Lecture Notes
on
Geometry of Numbers

Martin Henk

May 27, 2015, 9:53am

Contents

Once upon a time...

Acknowledgement. For helpful remarks, corrections and discussions I would like to thank André Apitzsch, Eva Linke, María de los Angeles Hernández Cifre, Eva Linke,....

1

Basic (and some more) facts from Convexity

In this chapter we will briefly collect some basic results and concepts concerning convex bodies, which may be considered as the main "continuous" component/ingredient of the Geometry of Numbers. For proofs of the results we will refer to the literature, mainly to the books [2, 5, 6, 7, 9, 11, 13, 14], which are also excellent sources for further information on the state of the art concerning convex bodies, polytopes and their applications. As usual we will start with setting up some basic notation; more specific ones will be introduced when needed.

Let $\mathbb{R}^n = \{\mathbf{x} = (x_1, \ldots, x_n)^\intercal : x_i \in \mathbb{R}\}$ be the *n-dimensional Euclidean space*, equipped with the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i\, y_i$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

For a subset $X \subseteq \mathbb{R}^n$, let $\operatorname{lin} X$ and $\operatorname{aff} X$ be its *linear* and *affine hull*, respectively, i.e., with respect to inclusions, the smallest linear or affine subspace containing X. If $X = \emptyset$ we set $\operatorname{lin} X = \{\mathbf{0}\}$, and $\operatorname{aff} X = \emptyset$. The *dimension* of a set X is the dimension of its affine hull and will be denoted by $\dim X$. By $X + Y$ we mean the vectorwise addition of subsets $X, Y \subseteq \mathbb{R}^n$, i.e.,

$$X + Y = \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in X,\, \mathbf{y} \in Y\}.$$

The multiplication $\lambda\, X$ of a set $X \subseteq \mathbb{R}^n$ with $\lambda \in \mathbb{R}$ is also declared vectorwise, i.e., $\lambda\, X = \{\lambda\, \mathbf{x} : \mathbf{x} \in X\}$. We write $-X = (-1)\, X$, $X - Y = X + (-1)\, Y$ and $\mathbf{x} + Y$ instead of $\{\mathbf{x}\} + Y$.

A subset $C \subseteq \mathbb{R}^n$ is called *convex*, iff for all $\mathbf{c}_1, \mathbf{c}_2 \in C$ and for all $\lambda \in [0, 1]$ hold $\lambda\, \mathbf{c}_1 + (1 - \lambda)\, \mathbf{c}_2 \in C$. A convex compact (i.e., bounded and closed) subset $K \subset \mathbb{R}^n$ is called a *convex body*. The set of all convex bodies in $\mathbb{R}^n$ is denoted by $\mathcal{K}^n$. $K \in \mathcal{K}^n$ is called *o-symmetric*, i.e., symmetric with respect to the origin, if $K = -K$. The set of all o-symmetric convex bodies in $\mathbb{R}^n$ is denoted by $\mathcal{K}^n_o$.

There is a simple one-to-one correspondence between n-dimensional o-symmetric convex bodies and norms on $\mathbb{R}^n$: given a norm $|\cdot| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$

we consider its *unit ball* $\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| \leq 1\}$, which is a o-symmetric convex body by the definition of a norm. For the reverse correspondence we need the notion of a distance function. For $K \in \mathcal{K}_o^n$, $\dim K = n$, the function

$$|\cdot|_K : \mathbb{R}^n \to \mathbb{R}_{\geq 0} \text{ given by } |\mathbf{x}|_K := \min\{\rho \in \mathbb{R}_{\geq 0} : \mathbf{x} \in \rho\, K\}$$

is called the *distance function* of K.

By the convexity and symmetry of K it is easily verified that $|\cdot|_K$ is indeed a norm. A well-studied family of o-symmetric convex bodies are the unit balls $B_n^p := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}|_p \leq 1\}$ associated to the p-norms

$$|\mathbf{x}|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, \; p \in [1, \infty), \text{ and } |\mathbf{x}|_\infty := \max_{1 \leq i \leq n} \{|x_i|\}.$$

In the Euclidean case, i.e., $p = 2$, we denote the *Euclidean norm* just by $|\cdot|$ and its unit ball by B_n.

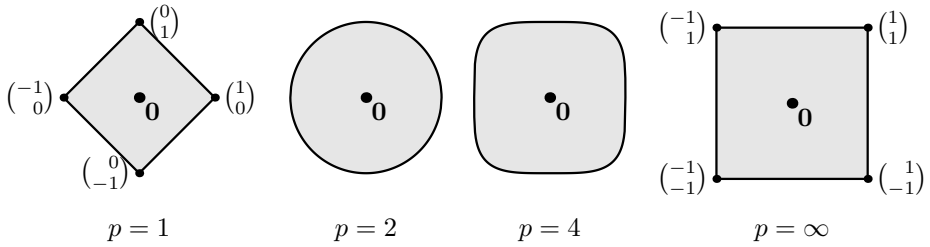

Figure 1.1: Unit balls of p-norms

$B_n^\infty = \{\mathbf{x} \in \mathbb{R}^n : -1 \leq x_i \leq 1\}$ is the o-symmetric *(regular) cube* with edge length 2, and $B_n^1 = \{\mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n |x_i| \leq 1\}$ is called the *(regular) crosspolytope*.

One big advantage of convex sets K is their property that points not belonging to K can easily be separated from K. In order to formulate this precisely and a bit more generally, we denote for $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $b \in \mathbb{R}$ by $H(\mathbf{a}, b) = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} \rangle = b\}$ the *hyperplane* with *normal vector* $\mathbf{a}$ and *right hand side* b. The associated two closed *halfspaces* are denoted by $H_\leq(\mathbf{a}, b)$ and $H_\geq(\mathbf{a}, b)$, respectively, i.e., $H_\leq(\mathbf{a}, b) = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, x \rangle \leq b\}$ and $H_\geq(\mathbf{a}, b) = H_\leq(-\mathbf{a}, -b)$.

1.1 Theorem [Separation Theorem]. *Let $C_1, C_2 \subset \mathbb{R}^n$ be convex with $C_1 \cap C_2 = \emptyset$.*

i) *Then there exists a* separating hyperplane $H(\mathbf{a}, b)$, *i.e., $C_1 \subseteq H_\leq(\mathbf{a}, b)$ and $C_2 \subseteq H_\geq(\mathbf{a}, b)$.*

ii) *If C_1 is compact, i.e., $C_1 \in \mathcal{K}^n$ and C_2 is closed then there exists a* strictly separating hyperplane $H(\mathbf{a}, b)$, *i.e., $C_1 \subset H_\leq(\mathbf{a}, b)$ and $C_2 \subseteq H_\geq(\mathbf{a}, b)$ and $C_1 \cap H(\mathbf{a}, b) = \emptyset = C_2 \cap H(\mathbf{a}, b)$.*

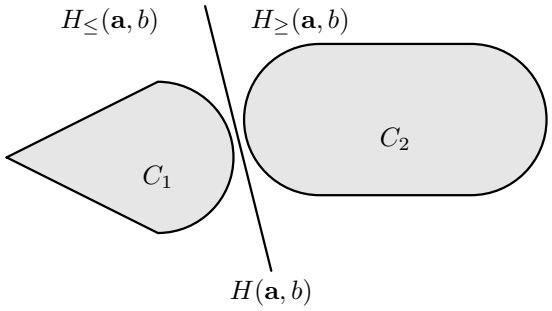

Figure 1.2: Strictly separating hyperplane of two convex bodies

A hyperplane $H(\mathbf{a}, b)$ passing through the boundary of a closed convex set C and containing C entirely in one of the two halfspaces, i.e., $C \cap H(\mathbf{a}, b) \neq \emptyset$ and $C \subseteq H_{\leq}(\mathbf{a}, b)$, say, is called a *supporting hyperplane* of C. The intersection of $F = C \cap H(\mathbf{a}, b)$ is called a *face* of C, or more precisely, a k-face if $\dim F = k$. If $k = 0$ then F is called an *extreme point* of C. Faces themselves are closed convex sets, and the interesection of faces of C is again a face of C, where we regard the empty set as the (-1)-dimensional face of C. The boundary of C, denoted by $\operatorname{bd} C$, is the union of its faces.

For a convex body $K \in \mathcal{K}^n$ and a given normal vector (direction) $\mathbf{a} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ the supporting hyperplane in direction $\mathbf{a}$ is given by $H(\mathbf{a}, h_K(\mathbf{a}))$ where

$$h_K : \mathbb{R}^n \to \mathbb{R} \text{ with } h_K(\mathbf{a}) := \max\{\langle \mathbf{a}, \mathbf{x} \rangle : \mathbf{x} \in K\}$$

is called the *support function* of K. The support function is *positive homogeneous*, i.e., $h_K(\lambda\, \mathbf{a}) = \lambda\, h_K(\mathbf{a})$, $\lambda \in \mathbb{R}_{\geq 0}$, and *sub-additive*, i.e., $h_K(\mathbf{x} + \mathbf{y}) \leq h_K(\mathbf{x}) + h_K(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Indeed, these two properties characterize support functions of convex bodies among the real-valued functions on $\mathbb{R}^n$.

A quite general concept in different branches of mathematics is that of duality and/or polarity. In our setting it is defined as follows: for $X \subseteq \mathbb{R}^n$, the set

$$X^\star := \left\{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{y} \rangle \leq 1 \text{ for all } \mathbf{x} \in X\right\}$$

is called the *polar set* of X. As the intersection of convex closed sets containing the origin, the polar set is always a convex closed set containing $\mathbf{0}$. Its actual shape, however, depends on its position relative to the origin. For instance, $X^\star$ is bounded if and only if $\mathbf{0}$ is an interior point of X.

On the other hand, as stated in the next proposition, its shape behaves properly with respect to linear transformations and a polar set of a o-symmetric convex body is again an o-symmetric convex body.

1.2 Proposition.

i) *Let $X \subseteq \mathbb{R}^n$, and let $M \in \mathrm{GL}(n, \mathbb{R})$. Then $(MX)^\star = M^{-\intercal} X^\star$. Here $\mathrm{GL}(n, \mathbb{R})$ denotes the* general linear group *of all invertible real $n \times n$*

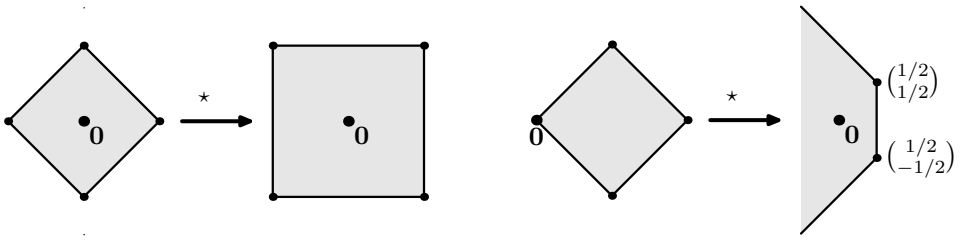

Figure 1.3: Polar bodies of B_2^1 and $\binom{1}{0} + B_2^1$

matrices.

ii) *Let $K \subseteq \mathbb{R}^n$ be convex and closed with $\mathbf{0} \in K$. Then $(K^\star)^\star = K$.*

iii) *Let $K \in \mathcal{K}_o^n$. Then $K^\star \in \mathcal{K}_o^n$ and $|\mathbf{x}|_{K^\star} = \max_{\mathbf{y}\in K} \langle \mathbf{x}, \mathbf{y}\rangle = h_K(\mathbf{x})$.*

By Proposition 1.2 iii) and Hölder's inequality for sums,

$$\sum_{i=1}^{n} |x_k\, y_k| \le |\mathbf{x}|_p\, |\mathbf{y}|_{p/(p-1)}\,, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\, p \in (1,\infty),$$

with equality if and only if $|x_i| = \mathrm{c}\,|y_i|^{p-1}$ for a constant c, the polar bodies of the p-norms are given by

$$(B_n^p)^\star = B_n^{\frac{p}{p-1}}, 1 < p < \infty, \text{ and } (B_n^1)^\star = B_n^\infty \text{ and thus } (B_n^\infty)^\star = B_n^1.$$

Since we are mainly dealing with simple structured bounded convex sets, we define the volume of sets via the Jordan-Peano measure: A subset $D \subset \mathbb{R}^n$ is called *(Jordan-Peano) measurable if its* indicator function $\chi_D : \mathbb{R}^n \to \{0,1\}$ *given by*

$$\chi_D(\mathbf{x}) := \begin{cases} 1, & \mathbf{x} \in D, \\ 0, & \mathbf{x} \notin D, \end{cases}$$

is Riemann integrable, i.e., the Riemann integral

$$\mathrm{vol}(D) := \int_{\mathbb{R}^n} \chi_D(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

exists. $\mathrm{vol}(D)$ is called the *volume* of D. So the volume of a set D is based on the principle of exhausting or approximating D by "small" n-dimensional boxes $\{\mathbf{x} \in \mathbb{R}^n : \alpha_i \le x_i \le \beta_i,\, 1 \le i \le n\}$, $\alpha_i < \beta_i \in \mathbb{R}$, for which the volume is defined as $\prod_{i=1}^n (\beta_i - \alpha_i)$. In particular, we have $\mathrm{V}(X) = 0$ for a measureable set of $\dim X < n$, and

1.3 Proposition. *Let $X \subset \mathbb{R}^n$ be a measurable set. Then*

$$\mathrm{vol}(X) = \lim_{m\to\infty} \frac{\#(X \cap \frac{1}{m}\mathbb{Z}^n)}{m^n},$$

where $\mathbb{Z}^n = \{\mathbf{z} \in \mathbb{R}^n : z_i \in \mathbb{Z}\}$ *denotes the set of all points with integral coordinates.*

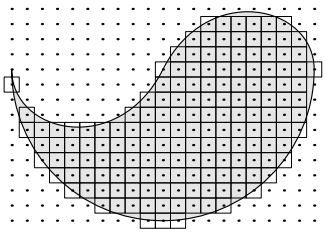

Figure 1.4: Approximation by "small" boxes centered at points of $\frac{1}{m}\mathbb{Z}^n$

The unit ball B_n^∞, i.e., the cube (box) of edge length 2 centered at the origin has $\mathrm{vol}(B_n^\infty) = 2^n$. In general, the volume of the unit balls B_n^p can be expressed via the Γ-function $\Gamma(x) := \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$ as (cf. Exercise 1.3)

$$\mathrm{vol}(B_n^p) = 2^n \frac{\Gamma(1+1/p)^n}{\Gamma(1+n/p)}, \tag{1.1}$$

where the $p = \infty$ case is covered by the limit of the formula above.

As a function $\mathrm{vol} : \mathcal{K}^n \to \mathbb{R}_{\geq 0}$ on the set of convex bodies, the volume shares the following properties: first of all it is *valuation*, i.e., it is an *additive* function on the space of all convex bodies, which means that for $K, L \in \mathcal{K}^n$ with $K \cup L \in \mathcal{K}^n$ it is $\mathrm{vol}(K \cup L) = \mathrm{vol}(K) + \mathrm{vol}(L) - \mathrm{vol}(K \cap L)$. The volume is *translation invariant*, i.e., $\mathrm{vol}(t+X) = \mathrm{vol}(X)$ for any measurable set X and $t \in \mathbb{R}^n$, and since $\mathrm{vol}(A\,X) = |\det A|\mathrm{vol}(X)$ for $A \in \mathrm{GL}(n, \mathbb{R})$ the volume is also *invariant with respect to rotations*, i.e., $\mathrm{vol}(A\,X) = \mathrm{vol}(X)$ for any orthogonal matrix A of determinant 1. Moreover, the volume is *homogeneous of degree* n, i.e., $\mathrm{vol}(\lambda\,K) = |\lambda|^n \mathrm{vol}(K)$ and (strictly) *monotone* in the sense that for $K \subseteq L \in \mathcal{K}^n$ we have $\mathrm{vol}(K) \leq \mathrm{vol}(L)$ with equality iff $\dim L < n$ or $\dim L = n$ and $K = L$. Finally, we mention that the volume is also a *continuous* function with respect to the *Hausdorff metric* induced by the *Hausdorff distance* which for $K, L \in \mathcal{K}^n$ is defined as

$$\mathrm{d_H}(K, L) := \min\{\rho \in \mathbb{R}_{\geq 0} : K \subseteq L + \rho\,B_n \text{ and } L \subseteq K + \rho\,B_n\}.$$

For instance, for the convex bodies B_n^p we have (cf. Exercise 1.4)

$$\mathrm{d_H}(B_n^p, B_n^q) = \sqrt{n}\,|n^{-1/p} - n^{-1/q}|.$$

$\mathrm{d_H}(\cdot,\cdot)$ defines a metric on $\mathcal{K}^n$ which makes $\mathcal{K}^n$ a complete space. The next theorem, known as Blaschke's selection theorem, is a very useful tool in guaranteeing the existence of solutions of extremum problems.

1.4 Theorem [Blaschke selection theorem]. *Any bounded sequence of convex bodies in* $\mathcal{K}^n$ *contains a convergent subsequence.*

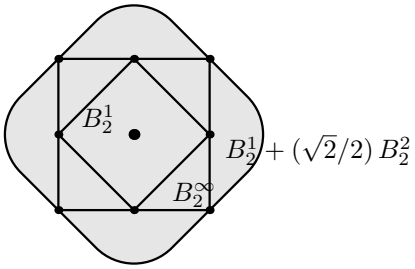

Figure 1.5: $\mathrm{d_H}(B_2^1, B_2^\infty) = \sqrt{2}/2$

Another theorem with an enormous aftermath is about the so-called Löwner-John ellipsoids. In its basic form it says

1.5 Theorem [Löwner-John ellipsoids]. *For each convex body $K \in \mathcal{K}^n$, $\dim K = n$, there is an unique ellipsoid $\mathbf{a} + A\, B_n$, $A \in \mathrm{GL}(n, \mathbb{R})$ of maximal volume contained in K and it holds*

$$\mathbf{a} + A\, B_n \subseteq K \subset \mathbf{a} + n\, A\, B_n.$$

If $K \in \mathcal{K}_o^n$ then we have $\mathbf{a} = \mathbf{0}$ and the factor n can be replaced by $\sqrt{n}$.

The dilatation factors n or $\sqrt{n}$ are also the smallest possible. For instance, in the symmetric case the bound is attained by the cube B_n^∞ or the crosspolytope B_n^1 whose maximal volume ellipsoid is a ball. In the general setting a simplex (see Figure 1.8) is an extremal case.

In some cases the Löwner-John ellipsoids provide an easy way to generalize inequalities valid for ellipsoids to arbitrary o-symmetric convex bodies. As an example we consider the so-called *Mahler volume* $\mathrm{M}(K) := \mathrm{vol}(K)\,\mathrm{vol}(K^\star)$ for $K \in \mathcal{K}_o^n$, $\dim K = n$. In view of Proposition 1.2 i) we know that the Mahler volume is invariant with respect to linear transformations, i.e., $\mathrm{M}(A\,K) = \mathrm{M}(K)$ for any $A \in \mathrm{GL}(n, \mathbb{R})$, and hence, in order to bound $\mathrm{M}(K)$ we may assume that B_n is the volume maximal ellipsoid contained in K. By Theorem 1.5 we have $B_n \subseteq K \subset \sqrt{n} B_n$ and thus $(1/\sqrt{n})\, B_n \subset K^\star \subseteq B_n$ (cf. Exercise 1.1). So by the monotonicity of the volume we get $\mathrm{vol}(K)\,\mathrm{vol}(K^\star) \leq \mathrm{vol}(\sqrt{n}\, B_n)\mathrm{vol}(B_n)$ and $\mathrm{vol}(K)\,\mathrm{vol}(K^\star) \geq \mathrm{vol}(B_n)\,\mathrm{vol}((1/\sqrt{n})\, B_n)$, or

$$\sqrt{n}^{-n}\,\mathrm{M}(B_n) \leq \mathrm{M}(K) \leq \sqrt{n}^{n}\,\mathrm{M}(B_n).$$

Of course, these bounds are not the best possible, but they show that $\mathrm{M}(K)$ is bounded. In order to improve these bounds, we firstly observe that convex bodies mini- and maximizing $\mathrm{M}(K)$ really exist. This, however, is a consequence of Theorem 1.4 and the fact that it suffices to consider convex bodies $K \in \mathcal{K}_o^n$ fulfilling $B_n \subseteq K \subset \sqrt{n} B_n$.

It was shown by Blaschke and Santaló that the maximum of $\mathrm{M}(K)$ is only attained for ellipsoids, and so we have $\mathrm{M}(K) \leq \mathrm{M}(B_n) = \mathrm{vol}(B_n)^2$. The lower bound leads to the so called *Mahler conjecture* claiming that

$$\mathrm{M}(K) \geq \mathrm{M}(B_n^\infty) = \mathrm{vol}(B_n^\infty)\,\mathrm{vol}(B_n^1) = \frac{4^n}{n!} = \left(\frac{4}{2\pi} + o(1)\right)^n \mathrm{M}(B_n).$$

This conjecture is open for $n \geq 3$, and it is also known that there is a whole family of (affinely inequivalent) convex bodies with $\mathrm{M}(K) = \mathrm{M}(B_n^\infty)$. The best known lower bounds are of the type $(\text{constant}^n) \cdot \mathrm{M}(B_n)$, and the first who established such a bound were Bourgain and Milman

1.6 Theorem [Bourgain-Milman]. *There exists a positive absolute constant C such that $\mathrm{vol}(K)\,\mathrm{vol}(K^\star) \geq C^n\,\mathrm{vol}(B_n)^2$ for any convex body $K \in \mathcal{K}_o^n$, $\dim K = n$.*

Most of the (classical) results in Geometry of Numbers are dealing with o-symmetric bodies because of their interpretation as unit balls of norms. On the other hand, most of the underlying geometric problems and structures can be formulated for arbitrary convex bodies and so there are many problems which have a symmetric and non-symmetric side. In particular, since there are already many results for o-symmetric bodies, one is interested in transformations making non-symmetric bodies symmetric such that the geometric sizes in question, e.g., volume, keeps controllable. A standard procedure in this respect is building the *central symmetral* $\mathrm{cs}(K)$ of $K \in \mathcal{K}^n$ defined by

$$\mathrm{cs} : \mathcal{K}^n \to \mathcal{K}_o^n \text{ with } \mathrm{cs}(K) := \frac{1}{2}(K - K). \tag{1.2}$$

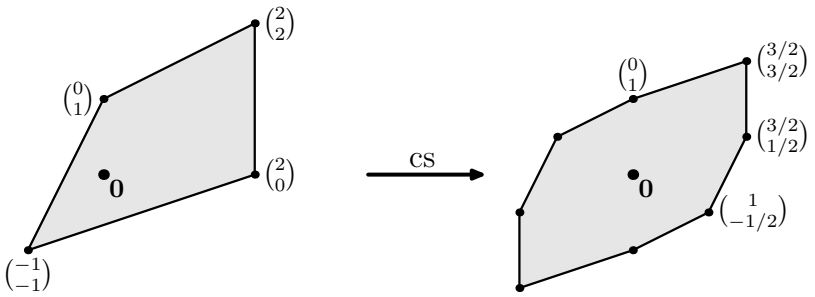

Figure 1.6: Central symmetrization of a convex body

Of course, $K \in \mathcal{K}_o^n$ remains invariant under this map and $\mathrm{cs}(\mathbf{t} + K) = \mathrm{cs}(K)$ for $\mathbf{t} \in \mathbb{R}^n$. The following classical theorem by Rogers and Shephard completely describes the behavior of the volume with respect to this map.

