Lattices in Optimization

Lecture Notes, Summer 2022

PD Dr. Andreas Paffenholz

September 2, 2022

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Version of September 2, 2022

Contents

Preface

These notes have been written for the class *Integer Points in Polyhedra* at TU Darmstadt in summer 2022. They are based on previous notes on a similar course at TU Berlin in 2019, and on the notes of a course on *Lattice Polytopes* at FU Berlin in 2007, that I prepared together with Benjamin Nill and Christian Haase.[1]

Sections marked with an asterix (*) in front have not been covered in the course.

Please contact me if you find errors or typos in the script, or if you have suggestions for improvements or additions.

Darmstadt, summer 2022 Andreas Paffenholz

[1] Haase, Nill, and Paffenholz, *Lattice Polytopes*.

Version of September 2, 2022

Timeline

	Date	Topics
01	April 13	Motivation summary of the course organisation
02	April 20	Lattices every lattice has a basis Hermite normal form
03	April 27	Hermite normal form Sublattices Smith normal form Dual lattices
04	May 04	Geometry of Numbers Blichfeldt's Theorem Minkowski's First and Second Theorem Short vectors
05	May 11	Minkowski's Second Theorem Dirichlet's Theorem Covering Packing
06	May 18	Covering Packing Flatness Shortest Vector
07	May 25	lecture postponed to June 15 and 22
08	June 01	lecture postponed to June 19 and July 06
09	June 08	lecture postponed to July 13
10	June 15	Shortest Vector Problem Reduced Bases LLL algorithm
11	June 22	Integer programming in fixed dimension Minimally infeasible subsets Subset-Sum Problems

12	June 29	Subset-Sum Problems Closest Vector Problem
13	July 06	Generating Functions Counting Lattice Points
14	July 13	Barvinok's Algorithm Lattice Points in Dilates

1. Introduction

In its most common form the basic problem in *Discrete* or *Integer Optimization* is the task to find integral solutions for a linear programming problem

$$\max\left(\mathbf{c}^t\mathbf{x} \;:\; A\mathbf{x} \;\leq\; \mathbf{b}\right)\,, \tag{LP}$$

for a system of linear inequalities $A\mathbf{x} \leq \mathbf{b}$ with a rational right hand side $\mathbf{b} \in \mathbb{Q}^m$ and a rational matrix $A \in \mathbb{Q}^{d\times m}$, that is to solve the problem

$$\max\left(\mathbf{c}^t\mathbf{x} \;:\; A\mathbf{x} \;\leq\; \mathbf{b} \text{ and } \mathbf{x} \in \mathbb{Z}^d\right)\,, \tag{IP}$$

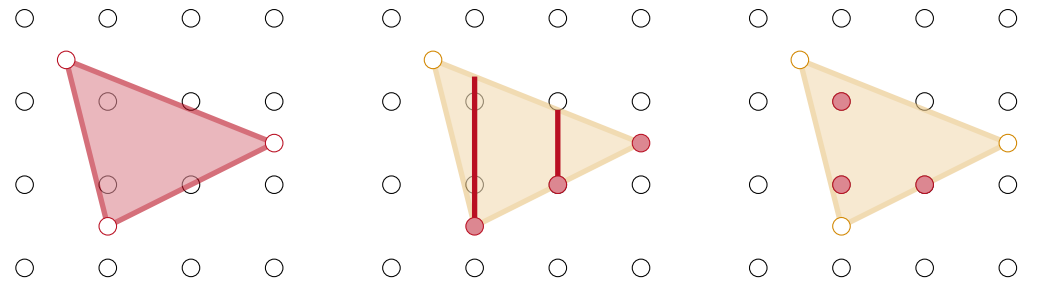
The set of *feasible points* of (LP) is a *polyhedron*

$$P \;=\; P(A,\mathbf{b}) \;:=\; \{\,\mathbf{x} \;:\; A\mathbf{x} \leq \mathbf{b}\,\}\,,$$

where A is a (rational) matrix in $Q^{m\times d}$ and $\mathbf{b} \in \mathbb{Q}^m$ is a (rational) vector. The feasible points of (IP) is the intersection of P with the set of integer points, *i.e.* $P \cap \mathbb{Z}^d$. One may also consider programs where only some of the variables are required to be integral, so called *mixed-integer programs*

$$\max\left(\mathbf{c}^t\mathbf{x} \;:\; A\mathbf{x} \;\leq\; \mathbf{b} \text{ and } \mathbf{x} \in \mathbb{Z}^k \times \mathbb{R}^{d-k}\right)\,, \tag{MIP}$$

for some $1 \leq k \leq d$. We will restrict to pure integer programs of the form (IP), but with some, often mostly technical, effort, corresponding results can also be obtained for these mixed-integer programs. See Figure 1.1 for an illustration of the sets of feasible points in each case.

