Chapter 2

Geometry of numbers

Literature:
W.M. Schmidt, Diophantine approximation, Lecture Notes in Mathematics 785, Springer Verlag 1980, Chap.II, §§1,2, Chap. IV, §1
J.W.S. Cassels, An Introduction to the Geometry of Numbers, Springer Verlag 1997, Classics in Mathematics series, reprint of the 1971 edition
C.L. Siegel, Lectures on the Geometry of Numbers, Springer Verlag 1989

2.1 Introduction

Geometry of numbers is concerned with the study of lattice points in certain bodies in $\mathbb{R}^n$, where $n \geqslant 2$. We discuss Minkowski's theorems on lattice points in central symmetric convex bodies. In this introduction we give the necessary definitions.

Discrete subgroups of $\mathbb{R}^n$. We call vectors $\mathbf{w}_1, \ldots, \mathbf{w}_r \in \mathbb{R}^n$ linearly independent if they are linearly independent over $\mathbb{R}$, i.e., there is no tuple $(\xi_1, \ldots, \xi_r) \in \mathbb{R}^r \setminus \{\mathbf{0}\}$ with $\xi_1\mathbf{w}_1 + \cdots + \xi_r\mathbf{w}_r = \mathbf{0}$. Notions such as boundedness, openness, closedness, etc. for subsets of $\mathbb{R}^n$ are all with respect to the usual Euclidean metric on $\mathbb{R}^n$ and the topology induced by it. We use $[x]$ to denote the largest integer $\leqslant x$.

A subset S of $\mathbb{R}^n$ is called *discrete* if $S \cap B$ is finite for every bounded subset B of $\mathbb{R}^n$. We consider discrete subgroups of $\mathbb{R}^n$, i.e., discrete subsets such that if $\mathbf{x}, \mathbf{y}$ belong to this set then so do $z\mathbf{x}+w\mathbf{y}$ for all $z, w \in \mathbb{Z}$. The *rank* of a discrete subgroup M of $\mathbb{R}^n$ is the maximal number r such that M contains r linearly independent

vectors.

The lemma below shows that a non-zero discrete subgroup M of $\mathbb{R}^n$ has a *basis*. A basis of M is a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, linearly independent over $\mathbb{R}$, such that

$$M = \{z_1\mathbf{v}_1 + \cdots + z_r\mathbf{v}_r : z_1, \ldots, z_r \in \mathbb{Z}\}.$$

We show how to construct a basis of M taking as starting point any linearly independent subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ of M of maximal cardinality.

Lemma 2.1. *Let M be a discrete subgroup of $\mathbb{R}^n$ and let $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$ be any linearly independent subset of M of maximal cardinality. Then M has a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ such that for $k = 1, \ldots, r$ we have*

$$\mathbf{v}_k = \xi_{k1}\mathbf{w}_1 + \cdots + \xi_{kk}\mathbf{w}_k \text{ with } \xi_{kj} \in \mathbb{R},\ 0 \leqslant \xi_{kj} \leqslant 1 \text{ for } j = 1, \ldots, k,\ \xi_{kk} \neq 0. \tag{2.1}$$

Proof. We first choose $\mathbf{v}_1, \ldots, \mathbf{v}_r$. For $k = 1, \ldots, r$ define S_k to be the set of vectors in M of the form

$$\sum_{j=1}^{k} \xi_j\mathbf{w}_j \text{ with } \xi_j \in \mathbb{R},\ 0 \leqslant \xi_j \leqslant 1 \text{ for } j = 1, \ldots, k,\ \xi_k \neq 0.$$

The set S_k is non-empty since $\mathbf{w}_k \in S_k$. Since M is discrete and S_k is a bounded subset of M, the set S_k is finite. Choose $\mathbf{v}_k$ from S_k with minimal $\mathbf{w}_k$-coordinate, that is, if we write $\mathbf{v}_k = \sum_{j=1}^{k} \xi_{kj}\mathbf{w}_j$, then

$$\sum_{j=1}^{k} \xi_j\mathbf{w}_j \in S_k \Rightarrow \xi_k \geqslant \xi_{kk}. \tag{2.2}$$

We have thus chosen a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ which satisfies (2.1), and from this it follows easily that it is linearly independent over $\mathbb{R}$. We show that it is a basis of M. In fact, we prove the following assertion by induction on k, for $k = 0, \ldots, r$:

let $M_k := M \cap \{\sum_{j=1}^{k} \xi_j\mathbf{w}_j : \xi_1, \ldots, \xi_k \in \mathbb{R}\}$. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a basis of M_k. For $k = 0$ this has to be interpreted as that $M_0 = \{\mathbf{0}\}$ and the empty set is a basis of M_0.

For $k = 0$ our assertion is trivially true. Let $k > 0$. The induction hypothesis is that our assertion is true for $k - 1$ replacing k. In the induction step we use the

following observation: if $\mathbf{x} = \sum_{j=1}^{k} \xi_j \mathbf{w}_j \in M_k$ and $\xi_k > 0$, then in fact $\xi_k \geqslant \xi_{kk}$. Indeed, suppose that $0 < \xi_k < \xi_{kk}$. Then

$$\mathbf{x} - \sum_{j=1}^{k-1} [\xi_j] \mathbf{w}_j = \sum_{j=1}^{k-1} (\xi_j - [\xi_j]) \mathbf{w}_j + \xi_k \mathbf{w}_k \in S_k.$$

But this contradicts (2.2).

Now we complete the induction step. Let $\mathbf{x} = \sum_{j=1}^{k} \xi_j \mathbf{w}_j \in M_k$. We must show that $\mathbf{x}$ is a $\mathbb{Z}$-linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Let $z_k := [\xi_k / \xi_{kk}]$; then $0 \leqslant \xi_k - z_k \xi_{kk} < \xi_{kk}$. Now

$$\mathbf{x}_1 := \mathbf{x} - z_k \mathbf{v}_k = \sum_{j=1}^{k} (\xi_j - z_k \xi_{kj}) \mathbf{w}_j \in M_k.$$

By the above observation we must have $\xi_k - z_k \xi_{kk} = 0$. So in fact, $\mathbf{x}_1 \in M_{k-1}$. By the induction hypothesis, $\mathbf{x}_1$ is a $\mathbb{Z}$-linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$ (or $\mathbf{0}$ if $k = 1$). Hence $\mathbf{x}$ is a $\mathbb{Z}$-linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. This completes the induction step. □

Lattices. A (full) *lattice* in $\mathbb{R}^n$ is a discrete subgroup of $\mathbb{R}^n$ of maximal rank n. By the above lemma this implies that L has a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, that is,

$$L = \{z_1 \mathbf{v}_1 + \cdots + z_n \mathbf{v}_n : \, z_1, \ldots, z_n \in \mathbb{Z}\}.$$

The *determinant* of L is defined by

$$d(L) := |\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)|,$$

that is, the absolute value of the determinant of the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

We show that the determinant of a lattice does not depend on the choice of the basis. Recall that $\mathrm{GL}(n, \mathbb{Z})$ is the multiplicative group of $n \times n$-matrices with entries in $\mathbb{Z}$ and determinant ± 1.

Lemma 2.2. *Let L be a lattice, and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ two bases of L. Then there is a matrix $U = (u_{ij}) \in \mathrm{GL}(n, \mathbb{Z})$ such that*

$$\mathbf{w}_i = \sum_{j=1}^{n} u_{ij} \mathbf{v}_j \quad \textit{for } i = 1, \ldots, n. \tag{2.3}$$

Consequently, $|\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)| = |\det(\mathbf{w}_1, \ldots, \mathbf{w}_n)|$.

Proof. Let U be the matrix expressing $\mathbf{w}_1, \ldots, \mathbf{w}_n$ into $\mathbf{v}_1, \ldots, \mathbf{v}_n$, that is, the matrix given by (2.3). A priori, U is just a non-singular matrix, but since $\mathbf{w}_1, \ldots, \mathbf{w}_n$ lie in the lattice generated by $\mathbf{v}_1, \ldots, \mathbf{v}_n$, it must have its entries in $\mathbb{Z}$.

Let $U^{-1} = (u^{ij})$. Then from linear algebra we know that $\mathbf{v}_i = \sum_{j=1}^{n} u^{ij}\mathbf{w}_j$ for $i = 1, \ldots, n$. Now U^{-1} has its entries in $\mathbb{Z}$ since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ lie in the lattice generated by $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Since both $\det U$ and $\det U^{-1}$ are integers and must be multiplicative inverses of one another, we have $\det U = \pm 1$, i.e., $U \in \mathrm{GL}(n, \mathbb{Z})$.

Finally, we observe that

$$|\det(\mathbf{w}_1, \ldots, \mathbf{w}_n)| = |\det U| \cdot |\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)| = |\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)|.$$

□

Let L, M be two lattices in $\mathbb{R}^n$ with $M \subseteq L$. Choose bases $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of L, $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ of M. Let $U = (u_{ij})$ be the matrix expressing $\mathbf{w}_1, \ldots, \mathbf{w}_n$ into $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then U has its entries in $\mathbb{Z}$. We define the *index* of M in L by

$$|L : M| := |\det U|.$$

The relation $\det(\mathbf{w}_1, \ldots, \mathbf{w}_n) = \det U \cdot \det(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ easily translates into

$$d(M) = |L : M| \cdot d(L)$$

and this shows that $|L : M|$ does not depend on the choices of the bases of L and M. [1]

Convex bodies. Recall that a subset C of $\mathbb{R}^n$ is convex if for any two points $\mathbf{x}, \mathbf{y} \in C$, also the line segment connecting them, i.e., $\{t\mathbf{x} + (1-t)\mathbf{y} : 0 \leqslant t \leqslant 1\}$, is contained in C. A *central symmetric convex body* in $\mathbb{R}^n$ is a closed, bounded, convex subset C of $\mathbb{R}^n$ having $\mathbf{0}$ as an interior point, and which is symmetric about $\mathbf{0}$, i.e. if $\mathbf{x} \in C$ then also $-\mathbf{x} \in C$.

