \ldots \boldsymbol{a}_{m_2}\}$, so that $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m_1}\}$ is a basis for L_1. If L_2 is an extension of L_1, then L_1 is said to be **extendable** to L_2. If L_1 is extendable to L_2 and $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m_1}\}$ is a basis for L_1 then the collection of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{m_1}$ is said to be **primitive in** L_2. In the case that $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ is primitive in $\mathbb{Z}^n$ we will simply say the collection is **primitive**.

Lattice extensions have been implicitly used in a variety of contexts such as in the construction of laminated lattices (see [16], [47]) which are studied in the context of lattice packings and coverings, and in the construction of Minkowski or HKZ reduced bases (see [37], [47]). However, we are unaware of any explicit study of lattice extensions similar to the approach we take in this paper.

Given a lattice Ω in $\mathbb{R}^n$ of rank $m < n$ with basis $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$ we are interested in constructing a lattice Λ as an extension of Ω, so that one of the above geometric invariants of Λ is controlled in relation to geometric invariants of Ω. We do this by selecting vectors $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d \in \mathbb{R}^n$ for $1 \leq d \leq n - m$ so that the augmented set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_d\}$ is a basis for Λ. Depending on the particular property we are interested in, we may require $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d$ to be elements of some fixed full rank ambient lattice.

Let $\Omega \subset \mathbb{Z}^n$ be a lattice of rank $m \leq n$ so that Ω is extendable to $\mathbb{Z}^n$. The first of our results asymptotically estimates the number of ways to extend a primitive

collection of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ with $1 \leq m < n$ to a basis of $\mathbb{Z}^n$ by selecting vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_d$, $1 \leq d \leq n-m$ so that $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_d$ is again primitive and so that each $|\boldsymbol{b}_i|$ is bounded for $1 \leq i \leq d$.

Theorem 1.1.1. *Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{Z}^n$ be a primitive collection of vectors.*

1. *if $m < n-1$, the number of vectors $\boldsymbol{b} \in \mathbb{Z}^n$ with $|\boldsymbol{b}| \leq T$ such that the collection $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}$ is again primitive is equal to $\Theta(T^n)$ as $T \to \infty$.*

2. *if $m = n-1$, the number of vectors $\boldsymbol{b} \in \mathbb{Z}$ with $|\boldsymbol{b}| \leq T$ such that the collection $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}$ is a basis for $\mathbb{Z}^n$ is equal to $\Theta(T^{n-1})$ as $T \to \infty$.*

As a result, for any $1 \leq k < n-m$ there exist $\Theta(T^{nk})$ collections of vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k \in \mathbb{Z}^n$ with $|\boldsymbol{b}_j| \leq T, 1 \leq j \leq k$, such that $\{\boldsymbol{a}_i, \boldsymbol{b}_j : 1 \leq i \leq m, 1 \leq j \leq k\}$ is again primitive. Further, there are $\Theta(T^{n^2-nm-1})$ such collections $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-m}\}$ so that

$$\mathbb{Z}^n = span_{\mathbb{Z}}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-m}\}.$$

The constants in the Θ-notation depend on the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$, n and m.

More generally, since any full rank lattice Λ in $\mathbb{R}^n$ is of the form $\Lambda = U\mathbb{Z}^n$ for some full rank $n \times n$ matrix U and therefore bases of Λ are in bijective correspondence with bases of $\mathbb{Z}^n$, we can generalize Theorem 1.1.1 to arbitrary lattices.

Corollary 1.1.2. *Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ be a primitive collection of vectors in a full-rank lattice $\Lambda \subset \mathbb{R}^n$ with $1 \leq m \leq n$. Then there exist*

$$\Theta\left(T^{n+min\{0,n-m-2\}}\right)$$

vectors $\boldsymbol{b} \in \Lambda$ with $|\boldsymbol{b}| \leq T$ so that the collection $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}$ is again primitive in Λ, and hence for any $1 \leq k \leq n-m$ there exist $\Theta(T^{nk})$ collection of vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k \in \Lambda$ with $|\boldsymbol{b}_i| \leq T, 1 \leq i \leq k$, such that $\{\boldsymbol{a}_i, \boldsymbol{b}_j : 1 \leq i \leq m, 1 \leq j \leq k\}$ is

again primitive. Further, there are $\Theta(T^{n^2-nm-1})$ *such collections* $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-m}\}$ *so that*

$$\Lambda = span_{\mathbb{Z}}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-m}\}.$$

The constants in the Θ*-notation depend on the lattice* Λ *and the vectors* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$*,* n *and* m*.*

Theorem 1.1.1 and Corollary 1.1.2 will be proved in Section 2.2. There we will also discuss the Θ constants in Theorem 1.1.1 and in Lemma 2.2.1.

Theorem 1.1.1 can also be interpreted as a statement on unimodular matrices. An $n \times m$ integer matrix A is said to be **unimodular** if there exists an $n \times n - m$ integer matrix B so that the augmented matrix $(A\ B) \in \mathrm{GL}_n(\mathbb{Z})$. Alternatively A is unimodular if the columns of A are a primitive collection of vectors. Thus Theorem 1.1.1 estimates the number of ways to augment a unimodular matrix A by a bounded integer matrix B so that $(A\ B)$ is again unimodular.

The next result deals with the construction of an integer lattice extension with small determinant.

Theorem 1.1.3. *Let* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ *be linearly independent vectors in* $\mathbb{Z}^n$ *and let*

$$\Omega = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subset \mathbb{Z}^n$$

be the sublattice of rank m *spanned by these vectors. Then there exists an extension* Ω' *of* Ω *in* $\mathbb{Z}^n$ *so that*

$$\det \Omega' = \gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m).$$

Further, if $m = n - 1$ *then there exists* $\boldsymbol{y} \in \mathbb{Z}^n$ *so that* $\Omega' = \mathrm{span}_{\mathbb{Z}}\{\Omega, \boldsymbol{y}\}$ *and*

$$\|\boldsymbol{y}\| \le \left\{ \left(\frac{\gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)}{\det \Omega} \right)^2 + \mu^2 \right\}^{1/2},$$

where μ is the covering radius of Ω and $\wedge$ is the Grassmann wedge product (see section 1.3)

We will prove this theorem in Section 2.1 and we will also explain how it can be generalized to sublattices of an arbitrary full rank lattice Λ in Remark 2.1.1.

Let Ω be a lattice of rank m in $\mathbb{R}^n$ with $m \leq n$ and let Λ be an extension of Ω. Λ is said to be a **successive minima extension of** Ω if the first m successive minima of Λ agree with the successive minima of Ω, that is

$$\lambda_i(\Omega) = \lambda_i(\Lambda) \quad \forall 1 \leq i \leq m.$$

In general it is easy to construct a successive minima extension of Ω in $\mathbb{R}^n$. Simply choose a vector $\boldsymbol{u} \in \mathbb{R}^n$ so that $\boldsymbol{u}$ is orthogonal to Ω with $\|\boldsymbol{u}\| \geq \lambda_m(\Omega)$ so that $\text{span}_{\mathbb{Z}}\{\Omega, \boldsymbol{u}\}$ is a successive minima extension. A more interesting case is when $\boldsymbol{u}$ is drawn from a fixed ambient lattice. The next theorem proves the existence of such successive minima extensions that are bounded in terms of the base lattice and the ambient lattice.

Theorem 1.1.4. *Let $\Lambda \subset \mathbb{R}^n$ be a lattice of full rank, and let $L_k \subset \Lambda$ be a sublattice of rank $1 \leq k < n$. There exists a sublattice $L_{k+1} \subset \Lambda$ of rank $k+1$ such that $L_k \subset L_{k+1}$, $\lambda_j(L_{k+1}) = \lambda_j(L_k)$ for all $1 \leq j \leq k$ and*

$$\lambda_{k+1}(L_{k+1}) \leq \frac{\lambda_k(L_k)(v_*^2 + \sqrt{1-v_*^2})}{\sqrt{1-v_*^4}} + 2\mu, \tag{1.1}$$

where μ is the covering radius of Λ and v_* is the smallest root of the polynomial

$$p(v) = \left(\frac{\mu^2}{\lambda_k^2}(1-v^4) - v^2(v^4 - v^2 + 1)\right)^2 - \left(\frac{2\mu^2}{\lambda_k^2}v(1-v^4) + 2v^4\right)^2(1-v^2)$$

in the interval $(0,1)$: such v_* necessarily exists.

Theorem 1.1.4 as well as Corollary 1.1.2 will be proved in Section 2.3.

The results of this section as well as Chapter II are based on joint work with L. Fukshansky and were originally presented in [19] and [20].

1.2 Planar Lattices

This section presents lattice extension results that are specific to planar lattices or only partially generalize to higher dimensions.

Let $\boldsymbol{e}_1$ denote the vector $(1,0)^T$ in $\mathbb{R}^2$. Let E_1 denote the lattice $\boldsymbol{e}_1\mathbb{Z}$. Planar lattice extensions of E_1 can then be constructed by selecting a vector $\boldsymbol{v} \in \mathbb{R}^2$ so that $\boldsymbol{e}_1, \boldsymbol{v}$ are linearly independent. Define such a lattice extension as $E_1(\boldsymbol{v}) = \text{span}_{\mathbb{Z}}\{\boldsymbol{e}_1, \boldsymbol{v}\}$. Define the set

$$F = \{(a,b) \in \mathbb{R}^2 : 0 \leq a < 1/2, b > 0, a^2 + b^2 \geq 1\}. \tag{1.2}$$

If $\boldsymbol{v} \in F$ then $\boldsymbol{e}_1, \boldsymbol{v}$ are vectors corresponding to the first and second successive minima of $E(\boldsymbol{v})$ respectively. Moreover, if $\{\boldsymbol{e}_1, \boldsymbol{v}\}$ is a basis for Λ with $\boldsymbol{v} \in F$, then $\{\boldsymbol{e}, \boldsymbol{v}\}$ is a minimal basis for Λ, that is

$$\lambda_1(\Lambda) = \|\boldsymbol{e}_1\| = 1, \quad \lambda_2(\Lambda) = \|\boldsymbol{v}\|. \tag{1.3}$$

The set F parameterizes all planar lattices up to similarity class. There is a particular advantage to working with planar lattices in that a pair of minimal vectors corresponding to the first and second successive minima are always guaranteed to be a basis. Moreover, these minimal vectors can always be chosen so that the angle between minimal vectors is in $(\pi/3, \pi/2]$.

Planar lattices also have connections to other branches of mathematics, for instance planar lattices have a natural correspondence with elliptic curves (see [58]). Another such connection is to Farey fractions which in turn, has a natural connection to our results in Theorem 1.1.1. We will discuss this connection in detail in Section 3.1.

Let $\Lambda \subset \mathbb{R}^n$ be a rank $m + k$ lattice extension of the rank m lattice Ω with

$1 \leq m < m+k \leq n$. Λ is said to be an **equal covering extension** of Ω if $\mu(\Lambda) = \mu(\Omega)$. In dimension 2 we can explicitly construct all equal covering extensions of E_1.

Theorem 1.2.1. *A lattice $\Lambda \subset \mathbb{R}^2$ is an equal covering extension of E_1 if and only if*

$$\Lambda = \Lambda(\alpha) := \begin{pmatrix} \alpha & \alpha - 1 \\ \sqrt{\alpha - \alpha^2} & \sqrt{\alpha - \alpha^2} \end{pmatrix} \mathbb{Z}^2 \tag{1.4}$$

for some real number $0 < \alpha < 1$. More generally, a lattice $\Lambda \subset \mathbb{R}^n$ of rank 2 is an equal covering extension of a rank-one lattice $L \subset \Lambda$ if and only if it is isometric to some lattice of the form $\det(L)\Lambda(\alpha)$, where $\Lambda(\alpha)$ is as in (1.4).

This theorem partially generalizes to higher dimensions:

Theorem 1.2.2. *Let $\Lambda_k \subset \mathbb{R}^n$ be an orthogonal lattice of rank $k < n$. There exists an orthogonal lattice $\Lambda_{k+1} \subset \mathbb{R}^n$ of rank $k+1$ so that $\Lambda_k \subset \Lambda_{k+1}$ and $\mu(\Lambda_{k+1}) = \mu(\Lambda_k)$. Further, if $\boldsymbol{z}$ is a deep hole of Λ_k it is also a deep hole of Λ_{k+1}.*

The construction used in Theorem 1.2.2 mirrors the construction of laminated lattices (see [16]) and thus relies on the structure of the deep holes of Λ_k. Let Λ be a lattice in $\mathbb{R}^n$, a vector $\boldsymbol{z} \in \mathbb{R}^n$ is said to be a **deep hole** of Λ if

$$\min_{\boldsymbol{x} \in \Lambda} \|\boldsymbol{z} - \boldsymbol{x}\| = \mu(\Lambda).$$

Our examination of deep holes led to a particularly interesting result that we report in Section 3.2.

Proofs of Theorems 1.2.1 and 1.2.2 will be presented in Chapter III. The results of this section as well as Chapter III are based on joint work with L. Fukshansky and were originally presented in [19] and [20].

1.3 Height Functions

For the next two sections as well as Chapters IV and V we will need to introduce the notion of height functions. Additional properties of heights specific to Chapters IV and V will be presented in Sections 4.1 and 5.1 respectively. Additionally, the Grassmann wedge product which we also introduce here is used in Chapters II and III.

For integers $1 \leq d \leq n$, let $[n] := \{1, \ldots, n\}$ and define

$$\mathcal{J}(n,d) := \{I \subseteq [n] : |I| = d\},$$

to be the set of size d subsets of $[n]$, so that $|\mathcal{J}(n,d)| = \binom{n}{d}$. For an $n \times d$ matrix A with coefficients in a number field K, we define $A_I, I \subseteq [n]$ to be the $|I| \times d$ submatrix of A so that the rows of A_I are the rows of A indexed by I. When A is a full rank $n \times d$ matrix we define the **Grassmann/Plücker coordinates** of A with

$$(\det(A_I))_{I \in \mathcal{J}(n,d)},$$

the vector of the determinants of the $d \times d$ minors of A. Likewise, if B is a $d \times n$ matrix with coefficients in K we define $B^I, I \subseteq [n]$ to be $d \times |I|$ submatrix of B so that the columns of B^I are the columns of B indexed by I, and if B is full rank, define the Grassmann coordinates of B to be the vector

$$\left(\det(B^I)\right)_{I \in \mathcal{J}(n,d)}.$$

For a set of column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_d \in K^n$ we identify the **Grassmann/Exterior wedge product**

$$\boldsymbol{a}_1 \wedge \ldots \wedge \boldsymbol{a}_d$$

with the Grassmann coordinates of the $n \times d$ matrix

$$A = (\boldsymbol{a}_1 \dots \boldsymbol{a}_d).$$

Remark 1.3.1. Throughout this paper we will largely be unconcerned with the sign and order of the Grassmann coordinates. However, for the sake of concreteness we will order each $I \in \mathcal{J}(n,d)$ in ascending order and order $\mathcal{J}(n,d)$ in lexicographic order.

When U is an $n \times m$ matrix with rational integer coefficients we will define $\gcd(U)$ to be the greatest common divisor of the entries of U. Likewise for $\boldsymbol{x} \in \mathbb{Z}^n$ we will define $\gcd(\boldsymbol{x})$ as the g.c.d. of the coordinates of $\boldsymbol{x}$.

One tool that we will need is the Brill-Gordan duality principle (see [36], as well as Theorem 1 on p. 294 of [41]; see also proof of Theorem 4.3 of [26], as well as [40], [6] for more contemporary accounts of this principle).

Lemma 1.3.1. *(Duality Principle). Let $1 \leq m < n$, and let A, B be respectively $n \times m$ and $(n-m) \times n$ integer matrices such that*

$$A\mathbb{Z}^m = \{\boldsymbol{x} \in \mathbb{Z}^n : B\boldsymbol{x} = 0\}.$$

Let $I \in \mathcal{J}(n,d)$ and let $I' = [n] \setminus I$ Then

$$\det A_I = (-1)^{i_1 + \dots + i_m} \gamma \det B^{I'} \tag{1.5}$$

for an appropriate constant $\gamma \in \mathbb{Q}$, where $i_1, \dots, i_m, i_{m+1}, \dots, i_n = 1, \dots, n$. If column vectors of A and row vectors of B can be extended to a basis of $\mathbb{Z}^n$, then $\gamma = 1$.

Now let K be a number field of degree $d := [K : \mathbb{Q}] \geq 1$ and let r_1, r_2 be the numbers of real and conjugate pairs of complex embeddings of K, respectively, so

that

$$d = r_1 + 2r_2.$$

Let Δ_K be the discriminant of K, and write $M(K)$ for the set of places of K. For each $v \in M(K)$ let $d_v = [K_v : \mathbb{Q}_v]$ be the local degree. Then for each $u \in M(\mathbb{Q})$, $\sum_{v|u} d_v = d$. We normalize the absolute value at each place so that for each nonzero $x \in K$ we have the product formula:

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

Let $n \geq 2$, and for any place $v \in M(K)$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in K^n$ define the corresponding sup-norm

$$|\boldsymbol{x}|_v = \max\{|x_1|_v, \ldots, |x_n|_v\}.$$

If $v \mid \infty$, we also define the Euclidean norm

$$\|\boldsymbol{x}\|_v = \left(\sum_{i=1}^{n} |x_i|_v^2 \right)^{1/2}.$$

For more on absolute values see [56].

Remark 1.3.2. In Chapters II and III we will be working over the real numbers rather than a number field, as such the only norms we will use are $|\cdot|_\infty$ and $\|\cdot\|_\infty$. For clarity of notation in these chapters we will write $|\cdot|$ for $|\cdot|_\infty$ and $\|\cdot\|$ for $\|\cdot\|_\infty$.

We then define two projective height functions $H, \mathcal{H} : K^n \to \mathbb{R}_{\geq 0}$ as follows:

$$H(\boldsymbol{x}) = \left(\prod_{v \in M(K)} |\boldsymbol{x}|_v^{d_v} \right)^{1/d}, \quad \mathcal{H}(\boldsymbol{x}) = \left(\prod_{v \nmid \infty} |\boldsymbol{x}|_v^{d_v} \times \prod_{v \mid \infty} \|\boldsymbol{x}\|_v^{d_v} \right)^{1/d}.$$

These heights are absolute, meaning that they are the same when computed over

any number field K containing the coordinates of $\boldsymbol{x}$. This is due to the normalizing exponent $1/d$ in the definition. Further, for each nonzero $\boldsymbol{x} \in K^n$,

$$1 \leq H(\boldsymbol{x}) \leq \mathcal{H}(\boldsymbol{x}) \leq \sqrt{n} H(\boldsymbol{x}). \tag{1.6}$$

We also define the inhomogeneous or Weil height $h : K^n \to \mathbb{R}_{\geq 1}$ to be

$$h(\boldsymbol{x}) = H(1, \boldsymbol{x}) \geq H(\boldsymbol{x})$$

for every $n \geq 1$, thus including a height on algebraic numbers. We also write $H(A)$ and $h(A)$ for the projective and inhomogeneous heights, respectively, of an $m \times n$ matrix A viewed as a vector in K^{mn} and $H(F)$, $h(F)$ for the respective heights of the coefficient vector of a polynomial F over K.

We also define the **Schmidt/Arakelov height** on the subspaces of K^n as follows: Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in K^n, m \leq n$ and let $V = \operatorname{span}_K\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ be the m-dimensional subspace of K^n spanned by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$. The wedge product of these basis vectors $\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m$ can be viewed as a vector in $K^{\binom{n}{m}}$ under lexicographic embedding: this is a vector of the Grassmann coordinates of V. Define

$$\mathcal{H}(V) = \mathcal{H}(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m).$$

The product formula along with the square-free property of the Grassmann wedge product guarantees that this definition is independent of the choice of a basis for V. We also define the Arakelov height on matrices as follows: for an $n \times m$ matrix A over K, $1 \leq m \leq n$ we let

$$\mathcal{H}(A) = \mathcal{H}\left(\operatorname{span}_K\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}\right),$$

where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ are the columns of A. If $m > n$, we define $\mathcal{H}(A)$ to be $\mathcal{H}(A^\top)$.

Suppose the m-dimensional vector subspace $V \subset K^n$ is described as

$$V = \{A\boldsymbol{x} : \boldsymbol{x} \in K^m\} = \{\boldsymbol{y} \in K^n : B\boldsymbol{y} = \boldsymbol{0}\},$$

for the $n \times m$ matrix A and $(n-m) \times n$ matrix B over K, respectively. Then Lemma 1.3.1 states that

$$\mathcal{H}(A) = \mathcal{H}(B) = \mathcal{H}(V). \tag{1.7}$$

We also review here several additional properties of height functions: If $1 \leq J < L$ then a generalization of Hadamard's determinant inequality asserts that

$$\begin{aligned} &\mathcal{H}\big(\boldsymbol{\xi}_1 \wedge\ \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_L\big) \\ &\qquad \leq \mathcal{H}\big(\boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_J\big)\mathcal{H}\big(\boldsymbol{\xi}_{J+1} \wedge \boldsymbol{\xi}_{J+2} \wedge \cdots \wedge \boldsymbol{\xi}_L\big). \end{aligned} \tag{1.8}$$

Alternatively, let Y and Z be the $N \times J$ and $N \times (L-J)$ matrices

$$Y = \big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_J\big), \quad \text{and} \quad Z = \big(\boldsymbol{\xi}_{J+1}\ \boldsymbol{\xi}_{J+2}\ \cdots\ \boldsymbol{\xi}_L\big).$$

Then (1.8) is also the inequality

$$\mathcal{H}(X) = \mathcal{H}(Y\ Z) \leq \mathcal{H}(Y)\mathcal{H}(Z) \tag{1.9}$$

for a partitioned matrix. By repeated application of this inequality we get

$$\mathcal{H}(\mathcal{Z}) = \mathcal{H}(X) \leq \prod_{\ell=1}^{L} \mathcal{H}(\boldsymbol{\xi}_\ell). \tag{1.10}$$

In the special case $L = N$ we have

$$\det\big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_N\big) = \boldsymbol{\xi}_1 \wedge \boldsymbol{\xi}_2 \wedge \cdots \wedge \boldsymbol{\xi}_N$$

In this case (1.10) becomes

$$\big|\det\big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_N\big)\big| \le \prod_{n=1}^{N} \mathcal{H}(\boldsymbol{\xi}_n), \tag{1.11}$$

which is Hadamard's upper bound for a determinant. As the origin of (1.8) and (1.9) is somewhat obscure, we will follow the terminology used in [18] and refer to all of these as *Hadamard's inequality*.

There is another type of inequality that is satisfied by the Arakelov height on subspaces. Let $\mathcal{Z} \subseteq K^N$ and $\mathcal{Y} \subseteq K^N$ be K-linear subspaces, and let $\langle \mathcal{Z}, \mathcal{Y} \rangle$ be the subspace spanned over K by $\mathcal{Z} \cup \mathcal{Y}$. Of course $\mathcal{Z} \cap \mathcal{Y}$ is also a subspace of K^N. Then these four subspaces satisfy the inequality

$$\mathcal{H}\big(\langle \mathcal{Z}, \mathcal{Y} \rangle\big)\mathcal{H}(\mathcal{Z} \cap \mathcal{Y}) \le \mathcal{H}(\mathcal{Z})\mathcal{H}(\mathcal{Y}). \tag{1.12}$$

This was proved in [59, Theorem 1], and it was proved independently and at about the same time by W. M. Schmidt. An immediate consequence of (1.12) is the inequality

$$\mathcal{H}(\mathcal{Z} \cap \mathcal{Y}) \le \mathcal{H}(\mathcal{Z})\mathcal{H}(\mathcal{Y}). \tag{1.13}$$

1.4 A New Version of Siegel's Lemma

The name Siegel's lemma is usually attributed to results on small-size nontrivial solutions to under-determined systems of homogeneous linear equations over a field or a ring of arithmetic interest. The original Siegel's lemma asserts that given an integer $M \times N$ matrix A, $M < N$, there exists a nonzero $\boldsymbol{x} \in \mathbb{Z}^N$ such that $A\boldsymbol{x} = \mathbf{0}$ and

$$|\boldsymbol{x}| \leq 2 + (N|A|)^{\frac{1}{N-M}}, \tag{1.14}$$

where $|A|$ here denotes the sup-norm of the matrix A. This result was originally used in transcendental number theory by Thue [60] and Siegel [57]. The exponent $\frac{1}{N-M}$ in (1.14) is known to be the best possible, however this bound lacks invariance under linear transformations: indeed, for any $M \times M$ integer matrix U,

$$(UA)\boldsymbol{x} = A\boldsymbol{x} = \mathbf{0},$$

however $|UA|$ and $|A|$ can be very different. It therefore makes sense to rephrase this fundamental principle as a result on the existence of points of bounded size in a vector space, i.e. the L-dimensional null-space of the given $N \times M$ linear system, where $L = N - M$. In this context, "size" is usually measured via a suitable height function.