1.7 Theorem [Rogers-Shephard]. *Let $K \in \mathcal{K}^n$. Then*

$$\mathrm{vol}(K) \leq \mathrm{vol}(\mathrm{cs}\,(K)) \leq \frac{1}{2^n}\binom{2\,n}{n}\mathrm{vol}(K).$$

In the lower bound equality is attained iff $K \in \mathcal{K}^n_o$ and in the upper bound iff K is a simplex.

In fact, the lower bound is just an application of the famous Brunn-Minkowski theorem

1.8 Theorem [Brunn-Minkowski]. *Let $K, L \in \mathcal{K}^n$ and $\lambda \in [0,1]$. Then*

$$\mathrm{vol}(\lambda\,K + (1-\lambda)\,L)^{1/n} \geq \lambda \mathrm{vol}(K)^{1/n} + (1-\lambda)\,\mathrm{vol}(L)^{1/n} \tag{1.3}$$

with equality iff K and L are contained in parallel hyperplanes, or K and L are homothetic, i.e., there exists a $\mu \in \mathbb{R}_{\geq 0}$ and $\mathbf{t} \in \mathbb{R}^n$ with $L = \mathbf{t} + \mu\,K$.

In words, the n-th root of the volume is a concave function on $\mathcal{K}^n$. Due to the (weighted) arithmetic-geometric mean inequality (1.3) also implies

$$\mathrm{vol}(\lambda K + (1-\lambda)L) \geq \mathrm{vol}(K)^{\lambda}\,\mathrm{vol}(L)^{1-\lambda}, \tag{1.4}$$

i.e., the volume is a log-concave function. In fact, if (1.4) holds for all $K, L \in \mathcal{K}^n$ then it also implies (1.3).

Next we turn to some more discrete aspects of convex sets. Let $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^n$ and let $\lambda_i \in \mathbb{R}_{\geq 0}$, $1 \leq i \leq m$, with $\sum_{i=1}^m \lambda_i = 1$. Then $\mathbf{x} = \sum_{i=1}^m \lambda_i\,\mathbf{x}_i$ is called a *convex combination* of $\mathbf{x}_1, \ldots, \mathbf{x}_m$. As linear or affine sets are closed with respect to taking finite linear or affine combinations, convex sets are closed with respect to taking finite convex combinations. More precisely (cf. Exercise 1.5), $K \subseteq \mathbb{R}^n$ is convex if and only if

$$K = \left\{ \sum_{i=1}^m \lambda_i\,\mathbf{x}_i : m \in \mathbb{N}, \mathbf{x}_i \in K,\, \lambda_i \geq 0,\, \sum_{i=1}^m \lambda_i = 1 \right\}.$$

In analogy to linear and affine hulls of a subset $X \subseteq \mathbb{R}^n$ we define the *convex hull*, $\mathrm{conv}\,X$, of X as the smallest – with respect to inclusions – convex set containing X, i.e.,

$$\mathrm{conv}\,X := \bigcap_{\substack{C \subseteq \mathbb{R}^n,\, C \text{ convex},\\ X \subseteq C}} C.$$

As in the linear and affine setting, the convex hull of a set X is the set of all finite convex combinations which can be formed by elements of X. In fact, due to a classical result of Carathéodory it suffices to consider convex combinations consisting of at most $\dim X + 1$ elements (cf. Exercise 1.6).

1.9 Theorem [Carathéodory]. *Let $X \subseteq \mathbb{R}^n$. Then*

$$\operatorname{conv} X = \left\{ \sum_{i=1}^{\dim X+1} \lambda_i \mathbf{x}_i : \mathbf{x}_i \in X, \lambda_i \geq 0, \sum_{i=1}^{\dim X+1} \lambda_i = 1 \right\}.$$

By Caratheodory's theorem one can easily argue that the convex hull of a compact set X is again compact, but observe, convex hulls of closed sets do not necessarily need to be closed. We also note that for $X, Y \subseteq \mathbb{R}^n$

$$\operatorname{conv} X + \operatorname{conv} Y = \operatorname{conv}(X + Y).$$

If $\#X < \infty$, i.e., $X \subset \mathbb{R}^n$ is finite, the convex body $\operatorname{conv} X$ is called a *polytope*, and we shall denote the set of all polytopes and all o-symmetric polytopes by $\mathcal{P}^n$ and $\mathcal{P}^n_o$, respectively. In order to emphasize the dimension of a polytope $P \in \mathcal{P}^n$, P will be called an m-polytope if $\dim P = m$. In the case of a polytope $P \in \mathcal{P}^n$, the extreme points, i.e., the 0-dimensional faces are called *vertices*, the 1-dimensional faces are called *edges* and factes are $(\dim P - 1)$-dimensional faces. Each polytope is the convex hul of its

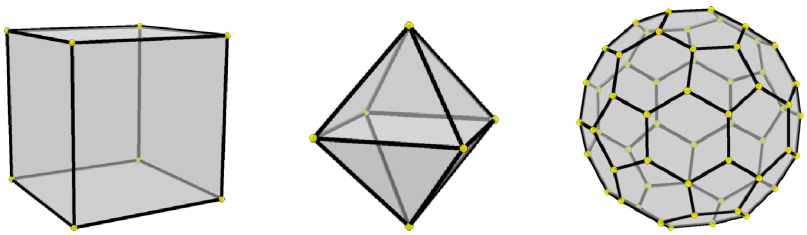

Figure 1.7: Polytopes: The 3-cube B_3^∞, the 3-crosspolytope B_3^1, and the "soccer ball" (truncated icosahedron)

vertices, in fact, each convex body is the convex hull of its extreme points (cf. Exercise 1.9). The vertices of the cube B_n^∞, for instance, are given by the 2^n vectors whose coordinate are plus or minus 1, and so we have

$$B_n^\infty = \operatorname{conv}\{(\epsilon_1, \ldots, \epsilon_n)^\intercal, \epsilon_i \in \{-1, 1\}\}. \tag{1.5}$$

Given a polytope $P = \operatorname{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \in \mathcal{P}^n$, the polar set of P is given by the system of inequalities $P^\star = \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{x}_i, \mathbf{y} \rangle \leq 1,\ 1 \leq i \leq m\}$. In general, such a set, i.e., a set given by finitely many linear (weak) inequalities is called a *polyhedron*. It was firstly shown by Minkowski and Weyl that bounded polyhedra are polytopes and vice versa.

1.10 Theorem [Minkowski-Weyl]. *$C \subset \mathbb{R}^n$ is a convex polytope if and only if it is a bounded polyhedron.*

So we always have two representations of polytopes, either as the convex hull of some points, e.g., the vertices or as the intersection of finitely many linear inequalities, e.g., the inequalities corresponding to facet defining supporting hyperplanes. For instance, the cube has also the presentation

$$B_n^\infty = \{\mathbf{x} \in \mathbb{R}^n : \langle \pm \mathbf{e}_i, \mathbf{x} \rangle \leq 1,\ 1 \leq i \leq n\}, \tag{1.6}$$

where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^\intercal$ with the 1 at the i-th position denotes the i-th unit vector. In comparison with the representation as the convex hull of 2^n points we only need $2\,n$ linear inequalities here for describing the cube. The convex hull of $n+1$ affinely independent points in $\mathbb{R}^n$ is called an n-simplex or just a simplex, e.g., $T_n = \operatorname{conv}\{\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the so-called *standard simplex*. By definition, any other simplex is an affine

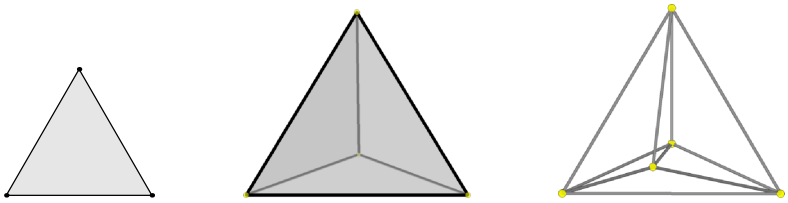

Figure 1.8: 2-simplex (triangle), 3-simplex (tetrahedron), and Schlegel-diagram of a 4-simplex

transformation of T_n. If the distance between any two points of a simplex is equal, the simplex will be called *regular*. Denoting by $f_i(P)$ the number of i-faces of a polytope $P \in \mathcal{P}^n$, we find for an n-simplex that $f_i(T_n) = \binom{n+1}{i+1}$, $i = 0, \ldots, n-1$, and for the cube B_n^∞ and the crosspolytope B_n^1 we have

$$f_{n-1-i}(B_n^1) = f_i(B_n^\infty) = 2^{n-i}\binom{n}{i},\ i = 0, \ldots, n-1.$$

Here the first equality follows from the general fact that the polarity map induces an inclusion reversing bijection between the i-faces of a polytope P and the $(n-1-i)$-faces of $P^\star$, where we assume that $\mathbf{0} \in \operatorname{int} P$, i.e., the origin is an *interior* point of P.

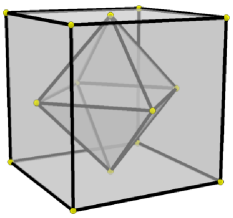

Figure 1.9: "Polarity" between the faces of B_3^∞ and B_3^1

In particular, for those n-polytopes P we have: if $P = \operatorname{conv}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ with vertices $\mathbf{v}_i$, $1 \leq i \leq m$, then $P^\star = \{x \in \mathbb{R}^n : \langle \mathbf{v}_i, x \rangle \leq 1,\ 1 \leq i \leq m\}$

with facets $P^\star \cap H(\mathbf{v}_i, 1)$, $1 \le i \le m$, and vice versa. Thus, in view of (1.6) and (1.5) we get the representations for the crosspolytope

$$\begin{aligned} B_n^1 &= \operatorname{conv}\{\pm\mathbf{e}_i : 1 \le i \le n\} \\ &= \{\mathbf{x} \in \mathbb{R}^n : \langle(\epsilon_1, \ldots, \epsilon_n)^\intercal, \mathbf{x}\rangle \le 1,\ \epsilon_i \in \{-1, 1\}\}. \end{aligned}$$

In principle, it is an easy task to calculate the volume of a polytope. To this end we note that the volume of an n-dimensional *pyramid* $P = \operatorname{conv}\{Q, \mathbf{w}\}$ over an $(n-1)$-dimensional polytope $Q \subset H(\mathbf{a}, b)$, say, and with apex $\mathbf{w} \notin H(\mathbf{a}, b)$ is given by $\operatorname{vol}(P) = (1/n)\operatorname{vol}_{n-1}(Q)\operatorname{dist}(Q, \mathbf{w})$. Here $\operatorname{vol}_{n-1}(\cdot)$ denotes the $(n-1)$-dimensional volume and $\operatorname{dist}(Q, \mathbf{w}) = |\langle \mathbf{a}, \mathbf{v}\rangle - b| / |\mathbf{a}|$ is the distance between $\mathbf{w}$ and $\operatorname{aff} Q = H(\mathbf{a}, b)$.

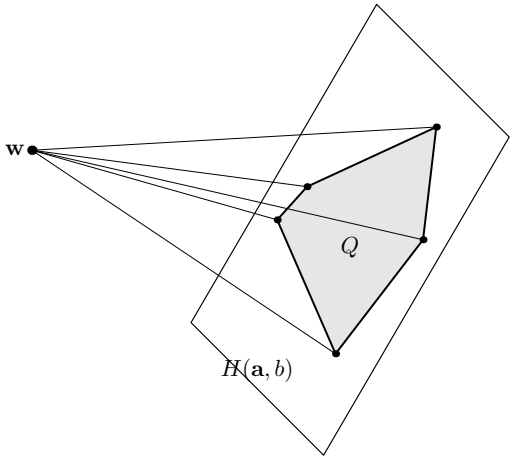

Figure 1.10: Pyramid over Q with apex $\mathbf{w}$

Hence, since the volume is additive we can dissect (without gaps) a polytope P into pyramids over its facets $F_1, \ldots, F_m$, and a common apex $\mathbf{w} \in P$, say, and get $\operatorname{vol}(P) = (1/n)\sum_{i=1}^m \operatorname{vol}_{n-1}(F_i)\operatorname{dist}(F_i, \mathbf{w})$. Since the volume is translation invariant the apex $\mathbf{w}$ can also lie outside of P, we just have to take the signed distances. For instance, let $\operatorname{aff} F_i = H(\mathbf{a}_i, b_i)$ with $|\mathbf{a}_i| = 1$ and $P \subset H_{\le}(\mathbf{a}_i, b_i)$, then with respect to the point $\mathbf{w} = \mathbf{0}$ we get

$$\operatorname{vol}(P) = \frac{1}{n}\sum_{i=1}^m \operatorname{vol}_{n-1}(F_i)\, b_i = \frac{1}{n}\sum_{i=1}^m \operatorname{vol}_{n-1}(F_i)\, h_P(\mathbf{a}_i). \tag{1.7}$$

This recursive formula, however, is apparently inapplicable for practical computations. In fact, it is known that computing the volume of a polytope is a $\#\mathcal{P}$- problem. Another ad hoc approach for computing the volume of a polytope is to dissect the polytope into simplices, since the volume of an n-simplex $S = \operatorname{conv}\{\mathbf{v}_0, \ldots, \mathbf{v}_n\}$ can easily be calculated as

$$\operatorname{vol}(S) = \frac{1}{n!}|\det(\mathbf{v}_1 - \mathbf{v}_0, \ldots, \mathbf{v}_n - \mathbf{v}_0)| = \frac{1}{n!}\left|\det\left(\begin{pmatrix}\mathbf{v}_0\\1\end{pmatrix}, \ldots, \begin{pmatrix}\mathbf{v}_{n+1}\\1\end{pmatrix}\right)\right|.$$

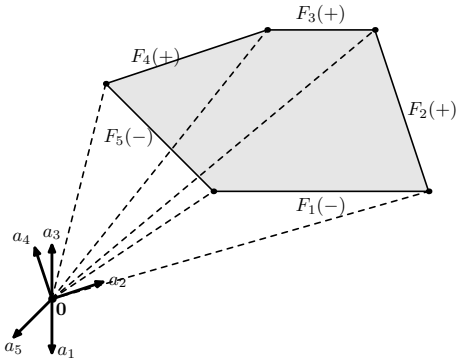

Figure 1.11: "Pyramid" formula for computing volume of polytopes

Here, however, the question is how to dissect P in as few as possible simplices, which is not as easy as it might look. For instance, "even" for the cube B_n^∞ the minimal number of simplices in a dissection is only known in small dimensions and the best bounds differ by an exponential factor.

A dissection of a polytope $P \in \mathcal{P}^n$, $\dim P = n$, into n-simplices S_i, $1 \leq i \leq m$, i.e., $P = \cup_{i=1}^m S_i$ and $\operatorname{int}(S_i) \cap \operatorname{int}(S_j) = \emptyset$, $i \neq j$, is called a *triangulation* if $S_i \cap S_j$ is a face of both, S_i and S_j. The theorem of Carathéodory (Theorem 1.9) implies that a polytope can be covered by the simplices generated by its vertices, the next theorem strengthens this statement in the sense that we can even trinagulate it with its vertices.

1.11 Theorem. *$P \in \mathcal{P}^n$, $\dim P = n$, can be triangulated into simplices without introducing new vertices, i.e., the vertices of each simplex of the triangulation are a subset of the vertices of P.*

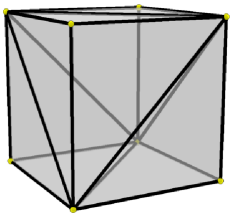

Figure 1.12: A triangulation of the 3-cube into 5 simplices

Exercises

1.1 Show that

i) If $X_1 \subseteq X_2 \subseteq \mathbb{R}^n$ then $X_2^\star \subseteq X_1^\star$.
ii) If $X_i \subseteq \mathbb{R}^n$, $i \in I$, then $\big(\bigcup_{i\in I} X_i\big)^\star = \bigcap_{i\in I} X_i^\star$.
iii) If $X \subseteq \mathbb{R}^n$ then $X \subseteq (X^\star)^\star$.
iv) If $X \subseteq \mathbb{R}^n$ then $X = X^\star$ if and only if $X = B_n$.

1.2 Let $X \subseteq \mathbb{R}^n$ and $\mathbf{t} \in \mathbb{R}^n$ such that $\mathbf{0} \in \operatorname{int} X \cap \operatorname{int}(\mathbf{t} + X)$. Show that

$$(\mathbf{t} + X)^\star = \left\{ \frac{1}{1 + \langle \mathbf{t}, \mathbf{y} \rangle}\, \mathbf{y} : \mathbf{y} \in X^\star \right\}.$$

1.3 i) Let $K \in \mathcal{K}_o^n$, $\dim K = n$, and let $p \in (1, \infty)$. Show that

$$\int_{\mathbb{R}^n} \mathrm{e}^{-|\mathbf{x}|_K^p}\, \mathrm{d}\mathbf{x} = \Gamma(n/p + 1)\, \mathrm{vol}(K).$$

ii) Show for $1 \le p < \infty$

$$\mathrm{vol}(B_n^p) = 2^n \frac{\Gamma(1 + 1/p)^n}{\Gamma(1 + n/p)},$$

and $\mathrm{vol}(B_n^\infty) = 2^n$.

1.4 Show that for $1 \le p, q \le \infty$

$$\mathrm{d_H}(B_n^p, B_n^q) = \sqrt{n}\, |n^{-1/p} - n^{-1/q}|.$$

1.5 Show that $K \subseteq \mathbb{R}^n$ is convex if and only if

$$K = \left\{ \sum_{i=1}^m \lambda_i\, \mathbf{x}_i : m \in \mathbb{N}, \mathbf{x}_i \in K,\ \lambda_i \ge 0,\ \sum_{i=1}^m \lambda_i = 1 \right\}.$$

1.6 Let $X \subseteq \mathbb{R}^n$. Show that

i)

$$\operatorname{conv} X = \left\{ \sum_{i=1}^m \lambda_i\, \mathbf{x}_i :\ m \in \mathbb{N},\ \mathbf{x}_i \in X, \lambda_i \ge 0, \sum_{i=1}^m \lambda_i = 1 \right\}.$$

ii) Show that the integer m in the identity above can be replaced by $\dim X + 1$.

1.7 Let $P \in \mathcal{P}^n$ be an n-dimensional polytope with facets $F_1, \dots, F_m$, and let $\mathbf{a}_i \in \mathbb{R}^n$, $|\mathbf{a}_i| = 1$, $b_i \in \mathbb{R}$, such that $\operatorname{aff} F_i = H(\mathbf{a}_i, b_i)$ and $P \subset H_{\leq}(\mathbf{a}_i, b_i)$, $1 \leq i \leq m$. Show that

$$\sum_{i=1}^{m} \operatorname{vol}_{n-1}(F_i)\, \mathbf{a}_i = \mathbf{0}.$$

(Hint: Formula (1.7) might help.)

1.8 Let $K, L \in \mathcal{K}^n$, $\dim K = \dim L = n$. Show that $\operatorname{int} K + \operatorname{int} L = \operatorname{int}(K + L)$.

1.9 Show that each convex body is the convex hull of its extreme points.

1.10 Show that the number of simplices in any dissection of the cube B_n^∞ into simplices can not be smaller than $2^n\, n!/(n+1)^{(n+1)/2}$.

Rough Notes

1) The number of i-faces of an n-simplex, $\binom{n+1}{i+1}$, is also the minimum number of i-faces which an n-polytope can have. In particular we have $\sum_{i=0}^{n-1} f_i(P) \geq 2^n - 2$ with equality if and only if P is an n-simplex. There is a very appealing conjecture by Gil Kalai claiming that for o-symmetric n-polytopes the lower bound $2^n - 2$ can be replaced by $3^n - 1$ as it is attained by the cube. The cube, however, is as in the case of the Mahler conjecture not the only potential minimizer. In fact, the known bodies for which the Mahler volume is minimal are (up to combinatorial equivalence) precisely those polytopes for which the sum of the i-faces seems to be minimal (see Kalai's paper [8]). The conjecture of Kalai has been confirmed in dimensions ≤ 4 [12], and the only result in the spirit of the conjecture for general dimensions is due to Figiel, Lindenstrauss and Milman [3]

$$\ln f_0(P) \ln f_{n-1}(P) \geq c\, n,$$

where c is an absolute constant. The interesting proof utilizes the Löwner-John ellipsoid of o-symmetric convex bodies.

2) The Mahler conjecture has also a "non-symmetric side". Here we consider for $K \in \mathcal{K}^n$, $\dim K = n$, the product

$$\overline{\mathrm{M}}(K) = \min_{z \in \operatorname{int} K} \operatorname{vol}(K) \operatorname{vol}((K-z)^\star).$$

It is known that, as in the o-symmetric case, it is maximized only for ellipsoids, but the lower bound is open for $n \geq 3$. Simplices are conjectured to be the only minimizer, i.e.,

$$\overline{\mathrm{M}}(K) \geq \overline{\mathrm{M}}(T_n) = \frac{(n+1)^{n+1}}{(n!)^2},$$

with equality if and only if K is an n-simplex. literature

3) The survey of Gardner [4] gives excellent access to the huge amount of classical and modern aspects of the Brunn-Minkowski inequality. A highly recommended article on modern aspects of convex geometry (in general) is (still) [1] by Ball.