[1] The index $|L : M|$ as defined above is equal to the index as defined in group theory, that is the order of the quotient group L/M. This can be seen as follows. By a general theorem for abelian groups, there are a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of L and positive integers $d_1, \ldots, d_n$ such that $\{d_1\mathbf{v}_1, \ldots, d_n\mathbf{v}_n\}$ is a basis of M. On the one hand, according to the above definition, $|L : M| = d_1 \cdots d_n$, on the other hand, $L/M \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_n\mathbb{Z}$, and so it has cardinality $d_1 \cdots d_n$.

Given a central symmetric convex body C in $\mathbb{R}^n$ and a real $\lambda \geqslant 0$, we define the dilation of C with factor λ by

$$\lambda C := \{\lambda \mathbf{x} : \mathbf{x} \in C\}.$$

In case that $\lambda > 0$, this is again a central symmetric convex body.

Exercise 2.1. *Let C be a central symmetric convex body.*
(i) Let λ, μ be reals with $0 \leqslant \lambda \leqslant \mu$. Prove that $\lambda C \subseteq \mu C$.
(ii) Let $\mathbf{x} \in \lambda C$, $\mathbf{y} \in \mu C$ where $\lambda, \mu \geqslant 0$. Prove that $\mathbf{x} + \mathbf{y} \in (\lambda + \mu)C$.
(iii) Let B be a bounded subset of $\mathbb{R}^n$. Then there is $\lambda > 0$ such that $B \subseteq \lambda C$.

Examples.
(i). Images under linear transformations: If C is a central symmetric convex body in $\mathbb{R}^n$ and ϕ a linear transformation of $\mathbb{R}^n$ (i.e., an invertible linear map from $\mathbb{R}^n$ to itself), then $\phi(C)$ is also a central symmetric convex body in $\mathbb{R}^n$.

(ii). Parallelepipeds, ellipsoids and octahedra: Let

$$K_n = \{\mathbf{x} \in \mathbb{R}^n : \max_{1 \leqslant i \leqslant n} |x_i| \leqslant 1\}, \quad B_n = \{\mathbf{x} \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 \leqslant 1\},$$
$$O_n = \{\mathbf{x} \in \mathbb{R}^n : |x_1| + \cdots + |x_n| \leqslant 1\}$$

be the n-dimensional unit cube, Euclidean unit ball, and unit octahedron, respectively, where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Then (central) parallelepipeds, ellipsoids and octahedra in $\mathbb{R}^n$ are the images of K_n, B_n and O_n respectively under linear transformations of $\mathbb{R}^n$. These are all central symmetric convex bodies.

(iii). Unit balls of norms: Recall that a *norm* on $\mathbb{R}^n$ is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ such that

- $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$;
- $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;
- $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$.

Then the unit ball $B_{\|\cdot\|} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leqslant 1\}$ is a central symmetric convex body. Indeed, recall that all norms on $\mathbb{R}^n$ induce the same topology, that is that the definitions of openness, closedness, interior points, boundedness, etc., do not depend on the choice of the norm. This implies directly that $B_{\|\cdot\|}$ is closed and bounded and has $\mathbf{0}$ as an interior point. The central symmetry and convexity follow easily from the first and second property of a norm.

In fact, every central symmetric convex body arises from a norm. Let again C be a central symmetric convex body in $\mathbb{R}^n$ and define for $\mathbf{x} \in \mathbb{R}^n$,

$$\|\mathbf{x}\|_C := \min\{\lambda \in \mathbb{R}_{\geqslant 0} : \mathbf{x} \in \lambda C\}.$$

Lemma 2.3. *(i)* $\|\cdot\|_C$ *is well defined.*
(ii) $\|\cdot\|_C$ *defines a norm on* $\mathbb{R}^n$*.*
(iii) $\lambda C = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_C \leqslant \lambda\}$ *for* $\lambda \geqslant 0$*.*

Proof. (i). Clearly, $\|\mathbf{0}\|_C = 0$. Let $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{x} \neq \mathbf{0}$. Consider the set

$$S := \{\lambda : \lambda \in \mathbb{R}_{\geqslant 0}, \mathbf{x} \in \lambda C\}.$$

We have to prove that S is non-empty and that it has a minimum. Our argument will imply also that this minimum is positive. Let r denote the (Euclidean) length of $\mathbf{x}$.

First, $\mathbf{0}$ is an interior point of C which means that there is $\delta > 0$ such that C contains all vectors in $\mathbb{R}^n$ of length at most δ. As a consequence $(r/\delta)C$ contains all vectors of length at most r, so in particular $\mathbf{x}$. Hence $S \neq \emptyset$. Thus, the set S has an *infimum*, which we denote by μ.

The definition of the infimum implies that $\mathbf{x} \in (\mu + \varepsilon)C$ for every $\varepsilon > 0$, hence $(\mu + \varepsilon)^{-1}\mathbf{x} \in C$ for every $\varepsilon > 0$. Since the set C is bounded, this implies $\mu > 0$. Further, since the set C is closed, it contains the limit of the sequence of points $\{(\mu + 1/m)^{-1}\mathbf{x} : m = 1, 2, \ldots\}$, which is $\mu^{-1}\mathbf{x}$. So $\mathbf{x} \in \mu C$, i.e., $\mu \in S$. Hence μ is the minimum of S. This shows that $\|\mathbf{x}\|_C$ is well-defined and positive.

(ii). We have shown above that $\|\mathbf{x}\|_C > 0$ if $\mathbf{x} \neq \mathbf{0}$. The proofs of the other two norm properties are left to the reader.

(iii). Left to the reader. □

Exercise 2.2. *Prove (ii) and (iii).*

2.2 Minkowski's first convex body theorem

Using Lebesgue theory, one can define an n-dimensional volume $\mathrm{vol}(S) \in \mathbb{R}_{\geqslant 0} \cup \{\infty\}$ (the so-called Lebesgue measure) for subsets S of $\mathbb{R}^n$ from a large class, the so-called

measurable subsets of $\mathbb{R}^n$. We do not need the precise definition of Lebesgue measure or measurable set. What is important to us is that all open sets and all closed sets are measurable, bounded measurable sets have finite volume, and the empty set has volume 0. The volume of S is equal to the Riemann integral $\int_S dx_1 \cdots dx_n$ for every set S for which this integral is defined. However, there are measurable sets S for which the Riemann integral is not defined. We mention some important properties of the volume:

1. Let S be a measurable subset of $\mathbb{R}^n$. Then every translate $\mathbf{a}+S := \{\mathbf{a}+\mathbf{x} : \mathbf{x} \in S\}$ is also measurable and $\mathrm{vol}(\mathbf{a}+S) = \mathrm{vol}(S)$. Further, if ϕ is a linear transformation of $\mathbb{R}^n$, then $\phi(S)$ is measurable and $\mathrm{vol}(\phi(S)) = |\det \phi| \cdot \mathrm{vol}(S)$.

2. Let $S \subset \mathbb{R}^n$ be measurable. Then $S^c := \mathbb{R}^n \setminus S$ is measurable.

3. Let S_n $(n = 1, 2, 3, \ldots)$ be a countable collection of measurable subsets of $\mathbb{R}^n$. Then $S = \bigcup_{n=1}^{\infty} S_n$ is measurable. Moreover, if the sets S_n are pairwise disjoint, then $\mathrm{vol}(S) = \sum_{n=1}^{\infty} \mathrm{vol}(S_n)$.

Theorem 2.4. (Minkowski's first convex body theorem, 1896). *Let C be a central symmetric convex body in $\mathbb{R}^n$ and L a lattice in $\mathbb{R}^n$ of rank n. Suppose that $\mathrm{vol}(C) \geqslant 2^n d(L)$. Then C contains a point from $L \setminus \{\mathbf{0}\}$.*

Choose a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of L. We call

$$F := \{x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n : x_i \in \mathbb{R},\ 0 \leqslant x_i < 1 \text{ for } i = 1, \ldots, n\}$$

a *fundamental parallelepiped* for L. Notice that F has volume $d(L)$ and that the translates $\mathbf{u} + F$ $(\mathbf{u} \in L)$ are pairwise disjoint and cover $\mathbb{R}^n$, that is,

$$\mathbb{R}^n = \bigcup_{\mathbf{u} \in L} (\mathbf{u} + F).$$

We present two proofs of Theorem 2.4: one based on computing volumes, and another one based on a lattice point counting result which is of interest in itself.

First proof of Theorem 2.4. We first assume that $\mathrm{vol}(C) > 2^n d(L)$. Then the set $\frac{1}{2}C = \{\frac{1}{2}\mathbf{x} : \mathbf{x} \in C\}$ has volume $> d(L)$. For $\mathbf{u} \in L$, define $S_{\mathbf{u}} := \frac{1}{2}C \cap (\mathbf{u}+F)$. Then the sets $S_{\mathbf{u}}$ $(\mathbf{u} \in L)$ are pairwise disjoint and their union is precisely $\frac{1}{2}C$. Hence

$$\sum_{\mathbf{u} \in L} \mathrm{vol}(S_{\mathbf{u}}) = \mathrm{vol}(\tfrac{1}{2}C) > d(L).$$

We shift the sets $S_{\mathbf{u}}$ into F, that is, we define

$$S^*_{\mathbf{u}} := -\mathbf{u} + S_{\mathbf{u}} = (-\mathbf{u} + \tfrac{1}{2}C) \cap F \quad \text{for } \mathbf{u} \in L.$$

Since $S^*_{\mathbf{u}}$ has the same volume as $S_{\mathbf{u}}$, we have

$$\sum_{\mathbf{u}\in L} \mathrm{vol}(S^*_{\mathbf{u}}) = \sum_{\mathbf{u}\in L} \mathrm{vol}(S_{\mathbf{u}}) > d(L) = \mathrm{vol}(F).$$

That is, we have a collection of subsets $S^*_{\mathbf{u}}$ ($\mathbf{u} \in L$) of F, the sum of whose volumes is larger than the volume of F. So there are two distinct $\mathbf{u}, \mathbf{v} \in L$ such that $S^*_{\mathbf{u}} \cap S^*_{\mathbf{v}} \neq \emptyset$.

Pick a point $\mathbf{a} \in S^*_{\mathbf{u}} \cap S^*_{\mathbf{v}}$. Then for certain $\mathbf{x}, \mathbf{y} \in \frac{1}{2}C$ we have $\mathbf{x} - \mathbf{u} = \mathbf{y} - \mathbf{v} = \mathbf{a}$. Hence $\mathbf{x} - \mathbf{y} = \mathbf{u} - \mathbf{v} \in L \setminus \{\mathbf{0}\}$.