The first subspace version of Siegel's lemma over an arbitrary number field k, providing a full small-height basis, was established by Bombieri and Vaaler in [8]. It asserts that, given an L-dimensional subspace V of k^N there exists a basis $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ for V such that

$$\prod_{i=1}^{L} h(\boldsymbol{x}_i) \leq N^{L/2} \left(\left(\frac{2}{\pi} \right)^{r_2} |\Delta_k| \right)^{\frac{L}{2d}} \mathcal{H}(V), \tag{1.15}$$

where $d = [k : \mathbb{Q}]$, r_2 is the number of pairs of complex conjugate embeddings of k, Δ_k is the discriminant of k.

Our goal is to establish a new version of Siegel's lemma with somewhat different bounds on the maximum heights of the basis vectors, independent of the field of definition. Throughout this section, we assume that k is an algebraic number field and work in the k-linear space of $N \times 1$ column vectors k^N. More generally, we could also work in an intermediate field K such that $\mathbb{Q} \subseteq K \subseteq \overline{\mathbb{Q}}$ where it may happen that $K/\mathbb{Q}$ is an extension of infinite degree. However, we are interested in simultaneous solutions to finitely many linear equations having algebraic numbers as coefficients. Such systems involve only finitely many algebraic numbers and these generate a finite extension of $\mathbb{Q}$. Several inequalities in the literature (for example, [8, Theorem 8 and Theorem 9] and [62, Theorem 1, and Theorem 2]) that bound the height of solutions to simultaneous systems of linear equations with coefficients in a number field k contain constants that depend on k. Usually these constants depend on the discriminant of k, as in (1.15) above. An exception to this situation can be found in the striking results of Roy and Thunder [53], and [54], on absolute forms of Siegel's lemma. Analogously to (1.15), they prove the existence of a basis $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ for an L-dimensional subspace $V \subset \overline{\mathbb{Q}}^N$ with

$$\prod_{i=1}^{L} \mathcal{H}(\boldsymbol{x}_i) \leq \left(2^{\frac{L(L-1)}{2}} + \varepsilon\right) \mathcal{H}(V), \tag{1.16}$$

for any $\varepsilon > 0$ (the choice of the basis depends on ε). While their bound does not depend on any number field, the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ are also not guaranteed to lie over a fixed number field.

Similar to the work of Roy and Thunder, we establish the existence of a small-height basis for an L-dimensional subspace of k^N (i.e., the space of solutions to a system of simultaneous linear equations), and the inequalities we prove are free of constants that depend on a number field. While we bound the individual heights of the vectors instead of the product, our basis lies over a fixed number field k and

our bound is particularly simple. In fact, we prove more than just the existence of a small-height basis for a subspace. Here is our main result.

Theorem 1.4.1. *Let $\mathcal{Z} \subseteq k^N$ be a subspace of dimension L where $1 \leq L < N$. There exists a basis*

$$\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_L\}$$

for $\mathcal{Z}$ over k with the following property: if $I \subseteq \{1, \ldots, L\}$ is a nonempty subset, and

$$\mathcal{Y}_I = \operatorname{span}_k \{\boldsymbol{\omega}_i : i \in I\} \tag{1.17}$$

is the k-linear subspace spanned by the basis vectors that are indexed by the elements in I, then

$$\mathcal{H}(\mathcal{Y}_I) \leq \mathcal{H}(\mathcal{Z}). \tag{1.18}$$

In particular, $\max_{1 \leq i \leq L} \mathcal{H}(\boldsymbol{\omega}_i) \leq \mathcal{H}(\mathcal{Z})$. *Moreover, if $I_1 \subsetneq I_2 \subseteq \{1, \ldots, L\}$, then*

$$\mathcal{H}(\mathcal{Y}_{I_1}) \leq \mathcal{H}(\mathcal{Y}_{I_2}) \tag{1.19}$$

Observe that, while our result does not imply the bounds (1.15) and (1.16), those bounds do not imply our result either. Further, notice that in situations when the height of the subspace $\mathcal{Z}$ is dominated by the constant depending on N and k in (1.15) or on L in (1.16) our bound may be better. In fact, the constant in (1.15) has been improved in [62], but even this optimal constant depends on a power of Δ_k. Additionally, our bound can be preferable in some applications due to its simplicity.

Our proof of this new form of Siegel's lemma does not use the Dirichlet box principle which was exploited in the earlier work of Baker [4], Roth [52], Siegel [57], and Thue [60]. Our approach also does not use methods from the geometry of numbers which were introduced in [8]; it is based solely on linear algebra. One application of our Theorem 1.4.1 together with a previous version of Siegel's lemma and certain

results on integer sensing matrices (see [23], [44], [45]) is the following observation.

Theorem 1.4.2. *Let $\mathcal{Z} \subseteq k^N$ be a subspace of dimension L where $1 \leq L < N$. For each integer $M > L$ there exists a collection of vectors*

$$S(M) = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$$

with the following properties:

1. *Every subcollection of L vectors from $S(M)$ forms a basis for $\mathcal{Z}$,*

2. *For every $\boldsymbol{y}_i \in S(T)$,*

$$\mathcal{H}(\boldsymbol{y}_i) \leq L^{3/2}(2M)^{\frac{L-1}{L}} \min\left\{\mathcal{H}(\mathcal{Z})^L, \gamma_k(L)^{L/2}\mathcal{H}(\mathcal{Z})\right\},$$

where $\gamma_k(L)^{1/2}$ is the generalized Hermite's constant discussed in Chapter IV.

We present the proofs of Theorems 1.4.1 and 1.4.2 in Section 4.1

The results of this section as well as Chapter III are based on joint work with L. Fukshansky and J. Vaaler and will also be presented in [22].

1.5 Multilinear Forms

Hilbert's 10th problem asks for an algorithm to decide whether a given Diophantine equation has an integer solution. By a celebrated result of Matiyasevich [48], such an algorithm does not exist in general. On the other hand, for linear Diophantine equations solutions are classically given by the Euclidean algorithm. Further, there are also known algorithms for quadratic polynomials (see, for instance [38]). An important approach to the problem of finding such algorithms for different classes of polynomials is through the use of *search bounds*, as described in [46]. Suppose we can prove that a given equation has integer solutions if and only if it has a solution of norm bounded by some explicit function of the coefficients of this equation. Then a search through a finite set of all integer points with norm bounded by this function provides an algorithm that decides whether a solution exists and finds at least one such solution if it exists.

In fact, search bounds for zeros of polynomial equations have been studied quite extensively over more general rings and fields as well: in these more general situations the role of a norm guaranteeing the finiteness of a searchable set is played by a height function. The subject of search bounds for quadratic polynomials has been started by a classical theorem of Cassels [15]; see [25] for a detailed overview of a large body of work on various extensions and generalizations of this important theorem. Additionally, there are search bounds for integral cubic forms in a sufficient number of variables [12], as well as for systems of integral forms under certain technical non-singularity conditions [42].

The results in this section investigate bounds on height of "small" solutions to polynomial equations, linear in some of the variables, over number fields. Let $F(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ and let $1 \leq k < n$. Let $I = \{i_1, \ldots, i_k\} \subset [n]$, and

$I' = [n] \setminus I$. Let $\boldsymbol{x}_{I'} = (x_j)_{j \in I'}$. We will say that F is linear in I-separated variables if

$$F(x_1, \ldots, x_n) = \sum_{j=1}^{k} x_{i_j} F_j(\boldsymbol{x}_{I'}) + F_{k+1}(\boldsymbol{x}_{I'}), \tag{1.20}$$

where $F_j(\boldsymbol{x}_{I'}) \in K[\boldsymbol{x}_{I'}]$ for $1 \leq j \leq k+1$ are any polynomials in $n-k$ variables indexed by I' with coefficients in K. For a polynomial $F(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$, we define its zero-set over K:

$$Z_K(F) = \{\boldsymbol{z} \in K^n : F(\boldsymbol{z}) = 0\}.$$

We also write $\mathcal{N}(F)$ for the number of nonzero monomials of F. With this notation, we can state our first main result.

Theorem 1.5.1. *Let I be as above and let*

$$F_l(x_1, \ldots, x_n) = \sum_{j=1}^{k} x_{i_j} F_{l,j}(\boldsymbol{x}_{I'}) + F_{l,k+1}(\boldsymbol{x}_{I'}),\ 1 \leq l \leq k$$

be polynomials over K of respective degrees $m_1, \ldots, m_k$ linear in I-separated variables as in (1.20)*. Consider the inhomogeneous system*

$$\left.\begin{aligned} F_1(x_1, \ldots, x_n) = \textstyle\sum_{j=1}^{k} x_{i_j} F_{1,j}(\boldsymbol{x}_{I'}) + F_{1,k+1}(\boldsymbol{x}_{I'}) &= 0 \\ &\vdots \\ F_k(x_1, \ldots, x_n) = \textstyle\sum_{j=1}^{k} x_{i_j} F_{k,j}(\boldsymbol{x}_{I'}) + F_{k,k+1}(\boldsymbol{x}_{I'}) &= 0 \end{aligned}\right\} \tag{1.21}$$

of linear equations in the variables $x_{i_1}, \ldots, x_{i_k}$ with coefficients $F_{l,j}(\boldsymbol{x}_{I'})$, $1 \leq l \leq k$, $1 \leq j \leq k+1$. Assume that the matrix $\mathcal{F} := (F_{l,j}(\boldsymbol{x}_{I'}))_{1 \leq l \leq k, 1 \leq j \leq k}$ of the corresponding homogeneous system has the same rank as the coefficient matrix of inhomogeneous system, i.e., $\mathcal{F}$ augmented by the column $(F_{l,k+1}(\boldsymbol{x}_{I'}))_{1 \leq l \leq k}$. Then $\bigcap_{l=1}^{k} Z_K(F_l) \neq \emptyset$

and there exists a point $\boldsymbol{z} \in \bigcap_{l=1}^{k} Z_K(F_l)$ *with*

$$h(\boldsymbol{z}) \le k^{k+1} |\Delta_K|^{\frac{1}{d}} \left(\frac{D+2}{2} \right)^{2km+1} (\mathcal{N}\mathfrak{H})^{2k},$$

where

$$D = \sum_{l=1}^{k} m_l, \ \ m = \max_{1 \le l \le k} m_l,$$

$$\mathcal{N} = \max_{1 \le l \le k} \mathcal{N}(F_l), \ \ \mathfrak{H} = \max_{1 \le l \le k} h(F_l).$$

We prove this theorem in Section 5.2, where we also show how our method of proof leads to an explicit algorithm for finding a simultaneous zero of the polynomial system in question. Our main tools are the Bombieri–Vaaler Siegel's lemma (Theorem 1.15), a non-vanishing lemma for polynomials related to Alon's Combinatorial Nullstellensatz (Lemma 5.1.4) and a collection of height inequalities that we discuss in Sections 1.3 and 5.1. One of these height inequalities that we prove, a bound on the height of the inverse of a nonsingular matrix (Lemma 5.1.1), is of some independent interest and may have other applications in Diophantine geometry.

In the case of a single polynomial, we can prove a similar result but with additional avoidance conditions.

Theorem 1.5.2. *Let* $n \ge 2$ *and* $F(\boldsymbol{x})$ *and* $P(\boldsymbol{x})$ *be polynomials in* n *variables over* K *of degrees* g *and* m*, respectively. Assume further that* F *is linear in at least one of the variables and* $Z_K(F) \not\subseteq Z_K(P)$*. Then there exists a point* $\boldsymbol{z} \in Z_K(F) \setminus Z_K(P)$ *such that*

$$h(\boldsymbol{z}) \le \mathcal{N}(F) \left(\frac{m(2g-1)+2}{2} \right)^{g+1} h(F), \tag{1.22}$$

where $\mathcal{N}(F)$ *is the number of monomials of* F*.*

We prove this theorem in Section 5.3. Finally, we separately discuss the case of

homogeneous multilinear polynomials. We refer to a homogeneous polynomial

$$F(x_1, \dots, x_n) \ \in \ K[x_1, \dots, x_n]$$

of degree g that is linear in every variable as an (n, g)-multilinear form over K. Such forms have many zeros; in particular, they vanish on all sufficiently sparse vectors, specifically on vectors with no more than $g - 1$ nonzero coordinates. We use this observation to obtain the following search bounds which we prove in Section 5.4.

Theorem 1.5.3. *Let $V \subseteq K^n$ be an m-dimensional subspace and F a multilinear (n, g)-form over K. Assume that $m + g - 1 > n$ and $g > 1$. Then V contains a basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ of vectors such that $F(\boldsymbol{x}_1) = \cdots = F(\boldsymbol{x}_m) = 0$ and*

$$H(\boldsymbol{x}_i) \le \mathcal{H}(V).$$

for each $1 \le i \le m$. Further, suppose that $P(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$ is a polynomial such that the dimension of the subspace of V that $Z_K(P, V) := Z_K(P) \cap V$ spans is

$$D(P, V) := \dim \operatorname{span}_K Z_K(P, V) < m.$$

Then there exists a point $\boldsymbol{z} \in V \setminus Z_K(P, V)$ such that $F(\boldsymbol{z}) = 0$ and

$$H(\boldsymbol{z}) \le \sqrt{2} m |\Delta_K|^{\frac{m+1}{2d}} \mathcal{H}(V).$$

This theorem is a direct consequence of our new version of Siegel's lemma (Theorem 1.4.1).

The results of this section as well as Chapter V are based on joint work with L. Fukshansky and will also be presented in [21].

Chapter II

On Lattice Extensions

This chapter deals with results related to the construction of lattice extensions. The results in this chapter apply to lattices in an arbitrary dimension. For similar results that are specific to planar lattices see Chapter III.

2.1 Small Determinant Extensions

We begin with a powerful tool to determine if a collection of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{Z}^n$, $1 \leq m \leq n$, is primitive. Such a collection is primitive if and only if

$$\gcd(\boldsymbol{a}_1 \wedge \ldots \wedge \boldsymbol{a}_m) = 1.$$

This result follows directly from Lemma 2, p.15, [13].

Proposition 2.1.1. *Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ be linearly independent vectors in $\mathbb{Z}^n$ and let*

$$L_m = \operatorname{span}_{\mathbb{Z}} \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subset \mathbb{Z}^n$$

be the sublattice of rank m spanned by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, $m < n$. Then there exists a rank n lattice extension, L_n, of L_m in $\mathbb{Z}^n$ so that

$$\det L_n = \gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m).$$

Proof. Let $\overline{L}_m = \mathbb{Z}^n \cap \operatorname{span}_{\mathbb{R}} L_m$, then $\overline{L}_m \subset \mathbb{Z}^n$ is a sublattice of rank m containing L_m such that $\mathbb{Z}^n/\overline{L}_m$ is torsion free. Hence any basis of $\overline{L}_m$ is extendable to $\mathbb{Z}^n$. Let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ be a basis for $\overline{L}_m$ extended to a basis for $\mathbb{Z}^n$ by $\boldsymbol{y}_{m+1}, \ldots, \boldsymbol{y}_n$. Since $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ are two collections of integer vectors spanning the same subspace of $\mathbb{R}^n$, the vectors of Grassmann coordinates represent the same rational projective point. Further, since the collection $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ is extendable to a basis of

$\mathbb{Z}^n$, the Grassmann coordinates of this collection must be relatively prime. Hence

$$\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m = g\left(\boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_m\right)$$

where $g = \gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)$. Define

$$L_n = \operatorname{span}_{\mathbb{Z}}\left\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \boldsymbol{y}_{m+1}, \ldots, \boldsymbol{y}_n\right\}.$$

By the bi-linearity of the wedge product,

$$\det L_n = \boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m \wedge \boldsymbol{y}_{m+1} \wedge \cdots \wedge \boldsymbol{y}_n = g\left(\boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_m \wedge \boldsymbol{y}_{m+1} \wedge \cdots \wedge \boldsymbol{y}_n\right),$$

and since $\boldsymbol{y}_1 \wedge \cdots \wedge \boldsymbol{y}_m \wedge \boldsymbol{y}_{m+1} \wedge \cdots \wedge \boldsymbol{y}_n = \det \mathbb{Z}^n = 1$, we have that $\det L_n = g$. $\square$

Corollary 2.1.2. *Let the notation be as in Proposition 2.1.1 with $m = n-1$. Then there exists $\boldsymbol{y} \in \mathbb{Z}^n$ so that $L_n = \operatorname{span}_{\mathbb{Z}}\{L_{n-1}, \boldsymbol{y}\}$ and*

$$\|\boldsymbol{y}\| \leq \left\{\left(\frac{\gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)}{\det L_{n-1}}\right)^2 + \mu^2\right\}^{1/2},$$

where μ is the covering radius of L_{n-1}.

Proof. Write $A = (\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n-1})$ for the corresponding basis matrix of L_{n-1} and let L_n be as given by Proposition 2.1.1. This means that there exists $\boldsymbol{z} \in \mathbb{Z}^n$ such that $L_n = \operatorname{span}_{\mathbb{Z}}\{L_{n-1}, \boldsymbol{z}\}$, so $\det(L_n) = \gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)$. Let $\rho_{L_{n-1}} = A(A^\top A)^{-1}A^\top$ be the orthogonal projector onto $\operatorname{span}_{\mathbb{R}} L_{n-1}$. Let

$$\begin{aligned}
\mathcal{P} &= \left\{\sum_{i=1}^{n-1} a_i \boldsymbol{x}_i : 0 \leq a_i < 1 \;\forall\; 1 \leq i \leq n-1\right\}, \\
\mathcal{P}' &= \{\boldsymbol{u} + a\boldsymbol{z} : \boldsymbol{u} \in \mathcal{P},\ 0 \leq a < 1\}
\end{aligned}$$

be fundamental parallelepipeds for L_{n-1} and L_n, respectively. Then

$$\begin{aligned}\gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m) &= \det L_n = \mathrm{Vol}_n(\mathcal{P}') \\ &= \mathrm{Vol}_{n-1}(\mathcal{P}) \left\|(I_n - \rho_{L_{n-1}})\boldsymbol{z}\right\| \\ &= \det L_{n-1} \left\|(I_n - \rho_{L_{n-1}})\boldsymbol{z}\right\|,\end{aligned}$$

hence

$$\left\|(I_n - \rho_{L_{n-1}})\boldsymbol{z}\right\| = \frac{\gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)}{\det L_{n-1}}.$$

On the other hand, $\rho_{L_{n-1}}\boldsymbol{z} \in \mathrm{span}_{\mathbb{R}} L_{n-1}$, and by definition of the covering radius μ of L_{n-1}, there exists $\boldsymbol{v} \in L_{n-1}$ such that $\|\rho_{L_{n-1}}\boldsymbol{z} - \boldsymbol{v}\| \leq \mu$. Let $\boldsymbol{y} = \boldsymbol{z} - \boldsymbol{v}$, then $\boldsymbol{y} \in \mathbb{Z}^n$ and

$$\rho_{L_{n-1}}\boldsymbol{y} = \rho_{L_{n-1}}\boldsymbol{z} - \rho_{L_{n-1}}\boldsymbol{v} = \rho_{L_{n-1}}\boldsymbol{z} - \boldsymbol{v},$$

since $\boldsymbol{v} \in \mathrm{span}_{\mathbb{R}} L_{n-1}$. Then $(I_n - \rho_{L_{n-1}})\boldsymbol{y} = (I_n - \rho_{L_{n-1}})\boldsymbol{z}$ and

$$L_n = \mathrm{span}_{\mathbb{Z}} \{L_{n-1}, \boldsymbol{z}\} = \mathrm{span}_{\mathbb{Z}} \{L_{n-1}, \boldsymbol{y}\}.$$

Therefore, by Pythagorean theorem,

$$\begin{aligned}\|\boldsymbol{y}\|^2 &= \left\|(I_n - \rho_{L_{n-1}})\boldsymbol{y}\right\|^2 + \|\rho_{L_{n-1}}\boldsymbol{y}\|^2 \\ &= \left\|(I_n - \rho_{L_{n-1}})\boldsymbol{z}\right\|^2 + \|\rho_{L_{n-1}}\boldsymbol{z} - \boldsymbol{v}\|^2 \\ &\leq \left(\frac{\gcd(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_m)}{\det L_{n-1}}\right)^2 + \mu^2.\end{aligned}$$

The result then follows. □

Now Theorem 1.1.3 follows by combining Proposition 2.1.1 with Corollary 2.1.2.

Remark 2.1.1. Let $\Lambda = A\mathbb{Z}^n \subset \mathbb{R}^k$ be a lattice of rank $n \leq k$ and let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m, m \leq n$, be linearly independent vectors in Λ. Then for each $1 \leq i \leq m, \boldsymbol{z}_i = A\boldsymbol{x}_i$, where

$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \mathbb{Z}^n$ are also linearly independent. Let

$$\Omega = \operatorname{span}_{\mathbb{Z}}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subset \mathbb{Z}^n \tag{2.1}$$

be the sublattice of rank n spanned by these vectors and let Ω' be an extension of Ω in $\mathbb{Z}^n$ guaranteed by Proposition 2.1.1. Then $A\Omega = \operatorname{span}\mathbb{Z}\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\} \subseteq \Lambda$ and $A\Omega' \subseteq \Lambda$ is an extension of $A\Omega$ with

$$\det A\Omega' = \sqrt{\det(A^T A)}\det\Omega' = \det\Lambda \det\Omega'.$$

Further, if $m = n - 1$ then there exists $\boldsymbol{y} \in \mathbb{Z}^n$ so that $A\Omega' = \operatorname{span}_{\mathbb{Z}}\{A\Omega, A\boldsymbol{y}\}$ and $\|\boldsymbol{y}\|$ is bounded as in Corollary 2.1.2.

2.2 Counting Lattice extensions

The goal of this section is to prove Theorem 1.1.1 and Corollary 1.1.2. We will approach this by breaking it into several steps, the first of which is to count the number of ways to extend a primitive collection of $m < n - 1$ vectors in $\mathbb{Z}^n$ by one vector.

For a real number $T > 0$ define the integer n-cube centered at the origin with sidelength $2T$ as

$$C_n(T) := \{\boldsymbol{x} \in \mathbb{Z}^n : |\boldsymbol{x}| \leq T\},$$

then $|C_n(T)| = (2\lfloor T \rfloor + 1)^n$.

Lemma 2.2.1. *Let $1 \leq m < n - 1$ and let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{Z}^n$ be a primitive collection of vectors. For $T > 0$, define*

$$f(T) = |\{\boldsymbol{x} \in C_n(T) : \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{x} \text{ is primitive}\}|.$$

Then as $T \to \infty$,

$$f(T) \leq (\zeta^{-1}(n) + \epsilon)(2T+1)^n, \tag{2.2}$$

where ζ is the Riemann zeta-function, $\epsilon > 0$. Additionally,

$$f(T) \leq \beta(n, m, A)T^n, \tag{2.3}$$

where $\beta(n, m, A)$ is a constant depending only on n, m and the A.

Proof. Since the collection $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ is primitive, the corresponding $n \times m$ matrix$A = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_m)$ is unimodular.

By the primitivity criterion above, we want to count $\boldsymbol{x} \in C_n(T)$ such that the extended $n \times (m+1)$ matrix $(A\ \boldsymbol{x})$ is still unimodular.

First notice that each $\boldsymbol{x}$ must itself be a primitive vector, as

$$\gcd(\boldsymbol{x}) | \gcd(\boldsymbol{a}_1 \wedge \ldots \wedge \boldsymbol{a}_m \wedge \boldsymbol{x}).$$

Therefore the total number of such vectors $\boldsymbol{x}$ is no bigger than the number of primitive vectors in $C_n(T)$. It is a well-known fact that the probability of a vector in $C_n(T)$ being primitive is $\zeta^{-1}(n)$ (this result has apparently first been proved by E. Cesàro in 1884, but has been re-discovered several times since; see [49] for the references). More specifically, a result of [51] asserts that

$$|\{\boldsymbol{x} \in C_n(T) : \boldsymbol{x} \text{ is primitive}\}| = \zeta(n)^{-1}T^n + O(T^{n-1}).$$

Taking any $\epsilon > 0$ then guarantees (2.2) for all sufficiently large T.

