The Dissertation Committee for Leonid Eugene Fukshansky
Certifies that this is the approved version of the following dissertation:

ALGEBRAIC POINTS OF SMALL HEIGHT WITH ADDITIONAL ARITHMETIC CONDITIONS

Committee:

Jeffrey D. Vaaler, Supervisor

Fernando Rodriguez-Villegas

David J. Saltman

John T. Tate

J. Felipe Voloch

Jeffrey L. Thunder

ALGEBRAIC POINTS OF SMALL HEIGHT WITH ADDITIONAL ARITHMETIC CONDITIONS

by

LEONID EUGENE FUKSHANSKY, B.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2004

To my parents Eugene and Polina and my wife Eugenia

Acknowledgments

I would first of all like to express my deepest gratitude to Professor Jeff Vaaler, whom I was extremely lucky to have as my advisor. This dissertation would have never been possible without his advice and guidance. The problems treated here were initially suggested by him. It is largely due to his support, patience, and our numerous wonderful conversations that I was able to succeed solving them.

I would also like to thank Professors John Tate, Fernando Rodriguez-Villegas, Felipe Voloch, and David Saltman for agreeing to serve on my Ph.D. committee, and for all the mathematics I have learned from them. I also want to thank Professors Paula Cohen, Sinnou David, and Bruce Reznick for their useful comments, and the referee of my paper [13] for pointing out possible simplifications of some proofs in Chapter 3. I thank Professor Jeff Thunder for agreeing to be an outside member on my dissertation committee.

Last, but not least, I would like to thank my family, friends, and all the people around me, whose constant support and understanding created the pleasant environment that I enjoyed all through my graduate school. These words are especially directed at my parents and my wife, to whom this dissertation is dedicated.

ALGEBRAIC POINTS OF SMALL HEIGHT WITH ADDITIONAL ARITHMETIC CONDITIONS

Publication No. _____

Leonid Eugene Fukshansky, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Jeffrey D. Vaaler

We treat a few related problems about the existence of algebraic points of small height that satisfy certain arithmetic conditions. All bounds on height of points in question are explicit. First we prove the existence of a small-height point over a fixed number field outside of a collection of subspaces; this includes a generalization and a converse of the celebrated Siegel's Lemma, [5]. Next, assuming that a quadratic form has a zero outside of a collection of subspaces over a fixed number field, we prove the existence of such a zero of bounded height; this generalizes a result of Masser, [19]. A corollary of this is an extension of Cassels' famous theorem on small zeros of quadratic forms (see [7]) to small non-singular zeros of quadratic forms. Finally, we prove a theorem about existence of small-height zeros of homogeneous polynomials of arbitrary degree over $\overline{\mathbb{Q}}$ outside of a collection of subspaces. This direction is similar in spirit to the so-called "absolute" results like, for instance, the absolute version of Siegel's Lemma of Roy and Thunder, [24].

Table of Contents

Chapter 1

Introduction

1.1 Brief overview

An important direction in the area of Diophantine Approximations is the range of problems that are formulated with the use of various height functions. A particularly interesting selection of such problems is connected with effective arithmetic geometry. Write $\overline{\mathbb{Q}}$ for the algebraic closure of $\mathbb{Q}$, and let $\mathcal{A}$ be a set of arithmetic conditions imposed on points of $\overline{\mathbb{Q}}^N$. Let $S(\mathcal{A})$ be the set of all points in $\overline{\mathbb{Q}}^N$ that satisfy conditions $\mathcal{A}$. One can state the following two general problems.

Problem 1. *Decide whether the set $S(\mathcal{A})$ is empty or not.*

Problem 2. *Assuming that $S(\mathcal{A})$ is not empty, prove the existence of a point in $S(\mathcal{A})$ of bounded height with explicit bounds on the height.*

A typical example of Problem 1 would be an inquiry as to whether a given polynomial F in N variables has non-trivial zeros over a given number field K. In this case there are two arithmetic conditions that a point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ needs to satisfy, namely we must have $\boldsymbol{x} \in K^N$ and $F(\boldsymbol{x}) = 0$. Notice that this is an example of a "non-effective" problem, since it is usually very difficult

to solve a polynomial equation, so most proofs that a set $S(\mathcal{A})$ like this is not empty are non-constructive. On the other hand, Problem 2 in this situation would require us to provide an explicit bound for the height of a point $\boldsymbol{x}$ like this. This is an example of an "effective" problem, since its solution provides an explicit "search-bound" for such points. This follows by the celebrated theorem of Northcott [20], which states that for any two positive numbers C_1 and C_2 the set of points $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ of degree no larger than C_1 and height no larger than C_2 is finite.

We will mostly be interested in examples of Problem 2. The first result along these lines is Siegel's Lemma. We present it in its most general formulation, which is due to Bombieri and Vaaler. Let K be a number field of degree d, O_K be its ring of integers, and $\mathcal{D}_K$ be its discriminant. We write $\mathcal{H}$ for an appropriate height function to be explicitly defined later.

Theorem 1.1.1 ([5]). *Let V be an M-dimensional subspace of K^N, $M < N$. Then there exists a basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_M \in O_K^N$ for V such that*

$$\prod_{i=1}^{M} \mathcal{H}(\boldsymbol{x}_i) \leq \left\{ N |\mathcal{D}_K|^{1/d} \right\}^{M/2} \mathcal{H}(V). \tag{1.1}$$

Notice that this in particular means that there exists a non-zero point $\boldsymbol{x} \in V$ such that

$$\mathcal{H}(\boldsymbol{x}) \leq \left\{ N |\mathcal{D}_K|^{1/d} \right\}^{1/2} \mathcal{H}(V)^{1/M}. \tag{1.2}$$

The exponent on $\mathcal{H}(V)$ in the upper bound of Theorem 1.1.1 is best possible, however the constant is not. The best possible constant for Siegel's Lemma was recently obtained by Vaaler in [35]. It turns out to be the generalized Hermite's constant as introduced by Thunder in [32].

Notice that if we are willing to consider the subspace V of Theorem 1.1.1 over $\overline{\mathbb{Q}}$ and look for a basis of small height in say any extension E of K, then the bound of (1.1) is not very good any longer since the constant grows as power of $|\mathcal{D}_E|$, which can be arbitrarily large. In this case one would want an "absolute" version of Siegel's Lemma, i.e. a bound that would not depend on the choice of a number field. Such a result was produced by Thunder and Roy.

Theorem 1.1.2 ([24]). *Let V be an M-dimensional subspace of $\overline{\mathbb{Q}}^N$ defined over K, $M < N$, and let $\epsilon > 0$. Then there exists a basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_M$ for V over $\overline{\mathbb{Q}}$ (which depends on the choice of ϵ) such that*

$$\prod_{i=1}^{M} \mathcal{H}(\boldsymbol{x}_i) \leq \left(2^{\frac{M(M-1)}{2}} + \epsilon\right) \mathcal{H}(V). \tag{1.3}$$

Another example of a result along the lines of Problem 2 is Cassels' theorem on small zeros of quadratic forms. We write H for an appropriate height function to be explicitly defined later.

Theorem 1.1.3 ([7]). *Let F be a quadratic form with integral coefficients in $N \geq 2$ variables which is isotropic over $\mathbb{Q}$. Then it has a non-trivial integral zero $\boldsymbol{x}$ with*

$$H(\boldsymbol{x}) \leq \left\{\sqrt{3}N\right\}^{N-1} H(F)^{\frac{N-1}{2}}. \tag{1.4}$$

The exponent in the upper bound of (1.4) is best possible as Cassels illustrates by an example due to Kneser. The constant in our presentation of Cassels' theorem follows from Lemma 8.1 on p. 87 of [9]. Cassels' result was generalized to number fields with the same exponent but different constant in the upper bound by Raghavan in [23]. More recently, Masser in [19] extended Cassels'

result to inhomogeneous quadratic polynomials over $\mathbb{Q}$ by means of considering small rational zeros of a rational quadratic form in one more variable with the condition that this additional variable is not zero.

Theorem 1.1.4 ([19]). *Let F be a quadratic form with rational coefficients in $N+1 \geq 2$ variables. Suppose that there exists $\boldsymbol{x} = (x_0, ..., x_N) \in \mathbb{Q}^{N+1}$ such that $F(\boldsymbol{x}) = 0$ and $x_0 \neq 0$, then there exists such $\boldsymbol{x}$ with*

$$H(\boldsymbol{x}) \leq \left\{\sqrt{3}(N+1)\right\}^{N+1} H(F)^{\frac{N+1}{2}}. \tag{1.5}$$

This implies that if an inhomogeneous quadratic polynomial in N variables with rational coefficients has a rational zero, then it has a rational zero whose height is bounded as in (1.5). The exponent in the upper bound of (1.5) is best possible as demonstrated by an example of Masser.

There are no further known examples of solutions of Problem 2 for polynomials of higher degree over a fixed number field. In fact, this seems to be a particularly difficult problem. In this dissertation we produce some further examples of solutions of Problem 2 with arithmetic conditions extending or complementing those of stated above well-known results. First we set some notation.

1.2 Definitions and notation

Throughout this dissertation, let K be a number field of degree d over $\mathbb{Q}$, O_K its ring of integers, $\mathcal{D}_K$ its discriminant, and $M(K)$ its set of places. For each place $v \in M(K)$ we write K_v for the completion of K at v and let

$d_v = [K_v : \mathbb{Q}_v]$ be the local degree of K at v, so that for each $u \in M(\mathbb{Q})$

$$\sum_{v \in M(K), v|u} d_v = d. \tag{1.6}$$

For each place $v \in M(K)$ we define the absolute value $\| \ \|_v$ to be the unique absolute value on K_v that extends either the usual absolute value on $\mathbb{R}$ or $\mathbb{C}$ if $v|\infty$, or the usual p-adic absolute value on $\mathbb{Q}_p$ if $v|p$, where p is a prime. We also define the second absolute value $| \ |_v$ for each place v by $|a|_v = \|a\|_v^{d_v/d}$ for all $a \in K$. Then for each non-zero $a \in K$ the *product formula* reads

$$\prod_{v \in M(K)} |a|_v = 1. \tag{1.7}$$

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

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

$$H_v(\boldsymbol{x}) = \max_{1 \leq i \leq N} |x_i|_v, \tag{1.8}$$

for each $\boldsymbol{x} \in K_v^N$. If $v|\infty$, we also define another local height $\mathcal{H}_v$ on K_v^N by

$$\mathcal{H}_v(\boldsymbol{x}) = \left(\sum_{1 \leq i \leq N} \|x_i\|_v^2 \right)^{1/2}, \tag{1.9}$$

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

$$H(\boldsymbol{x}) = \prod_{v \in M(K)} H_v(\boldsymbol{x}) \tag{1.10}$$

$$\mathcal{H}(\boldsymbol{x}) = \prod_{v \nmid \infty} H_v(\boldsymbol{x}) \times \prod_{v|\infty} \mathcal{H}_v(\boldsymbol{x})^{d_v/d} \tag{1.11}$$

$$h(\boldsymbol{x}) = \prod_{v \in M(K)} \max\{1, H_v(\boldsymbol{x})\} \tag{1.12}$$

for each $\boldsymbol{x} \in K^N$. We refer to H as *homogeneous height with sup-norms*, to $\mathcal{H}$ as *homogeneous height with L_2-norms at infinite places*, and to h as *inhomogeneous height*. The following inequalities for each $x \in K^N$ can be immediately seen:

$$H(\boldsymbol{x}) \leq \mathcal{H}(\boldsymbol{x}) \leq \sqrt{N} H(\boldsymbol{x}), \quad H(\boldsymbol{x}) \leq h(\boldsymbol{x}). \tag{1.13}$$

We will also need a height function on algebraic numbers, defined by

$$h(\alpha) = h(1, \alpha) = \prod_{v \in M(K)} \max\{1, |\alpha|_v\} \tag{1.14}$$

for each $\alpha \in K$.

It is not difficult to see that if $K = \mathbb{Q}$ and $\boldsymbol{x} = \left(\frac{x_1}{x_0}, ..., \frac{x_N}{x_0}\right) \in \mathbb{Q}^N$ with all x_i in $\mathbb{Z}$, then

$$H(\boldsymbol{x}) = \max\{|x_0|, |x_1|, ..., |x_N|\}, \tag{1.15}$$

and so in case $\boldsymbol{x} \in \mathbb{Z}^N$ we will often write $|\boldsymbol{x}| = \max\{|x_1|, ..., |x_N|\}$ instead of $H(\boldsymbol{x})$.

We will also need the following well known property of height functions.

Lemma 1.2.1. *Let $\boldsymbol{x}, \boldsymbol{y} \in K^N$, and m, n be positive integers, then*

$$H(m\boldsymbol{x} \pm n\boldsymbol{y}) \leq h(m\boldsymbol{x} \pm n\boldsymbol{y}) \leq (m+n) h(\boldsymbol{x}) h(\boldsymbol{y}). \tag{1.16}$$

Proof. Notice that for every $v \nmid \infty$, $|m|_v, |n|_v \leq 1$, and so

$$\begin{aligned} \max\{1, H_v(m\boldsymbol{x} \pm n\boldsymbol{y})\} &\leq \max\{1, H_v(\boldsymbol{x}), H_v(\boldsymbol{y})\} \\ &\leq \max\{1, H_v(\boldsymbol{x})\} \max\{1, H_v(\boldsymbol{y})\}. \end{aligned}$$

Also, for every $v|\infty$,

$$\begin{aligned} H_v(m\boldsymbol{x} \pm n\boldsymbol{y}) &\leq mH_v(\boldsymbol{x}) + nH_v(\boldsymbol{y}) \\ &\leq (m+n)\max\{H_v(\boldsymbol{x}), H_v(\boldsymbol{y})\} \\ &\leq (m+n)\max\{1, H_v(\boldsymbol{x})\}\max\{1, H_v(\boldsymbol{y})\}, \end{aligned}$$

and therefore

$$\begin{aligned} \max\{1, H_v(m\boldsymbol{x} \pm n\boldsymbol{y})\} &\leq \max\{1, (m+n)\max\{1, H_v(\boldsymbol{x})\}\max\{1, H_v(\boldsymbol{y})\}\} \\ &= (m+n)\max\{1, H_v(\boldsymbol{x})\}\max\{1, H_v(\boldsymbol{y})\}. \end{aligned}$$

The lemma follows by taking a product. □

Notice that all of the height functions we defined are *absolute*, that is independent of the field of definition. This is due to our normalization of absolute values $|\ |_v$ for each $v \in M(K)$ with the exponents $\frac{d_v}{d}$. Hence for any vector $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ we can define height of $\boldsymbol{x}$ over any number field that contains coordinates of $\boldsymbol{x}$ and all such definitions will be the same. Another important observation is that due to the product formula the height functions can be viewed as functions on the corresponding projective space $\mathbb{P}^N(K)$.

We extend all heights to polynomials by viewing them as height functions of the coefficient vector of a given polynomial. We also define two different heights on matrices. Suppose $A = (\boldsymbol{\alpha}_1\ ...\ \boldsymbol{\alpha}_M) = (a_{nm})$ is an $N \times M$ matrix with entries in K. We define the height functions $H_*(A)$ and $\mathcal{H}_*(A)$ on matrices by extending the heights H and $\mathcal{H}$ to matrices, i.e. by viewing each such matrix A as a vector $(a_{11}, ..., a_{NM})$ in K^{NM}.

We also define the heights $H(A)$ and $\mathcal{H}(A)$ by

$$H(A) = H(\boldsymbol{\alpha}_1 \wedge ... \wedge \boldsymbol{\alpha}_M),\ \mathcal{H}(A) = \mathcal{H}(\boldsymbol{\alpha}_1 \wedge ... \wedge \boldsymbol{\alpha}_M) \tag{1.17}$$

where $\wedge$ stands for the wedge product of two vectors, and so $\boldsymbol{\alpha}_1 \wedge ... \wedge \boldsymbol{\alpha}_M$ is a vector in $\bigwedge^M \left(K^N\right) \cong K^L$, where $L = \binom{N}{M}$. Using this last definition we can extend the notion of height to subspaces of K^N. Let V be an M-dimensional subspace of K^N where $1 \leq M \leq N$, and let $\boldsymbol{x}_1, ..., \boldsymbol{x}_M$ be a basis for V over K. Write $\mathcal{X} = (\boldsymbol{x}_1\ ...\ \boldsymbol{x}_M)$ for the basis matrix of V, and define heights of V by

$$H(V) = H(\mathcal{X}),\ \mathcal{H}(V) = \mathcal{H}(\mathcal{X}). \tag{1.18}$$

It is not difficult to see that this definition is independent of the choice of the basis. Suppose that $\boldsymbol{y}_1, ..., \boldsymbol{y}_M$ is another basis for V over K, and write $\mathcal{Y} = (\boldsymbol{y}_1\ ...\ \boldsymbol{y}_M)$ for the corresponding basis matrix. Then there exists a non-singular matrix U with entries in K such that $\mathcal{Y} = \mathcal{X}U$, and so

$$\boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_M = (\det U)\ \boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_M, \tag{1.19}$$

therefore

$$\begin{aligned} H(\boldsymbol{y}_1 \wedge ... \wedge \boldsymbol{y}_M) &= H(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_M) \prod_{v \in M(K)} |\det U|_v \\ &= H(\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_M), \end{aligned} \tag{1.20}$$

by the product formula (1.7), and similarly for $\mathcal{H}$. Thus $H(V)$ and $\mathcal{H}(V)$ are well-defined. In general, notice that height functions are well-defined on projective spaces, since height of $\boldsymbol{x}$ is the same as height of $a\boldsymbol{x}$ for any non-zero $a \in K$ by the product formula.

Another equivalent way to define heights on matrices and hence on subspaces $H(V)$ is the following. Let again $\mathcal{X} = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_J)$ be a basis matrix for a subspace V of K^N of dimension M as above. Let $\mathcal{I}$ be the collection of all subsets $I \subseteq \{1, ..., N\}$ of cardinality M. There is a total of $\binom{N}{M}$ such subsets. For each such $I \in \mathcal{I}$, let $\mathcal{X}_I$ be the $M \times M$ submatrix of $\mathcal{X}$ whose rows are those rows of $\mathcal{X}$ which are indexed by I. The vector of *Grassmann coordinates* of V is

$$Gr(V) = (\det(\mathcal{X}_I))_{I \in \mathcal{I}} \in K^{\binom{N}{M}}, \tag{1.21}$$

which is easily seen to be precisely the vector $\boldsymbol{x}_1 \wedge ... \wedge \boldsymbol{x}_M$. As discussed above, the height of this vector is an invariant of V, i.e. it does not depend on the choice of a basis. Therefore

$$H(V) = H(\mathcal{X}) = H(Gr(V)), \quad \mathcal{H}(V) = \mathcal{H}(\mathcal{X}) = \mathcal{H}(Gr(V)). \tag{1.22}$$

We will now describe an important *duality principle* as applied to heights on subspaces. Suppose that V as above is an M-dimensional subspace of K^N. There are two different ways to describe V. First let $\boldsymbol{x}_1, ..., \boldsymbol{x}_M$ be a basis for V in K^N, and write $\mathcal{X} = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_M)$ for the $N \times M$ basis matrix. Then $V = \mathcal{X}K^M$. On the other hand, there exists an $(N - M) \times N$ matrix A with entries in K such that V is the nullspace of A, that is

$$V = \{\boldsymbol{x} \in K^N : A\boldsymbol{x} = \boldsymbol{0}\}. \tag{1.23}$$

If $I \subseteq \{1, ..., N\}$, $|I| = M$, and $\overline{I} = \{1, ..., N\} \setminus I$ is its complement, then by a duality theorem of Brill-Gordan [14] (see also Theorem 1 on p. 294 of [15])

there exists a non-zero constant $\gamma \in K$

$$\det(\mathcal{X}_I) = (-1)^{\varepsilon(I)}\gamma \det({}_{\overline{I}}A), \tag{1.24}$$

where $\det(\mathcal{X}_I)$ and $\det({}_{\overline{I}}A)$ stand for the corresponding Grassmann coordinates of $\mathcal{X}$ and A respectively, and $\varepsilon(I) = \sum_{i\in I} i$. Then

$$H_v(\mathcal{X}) = |\gamma|_v H_v(A)\ \forall v \in M(K), \quad \mathcal{H}_v(\mathcal{X}) = |\gamma|_v \mathcal{H}_v(A)\ \forall v|\infty. \tag{1.25}$$

Therefore by definition of $H(V)$, $\mathcal{H}(V)$, and the product formula, we obtain the following important principle.

Lemma 1.2.2. *Let V be an M-dimensional subspace of K^N with an $N \times M$ basis matrix $\mathcal{X}$, so that V is also the nullspace of an $(N-M)\times N$ matrix A with entries in K, that is*

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

Then

$$H(V) = H(\mathcal{X}) = H(A),\ \mathcal{H}(V) = \mathcal{H}(\mathcal{X}) = \mathcal{H}(A). \tag{1.26}$$

We are now ready to state the main results of this dissertation.

1.3 Statement of main results

Let K be a number field, $N \geq 1$, $M \geq 0$ integers, and let $V_1, ..., V_M$ be a collection of proper subspaces of K^N. In this dissertation we study algebraic points of relatively small height outside of such a collection of subspaces that satisfy additional arithmetic conditions. In particular we prove the existence of algebraic points of bounded height that are outside of $\bigcup_{i=1}^{M} V_i$ and:

(I) In a subspace of K^N,

(II) In a quadratic variety over K,

(III) In a hypersurface over a finite extension of K.

Thus the condition that a point is outside of a collection of subspaces is the defining one here.

In Chapter 2 we study problem (I). It can be viewed as a generalization of Siegel's Lemma with additional conditions. Suppose that W is a non-zero subspace of K^N. Siegel's Lemma implies the existence of a non-zero point of bounded height in W. We consider a more general situation. Let $V_1, ..., V_M$ be proper subspaces of W. We prove the existence of a point of bounded height in $W \setminus \bigcup_{i=1}^{M} V_i$. Our main result in this direction reads as follows.

Theorem 1.3.1. *Let K be a number field of degree d. Let $N \geq 2$ be an integer, and let W be a non-zero subspace of K^N. Let $V_1, ..., V_M \subseteq W$ be proper non-zero subspaces of W. There exists a point $\boldsymbol{x} \in \left(W \setminus \bigcup_{i=1}^{M} V_i\right) \cap O_K^N$ such that*

$$H(\boldsymbol{x}) \leq \mathcal{C}_1 H(W) \left\{ \left(\sum_{i=1}^{M} \frac{\mathcal{C}_2^i}{H(V_i)^d} \right)^{\frac{1}{d}} + M^{\frac{1}{d+1}} \right\}, \tag{1.27}$$

where the constants $\mathcal{C}_1$ and $\mathcal{C}_2^i$ for each $1 \leq i \leq M$ depend on K, N, and either dimension of W or dimension of V_i respectively.

We provide explicit values for the constants $\mathcal{C}_1$ and C_2 in Chapter 2. The dependence of the upper bound in (1.27) on M and on heights of the subspaces is optimal, as demonstrated for instance by the classical case ($K = \mathbb{Q}$), which we consider separately producing sharper constants. We also identify two

important extremal cases of this problem: $M = 0$ and $W = K^N$. If $M = 0$, we produce a simple version of Siegel's Lemma with an upper bound which essentially agrees (up to a constant) with that of Bombieri and Vaaler (see (1.2)). Consider the other extremal case, that is suppose that $W = K^N$, and assume that $\dim_K(V_i) = N - 1$ for each $1 \leq i \leq M$. Then $V_1, ..., V_M$ can be viewed as nullspaces of linear forms $L_1, ..., L_M$ with coefficients in O_K. By the duality principle of Lemma 1.2.2 we have $H(V_i) = H(L_i)$ for each i, $1 \leq i \leq M$. Then we prove the existence of a point of small height outside of a collection of subspaces (or equivalently at which a collection of linear forms does not vanish). This result can be thought of as a converse of Siegel's Lemma. We produce slightly sharper constants in the upper bound in this particular case. We also produce some basic bounds for this problem that depend only on the number of linear forms, not on their heights.

In the last section of Chapter 2 we study a different problem related to Siegel's Lemma. We apply a classical version of Siegel's Lemma to produce an orthogonal integral basis of small height for a certain inner-product space. Although the inner-product that we work with is of particular interest, the same simple method can be applied to any quadratic space.

In Chapter 3 we treat a version of problem (II). Let $F(\boldsymbol{X}) \in K[X_0, ..., X_N]$ be a quadratic form and $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X}) \in K[X_0, ..., X_N]$ be M linear forms in $N + 1$ variables with coefficients in a number field K. Suppose that there exists a point in K^{N+1} at which the quadratic form vanishes and the linear forms do not. We want to prove the existence of such a point of bounded height In other words, we are searching for a small-height zero of a quadratic

form outside of a collection of subspaces. This problem is logically related to the problem of Chapter 2 also in the way that we use a basic bound of Chapter 2 (one that depends only on the number of linear forms) in the proof of the main theorem of Chapter 3. The main result of Chapter 3 is the following.

Theorem 1.3.2. *Suppose that there exists a point* $\boldsymbol{x} \in \mathbb{P}^N(K)$ *such that* $F(\boldsymbol{x}) = 0$*, and* $L_i(\boldsymbol{x}) \neq 0$ *for each* $1 \leq i \leq M$*. Then there exists such a point* $\boldsymbol{x}$ *with*

$$H(\boldsymbol{x}) \leq \mathcal{C}_3 H(F)^{\frac{N+2M}{2}+(M-1)(N+2)}, \tag{1.28}$$

as well as

$$H(\boldsymbol{x}) \leq \mathcal{C}_3 H(F)^{\frac{N+1}{2}+(M-1)(N+2)} \prod_{i=1}^{M} H(L_i)^{\frac{(2M-1)N}{M}}, \tag{1.29}$$

and finally

$$H(\boldsymbol{x}) \leq \mathcal{C}_3 H(F)^{\frac{2N+2M+1}{4}+(M-1)(N+2)} \prod_{i=1}^{M} H(L_i)^{\frac{(2M-1)N}{2M}}, \tag{1.30}$$

where the constant $\mathcal{C}_3$ *depends on* K*,* N*, and* M *only.*

We provide an explicit value for the constant $\mathcal{C}_3$ in Chapter 3. Notice that Theorem 1.3.2 is a generalization of Masser's Theorem 1.1.4: in the case $K = \mathbb{Q}$, $M = 1$, and $L_1(\boldsymbol{X}) = X_0$, Masser's upper bound (up to a constant) follows from (1.29).