Now $2\mathbf{x}, 2\mathbf{y} \in C$, by the symmetry of C we have $-2\mathbf{y} \in C$, and by the convexity of C we have $\frac{1}{2}(2\mathbf{x} - 2\mathbf{y}) = \mathbf{x} - \mathbf{y} \in C$. This shows that C contains a non-zero point from L.

Now assume that $\mathrm{vol}(C) = 2^n d(L)$. Suppose that C does not contain a non-zero point from L. Then for every integer $m \geqslant 1$, $(1 + m^{-1})C$ contains a non-zero point $\mathbf{x}_m$ from L since $\mathrm{vol}((1 + m^{-1})C) = (1 + m^{-1})^n \,\mathrm{vol}(C) > 2^n d(L)$. All points $\mathbf{x}_m$ lie in $2C$, and since $(2C) \cap L$ is finite, there can be only finitely many distinct ones among them. So there is a non-zero $\mathbf{x} \in L$ such that $\mathbf{x} \in (1 + m^{-1})C$ for infinitely many m. Hence $(1 + m^{-1})^{-1}\mathbf{x} \in C$ for infinitely many m. Taking the limit, using that C is closed, it follows that $\mathbf{x} \in C$. □

Exercise 2.3. *Prove the following theorem of Blichfeldt. Let S be a measurable, not necessarily convex, subset of $\mathbb{R}^n$ with $\mathrm{vol}(S) > d(L)$. Then there are $\mathbf{x}, \mathbf{y} \in S$ with $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{x} - \mathbf{y} \in L$.*

Before giving the second proof of Theorem 2.4, we derive a lattice point counting result. We use Landau's O-notation: for real functions f, g, h, defined on subintervals of $\mathbb{R}$, we write

$$f(x) = g(x) + O(h(x)) \quad \text{as } x \to \infty$$

if there are real numbers x_0 and c such that f, g, h are defined for $x \geqslant x_0$, and $|f(x) - g(x)| \leqslant ch(x)$ for $x \geqslant x_0$. This means that if we approximate $f(x)$ by $g(x)$ and let $x \to \infty$, then asymptotically our error has order of magnitude at most $h(x)$. For instance, for fixed n, a we have $(x + a)^n = x^n + O(x^{n-1})$ as $x \to \infty$.

The cardinality of a set S is denoted by $\#S$.

Lemma 2.5. *Let C be a central symmetric convex body and L a lattice in $\mathbb{R}^n$. Put $\alpha := \mathrm{vol}(C)/d(L)$. Then*

$$\#(\lambda C \cap L) = \alpha\lambda^n + O(\lambda^{n-1}) \quad \textit{as } \lambda \to \infty.$$

Proof. Let $N(\lambda) := \#(\lambda C \cap L)$. Consider the set $S := \bigcup_{\mathbf{u}\in\lambda C\cap L}(\mathbf{u}+F)$. Note that S is a disjoint union of precisely $N(C)$ parallelepipeds, each of volume $d(L)$. So S has volume $N(\lambda)\cdot d(L)$. Further, λC has volume $\lambda^n \mathrm{vol}(C)$. Suppose that all parallelepipeds $\mathbf{u}+F$ lie either completely inside, or completely outside λC. Then $S = \lambda C$, and $N(\lambda) = \alpha\lambda^n$. But of course, in general some of the parallelepipeds $\mathbf{u}+F$ lie partly inside, partly outside λC. So by approximating $N(\lambda)$ by $\alpha\lambda^n$ we make an error, which we have to estimate.

Since F is bounded, there is $a > 0$ such that $\|\mathbf{y}\|_C \leqslant a$ for $\mathbf{y} \in F$. We first prove that

$$(*) \qquad (\lambda - a)C \subseteq S \subseteq (\lambda + a)C \quad \text{for } \lambda > a.$$

Write $\mathbf{x} \in \mathbb{R}^n$ as $\mathbf{u}+\mathbf{y}$, with $\mathbf{u} \in L$ and $\mathbf{y} \in F$. So $\|\mathbf{x}-\mathbf{u}\|_C \leqslant a$. Translating (*) into norms via Lemma 2.3, what we have to show that $\|\mathbf{x}\|_C \leqslant \lambda - a \Rightarrow \|\mathbf{u}\|_C \leqslant \lambda$, and $\|\mathbf{u}\|_C \leqslant \lambda \Rightarrow \|\mathbf{x}\|_C \leqslant \lambda + a$. But this follows directly from the triangle inequality.

Taking volumes in (*), we obtain

$$(\lambda - a)^n \mathrm{vol}(C) \leqslant N(\lambda)\cdot d(L) \leqslant (\lambda + a)^n \mathrm{vol}(C),$$

and this shows that there are $c > 0, \lambda_0 > 0$ such that $|N(\lambda) - \alpha\lambda^n| \leqslant c\lambda^{n-1}$ for all $\lambda > \lambda_0$. □

Second proof of Theorem 2.4. We first consider again the case that $\mathrm{vol}(C) > 2^n d(L)$, so that $\alpha > 2^n$. Let m be a positive integer. The previous lemma implies that $\#(mC\cap L) = \alpha m^n + O(m^{n-1})$ is larger than $(2m)^n$, provided we choose m sufficiently large.

We divide L into congruence classes modulo $2m$ by setting $\mathbf{x} \equiv \mathbf{y} (\mathrm{mod}\, 2m)$ if $(2m)^{-1}(\mathbf{x}-\mathbf{y}) \in L$. Thus, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of L and if we write $\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$, $\mathbf{y} = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n$ with $x_i, y_i \in \mathbb{Z}$, we have $\mathbf{x} \equiv \mathbf{y} (\mathrm{mod}\, 2m)$ if and only if $x_i \equiv y_i (\mathrm{mod}\, 2m)$ for $i = 1, \ldots, n$. Hence L can be divided into precisely $(2m)^n$ congruence classes modulo $2m$.

Now by the box principle, there are two distinct $\mathbf{x}, \mathbf{y} \in mC \cap L$ with $\mathbf{x} \equiv \mathbf{y} (\mathrm{mod}\, 2m)$. So $\mathbf{u} := \frac{1}{2m}(\mathbf{x} - \mathbf{y})$ is a non-zero element of L. By the symmetry of mC, we have $-\mathbf{y} \in mC$, and by its convexity, $\frac{1}{2}(\mathbf{x} - \mathbf{y}) \in mC$. Hence $\mathbf{u} \in C$. Thus Theorem 2.4 follows in the case $\mathrm{vol}(C) > 2^n d(L)$. The case $\mathrm{vol}(C) = 2^n d(L)$ is treated in the same way as in the first proof. □

Exercise 2.4. *Let C, L be a central symmetric convex body and lattice in $\mathbb{R}^n$ and r a positive integer such that $\mathrm{vol}(C) > r \cdot 2^n d(L)$. Prove that there are $\mathbf{u}_1, \ldots, \mathbf{u}_r \in C \cap L$ such that $\mathbf{u}_i \neq \pm \mathbf{u}_j$ for $i, j = 1, \ldots, r$.*
Hint. *Given distinct points $\mathbf{x}_0, \ldots, \mathbf{x}_r \in \mathbb{R}^n$, prove that there is $i \in \{1, \ldots, r\}$ such that $2\mathbf{x}_i \neq \mathbf{x}_j + \mathbf{x}_k$ for all $j, k \in \{1, \ldots, r\} \setminus \{i\}$.*

We give some consequences of Theorem 2.4.

Corollary 2.6. *Let $l_i = \alpha_{i1} X_1 + \cdots + \alpha_{in} X_n$ $(i = 1, \ldots, n)$ be linear forms with real coefficients and with $\det(l_1, \ldots, l_n) \neq 0$. Let $A_1, \ldots, A_n$ be positive reals with*

$$A_1 \cdots A_n \geqslant |\det(l_1, \ldots, l_n)|.$$

Then there is a non-zero $\mathbf{x} \in \mathbb{Z}^n$ with

$$|l_1(\mathbf{x})| \leqslant A_1, \ldots, |l_n(\mathbf{x})| \leqslant A_n.$$

Proof. Recall that

$$K_n = \{\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n : |y_i| \leqslant 1 \text{ for } i = 1, \ldots, n\},$$

and define the lattice

$$L := \{(A_1^{-1} l_1(\mathbf{x}), \ldots, A_n^{-1} l_n(\mathbf{x})) : \mathbf{x} \in \mathbb{Z}^n\}.$$

Clearly, K_n is a central symmetric convex body with volume 2^n, while L is a lattice of determinant $|\det(l_1, \ldots, l_n)|/A_1 \cdots A_n \leqslant 1$. Theorem 2.4 implies that K_n contains a non-zero point from L. Corollary 2.6 follows. □

Exercise 2.5. *Prove the following refinement of Corollary 2.6. Let again $l_i = \alpha_{i1} X_1 + \cdots + \alpha_{in} X_n$ $(i = 1, \ldots, n)$ be linear forms with real coefficients and with $\det(l_1, \ldots, l_n) \neq 0$, and $A_1, \ldots, A_n$ positive reals with*

$$A_1 \cdots A_n \geqslant |\det(l_1, \ldots, l_n)|.$$

Then there is a non-zero $\mathbf{x} \in \mathbb{Z}^n$ *with*

$$|l_1(\mathbf{x})| < A_1, \ldots, |l_{n-1}(\mathbf{x})| < A_{n-1}, \ |l_n(\mathbf{x})| \leqslant A_n$$

(so $n-1$ *inequalities have a* $<$*-sign and one a* $\leqslant$*-sign).*

Hint. *You have to apply Corollary 2.6 to systems of inequalities*

$$|l_1(\mathbf{x})| \leqslant A_1', \ldots, |l_{n-1}(\mathbf{x})| \leqslant A_{n-1}', \ |l_n(\mathbf{x})| \leqslant A_n'$$

where $A_i' < A_i$ *for* $i = 1, \ldots, n-1$ *and* $A_n' > A_n$ *and let* $A_i' \nearrow A_i$ *for* $i = 1, \ldots, n-1$ *and* $A_n' \searrow A_n$.