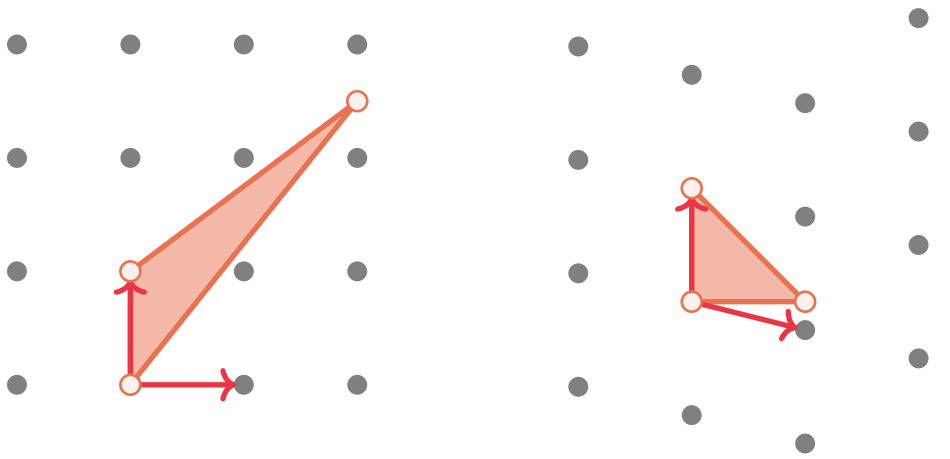


(a) The feasible region of a linear program

(b) The set of feasible points of a mixed integer program where $x_1 \in \mathbb{Z}$

(c) The set of feasible points of an integer program

Figure 1.1.: Feasible regions



(a) A triangle in the standard basis

(b) The same triangle normalized, with the transformed lattice

Figure 1.2.: An example

Sets of feasible points for integer programs, or their convex hull, the *integer hull* of the polyhedron P, are not yet well understood. In this notes we will discuss various geometric and algorithmic topics connected to such sets of integer points in polyhedra.

Despite the apparent similarity between (LP) and (IP), these two problems differ vastly in various respects. Solutions to (LP) may be arbitrarily far away from solutions to (IP), even in small dimensions. For a simple example, we can consider the triangle given by

$$P^k := \mathsf{conv}\left(\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right], \tfrac{1}{2}\left[\begin{smallmatrix}k-1\\k\end{smallmatrix}\right]\right)$$

See Figure 1.2(a) for $k = 5$. If we set $\mathbf{c}^t := (1, 1)$, then the solution to (LP) is $2k - \frac{1}{2}$, realized by the point $\frac{1}{2}\left[\begin{smallmatrix}k-1\\k\end{smallmatrix}\right]$. However, the solution to (IP) is 1, realized by $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$.

While solving linear programs it is often useful to *normalize* the problem in some way, *i.e.* to apply an affine transformation $\varphi(\mathbf{x}) := T\mathbf{x} + \mathbf{t}$ for some (non-singular) matrix T and a translation $\mathbf{t}$ to the polyhedron P of feasible points in such a way that P has a *nice* representation, *e.g.* in the sense that the sizes of the entries of the constraint matrix don't differ too much. We can solve the transformed problem and apply the inverse map φ^{-1} to the solution to obtain a solution of the original problem.

We cannot necessarily do the same with integer linear programs, as a bijective affine transformation φ need not preserve integrality. Hence, we either have to restrict to affine maps that map $\mathbb{Z}^d$ into $\mathbb{Z}^d$ or we have to generalize our notion of integrality to solutions in $T(\mathbb{Z}^d)$ instead of $\mathbb{Z}^d$. Such images $\Lambda := T(\mathbb{Z}^d)$ of $\mathbb{Z}^d$ are called *lattices*. See Figure 1.2(b) for the triangle of Figure 1.2(a) after an affine transformation that maps the triangle into a standard triangle. Sets $\Lambda := T(\mathbb{Z}^d)$ can also be seen as the set of all linear combinations of the vectors $T\mathbf{e}_i$ for the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_d$ with *integral*

coefficients,

$$\Lambda := \left\{ \sum_{i=1}^{d} \lambda_i \mathbf{e}_i \ : \ \lambda_i \in \mathbb{Z} \text{ for } 1 \leq i \leq d \right\}.$$

The set $\{ T\mathbf{e}_1, \ldots, T\mathbf{e}_d \}$ is called a *basis* of Λ in analogy with the bases in linear algebra.

More generally, we may face the optimization problem to find solutions in some discrete subset Λ that satisfy linear constraints $A\mathbf{x} \leq \mathbf{b}$ defining a polyhedron P. If Λ is a subgroup of $\mathbb{R}^d$ (*i.e.* it is closed under taking sums and inverses), then we say that Λ is a *lattice*. For example, we may require that some or all of the coordinates of the solution should be even, *i.e.* feasible solutions should be in $\Lambda := (2\mathbb{Z})^k \times \mathbb{Z}^{d-k}$ for some k. We will see that such sets are always of the form $\varphi(\mathbb{Z}^d)$ for some linear map $\varphi(\mathbf{x}) := T\mathbf{x}$. Thus, in principle we can reduce to the problem of finding integer solutions. Yet, as this also transforms the polyhedron P, which may vary its shape, its volume, or the size of the coordinates, this may not always be desirable.