Now, let $\Lambda := A\mathbb{Z}^m$ be a lattice of rank m. Since A is a unimodular matrix, there exists an $n \times (n-m)$ integer matrix B so that the augmented matrix $(AB) \in \mathrm{GL}_n(\mathbb{Z})$. In fact, such B can be chosen so that $\sqrt{\det(B^T B)}$ is bounded by a function of A, call

it $\alpha(A)$: this can be done, for instance, by a repetitive application of the search-bound presented in Section 1 of [11], after Theorem 1.3. Then

$$\Omega := B\mathbb{Z}^{n-m} \cong \mathbb{Z}^n/\Lambda \cong \mathbb{Z}^{n-m}$$

is a lattice, and

$$\det(\Omega) = \sqrt{\det(B^T B)} \leq \alpha(A), \det(\Lambda) = \sqrt{\det(A^T A)}. \tag{2.4}$$

Then $\mathbb{Z}^n = \Omega \oplus \Lambda$, so for every $\boldsymbol{x} \in \mathbb{Z}^n$ there exists a unique pair $\boldsymbol{y} \in \Omega, \boldsymbol{z} \in \Lambda$ such that $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, so $|\boldsymbol{x}| \leq |\boldsymbol{y}| + |\boldsymbol{z}|$. Then $\boldsymbol{x}$ is such that $(A\boldsymbol{x})$ is unimodular if and only if the corresponding $\boldsymbol{y} \in \Omega$ is primitive, i.e. extendable to a basis of Ω. Let $\gamma \in (0, 1]$ and notice that

$$g_\gamma(T) := |\{\boldsymbol{y} + \boldsymbol{z} : \boldsymbol{y} \in C_n(\gamma T) \cap \Omega, \boldsymbol{z} \in C_n((1-\gamma)T) \cap \Lambda\}| \leq f(T),$$

where Ω' stands for the set of primitive points in Ω. Now notice that

$$g_\gamma(T) = |C_n(\gamma T) \cap \Omega' \cdot |C_n((1-\gamma)T) \cap \Lambda|. \tag{2.5}$$

Assume $T \geq \max\left\{\frac{(n-m)\det(\Omega)}{2\gamma}, \frac{m\det(\Lambda)}{2(1-\gamma)}\right\}$. Then Lemma 3.1 of [30] guarantees that

$$|C_n(\gamma T) \cap \Omega| \geq \left(\frac{2\gamma T}{(n-m)\det(\Omega)}\right)\left(\frac{2\gamma T}{n-m} - 1\right)^{n-m-1}, \tag{2.6}$$

$$|C_n((1-\gamma)T) \cap \Lambda| \geq \left(\frac{2(1-\gamma)T}{m\det(\Lambda)} - 1\right)\left(\frac{2(1-\gamma)T}{m} - 1\right)^{m-1}. \tag{2.7}$$

Again, by Cesàro's Theorem the proportion of primitive points among all points in Ω is $\zeta(n-m)^{-1}$. Combining this observation with (2.4), (2.5), (2.6), (2.7) and taking

$\gamma = 1/2$, we obtain

$$f(T) \geq g_{1/2}(T) \geq \beta(n, m, A)T^n$$

for an appropriate constant $\beta(n, m, A)$. This proves (2.3). □

Next we prove a counting lemma on the number of integer lattice points in a section of the cube $C_n(T)$ by a hyperplane, building on a previous result for a section by a subspace. Let

$$L(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \sum_{i=1}^{n} c_i \boldsymbol{x}_i \in \mathbb{Z}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$$

be a linear form in $n \geq 2$ variables with coprime coefficients, and write $\boldsymbol{c} = (c_1, \ldots, c_n)$ for this coefficient vector. Let $b \in \mathbb{Z}$ and let $T > 0$ be a real number. Define the set

$$\mathbb{B}_{L,b}(T) = \{\boldsymbol{x} \in \mathbb{Z}^n : L(\boldsymbol{x}) = b, |\boldsymbol{x}| \leq T\} = C_n(T) \cap \{\boldsymbol{x} \in \mathbb{Z}^n : L(x) = b\}.$$

Since coefficients of L are coprime, the equation $L(\boldsymbol{z}) = b$ has infinitely many integer solutions for any $b \in \mathbb{Z}$, and so the set $\mathbb{B}_{L,b}(T)$ is not empty for a sufficiently large T. We want to estimate the size of $\mathbb{B}_{L,b}(T)$ as a function of the coefficients of L, b and T.

Theorem 2.2.2. *For any T,*

$$|\mathbb{B}_{L,b}(T)| \leq \left(\frac{2(T + \max\{|L|, |b|\}}{|L|} + 1\right)(2(T + \max\{|L|, |b|\}) + 1)^{n-2}, \tag{2.8}$$

and for $T \geq \max|L|, |b|$

$$|\mathbb{B}_{L,b}(T)| \leq \left(\frac{2(T - \max\{|L|, |b|\}}{|L|} - 1\right)(2(T - \max\{|L|, |b|\}) - 1)^{n-2}. \tag{2.9}$$

Therefore

$$|\mathbb{B}_{L,b}(T)| \sim \frac{(2T)^{n-1}}{|L|} \tag{2.10}$$

as $T \to \infty$.

Proof. Let

$$\Lambda = \{x \in \mathbb{Z}^n : L(x) = 0\} = \{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{c} \cdot \boldsymbol{x} = 0\},$$

then Λ is a sublattice of $\mathbb{Z}^n$ of rank $n-1$. Further, if A is any basis matrix for Λ then column vectors of A must be primitive, since Λ is the full intersection of $\mathbb{Z}^n$ with a subspace. We define the Grassmann coordinates of Λ to be the absolute values of Grassmann coordinates of A. This definition does not depend on the choice of a basis matrix for Λ_L, since for any two such basis matrices A_1, A_2 there exists a matrix $U \in \mathrm{GL}_{n-1}(\mathbb{Z})$ such that $A_2 = UA_1$, where $\det(U) = \pm 1$. Let us write Δ for the maximum of Grassmann coordinates of Λ, then by Lemma 1.3.1,

$$\Delta = |\boldsymbol{c}| = |L|. \tag{2.11}$$

For a fixed integer b, let

$$\Lambda_L(b) = \{\boldsymbol{x} \in \mathbb{Z}^n : L(\boldsymbol{x}) = b\},$$

so $\Lambda_L = \Lambda_L(0)$ and $\mathbb{B}_{L,b}(T) = \{x \in \Lambda_L(b) : |x| \leq T\}$. Pick any $\boldsymbol{z} \in \Lambda_L(b)$, then it is easy to notice that

$$\Lambda_L(b) = \{\boldsymbol{x} + \boldsymbol{z} : \boldsymbol{x} \in \Lambda\}, \tag{2.12}$$

i.e. $\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{z}$ is a bijective map between Λ_L and $\Lambda_L(b)$ for any fixed $\boldsymbol{z} \in \Lambda(b)$. The main Theorem of [10] guarantees that there exists $\boldsymbol{z} \in \Lambda(b)$ such that

$$|\boldsymbol{z}| \leq \max\{|L|, |b|\}, \tag{2.13}$$

so from now on we use description (2.12) for$\Lambda(b)$ with $\boldsymbol{z}$ satisfying (2.13). Hence for any $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z} \in \Lambda(b)$,

$$|\boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{z}| \leq |\boldsymbol{z}| + \max(|L|, |b|).$$

Combining Theorem 4.2 of [10] with (2.11), we have

$$|\mathbb{B}_{L,0}(T)| \leq \left(\frac{2T}{|L|} + 1\right)(2T+1)^{n-2}, \tag{2.14}$$

and combining Lemma 3.1 of [30] (see also equation (50)) with (2.11), we have for every $T \geq \frac{|L|}{2}$

$$|\mathbb{B}_{L,0}(T)| \leq \left(\frac{2T}{|L|} - 1\right)(2T-1)^{n-2}. \tag{2.15}$$

Suppose that $\boldsymbol{y} \in \mathbb{B}_{L,b}(T)$, then $|\boldsymbol{y}| \leq T$ and $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$ for a unique $\boldsymbol{x} \in \Lambda$, so

$$|\boldsymbol{x}| = |\boldsymbol{y} - \boldsymbol{z}| \leq T + |\boldsymbol{z}| = T + \max\{|L|, |b|\}.$$

Therefore $|\mathbb{B}_{L,b}(T)| \leq |\mathbb{B}_{L,0}(T + \max\{|L|, |b|\})|$, and combining this observation with (2.14), we obtain (2.8).

Next assume $\boldsymbol{x} \in \mathbb{B}_{L,0}(T - \max\{|L|, |b|\})$, which implicitly implies that $T \geq \max\{|L|, |b|\}$. Let $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z} \in \mathbb{B}_{L,b}(T)$, then

$$|\boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{z}| \leq T,$$

and so

$$|\mathbb{B}_{L,b}(T)| \geq |\mathbb{B}_{L,0}(T - \max\{|L|, |b|\})|.$$

Then combining this observation with (2.15), we obtain (2.9), since we have $T \geq \max\{|L|, |b|\} > \frac{|L|}{2}$.

Now notice that both, the upper bound (2.8) and the lower bound (2.9) when

expanded under the assumption $T \to \infty$ have the order of magnitude $\frac{(2T)^{n-1}}{|L|} + o(T^{n-1})$. Thus

$$\lim_{T\to\infty} \frac{|\mathbb{B}_{L,b}(T)}{(2T)^{n-1}/|L|} = 1$$

which implies (2.10). □

We are now ready to prove Theorem 1.1.1 and Corollary 1.1.2. To start with, let $n \geq 2, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1} \in \mathbb{Z}^n$ be a primitive collection of vectors, and let $A = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_{n-1})$ be the corresponding $n \times (n-1)$ unimodular matrix. In how many ways can this primitive collection be extended to a basis of $\mathbb{Z}^n$? More precisely, for a positive integer T let

$$\mathbb{B}_A(T) = \{z \in \mathbb{Z}^n : \mathbb{Z}^n = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1}, \boldsymbol{z}\}, |\boldsymbol{z}| \leq T\}. \tag{2.16}$$

We want to understand how big is the cardinality of this set, $|\mathbb{B}_A(T)|$, as a function of A and T. Notice that $\boldsymbol{z} \in \mathbb{B}_A(T)$ if and only if $|\boldsymbol{z}| \leq T$ and

$$\det(A\boldsymbol{z}) = \pm 1.$$

For each $1 \leq k \leq n$, let A_k be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting k-th row, then

$$L_A(\boldsymbol{z}) := \det(A\boldsymbol{z}) = \sum_{k=1}^{n} (-1)^{n+k} \det(A_k) z_k,$$

which is a linear form in the variables $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$. Since the collection of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1}$ is extendable to a basis for $\mathbb{Z}^n$, it must be true that

$$\gcd(\det(A_1), \ldots, \det(A_n)) = 1,$$

and hence the equation $L_A(\boldsymbol{z}) = \pm 1$ has infinitely many integer solutions. Define

$$\Delta_A := \max |\det(A_k)| : 1 \leq k \leq n,$$

then $|L_A| = \Delta_A \geq 1$, and so we can apply Theorem 2.2.2 with $b = 1$ and with $b = -1$ to obtain the following bound.

Corollary 2.2.3. *For any T,*

$$|\mathbb{B}_A(T)| \leq 2\left(\frac{2T}{\Delta_A} + 3\right)(2(T + \Delta_A) + 1)^{n-2},$$

and for $T \geq \Delta_A$

$$|\mathbb{B}_A(T)| \geq 2\left(\frac{2T}{\Delta_A} - 3\right)(2(T - \Delta_A) - 1)^{n-2}.$$

Therefore

$$|\mathbb{B}_A(T)| \sim 2\left(\frac{(2T)^{n-1}}{\Delta_A}\right)$$

as $T \to \infty$.

Proof. Since

$$\mathbb{B}_A(T) = \mathbb{B}_{L_A,1}(T) \cup \mathbb{B}_{L_A,-1}(T),$$

we are applying Theorem 2.2.2 twice, with $b = \pm 1$, and adding the results. This produces the factor of two in our bounds. □

Now we combine Corollary 2.2.3 with Lemma 2.2.1 to prove Theorem 1.1.1.

Proof of Theorem 1.1.1. Parts (1) and (2) of the Theorem are given by Lemma 2.2.1 and Corollary 2.2.3, respectively. Let us prove that there exist $\Theta(T^{n^2-nm-1})$ collections of vectors $\mathbb{B}_1, \ldots, \mathbb{B}_{n-m} \in \mathbb{Z}^n$ such that $|\mathbb{B}_i| \leq T$ for each $1 \leq i \leq n - m$ and $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \mathbb{B}_1, \ldots, \mathbb{B}_{n-m}\}$ is a basis for $\mathbb{Z}^n$. Let us argue by induction on $n - m \geq 1$.

If $n - m = 1$, then we only need to add one vector to this primitive collection, and by Corollary 2.2.3 there are $\Theta(T^{n-1})$ ways to do it. Notice that in this case

$$n^2 - nm - 1 = n^2 - n(n-1) - 1 = n - 1,$$

so the result follows. Then assume $n - m > 1$, and a result is proved for $n - m - 1$, i.e. for a primitive collection of $m + 1$ vectors. By Lemma 2.2.1 there are $\Theta(T^n)$ to extend this primitive collection by one vector. For each such vector, there are $\Theta(T^{n^2-n(m+1)-1})$ extensions to a basis by the induction hypothesis, and hence the total number of extensions of our primitive collection is

$$\Theta(T^n T^{n^2-n(m+1)-1}) = \Theta(T^{n^2-nm-n-1+n}) = \Theta(T^{n^2-nm-1}).$$

Finally, the argument for extending the primitive collection $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$ to a primitive collection $\{\boldsymbol{a}_i, \mathbb{B}_j : 1 \leq i \leq m, 1 \leq j \leq k\}, 1 \leq k < n - m$, is the same as above, but simpler: we do not need to account for the case of the last vector contributing only $\Theta(T^{n-1})$ possibilities, and hence the total number is simply $\Theta(T^{nk})$. This completes the proof. □

We now extend these observations to general lattices.

Proof of Corollary 1.1.2. Let Λ be a lattice of full rank in $\mathbb{R}^n$, and let

$$\lambda_1(\Lambda) := \min\{|\boldsymbol{z}| : \boldsymbol{z} \in \Lambda \setminus \{\boldsymbol{0}\}\}$$

be the first successive minimum of Λ with respect to the sup-norm. By Minkowski reduction (see, for instance, Theorem 2 on p.66 of [37] combined with Theorem 2 on

p.62 of the same book), there exists a basis $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ for Λ such that

$$\frac{1}{n!}\det(\Lambda) \leq \prod_{i=1}^{n} |\boldsymbol{z}_i| \leq \left(\frac{3}{2}\right)^{\frac{(n-1)(n-2)}{2}} \det(\Lambda).$$

Let U be the basis matrix for Λ with column vectors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$, ordered in order of increasing sup-norm, so $|U| = |\boldsymbol{z}_n|$, and thus

$$\left(\frac{\det(\Lambda)}{n!}\right)^{1/n} \leq |U| \leq \left(\frac{3}{2}\right)^{\frac{(n-1)(n-2)}{2}} \frac{\det(\Lambda)}{\lambda_1(\Lambda)^{n-1}}. \tag{2.17}$$

Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ be a primitive collection of vectors in $\Lambda, 1 \leq m < n$. Then for each $1 \leq i \leq m, \boldsymbol{a}_i = Ua_i'$ for some $a_i' \in \mathbb{Z}^n$. Let us write $A = (\boldsymbol{a}_1 \ldots \boldsymbol{a}_m)$, then there exists an $n \times (n-m)$ matrix B such that

$$(AB)\mathbb{Z}^n = \Lambda = U\mathbb{Z}^n,$$

hence $U^{-1}(AB) = ((U^{-1}A)(U^{-1}B)) \in \mathrm{GL}_n(\mathbb{Z})$, where $\boldsymbol{a}_1', \ldots, \boldsymbol{a}_m'$ are the column vectors of $A' := U^{-1}A$. This means that the collection of vectors $\boldsymbol{a}_1', \ldots, \boldsymbol{a}_m'$ is primitive in $\mathbb{Z}^n$, and hence we can apply Theorem 1.1.1 to it. By analogy with (2.16), let

$$\mathbb{B}^m_{A',\mathbb{Z}^n}(T) = \{\boldsymbol{z} \in \mathbb{Z}^n : \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{z} \text{ is primitive in } \mathbb{Z}^n, |\boldsymbol{z}| \leq T\},$$
$$\mathbb{B}^m_{A,\Lambda}(T) = \{\boldsymbol{z} \in \Lambda : \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{z} \text{ is primitive in } \Lambda, |\boldsymbol{z}| \leq T\}.$$

Suppose that $\mathbb{B} \in \mathbb{B}_{A,\Lambda}(T)$. Then $\boldsymbol{a}_1', \ldots, \boldsymbol{a}_m', \mathbb{B}'$ is primitive in $\mathbb{Z}^n$, where $\mathbb{B}' = U^{-1}\mathbb{B}$, and so

$$|\mathbb{B}'| \leq n|U^{-1}||\mathbb{B}| \leq n|U^{-1}|T.$$

Therefore

$$|\mathbb{B}^m_{A,\Lambda}(T)| \leq |\mathbb{B}^m_{A',\mathbb{Z}^n}(n|U^{-1}|T)| = \Theta\left(T^{n+\min\{0,n-m-2\}}\right), \tag{2.18}$$

since $\min\{0, n-m-2\} = 0$ if $m < n-1$ and $\min\{0, n-m-2\} = -1$ if $m = n-1$. On the other hand, assume that $\mathbb{B}' \in \mathbb{B}_{A',\mathbb{Z}^n}(T/n|U|)$. Then $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \mathbb{B}'$ is primitive in Λ, where $\mathbb{B} = U\mathbb{B}'$, and so

$$|\mathbb{B}| \leq n|U||\mathbb{B}'| \leq T.$$

Therefore

$$|\mathbb{B}^m_{A,\Lambda}(T)| \geq |\mathbb{B}^m_{A',\mathbb{Z}^n}(T/n|U|)| = \Theta\left(T^{n+\min\{0,n-m-2\}}\right). \tag{2.19}$$

Combining (2.18) and 2.19 and applying an argument identical to the one in the proof of Theorem 1.1.1 above yields the corollary. Since we choose U to be a Minkowski reduced basis for Λ with sup-norm bounded as in (2.17), the constants in Θ-notation depend intrinsically on Λ, not on the choice of a basis for Λ □

2.3 Successive minima extensions

In this section, we prove Theorem 1.1.4. Let Λ be a full rank lattice in $\mathbb{R}^n$. Our goal is to construct a sublattice $L_{k+1} \subset \Lambda$ of rank $k+1$ such that $L_k \subset L_{k+1}$, $\lambda_j(L_{k+1}) = \lambda_j(L_k)$ for all $1 \leq j \leq k$ and $\lambda_{k+1}(L_{k+1})$ is as small as possible. To prove the theorem, we first need an auxiliary lemma. Write $\lambda_1, \ldots, \lambda_k$ for the successive minima of L_k and let $V_k = \operatorname{span}_{\mathbb{R}} L_k$, $\theta \in (0, \pi/2]$, and

$$C_\theta(V_k) = \{\boldsymbol{x} \in \mathbb{R}^n : \mathfrak{a}(\boldsymbol{x}, \boldsymbol{y}) \in [\theta, \pi - \theta] \ \forall\ \boldsymbol{y} \in V_k\}, \tag{2.20}$$

where $\mathfrak{a}(\boldsymbol{x}, \boldsymbol{y})$ stands for the angle between two vectors.

Lemma 2.3.1. *If $\boldsymbol{x} \in C_\theta(V_k)$ and*

$$\|\boldsymbol{x}\| \geq \frac{\lambda_k(\cot\theta\cos\theta + 1)}{\sqrt{1 + \cos^2\theta}},$$

then $\|\boldsymbol{x} + \boldsymbol{y}\| \geq \lambda_k$ for every $\boldsymbol{y} \in V_k$.

Proof. For $\boldsymbol{x} \in C_\theta(V_k)$ and $\boldsymbol{y} \in V_k$, define

$$f(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| \cos \boldsymbol{a}(\boldsymbol{x}, \boldsymbol{y}).$$

We want to guarantee that $f(\boldsymbol{x}, \boldsymbol{y}) \geq \lambda_k^2$ for all $\boldsymbol{y} \in V_k$. Let us write $t = \|\boldsymbol{x}\|$, $z = \|\boldsymbol{y}\|$, and notice that

$$f(\boldsymbol{x}, \boldsymbol{y}) - \lambda_k^2 \geq g(t, z) := t^2 + z^2 - 2tz \cos\theta - \lambda_k^2.$$

Thus we want to find a lower bound on t that would guarantee $g(t, z) \geq 0$ for all $z > 0$. In other words, we want

$$t \geq h(z) := z \cos\theta + \sqrt{\lambda_k^2 - z^2 \sin^2\theta}$$

for all $z > 0$. Notice that $h(z)$ is real-valued if and only if $z \leq \frac{\lambda_k}{\sin\theta}$, then let us find the value of z that maximizes $h(z)$. Differentiating $h(z)$ and setting the derivative equal to zero, we obtain

$$z_* = \frac{\lambda_k \cot\theta}{\sqrt{1 + \cos^2\theta}},$$

the point at which $h(z)$ assumes its maximum value of

$$h(z_*) = \frac{\lambda_k(\cot\theta \cos\theta + 1)}{\sqrt{1 + \cos^2\theta}}.$$

Thus, taking $t = \|\boldsymbol{x}\|$ to be greater than or equal to this value ensures that $\|\boldsymbol{x} + \boldsymbol{y}\| \geq \lambda_k$ for every $\boldsymbol{y} \in V_k$, as required. □

Proof of Theorem 1.1.4. Let us write $B_n(r)$ for the ball of radius $r > 0$ centered at the origin in $\mathbb{R}^n$. Let $\theta \in (0, \pi/2]$ and

$$r(\theta) = \frac{\lambda_k(\cot\theta \cos\theta + 1)}{\sqrt{1 + \cos^2\theta}}. \tag{2.21}$$

Then Lemma 2.3.1 guarantees that for any vector $\boldsymbol{x} \in \Lambda \cap (C_\theta(V_k) \setminus B_n(r(\theta)))$ the lattice $M = \operatorname{span}_{\mathbb{Z}} \{L_k, \boldsymbol{x}\}$ satisfies $\lambda_j(M) = \lambda_j(L_k)$ for all $1 \leq j \leq k$ and $\lambda_{k+1}(M) \leq \|\boldsymbol{x}\|$. Hence we want to minimize

$$\lambda_{k+1}(\theta) := \min \{\|\boldsymbol{x}\| : \boldsymbol{x} \in \Lambda \cap (C_\theta(V_k) \setminus B_n(r(\theta))\}$$

as a function of θ.