A simple, but interesting corollary of Theorem 1.3.2 is the following.

Corollary 1.3.3. *Let* $F(\boldsymbol{X})$ *be a quadratic form in* $N + 1$ *variables with coefficients in the number field* K*, as above. Let*

$$\mathcal{V}_K(F) = \left\{ \boldsymbol{t} \in K^{N+1} : F(\boldsymbol{t}) = 0 \right\}.$$

Suppose that there exists a non-singular point $\boldsymbol{x} \in \mathcal{V}_K(F)$. *Then there exists such a point* $\boldsymbol{x}$ *with*

$$H(\boldsymbol{x}) \leq \mathcal{C}_4 H(F)^{\frac{N}{2}}, \tag{1.31}$$

where the constant $\mathcal{C}_4$ *depends on* K *and* N *only.*

We provide an explicit value for the constant $\mathcal{C}_4$ in Chapter 3. Notice that Corollary 1.3.3 is an extension of Cassels' Theorem 1.1.3: it guarantees the existence of a *non-singular* point with the same (up to a constant) bound on the height as in Theorem 1.1.3, provided that the given quadratic variety contains non-singular points.

A natural next step would be to prove the existence of a zero of small height over a fixed number field for a polynomial of degree $M > 2$, assuming that it has non-trivial zeros over this number field. This, however, seems to be a particularly difficult problem: much less is known about hypersurfaces of higher degree than about quadrics; in particular, the beautiful geometrical construction of Cassels [7] that produces a new point in the quadratic variety from a given one does not have a higher degree analogue. On the other hand, the problem becomes accessible if we relax the condition that the point in question must lie in the fixed number field K, and instead search in extensions of K of degree at most M. We study this problem in Chapter 4.

The main result of Chapter 4 treats problem (III). Let $F(\boldsymbol{X}) \in K[X_1, ..., X_N]$ be a homogeneous polynomial of degree M in N variables with coefficients in K, and let $A \in GL_N(K)$. Assuming a necessary algebraic condition, we prove the existence of a point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ of bounded height with $\deg_K(\boldsymbol{x}) =$

$[K(x_1, ..., x_N) : K] \leq M$ so that $F(\boldsymbol{x}) = 0$ and $A\boldsymbol{x} \in (\overline{\mathbb{Q}}^{\times})^N$. This can be restated as follows. Let $A = (a_{ij})_{1 \leq i,j \leq N}$, and define N linear forms $L_i(\boldsymbol{X}) = \sum_{j=1}^{N} a_{ij} X_j$, $1 \leq i \leq N$. Then we are proving the existence of a zero of F of bounded height at which none of the linear forms vanish. In other words, we are still (in a certain form) preserving the original arithmetic conditions of Chapter 2, and complementing them with a vanishing condition on a homogeneous polynomial of arbitrary degree. The actual theorem reads like this.

Theorem 1.3.4. *Let $F(X_1, ..., X_N)$ be a homogeneous polynomial in $N \geq 2$ variables of degree $M \geq 1$ over a number field K with $[K : \mathbb{Q}] = d$, and let $A \in GL_N(K)$. Then either there exists a non-zero point $\boldsymbol{y} \in K^N$ such that $F(\boldsymbol{y}) = 0$ and*

$$\mathcal{H}(\boldsymbol{y}) \leq \mathcal{H}_*(A^{-1}),$$

or there exists $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ with $\deg_K(\boldsymbol{x}) \leq M$ such that $F(\boldsymbol{x}) = 0$, $A\boldsymbol{x} \in (\overline{\mathbb{Q}}^{\times})^N$, and

$$\mathcal{H}(\boldsymbol{x}) \leq \mathcal{C}_5 \mathcal{H}_*(A^{-1})^2 \mathcal{H}(F)^{1/M}, \tag{1.32}$$

where the constant $\mathcal{C}_5$ depends on N and M only.

We provide an explicit value for the constant $\mathcal{C}_5$ in Chapter 4. Notice that the assertion of Theorem 1.3.4, as explained in Chapter 4, can be interpreted as follows. Consider the set of points $S = \{A^{-1}\boldsymbol{e}_i : 1 \leq i \leq N\} \subset K^N$, where $\boldsymbol{e}_1, ..., \boldsymbol{e}_N$ are the standard basis vectors. Then either F is a very "special" polynomial that has a zero in S, or F is a generic polynomial, and then it has a zero of bounded height over $\overline{\mathbb{Q}}$ which is outside of the union of nullspaces of row-vectors of A. The condition for F to be "generic" is important for the exponent on $H(F)$ in the upper bound (1.32) to be $1/M$; assuming only that F

is not a monomial one can obtain a rougher upper bound with exponent 1 on $H(F)$ in (1.32). Chapter 4 is concluded with a brief discussion of small-height simultaneous zeros of a collection of polynomials of arbitrary degrees over $\overline{\mathbb{Q}}$.

Chapter 2

Points of small height outside of a collection of subspaces

2.1 Introduction

The name Siegel's Lemma is usually used to denote results about small-height solutions of a system of linear equations. Such a result in a simple form was first proved by Thue in 1909 ([30], pp. 288-289) using the Dirichlet's box principle. Siegel ([27], Bd. I, p. 213, Hilfssatz) was the first to formally state this principle in the classical case. He proved the following.

Theorem 2.1.1 ([27]). *Let $A = (a_{ij})$ be an $M \times N$, matrix with integer entries, where $M < N$ and rank of A is M. Let $|A| = \max_{1\le i\le M, 1\le j\le N} |a_{ij}|$, then there exists a non-zero point $\boldsymbol{x} \in \mathbb{Z}^N$ such that $A\boldsymbol{x} = \mathbf{0}$ and*

$$|\boldsymbol{x}| \le (N|A|)^{\frac{M}{N-M}}. \tag{2.1}$$

Notice that by the duality of Lemma 1.2.2, we can view results of this kind as statements about points of small height in a given subspace, namely the nullspace of A. This principle in its general form due to Bombieri and Vaaler is represented by our Theorem 1.1.1. Results of this sort were originally treated as important technical lemmas used in transcendental number theory and diophantine approximations for the purpose of constructing a certain auxiliary

polynomial (see [5] and [3] for more information). Nowadays it has evolved as an important result in its own right; it can, for instance, be viewed in the context of Problem 2 that we stated in Chapter 1.

In this chapter we consider a generalization of this problem. Let K be a number field, and let W be a subspace of K^N, $N \geq 1$. Let $V_1, ..., V_M$ be proper subspaces of W. We want to prove the existence of a non-zero point of small height in $(W \cap O_K^N) \setminus \bigcup_{i=1}^{M} V_i$ providing an explicit upper bound on the height of such a point. We produce a version of Siegel's Lemma as a special case of this problem when $M = 0$. We separately discuss another interesting special case of our main result. Suppose that $W = K^N$, and let $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X})$ be M linear forms in N variables with coefficients in K. We want to prove the existence of a point $\boldsymbol{x}$ in O_K^N of relatively small height such that $L_i(\boldsymbol{x}) \neq 0$ for every $i = 1, ..., M$ (i.e. $\boldsymbol{x}$ is outside of the union of nullspaces of linear forms). Hence this less general problem can also be restated in the following form.

Problem 3. *Given a collection of $(N-1)$-dimensional subspaces $V_1, ..., V_M$ in K^N, prove the existence of a point of small height outside the union $\bigcup_{i=1}^{M} V_i$ with an explicit bound on height.*

In section 2.2 we consider a certain more general version of Problem 3 and produce some basic results. Namely, given a polynomial in N variables of degree M we prove existence of an integral (and algebraic) point, whose height is bounded above by an expression that depends only on N and M, at which this polynomial does not vanish. In sections 2.3 and 2.4 we present some results on the problem of counting integer lattice points in closed boxes in $\mathbb{R}^N$.

Results of section 2.3 are due to Vaaler and presented here with his permission for the purpose of self-containment. In section 2.5 we consider the classical case $K = \mathbb{Q}$. In section 2.6 we consider the general problem in the number field case and prove Theorem 2.6.1, which is an effective version of Theorem 1.3.1. This is the main result of this chapter. In section 2.7 we exhibit a simple application of our results by relating them to a discrete analogue of the Tarski plank problem.

In section 2.8 we consider a different problem, which is an application of Siegel's Lemma. We consider a certain inner-product space, which is quite important in harmonic analysis, and produce an orthogonal integral basis of small height for it. The simple technique we use in applying Siegel's lemma can easily be used to obtain similar results for other quadratic spaces as well.

2.2 Basic bounds

In this section we prove the existence of integral (and algebraic) points of small height (and length) at which a given polynomial (homogeneous and not) in N variables of degree M does not vanish. Our bounds on height depend on M and N only.

First let

$$\mathcal{M}' = \mathcal{M}'(N, M) = \{\boldsymbol{m} \in \mathbb{Z}_{\geq 0}^N : m_1 + ... + m_N \leq M\}.$$

Then let

$$F(X_1, ..., X_N) = \sum_{\boldsymbol{m} \in \mathcal{M}'} f_{\boldsymbol{m}} X_1^{m_1} ... X_N^{m_N} \in \mathbb{C}[X_1, ..., X_N],$$

be a polynomial (not necessarily homogeneous) in $N \geq 1$ variables of degree $M \geq 1$ with coefficients in $\mathbb{C}$. In the next lemma $\mathbb{C}$ can be replaced with any algebraically closed field in which F has its coefficients. We write $\deg_{X_i}(F)$ for degree of F in the variable X_i for each $1 \leq i \leq N$, and $\deg(F)$ for the total degree of F. Let

$$m(F) = \max_{1 \leq i \leq N} \deg_{X_i}(F),$$

then $m(F) \leq \deg(F) = M$.

Lemma 2.2.1. *Suppose $F(\boldsymbol{X})$ is not identically 0. Let $S \subseteq \mathbb{C}$ be a set of elements of cardinality at least $m(F)+1$. Then there exists $\boldsymbol{q} \in S^N$ such that $F(\boldsymbol{q}) \neq 0$.*

Proof. Idea for the following argument was suggeted to me by Professor Sinnou David. We argue by induction on N. First suppose $N = 1$. Then our polynomial is of the form

$$F(X) = f_{m(F)}X^{m(F)} + ... + f_1X + f_0 \in \mathbb{C}[X],$$

and F has at most $m(F)$ roots. Since $|S| > m(F)$, there must exist $q \in S$ such that $F(q) \neq 0$. Now suppose the lemma has been proved for all polynomials in k variables for any $1 \leq k < N$. Recall that for each $1 \leq i \leq N$, $\deg_{X_i}(F) \leq m(F)$.

Since F is not identically zero, there must exist $\boldsymbol{x} \in \mathbb{C}^{N-1}$ such that $F(\boldsymbol{x}, X_N)$ is not identically 0. Since $F(\boldsymbol{x}, X_N)$ is a polynomial in one variable, by the base of induction there exists $q_N \in S$ such that $F(\boldsymbol{x}, q_N) \neq 0$. Let

$$P(X_1, ..., X_{N-1}) = F(X_1, ..., X_{N-1}, q_N),$$

then P is not identically 0, and $m(P) \leq m(F)$. By induction hypothesis, there exists $\boldsymbol{q} \in S^{N-1}$ such that $P(\boldsymbol{q}) \neq 0$. Then $(\boldsymbol{q}, q_N) \in S^N$ and $F(\boldsymbol{q}, q_N) = P(\boldsymbol{q}) \neq 0$. □

The assertion of Lemma 2.2.1 can also be deduced as a simple corollary from Lemma 1 on p. 261 of [8]. Notice that the assertion of Lemma 2.2.1 is best possible (i.e. $|S|$ *must* be at least $m(F) + 1$) as seen on the following example. Let $S = \{\alpha_1, ..., \alpha_M\} \subset \mathbb{C}$, and let

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

Then for each $\boldsymbol{q} \in S^N$, we have $F(\boldsymbol{q}) = 0$.

Lemma 2.2.2. *Let F be as in Lemma 2.2.1. There exists $\boldsymbol{q} \in \mathbb{Z}^N$ with $\boldsymbol{q}_i \neq 0$ for all $1 \leq i \leq N$, $F(\boldsymbol{q}) \neq 0$, and*

$$|\boldsymbol{q}| \leq \frac{M+2}{2}. \tag{2.2}$$

Moreover, if F is homogeneous, then the upper bound of (2.2) can be replaced with $\frac{M+1}{2}$.

Proof. Recall that $M \geq m(F)$, and let

$$S = \left\{ -\left[\frac{M}{2}\right] - 1, ..., -1, 1, ..., \left[\frac{M}{2}\right] + 1 \right\},$$

then $|S| = 2\left(\left[\frac{M}{2}\right] + 1\right) \geq M + 1$. Hence, by Lemma 2.2.1, there must exist $\boldsymbol{q} \in S^N$ such that $F(\boldsymbol{q}) \neq 0$.

Now assume that F is homogeneous and $M \geq 1$. Notice that if for any $1 \leq i \leq N$ the "diagonal" coefficient $f_{M\boldsymbol{e}_i} \neq 0$, then $F(\boldsymbol{e}_i) = f_{M\boldsymbol{e}_i} \neq 0$, and we

are done. Hence assume $f_{M\boldsymbol{e}_i} = 0$ for all $1 \leq i \leq N$. Then each monomial of F has degree M and is a product of powers of at least two variables. Therefore $m(F) \leq M - 1$, and so we can take

$$S = \left\{ -\left[\frac{M-1}{2}\right] - 1, ..., -1, 1, ..., \left[\frac{M-1}{2}\right] + 1 \right\},$$

then $|S| = 2\left(\left[\frac{M-1}{2}\right] + 1\right) \geq M \geq m(F) + 1$. Hence, by Lemma 2.2.1, there must exist $\boldsymbol{q} \in S^N$ such that $F(\boldsymbol{q}) \neq 0$. This completes the proof. □

A better basic bound follows from Lemma 2.2.1 if we allow the point in question to have algebraic coordinates.

Lemma 2.2.3. *Let F be as above, and let K be a number field of degree d. There exists a constant $\mathcal{C}_K(N)$ and $\boldsymbol{x} \in O_K^N$ such that $F(\boldsymbol{x}) \neq 0$, and*

$$H(\boldsymbol{x}) \leq \mathcal{C}_K(N) M^{1/d}. \tag{2.3}$$

Proof. Let

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

where $\mathcal{C}(K)$ is a positive field constant to be specified later. By [17] (Theorem 0, p. 102) there exist constants $\mathcal{A}(K)$ and $\mathcal{B}(K)$ such that

$$\mathcal{A}(K)\mathcal{C}(K)^d M \leq |S_M(K)| \leq \mathcal{B}(K)\mathcal{C}(K)^d M. \tag{2.4}$$

Let

$$\mathcal{C}(K) = \left(\frac{2}{\mathcal{A}(K)}\right)^{1/d}, \tag{2.5}$$

so that $|S_M(K)| \geq 2M \geq M+1$. By Lemma 2.2.1 there must exist $\boldsymbol{x} \in S_M(K)^N$ such that $F(\boldsymbol{x}) \neq 0$, and so

$$H(\boldsymbol{x}) \leq \prod_{v|\infty} \left(\mathcal{C}(K)M^{1/d}\right)^{d_v/d} = \mathcal{C}(K)M^{1/d}. \tag{2.6}$$

This completes the proof. □

Another useful measure of "size" of an integral point is *length*. It is defined by

$$L(\boldsymbol{q}) = \sum_{i=1}^{N} |q_i|, \tag{2.7}$$

for each $\boldsymbol{q} \in \mathbb{Z}^N$. Consider $\boldsymbol{q}$ of Lemma 2.2.2. Notice that

$$L(\boldsymbol{q}) \leq N|\boldsymbol{q}| \leq \frac{N(M+2)}{2}. \tag{2.8}$$

This is a trivial bound. Next we want to produce a non-trivial bound on $L(\boldsymbol{q})$. Let $N \geq 2$, $M \geq 1$ be integers, and write

$$\mathcal{M} = \mathcal{M}(N, M) = \left\{\boldsymbol{m} \in \mathbb{Z}_+^N : m_1 + ... + m_N = M\right\}.$$

For the rest of this section, let

$$F(X_1, ..., X_N) = \sum_{\boldsymbol{m} \in \mathcal{M}} f_{\boldsymbol{m}} X_1^{m_1} ... X^{m_N} \in \mathbb{C}[X_1, ..., X_N],$$

be a non-zero homogeneous polynomial in N variables of degree M with coefficients in $\mathbb{C}$.

Lemma 2.2.4. *Let F be as above. There exists a point $\boldsymbol{q} \in \mathbb{Z}^N$ such that $F(\boldsymbol{q}) \neq 0$, and $L(\boldsymbol{q}) \leq \frac{(M+2)^2}{8}$.*

Proof. If $M = 1$, then F is just a linear form in $N \geq 2$ variables. Its nullspace has dimension $N - 1$, and so cannot contain all the standard basis vectors. Therefore there exists $\boldsymbol{x} \in \mathbb{Z}^N$ with $L(\boldsymbol{x}) = 1$ and $F(\boldsymbol{x}) \neq 0$. From now on assume that $M \geq 2$. We can also assume that for each $1 \leq i \leq N$, the coefficient $f_{M\boldsymbol{e}_i} = 0$, where $\boldsymbol{e}_1, ..., \boldsymbol{e}_N$ are the standard basis vectors, since if for some $1 \leq i \leq N$, $f_{M\boldsymbol{e}_i} \neq 0$, then $F(\boldsymbol{e}_i) = f_{M\boldsymbol{e}_i} \neq 0$.

We argue by induction on N. First suppose that $N = 2$, then we can write

$$F(X_1, X_2) = \sum_{i=1}^{M-1} f_i X_1^i X_2^{M-i} = X_1 X_2 \sum_{i=1}^{M-1} f_i X_1^{i-1} X_2^{M-i-1},$$

and so $\frac{1}{X_1} F(X_1, 1) = \sum_{i=1}^{M-1} f_i X_1^{i-1}$ is a polynomial in one variable of degree at most $M - 2$, therefore it can have at most $M - 2$ nonzero roots, and so there must exist an integer β with $|\beta| \leq \frac{M-2}{2} + 1$ such that $F(\beta, 1) \neq 0$. Then $\boldsymbol{q} = (\beta, 1)$ is the required point with

$$L(\boldsymbol{q}) \leq \frac{M+2}{2} \leq \frac{(M+2)^2}{8},$$

since $M \geq 2$.

Next assume $N > 2$. For each $1 \leq i \leq N$, define F_i, *i-th section of* F, to be the homogeneous polynomial in $N - 1$ variables of degree M obtained from F by setting i-th variable equal to 0. First suppose that all sections of F are identically zero, then

$$F(X_1, ..., X_N) = X_1 ... X_N G(X_1, ..., X_N),$$

where G is a homogeneous polynomial of degree $M - N$ (this is only possible if $N \leq M$). By Lemma 2.2.2 and (2.8), there exists $\boldsymbol{q} \in \mathbb{Z}^N$ such that $q_j \neq 0$

for all $1 \leq j \leq N$, $G(\boldsymbol{q}) \neq 0$, and

$$L(\boldsymbol{q}) \leq \frac{N}{2}(M - N + 2).$$

Then $F(\boldsymbol{q}) \neq 0$. Let

$$f(x) = \frac{x}{2}(M - x + 2) = -\frac{1}{2}x^2 + \frac{(M+2)}{2}x,$$

then f achieves its maximum when $x = \frac{M+2}{2}$, and this maximum value is $\frac{(M+2)^2}{8}$. Hence

$$L(\boldsymbol{q}) \leq \frac{(M+2)^2}{8}. \tag{2.9}$$

Next assume that for some $1 \leq i \leq N$, F_i is not identically zero. Then we are done by the inductive hypothesis. This completes the proof. □

Notice that observations of (2.8) and Lemma 2.2.4 can be summarized as follows.

Proposition 2.2.5. *Let F be as above. There exists a point $\boldsymbol{q} \in \mathbb{Z}^N$ such that $F(\boldsymbol{q}) \neq 0$, and*

$$L(\boldsymbol{q}) \leq \left[\frac{(M+2)}{2} \ \min\left\{ N, \frac{M+2}{4} \right\} \right], \tag{2.10}$$

where $[\ ]$, as above, stands for the integer part function.

Now suppose that F is irreducible. In that case the bound of Proposition 2.2.5 can be trivially improved.

Corollary 2.2.6. *Let F as above be irreducible. There exists a point $\boldsymbol{q} \in \mathbb{Z}^N$ such that $F(\boldsymbol{q}) \neq 0$, and*

$$L(\boldsymbol{q}) \leq \left[\frac{(M+2)}{2} \ \min\left\{ N-1, \frac{M+2}{4} \right\} \right]. \tag{2.11}$$

Proof. Since F is irreducible $X_N \nmid F$, therefore

$$F'(X_1, ..., X_{N-1}) = F(X_1, ..., X_{N-1}, 0)$$

is not identically 0. Applying Proposition 2.2.5 to F' finishes the proof. □

2.3 Lattice points in an aligned box

In the next two sections we produce estimates for the number of points of a sublattice of the integer lattice in a closed cube in $\mathbb{R}^N$. These estimates are later used to prove our main theorem.

All results of this section are due to Vaaler ([36]), and are presented here with his permission. Let $A = (a_{mn})$ be an $N \times N$, uppertriangular, nonsingular matrix with real entries. Let $u_m < v_m$ for $m = 1, 2, \ldots, N$ and write

$$R(\boldsymbol{u}, \boldsymbol{v}) = \{\boldsymbol{x} \in \mathbb{R}^N : u_m < x_m \leq v_m\}.$$

We will be interested in estimating the number of points $\boldsymbol{\xi}$ in $\mathbb{Z}^N$ such that $A\boldsymbol{\xi}$ belongs to the aligned box $R(\boldsymbol{u}, \boldsymbol{v})$. To begin with we have the following special result.

Lemma 2.3.1. *Assume that $a_{11} = a_{22} = \cdots = a_{NN} = 1$ and $v_m - u_m$ is a positive integer for each $m = 1, 2, \ldots, N$. Then*

$$|\{\boldsymbol{\xi} \in \mathbb{Z}^N : A\boldsymbol{\xi} \in R(\boldsymbol{u}, \boldsymbol{v})\}| = \prod_{m=1}^{N} (v_m - u_m). \tag{2.12}$$

Proof. We argue by induction on N. If $N = 1$ the result is trivial because one easily checks that the number of integer points ξ_1 such that $u_1 < \xi_1 \leq v_1$ is equal to $[v_1] - [u_1]$. As $v_1 - u_1$ is an integer we find that $[v_1] - [u_1] = v_1 - u_1$.

Now assume that $N \geq 2$. Let $\boldsymbol{\eta}$ be a point in $\mathbb{Z}^{N-1}$ with coordinates indexed by $n = 2, 3, \ldots, N$. Then define

$$I_{N-1} = \{\boldsymbol{\eta} \in \mathbb{Z}^{N-1} : u_m < \sum_{n=m}^{N} a_{mn}\eta_n \leq v_m \text{ for } m = 2, 3, \ldots, N\}.$$

By the inductive hypothesis we have

$$|I_{N-1}| = \prod_{m=2}^{N}(v_m - u_m). \tag{2.13}$$

If $\boldsymbol{\eta}$ is a point in I_{N-1} then the number of integer points ξ_1 such that

$$u_1 < \xi_1 + a_{12}\eta_2 + a_{13}\eta_3 + \cdots + a_{1N}\eta_N \leq v_1, \tag{2.14}$$

is $v_1 - u_1$. Clearly a point

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{pmatrix}$$

satisfies the condition $A\boldsymbol{\xi} \in R(\boldsymbol{u}, \boldsymbol{v})$ if and only if $\boldsymbol{\eta} \in I_{N-1}$ and ξ_1 satisfies (2.14). We have shown that the number of such points is

$$(v_1 - u_1)|I_{N-1}|,$$

and this proves the lemma. □

If we drop the condition that each edge length $v_m - u_m$ is an integer then we get the following estimates.

Lemma 2.3.2. *Assume that $a_{11} = a_{22} = \cdots = a_{NN} = 1$, then*

$$\prod_{m=1}^{N}[v_m - u_m] \leq |\{\boldsymbol{\xi} \in \mathbb{Z}^N : A\boldsymbol{\xi} \in R(\boldsymbol{u}, \boldsymbol{v})\}| \leq \prod_{m=1}^{N}([v_m - u_m] + 1). \tag{2.15}$$

Proof. When proving the lower bound on the left of (2.15) we can assume that $1 \leq v_m - u_m$ for each $m = 1, 2, \ldots, N$. Now select real numbers u'_m and v'_m so that

$$u_m \leq u'_m < v'_m \leq v_m \quad \text{and} \quad v'_m - u'_m = [v_m - u_m], \quad \text{for} \quad m = 1, 2, \ldots, N.$$

As $R(\boldsymbol{u}', \boldsymbol{v}') \subseteq R(\boldsymbol{u}, \boldsymbol{v})$ the inequality follows from Lemma 2.3.1. To obtain the upper bound on the right of (2.15) we argue in essentially the same way. Select real numbers u''_m and v''_m so that

$$u''_m \leq u_m < v_m \leq v''_m \quad \text{and} \quad v''_m - u''_m = [v_m - u_m] + 1, \quad \text{for} \quad m = 1, 2, \ldots, N.$$

Then $R(\boldsymbol{u}, \boldsymbol{v}) \subseteq R(\boldsymbol{u}'', \boldsymbol{v}'')$, and again the inequality follows from Lemma 2.3.1. □

Next we drop the condition that the diagonal entries of the matirx A are all equal to 1.

Corollary 2.3.3. *Assume that the diagonal entries $a_{11}, a_{22}, \ldots, a_{NN}$ are all positive. Then we have*

$$\prod_{m=1}^{N} \Big[\frac{v_m - u_m}{a_{mm}} \Big] \leq |\{\boldsymbol{\xi} \in \mathbb{Z}^N : A\boldsymbol{\xi} \in R(\boldsymbol{u}, \boldsymbol{v})\}| \leq \prod_{m=1}^{N} \Big(\Big[\frac{v_m - u_m}{a_{mm}} \Big] + 1 \Big). \quad (2.16)$$

2.4 Lattice points in cubes

In this section we focus on the case when the box of section 2.3 is actually a cube, and in this case extend the estimate of section 2.3 to lattices of not full rank.

