Siegel's Lemma and Minkowski's Theorems

Master's Thesis

Louna Seppälä

Department of Mathematical Sciences

University of Oulu

Autumn 2015

Contents

Introduction **3**

1 Preliminaries **5**

1.1 Some notations 5

1.2 Basics on algebraic number theory 6

2 Siegel's lemma **8**

2.1 Siegel's lemma in the field of rational numbers 9

2.2 Siegel's lemma in imaginary quadratic fields 11

3 Siegel's lemma in an arbitrary algebraic number field **22**

4 Minkowski's convex body theorems **29**

4.1 Convex bodies 29

4.2 Lattices 32

4.3 The convex body theorems 35

5 Some Diophantine inequalities **44**

5.1 A few Diophantine inequalities over $\mathbb{R}$ 44

5.2 On simultaneous Diophantine inequalities 51

5.3 A Diophantine inequality over $\mathbb{C}$ 53

6 Applying Siegel's lemma **57**

6.1 Approximations for $e^{\alpha t}$ 57

6.2 Siegel's lemma refined 61

6.3 Vandermonde determinants 65

Bibliography **71**

Introduction

Geometry of numbers is a branch of mathematics founded by the German Hermann Minkowski [1], which uses geometric arguments in n-dimensional euclidean space to prove arithmetic results. The diverse applications reach from coding theory to functional analysis, and have turned out to be especially useful in Diophantine approximation, the problem of approximating irrational numbers by rational ones. Siegel's [2] and Minkowski's existential theorems form the core of this work: When dealing with a group of linear equations where the number of unknowns exceeds the number of equations, Siegel's lemma confirms the existence of a non-trivial solution whose size is bounded by a certain positive function depending on the coefficients of the linear forms and the number of unknowns. Minkowski's theorems in turn concern convex bodies and lattices in the space $\mathbb{R}^n$: when a convex body satisfies a specific condition with respect to the lattice, it is bound to intersect the lattice in a non-zero point.

In the first part of the text three versions of Siegel's lemma are presented and proved. The use of geometry in the proof of the second version is particularly delightful, as one can visualize the mathematics behind the theorem. Geometric thinking comes into focus even more when moving to the second

[1] 1864 - 1909

[2] Carl Ludwig Siegel, 1896 - 1981, Germany

part, the purpose of which is to introduce two theorems of Minkowski and then explore the consequences by presenting a selection of Diophantine inequalities. The final part returns to Siegel's lemma again with an example on its use. A fourth version of the lemma is briefly mentioned because of the possible affordance the example provides: the ending concentrates on finding out whether this Bombieri-Vaaler version of Siegel's lemma could be applied to improve the result of the example in rational case.

Chapter 1, rather list-like as it is, should offer enough basics on algebraic number theory for the reader to be able to follow. The results of linear algebra are also used extensively throughout the text. Apart from these, there are not much serious prerequisites. The source material is listed in Bibliography and referred to in appropriate points of the study. Whenever a proof has some kind of a source, it is mentioned in the beginning.

Chapter 1

Preliminaries

The reader is assumed to be familiar with the fundamentals of algebraic numbers, but some important concepts are shortly covered here. For more details, consult [10] or [14], for example. The definitions in this chapter follow the lecture notes [10].

1.1 Some notations

The following sets appear in this text:

$\mathbb{Q}$	rational numbers
$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$	rational integers
$\mathbb{Z}^+ = \{k \in \mathbb{Z} \mid k \geq 1\}$	positive rational integers
$\mathbb{N} = \{0, 1, 2, \ldots\}$	natural numbers
$\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$	positive real numbers
$R[x]$	polynomials in x with coefficients in R
$\mathcal{M}_{k\times n}(R)$	$k \times n$ matrices with entries in R

Here R is a ring or field specified later. A column vector $\bar{x} \in R^n$ is denoted by

$(x_1, \ldots, x_n)^T$ and the standard basis vectors by $\bar{e}_1 = (1, 0, \ldots, 0)^T, \ldots, \bar{e}_n = (0, 0, \ldots, 1)^T$. For the length of a vector $\bar{x} \in R^n$ the norms

$$\|\bar{x}\|_1 = \sum_{k=1}^{n} |x_k|, \quad \|\bar{x}\|_2 = \|\bar{x}\| = \sqrt{\sum_{k=1}^{n} |x_k|^2} \quad \text{and} \quad \|\bar{x}\|_\infty = \max_{1 \leq k \leq n} |x_k|$$

are used. The volume $V(\mathcal{C})$ of a subset $\mathcal{C} \subseteq \mathbb{R}^n$ means the Riemann or Lebesgue integral over $\mathcal{C}$, when it exists.

1.2 Basics on algebraic number theory

Definition 1. A number $\alpha \in \mathbb{C}$ is *algebraic* if there exists a polynomial $p(x) \in \mathbb{Q}[x] \setminus \mathbb{Q}$ such that $p(\alpha) = 0$. Otherwise, α is *transcendental*.

Definition 2. The *minimum polynomial* of an algebraic number $\alpha \in \mathbb{C}$ is the monic polynomial $M_\alpha(x) \in \mathbb{Q}[x] \setminus \mathbb{Q}$ of lowest degree for which $M_\alpha(\alpha) = 0$.

Definition 3. The *degree* of an algebraic number $\alpha \in \mathbb{C}$ is the degree of its minimum polynomial: $\deg_{\mathbb{Q}} \alpha := \deg M_\alpha(x)$.

Definition 4. An algebraic number $\alpha \in \mathbb{C}$ is an *algebraic integer* when $M_\alpha(x) \in \mathbb{Z}[x]$. The ring of all algebraic integers is denoted by the symbol $\mathbb{B}$.

Definition 5. An *algebraic number field* $\mathbb{K}$ is a finite field extension of $\mathbb{Q}$.

Theorem 1. *If $\mathbb{K}$ is an algebraic number field, then $\mathbb{K} = \mathbb{Q}(\alpha)$ for some $\alpha \in \mathbb{K}$. In other words, all algebraic number fields are field extensions of $\mathbb{Q}$ generated by one element.*

Proof. See [10, p. 47] or [14, p. 40], for example. □

Definition 6. The *degree* of an algebraic number field $\mathbb{K}$ is $[\mathbb{K} : \mathbb{Q}] := \dim_{\mathbb{Q}} \mathbb{K}$.

Theorem 2. *If* $\mathbb{K} = \mathbb{Q}(\alpha)$ *for some algebraic number* $\alpha \in \mathbb{C}$ *and* $\deg_{\mathbb{Q}} \alpha = m$, *then* $[\mathbb{K} : \mathbb{Q}] = m$.

Proof. See [14, p. 23]. (Follows from the fact that the field $\mathbb{K}$ is a vector space over $\mathbb{Q}$ spanned by the elements $1, \alpha, \ldots, \alpha^{m-1}$.) □

Definition 7. A *quadratic field* $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$, where $d \in \mathbb{Z}$ is square-free, is an algebraic number field of degree two over $\mathbb{Q}$.

Definition 8. The set of integers of an algebraic number field $\mathbb{K}$ is $\mathbb{Z}_{\mathbb{K}} = \mathbb{K} \cap \mathbb{B}$.

Theorem 3. *Let* $\mathbb{K} = \mathbb{Q}(\sqrt{d})$, $d \in \mathbb{Z}$, *be a quadratic field. Then*

$$\mathbb{Z}_{\mathbb{K}} = \left\{ a + b\sqrt{d} \;\middle|\; a, b \in \mathbb{Z} \right\}, \quad d \equiv 2 \text{ or } 3 \pmod 4,$$
$$\mathbb{Z}_{\mathbb{K}} = \left\{ \frac{a + b\sqrt{d}}{2} \;\middle|\; a, b \in \mathbb{Z}, \quad a \equiv b \pmod 2 \right\}, \quad d \equiv 1 \pmod 4.$$

Proof. See [10, p. 63] or [14, p. 67]. □

Theorem 4. *Let* $\mathbb{K}$ *be an algebraic number field of degree* m. *Then there exist exactly* m *different monomorphisms* $\sigma_i : \mathbb{K} \to \mathbb{C}$, $i = 1, \ldots, m$.

Proof. See [14, p. 41]. □

Definition 9. Let $\mathbb{K}$ be an algebraic number field of degree m and σ_i, $i = 1, \ldots, m$, its field monomorphisms. The *conjugates* of an element $\alpha \in \mathbb{K}$ relative to $\mathbb{K}$ are the m complex numbers $\sigma_i(\alpha) =: \alpha^{(i)}$, $i = 1, \ldots m$.

Definition 10. The *field norm* of an algebraic number $\alpha \in \mathbb{K}$ is $N(\alpha) := \prod_{i=1}^{m} \alpha^{(i)}$, where $m = [\mathbb{K} : \mathbb{Q}]$.

Chapter 2

Siegel's lemma

Consider the system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1N}x_N = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2N}x_N = 0, \\ \vdots \\ a_{M1}x_1 + a_{M2}x_2 + \ldots + a_{MN}x_N = 0, \end{cases}$$

where the coefficients a_{mn} are (rational or more general algebraic) integers and $M < N$. Siegel's lemma gives an estimate for the size of a non-zero integer solution $\bar{x} = (x_1, x_2, \ldots, x_N)^T$. It is a consequence of the so-called *pigeonhole principle* or *Dirichlet's principle*, which states:

> If a set of $n + 1$ elements is subdivided into n subsets, then at least one of these subsets contains at least two elements.

In the following the cases of rationals and imaginary quadratic fields are studied, whereas Chapter 3 deals with the case of a general algebraic number field.

2.1 Siegel's lemma in the field of rational numbers

Theorem 5. *Let*

$$L_m(\bar{x}) = \sum_{n=1}^{N} a_{mn}x_n, \quad m = 1, \ldots, M,$$

be M non-trivial linear forms in N variables x_n with coefficients $a_{mn} \in \mathbb{Z}$. Define

$$A_m := \sum_{n=1}^{N} |a_{mn}| \in \mathbb{Z}^+, \quad m = 1, \ldots, M.$$

Suppose $M < N$; then the system of equations

$$L_m(\bar{x}) = 0, \quad m = 1, \ldots, M, \tag{2.1}$$

has a non-zero integer solution $\bar{z} = (z_1, \ldots, z_N)^T \in \mathbb{Z}^N \setminus \{\bar{0}\}$ with

$$1 \leq \max_{1 \leq n \leq N} |z_n| \leq \left\lfloor (A_1 \cdots A_M)^{\frac{1}{N-M}} \right\rfloor. \tag{2.2}$$

Proof. [9, p. 11] Since $N > M$, the homomorphism

$$\bar{L} = (L_1, \ldots, L_M)^T : \mathbb{Z}^N \to \mathbb{Z}^M.$$

between the additive groups $(\mathbb{Z}^N, +)$ and $(\mathbb{Z}^M, +)$ is not injective. Thus $\ker \bar{L} \neq \{\bar{0}\}$, and there exists a non-zero solution $\bar{z} = (z_1, \ldots, z_N) \in \mathbb{Z}^N \setminus \{\bar{0}\}$ to the system of equations (2.1). Denote $Z := \left\lfloor (A_1 \cdots A_M)^{\frac{1}{N-M}} \right\rfloor$. Then $(A_1 \cdots A_M)^{\frac{1}{N-M}} - Z < 1$, so $A_1 \cdots A_M < (Z+1)^{N-M}$, where $Z, A_1, \ldots, A_M$ are positive integers. From this it follows that

$$\begin{aligned} &(A_1 Z + 1) \cdots (A_M Z + 1) \\ &\leq A_1(Z+1) \cdots A_M(Z+1) \\ &= A_1 \cdots A_M (Z+1)^M \\ &< (Z+1)^N. \end{aligned} \tag{2.3}$$

Now define a set

$$\square_1 := \left\{ \bar{x} \in \mathbb{Z}^N \;\middle|\; 0 \leq x_n \leq Z \right\}.$$

The number of integer points in this box is $\#\square_1 = (Z+1)^N$, since each component of the vector $\bar{x}$ has $Z+1$ possibilities. The linear mappings

$$L_m(\bar{x}) = \sum_{n=1}^{N} a_{mn} x_n, \quad m = 1, \ldots, M,$$

are bounded in the box $\square_1$ by

$$\sum_{a_{mn}<0} a_{mn} x_n \leq L_m(\bar{x}) \leq \sum_{a_{mn}>0} a_{mn} x_n,$$

as the components x_n are non-negative there. Let $-b_m := \sum_{a_{mn}<0} a_{mn}$ be the sum of the negative coefficients and $c_m := \sum_{a_{mn}>0} a_{mn}$ that of the positive ones. Note that $b_m + c_m = A_m$. Further,

$$-b_m Z \leq L_m(\bar{x}) \leq c_m Z \tag{2.4}$$

in the box $\square_1$.

Next define a second set

$$\square_2 := \left\{ \bar{l} \in \mathbb{Z}^M \;\middle|\; -b_m Z \leq l_m \leq c_m Z \right\},$$

where each component of the vector $\bar{l}$ has $\#\{l_m\} = (b_m + c_m)Z + 1 = A_m Z + 1$ possibilities, and therefore the number of integer points in this second box is $\#\square_2 = (A_1 Z + 1) \cdots (A_M Z + 1)$. By the estimate (2.4) we have $\bar{L}(\square_1) \subseteq \square_2$, where

$$\#\square_2 = (A_1 Z + 1) \cdots (A_M Z + 1) < (Z+1)^N = \#\square_1,$$

using the estimate made in (2.3). Hence the mapping $\bar{L} : \square_1 \to \square_2$ is not injective on $\square_1$. Therefore there exist two different vectors $\bar{x}_1, \bar{x}_2 \in \square_1$ such that $\bar{L}(\bar{x}_1) = \bar{L}(\bar{x_2})$, which further gives $\bar{L}(\bar{x}_1 - \bar{x}_2) = \bar{0}$, where $\bar{x}_1 - \bar{x}_2 \in$

$\pm\square_1 \setminus \{\bar{0}\}$. By denoting $\bar{z} = (z_1, \ldots, z_N)^T := \bar{x}_1 - \bar{x}_2$ we get a non-zero solution to the system of equations (2.1) satisfying the estimate

$$|z_n| \leq Z = \left\lfloor (A_1 \cdots A_M)^{\frac{1}{N-M}} \right\rfloor, \quad n = 1, \ldots, N.$$

□

Remark 1. Because

$$A_m \leq N \max_{1 \leq n \leq N} |a_{mn}|,$$

then

$$A_1 \cdots A_M \leq (N \max_{1 \leq m,n \leq N} |a_{mn}|)^M$$

and the upper bound given in (2.2) can further be estimated to give a solution $\bar{z}$ with

$$1 \leq \max_{1 \leq n \leq N} |z_n| \leq \left(N \max_{1 \leq m,n \leq N} |a_{mn}| \right)^{\frac{M}{N-M}}.$$

2.2 Siegel's lemma in imaginary quadratic fields

Theorem 6. *Let $\mathbb{I}$ denote the field $\mathbb{Q}$ of rational numbers or an imaginary quadratic field $\mathbb{Q}(\sqrt{-D})$, where $D \in \mathbb{Z}^+$, and $\mathbb{Z}_{\mathbb{I}}$ its ring of integers. Let*

$$L_m(\bar{z}) = \sum_{n=1}^{N} a_{mn} z_n, \quad m = 1, \ldots, M, \tag{2.5}$$

be M non-trivial linear forms in N variables z_n with coefficients $a_{mn} \in \mathbb{Z}_{\mathbb{I}}$. Define

$$A_m := \sum_{n=1}^{N} |a_{mn}| \in \mathbb{Z}^+, \quad m = 1, \ldots, M. \tag{2.6}$$

Suppose $M < N$; then there exist positive constants $s_{\mathbb{I}}$ and $t_{\mathbb{I}}$ such that the system of equations

$$L_m(\bar{z}) = 0, \quad m = 1, \ldots, M, \tag{2.7}$$

has a non-zero solution $\bar{z} = (z_1, \ldots, z_N)^T \in \mathbb{Z}_{\mathbb{I}}^N \setminus \{\bar{0}\}$ *with*

$$1 \leq \max_{1 \leq n \leq N} |z_n| \leq \max \left\{ 4\sqrt{2}\sqrt{D},\ s_{\mathbb{I}} t_{\mathbb{I}}^{\frac{M}{N-M}} (A_1 \cdots A_M)^{\frac{1}{N-M}} \right\},$$

where $s_{\mathbb{Q}} = t_{\mathbb{Q}} = 1$,

$$s_{\mathbb{Q}(\sqrt{-D})} = \begin{cases} \frac{2\sqrt{2}}{\sqrt{\pi}} D^{\frac{1}{4}}, & D \equiv 1 \text{ or } 2 \pmod 4 \\ \frac{2}{\sqrt{\pi}} D^{\frac{1}{4}}, & D \equiv 3 \pmod 4 \end{cases}$$

and

$$t_{\mathbb{Q}(\sqrt{-D})} = \frac{5}{2\sqrt{2}}.$$

Proof. [4, p. 3] [1] Let $c\sqrt{D} \leq B \in \mathbb{R}^+$, where c is some positive real constant. Every integer $z = x + y\sqrt{-D} \in \mathbb{Z}_{\mathbb{I}}$ for which $|z| \leq B$ lies inside the (filled) ellipse

$$E := \left\{ (x, y)^T \in \mathbb{R}^2 \;\middle|\; x^2 + Dy^2 \leq B^2 \right\}.$$

It has B and $\frac{B}{\sqrt{D}}$ as its semi-major and semi-minor axes, respectively, and area $V(E) = \frac{\pi B^2}{\sqrt{D}}$.

Assume first $D \equiv 1$ or $2 \pmod 4$. According to Theorem 3, now

$$\mathbb{Z}_{\mathbb{I}} = \left\{ a + b\sqrt{-D} \;\middle|\; a, b \in \mathbb{Z} \right\}.$$

Allocate each integer point $(x_0, y_0)^T \in \mathbb{Z}^2$ to the square

$$\square_{(x_0, y_0)} := \left\{ (x, y)^T \in \mathbb{R}^2 \;\middle|\; \max\{|x - x_0|, |y - y_0|\} \leq \frac{1}{2} \right\},$$

[1] What is presented here is a slightly modified version of the original proof, developed together with Tapani Matala-aho. It seems that in the case $D \equiv 3 \pmod 4$ the value of the constant c that was claimed in the article cannot be reached, at least not without doing some further calculations.

which has area 1. Then the number of integer points inside the ellipse E can be estimated using the shape

$$F_1 := \left\{ (s,t)^T \in \mathbb{R}^2 \;\middle|\; s = x - \frac{1}{2},\, t = y - \frac{1}{2},\, x^2 + Dy^2 \leq B^2, \right.$$
$$\left. x \in \left[\frac{1}{2}, \sqrt{B^2 - \frac{D}{4}} \right],\, y \in \left[\frac{1}{2}, \frac{1}{\sqrt{D}} \sqrt{B^2 - \frac{1}{4}} \right] \right\} \subseteq E.$$

(See the picture on page 19. The dashed curve represents the boundary of the set F_1, showing where the integer point (x_0, y_0) can be located in order to have the square $\square_{(x_0,y_0)}$ stay inside E.) For any $(x_0, y_0)^T \in F_1$ we have $\square_{(x_0,y_0)} \subseteq E$, so the number of integer points inside E is at least as big as the area of F_1 multiplied by four. [2] One can estimate the area $4V(F_1)$ by noticing that the ellipse E consists of four pieces identical to F_1 plus some additional parts, which in turn can be estimated by rectangles of area $2B$ and $\frac{2B}{\sqrt{D}}$. [3]

[2] Due to symmetry, it is enough to consider only the first quarter of the ellipse E.