Let μ be the covering radius of Λ, then any translated copy $B'_n(\mu)$ of the ball of radius μ in $\mathbb{R}^n$ must contain a point of Λ. Let us choose $\theta \in (0, \pi/2]$ such that

$$B'_n(\mu) \subset (C_\theta(V_k) \cap B_n(r(\theta) + 2\mu)) \setminus B_n(r(\theta)),$$

then $C_\theta(V_k) \setminus B_n(r(\theta)$ would be guaranteed to contain a point $\boldsymbol{x}$ of Λ with

$$\|\boldsymbol{x}\| \leq r(\theta) + 2\mu, \tag{2.22}$$

so that we can take $L_{k+1} = \operatorname{span}_{\mathbb{Z}} \{L_k, \boldsymbol{x}\}$. For this to be true, we need the line segment from $\mathbf{0}$ to the center of the ball $B'_n(\mu)$ to be of length $r(\theta) + \mu$ and to make the angle $\pi/2 - \theta$ with any line in the boundary of $C_\theta(V_k)$ emanating from the center and tangent to the ball $B'_n(\mu)$. These conditions result in a right triangle with legs $r(\theta) + \mu$ and μ and the angle $\pi/2 - \theta$ opposite to the second leg (see Figure 2.1 for a graphical illustration of this argument). Hence we have the equation

$$\tan(\pi/2 - \theta) = \frac{\mu}{r(\theta) + \mu}.$$

□

Using (2.21), along with the fact that $\tan(\pi/2 - \theta) = \cot\theta$, writing $v = \cos\theta$ and

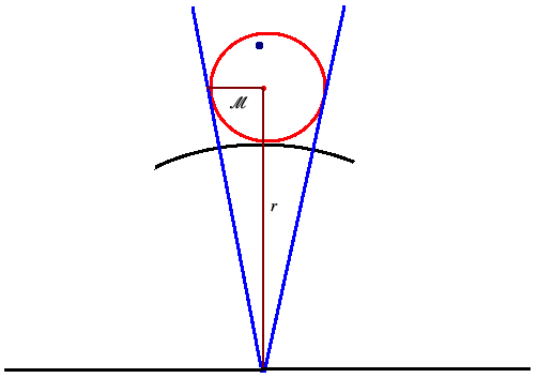

Figure 2.1: Cone construction with the lattice point (in blue) caught in the ball (red) of covering radius.

simplifying, we obtain the following relation in terms of v:

$$\mu\left(\sqrt{1-v^2}-v\right)=\frac{\lambda_k\left(v^2+\sqrt{1-v^2}\right)v}{\sqrt{1-v^4}},$$

which transforms into the following polynomial equation:

$$\left(\frac{\mu^2}{\lambda_k^2}(1-v^4)-v^2(v^4-v^2+1)\right)^2=\left(\frac{2\mu^2}{\lambda_k^2}v(1-v^4)+2v^4\right)^2(1-v^2). \tag{2.23}$$

It follows from our construction that this equation has at least one solution v in the interval $(0,1)$. Then $r(\theta)$ as a function of v becomes

$$r(v)=\frac{\lambda_k(v^2+\sqrt{1-v^2})}{\sqrt{1-v^4}},$$

which is an increasing function of v in the interval $(0,1)$. Hence, to minimize the bound (2.22), we can pick v_* to be the smallest root of the equation (2.23) in the interval $(0,1)$. The inequality (1.1) follows.

Remark 2.3.1. This result can be loosely compared to the construction of a *canonical filtration* of a lattice as originally defined by Grayson and Stuhler (see Casselman's

survey paper [13] for a detailed discussion). This is a unique flag of sublattices

$$\{\mathbf{0}\} = \Lambda_0 \subset \Lambda_1 \subset \cdots \subset \Lambda_n = \Lambda$$

in a lattice Λ such that $\mathrm{rk}(\Lambda_k) = k$ and $\det(\Lambda_n)^{1/n} > \det(\Lambda_k)^{1/k}$ for every $k < n$, where

$$\det(\Lambda_k) = \min\left\{\det(\Omega) : \Omega \subset \Lambda,\ \mathrm{rk}(\Omega) = k\right\}.$$

A lattice Λ is called *semi-stable* if the canonical filtration is $\Lambda_0 \subset \Lambda_n$, i.e. if for each sublattice $\Omega \subseteq \Lambda$,

$$\det(\Lambda)^{1/\mathrm{rk}(\Lambda)} \leq \det(\Lambda)^{1/\mathrm{rk}(\Omega)}. \tag{2.24}$$

This family of lattices is important in reduction theory. Y. Andre explains in [2]:

Reduction theory aims at estimating the length of short vectors, and more generally the (co)volumes of small sublattices of lower ranks, of lattices of given rank and (co)volume, and at combining lower and upper bounds to get finiteness results. A better grasp on lower bounds comes from the more recent part of reduction theory which deals with semistability and slope filtrations (heuristically, semistability means that the Minkowski successive minima are not far from each other, cf. [9]).

On the other hand, our Theorems 1.1.3 and 1.1.4 give constructions of lattice extensions of a given sublattice within an ambient lattice with small determinant and successive minima, respectively, while preserving the geometric properties of the sublattice that is being extended.

Chapter III

On Planar Lattices

3.1 Farey Fractions

In this section we focus on the 2-dimensional case of the problem considered in Theorem 1.1.1. Given a primitive vector in $(a, b) \in \mathbb{Z}^n$, in how many ways can it be extended to a basis of $\mathbb{Z}^2$ by a vector $(\boldsymbol{z}_1, \boldsymbol{z}_2)$ of sup-norm $\leq T$? This is equivalent to counting the number of integer solutions to

$$a\boldsymbol{z}_2 - b\boldsymbol{z}_1 = \pm 1 \tag{3.1}$$

with $|\boldsymbol{z}_1|, |\boldsymbol{z}_2| \leq T$, i.e. the number of points in $\mathbb{B}_A(T)$ where $A = (ab)$. Applying Corollary 2.2.2, we have

$$\frac{4T}{|A|} - 6 \leq |\mathbb{B}_A(T)| \leq \frac{4T}{|A|} + 6, \tag{3.2}$$

where $|A| = \max|a|, |b|$. Here we do not prove any new results, but instead show a connection of this problem to Farey fractions and Diophantine approximation. The set of rational numbers in the interval $[0, 1]$ can be organized into Farey series as follows. For each $n \geq 1$, let $\mathcal{F}_n$ be the set of all rationals $a/b \in [0, 1]$ with $\gcd(a, b) = 1$ and $b \leq n$ written in ascending order. For example,

$$\mathcal{F}_5 = \left\{ \frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5} \right\}.$$

The set $\mathcal{F}_n$ is called the Farey series of order n. The set $\mathbb{Q} \cap [0, 1]$ can then be viewed as the limit of $\mathcal{F}_n$ as $n \to \infty$, and this interpretation induces one possible enumeration on $\mathbb{Q} \cap [0, 1]$. A good source of information on Farey series is Chapter 3 of Hardy and Wright's classical book [39]. On the other hand, reduced fractions correspond to primitive integer points in the plane. Let

$$\mathbb{Z}^2_{pr} = \{(x, y) \in \mathbb{Z}^2 : \gcd(x, y) = 1\}.$$

Elements of this set are precisely primitive vectors in $\mathbb{Z}^2$, sometimes also called visible lattice points, the second name alluding to the property that the line segment connecting (x, y) to the origin contains no other lattice points on it, so (x, y) is not obstructed by anything, hence visible from the origin. If a pair of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{Z}^n$ forms a basis for the lattice $\mathbb{Z}^2$, then they both must be contained in $\mathbb{Z}^2_{pr}$ (we routinely identify vectors with their endpoints).

Lemma 3.1.1. *Let $\boldsymbol{x}_1 = (a, b)$ and $\boldsymbol{x}_2 = (c, d)$ be in $\mathbb{Z}^2_{pr}$ and let $n = \max\{b, d\}$. Then $\boldsymbol{x}_1, \boldsymbol{x}_2$ form a basis for $\mathbb{Z}^2$ if and only if $\frac{a}{b}$ and $\frac{c}{d}$ are consecutive elements in the Farey series $\mathcal{F}_n$; we call such elements Farey neighbors.*

Proof. First notice that $\boldsymbol{x}_1, \boldsymbol{x}_2$ form a basis for $\mathbb{Z}^n$ if and only if

$$\left| \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} \right| = |ad - bc| = 1.$$

Now, suppose that $\frac{a}{b}$ and $\frac{c}{d}$ are Farey neighbors in the Farey series $\mathcal{F}_n$. Then Theorem 28 of [39] guarantees that

$$bc - ad = 1,$$

and so $\boldsymbol{x}_1, \boldsymbol{x}_2$ are a basis for $\mathbb{Z}^2$. In the reverse direction, assume $\boldsymbol{x}_1, \boldsymbol{x}_2$ are a basis for $\mathbb{Z}^2$. Assume without loss of generality that $\frac{a}{b} < \frac{c}{d}$. Then $\frac{a}{b}, \frac{c}{d} \in \mathcal{F}_n$, and we only need to prove that there does not exist some $\frac{h}{k} \in \mathcal{F}_n$ such that

$$\frac{a}{b} < \frac{h}{k} < \frac{c}{d}.$$

Let P be the parallelogram spanned by the vectors $bx_1, \boldsymbol{x}_2$, then the vertices of P are $(0, 0), (a, b), (c, d), (a + c, b + d)$ and the area of P is the determinant $bc - ad = 1$. Further, P does not contain any integer lattice points in its interior, in particular $(a + c, b + d)$ is also a primitive lattice point. But since $b + d > n$, a primitive point

(h, k) satisfying 3.1 would have to be in the interior of P, hence such a point cannot exist. This proves the lemma. □

Let

$$C(T) = \left\{ \boldsymbol{z} \in \mathbb{Z}^2_{pr} : |\boldsymbol{z}|_\infty \leq T \right\}.$$

We can subdivide $C(T)$ into eight pieces $Q_i^{\pm}(T)$, where $1 \leq i \leq 4$ indicates a quadrant (numbered in the counterclockwise order) and $\pm$ indicates whether the region is above or below the corresponding line $\boldsymbol{y} = \pm\boldsymbol{x}$. For instance,

$$Q_1^+(T) = \{\boldsymbol{z} \in C(T) : 0 \leq \boldsymbol{z}_1 \leq \boldsymbol{z}_2\}, Q_2^-(T) = \{\boldsymbol{z} \in C(T) : 0 \leq \boldsymbol{z}_2 \leq -\boldsymbol{z}_1\}.$$

These pieces have equal cardinality, since they can be obtained from each other by an appropriate reflection. For instance

$$-Q_1^+(T) = \{-\boldsymbol{z} \in C(T) : 0 \leq \boldsymbol{z}_1 \leq \boldsymbol{z}_2\} = \{\boldsymbol{z} \in C(T) : 0 \leq -\boldsymbol{z}_1 \leq -\boldsymbol{z}_2\} = Q_3^-(T).$$

It is then easy to see that if some $(a, b) \in Q_i^{\pm}(T)$, then all the corresponding vectors extending (a, b) to a basis of $\mathbb{Z}^2$ are contained in $\pm Q_i^{\pm}(T)$: this follows from (3.1). In other words, $\mathbb{B}_A(T) \subseteq Q_i^{\pm}(T) \cup -Q_i^{\pm}(T)$, and $|\mathbb{B}_A(T) \cap Q_i^{\pm}(T)| = |\mathbb{B}_A(T) \cap -Q_i^{\pm}(T)|$. Further, these cardinalities do not depend on which $Q_i^{\pm}(T)$ the vector (a, b) belongs to. Hence we can assume that $(a, b) \in Q_1^+(T)$, so $|\mathbb{B}_A(T)| = 2|\mathbb{B}_A(T) \cap Q_1^+(T)|$. Then the fraction $\frac{a}{b}$ belongs to the Farey series $\mathcal{F}_n$ for every $n \geq b$. Further, in this case (assuming $T \geq b$)

$$|\mathbb{B}_A(T) \cap Q_1^+(T)| = |\{c/d \in \mathcal{F}_n : b \leq n \leq T, a/b \text{ and } c/d \text{ are neighbors in } \mathcal{F}_n\}|.$$

Assume that $\frac{a}{b}$ and $\frac{c}{d}$ are neighbors in some $\mathcal{F}_n$, then $n < b + d$ (Theorem 30 of [39]) and the next neighbor that will "squeeze in" between $\frac{a}{b}, \frac{c}{d}$ will be $\frac{a+c}{b+d}$ (Theorem 29

of [39]). When $T \gg b$, new neighbors will appear every time n grows by another b, and on this interval in n, say $(k-1)b \leq n \leq kb$ for some k, $\frac{a}{b}$ will acquire two new neighbors: on the left and on the right. This means that

$$|\mathbb{B}_A(T) \cap Q_1^+(T)| \sim \frac{2T}{b},$$

and hence $|\mathbb{B}_A(T)| \sim \frac{4T}{b}$ as $T \to \infty$. Since $a \leq b = |A|$, this agrees with (3.2), and also implies that the number of Farey neighbors of a given Farey fraction grows linearly with the denominator.

Farey fractions are also related to Diophantine approximations. Dirichlet's approximation theorem guarantees that for any irrational $\alpha \in \mathbb{R}$ there exist infinitely many primitive points $(p, q) \in \mathbb{Z}^2$ such that

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{q^2}. \tag{3.3}$$

Let

$$\mathcal{D}_n(\alpha) = \{p/q \in Q : p/q \text{ satisfies (3.3)}, q \leq n\}$$

be the set of all Dirichlet approximations to α with denominator no bigger than n. Farey fractions provide another method of approximating irrational numbers in the interval $(0, 1)$. Let $0 < \alpha < 1$ be irrational, and define the sequence of Farey approximations for α in the following manner: $F_0(\alpha) = \frac{0}{1}, F_1(\alpha) = \frac{1}{1}$ and for each $k \geq 2, F_k(\alpha) = \frac{a+c}{b+d}$, where

$$\frac{a}{b} = \min_{0 \leq j < k} \{F_j(\alpha) : F_j(\alpha) > \alpha\}, \frac{c}{d} = \max_{0 \leq j < k} \{F_j(\alpha) : F_j(\alpha) < \alpha\}, \tag{3.4}$$

and $\gcd(a, b) = \gcd(c, d) = 1$. Define

$$\mathcal{F}_n(\alpha) = \mathcal{F}_n \cap \{F_k(\alpha)\}_{k=0}^n \tag{3.5}$$

to be the set of all Farey approximations to α with denominator less than or equal to n. An element $\frac{a}{b}$ of $\mathcal{F}_n(\alpha)$ is not guaranteed to satisfy (3.3), but is the best upper or lower approximation to α with denominator $\leq b$. Moreover, if $\frac{c}{d} \in \mathcal{F}_n$ is not a Farey approximation, then there exists $\frac{a}{b} \in \mathcal{F}_n$ such that either $\frac{c}{d} < \frac{a}{b} < \alpha$ or $\alpha < \frac{a}{b} < \frac{c}{d}$. Since $b, d \leq n$,

$$\left|\alpha - \frac{a}{b}\right| > \left|\frac{c}{d} - \frac{a}{b}\right| \geq \frac{1}{n^2}. \tag{3.6}$$

Therefore

$$\mathcal{D}_n(\alpha) \subseteq \mathcal{F}_n(\alpha). \tag{3.7}$$

Now, let $\alpha = [a_0; a_1, a_2, ...]$ be the continued fraction expansion for α, and for each $n \geq 1$ let $\alpha_n = [a0; a1, a2, ..., an]$ be its n-th convergent. It is well known that

$$\alpha_k{}_{k=1}^{n} \subseteq \mathcal{D}_n(\alpha), \tag{3.8}$$

and the convergents alternate in the following sense: $\alpha_{k-1} < \alpha \Leftrightarrow \alpha_k > \alpha$. We can now prove that, unlike the number of Farey neighbors of a given Farey fraction, the number of Farey approximations of a given irrational number grows less than linearly with the denominator.

Lemma 3.1.2. *Let* $0 < \alpha < 1, \alpha \notin \mathbb{Q}$. *Then*

$$\lim_{n\to\infty} \frac{|Fn(\alpha)|}{n} = 0.$$

Proof. Let d_k be the denominator of $F_k(\alpha)$ expressed in lowest terms, where $d_1 = 1, d_2 = 1$ corresponding to $\frac{0}{1}, \frac{1}{1} \in \mathcal{F}_1(\alpha)$, respectively. Define $a_1 = 1, b_1 = 1$, then $d_{k+1} = a_k + b_k$, and

$$a_{k+1} = d_{k+1}, b_{k+1} = b_k, \tag{3.9}$$

if $F_k(\alpha) > \alpha$, or

$$a_{k+1} = a_k, b_{k+1} = d_{k+1}, \tag{3.10}$$

if $F_k(\alpha) < \alpha$, where a_k is the denominator of $\min_{0\leq j<k} |\{F_j(\alpha) : F_j(\alpha) > \alpha\}$ and b_k is the denominator of $\max_{0\leq j<k}\{F_j(\alpha) : F_j(\alpha) < \alpha\}$. Then observe that, with the exception of the first and second terms, the sequence $\{d_k\}$ is strictly increasing and the sequences $\{a_k\}, \{b_k\}$ are non-decreasing. Observe also that

$$|\mathcal{F}_n(\alpha)| = |\{d_k : d_k \leq n\}| = \max\{k : d_k \leq n\}. \tag{3.11}$$

For a fixed $N \in \mathbb{Z}_{>0}$, let $l = |\{d_k : d_k \leq N\}|$ and notice that for $k > l$

$$d_k = a_{k-1} + b_{k-1} \geq d_{k-1} + \min\{a_{k-1}, b_{k-1}\} \geq d_l + (k-l)\min\{a_l, b_l\}. \tag{3.12}$$

This implies

$$\lim_{n\to\infty} \frac{|\{d_k : d_k \leq n\}|}{n} = \lim_{n\to\infty} \left(\frac{|\{d_k : d_k \leq N\}|}{n} + \frac{|\{d_k : N < d_k \leq n\}|}{n} \right) \tag{3.13}$$

$$\leq \lim_{n\to\infty} \left(\frac{l}{n} + \frac{\frac{n-N}{\min\{a_l,b_l\}}}{n} \right) = \frac{1}{\min\{a_l, b_l\}}. \tag{3.14}$$

It remains to show that $\{a_k\}$, $\{b_k\}$ are unbounded. Observe that a_k increases whenever $F_k(\alpha) > \alpha$ and b_k increases whenever $F_k(\alpha) < \alpha$, so we must show that $F_k(\alpha)$ "switches sides" sufficiently often. By (3.8) and (3.7) we know that there exists a sub-sequence $\{k_j\}$ such that $\alpha_j = F_{k_j}(\alpha)$. Now,

$$F_{k_j}(\alpha) > \alpha \Rightarrow F_{k_{j+1}}(\alpha) < \alpha \Rightarrow F_{k_{j+2}}(\alpha) > \alpha, \tag{3.15}$$

since the continued fraction convergents alternate. By the recurrence relation on a_k

and b_k, if $F_{k_j}(\alpha) > \alpha$ then

$$b_{k_{j+1}} \geq (k_{j+1} - k_j)a_{k_j} + b_{k_j} \geq f_j, \tag{3.16}$$

where f_j is the j-th Fibonacci number. Likewise, if $F_{k_j}(\alpha) < \alpha$ then

$$a_{k_{j+1}} \geq (k_{j+1} - k_j)b_{k_j} + a_{k_j} \geq f_j. \tag{3.17}$$

Combining these observations with (3.13), we conclude that

$$\lim_{n\to\infty} \frac{\max\{k : d_k \leq n\}}{n} \leq \frac{1}{\min(a_{k_j}, b_{k_j})} \leq \frac{1}{f_j} \tag{3.18}$$

for any $l > 0$, and therefore

$$\lim_{n\to\infty} \frac{\max\{j : d_j \leq n\}}{n} = 0. \tag{3.19}$$

This completes the proof of the lemma. □

3.2 Deep Holes of Planar Lattices

We start this section with the following simple but useful technical lemma.

Lemma 3.2.1. *Consider the m-dimensional simplex with vertices $\mathbf{0}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ in $\mathbb{R}^n, m \leq n$. There exist points $\boldsymbol{z} \in \mathbb{R}^n$ so that $\|\boldsymbol{z}\|_\infty = \|\boldsymbol{z} - \boldsymbol{x}_i\|_\infty$ for all $1 \leq i \leq n$, and these points are solutions to*

$$\begin{pmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_m^T \end{pmatrix} \boldsymbol{z} = \frac{1}{2} \begin{pmatrix} \|\boldsymbol{x}_1\|_\infty^2 \\ \vdots \\ \|\boldsymbol{x}_m\|_\infty^2 \end{pmatrix}. \tag{3.20}$$

Proof. If $\boldsymbol{z}$ is equidistant from $\boldsymbol{x}_i$ and $\mathbf{0}$ then $\boldsymbol{z}$ lies in the hyperplane orthogonal to

$\boldsymbol{x}_i$ that passes through the point $(\boldsymbol{x}_i - \mathbf{0})/2$. That is

$$\boldsymbol{z} \cdot \boldsymbol{x}_i = \mathrm{proj}_{\boldsymbol{x}_i}(\boldsymbol{z}) \cdot \boldsymbol{x} = \frac{\|\boldsymbol{x}_i\|_\infty^2}{2}.$$

Since this is true for each $1 \leq i \leq n$ this gives the linear system in (3.2). □

Our main goal here is to describe some properties of the deep holes of lattices, focusing especially on the two-dimensional case. Our first basic observation is that if $\boldsymbol{z}$ is a deep hole of a lattice $\Lambda \subset \mathbb{R}^n$, then so is $-\boldsymbol{z}$: this follows by the fact that $-\Lambda = \Lambda$, since lattices are symmetric about the origin. Further, we have the following observation.

Lemma 3.2.2. *Let $\Lambda \subset \mathbb{R}^2$ be a lattice of rank 2 with minimal basis $\boldsymbol{x}, \boldsymbol{y}$ and angle $\theta \in [\pi/3, \pi/2]$ between these basis vectors. Write λ_1, λ_2 for the successive minima of Λ, so that $0 < \lambda_1 = \|x\|_\infty \leq \lambda_2 = \|y\|_\infty$. Then the fundamental parallelogram*

$$\mathcal{P} = \{s\boldsymbol{x} + t\boldsymbol{y} : 0 \leq s, t < 1\} \tag{3.21}$$

contains two deep holes $\boldsymbol{z}_1, \boldsymbol{z}_2$ and $bz_1 + \boldsymbol{z}_2 \in \Lambda$. If the angle $\theta = \pi/2$, then $z_1 = z_2$ is the center of $\mathcal{P}$, and we say that this deep hole has multiplicity 2.

Proof. Let us label the vertices of $\mathcal{P}$ as follows: O for the origin, X for the endpoint of the vector $\boldsymbol{x}$, Y for the endpoint of the vector $\boldsymbol{y}$, and Q for the endpoint of the vector $\boldsymbol{x} + \boldsymbol{y}$. The parallelogram $\mathcal{P}$ can be split into two congruent triangles: OXY and QYX. Then the endpoints of the deep holes of Λ contained in $\mathcal{P}$ are the centers of the circles circumscribed around these two triangles, call them Z_1 and Z_2, respectively, and let $\boldsymbol{z}_1, \boldsymbol{z}_2$ be vectors with the endpoints Z_1, Z_2 (see [43]). The two triangles are symmetric to each other about the center C of $\mathcal{P}$, which means that reflection with respect to C maps the line segment OZ_1 onto the line segment QZ_2. This means that OZ_1QZ_2 is a parallelogram with OQ as its longer diagonal, and hence

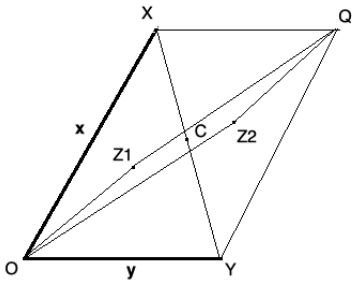

Figure 3.1: Fundamental parallelogram $\mathcal{P}$ of Λ with deep holes Z_1 and Z_2.

the corresponding vector is the sum $\boldsymbol{z}_1 + \boldsymbol{z}_2$. Since its endpoint is Q, a vertex of $\mathcal{P}$, this vector is in Λ (see Figure 3.1 for a graphical illustration of this argument). If $\theta = \pi/2$, then the deep hole of each of the triangles is in the center of the hypotenuse of its corresponding right triangle, i.e. at the center point C of $\mathcal{P}$; in this case, the two deep holes coincide, so $Z_1 = Z_2 = C$. □

An immediate implication of Lemma 3.2.2 is that deep holes z_1, z_2 are each other's inverses in the additive abelian group $\mathbb{R}^2/\Lambda$. Further, $\boldsymbol{z}_1 = \boldsymbol{z}_2$ is an element of order two in this group if and only if the angle $\theta = \pi/2$. On the other hand, z_1, z_2 can be elements of finite order in other situations too. For instance, in the hexagonal lattice

$$L_{\pi/3} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}$$

the deep holes are $\boldsymbol{z}_1 = (1/2, 1/(2\sqrt{3})), \boldsymbol{z}_2 = (1, 1/\sqrt{3})$ have order three in the group $\mathbb{R}^2/L_{\pi/3}$, while the lattice

$$L' = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \sqrt{3} \end{pmatrix} \mathbb{Z}^2$$

has deep a hole $\boldsymbol{z}_1 = (1/2, 11\sqrt{3}/24)$ satisfying the condition

$$48\boldsymbol{z}_1 = 13(1,0) + 22(1/2, \sqrt{3}) \in L',$$

which makes $\boldsymbol{z}_1$ an element of order dividing 48 in the group $\mathbb{R}^2/L'$. These observations raise a natural question: when does a deep hole of $\Lambda \subset \mathbb{R}^2$ have finite order as an element of the group $\mathbb{R}^2/\Lambda$?