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

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

to be a cube in $\mathbb{R}^N$ centered at the origin with sidelength $2R$. Given a lattice Λ of rank N and determinant Δ, we first want to estimate the number of points of Λ in C_R^N. If the uppertriangular matrix A with a fixed determinant Δ in Corollary 2.3.3 is such that all diagonal entries $a_{mm} \geq c$ for some positive constant c, then the right hand side of (2.16) takes its maximum value when $a_{mm} = c$ for $N-1$ distinct values of m. This leads to the following corollary.

Corollary 2.4.1. *Let Λ be a lattice of full rank in $\mathbb{R}^N$ of determinant Δ such that there exists a positive constant c and a basis matrix $A = (a_{mn})_{1 \leq m,n \leq N}$ of Λ with diagonal entries $a_{mm} \geq c$ for all $1 \leq m \leq N$ (in particular, this is true with $c = 1$ if $\Lambda \subseteq \mathbb{Z}^N$). Then for each point $\boldsymbol{z}$ in $\mathbb{R}^N$ we have*

$$\frac{(2R)^N}{\Delta} \leq |\Lambda \cap (C_R^N + \boldsymbol{z})| \leq \left(\frac{2Rc^{N-1}}{\Delta} + 1\right)\left(\frac{2R}{c} + 1\right)^{N-1}. \tag{2.17}$$

This upper bound is best possible as seen on the example of

$$\Lambda = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{e}_1, ..., \boldsymbol{e}_{N-1}, \Delta \boldsymbol{e}_N\}.$$

Now suppose that Λ of Corollary 2.4.1 is not of full rank.

Theorem 2.4.2. *Suppose that $\Lambda \subseteq \mathbb{Z}^N$ is a lattice of rank $N - l$, where $1 \leq l \leq N-1$. Let Δ be the maximum of absolute values of Grassmann coordinates of Λ. Then for each point $\boldsymbol{z}$ in $\mathbb{R}^N$ we have*

$$|\Lambda \cap (C_R^N + \boldsymbol{z})| \leq \left(\frac{2R}{\Delta} + 1\right)(2R+1)^{N-l-1}. \tag{2.18}$$

Proof. Pick $\boldsymbol{e}_{k_1}, ..., \boldsymbol{e}_{k_l}$ such that the lattice $\text{span}_{\mathbb{Z}}\{\Lambda, \boldsymbol{e}_{k_1}, ..., \boldsymbol{e}_{k_l}\} \subseteq \mathbb{Z}^N$ has rank N. Write $X = (\boldsymbol{x}_1 \dots \boldsymbol{x}_{N-l})$ for the $N \times (N-l)$ basis matrix of Λ. Let $\boldsymbol{k} = (k_1, ..., k_l)$, and let $\Delta_{\boldsymbol{k}}$ be absolute value of the $\boldsymbol{k}$-th Grassmann coordinate of X and so of Λ (i.e. $\Delta_{\boldsymbol{k}}$ is absolute value of the $(N-l) \times (N-l)$ subdeterminant of X obtained by removing the rows numbered $k_1, ..., k_l$; this is an invariant of the lattice). Let $L_{k_1}, ..., L_{k_l}$ be distinct prime numbers so that $L_{k_i} \nmid \Delta_{\boldsymbol{k}}$ for each $1 \leq i \leq l$. Define

$$\Omega_{\boldsymbol{k}} = \text{span}_{\mathbb{Z}}\{\boldsymbol{x}_1, ..., \boldsymbol{x}_{N-1}, L_{k_1}\boldsymbol{e}_{k_1}, ..., L_{k_l}\boldsymbol{e}_{k_l}\}. \tag{2.19}$$

Then $\Lambda \subset \Omega_{\boldsymbol{k}} \subseteq \mathbb{Z}^N$, $\Lambda \neq \Omega_{\boldsymbol{k}}$, and $\Omega_{\boldsymbol{k}}$ is a lattice of rank N. Notice that $\boldsymbol{e}_{k_1}, ..., \boldsymbol{e}_{k_l} \notin \Omega_{\boldsymbol{k}}$.

Choose an integer basis $\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_N$ for $\Omega_{\boldsymbol{k}}$ so that the $N \times N$ basis matrix $A = (\boldsymbol{\alpha}_1 \dots \boldsymbol{\alpha}_N)$ is upper triangular, and

$$0 \leq a_{nj} < a_{nn} \;\; \forall\, 1 \leq n \leq N,\; 1 \leq j \leq N,\; j \neq n. \tag{2.20}$$

Such a basis for $\Omega_{\boldsymbol{k}}$ exists uniquely by Corollary 1 on p. 13 of [8]. Notice that

$$\det(\Omega_{\boldsymbol{k}}) = \Delta_{\boldsymbol{k}} \prod_{i=1}^{l} L_{k_i} = |\det(A)| = \prod_{n=1}^{N} a_{nn}.$$

Fix an s, $1 \leq s \leq l$. Since L_{k_s} is prime, $L_{k_s} | a_{nn}$ for some $1 \leq n \leq N$, and since $L_{k_s} \nmid \Delta_{\boldsymbol{k}}$ this is the only a_{nn} that L_{k_s} divides. Since $L_{k_s}\boldsymbol{e}_{k_s} \in \Omega_{\boldsymbol{k}}$ and A is upper triangular, there must exist integers $\alpha_{s1}, ..., \alpha_{sN}$ such that

$$L_{k_s} = \sum_{i=k_s}^{N} \alpha_{si} a_{ik_s}, \quad 0 = \sum_{i=j}^{N} \alpha_{si} a_{ij}, \;\; \forall\; j \neq k_s.$$

Case 1. Suppose $k_s = N$. Then $L_N = \alpha_{sN} a_{NN}$, which implies that either $\alpha_{sN} = L_N,\; a_{NN} = 1$, or $\alpha_{sN} = 1,\; a_{NN} = L_N$. However, if $a_{NN} = 1$,

then by (2.20) $a_{Ni} = 0$ for all $1 \leq i \leq N-1$, and so $\boldsymbol{\alpha}_N = \boldsymbol{e}_N \in \Omega_{\boldsymbol{k}}$, which is a contradiction. Therefore $a_{NN} = L_N$.

Case 2. Suppose $k_s < N$. Then $\alpha_{sN} a_{NN} = 0$, and so $\alpha_{sN} = 0$. Then $\alpha_{s(N-1)} a_{(N-1)(N-1)} + \alpha_{sN} a_{N(N-1)} = 0$, which means that $\alpha_{s(N-1)} = 0$. Continuing in the same manner, we see that $\alpha_{si} = 0$ for each $i > k_s$. Hence $L_{k_s} = \sum_{i=k_s}^{N} \alpha_{si} a_{ik_s} = \alpha_{sk} a_{k_s k_s}$. By the same argument as in case 1, this means that $a_{k_s k_s} = L_{k_s}$.

Therefore we proved that $a_{k_s k_s} = L_{k_s}$ for all $1 \leq s \leq l$, and each L_{k_s} does not divide any other a_{nn}, hence

$$\prod_{n=1, n \neq k_1, \ldots, k_l}^{N} a_{nn} = \Delta_{\boldsymbol{k}}. \tag{2.21}$$

Applying Corollary 2.4.1, we see that for any $\boldsymbol{z} \in \mathbb{R}^N$,

$$\begin{aligned} |\Lambda \cap (C_R^N + \boldsymbol{z})| &\leq |\Omega_{\boldsymbol{k}} \cap (C_R^N + \boldsymbol{z})| \\ &\leq \prod_{s=1}^{l} \left(\frac{2R}{L_{k_s}} + 1 \right) \prod_{n=1, n \neq k_1, \ldots, k_l}^{N} \left(\frac{2R}{a_{nn}} + 1 \right). \end{aligned} \tag{2.22}$$

Since our choice of $L_{k_1}, \ldots, L_{k_l}$ was arbitrary, we will now let $L_{k_s} \to \infty$ for all $1 \leq s \leq l$, and so

$$\begin{aligned} |\Lambda \cap (C_R^N + \boldsymbol{z})| &\leq \prod_{n=1, n \neq k_1, \ldots, k_l}^{N} \left(\frac{2R}{a_{nn}} + 1 \right) \prod_{s=1}^{l} \lim_{L_{k_s} \to \infty} \left(\frac{2R}{L_{k_s}} + 1 \right) \\ &= \prod_{n=1, n \neq k_1, \ldots, k_l}^{N} \left(\frac{2R}{a_{nn}} + 1 \right). \end{aligned} \tag{2.23}$$

The right hand side of (2.23) takes its maximum value when $a_{nn} = 1$ for $N - l - 1$ distinct values of n. Therefore, applying (2.21) we obtain

$$|\Lambda \cap (C_R^N + \boldsymbol{z})| \leq \left(\frac{2R}{\Delta_{\boldsymbol{k}}} + 1 \right) (2R+1)^{N-l-1}. \tag{2.24}$$

We can now specify how we select $\boldsymbol{k}$. We want to do it so that the upper bound in (2.24) is minimized. For this, let $\boldsymbol{k}$ be such that $\Delta_{\boldsymbol{k}}$ is the maximal among all the Grassmann coordinates of Λ, and call this maximum value Δ (notice that if $\Delta_k \neq 0$, then the lattice $\text{span}_{\mathbb{Z}}\{\Lambda, \boldsymbol{e}_{k_1}, ..., \boldsymbol{e}_{k_l}\} \subseteq \mathbb{Z}^N$ has rank N). This completes the proof. □

Proposition 2.4.3. *Suppose that Λ is a lattice (i.e. a free $\mathbb{Z}$-module) of rank $N-l$ in $\mathbb{R}^N$, where $1 \leq l \leq N-1$. Let Δ be the maximum of absolute values of Grassmann coordinates of Λ. Then*

$$\frac{(2R)^{N-l}}{(N-l)^{N-l}\Delta} \leq |\Lambda \cap C_R^N|. \tag{2.25}$$

Proof. Pick a basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_{N-l}$ for Λ and write $X = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_{N-l})$ for the corresponding $N \times (N-l)$ basis matrix. Let A be an $l \times N$ matrix such that $A\boldsymbol{x} = \boldsymbol{0}$ for each $\boldsymbol{x} \in \Lambda$. Write $\Delta_{i_1 ... i_{N-l}}$ for the Grassmann coordinate of X, which is the determinant of the submatrix of X whose rows are indexed by $i_1, ..., i_{N-l} \in \{1, ..., N\}$. Write $\delta^{j_1 ... j_l}$ for the Grassmann coordinate of A, which is the determinant of the submatrix of A whose columns are indexed by $j_1, ..., j_l \in \{1, ..., N\}$. By (1.24) we have

$$\Delta_{i_1 ... i_{N-l}} = (-1)^{i_1 + ... + i_{N-l}} \ \gamma \ \delta^{i_{N-l+1} ... i_N}, \tag{2.26}$$

for an appropriate $\gamma \in \mathbb{R}$, where $\{i_1, ..., i_{N-l}, i_{N-l+1}, ..., i_N\} = \{1, ..., N\}$. We can assume without loss of generality that

$$\Delta = |\Delta_{1...(N-l)}| = |\gamma||\delta^{(N-l+1)...N}|. \tag{2.27}$$

Let X' be the $(N-l) \times (N-l)$ submatrix of X whose rows are indexed by $1, ..., (N-l)$, and let Λ' be a lattice of full rank in $\mathbb{R}^{N-l}$ generated by the

column vectors of X'. Then $\det(\Lambda') = \det(X') = \Delta$, and so, by the lower bound of Corollary 2.4.1

$$|\Lambda' \cap C_R^{N-l}| \geq \frac{(2R)^{N-l}}{\Delta}. \tag{2.28}$$

Suppose that $\boldsymbol{y} = (y_1, ..., y_{N-l}) \in \Lambda' \cap C_R^{N-l}$, then $|\boldsymbol{y}| = \max_{1 \leq i \leq N-l} |y_i| \leq R$, where $R \geq 1$. There exists $\boldsymbol{z} = (z_1, ..., z_l) \in \mathbb{R}^l$ such that $(\boldsymbol{y}, \boldsymbol{z}) = (y_1, ..., y_{N-l}, z_1, ..., z_l) \in \Lambda$. We want to establish an upper bound on $|\boldsymbol{z}|$. By equation (4) on page 293 of [15], every point $\boldsymbol{x} \in \Lambda$ must satisfy the following system of linear equations:

$$\sum_{j=1}^{N} \delta^{i_1 ... i_{l-1} j} x_j = 0, \tag{2.29}$$

where $i_1, ..., i_{l-1}$ assume all possible values; only $N - l$ of these equations are linearly independent. It is easy to see that the sum on the left side of each equation like (2.29) has only $N - l + 1$ terms: there are only $N - l + 1$ possibilities for j since the $l - 1$ values $i_1, ..., i_{l-1}$ have been preassigned. For each $N - l + 1 \leq i \leq N$ let $I_i = \{N - l + 1, ..., N\} \setminus \{i\}$, then the following l equations form a subset of equations in (2.29):

$$\sum_{j=1}^{N-l} \delta^{I_i j} x_j + \delta^{(N-l+1)...N} x_i = 0. \tag{2.30}$$

Substitute the coordinates of the point $(\boldsymbol{y}, \boldsymbol{z})$ into (2.30), then we see that for each $1 \leq i \leq l$

$$z_i = -\sum_{j=1}^{N-l} \frac{\delta^{I_{(N-l+i)} j}}{\delta^{(N-l+1)...N}} y_j, \tag{2.31}$$

and so

$$|z_i| \leq \sum_{j=1}^{N-l} \left| \frac{\delta^{I_{(N-l+i)} j}}{\delta^{(N-l+1)...N}} \right| |y_j| \leq (N - l)R, \tag{2.32}$$

since each $\left|\frac{\delta^{I(N-l+i)^j}}{\delta^{(N-l+1)\ldots N}}\right| \leq 1$ by construction, since $\delta^{(N-l+1)\ldots N}$ is the biggest in absolute value among all the Grassmann coordinates of A, and $|\boldsymbol{y}| \leq R$. Therefore for each $\boldsymbol{y} \in \Lambda' \cap C_R^{N-l}$ there exists $\boldsymbol{z} \in \mathbb{R}^l$ such that $(\boldsymbol{y}, \boldsymbol{z}) \in \Lambda \cap C^N_{(N-l)R}$, hence

$$|\Lambda \cap C_R^N| \geq |\Lambda' \cap C^{N-l}_{\frac{R}{N-l}}| \geq \frac{(2R)^{N-l}}{(N-l)^{N-l}\Delta}. \tag{2.33}$$

This completes the proof. □

Notice that Theorem 2.4.2 deals only with a sublattice of the integer lattice while Proposition 2.4.3 deals with any lattice in $\mathbb{R}^N$.

2.5 Classical case

In this section we consider the following problem. Given a subspace W of $\mathbb{R}^N$ and a collection of proper subspaces $V_1, ..., V_M$ of W we want to prove the existence of a non-zero integral point of small height in $W \setminus \bigcup_{i=1}^M V_i$.

We start with a discussion of a partial case of this problem, namely let $W = \mathbb{R}^N$ and let $V_1, ..., V_M$ be $(N-1)$-dimensional subspaces of $\mathbb{R}^N$, that is $V_1, ..., V_M$ are nullspaces of linear forms with integer coefficients. Then, given a collection of non-zero linear forms with integer coefficients, we want to prove the existence of an integer lattice point at which none of these linear forms would vanish.

A basic bound that depends only on the number of linear forms follows from results of section 2.2. If $L_1, .., L_M$ are non-zero linear forms in N variables with coefficients in $\mathbb{C}$, take $F(\boldsymbol{X}) = L_1(\boldsymbol{X}) \cdots L_M(\boldsymbol{X})$. By Lemma 2.2.2, there

exists an integer lattice point $\boldsymbol{x}$ such that $F(\boldsymbol{x}) \neq 0$, and so $L_i(\boldsymbol{x}) \neq 0$ for all $1 \leq i \leq N$, with

$$|\boldsymbol{x}| \leq \frac{M+1}{2}. \tag{2.34}$$

Moreover, if the linear forms have integer coefficients, then a bound of the form $|\boldsymbol{x}| \ll_N M^{(N-1)/N}$ follows from [2].

We want to produce a result that depends on the actual linear forms, not just on their number. We will relate this problem to the lattice point counting problem of sections 2.3 and 2.4.

Theorem 2.5.1. *Let $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X}) \in \mathbb{Z}[X_1, ..., X_N]$ be non-zero linear forms, given by*

$$L_i(\boldsymbol{X}) = \boldsymbol{q}_i \cdot \boldsymbol{X}, \quad \boldsymbol{q}_i \in \mathbb{Z}^N \quad \forall\ 1 \leq i \leq M,$$

so that $\boldsymbol{q}_i \neq \boldsymbol{0}$ for all $1 \leq i \leq M$, and each $\boldsymbol{q}_i$ has relatively prime coordinates. Then there exists $\boldsymbol{x} \in \mathbb{Z}^N$ such that

$$L_i(\boldsymbol{x}) \neq 0$$

for every $i = 1, ..., M$ and

$$|\boldsymbol{x}| \leq \sum_{i=1}^{M} \frac{1}{|\boldsymbol{q}_i|} + \sqrt{M}. \tag{2.35}$$

Proof. For any $1 \leq i \leq M$, let

$$V_i = \{\boldsymbol{x} \in \mathbb{R}^N : L_i(\boldsymbol{x}) = 0\},$$

and define $\Lambda_i = V_i \cap \mathbb{Z}^N$, then Λ_i is a lattice of rank $N-1$ in $\mathbb{R}^N$. Let Δ_i be the maximum of Grassmann coordinates of Λ_i. By Theorem 2.4.2,

$$|\Lambda_i \cap C_R^N| \leq \left(\frac{2R}{\Delta_i} + 1\right)(2R+1)^{N-2}. \tag{2.36}$$

By the duality principle of Brill-Gordan (Lemma 1.2.2)

$$\Delta_i = |\boldsymbol{q}_i|,$$

since $\boldsymbol{q}_i$ has relatively prime coordinates. Putting this together with (2.36), we obtain

$$|\Lambda_i \cap C_R^N| \leq \left(\frac{2R}{|\boldsymbol{q}_i|} + 1\right)(2R+1)^{N-2}.$$

Then

$$\begin{aligned}\left|\left(\bigcup_{i=1}^{M} \Lambda_i\right) \cap C_R^N\right| &\leq \sum_{i=1}^{M} |\Lambda_i \cap C_R^N| \\ &\leq (2R+1)^{N-2}\left(2R\sum_{i=1}^{M}\frac{1}{|\boldsymbol{q}_i|} + M\right) \qquad (2.37)\end{aligned}$$

Notice, on the other hand, that if R is a positive integer, then

$$|C_R^N \cap \mathbb{Z}^N| = (2R+1)^N.$$

Putting this together with (2.37), we see that if

$$(2R+1)^N - (2R+1)^{N-2}\left(2R\sum_{i=1}^{M}\frac{1}{|\boldsymbol{q}_i|} + M\right) > 0, \qquad (2.38)$$

then there must exist a point $\boldsymbol{x} \in C_R^N \cap \mathbb{Z}^N$, which is not in $\bigcup_{i=1}^{M} V_i$. Rewriting (2.38), we obtain

$$4R^2 - 2R\left(\sum_{i=1}^{M}\frac{1}{|\boldsymbol{q}_i|} - 2\right) - (M-1) > 0,$$

which implies that we can, for instance, take

$$R = \sum_{i=1}^{M}\frac{1}{|\boldsymbol{q}_i|} + \sqrt{M}.$$

This completes the proof. □

Theorem 2.5.1 produces a better bound than (2.34) for linear forms with suffciently large heights. Also, suppose that N is fixed and M grows. Then our collection must contain linear forms with relatively large heights, since there are only finitely many vectors of height $\leq C$ in $\mathbb{Z}^N$ for each C. This is definitely the more interesting situation, since if $M < N$ or if the two are comparable, there must exist integer lattice points of height $\ll N$ at which the linear forms do not vanish.

Next we consider the general problem of finding an integral point of small height in a real vector space outside of a collection of subspaces.

Theorem 2.5.2. *Let $W \subseteq \mathbb{R}^N$ be a subspace of dimension w, $1 \leq w \leq N$, so that $\dim_{\mathbb{Q}}(W \cap \mathbb{Q}^N) = w$. Let $V_1, ..., V_M$ be proper non-zero subspaces of W. Then there exists $\boldsymbol{x} \in \left(W \setminus \bigcup_{i=1}^{M} V_i\right) \cap \mathbb{Z}^N$ such that*

$$H(\boldsymbol{x}) \leq \left(\frac{3}{2}\right)^{w-1} w^w \left\{\sum_{i=1}^{M} \frac{1}{H(V_i)} + \sqrt{M}\right\} H(W). \tag{2.39}$$

Proof. For each $1 \leq i \leq M$ let l_i be the dimension of V_i, then $0 < l_i < w$. Define $\Omega = W \cap \mathbb{Z}^N$ and $\Lambda_i = \Omega \cap V_i$ for each $1 \leq i \leq M$. Then Ω is a sublattice of $\mathbb{Z}^N$ of rank w, and each Λ_i is a sublattice of Ω of rank $l_i \leq w-1$. Let Δ be the maximum of absolute values of the Grassmann coordinates of Ω, and hence of W. In the same manner define $\Delta_1, ..., \Delta_M$ for all subspaces $V_1, ..., V_M$ respectively. Define a function of a positive real variable R

$$f_W(R) = |\Omega \cap C_R^N| - \left|\bigcup_{i=1}^{M} \Lambda_i \cap C_R^N\right|, \tag{2.40}$$

then

$$\begin{aligned} f_W(R) &\geq |\Omega \cap C_R^N| - \sum_{i=1}^{M} |\Lambda_i \cap C_R^N| \\ &\geq \frac{(2R)^w}{w^w \Delta} - \sum_{i=1}^{M} \left(\frac{2R}{\Delta_i} + 1 \right) (2R+1)^{l_i - 1}, \end{aligned} \tag{2.41}$$

where the last inequality follows by Theorem 2.4.2 and Proposition 2.4.3. Notice that $f_W(R) > 0$ if and only if there exists a point $\boldsymbol{x} \in \left(W \setminus \bigcup_{i=1}^{M} V_i \right) \cap \mathbb{Z}^N$ with $|\boldsymbol{x}| \leq R$. Hence ideally we want to find the smallest R so that $f_W(R) > 0$. Recall that $l_i \leq w - 1$ for each i, and we can assume without loss of generality that $R \geq 1$. Then using (2.41), we have

$$f_W(R) \geq R^{w-2} \left\{ \left(\frac{2^w}{w^w \Delta} \right) R^2 - 3^{w-1} \left(\sum_{i=1}^{M} \frac{1}{\Delta_i} \right) R - 3^{w-2} M \right\}.$$

Therefore we want to solve for R a quadratic inequality

$$\left(\frac{2^w}{w^w \Delta} \right) R^2 - 3^{w-1} \left(\sum_{i=1}^{M} \frac{1}{\Delta_i} \right) R - 3^{w-2} M > 0. \tag{2.42}$$

It is not difficult to deduce from (2.42) that we can take

$$R = w^w \Delta \left(\frac{3}{2} \right)^{w-1} \left\{ \sum_{i=1}^{M} \frac{1}{\Delta_i} + \sqrt{M} \right\}. \tag{2.43}$$

Now notice that since $Gr(W), Gr(V_1), ..., Gr(V_M)$ are vectors with integer coordinates, we have

$$H(W) = \Delta, \ H(V_i) = \Delta_i, \ \forall \ 1 \leq i \leq M. \tag{2.44}$$

The conclusion of the theorem follows. □

Suppose that we have $M = 0$ in Theorem 2.5.2, that is we are looking for a point of small height in W without additional conditions. Then,

combining (2.40), (2.41), and (2.44) we can introduce a counting function

$$f_W(R) = |\Omega \cap C_R^N| \geq \frac{(2R)^w}{w^w \Delta} \geq \frac{(2R)^w}{w^w H(W)}.$$

We now want to find a value of $R \geq 1$ such that $f_W(R) \geq 2$; this would imply that there exists a non-zero point $\boldsymbol{x} \in \Omega \cap C_R^N$, so its height is $\leq R$. Hence we want $\frac{(2R)^w}{w^w H(W)} \geq 2$, which means that we can take

$$R = \left(\frac{w}{2^{\frac{w-1}{w}}} \right) H(W)^{1/w}.$$

This proves the following version of Siegel's Lemma.

Corollary 2.5.3. *Let $W \subseteq \mathbb{R}^N$ be a subspace of dimension w, $1 \leq w \leq N$, so that $\dim_{\mathbb{Q}}(W \cap \mathbb{Q}^N) = w$. There exists a non-zero point $\boldsymbol{x} \in W \cap \mathbb{Z}^N$ such that*

$$H(\boldsymbol{x}) \leq \left(\frac{w}{2^{\frac{w-1}{w}}} \right) H(W)^{1/w}.$$

On the other hand, if $w = N$ in Theorem 2.5.2, then we obtain a point $\boldsymbol{x} \in \mathbb{Z}^N$ such that $\boldsymbol{x} \notin \bigcup_{i=1}^M V_i$ and

$$H(\boldsymbol{x}) \leq \left(\frac{3}{2} \right)^{N-1} N^N \left\{ \sum_{i=1}^M \frac{1}{H(V_i)} + \sqrt{M} \right\},$$

i.e. a version of Theorem 2.5.1. Hence Theorem 2.5.2 should be thought of as a combination of Siegel's Lemma and an inverse problem of finding a point of small height outside of a collection of subspaces.

Another immediate corollary of Theorem 2.5.2 in case when $M = 1$ is a sharper bound for a special case of Faltings' version of Siegel's Lemma (see [12], [16], and [11]).