[3] A lower bound for the area $4V(F_1)$ can also be found by fitting a smaller ellipse E_1 inside E so that the first quarter of E_1 is inside F_1. This is a rather lengthy approach where the method of Lagrange multipliers is needed, so to save some space, we will settle for the block approximation.

(See the picture on page 20.) Subtracting these from $V(E)$ results in

$$\begin{aligned} 4V(F_1) &\geq \frac{\pi B^2}{\sqrt{D}} - 2B - \frac{2B}{\sqrt{D}} + 1 \\ &\geq \frac{\pi B^2}{\sqrt{D}} \left(1 - \frac{2\sqrt{D}}{\pi B} - \frac{2}{\pi B}\right) \\ &\geq \frac{\pi B^2}{\sqrt{D}} \left(1 - \frac{2\sqrt{D}}{\pi c\sqrt{D}} - \frac{2}{\pi c\sqrt{D}}\right) \\ &\geq \frac{\pi B^2}{\sqrt{D}} \left(1 - \frac{2}{\pi c} - \frac{2}{\pi c}\right) \\ &= \frac{\pi B^2}{\sqrt{D}} \left(1 - \frac{4}{\pi c}\right) \\ &\geq \frac{\pi B^2}{\sqrt{D}} \left(1 - \frac{\sqrt{2}}{c}\right). \end{aligned} \tag{2.8}$$

When $|z_n| \leq B,\ n = 1, \ldots, N$, for the linear forms (2.5) it holds

$$|L_m(\bar{z})| = \left| \sum_{n=1}^{N} a_{mn} z_n \right| \leq \sum_{n=1}^{N} |a_{mn} z_n| \leq \sum_{n=1}^{N} |a_{mn}|\, B = A_m B, \quad m = 1, \ldots, M.$$

So, if each component of the vector $\bar{z} = (z_1, \ldots, z_N)^T \in \mathbb{Z}_{\mathbb{I}}^N$ lies inside the ellipse E, then every component of its image in the mapping $\bar{L} := (L_1, \ldots, L_M) : \mathbb{Z}_{\mathbb{I}}^N \to \mathbb{Z}_{\mathbb{I}}^M$ has to lie inside the ellipse

$$E' := \left\{ (x, y)^T \in \mathbb{R}^2 \;\middle|\; x^2 + Dy^2 \leq A_m^2 B^2 \right\}.$$

With arguments similar to those justifying the estimate (2.8), one can define a shape F_2 (marked with the dashed curve in the figure on page 21) which consists of the set

$$\left\{ (s, t)^T \in \mathbb{R}^2 \;\middle|\; s = x + \frac{1}{2},\ t = y + \frac{1}{2},\ x^2 + Dy^2 \leq A_m^2 B^2, \right.$$
$$\left. x \in [0, A_m B]\,,\ y \in \left[0, \frac{A_m B}{\sqrt{D}}\right] \right\}$$

and the rectangles $\left[0,\frac{1}{2}\right]\times\left[0,\frac{A_mB}{\sqrt{D}}+\frac{1}{2}\right]$ and $\left[0,A_mB+\frac{1}{2}\right]\times\left[0,\frac{1}{2}\right]$. Then we see that the ellipse E' contains at most

$$\begin{aligned}
4V(F_2) &= \frac{\pi A_m^2B^2}{\sqrt{D}}+2A_mB+\frac{2A_mB}{\sqrt{D}}+1\\
&= \frac{\pi A_m^2B^2}{\sqrt{D}}\left(1+\frac{2\sqrt{D}}{\pi A_mB}+\frac{2}{\pi A_mB}+\frac{\sqrt{D}}{\pi A_m^2B^2}\right)\\
&\le \frac{\pi A_m^2B^2}{\sqrt{D}}\left(1+\frac{2}{\pi c}\left(1+\frac{1}{\sqrt{D}}\right)+\frac{1}{\pi c^2\sqrt{D}}\right)\\
&\le \frac{\pi A_m^2B^2}{\sqrt{D}}\left(1+\frac{2\sqrt{2}}{\pi c}+\frac{1}{\pi c^2}\right)\\
&\le \frac{\pi A_m^2B^2}{\sqrt{D}}\left(1+\frac{\sqrt{2}}{c}+\frac{1}{2c^2}\right)\\
&= \frac{\pi A_m^2B^2}{\sqrt{D}}\left(1+\frac{1}{c\sqrt{2}}\right)^2
\end{aligned}\tag{2.9}$$

integer points. (Here the facts that $A_m, D\ge 1$ and $B\ge c\sqrt{D}$ have been taken into account.) Therefore the number of values of the linear forms (2.5) with $|z_n|\le B,\ n=1,\dots,N$, is at most

$$\frac{\pi^MB^{2M}}{\left(\sqrt{D}\right)^M}\left(1+\frac{1}{c\sqrt{2}}\right)^{2M}\prod_{m=1}^{M}A_m^2.$$

Hence, the number of vectors $\bar{z}\in\mathbb{Z}_{\mathbb{I}}^N$ with $|z_n|\le B,\ n=1,\dots,N$, is greater than the number of values of the linear forms when

$$\frac{\pi^NB^{2N}}{\left(\sqrt{D}\right)^N}\left(1-\frac{\sqrt{2}}{c}\right)^N \ge \frac{\pi^MB^{2M}}{\left(\sqrt{D}\right)^M}\left(1+\frac{1}{c\sqrt{2}}\right)^{2M}\prod_{m=1}^{M}A_m^2. \tag{2.10}$$

It follows that

$$B\ge \frac{D^{\frac{1}{4}}}{\sqrt{\pi}}\left(1-\frac{\sqrt{2}}{c}\right)^{-\frac{N}{2N-2M}}\left(1+\frac{1}{c\sqrt{2}}\right)^{\frac{M}{N-M}}\left(\prod_{m=1}^{M}A_m\right)^{\frac{1}{N-M}}. \tag{2.11}$$

Now choose $c = 2\sqrt{2}$, which gives

$$\begin{aligned} B &\geq \frac{D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{1}{2}\right)^{-\frac{N}{2N-2M}} \left(\frac{5}{4}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}} \\ &= \frac{D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{1}{\sqrt{2}}\right)^{-\frac{N}{N-M}} \left(\frac{1}{\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}} \qquad (2.12) \\ &= \frac{\sqrt{2}D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}. \end{aligned}$$

At the beginning the assumption $B \geq c\sqrt{D} = 2\sqrt{2D}$ was made, so let

$$B' := \max\left\{2\sqrt{2D},\ \frac{\sqrt{2}D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}\right\}.$$

By the pigeonhole principle, there exist two different vectors $\bar{z}^{(1)}, \bar{z}^{(2)} \in \mathbb{Z}_{\mathbb{I}}^N$ with $\left|z_n^{(i)}\right| \leq B',\ n = 1, \ldots, N,\ i = 1, 2$, and $\bar{L}\big(\bar{z}^{(1)}\big) = \bar{L}\big(\bar{z}^{(2)}\big)$. Thus we have a solution $\bar{z} = (z_1, \ldots, z_N)^T := \bar{z}^{(1)} - \bar{z}^{(2)} \in \mathbb{Z}_{\mathbb{I}}^N \setminus \{\bar{0}\}$ to the system of equations (2.7) such that

$$\begin{aligned} |z_n| &\leq \left|z_n^{(1)}\right| + \left|z_n^{(2)}\right| \\ &\leq 2B' \\ &= \max\left\{4\sqrt{2D},\ \frac{2\sqrt{2}D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}\right\} \end{aligned}$$

for all $n = 1, \ldots, N$. Here the coefficients $s_{\mathbb{I}}$ and $t_{\mathbb{I}}$ can be identified.

Secondly, let $D \equiv 3 \pmod 4$, in which case Theorem 3 gives

$$\mathbb{Z}_{\mathbb{I}} = \left\{\frac{a + b\sqrt{-D}}{2} \;\middle|\; a, b \in \mathbb{Z},\quad a \equiv b \pmod 2\right\}.$$

Again the number of points corresponding to integers inside the ellipse E is studied, but now by identifying each point $\left(\frac{x_0}{2}, \frac{y_0}{2}\right)^T \in (\frac{1}{2}\mathbb{Z})^2$, $x_0 \equiv y_0 \pmod 2$, with the square

$$\diamondsuit_{\left(\frac{x_0}{2}, \frac{y_0}{2}\right)} := \left\{(x, y)^T \in \mathbb{R}^2 \;\middle|\; \left|x - \frac{x_0}{2}\right| + \left|y - \frac{y_0}{2}\right| \leq \frac{1}{2}\right\},$$

whose area is $\frac{1}{2}$. (This is because the integer points are in this case somewhat tighter packed on the plane than they were when $D \equiv 1, 2 \pmod 4$.) Now $\diamondsuit_{\left(\frac{x_0}{2}, \frac{y_0}{2}\right)} \subseteq \square_{\left(\frac{x_0}{2}, \frac{y_0}{2}\right)}$, so we can use the previous estimates (2.8) and (2.9) — they just need to be multiplied by two since each point now corresponds to a square of area $\frac{1}{2}$. Hence the ellipse E contains at least

$$\frac{2\pi B^2}{\sqrt{D}} \left(1 - \frac{\sqrt{2}}{c}\right)$$

and the ellipse E' at most

$$\frac{2\pi A_m^2 B^2}{\sqrt{D}} \left(1 + \frac{1}{c\sqrt{2}}\right)^2$$

integer points. As in (2.10), demanding the number of N-tuples $\bar{z} \in \mathbb{Z}_{\mathbb{I}}^N$ to be greater than the number of possible values of the linear forms (2.5) leads to the inequality

$$\frac{2^N \pi^N B^{2N}}{\left(\sqrt{D}\right)^N} \left(1 - \frac{\sqrt{2}}{c}\right)^N \geq \frac{2^M \pi^M B^{2M}}{\left(\sqrt{D}\right)^M} \left(1 + \frac{1}{c\sqrt{2}}\right)^{2M} \prod_{m=1}^{M} A_m^2.$$

We get (2.11) with 2π instead of π:

$$B \geq \frac{D^{\frac{1}{4}}}{\sqrt{2\pi}} \left(1 - \frac{\sqrt{2}}{c}\right)^{-\frac{N}{2N-2M}} \left(1 + \frac{1}{c\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}.$$

Again choose $c = 2\sqrt{2}$, and comparison to (2.12) gives

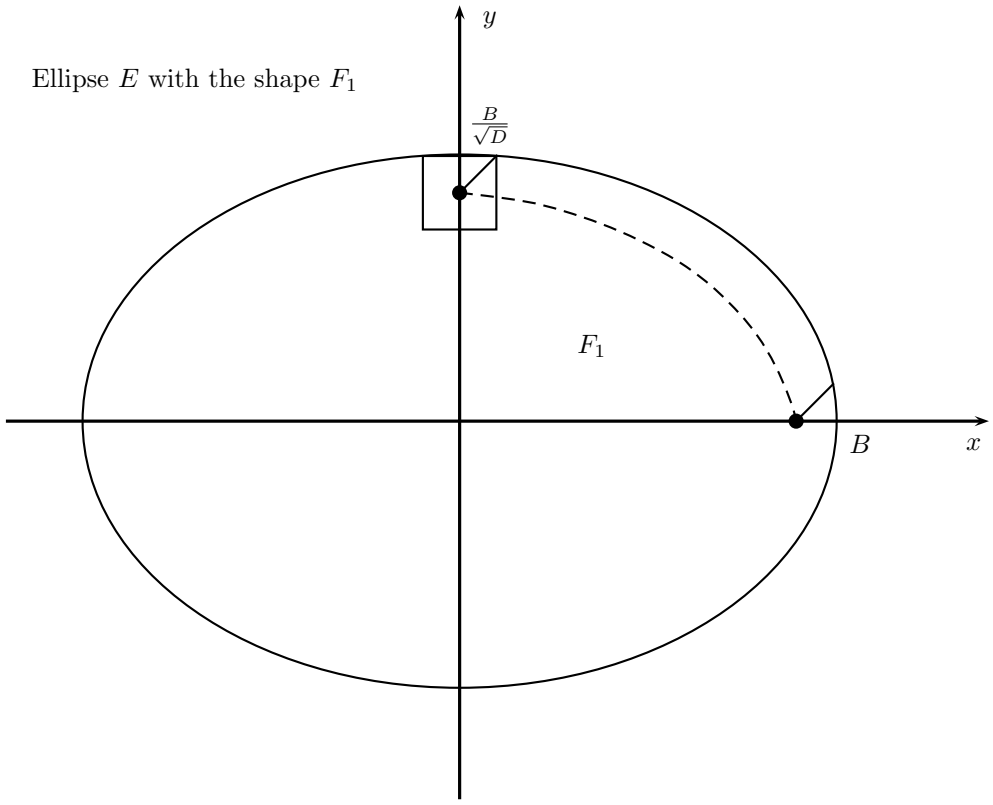
$$B \geq \frac{D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}.$$

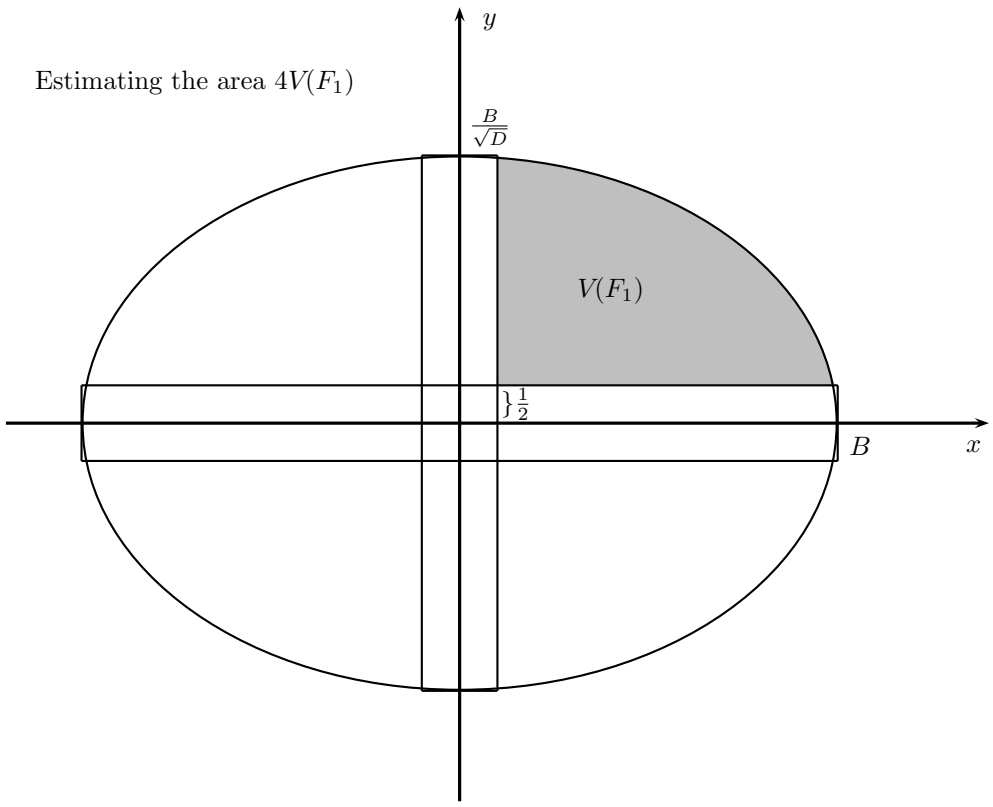
Analogously to the case $D \equiv 1, 2 \pmod 4$, there exists a solution $\bar{z} \in \mathbb{Z}_{\mathbb{I}}^N \backslash \{\bar{0}\}$ to the system of equations (2.7) with

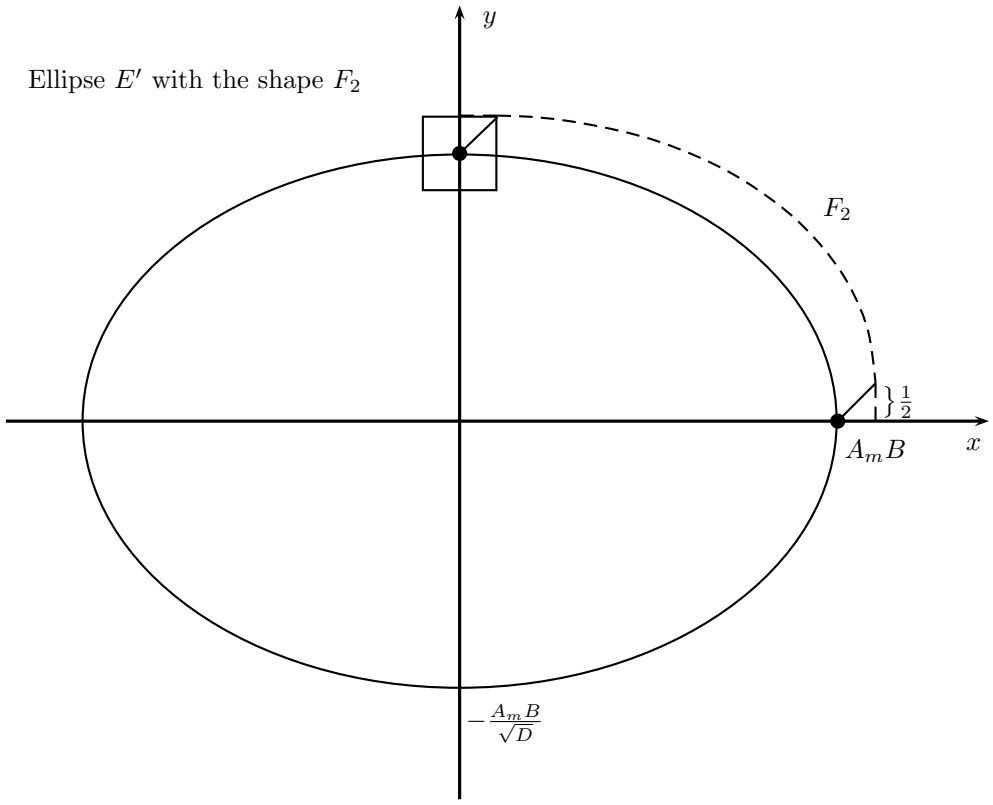
$$\|\bar{z}\|_\infty \leq \max\left\{4\sqrt{2D},\ \frac{2D^{\frac{1}{4}}}{\sqrt{\pi}} \left(\frac{5}{2\sqrt{2}}\right)^{\frac{M}{N-M}} \left(\prod_{m=1}^{M} A_m\right)^{\frac{1}{N-M}}\right\}.$$

□

Remark 2. The necessary assumption $B \geq 2\sqrt{2D}$ only matters when the values of the row sums A_m are small.


Ellipse E with the shape F_1
y
$\frac{B}{\sqrt{D}}$
F_1
B
x


Estimating the area $4V(F_1)$
y
$\frac{B}{\sqrt{D}}$
$V(F_1)$
$\frac{1}{2}$
B
x


Ellipse E' with the shape F_2
y
F_2
$\frac{1}{2}$
$A_m B$
x
$-\frac{A_m B}{\sqrt{D}}$

Chapter 3

Siegel's lemma in an arbitrary algebraic number field

Let $\mathbb{K}$ be an algebraic number field of degree K and $\mathbb{Z}_{\mathbb{K}}$ its ring of integers with the fixed basis $\{\omega_1, \ldots, \omega_K\} \subseteq \mathbb{Z}_{\mathbb{K}}$.