Sometimes we are given the discrete structure of the feasible solutions not with the standard basis of $\mathbb{Z}^d$, but with some other basis B, and a solution is feasible if it can be written as an integer linear combination in this basis (and satisfies some linear constraints). We then face the problem to find a suitable *simple* basis for the discrete set of solutions, or the problem to decide if our basis actually generates $\mathbb{Z}^d$ or only some proper subset. For example, consider the two bases in Figure 1.3. Both in fact generate $\mathbb{Z}^2$, but you may feel that this is easier to check for the basis in Figure 1.3(a).

There are two tasks connected with this. The first is to decide if both bases generate the same lattice, which, as we will see, can be solved efficiently with the computation of the Hermite normal form.

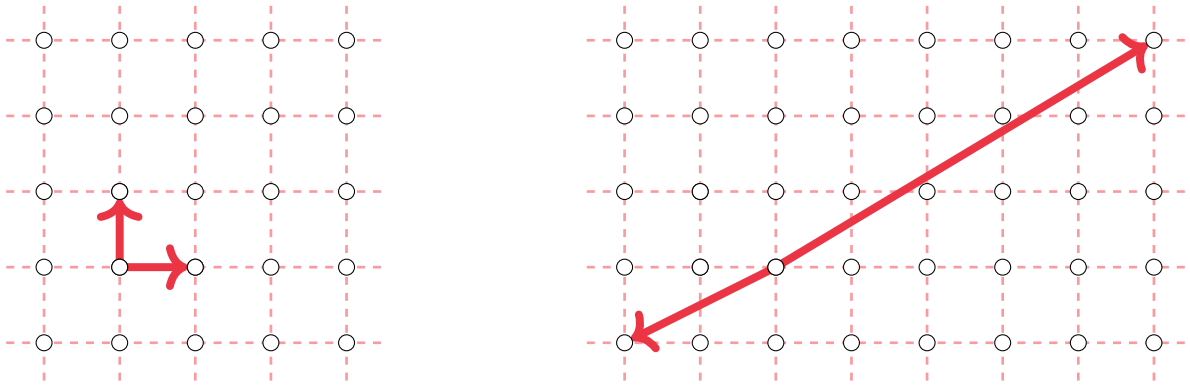
The second task turns out to be harder. Comparing the two bases in Figure 1.3 we see that the first has *shorter* vectors and they are orthogonal. In the linear setting we can easily achieve this, *e.g.*, with Gram-Schmidt-Orthogonalization. For lattices, such a basis need not exist, and finding at least some approximation with *short* and *almost* orthogonal vectors is surprisingly difficult. We will discuss the celebrated LLL-algorithm of A.K. Lenstra, H.W. Lentra Jr., and L. Lovász,[1] that computes such a basis and is the foundation of many other algorithms in this field.

If we want to have short vectors in our basis we may, as a first step, ask for one *shortest* nontrivial vector (say, in the Euclidean norm) the is in the integral span of our basis. This is the *Shortest Vector Problem*. Already this problem turns out to be difficult to solve, and we can efficiently only produce approximations, based on the LLL-algorithm.

Given some point $\mathbf{x} \in \mathbb{R}^d$ we may also ask for the *closest* point in the integral span, which leads to the *Closest Vector Problem*. Although seemingly similar to the shortest vector problem it is not quite, as we cannot just translate $\mathbf{x}$ into the origin. We would need to translate Λ as well, but then 0 is not contained in the lattice anymore, unless $\mathbf{x} \in \Lambda$.

The problem (IP) may not have a solution at all, even if the first is feasible. Strips of the form $P := [1/3, 2/3] \times [a, b]$ for $a < b$ show that this can happen for arbitrarily large polyhedra. Note, however, that P appears to be *thin* in some direction. We will see that

[1] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".



(a) The standard *short* basis of $\mathbb{Z}^2$.

(b) Another basis of $\mathbb{Z}^2$.

Figure 1.3.: Bases of $\mathbb{Z}^2$

this is, in some sense, true for any such example.

More precisely, we will consider polyhedra $P := \{\, \mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b} \,\}$ such that $P \cap \Lambda$ is empty. The *Flatness Theorem* of Khinchine then tells us that there is a universal bound $c(d)$ only depending on the dimension, such that the *lattice width* $\operatorname{width}_{\mathbb{Z}^d}(P)$ of P satisfies

$$\operatorname{width}_{\mathbb{Z}^d}(P) \;:=\; \min_{\mathbf{c}^t \in (\mathbb{Z}^d)^\star} \left(\max\left(\mathbf{c}^t\mathbf{x} \,:\, \mathbf{x} \in P\right) \,-\, \min\left(\mathbf{c}^t\mathbf{x} \,:\, \mathbf{x} \in P\right)\right) \;\leq\; c(d) \quad (1.1)$$

We will see that the currently best bound for the width is $\mathcal{O}(d^{4/3} \log^{a} d)$ for some integer $a > 0$,[2] but it is conjectured to be $\mathcal{O}(d)$. There is much research activity around this question. Also, people are interested in precise bounds for certain families of lattice polytopes, and for examples of polytopes of large width.

The Flatness Theorem follows from results in *Geometry of Numbers*, an area of mathematics that connects results from number theory with lattice points and convex sets. The initial result in this area by Minkowski from 1898 shows that any *centrally symmetric* convex body contains a non-zero integral point (it always contains 0 if it is not empty) if the volume is large enough.