Theorem 3.2.3. *Let $\Lambda \subset \mathbb{R}^2$ be a full-rank lattice with successive minima λ_1, λ_2 and corresponding minimal basis vectors $\boldsymbol{x}_1, \boldsymbol{x}_2$. A deep hole $\boldsymbol{z}$ of Λ has finite order in the group $\mathbb{R}^2/\Lambda$ if and only if Λ is orthogonal or there exist rational numbers p, q so that $p\lambda_1^2 = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = q\lambda_2^2$. Moreover, if $\lambda_1^2, \lambda_2^2, \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 \in \mathbb{Z}$ then the order of $\boldsymbol{z}$ in $\mathbb{R}^2/\Lambda$ is less than or equal to $12\sqrt{3}\lambda_2^4$.*

Proof. As we discussed above, if Λ is orthogonal then the deep hole always has order 2 in $\mathbb{R}^2/\Lambda$, hence we assume Λ is not orthogonal. Further, we can assume that the minimal basis vectors $\boldsymbol{x}_1, \boldsymbol{x}_2$ are chosen so that the angle θ between them satisfies $\pi/3 \leq \theta \leq \pi/2$. If $\boldsymbol{z}$ is the equidistant from $\boldsymbol{x}_1, \boldsymbol{x}_2$ and the origin then $\boldsymbol{z}$ is a deep hole of Λ and is contained in the convex hull of $\{0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$. By Lemma 3.2.2,

$$\boldsymbol{z} \cdot \boldsymbol{x}_1 = \frac{\lambda_1^2}{2}, \boldsymbol{z} \cdot \boldsymbol{x}_2 = \frac{\lambda_2^2}{2}. \tag{3.22}$$

Now suppose that $\boldsymbol{z}$ has finite order in $\mathbb{R}^2/\Lambda$. Then there integers a, b, c so that $c \neq 0$ and

$$a\boldsymbol{x}_1 + b\boldsymbol{x}_2 = c\boldsymbol{z}.$$

In fact, the pairs $\boldsymbol{z}, \boldsymbol{x}_1$ and $\boldsymbol{z}, \boldsymbol{x}_2$ are linearly independent so a, b, c are all nonzero. Taking scalar products of both sides of this equation with $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and applying (3.22), we obtain

$$a\lambda_1^2 + b\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 2c\lambda_1^2$$
$$a\lambda_2^2 + a\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 2c\lambda_2^2.$$

Notice that since Λ is not orthogonal, $\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 \neq 0$ and

$$x_1 \cdot x_2 = \frac{2c-a}{b}\lambda_1^2 = \frac{2c-b}{a}\lambda_2^2. \tag{3.23}$$

Now suppose that there are there are rational numbers p, q so that

$$p\lambda_1^2 = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = q\lambda_2^2.$$

Then, by (3.22), there exist rational, and hence integer solutions a, b, c to the linear system

$$\begin{cases} a\boldsymbol{x}_1 \cdot \boldsymbol{x}_1 + b\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 + c\boldsymbol{x}_1 \cdot \boldsymbol{z} = 0 \\ a\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 + b\boldsymbol{x}_2 \cdot \boldsymbol{x}_2 + c\boldsymbol{x}_2 \cdot \boldsymbol{z} = 0, \end{cases} \tag{3.24}$$

which factors as

$$\begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{0} \tag{3.25}$$

Since the matrix $\begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \end{pmatrix}$ is of full rank, $(a, b, c)^T$ solves (3.25) if and only if it solves

$$\begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{0} \tag{3.26}$$

On the other hand, $(a, b, c)^T$ being an integer solution of (3.26) is equivalent to $\boldsymbol{z}$ having finite order in $\mathbb{R}^2/\Lambda$. In fact, the order of $\boldsymbol{z}$ in $\mathbb{R}^2/\Lambda$ is $\leq |c|_\infty$. By combining

(3.25) with (3.23), we obtain the linear system

$$\begin{pmatrix} 2\lambda_1^2 & 2\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 \lambda_1^2 & \\ 2\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 & 2\lambda_2^2 & \lambda_2^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{0} \tag{3.27}$$

If $\lambda_1^2, \lambda_2^2, \boldsymbol{x}_1 \cdot \boldsymbol{x}_2 \in \mathbb{Z}$, then by Siegel's lemma (1.14) there exists a nontrivial integer solution to this system with

$$\max\{|a|_\infty, |b|_\infty, |c|_\infty\} \leq 3\sqrt{3}(\max\left\{2\lambda_1^2, 2\lambda_2^2, |\boldsymbol{x}_1 \cdot \boldsymbol{x}_2|_\infty\right\} = 12\sqrt{3}\lambda_2^4,$$

since $\lambda_1^2 \leq |\boldsymbol{x}_1 \cdot \boldsymbol{x}_2|_\infty \leq \lambda_2^2$. □

3.3 Covering radius of planar lattices

In this section we investigate the covering radii of lattices in the plane, in particular proving Theorem 1.3 and its corollaries. Let $\Lambda \subset \mathbb{R}^2$ be a lattice of rank 2 with minimal basis $\boldsymbol{x}, \boldsymbol{y}$ and angle $\theta \in [\pi/3, \pi/2]$ between these basis vectors. Then the successive minima of Λ are

$$0 < \lambda_1 = |\boldsymbol{x}|_\infty \leq \lambda_2 = |\boldsymbol{y}|_\infty. \tag{3.28}$$

See, for instance, [27] for the details on the existence of such a minimal basis.

Lemma 3.3.1. *The covering radius of* Λ *is*

$$\mu = \frac{\sqrt{\lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2\cos\theta}}{2\sin\theta} \tag{3.29}$$

Proof. The vectors $\boldsymbol{x}, \boldsymbol{y}$ correspond to successive minima in Λ, and hence form a reduced basis. Then Theorem 3.2 of [43] asserts that the covering radius of Λ is equal

to the circumradius of the triangle with sides corresponding to the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. The length of the third side of this triangle is

$$\sqrt{\lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2 \cos\theta} \tag{3.30}$$

and the area of this triangle is

$$A = \frac{1}{2}\lambda_1\lambda_2 \sin\theta. \tag{3.31}$$

Now, the circumradius of a triangle with sides a, b, c and area A is given by the formula

$$R = \frac{abc}{4A}. \tag{3.32}$$

Putting together (3.30), (3.31) and (3.32) produces (3.29) □

The similarity classes of WR lattices in the plane are parameterized by the angle $\theta \in [\pi/3, \pi/2]$, and each similarity class is represented by

$$L_\theta = \begin{pmatrix} 1 & \cos\theta \\ 0 & \sin\theta \end{pmatrix} \mathbb{Z}^2.$$

The following corollary follows immediately from Lemma 3.3.1 by substituting $\lambda_1 = \lambda_2 = 1$ into (3.29).

Corollary 3.3.2. *The covering radius $\mu = \mu(\theta)$ of the lattices L_θ is a continuous function on the interval $[\pi/3, \pi/2]$, given by*

$$\mu(\theta) = \frac{\sqrt{1 - \cos\theta}}{\sqrt{2}\sin\theta}.$$

The endpoints of the interval are represented by the hexagonal lattice and the square lattice $\mathbb{Z}^2$ with the covering radii $1/\sqrt{3}$ and $1/\sqrt{2}$, respectively.

We are now ready to prove Theorem 1.2.1. We first want to build an extension $E_1 \subset \Lambda \subset \mathbb{R}^2$ with rank $(\Lambda) = 2$ so that $\mu(\Lambda) = \mu(E_1)$. Our argument characterizes all possible such extensions, showing that they must be rectangular lattices, i.e. lattices containing an orthogonal basis.

Proof of Theorem 1.2.1. First notice that each $\Lambda(\alpha)$ as in (1.4) is a rectangular lattice, thus its successive minima are

$$\lambda_{1,2} = \sqrt{\alpha}, \sqrt{1-\alpha},$$

i.e. norms of the orthogonal basis vectors given in (1.4). By Lemma 3.3.1, the covering of $\Lambda(\alpha)$ is

$$\mu = \frac{\sqrt{\alpha + (1-\alpha) - 2\sqrt{\alpha(1-\alpha)cos(\pi/2)}}}{2\sin(\pi/2)} = \frac{1}{2}.$$

In the reverse direction, assume $\Lambda \subset \mathbb{R}^2$ is a full rank lattice so that $e_1 \in \Lambda$ and $\mu(\Lambda) = 1/2$. The vector $a := \frac{1}{2}e_1$ is a deep hole of E_1. First we show that a is a deep hole of the lattice Λ as well. Suppose not, then there exists a point $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \Lambda$ such that

$$\|x - a\|_\infty < 1/2.$$

Then the vector $\boldsymbol{z} = e_1 - \boldsymbol{x} = \begin{pmatrix} 1 - \boldsymbol{x}_1 \\ -x_2 \end{pmatrix}$ is also in Λ, and

$$\|\boldsymbol{z} - \boldsymbol{a}\|_\infty = \|\boldsymbol{x} - \boldsymbol{a}\|_\infty < 1/2.$$

The vectors $\boldsymbol{x}, \boldsymbol{z}$ form a basis for Λ, and the triangle with sides corresponding to these vectors is contained in the interior of the circle of radius $1/2$ with center at $\boldsymbol{a}$, thus the circumradius R of this triangle is $< 1/2$. On the other hand, by Theorem 3.2 of [43] the covering radius of Λ is equal to the circumradius of the triangle with sides

corresponding to the shortest basis vectors, which has to be $\leq R$. Hence $\mu(\Lambda) < 1/2$, which is a contradiction, so a is a deep hole of the lattice Λ. This means that there exists a basis $\boldsymbol{x}, \boldsymbol{z} \in \Lambda$ with $|\boldsymbol{x}|_\infty = \lambda_1, |\boldsymbol{z}|_\infty = \lambda_2$ so that the point a is the center of the circle circumscribed around the triangle with sides $\boldsymbol{x}, \boldsymbol{z}$, meaning in particular that

$$(1/2)^2 = \|\boldsymbol{x} - \boldsymbol{a}\|_\infty^2 = (x_1 - 1/2)^2 2 + x_2^2 = x_1^2 + x_2^2 - x_1 + (1/2)^2. \tag{3.33}$$

Also, $2\boldsymbol{a} = \boldsymbol{e}_1$ is a diagonal of the fundamental parallelogram of Λ spanned by $\boldsymbol{x}, \boldsymbol{z}$, meaning that

$$\boldsymbol{x} + \boldsymbol{z} = \boldsymbol{e}_1.$$

Hence $\boldsymbol{z} = \begin{pmatrix} 1 - x_1 \\ -x_2 \end{pmatrix}$, and

$$\cos\theta = \frac{\boldsymbol{x} \cdot \boldsymbol{z}}{\|x\|_\infty \|z\|_\infty} = \frac{x_1 - x_1^2 - x_2^2}{\lambda_1 \lambda_2}, \tag{3.34}$$

where θ is the angle between $\boldsymbol{x}$ and $\boldsymbol{z}$, which lies in the interval $[\pi/3, \pi/2]$. Hence $\cos\theta = 0$ by (3.33). Letting $x_1 = \alpha$, we obtain

$$x_2 = \sqrt{\alpha - \alpha^2},$$

and (1.4) follows by replacing $\boldsymbol{z}$ with $-\boldsymbol{z}$. Next, suppose that $L \subset \mathbb{R}^n$ be a lattice of rank 1 and let $\boldsymbol{u} \in \mathbb{R}^n$ is such that $L = \mathbb{Z}\boldsymbol{u}$, so the covering radius of L is $\mu = \|u\|_\infty/2$. Let Λ be a lattice of rank 2 in $\mathbb{R}^n$ containing L and let $V = \operatorname{span}_{\mathbb{R}} \Lambda$ be the 2-dimensional subspace spanned by this lattice. Applying a suitable isometry of $\mathbb{R}^n$, we can identify V with $C_2 := \{\boldsymbol{x} \in \mathbb{R}^n : x_i = 0 \forall 2 < i \leq n\}$. In fact, we can choose such an isometry τ so that $\boldsymbol{u}$ maps to $\beta \boldsymbol{e}_1$ for $\beta = \|\boldsymbol{u}\|_\infty$. Then $\Lambda' = \frac{1}{\beta}\tau(\Lambda)$ is a lattice isometric to $\frac{1}{\beta}\Lambda$ in $\mathbb{R}^n$, and Λ' contains $\boldsymbol{e}_1$. Identifying C_2 with $\mathbb{R}^2$ we see

that Theorem 1.2.1 implies that Λ' is an equal covering extension of $\mathbb{Z}\boldsymbol{e}_1$ in $\mathbb{R}^2$ if and only if it is of the form (1.4). Finally, notice that

$$\det(L) = \sqrt{\det(\boldsymbol{u}T\boldsymbol{u})} = \|\boldsymbol{u}\|_\infty = \beta.$$

This completes the proof of the theorem. □

Remark 3.3.1. An immediate implication of Theorem 1.2.1 is that the only well-rounded equal covering extension of E_1 is

$$\Lambda(1/2) = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \mathbb{Z}^2, \tag{3.35}$$

which is a square lattice in the plane containing $\mathbb{Z}^2$ as a sublattice of index 2. More generally, a rank-two equal covering extension $\Lambda \subset \mathbb{R}^n$ of a rank-one lattice $L \subset \Lambda$ is well-rounded if and only if it is isometric to $\det(L)\Lambda(1/2)$. As in (1.2) the set of all similarity classes of planar lattices is parameterized by

$$F = \{(a, b) \in \mathbb{R}^2 : 0 \leq a < 1/2, b > 0, a^2 + b^2 \geq 1\},$$

see Figure 3.2. The set of semi-stable classes (see Remark 2.3.1)in $\mathbb{R}^2$ contains the WR classes: from (2.24) it follows that a lattice Λ in $\mathbb{R}^2$ is semi-stable if and only if $\lambda_1(\Lambda) \geq \det(\Lambda)^{1/2}$ (see [28] for more details). Thus the only semi-stable equal covering extensions are also those similar to $\Lambda(1/2)$ as in (3.35), i.e. similar to $\mathbb{Z}^2$ as demonstrated in Figure 3.2. Let $K = \mathbb{Q}(\sqrt{D})$ be a quadratic number field with embeddings $\sigma_1, \sigma_2 : K \to \mathbb{C}$, and let $\Sigma_K : K \to \mathbb{R}^2$ be the Minkowski embedding of

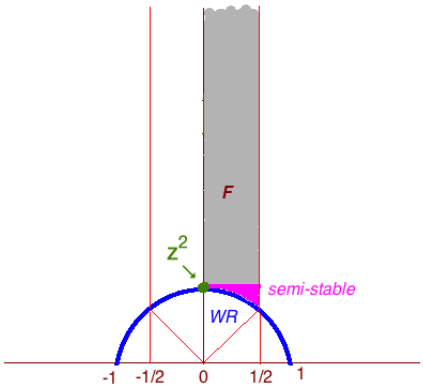

Figure 3.2: Similarity classes of planar lattices with $\mathbb{Z}^2$ representing the only equal covering extension class that is WR and semi-stable.

K, defined for every $\boldsymbol{x} \in K$ as

$$\Sigma_K(x) = \begin{pmatrix} \sigma_1(x) \\ \sigma_2(x) \end{pmatrix},$$

if the squarefree integer $D > 0$ and

$$\Sigma_K(x) = \begin{pmatrix} \Re(\sigma_1(x)) \\ \Im(\sigma_1(x)) \end{pmatrix},$$

if $D < 0$. We write Ω_K for the planar lattice $\Sigma_K(\mathcal{O}_K)$, where $\mathcal{O}_K$ is the ring of integers of the number field K.

Corollary 3.3.3. *Assume $D \not\equiv 1 (\mathrm{mod}\, 4)$. Then Ω_K is an equal covering extension of the rank-one lattice $\mathbb{Z}\Sigma_K(1 + \sqrt{D})$.*

Proof. If $D \not\equiv 1 (\mathrm{mod}\, 4)$ then $\mathcal{O}_K = \mathbb{Z}[\sqrt{D}]$, and so

$$\Omega_K = \begin{pmatrix} 1 & \sqrt{D} \\ 1 & -\sqrt{D} \end{pmatrix} \mathbb{Z}^2 \text{ if } D > 0, \Omega_K = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{|D|} \end{pmatrix} \mathbb{Z}^2 \text{ if } D < 0.$$

In either case, the lattice Ω_K is rectangular. If $D > 0$, then $\lambda_1 = \sqrt{2}$, $\lambda_2 = \sqrt{2D}$, and so (3.29) implies that $\mu(\Omega_K) = \sqrt{D+1}/\sqrt{2}$, while

$$\Sigma_K(1+\sqrt{D}) = \begin{pmatrix} 1+\sqrt{D} \\ 1-\sqrt{D} \end{pmatrix},$$

and so $\mu\left(\mathbb{Z}\Sigma_K(1+\sqrt{D})\right) = \sqrt{D+1}/\sqrt{2}$. If $D < 0$, then $\lambda_1 = 1, \lambda_2 = \sqrt{|D|}$, and so (3.29) implies that $\mu(\Omega_K) = \sqrt{|D|+1}/2$, while

$$\Sigma_K(1+\sqrt{D}) = \begin{pmatrix} 1 \\ \sqrt{|D|} \end{pmatrix},$$

and so $\mu\left(\mathbb{Z}\Sigma_K(1+\sqrt{D})\right) = \sqrt{|D|+1}/2$. □

Corollary 3.3.4. *Assume $D \equiv 1(mod 4)$, then Ω_K is not an equal covering extension of any rank-one lattice.*

Proof. If $D \equiv 1(\mathrm{mod}\, 4)$ then $\mathcal{O}_K = \mathbb{Z}\left[\frac{1+\sqrt{D}}{2}\right]$, and so

$$\Omega_K = \begin{pmatrix} 1 & \frac{1+\sqrt{D}}{2} \\ 1 & \frac{1-\sqrt{D}}{2} \end{pmatrix} \mathbb{Z}^2 \text{ if } D > 0, \Omega_K = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{|D|}}{2} \end{pmatrix} \mathbb{Z}^2 \text{ if } D < 0. \tag{3.36}$$

In both cases it is not difficult to check that Ω_K does not have an orthogonal basis, and hence cannot be similar to a lattice of the form $\Lambda(\alpha)$ as in (1.4). The conclusion follows from Theorem 1.2.1. □

Finally, we discuss a construction of orthogonal equal covering extensions in any dimension.

Proof of Theorem 1.2.2. We will argue by induction on $k \geq 1$. Theorem 1.2.1 establishes the base of induction, so let $k \geq 2$. Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subset \mathbb{R}^n$ be an orthogonal

basis for Λ_k and let $e_{k+1} \in \mathbb{R}^n$ be a vector orthogonal to Λ_k. Let

$$P_k = \left\{ \sum_{i=1}^{k} a_i \boldsymbol{x}_i : a_i \in \{0,1\}, 1 \leq i \leq k \right\}$$

be the set of vertices of the fundamental parallelepiped spanned by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$. The circumcenter of this orthogonal parallelepiped is the point $\mathbb{Z} \in \mathbb{R}^n$ which is equidistant from the points of P_k by Lemma 3.2.1, and hence is a deep hole of Λ_k. Let

$$B_k = \{\boldsymbol{y} \in \operatorname{span}_{\mathbb{R}}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} : \|\boldsymbol{y} - \boldsymbol{z}\|_\infty = \mu(\Lambda_k)\}.$$

Let $\Lambda_{k-1} = \operatorname{span} \mathbb{Z}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1}\}$ and let

$$P_{k-1} = \left\{ \sum_{i=1}^{k-1} a_i \boldsymbol{x}_i : a_i \in 0, 1, 1 \leq i \leq k-1 \right\}.$$

Now define

$$B_{k-1} = B_k \cap \operatorname{span}_{\mathbb{R}} \Lambda_{k-1}. \tag{3.37}$$

By construction, $P_{k-1} \subset B_{k-1}$, while B_{k-1} is the surface of $(k-1)$-dimensional ball in a $(k-1)$-dimensional subspace and the points of P_{k-1} are elements of an orthogonal lattice in that subspace. Let $\boldsymbol{z}'$ be the orthogonal projection of $\boldsymbol{z}$ onto $\operatorname{span}_{\mathbb{R}}(\Lambda_{k-1})$. Since $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is an orthogonal set, $\boldsymbol{z}' = \boldsymbol{z} - \operatorname{proj}_{x_k}(\boldsymbol{z})$. Moreover, $\boldsymbol{z}$ is equidistant from $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$, so by Lemma 3.2.1

$$\begin{pmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_k^T \end{pmatrix} \boldsymbol{z} = \frac{1}{2} \begin{pmatrix} \|x_1\|_\infty \\ \vdots \|x_k\|_\infty \end{pmatrix},$$

and therefore

$$\begin{pmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_{k-1}^T \end{pmatrix} \boldsymbol{z}' = \frac{1}{2} \begin{pmatrix} \|x_1\|_\infty \\ \vdots \|x_{k-1}\|_\infty \end{pmatrix}.$$

Then $\mathbb{Z}'$ is equidistant from $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1}\}$ and Λ_{k-1} is an orthogonal lattice contained in the k-dimensional subspace $V = \text{span}_{\mathbb{R}} \Lambda_{k-1}, e_{k+1}$. By the induction hypothesis, there exists a rank k orthogonal lattice $\Lambda'_k \subset V$ so that $\Lambda_{k-1} \subset \Lambda'_k$ and $\boldsymbol{z}'$ is a deep hole of Λ'_k. Let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ be an orthogonal basis for Λ'_k so that $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ are equidistant from $\boldsymbol{z}'$. Since $\boldsymbol{z} = \boldsymbol{z}' - \text{proj}_{x_k}(\boldsymbol{z})$, and $\boldsymbol{x}_k$ is orthogonal to $V, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ are also equidistant from $\boldsymbol{z}$. Let $\Lambda_{k+1} = \text{span}_{\mathbb{Z}}\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{x}_k\}$. Then Λ_{k+1} is an orthogonal lattice that contains Λ_k and

$$\begin{pmatrix} \boldsymbol{y}_1^T \\ \vdots \\ \boldsymbol{y}_k^T \\ \boldsymbol{x}_k \end{pmatrix} \boldsymbol{z} = \frac{1}{2} \begin{pmatrix} \|\boldsymbol{y}_1\|_\infty \\ \vdots \\ \|\boldsymbol{y}_k\|_\infty \\ \|\boldsymbol{x}_k\|_\infty \end{pmatrix},$$

Thus $\boldsymbol{z}$ is a deep hole of Λ_k. □

Chapter IV

On Heights and Siegel's Lemma

4.1 Additional Properties of Heights

Let k be a number field and let U_1, U_2 be subspaces of k^N. Recall that from (1.11);

$$\big|\det\big(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2\ \cdots\ \boldsymbol{\xi}_N\big)\big| \le \prod_{n=1}^{N} \mathcal{H}(\boldsymbol{\xi}_n).$$

A natural question suggested by the inequality (1.11) is this: if $\mathcal{Z} \subseteq k^N$ is a subspace of dimension L, does there exist a basis for $\mathcal{Z}$ such that there is nearly equality in (1.11)? As explained in [8], an answer to this question is given by a result that is dual to Siegel's lemma. This was proved as [8, Theorem 8] using the Weil height on vectors. An analogous result for the Arakelov height was proved as [62, Theorem 2]. This later result asserts that there exists a basis $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ for $\mathcal{Z}$ such that

$$\prod_{\ell=1}^{L} \mathcal{H}\big(\boldsymbol{\eta}_\ell\big) \le \gamma_k(L)^{L/2} \mathcal{H}(\mathcal{Z}), \tag{4.1}$$

where $\gamma_k(L)^{1/2}$ is a positive constant that depends on the field k and the parameter L. The field constant $\gamma_k(L)^{1/2}$ is a generalization of Hermite's constant. It was first defined and used by J. Thunder in [61].