Definition 11. Let $\alpha \in \mathbb{K}$ and denote with $\alpha^{(1)} := \alpha, \alpha^{(2)}, \ldots, \alpha^{(K)}$ its conjugates relative to $\mathbb{K}$. The *house function* $\left\boxed{}\right\rvert : \mathbb{K} \to \mathbb{R}$ is defined as follows:

$$\left\boxed{\alpha}\right\rvert := \max\left\{|\alpha|, \left|\alpha^{(2)}\right|, \ldots, \left|\alpha^{(K)}\right|\right\}.$$

Lemma 1. *Let $\alpha, \beta \in \mathbb{K}$. Then*

$$\left\boxed{\alpha \pm \beta}\right\rvert \leq \left\boxed{\alpha}\right\rvert + \left\boxed{\beta}\right\rvert$$

and

$$\left\boxed{\alpha\beta}\right\rvert \leq \left\boxed{\alpha}\right\rvert \left\boxed{\beta}\right\rvert.$$

Proof. Since the field monomorphisms $\sigma_i : \mathbb{K} \to \mathbb{C},\ i = 1, \ldots, K$, preserve addition and multiplication, for the conjugates it holds

$$(\alpha \pm \beta)^{(i)} = \sigma_i(\alpha \pm \beta) = \sigma_i(\alpha) \pm \sigma_i(\beta) = \alpha^{(i)} \pm \beta^{(i)}$$

and

$$(\alpha\beta)^{(i)} = \sigma_i(\alpha\beta) = \sigma_i(\alpha)\sigma_i(\beta) = \alpha^{(i)}\beta^{(i)},$$

where $i = 1, \ldots, K$. Directly

$$\begin{aligned}
\left|\alpha \pm \beta\right| &= \max_{1\le i\le K}\left\{\left|(\alpha \pm \beta)^{(i)}\right|\right\} \\
&= \max_{1\le i\le K}\left\{\left|\alpha^{(i)} \pm \beta^{(i)}\right|\right\} \\
&\le \max_{1\le i\le K}\left\{\left|\alpha^{(i)}\right| + \left|\beta^{(i)}\right|\right\} \\
&\le \max_{1\le i\le K}\left\{\left|\alpha^{(i)}\right|\right\} + \max_{1\le i\le K}\left\{\left|\beta^{(i)}\right|\right\} \\
&= \left|\alpha\right| + \left|\beta\right|,
\end{aligned}$$

and similarly

$$\begin{aligned}
\left|\alpha\beta\right| &= \max_{1\le i\le K}\left\{\left|(\alpha\beta)^{(i)}\right|\right\} \\
&= \max_{1\le i\le K}\left\{\left|\alpha^{(i)}\right|\left|\beta^{(i)}\right|\right\} \\
&\le \max_{1\le i\le K}\left\{\left|\alpha^{(i)}\right|\right\} \max_{1\le i\le K}\left\{\left|\beta^{(i)}\right|\right\} \\
&= \left|\alpha\right|\left|\beta\right|.
\end{aligned}$$

□

Lemma 2. *Let $\gamma \in \mathbb{Z}_{\mathbb{K}}$ and write $\gamma = g_1\omega_1 + \ldots + g_K\omega_K$, where $g_1, \ldots, g_K \in \mathbb{Z}$ are unique. There exist two positive constants c_1 and c_2 depending solely on the ring $\mathbb{Z}_{\mathbb{K}}$, such that*

$$\left|\gamma\right| \le c_1 \max\{|g_1|, \ldots, |g_K|\} \tag{3.1}$$

and

$$\max\{|g_1|, \ldots, |g_K|\} \le c_2\left|\gamma\right|. \tag{3.2}$$

Proof. Mahler gives the outlines for this proof in [8, p. 91].

First, using Lemma 1 and the fact that $g_1, \ldots, g_K \in \mathbb{Z}$,

$$\begin{aligned}
\left|\gamma\right| &= \left|g_1\omega_1 + \ldots + g_K\omega_K\right| \\
&\leq \left|g_1\right|\left|\omega_1\right| + \ldots + \left|g_K\right|\left|\omega_K\right| \\
&= |g_1|\left|\omega_1\right| + \ldots + |g_K|\left|\omega_K\right| \\
&\leq \left(\left|\omega_1\right| + \ldots + \left|\omega_K\right|\right) \max\{|g_1|, \ldots, |g_K|\}.
\end{aligned}$$

Choose $c_1 := \left|\omega_1\right| + \ldots + \left|\omega_K\right| > 0$, which does not depend on γ.

From the properties of the field monomorphisms it follows that

$$\gamma^{(i)} = g_1\omega_1^{(i)} + \ldots + g_K\omega_K^{(i)}, \quad i = 1, \ldots, K. \tag{3.3}$$

Since $\{\omega_1, \ldots, \omega_K\}$ is a basis for the ring $\mathbb{Z}_\mathbb{K}$ (over $\mathbb{Z}$), it is also a basis for the field $\mathbb{K}$ (over $\mathbb{Q}$) due to the fact that for any $\alpha \in \mathbb{K}$ there exists $d \in \mathbb{Z}^+$ such that $d\alpha \in \mathbb{Z}_\mathbb{K}$. Hence the discriminant

$$\begin{vmatrix}
\omega_1^{(1)} & \cdots & \omega_K^{(1)} \\
\vdots & & \vdots \\
\omega_1^{(K)} & \cdots & \omega_K^{(K)}
\end{vmatrix}^2$$

of this basis is non-zero (see [14, p. 44]), so the determinant

$$\begin{vmatrix}
\omega_1^{(1)} & \cdots & \omega_K^{(1)} \\
\vdots & & \vdots \\
\omega_1^{(K)} & \cdots & \omega_K^{(K)}
\end{vmatrix}$$

does not vanish. Therefore the equations (3.3) can be solved in the form

$$g_j = \sum_{i=1}^{K} \Omega_{ij}\gamma^{(i)}, \quad j = 1, \ldots, K,$$

where the coefficients $\Omega_{ij} \in \mathbb{K}$ depend only on the basis $\{\omega_1, \ldots, \omega_K\}$ of $\mathbb{Z}_{\mathbb{K}}$. As a result,

$$\begin{aligned}
\max\{|g_1|, \ldots, |g_K|\} &= \max_{1 \le j \le K} \left\{ \left| \sum_{i=1}^{K} \Omega_{ij} \gamma^{(i)} \right| \right\} \\
&\le \max_{1 \le j \le K} \left\{ \sum_{i=1}^{K} |\Omega_{ij}| \left|\gamma^{(i)}\right| \right\} \\
&\le \max_{1 \le j \le K} \left\{ \max_{1 \le i \le K} \left\{ \left|\gamma^{(i)}\right| \right\} \sum_{i=1}^{K} |\Omega_{ij}| \right\} \\
&= \max_{1 \le j \le K} \left\{ \left\lceil |\gamma| \right\rceil \sum_{i=1}^{K} |\Omega_{ij}| \right\} \\
&= \left\lceil |\gamma| \right\rceil \max_{1 \le j \le K} \left\{ \sum_{i=1}^{K} |\Omega_{ij}| \right\}.
\end{aligned}$$

Choose $c_2 := \max_{1 \le j \le K} \left\{ \sum_{i=1}^{K} |\Omega_{ij}| \right\} > 0$, which too is independent of γ. $\square$

Now we can generalize Siegel's lemma.

Theorem 7. *Let*

$$y_m = \sum_{n=1}^{N} a_{mn} x_n, \quad m = 1, \ldots, M, \tag{3.4}$$

be M non-trivial linear forms in N variables x_n with coefficients $a_{mn} \in \mathbb{Z}_{\mathbb{K}}$. Define

$$A := \max_{m,n} \left\lceil |a_{mn}| \right\rceil > 0, \tag{3.5}$$

and suppose $M < N$. Then the system of equations

$$y_m = 0, \quad m = 1, \ldots, M,$$

has a non-zero integer solution $\bar{z} = (z_1, \ldots, z_N)^T \in \mathbb{Z}_{\mathbb{K}}^N \setminus \{\bar{0}\}$ with

$$\max\left\{ \left\lceil |z_1| \right\rceil, \ldots, \left\lceil |z_N| \right\rceil \right\} \le c(cNKA)^{\frac{M}{N-M}},$$

where c is some positive constant.

Proof. [8, p. 92] Since $a_{mn} \in \mathbb{Z}_{\mathbb{K}}$, the products $a_{mn}\omega_j,\ j = 1, \ldots, K$, are also in $\mathbb{Z}_{\mathbb{K}}$ and thus can be represented as linear combinations of the basis elements $\omega_1, \ldots, \omega_K$:

$$a_{mn}\omega_j = \sum_{k=1}^{K} a_{mnjk}\omega_k, \tag{3.6}$$

where $m = 1, \ldots, M,\ n = 1, \ldots, N,\ j = 1, \ldots, K$, and the coefficients a_{mnjk} are unique rational integers. Similarly, for unknown variables $x_1, \ldots, x_N \in \mathbb{Z}_{\mathbb{K}}$ there exist unique rational integers x_{nj} such that

$$x_n = \sum_{j=1}^{K} x_{nj}\omega_j, \quad n = 1, \ldots, N. \tag{3.7}$$

The set of N variables $x_n \in \mathbb{Z}_{\mathbb{K}}$ has in this way been turned into a set of NK variables $x_{nj} \in \mathbb{Z}$. Combining the formulae (3.4), (3.6) and (3.7) gives

$$\begin{aligned}
y_m &= \sum_{n=1}^{N} a_{mn}x_n \\
&= \sum_{n=1}^{N} a_{mn} \sum_{j=1}^{K} x_{nj}\omega_j \\
&= \sum_{n=1}^{N} \sum_{j=1}^{K} a_{mn}\omega_j x_{nj} \\
&= \sum_{n=1}^{N} \sum_{j=1}^{K} \sum_{k=1}^{K} a_{mnjk}\omega_k x_{nj} \\
&= \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{j=1}^{K} a_{mnjk} x_{nj}\omega_k \\
&= \sum_{k=1}^{K} y_{mk}\omega_k,
\end{aligned} \tag{3.8}$$

where we have defined

$$y_{mk} := \sum_{n=1}^{N} \sum_{j=1}^{K} a_{mnjk}x_{nj} \in \mathbb{Z}, \quad m = 1, \ldots, M,\ k = 1, \ldots, K. \tag{3.9}$$

Let now c_1 and c_2 be the constants given by Lemma 2, and define

$$c_3 := \left\lfloor c_2 \max_{1\le j\le K} \left\lceil |\omega_j| \right\rceil \right\rfloor + 1. \tag{3.10}$$

Then c_3, like c_1 and c_2, only depends on the ring $\mathbb{Z}_{\mathbb{K}}$. Next pick $H \in \mathbb{Z}^+$ in such a way that

$$0 \le (NKc_3A)^{\frac{M}{N-M}} - 1 < 2H \le (NKc_3A)^{\frac{M}{N-M}} + 1. \tag{3.11}$$

Then H satisfies also the inequality

$$\begin{aligned} (2H+1)^{NK} &= (2H+1)^{(N-M)K}(2H+1)^{MK} \\ &> (NKc_3A)^{MK}(2H+1)^{MK} \\ &\ge (2NKc_3A+1)^{MK}, \end{aligned} \tag{3.12}$$

the last row following from the fact that $H, N, K, c_3, A \ge 1$.

Consider the linear forms (3.9). If the possible values of the variables $x_{nj} \in \mathbb{Z}$ are restricted by the condition $|x_{nj}| \le H$, the NK-tuple (x_{nj}) has $(2H+1)^{NK}$ distinct possibilities. By the representation (3.6) and the second inequality (3.2) of Lemma 2 it follows in turn that

$$\max_{m,n,j,k} |a_{mnjk}| \le c_2 \max_{m,n,j} \left\lceil |a_{mn}\omega_j| \right\rceil \le c_2 \max_{m,n} \left\lceil |a_{mn}| \right\rceil \max_j \left\lceil |\omega_j| \right\rceil \le c_3 A,$$

recalling the definitions (3.5) and (3.10) of the constants A and c_3. Hence there is a limitation for the values of the linear forms (3.9) too:

$$\begin{aligned} \max_{m,k} |y_{mk}| &= \max_{m,k} \left| \sum_{n=1}^{N} \sum_{j=1}^{K} a_{mnjk} x_{nj} \right| \\ &\le \max_{m,k} \sum_{n=1}^{N} \sum_{j=1}^{K} |a_{mnjk}||x_{nj}| \\ &\le NK \max_{m,n,j,k} |a_{mnjk}| \max_{n,j} |x_{nj}| \\ &\le NKc_3AH. \end{aligned}$$

The MK-tuple (y_{mk}) has thus only $(2NKc_3AH+1)^{MK}$ different possibilites, which by the estimate (3.12) is less than the number of possibilities the vector (x_{nj}) has. Therefore two distinct vectors (x'_{nj}) and (x''_{nj}) must be mapped to the same vector (y_{mk}) in the linear forms (3.9). By putting $z_{nj} := x'_{nj} - x''_{nj} \in \mathbb{Z}$ we have a non-zero NK-tuple (z_{nj}) with

$$\max_{n,j} |z_{nj}| \le \max_{n,j} |x'_{nj}| + \max_{n,j} |x''_{nj}| \le H + H \overset{(3.11)}{\le} (NKc_3A)^{\frac{M}{N-M}} + 1$$

satisfying $y_{mk} = 0$ for all $m = 1, \dots, M,\ k = 1, \dots, K$, due to linearity. It follows from (3.8) that the corresponding M algebraic integers y_m also vanish.

Finally, let $c := \max\{2c_1, c_3\}$. Using the representation (3.7) and the first inequality (3.1) of Lemma 2, we have a solution $\bar{z} := (z_1, \dots, z_N)^T \in \mathbb{Z}_{\mathbb{K}}^N \setminus \{\bar{0}\}$ to the system of equations (3.4) with

$$\begin{aligned}
\max_{1 \le n \le N} \boxed{z_n} &\le c_1 \max_{n,j} |z_{nj}| \\
&\le c_1 \left((NKc_3A)^{\frac{M}{N-M}} + 1 \right) \\
&\le 2c_1 (NKc_3A)^{\frac{M}{N-M}} \\
&\le c(cNKA)^{\frac{M}{N-M}}.
\end{aligned}$$

□

Chapter 4

Minkowski's convex body theorems

In this chapter two important geometric theorems are presented, first of which has consequences that are explored in chapter 5. Some amount of background theory is needed before that, however; it is mostly based on Sections 2 and 3 of the lecture notes [9] and Chapter 2 of the lecture notes [5].

For this chapter, let $n \in \mathbb{Z}^+$.

4.1 Convex bodies

Definition 12. A non-empty subset $\mathcal{C} \subseteq \mathbb{R}^n$ is *convex*, if for any pair of points $\bar{a}, \bar{b} \in \mathcal{C}$ it holds $\left\{ s\bar{a} + (1-s)\bar{b} \;\middle|\; 0 \leq s \leq 1 \right\} \subseteq \mathcal{C}$. In other words, a line segment connecting any two points of a convex set $\mathcal{C}$ has to be included in $\mathcal{C}$.

A bounded convex subset of $\mathbb{R}^n$ is called a *convex body*.

Definition 13. A subset $\mathcal{C} \subseteq \mathbb{R}^n$ is *central symmetric* (symmetric with respect to origin), if $\mathcal{C} = -\mathcal{C}$.

The following two properties are simple to prove, but good to keep in mind for later use.

Lemma 3. *If $\lambda \in \mathbb{R}^+$ and $\mathcal{C} \subseteq \mathbb{R}^n$ is a central symmetric convex body, then also the dilation $\lambda\mathcal{C} := \{\lambda\bar{c} \mid \bar{c} \in \mathcal{C}\}$ is a central symmetric convex body.*

Proof. Immediately $\lambda\mathcal{C} = \lambda(-\mathcal{C}) = -\lambda\mathcal{C}$, since the set $\mathcal{C}$ is central symmetric. Next, let $\lambda\bar{a}, \lambda\bar{b} \in \lambda\mathcal{C}$. Now

$$\left\{s\lambda\bar{a} + (1-s)\lambda\bar{b} \;\middle|\; 0 \leq s \leq 1\right\} = \left\{\lambda\left(s\bar{a} + (1-s)\bar{b}\right) \;\middle|\; 0 \leq s \leq 1\right\} \subseteq \lambda\mathcal{C},$$

for $\mathcal{C}$ is convex. Lastly, $\mathcal{C}$ being bounded, the set $\lambda\mathcal{C}$ is certainly bounded as well, for the constant λ is finite. □

Lemma 4. *If $\mathcal{C} \subseteq \mathbb{R}^n$ is a central symmetric convex body and $\bar{L} : \mathbb{R}^n \to \mathbb{R}^n$ a linear mapping, then also the transformed set $\bar{L}(\mathcal{C})$ is a central symmetric convex body.*

Proof. Again we see at once that $\bar{L}(\mathcal{C}) = \bar{L}(-\mathcal{C}) = -\bar{L}(\mathcal{C})$, since the set $\mathcal{C}$ is central symmetric and the mapping $\bar{L}$ linear. Let then $\bar{L}\bar{a}, \bar{L}\bar{b} \in \bar{L}(\mathcal{C})$ for some $\bar{a}, \bar{b} \in \mathcal{C}$. Just like in the previous proof,

$$\left\{s\bar{L}\bar{a} + (1-s)\bar{L}\bar{b} \;\middle|\; 0 \leq s \leq 1\right\} = \left\{\bar{L}\left(s\bar{a} + (1-s)\bar{b}\right) \;\middle|\; 0 \leq s \leq 1\right\} \subseteq \bar{L}(\mathcal{C}),$$

using the linearity of $\bar{L}$ and the convexity of $\mathcal{C}$. Moreover, every linear transformation on a finite-dimensional inner product space is bounded, so the boundedness of $\mathcal{C}$ implies boundedness of its image $\bar{L}(\mathcal{C})$. □

Intuitively it is clear that the smallest central symmetric convex body containing given points $\bar{v}_1, \ldots, \bar{v}_r \in \mathbb{R}^n$ must be the polyhedron that has the points $\pm\bar{v}_1, \ldots, \pm\bar{v}_r \in \mathbb{R}^n$ as its vertices. Lemma 5 confirms this formally.

Lemma 5. *Let $\bar{v}_1, \ldots, \bar{v}_r \in \mathbb{R}^n$. Then the set*

$$\mathcal{V}_r := \left\{ \sum_{i=1}^{r} x_i \bar{v}_i \;\middle|\; x_1, \ldots, x_r \in \mathbb{R},\ \sum_{i=1}^{r} |x_i| \leq 1 \right\} \subseteq \mathbb{R}^n$$

is the smallest central symmetric convex body containing the vectors $\bar{v}_1, \ldots, \bar{v}_r$.