Corollary 2.5.4. *Let V and W be real vector spaces of respective dimensions d_1 and d_2. Let $\Omega_1 = V \cap \mathbb{Z}^{d_1}$ and $\Omega_2 = W \cap \mathbb{Z}^{d_2}$. Let $\rho : V \longrightarrow W$ be a linear map such that $\rho(\Omega_1) \subseteq \Omega_2$. Let $U = \ker(\rho)$, and let $\Omega = U \cap \Omega_1$. Let J be the rank of Ω. Then for any subspace U_0 of U which does not contain Ω there exists a point $\boldsymbol{x} \in \Omega \setminus U_0$ such that*

$$H(\boldsymbol{x}) \leq 2J^J \left(\frac{3}{2}\right)^{J-1} H(\Omega). \tag{2.45}$$

Faltings' lemma is more general: it works with Ω_1, Ω_2, and Ω being any lattices in V, W, and U respectively, as well as any choice of norms on V and W (our height is the sup-norm). However, the upper bound on $H(\boldsymbol{x})$ which follows from Faltings' lemma is

$$H(\boldsymbol{x}) \leq d_1^{\frac{1}{d_3 - d_0}} H(\Omega)^{\frac{3d_1}{d_3 - d_0}}, \tag{2.46}$$

where $d_3 = \dim_{\mathbb{R}}(U)$ and $d_0 = \dim_{\mathbb{R}}(U_0)$, so that $d_1 > d_3 > d_0$ and $d_3 \geq J$. Faltings' method of proof is different from ours: it relies on Minkowski's theorem about successive minima.

2.6 Number field case

In this section we generalize the discussion of section 2.5 to the number field case. Let K be a number field of degree d. Let $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X}) \in K[X_1, ..., X_N]$ be non-zero linear forms. We want to prove the existence of a point of small height in K^N at which none of these linear forms would vanish.

A basic bound that depends only on the number of linear forms follows from results of section 2.2. Take $F(\boldsymbol{X}) = L_1(\boldsymbol{X}) \cdots L_M(\boldsymbol{X})$. By Lemma 2.2.3,

there exists a point $\boldsymbol{x} \in O_K^N$ such that $F(\boldsymbol{x}) \neq 0$, and so $L_i(\boldsymbol{x}) \neq 0$ for all $1 \leq i \leq N$, with

$$H(\boldsymbol{x}) \leq \mathcal{C}_K(N) M^{1/d}, \tag{2.47}$$

for an appropriate field constant $\mathcal{C}_K(N)$. Next we want to produce a bound that would depend on the heights of linear forms, not just on their number. This is a precise version of Theorem 1.3.1. In fact, we will again consider a more general problem: given a subspace W of K^N and a collection $V_1, ..., V_M$ of subspaces of W, we want to prove the existence of a point of small height in $W \setminus \bigcup_{i=1}^M V_i$. We again relate this problem to the lattice point counting problem of sections 2.3 and 2.4.

Theorem 2.6.1. *Let K be a number field of degree d with discriminant $\mathcal{D}_K$ and r_2 complex places. Let $N \geq 2$ be an integer, and let W be a subspace of K^N of dimension w, $1 \leq w \leq N$. Let $V_1, ..., V_M \subseteq W$ be proper subspaces of W of corresponding dimensions $l_1, ..., l_M \geq 1$. There exists a point $\boldsymbol{x} \in \left(W \setminus \bigcup_{i=1}^M V_i\right) \cap O_K^N$ such that*

$$H(\boldsymbol{x}) \leq \left(\mathcal{C}^1_{K,N}(W) H(W) + 1\right) \left\{ \left(\sum_{i=1}^M \frac{\mathcal{C}^2_{K,N}(V_i)}{H(V_i)^d} \right)^{\frac{1}{d}} + M^{\frac{1}{d+1}} \right\}, \tag{2.48}$$

where

$$\mathcal{C}^1_{K,N}(W) = 2^{w - \frac{(wr_2 - 1)}{d}} (wd)^w |\mathcal{D}_K|^{w/2d} \binom{N}{w}^{1/2}, \qquad \mathcal{C}^2_{K,N}(V_i) = \frac{2^{l_i r_2} \binom{Nd}{l_i d}^{1/2}}{|\mathcal{D}_K|^{l_i/2}}. \tag{2.49}$$

Proof. We write $\mathcal{D}_K$ for the discriminant of the number field K everywhere below. Let

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

be the embeddings of K into $\mathbb{C}$ with $\sigma_1, ..., \sigma_{r_1}$ being real embeddings and $\tau_i, \tau_{r_2+i} = \bar{\tau}_i$ for each $1 \leq i \leq r_2$ being the pairs of complex conjugate embeddings. For each $\alpha \in K$ and each complex embedding τ_i, write $\tau_{i1}(\alpha) = \Re(\tau_i(\alpha))$ and $\tau_{i2}(\alpha) = \Im(\tau_i(\alpha))$, where $\Re$ and $\Im$ stand respectively for real and imaginary parts of a complex number. We will view $\tau_i(\alpha)$ as a pair $(\tau_{i1}(\alpha), \tau_{i2}(\alpha)) \in \mathbb{R}^2$. Then $d = r_1 + 2r_2$, and for each $N \geq 1$ we define an embedding

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

where

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

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

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

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

then, by Theorem 2 of [31], $\Lambda(V)$ is a lattice in $\mathbb{R}^{Nd}$ of rank ld, and

$$\left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l H(V)^d \leq \det(\Lambda(V)) \leq \binom{N}{l}^{d/2} \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l H(V)^d. \qquad (2.50)$$

Notice that we obtain inequalities in (2.50) instead of equality as in Theorem 2 of [31] because we use a different height; the exponent d on $H(V)$ appears because our height is absolute unlike the one in Theorem 2 of [31]. Finally,

the constant 2^{-r_2} appears because we use a slightly different embedding into $\mathbb{R}^{Nd}$ than that in Theorem 2 of [31] (see Lemma 2 on p. 115 of [17]).

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

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

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

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

By the Cauchy-Binet formula,

$$\begin{aligned}\max_{|I|=ld} |\det(X_I)| &\leq \det(\Lambda(V)) \\ &= \left(\sum_{|I|=ld} |\det(X_I)|^2\right)^{1/2} \\ &\leq \binom{Nd}{ld}^{1/2} \max_{|I|=ld} |\det(X_I)|. \qquad (2.51)\end{aligned}$$

Let $J \subset \{1, ..., Nd\}$ with $|J| = ld$ be such that $|\det(X_J)| = \max_{|I|=ld} |\det(X_I)|$, and let $\Omega(V)$ be the lattice of full rank in $\mathbb{R}^{ld}$ spanned over $\mathbb{Z}$ by the column vectors of X_J. By combining (2.50) and (2.51), we see that

$$\begin{aligned}\binom{Nd}{ld}^{-1/2} \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l H(V)^d &\leq \binom{Nd}{ld}^{-1/2} \det(\Lambda(V)) \\ &\leq \det(\Omega(V)) = |\det(X_J)| \\ &\leq \det(\Lambda(V)) \\ &\leq \binom{N}{l}^{d/2} \left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^l H(V)^d. \qquad (2.52)\end{aligned}$$

For convenience, we denote $\det(\Omega(V))$ by $\Delta(V)$. By Corollary 1 on p. 13 of [8], we can select a basis for $\Omega(V)$ so that the basis matrix is upper triangular, all of its nonzero entries are positive, and maximum entry of each row occurs on the diagonal. Each of these maximum values is at least $\frac{1}{\sqrt{2}}$, since each row still consists of blocks of all conjugates of l algebraic integers from O_K. Therefore the lattice $\Omega(V)$ satisfies the conditions of Corollary 2.4.1 with $c = \frac{1}{\sqrt{2}}$. Hence

$$|\Omega(V) \cap C_R^{ld}| \leq \left(\frac{R}{2^{\frac{ld-3}{2}}\Delta(V)} + 1\right) (2^{\frac{3}{2}}R + 1)^{ld-1}. \qquad (2.53)$$

On the other hand, by Proposition 2.4.3,

$$|\Lambda(V) \cap C_R^{Nd}| \geq \frac{(2R)^{ld}}{(ld)^{ld}\Delta(V)}, \qquad (2.54)$$

since $\Delta(V)$ is the maximum of absolute values of Grassmann coordinates of $\Lambda(V)$.

For future use, we also need to define a projection $\varphi_V : \Lambda(V) \longrightarrow \Omega(V)$, given by our construction. Namely, if $X\boldsymbol{y} \in \Lambda(V)$ for some $\boldsymbol{y} \in \mathbb{Z}^{Nd}$, then $\varphi_V(X\boldsymbol{y}) = X_J\boldsymbol{y}_J$, where $\boldsymbol{y}_J \in \mathbb{Z}^{ld}$ is obtained from $\boldsymbol{y}$ by removing all the coordinates which are not indexed by J. It is quite easy to see that φ_V is a $\mathbb{Z}$-module isomorphism.

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

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

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

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

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

Notice that for each $\boldsymbol{x} \in K^N$,

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

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

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

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

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

where the notation is as above. Applying (2.53) and (2.54) we obtain

$$\begin{aligned}
f_W(R) &\geq \frac{(2R)^{wd}}{(wd)^{wd}\Delta(W)} - \sum_{i=1}^{M} \left(\frac{R}{2^{\frac{l_i d-3}{2}}\Delta(V_i)} + 1 \right) (2^{\frac{3}{2}}R + 1)^{l_i d-1} \\
&\geq \frac{(2R)^{wd}}{(wd)^{wd}\Delta(W)} - (2^{\frac{3}{2}}R + 1)^{(w-1)d-1} \sum_{i=1}^{M} \left(\frac{R}{2^{\frac{d-3}{2}}\Delta(V_i)} + 1 \right) \\
&\geq R^{(w-1)d-1} \times \\
&\times \left\{ \left(\frac{2^{wd}}{(wd)^{wd}\Delta(W)} \right) R^{d+1} - 4^{\left(w-\frac{5}{4}\right)d-\frac{1}{4}} \left(\sum_{i=1}^{M} \frac{1}{\Delta(V_i)} \right) R - 4^{(w-1)d-1} M \right\} \\
&\geq \left(\frac{2^{wd}}{(wd)^{wd}\Delta(W)} \right) R^{(w-1)d-1} \times \\
&\times \left\{ R^{d+1} - (2wd)^{wd}\Delta(W) \left(\sum_{i=1}^{M} \frac{1}{\Delta(V_i)} \right) R - (2wd)^{wd}\Delta(W) M \right\}. \quad (2.55)
\end{aligned}$$

Let $x = \sum_{i=1}^{M} \frac{1}{\Delta(V_i)}$, and let $\mathcal{A}_W = (2wd)^{wd}\Delta(W)$, and define

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

so that $f_W(R) \geq \left(\frac{2^{wd}}{(wd)^{wd}\Delta(W)} \right) R^{(w-1)d-1} g_W(R)$. Hence we want to determine a value of R for which $g_W(R) > 0$. Let $\mathcal{B}_W$ be a positive number to be specified

later. Then

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

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

$$\begin{aligned} \mathcal{B}_W &= (2\mathcal{A}_W)^{1/d}+1 = 2^{w+\frac{1}{d}}(wd)^w\Delta(W)^{1/d}+1 \\ &\leq 2^{w-\frac{(wr_2-1)}{d}}(wd)^w|\mathcal{D}_K|^{w/2d}\binom{N}{w}^{1/2}H(W)+1, \end{aligned} \tag{2.56}$$

where the last inequality follows by (2.52). Therefore, by (2.52), if we select

$$\begin{aligned} R &= \left(2^{w-\frac{(wr_2-1)}{d}}(wd)^w|\mathcal{D}_K|^{w/2d}\binom{N}{w}^{1/2}H(W)+1\right)\times \\ &\times \left\{\left(\sum_{i=1}^{M}\frac{1}{\Delta(V_i)}\right)^{\frac{1}{d}}+M^{\frac{1}{d+1}}\right\} \\ &\leq \left(2^{w-\frac{(wr_2-1)}{d}}(wd)^w|\mathcal{D}_K|^{w/2d}\binom{N}{w}^{1/2}H(W)+1\right)\times \\ &\times \left\{\left(\sum_{i=1}^{M}\frac{2^{l_i r_2}\binom{Nd}{l_i d}^{1/2}}{|\mathcal{D}_K|^{l_i/2}H(V_i)^d}\right)^{\frac{1}{d}}+M^{\frac{1}{d+1}}\right\}, \end{aligned} \tag{2.57}$$

then $f_W(R) > 0$. This completes the proof. □

Notice that in case $K = \mathbb{Q}$ the bound of Theorem 2.6.1 becomes

$$\left(2^{w+1}w^w\binom{N}{w}^{1/2}H(W)+1\right)\left\{\left(\sum_{i=1}^{M}\frac{\binom{N}{l_i}^{1/2}}{H(V_i)}\right)+\sqrt{M}\right\}, \tag{2.58}$$

which is essentialy (up to a constant) the bound of Theorem 2.5.2. Hence Theorem 2.6.1 is truly a generalization of Theorem 2.5.2. Another interesting observation is that in the case when $W = K^N$ and $V_1, ..., V_M$ is a collection of nullspaces of linear forms (i.e. $w = N$ and $l_i = N - 1$ for each $1 \leq i \leq M$), Theorem 2.6.1 produces a better bound than (2.47) when linear forms have sufficiently large heights. In fact, an effective version of (2.47) can be derived from the bound of Theorem 2.6.1:

$$\begin{aligned} H(\boldsymbol{x}) &\leq \left(2^{N-\frac{(Nr_2-1)}{d}}(Nd)^N|\mathcal{D}_K|^{N/2d}+1\right)\times \\ &\times \left\{\left(\frac{2^{(N-1)r_2}\binom{Nd}{Nd-d}^{1/2}}{|\mathcal{D}_K|^{(N-1)/2}}\right)\left(\sum_{i=1}^{M}\frac{1}{H(V_i)^d}\right)^{\frac{1}{d}}+M^{\frac{1}{d+1}}\right\} \\ &\leq 2^{N(d+1)}(Nd)^N|\mathcal{D}_K|^{1/2d}\binom{Nd}{Nd-d}^{1/2}M^{1/d}. \end{aligned} \tag{2.59}$$

Also notice that we can produce the following simple version of Siegel's Lemma over a number field by combining (2.55) and (2.52): the upper bound exhibits the best possible exponent on $H(W)$; the constant is of course not best possible. This is an alternative way to produce a Siegel's Lemma over a number field using an elementary construction.

Corollary 2.6.2. *Let $W \subseteq K^N$ be a subspace of dimension w, $1 \leq w \leq N$. There exists a non-zero point $\boldsymbol{x} \in W \cap O_K^N$ such that*

$$H(\boldsymbol{x}) \leq \left(\frac{wd}{2^{\frac{wd-1}{wd}}}\right)|\mathcal{D}_K|^{1/2d}\binom{N}{w}^{1/2w}H(W)^{1/w}. \tag{2.60}$$

Proof. Let $M = 0$, then the inequality (2.55) for the counting function $f_W(R)$

becomes

$$f_W(R) \geq \frac{(2R)^{wd}}{(wd)^{wd}\Delta(W)}.$$

Notice that if R is such that $f_W(R) \geq 2$, then there must exist a non-zero point $\boldsymbol{x} \in W \cap O_K^N$ with $H(\boldsymbol{x}) \leq R$. Thus we will look for $R \geq 1$ so that $\frac{(2R)^{wd}}{(wd)^{wd}\Delta(W)} \geq 2$, meaning that we can take

$$R \geq \left(\frac{wd}{2}\right)(2\Delta(W))^{1/wd}.$$

Using (2.52) we see that

$$\begin{aligned}\Delta(W)^{1/wd} &\leq \left(\binom{N}{w}^{d/2}\left(\frac{|\mathcal{D}_K|^{1/2}}{2^{r_2}}\right)^{w} H(W)^d\right)^{1/wd} \\ &\leq |\mathcal{D}_K|^{1/2d}\binom{N}{w}^{1/2w} H(W)^{1/w},\end{aligned}$$

and so we can take

$$R = \left(\frac{wd}{2^{\frac{wd-1}{wd}}}\right)|\mathcal{D}_K|^{1/2d}\binom{N}{w}^{1/2w} H(W)^{1/w}.$$

This completes the proof. □

Notice that the upper bound of Corollary 2.6.2 reduces to the upper bound of Corollary 2.5.3 in case $K = \mathbb{Q}$ up to multiplication by a constant $\binom{N}{w}^{1/2w}$.

Another interesting immediate corollary of Theorem 2.6.1 in the case $M = 1$ is the following subspace extension lemma.

Corollary 2.6.3. *Let K be a number field as in Theorem 2.6.1. Let $N \geq 2$ be an integer, and let W be a subspace of K^N of dimension w, $1 < w \leq N$. Let*

$V \subseteq W$ be a proper subspace of W of dimension $(w-1) \geq 1$. There exists a point $\boldsymbol{x} \in O_K^N$ such that $W = \mathrm{span}_K\{V, \boldsymbol{x}\}$, and

$$H(\boldsymbol{x}) \leq \left(\mathcal{C}_{K,N}^1(W)H(W)+1\right)\left\{\frac{\mathcal{C}_{K,N}^2(V)^{1/d}}{H(V)}+1\right\}, \tag{2.61}$$

where the constants $\mathcal{C}_{K,N}^1(W)$ and $\mathcal{C}_{K,N}^2(V)$ are as in (2.49).

Finally notice that one can produce the full power of Bombieri - Vaaler version of Siegel's Lemma (Theorem 1.1.1) as a corollary of Theorem 2.6.1. The upper bound may be weaker, but this demonstrates a new approach to the well-known principle. Here is the idea. If W is a subspace of K^N of dimension w, then Corollary 2.6.2 yields a non-zero point $\boldsymbol{x}_1 \in W \cap O_K^N$ of bounded height. Let $V_1 = K\boldsymbol{x}_1 \subseteq W$, $\dim_K(V_1) = 1$. By Theorem 2.6.1, there exists $\boldsymbol{x}_2 \in (W \setminus V_1) \cap O_K^N$ of bounded height. Let $V_2 = \mathrm{span}_K\{\boldsymbol{x}_1, \boldsymbol{x}_2\} \subseteq W$, $\dim_K(V_2) = 2$. Continue iteratively applying Theorem 2.6.1 in the same manner, obtaining a filtration of subspaces

$$V_1 = K\boldsymbol{x}_1 \subset V_2 = \mathrm{span}_K\{\boldsymbol{x}_1, \boldsymbol{x}_2\} \subset \ \ldots \ \subset V_w = \mathrm{span}_K\{\boldsymbol{x}_1, ..., \boldsymbol{x}_w\} = W.$$

This provides a full basis $\boldsymbol{x}_1, ..., \boldsymbol{x}_w \in O_K^N$ for W. In order to bound the height of this basis one needs to estimate heights of the subspaces at each step of the filtration; this may possibly be done by an induction argument.

2.7 Tarski plank problem

We now consider a simple application of our results in the classical case to a certain analogue of the discrete version of the Tarski plank problem. First we provide some background. By a *plank* of width h in $\mathbb{R}^N$ we mean a strip of

space of width h between two parallel $(N-1)$-dimensional hyperplanes. Let C be a convex body of minimal width w in $\mathbb{R}^N$. If C is covered by p planks of widths $h_1, ..., h_p$ respectively, is it true that $h_1 + ... + h_p \geq w$? This question was originally asked by Tarski in [29]. It was answered affirmitively by Bang in [1]. One can also ask for the minimal number of planks with prescribed widths that would cover C. The discrete version of this problem (see for instance [10]) asks for the minimal number of $(N-1)$-dimensional hyperplanes that would cover a convex set of lattice points in $\mathbb{R}^N$. We ask a somewhat different, but analogous question. Consider the set of all integer lattice points in $\mathbb{R}^N$ that are contained in the closed cube C_R^N, where R is a positive integer as above. This set has cardinality $(2R+1)^N$. What is the minimal number of $(N-1)$-dimensional subspaces of $\mathbb{R}^N$ that cover this set? Let M be this number. Then the inequality

$$M \geq 2R - 1 \tag{2.62}$$

follows immediately from (2.34). Further notice that if V is an $(N-1)$-dimensional subspace of $\mathbb{R}^N$ that contains a sublattice of $\mathbb{Z}^N$ of rank $N-1$, then there exists uniquely a linear form $L(\boldsymbol{X}) = \boldsymbol{q} \cdot \boldsymbol{X} \in \mathbb{Z}[X_1, ..., X_N]$ with relatively prime coefficients such that $V = \{\boldsymbol{x} \in \mathbb{R}^N : L(\boldsymbol{x}) = 0\}$. An analogue of width of a plank in this case would be the quantity $|\boldsymbol{q}|^{-1}$, and the sidelength of the cube C_R^N which is equal to $2R$ is an analogue of the width of a convex body. Then we can state the following result, which is an immediate corollary of Theorem 2.5.1.

Corollary 2.7.1. *Let $V_1, ..., V_M$ be $(N-1)$-dimensional subspaces of $\mathbb{R}^N$ which are the nullspaces of the linear forms $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X})$ in N variables with*

relatively prime integer coefficients and coefficient vectors $\boldsymbol{q}_1, ..., \boldsymbol{q}_M$ *respectively. Suppose that* $V_1, ..., V_M$ *cover the set of all integer lattice points contained in the closed cube* C_R^N*. Then*

$$\sum_{i=1}^{M} |\boldsymbol{q}_i|^{-1} \geq R - \sqrt{M}. \tag{2.63}$$

Notice that an analogous statement in the number field case can be easily derived from the bound of Theorem 2.6.1.

A similar problem is treated in [2]. Let C be a compact convex body, which is symmetric with respect to the origin in $\mathbb{R}^N$. Suppose that C can be inscribed into a cube C_R^N as above. How many $(N-1)$-dimensional subspaces of $\mathbb{R}^N$ does it take to cover $C \cap \mathbb{Z}^N$, the set of integer lattice points contained in C? Call this number M. Theorem 2 of [2] provides an upper bound for M in terms of the successive minima of C with respect to $\mathbb{Z}^N$, which implies that M is of the order of magnitude $O(R^{N/(N-1)})$, which is better than (2.62). However, the actual constants in the inequalities of [2] are not effectively computable, since they rely on successive minima. More precisely, Theorem 2 of [2] states that

$$M \leq c 2^N N^2 \log N \min_{0<m<N} (\lambda_m \cdots \lambda_N)^{-\frac{1}{N-m}}, \tag{2.64}$$

where $0 < \lambda_1 \leq ... \leq \lambda_N \leq 1$ are the successive minima of C with respect to $\mathbb{Z}^N$ (the case $\lambda_N > 1$ is trivial: $M = 1$), and c is an absolute constant.

2.8 A system of short integral orthogonal polynomials

In this section we consider a certain application of Siegel's Lemma. Let $M \geq 1$, $N \geq 2$ be integers, and write

$$\mathcal{M} = \mathcal{M}(N, M) = \{\boldsymbol{m} \in \mathbb{Z}_+^N : \sum_{i=1}^{N} m_i = M\}.$$

Then $L = |\mathcal{M}| = \binom{M+N-1}{N-1}$. Write $\mathcal{P}_N^M = \mathbb{R}[X_1, ..., X_N]_M$ for the space of all homogeneous polynomials of degree M in N variables with real coefficients. Each such polynomial $F(X_1, ..., X_M) \in \mathcal{P}_N^M$ can be written as

$$F(\boldsymbol{X}) = \sum_{\boldsymbol{m} \in \mathcal{M}} c(\boldsymbol{m}) \boldsymbol{X}^{\boldsymbol{m}},$$

where $\boldsymbol{X}^{\boldsymbol{m}} = X_1^{m_1} ... X_N^{m_N}$, $c(\boldsymbol{m}) \in \mathbb{R}$ for each $\boldsymbol{m} \in \mathcal{M}$. We also write $\boldsymbol{c} = (c(\boldsymbol{m}))_{\boldsymbol{m} \in \mathcal{M}} \in \mathbb{R}^L$ for the vector of coefficients of F. There is a cannonical isomorphism $\varphi : \mathcal{P}_N^M \longrightarrow \mathbb{R}^L$ given by $\varphi(F) = \boldsymbol{c}$.

We will define a map $I : \mathcal{P}_N^M \times \mathcal{P}_N^M \longrightarrow \mathbb{R}$ given by

$$I(F, G) = \int_{\Sigma_{N-1}} F(\boldsymbol{y}) G(\boldsymbol{y}) d\boldsymbol{y}, \tag{2.65}$$

for each $F, G \in \mathcal{P}_N^M$, where Σ_{N-1} is the unit sphere in $\mathbb{R}^N$ with respect to the L_2-norm, i.e.

$$\Sigma_{N-1} = \left\{ \boldsymbol{y} \in \mathbb{R}^N : \left(\sum_{i=1}^{N} y_i^2 \right)^{1/2} = 1 \right\},$$

and we integrate with respect to the usual Lebesgue measure on $\mathbb{R}^N$.

It is not difficult to see that I is an inner product on $\mathcal{P}_N^M$. In fact, this inner product is particularly important in Fourier Analysis and related subjects (see, for instance Chapter IV of [28]). From now on we write $\mathcal{P}_N^M$

to mean the inner-product space $(\mathcal{P}_N^M, I)$. We will also write $I(F) = I(F, F)$, so that $I^{1/2}(F) = I(F)^{1/2}$ is the norm of $F \in \mathcal{P}_N^M$. In the general spirit of Problem 2, we will consider the following question.

Question 1. *What is a "natural" small-height basis for the inner-product space $\mathcal{P}_N^M$?*

Clearly, one such possibility is just the collection of all monomials in N variables of degree M. This is an integral basis with elements having height equal to 1, however it is not really the most natural basis for $\mathcal{P}_N^M$ as an inner-product space, since it is not orthogonal with respect to the inner product I. In this section (Theorem 2.8.3) we will construct an orthogonal basis for $\mathcal{P}_N^M$ consisting of polynomials of bounded height with integral coefficients.