For each positive integer L and each archimedean place v of k we define positive real numbers

$$r_v(L) = \begin{cases} \pi^{-1/2}\{\Gamma(L/2+1)\}^{1/L} & \text{if } v \text{ is a real place} \\ (2\pi)^{-1/2}\{\Gamma(L+1)\}^{1/2L} & \text{if } v \text{ is a complex place,} \end{cases}$$

and then we define

$$c_k(L) = 2|\Delta_k|^{1/2d} \prod_{v|\infty} \{r_v(L)\}^{d_v/d},$$

where Δ_k is the discriminant of k. The constant $c_k(L)$ is the normalized Haar measure of a naturally occurring subset of the L-fold product of the adele ring $k_{\mathcal{A}}^L$, and appears

in applications of the geometry of numbers over $k_{\mathcal{A}}^L$. In particular, it leads to the simple upper bound for $\gamma_k(L)^{1/2}$ given by

$$\gamma_k(L)^{1/2} \le c_k(L). \tag{4.2}$$

A more complicated lower bound was proved in [61, Theorem 2]. And it was shown in [62, Theorem 3] that the constant $\gamma_k(L)^{L/2}$ which appears in (4.1) is best possible for the inequality (4.1). From the lower bound proved in [61, Theorem 2] it follows that for fixed $d = [k : \mathbb{Q}]$ and fixed L, the value of $\gamma_k(L)^{1/2}$ grows like a small positive power of the absolute discriminant $\big|\Delta_k\big|$. In particular, for fixed $d = [k : \mathbb{Q}]$ and fixed L, there are only finitely many fields k such that $\gamma_k(L)^{1/2}$ is less than or equal to a positive real number.

If we arrange the basis vectors $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_L$ in (4.1) so that

$$\mathcal{H}\big(\boldsymbol{\eta}_1\big) \le \mathcal{H}\big(\boldsymbol{\eta}_2\big) \le \cdots \le \mathcal{H}\big(\boldsymbol{\eta}_L\big),$$

we find that

$$\mathcal{H}\big(\boldsymbol{\eta}_1\big) \le \gamma_k(L)^{1/2}\mathcal{H}(\mathcal{Z})^{1/L}.$$

As the Arakelov height of a nonzero vector is always greater than or equal to 1, a trivial consequence of (4.1) is the inequality

$$\mathcal{H}\big(\boldsymbol{\eta}_L\big) \le \gamma_k(L)^{L/2}\mathcal{H}(\mathcal{Z}). \tag{4.3}$$

However, if one works over the algebraic closure $\overline{\mathbb{Q}}$, then it follows from the results of Roy and Thunder [53], [54], that the dependence on the number field k can be removed. Suppose, for example, that $\mathcal{Z} \subseteq \overline{\mathbb{Q}}^N$ is a $\overline{\mathbb{Q}}$-linear subspace of dimension L. As a basis for $\mathcal{Z}$ requires only finitely many algebraic numbers as coordinates, there exists a number field k and a collection of basis vectors for $\mathcal{Z}$ such that the basis

vectors belong to k^N. It follows that the Arakelov height of $\mathcal{Z}$ can be computed as before by using such a number field k. Because of the way we normalize the absolute values $|\ |_v$, where v is a place of k, it follows in a standard manner that the Arakelov height of $\mathcal{Z}$ does not depend on the choice of the field k. For such a subspace $\mathcal{Z} \subseteq \overline{\mathbb{Q}}^N$ and $\varepsilon > 0$, it follows from [54, Theorem 1] that there exists a basis

$$\{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_L\} \subseteq \mathcal{Z}, \tag{4.4}$$

such that

$$\prod_{\ell=1}^{L} \mathcal{H}(\boldsymbol{\zeta}_\ell) \leq \big(\exp(L(L-1)/4) + \varepsilon\big)\mathcal{H}(\mathcal{Z}).$$

Again the basis vectors (4.4) can be arranged so that

$$\mathcal{H}(\boldsymbol{\zeta}_1) \leq \mathcal{H}(\boldsymbol{\zeta}_2) \leq \cdots \leq \mathcal{H}(\boldsymbol{\zeta}_L).$$

Then it follows as before that

$$\mathcal{H}(\boldsymbol{\zeta}_1) \leq \big(\exp(L(L-1)/4) + \varepsilon\big)^{1/L}\mathcal{H}(\mathcal{Z})^{1/L},$$

and

$$\mathcal{H}(\boldsymbol{\zeta}_L) \leq \big(\exp(L(L-1)/4) + \varepsilon\big)\mathcal{H}(\mathcal{Z}).$$

4.2 Proof of Theorem 1.4.1

Throughout this section let $\mathbf{1}_M$ denote an $M \times M$ identity matrix for $M > 0$ and let $\mathbf{0}_{N\times M}$ denote an $N \times M$ matrix of all zeros for positive integers N, M. Additionally, let $[N] = \{1, \ldots, N\}$ and define:

$$\mathcal{J}(N, M) = \{I \subseteq [N] : |I| = M\}.$$

For an $N \times M$ matrix A with coefficients in the number field k and for each $I \in \mathcal{J}(N, M)$ define A_I to be the $M \times M$ minor of A whose rows are the rows of A indexed by I.

Lemma 4.2.1. *Let*

$$A = \begin{pmatrix} \mathbf{1}_{M-1} & \mathbf{0}_{(M-1)\times 1} \\ \mathbf{0}_{1\times(M-1)} & 1 \\ U & V \end{pmatrix} \tag{4.5}$$

be an $N \times M$ matrix with coefficients in k where U is a $(N-M) \times M$ matrix and V is an $(N-M) \times 1$ matrix over k. Let

$$A' = \begin{pmatrix} \mathbf{1}_{M-1} \\ \mathbf{0}_{1\times(M-1)} \\ U \end{pmatrix}. \tag{4.6}$$

Then

$$\mathcal{H}(A) \geq \mathcal{H}(A') \tag{4.7}$$

with equality if and only if $V = \boldsymbol{0}_{(N-M)\times 1}$.

Proof. We will first partition the Grassmann coordinates of A' into two sets. Define

$$R' = \{I' \in \mathcal{J}(N, M-1) : M \notin I'\}$$

and

$$T' = \mathcal{J}(N, M-1) \setminus R'.$$

Define

$$A'_{R'} = (\det A'_{I'})_{I' \in R'}$$

to be the vector of Grassmann coordinates of A' indexed by R'. Likewise define

$$A'_{T'} = (\det A'_{I'})_{I' \in T}$$

to be the vector of Grassmann coordinates of A' indexed by T'. Notice that the Mth row of A' is identically zero so $\det(A'_{I'}) = 0$ if $M \in I'$. Therefore for any $I' \in T'$, $\det A'_{I'} = 0$ and $A'_{T'} = \mathbf{0}$. Let $(A'_{R'}, A'_{T'})$ denote the vector formed by concatenating the vectors $A'_{R'}$ and $A'_{T'}$, so that $(A'_{R'}, A'_{T'})$ is the vector of Grassmann coordinates of A'. Since $A'_{T'}$ is identically zero and appending additional zeros to a vector does not affect the height we have

$$\mathcal{H}(A') = \mathcal{H}((A'_{R'}, A'_{T'})) = \mathcal{H}(A'_{R'})$$

We will similarly partition the Grassmann coordinates of A. Define

$$R = \{I \in \mathcal{J}(N, M) : M \in I\},$$

$$S = \{\{1, 2, \ldots, M-1, j\} : M+1 \leq j \leq N\},$$

and

$$T = J(N, M) \setminus (R \cup S)$$

Notice that for each $I' \in R'$ there exists $I \in R$ such that $I = I' \cup \{M\}$. Therefore, there is a one-to-one correspondence between elements of R' and elements R. In fact there is also a one-to-one correspondence between the Grassmann coordinates of A' indexed by R' and the Grassmann coordinates of A indexed by R. Let $I'_1 = I' \cap \{1, \ldots, M\}$ be the set of indices of I' less than M and let $I'_2 = I' \cap \{M+1, \ldots, N\}$

be the set indices of I' greater then M so that

$$A'_{I'} = \begin{pmatrix} A'_{I'_1} \\ A'_{I'_2} \end{pmatrix}.$$

Notice then that A_I is of the form

$$A_I = \begin{pmatrix} A'_{I'_1} & \mathbf{0}_{|I'_1|\times 1} \\ \mathbf{0}_{1\times M-1} & 1 \\ A'_{I'_2} & \mathbf{0}_{|I'_2|\times 1} \end{pmatrix}.$$

Then

$$\det A'_{I'} = \pm \det A_I. \tag{4.8}$$

Since multiplying the coordinates of a vector by ± 1 does not affect the height, (4.8) implies

$$\mathcal{H}(A'_{R'}) = \mathcal{H}(A_R).$$

There is also a correspondence between the Grassmann coordinates of A indexed by S and the entries of V. For each $M+1 \leq i \leq N$ and thus each $I = \{1, \ldots, M-1, i\} \in S$, A_I is of the form

$$A_I = \begin{pmatrix} \mathbf{1}_{M-1} & \mathbf{0}_{M-1\times 1} \\ \mathbf{0}_{1\times M-1} & V_{i-M} \end{pmatrix},$$

where V_{i-M} is the $i-M$th coordinate of V. Thus $\det A_I = V_{i-M}$ and

$$\mathcal{H}(A_S) = \mathcal{H}(V). \tag{4.9}$$

Let (A_R, A_S, A_T) be the vector formed by concatenating A_R, A_S and A_T so that

$H(A) = H((A_R, A_S, A_T))$. Then at each non-Archimedean place $\nu \in M(k)$

$$|(A_R, A_S, A_T)|_\nu = \max\{|A_R|_\nu, |A_S|_\nu, |A_T|_\nu\} \geq |A_S|_\nu = |A'_{S'}|_\nu, \tag{4.10}$$

and at each Archimedean place

$$\|(A_R, A_S, A_T)\|_\nu^2 = \|A_R\|_\nu^2 + \|A_S\|_\nu^2 + \|A_T\|_\nu^2 \geq \|A_S\|_\nu^2 = \|A'_{S'}\|_\nu^2, \tag{4.11}$$

with equality in (4.11) if and only if $\|A_S\|_\nu^2 = \|A_T\|_\nu^2 = 0$.

Combining (4.10) and (4.11) gives (1.19) where equality implies V is identically zero. As it is easy to verify that $H(A) = H(A')$ when $V = \mathbf{0}_{N-M}$ this completes the proof. □

Corollary 4.2.2. *Let A be a full rank $N \times M$ matrix with $N > M$ and coefficients in k. Let $\boldsymbol{\omega}_1$ through $\boldsymbol{\omega}_M$ be the columns of A so that*

$$A = \begin{pmatrix} \boldsymbol{\omega}_1 & \dots & \boldsymbol{\omega}_M \end{pmatrix},$$

and for some $1 \leq i \leq M$ let

$$A' = \begin{pmatrix} \boldsymbol{\omega}_1 \dots \boldsymbol{\omega}_{i-1} \boldsymbol{\omega}_{i+1} \dots \boldsymbol{\omega}_M \end{pmatrix},$$

be the $N \times M - 1$ matrix obtained by removing the i-th column of A. Let $I \in \mathcal{J}(N, M)$ and suppose that A_I is a permutation matrix. Then for any $1 \leq i \leq M$

$$\mathcal{H}(A) \geq \mathcal{H}(A')$$

with equality if and only if $\boldsymbol{\omega}_i$ is a standard basis vector.

Proof. Let Q be an $M \times M$ permutation matrix so that

$$AQ = \begin{pmatrix} \boldsymbol{\omega}_1 & \dots & \boldsymbol{\omega}_{i-1} & \boldsymbol{\omega}_{i+1} & \dots & \boldsymbol{\omega}_M & \boldsymbol{\omega}_i \end{pmatrix} = \begin{pmatrix} A' & \boldsymbol{\omega}_i \end{pmatrix}.$$

Multiplying by an $N \times N$ permutation matrix P_1 on the left we can permute the rows of A so that rows of A indexed by I are swapped with the rows of A indexed by $\{1, \dots, M\}$. Then P_1AQ is a matrix such that the first M rows form a permutation matrix. Again by multiplying by an $N \times N$ permutation matrix P_2 on the left we can further permute the rows so that P_2P_1AQ is a matrix so that the first M rows form an $M \times M$ identity matrix and thus P_2P_1AQ is of the same form as the matrix in (4.5) and

$$P_2P_1AQ = \begin{pmatrix} P_2P_1A' & P_2P_1\boldsymbol{\omega}_i \end{pmatrix}. \tag{4.12}$$

where P_2P_1A' is a matrix of the same form as (4.6). Thus we can apply Lemma 4.2.1 to show that

$$\mathcal{H}(P_2P_1AQ) \geq \mathcal{H}(P_2P_1A') \tag{4.13}$$

With equality if and only if $\boldsymbol{\omega}_i$ is a standard basis vector. Recall that multiplying be a permutation matrix does not change the height of a matrix thus, (4.13) implies

$$\mathcal{H}(P_2P_1AQ) = \mathcal{H}(A) \geq \mathcal{H}(A') = \mathcal{H}(P_2P_1A')$$

which completes the proof.

□

We are now ready to prove Theorem 1.4.1.

Proof of Theorem 1.4.1. Let A be an $N \times M$ basis matrix for $\mathcal{Z}$ and let $J \in \mathcal{J}(N, M)$ so that A_J is a full rank minor of A. Let $X = AA_J^{-1} = (\omega_1, \dots, \omega_m$ be another basis matrix for $\mathcal{Z}$ and let $\boldsymbol{\omega}_1$ through $\boldsymbol{\omega}_M$ the columns of X. By construction X_J, the

$M \times M$ minor of X indexed by J is a permutation matrix and thus we can apply Corollary 4.2.2 to X. In fact, for any $I \subseteq \{1, \ldots, M\}$ define X^I to the matrix whose columns are the columns of X indexed by I, then X^I will be a matrix such that one of the $|I| \times |I|$ minors of X^I will be a permutation matrix. Therefore, for any $I_1 \subsetneq I_2 \subset [M]$ we can construct a sequence of subsets

$$I_1 = L_1 \subsetneq L_2 \subsetneq \ldots \subsetneq L_{|I_2 \setminus I_1|} = I_2,$$

where $|L_{i+1} \setminus L_i| = 1$ for each $1 \leq i < |I_2 \setminus I_1|$. Then for each X^{L_i} we can apply Corollary 4.2.2 to conclude that

$$\mathcal{H}(X^{I_1}) = \mathcal{H}(X^{L_1}) \leq \ldots \leq \mathcal{H}(X^{L_{|I_2 \setminus I_1|}}) = \mathcal{H}(X^{I_2}),$$

With equality at each step if $\boldsymbol{\omega}_{L_{i+1} \setminus L_i}$ is a standard basis vector. Since $\mathcal{H}(\mathcal{Y}_{I_1}) = \mathcal{H}(X^{I_1})$ and $\mathcal{H}(\mathcal{Y}_{I_2}) = \mathcal{H}(X^{I_2})$ this implies (1.19) which in turn implies (1.18) and completes the proof. □

Remark 4.2.1. The basis constructed in Theorem 1.4.1 satisfies an additional sparseness property. We say that a vector $\boldsymbol{x} \in k^n$ is s-**sparse** for some $1 \leq s \leq n$ if $\boldsymbol{x}$ has no more than s nonzero coordinates. From the above proof, $X = AA_J^{-1}$ is a basis matrix for $\mathcal{Z}$. Notice that $X_J = A_J A_J^{-1}$ is the $M \times M$ identity matrix. This means that each column of X has at least $M - 1$ zero coordinates. Since the ω_i vectors for $1 \leq i \leq M$ are the columns of X the basis constructed in Theorem 1.4.1 is a basis of $(N - M + 1)$-sparse vectors.

We can additionally prove a relative version of Theorem 1.4.1, analogous to Theorem 12 of [8]. Let K be a finite extension of the number field k with $[K : k] = r \geq 2$. Let F be a number field such that $k \subseteq K \subseteq F$, F is a Galois extension of k with Galois group $G(F/k)$ and F is a Galois extension of K with Galois group $G(F/K)$.

Then $G(F/K)$ is a subgroup of $G(F/k)$ of index r. Let $\sigma_1, \ldots, \sigma_r \in G(F/k)$ be a set of distinct representatives of the cosets of $G(F/K)$ in $G(F/k)$. If A is an $M \times N$ matrix with coefficients a_{mn} in K for $1 \leq m \leq M, 1 \leq n \leq N$, define $\sigma_i(A) = (\sigma_i(a_{mn}))$. Then define

$$\mathcal{A} = \begin{pmatrix} \sigma_1(A) \\ \vdots \\ \sigma_r(A) \end{pmatrix}. \tag{4.14}$$

Theorem 4.2.3. *Let F, K, k be number fields as above. Let A be an $M \times N$ matrix with $rM < N$ and coefficients in K and define $\mathcal{A}$ as above. Suppose that* $\operatorname{rank} \mathcal{A} = rM$ *and let $L = N - rM$. Let A_m denote the m-th row of A. Then $\mathcal{Z} := \ker(A) \cap k^N$ is an L-dimensional k-vector space so that*

$$\mathcal{H}(\mathcal{Z}) = \mathcal{H}(\mathcal{A}) \leq \prod_{m=1}^{M} \mathcal{H}(A_m)^r. \tag{4.15}$$

Moreover, there exists a basis

$$\left\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_L\right\}$$

for $\mathcal{Z}$ with the following property: if $I \subseteq \{1, \ldots, L\}$ is a nonempty subset, and

$$\mathcal{Y}_I = \operatorname{span}_k \left\{\boldsymbol{\omega}_i : i \in I\right\}$$

is the k-linear subspace spanned by the basis vectors indexed by I, then

$$\mathcal{H}(\mathcal{Y}_I) \leq \prod_{m=1}^{M} \mathcal{H}(A_m)^r.$$

In particular, $\max_{1 \leq i \leq L} \mathcal{H}(\boldsymbol{\omega}_i) \leq \prod_{m=1}^{M} \mathcal{H}(A_m)^r$. *Moreover, if $I_1 \subsetneq I_2 \subseteq \{1, \ldots, L\}$,*

then

$$\mathcal{H}(\mathcal{Y}_{I_1}) \leq \mathcal{H}(\mathcal{Y}_{I_2}).$$

Proof. Let $\{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_r\}$ be a k-basis for K. Let $A = (a_{mn})$ for $1 \leq m \leq M, 1 \leq n \leq N$, where

$$a_{mn} = \sum_{j=1}^{r} a_{mn}^{(j)} \boldsymbol{\zeta}_j,$$

where $a_{mn}^{(j)} \in k$ for all m, n, j. For each $1 \leq j \leq r$, let $A^{(j)} = (a_{mn}^{(j)})$ and let

$$A' = \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(r)} \end{pmatrix}.$$

Then $\boldsymbol{x} \in k^n$ solves $A\boldsymbol{x} = \mathbf{0}$ if and only if it solves $A'\boldsymbol{x} = \mathbf{0}$, hence

$$V = \ker_k(A) = \ker_k(A').$$

Further, A' is a full-rank matrix with coefficients in k so

$$\mathcal{H}(\mathcal{Z}) = \mathcal{H}(A').$$

Now let

$$\Omega = (\sigma_i(\boldsymbol{\omega}_j)\mathbf{1}_M)_{\substack{i=1,\ldots,r \\ j=1,\ldots,r}}$$

be an $rM \times rM$ matrix organized into $M \times M$ blocks so that the (i, j) block is $\sigma_i(\boldsymbol{\omega}_j)\mathbf{1}_M$. Then

$$\Omega A' = \mathcal{A}.$$

Since Ω is a full-rank matrix,

$$\mathcal{H}(A') = \mathcal{H}(\mathcal{A}) = \mathcal{H}(\mathcal{Z}). \tag{4.16}$$

Writing $\mathcal{A}_i$ for the i-th row of $\mathcal{A}$, we see that (1.11) gives

$$\mathcal{H}(\mathcal{A}) \leq \prod_{i=1}^{rM} \mathcal{H}(\mathcal{A}_i) = \prod_{i=1}^{M} \prod_{j=1}^{r} \mathcal{H}(\sigma_j(A_i)). \tag{4.17}$$

Since $\mathcal{H}(\sigma_j(A_i)) = \mathcal{H}(A_i)$ combining (4.16) with (4.17) gives (4.15). We can now complete the proof by applying Theorem 1.4.1 to $\mathcal{A}$. □

4.3 Proof of Theorem 1.4.2

To prove Theorem 1.4.2, let us first recall some results on the existence of integer sensing matrices. An $L \times M$ integer matrix A, $L < M$, is called an *integer sensing matrix for L-sparse signals* if every $L \times L$ submatrix of A is nonsingular. This notation comes from the study of sparse signal recovery in the compressive sensing signal processing paradigm: the defining condition for A ensures that for a vector $\boldsymbol{x} \in \mathbb{Z}^M$ with no more than L nonzero coordinates (L-sparse), $A\boldsymbol{x} = \mathbf{0}$ if and only if $\boldsymbol{z} = \mathbf{0}$, which allows for a unique recovery of the original sparse signal from its image under A. The problem of existence of such matrices $A = (a_{ij})_{1 \leq i \leq L, 1 \leq j \leq M}$ with bounded sup-norm

$$|A| := \max\{|a_{ij}| : 1 \leq i \leq L, 1 \leq j \leq M\}$$

and M as large as possible as a function of L and $|A|$ has been considered recently in [23] with further improvements in [44] and [45]. In particular, Theorem 2.2 of [23] establishes the existence of an $L \times M$ integer sensing matrix A for L-sparse signals with $|A| = T$ and

$$M \geq CL\sqrt{T} \tag{4.18}$$

for an absolute constant C. Theorem 1.3 of [45] then establishes existence of such matrices with

$$M \geq \max\left\{T + 1, \frac{T^{\frac{L}{L-1}}}{2}\right\}, \tag{4.19}$$

which is an improvement on (4.18) when $L = o(\sqrt{T})$. Combining (4.18) with (4.19), we obtain an $L \times M$ integer matrix A with

$$|A| = T \leq \min\left\{\left(\frac{M}{CL}\right)^2, (2M)^{\frac{L-1}{L}}, M-1\right\}, \tag{4.20}$$

so that any subcollection of L column vectors of A is linearly independent.

Proof of Theorem 1.4.2. Now, let the notation be as in the statement of Theorem 1.4.2 and let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ be a basis for $\mathcal{Z}$, a specific choice will be made later. Dividing by a nonzero coordinate, if necessary, we can assume that each $\boldsymbol{\omega}_i$ has a coordinate equal to 1: this operation does not change the height of these basis vectors due to the product formula. This implies that for each $1 \leq i \leq L$,

$$\prod_{v \in M(K)} |(1, \boldsymbol{\omega}_i)|_v^{\frac{d_v}{d}} = \prod_{v \in M(K)} |\boldsymbol{\omega}_i|_v^{\frac{d_v}{d}} \leq \mathcal{H}(\boldsymbol{\omega}_i). \tag{4.21}$$

Write $W = (\boldsymbol{\omega}_1 \ \ldots \ \boldsymbol{\omega}_L)$ for the corresponding basis matrix and let $B = WA$. Let $S(M) = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$ be the set of column vectors of B. Then conclusion (1) of Theorem 1.4.2 are automatically satisfied, and we only need to prove (2). Notice that for each $1 \leq i \leq M$,

$$\boldsymbol{y}_i = \sum_{j=1}^{L} a_{ij} \boldsymbol{\omega}_j,$$

where all $a_{ij} \in \mathbb{Z}$ with $|a_{ij}| \leq T$. Then Lemma 2.1 of [24] combined with (4.21) implies that for each $1 \leq i \leq M$,

$$\mathcal{H}(\boldsymbol{y}_i) \leq L^{3/2} T \prod_{i=1}^{L} \mathcal{H}(\boldsymbol{\omega}_i). \tag{4.22}$$

Making a choice of the basis $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_L$ as in Theorem 1.4.1, we obtain a bound $\mathcal{H}(\boldsymbol{y}_i) \leq L^{3/2} T \mathcal{H}(\mathcal{Z})^L$. On the other hand, making a choice of the basis as in (4.1), we obtain a bound $\mathcal{H}(\boldsymbol{y}_i) \leq L^{3/2} \gamma_k(L)^{L/2} T \mathcal{H}(\mathcal{Z})$. The result of Theorem 1.4.2 now

follows by combining these observations with (4.20), where we are choosing the bound $(2M)^{\frac{L-1}{L}}$ for T since it is the smallest of the three for large M. $\square$

There are further results on $L \times M$ integer sensing matrix for s-sparse signals where $s < L$. Specifically, Theorem 1.1 of [29] asserts that for all sufficiently large L there exist $L \times M$ integer sensing matrices A for s-sparse signals with $1 \leq s \leq L-1$ such that $|A| = 2$ and

$$M \geq \left(\frac{L+2}{2}\right)^{1+\frac{2}{3s-2}}.$$

Using the same reasoning as in our argument above with $T = 2$, we immediately obtain the following observation.