Proof. The fact that $\mathcal{V}_r$ is a central symmetric convex body is trivial to prove: Let $\bar{x} = \sum_{i=1}^{r} x_i \bar{v}_i \in \mathcal{V}_r$. Then $\sum_{i=1}^{r} |-x_i| = \sum_{i=1}^{r} |x_i| \leq 1$, so $-\bar{x} \in \mathcal{V}_r$. Also,

$$\|\bar{x}\| = \left\| \sum_{i=1}^{r} x_i \bar{v}_i \right\| \leq \sum_{i=1}^{r} |x_i| \|\bar{v}_i\| \leq \max_{1 \leq i \leq r} \|\bar{v}_i\| \sum_{i=1}^{r} |x_i| \leq \max_{1 \leq i \leq r} \|\bar{v}_i\|,$$

showing that $\mathcal{V}_r$ is bounded. If $\bar{y} = s\bar{a} + (1-s)\bar{b}$ for some $\bar{a} = \sum_{i=1}^{r} a_i \bar{v}_i$, $\bar{b} = \sum_{i=1}^{r} b_i \bar{v}_i \in \mathcal{V}$, $0 \leq s \leq 1$, then

$$\sum_{i=1}^{r} |y_i| = \sum_{i=1}^{r} |s a_i + (1-s) b_i| \leq s \sum_{i=1}^{r} |a_i| + (1-s) \sum_{i=1}^{r} |b_i| \leq s + (1-s) = 1,$$

meaning that $\bar{y} \in \mathcal{V}_r$ and $\mathcal{V}_r$ is convex.

To prove the second claim, let $\mathcal{C} \subseteq \mathbb{R}^n$ be any central symmetric convex body that contains the points $\bar{v}_1, \ldots, \bar{v}_r$. If $r = 1$, clearly $\mathcal{V}_1 \subseteq \mathcal{C}$ since the set $\mathcal{V}_1$ is precisely the line segment connecting the points $\bar{v}_1$ and $-\bar{v}_1$. Suppose now that the statement is true for $r = k \in \mathbb{Z}^+$, and pick an arbitary point $\bar{x} = \sum_{i=1}^{k+1} x_i \bar{v}_i \in \mathcal{V}_{k+1}$. It suffices to study the case where $|x_i| < 1$ for all $i = 1, \ldots, k+1$ and $x_i \neq 0$ for some $i = 1, \ldots, k+1$, since otherwise immediately $\bar{x} \in \mathcal{C}$. By assumption $\mathcal{V}_k \subseteq \mathcal{C}$, so $\bar{z} := \sum_{i=1}^{k} x_i \bar{v}_i \in \mathcal{C}$. Denote

$$c := \frac{1}{1 - |x_{k+1}|} \leq \frac{1}{\sum_{i=1}^{k} |x_i|}.$$

Then also $c\bar{z} \in \mathcal{C}$, for $\sum_{i=1}^{k} |c x_i| = c \sum_{i=1}^{k} |x_i| \leq 1$ implies $c\bar{z} \in \mathcal{V}_k$. Because the set $\mathcal{C}$ is central symmetric, we have $\operatorname{sign}(x_{k+1})\bar{v}_{k+1} \in \mathcal{C}$, and the convexity of $\mathcal{C}$ further implies

$$|x_{k+1}| \operatorname{sign}(x_{k+1}) \bar{v}_{k+1} + (1 - |x_{k+1}|) c\bar{z} = x_{k+1} \bar{v}_{k+1} + \bar{z} = \sum_{i=1}^{k+1} x_i \bar{v}_i = \bar{x} \in \mathcal{C},$$

indicating $\mathcal{V}_{k+1} \subseteq \mathcal{C}$. As a result of the induction principle, the statement holds true for all $r \in \mathbb{Z}^+$. □

4.2 Lattices

In the following, let $(R, +, \times)$ be a ring formed from a non-empty set R with ring product $\times$ and addition $+$. The ring R has zero and identity elements $0, 1 \in R$, $0 \neq 1$.

Definition 14. Let R be a commutative ring. Then $(M, +, \cdot)$ is an *R-module*, if

1. $(M, +)$ is an abelian group,

and the scalar product $\cdot : R \times M \to M$ satisfies the following axioms:

2. $1 \cdot m = m$, $1 \in R$;
3. $(rs) \cdot m = r \cdot (s \cdot m)$;
4. $(r + s) \cdot m = r \cdot m + s \cdot m$;
5. $r \cdot (m + n) = r \cdot m + r \cdot n$

for all $r, s \in R$ and $m, n \in M$.

The elements of R are called *scalars*.

Let now M be an R-module, S a subring of R and $B \subseteq M$.

Definition 15. A non-empty subset $N \subseteq M$ is an *S-submodule* of M, if $(N, +)$ is a subgroup of $(M, +)$, and $s \cdot n \in N$ for every $s \in S$, $n \in N$.

Definition 16. The smallest S-submodule containing the set $B \subseteq M$, denoted by

$$\langle B \rangle_S := \bigcap_{B \subseteq N} N, \quad N \text{ is an } S\text{-submodule of } M,$$

is called a *linear hull* over S generated by B. In particular, the set

$$\langle m_1, \ldots, m_k \rangle_S := Sm_1 + \ldots + Sm_k$$

is called a linear hull over S generated by the elements $m_1, \ldots, m_k \in M$.

Remark 3. When K is a field, a K-module is the same thing as a vector space over K. Taking $M = K^n$, we may write

$$M = K^n = \langle \bar{e}_1, \ldots, \bar{e}_n \rangle_K = K\bar{e}_1 + \ldots + K\bar{e}_n;$$

the vector space K^n is a linear hull over the field K generated by the standard basis vectors $\bar{e}_1, \ldots, \bar{e}_n$.

Definition 17. Let $\bar{l}_1, \ldots, \bar{l}_r \in \mathbb{R}^n$ be linearly independent over $\mathbb{R}$. Then the linear hull

$$\Lambda := \left\langle \bar{l}_1, \ldots, \bar{l}_r \right\rangle_{\mathbb{Z}} = \mathbb{Z}\bar{l}_1 + \ldots + \mathbb{Z}\bar{l}_r \subseteq \mathbb{R}^n$$

over $\mathbb{Z}$ forms a *lattice*. The set $\left\{\bar{l}_1, \ldots, \bar{l}_r\right\}$ forms a *basis* for the lattice Λ, with $\operatorname{rank} \Lambda = r$. If $\operatorname{rank} \Lambda = n$, the Λ is called a *full lattice*.

The *determinant* of Λ is defined by $\det \Lambda := \sqrt{\det(M^T M)}$, where $M := \begin{bmatrix} \bar{l}_1 & \cdots & \bar{l}_r \end{bmatrix}$ is an $n \times r$ matrix formed from the basis vectors $\bar{l}_1, \ldots, \bar{l}_r$ of the lattice Λ.

The set

$$\mathcal{F} := \left\{ x_1\bar{l}_1 + \ldots + x_r\bar{l}_r \;\middle|\; x_i \in \mathbb{R},\, 0 \leq x_i < 1,\, i = 1, \ldots, r \right\}$$

is called the *fundamental parallelepiped* of the lattice Λ.

Remark 4. If $\Lambda \subseteq \mathbb{R}^n$ is a full lattice, the determinant becomes $\det \Lambda = |\det M|$, since M is now a square matrix. Also $V(\mathcal{F}) = \det \Lambda$, and we can write

$$\mathbb{R}^n = \bigcup_{\bar{u} \in \Lambda} (\bar{u} + \mathcal{F}),$$

where the translates $\bar{u} + \mathcal{F} := \left\{ \bar{u} + \bar{x} \;\middle|\; \bar{x} \in \mathcal{F} \right\}, \bar{u} \in \Lambda$, are pairwise disjoint.

Remark 5. The integer lattice $\Lambda = \mathbb{Z}^n = \mathbb{Z}\bar{e}_1 + \ldots + \mathbb{Z}\bar{e}_n$ has determinant $\det \Lambda = 1$, as $M = \begin{bmatrix} \bar{e}_1 & \cdots & \bar{e}_n \end{bmatrix}$ is now simply the identity matrix.

Lemma 6. *If $\Lambda \subseteq \mathbb{R}^2$ is a full lattice, then the lattice $\Lambda^m \subseteq \mathbb{R}^{2m}, m \in \mathbb{Z}^+$, has determinant $(\det \Lambda)^m$.*

Proof. Let $\Lambda = \mathbb{Z}\bar{l}_1 + \mathbb{Z}\bar{l}_2 =: \mathbb{Z}(l_{11}, l_{12})^T + \mathbb{Z}(l_{21}, l_{22})^T$ and denote $M := \begin{bmatrix} \bar{l}_1 & \bar{l}_2 \end{bmatrix}$. Then

$$\begin{aligned}
\Lambda^m &= \left\{ (\bar{x}_1, \ldots, \bar{x}_m)^T \in \mathbb{R}^{2m} \;\middle|\; \bar{x}_i \in \Lambda,\ i = 1, \ldots, m \right\} \\
&= \Big\{ (x_{11}l_{11} + x_{12}l_{21},\ x_{11}l_{12} + x_{12}l_{22},\ x_{21}l_{11} + x_{22}l_{21}, \ldots, \\
&\quad x_{m1}l_{12} + x_{m2}l_{22})^T \in \mathbb{R}^{2m} \;\middle|\; x_{ij} \in \mathbb{Z},\ i = 1, \ldots, m,\ j = 1, 2 \Big\} \\
&= \big\{ x_{11}(l_{11}, l_{12}, 0, \ldots, 0)^T + x_{12}(l_{21}, l_{22}, 0, \ldots, 0)^T + \\
&\quad x_{21}(0, 0, l_{11}, l_{12}, 0, \ldots, 0)^T + \ldots + x_{m2}(0, \ldots, 0, l_{21}, l_{22})^T \;\big|\; x_{ij} \in \mathbb{Z} \big\} \\
&= \mathbb{Z}(l_{11}, l_{12}, 0, \ldots, 0)^T + \mathbb{Z}(l_{21}, l_{22}, 0, \ldots, 0)^T + \\
&\quad \mathbb{Z}(0, 0, l_{11}, l_{12}, 0, \ldots, 0)^T + \ldots + \mathbb{Z}(0, \ldots, 0, l_{21}, l_{22})^T.
\end{aligned}$$

Using the notation of Definition 17, we can write $\det(\Lambda^m) = |\det M_m|$, where

$$\det M_m = \begin{vmatrix} l_{11} & l_{21} & 0 & 0 & \cdots & 0 & 0 \\ l_{12} & l_{22} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & l_{11} & l_{21} & \cdots & 0 & 0 \\ 0 & 0 & l_{12} & l_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & l_{11} & l_{21} \\ 0 & 0 & 0 & 0 & \cdots & l_{12} & l_{22} \end{vmatrix}_{2m \times 2m} = \begin{vmatrix} l_{11} & l_{21} \\ l_{12} & l_{22} \end{vmatrix}^m = (\det M)^m .$$

Hence $\det(\Lambda^m) = |\det M_m| = |\det M|^m = (\det \Lambda)^m$. □

Remark 6. Obviously Lemma 6 generalizes to the case $\Lambda \subseteq \mathbb{R}^n$, but to avoid more messing with indices, we stick to $n = 2$, which is enough for the purposes of Chapter 5.

4.3 The convex body theorems

For the main theorem of this chapter two different proofs are presented, both of which cleverly exploit the ever so familiar pigeonhole principle again.

Theorem 8 (The first Minkowski's convex body theorem)**.** *Suppose $\Lambda \in \mathbb{R}^n$ is a full lattice and $\mathcal{C} \subseteq \mathbb{R}^n$ is a central symmetric convex body with*

$$\begin{cases} V(\mathcal{C}) > 2^n \det \Lambda & \text{or} \\ V(\mathcal{C}) \geq 2^n \det \Lambda, & \text{if } \mathcal{C} \text{ is compact.} \end{cases}$$

Then there exists a non-zero lattice point in the set $\mathcal{C}$.

First proof. [13, pp. 121, 123] We start by examining the case $\Lambda = \mathbb{Z}^n$, which can then be generalized using a linear transformation.

Fix $t \in \mathbb{Z}^+$. The hyperplanes

$$x_j = \frac{2z_j}{t}, \quad j = 1, \ldots, n,\ z_j \in \mathbb{Z},$$

divide the space into hypercubes of volume $\left(\frac{2}{t}\right)^n$. Let $N(t)$ denote the number of corners of these cubes inside the set $\mathcal{C}$. Because $\mathcal{C}$ is bounded, this number is finite. The smaller the hypercubes are, the bigger portion of the vertices inside $\mathcal{C}$ are shared by several cubes contained in $\mathcal{C}$. Therefore the number of corners $N(t)$ approximates the number of cubes in $\mathcal{C}$, and we can write

$$V(\mathcal{C}) = \lim_{t\to\infty} \left(\frac{2}{t}\right)^n N(t).$$

Since $V(\mathcal{C}) > 2^n$ by the assumption, for the limit we have $\lim_{t\to\infty} \frac{N(t)}{t^n} > 1$, implying $N(t) > t^n$ for sufficiently large t. However, when n integers $z_1, \ldots, z_n$ are divided by t, the n-tuple of remainders has only t^n different possibilities. Consequently, the set $\mathcal{C}$ contains two distinct vertex points

$$\bar{P}_1 := \left(\frac{2z_1^{(1)}}{t}, \ldots, \frac{2z_n^{(1)}}{t}\right) \quad \text{and} \quad \bar{P}_2 := \left(\frac{2z_1^{(2)}}{t}, \ldots, \frac{2z_n^{(2)}}{t}\right)$$

giving exactly the same n-tuple of remainders. This leads to the difference $z_j^{(1)} - z_j^{(2)}$ being divisible by t for all $j = 1, \ldots, n$. Because $\mathcal{C}$ is central symmetric, we have $-\bar{P}_2 \in \mathcal{C}$. Hence the midpoint

$$\bar{M} := \frac{1}{2}\left(\bar{P}_1 - \bar{P}_2\right) = \left(\frac{z_1^{(1)} - z_1^{(2)}}{t}, \ldots, \frac{z_n^{(1)} - z_n^{(2)}}{t}\right)$$

of the line segment connecting the points $\bar{P}_1$ and $-\bar{P}_2$ has integer coordinates and, by the convexity of $\mathcal{C}$, lies inside $\mathcal{C}$. Lastly, $\bar{P}_1 \neq \bar{P}_2$ implies $\bar{M} \neq \bar{0}$. Hence we have a non-zero vector $\bar{M} \in \mathcal{C} \cap \Lambda$.

As for the second case, let $\mathcal{C} \subseteq \mathbb{R}^n$ now be a closed central symmetric convex body with $V(\mathcal{C}) = 2^n$. Then $\lambda\mathcal{C} \supsetneq \mathcal{C}$ and $V(\lambda\mathcal{C}) > V(\mathcal{C}) = 2^n$ for all $\lambda > 1$. In view of the first case, there is a non-zero lattice point in $\lambda\mathcal{C}$ for all

$\lambda > 1$. Since any particular set $\lambda\mathcal{C}$ is bounded, it cannot contain infinitely many distinct lattice points. Therefore there has to be a lattice point in the intersection $\bigcap_{\lambda>1} \lambda\mathcal{C} = \mathcal{C}$. (Here it is essential that $\mathcal{C}$ is closed, for actually $\bigcap_{\lambda>1} \lambda\mathcal{C} = \mathcal{C} \cup \partial\mathcal{C}$ and the point in question could be on the boundary $\partial\mathcal{C}$ of the set $\mathcal{C}$. However, $\mathcal{C}$ being closed means $\partial\mathcal{C} \subseteq \mathcal{C}$.)

Finally the generalization: Let $\Lambda \in \mathbb{R}^n$ be a full lattice with the basis $\{\bar{l}_1, \ldots, \bar{l}_n\} \subseteq \mathbb{R}^n$ and $\mathcal{C} \subseteq \mathbb{R}^n$ a central symmetric convex body with $V(\mathcal{C}) > 2^n \det\Lambda$. Define a bijective linear mapping $\bar{L} : \mathbb{R}^n \to \mathbb{R}^n$ by setting $\bar{L}\bar{x} = x_1\bar{l}_1 + \ldots + x_n\bar{l}_n$. Then

$$\begin{aligned} \bar{L}(\mathbb{Z}^n) &= \bar{L}(\mathbb{Z}\bar{e}_1 + \ldots + \mathbb{Z}\bar{e}_n) \\ &= \mathbb{Z}\bar{L}\bar{e}_1 + \ldots + \mathbb{Z}\bar{L}\bar{e}_n \\ &= \mathbb{Z}\bar{l}_1 + \ldots + \mathbb{Z}\bar{l}_n \\ &= \Lambda. \end{aligned} \tag{4.1}$$

The volume of the transformed convex body $\bar{L}(\mathcal{C})$ obeys the rule

$$V\big(\bar{L}(\mathcal{C})\big) = \left|\det\begin{bmatrix}\bar{l}_1 & \cdots & \bar{l}_n\end{bmatrix}\right| V(\mathcal{C}) = \det(\Lambda)V(\mathcal{C}),$$

the matrix of the mapping $\bar{L}$ being $\begin{bmatrix}\bar{l}_1 & \cdots & \bar{l}_n\end{bmatrix} =: L$. Since $\bar{L}$ is bijective, it has an inverse $\bar{L}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$, whose matrix L^{-1} has determinant $\frac{1}{\det L}$. According to Lemma 4 the set $\bar{L}^{-1}(\mathcal{C}) \subseteq \mathbb{R}^n$ is a central symmetric convex body with volume

$$V\big(\bar{L}^{-1}(\mathcal{C})\big) = \left|\frac{1}{\det L}\right| V(\mathcal{C}) = \frac{V(\mathcal{C})}{\det\Lambda} > 2^n.$$

By what was shown above, there exists a non-zero vector $\bar{z} \in \bar{L}^{-1}(\mathcal{C}) \cap \mathbb{Z}^n$. From the injectivity of $\bar{L}$ and the equation (4.1) it follows that $\bar{0} \neq \bar{L}\bar{z} \in \mathcal{C} \cap \Lambda$.