A *convex body* is a closed convex set K in $\mathbb{R}^d$, and it is centrally symmetric if $-\mathbf{x} \in K$ for all $\mathbf{x} \in K$. The result of Minkowski states that there is $\mathbf{a} \in (K \cap \mathbb{Z}^d) \setminus \{\mathbf{a}\}$ if $\operatorname{vol} K > 2^d$. Equivalently, $\lambda_1 \cdot K$ contains a non-zero integral point if $\lambda_1 > \frac{2}{\sqrt[d]{\operatorname{vol} K}}$. This is the *first successive minimum* of K. One can further extend this and ask for the smallest scaling factor λ_k such that $\lambda_k \cdot K$ contains k linearly independent integer points. These are the *successive minima* of K, see Figure 1.4. We may extend this to general lattices. Then the first successive minimum of the unit ball is the length of the shortest vector.

This leads to results on the *packing radius* of the lattice, which is the largest radius such that balls with this radius around lattice points at most touch on the boundary, and the *covering radius*, which is the smallest radius such that the corresponding balls around all lattice points cover the space.

[2] Rudelson, "Distances between non-symmetric convex bodies and the MM^*-estimate", Cor. 2.

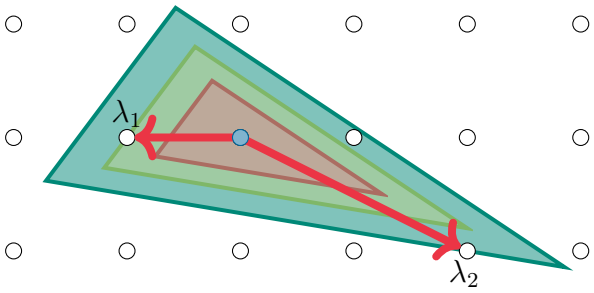


Figure 1.4.: Successive Minima

Given any convex body P the task of actually computing a short lattice direction, *i.e.* a direction which realizes the bound of (1.1), is difficult. If we want to compute such a direction in polynomial time, then we currently only know algorithms that give an approximate solution, and we have to be satisfied with the much weaker bound of $\mathcal{O}(2^d)$ on the width in this direction. Nevertheless, we will discuss an algorithm for this, as it turns out that this is a crucial ingredient to gain more insight into the complexity of integer linear programming.

For this recall first that the problems (LP) and (IP) also differ algorithmically. While we know that the first is algorithmically solvable in polynomial time[3], the second problem is only known to be in NP. It is even NP-complete,[4] so there is no hope to solve it efficiently[5], unless actually $\mathrm{P} = \mathrm{NP}$.

To learn more about what makes (IP) difficult to solve we may look at the contribution of the various input parameters, *e.g.* the number of variables (the dimension), or the size of the constraints. We will see that the problem can be solved in polynomial time if we fix the number of variables, which is the dimension of the polyhedron. This has been discovered by H.W. Lenstra Jr..[6] This is essentially based on the observation, that, given a polyhedron P, we can, in polynomial time, either find a lattice point in P or we obtain a direction in which P is flat in the sense that we can slice P with a polynomial number of parallel hyperplanes containing all lattice points in P. We can then use recursion in the dimension to solve the problem.

This result of Lenstra gave rise to the LLL-algorithm that we have already seen above, and which is now part of the proof of Lenstra's Theorem, but also found many other applications in different branches of mathematics, in particular also for algorithmic results in the geometry of numbers.

The Flatness Theorem considers lattice polyhedra without interior lattice points. It is easy to see that in dimensions $d \geq 2$ there are infinitely many such lattice polyhedra, even, if we, as is sensible to obtain a meaningful result, identify lattice polyhedra that can be mapped into each other with a unimodular transformation.

[3] Although not via the commonly employed simplex algorithm, which is not known to be polynomial. Yet, the ellipsoid method is, and we will reuse some of its ideas later.

[4] Alexander Schrijver, *Theory of linear and integer programming.* Thm. 18.1.

[5] *Efficiently* here means in terms of its theoretical worst case complexity. As (IP) is such an important problem in applications, people developed many algorithms that run efficiently on certain instances that appear in these applications. See, *e.g.*, the book of Nemhauser and Wolsey (Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*)

[6] Hendrik W. Lenstra, "Integer Programming with a fixed number of variables".