In the introduction we showed that if α is a real irrational number, then there are infinitely many pairs of integers (x, y) with $\gcd(x, y) = 1$, $y > 0$ and

$$\left|\alpha - \frac{x}{y}\right| \leqslant y^{-2}. \tag{2.4}$$

We prove some generalizations.

Corollary 2.7 (Dirichlet, 1842)**.** *(i) Let* $\alpha_1, \ldots, \alpha_n$ *be real numbers, at least one of which is irrational. Then there are infinitely many tuples of integers* $(x_1, \ldots, x_n, y)$ *with* $\gcd(x_1, \ldots, x_n, y) = 1$, $y > 0$ *and*

$$\left|\alpha_i - \frac{x_i}{y}\right| \leqslant y^{-1-1/n} \text{ for } i = 1, \ldots, n. \tag{2.5}$$

(ii) Let $\alpha_1, \ldots, \alpha_n$ *be real numbers such that* $1, \alpha_1, \ldots, \alpha_n$ *are linearly independent over* $\mathbb{Q}$*. Then there are infinitely many tuples of integers* $(x, y_1, \ldots, y_n)$ *with* $(y_1, \ldots, y_n) \neq (0, \ldots, 0)$ *and*

$$|\alpha_1 y_1 + \cdots + \alpha_n y_n - x| \leqslant \max(|y_1|, \ldots, |y_n|)^{-n}.$$

Proof. We prove only (i). We consider instead of (2.5) the system of inequalities

$$|x_i - \alpha_i y| \leqslant Q^{-1/n} \ (i = 1, \ldots, n), \ 0 < y \leqslant Q, \quad \gcd(x_1, \ldots, x_n, y) = 1 \tag{2.6}$$

for any integer $Q > 1$, and let Q vary. If $(x_1, \ldots, x_n, y)$ is a tuple of integers satisfying this system, then it is also a solution of

$$|x_i - \alpha_i y| \leqslant y^{-1/n} \text{ for } i = 1, \ldots, n$$

with $y > 0$, $\gcd(x_1, \ldots, x_n, y) = 1$, and hence a solution of (2.5) with these properties.

Notice that the system of linear forms $X_1 - \alpha_1 X_{n+1}, \ldots, X_n - \alpha_n X_{n+1}, X_{n+1}$ has determinant 1. So by Corollary 2.6, for every integer $Q > 1$ there is a non-zero tuple of integers $(x_1, \ldots, x_n, y)$ satisfying $|x_i - \alpha_i y| \leqslant Q^{-1/n}$ for $i = 1, \ldots, n$ and $|y| \leqslant Q$. If $y = 0$ then $x_1 = \cdots = x_n = 0$ which is impossible. Hence $y \neq 0$. By changing signs and dividing out the gcd of $x_1, \ldots, x_n, y$ if necessary, we obtain a solution $\mathbf{x}_Q = (x_1, \ldots, x_n, y)$ of (2.6).

We claim that if we let $Q \to \infty$, then $\mathbf{x}_Q$ runs through an infinite set. Indeed, suppose the contrary. Then there is an infinite sequence of integers $Q_i \to \infty$ such that for each Q_i the point $\mathbf{x}_{Q_i}$ is equal to some fixed tuple of integers $\mathbf{x} = (x_1, \ldots, x_n, y)$ independent of i. But then, $\alpha_i = x_i/y \in \mathbb{Q}$ for $i = 1, \ldots, n$, against our assumption.

Thus, as observed above, the vectors $\mathbf{x}_Q$ give infinitely many solutions of (2.5) with $y > 0$ and $\gcd(x_1, \ldots, x_n, y) = 1$. □

Exercise 2.6. *Prove the following common generalization of both (i) and (ii). Let m, n be positive integers and $l_i = \alpha_{i1} X_1 + \cdots + \alpha_{in} X_n$ $(i = 1, \ldots, m)$ linear forms with real coefficients satisfying*

$$\{\mathbf{y} \in \mathbb{Z}^n : l_i(\mathbf{y}) \in \mathbb{Z} \text{ for } i = 1, \ldots, m\} = \{\mathbf{0}\}.$$

Then there are infinitely many tuples $(\mathbf{x}, \mathbf{y})$, with $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{Z}^m$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{Z}^n$, $\mathbf{y} \neq \mathbf{0}$, such that

$$|l_i(\mathbf{y}) - x_i| \leqslant \Big(\max_{1 \leqslant j \leqslant n} |y_j| \Big)^{-n/m} \text{ for } i = 1, \ldots, m.$$

Given a real number θ, we denote by $\|\theta\|$ the distance of θ to the nearest integer, i.e., $\|\theta\| = \min\{|\theta - m| : m \in \mathbb{Z}\}$. Corollary 2.7 implies that for any two real numbers α_1, α_2, not both in $\mathbb{Q}$, there are infinitely many positive integers y such that

$$\|\alpha_1 y\| \leqslant y^{-1/2}, \quad \|\alpha_2 y\| \leqslant y^{-1/2}.$$

This implies that there are infinitely many positive integers y such that

$$y \|\alpha_1 y\| \cdot \|\alpha_2 y\| \leqslant 1.$$

In fact, this is true also if both $\alpha_1, \alpha_2 \in \mathbb{Q}$, since then, there are infinitely many integers y with $\|\alpha_1 y\| = 0$, $\|\alpha_2 y\| = 0$. The following famous conjecture, due to

Littlewood, is still open:

Littlewood's Conjecture. *Let α_1, α_2 be any two real numbers. Then for every $\varepsilon > 0$ there exists a positive integer y such that*

$$y\|\alpha_1 y\| \cdot \|\alpha_2 y\| < \varepsilon.$$

Note that $\|x\| \leqslant \frac{1}{2}$ for every $x \in \mathbb{R}$. So Littlewood's conjecture would imply also that for any $n \geqslant 3$ reals $\alpha_1, \ldots, \alpha_n$, and for any $\varepsilon > 0$ there is a positive integer y with $y\|\alpha_1 y\| \cdots \|\alpha_n y\| < \varepsilon$.

Exercise 2.7. *Let d be a positive integer that is not a square. Prove that there is a constant $c(d) > 0$ such that*

$$y \cdot \|\sqrt{d}y\| \geqslant c(d) \quad \textit{for all } y \in \mathbb{Z}_{>0}$$

(that is, there is no one-dimensional analogue of Littlewood's Conjecture).

In general, a real, irrational number α for which there exists $c > 0$ such that $y \cdot \|\alpha y\| \geqslant c$ for all positive integers y is called **badly approximable**. It can be shown that there are uncountably many badly approximable numbers.

2.3 Minkowski's second convex body theorem

Let L be a lattice in $\mathbb{R}^n$ and C a central symmetric convex body in $\mathbb{R}^n$.

Definition. The n *successive minima* $\lambda_1, \ldots, \lambda_n$ of C with respect to L are defined as follows:

λ_i is the minimum of all positive reals λ such that $\lambda C \cap L$ contains at least i linearly independent points.

Lemma 2.8. *The successive minima $\lambda_1, \ldots, \lambda_n$ of C with respect to L are well-defined, and $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n < \infty$.*

Further, there are linearly independent $\mathbf{v}_1, \ldots, \mathbf{v}_n \in L$ with $\mathbf{v}_i \in \lambda_i C$ for $i = 1, \ldots, n$.

Proof. Let $\|\cdot\|_C$ be the norm associated with C, defined by $\|\mathbf{x}\|_C = \min\{\lambda \in \mathbb{R}_{\geqslant 0} : \mathbf{x} \in \lambda C\}$. Recall that $\lambda C = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_C \leqslant \lambda\}$.

We can order the points of L as a sequence $\mathbf{x}_0 = \mathbf{0}, \mathbf{x}_1, \mathbf{x}_2, \ldots$ such that $0 = \|\mathbf{x}_0\|_C < \|\mathbf{x}_1\|_C \leqslant \|\mathbf{x}_2\|_C \leqslant \cdots$. To see this, consider for each positive integer m the points $\mathbf{x} \in L$ with $m - 1 < \|\mathbf{x}\|_C \leqslant m$, these are the points with $\mathbf{x} \in mC$, $\mathbf{x} \notin (m-1)C$. Since L is discrete and mC is closed and bounded, there are only finitely many such points and these can be ordered according to their $\|\cdot\|_C$-values.

Define $\lambda_1 := \|\mathbf{x}_1\|_C$ and put $k_1 := 1$, $\mathbf{v}_1 := \mathbf{x}_1$. For $i = 2, \ldots, n$, let k_i be the first index k such that $\operatorname{rank}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = i$, and put $\lambda_i := \|\mathbf{x}_{k_i}\|_C$ and $\mathbf{v}_i := \mathbf{x}_{k_i}$. Clearly, $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n < \infty$, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, and $\mathbf{v}_i \in \lambda_i C$ for $i = 1, \ldots, n$.

It remains to show that λ_i is the i-th successive minimum of C with respect to L, for $i = 1, \ldots, n$. Clearly, $\lambda_i C \cap L$ contains the i linearly independent points $\mathbf{v}_1, \ldots, \mathbf{v}_i$. We have to show that $\lambda C \cap L$ does not contain i linearly independent points if $0 < \lambda < \lambda_i$. Take such λ. Note that $\lambda C \cap L$ contains precisely the points $\mathbf{x} \in L$ with $\|\mathbf{x}\|_C \leqslant \lambda$, i.e., the points $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k$ where $\|\mathbf{x}_k\|_C \leqslant \lambda < \|x_{k+1}\|_C$. Clearly, $k < k_i$, so there cannot be i linearly independent points among $\mathbf{x}_0, \ldots, \mathbf{x}_k$. This proves our lemma. □

Remark. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ from the above lemma need not form a basis of L.

Exercise 2.8. *Prove that L has a basis $\{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ such that*

$$\mathbf{v}'_i \in (\lambda_1 + \cdots + \lambda_i)C \quad \textit{for } i = 1, \ldots, n.$$

Hint. *Use Lemma 2.1.*

Minkowski's second convex body theorem gives an optimal upper and lower bound for the product of the successive minima of a central symmetric convex body C with respect to a lattice L.