Suppose in addition that $\mathcal{C}$ is closed and $V(\mathcal{C}) = 2^n \det\Lambda$. The linear mapping $\bar{L}$ is continuous, so the preimage $\bar{L}^{-1}(\mathcal{C})$ is closed as well. The

dilation $\lambda\bar{L}^{-1}(\mathcal{C}) \supsetneq \bar{L}^{-1}(\mathcal{C}),\ \lambda > 1$, has volume

$$V\big(\lambda\bar{L}^{-1}(\mathcal{C})\big) > V\big(\bar{L}^{-1}(\mathcal{C})\big) = \frac{V(\mathcal{C})}{\det\Lambda} = 2^n.$$

Again the intersection $\bigcap_{\lambda>1}\lambda\bar{L}^{-1}(\mathcal{C}) = \bar{L}^{-1}(\mathcal{C})$ has to contain a non-zero lattice point $\bar{z} \in \mathbb{Z}^n$, as the intersecting sets $\lambda\bar{L}^{-1}(\mathcal{C})$ all contain one. Hence $\bar{0} \neq \bar{L}\bar{z} \in \mathcal{C} \cap \Lambda$. □

Second proof. [5, p. 11] Assume $V(\mathcal{C}) > 2^n \det\Lambda$, and let $\mathcal{F}$ stand for the fundamental parallelepiped of the lattice Λ. Now the set $\frac{1}{2}\mathcal{C}$ has volume $V\big(\frac{1}{2}\mathcal{C}\big) = \frac{1}{2^n}V(\mathcal{C}) > \det\Lambda$. For $\bar{u} \in \Lambda$, define $\mathcal{S}_{\bar{u}} := \frac{1}{2}\mathcal{C} \cap (\bar{u} + \mathcal{F})$. Then the sets $\mathcal{S}_{\bar{u}}, \bar{u} \in \Lambda$, are pairwise disjoint and $\bigcup_{\bar{u}\in\Lambda}\mathcal{S}_{\bar{u}} = \frac{1}{2}\mathcal{C}$. Hence

$$\sum_{\bar{u}\in\Lambda} V(\mathcal{S}_{\bar{u}}) = V\left(\frac{1}{2}\mathcal{C}\right) > \det\Lambda.$$

Now all the sets $\mathcal{S}_{\bar{u}}$ are shifted into the fundamental parallelepiped $\mathcal{F}$ by defining

$$\mathcal{S}^*_{\bar{u}} := -\bar{u} + \mathcal{S}_{\bar{u}} = \left(-\bar{u} + \frac{1}{2}\mathcal{C}\right) \cap \mathcal{F}, \quad \bar{u} \in \Lambda.$$

This doesn't change the volumes of the sets, so

$$\sum_{\bar{u}\in\Lambda} V(\mathcal{S}^*_{\bar{u}}) = \sum_{\bar{u}\in\Lambda} V(\mathcal{S}_{\bar{u}}) > \det\Lambda = V(\mathcal{F}) \geq V\left(\bigcup_{\bar{u}\in\Lambda}\mathcal{S}^*_{\bar{u}}\right).$$

This implies that there exist two distinct points $\bar{u}, \bar{v} \in \Lambda$ such that $\mathcal{S}^*_{\bar{u}} \cap \mathcal{S}^*_{\bar{v}} \neq \varnothing$. Pick a point $\bar{a} \in \mathcal{S}^*_{\bar{u}} \cap \mathcal{S}^*_{\bar{v}}$. Then $\bar{a} = \bar{x} - \bar{u} = \bar{y} - \bar{v}$ for some $\bar{x}, \bar{y} \in \frac{1}{2}\mathcal{C}$, whence $\bar{z} := \bar{x} - \bar{y} = \bar{u} - \bar{v} \in \Lambda \setminus \{\bar{0}\}$. Now $2\bar{x}, 2\bar{y} \in \mathcal{C}$, and because $\mathcal{C}$ is central symmetric and convex, we have $-2\bar{y} \in \mathcal{C}$ and $\frac{1}{2}\cdot 2\bar{x} + \frac{1}{2}(-2\bar{y}) = \bar{x} - \bar{y} = \bar{z} \in \mathcal{C}$. Thus the set $\mathcal{C}$ contains a non-zero lattice point.

For the case that $\mathcal{C}$ is closed, see the first proof. □

Remark 7. The second proof is essentially about cutting the set $\frac{1}{2}\mathcal{C}$ with the lattice Λ and then rearranging the pieces into the fundamental parallelepiped

$\mathcal{F}$. Since the added volume of the pieces (which is the volume of $\frac{1}{2}\mathcal{C}$) is greater than the volume of $\mathcal{F}$, some intersecting must happen.

Remark 8. Actually $\#(\mathcal{C} \cap \Lambda) \geq 3$, for $\left\{\bar{0}, \bar{M}, -\bar{M}\right\} \subseteq \mathcal{C} \cap \Lambda$.

Definition 18. Let $\mathcal{C} \subseteq \mathbb{R}^n$ be non-empty. The *successive minima* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ of the set $\mathcal{C}$ with respect to a lattice $\Lambda \subseteq \mathbb{R}^n$ are given by

$$\lambda_j = \inf\left\{\lambda > 0 \mid \operatorname{rank}\langle \lambda\mathcal{C} \cap \Lambda\rangle_{\mathbb{Z}} \geq j\right\}, \quad j = 1, \ldots, n.$$

That is, the number λ_j is the lower bound of all $\lambda > 0$ such that the dilation $\lambda\mathcal{C}$ contains at least j linearly independent lattice points.

Lemma 7. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a central symmetric convex body with $\bar{0}$ as an interior point, and $\Lambda \subseteq \mathbb{R}^n$ a full lattice. Then the successive minima of $\mathcal{C}$ with respect to Λ satisfy the inequality chain $0 < \lambda_1 \leq \ldots \leq \lambda_n < \infty$.*

Proof. From the obvious inclusion

$$\{\lambda > 0 \mid \operatorname{rank}\langle \lambda\mathcal{C} \cap \Lambda\rangle_{\mathbb{Z}} \geq j+1\} \subseteq \{\lambda > 0 \mid \operatorname{rank}\langle \lambda\mathcal{C} \cap \Lambda\rangle_{\mathbb{Z}} \geq j\}$$

and the previous definition it follows that $\lambda_j \leq \lambda_{j+1}$, when $j = 1, \ldots, n-1$. Letting $\lambda \to 0$ we have $\lambda\mathcal{C} \to \{\bar{0}\}$, since $\mathcal{C}$ is bounded, so ultimately the only lattice point contained in the intersection $\lambda\mathcal{C} \cap \Lambda$ is $\bar{0}$. Therefore $\lambda_1 \neq 0$.

Take now n linearly independent lattice points $\bar{l}_1, \ldots, \bar{l}_n$. By assumption there exists a ball $\mathcal{B}(\bar{0}, r) := \{\bar{x} \in \mathbb{R}^n \mid \|\bar{x}\| < r\}$ such that $\mathcal{B}(\bar{0}, r) \subseteq \mathcal{C}$ for some $r \in \mathbb{R}$. Therefore we can pick n vectors $\bar{c}_1, \ldots, \bar{c}_n \in \mathcal{C}$ such that $\bar{l}_k = \tilde{\lambda}_k \bar{c}_k$ for some $\tilde{\lambda}_k \in \mathbb{R}^+,\ k = 1, \ldots, n$. Choose $\tilde{\lambda} = \max_{1 \leq k \leq n}\left\{\tilde{\lambda}_k\right\}$, whereupon it follows from the convexity of the set $\tilde{\lambda}\mathcal{C}$ (see Lemma 3) that $\bar{l}_1, \ldots, \bar{l}_n \in \tilde{\lambda}\mathcal{C}$ (the n line segments connecting $\bar{0}$ with the points $\tilde{\lambda}\bar{c}_1, \ldots, \tilde{\lambda}\bar{c}_n$ contain each of the chosen n lattice points $\bar{l}_k$). Thus $\{\lambda > 0 \mid \operatorname{rank}\langle \lambda\mathcal{C} \cap \Lambda\rangle_{\mathbb{Z}} \geq n\} \neq \varnothing$ and therefore the infimum λ_n exists. □

Lemma 8. *The successive minima of a closed, central symmetric convex body $\mathcal{C} \subseteq \mathbb{R}^n$ are actual minima.*

Proof. Similar deduction as in the proof of Theorem 8 applies. Let $\lambda_k \in \mathbb{R}^+$ be a successive minimum of the set $\mathcal{C}$. Then $\lambda\mathcal{C}$ contains k linearly independent lattice points for all $\lambda > \lambda_k$. Because the sets $\lambda\mathcal{C}$ are bounded, they contain only finitely many lattice points. Therefore the intersection $\bigcap_{\lambda>\lambda_k} \lambda\mathcal{C} = \lambda_k\mathcal{C}$ has to contain k linearly independent lattice points as well. □

Theorem 9 (The second Minkowski's convex body theorem)**.** *Let $\Lambda \subseteq \mathbb{R}^n$ be a full lattice and $\mathcal{C} \subseteq \mathbb{R}^n$ a central symmetric convex body with successive minima $\lambda_1, \ldots, \lambda_n \in \mathbb{R}^+$. Then*

$$\frac{2^n}{n!} \det \Lambda \leq \lambda_1 \cdots \lambda_n V(\mathcal{C}) \leq 2^n \det \Lambda.$$

Partial proof. [5, p. 19] We follow Chapter 2 of Evertse's notes and prove the lower bound assuming that $\mathcal{C}$ is closed, and the upper bound for the euclidean unit ball $\mathcal{B}_n := \{\bar{x} \in \mathbb{R}^n \mid \|\bar{x}\| \leq 1\}$.

Let first $\{\bar{l}_1, \ldots, \bar{l}_n\} \subseteq \mathbb{R}^n$ be a basis for the lattice Λ. In light of Lemma 8, there exist n linearly independent vectors $\bar{v}_1, \ldots, \bar{v}_n \in \Lambda$ such that $\bar{v}_i \in \lambda_i\mathcal{C}$ for each $i = 1, \ldots, m$. Since each of these vectors has a representation $\bar{v}_i = \sum_{k=1}^n a_{ik}\bar{l}_k,\ a_{ik} \in \mathbb{Z},\ i = 1, \ldots, n$, in the basis $\{\bar{l}_1, \ldots, \bar{l}_n\} \subseteq \mathbb{R}^n$, we can write

$$\begin{bmatrix} \bar{v}_1 & \cdots & \bar{v}_n \end{bmatrix} = \begin{bmatrix} \bar{l}_1 & \cdots & \bar{l}_n \end{bmatrix} A,$$

where $A \in \mathcal{M}_{n\times n}(\mathbb{Z})$ is a non-singular matrix. Moreover, by Lemma 5 the set

$$\mathcal{S} := \left\{ \frac{x_1}{\lambda_1}\bar{v}_1 + \ldots + \frac{x_n}{\lambda_n}\bar{v}_n \;\middle|\; x_1, \ldots, x_n \in \mathbb{R},\ \sum_{i=1}^n |x_i| \leq 1 \right\}$$

is the smallest central symmetric convex body containing the points $\frac{1}{\lambda_i}\bar{v}_i \in \mathcal{C},\ i = 1, \ldots, n$. Hence $\mathcal{S} \subseteq \mathcal{C}$.

The set $\mathcal{S}$ is the image of the n-dimensional octahedron

$$\mathcal{O}_n := \left\{ \bar{x} \in \mathbb{R}^n \;\middle|\; \sum_{i=1}^{n} |x_i| \leq 1 \right\}$$

under the linear transformation $\bar{L} : \mathbb{R}^n \to \mathbb{R}^n,\ \bar{L}\bar{x} = \frac{x_1}{\lambda_1}\bar{v}_1 + \ldots + \frac{x_n}{\lambda_n}\bar{v}_n$. As it is explained in [1, p. 429], the volume of an n-dimensional pyramid $\mathcal{P} \subseteq \mathbb{R}^n$ is $V(\mathcal{P}) = \frac{h}{n}V(\mathcal{B})$, where h is the height of the pyramid and $\mathcal{B} \subseteq \mathbb{R}^{n-1}$ denotes its $(n-1)$-dimensional base. Each hyperoctahedron $\mathcal{O}_n$ is formed from two pyramids of height 1 with $\mathcal{O}_{n-1}$ as their common base. Therefore the formula for the volume becomes recursive, yielding

$$V(\mathcal{O}_n) = 2 \cdot \frac{1}{n}V(\mathcal{O}_{n-1}) = \frac{2}{n} \cdot 2 \cdot \frac{1}{n-1}V(\mathcal{O}_{n-2}) = \ldots = \frac{2^n}{n!},$$

since $V(\mathcal{O}_1) = V([-1, 1]) = 2$.

Consequently (L denoting, as usual, the matrix of the mapping $\bar{L}$),

$$\begin{aligned}
V(\mathcal{C}) &\geq V(\mathcal{S}) \\
&= |\det L|\, V(\mathcal{O}_n) \\
&= \left|\det \left[\begin{array}{ccc} \frac{1}{\lambda_1}\bar{v}_1 & \cdots & \frac{1}{\lambda_n}\bar{v}_n \end{array}\right]\right| \cdot \frac{2^n}{n!} \\
&= \frac{1}{\lambda_1 \cdots \lambda_n} \left|\det \left[\begin{array}{ccc} \bar{v}_1 & \cdots & \bar{v}_n \end{array}\right]\right| \cdot \frac{2^n}{n!} \\
&= \frac{2^n}{n!} \cdot \frac{1}{\lambda_1 \cdots \lambda_n} \left|\det \left(\left[\begin{array}{ccc} \bar{l}_1 & \cdots & \bar{l}_n \end{array}\right] A\right)\right| \\
&= \frac{2^n}{n!} \cdot \frac{|\det A| \det \Lambda}{\lambda_1 \cdots \lambda_n} \\
&\geq \frac{2^n}{n!} \cdot \frac{\det \Lambda}{\lambda_1 \cdots \lambda_n}.
\end{aligned}$$

As for the upper bound, let $\mathcal{C} = \mathcal{B}_n$ and define

$$\mu := \left(\frac{2^n \det \Lambda}{V(\mathcal{B}_n)\lambda_1 \cdots \lambda_n}\right)^{\frac{1}{n}}.$$

Then it suffices to prove that $\mu \geq 1$.

Pick again n linearly independent lattice vectors $\bar{v}_1, \ldots, \bar{v}_n \in \Lambda$ such that $\bar{v}_i \in \lambda_i \mathcal{B}_n$, meaning $\|\bar{v}_i\| = \lambda_i$ for $i = 1, \ldots, n$. Using the Gram-Schmidt orthogonalization procedure, an orthonormal basis $\{\bar{e}_1, \ldots, \bar{e}_n\}$ of the space $\mathbb{R}^n$ can be constructed so that $\langle \bar{e}_1, \ldots, \bar{e}_i \rangle_{\mathbb{R}} = \langle \bar{v}_1, \ldots, \bar{v}_i \rangle_{\mathbb{R}}$ for $i = 1, \ldots, n$. Let $\bar{L} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation defined by $\bar{L}\bar{e}_i = \lambda_i \bar{e}_i$, $i = 1, \ldots, n$; it has determinant $\det L = \lambda_1 \cdots \lambda_n$. When the unit ball is transformed and dilated, its volume becomes

$$V\big(\mu \bar{L}(\mathcal{B}_n)\big) = \mu^n \lambda_1 \cdots \lambda_n V(\mathcal{B}_n) = 2^n \det \Lambda.$$

Hence, according to Minkowski's first convex body theorem, the set $\mu \bar{L}(\mathcal{B}_n)$ contains a non-zero lattice point $\bar{0} \neq \bar{z} \in \Lambda$.

First, $\bar{L}^{-1}(\bar{z}) \in \mu \mathcal{B}_n$, so

$$\big\|\bar{L}^{-1}(\bar{z})\big\| \leq \mu. \tag{4.2}$$

Then, since $\bar{z} \in \Lambda \setminus \{\bar{0}\}$, we have $\|\bar{z}\| \geq \lambda_1$. Let

$$j := \max\{i \in \{1, \ldots, n\} \mid \|\bar{z}\| \geq \lambda_i\}.$$

This means that $\bar{z} \in \langle \bar{v}_1, \ldots, \bar{v}_j \rangle_{\mathbb{R}}$. Indeed, if $j = n$, then $\langle \bar{v}_1, \ldots, \bar{v}_n \rangle_{\mathbb{R}} = \mathbb{R}^n \ni \bar{z}$. If $j < n$, it holds $\|\bar{z}\| < \lambda_{j+1}$. Now the vectors $\bar{v}_1, \ldots, \bar{v}_j \in \Lambda$ are linearly independent and $\|\bar{v}_1\|, \ldots, \|\bar{v}_j\| \leq \lambda_j < \lambda_{j+1}$ too. As there cannot be $j + 1$ linearly independent lattice points having euclidean norm strictly less than λ_{j+1}, we must have $\bar{z} \in \langle \bar{v}_1, \ldots, \bar{v}_j \rangle_{\mathbb{R}}$.

As mentioned before, $\langle \bar{v}_1, \ldots, \bar{v}_j \rangle_{\mathbb{R}} = \langle \bar{e}_1, \ldots, \bar{e}_j \rangle_{\mathbb{R}}$, so we can write $\bar{z} = z_1 \bar{e}_i + \ldots + z_j \bar{e}_j$ for some $z_1, \ldots, z_j \in \mathbb{R}$. Accordingly, $\bar{L}^{-1}(\bar{z}) = \sum_{k=1}^{j} \frac{z_k}{\lambda_k} \bar{e}_k$. Using the definition of j and the fact that $\lambda_1 < \ldots < \lambda_j$ results in

$$\big\|\bar{L}^{-1}(\bar{z})\big\|^2 = \sum_{k=1}^{j} \left(\frac{z_k}{\lambda_k}\right)^2 \geq \frac{1}{\lambda_j^2} \sum_{k=1}^{j} z_k^2 = \frac{\|\bar{z}\|^2}{\lambda_j^2} \geq 1. \tag{4.3}$$

Combining the inequalities (4.2) and (4.3) shows that $\mu \geq 1$, thus completing the proof. □

Remark 9. Proving the upper bound for the euclidean unit ball actually proves it for any *ellipsoid* $\mathcal{E} = \bar{L}(\mathcal{B}_n)$, where $\bar{L} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation: Denote $\Lambda' := \bar{L}(\Lambda)$. Then the successive minima of the ellipsoid $\mathcal{E}$ with respect to the lattice Λ' are equal to the successive minima of the unit ball $\mathcal{B}_n$ with respect to the lattice Λ. Furthermore, $V(\mathcal{E}) = |\det L| V(\mathcal{B}_n)$ and $\det \Lambda' = |\det L| \det \Lambda$. [1] Therefore

$$V(\mathcal{E}) = \frac{\det \Lambda'}{\det \Lambda} V(\mathcal{B}_n) \leq \frac{\det \Lambda'}{\det \Lambda} \cdot \frac{2^n \det \Lambda}{\lambda_1 \cdots \lambda_n} = \frac{2^n \det \Lambda'}{\lambda_1 \cdots \lambda_n}.$$

The full proof of Minkowski's second theorem, requiring more advanced theory, can be found in Chapter 8 of Cassels' book [3].

[1] If the vectors $\bar{l}_1, \ldots, \bar{l}_n$ form a base for the lattice Λ, then the vectors $\bar{L}(\bar{l}_1), \ldots, \bar{L}(\bar{l}_n)$ form a base for the image Λ'. The latter equality follows directly from this.

Chapter 5

Some Diophantine inequalities

The first theorem of Minkowski from the preceding chapter is now put to use.