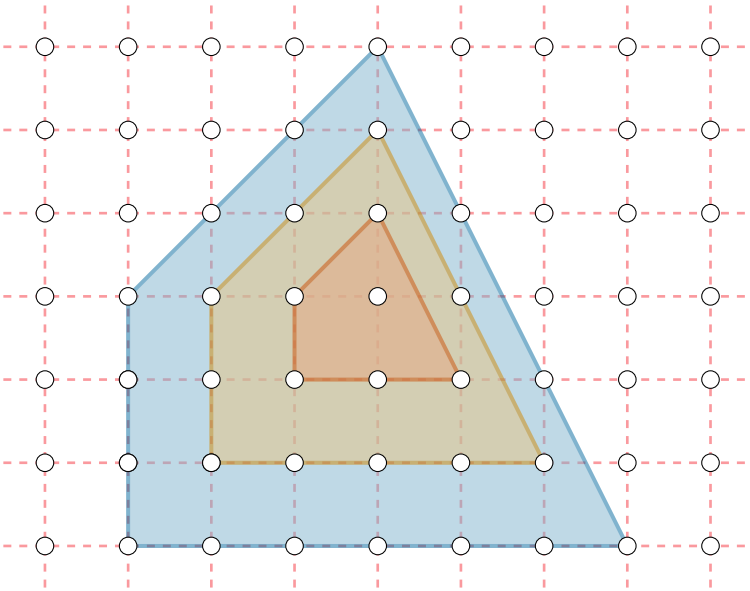


Figure 1.5.: Multiples of a lattice polytope. The polytope P contains 6 integral points, $2P$ contains 16, and $3P$ contains 31. These values are given by the polynomial $\frac{5}{2}k^2 + \frac{5}{2}k + 1$.

The situation drastically changes once we consider the class of lattice polyhedra with a fixed number $k > 0$ of lattice points in the interior. For any given k there are only finitely many such polyhedra (of course, we again have to identify polyhedra that can be mapped into each other via a unimodular transformation).

This follows from volume bounds for the lattice polyhedron that depend on the dimension and the number of interior lattice points. For $d = 2$ this is a result of Scott,[7] and in the general case a result of Pikhurko.[8]

We can take a different approach to determine feasibility of (IP) and ask, whether we can, given a polytope P, *enumerate* or *count* the number of lattice points inside P, *i.e.* whether we can compute

$$P \cap \Lambda \qquad \text{or} \qquad |P \cap \Lambda|\,.$$

It turns out that instead of this question one should consider the function

$$f(k) := |k \cdot P \cap \Lambda| \qquad \text{for} \qquad k \in \mathbb{Z}_{>0}$$

that counts the lattice points in integer dilates of the polytope. The Theorem of Ehrhart shows that this function is indeed given by a polynomial of degree d, the *Ehrhart Polynomial*.

Its generating function

$$F(x) := 1 + \sum_{k \geq 1} f(k) x^k$$

[7] Scott, "On convex lattice polygons".
[8] Pikhurko, "Lattice points in lattice polytopes".

can be represented by a *short* rational function

$$1 + \sum_{k \geq 1} f(k) x^k = \frac{h(x)}{g(x)}$$

that can be computed with the algorithm of Barvinok.[9] Evaluation at 1 gives the number of lattice points inside P. This is an efficient algorithm that can be used in applications. The polynomials $f(x)$ and $h(x)$ encode a lot of information about the polytope, with applications in number theory, algebra, and algebraic geometry. Much of this is still a topic of current research.

Barvinok and Woods[10] have later extended the counting algorithm to integer points in *integer projections* of polyhedra in fixed dimension, *i.e.* to sets of the form

$$\left\{ \mathbf{x} \in \mathbb{Z}^k \ : \ \exists \mathbf{y} \in \mathbb{Z}^{d-k} \text{ s.th. } (\mathbf{x}, \mathbf{y}) \in P \right\} .$$

This is based on results of Kannan on *parametric integer linear programming*, which is solvable in polynomial time in fixed dimension.[11]

The Ehrhart polynomial and its generating function have many interesting properties, with applications not only in optimization, but also in number theory and algebra. For example, we can count lattice points in the relative interior of $k \cdot P$ by evaluating the polynomial of f at $-k$.

Polyhedra with no interior lattice points can be employed as a tool in algorithms to solve (IP). This generalizes the notion of split cuts. The feasible region of the linear relaxation of (IP) is the polyhedron

$$P := \{ \mathbf{x} \ : \ A\mathbf{x} \leq \mathbf{b} \} ,$$

and the general idea of *cuts*is the approximation of the *integer hull* of P, *i.e.* the set

$$P_I := \left\{ \mathbf{z} \ : \ A\mathbf{x} \leq \mathbf{b} \ \mathbf{x} \in \mathbb{Z}^d \right\}$$

with polyhedra Q that satisfy

$$P_I \subseteq Q \subsetneq P$$

using new inequalities $\mathbf{c}^t \mathbf{x} \leq \delta$, called *cuts*,that separate some points in $P \setminus P_I$ from P_I. You can find various ways to generate such cuts in the literature, *e.g.* the *Chvatál-Gomory-Cuts*, which where among the first to be considered.[12]

Given an integer program as in (IP), a classical *split cut* is a cut derived from the maximally lattice free set

$$S := \left\{ \mathbf{x} \ : \ \pi_0 \leq \boldsymbol{\pi}^t \mathbf{x} \leq \pi_0 + 1 \right\}$$

[9] A. Barvinok and Pommersheim, "An algorithmic theory of lattice points in polyhedra".
[10] A. Barvinok and Woods, "Short rational generating functions for lattice point problems."
[11] Ravi Kannan, "Lattice translates of a polytope and the Frobenius problem"; Ravi Kannan, "Test sets for integer programs, ∀∃ sentences".
[12] Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*.