Theorem 2.9 (**Minkowski's second convex body theorem, 1910).** *Let L be a lattice and C a central symmetric convex body, both in $\mathbb{R}^n$, and let $\lambda_1, \ldots, \lambda_n$ be the successive minima of C with respect to L. Then*

$$\frac{2^n}{n!} \cdot \frac{d(L)}{\operatorname{vol}(C)} \leqslant \lambda_1 \cdots \lambda_n \leqslant 2^n \cdot \frac{d(L)}{\operatorname{vol}(C)}.$$

Remark. Theorem 2.9 is invariant under linear transformations in the following sense. Let $C, L, \lambda_1, \ldots, \lambda_n$ be as in Theorem 2.9. Let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be a linear

transformation. Then $\phi(C)$ is a central symmetric convex body, $\phi(L)$ is a lattice, and one easily shows that $\lambda_1, \ldots, \lambda_n$ are the successive minima of $\phi(C)$ with respect to $\phi(L)$. Further,

$$\frac{d(\phi(L))}{\mathrm{vol}(\phi(C))} = \frac{|\det(\phi)| \cdot d(L)}{|\det(\phi)| \cdot \mathrm{vol}(C)} = \frac{d(L)}{\mathrm{vol}(C)}.$$

For every lattice L of $\mathbb{R}^n$ there is a linear transformation ϕ of $\mathbb{R}^n$ such that $\phi(L) = \mathbb{Z}^n$. This observation shows that the general Minkowski's second convex body theorem with arbitrary lattices L follows from the special case where $L = \mathbb{Z}^n$.

We show that Minkowski's second convex body theorem implies his first.

Second convex body theorem $\Rightarrow$ *First convex body theorem.* Minkowski's second convex body theorem implies that $\lambda_1^n \leqslant 2^n d(L)/\mathrm{vol}(C)$. Assume that $\mathrm{vol}(C) \geqslant 2^n d(L)$; then $\lambda_1 \leqslant 1$. Now $\lambda_1 C$ contains a non-zero point from L and $\lambda_1 C \subseteq C$; hence C contains a non-zero point from L. □

Example 1. Let B_n be the Euclidean ball in $\mathbb{R}^n$, given by $x_1^2 + \cdots + x_n^2 \leqslant 1$. Let L be a lattice in $\mathbb{R}^n$, and let $\lambda_1, \ldots, \lambda_n$ be the successive minima of B_n with respect to L. It is clear that $\mathbf{x} \in \lambda B_n$ if and only if $\|\mathbf{x}\|_2 \leqslant \lambda$, where $\|x\|_2 = \left(\sum_{j=1}^n x_j^2\right)^{1/2}$ is the Euclidean norm. There are linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in L$ with $\|\mathbf{v}_i\|_2 = \lambda_i$ for $i = 1, \ldots, n$. In fact, $\mathbf{v}_1$ is a (not necessarily unique) shortest non-zero vector in L, and for $i = 2, \ldots, n$, $\mathbf{v}_i$ is a shortest vector in L outside the linear subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$.

Now Theorem 2.9 implies that

$$\prod_{i=1}^n \|\mathbf{v}_i\|_2 \leqslant 2^n V(n)^{-1} d(L),$$

where $V(n) = \mathrm{vol}(B_n)$. Recall that $V(1) = 2$, $V(2) = \pi$, and $V(n) = \frac{2\pi}{n} V(n-2)$ for $n \geqslant 3$. We mention once more that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ need not be a basis of L.

Example 2. We prove that the constant 2^n in the upper bound of Theorem 2.9 is best possible, i.e., the theorem becomes false if 2^n is replaced by a smaller quantity. Moreover, we show that every sequence of reals $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n$ may occur as successive minima. For the lattice we take $\mathbb{Z}^n$. Let $\mathbf{e}_1 = (1, 0, \ldots, 0), \ldots, \mathbf{e}_n =$

$(0,\dots,0,1)$ be the standard basis of $\mathbb{Z}^n$. Further, let $\lambda_1,\dots,\lambda_n$ be reals with $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_n$. Define

$$C_1 := \left\{\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n : |x_i| \leqslant \lambda_i^{-1} \text{ for } i = 1,\dots,n\right\}.$$

Clearly, C_1 is a central symmetric convex body with volume $2^n(\lambda_1\cdots\lambda_n)^{-1}$. Thus, $\lambda_1\cdots\lambda_n = 2^n d(\mathbb{Z}^n)\,\mathrm{vol}(C_1)^{-1}$.

We now show that λ_j is the j-th the successive minima of C_1 with respect to $\mathbb{Z}^n$, for $j = 1,\dots,n$. For $\lambda \geqslant 0$ we have

$$\lambda C_1 = \{\mathbf{x} \in \mathbb{R}^n : |x_i| \leqslant \lambda/\lambda_i \text{ for } i = 1,\dots,n\}.$$

This implies that $\lambda_j C_1$ contains the j linearly independent points $\mathbf{e}_1,\dots,\mathbf{e}_j$. Let $\lambda < \lambda_j$ and let $\mathbf{y} = (y_1,\dots,y_n) \in \mathbb{Z}^n$ be a point in λC_1. Then $|y_j| < 1,\dots,|y_n| < 1$, implying that $y_j = \cdots = y_n = 0$. So all lattice points in λC_1 lie in the $(j-1)$-dimensional space spanned by $\mathbf{e}_1,\dots,\mathbf{e}_{j-1}$, and this space cannot contain j linearly independent points. So λ_j is the j-th successive minimum of C_1 with respect to $\mathbb{Z}^n$.

Example 3. We prove that the factor $2^n/n!$ in the lower bound of Theorem 2.9 is best possible. For our lattice we take again $\mathbb{Z}^n$. Let

$$C_2 := \left\{\mathbf{x} = (x_1,\dots,x_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \lambda_i |x_i| \leqslant 1\right\}.$$

Then C_2 is a central symmetric convex body of volume $\frac{2^n}{n!}(\lambda_1\cdots\lambda_n)^{-1}$ (verify this!). Hence $\lambda_1\cdots\lambda_n = \frac{2^n}{n!} d(L)/\,\mathrm{vol}(C_2)$.

Exercise 2.9. *Prove that $\lambda_1,\dots,\lambda_n$ are the successive minima of C_2 with respect to $\mathbb{Z}^n$.*

We deduce the lower bound for $\lambda_1\cdots\lambda_n$ in Theorem 2.9. For a proof of the upper bound, which is much more involved, we refer to the book of Cassels, Chapter 8.

We need a lemma.

Lemma 2.10. *Let $\mathbf{w}_1,\dots,\mathbf{w}_r \in \mathbb{R}^n$. Then*

$$\left\{\sum_{i=1}^{r} x_i \mathbf{w}_i : \ x_i \in \mathbb{R} \text{ for } i = 1,\dots,n, \ \sum_{i=1}^{r} |x_i| \leqslant 1\right\}$$

is the smallest convex subset in $\mathbb{R}^n$, symmetric about $\mathbf{0}$, that contains $\mathbf{w}_1, \ldots, \mathbf{w}_r$, that is, the set itself is convex and symmetric about $\mathbf{0}$, and it is contained in every other convex set which is symmetric about $\mathbf{0}$ and contains $\mathbf{w}_1, \ldots, \mathbf{w}_r$.

Exercise 2.10. *Prove this lemma.*

Proof of the lower bound in Theorem 2.9. Choose linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of L such that $\mathbf{v}_i \in \lambda_i C$ for $i = 1, \ldots, n$. Then $\lambda_i^{-1}\mathbf{v}_i \in C$ for $i = 1, \ldots, n$. Consider the set

$$D := \Big\{ \sum_{i=1}^{n} x_i \cdot \lambda_i^{-1}\mathbf{v}_i : \ x_i \in \mathbb{R} \text{ for } i = 1, \ldots, n, \ \sum_{i=1}^{n} |x_i| \leqslant 1 \Big\}.$$

By Lemma 2.10, this is the smallest symmetric convex set containing the points $\lambda_i^{-1}\mathbf{v}_i \in C$ $(i = 1, \ldots, n)$. Hence $D \subseteq C$.

Note that D is the image of the n-dimensional octahedron

$$O_n := \Big\{ \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^{n} |x_i| \leqslant 1 \Big\}$$

under the linear transformation ϕ given by $\phi(\mathbf{e}_i) = \lambda_i^{-1}\mathbf{v}_i$ for $i = 1, \ldots, n$. Hence

$$\begin{aligned} \mathrm{vol}(D) &= |\det(\phi)| \cdot \mathrm{vol}(O_n) = \frac{|\det(\mathbf{v}_1, \ldots, \mathbf{v}_n))|}{\lambda_1 \cdots \lambda_n} \cdot \frac{2^n}{n!} \\ &= \frac{d(M)}{\lambda_1 \cdots \lambda_n} \cdot \frac{2^n}{n!}, \end{aligned}$$

where M is the lattice with basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Clearly, M is a sublattice of L, therefore, $d(M) = |L : M| \cdot d(L) \geqslant d(L)$. By combining this with what we obtained above, we obtain

$$\mathrm{vol}(C) \geqslant \mathrm{vol}(D) \geqslant \frac{2^n}{n!} d(L)(\lambda_1 \cdots \lambda_n)^{-1}.$$

This implies the lower bound for $\lambda_1 \cdots \lambda_n$ from Theorem 2.9. □

We prove a weaker version of the upper bound in the special case that $C = B_n$ is the n-dimensional Euclidean unit ball. Recall that the associate norm is $\|\cdot\|_2$. In fact, we prove the following theorem which goes back to Hermite.

Theorem 2.11. *Let L be a lattice in $\mathbb{R}^n$. Then L has a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ with*

$$\|\mathbf{v}_1\|_2 \cdots \|\mathbf{v}_n\|_2 \leqslant (4/3)^{n(n-1)/4} \cdot d(L).$$

Corollary 2.12. *let $\lambda_1, \ldots, \lambda_n$ be the successive minima of B_n with respect to L. Then $\lambda_1 \cdots \lambda_n \leqslant (4/3)^{n(n-1)/4} \cdot d(L)$.*

For suppose that $\|\mathbf{v}_1\|_2 \leqslant \cdots \leqslant \|\mathbf{v}_n\|_2$. Then clearly, $\lambda_i \leqslant \|\mathbf{v}_i\|_2$ for $i = 1, \ldots, n$.