Proposition 4.3.1. *For sufficiently large integers $L < N$, any integer $1 \leq s \leq L-1$, and an L-dimensional subspace $\mathcal{Z} \subseteq k^N$, there exists a collection of vectors*

$$S = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\} \subset \mathcal{Z}$$

with the following properties:

1. *For every $\boldsymbol{y}_i \in S$,*

$$\mathcal{H}(\boldsymbol{y}_i) \leq 2L^{3/2} \min\left\{\mathcal{H}(\mathcal{Z})^L, \gamma_k(L)^{L/2}\mathcal{H}(\mathcal{Z})\right\},$$

   where $\gamma_k(L)^{1/2}$ is the generalized Hermite's constant discussed in Section 1.3,

2. *The cardinality of the set S is*

$$M \geq \left(\frac{L+2}{2}\right)^{1+\frac{2}{3s-2}},$$

3. *Every subcollection of s vectors from S is linearly independent.*

There is also a connection between integer matrices and the basis constructed in Theorem 1.4.1.

Corollary 4.3.2. *Let A be an $L \times M$ integer sensing matrix with $L \leq M$. Let $\mathcal{Z}$ be the subspace spanned by A^T and let $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_M$ be the basis for $\mathcal{Z}$ constructed in Theorem 1.4.1. Let $\mathcal{Y}_I$ be as in* (1.17) *for each $I \subseteq [M]$. Then for every $I_1 \subsetneq I_2 \subset [M]$*

$$\mathcal{H}(\mathcal{Y}_{I_1}) < \mathcal{H}(\mathcal{Y}_{I_2}).$$

Proof. As demonstrated in the proofs of Lemmas 4.2.1 and 4.2.2 if $\mathcal{H}(\mathcal{Y}_{I_1}) = \mathcal{H}(\mathcal{Y}_{I_2})$ then there must exist an $M \times M$ minor of A^T and thus of A of rank less than M. This contradicts the fact that A is integer sensing. □

Chapter V

On Multilinear Forms

5.1 Height Inequalities

Here we prove several height inequalities we will need. The first lemma gives a bound on the height H of the inverse of a square nonsingular matrix (viewed as a vector) in terms of the height of the matrix itself.

Lemma 5.1.1. *Let $A \in \mathrm{GL}_n(K)$, then*

$$H(A^{-1}) \leq \left(\sqrt{n}H(A)\right)^{n-1}, \tag{5.1}$$

and

$$h(A^{-1}) \leq n^n |\Delta_K|^{\frac{1}{d}} h(A)^{2n-1}. \tag{5.2}$$

Proof. Let us write $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ for the row vectors of A, and for each $1 \leq j \leq n$ let A_j be the $(n-1) \times n$ submatrix of A obtained by deleting the j-th row $\boldsymbol{a}_j$. For each j,

$$V_j = \{\boldsymbol{x} \in K^n : A_j \boldsymbol{x} = \boldsymbol{0}\}$$

is a 1-dimensional subspace. Let $\boldsymbol{x}_j$ be any nonzero point in V_j and notice that $\boldsymbol{a}_j \boldsymbol{x}_j \neq 0$: if it was equal to 0, then $A\boldsymbol{x}_j = \boldsymbol{0}$, contradicting the assumption that A is nonsingular. Then take $\boldsymbol{b}_j = \frac{1}{\boldsymbol{a}_j \boldsymbol{x}_j} \boldsymbol{x}_j$, then $\boldsymbol{a}_i \boldsymbol{b}_j = 0$ for every $i \neq j$ and $\boldsymbol{a}_j \boldsymbol{b}_j = 1$. Take B to be the $n \times n$ matrix whose columns are the vectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$, then AB is the identity matrix, so $B = A^{-1}$. Notice also that for every j,

$$H(\boldsymbol{b}_j) \leq \mathcal{H}(\boldsymbol{b}_j) = \mathcal{H}(V_j) = \mathcal{H}(A_j) \leq \prod_{i=1, i\neq j}^{n} \mathcal{H}(\boldsymbol{a}_i) \leq n^{\frac{n-1}{2}} \prod_{i=1, i\neq j}^{n} H(\boldsymbol{a}_i), \tag{5.3}$$

by (1.7) combined with Lemma 1.11 and (1.6). Since $H(B) = \max_{1 \leq j \leq n} H(\boldsymbol{b}_j)$ and $H(\boldsymbol{a}_i) \leq H(A)$ for every $1 \leq i \leq n$, (5.1) follows.

To obtain the second inequality, we will use the Bombieri-Vaaler version of Siegel's

lemma (1.15) to choose a specific vector $\boldsymbol{x}_j \in V_j$ such that

$$h(\boldsymbol{x}_j) \leq \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{1/2d} \mathcal{H}(V_j) \leq |\Delta_K|^{\frac{1}{2d}} \mathcal{H}(V_j). \tag{5.4}$$

Now take $\boldsymbol{b}_j = \frac{1}{\boldsymbol{a}_j \boldsymbol{x}_j} \boldsymbol{x}_j$, then

$$\begin{aligned} h(\boldsymbol{b}_j) &\leq h\left(\sum_{l=1}^{n} a_{jl} x_{jl}, \boldsymbol{x}_j \right) \leq h\left(\sum_{l=1}^{n} a_{jl} x_{jl} \right) h(\boldsymbol{x}_j) \leq n h(\boldsymbol{a}_j) h(\boldsymbol{x}_j)^2 \\ &\leq n |\Delta_K|^{\frac{1}{d}} h(\boldsymbol{a}_j) \mathcal{H}(V_j)^2, \end{aligned} \tag{5.5}$$

where the last inequality follows by (5.4). Combining (5.5) with (5.3) and observing that $h(B) = \max_{1 \leq j \leq n} h(\boldsymbol{b}_j)$ and $h(\boldsymbol{a}_i) \leq h(A)$ for every $1 \leq i \leq n$, we obtain (5.2).

□

Next we present a useful bound on the height of a vector whose coordinates are images of a given point under a collection of polynomials.

Lemma 5.1.2. *Let $F_1, \ldots, F_k$ be polynomials of respective degrees $m_1, \ldots, m_k$ in $K[x_1, \ldots, x_n]$ and $\boldsymbol{z} \in K^n$. Then*

$$h(F_1(\boldsymbol{z}), \ldots, F_k(\boldsymbol{z})) \leq \mathcal{N} \mathfrak{H} h(\boldsymbol{z})^m, \tag{5.6}$$

where $\mathcal{N} = \max_{1 \leq i \leq k} \mathcal{N}(F_i)$, $\mathfrak{H} = \max_{1 \leq i \leq k} h(F_i)$ and $m = \max_{1 \leq i \leq k} m_i$.

Proof. For each $1 \leq i \leq k$, let us write

$$F_i(\boldsymbol{x}) = \sum_J f_J \boldsymbol{x}^J,$$

where $J = (j_1, \ldots, j_n)$ is a multi-index with $0 \leq j_l \leq m_i$ for each $1 \leq l \leq n$. Let

$\boldsymbol{z} \in K^n$, then for every $v \nmid \infty$ in $M(K)$,

$$|F_i(\boldsymbol{z})|_v \leq \max_J \left\{ |f_J|_v \prod_{l=1}^{n} |z_l|_v^{j_l} \right\} \leq \max_J \{|f_J|_v\} \max\{1, |z_1|_v, \ldots, |z_n|_v\}^{m_i},$$

and for $v \mid \infty$ in $M(K)$,

$$\begin{aligned} |F_i(\boldsymbol{z})|_v &\leq \mathcal{N}(F_i) \max_J \left\{ |f_J|_v \prod_{l=1}^{n} |z_l|_v^{j_l} \right\} \\ &\leq \mathcal{N}(F_i) \max_J \{|f_J|_v\} \max\{1, |z_1|_v, \ldots, |z_n|_v\}^{m_i}. \end{aligned}$$

Then

$$\begin{aligned} h(F_1(\boldsymbol{z}), \ldots, F_k(\boldsymbol{z})) &\leq \prod_{v \in M(K)} \max\{1, |F_1(\boldsymbol{z})|_v, \ldots, |F_k(\boldsymbol{z})|_v\}^{\frac{d_v}{d}} \\ &\leq \left(\max_{1 \leq i \leq k} \mathcal{N}(F_i) \right) \left(\max_{1 \leq i \leq k} h(F_i) \right) h(\boldsymbol{z})^{\max_{1 \leq i \leq k} m_i}. \end{aligned}$$

This is precisely (5.6). □

The next lemma bounds the height of a polynomial under a linear transformation.

Lemma 5.1.3. *Let $F(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ be a polynomial of degree m and let A be an $n \times k$ matrix over K. For a variable vector $\boldsymbol{y} = (y_1, \ldots, y_k)$, define $G(\boldsymbol{y}) = F(A\boldsymbol{y}^\top)$, then G is a polynomial in k variables over K of degree $\leq m$. Further,*

$$h(G) \leq k^m \mathcal{N}(F) h(F) h(A)^m.$$

Proof. Write $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq k}$ and notice that

$$A\boldsymbol{y}^\top = \left(\sum_{j=1}^{k} a_{1j} y_j, \ldots, \sum_{j=1}^{k} a_{nj} y_j \right)^\top.$$

Then $G(\boldsymbol{y}) = F\left(\sum_{j=1}^{k} a_{1j} y_j, \ldots, \sum_{j=1}^{k} a_{nj} y_j \right)$. Each linear form $\sum_{j=1}^{k} a_{ij} y_j$ has coef-

ficient vector of inhomogeneous height $\leq h(A)$, and F has $\mathcal{N}(F)$ monomials each of degree no bigger than m and height no bigger than $h(F)$. Therefore

$$h(G) \leq \mathcal{N}(F)h(F)(kh(A))^m,$$

where k^m is an upper bound on the binomial coefficients which occur when taking m-th power of a linear form. This is precisely the result of the lemma. □

The following lemma that will be quite important to us is a rigorous form of the basic principle that a polynomial which is not identically zero cannot vanish "too much". Somewhat different formulations of this principle can be found in [14] (Lemma 1 on p. 261) as well as in the context of N. Alon's celebrated Combinatorial Nullstellensatz [1]. The following formulation, which is most convenient for our purposes, follows easily from Lemma 2.2 of [26].

Lemma 5.1.4. *Suppose $P(\boldsymbol{x}) \in K[x_1, \ldots, x_n]$ is a polynomial of degree m which is not identically 0. There exists a point $\boldsymbol{z} \in \mathbb{Z}^n$ such that $P(\boldsymbol{z}) \neq 0$ and*

$$h(\boldsymbol{z}) \leq \frac{m+2}{2}.$$

5.2 Zeros of multiple polynomials

The main goal of this section is to prove Theorem 1.5.1. Let the notation be as in the statement of the Theorem.

Proof of Theorem 1.5.1. Let r be the rank of the linear system (1.21), which is the same as the rank of the homogeneous system as in the statement of the theorem. If $r < k$, then $k - r$ equations of (1.21) are dependent on the rest of them, and so every solution to the rest of the equations is automatically a solution to the whole system. Since r equations are linearly independent, there must exist some $r \times r$ submatrix of

the coefficient matrix of these equations with entries $F_{ij}(\boldsymbol{x}_{I'})$ which is nonsingular for some $\boldsymbol{x}_{I'} \in K^{n-k}$, and hence the determinant of this matrix is not identically zero as a polynomial in the variables of $\boldsymbol{x}_{I'}$. We can set all the $k-r$ variables x_{i_j} not corresponding to the columns of this submatrix equal to 0, and hence reduce to the case of r equations in r variables. Hence we can assume without loss of generality that $r = k$.

Let us rewrite (1.21) as

$$\mathcal{F}(\boldsymbol{x}_{I'})\boldsymbol{x}_I = \boldsymbol{f}(\boldsymbol{x}_{I'}), \tag{5.7}$$

where $\mathcal{F}(\boldsymbol{x}_{I'})$ is the $k \times k$ matrix with entries $F_{ij}(\boldsymbol{x}_{I'})$, $\boldsymbol{f}(\boldsymbol{x}_{I'})$ is the k-dimensional column vector with coordinates $-F_{i(k+1)}(\boldsymbol{x}_{I'})$ and $\boldsymbol{x}_I = (x_{i_1}, \ldots, x_{i_k})^\top$ is the variable vector. Then $\mathcal{F}(\boldsymbol{x}_{I'})$ is nonsingular for some choice of $\boldsymbol{x}_{I'} \in K^{n-k}$, hence $P(\boldsymbol{x}_{I'}) := \det(\mathcal{F}(\boldsymbol{x}_{I'}))$ is not identically zero as a polynomial in the variables $\boldsymbol{x}_{I'}$. Notice that $\deg(P) \leq \sum_{i=1}^{k} \deg(F_i) = D$, and hence by Lemma 5.1.4 there exists a point $\boldsymbol{z}_{I'} \in \mathbb{Z}^{n-k}$ such that $P(\boldsymbol{z}_{I'}) \neq 0$ and

$$h(\boldsymbol{z}_{I'}) \leq \frac{D+2}{2}. \tag{5.8}$$

Plugging in $\boldsymbol{z}_{I'}$ for $\boldsymbol{x}_{I'}$ into (5.7), we obtain a nonsingular linear system of k equations in k variables, and hence have a unique solution $\boldsymbol{z}_I = \mathcal{F}(\boldsymbol{z}_{I'})^{-1}\boldsymbol{f}(\boldsymbol{z}_{I'})$. Combining $\boldsymbol{z}_I$ with $\boldsymbol{z}_{I'}$ into appropriately indexed coordinates, we obtain a vector $\boldsymbol{z} \in \bigcap_{j=1}^{k} Z_K(F_j)$ and $H(\boldsymbol{z}) = H(\boldsymbol{z}_I, \boldsymbol{z}_{I'})$.

We now estimate the height of $\boldsymbol{z}_I$. First notice that, by Lemma 5.1.1,

$$h(\mathcal{F}(\boldsymbol{z}_{I'})^{-1}) \leq k^k |\Delta_K|^{\frac{1}{d}} h(\mathcal{F}(\boldsymbol{z}_{I'}))^{2k-1}. \tag{5.9}$$

On the other hand,

$$h(\mathcal{F}(\boldsymbol{z}_{I'})) = H(1, F_{11}(\boldsymbol{z}_{I'}), \ldots, F_{kk}(\boldsymbol{z}_{I'})) \leq \mathcal{N} h(\boldsymbol{z}_{I'})^m \max_{1\leq i\leq k} h(F_i), \tag{5.10}$$

as well as

$$h(\boldsymbol{f}(\boldsymbol{z}_{I'})) = H(1, -F_{1(k+1)}(\boldsymbol{z}_{I'}), \ldots, -F_{k(k+1)}(\boldsymbol{z}_{I'})) \leq \mathcal{N} h(\boldsymbol{z}_{I'})^m \max_{1\leq i\leq k} h(F_i), \tag{5.11}$$

both by Lemma 5.1.2. Combining (5.9), (5.10), (5.11) and (5.8), we obtain:

$$\begin{aligned} h(\boldsymbol{z}_I) &= h\left(\mathcal{F}(\boldsymbol{z}_{I'})^{-1}\boldsymbol{f}(\boldsymbol{z}_{I'})\right) \leq kh\left(\mathcal{F}(\boldsymbol{z}_{I'})^{-1})h(\boldsymbol{f}(\boldsymbol{z}_{I'}))\right) \\ &\leq k^{k+1}|\Delta_K|^{\frac{1}{d}} h(\mathcal{F}(\boldsymbol{z}_{I'}))^{2k-1} h(\boldsymbol{f}(\boldsymbol{z}_{I'})) \\ &\leq k^{k+1}|\Delta_K|^{\frac{1}{d}} \left(\mathcal{N} h(\boldsymbol{z}_{I'})^m \max_{1\leq i\leq k} h(F_i)\right)^{2k} \\ &\leq k^{k+1}|\Delta_K|^{\frac{1}{d}} \left(\mathcal{N}\left(\frac{D+2}{2}\right)^m \max_{1\leq i\leq k} h(F_i)\right)^{2k}. \end{aligned} \tag{5.12}$$

Now notice that

$$\begin{aligned} h(\boldsymbol{z}) &= H(1, \boldsymbol{z}_I, \boldsymbol{z}_{I'}) = \prod_{v\in M(K)} \max\left\{1, |\boldsymbol{z}_I|_v, |\boldsymbol{z}_{I'}|_v\right\}^{\frac{d_v}{d}} \\ &\leq \prod_{v\in M(K)} \left(\max\left\{1, |\boldsymbol{z}_I|_v\right\}^{\frac{d_v}{d}} \max\left\{1, |\boldsymbol{z}_{I'}|_v\right\}^{\frac{d_v}{d}}\right) = h(\boldsymbol{z}_I) h(\boldsymbol{z}_{I'}), \end{aligned}$$

and hence the theorem follows from (5.8) and (5.12). □

Remark 5.2.1. Our argument leads to an algorithm for finding a simultaneous zero $\boldsymbol{z}$ of the polynomial system (1.21) under the assumption that it exists:

1. Compute $P(\boldsymbol{x}_{I'}) = \det(\mathcal{F}(\boldsymbol{x}_{I'}))$ as a polynomial in the variables $\boldsymbol{x}_{I'}$.

2. Search through the set of integer points of sup-norm $\leq \deg(P)$ to find $\boldsymbol{z}_{I'}$ such that $P(\boldsymbol{z}_{I'}) \neq 0$.

3. For this choice of $\boldsymbol{z}_{I'}$, compute $\mathcal{F}(\boldsymbol{z}_{I'})$ and $\boldsymbol{f}(\boldsymbol{z}_{I'})$.

4. Compute $\mathcal{F}(\boldsymbol{z}_{I'})^{-1}$.

5. Compute $\boldsymbol{z}_I = \mathcal{F}(\boldsymbol{z}_{I'})^{-1}\boldsymbol{f}(\boldsymbol{z}_{I'})$.

6. Combine $\boldsymbol{z}_I$ and $\boldsymbol{z}_{I'}$ according to the indices of the coordinates to obtain the vector $\boldsymbol{z}$.

5.3 Zeros of one polynomial

In this section, we prove theexistence of a zero of bounded height for a polynomial F, linear in at least one variable, outside of an algebraic set not containing the entire zero locus of F. Our main goal is to prove Theorem 1.5.2.

Proof of Theorem 1.5.2. Assume without loss of generality that F is linear in x_n. Define $\boldsymbol{x}' := (x_1, \ldots, x_{n-1})$, so $\boldsymbol{x} = (\boldsymbol{x}', x_n)$. We can write

$$F(x_1, \ldots, x_n) = x_n F_1(x_1, \ldots, x_{n-1}) + F_2(x_1, \ldots, x_{n-1}),$$

where $\deg F_1 \leq g-1$, $\deg F_2 \leq g$ and both of them are polynomials in $n-1$ variables $\boldsymbol{x}'$ with at least F_1 not identically zero. Then we can describe $Z_K(F)$ as the union $Z_K^1(F) \cup Z_K^2(F)$, where

$$\begin{aligned} Z_K^1(F) &= \{\boldsymbol{z} \in K^n : F_1(\boldsymbol{z}') = F_2(\boldsymbol{z}') = 0\}, \\ Z_K^2(F) &= \{\boldsymbol{z} \in K^n : F_1(\boldsymbol{z}') \neq 0, z_n = -F_2(\boldsymbol{z}')/F_1(\boldsymbol{z}')\}. \end{aligned} \tag{5.13}$$

Let $Z_K^2(F)' = \{\boldsymbol{z}' : \boldsymbol{z} \in Z_K^2(F)\}$, and define

$$Q(\boldsymbol{x}') = P\left(x_1, \ldots, x_{n-1}, -\frac{F_2(\boldsymbol{x}')}{F_1(\boldsymbol{x}')}\right) F_1(\boldsymbol{x}')^m.$$

Since the degree of P is m, Q is a polynomial in $n-1$ variables $\boldsymbol{x}'$ on $Z_K^2(F)'$. Let us show that Q is not identically 0. Suppose it is, then $P(Z_K^2(F)) = 0$. For each $\boldsymbol{z}' \in Z_K^2(F)'$ view $P(\boldsymbol{z}', x_n)$ as a polynomial in one variable x_n. Since

$$x_n = -F_2(\boldsymbol{z}')/F_1(\boldsymbol{z}')$$

is a root of this polynomial, the linear factor $x_n + F_2(\boldsymbol{z}')/F_1(\boldsymbol{z}')$ must divide it, and since we are working over a field, we can say that $F_1(\boldsymbol{z}')x_n + F_2(\boldsymbol{z}')$ divides $P(\boldsymbol{z}', x_n)$ for every $\boldsymbol{z}' \in Z_K^2(F)'$. This implies that $P(\boldsymbol{x})$ must be divisible by $x_n F_1(\boldsymbol{x}') + F_2(\boldsymbol{x}') = F(\boldsymbol{x})$, and therefore $Z_K(F) \subseteq Z_K(P)$. This is a contradiction, and hence there exists $\boldsymbol{z}' \in Z_K^2(F)$ such that $Q(\boldsymbol{z}') \neq 0$. We want to find such $\boldsymbol{z}'$ of bounded height. Notice that

$$\deg Q \leq m \deg F_2 + m \deg F_1 \leq m(g + g - 1) = m(2g - 1).$$

By Lemma 5.1.4, there exists $\boldsymbol{z}' \in \mathbb{Z}^{n-1}$ such that $Q(\boldsymbol{z}') \neq 0$ and

$$h(\boldsymbol{z}') \leq \frac{\deg Q + 1}{2} = \frac{m(2g-1)+2}{2}. \tag{5.14}$$

Then we can estimate the height of the corresponding point $\boldsymbol{z} = \left(\boldsymbol{z}', -\frac{F_2(\boldsymbol{z}')}{F_1(\boldsymbol{z}')}\right)$. Notice that

$$\begin{aligned}
h(\boldsymbol{z}) &\leq H(1, F_1(\boldsymbol{z}')\boldsymbol{z}', F_2(\boldsymbol{z}')) \\
&= \prod_{v \in M(K)} \max\left\{1, |F_1(\boldsymbol{z}')z_1|_v, \ldots, |F_1(\boldsymbol{z}')z_{n-1}|_v, |F_2(\boldsymbol{z}')|_v\right\}^{\frac{d_v}{d}} \\
&\leq \prod_{v \in M(K)} \left(\max\left\{1, |F_1(\boldsymbol{z}')|_v, |F_2(\boldsymbol{z}')|_v\right\} \max\{1, |z_1|_v, \ldots, |z_{n-1}|_v\}\right)^{\frac{d_v}{d}} \\
&\leq H\left(1, F_1(\boldsymbol{z}'), F_2(\boldsymbol{z}')\right) h(\boldsymbol{z}') \leq \mathcal{N}(F) h(F) h(\boldsymbol{z}')^g h(\boldsymbol{z}'),
\end{aligned} \tag{5.15}$$

where the last inequality follows by Lemma 5.1.2. Combining (5.14) and (5.15) yields (1.22). □

Remark 5.3.1. Notice that description (5.13) immediately implies that $Z_K(F)$ is not empty, and therefore taking P to be a nonzero constant polynomial we see that F has a zero $\boldsymbol{z} \in K^n$ with $h(\boldsymbol{z}) \leq \mathcal{N}(F)h(F)$. Further, our argument allows for an explicit algorithm to find the point $\boldsymbol{z}$ in question, similar to the procedure described in Remark 5.2.1 for Theorem 1.5.1: here, we just need to do a finite search for an integer point $\boldsymbol{z}'$ so that $Q(\boldsymbol{z}') \neq 0$ and define $\boldsymbol{z} = \left(\boldsymbol{z}', -\frac{F_2(\boldsymbol{z}')}{F_1(\boldsymbol{z}')}\right)$.