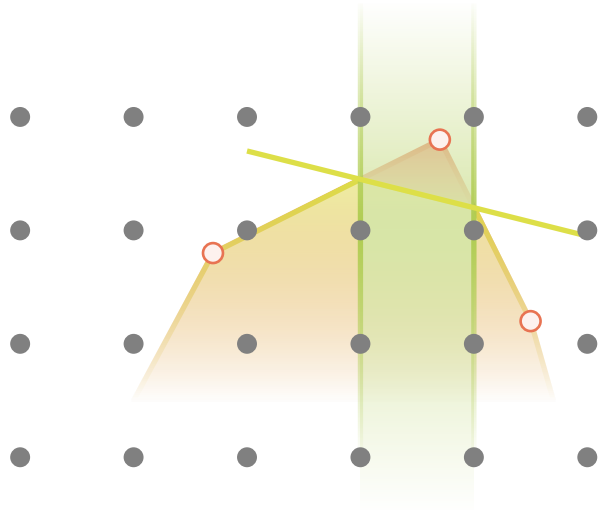


Figure 1.6.: A split cut of width 1

for some $\boldsymbol{\pi} \in \mathbb{Z}^d$ and $\pi_0 \in \mathbb{Z}$. Here, we say that a set S is *lattice free* if the relative interior of S does not contain any lattice points. We obtain an approximation of P_I via

$$Q(P,S) \;:=\; \mathsf{conv}\,(P \setminus S)\;.$$

This is again a polyhedron, and if S contains a vertex of P in its interior, then also $Q(P,S) \subsetneq P$. See Figure 1.6 for an example.

This can be generalized to mixed integer programs with solutions in $P \cap (\mathbb{Z}^p \times \mathbb{R}^{d-p})$ by considering lattice free sets of the form $S \times \mathbb{R}^{d-p}$ for some lattice free set $S \subseteq \mathbb{R}^p$. However, for simplicity, we only consider pure integer programs in the following.

The lattice free set S we have considered is of the form $S = I \times R^{d-1}$ for some interval I of length 1. We can generalize this idea and use sets of the form $S := L \times R^{d-k}$ for a convex set $L \subseteq \mathbb{R}^k$ that is lattice free, *i.e.* that does not contain lattice points in its relative interior. We then consider the relaxation

$$\begin{aligned} Q(P,S) \;&:=\; \mathsf{conv}\,(P \setminus S) \\ &=\; \mathsf{conv}\left((\mathbf{x}_1,\mathbf{x}_2) \in \mathbb{R}^k \times \mathbb{R}^{d-k} \;:\; (\mathbf{x}_1,\mathbf{x}_2) \in P,\; \mathbf{x}_1 \notin \operatorname{int} L\right)\;. \end{aligned}$$

Clearly, $P_I \subseteq Q(P,S) \subseteq P$ and for any lattice free convex sets $S' \subseteq S$ we have

$$Q(P,S) \;\subseteq\; Q(P,S')\,.$$

Hence, we are interested in *maximally lattice free sets*, *i.e.* inclusion maximal convex lattice free sets. Such sets are polyhedra,[13] so we are looking for inclusion maximal lattice free polyhedra.

[13] Lovász, "Geometry of numbers and integer programming".

2. The Geometry of Lattices

The main player in these notes are *polyhedra* and *lattices*, which we can consider as a generalization of the set of integer points in $\mathbb{R}^d$.[1] We will introduce lattices in this chapter while we assume that the reader is familiar with the basics of polyhedral theory.[2] We will briefly repeat some some of this, together with some additional material from convex geometry, in Appendix A.

Although trivial to define, lattices have a surprisingly rich geometric structure and many, even basic, theoretic and algorithmic questions are still unsolved. Here we discuss basic properties, while the more interesting topics will be spread over the following chapters.

2.1. Lattices

We can define lattices in two ways, either as the integral span of a finite set of vectors in $\mathbb{R}^d$, or, equivalently, as discrete additive subgroups of $\mathbb{R}^d$. We start with the latter version and introduce that second and prove equivalence in Theorem 2.12.

Recall that a subset $\Lambda \subseteq \mathbb{R}^d$ is an *additive subgroup* of $\mathbb{R}^d$ if

(1)	$\mathbf{0} \in \Lambda$	
(2)	$\mathbf{x}+\mathbf{y} \in \Lambda$	for any $\mathbf{x}, \mathbf{y} \in \Lambda$
(3)	$-\mathbf{x} \in \Lambda$	for any $\mathbf{x} \in \Lambda$.

A subset $\Lambda \subseteq \mathbb{R}^d$ is *discrete* if for all $\mathbf{x} \in \Lambda$ there is $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon(\mathbf{x}) \cap \Lambda = \{\mathbf{x}\}$, where we define the *ball* of radius ε in the given norm as

$$\mathcal{B}_\varepsilon(\mathbf{x}) := \{\mathbf{y} \in V \ : \ \|\mathbf{x}-\mathbf{y}\| \leq \varepsilon\} .$$

Definition 2.1. A *lattice* in $\mathbb{R}^d$ is a discrete additive subgroup $\Lambda \subseteq \mathbb{R}^d$.