Corollary 2.13. *Let E be a central ellipsoid in $\mathbb{R}^n$ and L a lattice in $\mathbb{R}^n$. Then for the successive minima $\lambda_1, \ldots, \lambda_n$ of E with respect to L we have*

$$\lambda_1 \cdots \lambda_n \leqslant (4/3)^{n(n-1)/4} \cdot V(n) \cdot \frac{d(L)}{\mathrm{vol}(E)},$$

where $V(n) := \mathrm{vol}(B_n)$.

For this is clearly true for $E = B_n$, and the assertion for an arbitrary ellipsoid E follows by taking a linear transformation ϕ such that $E = \phi(B_n)$ and using the invariance of Corollary 2.13 under linear transformations.

In fact, by applying a theorem from 1949 of the German mathematician Fritz John, one can proceed further, and prove a weaker version of Minkowski's theorem for arbitrary central symmetric convex bodies. For a proof, we refer to Schmidt's lecture notes, p. 87 (there called 'Jordan's Theorem').

Theorem 2.14. *Let C be a central symmetric convex body in $\mathbb{R}^n$. Then there is a central ellipsoid E such that $E \subseteq C \subseteq \sqrt{n}E$.*

As has also been explained in Schmidt's lecture notes (and you should be able to prove this yourself), together with Corollary 2.13, this implies the following weaker version of Minkowski's second convex body theorem:

Corollary 2.15. *There is a number $c(n) > 0$, depending only on n, with the following property. Let C be a central symmetric convex body, and L a lattice in $\mathbb{R}^n$. Then for the successive minima $\lambda_1, \ldots, \lambda_n$ of C with respect to L we have*

$$\lambda_1 \cdots \lambda_n \leqslant c(n) \cdot \frac{d(L)}{\mathrm{vol}(C)}.$$

Proof of Theorem 2.11. We need some first year linear algebra. We proceed by induction on n. For $n = 1$ the assertion is easily verified. Let $n \geqslant 2$ and assume Theorem 2.11 is true for lattices of dimension $n - 1$.

The first successive minimum λ_1 of B_n with respect to L is the length of a shortest non-zero vector in L. Exercise 2.8 implies that L has a basis whose first vector $\mathbf{v}_1$ has $\|\mathbf{v}_1\|_2 \leqslant \lambda_1$. This necessarily shows that $\|\mathbf{v}_1\|_2 = \lambda_1$.

Put $\mathbf{e}_1 := \lambda_1^{-1}\mathbf{v}_1$. Then $\mathbf{e}_1$ has length 1. From first year linear algebra it follows that we can augment $\mathbf{e}_1$ to an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$. Then a vector $\mathbf{x} \in \mathbb{R}^n$ can be expressed uniquely as $\sum_{i=1}^n x_i\mathbf{e}_i$ with $x_i \in \mathbb{R}$ for all i, and $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$. We define a linear map

$$\rho : \mathbb{R}^n \to \mathbb{R}^{n-1} : \quad \sum_{i=1}^n x_i\mathbf{e}_i \mapsto (x_2, \ldots, x_n).$$

Define $L' := \rho(L)$. We need a few lemmas.

Lemma 2.16. *L' is a lattice in $\mathbb{R}^{n-1}$.*
More precisely, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is any basis of L containing $\mathbf{v}_1$, then $\{\rho(\mathbf{v}_2), \ldots, \rho(\mathbf{v}_n)\}$ is a basis of L'.

Proof. Left to the reader. □

Lemma 2.17. *Let $\{\mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ be a basis of L' and let $\mathbf{v}_i \in L$ with $\rho(\mathbf{v}_i) = \mathbf{v}'_i$ for $i = 2, \ldots, n$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of L.*

Proof. Let $\mathbf{x}$ be any element of L. We have to show that $\mathbf{x} = \sum_{i=1}^n z_i\mathbf{v}_i$ with $z_i \in \mathbb{Z}$ for $i = 1, \ldots, n$. Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of $\mathbb{R}^n$, as can be easily verified, we know that $\mathbf{x}$ can be expressed as such, but with all $z_i \in \mathbb{R}$. By applying ρ, we get $z_i \in \mathbb{Z}$ for $i = 2, \ldots, n$. Let m be an integer with $|z_1 - m| \leqslant \frac{1}{2}$. Then $(z_1 - m)\mathbf{v}_1 = \mathbf{x} - m\mathbf{v}_1 - \sum_{i=2}^n z_i\mathbf{v}_i \in L$. Since $\mathbf{v}_1$ is a non-zero vector of minimal length in L, we must have $z_1 = m \in \mathbb{Z}$. □

Lemma 2.18. $d(L) = \lambda_1 \cdot d(L')$.

Proof. Pick a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of L. Then $\mathbf{v}_1 = \lambda_1\mathbf{e}_1$ and $\mathbf{v}_i = \sum_{j=1}^n a_{ij}\mathbf{e}_j$ with

$a_{ij} \in \mathbb{R}$ for $i = 2, \ldots, n$. Recall that $\rho(\mathbf{v}_i) = (a_{i2}, \ldots, a_{in})$. Now we get

$$\begin{aligned} d(L) &= |\det(\mathbf{e}_1, \ldots, \mathbf{e}_n)| \cdot \left| \det \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \right| \\ &= \lambda_1 \cdot |\det(\rho(\mathbf{v}_2), \ldots, \rho(\mathbf{v}_n))| = \lambda_1 \cdot d(L'). \end{aligned}$$

□

Lemma 2.19. *Let $\mathbf{v}' \in L'$. Then there is $\mathbf{v} \in L$ with $\rho(\mathbf{v}) = \mathbf{v}'$ and $\|\mathbf{v}\|_2^2 \leqslant \frac{4}{3} \cdot \|\mathbf{v}'\|_2^2$.*

Proof. If $\mathbf{v}' = \mathbf{0}$ we may take $\mathbf{v} = \mathbf{0}$. Assume $\mathbf{v}' \neq \mathbf{0}$. Write $\mathbf{v}' = (x_2, \ldots, x_n)$. Take $\mathbf{w} \in L$ with $\rho(\mathbf{w}) = \mathbf{v}'$. Then $\mathbf{w} = x\mathbf{e}_1 + \sum_{i=2}^n x_i \mathbf{e}_i$ with $x \in \mathbb{R}$. Let m be an integer such that $|x - m\lambda_1| \leqslant \frac{1}{2}\lambda_1$ and put $\mathbf{v} := \mathbf{w} - m\mathbf{v}_1$, $x_1 := x - m\lambda_1$. Then $\mathbf{v} \in L$, $\rho(\mathbf{v}) = \mathbf{v}'$ and $\mathbf{v} = \sum_{i=1}^n x_i \mathbf{e}_i$ with $|x_1| \leqslant \frac{1}{2}\lambda_1 = \frac{1}{2}\|\mathbf{v}_1\|_2$. Now using that $\mathbf{v}_1$ is a vector of minimal length in L we get

$$\|\mathbf{v}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2 = x_1^2 + \|\mathbf{v}'\|^2 \leqslant \tfrac{1}{4}\|\mathbf{v}_1\|_2^2 + \|\mathbf{v}'\|_2^2 \leqslant \tfrac{1}{4}\|\mathbf{v}\|_2^2 + \|\mathbf{v}'\|_2^2.$$

Hence $\frac{3}{4} \cdot \|\mathbf{v}\|_2^2 \leqslant \|\mathbf{v}'\|_2^2$. □

Completion of the induction step. By the induction hypothesis, L' has a basis $\{\mathbf{v}'_2, \ldots, \mathbf{v}'_n\}$ such that

$$\prod_{i=2}^n \|\mathbf{v}'_i\|_2 \leqslant (4/3)^{(n-1)(n-2)/4} \cdot d(L').$$

By Lemma 2.19, for $i = 2, \ldots, n$, there exists $\mathbf{v}_i \in L$ such that $\rho(\mathbf{v}_i) = \mathbf{v}'_i$ and $\|\mathbf{v}_i\|_2 \leqslant \sqrt{\frac{4}{3}}\|\mathbf{v}'_i\|_2$. By Lemma 2.17, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of L. For this basis we have

$$\begin{aligned} \prod_{i=1}^n \|\mathbf{v}_i\|_2 &\leqslant (4/3)^{(n-1)/2}\|\mathbf{v}_1\|_2 \cdot \prod_{i=2}^n \|\mathbf{v}'_i\|_2 \\ &\leqslant (4/3)^{(n-1)/2 + (n-1)(n-2)/4} \cdot \lambda_1 \cdot d(L') = (4/3)^{n(n-1)/4} d(L), \end{aligned}$$

where in the last step we used Lemma 2.18. □

Exercise 2.11. *Prove Hadamard's inequality: if L is a lattice in $\mathbb{R}^n$ with basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, then $\|\mathbf{v}_1\|_2 \cdots \|\mathbf{v}_n\|_2 \geqslant d(L)$.*
Hint. *From the Gram-Schmidt orthogonalization process, one computes an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ such that $\mathbf{v}_i = \sum_{j=1}^{i} a_{ij}\mathbf{e}_j$ with $a_{ij} \in \mathbb{R}$ for $i = 1, \ldots, n$, $j = 1, \ldots, i$.*

Exercise 2.12. *Let $l_i = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n$ $(i = 1, \ldots, n)$ be linear forms with real coefficients and with $\det(l_1, \ldots, l_n) \neq 0$. Let $A_1, \ldots, A_n$ be positive reals. Denote by $\lambda_1, \ldots, \lambda_n$ the successive minima of the central symmetric convex body*

$$C := \{\mathbf{x} \in \mathbb{R}^n : |l_1(\mathbf{x})| \leqslant A_1, \ldots, |l_n(\mathbf{x})| \leqslant A_n\}$$

with respect to $\mathbb{Z}^n$. Using Theorem 2.11, prove that

$$\lambda_1 \cdots \lambda_n \leqslant (4/3)^{n(n-1)/4} \frac{|\det(l_1, \ldots, l_n)|}{A_1 \cdots A_n}.$$

For many applications, for instance to factorization of polynomials, cryptography, determining all solutions of Diophantine equations from certain classes, it is desirable to have a computationally efficient algorithm, which for a given lattice computes a basis such as in Theorem 2.11. In 1982, Arjen Lenstra, Hendrik Lenstra and László Lovász developed a very efficient, fundamental algorithm, now known as the *LLL lattice basis reduction algorithm* which, from input an arbitrary basis of a given lattice L, computes a so-called LLL-reduced basis of L. Such a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ has various properties, among which

$$\|\mathbf{v}_1\|_2 \cdots \|\mathbf{v}_n\|_2 \leqslant 2^{n(n-1)/4} \cdot d(L).$$

So in certain respects it is slightly worse than the one from Theorem 2.11, but good enough for most purposes. For more information on the LLL-algorithm, see for instance the paper where it was introduced, A.K. Lenstra, H.W. Lenstra, L. Lovász, Factoring polynomials with rational coefficients, Mathematische Annalen 261 (1982), 515–534.