One can further ask if a result similar to Theorem 1.5.2 holds with a restriction to a subspace V of K^n. The problem here is that the restriction of our polynomial F to V may no longer be linear in any of the variables. For example, if $F(x_1, x_2) = x_1x_2$ and $V = K\begin{pmatrix}1\\1\end{pmatrix} \subset K^2$, then the restriction of F to V is

$$F_V(x) = F(x, x) = x^2,$$

and hence is not linear. On the other hand, we can prove a simple lemma in case $\dim_K V = 1$.

Lemma 5.3.1. *Let F, P be as in Theorem 1.5.2 and suppose V is a one-dimensional subspace of K^n such that*

$$Z_K(F, V) := Z_K(F) \cap V \not\subseteq Z_K(P).$$

Then there exists $\boldsymbol{z} \in Z_K(F, V) \setminus Z_K(P)$ such that

$$h(\boldsymbol{z}) \leq \left(\left(\frac{2}{\pi}\right)^{2r_2} |\Delta_K|\right)^{\frac{m+1}{2d}} \mathcal{N}(F)^{\frac{3}{2}} h(F) \mathcal{H}(V)^{m+1}.$$

Proof. Now suppose that V is a one-dimensional subspace of K^n, i.e., $\ell = 1$. Then $V = K\boldsymbol{y}$ for some vector $\boldsymbol{y} \in K^n$ and $F(\alpha\boldsymbol{y}) = 0$ for some $\alpha \in K$ such that $\boldsymbol{z} := \alpha\boldsymbol{y} \notin Z_K(P)$, since $Z_K(F, V) \not\subseteq Z_K(P)$. By Theorem 1.15, we can choose $\boldsymbol{y}$

such that

$$h(\boldsymbol{y}) \le \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{1}{2d}} \mathcal{H}(V). \tag{5.16}$$

Then F_V is a polynomial in one variable of degree $\le m$ with

$$h(F_V) \le \mathcal{N}(F) h(F) h(\boldsymbol{y})^m \le \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{m}{2d}} \mathcal{N}(F) h(F) \mathcal{H}(V)^m, \tag{5.17}$$

by Lemma 5.1.3 combined with (5.16). Let $\alpha_1, \dots, \alpha_k$ be the roots of F_V, $k \le m$, with repetition if necessary. Then Lemma 2 of [63] combined with (1.6) guarantees that

$$\prod_{i=1}^{k} h(\alpha_i) \le \sqrt{\mathcal{N}(F)} h(F_V). \tag{5.18}$$

Observing that $\max_{1 \le i \le k} h(\alpha_i) \le \prod_{i=1}^{k} h(\alpha_i)$ and combining (5.17) with (5.18), we obtain

$$\max_{1 \le i \le k} h(\alpha_i \boldsymbol{y}) \le h(\boldsymbol{y}) \max_{1 \le i \le k} h(\alpha_i) \le \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{m+1}{2d}} \mathcal{N}(F)^{\frac{3}{2}} h(F) \mathcal{H}(V)^{m+1},$$

which is the bound of the lemma. This completes the proof. □

5.4 Zeros of multilinear forms

In this section we prove Theorem 1.5.3. For $n \ge 2$, let $n \ge g \ge 1$ and let $F(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$ be a multilinear (n, g)-form, which is not identically zero.

Lemma 5.4.1. *There exists a nonzero point $\boldsymbol{z} \in K^n$ such that $F(\boldsymbol{z}) = 0$ and*

$$H(\boldsymbol{z}) \le H(F). \tag{5.19}$$

Proof. We argue by induction on n. If $n = 2$, then

$$F(x_1, x_2) = ax_1 + bx_2, \text{ or } F(x_1, x_2) = cx_1x_2,$$

for some $a, b, c \in K$. In the first case, $F(-b, a) = 0$, and in the second $F(0, 1) = 0$. In either case, the nontrivial zero $\boldsymbol{z} = (-b, a)$ or $\boldsymbol{z} = (0, 1)$ satisfies the bound (5.19).

Suppose now $n > 2$. If $n = g$, then

$$F(x_1, \ldots, x_n) = cx_1 \cdots x_n,$$

and so $F(0, \ldots, 0, 1) = 0$. Hence assume $n > g$, then for some $1 \leq i \leq n$ we can write

$$F(x_1, \ldots, x_n) = x_i F_1(\boldsymbol{x}_i') + F_2(\boldsymbol{x}_i'),$$

where

$$\boldsymbol{x}_i' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

is the vector of $n - 1$ variables excluding x_i and F_1, F_2 are multilinear forms in $n - 1$ variables, not identically zero. By the induction hypothesis, there exists a nonzero point $\boldsymbol{z}' \in K^{n-1}$ such that $F_2(\boldsymbol{z}') = 0$ satisfying (5.19). Define $\boldsymbol{z}$ by inserting 0 for the i-th coordinate in $\boldsymbol{z}'$, then $F(\boldsymbol{z}) = 0$ and $\boldsymbol{z}$ satisfies (5.19). □

On the other hand, if $V \subseteq K^n$ is a subspace of K^n, then F may not necessarily have nontrivial zeros on V. Indeed, the form $x_1 + x_2$ has no nontrivial zeros on the subspace $\{(a, 2a) : a \in \mathbb{Q}\}$ of $\mathbb{Q}^2$. There are, however, some situations when F is guaranteed to have nontrivial zeros on V, and then we can find such zeroes of small height.

Lemma 5.4.2. *Assume $g > 1$ and let F be a multilinear (n, g)-form over K and $\boldsymbol{x} \in K^n$ a $(g - 1)$-sparse vector. Then $F(\boldsymbol{x}) = 0$.*

Proof. The vector $\boldsymbol{x}$ has no more than $g-1$ nonzero coordinates. Hence $F(\boldsymbol{x})=0$, since every monomial of F has degree g and is linear in each variable, hence is a product of $g>1$ distinct variables. □

Corollary 5.4.3. *Let $V \subseteq K^n$ be an m-dimensional subspace of K^n and F a multilinear (n,g)-form over K. Assume that $m+g-1>n$ and $g>1$. Then V contains a basis $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$ of vectors such that $F(\boldsymbol{x}_1)=\cdots=F(\boldsymbol{x}_m)=0$ and*

$$H(\boldsymbol{x}_i) \leq \mathcal{H}(V). \tag{5.20}$$

for each $1 \leq i \leq m$.

Proof. By Theorem 1.4.1, see Remark 4.2.1, V contains a basis $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$ of $(n-m+1)$-sparse vectors satisfying (5.20). Since $m+g-1>n$, these vectors have no more than $g-1 \geq n-m+1$ nonzero coordinates, and hence $F(\boldsymbol{x}_i)=0$ for each $1 \leq i \leq m$. □

Corollary 5.4.4. *Let $V \subseteq K^n$ be an m-dimensional subspace of K^n and F a multilinear (n,g)-form over K. Assume that $m+g-1>n$ and $g>1$. Suppose also that $P(x_1,\ldots,x_n) \in K[x_1,\ldots,x_n]$ is a polynomial such that $D(P,V)<m$. Then there exists a point $\boldsymbol{z} \in V \setminus Z_K(P,V)$ such that $F(\boldsymbol{z})=0$ and*

$$H(\boldsymbol{z}) \leq \sqrt{2}m|\Delta_K|^{\frac{m+1}{2d}}\mathcal{H}(V).$$

Proof. Let $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$ be the $(n-m+1)$-sparse basis for V guaranteed by Theorem 1.4.1 and Remark 4.2.1. Since $m+g-1>n$, we see that $g>n-m+1$ and hence $F(\boldsymbol{x}_i)=0$ for each $1 \leq i \leq m$, by the same argument as in the proof of Corollary 5.4.3. Since $D(P,V)<m$, at least one of these vectors is not in $Z_K(P,V)$. Call this vector $\boldsymbol{z}$, and the result follows. □

Proof of Theorem 1.5.3. The theorem now follows by combining Corollaries 5.4.3 and 5.4.4. □

The case $g = 1$ of linear forms has to be considered separately: this is just a simple case of Theorem 1.4 of [24], which we present here in a simplified form. Suppose that

$$F(x_1, \dots, x_n) = \sum_{i=1}^{n} a_i x_i \in K[x_1, \dots, x_n],$$

and $V \subseteq K^n$ is an m-dimensional subspace.

Lemma 5.4.5. *Let $V \subseteq K^n$ be an m-dimensional subspace with $2 \le m \le n$. Then there exists $\mathbf{0} \neq \boldsymbol{z} \in V$ such that $F(\boldsymbol{z}) = 0$ and*

$$H(\boldsymbol{z}) \le \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{1}{2d}} \left(\sqrt{n} \mathcal{H}(V) H(F) \right)^{\frac{1}{m-1}}. \tag{5.21}$$

Further, if $P(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$ is a polynomial such that $D(P, V) < m - 1$, then there exists a point $\boldsymbol{z} \in V \setminus Z_K(P, V)$ such that $F(\boldsymbol{z}) = 0$ and

$$H(\boldsymbol{z}) \le \sqrt{n} \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{m-1}{2d}} \mathcal{H}(V) H(F). \tag{5.22}$$

Proof. Define

$$U(F) = \{ \boldsymbol{x} \in K^n : F(\boldsymbol{x}) = 0 \},$$

then, by (1.7),

$$\mathcal{H}(U(F)) = \mathcal{H}(F) \le \sqrt{n} H(F), \tag{5.23}$$

where $\mathcal{H}(F)$ is the $\mathcal{H}$ height applied to the coefficient vector of F, which has n coordinates. This is an $(n-1)$-dimensional subspace of K^n and $\dim V \ge 2$, hence

$$\ell := \dim(V \cap U(F)) \ge m - 1 \ge 1.$$

By Theorem 1.15, there is a basis $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell \in V \cap U(F)$ such that

$$\begin{aligned} \prod_{i=1}^{\ell} H(\boldsymbol{x}_i) &\leq \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{\ell}{2d}} \mathcal{H}(V \cap U(F)) \\ &\leq \left(\left(\frac{2}{\pi} \right)^{2r_2} |\Delta_K| \right)^{\frac{\ell}{2d}} \mathcal{H}(V)\mathcal{H}(U(F)), \end{aligned} \tag{5.24}$$

by Lemma 1.13. Then (5.21) follows by combining (5.23) with (5.24) and taking $\boldsymbol{z}$ to be the vector with smallest height among $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell$. Since $D(P, V) < m - 1$, at least one of these vectors is not in $Z_K(P, V)$, and this implies (5.22). □

We will finish with a partial result on a special class of multilinear forms referred to as strong multilinear forms. Let $[n] := \{1, \ldots, n\}$ and for an integer $1 \leq d \leq n$ and let $J_d(n) := \{J \subseteq [n] : |I| = d\}$. An (n, d) multilinear form F is said to be **strongly multilinear** if there exists a partition $\{x_1, \ldots, x_n\} = \bigsqcup_{k=1}^{d} X_k$ such that F is linear with respect to each partition treated as a vector X_k. For each partition subset X_k let $n_k = \text{length}(X_k)$ and index the elements of $X_k = \{x_{k_1}, \ldots, x_{k_{nk}}\}$

Strongly multilinear forms are of the form

$$F(x_1, \ldots, x_n) = F(X_1, \ldots, X_d) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} f_I x_{i_1} \ldots x_{i_d}$$

where $I = \{i_1, \ldots, i_d\}$ and $f_I \in \mathbb{Z}$.

Alternatively

$$F(X_1, \ldots, X_d) = A^T \otimes X_1 \otimes \ldots \otimes X_d$$

where A is an appropriately sized coefficient tensor. For the sake of convenience we represent $X_1 \otimes \ldots \otimes X_d$ as an $\prod_{i=1}^{d} n_i$ column vector.

We now recall that any $m \times n$ matrix A put into Smith normal form:

$$A = QAP^{-1}, \tag{5.25}$$

Where $Q \in \mathrm{GL}_m(\mathbb{Z})$, $P \in \mathrm{GL}_n(\mathbb{Z})$ and A' is of the form

$$A' = \begin{bmatrix} \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_r \end{bmatrix} & \\ & 0 \end{bmatrix},$$

Such that $d_1, \ldots, d_r > 0$ and for $1 \leq i < j \leq r$ $d_i | d_j$. For more information on Smith normal form as well as an algorithm to find Smith normal form of a matrix see section 12.4 [3].

This leads to the corollary:

Corollary 5.4.6. *Let A be an $m \times n$ matrix with integer coefficients. Then there exists $\boldsymbol{y} \in \mathbb{Z}^n$ such that* $\gcd(A\boldsymbol{y}) = \gcd(A)$.

Proof. We begin by showing $d_1 = \gcd(A)$. Define A' and $d_1, \ldots, d_r$ as in Theorem 2.1. Since $d_1 | d_i$ for all i, $d_1 | A'$. Then $d_1 | Q^{-1}A'P = A$. Then $d_1 | \gcd(A)$. Likewise, $\gcd(A) | A = Q^{-1}A'P$. $\gcd(A)$ cannot divide Q^{-1} or P as these matrices are in general linear groups over $\mathbb{Z}$ therefore they must have primitive columns and therefore $\gcd(P) = \gcd(Q^{-1}) = 1$. Therefore $\gcd(A) | \gcd(A')$ which implies $\gcd(A) | \gcd(A') = d_1$. Therefore $\gcd(A) = d_1$.

Now let $\boldsymbol{y} = P_1^{-1}$ be the first column of P^{-1} and let Q_1^{-1} be the first column of Q^{-1}. Let $e_1(n)$ denote the first standard basis vector of $\mathbb{Z}^n$ Then

$$A\boldsymbol{y} = Q^{-1}A'PY = Q^{-1}A'e_1(n) = Q^{-1}d_1e_1(m) = d_1Q_1^{-1}$$

Finally note that $\gcd(Q_1^{-1}) = 1$ since $Q^{-1} \in GL_m(\mathbb{Z})$. □

We are now ready to prove a result on strong integer multilinear forms:

Theorem 5.4.7. *Let $F(X_1, \ldots, X_d)$ be an $(n_1 \ldots n_d, d)$ strongly multilinear form with coefficients in $\mathbb{Z}$. Then for any b such that $\gcd(F)$ divides b, there exists $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d)$ such that $F(\boldsymbol{y}) = b$.*

Proof. By induction on the degree d:

Suppose that $d = 1$. Then F is a linear form and F represents $\gcd(F)\mathbb{Z}$ by the Euclidean algorithm. This establishes the base of induction.

Now suppose that $d > 1$ and that $n_d = \text{length}(X_d) = 1$. Then x_{d_1} divides every monomial in F. Set $\boldsymbol{y}_d = 1$ then $F' = F(X_1, \ldots, X_{d-1}, \boldsymbol{y}_d)$ is an $(n_1 \ldots n_{d-1}, d-1)$ strongly multilinear form with $\gcd(F') = \gcd(F)$.

Finally, suppose that $d > 1$ and $n_d > 1$. Then factor F as

$$F(X_1, \ldots, X_d) = (X_1 \otimes, \ldots, \otimes X_{d-1})^T A X_d$$

Where A is the $n_1 \ldots, n_{d-1} \times n_d$ coefficient matrix for F. Then by Corollary 5.4.6 there exists Y_d such that $\gcd(A\boldsymbol{y}_d) = \gcd(A)$. Then $F' = F(X_1, \ldots, X_{d-1}, \boldsymbol{y}_d)$ is an $(n_1 \ldots n_{d-1}, d-1)$ strongly multilinear form with $\gcd(F') = \gcd(F)$. Thus by the induction hypothesis there exist $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{d-1}$ so that $F(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_d) = b$. □

Unfortunately, we are currently unable to bound the size of $\boldsymbol{y}$ in Theorem 5.4.7. This is due to the use of Smith normal form as we cannot currently bound the sizes of Q and P^{-1} in (5.25). Theorem 5.4.7 can be generalized to multilinear forms with coefficients in a PID but not to an abitrary ring of algebraic integers as Smith normal form is only defined for PIDs.

Bibliography

[1] N. Alon. Combinatorial Nullstellensatz. *Combin. Probab. Comput.*, 8 (1999), no. 1-2, pp. 7–29.

[2] Y. André. On nef and semistable hermitian lattices, and their behaviour under tensor product. *Tohoku Math. J. (2)*, 63(4):629–649, 2011.

[3] M. Artin Algebra. Prentice-Hall, 1991.

[4] A. Baker, Linear forms in the logarithms of algebraic numbers, I, II, III, IV, *Mathematika*, 13 (1966), 204–216; 14 (1967), 102–107, 220–228; 15 (1968), 204–216.

[5] R. Baraniuk, S. Dash, and R. Neelamani. On nearly orthogonal lattice bases. *SIAM J. Discrete Math*, 21(1):199–219, 2007.

[6] D. Bertrand. Duality on tori and multiplicative dependence relations. *J. Austral. Math. Soc. Ser. A*, 62 (1997), no. 2, 198–2

[7] E. Bombieri and W. Gubler, *Heights in Diophantine Geometry*, Cambridge U. Press, New York, 2006.

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

[9] T. Borek. Successive minima and slopes of Hermitian vector bundles over number fields. *J. Number Theory*, 113(2):380–388, 2005.

[10] I. Borosh and M. Flahive and D. Rubin and B. Treybig. A sharp bound for solutions of linear Diophantine equations. *Proc. Amer. Math. Soc.*, 105 (1989), no. 4, 844–8

[11] A. Böttcher and L. Fukshansky. Representing integers by multilinear polynomials. *Res. Number Theory*, 6 (2020), no. 4, Paper No. 38, 8 pp.

[12] T. D. Browning, R. Dietmann and P. D. T. A Elliott. Least zero of a cubic form. *Math. Ann.*, 352 (2012) no. 3, pp. 745–778.

[13] B. Casselman. Stability of lattices and the partition of arithmetic quotients. *Asian J. Math.*, 8(4):607–637, 2004.

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

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

[16] J. H. Conway and N. J. A. Sloane. Sphere packings, lattices and groups, 3rd edition, Springer-Verlag, 1999.

[17] M. Fang. On the completion of a partial integral matrix to a unimodular matrix. *Linear Algebra Appl.*, 422:291–294, 2007.

[18] E. Fischer, Über den Hadamardschen Determinantensatz, *Arch. Math. (Basel)*, 13, (1908), 32–40.

[19] M. Forst and L. Fukshansky. Counting basis extensions in a lattice. *Proc. Amer. Math. Soc.*, 150(8):3199–3213, 2022.

[20] M. Forst and L. Fukshansky. On zeros of multilinear forms. J. Number Theory, 245:169–186, 2023.

[21] M. Forst and L. Fukshansky. On lattice extensions. in submission.

[22] M. Forst, L. Fukshansky and J. Vaaler. On a new absolute version of Siegel's lemma. in preparation.

[23] L. Fukshansky, D. Needell, and B. Sudakov. An algebraic perspective on integer sparse recovery. *Appl. Math. Comput.*, 340:31–42, 2019.

[24] L. Fukshansky. Algebraic points of small height missing a union of varieties. *J. Number Theory*, 130 (2010), no. 10, pp. 2099–2118.

[25] L. Fukshansky. Heights and quadratic forms: Cassels' theorem and its generalizations. Diophantine methods, lattices, and arithmetic theory of quadratic forms, 77–93, Contemp. Math., 587, Amer. Math. Soc., Providence, RI, 2013.

[26] L. Fukshansky. Integral points of small height outside of a hypersurface. *Monatsh. Math.*, 147 (2006), no. 1, pp. 25–41.

[27] L. Fukshansky. Revisiting the hexagonal lattice: on optimal lattice circle packing. *Elem. Math.*, 66(1):1–9, 2011.

[28] L. Fukshansky, P. Guerzhoy and F. Luca. On arithmetic lattices in the plane. *Proc. Amer. Math. Soc.*, 145(4):1453–1465, 2017.

[29] L. Fukshansky and A. Hsu. Covering point-sets with parallel hyperplanes and sparse signal recovery. *Discrete Comput. Geom.*, 2022, https://doi.org/10.1007/s00454-022-00375-y.

[30] L. Fukshansky and G. Henshaw. Lattice point counting and height bounds over number fields and quaternion algebras. *Online J. Anal. Comb.*, 8 (2006), 20 pp

[31] L. Fukshansky and D. Kogan. On the geometry of nearly orthogonal lattices. *Linear Algebra Appl.*, 629:112–137, 2021.

[32] F. R. Gantmacher, The theory of matrices. Vol. 1. Chelsea Publishing Co., New York 1959

[33] É. Gaudron and G. Rémond, Lemmes de Siegel df'évitement, *Acta Arith.* 154, no. 2 (2012), 125–136.

[34] É. Gaudron and G. Rémond, Corps de Siegel, *J. reine angew. Math.* 726 (2017), 187–247.

[35] P. Gordan, Über einige Anwendungen diophantischer Approximationen, *Math. Annalen*, 7, (1873), 443–448.

[36] P. Gordan. Uber den grossten gemeinsamen factor. *Math. Ann.*, 7 (1873), pp. 443–448.

[37] P. M. Gruber and C. G. Lekkerkerker, Geometry of Numbers. North-Holland Publishing Co., 1987.

[38] F. J. Grunewald and D. Segal. How to solve a quadratic equation in integers. *Math. Proc. Cambridge Philos. Soc.*, 89 (1981) no. 1, pp. 1–5.

[39] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers. 5th edition. The Clarendon Press, Oxford University Press, New York, 1979.

[40] D. R. Heath-Brown. Diophantine approximation with square-free numbers. *Math. Z., 187* (1984), no. 3, 335–344

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

[42] J.-W. M. van Ittersum. Quantitative results on Diophantine equations in many variables. *Acta Arith.*, 194 (2020), pp. 219–240.

[43] Y. Jiang, Y. Deng and Y. Pan. Covering radius of two-dimensional lattices. *J. Systems Sci. Math. Sci.*, 32 (2012), no. 7, pp. 908–914.

[44] S. V. Konyagin. On the recovery of an integer vector from linear measurements. *Mat. Zametki*, 104(6):863–871, 2018.

[45] S. V. Konyagin and B. Sudakov. An extremal problem for integer sparse recovery. *Linear Algebra Appl.*, 586:1–6, 2020.

[46] D. W. Masser. Search bounds for Diophantine equations. A panorama of number theory or the view from Baker's garden (Zürich, 1999), 247–259, Cambridge Univ. Press, Cambridge, 2002.

[47] J. Martinet, Perfect lattices in Euclidean spaces, Springer-Verlag, 2003.

[48] Y. V. Matiyasevich. The Diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR*, 191 (1970), pp. 279–282.

[49] G. Maze, J. Rosenthal, and U. Wagner. Natural density of rectangular unimodular integer matrices. *Linear Algebra Appl.*, 434:1319–1324, 2011.

[50] D. Micciancio and S. Goldwasser. Complexity of lattice problems: a cryptographic perspective. *Kluwer Academic Publishers*, Boston, 2002.

[51] J. E. Nymann. On the probability that k positive integers are relatively prime. *J. Number Theory*, 4:469–473, 1972.

[52] K. F. Roth, Rational approximations to algebraic numbers, *Mathematika,* 2 (1955), 1–20; corrigendum, *ibid* 168.

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

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

[55] W. M. Schmidt, On heights of algebraic subspaces and Diophantine approximations, *Ann. Math.*, 85, (1967), 430–472.

[56] W. M. Schmidt, Diophantine approximations and Diophantine equations, Springer-Verlag, 1991.

[57] C. L. Siegel. Über einige Anwendungen diophantischer Approximationen. *Abh. Preuss. Akad. Wiss. Phys. Math. Kl.*, (1929), pp. 41–69

[58] J. Silverman The Arithmetric of Elliptic Curves Springer-Verlag, 2009.

[59] 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 (1990), pp. 493–528.

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

[61] J. L. Thunder Higher dimensional analogues of Hermite's constant *Mich. Math. Jour.*, 45 (1998), 301–314

[62] J. D. Vaaler, The Best Constant in Siegel's Lemma, Monatsh. Math. 140, (2003), 71–89.

[63] C. G. Pinner and J. D. Vaaler. The number of irreducible factors of a polynomial. I. *Trans. Amer. Math. Soc.*, 339 (1993) no. 2, pp. 809–834.