4) For more information on the surprising and marvellous world of triangulations&friends we refer to the recent book [10] of De Loera, Rambau and Santos.

2

Basic facts about Lattices

In this chapter we will start to introduce the second and discrete ingredient of the Geometry of Numbers: lattices. Although (or should one say *because*) trivial to define, they have a surprisingly rich geometric structure and even basic theoretical as well as elementary algorithmic questions are still unsolved.

2.1 Definition [Lattice]. *Let* $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^n$ *be linearly independent. The set*

$$\Lambda = \{z_1\mathbf{b}_1 + z_2\mathbf{b}_2 + \cdots + z_k\mathbf{b}_k : z_i \in \mathbb{Z},\, 1 \leq i \leq n\}$$

is called an (n-dimensional) lattice. *The set of generating vectors* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ *or the matrix* $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ *with columns* $\mathbf{b}_i$ *is called the* basis *of* Λ. *An element* $\mathbf{b} \in \Lambda$ *is called the* lattice point *of* Λ. *The set of all n-dimensional lattices in* $\mathbb{R}^n$ *is denoted by* $\mathcal{L}^n$.

In other words, a lattice is of the form $\Lambda = B\,\mathbb{Z}^n$, where B is a regular $n \times n$-matrix, i.e., $B \in \mathrm{GL}(n, \mathbb{R})$. In the case that B is the identity matrix we obtain the lattice $\mathbb{Z}^n$ itself, which is also called *integral lattice* or (*standard lattice*). The unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are a (canonical) basis of $\mathbb{Z}^n$, but, of course, there are many (infinitely many) more bases of $\mathbb{Z}^n$. For instance, the vectors $\mathbf{b}_1 = (1, 1)^\intercal$ and $\mathbf{b}_2 = (-1, 0)^\intercal$ or $\mathbf{b}_1 = (25, 16)^\intercal$ and $\mathbf{b}_2 = (64, 41)^\intercal$ are also bases of $\mathbb{Z}^2$ (since $\mathbf{e}_1$ and $\mathbf{e}_2$ are integral linear combinations of these vectors). In order to characterize all bases of $\mathbb{Z}^n$ we need the following definition.

2.2 Definition [Unimodular matrix]. *An integral matrix* $U \in \mathbb{Z}^{n\times n}$ *is called* unimodular *iff* $|\det U| = 1$. *The group of all unimodular matrices is denoted by* $\mathrm{GL}(n, \mathbb{Z})$.

Oberserve that a matrix is unimodular if and only if the matrix and its inverse are integral matrices, which implies that $\mathrm{GL}(n, \mathbb{Z})$ is precisely the set of all bases of $\mathbb{Z}^n$ as pointed out in the next proposition.

2.3 Proposition. $\mathrm{GL}(n,\mathbb{Z}) = \{U \in \mathbb{R}^{n\times n} : U\,\mathbb{Z}^n = \mathbb{Z}^n\}$.

Proof. $U \in \mathrm{GL}(n,\mathbb{Z})$ if and only if $U, U^{-1} \in \mathbb{Z}^{n\times n}$, which is equivalent to $U\,\mathbb{Z}^n \subseteq \mathbb{Z}^n$ and $U^{-1}\,\mathbb{Z}^n \subseteq \mathbb{Z}^n$. Since the last inclusion is equivalent to $\mathbb{Z}^n \subseteq U\,\mathbb{Z}^n$ we are done. Observe that $\mathbb{Z}^n \subseteq U\,\mathbb{Z}^n$ implies that U is a regular matrix. □

By this proposition we can easily describe all bases of a given lattice.

2.4 Lemma. *Let $\Lambda = B\,\mathbb{Z}^n \in \mathcal{L}^n$. $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ is a basis of Λ if and only if there exists a $U \in \mathrm{GL}(n,\mathbb{Z})$ such that $A = B\,U$.*

Proof. A is a basis of Λ if and only if $A\,\mathbb{Z}^n = \Lambda = B\,\mathbb{Z}^n$. Since B is regular this identity is equivalent to $B^{-1}\,A\,Z^n = \mathbb{Z}^n$ which by Proposition 2.3 is equivalent to $B^{-1}A \in \mathrm{GL}(n,\mathbb{Z})$. □

So two different bases of a lattice Λ are related by an unimodular transformation, and thus $|\det B|$ is independent of the chosen basis B of Λ. Hence it makes sense to associate this size to the lattice.

2.5 Definition [Determinant, fundamental cell]. *Let $\Lambda \in \mathcal{L}^n$ with basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$.*

i) $\det\Lambda = |\det B|$ *is called* determinant *of* Λ.

ii) $P_B = \{\rho_1\mathbf{b}_1 + \cdots + \rho_n\mathbf{b}_n : 0 \le \rho_i < 1,\ 1 \le i \le n\} = B\,[0,1)^n$ *is called* fundamental cell *or* fundamental parallelepiped *of* Λ *(with respect to the basis B).*

Since $\mathrm{vol}(B\,[0,1)^n) = |\det B|\mathrm{vol}([0,1)^n) = |\det B|$ we have for $\Lambda = B\,\mathbb{Z}^n \in \mathcal{L}^n$ that

$$\mathrm{vol}(P_B) = \det\Lambda, \tag{2.5.1}$$

i.e., the determinant is the volume of a fundamental cell. Another apparent, but useful property of a fundamental cell is expressed via the identity

$$(P_B - P_B) \cap \Lambda = \{\mathbf{0}\}. \tag{2.5.2}$$

Observe that $P_B - P_B = B\,(-1,1)^n$ and thus, the only vector in $P_B - P_B$ which has a representation as an integral combination of the columns of B is $\mathbf{0}$.

Before we state the main property of such a fundamental cell, we introduce a notation which can be interpreted as a kind of rounding of a vector $\mathbf{x} \in \mathbb{R}^n$ with respect to a lattice, or more precisely with respect to a basis B of a lattice Λ. To this end let $B \in \mathrm{GL}(n,\mathbb{R})$ and for $\mathbf{x} = \sum_{i=1}^n \rho_i\mathbf{b}_i \in \mathbb{R}^n$, $\rho_i \in \mathbb{R}$, we write

$$\lfloor\mathbf{x}\rceil_B = \sum_{i=1}^n \lfloor\rho_i\rceil\mathbf{b}_i,$$

where $\lfloor \rho \rfloor$ is the largest integer not bigger than ρ. In particular, $\lfloor \mathbf{x} \rfloor_B \in B\mathbb{Z}^n$ and $\mathbf{x} - \lfloor \mathbf{x} \rfloor_B \in P_B$.

2.6 Proposition. *Let $\Lambda = B\mathbb{Z}^n \in \mathcal{L}^n$. Then*

$$\mathbb{R}^n = \dot{\bigcup}_{\mathbf{b}\in\Lambda} (\mathbf{b} + P_B),$$

i.e., $\mathbb{R}^n$ is the pairwise disjoint union of the lattice translates $\mathbf{b} + P_B$.

Proof. Each $\mathbf{x} \in \mathbb{R}^n$ can be decomposed as $\mathbf{x} = (\mathbf{x} - \lfloor \mathbf{x} \rfloor_B) + \lfloor \mathbf{x} \rfloor_B$, where the first summand is in P_B and the latter is a lattice point of Λ. To show that the union is pairwise disjoint, we note that the intersection of two lattice translates $\mathbf{b} + P_B, \overline{\mathbf{b}} + P_B$, $\mathbf{b}, \overline{\mathbf{b}} \in \Lambda$, is non-empty, if and only if $\mathbf{b} - \overline{\mathbf{b}} \in (P_B - P_B) \cap \Lambda$. By observation (2.5.2) this is equivalent to $\mathbf{b} = \overline{\mathbf{b}}$. □

The next theorem gives a characterization of lattices as the discrete subgroups S of $\mathbb{R}^n$ which spans $\mathbb{R}^n$ as a vectorspace. Here "discrete" just means that there exists an $\epsilon > 0$ such that $|s_1 - \mathbf{s}_2| \geq \epsilon$ for any $\mathbf{s}_1, \mathbf{s}_2 \in S$. The main part of the proof of this theorem is covered by the next lemma which will also be used later on.

2.7 Lemma. *Let $S \subset \mathbb{R}^n$ be a discrete subgroup containing n linearly independent elements $\mathbf{s}_1, \ldots, \mathbf{s}_n \in S$. There exist $\mathbf{b}_1, \ldots, \mathbf{b}_n \in S$ such that for $1 \leq k \leq n$*

$$\operatorname{lin} \{\mathbf{s}_1, \ldots, \mathbf{s}_k\} \cap S = \{z_1\mathbf{b}_1 + \cdots + z_k\mathbf{b}_k : z_i \in \mathbb{Z}\}. \tag{2.7.1}$$

Proof. We will construct the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ inductively. For $k = 1$ let $\mathbf{b}_1 \neq 0$ be the (Euclidean) shortest vector in $\operatorname{conv}\{\mathbf{0}, \mathbf{s}_1\} \cap S$. Since S is discrete such a choice is possible, and since S is a subgroup we certainly have $\{z_1\mathbf{b}_1 : z_1 \in \mathbb{Z}\} \subseteq \operatorname{lin}\{\mathbf{s}_1\} \cap S$. For the reverse inclusion, let $\mathbf{s} \in \operatorname{lin}\{\mathbf{s}_1\} \cap S$ and let $\lambda \in \mathbb{R}$ such that $\mathbf{s} = \lambda\,\mathbf{b}_1$. Then $\mathbf{s} - \lfloor\lambda\rfloor\,\mathbf{b}_1 = (\lambda - \lfloor\lambda\rfloor)\mathbf{b}_1$, and by the minimality of $\mathbf{b}_1$ we conclude $\lambda = \lfloor\lambda\rfloor \in \mathbb{Z}$. Hence (2.7.1) holds for $k = 1$.

Let us assume that we have already found $\mathbf{b}_1, \ldots, \mathbf{b}_m$, $m \geq 1$, satisfying (2.7.1) for $1 \leq k \leq m < n$. Next we consider the $(m+1)$-dimensional parallelepiped

$$P_{m+1} = \left\{ \sum_{i=1}^{m} \alpha_i\,\mathbf{b}_i + \alpha_{m+1}\mathbf{s}_{m+1} : 0 \leq \alpha_i \leq 1 \right\}.$$

Let $\mathbf{b}_{m+1} \in P_{m+1} \cap S$ having minimum positive distance to $\operatorname{lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$, i.e.,

$$\mathbf{b}_{m+1} = \sum_{i=1}^{m} \overline{\alpha}_i\,\mathbf{b}_i + \overline{\alpha}_{m+1}\mathbf{s}_{m+1}, \tag{2.7.2}$$

and $\overline{\alpha}_{m+1} > 0$ is minimal among all points in $P_{m+1} \cap S$. Obviously, by our inductive construction we have

$$\operatorname{lin}\{\mathbf{b}_1, \dots, \mathbf{b}_{m+1}\} = \operatorname{lin}\{\mathbf{s}_1, \dots, \mathbf{s}_{m+1}\}, \tag{2.7.3}$$

and together with the discrete subgroup property of S we get $\{z_1\mathbf{b}_1 + \cdots + z_{m+1}\mathbf{b}_{m+1} : z_i \in \mathbb{Z}\} \subseteq \operatorname{lin}\{\mathbf{s}_1, \dots, \mathbf{s}_{m+1}\} \cap S$. Again, in order to establish the reverse inclusion, let $\mathbf{s} \in \operatorname{lin}\{\mathbf{s}_1, \dots, \mathbf{s}_{m+1}\} \cap S$ which, by (2.7.3), can be written as $\mathbf{s} = \sum_{i=1}^{m+1} \beta_i \mathbf{b}_i$ for some scalars $\beta_i \in \mathbb{R}$, and we have to show that $\beta_i \in \mathbb{Z}$. On account of (2.7.2), we can write

$$\begin{aligned}\mathbf{s} - \sum_{i=1}^{m+1} \lfloor \beta_i \rfloor \mathbf{b}_i &= \sum_{i=1}^{m+1} (\beta_i - \lfloor \beta_i \rfloor) \mathbf{b}_i \\ &= \sum_{i=1}^{m} \big((\beta_i - \lfloor \beta_i \rfloor) + \overline{\alpha}_i (\beta_{m+1} - \lfloor \beta_{m+1} \rfloor) \big) \mathbf{b}_i \\ &\qquad + \overline{\alpha}_{m+1} (\beta_{m+1} - \lfloor \beta_{m+1} \rfloor) \mathbf{b}_{m+1}.\end{aligned}$$

For abbreviation we write μ_i for all of these scalars in front of the vectors $\mathbf{b}_i$ in the last sum. Then $0 \le \mu_{m+1} < \overline{\alpha}_{m+1} \le 1$, and by the discrete subgroup property of S and the definition of P_{m+1} we also know

$$\mathbf{s} - \sum_{i=1}^{m+1} \lfloor \beta_i \rfloor \mathbf{b}_i - \sum_{i=1}^{m} \lfloor \mu_i \rfloor \mathbf{b}_i = \sum_{i=1}^{m} (\mu_i - \lfloor \mu_i \rfloor) \mathbf{b}_i + \mu_{m+1} \mathbf{b}_{m+1} \in S \cap P_{m+1}.$$

By the minimality of $\overline{\alpha}_{m+1}$ we must have $\mu_{m+1} = 0$ and so $\beta_{m+1} = \lfloor \beta_{m+1} \rfloor \in \mathbb{Z}$. Thus

$$\mathbf{s} - \beta_{m+1} \mathbf{b}_{m+1} = \sum_{i=1}^{m} \beta_i \mathbf{b}_i \in \operatorname{lin}\{\mathbf{s}_1, \dots, \mathbf{s}_m\} \cap S,$$

and by our inductive hypothesis, (2.7.1) implies the integrality of the other scalars β_i, $1 \le i \le m$. □

2.8 Theorem. *$S \subset \mathbb{R}^n$ is a lattice if and only if S a discrete subgroup of $\mathbb{R}^n$ and it contains n linearly independent points.*

Proof. Obviously, a lattice $\Lambda = B\,\mathbb{Z}^n$ is a subgroup of $\mathbb{R}^n$ containing n linearly independent points. Since $\mathbf{0}$ is an interior point of the convex set $P_B - P_B = B(-1, 1)^n$, there exists an $\epsilon > 0$ such that $\epsilon\, B_n \subset P_B - P_B$. Since $(P_B - P_B) \cap \Lambda = \{\mathbf{0}\}$ (cf. (2.5.2)) and since Λ is a subgroup we conclude $|\mathbf{b}_1 - \mathbf{b}_2| \ge \epsilon$ for all $\mathbf{b}_1 \ne \mathbf{b}_2 \in \Lambda$, i.e., Λ is discrete.

For the reverse implication, we observe that Lemma 2.7, more precisely equation (2.7.1) for $k = n$, implies that there exists n linearly independent $\mathbf{b}_1, \dots, \mathbf{b}_n \in S$ such that

$$S = \{z_1\mathbf{b}_1 + \cdots + z_n\mathbf{b}_n : z_i \in \mathbb{Z}\},$$

i.e., S is a lattice. □

Vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of a lattice Λ fulfilling 2.7.1, i.e., they generate all lattice point in a certain plane (here $\operatorname{lin}\{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$) are of particular interest and we define

2.9 Definition [Primitive set of lattice vectors]. *Let* $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \Lambda \in \mathcal{L}^n$ *be linearly independent. The set* $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ *is called a* primitive set of lattice vectors *if*

$$\operatorname{lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\} \cap \Lambda = (\mathbf{b}_1, \ldots \mathbf{b}_k)\mathbb{Z}^k.$$

So lattice vectors are primitive if each lattice point in their linear hull can be represented as an integral combination of these vectors.

2.10 Lemma. *Let* $\Lambda \in \mathcal{L}^n$.

i) *Each primitive set of vectors can be supplemented to a basis of the lattice, and each subset of a basis is a primitive set.*

ii) *Let* $\mathbf{b}_1, \ldots, \mathbf{b}_n$ *be a basis of* Λ, *let* $\mathbf{c} = \sum_{i=1}^n z_i \mathbf{b}_i \in \Lambda$, $z_i \in \mathbb{Z}$, *and let* $j \in \{1, \ldots, n\}$. *Then it holds:*

$$\{\mathbf{b}_1, \ldots, \mathbf{b}_{j-1}, \mathbf{c}\} \text{ is primitive } \Leftrightarrow \gcd(z_j, \ldots, z_n) = 1.$$

Proof. In order to show that a primitive set of vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ can be supplemented to a basis we first supplement it to a set $\{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{s}_{k+1}, \ldots, \mathbf{s}_n\}$ of linearly independent lattice vectors. Since the first k vectors of this set satisfy (2.7.1) the inductive proof of of Lemma 2.7 finds basis containing $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$. If $b_1, \ldots, b_n$ is a basis of Λ, then any $b \in \Lambda$ has an unique and integral representation as linear combination of $\{b_1, \ldots, b_k\} \cap \Lambda$. Hence any subset of a basis is primitive.

For ii) assume first that $\gcd(z_j, \ldots, z_n) \neq 1$. If the gcd is ∞, i.e., all $z_i = 0$, $i = j \ldots, n$, then $\{\mathbf{b}_1, \ldots, \mathbf{b}_{j-1}, \mathbf{c}\}$ are linearly dependent. So let $\gcd(z_j, \ldots, z_n) = m \geq 2$ and let $\mathbf{a} = \sum_{i=j}^n z_i \, \mathbf{b}_i = c - \sum_{i=1}^{j-1} z_i \, \mathbf{b}_i$. Then

$$\frac{1}{m}\mathbf{a} = \sum_{i=j}^{n} \frac{z_i}{m} \, \mathbf{b}_i = \frac{1}{m} \, \mathbf{c} - \sum_{i=1}^{j-1} \frac{z_i}{m} \, \mathbf{b}_i.$$

The first identity shows that $\frac{1}{m}\mathbf{a} \in \Lambda$ and the second says that $\frac{1}{m}\mathbf{a}$ does not have an integral representation by the the linearly independent vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_{j-1}, \mathbf{c}\}$. Hence these vectors are not primitive. For the reverse direction let $\mathbf{a} \in \operatorname{lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_{j-1}, \mathbf{c}\} \cap \Lambda$, i.e., $\mathbf{a} = \sum_{i=1}^{j-1} \mu_i \, \mathbf{b}_i + \mu_j \, \mathbf{c}$ for

some $\mu_i \in \mathbb{R}$ and $\mathbf{a} = \sum_{i=1}^{n} \overline{z_i}\,\mathbf{b}_i$ with $\overline{z_i} \in \mathbb{Z}$. We have to show $\mu_i \in \mathbb{Z}$, $1 \le i \le j$. Substituting $\mathbf{c} = \sum_{i=1}^{n} z_i \mathbf{b}_i$ yields

$$\sum_{i=1}^{j-1}(\mu_i + \mu_j\, z_i)\mathbf{b}_i + \sum_{i=j}^{n} \mu_j\, z_i\, \mathbf{b}_i = \sum_{i=1}^{n} \overline{z_i}\,\mathbf{b}_i. \tag{2.10.1}$$

Hence $\mu_j\, z_i = \overline{z_i}$, $i = j, \ldots, n$, and since $\gcd(z_j, \ldots, z_n) = 1$ we conclude $\mu_j \in \mathbb{Z}$. Comparing the frist $(j-1)$ scalars in (2.10.1) then also shows $\mu_i \in \mathbb{Z}$, $1 \le i \le j-1$. □

Lemma 2.7 applied to a lattice Λ, in particular, shows that for any set of n linearly independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \Lambda$ we can find a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ such that $\mathbf{a}_k = \sum_{i=1}^{k} z_{i,k}\,\mathbf{b}_i$ for some integral scalars $z_{i,k}$ and $1 \le k \le n$, or in matrix notation $A = B\,Z$, where $Z \in \mathbb{Z}^{n\times n}$ is an upper triangular matrix. Via some "integral Yoga" on the basis vectors $\mathbf{b}_i$ one can achieve the following unique representation of this type which is the theory of matrices and is also known as the *Hermite Normalform*.

2.11 Theorem [Hermite Normalform]. *Let $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ be linearly independent lattice vectors of a lattice $\Lambda \in \mathcal{L}^n$. Then there exists a basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ of Λ and an upper triangular matrix $Z \in \mathbb{Z}^{n\times n}$ such that $A = B\,Z$ and for $1 \le k \le n$*

$$0 \le z_{i,k} < z_{k,k},\ 1 \le i \le k-1. \tag{2.11.1}$$

In words, Z is a non-negative upper triangular integral matrix and for each column of Z the unique maximal element is the diagonal element. Moreover, for given A such a basis B is uniquely determined.

Proof. First we show the existence: As mentioned above, on account of Lemma 2.7 there exists a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of Λ such that

$$\mathbf{a}_k = \sum_{i=1}^{k} \overline{z}_{i,k}\,\mathbf{b}_i, \quad \overline{z}_{i,k} \in \mathbb{Z}, 1 \le k \le n.$$

In the following we modify successively this basis in order to find a basis satisfying (2.11.1). If $\overline{z}_{1,1} < 0$, we replace $\mathbf{b}_1$ by $-\mathbf{b}_1$ and (2.11.1) is fulfilled for $k = 1$. Let us assume that (2.11.1) holds for $1 \le i \le l < n$, say. Obviously, for arbitrary scalars $\widetilde{z}_{l+1,l+1} \in \{-1, 1\}$, $\widetilde{z}_{i,l+1} \in \mathbb{Z}$, $1 \le i \le l$, the vectors

$$\mathbf{b}_1, \ldots, \mathbf{b}_l, \left(\sum_{i=1}^{l+1} \widetilde{z}_{i,l+1}\,\mathbf{b}_i\right), \mathbf{b}_{l+2}, \ldots, \mathbf{b}_n \tag{2.11.2}$$

form a basis of Λ. Substituting $\mathbf{b}_{l+1}$ by $\left(\sum_{i=1}^{l+1} \widetilde{z}_{i,l+1}\, \mathbf{b}_i\right)$ yields for $\mathbf{a}_{l+1}$

$$\begin{aligned}
\mathbf{a}_{l+1} &= \sum_{i=1}^{l+1} \overline{z}_{i,l+1}\, \mathbf{b}_i \\
&= \sum_{i=1}^{l} \overline{z}_{i,l+1}\, \mathbf{b}_i + \overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1} \left(\sum_{i=1}^{l+1} \widetilde{z}_{i,l+1}\mathbf{b}_i - \sum_{i=1}^{l} \widetilde{z}_{i,l+1}\mathbf{b}_i \right) \\
&= \sum_{i=1}^{l} \left(\overline{z}_{i,l+1} - \overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1}\widetilde{z}_{i,l+1}\right)\, \mathbf{b}_i \\
&\qquad + \overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1} \left(\sum_{i=1}^{l+1} \widetilde{z}_{i,l+1}\mathbf{b}_i \right).
\end{aligned}$$

Now we choose the sign $\widetilde{z}_{l+1,l+1}$ such that $\overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1} > 0$ and then we choose $\widetilde{z}_{i,l+1}$ such that

$$\overline{z}_{i,l+1} - \overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1}\widetilde{z}_{i,l+1} < \overline{z}_{l+1,l+1}\widetilde{z}_{l+1,l+1}, \quad 1 \le i \le l.$$

The new basis (2.11.2) satisfies now (2.11.1) for $1 \le k \le l+1$, and, of course, in this iterative way we obtain the desired basis. It remains to show the uniqueness of such a basis.