2.4 Polar lattices

Henceforth, vectors in $\mathbb{R}^n$ will be column vectors, unless otherwise stated. By A^T we denote the transpose of a matrix A. As usual, the standard inner product of

$\mathbf{x} = (x_1, \ldots, x_n)^T, \mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ is given by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = \mathbf{x}^T \mathbf{y}$. We denote as before by B_n the n-dimensional Euclidean ball given by $\|\mathbf{x}\|_2 = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} \leqslant 1$. We will need the *Cauchy-Schwarz inequality:*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Let L be a lattice in $\mathbb{R}^n$ The *polar* (or *reciprocal*) of L is given by

$$L^* := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z} \text{ for all } \mathbf{y} \in L\}.$$

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of L, and V the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Thus, $L = \{V\mathbf{z} : \mathbf{z} \in \mathbb{Z}^n\}$. Using $\langle \mathbf{x}, V\mathbf{y} \rangle = \langle V^T\mathbf{x}, \mathbf{y} \rangle$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we obtain

$$\begin{aligned} L^* &= \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, V\mathbf{z} \rangle \in \mathbb{Z} \ \forall \mathbf{z} \in \mathbb{Z}^n\} \\ &= \{\mathbf{x} \in \mathbb{R}^n : \langle V^T\mathbf{x}, \mathbf{z} \rangle \in \mathbb{Z} \ \forall \mathbf{z} \in \mathbb{Z}^n\} \\ &= \{\mathbf{x} \in \mathbb{R}^n : V^T\mathbf{x} \in \mathbb{Z}^n\} = \{(V^T)^{-1}\mathbf{w} : \mathbf{w} \in \mathbb{Z}^n\}. \end{aligned}$$

Hence L^* is a lattice in $\mathbb{R}^n$, with basis the columns $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ of $(V^T)^{-1}$. Note that

$$d(L^*) = |\det(V^T)^{-1}| = |\det V|^{-1} = d(L)^{-1}.$$

Theorem 2.20. *Let L be a lattice in $\mathbb{R}^n$ and L^* its polar. Further, let $\lambda_1, \ldots, \lambda_n$ be the successive minima of B_n with respect to L, and $\lambda_1^*, \ldots, \lambda_n^*$ the successive minima of B_n with respect to L^*. Then*

$$1 \leqslant \lambda_i \lambda_{n+1-i}^* \leqslant c(n) \quad \textit{for } i = 1, \ldots, n,$$

where $c(n)$ depends only on n.

Remark. If we use Minkowski's second convex body theorem, we obtain the above theorem with $c(n) = 4^n \operatorname{vol}(B_n)^{-2}$. If we use instead Theorem 2.11, we can prove the above theorem with $c(n) = (4/3)^{n(n-1)/2}$. For the application we have in mind, the precise value of $c(n)$ doesn't matter.

Proof. We first deduce the lower bound for $\lambda_i \lambda_{n+1-i}^*$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be linearly independent vectors from L such that $\mathbf{v}_i \in \lambda_i B_n$, i.e., $\|\mathbf{v}_i\|_2 = \lambda_i$ for $i = 1, \ldots, n$. Likewise, let $\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*$ be linearly independent vectors from L^* such that $\|\mathbf{v}_i^*\|_2 = \lambda_i^*$ for $i = 1, \ldots, n$.

Take $i \in \{1, \ldots, n\}$, and consider the set of vectors

$$\{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{v}_k, \mathbf{x} \rangle = 0 \text{ for } k = 1, \ldots, i\}.$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_i$ are linearly independent, this is a linear subspace of $\mathbb{R}^n$ of dimension $n - i$. Hence at least one of the vectors $\mathbf{v}_1^*, \ldots, \mathbf{v}_{n+1-i}^*$ does not lie in this space. It follows that there are indices $k \leqslant i$, $l \leqslant n + 1 - i$, such that $\langle \mathbf{v}_k, \mathbf{v}_l^* \rangle \neq 0$.

But $\langle \mathbf{v}_k, \mathbf{v}_l^* \rangle \in \mathbb{Z}$, since $\mathbf{v}_k \in L$, $\mathbf{v}_l^* \in L^*$. Hence $|\langle \mathbf{v}_k, \mathbf{v}_l^* \rangle| \geqslant 1$. Now by the Cauchy-Schwarz inequality,

$$1 \leqslant |\langle \mathbf{v}_k, \mathbf{v}_l^* \rangle| \leqslant \|\mathbf{v}_k\|_2 \|\mathbf{v}_l^*\|_2 \leqslant \lambda_k \lambda_l^* \leqslant \lambda_i \lambda_{n+1-i}^*.$$

This establishes the lower bound. To prove the upper bound, recall that by Theorem 2.11, we have

$$\lambda_1 \cdots \lambda_n \leqslant c'(n) d(L), \quad \lambda_1^* \cdots \lambda_n^* \leqslant c'(n) d(L^*),$$

where $c'(n)$ depends on n only. Further, $d(L^*) = d(L)^{-1}$. Hence

$$\prod_{i=1}^{n} (\lambda_i \lambda_{n+1-i}^*) \leqslant c'(n)^2 d(L) d(L^*) = c'(n)^2 =: c(n).$$

It follows that for $i = 1, \ldots, n$,

$$\lambda_i \lambda_{n+1-i}^* \leqslant \frac{c(n)}{\prod_{j \neq i} \lambda_j \lambda_{n+1-j}^*} \leqslant c(n).$$

This proves Theorem 2.20. □

As an application we show that if the polar lattice L^* does not have small non-zero vectors, then every point of $\mathbb{R}^n$ can be approximated closely by a point from L.

Corollary 2.21. *Let L be a lattice of $\mathbb{R}^n$ and $R > 0$. Assume that $\|\mathbf{y}\|_2 \geqslant R$ for every non-zero $\mathbf{y} \in L^*$. Then for every $\mathbf{b} \in \mathbb{R}^n$ there is $\mathbf{x} \in L$ such that $\|\mathbf{x} - \mathbf{b}\|_2 \leqslant n \cdot c(n)/2R$.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the successive minima of B_n with respect to L, and let $\lambda_1^*, \ldots, \lambda_n^*$ be the successive minima of B_n with respect to L^*. By assumption,

$\lambda_1^* \geqslant R$. So by Theorem 2.20, $\lambda_n \leqslant c(n)/R$. Choose linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in L$ with $\|\mathbf{v}_i\|_2 = \lambda_i$ for $i = 1, \ldots, n$. So $\|\mathbf{v}_i\|_2 \leqslant \lambda_n \leqslant c(n)/R$ for $i = 1, \ldots, n$.

Let $\mathbf{b} \in \mathbb{R}^n$. Then $\mathbf{b} = \xi_1\mathbf{v}_1 + \cdots + \xi_n\mathbf{v}_n$ with $\xi_1, \ldots, \xi_n \in \mathbb{R}$. There exist integers $z_1, \ldots, z_n$ with $|z_i - \xi_i| \leqslant \frac{1}{2}$ for $i = 1, \ldots, n$. Put $\mathbf{x} := z_1\mathbf{v}_1 + \cdots + z_n\mathbf{v}_n$. Then $\mathbf{x} \in L$ and

$$\begin{aligned} \|\mathbf{x} - \mathbf{b}\|_2 &= \|(z_1 - \xi_1)\mathbf{v}_1 + \cdots + (z_n - \xi_n)\mathbf{v}_n\|_2 \leqslant \tfrac{1}{2}(\|\mathbf{v}_1\|_2 + \cdots + \|\mathbf{v}_n\|_2) \\ &\leqslant n \cdot c(n)/2R. \end{aligned}$$

□

2.5 Kronecker's approximation theorem

Recall that by Dirichlet's Theorem, if $\alpha_1, \ldots, \alpha_n$ are real numbers of which at least one is irrational, then there are infinitely many tuples of integers $x_1, \ldots, x_n, y$ such that

$$|\alpha_i - x_i/y| \leqslant y^{-1-1/n} \text{ for } i = 1, \ldots, n, \quad y > 0.$$

This implies that for every $\varepsilon > 0$, there exists $(x_1, \ldots, x_n, y) \in \mathbb{Z}^{n+1}$ such that

$$|\alpha_i y - x_i| \leqslant \varepsilon \text{ for } i = 1, \ldots, n, \quad y > 0.$$

Kronecker's approximation theorem deals with systems of inhomogeneous inequalities of the shape

$$|\alpha_i y - x_i - \theta_i| \leqslant \varepsilon \ (i = 1, \ldots, n) \quad \text{in } x_1, \ldots, x_n, y \in \mathbb{Z} \tag{2.7}$$

where $\theta_1, \ldots, \theta_n$ are any real numbers.