Define a map $P : \mathcal{M} \longrightarrow \mathbb{Z}$ by

$$P(\boldsymbol{m}) = \prod_{i=1}^{N} \prod_{k=0}^{m_i - 1} (2k+1), \tag{2.66}$$

for each $\boldsymbol{m} = (m_1, ..., m_N) \in \mathcal{M}$. It is not difficult to see that

$$P(\boldsymbol{m}) \leq P(M, 0, ..., 0) = \prod_{k=0}^{M-1} (2k+1),$$

and

$$P(\boldsymbol{m}) \geq \begin{cases} P(1, ..., 1, 0, ..., 0) = 1 & \text{if } N \geq M \\ P(M - N + 1, 1, ..., 1) = \prod_{k=0}^{M-N} (2k+1) & \text{if } N < M \end{cases}$$

for all $\boldsymbol{m} \in \mathcal{M}$. Denote this minimal value of $P(\boldsymbol{m})$ by $P_{\min}$. Let $2\mathcal{M} = \{2\boldsymbol{m} : \boldsymbol{m} \in \mathcal{M}\}$, and let $E = \{(\boldsymbol{m}_1, \boldsymbol{m}_2) \in \mathcal{M} \times \mathcal{M} : \boldsymbol{m}_1 + \boldsymbol{m}_2 \in 2\mathcal{M}\}$.

Let us think of $\mathbb{R}^L$ as $\mathbb{R}^{|\mathcal{M}|}$. Then for each $\boldsymbol{\alpha} \in \mathbb{R}^L$, we write $\boldsymbol{\alpha} = (a(\boldsymbol{m}))_{\boldsymbol{m}\in\mathcal{M}}$, where $\mathcal{M}$ is arranged in lexicographic order. Then define a bilinear map $\mathfrak{L} : \mathbb{R}^L \times \mathbb{R}^L \longrightarrow \mathbb{R}$ by

$$\mathfrak{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{(\boldsymbol{m}_1, \boldsymbol{m}_2)\in E} P\left(\frac{\boldsymbol{m}_1 + \boldsymbol{m}_2}{2}\right) a(\boldsymbol{m}_1) b(\boldsymbol{m}_2), \tag{2.67}$$

for each $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^L \times \mathbb{R}^L$. Also write $\mathfrak{L}(\boldsymbol{\alpha})$ for the corresponding quadratic form $\mathfrak{L}(\boldsymbol{\alpha}, \boldsymbol{\alpha})$.

Lemma 2.8.1. *For each $F \in \mathcal{P}_N^M$,*

$$I(F) = \frac{\pi^{N/2}}{2^{M-1}\Gamma\left(\frac{N}{2} + M\right)} \mathfrak{L}(\varphi(F)), \tag{2.68}$$

where Γ stands for the Gamma-function.

Proof. Notice that

$$I(F) = \sum_{\boldsymbol{m}_1\in\mathcal{M}} \sum_{\boldsymbol{m}_2\in\mathcal{M}} c(\boldsymbol{m}_1) c(\boldsymbol{m}_2) S(\boldsymbol{m}_1, \boldsymbol{m}_2), \tag{2.69}$$

where

$$S(\boldsymbol{m}_1, \boldsymbol{m}_2) = \int_{\Sigma_{N-1}} \boldsymbol{x}^{\boldsymbol{m}_1 + \boldsymbol{m}_2} d\boldsymbol{x} = \int_{\Sigma_{N-1}} \prod_{i=1}^{N} x_i^{\varepsilon_i} d\boldsymbol{x}, \tag{2.70}$$

where $\sum_{i=1}^N \varepsilon_i = 2M$, $\varepsilon_i \in \mathbb{Z}_+$ for all $1 \leq i \leq N$. Consider a change to spherical coordinates $0 \leq \theta_i \leq \pi$ for all $1 \leq i \leq N-2$, $0 \leq \theta_{N-1} \leq 2\pi$, given by

$$x_i = \cos\ \theta_i \prod_{j=1}^{i-1} \sin\ \theta_j, \tag{2.71}$$

for all $1 \leq i \leq N-1$, and $x_N = \prod_{j=1}^{N-1} \sin\ \theta_j$. The Jacobian of this coordinate change is

$$J = \prod_{i=1}^{N-2} \sin^{N-1-i} \theta_i. \tag{2.72}$$

Then

$$S(\boldsymbol{m}_1, \boldsymbol{m}_2) = \left(\prod_{i=1}^{N-2} \int_0^{\pi} \cos^{\varepsilon_i} \theta_i \ \sin^{\beta_i} \theta_i \ d\theta_i \right) \times$$
$$\times \int_0^{2\pi} \cos^{\varepsilon_{N-1}} \theta_{N-1} \ \sin^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1}, \tag{2.73}$$

where $\beta_i = N-1-i+\sum_{j=i+1}^{N} \varepsilon_i$, for all $1 \leq i \leq N-2$. For each $1 \leq i \leq N-2$,

$$\int_0^{\pi} \cos^{\varepsilon_i} \theta_i \ \sin^{\beta_i} \theta_i \ d\theta_i = (-1)^{\beta_i}(1+(-1)^{\varepsilon_i}) \times$$
$$\times \int_0^{\pi/2} \sin^{\varepsilon_i} \theta_i \ \cos^{\beta_i} \theta_i \ d\theta_i = 0, \tag{2.74}$$

unless ε_i is even. Similarly,

$$\begin{aligned}
&\int_0^{2\pi} \cos^{\varepsilon_{N-1}} \theta_{N-1} \ \sin^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1} \\
&\quad = (-1)^{\varepsilon_{N-1}+\varepsilon_N} \times \\
&\quad \times \int_{-\pi}^{\pi} \cos^{\varepsilon_{N-1}} \theta_{N-1} \ \sin^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1} \\
&\quad = (-1)^{\varepsilon_{N-1}}(1+(-1)^{\varepsilon_N}) \times \\
&\quad \times \int_0^{\pi} \cos^{\varepsilon_{N-1}} \theta_{N-1} \ \sin^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1} \\
&\quad = (-1)^{\varepsilon_{N-1}+\varepsilon_N}(1+(-1)^{\varepsilon_N}) \times \\
&\quad \times \int_{-\pi/2}^{\pi/2} \sin^{\varepsilon_{N-1}} \theta_{N-1} \ \cos^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1} \\
&\quad = (1+(-1)^{\varepsilon_{N-1}}+(-1)^{\varepsilon_N}+(1)^{\varepsilon_{N-1}+\varepsilon_N}) \times \\
&\quad \times \int_0^{\pi/2} \sin^{\varepsilon_{N-1}} \theta_{N-1} \ \cos^{\varepsilon_N} \theta_{N-1} \ d\theta_{N-1} \\
&\quad = 0,
\end{aligned} \tag{2.75}$$

unless ε_{N-1} and ε_N are both even. So assume that for each $1 \leq i \leq N$, $\varepsilon_i = 2t_i$ for some $t_i \in \mathbb{Z}_+$. Then $\beta_i = N-1+i+2\sum_{j=i+1}^{N} t_j$, and so $(-1)^{\beta_i} = (-1)^{N-1-i}$.

Putting things together, we see that $S(\boldsymbol{m}_1, \boldsymbol{m}_2) = 0$ unless $(\boldsymbol{m}_1, \boldsymbol{m}_2) \in E$, in which case combining (2.73), (2.74), and (2.75) produces

$$\begin{aligned}
S(\boldsymbol{m}_1, \boldsymbol{m}_2) &= 2^N \left(\prod_{i=1}^{N-2} \int_0^{\pi/2} \sin^{2t_i} \theta_i \ \cos^{\beta_i} \theta_i \ d\theta_i \right) \times \\
&\times \int_0^{\pi/2} \sin^{2t_{N-1}} \theta_{N-1} \ \cos^{2t_N} \theta_{N-1} \ d\theta_{N-1} \\
&= 2 \left(\prod_{i=1}^{N-2} \frac{\Gamma\left(\frac{2t_i+1}{2}\right) \Gamma\left(\frac{\beta_i+1}{2}\right)}{\Gamma\left(\frac{2t_i+\beta_i}{2}+1\right)} \right) \times \\
&\times \frac{\Gamma\left(\frac{2t_{N-1}+1}{2}\right) \Gamma\left(\frac{2t_N+1}{2}\right)}{(t_{N-1}+t_N+1)!} \\
&= \frac{\pi^{N/2}}{2^{M-1}\Gamma\left(\frac{N}{2}+M\right)} \prod_{i=1}^{N} \prod_{k=0}^{t_i-1} (1+2k) \\
&= \frac{\pi^{N/2}}{2^{M-1}\Gamma\left(\frac{N}{2}+M\right)} P\left(\frac{\boldsymbol{m}_1+\boldsymbol{m}_2}{2}\right), \qquad (2.76)
\end{aligned}$$

since $2t_i = m_{1i} + m_{2i}$. The result follows by combining (2.69) and (2.76). □

For each $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^L \times \mathbb{R}^L$, let

$$\mathcal{L}_1(\alpha, \beta) = \frac{\pi^{N/2}}{2^{M-1}\Gamma\left(\frac{N}{2}+M\right)} \mathcal{L}(\alpha, \beta), \qquad (2.77)$$

and also write $\mathcal{L}_1$ for the associated quadratic form. An immediate consequence of Lemma 2.8.1 is that the quadratic form $\mathcal{L}_1$ is positive definite, and so defines a norm $\mathcal{L}_1^{1/2}$ on $\mathbb{R}^L$. Then the bilinear form $\mathcal{L}_1$ defines an inner product on $\mathbb{R}^L$. This corresponds to the inner product I on $\mathcal{P}_N^M$. Then Lemma 2.8.1 can be restated as follows.

Corollary 2.8.2. *The isomorphism φ as defined above is an isometry of the inner product spaces $(\mathcal{P}_N^M, I)$ and $(\mathbb{R}^L, \mathcal{L}_1)$.*

Notice that the matrix of the bilinear form $\mathcal{L}$ is non-singular, and its entries are positive rational numbers with denominators ≤ 2 and numerators divisible by $P_{\min}$. Then matrix of $\frac{2}{P_{\min}}\mathcal{L}$ has positive integer entries, call this matrix $T = (t_{ij})_{1\leq i,j\leq L}$. For any polynomial F we can define a norm $|F|$ to be the maximum of absolute values of its coefficients, then

$$|\mathcal{L}| = |T| = \max_{1\leq i,j\leq L} |t_{ij}|, \tag{2.78}$$

for all $1 \leq i \leq L$. In fact, it is easy to see that

$$|\mathcal{L}| = \frac{1}{P_{\min}} \prod_{k=0}^{M-1} (2k+1). \tag{2.79}$$

Next we construct an orthogonal integral basis of small $|\ |$-norm for the inner product space $(\mathbb{R}^L, \mathcal{L})$ by an iterative application of the classical version of Siegel's Lemma (Theorem 2.1.1).

Write $\boldsymbol{X}$ for the variable vector $(X_1, ..., X_L)$. Let $\boldsymbol{w}_1 = \boldsymbol{e}_1 = (1, 0, ..., 0) \in \mathbb{R}^L$. Define the linear form

$$T_1(\boldsymbol{X}) = \boldsymbol{w}_1 T \boldsymbol{X},$$

so $|T_1| \leq |\mathcal{L}|$. By Theorem 2.1.1, there exists $\boldsymbol{0} \neq \boldsymbol{w}_2 \in \mathbb{Z}^L$ such that

$$T_1(\boldsymbol{w}_2) = \boldsymbol{w}_1 T \boldsymbol{w}_2 = \frac{2}{P_{\min}} \mathcal{L}(\boldsymbol{w}_1, \boldsymbol{w}_2) = 0,$$

and $|\boldsymbol{w}_2| \leq (L|\mathcal{L}|)^{\frac{1}{L-1}}$. Also notice that $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ are linearly independent, since otherwise T would have to be a singular matrix. Let

$$T_2(\boldsymbol{X}) = \boldsymbol{w}_2 T \boldsymbol{X},$$

so $|T_2| \leq L|\boldsymbol{w}_2||\mathcal{L}| \leq (L|\mathcal{L}|)^{\frac{L}{L-1}}$. By Theorem 2.1.1, there exists $\mathbf{0} \neq \boldsymbol{w}_3 \in \mathbb{Z}^L$ such that

$$T_1(\boldsymbol{w}_3) = \boldsymbol{w}_1 T \boldsymbol{w}_3 = \frac{2}{P_{\min}} \mathcal{L}(\boldsymbol{w}_1, \boldsymbol{w}_3) = 0,$$
$$T_2(\boldsymbol{w}_3) = \boldsymbol{w}_2 T \boldsymbol{w}_3 = \frac{2}{P_{\min}} \mathcal{L}(\boldsymbol{w}_2, \boldsymbol{w}_3) = 0,$$

and

$$|\boldsymbol{w}_3| \leq (L \max\{|T_1|, |T_2|\})^{\frac{2}{L-2}} \leq L^{\frac{2(2L-1)}{(L-1)(L-2)}} |\mathcal{L}|^{\frac{2L}{(L-1)(L-2)}}.$$

Again, $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, and $\boldsymbol{w}_3$ are linearly independent, since otherwise T would have to be a singular matrix. Continuing in the same manner, we produce a collection of linearly independent vectors $\boldsymbol{w}_1, ..., \boldsymbol{w}_L \in \mathbb{Z}^L$ such that $\mathcal{L}(\boldsymbol{w}_i, \boldsymbol{w}_j) = 0$ for every $1 \leq i \neq j \leq L$, and

$$|\boldsymbol{w}_i| \leq L^{\prod_{k=1}^{i-1} \frac{2L-k}{L-k}} |\mathcal{L}|^{\prod_{k=1}^{i-1} \frac{L}{L-k}}, \tag{2.80}$$

for each $1 \leq i \leq L$. Let $f_i(\boldsymbol{X}) \in \mathcal{P}_N^M$ be given by $f_i(\boldsymbol{X}) = \varphi^{-1}(\boldsymbol{w}_i)$ for each $1 \leq i \leq L$, then all f_i have integer coefficients, and $I(f_i, f_j) = 0$ whenever $i \neq j$. Also $|f_i| = |\boldsymbol{w}_i|$. Combining (2.79) and (2.80), we have proved the following result.

Theorem 2.8.3. *There exists an orthogonal basis for the inner product space $(\mathcal{P}_N^M, I)$, consisting of polynomials $f_1, ..., f_L$ with integer coefficients such that*

$$|f_i| \leq L^{\prod_{k=1}^{i-1} \frac{2L-k}{L-k}} \left(\frac{1}{P_{\min}} \prod_{k=0}^{M-1} (2k+1) \right)^{\prod_{k=1}^{i-1} \frac{L}{L-k}}, \tag{2.81}$$

for each $1 \leq i \leq L$.

It is easy to see that the same technique can be applied to any inner product on $\mathcal{P}_N^M$ (and other spaces) to produce short orthogonal integral bases.

Let K be a number field, and let $M(K)$ be its set of places. For each $v \in M(K)$, write Ω_v for the completion of the algebraic closure of the completion of K at v, i.e. $\Omega_v = (\overline{K_v})_v$. Hence if v is archimedean, $\Omega_v = \mathbb{C}$. Define

$$\Sigma_{N-1}^v = \left\{ \boldsymbol{y} \in \Omega_v^N : \max_{1 \le i \le N} |y_i|_v \le 1 \right\},$$

for all $v \nmid \infty$, and

$$\Sigma_{N-1}^v = \left\{ \boldsymbol{y} \in \Omega_v^N : \left(\sum_{i=1}^N \|y_i\|_v^2 \right)^{1/2} = 1 \right\},$$

if $v|\infty$.

If F is a homogeneous polynomial of degree M in N variables with coefficients in Ω_v, a convenient way to define a norm $\mathbb{N}_v(F)$ of F (see for example [18] and [34]) is

$$\mathbb{N}_v(F) = \left(\int_{\Sigma_{N-1}^v} |F(\boldsymbol{z})|_v^2 d\boldsymbol{z} \right)^{1/2}, \tag{2.82}$$

if $v|\infty$, and

$$\mathbb{N}_v(F) = \sup \left\{ |F(\boldsymbol{y})|_v : \boldsymbol{y} \in \Sigma_{N-1}^v \right\}, \tag{2.83}$$

if $v \nmid \infty$. Lemma 2.8.1 above was inspired by the following two well known identities (see [25], pp. 16-17, and [34]). If $v|\infty$, then

$$\mathbb{N}_v(F) = \binom{N+M}{N}^{-1} \sum_{\boldsymbol{m} \in \mathcal{M}} \binom{M}{\boldsymbol{m}}^{-1} |c(\boldsymbol{m})|^2, \tag{2.84}$$

where $\binom{M}{\boldsymbol{m}} = \frac{M!}{m_1! \ldots m_N!}$. If $v \nmid \infty$, then

$$\mathbb{N}_v(F) = \max_{\boldsymbol{m} \in \mathcal{M}} |c(\boldsymbol{m})|_v. \tag{2.85}$$

Chapter 3

Small zeros of quadratic forms with linear conditions

3.1 Introduction and notation

Throughout this chapter let

$$F(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=0}^{N} \sum_{j=0}^{N} f_{ij} X_i Y_j \tag{3.1}$$

be a symmetric bilinear form in $N+1$ variables with coefficients $f_{ij} = f_{ji}$. We write $F = (f_{ij})$ for the associated $(N+1) \times (N+1)$ matrix, and $F(\boldsymbol{X}) = F(\boldsymbol{X}, \boldsymbol{X})$ for the associated quadratic form. First assume that the coefficients f_{ij} are in $\mathbb{Q}$. Suppose there exists a point $\boldsymbol{x} \in \mathbb{Q}^{N+1}$ such that $x_0 \neq 0$ and $F(\boldsymbol{x}) = 0$. In [19] Masser shows that in this case there exists such a point $\boldsymbol{x}$ with

$$H(\boldsymbol{x}) \ll_N H(F)^{(N+1)/2}.$$

We presented an exact formulation of this result in Theorem 1.1.4. This generalizes a well known result of Cassels [7] (presented here as Theorem 1.1.3) about the existence of small zeros of quadratic forms with rational coefficients to the existence of small zeros of quadratic polynomials with rational coefficients.

In this chapter we generalize Masser's result in the following way. Let K be a number field of degree d over $\mathbb{Q}$. Let the coefficients f_{ij} be in K. Let M be a positive integer. Let $L_1(\boldsymbol{X}), ..., L_M(\boldsymbol{X})$ be linear forms in $N+1$ variables with coefficients in K. Suppose there exists a point $\boldsymbol{t} \in K^{N+1}$ such that $F(\boldsymbol{t}) = 0$, and $L_i(\boldsymbol{t}) \neq 0$ for each $1 \leq i \leq M$. Then we prove that there exists such a point of bounded height. The bound on height is in terms of the heights of quadratic and linear forms, and reduces (up to a constant) to Masser's type result over a number field in case $M = 1$ and $L_1(\boldsymbol{X}) = X_0$.

First we set some additional notation as in [33]. Let j be a positive integer. For each $v|\infty$, define

$$r_v(j) = \begin{cases} \pi^{-1/2}\Gamma(j/2+1)^{1/j} & \text{if } v|\infty \text{ is real} \\ (2\pi)^{-1/2}\Gamma(j+1)^{1/2j} & \text{if } v|\infty \text{ is complex} \end{cases}$$

For each j, define a field constant

$$C_K(j) = 2|\mathcal{D}_K|^{1/2d} \prod_{v|\infty} r_v(j)^{d_v/d}, \tag{3.2}$$

where $\mathcal{D}_K$ is the discriminant of K. It will also be useful to define another constant for each positive integer j

$$\begin{aligned} A_K(j) &= (8(j+1))^{j/2}|\mathcal{D}_K|^{1/2d} C_K(j)^j \\ &= \left\{2^{5j}(j+1)^j|\mathcal{D}_K|^{\frac{j+1}{d}}\right\}^{1/2} \prod_{v \in M(K)} r_v(j)^{\frac{jd_v}{d}}. \end{aligned} \tag{3.3}$$

Recall that for each $v \nmid \infty$ in $M(K)$, we write O_v for the ring of v-adic integers of K, i.e.

$$O_v = \{x \in K : |x|_v \leq 1\},$$

then O_v is a local ring, and $O_K = \bigcap_{v \nmid \infty} O_v$. We also define the ring of *adeles* of K, denoted $K_{\mathbb{A}}$, in the following manner. Let $P \subseteq M(K)$ be a finite set of places, containig all arhcimidean places. Let

$$K_{\mathbb{A}}(P) = \prod_{v \in P} K_v \times \prod_{v \notin P} O_v, \tag{3.4}$$

and put the usual topology on it. Define addition and multiplication on $K_{\mathbb{A}}(P)$ componentwise, thus making it into a topological ring. Let

$$K_{\mathbb{A}} = \bigcup_{P} K_{\mathbb{A}}(P), \tag{3.5}$$

where union is taken over all sets P as above, i.e. it consists of all the elements $a = (a_v) \in \prod_{v \in M(K)} K_v$ which satisfy $|a_v|_v \leq 1$ for almost all v. It also is a topological ring with each $K_{\mathbb{A}}(P)$ being an open subring of $K_{\mathbb{A}}$. Notice that K can be viewed as contained in $K_{\mathbb{A}}$ under the diagonal embedding $a \longmapsto (a)$ for every $a \in K$. We can now define Haar measure on $K_{\mathbb{A}}$. First we normalize Haar measure γ_v on each completion K_v as follows:

(i) if $v \nmid \infty$, then $\gamma_v(O_v) = |\mathcal{D}_v|_v^{d/2}$, where $\mathcal{D}_v$ is the local different of K at v,

(ii) if $K_v = \mathbb{R}$, then γ_v is the Lebesgue measure on $\mathbb{R}$,

(iii) if $K_v = \mathbb{C}$, then γ_v is twice the Lebesgue measure on $\mathbb{C}$.

Then for each $v|\infty$,

$$\gamma_v\left(\left\{x \in K_v : |x|_v < r_v(1)^{d_v/d}\right\}\right) = 1. \tag{3.6}$$

Define the Haar measure on $K_{\mathbb{A}}$ to be the product measure

$$\gamma = \prod_{v \in M(K)} \gamma_v, \tag{3.7}$$

or more precisely γ is the Haar measure whose restriction to each $K_{\mathbb{A}}(P)$ as above is the product measure. A detailed discussion of adeles can be found for instance in [37].

We also define the ring of *S-integers* as in [21] and introduce the *S-height*. Let $S \subseteq M(K)$ be a finite set of place of K which contains all the archimedean places. The ring of S-integers is given by

$$O_S = \{\alpha \in K : |\alpha|_v \leq 1 \ \ \forall\, v \notin S\}. \tag{3.8}$$

We write O_S^N for the N-fold product of O_S. For each $\boldsymbol{x} \in O_S^N$ define

$$H_S(\boldsymbol{x}) = \prod_{v \in S} H_v(\boldsymbol{x}), \tag{3.9}$$

then for each $\boldsymbol{x} \in O_S^N$

$$H(\boldsymbol{x}) \leq H_S(\boldsymbol{x}). \tag{3.10}$$

Notice that if S is simply the set of all archimedean places of K, then $O_S = O_K$, and $H_S(\boldsymbol{x}) = H_\infty(\boldsymbol{x}) = \prod_{v|\infty} H_v(\boldsymbol{x})$.

Now we can rigorously state the main result of this chapter, which is a precise version of Theorem 1.3.2.

Theorem 3.1.1. *Suppose there exists a point $\boldsymbol{u} \in K^{N+1}$ such that $F(\boldsymbol{u}) = 0$, and $L_i(\boldsymbol{u}) \neq 0$ for each $1 \leq i \leq M$. Then there exists such a point $\boldsymbol{u}$ so that*

$$H(\boldsymbol{u}) \leq B_K(N, M) H(F)^{\frac{N+2M}{2} + (M-1)(N+2)}, \tag{3.11}$$

as well as

$$H(\boldsymbol{u}) \leq B_K(N, M) H(F)^{\frac{N+1}{2} + (M-1)(N+2)} \prod_{i=1}^{M} H(L_i)^{\frac{(2M-1)N}{M}}, \tag{3.12}$$

and finally

$$H(\boldsymbol{u}) \leq B_K(N,M) H(F)^{\frac{2N+2M+1}{4}+(M-1)(N+2)} \prod_{i=1}^{M} H(L_i)^{\frac{(2M-1)N}{2M}}, \tag{3.13}$$

where the constant $B_K(N,M)$ *is given by*

$$\begin{aligned} B_K(N,M) &= \frac{1}{192}(N+1)^2 A_K(N) \left\{486\ (N+1)^6 A_K(N)^2\right\}^{M-1} \times \\ &\times\ (M+2)!\{(M+3)!\}^2, \end{aligned} \tag{3.14}$$

with $A_K(N)$ *as in (3.3).*

The following result is a simple, but useful corollary of Theorem 3.1.1 in the case $M = 1$. This is a precise version of Corollary 1.3.3.

Corollary 3.1.2. *Let* $F(\boldsymbol{X})$ *be a quadratic form in* $N+1$ *variables with coefficients in the number field* K*, as above. Let*

$$\mathcal{V}_K(F) = \left\{\boldsymbol{t} \in K^{N+1} : F(\boldsymbol{t}) = 0\right\}.$$

Suppose that there exists a non-singular point $\mathbf{0} \neq \boldsymbol{x} \in \mathcal{V}_K(F)$*. Then there exists such a point* $\boldsymbol{x}$ *with*

$$H(\boldsymbol{x}) \leq \max\{3, A_K(N)\}\ H(F)^{\frac{N}{2}}. \tag{3.15}$$

The structure of this chapter is the following. In section 3.2 we produce a solution to the problem in case there is only one linear form, obtaining upper bounds for the inhomogeneous height of the point in question, and proving Corollary 3.1.2. Our line of argument here follows that of Masser [19]. In the process of proof we state a generalization of Cassels' result on small zeros of quadratic forms, that we use to construct auxiliary points. In section 3.3 we

prove Theorem 3.1.1 in its general form. It is derived from a slightly more technical result. Our argument is by induction on the number of linear forms, so we use the results of section 3.2 for the base case of the induction, and we use a basic bound of chapter 2 to construct certain auxiliary points. Then we compute bounds on the height. In section 3.4 we produce a solution to our problem with coordinates in S-integers and provide an upper bound on its S-height. Results of this chapter also appear in [13].

3.2 The problem with one linear form

Let $L(\boldsymbol{X})$ be a linear form in $N+1$ variables with coefficients in K, and suppose there exists a point $\boldsymbol{t} \in K^{N+1}$ so that $F(\boldsymbol{t}) = 0$ and $L(\boldsymbol{t}) \neq 0$. We want to show the existence of such a point of small height. The argument of this section parallels that of Masser [19]. We argue by induction on N.