To this end let B' be another basis of Λ and Z' be another upper triangular integral matrix satisfying (2.11) and such that $A = B'Z'$. Then $B = B'\,U$ for some $U \in \mathrm{GL}(n,\mathbb{Z})$ and thus $Z = U\,Z'$. Hence, the rows of Z and Z' generate the same lattice $\bar{\Lambda} \in \mathcal{L}^n$, say. Suppose there Let k be a row index such that the kth row of Z and Z' differ and let l be the minimal column index such that $z_{k,l} \neq z'_{k,l}$, $l \in \{k, \ldots, n\}$. We may assume $z_{k,l} > z'_{k,l}$. Therefore the difference vector $\mathbf{h}$ of the kth row of Z and the kth row of Z' has its frist nonzero coordinate in $h_l > 0$. Since it also belongs to the lattice $\bar{\Lambda}$ we find that

$$z_{k,j} - z'_{k,l} = h_l = m\, z_{l,l}$$

for a non-negative integer m. In view of (2.11) we conclude $m = 0$ which shows $B = B'$. □

With respect to the lattice $\mathbb{Z}^n$ and an integral matrix $A \in \mathbb{Z}^{n\times n}$ the theorem implies that there exists an unimodular matrix $U \in \mathrm{GL}(n,\mathbb{Z})$ and a uniquely determined upper triangular matrix $Z \in \mathbb{Z}^{n\times n}$ with entries $z_{i,k}$ satisfying (2.11.1) such that $A = U\,Z$. Hence, if we start with $A^\intercal$ then we find that for any integral matrix $A \in \mathbb{Z}^{n\times n}$ there exists a unique lower triangular matrix Z such that the entries of $Z^\intercal$ satisfies (2.11.1) and

$$A = Z\,U.$$

Z is usually called the Hermite Normalform of A. In particular, we see that the columns of Z and A generate the same lattice (cf. Lemma (2.4)), and sometimes, it is preferable to work with a basis in triangular form.

2.12 Definition [Sublattice, Index of a sublattice]. *Let $\Lambda \in \mathcal{L}^n$ and let $\mathbf{a}_1, \dots, \mathbf{a}_n \in \Lambda$ be linearly independent. The lattice*

$$\Lambda_0 = \{z_1\mathbf{a}_1 + \cdots + z_n\mathbf{a}_n : z_i \in \mathbb{Z}\}$$

is called a sublattice *with basis $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$. The number of cosets of the subgroup Λ_0 in Λ is called the* index of Λ_0 in Λ, *and it is denoted by* $|\Lambda : \Lambda_0|$.

Of course, the index is a positive integer and the next lemma states that it can easily be determined, either by counting lattice points or by the ratio of the determinants of the involved lattices.

2.13 Lemma. *Let $\Lambda_0 \subseteq \Lambda \in \mathcal{L}^n$ be a sublattice of Λ. Then*

i) $|\Lambda : \Lambda_0| = \#(P_A \cap \Lambda)$ *for any basis A of Λ_0.*

ii) $|\Lambda : \Lambda_0| = \det \Lambda_0 / \det \Lambda$.

Proof. By the definition of cosets, for i) we have to show that Λ is the disjoint union of the cosets $\mathbf{c} + \Lambda_0$, $\mathbf{c} \in P_A \cap \Lambda$, i.e.,

$$\Lambda = \dot{\bigcup}_{\mathbf{c} \in P_A \cap \Lambda} (\mathbf{c} + \Lambda_0) .$$

Let $\mathbf{b} \in \Lambda$, then $\lfloor \mathbf{b} \rfloor_A \in \Lambda_0 \subseteq \Lambda$ and so $(\mathbf{b} - \lfloor \mathbf{b} \rfloor_A) \in P_A \cap \Lambda$. Thus

$$\mathbf{b} = (\mathbf{b} - \lfloor \mathbf{b} \rfloor_A) + \lfloor \mathbf{b} \rfloor_A \in (P_A \cap \Lambda) + \Lambda_0,$$

and it remains to show that the cosets are different. Let $\mathbf{b}_1, \mathbf{b}_2 \in P_A \cap \Lambda$ such that $(\mathbf{b}_1 + \Lambda_0) \cap (\mathbf{b}_2 + \Lambda_0) \neq \emptyset$. Then $\mathbf{b}_1 - \mathbf{b}_2 \in (P_A - P_A) \cap \Lambda_0 = \{\mathbf{0}\}$ (cf. (2.5.2)) and i) is shown.

For ii) we firstly observe that

$$m\, P_A = \dot{\bigcup}_{0 \le m_i < m} (m_1\mathbf{a}_1 + \cdots + m_n\mathbf{a}_n + P_A) ,$$

where m_i, $m \in \mathbb{N}$. Since for every $\mathbf{b} \in \Lambda$ we have $\#\big((\mathbf{b} + P_A) \cap \Lambda\big) = \#(P_A \cap \Lambda)$ we get with i)

$$\#\big(mP_A \cap \Lambda\big) = m^n \, \#(P_A \cap \Lambda) = m^n \, |\Lambda : \Lambda_0|.$$

Finally, since P_A is measureable we may write (cf. Proposition 1.3)

$$\begin{aligned} \det \Lambda_0 = \mathrm{vol}(P_A) &= \lim_{m\to\infty} \# \left(P_A \cap \frac{1}{m}\Lambda \right) \frac{\det \Lambda}{m^n} = \det \Lambda \lim_{m\to\infty} \frac{\#\,(mP_A \cap \Lambda)}{m^n} \\ &= \det \Lambda \ |\Lambda : \Lambda_0|. \end{aligned}$$

□

A simple and useful application of the last lemma is the fact that the determinant of an integral matrix is equal to the number of lattice points contained in the associated halfopen parallelepiped and, by definiton, it is also equal to the volume of this parallelepiped.

2.14 Corollary. *Let $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in \mathrm{GL}(n, \mathbb{Z})$, and let $P_A = A[0,1)^n$ be the half open parallelepiped generated by the columns of A. Then $|\det A| = \#(P_A \cap \mathbb{Z}^n) = \mathrm{vol}(P_A)$.*

Proof. We just apply Lemma 2.13 i), ii) with $\Lambda = \mathbb{Z}^n$ and Λ_0 being the lattice generated by A. □

In some cases we are interested in subsets $\mathbb{Z}^n$ fullfiling certain (homogeneous) linear congruences, for instance, $\{\mathbf{z} \in \mathbb{Z}^n : z_1 + \cdots + z_n \equiv 0 \bmod 2\}$. The next application of Theorem 2.8 shows that those sets are lattices.

2.15 Corollary. *Let $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{Z}^n$ and let $k_i \in \mathbb{N}_{\geq 1}$, $1 \leq i \leq m$. The set*

$$\Lambda = \{\mathbf{z} \in \mathbb{Z}^n : \langle \mathbf{u}_i, \mathbf{z} \rangle \equiv 0 \bmod k_i,\ 1 \leq i \leq m\}$$

is a sublattice of $\mathbb{Z}^n$ with $\det \Lambda \leq k_1 k_2 \cdot \ldots \cdot k_m$.

Proof. By definition Λ is a discrete subgroup of $\mathbb{R}^n$, actually of $\mathbb{Z}^n$. Since the n linearly independent vectors $(k_1 \cdots k_m)\,\mathbf{e}_i$, $1 \leq i \leq n$, belong to Λ, Theorem 2.8 shows that Λ is a lattice. Since Λ is a sublattice of $\mathbb{Z}^n$ we may consider the different cosets of Λ w.r.t. $\mathbb{Z}^n$. Two points $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}^n$ belong to different cosets if and only if $(\mathbf{z}_1 - \mathbf{z}_2) \notin \Lambda$, i.e., there exists an i such that $\langle \mathbf{u}_i, \mathbf{z}_1 - \mathbf{z}_2 \rangle \not\equiv 0 \bmod k_i$.

In other words, for some $\mathbf{u}_i$ the numbers $\langle \mathbf{u}_i, \mathbf{z}_1 \rangle$ and $\langle \mathbf{u}_i, \mathbf{z}_2 \rangle$ belong to different residue classes mod k_i. For each $\mathbf{u}_i$ there are at most k_i different residue classes and thus the number of different cosets is bounded by the product $k_1 \cdot \ldots \cdot k_m$. Hence we get by Lemma (2.13)

$$\det \Lambda = |\mathbb{Z}^n : \Lambda| \det \mathbb{Z}^n \leq k_1 k_2 \cdot \ldots \cdot k_m.$$

□

Given linearly independent vectors $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ of a lattice $\Lambda \in \mathcal{L}^n$ we can easily decide when these vectors form a basis of Λ. Namely by Lemma 2.13 we know that

$$A \text{ is basis of } \Lambda \Leftrightarrow |\Lambda : A\mathbb{Z}^n| = 1 \Leftrightarrow \Lambda \cap P_A = \{\mathbf{0}\}. \tag{2.15.1}$$

So we just have to check if the parallelepiped P_A contains a non-trivial lattice point of Λ. In the planar case we even have

2.16 Proposition. *Let $\Lambda \in \mathcal{L}^2$ and let $\mathbf{a}_1, \mathbf{a}_2 \in \Lambda$ be linearly independent. Then*

$$\mathbf{a}_1,\, \mathbf{a}_2 \text{ is basis of } \Lambda \Leftrightarrow \operatorname{conv}\{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\} \cap \Lambda = \{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\}.$$

Proof. The main and only observation is that in the plane each parallelogram P can be dissected into two triangles which are reflections of each other with respect to the midpoint of the main diagonal. With $A = (\mathbf{a}_1, \mathbf{a}_2)$ this becomes for the halfopen parallelogram P_A

$$P_A = \operatorname{conv}\{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\} \setminus \{\mathbf{a}_!, \mathbf{a}_2\} \cup ((\mathbf{a}_1 + \mathbf{a}_2) - \operatorname{int}\operatorname{conv}\{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\})\,.$$

Hence $P_A \cap \Lambda = \{\mathbf{0}\}$ if and only if $\operatorname{conv}\{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\} \cap \Lambda = \{\mathbf{0}, \mathbf{a}_1, \mathbf{a}_2\}$ and with (2.15.1) we are done. □

As one can guess from the arguments in the proof of Proposition 2.16 an analogous statement does not exist in dimension ≥ 3. For instance, for $n \geq 3$ and $m \in \mathbb{N}$ let $\mathbf{b}(m) = (1, \ldots, 1, m)^\intercal \in \mathbb{R}^n$ and let

$$T_n(m) = \operatorname{conv}\{\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{b}(m)\}.$$

Then the only integral points of the simplices $T_n(m)$ are the vertices, i.e., $T_n(m) \cap \mathbb{Z}^n = \{\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{b}(m)\}$, but $\mathbf{e}_1, \ldots, \mathbf{e}_{n-1}, \mathbf{b}(m)$ do not build a basis of $\mathbb{Z}^n$ for $m \geq 2$.These simplices are called *Reeve simplices* and will also play a role later on.

Next we want to extend the definition of lattices to k-dimensional structures, more precisely, to k-dimensional linear subspaces.

2.17 Definition [Lattice planes, k-dimensional (sub)lattices]. *Let $\Lambda \in \mathcal{L}^n$.*

i) *A linear subspace $L \subseteq \mathbb{R}^n$ is called a* lattice plane *of Λ if $\dim(L \cap \Lambda) = \dim L$. The set of all k-dimensional lattice planes L of Λ, i.e., $\dim L = k$, is denoted by $\mathcal{L}(k, \Lambda)$.*

ii) *For $L \in \mathcal{L}(k, \Lambda)$, the set $\Lambda_k = L \cap \Lambda$ is called a* (k-dimensional) (sub)lattice *of Λ.*

We note that Theorem 2.8, applied to the Euclidean space $L_k \in \mathcal{L}(k, \Lambda)$, shows that Λ_k is indeed a (k-dimensional) lattice. Thus there exists a basis $\mathbf{b}_1, \ldots, \mathbf{b}_k$ of Λ_k, which can also be seen directly from (2.7.1). As in the full dimensional case, the determinant of such a k-dimensional lattice Λ_k, denoted by $\det \Lambda_k$, is the k-dimensional volume of a fundamental cell $\{\rho_1 \mathbf{b}_1 + \cdots + \rho_k \mathbf{b}_k : \rho_i \in [0, 1)\}$ which can be calculated as

$$\det \Lambda_k = \sqrt{\det\left((\mathbf{b}_1, \ldots, \mathbf{b}_k)^\intercal (\mathbf{b}_1, \ldots, \mathbf{b}_k)\right)}. \tag{2.17.1}$$

2.18 Theorem [Smith Normalform]. *Let $\Lambda_0 \subseteq \Lambda \in \mathcal{L}^n$ be a sublattice of Λ. Then there exists a Basis $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ of Λ_0 and a basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ of Λ such that $\mathbf{a}_i = z_i\, \mathbf{b}_i$ with $z_i \in \mathbb{Z}$, $1 \le i \le n$, i.e., $A = B\, Z$ and Z is a diagonal matrix.*

Proof. We use induction on the dimension n, and the case $n = 1$ is obvious. So let $n > 1$. First we show the assertion under the assumption that

$$\text{there exists a } \mathbf{b}_1 \in \Lambda_0 \text{ which is primitive w.r.t. } \Lambda. \tag{2.18.1}$$

Then we can find a basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of Λ (cf. Lemma 2.10 i)) and let $L = \operatorname{lin}\{\mathbf{b}_2, \dots, \mathbf{b}_n\}$. Then $\Lambda = \mathbb{Z}\,\mathbf{b}_1 + (L \cap \Lambda)$ and since $\Lambda_0 \subseteq \Lambda$ and $\mathbf{b}_1 \in \Lambda_0$ we also have $\Lambda_0 = \mathbb{Z}\,\mathbf{b}_1 + (L \cap \Lambda_0)$. Applying our inductive argument to the $(n-1)$-dimensional sublattice $L \cap \Lambda_0$ of $L \cap \Lambda$ gives the assertion.

Next we will scale our lattice Λ_0 such that (2.18.1) holds. To this end we consider for any $\mathbf{a} \in \Lambda_0$ the unique $\alpha_a \in \mathbb{N}$ such that $\frac{1}{\alpha_a}\mathbf{a} \in \Lambda$ is primitive. Let $\overline{\mathbf{a}} \in \Lambda_0$ having minimal factor $\overline{\alpha} = \alpha_{\overline{a}}$, say, and let $\overline{\mathbf{a}} = \overline{\alpha}\,\overline{\mathbf{b}}$, with $\overline{\mathbf{b}} \in \Lambda$ primitive. In the following we will show

$$\overline{\alpha} | \alpha_a \text{ for every } \mathbf{a} \in \Lambda_0. \tag{2.18.2}$$

Let $\mathbf{a} \in \Lambda_0$ and we may assume that $\mathbf{a}$ is linearly independent of $\overline{\mathbf{a}}$. Hence $\Lambda_0' = (\overline{\mathbf{a}}, \mathbf{a})\mathbb{Z}^2$ is a 2-dimensional sublattice of $\Lambda' = \operatorname{lin}\{\overline{\mathbf{a}}, \mathbf{a}\} \cap \Lambda = (\overline{\mathbf{b}}, \mathbf{b})\mathbb{Z}^2$ for a suitable $\mathbf{b} \in \Lambda$ (cf. Lemma 2.10 i)). Hence we may write

$$\mathbf{a} = \alpha_a\,(z_1\,\mathbf{b}_1 + z_2\,\mathbf{b})$$

with $\gcd(z_1, z_2) = 1$, and it is enough to show that $\overline{\alpha} | (\alpha_a\, z_1)$ and $\overline{\alpha} | (\alpha_a\, z_2)$. To this end let $\gamma \in \mathbb{Z}$ such that $0 \le (\alpha_a\, z_1 - \overline{\alpha}\,\gamma) < \overline{\alpha}$. Then

$$\mathbf{c} = \mathbf{a} - \gamma\,\overline{\mathbf{a}} = (\alpha_a\, z_1 - \overline{\alpha}\,\gamma)\,\overline{\mathbf{b}} + \alpha_a\, z_2\, \mathbf{b} = \alpha_c\,(s_1\,\overline{\mathbf{b}} + s_2\,\mathbf{b})$$

for suitable relatively prime integers s_1, s_2. By the choice of γ and $\overline{\alpha}$ we have $0 \le \alpha_c\, s_1 < \overline{\alpha} \le \alpha_c$, and thus $s_1 = 0$. Hence $\alpha_a\, z_1 - \overline{\alpha}\,\gamma = 0$ and so

$$\overline{\alpha} | (\alpha_a\, z_1). \tag{2.18.3}$$

Moreover we hve $\mathbf{c} = \alpha_a\, z_2\, \mathbf{b}$ and so $\overline{\mathbf{a}} + \mathbf{c} = \overline{\alpha}\,\overline{\mathbf{b}} + \alpha_a\, z_2\, \mathbf{b} \in \Lambda_0 \setminus \{0\}$. Hence $\alpha_{\overline{a}+c} | \overline{\alpha}$ and by the minimality of $\overline{\alpha}$ we get $\alpha_{\overline{a}+c} = \overline{\alpha}$ and thus

$$\overline{\alpha} | (\alpha_a\, z_2). \tag{2.18.4}$$

(2.18.3) and (2.18.4) imply (2.18.2), which shows that $\overline{\Lambda_0} = \frac{1}{\overline{\alpha}}\Lambda_0$ is a sublattice of Λ. Since $\overline{\mathbf{b}} \in \overline{\Lambda_0}$ is primitive w.r.t. Λ, (2.18.1) is satisfied. Hence $\overline{\Lambda_0}, \Lambda$, and thus Λ_0, Λ, have bases with the desired property. □

2.19 Definition [Polar lattice]. *Let $\Lambda \in \mathcal{L}^n$ with basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Let $\mathbf{b}_i^\star \in \mathbb{R}^n$, $1 \le i \le n$ be given by*

$$\langle \mathbf{b}_j, \mathbf{b}_i^\star \rangle = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j, \end{cases}, \quad 1 \le j \le n.$$

The lattice $\Lambda^\star$ with basis $\mathbf{b}_1^\star, \ldots, \mathbf{b}_n^\star$ is called the polar lattice *of Λ.*

2.20 Remark.

i) *If $B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is a basis of Λ then $(\mathbf{b}_1^\star, \ldots, \mathbf{b}_n^\star) = B^{-\intercal}$. In particular, the definition of the polar lattice is independent of the basis B of Λ.*

ii) $\det \Lambda^\star = \det(\Lambda)^{-1}$.