5.1 A few Diophantine inequalities over $\mathbb{R}$

Theorem 10. *Let $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and $h_1, \ldots, h_m \in \mathbb{Z}^+$ be given. Then there exist $m+1$ numbers $p, q_1, \ldots, q_m \in \mathbb{Z}$, with some $q_i \neq 0$, satisfying the conditions*

$$|q_i| \leq h_i, \quad i = 1, \ldots, m \tag{5.1}$$

and

$$|p + q_1\alpha_1 + \ldots + q_m\alpha_m| < \frac{1}{h_1 \cdots h_m}. \tag{5.2}$$

Proof. [9, p. 5] Writing $L_0\bar{x} := x_0 + \alpha_1 x_1 + \ldots + \alpha_m x_m$ and $L_k\bar{x} := x_k$, $k = 1, \ldots, m$, defines $m+1$ linear mappings $L_i : \mathbb{Z}^{m+1} \to \mathbb{R}$, $i = 0, 1, \ldots, m+1$. By transforming the integer lattice $\mathbb{Z}^{m+1} = \mathbb{Z}\bar{e}_1 + \ldots + \mathbb{Z}\bar{e}_{m+1}$ with the mapping

$\bar{L} := (L_0, L_1, \ldots, L_m) : \mathbb{Z}^{m+1} \to \mathbb{R}^{m+1}$ we get a full lattice

$$\begin{aligned}
\Lambda &:= \bar{L}(\mathbb{Z}^{m+1}) \\
&= \mathbb{Z}\bar{L}\bar{e}_1 + \ldots + \mathbb{Z}\bar{L}\bar{e}_{m+1} \\
&= \mathbb{Z}(1, 0, \ldots, 0)^T + \mathbb{Z}(\alpha_1, 1, 0, \ldots, 0)^T + \ldots + \mathbb{Z}(\alpha_m, 0, \ldots, 1)^T \\
&=: \mathbb{Z}\bar{l}_0 + \mathbb{Z}\bar{l}_1 + \ldots + \mathbb{Z}\bar{l}_m \\
&\subseteq \mathbb{R}^{m+1}
\end{aligned}$$

with determinant

$$\det \Lambda = \left|\det \begin{bmatrix} \bar{l}_0 & \bar{l}_1 & \cdots & \bar{l}_m \end{bmatrix}\right| = \begin{vmatrix} 1 & \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

Now put

$$|L_0\bar{x}| < \frac{1}{h} := \frac{1}{h_1 \cdots h_m}$$

and

$$|L_k\bar{x}| \leq h_k + \frac{1}{2}, \quad k = 1, \ldots, m.$$

These conditions determine a central symmetric convex body

$$\mathcal{C} := \left\{ (y_0, y_1, \ldots, y_m)^T \in \mathbb{R}^{m+1} \;\middle|\; |y_0| < \frac{1}{h},\ |y_k| \leq h_k + \frac{1}{2} \right\} \subseteq \mathbb{R}^{m+1},$$

where $k = 1, \ldots, m$. The volume of this hyperrectangle is

$$\begin{aligned}
V(\mathcal{C}) &= \frac{2}{h} \cdot 2\left(h_1 + \frac{1}{2}\right) \cdot \ldots \cdot 2\left(h_m + \frac{1}{2}\right) \\
&= \frac{2^{m+1} \prod_{k=1}^{m} \left(h_k + \frac{1}{2}\right)}{h_1 \cdots h_m} \\
&> 2^{m+1} \\
&= 2^{m+1} \det \Lambda.
\end{aligned}$$

By the first Minkowski's convex body theorem, there exists a non-zero vector

$$\bar{0} \neq \bar{y} := (p + q_1\alpha_1 + \ldots + q_m\alpha_m, q_1, \ldots, q_m)^T \in \mathcal{C} \cap \Lambda,$$

where $(p, q_1, \ldots, q_m)^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$, $|q_i| \leq h_i$ for all $i = 1, \ldots, m$, and

$$|p + q_1\alpha_1 + \ldots + q_m\alpha_m| < \frac{1}{h_1 \cdots h_m}. \tag{5.3}$$

If $q_i = 0$ for all $i = 1, \ldots, m$, then the equation (5.3) above implies $p = 0$ too, which is a contradiction. □

Remark 10. If $-\alpha \in \mathbb{R}$ is an irrational number and $h \in \mathbb{Z}^+$, according to Theorem 10 there exist $p, q \in \mathbb{Z},\ q \neq 0$, such that

$$\left|\frac{p}{q} - \alpha\right| = \frac{1}{|q|}|p - q\alpha| \leq |p - q\alpha| < \frac{1}{h}.$$

This is a proof for the fact that any real irrational number can be approximated by rationals with accuracy as good as we please. Since $|q| \leq h$, actually we get the result first discovered by Dirichlet in 1842: *For any irrational number $\alpha \in \mathbb{R}$ there exists infinitely many rational numbers $\frac{p}{q} \in \mathbb{Q}$ with*

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

The fact that there are indeed infinitely many solutions to this equation will be demonstrated later in Theorem 12.

Definition 19. An integer vector $(r_1, \ldots, r_n)^T \in \mathbb{Z}^n$ is said to be *primitive*, if $\gcd(r_1, \ldots, r_n) = 1$.

With this definition a tuned version of the previous theorem can easily be achieved:

Theorem 11. *Let $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and $h_1, \ldots, h_m \in \mathbb{Z}^+$ be given. Then there exists a primitive vector $(s, r_1, \ldots, r_m) \in \mathbb{Z}^{m+1}$, with some $r_i \neq 0$, satisfying the conditions*

$$|r_i| \leq h_i, \quad i = 1, \ldots, m$$

and

$$|s + r_1\alpha_1 + \ldots + r_m\alpha_m| < \frac{1}{h_1 \cdots h_m}.$$

Proof. [9, p. 6] Theorem 10 gives an integer vector $(p, q_1, \ldots, q_m)^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$, with some $q_i \neq 0$, satisfying the conditions (5.1) and (5.2). Denoting $d := \gcd(p, q_1, \ldots, q_m) \in \mathbb{Z}^+$ we have $(p, q_1, \ldots, q_m)^T = d(s, r_1, \ldots, r_m)^T$, where $(s, r_1, \ldots, r_m)^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$ is now primitive. Furthermore,

$$d\,|s + r_1\alpha_1 + \ldots + r_m\alpha_m| = |p + q_1\alpha_1 + \ldots + q_m\alpha_m| < \frac{1}{h_1 \cdots h_m},$$

meaning

$$|s + r_1\alpha_1 + \ldots + r_m\alpha_m| < \frac{1}{dh_1 \cdots h_m} \leq \frac{1}{h_1 \cdots h_m}.$$

Also $|r_i| \leq |q_i| \leq h_i$ for all $i = 1, \ldots, m$, and $r_i = \frac{1}{d}q_i \neq 0$ for some $i = 1, \ldots, m$. □

Theorem 12. *Let $1, \alpha_1, \ldots, \alpha_m \in \mathbb{R}$ be linearly independent over $\mathbb{Q}$. Then there exist infinitely many primitive integer vectors $\bar{v}_k := (p_k, q_{1,k}, \ldots, q_{m,k})^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$ satisfying the condition*

$$|p_k + q_{1,k}\alpha_1 + \ldots + q_{m,k}\alpha_m| < \frac{1}{h_{1,k} \cdots h_{m,k}} =: \frac{1}{h_k}, \tag{5.4}$$

where $h_{i,k} := \max\{1, |q_{i,k}|\}$, $i = 1, \ldots, m$.

Proof. [9, p. 7] Suppose on the contrary that there exist only finitely many primitive vectors $\bar{v}_k \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$ satisfying the inequality (5.4). As these

vectors are non-zero and the elements $1, \alpha_1, \ldots, \alpha_m$ linearly independent, a positive minimum exists: denote

$$\frac{1}{R} := \min_k |p_k + q_{1,k}\alpha_1 + \ldots + q_{m,k}\alpha_m| > 0. \tag{5.5}$$

Choose then m such numbers $\hat{h}_i \in \mathbb{Z}^+,\ i = 1, \ldots, m$, that for their product $\hat{h}_1 \cdots \hat{h}_m =: \hat{h}$ it holds $\frac{1}{\hat{h}} \leq \frac{1}{R}$. Theorem 11 gives then a primitive vector $(\hat{p}, \hat{q}_1, \ldots, \hat{q}_m) \in \mathbb{Z}^{m+1}$ with $\max\{1, |\hat{q}_i|\} \leq \hat{h}_i,\ i = 1, \ldots, m$, satisfying the condition

$$|\hat{p} + \hat{q}_1\alpha_1 + \ldots + \hat{q}_m\alpha_m| < \frac{1}{\hat{h}} \leq \frac{1}{R}.$$

This contradicts the minimum assumption (5.5). □

As a corollary of Theorem 12 we get

Theorem 13. *Let $1, \alpha_1, \ldots, \alpha_m \in \mathbb{R}$ be linearly independent over $\mathbb{Q}$. If there exist positive constants $c, \omega \in \mathbb{R}^+$ such that the inequality*

$$|\beta_0 + \beta_1\alpha_1 + \ldots + \beta_m\alpha_m| \geq \frac{c}{(h_1 \cdots h_m)^\omega} \tag{5.6}$$

holds for all vectors $(\beta_0, \beta_1, \ldots, \beta_m)^T \in \mathbb{Z}^{m+1}\backslash\{\bar{0}\}$ with $h_k := \max\{1, |\beta_k|\},\ k = 1, \ldots, m$, then $\omega \geq 1$.

Proof. [9, p. 7] Again we prove by contradiction: assume $\omega < 1$. By Theorem 12 and the assumption (5.6) there exists an infinite amount of primitive vectors $\bar{v}_k = (p_k, q_{1,k}, \ldots, q_{m,k})^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$ satisfying

$$\frac{c}{(h_{1,k} \cdots h_{m,k})^\omega} \leq |p_k + q_{1,k}\alpha_1 + \ldots + q_{m,k}\alpha_m| < \frac{1}{h_{1,k} \cdots h_{m,k}},$$

where $h_{i,k} = \max\{1, |q_{i,k}|\},\ i = 1, \ldots, m$. This results in

$$\frac{c}{(h_{1,k} \cdots h_{m,k})^\omega} \cdot \left(\frac{\frac{1}{c}}{(h_{1,k} \cdots h_{m,k})^{1-\omega}} - 1 \right) > 0,$$

and the fact that $c, h_{1,k}, \ldots, h_{m,k} > 0$ further implying

$$\frac{\frac{1}{c}}{(h_{1,k} \cdots h_{m,k})^{1-\omega}} > 1.$$

Consequently $\frac{1}{c} > (h_{1,k} \cdots h_{m,k})^{1-\omega}$, meaning

$$\log(h_{1,k} \cdots h_{m,k}) < \frac{\log \frac{1}{c}}{1-\omega}.$$

However, from the infinity of solutions such a vector $\bar{v}_k$ can be chosen that the numbers $h_{1,k}, \ldots, h_{m,k} \in \mathbb{Z}^+$ are large enough to fulfill the condition

$$\log(h_{1,k} \cdots h_{m,k}) \geq \frac{\log \frac{1}{c}}{1-\omega}.$$

Therefore we must have $\omega \geq 1$. □

The definitions 3, 4, 9, 10 and 11 from Chapters 1 and 3 should be recalled before the following theorem, which presents an inequality on the powers of an algebraic integer.

Theorem 14. *Let $\alpha \in \mathbb{B}$ with $\deg_{\mathbb{Q}} \alpha = m+1$, and denote by $\alpha^{(0)} := \alpha, \alpha^{(1)}, \ldots, \alpha^{(m)}$ its conjugates relative to the field $\mathbb{Q}(\alpha)$. Then the inequality*

$$|p + q_1\alpha + \ldots + q_m\alpha^m| > \frac{1}{(3mH)^m A^{m^2}}, \quad A := \max\left\{1, \left\lceil \alpha \right\rceil\right\},$$

holds true for all rational integer vectors $(p, q_1, \ldots, q_m)^T \in \mathbb{Z}^{m+1}$ with $1 \leq H := \max_{1 \leq i \leq m} |q_i|$.

Proof. [9, p. 8] Let $(p, q_1, \ldots, q_m)^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$. First note that $m, H, A \geq 1$. So, if $|p + q_1\alpha + \ldots + q_m\alpha^m| \geq 1$, the claim holds. Hence it suffices to study the case $|p + q_1\alpha + \ldots + q_m\alpha^m| < 1$. The triangle inequality gives

$$\begin{aligned} & |p| - |q_1\alpha + \ldots + q_m\alpha^m| \\ \leq & \big||p| - |q_1\alpha + \ldots + q_m\alpha^m|\big| \\ \leq & |p + q_1\alpha + \ldots + q_m\alpha^m| \\ < & 1, \end{aligned}$$

leading to

$$\begin{aligned}
|p| &< 1 + |q_1\alpha + \ldots + q_m\alpha^m| \\
&\leq 1 + \sum_{i=1}^{m} |q_i||\alpha|^i \\
&\leq 1 + H\sum_{i=1}^{m} |\alpha|^i \\
&\leq 1 + H\sum_{i=1}^{m} \max\left\{1, |\alpha|^i\right\} \\
&\leq 1 + mH(\max\{1, |\alpha|\})^m.
\end{aligned}$$

Then

$$\begin{aligned}
&\left|p + q_1\alpha^{(i)} + \ldots + q_m\left(\alpha^{(i)}\right)^m\right| \\
&< |p| + \left|q_1\alpha^{(i)} + \ldots + q_m\left(\alpha^{(i)}\right)^m\right| \\
&\leq 1 + mH(\max\{1, |\alpha|\})^m + mH\left(\max\left\{1, \left|\alpha^{(i)}\right|\right\}\right)^m \\
&\leq 3mH\left(\max_{0\leq i\leq m}\left\{1, \left|\alpha^{(i)}\right|\right\}\right)^m \\
&= 3mH\left(\max\left\{1, \left|\overline{\alpha}\right|\right\}\right)^m \\
&= 3mHA^m.
\end{aligned}$$

Because α is an algebraic integer of degree $\deg_{\mathbb{Q}}\alpha = m+1$, the polynomial expression $\Theta := p + q_1\alpha + \ldots + q_m\alpha^m \in \mathbb{Z}[\alpha] \setminus \{\bar{0}\}$ is a non-zero algebraic integer and therefore $N(\Theta) \in \mathbb{Z} \setminus \{0\}$. Adopting, as with α, the notation

$\Theta^{(0)} := \Theta$ we have

$$\begin{aligned} 1 &\leq |N(\Theta)| \\ &= \left| \prod_{i=0}^{m} \Theta^{(i)} \right| \\ &\leq |\Theta| \prod_{i=1}^{m} \left| p + q_1 \alpha^{(i)} + \ldots + q_m \left(\alpha^{(i)} \right)^m \right| \\ &< |\Theta| \, (3mH)^m A^{m^2}. \end{aligned}$$

□

5.2 On simultaneous Diophantine inequalities

Before concluding the chapter with a Diophantine inequality over $\mathbb{C}$, a couple of results on simultaneous Diophantine approximations are introduced.

Theorem 15. *Let the numbers $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and $f_1, \ldots, f_m \in \mathbb{R}_{\geq 0}$ with $f_1 + \ldots + f_m = 1$ be given. Then there exist $p \in \mathbb{Z}^+$ and $q_1, \ldots, q_m \in \mathbb{Z}$ satisfying the equations*

$$|p\alpha_i + q_i| < \frac{1}{p^{f_i}}, \quad i = 1, \ldots, m.$$

Proof. [9, p. 10] Let $q \in \mathbb{Z}^+$ and define a set

$$\mathcal{C} := \left\{ (x_0, x_1, \ldots, x_m)^T \in \mathbb{R}^{m+1} \;\middle|\; |x_0| \leq q + \frac{1}{2},\ |x_0\alpha_i + x_i| < \frac{1}{q^{f_i}} \right\},$$

where $i = 1, \ldots, m$. This is a central symmetric convex body with volume

$$V(\mathcal{C}) = (2q+1)\frac{2}{q^{f_1}} \cdots \frac{2}{q^{f_m}} = \frac{2^{m+1}q}{q} + \frac{2^m}{q} > 2^{m+1}.$$

Hence the first Minkowski's convex body theorem with the integer lattice $\Lambda = \mathbb{Z}^{m+1}$ gives a non-zero integer vector $\bar{0} \neq (p, q_1, \ldots, q_m)^T \in \mathcal{C} \cap \Lambda$.

Here $p \neq 0$, for otherwise the equation $|p\alpha_i + q_i| < \frac{1}{q^{f_i}}$ would imply $q_1 = \ldots = q_m = 0$ as well. Because the set $\mathcal{C}$ is central symmetric, we have $(-p, -q_1, \ldots, -q_m)^T \in \mathcal{C} \cap \Lambda$ too, so we can choose p to be positive. Since now $p \leq q$, we have also

$$|p\alpha_i + q_i| < \frac{1}{q^{f_i}} \leq \frac{1}{p^{f_i}}, \quad i = 1, \ldots, m,$$

as was to be proved. □

Remark 11. If $(p, q_1, \ldots, q_m)^T = d(s, r_1, \ldots, r_m)^T \in \mathbb{Z}^{m+1}$ is the integer vector given by Theorem 15 with $d := \gcd(p, q_1, \ldots, q_m)$, then

$$|s\alpha_i + r_i| \leq |p\alpha_i + q_i| < \frac{1}{p^{f_i}} \leq \frac{1}{s^{f_i}}, \quad i = 1, \ldots, m,$$

so there exists a primitive solution as well.

Theorem 16. *Let at least one of the numbers $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ be irrational. Then there exist infinitely many primitive vectors $\bar{v}_k := (p_k, q_{1,k}, \ldots, q_{m,k})^T \in \mathbb{Z}^{m+1} \setminus \{\bar{0}\}$ with $p_k \in \mathbb{Z}^+$ satisfying the conditions*

$$|p_k\alpha_i + q_{i,k}| < \frac{1}{p_k^{\frac{1}{m}}}, \quad i = 1, \ldots, m.$$

Proof. We proceed as in the proof of Theorem 12, and suppose that there are only finitely many primitive solutions $\bar{v}_k$. Since these are non-zero and some number α_j is irrational, there is a positive minimum

$$\frac{1}{R} := \min_k |p_k\alpha_j + q_{j,k}| > 0.$$

Choose then $q \geq R^m$ and $f_i = \frac{1}{m}$, $i = 1, \ldots, m$, in the proof of Theorem 15. It follows that there exists a primitive vector $\bar{0} \neq (\hat{p}, \hat{q}_1, \ldots, \hat{q}_m)^T \in \mathbb{Z}^{m+1}$, $\hat{p} \in \mathbb{Z}^+$, such that

$$|\hat{p}\alpha_i + \hat{q}_i| < \frac{1}{q^{\frac{1}{m}}} \leq \frac{1}{\hat{p}^{\frac{1}{m}}}, \quad i = 1, \ldots, m.$$

In particular,

$$|\hat{p}\alpha_j + \hat{q}_j| < \frac{1}{q^{\frac{1}{m}}} \leq \frac{1}{R},$$

which is a contradiction. □

5.3 A Diophantine inequality over $\mathbb{C}$

Theorem 17. *Let $\mathbb{I} = \mathbb{Q}(\sqrt{-D})$, where $D \in \mathbb{Z}^+$, denote an imaginary quadratic field, and $\mathbb{Z}_\mathbb{I}$ its ring of integers. Let the numbers $\Theta_1, \ldots, \Theta_m \in \mathbb{C}$ and $H_1, \ldots, H_m \in \mathbb{Z}^+$ be given. Then there exists a non-zero integer vector $(\beta_0, \beta_1, \ldots, \beta_m)^T \in \mathbb{Z}_\mathbb{I}^{m+1} \setminus \{\bar{0}\}$ with $|\beta_j| \leq H_j,\ j = 1, \ldots, m$, satisfying the inequality*

$$|\beta_0 + \beta_1\Theta_1 + \ldots + \beta_m\Theta_m| \leq \left(\frac{2^\tau D^{\frac{1}{4}}}{\sqrt{\pi}}\right)^{m+1} \frac{1}{H_1 \cdots H_m},$$

where

$$\tau = \begin{cases} 1, & D \equiv 1 \text{ or } 2 \pmod 4 \\ \frac{1}{2}, & D \equiv 3 \pmod 4 \end{cases}.$$

Proof. [11, p. 8] The ring of integers $\mathbb{Z}_\mathbb{I}$ of the field $\mathbb{I}$, given by Theorem 3, can be written in the form $\mathbb{Z}_\mathbb{I} = \mathbb{Z} + \mathbb{Z}\big(h + l\sqrt{-D}\big)$, where

$$h = \begin{cases} 0, & D \equiv 1 \text{ or } 2 \pmod 4 \\ \frac{1}{2}, & D \equiv 3 \pmod 4 \end{cases} \quad \text{and} \quad l = \begin{cases} 1, & D \equiv 1 \text{ or } 2 \pmod 4 \\ \frac{1}{2}, & D \equiv 3 \pmod 4 \end{cases}.$$

First define a full lattice $\lambda = \mathbb{Z}(1,0) + \mathbb{Z}\Big(h, l\sqrt{D}\Big) \subseteq \mathbb{R}^2$, which has determinant

$$\det\lambda = \begin{vmatrix} 1 & h \\ 0 & l\sqrt{D} \end{vmatrix} = l\sqrt{D} = \sqrt{D}2^{-2h}. \tag{5.7}$$

Define also a complex disc

$$\mathcal{D}_R := \Big\{ x + y\Big(h + l\sqrt{-D}\Big) \in \mathbb{C} \;\Big|\; x, y \in \mathbb{R}, \Big|x + y\Big(h + l\sqrt{-D}\Big)\Big| \leq R \Big\}$$

of radius $R > 0$, and a corresponding real disc

$$\mathcal{C}_R := \left\{ (v, w)^T \in \mathbb{R}^2 \;\middle|\; v^2 + w^2 \leq R^2 \right\}.$$

Here $V(\mathcal{C}_R) = \pi R^2$. Now $x + y\big(h + l\sqrt{-D}\big) \in \mathcal{D}_R \cap \mathbb{Z}_{\mathbb{I}}$ means that $x, y \in \mathbb{Z}$ and $(x + yh)^2 + \left(yl\sqrt{D}\right)^2 \leq R^2$, implying $\left(x + yh, yl\sqrt{D}\right)^T \in \mathcal{C}_R \cap \lambda$. The other way round, $\left(x + yh, yl\sqrt{D}\right)^T \in \mathcal{C}_R \cap \lambda$ means exactly that $x + y\big(h + l\sqrt{-D}\big) \in \mathcal{D}_R \cap \mathbb{Z}_{\mathbb{I}}$. In this way the intersection $\mathcal{C}_R \cap \lambda$ is a real counterpart for the set $\mathcal{D}_R \cap \mathbb{Z}_{\mathbb{I}}$.