First suppose that $N = 1$, then

$$F(X_0, X_1) = aX_0^2 + bX_0X_1 + cX_1^2,$$

$$L(X_0, X_1) = q_0X_0 + q_1X_1,$$

where $a, b, c, q_0, q_1 \in K$, and not both of q_0, q_1 are zero. If F has only one nonzero coefficient the result is obvious. Hence suppose that two of the coefficients of F are not zero, then at least one of a and c must be nonzero, so assume without loss of generality that $a \neq 0$.

Proposition 3.2.1. *There exists a point $\boldsymbol{x} \in K^2$ such that $F(\boldsymbol{x}) = 0$, $L(\boldsymbol{x}) \neq 0$, and*

$$H(\boldsymbol{x}) \leq h(\boldsymbol{x}) \leq 3H(F). \tag{3.16}$$

Proof. Let $\boldsymbol{x} = (x_0, x_1)$ be a non-trivial zero of F, so $x_0, x_1 \neq 0$. Then

$$(x_0, x_1) = x_1 \left(\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, 1 \right),$$

in other words the zero set of F consists of only two projective points. Hence L must not vanish at one of them. Thus we just have to estimate the heights of these two points. We can assume that $x_1 = 1$, and so for each $v \in M(K)$

$$\max\{1, H_v(x_0, x_1)\} = H_v(x_0, x_1). \tag{3.17}$$

Suppose $v \nmid \infty$. Then

$$H_v(x_0, x_1) = \max\left\{ \left| \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \right|_v, 1 \right\} \leq \max\left\{ \left| \frac{b}{2a} \right|_v, \left| \frac{c}{a} \right|_v^{1/2}, 1 \right\}.$$

On the other hand,

$$\begin{aligned} H_v(F) &= \max\left\{ |a|_v, \left| \frac{b}{2} \right|_v, |c|_v \right\} \\ &= |a|_v \max\left\{ 1, \left| \frac{b}{2a} \right|_v, \left| \frac{c}{a} \right|_v \right\} \\ &\geq |a|_v H_v(x_0, x_1). \end{aligned}$$

Now suppose $v | \infty$. Then

$$H_v(x_0, x_1) = \max\left\{ \left| \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \right|_v, 1 \right\} \leq \max\left\{ \left| \frac{b}{a} \right|_v + \left| \frac{c}{a} \right|_v^{1/2}, 1 \right\}.$$

On the other hand

$$3^{\frac{d_v}{d}} H_v(F) = 3^{\frac{d_v}{d}} |a|_v \max\left\{ 1, \frac{1}{2} \left| \frac{b}{a} \right|_v, \left| \frac{c}{a} \right|_v \right\} \geq |a|_v H_v(x_0, x_1).$$

Then (3.16) follows by using (3.17) and taking a product. □

Next we state a generalized form of Cassels' theorem on small zeros of quadratic forms (Theorem 1.1.3), which we will use in the proof. The following version is due to Vaaler.

Theorem 3.2.2. *If a quadratic form F has a nontrivial zero in K^{N+1}, then there exists $\mathbf{0} \neq \boldsymbol{x} \in O_K^{N+1}$ such that $F(\boldsymbol{x}) = 0$, and*

$$H(\boldsymbol{x}) \leq h(\boldsymbol{x}) \leq A_K(N) H(F)^{N/2}, \tag{3.18}$$

where

$$A_K(N) = (8(N+1))^{N/2} |\mathcal{D}_K|^{1/2d} C_K(N)^N,$$

as in (3.3).

This follows by combining Theorem 1, Corollary 2 and remark after it of [33] with Corollary 11 of [5].

A theorem like this has first been proved for the case $K = \mathbb{Q}$ by Cassels in [7], and later generalized to number fields by Raghavan [23] (various other important generalizations of Cassels' result were also carried out by Birch, Davenport, Chalk, Schmidt, Schlickewei, and Vaaler, just to name a few; see [33] for a more detailed account and bibliography).

We return to the proof. Now assume that $N \geq 2$. Then

$$L(\boldsymbol{X}) = \boldsymbol{q} \,.\, \boldsymbol{X} = \sum_{i=0}^{N} q_i X_i \in K[X_0, ..., X_N]. \tag{3.19}$$

By Theorem 3.2.2, there exists $\mathbf{0} \neq \boldsymbol{x} \in K^{N+1}$ such that $F(\boldsymbol{x}) = 0$ and

$$H(\boldsymbol{x}) \leq h(\boldsymbol{x}) \leq A_K(N) H(F)^{N/2}, \tag{3.20}$$

where $A_K(N)$ is as in (3.3). If $L(\boldsymbol{x}) \neq 0$, we are done, so assume $L(\boldsymbol{x}) = 0$. Again, since $L(\boldsymbol{X})$ is not identically zero, we can assume that for instance $q_0 \neq 0$. This implies that

$$x_0 = -\frac{1}{q_0} \sum_{i=1}^{N} q_i x_i,$$

hence

$$0 = F(\boldsymbol{x}) = \sum_{i=1}^{N}\sum_{j=1}^{N} f_{ij}x_ix_j + 2\sum_{i=1}^{N} f_{0i}x_0x_i + f_{00}x_0^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} g_{ij}x_ix_j,$$

where for each $1 \leq i, j \leq N$, $g_{ij} = f_{ij} - \frac{2q_j}{q_0}f_{0i} + \frac{f_{00}}{q_0^2}q_iq_j$. Then define a quadratic form G in N variables:

$$G(\boldsymbol{X}) = \sum_{i=1}^{N}\sum_{j=1}^{N} g_{ij}X_iX_j.$$

Notice that $\mathbf{0} \neq (x_1, ..., x_N) \in K^N$, and $G(x_1, ..., x_N) = 0$, hence by Theorem 3.2.2, there exists $\mathbf{0} \neq \boldsymbol{z} \in K^N$ such that $G(\boldsymbol{z}) = 0$ and

$$H(\boldsymbol{z}) \leq h(\boldsymbol{z}) \leq A_K(N-1)H(G)^{(N-1)/2}.$$

We will now establish a bound on $H(G)$ in terms of $H(F)$ and $H(L)$. Suppose that $v \nmid \infty$, and notice that $|2|_v \leq 1$. Then

$$H_v(G) \leq H_v(F)\max\left\{1, \left|\frac{2}{q_0}\right|_v H_v(L), \frac{H_v(L)^2}{|q_0|_v^2}\right\} = H_v(F)\frac{H_v(L)^2}{|q_0|_v^2},$$

since $H_v(L) \geq |q_0|_v$. Now suppose that $v|\infty$, then

$$\begin{aligned} H_v(G) &\leq H_v(F)\left\{1 + \frac{2^{\frac{d_v}{d}}H_v(L)}{|q_0|_v} + \frac{H_v(L)^2}{|q_0|_v^2}\right\} \\ &\leq 3^{\frac{d_v}{d}}H_v(F)\max\left\{1, \frac{2^{\frac{d_v}{d}}H_v(L)}{|q_0|_v}, \frac{H_v(L)^2}{|q_0|_v^2}\right\} \\ &\leq 6^{\frac{d_v}{d}}H_v(F)\frac{H_v(L)^2}{|q_0|_v^2}. \end{aligned}$$

Therefore

$$H(z) \leq h(\boldsymbol{z}) \leq A_K(N-1)H(F)^{(N-1)/2}H(L)^{N-1}. \tag{3.21}$$

Define

$$y_0 = -\frac{1}{q_0}\sum_{i=1}^{N} q_iz_i,$$

and let $\mathbf{0} \neq \boldsymbol{y} = (y_0, \boldsymbol{z}) \in K^{N+1}$. By construction, $F(\boldsymbol{y}) = L(\boldsymbol{y}) = 0$. Then using (3.21), we obtain

$$\begin{aligned} H(\boldsymbol{y}) \leq h(\boldsymbol{y}) &\leq N \prod_{v \in M(K)} \frac{H_v(L)}{|q_0|_v} \max\{1, H_v(\boldsymbol{z})\} \\ &\leq 6NA_K(N-1)H(F)^{(N-1)/2}H(L)^N. \end{aligned} \tag{3.22}$$

Since the bilinear form F is not identically zero, there must exist a coefficient $f_{ij} \neq 0$. This implies that

$$\max\{1, H_v(F)\} = H_v(F), \tag{3.23}$$

for each $v \in M(K)$.

Next let $\mathbf{0} \neq \boldsymbol{t}_1, \boldsymbol{t}_2 \in K^{N+1}$, and define

$$\boldsymbol{u}_1 = F(\boldsymbol{t}_1)\boldsymbol{x} - 2F(\boldsymbol{t}_1, \boldsymbol{x})\boldsymbol{t}_1, \tag{3.24}$$

and

$$\boldsymbol{u}_2 = F(\boldsymbol{t}_2)\boldsymbol{y} - 2F(\boldsymbol{t}_2, \boldsymbol{y})\boldsymbol{t}_2. \tag{3.25}$$

It is easy to check that $F(\boldsymbol{u}_1) = F(\boldsymbol{u}_2) = 0$. Let

$$\mathcal{V}_K(F) = \left\{\boldsymbol{t} \in K^{N+1} : F(\boldsymbol{t}) = 0\right\}.$$

Lemma 3.2.3. *Suppose that $\boldsymbol{x}, \boldsymbol{y}$ are non-singular points in the variety $\mathcal{V}_K(F)$. Then there exist non-zero points $\boldsymbol{t}_1$, $\boldsymbol{t}_2$ in K^{N+1} with coordinates $0, \pm 1$ such that*

$$L(\boldsymbol{u}_1),\ L(\boldsymbol{u}_2) \neq 0.$$

Proof. We will go through the construction of $\boldsymbol{t}_1$, and the construction of $\boldsymbol{t}_2$ is identical. Since $L(\boldsymbol{x}) = 0$, we want to construct $\boldsymbol{t}_1 \in K^{N+1}$ such that the following holds:

(i) $t_{10} \neq -\frac{1}{q_0}\sum_{i=1}^{N} q_i t_{1i}$,

(ii) $F(\boldsymbol{t}_1, \boldsymbol{x}) \neq 0$,

(iii) $t_{1i} = 0, \pm 1 \ \ \forall \ \ 0 \leq i \leq N$.

Notice that (i) is equivalent to $L(\boldsymbol{t}_1) \neq 0$, and (ii) is possible since $\boldsymbol{x}$ is non-singular in $\mathcal{V}_K(F)$. Write $\boldsymbol{e}_0, ..., \boldsymbol{e}_N$ for the standard basis vectors. Each $\boldsymbol{e}_i$ satisfies (iii). There exists $\boldsymbol{e}_i$ satisfying (i). If $\boldsymbol{e}_i$ satisfies (ii), let $\boldsymbol{t}_1 = \boldsymbol{e}_i$. Otherwise, there exists $\boldsymbol{e}_j$ satisfying (ii), and $i \neq j$. If $\boldsymbol{e}_j$ satisfies (i), let $\boldsymbol{t}_1 = \boldsymbol{e}_j$. If not, then let $\boldsymbol{t}_1 = \boldsymbol{e}_i + \boldsymbol{e}_j$, and we are done. □

Assume $\boldsymbol{x}, \boldsymbol{y}$ are non-singular points in the variety $\mathcal{V}_K(F)$. Make the choice of $\boldsymbol{t}_1, \boldsymbol{t}_2$ in (3.24), (3.25) as in Lemma 3.2.3. Then $F(\boldsymbol{u}_1) = F(\boldsymbol{u}_2) = 0$, $L(\boldsymbol{u}_1), L(\boldsymbol{u}_2) \neq 0$. We want to estimate heights of $\boldsymbol{u}_1, \boldsymbol{u}_2$.

Lemma 3.2.4. *If $\boldsymbol{t}, \boldsymbol{w} \in K^{N+1}$, and $\boldsymbol{u} = F(\boldsymbol{t})\boldsymbol{w} - 2F(\boldsymbol{t}, \boldsymbol{w})\boldsymbol{t}$, then*

$$H(\boldsymbol{u}) \leq h(\boldsymbol{u}) \leq 3(N+1)^2 H(F) h(\boldsymbol{w}) h(\boldsymbol{t})^2. \tag{3.26}$$

Proof. If $v \nmid \infty$, then $|2|_v \leq 1$, and so

$$\begin{aligned}
\max\{1, H_v(\boldsymbol{u})\} &\leq \max\{1, |F(\boldsymbol{t})|_v H_v(\boldsymbol{w}), |2|_v |F(\boldsymbol{t}, \boldsymbol{w})|_v H_v(\boldsymbol{t})\} \\
&\leq \max\{1, H_v(F) H_v(\boldsymbol{w}) H_v(\boldsymbol{t})^2\} \\
&\leq \max\{1, H_v(F)\} \max\{1, H_v(\boldsymbol{w})\} \max\{1, H_v(\boldsymbol{t})\}^2 \\
&= H_v(F) \max\{1, H_v(\boldsymbol{w})\} \max\{1, H_v(\boldsymbol{t})\}^2,
\end{aligned}$$

where the last equality follows by (3.23). If $v|\infty$, then

$$\begin{aligned} H_v(\boldsymbol{u}) &\leq |F(\boldsymbol{t})|_v H_v(\boldsymbol{w}) + 2|F(\boldsymbol{t},\boldsymbol{w})|_v H_v(\boldsymbol{t}) \\ &\leq \{3(N+1)^2\}^{d_v/d} H_v(F) H_v(\boldsymbol{w}) H_v(\boldsymbol{t})^2, \end{aligned}$$

and so

$$\begin{aligned} \max\{1, H_v(\boldsymbol{u})\} &\leq \{3(N+1)^2\}^{d_v/d} \max\{1, H_v(F) H_v(\boldsymbol{w}) H_v(\boldsymbol{t})^2\} \\ &\leq \{3(N+1)^2\}^{d_v/d} \max\{1, H_v(F)\} \times \\ &\times \max\{1, H_v(\boldsymbol{w})\} \max\{1, H_v(\boldsymbol{t})\}^2 \\ &= \{3(N+1)^2\}^{d_v/d} H_v(F) \max\{1, H_v(\boldsymbol{w})\} \max\{1, H_v(\boldsymbol{t})\}^2, \end{aligned}$$

where the last equality follows by (3.23). Then (3.26) follows by taking a product. □

By Lemma 3.2.3, $h(\boldsymbol{t}_1) = h(\boldsymbol{t}_2) = 1$, and so by Lemma 3.2.4, (3.20), and (3.22) we have

$$\begin{aligned} h(\boldsymbol{u}_1) &\leq 3(N+1)^2 H(F) h(\boldsymbol{x}) \\ &\leq 3(N+1)^2 A_K(N) H(F)^{(N+2)/2}, \end{aligned} \tag{3.27}$$

and

$$\begin{aligned} h(\boldsymbol{u}_2) &\leq 3(N+1)^2 H(F) h(\boldsymbol{y}) \\ &\leq 18N(N+1)^2 A_K(N-1) H(F)^{(N+1)/2} H(L)^N. \end{aligned} \tag{3.28}$$

Next we consider the "singular" case.

Proposition 3.2.5. *Assume that $\boldsymbol{x}$ is a singular point in the variety $\mathcal{V}_K(F)$. Then there exists a point $\boldsymbol{s} \in K^{N+1}$ so that $F(\boldsymbol{s}) = 0$, $L(\boldsymbol{s}) \neq 0$, and*

$$H(\boldsymbol{s}) \leq h(\boldsymbol{s}) \leq 3H(F)^{N/2}. \tag{3.29}$$

Proof. Here the idea is as in [19], to reduce to fewer variables keeping coefficients under control and to use induction. If $N = 1$, (3.29) is just (3.16). Then assume that $N \geq 2$, and that (3.29) has been proved for $N - 1$. Without loss of generality, assume that $x_N \neq 0$. Then $\boldsymbol{x}$ is linearly independent of the first N standard unit vectors $\boldsymbol{e}_0, ..., \boldsymbol{e}_{N-1}$, so we can define new variables $Y_0, ..., Y_N$ by

$$\boldsymbol{X} = (X_0, ..., X_N) = Y_0\boldsymbol{e}_0 + ... + Y_{N-1}\boldsymbol{e}_{N-1} + Y_N\boldsymbol{x}. \tag{3.30}$$

We have

$$F(\boldsymbol{X}) = F\left(\sum_{i=0}^{N-1} Y_i\boldsymbol{e}_i\right) + Y_N^2F(\boldsymbol{x}) + 2F\left(\sum_{i=0}^{N-1} Y_i\boldsymbol{e}_i, Y_N\boldsymbol{x}\right) = F\left(\sum_{i=0}^{N-1} Y_i\boldsymbol{e}_i\right),$$

since $F(\boldsymbol{x}) = 0$, and $\boldsymbol{x}$ is a singular point in V, i.e. $F(\boldsymbol{t}, \boldsymbol{x}) = 0$ for all $\boldsymbol{t} \in K^{N+1}$. Then define a new quadratic form Q in N variables $Y_0, ..., Y_{N-1}$ by

$$Q(\boldsymbol{Y}) = F\left(\sum_{i=0}^{N-1} Y_i\boldsymbol{e}_i\right),$$

and so $F(\boldsymbol{X}) = Q(\boldsymbol{Y})$. Clearly, the coefficients of Q form a subset of coefficients of F, and hence

$$H(Q) \leq H(F). \tag{3.31}$$

There exists a $\boldsymbol{t} \in K^{N+1}$ so that $F(\boldsymbol{t}) = 0$, and $L(\boldsymbol{t}) \neq 0$. Let $\boldsymbol{w} = (w_0, ..., w_{N-1})$ be the vector that corresponds to $\boldsymbol{t}$ under the coordinate change (3.30) and reduction to N variables. Then

$$0 \neq L(\boldsymbol{t}) = L\left(\sum_{i=0}^{N-1} w_i\boldsymbol{e}_i\right) + \frac{t_N}{x_N}L(\boldsymbol{x}) = L\left(\sum_{i=0}^{N-1} w_i\boldsymbol{e}_i\right),$$

since $L(\boldsymbol{x}) = 0$. Then define a new linear form L_1 in N variables $Y_0, ..., Y_{N-1}$ by

$$L_1(\boldsymbol{Y}) = L\left(\sum_{i=0}^{N-1} Y_i \boldsymbol{e}_i\right),$$

and so $L_1(\boldsymbol{w}) \neq 0$, and

$$H(L_1) \leq H(L),$$

since coefficents of L_1 form a subset of coefficients of L. We also know that $Q(\boldsymbol{w}) = F(\boldsymbol{t}) = 0$. Therefore, by induction hypothesis, there exists $\boldsymbol{u} \in K^N$ such that $Q(\boldsymbol{u}) = 0$, $L_1(\boldsymbol{u}) \neq 0$, and

$$h(\boldsymbol{u}) \leq 3H(Q)^{N/2} \leq 3H(F)^{N/2},$$

by (3.31). Define $\boldsymbol{s} = (\boldsymbol{u}, 0) \in K^{N+1}$, and then $F(\boldsymbol{s}) = Q(\boldsymbol{u}) = 0$, $L(\boldsymbol{s}) = L_1(\boldsymbol{u}) \neq 0$, and $h(\boldsymbol{s}) = h(\boldsymbol{u})$. This completes the proof. □

Now notice that if $N \geq 1$, $3(N+1)^2 A_K(N) > 3$, as well as for each $N \geq 2$, $(N+1)^2 A_K(N) \geq N(N+1)^2 A_K(N-1)$. Putting this together with (3.25), (3.27), and (3.28), we have proved the following theorem.

Theorem 3.2.6. *Let the notation be as above. Suppose there exists a point $\boldsymbol{x} \in K^{N+1}$ such that $F(\boldsymbol{x}) = 0$, and $L(\boldsymbol{x}) \neq 0$. Then there exists such $\boldsymbol{x}$ with*

$$\begin{aligned} H(\boldsymbol{x}) \leq h(\boldsymbol{x}) &\leq 18(N+1)^2 A_K(N) H(F)^{(N+1)/2} \times \\ &\times \min\left\{H(F)^{1/2}, H(L)^N\right\}. \end{aligned} \tag{3.32}$$

Proof of Corollary 3.1.2. Let $\boldsymbol{x}$ be the zero of F guranteed by Theorem 3.2.2. If $\boldsymbol{x}$ is non-singular, we are done. If $\boldsymbol{x}$ is singular, let $L(\boldsymbol{X}) = \frac{\partial F}{\partial X_i}$ for some

$0 \leq i \leq N$, so $L(\boldsymbol{x}) = 0$. Then by Proposition 3.2.5, there must exist $\boldsymbol{s} \in K^{N+1}$ so that $F(\boldsymbol{s}) = 0$, $L(\boldsymbol{s}) \neq 0$, and

$$H(\boldsymbol{s}) \leq h(\boldsymbol{s}) \leq 3H(F)^{N/2}.$$

□

3.3 Proof of Theorem 3.1.1

Let M and N be positive integers. Let F be a quadratic form in $N+1$ variables with coefficients in a number field K of degree d, as above. Let $L_1, ..., L_M$ be linear forms in $N+1$ variables with coefficients in K.

Theorem 3.3.1. *Suppose there exists a point $\boldsymbol{u} \in K^{N+1}$ such that $F(\boldsymbol{u}) = 0$, and $L_i(\boldsymbol{u}) \neq 0$ for each $1 \leq i \leq M$. Then there exists such $\boldsymbol{u}$ with*

$$H(\boldsymbol{u}) \leq h(\boldsymbol{u}) \leq B_K(N, M)H(F)^{\frac{N+1}{2}+(M-1)(N+2)} \prod_{i=1}^{M} \mathcal{M}_i^{2-\frac{1}{M}}, \tag{3.33}$$

where

$$\mathcal{M}_i = \min\left\{H(F)^{1/2}, H(L_i)^N\right\}, \tag{3.34}$$

and the constant $B_K(N, M)$ is given by

$$\begin{aligned} B_K(N, M) &= \frac{1}{192}(N+1)^2 A_K(N)\left\{486\ (N+1)^6 A_K(N)^2\right\}^{M-1} \times \\ &\times\ (M+2)!\{(M+3)!\}^2, \end{aligned} \tag{3.35}$$

as in (3.14).

Proof. We will actually prove a slightly stronger upper bound:

$$h(\boldsymbol{u}) \leq B_K(N, M)H(F)^{\frac{N+1}{2}+(M-1)(N+2)}\mathcal{M}_1 \prod_{i=2}^{M} \mathcal{M}_i^2. \tag{3.36}$$

We argue by induction on M. If $M = 1$, then Theorem 3.3.1 follows from Theorem 3.2.6. So suppose $M \geq 2$, and that theorem has been proved for any subset of $L_1, ..., L_M$ of k linear forms, where $1 \leq k \leq M-1$. Then there exist points $\boldsymbol{x}, \boldsymbol{y} \in K^{N+1}$ such that $F(\boldsymbol{x}) = F(\boldsymbol{y}) = 0$, $L_i(\boldsymbol{x}) \neq 0$ for every $1 \leq i \leq M-1$, $L_M(\boldsymbol{y}) \neq 0$, and

$$h(\boldsymbol{x}) \leq B_K(N, M-1)H(F)^{\frac{N+1}{2}+(M-2)(N+2)}\mathcal{M}_1 \prod_{i=2}^{M-1} \mathcal{M}_i^2, \tag{3.37}$$

$$h(\boldsymbol{y}) \leq 18(N+1)^2 A_K(N)H(F)^{(N+1)/2}\mathcal{M}_M. \tag{3.38}$$

Notice that if $b < a$ are positive integers, we interpret $\prod_{i=a}^{b}$ as 1. If $L_M(\boldsymbol{x}) \neq 0$ or $L_i(\boldsymbol{y}) \neq 0$ for all $1 \leq i \leq M-1$, then we are done. So assume it is not so. Then there exists a k, such that $1 \leq k < M-1$ and by reordering the linear forms if necessary we have

(i) $L_i(\boldsymbol{x}) \neq 0$, $L_i(\boldsymbol{y}) \neq 0$, for all $1 \leq i \leq k$,

(ii) $L_i(\boldsymbol{x}) \neq 0$, $L_i(\boldsymbol{y}) = 0$, for all $k < i \leq M-1$,

(iii) $L_M(\boldsymbol{x}) = 0$, $L_M(\boldsymbol{y}) \neq 0$.

Notice that for every $k < i \leq M$, $L_i(\boldsymbol{x}+\boldsymbol{y}) \neq 0$. In fact, there exists a positive integer β such that for all $1 \leq i \leq M$,

$$L_i(\boldsymbol{x} \pm \beta\boldsymbol{y}) \neq 0,$$

for the same choice of $\pm$. For this, β needs to be such that for the same choice of $\pm$ none of the linear equations in β

$$L_i(\boldsymbol{x}) \pm \beta L_i(\boldsymbol{y}) = 0, \quad 1 \leq i \leq k \leq M-2,$$

are true. There are at most $M-2$ such equations, and since we can also choose $\pm$, there exists such a β so that

$$1 \leq \beta \leq \left[\frac{M-2}{2}\right] + 1 \leq \frac{M}{2}. \tag{3.39}$$

Define

$$\boldsymbol{u} = \boldsymbol{x} \pm \beta \boldsymbol{y},$$

for this choice of $\pm$ and β.