2.21 Proposition. *Let $\Lambda \subset \mathbb{R}^n$ be a lattice. Then*

$$\Lambda^\star = \{\mathbf{y} \in \mathbb{R}^n : \langle \mathbf{b}, \mathbf{y} \rangle \in \mathbb{Z} \text{ for all } \mathbf{b} \in \Lambda\}.$$

Proof. Let B be a basis of Λ and let $\mathbf{y} = B^{-\intercal}\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{x} \in \mathbb{R}^n$. Then

$$\begin{aligned} \mathbf{y} \in \Lambda^\star &\Leftrightarrow \mathbf{x} \in \mathbb{Z}^n \Leftrightarrow \mathbf{x}^\intercal \mathbf{z} \in \mathbb{Z} \text{ for all } z \in \mathbb{Z}^n \\ &\Leftrightarrow (B^{-\intercal}\mathbf{x})^\intercal B\mathbf{z} \in \mathbb{Z} \text{ for all } z \in \mathbb{Z}^n \Leftrightarrow (\mathbf{y})^\intercal B\mathbf{z} \in \mathbb{Z} \text{ for all } z \in \mathbb{Z}^n. \end{aligned}$$

□

2.22 Lemma. *Let $L \in \mathcal{L}(k, \Lambda)$ and let $L^\perp$ be the $(n-k)$-dimensional orthogonal complement of L. Then*

i) *the orthogonal projection of Λ onto $L^\perp$, denoted by $\Lambda|L^\perp$, is an $(n-k)$-dimensional lattice with*

$$\det(\Lambda|L^\perp) \cdot \det(L \cap \Lambda) = \det \Lambda. \tag{2.22.1}$$

ii) *$L^\perp \in \mathcal{L}(n-k, \Lambda^\star)$ and it holds $(L^\perp \cap \Lambda^\star) = (\Lambda|L^\perp)^\star$ (w.r.t. the space $L^\perp$). In particular, we have*

$$\det(L \cap \Lambda) = \det \Lambda \cdot \det(L^\perp \cap \Lambda^\star). \tag{2.22.2}$$

Proof. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of Λ such that $\mathbf{b}_1, \ldots, \mathbf{b}_k$ is a basis of $L \cap \Lambda$ (cf. Lemma (2.10) i)). Furthermore let $\overline{\mathbf{b}}_{k+1}, \ldots, \overline{\mathbf{b}}_n$ be the orthogonal projection of $\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n$ onto $L^\perp$, i.e., $\overline{\mathbf{b}}_i = \mathbf{b}_i|L^\perp$, $i = k+1, \ldots, n$. Then for every $\mathbf{b} = \sum_{i=1}^n z_i \mathbf{b}_i \in \Lambda$ we have

$$\mathbf{b}|L^\perp = \sum_{i=1}^n z_i (\mathbf{b}_i|L^\perp) = \sum_{i=k+1}^n z_i \overline{\mathbf{b}}_i.$$

This shows that $\Lambda|L^\perp$ is an $(n-k)$-dimensional lattice with basis $\overline{\mathbf{b}}_{k+1},\dots,\overline{\mathbf{b}}_n$. Now let $B_k=(\mathbf{b}_1,\dots,\mathbf{b}_k)$ and $\overline{B}_{n-k}=(\overline{\mathbf{b}}_{k+1},\dots,\overline{\mathbf{b}}_n)$. Then we can write

$$\begin{aligned}\det\Lambda &= |\det(\mathbf{b}_1,\dots,\mathbf{b}_n)| = |\det(B_k,\overline{B}_{n-k})| = \sqrt{\det\big((B_k,\overline{B}_{n-k})^\intercal(B_k,\overline{B}_{n-k})\big)}\\ &= \left[\det\begin{pmatrix}(B_k)^\intercal B_k & 0\\ 0 & (\overline{B}_{n-k})^\intercal\overline{B}_{n-k}\end{pmatrix}\right]^{1/2}\\ &= \sqrt{\det((B_k)^\intercal B_k)}\cdot\sqrt{\det((\overline{B}_{n-k})^\intercal\overline{B}_{n-k})} = \det(L\cap\Lambda)\cdot\det(\Lambda|L^\perp).\end{aligned}$$

Next let $\mathbf{b}_1^\star,\dots,\mathbf{b}_n^\star$ be the associated basis of $\Lambda^\star$, i.e., $\langle\mathbf{b}_j,\mathbf{b}_i^\star\rangle=\delta_{i\,j}$. This immediately implies that $\mathbf{b}_{k+1}^\star,\dots,\mathbf{b}_n^\star\in L^\perp\cap\Lambda^\star$ and since

$$\big\langle\mathbf{b}_{k+i}^\star,\overline{\mathbf{b}}_{k+j}\big\rangle=\big\langle\mathbf{b}_{k+i}^\star,\mathbf{b}_{k+j}\big\rangle=\delta_{i\,j}$$

we see that $L^\perp\cap\Lambda^\star$ is an $(n-k)$-dimensional lattice, which is polar to $\Lambda|L^\perp$. Thus we have $\det(L^\perp\cap\Lambda^\star)=\det(\Lambda|L^\perp)^{-1}$ and from (2.22.1) we get (2.22.2). □

3

Minkowski's successive minima

In this chapter we start to study the interaction of convex bodies with lattices, and the beginning we address the problem to bound the volume of a $\mathbf{0}$-symmetric convex body K in terms of functionals related to the lattices points of K.

The first lemma, although it is more or less just an observation on so volume of so-called packing and covering sets, is of particular interest.

3.1 Lemma. *Let $X \subset \mathbb{R}^n$ be a bounded Jordan-measurable set and $\Lambda \in \mathcal{L}^n$.*

i) *If $(\mathbf{b}_1 + X) \cap (\mathbf{b}_2 + X) = \emptyset$, for all $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda$, $\mathbf{b}_1 \neq \mathbf{b}_2$, then* $\mathrm{vol}(X) \leq \det \Lambda$.

ii) *If $\Lambda + X = \mathbb{R}^n$ then* $\mathrm{vol}(\Lambda) \geq \det \Lambda$.

Proof. Let P be the fundamental cell of Λ. Since $\Lambda + P$ is a tiling of $\mathbb{R}^n$ (cf. Proposition 2.6) and the volume is a translation invariant functional we may write

$$\begin{aligned}\mathrm{vol}(X) = \mathrm{vol}\left((\Lambda + P) \cap X\right) &= \sum_{\mathbf{b}\in\Lambda} \mathrm{vol}\left((\mathbf{b} + P) \cap X\right) \\ &= \sum_{\mathbf{b}\in\Lambda} \mathrm{vol}\left(P \cap (X - \mathbf{b})\right).\end{aligned}$$

In the first case i) we have that two different translates $P \cap (X - \mathbf{b}_1)$, $P \cap (X - \mathbf{b}_2)$, $\mathbf{b}_1 \neq \mathbf{b}_2$, do not overlap and thus

$$\sum_{\mathbf{b}\in\Lambda} \mathrm{vol}\left(P \cap (X - \mathbf{b})\right) = \mathrm{vol}\left(\cup_{\mathbf{b}\in\Lambda}(P \cap (X - \mathbf{b}))\right) \leq \mathrm{vol}(P) = \det \Lambda.$$

In the second case ii), where $\Lambda + X = \mathbb{R}^n$ we find

$$\sum_{\mathbf{b}\in\Lambda} \mathrm{vol}\left(P \cap (X - \mathbf{b})\right) \geq \mathrm{vol}\left(\cup_{\mathbf{b}\in\Lambda}(P \cap (X - \mathbf{b}))\right) = \mathrm{vol}(P) = \det \Lambda.$$

□

3.2 Corollary. *Let $X \subset \mathbb{R}^n$ with $\mathrm{vol}(X) > \det\Lambda$ and let $\Lambda \in \mathcal{L}^n$. Then $(X - X) \cap \Lambda \setminus \{\mathbf{0}\} \neq \emptyset$.*

Proof. By Lemma 3.1 i) there exist $\mathbf{b}_1, \mathbf{b}_2 \in \Lambda$, $\mathbf{b}_1 \neq \mathbf{b}_2$, and a $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \in (\mathbf{b}_1 + X) \cap (\mathbf{b}_2 + X)$. Thus $\mathbf{x} - \mathbf{b}_1, \mathbf{x} - \mathbf{b}_2 \in X$ and so $\mathbf{b}_2 - \mathbf{b}_1 = (\mathbf{x} - \mathbf{b}_1) - (\mathbf{x} - \mathbf{b}_2) \in (X - X) \cap \Lambda \setminus \{\mathbf{0}\}$. □

Another way to state the conclusion of the corollary above is to say that there exists a $\mathbf{t} \in \mathbb{R}^n$ such that $\#((\mathbf{t} + X) \cap \Lambda) \geq 2$.

3.3 Theorem [Minkowski]. *Let $K \in \mathcal{K}_o^n$, $\Lambda \in \mathcal{L}^n$, and let $\mathrm{vol}(K) \geq 2^n \det\Lambda$. Then*

$$K \cap \Lambda \setminus \{\mathbf{0}\} \neq \emptyset,$$

i.e., a $\mathbf{0}$-symmetric convex body of volume at least $2^n \det\Lambda$ contains a non-trivial lattice point.

Proof. First we assume $\mathrm{vol}(K) > 2^n \det\Lambda$. Then $\mathrm{vol}(\frac{1}{2}K) > \det\Lambda$ and by Corollary 3.2 there exists $\mathbf{b} \in \Lambda \setminus \{\mathbf{0}\}$ with $\mathbf{b} \in \frac{1}{2}K - \frac{1}{2}K = K$.

Now let $\mathrm{vol}(K) = 2^n \det\Lambda$. Since K compact and Λ discrete there exists a $\lambda > 1$ with $\lambda K \cap \Lambda = K \cap \Lambda$. However $\mathrm{vol}(\lambda K) > 2^n$ and thus the conclusion follows from the first case. □

The cube $[-1, 1]^n$ shows that the statement in Minkowski's theorem is best possible.

3.4 Corollary [Theorem on linear forms]. *Let $l_1(\mathbf{x}), \ldots, l_n(\mathbf{x})$ be n homogeneous linear forms given by*

$$l_i(\mathbf{x}) = a_{i\,1}\mathbf{x}_1 + \cdots + a_{i\,n}\mathbf{x}_n, \quad a_{i\,j} \in \mathbb{R}, \quad 1 \leq i, j \leq n.$$

Let $A \in \mathbb{R}^{n\times n}$ be the matrix with entries $a_{i\,j}$ and let $\det A \neq 0$. For any choice of positive numbers τ_i, $1 \leq i \leq n$, satisfying $\tau_1\tau_2 \cdot \ldots \cdot \tau_n \geq |\det A|$ there exists a $\mathbf{z} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ with

$$|l_i(\mathbf{z})| \leq \tau_i, \quad 1 \leq i \leq n.$$

Proof. Let $P = \{\mathbf{x} \in \mathbb{R}^n : |l_i(\mathbf{x})| \leq \tau_i, \quad 1 \leq i \leq n\}$. In order to calculate the volume of that $\mathbf{0}$-symmetric parallelepiped we observe that $P = A^{-1}\{\mathbf{x} \in \mathbb{R}^n : |x_i| \leq \tau_i, 1 \leq i \leq n\}$. Thus

$$\mathrm{vol}(P) = 2^n \frac{\tau_1 \cdot \ldots \cdot \tau_n}{|\det A|} \geq 2^n.$$

The assertion follows from Theorem 3.3. □

3.5 Corollary [Dirichlet, 1842]. *Let $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and let $0 < \epsilon < 1$. Then there exist $p_1, \dots, p_n, q \in \mathbb{Z}$ with $1 \le q \le \epsilon^{-n}$ such that*

$$\left| \alpha_i - \frac{p_i}{q} \right| < \frac{\epsilon}{q}, \quad 1 \le i \le n.$$

Proof. Let $l_i(\mathbf{x}) = \alpha_i \mathbf{x}_{n+1} - \mathbf{x}_i$, $1 \le i \le n$, and $l_{n+1}(x) = \mathbf{x}_{n+1}$. By Corollary 3.4 there exists for every $\tau > 0$ an integral point $\mathbf{z} = (p_1, \dots, p_n, q)^\intercal \in \mathbb{Z}^{n+1} \setminus \{\mathbf{0}\}$ (depending on τ) satisfying the system of linear forms

$$|l_i(\mathbf{z})| = |\alpha_i\, q - p_i| \le \tau^{-1/n},\ 1 \le i \le n, \quad \text{and} \quad |l_{n+1}(\mathbf{z})| = |q| \le \tau.$$

Note that the product of the right hand sides as well as the absolute value of the determinant of the system is 1. Now we choose a $\tau > \epsilon^{-n}$ such that $\lfloor \tau \rfloor \le \epsilon^{-n}$ and we get

$$|l_i(\mathbf{z})| = |\alpha_i\, q - p_i| < \epsilon,\ 1 \le i \le n, \quad \text{and} \quad |l_{n+1}(\mathbf{z})| = |q| \le \lfloor \tau \rfloor \le \epsilon^{-n}.$$

Finally, we observe that $q \neq 0$, because otherwise $p_i = 0$ for $i = 1, \dots, n$ since $\epsilon < 1$ and we would get the contradiction $\mathbf{z} = \mathbf{0}$. Thus we may assume $q \ge 1$.

□

3.6 Proposition. *Let p be prime. Then there exist $a, b \in \mathbb{N}$ with*

$$a^2 + b^2 + 1 \equiv 0 \bmod p.$$

Proof. For $p = 2$ the statement is certainly true. So let p be odd. For $0 \le a \le \frac{1}{2}(p-1)$ the numbers a^2 belong to pairwise distinct residue classes $\bmod p$, because

$$a^2 \equiv \bar{a}^2 \bmod p \Leftrightarrow (a - \bar{a})(a + \bar{a}) \equiv 0 \bmod p.$$

The same is true if we look at the residue classes of $-b^2 - 1$ for $0 \le b \le \frac{1}{2}(p-1)$. Since there are precisely p different residue classes $\bmod p$ we can find integers $0 \le a, b \le \frac{1}{2}(p-1)$ such that a^2 and $-(b^2+1)$ belong to the same residue class $\bmod p$ which proves the proposition. □

3.7 Corollary [Fermat, Lagrange]. *Every positive number $m \in \mathbb{N}$ can be written as the sum of four integer squares, i.e., there exist $m_1, m_2, m_3, m_4 \in \mathbb{N}$ such that*

$$m = (m_1)^2 + (m_2)^2 + (m_3)^2 + (m_4)^2.$$

Proof. First we observe that it suffices to prove the theorem for integers m which are not divisible by a square other than 1. Hence let $m = p_1 \cdot \ldots \cdot p_k$ with distinct primes p_i, $1 \le i \le k$. For each p_i we choose a_i, b_i, $1 \le i \le k$, according to Proposition 3.6 such that

$$(a_i)^2 + (b_i)^2 + 1 \equiv 0 \bmod p_i, \quad 1 \le i \le k. \tag{3.7.1}$$

With this notation we set

$$\Lambda = \left\{ \mathbf{z} \in \mathbb{Z}^4 : \mathbf{z}_1 \equiv (a_i \mathbf{z}_3 + b_i \mathbf{z}_4) \bmod p_i, \mathbf{z}_2 \equiv (b_i \mathbf{z}_3 - a_i \mathbf{z}_4) \bmod p_i,\ 1 \le i \le k \right\}.$$

By Corollary 2.15 we know that $\Lambda \subset \mathbb{R}^4$ is a lattice with $\det \Lambda \le (p_1)^2 \cdot \ldots \cdot (p_k)^2 = m^2$. Thus we get

$$\operatorname{vol}(\sqrt{2m}\, B_4) = (2m)^2 \kappa_4 = (2m)^2 \frac{\pi^2}{2} > 2^4 m^2 \ge 2^4 \det \Lambda.$$

Hence by Theorem 3.3 there exists a $\mathbf{z} \in \Lambda$ with

$$0 < (\mathbf{z}_1)^2 + (\mathbf{z}_2)^2 + (\mathbf{z}_3)^2 + (\mathbf{z}_4)^2 < 2m.$$

In order to prove the theorem it suffices to show that m is a divisor of the sum $(\mathbf{z}_1)^2 + (\mathbf{z}_2)^2 + (\mathbf{z}_3)^2 + (\mathbf{z}_4)^2$. By the choice of Λ we get for $1 \le i \le k$

$$\begin{aligned}
&(\mathbf{z}_1)^2 + (\mathbf{z}_2)^2 + (\mathbf{z}_3)^2 + (\mathbf{z}_4)^2 \\
&\qquad \equiv \left((a_i \mathbf{z}_3 + b_i \mathbf{z}_4)^2 + (b_i \mathbf{z}_3 - a_i \mathbf{z}_4)^2 + (\mathbf{z}_3)^2 + (\mathbf{z}_4)^2 \right) \bmod p_i \\
&\qquad \equiv \left((\mathbf{z}_3)^2 ((a_i)^2 + (b_i)^2 + 1) + (\mathbf{z}_4)^2 ((a_i)^2 + (b_i)^2 + 1) \right) \bmod p_i \\
&\qquad \equiv 0 \bmod p_i,
\end{aligned}$$

where the last relation is a consequence of (3.7.1). Thus all the distinct p_i are divisors of $(\mathbf{z}_1)^2 + (\mathbf{z}_2)^2 + (\mathbf{z}_3)^2 + (\mathbf{z}_4)^2$ and so is m. □

3.8 Theorem. *Let $k \in \mathbb{N}$ and let $X \subset \mathbb{R}^n$ be a Jordan measurable set with $\operatorname{vol}(X) > k$. Then there exist $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1} \in X$, $\mathbf{x}_i \ne \mathbf{x}_j$, $1 \le i \ne j \le k+1$, such that $\mathbf{x}_i - \mathbf{x}_j \in \mathbb{Z}^n$. (In other words: there exists a $\mathbf{t} \in \mathbb{R}^n$ such that $\mathbf{t} + X$ contains at least $k+1$ lattice points of $\mathbb{Z}^n$).*

Proof. Since we have a Jordan-measurable set we know

$$k < \operatorname{vol}(X) = \lim_{m \to \infty} \# \left(X \cap \frac{1}{m} \mathbb{Z}^n \right) \frac{1}{m^n}.$$

Thus there exists an $m \in \mathbb{N}$ such that $\# \left(X \cap \frac{1}{m} \mathbb{Z}^n \right) > k\, m^n$. Since there are m^n cosets of the sublattice $\mathbb{Z}^n$ with respect to $\frac{1}{m}\mathbb{Z}^n$ there exist at least $(k+1)$ different $\mathbf{b}_1, \ldots, \mathbf{b}_{k+1} \in X \cap \frac{1}{m}\mathbb{Z}^n$ belonging to the same coset and thus $\mathbf{b}_i - \mathbf{b}_j \in \mathbb{Z}^n$.

□

3.9 Corollary. *Let* $K \in \mathcal{K}_o^n$, $\Lambda \in \mathcal{L}^n$ *and let* $\mathrm{vol}(K) \geq k\, 2^n \det \Lambda$. *Then*

$$\#\,(K \cap \Lambda) \geq 2k + 1.$$

Proof. Without loss of generality let $\Lambda = \mathbb{Z}^n$ and $\mathrm{vol}(K) > k\, 2^n$ (cf. the arguments in Corollary **??** and Theorem 3.3). By Theorem 3.8 there exist $(k+1)$ different points $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1} \in \frac{1}{2}K$ with $\mathbf{x}_i - \mathbf{x}_j \in \mathbb{Z}^n$. Let us assume that $\mathbf{x}_1$ is one with maximal Euclidean length among these $(k+1)$ points and let $\mathbf{z}_i = \mathbf{x}_{i+1} - \mathbf{x}_1$, $1 \leq i \leq k$. Then we have $\mathbf{z}_i \neq \mathbf{z}_j$, $i \neq j$, and $\mathbf{z}_i \in K \cap \mathbb{Z}^n \setminus \{\mathbf{0}\}$. Furthermore, by the choice of $\mathbf{x}_1$ all these points satisfy $\langle \mathbf{x}_1, \mathbf{z}_i \rangle < 0$ which implies that the $2k$ points $\pm\mathbf{z}_i$, $1 \leq i \leq k$, are pairwise distinct. Together with the point $\mathbf{0}$ this gives the desired lower bound. □

3.10 Definition [Successive minima]. *Let* $K \in \mathcal{K}_o^n$, $\Lambda \in \mathcal{L}^n$. *For* $1 \leq i \leq n$,

$$\lambda_i(K, \Lambda) = \min\{\lambda > 0 : \dim(\lambda K \cap \Lambda) \geq i\}$$

is called the i-th successive minimum *(of* K *w.r.t.* Λ*).*

3.11 Remark.

i) $\lambda_i(K, \Lambda) \geq \lambda_{i-1}(K, \Lambda)$, $2 \leq i \leq n$.

ii) $\lambda_i(K, \Lambda) = \lambda_i(AK, A\Lambda)$, $A \in \mathrm{GL}(n, \mathbb{R})$, *i.e.,* $A \in \mathbb{R}^{n \times n}$ *with* $\det A \neq 0$.

iii) $\lambda_i(\mu\, K, \Lambda) = \frac{1}{\mu}\lambda_i(K, \Lambda) = \lambda_i(K, \frac{1}{\mu}\Lambda)$, $\mu \in \mathbb{R}_{>}0$.

iv) $\mathrm{int}\, K \cap \Lambda \setminus \{\mathbf{0}\} = \emptyset \Leftrightarrow \lambda_1(K, \Lambda) \geq 1$.

v) $\lambda_1(K, \Lambda) = \min\big\{|\mathbf{b}|_K : \mathbf{b} \in \Lambda \setminus \{\mathbf{0}\}\big\}$.

3.12 Proposition. *Let* $K \in \mathcal{K}_o^n$, $\Lambda \in \mathcal{L}^n$, *and let* $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \Lambda$ *be linearly independent such that* $\mathbf{a}_i \in \lambda_i(K, \Lambda)\, K$, $1 \leq i \leq n$. *Then*

$$\mathrm{int}\,(\lambda_i(K, \Lambda)\, K) \cap \Lambda \subset \mathrm{lin}\,\{\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}\},\ 1 \leq i \leq n,$$

where we set $\mathrm{lin}\,\emptyset = \{\mathbf{0}\}$.

Proof. For short we write $\lambda_i = \lambda_i(K, \Lambda)$, and let $k \leq i$ be the minimal index with $\lambda_k = \lambda_i$. Due to the definition of the successive minima, $\mathrm{int}\,(\lambda_i\, K) = \mathrm{int}\,(\lambda_k\, K)$ contains at most $k-1$ linearly independent lattice points and due to teh choice of k, it contains the lattice points $\mathbf{a}_i$, $1 \leq i \leq k$.

□

3.13 Theorem [Minkowski's first theorem on successive minima]. *Let* $K \in \mathcal{K}_o^n$ *and* $\Lambda \in \mathcal{L}^n$. *Then*

$$\lambda_1(K, \Lambda)^n \mathrm{vol}(K) \leq 2^n \det \Lambda.$$

Proof. By the definition of $\lambda_1(K,\Lambda)$ it is $\operatorname{int}(\lambda_1(K,\Lambda)\,K)\cap\Lambda\setminus\{\mathbf{0}\}=\emptyset$ and so by Theorem 3.3 we get $\operatorname{vol}(\lambda_1(K,\Lambda)\,K)\le 2^n\det\Lambda$. □

Actually, the last statement is equivalent to Theorem 3.3.