Next define a lattice $\Lambda := \lambda^{m+1} \subseteq \mathbb{R}^{2m+2}$. Recalling Lemma 6, it has determinant $\det \Lambda = \left(\sqrt{D} 2^{-2h}\right)^{m+1}$. By using the following notations

$$\begin{cases} a + b\big(h + l\sqrt{-D}\big) = -(z_1\Theta_1 + \ldots + z_m\Theta_m),\ a, b \in \mathbb{Z} \\ z_k = x_k + y_k\big(h + l\sqrt{-D}\big)\,,\ x_k, y_k \in \mathbb{R},\ k = 0, 1, \ldots, m, \\ v_k = x_k + y_k h = \mathrm{Re}(z_k), \\ w_k = y_k l\sqrt{D} = \mathrm{Im}(z_k), \\ R_0 := \left(\frac{2^\tau D^{\frac{1}{4}}}{\sqrt{\pi}}\right)^{m+1} \frac{1}{H_1 \cdots H_m}, \\ \tau = \begin{cases} 1, & D \equiv 1 \text{ or } 2 \pmod 4 \\ \frac{1}{2}, & D \equiv 3 \pmod 4 \end{cases} \end{cases}$$

two more sets are defined:

$$\mathcal{D} := \left\{ (z_0, z_1, \ldots, z_m)^T \in \mathbb{C}^{m+1} \;\middle|\; \left| z_0 - \left(a + b\left(h + l\sqrt{-D}\right)\right)\right| \leq R_0, \right.$$
$$\left. |z_k| \leq H_k,\ k = 1, \ldots, m \right\}$$

and

$$\mathcal{C} := \left\{ (v_0, w_0, v_1, w_1, \ldots, v_m, w_m)^T \in \mathbb{R}^{2m+2} \;\middle|\; \right.$$
$$\left. (v_0 - (a + bh))^2 + \left(w_0 - bl\sqrt{D}\right)^2 \leq R_0^2,\ v_k^2 + w_k^2 \leq H_k^2,\ k = 1, \ldots, m \right\}.$$

The set $\mathcal{C}$ is a compact convex body, and if the coordinate axes are shifted so that the origin lies at the point $(a+bh, bl\sqrt{D}, 0, \ldots, 0)^T \in \Lambda$ (this doesn't change the lattice), it is also central symmetric. The volume of the set $\mathcal{C}$ is by definition

$$\begin{aligned}
V(\mathcal{C}) &= \iint \cdots \iint \left(\iint_{(v_0-(a+bh))^2+\left(w_0-bl\sqrt{D}\right)^2 \le R_0^2} \mathrm{d}v_0 \mathrm{d}w_0 \right) \mathrm{d}v_1 \mathrm{d}w_1 \cdots \mathrm{d}v_m \mathrm{d}w_m \\
&= \pi R_0^2 \iint \cdots \iint \left(\iint_{v_1^2+w_1^2 \le H_1^2} \mathrm{d}v_1 \mathrm{d}w_1 \right) \mathrm{d}v_2 \mathrm{d}w_2 \cdots \mathrm{d}v_m \mathrm{d}w_m \\
&= \ldots \\
&= \pi^{m+1} H_1^2 \cdots H_m^2 R_0^2 \\
&= \pi^{m+1} H_1^2 \cdots H_m^2 \left(\frac{2^{2\tau}\sqrt{D}}{\pi} \right)^{m+1} \frac{1}{H_1^2 \cdots H_m^2} \\
&= \left(2^{2m+2}\right)^{\tau} \left(\sqrt{D}\right)^{m+1} \\
&= 2^{2m+2} \left(\frac{\sqrt{D}}{2^{2h}} \right)^{m+1} \\
&= 2^{2m+2} \det \Lambda.
\end{aligned}$$

The neat result explains the seemingly random definition of the radius R_0. Pay also attention to how the volume, like the determinant in (5.7), can be expressed in terms of h due to the way the numbers h, l and τ are defined. Now, by the first Minkowski's convex body theorem, there exists a non-zero lattice point

$$\left(x_0 + y_0 h, y_0 l\sqrt{D}, \ldots, x_m + y_m h, y_m l\sqrt{D}\right)^T \in \mathcal{C} \cap \Lambda \setminus \{\bar{0}\}.$$

The correspondence between the sets $\mathcal{C}_R \cap \lambda$ and $\mathcal{D}_R \cap \mathbb{Z}_{\mathbb{I}}$ implies then the

existence of a non-zero integer vector

$$\begin{aligned}&(\beta_0, \beta_1, \ldots, \beta_m)^T \\ &:= \left(x_0 + y_0\left(h + l\sqrt{-D}\right), \ldots, x_m + y_m\left(h + l\sqrt{-D}\right)\right)^T \in \mathcal{D} \cap \mathbb{Z}_{\mathbb{I}}^{m+1} \setminus \{\bar{0}\},\end{aligned}$$

with $|\beta_k| \leq H_k,\ k = 1, \ldots, m$, satisfying the condition

$$|\beta_0 + \beta_1\Theta_1 + \ldots + \beta_m\Theta_m| \leq \left(\frac{2^\tau D^{\frac{1}{4}}}{\sqrt{\pi}}\right)^{m+1} \frac{1}{H_1 \cdots H_m}.$$

□

Chapter 6

Applying Siegel's lemma

Let again $\mathbb{I}$ stand for an imaginary quadratic field or the field $\mathbb{Q}$ of rational numbers, and denote with $\mathbb{Z}_{\mathbb{I}}$ its ring of integers.

6.1 Approximations for $e^{\alpha t}$

In the article [4] the writers study linear forms of exponential values, and the following theorem is one of the lemmas they need in order to prove a lower bound for the expression $|\beta_0 + \beta_1 e^{\alpha_1} + \ldots + \beta_m e^{\alpha_m}|$, where $(\beta_0, \beta_1, \ldots, \beta_m)^T \in \mathbb{Z}_{\mathbb{I}}^{m+1}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{I}$. Here it serves as an enlightening example on how Siegel's lemma (Theorem 6) can be used.

Definition 20. The *order* of a power series $P(x) = \sum_{k=0}^{\infty} a_k (x-c)^k$ is $\operatorname{ord}_{x=c} P(x) := \inf \{k \in \mathbb{N} \mid a_k \neq 0\}$.

Theorem 18. *Let $\alpha_j = \frac{x_j}{y_j} \in \mathbb{I} \setminus \{0\}$ be m different numbers with $x_j \in \mathbb{Z}_{\mathbb{I}} \setminus \{0\}$, $y_j \in \mathbb{Z}^+$ and $\gcd(x_j, y_j) = 1$ for $j = 1, \ldots, m$. Let also $\nu_1, \ldots, \nu_m \in \mathbb{Z}^+$ and $l_1, \ldots, l_m \in \mathbb{Z}^+$ be such that $1 \leq \nu_j \leq l_j$ for $j = 1, \ldots, m$, and denote $M := \nu_1 + \ldots + \nu_m \leq L := l_1 + \ldots + l_m$. Write $a := \max_{1 \leq j \leq m} \{|x_j| + |y_j|\}$*

and $b := \max_{1\le j\le m}\left\{1+\frac{|x_j|}{|y_j|}\right\}$, *as well as* $\bar{\alpha} := (\alpha_1,\dots,\alpha_m)^T \in \mathbb{I}^m$ *and* $\bar{\nu} := (\nu_1,\dots,\nu_m)^T \in \mathbb{Z}^m$. *Then there exists a non-zero polynomial*

$$P_{0,0}(t) = \sum_{h=0}^{L} c_h \frac{L!}{h!} t^h \in \mathbb{Z}_{\mathbb{I}}[t]$$

depending on L, *with*

$$|c_h| \le \max\left\{4\sqrt{2}\sqrt{D}, s_{\mathbb{I}} t_{\mathbb{I}}^{\frac{M}{L+1-M}} \left(a^{ML} b^{\frac{M^2}{2}}\right)^{\frac{1}{L+1-M}}\right\}, \quad h = 0,1,\dots,L. \quad (6.1)$$

Furthermore, there exist non-zero polynomials $P_{0,j}(t) \in \mathbb{Z}_{\mathbb{I}}[t,\bar{\alpha}], j = 1,\dots,m$, *depending on* L *and* $\bar{\nu}$, *such that*

$$P_{0,0}(t)e^{\alpha_j t} - P_{0,j}(t) = R_{0,j}(t), \quad j = 1,\dots,m,$$

where

$$\begin{cases} \deg_t P_{0,j}(t) \le L, \quad j = 0,1,\dots,m, \\ L + \nu_j + 1 \le \operatorname*{ord}_{t=0} R_{0,j}(t) < \infty, \quad j = 1,\dots,m. \end{cases}$$

Proof. [4, p. 11] Let

$$P_{0,0}(t)e^{\alpha_j t} = \sum_{N=0}^{\infty} r_{N,j} t^N, \quad j = 1,\dots,m,$$

where

$$r_{N,j} := \sum_{\substack{h+n=N \\ 0\le h\le L}} c_h \frac{L!}{h!}\frac{\alpha_j^n}{n!}.$$

Cut this series after $L+1$ terms and let

$$P_{0,j}(t) := \sum_{N=0}^{L} r_{N,j} t^N, \quad j = 1,\dots,m.$$

Set $r_{L+i_j,j} = 0$ for $i_j = 1, \ldots, \nu_j, j = 1, \ldots, m$. Then also

$$\begin{aligned}
0 &= \frac{(L+i_j)!}{L!} \frac{y_j^{L+i_j}}{x_j^{i_j}} r_{L+i_j,j} \\
&= \frac{(L+i_j)!}{L!} \frac{y_j^{L+i_j}}{x_j^{i_j}} \sum_{\substack{h+n=L+i_j \\ 0 \le h \le L}} c_h \frac{L!}{h!n!} \frac{x_j^n}{y_j^n} \\
&= \sum_{\substack{h+n=L+i_j \\ 0 \le h \le L}} \frac{(L+i_j)!}{h!n!} x_j^{n-i_j} y_j^{L+i_j-n} c_h \qquad (6.2) \\
&= \sum_{h=0}^{L} \frac{(L+i_j)!}{h!(L+i_j-h)!} x_j^{L-h} y_j^h c_h \\
&= \sum_{h=0}^{L} \binom{L+i_j}{h} x_j^{L-h} y_j^h c_h
\end{aligned}$$

for $i_j = 1, \ldots, \nu_j$, $j = 1, \ldots, m$, meaning that we have M equations in $L+1$ unknowns c_h, $h = 0, 1, \ldots, L$, with coefficients $\binom{L+i_j}{h} x_j^{L-h} y_j^h \in \mathbb{Z}_{\mathbb{I}}$. These coefficients satisfy

$$\begin{aligned}
A_{j,i_j} &:= \sum_{h=0}^{L} \left| \binom{L+i_j}{h} x_j^{L-h} y_j^h \right| \\
&= \frac{1}{|x_j|^{i_j}} \sum_{h=0}^{L} \binom{L+i_j}{h} |x_j|^{L+i_j-h} |y_j|^h \\
&< \frac{1}{|x_j|^{i_j}} \sum_{h=0}^{L+i_j} \binom{L+i_j}{h} |x_j|^{L+i_j-h} |y_j|^h \\
&= \frac{1}{|x_j|^{i_j}} (|x_j| + |y_j|)^{L+i_j} \\
&= (|x_j| + |y_j|)^L \left(1 + \frac{|y_j|}{|x_j|}\right)^{i_j},
\end{aligned}$$

when $i_j = 1, \ldots, \nu_j, j = 1, \ldots, m$. (Recall that in order to apply Theorem 6,

the product of the row sums (2.6) is needed.) Then

$$\begin{aligned}\prod_{j,i_j} A_{j,i_j} &< \prod_{j,i_j} (|x_j| + |y_j|)^L \prod_{j,i_j} \left(1 + \frac{|y_j|}{|x_j|}\right)^{i_j} \\ &= \prod_{j} (|x_j| + |y_j|)^{\nu_j L} \prod_{j} \left(1 + \frac{|y_j|}{|x_j|}\right)^{\frac{\nu_j(\nu_j+1)}{2}} \\ &\leq a^{ML} b^{\frac{M^2}{2}}.\end{aligned}$$

The last step follows by employing the definition $M = \nu_1 + \ldots + \nu_m$, which implies $\sum_{j=1}^m \nu_j(\nu_j+1) = \sum_{j=1}^m \nu_j^2 + \sum_{j=1}^m \nu_j \leq M^2$. Thus by Theorem 6 there exists a solution $(c_0, c_1, \ldots, c_L)^T \in \mathbb{Z}_{\mathbb{I}}^{L+1} \setminus \{\bar{0}\}$ to the group of M equations derived in (6.2) with

$$|c_h| \leq \max\left\{4\sqrt{2}\sqrt{D}, s_{\mathbb{I}} t_{\mathbb{I}}^{\frac{M}{L+1-M}} \left(a^{ML} b^{\frac{M^2}{2}}\right)^{\frac{1}{L+1-M}}\right\}, \quad h = 0, 1, \ldots, L.$$

Writing

$$R_{0,j}(t) := \sum_{N=L+\nu_j+1}^{\infty} r_{N,j} t^N$$

we get

$$P_{0,0}(t) e^{\alpha_j t} - P_{0,j}(t) = R_{0,j}(t), \quad j = 1, \ldots, m.$$

Here $P_{0,j}(t)$ are non-zero polynomials for all $j = 0, 1, \ldots, m$, since the solution $(c_0, c_1, \ldots, c_L)^T$ is a non-zero vector. In addition

$$\begin{cases} \deg_t P_{0,j}(t) \leq L, \quad j = 0, 1, \ldots, m, \\ L + \nu_j + 1 \leq \operatorname*{ord}_{t=0} R_{0,j}(t) < \infty, \quad j = 1, \ldots, m, \end{cases}$$

as the series $R_{0,j}(t)$ is non-zero as well. □

6.2 Siegel's lemma refined

Suppose now that $\mathbb{I} = \mathbb{Q}$ and $\alpha_j \in \mathbb{Z} \setminus \{0\}$ $(x_j \in \mathbb{Z} \setminus \{0\},\, y_j = 1)\,, j = 1, \ldots, m$, $m \geq 2$, are m different rational integers. Then the equations

$$\sum_{h=0}^{L} \binom{L+i_j}{h} x_j^{L-h} y_j^h c_h = 0, \quad i_j = 1, \ldots, \nu_j,\ j = 1, \ldots, m,$$

derived in (6.2) can be written in matrix form:

$$\begin{pmatrix} \binom{L+1}{0}x_1^L & \binom{L+1}{1}x_1^{L-1} & \cdots & \binom{L+1}{L-1}x_1 & \binom{L+1}{L} \\ \binom{L+2}{0}x_1^L & \binom{L+2}{1}x_1^{L-1} & \cdots & \binom{L+2}{L-1}x_1 & \binom{L+2}{L} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \binom{L+\nu_1}{0}x_1^L & \binom{L+\nu_1}{1}x_1^{L-1} & \cdots & \binom{L+\nu_1}{L-1}x_1 & \binom{L+\nu_1}{L} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \binom{L+\nu_m}{0}x_m^L & \binom{L+\nu_m}{1}x_m^{L-1} & \cdots & \binom{L+\nu_m}{L-1}x_m & \binom{L+\nu_m}{L} \end{pmatrix}_{M\times(L+1)} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{L-1} \\ c_L \end{pmatrix} = \bar{0}. \tag{6.3}$$

Denote with B the above matrix of coefficients.

The interest in this matrix representation is explained when looking at the following improved version of Siegel's lemma, proved by Bombieri and Vaaler in 1983:

Theorem 19. *Let $A := [a_{mn}] \in \mathcal{M}_{M\times N}(\mathbb{Z})$ with $\operatorname{rank} A = M < N$. Then the equation $A\bar{x} = \bar{0}$ has $N - M$ linearly independent integer solutions*

$$\bar{x}_1, \ldots, \bar{x}_{N-M} \in \mathbb{Z}^N \setminus \{\bar{0}\} \tag{6.4}$$

satisfying

$$\prod_{k=1}^{N-M} \|\bar{x}_k\|_\infty \leq \frac{\sqrt{\det(AA^T)}}{D},$$

where D is the greatest common divisor of all $M \times M$ minors of the matrix A.

Proof. The proof, which is beyond the scope of this work, can be found in the article [2]. □

Write

$$A = [a_{mn}] = \begin{bmatrix} \bar{a}_1^T \\ \vdots \\ \bar{a}_M^T \end{bmatrix},$$

where $\bar{a}_m^T := (a_{m1}, \ldots, a_{mN})$ denotes the mth row of A. The *Gram determinant* $\det(AA^T)$ [1] satisfies

$$\begin{aligned}
\sqrt{\det(AA^T)} &= \sqrt{\det\left[\bar{a}_l \cdot \bar{a}_m\right]_{1\le l,m\le M}} \\
&\le \sqrt{\prod_{m=1}^{M} \bar{a}_m \cdot \bar{a}_m} \\
&= \prod_{m=1}^{M} \|\bar{a}_m\| \\
&\le \prod_{m=1}^{M} \|\bar{a}_m\|_1 \\
&\le \prod_{m=1}^{M} N \max_{1\le n\le N} |a_{mn}| \\
&\le \left(N \max_{1\le m,n\le N} |a_{mn}| \right)^M .
\end{aligned}$$

For the solutions (6.4) we can write $\|\bar{x}_1\|_\infty \le \|\bar{x}_2\|_\infty \le \ldots \le \|\bar{x}_{N-M}\|_\infty$.