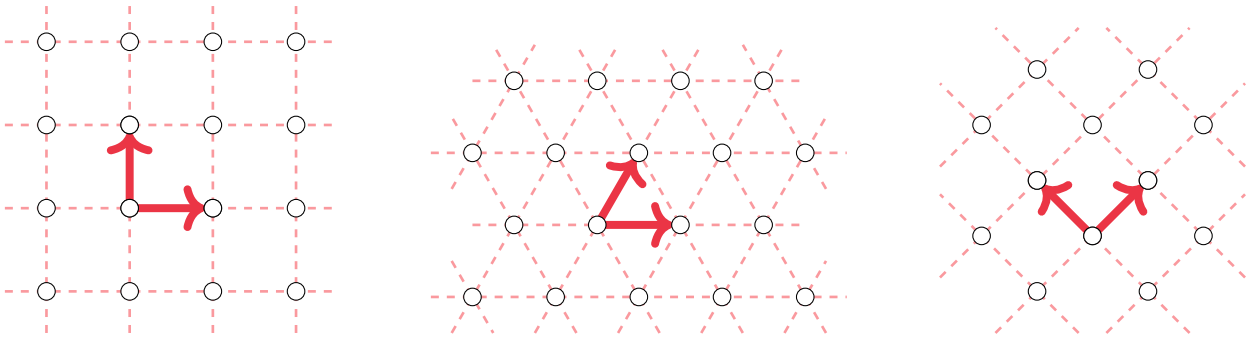
The *rank* of a lattice Λ is the dimension of its linear span, that is,

$$\operatorname{rank} \Lambda \ := \ \dim \operatorname{lin} \Lambda .$$

The lattice has *full rank* if it has rank d, *i.e.* the dimension of its ambient space.

[1] One can define lattices in any finite dimensional vector space with a norm, but $\mathbb{R}^d$ is sufficient for our applications.

[2] *e.g.* from the courses *Einführung in die Optimierung* and *Discrete Optimization*. If needed we may repeat the relevant notions in the exercises.



(a) The lattice $\mathbb{Z}^2$ (b) The lattice A_2 (c) The lattice D_2

Figure 2.1.: Some lattices

You will show in Problem 2.1 that for lattices one can choose the same radius ε for the ball for all lattice points. Any discrete additive subgroup is closed in $\mathbb{R}^d$ in the usual topology induced by the scalar product (Problem 2.2), and the intersection of any bounded subset with the lattice is a finite set (Problem 2.3).

Problem 2.1

Problem 2.2

Problem 2.3

Example 2.2. (i) The set $\mathbb{Z}^d$ of points with integral coordinates in $\mathbb{R}^d$ is a lattice of full rank, the *standard integer lattice*. We can write this as the set of all linear combinations of the d standard unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_d$ with integral coefficients. See Figure 2.1(a) for an example in dimension 2.

We will later see that any lattice has such a generating set, and thus essentially any lattice looks like this integer lattice.

(ii) For any lattice Λ and linear subspace $L \subset \mathbb{R}^d$ the set $\Lambda \cap L$ is a lattice in L.

(iii) The set

$$\Lambda_{2;3} := \left\{ \mathbf{x} \in \mathbb{Z}^2 : x_1 + x_2 \equiv 0 \mod 3 \right\}$$

is a lattice and a subgroup of the lattice $\mathbb{Z}^2$. More generally, any subgroup of a lattice is again a lattice.

(iv) We can identify $\mathbb{R}^d$ with the linear subspace

$$L := \left\{ \mathbf{x} \in \mathbb{R}^{d+1} : \sum_{i=0}^{d} x_i = 0 \right\}.$$

See Figure 2.1(b) for $d = 3$ (note that the lattices lives in a 2-dimensional subspace). We claim that the set

$$\mathrm{A}_d := L \cap \mathbb{Z}^{d+1}$$

is a lattice in L. To check this we first observe that A_d is clearly discrete, as it is a subset of a discrete set. Further, the addition of any two elements in A stays in L, as this is a linear subspace. The same is true for the multiplication by -1. Hence, it is also an additive subgroup.

This lattice is the *root lattice* A_d. We will discuss this again later.

(v) Let D_d be the set

$$\mathrm{D}_d := \left\{ \mathbf{x} \in \mathbb{Z}^d \,:\, \sum_{i=1}^{d} x_i \text{ is even} \right\}. \tag{2.1}$$

Again, this is a discrete set and addition and multiplication by -1 stay inside the set. Hence, it is a lattice, the so called *root lattice* D_d. See Figure 2.1(c) for the case $d = 2$.

For computational purposes the abstract definition via discrete subgroups is not very useful. As for linear spaces, where we express elements in terms of coordinate vectors *w.r.t.* to a chosen basis, we would prefer to have some kind of generating set for a lattice. Our examples suggest that this should be possible. For the lattice $\mathbb{Z}^d$ of integer points we have already seen in its definition, that we can obtain the lattice as the set of linear combinations of the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_d$ with coefficients in $\mathbb{Z}$ (instead of $\mathbb{R}$ when considering this basis as a basis of the linear space $\mathbb{R}^d$). We want to formalize this approach. Let $\mathcal{A} = \{ \mathbf{a}_1, \dots, \mathbf{a}_m \} \subseteq \mathbb{R}^d$ be a finite set of vectors and define the additive subgroup

$$\Lambda_{\mathcal{A}} := \left\{ \sum_{i=1}^{m} \lambda_i \mathbf{a}_i \mid \lambda_i \in \mathbb{Z},\, 1 \le i \le m \right\} = \bigoplus_{i=1}^{m} \mathbb{Z}\mathbf{a}_i$$

We will show that any such set is in fact a lattice if the elements of $\mathcal{A}$ are linearly independent. We make the following definition.