3.14 Theorem [Minkowski's second theorem on successive minima]. *Let $K\in\mathcal{K}_o^n$ and $\Lambda\in\mathcal{L}^n$. Then*

$$\frac{2^n}{n!}\det\Lambda\le\lambda_1(K,\Lambda)\,\lambda_2(K,\Lambda)\cdot\ldots\cdot\lambda_n(K,\Lambda)\operatorname{vol}(K)\le 2^n\det\Lambda.$$

Proof. Without loss of generality let $\Lambda=\mathbb{Z}^n$. For convenience we write $\lambda_i=\lambda_i(K,\mathbb{Z}^n)$, and let $\mathbf{z}_1,\ldots,\mathbf{z}_n\in\mathbb{Z}^n$ be n linearly independent lattice points with

$$\mathbf{z}_i\in\lambda_i\,K\cap\mathbb{Z}^n,\ 1\le i\le n. \tag{3.14.1}$$

We start with the proof of the upper bound. Let $K_i=\frac{\lambda_i}{2}K$. After a suitable unimodular transformation we may assume (cf. e.g., Theorem 2.11) that $\mathbf{z}_i\in L_i$, where L_i denotes the i-dimensional coordinate plane $L_i=\{\mathbf{x}\in\mathbb{R}^n : x_{i+1}=\cdots=x_n=0\}$. For an integer $q\in\mathbb{N}$ let $M_q^n=\{\mathbf{z}\in\mathbb{Z}^n : |\mathbf{z}_i|\le q\}=q\,C_n\cap\mathbb{Z}^n$ and for $1\le j\le n-1$ let $M_q^j=M_q^n\cap L_j$. Since K is a bounded set there exists a positive constant γ, only depending on K, such that

$$\operatorname{vol}(M_q^n+K_n)\le(2q+2\,\gamma)^n. \tag{3.14.2}$$

By the definition of λ_1 we have $(\mathbf{z}+\operatorname{int}(K_1))\cap(\overline{\mathbf{z}}+\operatorname{int}(K_1))=\emptyset$ for two different lattice point $\mathbf{z},\overline{\mathbf{z}}$, because otherwise we would get the contradiction $\mathbf{z}-\overline{\mathbf{z}}\in\operatorname{int}(K_1)-\operatorname{int}(K_1)=\operatorname{int}(K_1-K_1)=\operatorname{int}(\lambda_1\,K)$. Thus we have

$$\operatorname{vol}(M_q^n+K_1)=(2q+1)^n\operatorname{vol}(K_1)=(2q+1)^n\left(\frac{\lambda_1}{2}\right)^n\operatorname{vol}(K). \tag{3.14.3}$$

In the following we will show that for $1\le i\le n-1$

$$\operatorname{vol}(M_q^n+K_{i+1})\ge\left(\frac{\lambda_{i+1}}{\lambda_i}\right)^{n-i}\operatorname{vol}(M_q^n+K_i). \tag{3.14.4}$$

To this end we may assume $\lambda_{i+1}>\lambda_i$ and $\mathbf{z},\overline{\mathbf{z}}\in\mathbb{Z}^n$ with $(\mathbf{z}_{i+1},\ldots,\mathbf{z}_n)\ne(\overline{\mathbf{z}_{i+1}},\ldots,\overline{\mathbf{z}}_n)$. Then

$$\mathbf{z}+\operatorname{int}(K_{i+1})\cap\overline{\mathbf{z}}+\operatorname{int}(K_{i+1})=\emptyset.$$

Otherwise, in view of Proposition 3.12 we get the contradiction

$$\mathbf{z}-\overline{\mathbf{z}}\in\operatorname{int}\lambda_{i+1}\,K\subset\operatorname{lin}\{\mathbf{z}_1,\ldots,\mathbf{z}_i\}\subset L_i.$$

Hence, for $\mathbf{z} = (0,\dots,0,z_{i+1},\dots,z_n)^\intercal \neq \overline{\mathbf{z}} = (0,\dots,0,\overline{z}_{i+1},\dots,\overline{z}_n)^\intercal \in \mathbb{Z}^n$ we have $(\mathbf{z} + M_q^i + \operatorname{int}(K_{i+1})) \cap (\overline{\mathbf{z}} + M_q^i + \operatorname{int}(K_{i+1})) = \emptyset$. Thus

$$\begin{aligned}\operatorname{vol}\left(M_q^n + K_{i+1}\right) &= (2q+1)^{n-i} \operatorname{vol}\left(M_q^i + K_{i+1}\right),\\ \operatorname{vol}\left(M_q^n + K_i\right) &= (2q+1)^{n-i} \operatorname{vol}\left(M_q^i + K_i\right),\end{aligned}$$

and in order to verify (3.14.4) it suffices to show

$$\operatorname{vol}\left(M_q^i + K_{i+1}\right) \geq \left(\frac{\lambda_{i+1}}{\lambda_i}\right)^{n-i} \operatorname{vol}(M_q^i + K_i). \tag{3.14.5}$$

Now let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}^n$ be the linear maps given by

$$\begin{aligned}f_1(\mathbf{x}) &= \left(\frac{\lambda_{i+1}}{\lambda_i}\mathbf{x}_1,\dots,\frac{\lambda_{i+1}}{\lambda_i}\mathbf{x}_i,\mathbf{x}_{i+1},\dots,\mathbf{x}_n\right)^\intercal,\\ f_2(\mathbf{x}) &= \left(\mathbf{x}_1,\dots,\mathbf{x}_i,\frac{\lambda_{i+1}}{\lambda_i}\mathbf{x}_{i+1},\dots,\frac{\lambda_{i+1}}{\lambda_i}\mathbf{x}_n\right)^\intercal.\end{aligned}$$

Since $M_q^i + K_{i+1} = f_2(M_q^i + f_1(K_i))$ we get

$$\operatorname{vol}\left(M_q^i + K_{i+1}\right) = \left(\frac{\lambda_{i+1}}{\lambda_i}\right)^{n-i} \operatorname{vol}(M_q^i + f_1(K_i))$$

and for the proof of (3.14.5) we have to show

$$\operatorname{vol}\left(M_q^i + f_1(K_i)\right) \geq \operatorname{vol}\left(M_q^i + K_i\right). \tag{3.14.6}$$

To this end let $L_i^\perp$ be the $(n-i)$-dimensional orthogonal complement of L_i. Then it is easy to see that for every $\mathbf{x} \in L_i^\perp$ there exists a $t(\mathbf{x}) \in L_i$ with $K_i \cap (\mathbf{x} + L_i) \subset (f_1(K_i) \cap (\mathbf{x} + L_i)) + t(\mathbf{x})$ and so

$$\left(M_q^i + K_i\right) \cap (\mathbf{x} + L_i) \subset \left[\left(M_q^i + f_1(K_i)\right) \cap (\mathbf{x} + L_i)\right] + t(\mathbf{x}).$$

Thus we get

$$\begin{aligned}\operatorname{vol}(M_q^i + K_i) &= \int_{\mathbf{x}\in L_i^\perp} \operatorname{vol}_i\left((M_q^i + K_i) \cap (\mathbf{x} + L_i)\right) \mathrm{d}\,\mathbf{x}\\ &\leq \int_{\mathbf{x}\in L_i^\perp} \operatorname{vol}_i\left((M_q^i + f_1(K_i)) \cap (\mathbf{x} + L_i)\right) \mathrm{d}\,\mathbf{x}\\ &= \operatorname{vol}(M_q^i + f_1(K_i)),\end{aligned}$$

where $\operatorname{vol}_i(\cdot)$ denotes the i-dimensional volume. This shows (3.14.6) and so we have verified (3.14.4). Finally, it follows from (3.14.2), (3.14.3) and (3.14.4)

$$\begin{aligned}(2q + 2\,\gamma)^n &\geq \operatorname{vol}\left(M_q^n + K^n\right) \geq \left(\frac{\lambda_n}{\lambda_{n-1}}\right) \operatorname{vol}\left(M_q^n + K_{n-1}\right) \geq \dots\\ &\geq \left(\frac{\lambda_n}{\lambda_{n-1}}\right)\left(\frac{\lambda_{n-1}}{\lambda_{n-2}}\right)^2 \cdot \ldots \cdot \left(\frac{\lambda_2}{\lambda_1}\right)^{n-1} \operatorname{vol}\left(M_q^n + K_1\right)\\ &= \lambda_n \cdot \ldots \cdot \lambda_1 \,\frac{\operatorname{vol}(K)}{2^n}\,(2q+1)^n\end{aligned}$$

and so

$$\lambda_n \cdot \ldots \cdot \lambda_1 \operatorname{vol}(K) \le 2^n \left(\frac{2q + 2\,\gamma}{2q+1}\right)^n .$$

Since this holds for all $q \in \mathbb{N}$ the upper bound is proven.

The lower bound follows just by inclusion. To this end we observe that (3.14.1) implies that the cross-polytope $\operatorname{conv}\{\pm\frac{1}{\lambda_i}\,\mathbf{z}_i : 1 \le i \le n\} \subseteq K$. Thus

$$\begin{aligned}\operatorname{vol}(K) &\ge \operatorname{volconv}\left(\pm\frac{1}{\lambda_i}\,\mathbf{z}_i : 1 \le i \le n\right) = \frac{2^n}{n!}\left|\det\left(\frac{1}{\lambda_1}\,\mathbf{z}_1, \ldots, \frac{1}{\lambda_n}\,\mathbf{z}_n\right)\right| \\ &\ge \frac{2^n}{n!}\frac{1}{\lambda_1 \cdot \ldots \cdot \lambda_n}.\end{aligned}$$

□

Obviously, both bounds of Minkowski's second theorem are best possible. For instance, the lower bound is attained by the regular cross-polytope $C_n^\star$ and the upper bound by any o-symmetric box with edges parallel to the coordinate edges.

3.15 Proposition. *Let $K \in \mathcal{K}_o^n$ and $\Lambda \in \mathcal{L}^n$. Then*

$$\#(K \cap \Lambda) \le \left\lfloor \frac{2}{\lambda_1(K,\Lambda)} + 1 \right\rfloor^n .$$

In particular: Let $K \in \mathcal{K}_o^n$ with $\operatorname{int} K \cap \Lambda = \{\mathbf{0}\}$. Then $\#(K \cap \Lambda) \le 3^n$.

Proof. Without loss of generality let $\Lambda = \mathbb{Z}^n$ and we write λ_1 instead of $\lambda_1(K, \mathbb{Z}^n)$. Let $k = \lfloor 2/\lambda_1 + 1 \rfloor$. Suppose there exist $\mathbf{z}, \bar{\mathbf{z}} \in K \cap \mathbb{Z}^n$ such that $\mathbf{z} \equiv \bar{\mathbf{z}} \bmod k$. Then

$$\mathbb{Z}^n \ni \frac{1}{k}\,(\mathbf{z} - \bar{\mathbf{z}}) = \frac{2}{k}\left(\frac{1}{2}\mathbf{z} - \frac{1}{2}\bar{\mathbf{z}}\right) \in \frac{2}{k}K \subset \operatorname{int}\,(\lambda_1\,K)\,,$$

since $2/k < \lambda_1$. By definition of λ_1 we conclude $\mathbf{z} = \bar{\mathbf{z}}$. This shows that for two different lattice points $\mathbf{z}, \bar{\mathbf{z}} \in K$ belongs to different cosets mod $\frac{1}{k}\mathbb{Z}^n$. Thus the number of lattice points in K can not exceed k^n. □

3.16 Conjecture. *Let $K \in \mathcal{K}_o^n$ and $\Lambda \in \mathcal{L}^n$. Then*

$$\#(K \cap \Lambda) \le \prod_{i=1}^{n} \left\lfloor \frac{2}{\lambda_i(K,\Lambda)} + 1 \right\rfloor .$$

3.17 Remark. *The upper bounds on the lattice points in Proposition 3.15 and in Conjecture 3.16 imply the corresponding bounds on the volume as*

stated in the Theorems 3.13, 3.14. For instance, for $\Lambda = \mathbb{Z}^n$ *and with* $\lambda_i(K) = \lambda_i(K, \mathbb{Z}^n)$ *we (would) get by Conjecture 3.16*

$$\begin{aligned}\mathrm{vol}(K) &= \lim_{m\to\infty} \left(\frac{1}{m}\right)^n \#\left(K \cap \frac{1}{m}\mathbb{Z}^n\right) \\ &\leq \lim_{m\to\infty} \left(\frac{1}{m}\right)^n \prod_{i=1}^{n} \left(\frac{2}{\lambda_i(m\,K)} + 1\right) \\ &= \lim_{m\to\infty} \prod_{i=1}^{n} \left(\frac{2}{\lambda_i(K)} + \frac{1}{m}\right) \\ &= \prod_{i=1}^{n} \frac{1}{\lambda_i(K)}.\end{aligned}$$

The second theorem of Minkowski can easily be extended to arbitrary (not necessarily o-symmetric) convex body, if the successive minima are measured with respect to central symmetral $\mathrm{cs}\,(K) = \frac{1}{2}(K - K)$ of K (cf. (1.2)).

3.18 Theorem. *Let* $K \in \mathcal{K}^n$, $\Lambda \in \mathcal{L}^n$. *Then*

$$\frac{2^n}{n!} \det\Lambda \leq \lambda_1(\mathrm{cs}\,(K), \Lambda)\, \lambda_2(\mathrm{cs}\,(K), \Lambda) \cdot \ldots \cdot \lambda_n(\mathrm{cs}\,(K), \Lambda)\, \mathrm{vol}(K) \leq 2^n \det\Lambda.$$

Proof. Without loss of generality $\Lambda = \mathbb{Z}^n$ and for short, we write $\lambda_i = \lambda_i(\mathrm{cs}\,(K), \mathbb{Z}^n)$, $1 \leq i \leq n$. We begin with the upper bound. Due to Theorem 3.14 we have

$$\lambda_1 \cdot \ldots \cdot \lambda_n \mathrm{vol}(\mathrm{cs}\,(K)) \leq 2^n.$$

By the Brunn-Minkowski inequality we also know (cf. e.g. (1.3))

$$\mathrm{vol}(\mathrm{cs}\,(K)) \geq \mathrm{vol}(K),$$

which shows the upper bound.

For the lower bound let $\mathbf{z}_i \in \lambda_i \mathrm{cs}\,(K)$, $1 \leq i \leq n$, be linearly independent and let $\mathbf{u}_i, \mathbf{w}_i \in \frac{1}{2}K$ such hat $\mathbf{v}_i = \frac{1}{\lambda_i}\mathbf{z}_i = \mathbf{u}_i - \mathbf{w}_i$, $1 \leq i \leq n$. Let $C = \mathrm{conv}\,\{\mathbf{u}_i, \mathbf{w}_i : 1 \leq i \leq n\} \subseteq \frac{1}{2}K$. In the following we show

$$\mathrm{vol}(C) \geq \frac{1}{n!} |\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)|$$

which implies the assertion. For the proof we use induction on n. For $n = 1$ nothing is to show and so let $n > 1$. Let $L = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{v}_1, \mathbf{x}\rangle = 0\}$ and for $\mathbf{x} \in \mathbb{R}^n$ we denote by $\bar{\mathbf{x}}$ its orthogonal projection onto L. Application of the Steiner symmetrization to C with respect to the hyperplane H gives a convex body C_S with $\mathrm{vol}(C_S) = \mathrm{vol}(C)$, and by definition of this symmetrization

C_S contains the points $\bar{\mathbf{u}}_1 + \frac{1}{2}\mathbf{v}_1, \bar{\mathbf{u}}_1 - \frac{1}{2}\mathbf{v}_1, \bar{\mathbf{u}}_2, \bar{\mathbf{w}}_2, \ldots, \bar{\mathbf{u}}_n, \bar{\mathbf{w}}_n$. Hence we have

$$\mathrm{vol}(C) = \mathrm{vol}(C_S) \geq \frac{|\mathbf{v}_1|}{n} \cdot \mathrm{vol}_{n-1}(\mathrm{conv}\,\{\bar{\mathbf{u}}_2, \bar{\mathbf{w}}_2, \ldots, \bar{\mathbf{u}}_n, \bar{\mathbf{w}}_n\}).$$

Hence, via our inductive approach we get

$$\begin{aligned}\mathrm{vol}(C) = \mathrm{vol}(C_S) &\geq \frac{|\mathbf{v}_1|}{n} \cdot \frac{1}{(n-1)!}\widetilde{\det}|(\bar{\mathbf{u}}_2 - \bar{\mathbf{w}}_2, \cdots, \bar{\mathbf{u}}_n - \bar{\mathbf{w}}_n)| \\ &= \frac{|\mathbf{v}_1|}{n!}\widetilde{\det}|(\bar{\mathbf{v}}_2, \cdots, \bar{\mathbf{v}}_n)| = \frac{1}{n!}|\det(\mathbf{v}_1, \ldots, \mathbf{v}_n)|.\end{aligned}$$

Here $\widetilde{\det}()$ denotes the determinant in the $(n-1)$-dimensional setting, i.e., the $(n-1)$-dimensional volume of the parallelepiped generated by the $n-1$ vectors. □

4

Reduced bases

4.1 Definition [Gram-Schmidt Orthogonalization]. *For a basis* $\mathbf{b}_1, \ldots, \mathbf{b}_n$ *of* $\mathbb{R}^n$*, the* Gram-Schmidt orthogonalized basis (GSO-basis) $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$ *is recursively given by*

$$\overline{\mathbf{b}}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{i,j} \overline{\mathbf{b}}_j \quad \textit{with} \quad \mu_{i,j} = \frac{\langle \mathbf{b}_i, \overline{\mathbf{b}}_j \rangle}{\langle \overline{\mathbf{b}}_j, \overline{\mathbf{b}}_j \rangle}, \quad 1 \leq i \leq n,\ 1 \leq j < i.$$

4.2 Remark.

i) $\overline{\mathbf{b}}_i = \mathbf{b}_i \big| \lin \{\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}\}^\perp$*, i.e.,* $\overline{\mathbf{b}}_i$ *is the orthogonal projection of* $\mathbf{b}_i$ *onto the orthogonal complement of* $\lin \{\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}\}$.

ii) $\langle \overline{\mathbf{b}}_i, \overline{\mathbf{b}}_j \rangle = 0$ *for* $i \neq j$.

iii) $\lin \{\mathbf{b}_1, \ldots, \mathbf{b}_i\} = \lin \{\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_i\}$, $1 \leq i \leq n$.

iv) $\big|\det(\mathbf{b}_1, \ldots, \mathbf{b}_n)\big| = \big|\det(\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n)\big| = \prod_{i=1}^n \big|\overline{\mathbf{b}}_i\big|$.

v) *Let* $k \geq i \geq 1$*. Then*

$$\mathbf{b}_k \big| \lin \{\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}\}^\perp = \overline{\mathbf{b}}_k + \sum_{j=i}^{k-1} \mu_{k,j} \overline{\mathbf{b}}_j = \mathbf{b}_k - \sum_{j=1}^{i-1} \mu_{k,j} \overline{\mathbf{b}}_j.$$

4.3 Proposition. *Let* $\mathbf{b}_1, \ldots, \mathbf{b}_n$ *be a basis of a lattice* $\Lambda \in \mathcal{L}^n$ *with* GSO-*basis* $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$*. Let* $L_i = \lin \{\mathbf{b}_1, \ldots, \mathbf{b}_i\}$ *and* $\Lambda_{n-i} = \Lambda | L_i^\perp$, $0 \leq i \leq n-1$*. Then* $\mathbf{b}_k | L_i^\perp$, $k = i+1, \ldots, n$*, is a basis of the lattice* Λ_{n-i} *and* $\det \Lambda_{n-i} = \big|\overline{\mathbf{b}}_{i+1}\big| \cdot \ldots \cdot \big|\overline{\mathbf{b}}_n\big|$.

Proof. For $\mathbf{a} = \sum_{j=1}^{n} z_j \, \mathbf{b}_j \in \Lambda$ we have $\mathbf{a}|L_i^{\perp} = \sum_{j=i+1}^{n} z_j \, (\mathbf{b}_j|L_i^{\perp})$ which shows that Λ_{n-i} is a lattice with basis $\mathbf{b}_j|L_i^{\perp}$, $j = i+1, \ldots, n$. Rewriting Remark 4.2 v) in matrix notation we have

$$\left(\mathbf{b}_{i+1}|L_i^{\perp}, \ldots, \mathbf{b}_n|L_i^{\perp}\right) = \left(\overline{\mathbf{b}}_{i+1}, \ldots, \overline{\mathbf{b}}_n\right) T_{n-i},$$

where T_{n-i} is an upper triangular matrix with 1s on the diagonal. Hence

$$\begin{aligned} \det \Lambda_{n-i} &= \det \left(T_{n-i}^{\mathsf{T}} \left(\overline{\mathbf{b}}_{i+1}, \ldots, \overline{\mathbf{b}}_n\right)^{\mathsf{T}} \left(\overline{\mathbf{b}}_{i+1}, \ldots, \overline{\mathbf{b}}_n\right) T_{n-i}\right)^{\frac{1}{2}} \\ &= \left|\overline{\mathbf{b}}_{i+1}\right| \cdot \ldots \cdot \left|\overline{\mathbf{b}}_n\right|. \end{aligned}$$

□

4.4 Definition [Hermite-Korkine-Zolotarev reduced basis, 1872]. *A basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of a lattice $\Lambda \in \mathcal{L}^n$ is called* Hermite-Korkine-Zolotarev (reduced) basis (HKZ-(reduced) basis) *if*

i) $|\mathbf{b}_1| = \lambda_1(B_n, \Lambda)$.

ii) $\mathbf{b}_2\big|\lin\{\mathbf{b}_1\}^{\perp}, \ldots, \mathbf{b}_n\big|\lin\{\mathbf{b}_1\}^{\perp}$ *is a* HKZ*-reduced basis of the lattice* $\Lambda\big|\lin\{\mathbf{b}_1\}^{\perp}$.

iii) *The coefficients $\mu_{i,1}$ of the* GSO*-basis satisfy $|\mu_{i,1}| \leq 1/2$, for $2 \leq i \leq n$.*

4.5 Theorem.

i) *Every lattice $\Lambda \in \mathcal{L}^n$ has a* HKZ*-reduced basis.*

ii) *Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis of $\Lambda \in \mathcal{L}^n$ and let $L_i = \lin\{\mathbf{b}_1, \ldots, \mathbf{b}_i\}$ for $i = 0, \ldots, n$. $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is* HKZ*-reduced if and only if $\overline{\mathbf{b}}_k = \mathbf{b}_k|L_{k-1}^{\perp}$ is the shortest non-zero lattice vector of the lattice $\Lambda|L_{k-1}^{\perp}$, $1 \leq k \leq n$, and $|\mu_{i,j}| \leq 1/2$, for $1 \leq j < i \leq n$.*

Proof. For i) we use induction with respect to the dimension n. For $n = 1$ nothing is to prove and so let $n > 1$. Let $\mathbf{b}_1$ be the shortest vector of the lattice Λ and let $\mathbf{v}_2, \ldots, \mathbf{v}_n$ be a HKZ-reduced basis of the lattice $\Lambda_{n-1} = \Lambda\big|\lin\{\mathbf{b}_1\}^{\perp}$. For each $\mathbf{v}_i$ let $\mu_{i,1} \in \mathbb{R}$ be of minimal absolute value such that $\mathbf{v}_i + \mu_{i,1}\,\mathbf{b}_1 \in \Lambda$. Then $|\mu_{i,1}| \leq 1/2$ and with $\mathbf{b}_i = \mathbf{v}_i + \mu_{i,1}\,\mathbf{b}_1$, $2 \leq i \leq n$, we may write

$$\mu_{i,1} = \frac{\langle (\mathbf{v}_i + \mu_{i,1}\,\mathbf{b}_1), \mathbf{b}_1 \rangle}{|\mathbf{b}_1|^2} = \frac{\langle \mathbf{b}_i, \mathbf{b}_1 \rangle}{|\mathbf{b}_1|^2}, \quad i = 2, \ldots, n.$$