Case 1. Suppose $F(\boldsymbol{x}, \boldsymbol{y}) = 0$. Then

$$F(\boldsymbol{u}) = F(\boldsymbol{x}) + \beta^2 F(\boldsymbol{y}) \pm 2\beta F(\boldsymbol{x}, \boldsymbol{y}) = 0,$$

and

$$L_i(\boldsymbol{u}) \neq 0, \quad \forall\ 1 \leq i \leq M.$$

Combining Lemma 1.2.1 and (3.39) we obtain

$$h(\boldsymbol{u}) \leq (\beta+1)h(\boldsymbol{x})h(\boldsymbol{y}) \leq \left(\frac{M+2}{2}\right) h(\boldsymbol{y})h(\boldsymbol{x}). \tag{3.40}$$

Case 2. Suppose $F(\boldsymbol{x}, \boldsymbol{y}) \neq 0$. By Lemma 2.2.2, there exists $\boldsymbol{w} \in K^{N+1}$ such that $L_i(\boldsymbol{w}) \neq 0$ for each $1 \leq i \leq M$ and

$$h(\boldsymbol{w}) \leq \frac{M+2}{2}. \tag{3.41}$$

If $F(\boldsymbol{w}) = 0$, we are done. Assume it is not so. Let β be a positive integer, and define

$$\boldsymbol{u} = F(\boldsymbol{y} \pm \beta\boldsymbol{w})\boldsymbol{x} - 2F(\boldsymbol{x}, \boldsymbol{y} \pm \beta\boldsymbol{w})(\boldsymbol{y} \pm \beta\boldsymbol{w}).$$

Notice that $F(\boldsymbol{u}) = 0$. We want to choose $\pm\beta$ in such a way that the following is true:

(i) $F(\boldsymbol{y} \pm \beta\boldsymbol{w}) = \beta(\beta F(\boldsymbol{w}) \pm 2F(\boldsymbol{y}, \boldsymbol{w})) \neq 0$,

(ii) $F(\boldsymbol{x}, \boldsymbol{y} \pm \beta\boldsymbol{w}) = F(\boldsymbol{x}, \boldsymbol{y}) \pm \beta F(\boldsymbol{x}, \boldsymbol{w}) \neq 0$,

(iii) $L_i(\boldsymbol{u}) = F(\boldsymbol{y} \pm \beta\boldsymbol{w})L_i(\boldsymbol{x}) - 2F(\boldsymbol{x}, \boldsymbol{y} \pm \beta\boldsymbol{w})(L_i(\boldsymbol{y}) \pm \beta L_i(\boldsymbol{w})) \neq 0$, for each $1 \leq i \leq M$.

It is not difficult to see that (i), (ii), (iii) amount to a total of 2 linear and M quadratic expressions in β. Selecting $\pm$ appropriately we see that there exists a positive integer β such that (i), (ii), (iii) are satisfied, and

$$\beta \leq M + 2. \tag{3.42}$$

By the same argument as in section 3.2, we can assume without loss of generality that $f_{00} = 1$. Then, for this choice of $\pm\beta$, Lemma 3.2.4, Lemma 1.2.1, (3.41), and (3.42) imply that

$$\begin{aligned} h(\boldsymbol{u}) &\leq 3(N+1)^2 H(F) h(\boldsymbol{x}) h(\boldsymbol{y} \pm \beta\boldsymbol{w})^2 \\ &\leq 3(N+1)^2(\beta+1)^2 \left(\frac{M+2}{2}\right) H(F) h(\boldsymbol{x}) h(\boldsymbol{y})^2 \\ &\leq \frac{3}{2}(N+1)^2(M+2)(M+3)^2 H(F) h(\boldsymbol{x}) h(\boldsymbol{y})^2. \end{aligned} \tag{3.43}$$

Combining (3.38), (3.40), and (3.43), we have proved that there exists $\boldsymbol{u} \in K^{N+1}$ such that $F(\boldsymbol{u}) = 0$, $L_i(\boldsymbol{u}) \neq 0$ for each $1 \leq i \leq M$, and

$$h(\boldsymbol{u}) \leq 486(N+1)^6 A_K(N)^2 (M+2)(M+3)^2 H(F)^{N+2} \mathcal{M}_M^2 h(\boldsymbol{x}). \tag{3.44}$$

This proves (3.36). Notice that the ordering of linear forms was arbitrary, so assume that $\mathcal{M}_1 = \max_{1 \leq i \leq M} \mathcal{M}_i$. Then $\mathcal{M}_1 \geq \mathcal{M}_1^{1/M} ... \mathcal{M}_M^{1/M}$, and so

$$\mathcal{M}_1 \prod_{i=2}^{M} \mathcal{M}_i^2 \leq \prod_{i=1}^{M} \mathcal{M}_i^{2-\frac{1}{M}}. \tag{3.45}$$

The theorem follows by combining (3.44) with (3.37) and (3.45). $\square$

To derive Theorem 3.1.1, notice that for each $1 \leq i \leq M$, the following inequalities hold:

$$\mathcal{M}_i \leq H(F)^{1/2}, \quad \mathcal{M}_i \leq H(L_i)^N, \quad \mathcal{M}_i \leq H(F)^{1/4} H(L_i)^{N/2}.$$

Combining these with the inequality of Theorem 3.3.1 produces (3.11), (3.12), and (3.13) respectively.

3.4 Solution in S-integers

Let S be any finite set of places of K which contains all the archimedean places. Here we prove that there exists a point $\boldsymbol{u}$ as in Theorem 3.1.1 with coordinates in S-integers and with an explicit bound on S-height.

Lemma 3.4.1. *Let $\mathbf{0} \neq \boldsymbol{x} \in K^{N+1}$. Then there exists $\mathbf{0} \neq \alpha \in K$ such that $\alpha \boldsymbol{x} \in O_S^{N+1}$.*

Proof. We need to construct $\mathbf{0} \neq \alpha \in K$ such that for every $v \notin S$

$$|\alpha|_v \leq H_v(\boldsymbol{x})^{-1}.$$

Since $H_v(\boldsymbol{x}) = 1$ for all but a finite number of places of K, the existence of such α is guaranteed by the Strong Approximation Theorem. However, we will carry out a construction that allows us to produce a bound on $|\alpha|_v$ for all $v \in M(K)$. If $v \nmid \infty$, let

$$\mathfrak{L}_v = \{x \in K_v : |x|_v \leq H_v(\boldsymbol{x})^{-1}\},$$

and if $v|\infty$, let

$$\mathfrak{L}_v = \{x \in K_v : |x|_v^{d/d_v} < r_v(1)\}.$$

Let $\mathfrak{L} = \prod_{v \in M(K)} \mathfrak{L}_v$, then $\mathfrak{L}$ is an admissible subset of $K_{\mathbb{A}}$ in the sence of the geometry of numbers. Let $0 < \lambda_1$ be the successive minimum of $\mathfrak{L}$ with respect to K. Then for every $\lambda > \lambda_1$, there exists $\alpha \in K$ such that $\alpha \in \lambda\mathfrak{L}$. By Minkowski's Convex Body Theorem,

$$\lambda_1 \leq 2\gamma(\mathfrak{L})^{-1/d},$$

where $d = [K : \mathbb{Q}]$. With our normalization of Haar measure, we have

$$\prod_{v \nmid \infty} \gamma_v(O_v) = \prod_{v \nmid \infty} |\mathcal{D}_v|_v^{d/2} = |\mathcal{D}_K|^{-1/2},$$

so that $\gamma(K_{\mathbb{A}}/K) = 1$. Therefore

$$\gamma(\mathfrak{L}) = |\mathcal{D}_K|^{-1/2} \prod_{v \nmid \infty} H_v(\boldsymbol{x})^{-1},$$

and so

$$\lambda_1 \leq 2\ |\mathcal{D}_K|^{1/2d} \prod_{v \nmid \infty} H_v(\boldsymbol{x})^{1/d}.$$

Then for every $v \nmid \infty$,

$$|\alpha|_v \leq H_v(\boldsymbol{x})^{-1},$$

and for every $v|\infty$,

$$|\alpha|_v \leq (\lambda_1 r_v(1))^{d_v/d} \leq (2\ |\mathcal{D}_K|^{1/2d} r_v(1))^{d_v/d} \prod_{u \nmid \infty} H_u(\boldsymbol{x})^{1/d}.$$

With this choice of α, we have

$$H_v(\alpha\boldsymbol{x}) \leq 1, \quad \forall\ v \nmid \infty, \tag{3.46}$$

$$H_v(\alpha\boldsymbol{x}) \leq (2\ |\mathcal{D}_K|^{1/2d} r_v(1))^{d_v/d} H_v(\boldsymbol{x}) \prod_{u \nmid \infty} H_u(\boldsymbol{x})^{1/d}, \quad \forall\ v|\infty. \tag{3.47}$$

Therefore $\alpha\boldsymbol{x} \in O_S^{N+1}$. □

Now let $\boldsymbol{u}$ be as in Theorem 3.1.1, and α as in Lemma 3.4.1 be such that $\alpha\boldsymbol{u} \in O_S^{N+1}$. Notice that by (3.46)

$$H_S(\alpha\boldsymbol{u}) \leq h(\alpha\boldsymbol{u}) = \prod_{v|\infty} \max\{1, H_v(\alpha\boldsymbol{u})\}^{d_v/d}. \tag{3.48}$$

Applying (3.47), we see that for each $v|\infty$

$$\begin{aligned} \max\{1, H_v(\alpha\boldsymbol{u})\} &\leq (2\ |\mathcal{D}_K|^{1/2d} r_v(1))^{d_v/d} \times \\ &\times \max\{1, H_v(\boldsymbol{u})\} \prod_{u\nmid\infty} \max\{1, H_u(\boldsymbol{x})\}^{1/d}, \end{aligned}$$

and since $\sum_{v|\infty} d_v = d$, we obtain

$$H_S(\alpha\boldsymbol{u}) \leq C_K(1) h(\boldsymbol{u}),$$

where $C_K(1) = (2\ |\mathcal{D}_K|^{1/2d} r_v(1))^{d_v/d}$ as in (3.2). Hence we have proved the following.

Theorem 3.4.2. *Let the notation be as in Theorem 3.1.1. Then $\boldsymbol{u}$ can be selected so that $\boldsymbol{u} \in O_S^{N+1}$, where S is any finite set of places of K containing all archimedean places, and the upper bound on $H_S(\boldsymbol{u})$ is $C_K(1)$ times the upper bound on $H(\boldsymbol{u})$ in Theorem 3.1.1.*

In particular this means that we can find a solution $\boldsymbol{u}$ as in Theorem 3.1.1 with $\boldsymbol{u} \in O_K^{N+1}$ and the upper bound on $H_\infty(\boldsymbol{u})$ of the form $C_K(1)$ times the upper bound of Theorem 3.1.1.

Chapter 4

Small zeros of polynomials over $\overline{\mathbb{Q}}$

4.1 Introduction and notation

The celebrated theorem of Cassels [7] and its various generalizations, some of which we discussed in Chapter 3, show that given an isotropic quadratic form over a fixed number field K one can find a zero of bounded height with coordinates in K, satisfying perhaps some additional arithmetic conditions. The next natural step would be to prove an analogous result for zeros of polynomials of higher degree, or for simultaneous zeros of collections of polynomials. This, however, seems to be out of reach at the present time. On the other hand, if we relax the condition that zeros in question have to have coordinates in the fixed number field K, and search for zeros over $\overline{\mathbb{Q}}$ instead, the problem becomes quite accessible. This approach is analogous to the so called "absolute" results, like the absolute Siegel's Lemma of Roy and Thunder, [24]. In this chapter we study the following problem. Given a polynomial or a collection of polynomials with coefficients in a fixed number field K, we want to prove the existence of a zero (common zero) over $\overline{\mathbb{Q}}$ of relatively small height with some additional arithmetic conditions. We use the notation of Chapter 1 and also introduce a few additional conventions.

Recall that all the height functions we are using are absolute, hence well defined over $\overline{\mathbb{Q}}$. For a polynomial

$$g(X) = \sum_{i=0}^{M} a_i X^i \in K[X],$$

in one variable over K we will need one more height function, namely the *Mahler measure*. Let α_1,...,$\alpha_M \in \overline{\mathbb{Q}}$ be the roots of g, and let $E = K(\alpha_1, ..., \alpha_M)$. Notice that each α_i has degree at most M over K. For each $v \in M(E)$, define local Mahler measure of g to be

$$\mu_v(g) = |a_M|_v \prod_{i=1}^{M} \max\{1, |\alpha_i|_v\},$$

and define the *absolute* global Mahler measure of g to be

$$\mu(g) = \prod_{v \in M(E)} \mu_v(g).$$

It is easy to see that

$$\mu(g) = \prod_{i=1}^{M} h(\alpha_i). \tag{4.1}$$

We also recall a well known theorem (see for instance Lemma 2 of [22]), which states that

$$\mu(g) \leq \mathcal{H}(g). \tag{4.2}$$

Putting (4.1) and (4.2) together we have the following lemma.

Lemma 4.1.1. *Let $g(X) \in K[X]$ be a polynomial of degree M in one variable with coefficients over K. There exists $\alpha \in \overline{\mathbb{Q}}$ of degree at most M over K such that $g(\alpha) = 0$, and*

$$h(\alpha) \leq \mathcal{H}(g)^{1/M}. \tag{4.3}$$

Throughout this chapter, let M, N be positive integers, and define

$$\mathcal{M}(N,M) = \left\{ (i_1, ..., i_N) \in \mathbb{Z}_+^N : \sum_{j=1}^{N} i_j = M \right\}, \tag{4.4}$$

and

$$\mathcal{M}'(N,M) = \left\{ (i_1, ..., i_N) \in \mathbb{Z}_+^N : \sum_{j=1}^{N} i_j \leq M \right\}, \tag{4.5}$$

where $\mathbb{Z}_+$ is the set of all *non-negative* integers. Then

$$F(X_1, ..., X_N) = \sum_{\boldsymbol{i} \in \mathcal{M}(N,M)} f_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in K[X_1, ..., X_N],$$

is a *homogeneous* polynomial of degree M in N variables over K, and

$$F(X_1, ..., X_N) = \sum_{\boldsymbol{i} \in \mathcal{M}'(N,M)} f_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in K[X_1, ..., X_N],$$

is an *inhomogeneous* polynomial of degree M in N variables over K.

For a point $\boldsymbol{z} = (z_1, ..., z_N) \in \overline{\mathbb{Q}}^N$, we write $\deg_K(\boldsymbol{z})$ to mean the degree of the extension $K(z_1, ..., z_N)$ over K, i.e.

$$\deg_K(\boldsymbol{z}) = [K(z_1, ..., z_N) : K].$$

We are now ready to state and prove our results. The main result of this chapter is Theorem 4.2.4, which is a precise version of Theorem 1.3.4.

4.2 One polynomial

All constants in the upper bounds of this section have a dependence on the number field K, however it is easily seen that they can be bounded absolutely; we present them in this form for sharpness only. We start with proving a basic bound for zeros of polynomials over $\overline{\mathbb{Q}}$.

Proposition 4.2.1. *Let $M \geq 1$, $N \geq 2$, and $F(X_1, ..., X_N)$ be a non-zero polynomial (homogeneous or not) in N variables of degree M over a number field K with $[K : \mathbb{Q}] = d$. There exists $\mathbf{0} \neq \boldsymbol{z} \in \overline{\mathbb{Q}}^N$ with $\deg_K(\boldsymbol{z}) \leq M$ such that $F(\boldsymbol{z}) = 0$ and*

$$\mathcal{H}(\boldsymbol{z}) \leq \sqrt{2}\, \mathcal{H}(F)^{1/M}. \tag{4.6}$$

Proof. First suppose that F is homogeneous. Write $\boldsymbol{e}_1, ..., \boldsymbol{e}_N$ for the standard basis vectors for $\overline{\mathbb{Q}}^N$ over $\overline{\mathbb{Q}}$. First suppose that for some $1 \leq i \leq N$, $\deg_{X_i} F < M$, then it is easy to see that $F(\boldsymbol{e}_i) = 0$, and $\mathcal{H}(\boldsymbol{e}_i) = 1$.

If $N > 2$, let

$$F_1(X_1, X_2) = F(X_1, X_2, 0, ..., 0),$$

and a point $\boldsymbol{x} = (x_1, x_2) \in \overline{\mathbb{Q}}^2$ is a zero of F_1 if and only if $(x_1, x_2, 0, ..., 0)$ is a zero of F, and

$$\mathcal{H}(x_1, x_2) = \mathcal{H}(x_1, x_2, 0, ..., 0).$$

In particular, if $F_1(X_1, X_2) = 0$, then $F(\boldsymbol{e}_1) = 0$.

Hence we can assume that $N = 2$, $F(X_1, X_2) \neq 0$, and $\deg_{X_1} F = \deg_{X_2} F = M$. Write

$$F(X_1, X_2) = \sum_{i=0}^{M} f_i X_1^i X_2^{M-i},$$

where $f_0, f_M \neq 0$. Let

$$g(X_1) = F(X_1, 1) = \sum_{i=0}^{M} f_i X_1^i \in K[X_1],$$

be a polynomial in one variable of degree M with coefficients in K. Notice that since coefficients of g are those of F, we have $\mathcal{H}(g) = \mathcal{H}(F)$. By Lemma

4.1.1, there must exist $\alpha \in \overline{\mathbb{Q}}$ with $\delta = \deg_K(\alpha) \leq M$ such that $g(\alpha) = 0$, and

$$\mathcal{H}(\alpha, 1) \leq \sqrt{2}\ h(\alpha) \leq \sqrt{2}\ \mathcal{H}(g)^{1/M} = \sqrt{2}\ \mathcal{H}(F)^{1/M}.$$

Taking $\boldsymbol{z} = (\alpha, 1)$, completes the proof in the homogeneous case.

Next assume that F is inhomogeneous. First we homogenize it by introducing an additional variable X_0. If

$$F(X_1, ..., X_N) = \sum_{\boldsymbol{i} \in \mathcal{M}'(N,M)} f_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in K[X_1, ..., X_N],$$

then let

$$F'(X_0, ..., X_N) = \sum_{\boldsymbol{i} \in \mathcal{M}'(N,M)} f_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in K[X_0, ..., X_N],$$

where $i_0 = M - \sum_{j=1}^{N} i_j$, so F' is a homogeneous polynomial in $N+1$ variables of degree M with coefficients in K, such that

$$F(X_1, ..., X_N) = F'(1, X_1, ..., X_N),$$

hence $\mathcal{H}(F') = \mathcal{H}(F)$. There exists $\boldsymbol{x} = (x_0, ..., x_N) \in \overline{\mathbb{Q}}^{N+1}$ so that $x_0 \neq 0$, and

$$F'(x_0, ..., x_N) = F(x_1/x_0, ..., x_N/x_0) = 0.$$

Notice that

$$\mathcal{H}(x_1/x_0, ..., x_N/x_0) = \mathcal{H}(x_1, ..., x_N) \leq \mathcal{H}(x_0, ..., x_N) = \mathcal{H}(\boldsymbol{x}),$$

hence it is sufficient to prove that there exists a zero $\boldsymbol{z} \in \overline{\mathbb{Q}}^{N+1}$ of F' so that $z_0 \neq 0$ and $\boldsymbol{z}$ is of bounded height.

Notice that since the variable X_0 was introduced to homogenize F, we have $\deg(F) = \deg(F') = M$, and so $X_0 \nmid F(X_0, ..., X_N)$. Now, same way as in the

homogeneous case, we can reduce our consideration to the case $N+1=2$, so that we have a polynomial $F'(X_0, X_1) \neq 0$, and $\deg_{X_0} F' = \deg_{X_1} F' = M$. In this case, the argument in the homogeneous case produces a zero of required height with $X_0 \neq 0$. This completes the proof. $\square$

Notice that if $N=2$, then the bound (4.6) is best possible with respect to the exponent in both, homogeneous and inhomogeneous cases. We can demonstrate it by the following examples. In the homogeneous case, take

$$F(X_1, X_2) = X_1^M - CX_2^M,$$

for some $0 \neq C \in K$. Then if $\boldsymbol{z} \neq \mathbf{0}$ is a zero of F, it must be true that

$$z_1 = C^{1/M} z_2,$$

and so the smallest possible zero of F would be

$$\boldsymbol{z} = (C^{1/M}, 1).$$

Let $E = K(C^{1/M})$, and write d'_v, d' for local and global degrees of E over $\mathbb{Q}$ respectively, then $\boldsymbol{z} \in E^2$ and

$$\begin{aligned} \mathcal{H}(\boldsymbol{z}) &= \prod_{v \nmid \infty} H_v(F)^{1/M} \times \prod_{v \mid \infty} \left(1 + \|C\|_v^{2/M}\right)^{d'_v/2d'} \\ &\geq \prod_{v \nmid \infty} H_v(F)^{1/M} \times \prod_{v \mid \infty} \max\{1, \|C\|_v\}^{d'_v/d'M} \\ &\geq \frac{1}{\sqrt{2}} \mathcal{H}(F)^{1/M}. \end{aligned} \tag{4.7}$$

In the inhomogeneous case, take

$$F(X_1, X_2) = X_1 - CX_2^M,$$

for some $0 \neq C \in K$. Then if $\boldsymbol{z} \neq \mathbf{0}$ is a zero of F, it must be true that

$$z_1 = C z_2^M,$$

and so the smallest possible zero of F would be

$$\boldsymbol{z} = \left(1, \frac{1}{C^{1/M}}\right) = \frac{1}{C^{1/M}}(C^{1/M}, 1).$$

Let $E = K(C^{1/M})$, then $\boldsymbol{z} \in E^2$ and

$$\mathcal{H}(\boldsymbol{z}) = \mathcal{H}(C^{1/M}, 1) \geq \frac{1}{\sqrt{2}} \mathcal{H}(F)^{1/M},$$

by the estimate (4.7).

Another remark is that in the proof of the homogeneous case of Proposition 4.2.1 we set all variables except for X_1 and X_2 equal to 0. Of course, the same argument is possible if we set all variables except for X_i and X_j equal to 0, where $1 \leq i \neq j \leq N$. In particular, we can produce N linearly independent points (i.e. a basis for $\overline{\mathbb{Q}}^N$) $\boldsymbol{\alpha}_1 = (1, a_1, 0, ..., 0)$, $\boldsymbol{\alpha}_2 = (0, 1, a_2, 0, ..., 0)$, ..., $\boldsymbol{\alpha}_N = (a_N, 0, ..., 0, 1)$ such that $F(\boldsymbol{\alpha}_i) = 0$ and $\mathcal{H}(\boldsymbol{\alpha}_i) \leq \sqrt{2}\ \mathcal{H}(F)^{1/M}$ for each $1 \leq i \leq N$.

Next we consider the problem of Proposition 4.2.1 with additional arithmetic conditions, similar to those of Chapters 2 and 3. We wonder what can be said about zeros of a polynomial over $\overline{\mathbb{Q}}$ outside of a collection of subspaces? For instance, under which conditions does a polynomial F vanish at a point with non-zero coordinates? Here is a very simple criterion.

Lemma 4.2.2. *Let $N \geq 1$, and let $F(X_1, ..., X_N) \in K[X_1, ..., X_N]$. There exists $\boldsymbol{z} \in \overline{\mathbb{Q}}^N$ such that $F(\boldsymbol{z}) = 0$ and $z_i \neq 0$ for all $1 \leq i \leq N$ if and only if F is not a monomial.*

Proof. Clearly a monomial does not vanish at a point with all coordinates non-zero. We prove the implication in the opposite direction. Let L be the number of monomials of F, so we assume that $L \geq 2$. We argue by induction on L. Assume $L = 2$, then

$$F(X_1, ..., X_N) = A(X_1, ..., X_N) + B(X_1, ..., X_N),$$

where A and B are monomials. There must exist $1 \leq j \leq N$ such that $\deg_{X_j}(A) \neq \deg_{X_j}(B)$. Let $G(X_j) = F(1, ..., 1, X_j, 1, ..., 1)$, then G is a polynomial in one variable which is a sum of two monomials, hence it has a non-zero root $\alpha \in \overline{\mathbb{Q}}$, and this completes the argument in case $L = 2$.

Assume $L > 2$. For each $1 \leq i \leq N - 1$ define

$$F_i(X_{i+1}, ..., X_N) = F(1, ..., 1, X_{i+1}, ..., X_N),$$

and write L_i for the number of monomials of F_i. Then

$$0 \leq L_{N-1} \leq L_{N-2} \leq ... \leq L_1 \leq L.$$

First suppose that $L_{N-1} = L > 2$, then $F_{N-1}(X_N)$ has a non-zero root in $\overline{\mathbb{Q}}$, and we are done. Hence suppose that $L_{N-1} < L$. Then there must exist $1 \leq j \leq N - 1$ such that $L_j = L - 1 > 1$. By the inductive hypothesis, F_j has a zero $\boldsymbol{x} \in \overline{\mathbb{Q}}^{N-j}$ with $x_i \neq 0$ for all $j + 1 \leq i \leq N$, and then $(1, ..., 1, \boldsymbol{x})$ is a required zero of F. This completes the proof. □

The lemma above can be made effective, i.e. one can force the point in question to be of small height with degree $\leq \deg(F)$, where the bound on height will be $O(\mathcal{H}(F))$. Under slightly stronger conditions we can find a zero of F of much smaller height all coordinates of which are non-zero.