[1] The Gram determinant is always non-negative, and since $\operatorname{rank} A = M$, in this case it is strictly positive. We have also the so-called *Hadamard inequality*

$$\det\left[\bar{a}_l \cdot \bar{a}_m\right]_{1\le l,m\le M} \le \prod_{m=1}^{M} \bar{a}_m \cdot \bar{a}_m.$$

Then

$$\|\bar{x}_1\|_\infty \le \left(\prod_{k=1}^{N-M} \|\bar{x}_k\|_\infty\right)^{\frac{1}{N-M}} \le \left(\frac{\sqrt{\det(AA^T)}}{D}\right)^{\frac{1}{N-M}}$$
$$\le \frac{1}{D^{\frac{1}{N-M}}} \left(N \max_{1\le m,n\le N} |a_{mn}|\right)^{\frac{M}{N-M}}.$$

In view of Remark 1, this means that in the rational case the estimate (6.1) could possibly be sharpened if the number D for the matrix B in (6.3) happened to be some other than one.

Example. Take $m = M = L = 2$ in (6.3), whereupon

$$B = \begin{pmatrix} \binom{3}{0}x_1^2 & \binom{3}{1}x_1 & \binom{3}{2} \\ \binom{3}{0}x_2^2 & \binom{3}{1}x_2 & \binom{3}{2} \end{pmatrix} = \begin{pmatrix} x_1^2 & 3x_1 & 3 \\ x_2^2 & 3x_2 & 3 \end{pmatrix},$$

and its 2×2 minors are

$$\begin{vmatrix} x_1^2 & 3x_1 \\ x_2^2 & 3x_2 \end{vmatrix} = 3x_1^2x_2 - 3x_1x_2^2 = 3x_1x_2(x_1 - x_2),$$
$$\begin{vmatrix} 3x_1 & 3 \\ 3x_2 & 3 \end{vmatrix} = 9x_1 - 9x_2 = 9(x_1 - x_2),$$
$$\begin{vmatrix} x_1^2 & 3 \\ x_2^2 & 3 \end{vmatrix} = 3x_1^2 - 3x_2^2 = 3(x_1 + x_2)(x_1 - x_2).$$

The greatest common divisor of these appears to be $D = 3|x_1 - x_2| > 1$.

Let now

$$0 \le k_1 < k_2 < \ldots < k_M \le L \tag{6.5}$$

and consider an arbitrary $M \times M$ minor

$$\begin{vmatrix} \binom{L+1}{k_1} x_1^{L-k_1} & \binom{L+1}{k_2} x_1^{L-k_2} & \cdots & \binom{L+1}{k_M} x_1^{L-k_M} \\ \binom{L+2}{k_1} x_1^{L-k_1} & \binom{L+2}{k_2} x_1^{L-k_2} & \cdots & \binom{L+2}{k_M} x_1^{L-k_M} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{L+\nu_1}{k_1} x_1^{L-k_1} & \binom{L+\nu_1}{k_2} x_1^{L-k_2} & \cdots & \binom{L+\nu_1}{k_M} x_1^{L-k_M} \\ \binom{L+1}{k_1} x_2^{L-k_1} & \binom{L+1}{k_2} x_2^{L-k_2} & \cdots & \binom{L+1}{k_M} x_2^{L-k_M} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{L+\nu_m}{k_1} x_m^{L-k_1} & \binom{L+\nu_m}{k_2} x_m^{L-k_2} & \cdots & \binom{L+\nu_m}{k_M} x_m^{L-k_M} \end{vmatrix} . \tag{6.6}$$

Interchange then the second and the (ν_1+1)th row, so that the first two rows of the minor are

$$\begin{pmatrix} \binom{L+1}{k_1} x_1^{L-k_1} & \binom{L+1}{k_2} x_1^{L-k_2} & \cdots & \binom{L+1}{k_M} x_1^{L-k_M} \\ \binom{L+1}{k_1} x_2^{L-k_1} & \binom{L+1}{k_2} x_2^{L-k_2} & \cdots & \binom{L+1}{k_M} x_2^{L-k_M} \end{pmatrix} . \tag{6.7}$$

(This of course only affects the sign of the minor.) Expanding this permutated determinant with the Laplace formula, each time with respect to the bottom row, means that eventually we are calculating 2×2 minors of the first two rows (6.7). Let $1 \le r < s \le M$. Then $0 \le k_r < k_s \le L$ by (6.5) and

$$\begin{aligned} &\begin{vmatrix} \binom{L+1}{k_r} x_1^{L-k_r} & \binom{L+1}{k_s} x_1^{L-k_s} \\ \binom{L+1}{k_r} x_2^{L-k_r} & \binom{L+1}{k_s} x_2^{L-k_s} \end{vmatrix} \\ &= \binom{L+1}{k_r}\binom{L+1}{k_s} x_1^{L-k_r} x_2^{L-k_s} - \binom{L+1}{k_r}\binom{L+1}{k_s} x_1^{L-k_s} x_2^{L-k_r} \\ &= \binom{L+1}{k_r}\binom{L+1}{k_s} x_1^{L-k_s} x_2^{L-k_s} \left(x_1^{k_s-k_r} - x_2^{k_s-k_r} \right) . \end{aligned}$$

Since $(x_1 - x_2)$ divides $\left(x_1^{k_s-k_r} - x_2^{k_s-k_r}\right)$, it is a common factor of all the 2×2 minors of the first two rows. Therefore $(x_1 - x_2)$ is a factor of the minor (6.6), implying $D \ge |x_1 - x_2|$.

Even more: Because any two rows of the determinant (6.6) can be changed to the top before expanding, actually we have $|x_i - x_j| \big| D$ for all $i, j =$

$1, \ldots, m$, $i \neq j$. Notably,

$$D \geq \max_{1 \leq i,j \leq m} |x_i - x_j|.$$

This raises the question of whether one could get the whole product of these differences out of the minor (6.6). Because the determinant essentially consists of decreasing powers of some integers, we shall next explore something similar.

6.3 Vandermonde determinants

The Vandermonde determinant formula, about to be proved in Lemma 10, is well worth studying as it so closely resembles our case. A property of Vandermonde matrices is needed first, though.

Lemma 9. *Let $a_0, a_1, \ldots, a_n \in \mathbb{C}$, $n \in \mathbb{Z}^+$, be $n+1$ distinct numbers. Then the* Vandermonde matrix

$$\left[a_i^j\right] = \begin{bmatrix} 1 & a_0 & \cdots & a_0^n \\ 1 & a_1 & \cdots & a_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^n \end{bmatrix}$$

is non-singular.

Proof. Denote $A := \left[a_i^j\right]$. Following the idea of Example 3 in [6, p. 84], we show that the only solution to the equation $A\bar{c} = \bar{0}$ is $\bar{c} = \bar{0} \in \mathbb{C}^{n+1}$. Let $\bar{c} = (c_0, c_1, \ldots, c_n)^T \in \mathbb{C}^{n+1}$ and $q(x) = c_0 + c_1 x + \ldots + c_n x^n \in \mathbb{C}[x]$. The equation $A\bar{c} = \bar{0}$ implies precisely that $q(a_0) = q(a_1) = \ldots = q(a_n) = 0$; in other words, the polynomial $q(x)$ has $n+1$ distinct zeros. But $\deg q \leq n$, so it must be that $q(x) = 0(x)$. Then $\bar{c} = \bar{0}$, which means that $\det A \neq 0$. $\square$

Lemma 10 (Vandermonde determinant)**.** *Let* $a_0, a_1, \ldots, a_n \in \mathbb{C}$, $n \in \mathbb{Z}^+$. *Then*

$$\det_{0\leq i,j\leq n}\left[a_i^j\right] = \begin{vmatrix} 1 & a_0 & \cdots & a_0^n \\ 1 & a_1 & \cdots & a_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^n \end{vmatrix} = \prod_{0\leq i<j\leq n}(a_j - a_i).$$

Proof. [7, p. 5] [14, p. 44] If $a_i = a_j$ for some $i \neq j$, it follows that

$$\det_{0\leq i,j\leq n}\left[a_i^j\right] = 0$$

and the claim holds. Suppose then that the numbers $a_0, a_1, \ldots, a_n$ are distinct. Denote $\det_{0\leq i,j\leq n}\left[a_i^j\right] =: \Delta(a_0, a_1, \ldots, a_{n-1}, a_n)$. Then the expression $\Delta(a_0, a_1, \ldots, a_{n-1}, x)$ is a polynomial in x of degree n with the leading coefficient

$$\begin{vmatrix} 1 & a_0 & \cdots & a_0^{n-1} \\ 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{n-1} & \cdots & a_{n-1}^{n-1} \end{vmatrix} = \Delta(a_0, a_1, \ldots, a_{n-1}) \neq 0.$$

This polynomial has the distinct zeros $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$, since inserting $x = a_i$ for any $i = 0, 1, \ldots, n-1$ results in two same rows in the determinant $\Delta(a_0, \ldots, a_{n-1}, x)$. Hence the polynomial $\Delta(a_0, a_1, \ldots, a_{n-1}, x)$ factorizes as

$$\Delta(a_0, a_1, \ldots, a_{n-1}, x) = \Delta(a_0, a_1, \ldots, a_{n-1}) \cdot (x - a_{n-1}) \cdots (x - a_1)(x - a_0).$$

So it follows inductively (keeping Lemma 9 in mind)

$$\begin{aligned} &\Delta(a_0, a_1, \ldots, a_{n-1}, a_n) \\ &= \Delta(a_0, a_1, \ldots, a_{n-1})(a_n - a_{n-1}) \cdots (a_n - a_1)(a_n - a_0) \\ &= \ldots \\ &= \prod_{0\leq i<j\leq n}(a_j - a_i). \end{aligned}$$

□

Now rearrange the minor (6.6) in such a way that the first m rows are

$$\begin{pmatrix} \binom{L+1}{k_1}x_1^{L-k_1} & \binom{L+1}{k_2}x_1^{L-k_2} & \dots & \binom{L+1}{k_M}x_1^{L-k_M} \\ \binom{L+1}{k_1}x_2^{L-k_1} & \binom{L+1}{k_2}x_2^{L-k_2} & \dots & \binom{L+1}{k_M}x_2^{L-k_M} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{L+1}{k_1}x_m^{L-k_1} & \binom{L+1}{k_2}x_m^{L-k_2} & \dots & \binom{L+1}{k_M}x_m^{L-k_M} \end{pmatrix}_{m\times M}. \tag{6.8}$$

Let $1 \le q_1 < q_2 < \ldots < q_m \le M$, whereupon $0 \le k_{q_1} < k_{q_2} < \ldots < k_{q_m} \le L$ by (6.5). Examine an arbitrary $m \times m$ minor of the first m rows in (6.8) (by definition, $M = \nu_1 + \ldots + \nu_m \ge m$):

$$\begin{aligned} &\begin{vmatrix} \binom{L+1}{k_{q_1}}x_1^{L-k_{q_1}} & \binom{L+1}{k_{q_2}}x_1^{L-k_{q_2}} & \dots & \binom{L+1}{k_{q_m}}x_1^{L-k_{q_m}} \\ \binom{L+1}{k_{q_1}}x_2^{L-k_{q_1}} & \binom{L+1}{k_{q_2}}x_2^{L-k_{q_2}} & \dots & \binom{L+1}{k_{q_m}}x_2^{L-k_{q_m}} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{L+1}{k_{q_1}}x_m^{L-k_{q_1}} & \binom{L+1}{k_{q_2}}x_m^{L-k_{q_2}} & \dots & \binom{L+1}{k_{q_m}}x_m^{L-k_{q_m}} \end{vmatrix}_{m\times m} \\ &= \underbrace{\begin{vmatrix} x_1^{L-k_{q_1}} & x_1^{L-k_{q_2}} & \dots & x_1^{L-k_{q_m}} \\ x_2^{L-k_{q_1}} & x_2^{L-k_{q_2}} & \dots & x_2^{L-k_{q_m}} \\ \vdots & \vdots & \ddots & \vdots \\ x_m^{L-k_{q_1}} & x_m^{L-k_{q_2}} & \dots & x_m^{L-k_{q_m}} \end{vmatrix}}_{=:\Delta} \prod_{i=1}^{m} \binom{L+1}{k_{q_i}} \\ &= \Delta \prod_{i=1}^{m} \binom{L+1}{k_{q_i}}. \end{aligned} \tag{6.9}$$

Denote $x := (x_1, x_2, \ldots, x_m)$. Then the determinant $\Delta = \Delta(x) \in \mathbb{Z}[x]$ is an alternating polynomial in m variables $x_1, \ldots, x_m$, for it changes sign when interchanging any two of them.

Lemma 11. *Every alternating polynomial $A(x) \in \mathbb{Q}[x]$, $x = (x_1, x_2, \ldots, x_m)$, is a product of some symmetric polynomial* [2] *and the Vandermonde polyno-*

[2] A *symmetric polynomial* in the variables $x_1, x_2, \ldots, x_m$ stays invariant under permutation of these variables.

mial

$$V(x) := \prod_{1 \le i < j \le m} (x_j - x_i).$$

Proof. [12, p. 28] By definition,

$$A(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) = -A(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m), \qquad (6.10)$$

when $i, j \in \{1, \ldots, m\}$, $i \neq j$,. Let $c \cdot a x_i^s x_j^t$ be any term of the polynomial $A(x)$. Here $c \in \mathbb{Q}$, $s, t \in \mathbb{N}$ and a denotes the product of the powers of the other variables x_k, $k \neq i, j$. Without loss of generality we may assume that $s \geq t$. From (6.10) it follows that

$$\begin{aligned}
&2A(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) \\
&= A(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) - A(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m) \\
&= \ldots + c \cdot a x_i^s x_j^t - c \cdot a x_i^t x_j^s + \ldots \\
&= \ldots + c \cdot a x_i^t x_j^t \left(x_i^{s-t} - x_j^{s-t}\right) + \ldots
\end{aligned}$$

Now $(x_i - x_j) \big| \left(x_i^{s-t} - x_j^{s-t}\right)$. Since the chosen term $c \cdot a x_i^s x_j^t$ was an arbitrary one, we have $(x_i - x_j) | 2A(x)$ and so $(x_i - x_j) | A(x)$ in the ring $\mathbb{Q}[x]$. Furthermore, $V(x) | A(x)$ as the factors of the Vandermonde polynomial are coprime in the ring $\mathbb{Q}[x]$, which is a UFD. The quotient $\frac{A(x)}{V(x)} \in \mathbb{Q}[x]$ is symmetric since interchanging any two variables changes the signs of both the numerator and the denominator. □

Thus $\Delta(x) = V(x)S(x)$ for some symmetric polynomial $S(x) \in \mathbb{Z}[x]$. One can argue inductively that the polynomial $S(x)$ really has integer coefficients: First write $\Delta(x) = (x_m - x_{m-1})S_1(x)$. Suppose not all the coefficients of the polynomial

$$S_1(x) = \sum_{i \geq 0} c_i \cdot x_1^{n_1^{(i)}} \cdots x_m^{n_m^{(i)}} \in \mathbb{Q}[x], \quad n_1^{(i)}, \ldots, n_m^{(i)} \in \mathbb{N},$$

are integers, and let $c \cdot a x_m^k$ be a (not necessarily unique) term of $S_1(x)$ for which $c \in \mathbb{Q} \setminus \mathbb{Z}$ and the power of x_m is the highest possible. Here $k \in \mathbb{N}$ and a denotes the product of the powers of the other variables $x_1, \ldots, x_{m-1}$. Now the product $x_m S_1(x)$ contains the term $c \cdot a x_m^{k+1}$, but the product $x_{m-1} S_1(x)$ can only contain an integer multiple of $a x_m^{k+1}$, for otherwise there would be a contradiction with the choice of the term $c \cdot a x_m^k$. As a result, the polynomial

$$\Delta(x) = x_m S_1(x) - x_{m-1} S_1(x) \in \mathbb{Z}[x]$$

contains a term $c' \cdot a x_m^k$, where $c' \in \mathbb{Q} \setminus \mathbb{Z}$. This is a contradiction, implying $S_1(x) \in \mathbb{Z}[x]$.

Next, $S_1(x) = (x_m - x_{m-2}) S_2(x)$, where again $S_2(x) \in \mathbb{Z}[x]$ by repeating the reasoning above. Continuing this way we have

$$\begin{aligned}
\Delta(x) &= (x_m - x_{m-1}) S_1(x) \\
&= (x_m - x_{m-1})(x_m - x_{m-2}) S_2(x) \\
&= \ldots \\
&= (x_m - x_{m-1})(x_m - x_{m-2}) \ldots (x_3 - x_2)(x_3 - x_1) S_{\frac{(m-1)m-2}{2}}(x) \\
&= V(x) S(x),
\end{aligned}$$

where $S_{\frac{(m-1)m-2}{2}}(x) = (x_2 - x_1) S(x) \in \mathbb{Z}[x]$, and so $S(x) \in \mathbb{Z}[x]$. [3]

Returning to the equation (6.9) we see that the product

$$\prod_{1 \le i < j \le m} (x_j - x_i)$$

is a factor of all the $m \times m$ minors of the first m rows in (6.8), which means that it divides the minor (6.6). (As seen before, this is a consequence of the Laplace expansion formula.) Since the chosen minor was an arbitrary one,

[3] Actually $S(x) = \frac{\Delta(x)}{V(x)}$ is a so-called *Schur polynomial*. More on these in Chapter 2 of Procesi's book [12].

the greatest common divisor D of all the $M \times M$ minors of the matrix B in (6.3) indeed satisfies

$$D \geq \prod_{1 \leq i < j \leq m} |x_j - x_i|.$$

Bibliography

[1] Berger, Marcel: *Geometry Revealed*, Springer, Berlin, 2010.

[2] Bombieri, E.; Vaaler, J.: *On Siegel's Lemma*, Invent. math., vol 73, pp. 11 - 32, 1983.

[3] Cassels, J. W. S.: *An Introduction to the Geometry of Numbers*, Springer-Verlag, Berlin, 1997.

[4] Ernvall-Hytönen, Anne-Maria; Leppälä, Kalle; Matala-aho, Tapani: *An explicit Baker type lower bound of exponential values*, Proc. Roy. Soc. Edinburgh Sect. A. (In press.) `http://cc.oulu.fi/~tma/JULKAISUJA.html`

[5] Evertse, Jan-Hendrik: *Diophantine approximation*, lecture notes, Leiden University, 2015. `http://www.math.leidenuniv.nl/~evertse/dio.shtml`

[6] Johnson, Lee W.; Riess, R. Dean; Arnold, Jimmy T.: *Introduction to Linear Algebra*, Addison-Wesley, Reading, 1993.

[7] Krattenthaler, C.: *Advanced determinant calculus*, Séminaire Lotharingien Combin. 42 (1999) ("The Andrews Festschrift"), Article B42q, 67 pp. math.CO/9902004. `http://www.mat.univie.ac.at/~kratt/papers.html`

[8] Mahler, Kurt: *Lectures on Transcendental Numbers*, Springer-Verlag, Berlin, 1976.

[9] Matala-aho, Tapani: *A geometric face of Diophantine analysis*, lecture notes given at Summer School for Master and PhD students in Diophantine Analysis, Würzburg 2014. `http://cc.oulu.fi/~tma/JULKAISUJA.html`

[10] Matala-aho, Tapani: *Algebralliset luvut*, lecture notes, University of Oulu, 2014. `http://cc.oulu.fi/~tma/ALGEBRALLISET.html`

[11] Matala-aho, Tapani: *On Baker type lower bounds for linear forms*. (Submitted.) `http://cc.oulu.fi/~tma/JULKAISUJA.html`

[12] Procesi, Claudio: *Lie Groups: An Approach through Invariants and Representations*, Springer, New York, 2007.

[13] Steuding, Jörn: *Diophantine Analysis*, Chapman & Hall / CRC, Boca Raton, 2005.

[14] Stewart, I. N.; Tall, D. O.: *Algebraic Number Theory*, Chapman & Hall / CRC, London, 1978.