Hence, part iii) of the definition of HKZ-reduced basis is satisfied by the lattice vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$. It remains to be shown that these vectors form a basis, which follows from the identity (cf. Lemma 2.22 i))

$$\left|\det(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)\right| = \left|\det(\mathbf{b}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)\right| = |\mathbf{b}_1| \, \det \Lambda_{n-1} = \det \Lambda.$$

The second statement is certainly true for $n = 1$ and so let $n > 1$. By definition we have

$$\begin{aligned} \mathbf{b}_1, \ldots, \mathbf{b}_n \text{ HKZ-reduced} \Leftrightarrow \text{i)} \quad & \mathbf{b}_1 \text{ is a shortest lattice vector of } \Lambda \\ \text{ii)} \quad & \mathbf{v}_2 = \mathbf{b}_2 - \mu_{2,1}\mathbf{b}_1, \ldots, \mathbf{v}_n = \mathbf{b}_n - \mu_{n,1}\mathbf{b}_1 \\ & \text{is a HKZ-basis of } \Lambda_{n-1} = \Lambda \big| \operatorname{lin}\{\mathbf{b}_1\}^{\perp}, \\ \text{iii)} \quad & |\mu_{i,1}| \le \frac{1}{2}, \quad i = 2, \ldots, n. \end{aligned} \tag{4.5.1}$$

Applying our induction hypothesis to $\mathbf{v}_2, \ldots, \mathbf{v}_n$ yields

$$\begin{aligned} \mathbf{v}_2, \ldots, \mathbf{v}_n \text{ HKZ-reduced} \Leftrightarrow \text{a)} \quad & \overline{\mathbf{v}}_k = \mathbf{v}_k | \tilde{L}_{k-2}^{\perp} \text{ is a shortest lattice vector} \\ & \text{of } \Lambda_{n-1} | \tilde{L}_{k-2}^{\perp}, \quad k = 2, \ldots, n \\ \text{b)} \quad & \left| \frac{\langle \mathbf{v}_i, \overline{\mathbf{v}}_j \rangle}{|\overline{\mathbf{v}}_j|^2} \right| \le \frac{1}{2}, \quad 2 \le i \le n,\ 2 \le j < i. \end{aligned} \tag{4.5.2}$$

Here $\tilde{L}_{k-2} = \operatorname{lin}\{\mathbf{v}_2, \ldots, \mathbf{v}_{k-1}\}$, $k = 2, \ldots, n$, and as before, $\overline{\mathbf{v}}_k$ denote the vectors of the Gram-Schmidt orthogonalized basis of $\mathbf{v}_2, \ldots, \mathbf{v}_n$. Next we observe that $\tilde{L}_{k-2} = L_{k-1} \big| \operatorname{lin}\{\mathbf{b}_1\}^{\perp}$ and thus $\Lambda | L_{k-1}^{\perp} = \Lambda_{n-1} | \tilde{L}_{k-2}$. Moreover we have

$$\begin{aligned} \overline{\mathbf{b}}_k &= \mathbf{b}_k | L_{k-1}^{\perp} = \mathbf{v}_k | \tilde{L}_{k-2}^{\perp} = \overline{\mathbf{v}}_k, \quad k = 2, \ldots, n, \\ \mu_{i,j} &= \frac{\langle \mathbf{b}_i, \overline{\mathbf{b}}_j \rangle}{|\overline{\mathbf{b}}_j|^2} = \frac{\langle \mathbf{b}_i, \overline{\mathbf{v}}_j \rangle}{|\overline{\mathbf{v}}_j|^2} = \frac{\langle \mathbf{v}_i, \overline{\mathbf{v}}_j \rangle}{|\overline{\mathbf{v}}_j|^2}, \quad 2 \le i \le n,\ 2 \le j < i. \end{aligned}$$

Hence, replacing ii) of (4.5.1) by (4.5.2) rewritten in terms of the vectors $\mathbf{b}_i$ gives the desired equivalence. □

4.6 Theorem. *Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a* HKZ*-basis of $\Lambda \in \mathcal{L}^n$.*

$$\frac{2}{\sqrt{i+3}} \lambda_i(B_n, \Lambda) \le |\mathbf{b}_i| \le \frac{\sqrt{i+3}}{2} \lambda_i(B_n, \Lambda), \quad 1 \le i \le n.$$

Proof. For short we write $\lambda_i = \lambda_i(B_n, \Lambda)$, $1 \le i \le n$. First we note that the associated Gram-Schmidt orthogonalized basis $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$ satisfies

$$|\mathbf{b}_i|^2 = \left| \overline{\mathbf{b}}_i + \sum_{j=1}^{i-1} \mu_{i,j} \overline{\mathbf{b}}_j \right|^2 \le |\overline{\mathbf{b}}_i|^2 + \frac{1}{4} \sum_{j=1}^{i-1} |\overline{\mathbf{b}}_j|^2 . \tag{4.6.1}$$

For the upper bound let $\mathbf{a}_1, \ldots, \mathbf{a}_i \in \Lambda$ be linearly independent with $|\mathbf{a}_j| \le \lambda_i$, $1 \le j \le i$, and let $\mathbf{a}_k \in \{\mathbf{a}_1, \ldots, \mathbf{a}_i\}$ such that $\mathbf{a}_k \big| \operatorname{lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}\}^{\perp} \neq \mathbf{0}$.

By Theorem 4.5 ii) we conclude $\left|\overline{\mathbf{b}}_i\right| \le \left|\mathbf{a}_k\right| \operatorname{lin}\{\mathbf{b}_1,\dots,\mathbf{b}_{i-1}\}^\perp\big| \le |\mathbf{a}_k|$ and thus

$$\left|\overline{\mathbf{b}}_j\right| \le \lambda_j, \quad 1 \le j \le n.$$

Applying these bounds in the right hand side of (4.6.1) results in

$$|\mathbf{b}_i|^2 \le \lambda_i^2 + \frac{1}{4}\sum_{j=1}^{i-1}\lambda_j^2 \le \left(1+\frac{i-1}{4}\right)\lambda_i^2 = \left(\frac{i+3}{4}\right)\lambda_i^2.$$

For the lower bound we observe that

$$\left|\overline{\mathbf{b}}_i\right| \le \Big|\mathbf{b}_k\big| \operatorname{lin}\{\mathbf{b}_1,\dots,\mathbf{b}_{i-1}\}^\perp\Big| \le |\mathbf{b}_k|, \quad 1 \le i \le k \le n.$$

Hence (4.6.1) gives

$$|\mathbf{b}_i|^2 \le \left|\overline{\mathbf{b}}_i\right|^2 + \frac{1}{4}\sum_{j=1}^{i-1}\left|\overline{\mathbf{b}}_j\right|^2 \le \frac{i+3}{4}\,|\mathbf{b}_k|^2, \quad 1 \le i \le k \le n.$$

Thus for a fixed index k, say, we conclude

$$\max\{|\mathbf{b}_i|^2 : 1 \le i \le k\} \le \frac{k+3}{4}\,|\mathbf{b}_k|^2.$$

Together with $\lambda_k \le \max\big\{|\mathbf{b}_i| : 1 \le i \le k\big\}$ we get the lower bound. □

4.7 Corollary. *Let $\Lambda \in \mathcal{L}^n$ be a lattice. There exists a basis $\mathbf{b}_1,\dots,\mathbf{b}_n$ of Λ such that*

$$\left(\frac{|\mathbf{b}_1|\cdot\ldots\cdot|\mathbf{b}_n|}{\det\Lambda}\right)^{1/n} \le c\,n,$$

where c is a universal constant.

Proof. Let $\mathbf{b}_1,\dots,\mathbf{b}_n$ be a HKZ-reduced basis of Λ. From Theorem 4.6 and Theorem 3.14 we get

$$\begin{aligned}\left(\frac{|\mathbf{b}_1|\cdot\ldots\cdot|\mathbf{b}_n|}{\det\Lambda}\right)^{1/n} &\le \left(\prod_{i=1}^{n}\frac{\sqrt{i+3}}{2}\right)^{1/n}\left(\frac{\prod_{i=1}^n\lambda_i(B_n,\Lambda)}{\det\Lambda}\right)^{1/n}\\ &< \sqrt{n}\left(\frac{2^n}{\kappa_n}\right)^{1/n} \le c\,n,\end{aligned}$$

for some universal constant c. □

4.8 Proposition. *Let $B = (\mathbf{b}_1,\dots,\mathbf{b}_n)$ be a HKZ-basis of $\Lambda \in \mathcal{L}^n$, and let $q_B(\mathbf{x}) = \mathbf{x}^\intercal B^\intercal B\,\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$. Then*

$$q_B(\mathbf{x}) = \sum_{i=1}^{n}\left(\left|\overline{\mathbf{b}}_i\right|^2\left(x_i + \sum_{j=i+1}^{n}\mu_{j,i}\,x_j\right)^2\right),$$

with

$$\text{i) } |\mu_{j,i}| \le \frac{1}{2}, \text{ and for } 1 \le i < j \le n$$

$$\text{ii) } \left|\overline{\mathbf{b}}_k\right|^2 = \min_{(z_k,\ldots,z_n)^\intercal \in \mathbb{Z}^{n-k+1}\setminus\{\mathbf{0}\}} \sum_{i=k}^{n} \left(\left|\overline{\mathbf{b}}_i\right|^2 \left(z_i + \sum_{j=i+1}^{n} \mu_{j,i}\, z_j \right)^2 \right). \tag{4.8.1}$$

Proof. Let $M \in \mathbb{R}^{n\times n}$ be the upper triangular matrix with diagonal elements 1, and non-trival entries $m_{i,j} = \mu_{j,i}$, $1 \le i < j \le n$. Then $B = \overline{B}\, M$, where $\overline{B} = (\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n)$, and thus

$$q_B(\mathbf{x}) = \mathbf{x}^\intercal\, M^\intercal \overline{B}^\intercal\, \overline{B}\; M\, x = \sum_{i=1}^{n} \left(\left|\overline{\mathbf{b}}_i\right|^2 \left(x_i + \sum_{j=i+1}^{n} \mu_{j,i}\, x_j \right)^2 \right).$$

By Theorem 4.5 ii) we have $|\mu_{j,i}| \le \frac{1}{2}$, and moreover, we know that $\overline{\mathbf{b}}_k$ is a shortest non-trivial vector of the lattice $\Lambda|\mathrm{lin}\,\{\mathbf{b}_1, \ldots, \mathbf{b}_{k-1}\}^\perp$. A basis of this lattice is given by

$$\mathbf{c}_l = \overline{\mathbf{b}}_l + \sum_{j=k}^{l-1} \mu_{l,j} \overline{\mathbf{b}}_j, \quad l = k, \ldots, n,$$

(cf. Proposition 4.3 v)). Let $C = (\mathbf{c}_k, \ldots, \mathbf{c}_n)$, and let $M_k \in \mathbb{R}^{(n-k)\times(n-k)}$ be the upper triangular matrix consisting of the last $(n-k)$ rows and columns of M. Then $C = (\overline{\mathbf{b}}_k, \ldots, \overline{\mathbf{b}}_n)\, M_k$ and the relation $\left|\overline{\mathbf{b}}_k\right| \le |\sum_{i=k}^{n} z_i\, \mathbf{c}_i|$ for any non-trivial integral vector $(z_k, \ldots, z_n)^\intercal \in \mathbb{Z}^{n-k+1}$ translates into the second condition ii). □

4.9 Definition [HKZ-(reduced) quadratic p.d. form]. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix (p.d. matrix). The p.d. quadratic form*

$$q_A(\mathbf{x}) = \mathbf{x}^\intercal A\, \mathbf{x} = \sum_{i=1}^{n} \left(\alpha_i \Big(x_i + \sum_{j=i+1}^{n} \alpha_{j,i}\, x_j \Big)^2 \right)$$

is called a HKZ (reduced) quadratic p.d. form *if the (positive)* outer coefficients α_i *and the* inner coefficients $\alpha_{j,i}$ *satisfy* (4.8.1).

We recall that for each p.d. symmetric matrix A there exists a diagonal matrix D and an upper triangular matrix M such that $A = (D\, M)^\intercal D\, M$. Hence $q_A(x) = x^\intercal A\, x = \sum_{i=1}^{n} (\alpha_i (x_i + \sum_{j=i+1}^{n} \alpha_{j,i}\, x_j)^2)$ for suitable scalars α_i and $\alpha_{j,i}$.

4.10 Definition. *Two quadratic forms q_A, q_B are called (unimodular) equivalent if there exists an $U \in \mathrm{GL}(n,\mathbb{Z})$ such that $q_A(\mathbf{z}) = q_B(U\mathbf{z})$ for all $\mathbf{z} \in \mathbb{Z}^n$, i.e., $A = U^\intercal B\, U$.*

4.11 Corollary. *Each p.d. quadratic form q_A is equivalent to a HKZ reduced quadratic p.d. form.*

Proof. Let $A = L^\intercal L$ for some matrix $L \in \mathbb{R}^{n\times n}$, and let Λ be the lattice generated by the columns of L. According to Theorem 4.5 i) there exists a unimodular matrix U such that $L\,U$ is HKZ-reduced. Hence, q_B with $B = U^\intercal A\, U$ is a HKZ reduced quadratic p.d. form which is equivalent to q_A. □

4.12 Theorem*. *Let $q_A(\mathbf{x}) = \sum_{i=1}^n(\alpha_i(x_i + \sum_{j=i+1}^n \alpha_{j,i}\, x_j)^2)$ be a HKZ reduced quadratic p.d. form. Then*

$$\text{i) } \alpha_{k+1} \geq \frac{3}{4}\alpha_k \quad \textit{and} \quad \text{ii) } \alpha_{k+2} \geq \frac{2}{3}\alpha_k.$$

Proof. Let $q_{A,k}(x_k,\ldots,x_n) = \sum_{i=k}^n(\alpha_i(x_i + \sum_{j=i+1}^n \alpha_{j,i}\, x_j)^2)$. For i) we just observe that by definition (cf. (4.8.1) ii))

$$\alpha_k \leq q_{A,k}(0,1,0\ldots,0) = \alpha_{k+1,k}^2\,\alpha_k + \alpha_{k+1}. \tag{4.12.1}$$

Since $|\alpha_{k+1,k}| \leq \frac{1}{2}$ (cf. (4.8.1) i)), we have shown i). Unfortunately, ii) is much more involved and ...not here

□

4.13 Definition [$\mathbf{L^3}$-reduced basis, 1982]. *Let $\Lambda \in \mathcal{L}^n$. A basis $\mathbf{b}_1,\ldots,\mathbf{b}_n$ of Λ is called L^3-*reduced *if*

$$\begin{aligned} &\text{i) } \left|\overline{\mathbf{b}}_{i+1} + \mu_{i+1,i}\,\overline{\mathbf{b}}_i\right|^2 \geq \frac{3}{4}\left|\overline{\mathbf{b}}_i\right|^2, \quad 1 \leq i \leq n-1,\\ &\text{ii) } |\mu_{i,j}| \leq \frac{1}{2}, \quad 1 \leq j < i \leq n. \end{aligned} \tag{4.13.1}$$

4.14 Remark.

i) *$\overline{\mathbf{b}}_i$, $\overline{\mathbf{b}}_{i+1} + \mu_{i+1,i}\,\overline{\mathbf{b}}_i$ are both members of the $(n-i+1)$-lattice $\Lambda_{n-i+1} = \Lambda\big|\,\mathrm{lin}\,\{\mathbf{b}_1,\ldots,\mathbf{b}_{i-1}\}^\perp$. In the case of a HKZ-reduced basis we demand that $\overline{\mathbf{b}}_i$ is a shortest lattice vector of the lattice Λ_{n-i+1} and therefore the first condition of a L^3-reduced basis is a relaxation of this condition, since we just demand that $\overline{\mathbf{b}}_i$ is not too large in comparison with the vector $\overline{\mathbf{b}}_{i+1} + \mu_{i+1,i}\,\overline{\mathbf{b}}_i$.*

ii) *Obviously, a HKZ-basis is also an L^3-basis.*

4.15 Lemma. *Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be a basis of $\Lambda \in \mathcal{L}^n$. Then we have*

$$\lambda_1(B_n, \Lambda) \geq \min\left\{\left|\overline{\mathbf{b}}_1\right|, \dots, \left|\overline{\mathbf{b}}_n\right|\right\}.$$

Proof. Let $\mathbf{b} \in \Lambda \setminus \{0\}$. Then we may write it as a non-trivial integral combination of the basis vectors, i.e., $\mathbf{b} = \sum_{i=1}^{k} z_i \mathbf{b}_i$ with $z_i \in \mathbb{Z}$. Let k be the largest index with $z_k \neq 0$. Then

$$|\mathbf{b}| \geq \left|\mathbf{b}|\mathrm{lin}\,\{\mathbf{b}_1, \dots, b_{k-1}\}^{\perp}\right| = |z_k| \left|\overline{\mathbf{b}}_k\right| \geq \left|\overline{\mathbf{b}}_k\right|,$$

and the lemma is shown. □

4.16 Lemma. *Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be an L^3-reduced basis of $\Lambda \in \mathcal{L}^n$ with associated* GSO*-basis $\overline{\mathbf{b}}_1, \dots, \overline{\mathbf{b}}_n$. Then we have*

$$\left|\overline{\mathbf{b}}_i\right|^2 \leq 2^{j-i} \left|\overline{\mathbf{b}}_j\right|^2, \quad i \leq j \leq n.$$

Proof. Obviously, it suffices to show this for $j = i + 1$. From (4.13.1) i) we get

$$\frac{3}{4}\left|\overline{\mathbf{b}}_i\right|^2 \leq \left|\overline{\mathbf{b}}_{i+1}\right|^2 + (\mu_{i+1,i})^2 \left|\overline{\mathbf{b}}_i\right|^2,$$

and with (4.13.1) ii) we can conclude that $\left|\overline{\mathbf{b}}_i\right|^2 \leq 2\left|\overline{\mathbf{b}}_{i+1}\right|^2$. □

4.17 Theorem. *Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be an L^3-reduced basis of $\Lambda \in \mathcal{L}^n$. Then*

i) $|\mathbf{b}_1| \leq 2^{(n-1)/2}\lambda_1(B_n, \Lambda)$,

ii) $|\mathbf{b}_1| \leq 2^{(n-1)/4}(\det \Lambda)^{1/n}$,

iii) $|\mathbf{b}_1| \cdot \ldots \cdot |\mathbf{b}_n| \leq 2^{n(n-1)/4} \det \Lambda$.

Proof. From Lemma 4.16 we conclude

$$\left|\overline{\mathbf{b}}_k\right|^2 \geq \left(\frac{1}{2}\right)^{k-i} \left|\overline{\mathbf{b}}_i\right|^2, \quad 1 \leq i \leq k \leq n, \tag{4.17.1}$$

and, in particular, for $i = 1$ we get

$$\left|\overline{\mathbf{b}}_k\right|^2 \geq \left(\frac{1}{2}\right)^{k-1} \left|\overline{\mathbf{b}}_1\right|^2, \quad 1 \leq k \leq n. \tag{4.17.2}$$

In view of Lemma 4.15 we obtain

$$\lambda_1(B_n, \Lambda) \geq \min\left\{\left|\overline{\mathbf{b}}_1\right|, \dots, \left|\overline{\mathbf{b}}_n\right|\right\} \geq \left(\frac{1}{2}\right)^{(n-1)/2} \left|\overline{\mathbf{b}}_1\right|,$$

which shows i). Moreover, by (4.17.2) we may also write

$$(\det \Lambda)^2 = \prod_{k=1}^{n} |\bar{\mathbf{b}}_k|^2 \geq \prod_{k=1}^{n} \left(\frac{1}{2}\right)^{k-1} |\bar{\mathbf{b}}_1|^2 = \left(\frac{1}{2}\right)^{\binom{n}{2}} |\mathbf{b}_1|^{2n},$$

which gives ii). Finally, by the definition of the GSO-basis and taking into account (4.17.1) we find

$$|\mathbf{b}_k|^2 \leq |\bar{\mathbf{b}}_k|^2 + \frac{1}{4}\sum_{i=1}^{k-1} |\bar{\mathbf{b}}_i|^2 \leq |\bar{\mathbf{b}}_k|^2 + \left(\frac{1}{4}\sum_{i=1}^{k-1} 2^{k-i}\right) |\bar{\mathbf{b}}_k|^2 \leq 2^{k-1} |\bar{\mathbf{b}}_k|^2.$$

Hence

$$\prod_{k=1}^{n} |\mathbf{b}_k|^2 \leq \prod_{k=1}^{n} |\bar{\mathbf{b}}_k|^2 \prod_{k=1}^{n} 2^{k-1} = (\det \Lambda)^2 \, 2^{\binom{n}{2}},$$

and iii) is proved. □

4.18 Remark. *Some applications of L^3-reduced bases:*

i) *Factorization of a polynomial $f \in \mathbb{Q}[x]$ into irreducible factors (Lenstra, A.K., Lenstra, H.W. Jr., Lovász, 1982).*

ii) *Integer programming problems can be solved in polynomial time* in fixed dimension *(Lenstra, H.W. Jr., 1983).*

iii) *Disproof of the Mertens conjecture: $M(n) = \sum_{k=1}^{n} \mu(k) < \sqrt{n}$, where $\mu(k)$ is the Möbius function, i.e., $\mu(k) = 1$ if $k = 1$, $\mu(k) = 0$ if k has a repeated prime factor and $\mu(k) = (-1)^l$ if k is the product of l distinct primes (Odlyzko, te Riele, 1985). Any bound of the type $M(n) \leq c\sqrt{n}$, where c is a constant, would imply the Riemann hypothesis.*

iv) *Let $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}$, $0 < \varepsilon < 1$, and let $N \geq 2^{n(n+1)/4}\varepsilon^{-n}$. Then there exists a polynomial time algorithm (cf. Theorem 4.22) which computes $p_1, \ldots, p_n, q \in \mathbb{Z}$, $1 \leq q \leq N$ such that*

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{\varepsilon}{q}, \quad 1 \leq i \leq n.$$

v) *Let $\Lambda \subset \mathbb{Q}^n$ be a lattice and $\mathbf{w} \in \mathbb{Q}^n$. Then there exists a polynomial time algorithm (cf. Theorem 4.22) which computes an $\mathbf{a} \in \Lambda$ such that*

$$|\mathbf{w} - \mathbf{a}| \leq 2^{n/2} \min\{|\mathbf{w} - \mathbf{b}| : \mathbf{b} \in \Lambda\}$$

(Babai, 1986).