Proposition 4.2.3. *Let $F(X_1, ..., X_N)$ be a homogeneous polynomial in $N \geq 2$ variables of degree $M \geq 1$ over a number field K with $[K : \mathbb{Q}] = d$. Suppose that F does not vanish at any of the standard basis vectors $\boldsymbol{e}_1, ..., \boldsymbol{e}_N$. Then there exists $\boldsymbol{z} \in \overline{\mathbb{Q}}^N$ with $\deg_K(\boldsymbol{z}) \leq M$ such that $F(\boldsymbol{z}) = 0$, $z_i \neq 0$ for all $1 \leq i \leq N$, and*

$$\mathcal{H}(\boldsymbol{z}) \leq C(N, M)\ \mathcal{H}(F)^{1/M}, \tag{4.8}$$

where

$$C(N, M) = 2^{\frac{N-1}{2}} \left(\frac{M+2}{2}\right)^{\frac{(4M+1)(N-2)}{2M}} \prod_{j=2}^{N} \binom{M+j-2}{j-2}^{\frac{1}{2M}}. \tag{4.9}$$

Proof. We argue by induction on N. If $N = 2$, then the result follows from the argument in case F is homogeneous and $N = 2$ in the proof of Proposition 4.2.1. Assume $N > 2$. Let β be a positive integer, and let

$$F'_{\pm\beta}(X_1, ..., X_{N-1}) = F(X_1, ..., X_{N-1}, \pm\beta X_{N-1}),$$

in other words set $X_N = \pm\beta X_{N-1}$, where we will specify the choice of $\pm\beta$ later. Let $\boldsymbol{e}'_1, ..., \boldsymbol{e}'_{N-1}$ be the standard basis vectors for $\overline{\mathbb{Q}}^{N-1}$. Notice that if $F'_{\pm\beta}$ vanishes at $\boldsymbol{e}'_i$ for $1 \leq i \leq N-2$, then F vanishes at $\boldsymbol{e}_i$, which is a contradiction. In particular, $F'_{\pm\beta}$ cannot be a monomial and cannot be identically zero. Suppose that $F'_{\pm\beta}(\boldsymbol{e}'_{N-1}) = 0$. This means that $F'_{\pm\beta}(0, ..., 0, X_{N-1})$ is identically zero. Write $\boldsymbol{u}_i = (0, ..., 0, i, M-i) \in \mathbb{Z}^N$ for each $0 \leq i \leq M$. Let

$$\begin{aligned} G(X_{N-1}, X_N) &= F(0, ..., 0, X_{N-1}, X_N) \\ &= \sum_{i=0}^{M} f_{\boldsymbol{u}_i} X_{N-1}^i X_N^{M-i}, \end{aligned}$$

then

$$\begin{aligned} F'_{\pm\beta}(0,...,0,X_{N-1}) &= G(X_{N-1},\pm\beta X_{N-1}) \\ &= \left(\sum_{i=0}^{M} f_{\boldsymbol{u}_i}(\pm\beta)^{M-i}\right) X_{N-1}^{M} = 0, \end{aligned}$$

that is

$$\sum_{i=0}^{M} f_{\boldsymbol{u}_i}(\pm\beta)^{M-i} = 0. \tag{4.10}$$

Notice that $f_{\boldsymbol{u}_0} \neq 0$ and $f_{\boldsymbol{u}_M} \neq 0$, since otherwise $F(\boldsymbol{e}_N) = 0$ or $F(\boldsymbol{e}_{N-1}) = 0$. Therefore the left hand side of (4.10) is a non-zero polynomial of degree M in β, and 0 is not one of its roots, so it has M non-zero roots. Therefore for the appropriate choice of $\pm$ we can select $\beta \in \mathbb{Z}_+$ such that (4.10) is *not* true and

$$0 < \beta \leq \frac{M}{2} + 1 = \frac{M+2}{2}. \tag{4.11}$$

Then for this choice of $\pm\beta$, $F'_{\pm\beta}$ is a polynomial in $N-1$ variables of degree M which does not vanish at any of the standard basis vectors. From now on we will write F'_β instead of $F'_{\pm\beta}$ for this fixed choice of $\pm\beta$.

Next we want to estimate height of such F'_β. For each vector $\boldsymbol{l} \in \mathbb{Z}_+^{N-1}$ such that $\sum_{i=1}^{N-1} l_i = M$ there exist $l_{N-1}+1 \leq M+1$ vectors $\boldsymbol{m}_j \in \mathbb{Z}_+^N$ such that $m_{ji} = l_i$ for each $1 \leq i \leq N-2$ and $m_{j(N-1)} + m_{jN} = l_{N-1}$, where $0 \leq j \leq l_{N-1}$. Therefore the monomial of F'_β which is indexed by $\boldsymbol{l}$ will have coefficient

$$\alpha_{\boldsymbol{l}} = \sum_{j=0}^{l_{N-1}} f_{\boldsymbol{m}_j}(\pm\beta)^{l_{N-1}-j}. \tag{4.12}$$

Then for each $v \nmid \infty$

$$|\alpha_{\boldsymbol{l}}|_v \leq H_v(F), \tag{4.13}$$

and for each $v|\infty$

$$\begin{aligned}
\|\alpha_{\boldsymbol{l}}\|_v^2 &\leq \sum_{i=0}^{l_{N-1}}\sum_{j=0}^{l_{N-1}} \beta^{2l_{N-1}-i-j}\|f_{\boldsymbol{m}_i}\|_v\|f_{\boldsymbol{m}_j}\|_v \\
&\leq \left(\frac{\beta^{2l_{N-1}}}{2}\right)\sum_{i=0}^{l_{N-1}}\sum_{j=0}^{l_{N-1}}(\|f_{\boldsymbol{m}_i}\|_v^2+\|f_{\boldsymbol{m}_j}\|_v^2) \\
&\leq \left(\frac{\beta^{2l_{N-1}}}{2}\right)\sum_{i=0}^{l_{N-1}}\left(\|f_{\boldsymbol{m}_i}\|_v^2+\sum_{j=0}^{l_{N-1}}\|f_{\boldsymbol{m}_j}\|_v^2\right) \\
&\leq \left(\frac{\beta^{2l_{N-1}}(l_{N-1}+2)}{2}\right)\sum_{i=0}^{l_{N-1}}\|f_{\boldsymbol{m}_i}\|_v^2 \\
&\leq \left(\frac{\beta^{2M}(M+2)}{2}\right)\mathcal{H}_v(F)^2 \\
&\leq \left(\frac{M+2}{2}\right)^{2M+1}\mathcal{H}_v(F)^2, \qquad (4.14)
\end{aligned}$$

where the last inequality follows by (4.11). Therefore, by (4.13) and (4.14), we have for each $v\nmid\infty$,

$$H_v(F'_\beta)\leq H_v(F), \qquad (4.15)$$

and for each $v|\infty$,

$$\begin{aligned}
\mathcal{H}_v(F'_\beta) &= \left(\sum_{\boldsymbol{l}\in\mathcal{M}(N-1,M)}\|\alpha_{\boldsymbol{l}}\|_v^2\right)^{1/2} \\
&\leq |\mathcal{M}(N-1,M)|^{1/2}\left(\frac{M+2}{2}\right)^{\frac{2M+1}{2}}\mathcal{H}_v(F) \\
&\leq \binom{M+N-2}{N-2}^{1/2}\left(\frac{M+2}{2}\right)^{\frac{2M+1}{2}}\mathcal{H}_v(F) \qquad (4.16)
\end{aligned}$$

Putting (4.15) and (4.16) together implies that

$$\mathcal{H}(F'_\beta)\leq\binom{M+N-2}{N-2}^{1/2}\left(\frac{M+2}{2}\right)^{\frac{2M+1}{2}}\mathcal{H}(F). \qquad (4.17)$$

By induction hypothesis, there exists $\boldsymbol{x} \in \overline{\mathbb{Q}}^{N-1}$ with $\deg_K(\boldsymbol{x}) \leq M$ such that $F'_\beta(\boldsymbol{x}) = 0$, $x_i \neq 0$ for all $1 \leq i \leq N-1$, and

$$\begin{aligned} \mathcal{H}(\boldsymbol{x}) &\leq C(N-1,M)\ \mathcal{H}(F'_\beta)^{1/M} \\ &\leq C(N-1,M) \binom{M+N-2}{N-2}^{1/2M} \times \\ &\times \left(\frac{M+2}{2}\right)^{\frac{2M+1}{2M}} \mathcal{H}(F)^{1/M}. \end{aligned} \tag{4.18}$$

Let $E = K(x_1, ..., x_{N-1})$. Set $\boldsymbol{z} = (\boldsymbol{x}, \pm\beta x_{N-1}) \in E^N$, then $\delta = \deg_K(\boldsymbol{z}) = [E : K] \leq M$, $F(\boldsymbol{z}) = 0$, $z_i \neq 0$ for all $1 \leq i \leq N$, and applying (4.11) and (4.18) we have

$$\begin{aligned} \mathcal{H}(\boldsymbol{z}) &\leq \prod_{v\nmid\infty} H_v(\boldsymbol{x}) \times \prod_{v|\infty} \left(\beta^2 \|x_{N-1}\|^2 + \mathcal{H}_v(\boldsymbol{x})^2\right)^{d'_v/2d'} \\ &\leq \left(\beta^2+1\right)^{1/2} \mathcal{H}(\boldsymbol{x}) \\ &\leq \sqrt{2} \binom{M+N-2}{N-2}^{\frac{1}{2M}} \left(\frac{M+2}{2}\right)^{\frac{1}{2M}+2} \times \\ &\times C(N-1,M)\ \mathcal{H}(F)^{1/M}, \end{aligned} \tag{4.19}$$

where the product in (4.19) is taken over all places in $M(E)$, and d'_v, d' stand for local and global degrees of E over $\mathbb{Q}$ respectively. The result follows. □

We are now ready to state the main result of this chapter, which is a precise version of Theorem 1.3.4. Recall that for an $N \times N$ matrix A with entries in K we defined height function $\mathcal{H}_*(A)$ by viewing A as a vector in K^{N^2}.

Theorem 4.2.4. *Let $F(X_1, ..., X_N)$ be a homogeneous polynomial in $N \geq 2$ variables of degree $M \geq 1$ over a number field K with $[K : \mathbb{Q}] = d$, and let*

$A \in GL_N(K)$. Then either there exists $\mathbf{0} \neq \boldsymbol{y} \in K^N$ such that $F(\boldsymbol{y}) = 0$ and

$$\mathcal{H}(\boldsymbol{y}) \leq \mathcal{H}_*(A^{-1}), \tag{4.20}$$

or there exists $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ with $\deg_K(\boldsymbol{x}) \leq M$ such that $F(\boldsymbol{x}) = 0$, $A\boldsymbol{x} \in (\overline{\mathbb{Q}}^\times)^N$, and

$$\mathcal{H}(\boldsymbol{x}) \leq C(N,M)\binom{N+M}{N}^{1/2M} \mathcal{H}_*(A^{-1})^2\mathcal{H}(F)^{1/M}, \tag{4.21}$$

where $C(N,M)$ is as in (4.9).

Proof. Let $K[\boldsymbol{X}]_M$ be the space of homogeneous polynomials of degree M in N variables over K. For any element $B \in GL_N(K)$ define a map $\rho_B : K[\boldsymbol{X}]_M \longrightarrow K[\boldsymbol{X}]_M$ (as in [4], [34]), given by $\rho_B(P)(\boldsymbol{X}) = P(B^{-1}\boldsymbol{X})$ for each $P \in K[\boldsymbol{X}]_M$. It is easy to see that each such ρ_B is a representation of $GL_N(K)$ in $GL(K[\boldsymbol{X}]_M)$.

Let $G(\boldsymbol{X}) = \rho_A(F)(\boldsymbol{X})$. First suppose that $G(\boldsymbol{e}_i) = F(A^{-1}\boldsymbol{e}_i) = 0$ for some $1 \leq i \leq N$. Since $\mathbf{0} \neq \boldsymbol{y} = A^{-1}\boldsymbol{e}_i \in K^N$ is a row of A^{-1}, it is easy to see that

$$\mathcal{H}(\boldsymbol{y}) \leq \mathcal{H}_*(A^{-1}),$$

which is (4.20). Next assume that $G(\boldsymbol{e}_i) \neq 0$ for each $1 \leq i \leq N$. By Proposition 4.2.3, there exists $\boldsymbol{z} \in (\overline{\mathbb{Q}}^\times)^N$ such that $G(\boldsymbol{z}) = 0$, $\deg_K(\boldsymbol{z}) \leq M$, and

$$\mathcal{H}(\boldsymbol{z}) \leq C(N,M)\ \mathcal{H}(G)^{1/M}.$$

Then $\boldsymbol{x} = A^{-1}\boldsymbol{z}$ is such that $F(\boldsymbol{x}) = 0$, $\deg_K(\boldsymbol{x}) \leq M$, and $A\boldsymbol{x} = \boldsymbol{z} \in (\overline{\mathbb{Q}}^\times)^N$. It is easy to see that

$$\mathcal{H}(\boldsymbol{x}) \leq \mathcal{H}_*(A^{-1})\mathcal{H}(\boldsymbol{z}) \leq C(N,M)\ \mathcal{H}_*(A^{-1})\mathcal{H}(G)^{1/M}. \tag{4.22}$$

We now want to estimate $\mathcal{H}(G)$. Let $v \in M(K)$, and suppose $v \nmid \infty$. Then combining (1.4) and (1.5) of [34] with (2.2) of [4], we have

$$H_v(G) \leq H_v(A^{-1})^M H_v(F).$$

Now assume $v|\infty$. Then combining (1.4) and (1.5) of [34], we have

$$\mathcal{H}_v(G) \leq \binom{N+M}{N}^{d_v/2d} \mathcal{H}_v(A^{-1})^M \mathcal{H}_v(F).$$

Therefore

$$\mathcal{H}(G) \leq \binom{N+M}{N}^{1/2} \mathcal{H}_*(A^{-1})^M \mathcal{H}(F). \tag{4.23}$$

The result follows by combining (4.22) and (4.23). □

Notice that Theorem 4.2.4 can be thought of as a statement about the existence of a point of bounded height at which F vanishes and which is outside of the union of nullspaces of row vectors of A.

The result of Theorem 4.2.4 shows that we can place a certain number of additional arithmetic conditions on a polynomial zero over $\overline{\mathbb{Q}}$ and still keep the exponent in the upper bound to be $1/M$. This suggests that perhaps the optimal exponent in the upper bound of Proposition 4.2.1 should be smaller than $1/M$. In fact, the exponent $1/M$ in the upper bound of Proposition 4.2.1 only seems to be optimal when $N = 2$, so one may expect an exponent that would depend on N as well as on M.

Conjecture 1. *Let F be a homogeneous polynomial in $N \geq 2$ variables of degree $M \geq 1$ with coefficients in a number field K. Then there exists a non-zero point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ such that $F(\boldsymbol{x}) = 0$ and $\mathcal{H}(\boldsymbol{x}) \ll_{N,M} \mathcal{H}(F)^{\frac{1}{M(N-1)}}$.*

This conjecture is easy to verify for diagonal forms.

Proposition 4.2.5. *Let K be a number field, and $N \geq 2$, $M \geq 1$ integers. Let $F(X_1, ..., X_N) = \sum_{i=1}^{N} f_i X_i^M \in K[X_1, ..., X_N]$. For each $\epsilon > 0$, there exists a non-zero point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ (which depends on the choice of ϵ) such that $F(\boldsymbol{x}) = 0$, and*

$$\mathcal{H}(\boldsymbol{x}) \leq \left(2^{\frac{(N-1)(N-2)}{2}} + \epsilon\right)^{\frac{1}{M(N-1)}} \mathcal{H}(F)^{\frac{1}{M(N-1)}}. \tag{4.24}$$

Moreover, the exponent in the upper bound of (4.24) is best possible.

Proof. For each $1 \leq i \leq N$, let $Y_i = X_i^M$, and let

$$G(Y_1, ..., Y_N) = \sum_{i=1}^{N} f_i Y_i,$$

then G is a linear form in N variables $Y_1, ..., Y_N$ over K, and $\mathcal{H}(G) = \mathcal{H}(F)$. Let V be the nullspace of G in $\overline{\mathbb{Q}}^N$, then V is also defined over K, and dimension of V is $N-1$. By the duality principle of Lemma 1.2.2, $\mathcal{H}(V) = \mathcal{H}(G) = \mathcal{H}(F)$. Therefore, by the absolute version of Siegel's Lemma due to Thunder and Roy (Theorem 1.1.2), for each $\epsilon > 0$ there exists a non-zero point $\boldsymbol{y} \in V$ over $\overline{\mathbb{Q}}$ (which depends on the choice of ϵ) such that

$$\mathcal{H}(\boldsymbol{y}) \leq \left(2^{\frac{(N-1)(N-2)}{2}} + \epsilon\right)^{\frac{1}{N-1}} \mathcal{H}(F)^{\frac{1}{N-1}}. \tag{4.25}$$

Then $G(\boldsymbol{y}) = 0$. Let $\boldsymbol{x} = (x_1, ..., x_N) \in \overline{\mathbb{Q}}^N$ be such that $y_i = x_i^M$ for each $1 \leq i \leq N$, then $0 = G(\boldsymbol{y}) = F(\boldsymbol{x})$. Also, $\mathcal{H}(\boldsymbol{x}) = \mathcal{H}(\boldsymbol{y})^{\frac{1}{M}}$. Combining this with (4.25) yields (4.24). The exponent in this upper bound is best possible since the exponent of Theorem 1.1.2 is best possible (see [24], Lemma 4.7). □

4.3 Many polynomials

Suppose that $F_1, ..., F_k$ are k homogeneous polynomials in $N \geq k+1$ variables of degrees $M_1, ..., M_k$ respectively with coefficients in a number field K as above. By classical Bezout's theorem they must have a non-trivial common zero over $\overline{\mathbb{Q}}$. In fact, without loss of generality we can set $X_{k+2} = ... = X_N = 0$, and consider k polynomials in $k+1$ variables. For each $1 \leq i \leq k$, write

$$\mathcal{V}(F_i) = \{P \in \mathbb{P}(\overline{\mathbb{Q}})^k : F_i(P) = 0\},$$

then $\cap_{i=1}^k \mathcal{V}(F_i)$ is a 0-dimensional variety of degree $M = M_1...M_k$, i.e. a set of $M = M_1...M_k$ points $P_1, ..., P_M$ up to multiplicity (i.e. some of the points may be the same). What can be said about height of those points, in particular is it possible to prove that at least one of them must be of relatively small height? A basic result of this kind follows immediately from the so called Arithmetic Bezout's Theorem (see for instance Proposition 4.9 of [18] and Theorem 4.2.3 of [6]). This was pointed out to me by Professor Felipe Voloch.

Theorem 4.3.1. *Let $F_1, ..., F_k$ be k homogeneous polynomials in $k+1$ variables of degrees $M_1, ..., M_k$ respectively with coefficients in K such that none of the polynomials are multiples of the others. Then there exists a point $\mathbf{0} \neq \boldsymbol{x} \in \overline{\mathbb{Q}}^{k+1}$ such that $F_1(\boldsymbol{x}) = ... = F_N(\boldsymbol{x}) = 0$, and*

$$\mathcal{H}(\boldsymbol{x}) \leq C'(k, M) \prod_{i=1}^{k} \mathcal{H}(F_i)^{1/M_i}, \tag{4.26}$$

where an explicit value for $C'(N, M)$ can be deduced from [18] or from [6].

Notice that (as we pointed out above) if we have polynomials in more than $N > k+1$ variables, then we set all but $k+1$ variables equal to 0 and apply

Theorem 4.3.1. This is the same principle as in the proof of Proposition 4.2.1, when we set all but 2 variables equal to zero for 1 polynomial. This principle is essentially provided by the classical Bezout's theorem, which guarantees the existence of zeros. This means that the bound of Theorem 4.3.1 in this case is again basic, and the optimal bound should be sharper.

Conjecture 2. *Let $F_1, ..., F_k$ be homogeneous polynomials in $N \geq k+1$ variables of respective degrees $M_1, ..., M_k \geq 1$ with coefficients in a number field K. Then there exists a non-zero point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ such that $F_i(\boldsymbol{x}) = 0$ for each $1 \leq i \leq k$, and $\mathcal{H}(\boldsymbol{x}) \ll_{N,M} \prod_{i=1}^{k} \mathcal{H}(F_i)^{\frac{1}{M_i(N-k)}}$.*

A bound like this can no longer be obtained from Arithmetic Bezout's Theorem, since the intersection cycle of k hypersurfaces in $(N-1)$-dimensional projective space when $N > k+1$ is no longer 0-dimensional. On the other hand, proving an upper bound for the height of points on general projective varieties and intersection cycles is presently out of reach. Conjecture 2, same as Conjecture 1, is easy to verify for a collection of diagonal forms of the same degree.

Proposition 4.3.2. *Let K be a number field, and $k \geq 1$, $N \geq k+1$, $M \geq 1$ integers. For each $1 \leq i \leq k$, let $F_i(X_1, ..., X_N) = \sum_{j=1}^{N} f_{ij} X_j^M \in K[X_1, ..., X_N]$. Write F for the $k \times N$ matrix (f_{ij}). For each $\epsilon > 0$, there exists a non-zero point $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ (which depends on the choice of ϵ) such that $F_i(\boldsymbol{x}) = 0$ for each $1 \leq i \leq k$, and*

$$\begin{aligned} \mathcal{H}(\boldsymbol{x}) &\leq \left(2^{\frac{(N-k)(N-k-1)}{2}} + \epsilon\right)^{\frac{1}{M(N-k)}} \mathcal{H}(F)^{\frac{1}{M(N-k)}} \\ &\leq \left(2^{\frac{(N-k)(N-k-1)}{2}} + \epsilon\right)^{\frac{1}{M(N-k)}} \prod_{i=1}^{k} \mathcal{H}(F_i)^{\frac{1}{M(N-k)}}. \end{aligned} \tag{4.27}$$

Moreover, the exponent in the upper bound of (4.27) is best possible.

Proof. The proof is essentially identical to that of Proposition 4.2.5. For each $1 \leq j \leq N$, let $Y_j = X_j^M$. Let

$$V = \{\boldsymbol{y} \in \overline{\mathbb{Q}}^N : F\boldsymbol{y} = \boldsymbol{0}\},$$

then V is a subspace of $\overline{\mathbb{Q}}^N$ also defined over K, and dimension of V is $N-k$. By the duality principle of Lemma 1.2.2, $\mathcal{H}(V) = \mathcal{H}(F)$. By Theorem 1.1.2, for each $\epsilon > 0$ there exists a non-zero point $\boldsymbol{y} \in V$ over $\overline{\mathbb{Q}}$ (which depends on the choice of ϵ) such that

$$\mathcal{H}(\boldsymbol{y}) \leq \left(2^{\frac{(N-k)(N-k-1)}{2}} + \epsilon\right)^{\frac{1}{N-k}} \mathcal{H}(F)^{\frac{1}{N-k}}. \tag{4.28}$$

Let $\boldsymbol{x} = (x_1, ..., x_N) \in \overline{\mathbb{Q}}^N$ be such that $y_i = x_i^M$ for each $1 \leq i \leq N$, then $F_i(\boldsymbol{x}) = 0$ for each $1 \leq i \leq k$. Also, $\mathcal{H}(\boldsymbol{x}) = \mathcal{H}(\boldsymbol{y})^{\frac{1}{M}}$. Combining this with (4.28) yields the first inequality of (4.27). The exponent in this upper bound is best possible since the exponent of Theorem 1.1.2 is best possible (see [24], Lemma 4.7). The second inequality of (4.27) follows by a basic inequality for heights (see for instance Lemma 4.7 of [24]). □

Bibliography

[1] T. Bang. A solution of the "Plank Problem". *Amer. Math. Soc. Proceedings*, 2:990–993, 1951.

[2] I. Bárány, G. Harcos, J. Pach, and G. Tardos. Covering lattice points by subspaces. *Period. Math. Hungar.*, 43:93–103, 2001.

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

[4] E. Bombieri, A. J. Van Der Poorten, and J. D. Vaaler. Effective measures of irrationality for cubic extensions of number fields. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 23(2):211–248, 1996.

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

[6] J.-B. Bost, H. Gillet, and C. Soulé. Heights of projective varieties and positive Green forms. *J. Amer. Math. Soc.*, 7(4):903–1027, 1994.

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

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

[9] J. W. S. Cassels. *Rational quadratic forms.* Academic Press, Inc., 1978.

[10] C. Corzatt. Covering convex sets of lattice points with straight lines. *Proceedings of the Sundance conference on combinatorics and related topics. Congr. Numer.*, 150:129–135, 1985.

[11] B. Edixhoven. Arithmetic part of Faltings's proof. *Diophantine approximation and abelian varieties (Soesterberg, 1992)*, Lecture Notes in Math.(1566):97–110, 1993.

[12] G. Faltings. Diophantine approximation on abelian varieties. *Ann. of Math.*, 133(2):549–576, 1991.

[13] L. Fukshansky. Small zeros of quadratic forms with linear conditions. *to appear in J. Number Theory.*

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

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

[16] R. J. Kooman. Faltings's version of Siegel's lemma. *Diophantine approximation and abelian varieties (Soesterberg, 1992)*, Lecture Notes in Math.(1566):93–96, 1993.

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

[18] M. Laurent and D. Roy. Criteria of algebraic independence with multiplicities and approximation by hypersurfaces. *J. Reine Angew. Math.*, 538:65–114, 2001.

[19] D. W. Masser. How to solve a quadratic equation in rationals. *Bull. London Math. Soc.*, 30(1):24–28, 1998.

[20] D. G. Northcott. An inequality in the theory of arithmetic on algebraic varieties. *Proc. Camb. Phil. Soc.*, 45:502–509 and 510–518, 1949.

[21] R. O'Leary and J. D. Vaaler. Small solutions to inhomogeneous linear equations over number fields. *Trans. Amer. Math. Soc.*, 336(2):915–931, 1993.

[22] C. G. Pinner and J. D. Vaaler. The number of irreducible factors of a polynomial. I. *Trans. Amer. Math. Soc.*, 339(2):809–834, 1993.

[23] S. Raghavan. Bounds of minimal solutions of diophantine equations. *Nachr. Akad. Wiss. Gottingen, Math. Phys. Kl.*, 9:109–114, 1975.

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

[25] W. Rudin. *Function theory in the unit ball of* $\mathbb{C}^N$. Springer-Verlag, 1980.

[26] W. M. Schmidt. On heights of algebraic subspaces and Diophantine approximation. *Ann. of Math. (2)*, 85:430–472, 1967.

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

[28] E. Stein and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, 1971.

[29] A. Tarski. Further remarks about the degree of equivalence of polygons (in Polish). *Odbitka Z. Parametru.*, 2:310–314, 1932.

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

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

[32] J. L. Thunder. Higher-dimensional analogues of Hermite's constant. *Michigan Math. J.*, 45(2):301–314, 1998.

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

[34] J. D. Vaaler. An inequality for the distance to a projective variety. *unpublished notes*, 1992.

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

[36] J. D. Vaaler. Notes on lattice points in spheres and cubes. *unpublished manuscript*, 2003.

[37] A. Weil. *Basic number theory (Reprint of the second (1973) edition).* Springer-Verlag (Classics in Mathematics), 1995.

Vita

Leonid Eugene Fukshansky was born in Leningrad, Russia, on July 1, 1973, son of Eugene and Polina Fukshansky. After completing his High School studies in Leningrad, Russia, in 1990, and briefly attending Leningrad Polytechnic Institute, he, along with his family, emmigrated to Palo Alto, CA, USA. He attended Foothill and DeAnza Community Colleges, and then entered University of California at Los Angeles (UCLA) in 1992. He received B. S. in Applied Mathematics with Specialization in Computing from UCLA in 1995. He then worked in the computer industry as a software engineer and attended part-time San Jose State University until January of 1998, when he moved for a semester to the University of Nebraska at Lincoln. He entered the Graduate School of the University of Texas at Austin in September of 1998.

Permanent address: 4100 Avenue C #311
Austin, Texas 78751

This dissertation was typeset with LaTeX[‡] by the author.

[‡]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.