Theorem 2.22. *Let $\alpha_1, \ldots, \alpha_n, \theta_1, \ldots, \theta_n$ be real numbers. Suppose that $1, \alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$. Then for every $\varepsilon > 0$ there are infinitely many $(x_1, \ldots, x_n, y) \in \mathbb{Z}^{n+1}$ with (2.7).*

Remark. The condition that $1, \alpha_1, \ldots, \alpha_n$ be linearly independent over $\mathbb{Q}$ can not be removed. For suppose that $1, \alpha_1, \ldots, \alpha_n$ are linearly dependent over $\mathbb{Q}$. Then

there are integers $a_1, \dots, a_n, a_0$, not all 0, such that $a_1\alpha_1 + \cdots + a_n\alpha_n = a_0$. In fact, at least one of $a_1, \dots, a_n$ is non-zero. Choose $\theta_1, \dots, \theta_n \in \mathbb{R}$ such that

$$a_1\theta_1 + \cdots + a_n\theta_n \notin \mathbb{Z}.$$

Let δ be the distance from $a_1\theta_1 + \cdots + a_n\theta_n$ to the nearest integer. We show that for sufficiently small $\varepsilon > 0$, (2.7) is not solvable. Indeed, suppose (2.7) is solvable and let $(x_1, \dots, x_n, y) \in \mathbb{Z}^{n+1}$ be a solution. Then

$$\Big|\sum_{i=1}^{n} a_i(\alpha_i y - x_i - \theta_i)\Big| \leqslant \sum_{i=1}^{n} |a_i| \cdot |\alpha_i y - x_i - \theta_i| \leqslant \varepsilon \sum_{i=1}^{n} |a_i|.$$

But on the other hand,

$$\Big|\sum_{i=1}^{n} a_i(\alpha_i y - x_i - \theta_i)\Big| = \Big|a_0 y - \sum_{i=1}^{n} a_i x_i - \sum_{i=1}^{n} a_i\theta_i\Big| \geqslant \delta.$$

Hence (2.7) is unsolvable for $\varepsilon < \delta / \sum_{i=1}^{n} |a_i|$.

Proof of Theorem 2.22. We apply Corollary 2.21 with an astutely chosen lattice. Let M be a large positive integer, to be chosen later. Consider the lattice in $\mathbb{R}^{n+1}$,

$$\begin{aligned} L_M &:= \big\{\big(x_1 - \alpha_1 y, \dots, x_n - \alpha_n y, M^{-1}y\big) : x_1, \dots, x_n, y \in \mathbb{Z}\big\} \\ &= \{A\mathbf{z} : \mathbf{z} \in \mathbb{Z}^{n+1}\}, \end{aligned}$$

where

$$A = \begin{pmatrix} 1 & & 0 & -\alpha_1 \\ & \ddots & & \vdots \\ 0 & & 1 & -\alpha_n \\ 0 & & 0 & M^{-1} \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y \end{pmatrix}.$$

Put

$$\mathbf{b} := (\theta_1, \dots, \theta_n, 2M\varepsilon)^T.$$

We want to show for appropriate M that there is $\mathbf{u} \in L_M$ such that $\|\mathbf{u} - \mathbf{b}\|_2 \leqslant \varepsilon$. Writing $\mathbf{u} = A\mathbf{z}$ with $\mathbf{z} = (x_1, \dots, x_n, y)^T \in \mathbb{Z}^{n+1}$, this translates into

$$\Big(\sum_{i=1}^{n}(x_i - \alpha_i y - \theta_i)^2 + M^{-2}(y - 2M\varepsilon)^2\Big)^{1/2} \leqslant \varepsilon,$$

and this certainly implies that $x_1, \ldots, x_n, y$ satisfy (2.7) and moreover that $|y - 2M\varepsilon| \leqslant M\varepsilon$. The latter implies that $y \geqslant M\varepsilon$. If we can choose M arbitrarily large, then it follows that (2.7) has solutions with arbitrarily large values of y, and thus, that (2.7) has infinitely many solutions.

By Corollary 2.21, we have to show that we can choose arbitrarily large M in such a way that every non-zero vector in the polar lattice L_M^* has length at least $R := (n+1) \cdot c(n+1)/2\varepsilon$.

It is easy to verify that

$$(A^T)^{-1} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ M\alpha_1 & \ldots & M\alpha_n & M \end{pmatrix}.$$

Hence

$$\begin{aligned} L_M^* &= \left\{(A^T)^{-1}\mathbf{z} : \mathbf{z} \in \mathbb{Z}^{n+1}\right\} \\ &= \left\{\big(x_1, \ldots, x_n, M(\alpha_1 x_1 + \cdots + \alpha_n x_n + y)\big) : x_1, \ldots, x_n, y \in \mathbb{Z}\right\}. \end{aligned}$$

Let μ be the minimum of all numbers $|\alpha_1 x_1 + \cdots + \alpha_n x_n + y|$, taken over all integers $x_1, \ldots, x_n, y$ such that

$$\begin{aligned} &|x_i| < R \text{ for } i = 1, \ldots, n, \quad |\alpha_1 x_1 + \cdots + \alpha_n x_n + y| \leqslant 1, \\ &(x_1, \ldots, x_n, y) \neq \mathbf{0}. \end{aligned}$$

Then μ is the minimum of finitely many real numbers which are all positive, since $1, \alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$. Hence $\mu > 0$.

Now let M be any integer with

$$M \geqslant \max(R/\mu,\ R). \tag{2.8}$$

Then for every non-zero $\mathbf{u} \in L_M^*$ we have indeed

$$\|\mathbf{u}\|_2 \geqslant R$$

since at least one of the numbers $x_1, \ldots, x_n$, $M(\alpha_1 x_1 + \cdots + \alpha_n x_n + y)$ has absolute value at least R. Further, M can be chosen arbitrarily large. This completes our proof. □

In fact, Kronecker proved a much more general approximation theorem, of which Theorem 2.22 is just a special case. As usual, $\|\mathbf{x}\|_2$ denotes the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^n$.

Theorem 2.23 (Kronecker, 1887)**.** *Let A be an $m \times n$-matrix with real entries, and $\mathbf{b} \in \mathbb{R}^m$ a column vector. Then the following two assertions are equivalent:*

(i) For every $\mathbf{y} \in \mathbb{R}^m$ with $A^T\mathbf{y} \in \mathbb{Z}^n$ we have $\langle \mathbf{b}, \mathbf{y} \rangle \in \mathbb{Z}$;

(ii) For every $\varepsilon > 0$ there is $\mathbf{z} \in \mathbb{Z}^n$ such that

$$\|A\mathbf{z} - \mathbf{b}\|_2 \leqslant \varepsilon.$$

For a proof, we refer to Siegel, Chapter II.

Exercise 2.13. *a) Prove (ii)$\Longrightarrow$(i).*
b) Deduce Theorem 2.22 from Theorem 2.23.

2.6 Further exercises

Exercise 2.14. *Let p be a prime number with $p \equiv 1 \pmod 4$.*

(i) Prove that there is an integer x_0 with $x_0^2 \equiv -1 \pmod p$.
Hint. *You may use that the group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic. Prove that it has an element of order 4.*

(ii) Let L be the lattice $\{(x, y) \in \mathbb{Z}^2 : x \equiv x_0 y \pmod p\}$. Prove that $x^2 + y^2 \equiv 0 \pmod p$ for $(x, y) \in L$.

(iii) Apply Minkowski's theorem with $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant A\}$ for appropriate A and the lattice L from b) and deduce that there is $(x, y) \in \mathbb{Z}^2$ with $x^2 + y^2 = p$.

Exercise 2.15. *(Dirichlet's theorem for Gaussian numbers). Let ζ be a complex number not belonging to the field $\mathbb{Q}(i) = \{x + iy : x, y \in \mathbb{Q}\}$. Prove that there are infinitely many pairs $(z, w) \in \mathbb{Z}[i] \times \mathbb{Z}[i]$ (where $\mathbb{Z}[i] = \{x+iy : x, y \in \mathbb{Z}\}$) such that*

$$|z - \zeta w| \leqslant \frac{4}{\pi}|w|^{-1}, \ w \neq 0.$$

Hint. *Prove that for every integer $Q \geqslant 2$ there are $z, w \in \mathbb{Z}[i]$ with*

$$|z - \zeta w| \leqslant \tfrac{4}{\pi}Q^{-1}, \quad |w| \leqslant Q, \quad w \neq 0.$$

To this end, consider the lattice in $\mathbb{R}^4$,

$$\big\{(\operatorname{Re}(z-\zeta w), \operatorname{Im}(z-\zeta w), \operatorname{Re} w, \operatorname{Im} w) : z, w \in \mathbb{Z}[i]\big\}$$

and apply Minkowski's first convex body theorem.

Exercise 2.16. *Determine the two successive minima of*

$$C := \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : |x_1 - \sqrt{2}x_2| \leqslant 1, |x_1 - \sqrt{3}x_2| \leqslant 1\}$$

with respect to $\mathbb{Z}^2$.

Exercise 2.17. *(i) Deduce the following result from Theorem 2.22.*
Let $\alpha_1, \ldots, \alpha_n$ *be real numbers, linearly independent over* $\mathbb{Q}$. *Then for every* $t_0 > 0$, $\varepsilon > 0$, $\theta_1, \ldots, \theta_n \in \mathbb{R}$, *there are* $t \in \mathbb{R}$ *with* $t > t_0$, *and* $x_1, \ldots, x_n \in \mathbb{Z}$, *such that*

$$|\alpha_1 t - x_1 - \theta_1| < \varepsilon, \ldots, |\alpha_n t - x_n - \theta_n| < \varepsilon$$

(so compared with Theorem 2.22, we have weakened the condition that $\{1, \alpha_1, \ldots, \alpha_n\}$ *be linearly independent over* $\mathbb{Q}$ *to* $\{\alpha_1, \ldots, \alpha_n\}$ *linearly independent over* $\mathbb{Q}$, *but instead of an unknown* y *assuming integer values we have an unknown* t *assuming real values).*
Hint. *Write* $t = (x_1 + \theta_1)/\alpha_1$ *with* $x_1 \in \mathbb{Z}_{>0}$ *so that the first inequality is satisfied and substitute this into the other inequalities.*

(ii) A star has n *planets, all whose orbits are circular with the star in the center and lie in the same plane. Each planet has a constant angular velocity with which it traverses its orbit. Prove that the planets are in almost the same direction infinitely often (i.e., for every* $\varepsilon > 0$ *there are arbitrarily large* t *such that at time* t, *seen from the star the directions of the planets are within an angle* $\varepsilon > 0$ *from each other) in each of the following two cases:*
(a) they once have been in the same direction;
(b) their angular velocities are linearly independent over $\mathbb{Q}$.