Proof. We prove iv) and v). In order to show iv) let $\Lambda \subset \mathbb{R}^{n+1}$ be the lattice with basis

$$\mathbf{e}_1, \ldots, \mathbf{e}_n, (\alpha_1, \ldots, \alpha_n, \epsilon/N)^\mathsf{T},$$

and let $\mathbf{b}_1, \ldots, \mathbf{b}_{n+1}$ be an L^3-reduced basis of Λ. From Theorem 4.17 ii) and the choice of N we get

$$|\mathbf{b}_1| \leq 2^{n/4} (\det \Lambda)^{1/(n+1)} \leq \varepsilon. \tag{4.18.1}$$

If we write $\mathbf{b}_1$ as integral combination of the given basis we obtain $\mathbf{b}_{1,i} = p_i - q\,\alpha_i$, $1 \leq i \leq n$, and $\mathbf{b}_{1,n+1} = -q\,\epsilon/N$, where $p_1, \ldots, p_n, q$ are integers. Since $\varepsilon < 1$ we conclude $q \neq 0$, and so we may assume $q \geq 1$. In view of (4.18.1) we have $|p_i - q\,\alpha_i| < \varepsilon$ and $|\mathbf{b}_{1,n+1}| \leq \varepsilon$ we get $q \leq N$.

Now we prove v). Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be an L^3-reduced basis of Λ with associated GSO-basis $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$. First we construct a point $a \in \Lambda$ such that

$$\mathbf{w} - \mathbf{a} = \sum_{i=1}^{n} \nu_i\, \overline{\mathbf{b}}_i, \quad \nu_i \in \mathbb{R},\ |\nu_i| \leq \frac{1}{2},\ 1 \leq i \leq n. \tag{4.18.2}$$

Let $\mathbf{w} = \sum_{i=1}^{n} \nu_{i,n} \overline{\mathbf{b}}_i$, for some suitable $\nu_{i,n} \in \mathbb{R}$. Then we have

$$\mathbf{w} - [\nu_{n,n}]\, \mathbf{b}_n = \left(\sum_{i=1}^{n-1} \nu_{i,n-1} \overline{\mathbf{b}}_i \right) + (\nu_{n,n} - [\nu_{n,n}])\, \overline{\mathbf{b}}_n,$$

for some scalars $\nu_{i,n-1} \in \mathbb{R}$ with $1 \leq i \leq n-1$. In the next step we subtract $[\nu_{n-1,n-1}]\mathbf{b}_{n-1}$ from the left hand side which results in a representation

$$\begin{aligned} \mathbf{w} - [\nu_{n,n}]\, \mathbf{b}_n - [\nu_{n-1,n-1}]\, \mathbf{b}_{n-1} = &\left(\sum_{i=1}^{n-2} \nu_{i,n-2}, \overline{\mathbf{b}}_i \right) \\ &+ (\nu_{n-1,n-1} - [\nu_{n-1,n-1}])\, \overline{\mathbf{b}}_{n-1} + (\nu_{n,n} - [\nu_{n,n}])\, \overline{\mathbf{b}}_n, \end{aligned}$$

with $\nu_{i,n-2} \in \mathbb{R}$. Continuing this procedure for all $\mathbf{b}_{n-2}, \ldots, \mathbf{b}_1$ we obtain a point $a = \sum_{i=1}^{n} [\nu_{i,i}] \mathbf{b}_i \in \Lambda$ satisfying (4.18.2).

Now let $\mathbf{c} \in \Lambda$ such that $|\mathbf{w} - \mathbf{c}| = \min \{ |\mathbf{w} - \mathbf{b}| : \mathbf{b} \in \Lambda \}$ and let

$$\mathbf{w} - \mathbf{c} = \sum_{i=1}^{n} \mu_i\, \overline{\mathbf{b}}_i.$$

Let $k \in \{1, \ldots, n\}$ be the largest index with $\nu_k \neq \mu_k$. Then we have

$$\mathbf{c} - \mathbf{a} = \sum_{i=1}^{k} (\nu_i - \mu_i)\, \overline{\mathbf{b}}_i = \sum_{i=1}^{k-1} \rho_i\, \overline{\mathbf{b}}_i + (\nu_k - \mu_k)\, \mathbf{b}_k.$$

Since the left hand side lies in Λ we conclude that $|\nu_k - \mu_k| \geq 1$, which implies $|\mu_k| \geq 1/2$. Together with Lemma 4.16 we finally get

$$|\mathbf{w}-\mathbf{a}|^2 \leq \sum_{i=1}^{k} \frac{1}{4}\left|\overline{\mathbf{b}}_i\right|^2 + \sum_{i=k+1}^{n} \nu_i^2 \left|\overline{\mathbf{b}}_i\right|^2 \leq \frac{1}{4}\sum_{i=1}^{k} 2^{k-i}\left|\overline{\mathbf{b}}_k\right|^2 + \sum_{i=k+1}^{n} \mu_i^2 \left|\overline{\mathbf{b}}_i\right|^2$$
$$< 2^k\left(\frac{1}{4}\left|\overline{\mathbf{b}}_k\right|^2 + \sum_{i=k+1}^{n} \mu_i^2 \left|\overline{\mathbf{b}}_i\right|^2\right) \leq 2^k |\mathbf{w}-\mathbf{c}|^2 .$$

□

One of the main advantages of a L^3-reduced basis is that it can be found in polynomial time. Here, we will call an algorithm *polynomial* if the number of elementary operations (addition, multiplication, division) needed by the algorithm is bounded by polynomial function in the *input size* of the algorithm. Thereby, the input size is the *encoding length* of the input data of the algorithm. The encoding length $\langle m\rangle$ of an integer $m \in \mathbb{Z}$ is defined by

$$\langle m\rangle = 1 + \left\lceil \log_2\bigl(|m|+1\bigr)\right\rceil,$$

and may be regarded as the number of binary bits (including the sign) needed to store the number. For a rational number r which can be uniquely written as p/q with $q > 0$ and p and q co-prime integers, the encoding length $\langle r\rangle$ is given by $\langle r\rangle = \langle p\rangle + \langle q\rangle$. The encoding length $\langle b\rangle$ of a vector $b \in \mathbb{Q}^n$ is just the sum of the encoding lengths of its components, and for a system of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{Q}^n$ the encoding length is given by $\langle \mathbf{b}_1, \ldots, \mathbf{b}_k\rangle = \langle \mathbf{b}_1\rangle + \cdots + \langle \mathbf{b}_k\rangle$. There are some useful relations between encoding length and other functionals.

4.19 Lemma.

i) *Let $r \in \mathbb{Q}$. Then $|r| \leq 2^{\langle r\rangle - 1} - 1$.*

ii) *Let $\mathbf{b} \in \mathbb{Q}^n$. Then $|\mathbf{b}| \leq 2^{\langle \mathbf{b}\rangle - n} - 1$.*

iii) *Let $D \in \mathbb{Q}^{n\times n}$. Then $|\det D| \leq 2^{\langle D\rangle - n^2} - 1$.*

We remark that, actually, an algorithm is called polynomial if the number of bit operations carried by the algorithm is bounded by a polynomial in the input size. So this definition also involves the binary size of the numbers produced during the performance of the algorithm, but for sake of simplification we just count the number of elementary operations. From this point of view it is easy to see that

4.20 Remark. *Let $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Q}^n$ be linearly independent. The associated Gram-Schmidt orthogonalized basis $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$ with coefficients $\mu_{i,j}$ can be computed by carrying out $O(n^3)$ elementary operations.*

Before presenting an algorithm for computing an L^3-reduced basis we show how one can find a basis of a lattice such that the second condition (4.13.1) ii) of a L^3-reduced basis is satisfied.

4.21 Proposition. *Let $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Q}^n$ be a basis of a lattice $\Lambda \in \mathcal{L}^n$ with* GSO*-basis $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$. Then a basis $\mathbf{c}_1, \ldots, \mathbf{c}_n$ of Λ can be computed in time $O(n^4)$ such that for the associated* GSO*-basis $\overline{\mathbf{c}}_1, \ldots, \overline{\mathbf{c}}_n$ and its coefficients $\gamma_{i,j}$ hold*

$$\text{i)} \quad |\gamma_{i,j}| \leq \frac{1}{2},\ 1 \leq j < i \leq n, \qquad\qquad \text{ii)} \quad \bar{\mathbf{c}}_i = \bar{\mathbf{b}}_i,\ 1 \leq i \leq n.$$

Proof. Let $\mu_{i,j}$ be the coefficients of the GSO-basis $\overline{\mathbf{b}}_1, \ldots, \overline{\mathbf{b}}_n$. If $|\mu_{i,j}| \leq 1/2$, $1 \leq j < i \leq n$, then there is nothing to show. So we suppose $|\mu_{k,l}| > 1/2$ and let

$$\mathbf{b}_i^{(1)} = \mathbf{b}_i,\ 1 \leq i \neq k \leq n \quad \text{and} \quad \mathbf{b}_k^{(1)} = \mathbf{b}_k - [\mu_{k,l}]\,\mathbf{b}_l.,$$

where, for $\mu \in \mathbb{R}$, $[\mu]$ denotes a nearest integer. Obviously, $\mathbf{b}_1^{(1)}, \ldots, \mathbf{b}_n^{(1)}$ is a basis of Λ and moreover, $\operatorname{lin}\{\mathbf{b}_1^{(1)}, \ldots, \mathbf{b}_i^{(1)}\} = \operatorname{lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_i\}$, $1 \leq i \leq n$, since $l < k$. Hence $\overline{\mathbf{b}}_i^{(1)} = \overline{\mathbf{b}}_i$ for $1 \leq i \leq n$. Next we study the coefficients $\mu_{i,j}^{(1)}$ of the Gram-Schmidt orthogonalized basis $\overline{\mathbf{b}}_1^{(1)}, \ldots, \overline{\mathbf{b}}_n^{(1)}$. By their definition we get

$$\begin{aligned}
\mu_{i,j}^{(1)} &= \frac{\left\langle \mathbf{b}_i^{(1)}, \overline{\mathbf{b}}_j^{(1)} \right\rangle}{\left|\overline{\mathbf{b}}_j^{(1)}\right|^2} = \frac{\langle \mathbf{b}_i, \overline{\mathbf{b}}_j \rangle}{\left|\overline{\mathbf{b}}_j\right|^2} = \mu_{i,j}, \quad i \neq k,\ 1 \leq j < i \leq n, \\
\mu_{k,j}^{(1)} &= \frac{\left\langle \mathbf{b}_k^{(1)}, \overline{\mathbf{b}}_j^{(1)} \right\rangle}{\left|\overline{\mathbf{b}}_j^{(1)}\right|^2} = \frac{\langle (\mathbf{b}_k - [\mu_{k,l}]\,\mathbf{b}_l), \overline{\mathbf{b}}_j \rangle}{\left|\overline{\mathbf{b}}_j\right|^2} = \mu_{k,j}, \quad l < j < k, \\
\mu_{k,l}^{(1)} &= \mu_{k,l} - [\mu_{k,l}].
\end{aligned}$$

So we have $\left|\mu_{k,l}^{(1)}\right| \leq 1/2$ and the coefficients $\mu_{k,j}^{(1)}$, $1 \leq j \leq l$, are the only ones which can change by replacing the vector $\mathbf{b}_k$ by $\mathbf{b}_k - [\mu_{k,l}]\,\mathbf{b}_l$. Therefore, the output of the following algorithm gives a basis $\mathbf{c}_1, \ldots, \mathbf{c}_n$ of Λ with the required properties:

```
for k = 2, n do
 for l = k − 1, 1 do
  if |μ_{k,l}| > 1/2 then
   replace b_k by b_k − [μ_{k,l}] b_l;
   update the GSO-coefficients μ_{k,j},  1 ≤ j ≤ l
  endif
```

`continue; continue;`
`set` $\mathbf{c}_i = \mathbf{b}_i$,
$\gamma_{i,j} = \mu_{i,j}$, $1 \leq \ j < i \leq n$.

If $|\mu_{k,l}| > 1/2$ then we have to carry out $O(l\,n)$ operations in order to update our basis and hence the running time of the algorithm is $O(n^4)$. □

4.22 Theorem. *Given a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Q}^n$ of a lattice $\Lambda \in \mathcal{L}^n$, there exists a polynomial time algorithm which computes an L^3-reduced basis.*

Proof. First we state the algorithm that produces a L^3-reduced basis. For the sake of a simple presentation we subdivide the algorithm into 3 basic steps.

(0) Let $\mathbf{b}_k^{(0)} = \mathbf{b}_k$, $1 \leq k \leq n$, and set $l = 0$.

(1) For a basis $\mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_n^{(l)}$ of the lattice Λ we compute the Gram-Schmidt orthogonalized basis $\overline{\mathbf{b}}_1^{(l)}, \ldots, \overline{\mathbf{b}}_n^{(l)}$ with coefficients $\mu_{i,j}^{(l)}$.

(2) Compute a basis $\mathbf{c}_1, \ldots, \mathbf{c}_n$, say, of Λ such that for the Gram-Schmidt orthogonalized basis $\overline{\mathbf{c}}_1, \ldots, \overline{\mathbf{c}}_n$ and its coefficients $\gamma_{i,j}$ verify

$$\overline{\mathbf{c}}_i = \overline{\mathbf{b}}_i^{(l)},\ 1 \leq i \leq n, \quad \text{and} \quad |\gamma_{i,j}| \leq 1/2,\ 1 \leq j < i \leq n.$$

Set $\mathbf{b}_k^{(l)} = c_k$, $1 \leq k \leq n$, and $\mu_{i,j}^{(l)} = \gamma_{i,j}$, $1 \leq j < i \leq n$.

(3) If for an $i \in \{1, \ldots, n-1\}$

$$\left|\overline{\mathbf{b}}_{i+1}^{(l)} + \mu_{i+1,i}^{(l)} \overline{\mathbf{b}}_i^{(l)}\right|^2 < (3/4)\left|\overline{\mathbf{b}}_i^{(l)}\right|^2 \text{ then}$$

set $\mathbf{b}_k^{(l+1)} = \mathbf{b}_k^{(l)}$ for $1 \leq k \leq n$ but $k \notin \{i, i+1\}$, and set $\mathbf{b}_i^{(l+1)} = \mathbf{b}_{i+1}^{(l)}$, $\mathbf{b}_{i+1}^{(l+1)} = \mathbf{b}_i^{(l)}$. Set $l = l + 1$ and goto Step (1).

Otherwise $\mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_n^{(l)}$ is an L^3-reduced basis.

After performing step (2) the basis $\mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_n^{(l)}$ satisfies the condition ii) of (4.13.1). In step (3) we take care of the first condition of (4.13.1) and thus, if the algorithms stops at all, we have found an L^3-reduced basis. Moreover, we have already seen in Proposition 4.21 how we can perform step (2) and thus we just have to check the finiteness of the algorithm. To this end we have to analyze the consequences of step (3) and without loss of generality we assume that the vectors $\mathbf{b}_1^{(0)}, \ldots, \mathbf{b}_n^{(0)}$ are integral. We suppose that

$$\left|\overline{\mathbf{b}}_{i+1}^{(l)} + \mu_{i+1,i}^{(l)} \overline{\mathbf{b}}_i^{(l)}\right|^2 < \frac{3}{4}\left|\overline{\mathbf{b}}_i^{(l)}\right|^2, \tag{4.22.1}$$

and let $\mathbf{b}_k^{(l+1)} = \mathbf{b}_k^{(l)}$, $k \notin \{i, i+1\}$, $\mathbf{b}_i^{(l+1)} = \mathbf{b}_{i+1}^{(l)}$ and $\mathbf{b}_{i+1}^{(l+1)} = \mathbf{b}_i^{(l)}$. With respect to the GSO-bases we find that $\overline{\mathbf{b}}_k^{(l+1)} = \overline{\mathbf{b}}_k^{(l)}$ for $k \notin \{i, i+1\}$ and

$$\overline{\mathbf{b}}_i^{(l+1)} = \mathbf{b}_{i+1}^{(l)} \big| \operatorname{lin}\{\mathbf{b}_1^{(l)}, \ldots, \mathbf{b}_{i-1}^{(l)}\}^{\perp} = \overline{\mathbf{b}}_{i+1}^{(l)} + \mu_{i+1,i}^{(l)} \overline{\mathbf{b}}_i^{(l)}.$$

Hence by (4.22.1) we have

$$\left|\overline{\mathbf{b}}_i^{(l+1)}\right|^2 < \frac{3}{4}\left|\overline{\mathbf{b}}_i^{(l)}\right|^2. \tag{4.22.2}$$

Since $(\det \Lambda)^2 = \left|\overline{\mathbf{b}}_1^{(l)}\right|^2 \cdot \ldots \cdot \left|\overline{\mathbf{b}}_n^{(l)}\right|^2 = \left|\overline{\mathbf{b}}_1^{(l+1)}\right|^2 \cdot \ldots \cdot \left|\overline{\mathbf{b}}_n^{(l+1)}\right|^2$ we also have

$$\left|\overline{\mathbf{b}}_i^{(l)}\right|^2 \left|\overline{\mathbf{b}}_{i+1}^{(l)}\right|^2 = \left|\overline{\mathbf{b}}_i^{(l+1)}\right|^2 \left|\overline{\mathbf{b}}_{i+1}^{(l+1)}\right|^2$$

and together with (4.22.2) we get

$$\begin{aligned}\left|\overline{\mathbf{b}}_i^{(l+1)}\right|^{2(n-i+1)} \left|\overline{\mathbf{b}}_{i+1}^{(l+1)}\right|^{2(n-i)} &= \left|\overline{\mathbf{b}}_i^{(l+1)}\right|^2 \left|\overline{\mathbf{b}}_i^{(l)}\right|^{2(n-i)} \left|\overline{\mathbf{b}}_{i+1}^{(l)}\right|^{2(n-i)} \\ &< \frac{3}{4}\left|\overline{\mathbf{b}}_i^{(l)}\right|^{2(n-i+1)} \left|\overline{\mathbf{b}}_{i+1}^{(l)}\right|^{2(n-i)}.\end{aligned} \tag{4.22.3}$$

Now let $D^{(l)} = \prod_{j=1}^{n} \left|\overline{\mathbf{b}}_j^{(l)}\right|^{2(n-j+1)}$. Then (4.22.3) shows that

$$D^{(l+1)} < \frac{3}{4} D^{(l)},$$

whenever we have to execute the swapping in step (3). In other words, as long as the algorithm does not stop, $D^{(l)}$ is diminished by a factor of at least 3/4. Now since

$$D^{(l)} = \prod_{j=1}^{n} \prod_{k=1}^{j} \left|\overline{\mathbf{b}}_k^{(l)}\right|^2$$

and since $\prod_{k=1}^{j} \left|\overline{\mathbf{b}}_k^{(l)}\right|^2$ is the square of the determinant of the j-dimensional lattice spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_j$, we conclude that $D^{(l)}$ is a positive integer. Hence step (3) can be executed at most $\log D^{(0)} / \log(4/3)$-times. Finally we observe that (cf. Lemma 4.19 ii))

$$\log_2 D^{(0)} \le n \sum_{k=1}^{n} \log_2 \left|\mathbf{b}_k^{(0)}\right|^2 < 2\,n \sum_{k=1}^{n} \langle \mathbf{b}_k^{(0)} \rangle = 2\,n \langle \mathbf{b}_1^{(0)}, \ldots, \mathbf{b}_n^{(0)} \rangle.$$

Therefore, in view of Remark 4.20 and Proposition 4.21, the number of elementary operations is bounded by $O(n^5 \langle \mathbf{b}_1^{(0)}, \ldots, \mathbf{b}_n^{(0)} \rangle)$ and thus it is bounded by $O\left(\langle \mathbf{b}_1^{(0)}, \ldots, \mathbf{b}_n^{(0)} \rangle^5\right)$. Hence, the number of elemenatry operations is polynomial in the input size $\langle \mathbf{b}_1^{(0)}, \ldots, \mathbf{b}_n^{(0)} \rangle$ of the algorithm. □

Bibliography

[1] Keith Ball. An elementary introduction to modern convex geometry. *Cambridge University Press. Math. Sci. Res. Inst. Publ.*, 31:1 – 58, 1997.

[2] Alexander Barvinok. A course in convexity. *Graduate Studies in Mathematics*, 54, 2002.

[3] T Figiel, J Lindenstrauss, and V.D Milman. The dimension of almost spherical sections of convex bodies. *Acta Math*, 139:53–94, 1977.

[4] R. J Gardner. The brunn-minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405 (electronic), 2002.

[5] Richard J Gardner. *Geometric Tomography*. 1995.

[6] Peter Manfred Gruber. *Convex and Discrete Geometry*, volume 336. 2007.

[7] Branko Grünbaum. *Convex polytopes*. 2003.

[8] G Kalai. The number of faces of centrally-symmetric polytopes. *Graphs and Combinatorics*, pages 389–391, Sep 1989.

[9] Alexander Koldobsky. *Fourier Analysis in Convex Geometry*, volume 116. 2005.

[10] Jesus De Loera, Jörg Rambau, and Francisco Santos. Triangulations: Structures for algorithms and applications. *Book*, pages 1–545, Apr 2010.

[11] Jiri Matousek. *Lectures on Discrete Geometry*, volume 212.

[12] R Sanyal, A Werner, and G.M Ziegler. On kalai's conjecture concerning centrally symmetric polytopes. *Discrete & Computational Geometry*, 41:183—198, Sep 2009. doi: 10.1007/s0054-008-9104-8.

[13] R Schneider. *Convex bodies: The Brunn-Minkowski theory*, volume 44. 1993.

[14] G M Ziegler. *Lectures on polytopes*, volume 152. 1995. Revised sixth printing 2006.

Index