Definition 2.3. Any linearly independent subset $\mathrm{B} \subseteq \mathbb{R}^d$ that generates a lattice Λ, *i.e.* $\Lambda = \Lambda_{\mathrm{B}}$ (as subsets of $\mathbb{R}^d$), is a *lattice basis* (or Λ-*basis*) of Λ.

We will also consider more general sets $\mathcal{A}$, but you will see in Problem 2.4 that we need some restriction on $\mathcal{A}$.

Problem 2.4

Example 2.4. Before we prove that all sets of the form $\Lambda_{\mathcal{A}}$ for a linearly independent set $\mathcal{A}$ are a lattice we want to construct explicit sets for the examples in Example 2.2.

(i) The set

$$\mathrm{B} := \left\{ \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

is a basis of the lattice $\Lambda_{2;3}$ that we have seen in Example 2.2.

(ii) The set $\mathrm{B} := \{ \mathbf{e}_i - \mathbf{e}_{i+1} \,:\, 1 \le i \le d \}$ generates the lattice A_d, and for D_d we may choose the roots of A_{d-1} together with $\mathbf{e}_d + \mathbf{e}_{d-1}$. You will prove this in Problem 2.5.

These sets of vectors are sets of simple roots of the *root systems* of type A_d and D_d. Generally, root systems are finite vector configurations $\Phi \subseteq \mathbb{R}^d$ with the property that

- ▷ their linear span is $\mathbb{R}^d$,

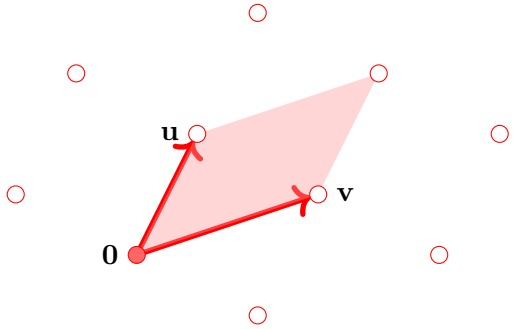


Figure 2.2.: A fundamental parallelepiped Π. Note that the points $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u}+\mathbf{v}$ are not contained in Π.

- ▷ for any $\mathbf{v} \in \Phi$ also $-\mathbf{v} \in \Phi$, and this is the only linear multiple of $\mathbf{v}$ contained in Φ,
- ▷ Φ is closed under reflection at any hyperplane $\{\, \mathbf{x} \,:\, \langle \mathbf{x}, \mathbf{v} \rangle \,=\, 0 \,\}$ for some $\mathbf{v} \in \Phi$, and
- ▷ for any two $\mathbf{u}, \mathbf{v} \in \Phi$ the projection of $\mathbf{u}$ onto the line spanned by $\mathbf{v}$ is an integer or half-integer multiple of $\mathbf{v}$.

A *root lattice* is the lattice spanned by a root system. A root system is *irreducible* if it cannot be decomposed into a direct sum of root systems. There is only a finite set of families of *irreducible root systems*, which give rise to corresponding lattices (You will consider them in Problem 2.6).

Problem 2.5
Problem 2.6

For the proof that any lattice has a basis we introduce the following notion of a fundamental zonotope $\Pi(\mathcal{A})$ of $\mathcal{A}$ via

$$\Pi(\mathcal{A}) \;:=\; \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{v}_i \mid 0 \le \lambda_i < 1 \text{ for } 1 \le i \le k \right\} .$$

The (half-open) zonotope is a (half-open) *parallelepiped* if $\mathcal{A}$ is linearly independent. See Figure 2.2 for an example.

Definition 2.5. Let Λ be a lattice generated by a basis B. The parallelepiped $\Pi(\mathrm{B})$ is the *fundamental parallelepiped* of Λ *w.r.t.* B.

Lemma 2.6. *Let* $\mathrm{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\} \subseteq \mathbb{R}^d$ *be a finite set of linearly independent vectors. Any* $\mathbf{x} \in \operatorname{lin} \Lambda_{\mathrm{B}}$ *has a unique representation* $\mathbf{x} = \mathbf{a} + \mathbf{y}$ *for some* $\mathbf{a} \in \Lambda_{\mathrm{B}}$ *and* $\mathbf{y} \in \Pi(\mathrm{B})$.

Proof. There are unique $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$ such that $\mathbf{x} = \sum_{i=1}^{d} \lambda_i \mathbf{b}_i$. Set

$$\mathbf{a} \;:=\; \sum_{i=1}^{d} \lfloor \lambda_i \rfloor \, \mathbf{b}_i \qquad \text{and} \qquad \mathbf{y} \;:=\; \sum_{i=1}^{d} \{\lambda_i\} \, \mathbf{b}_i \,.$$

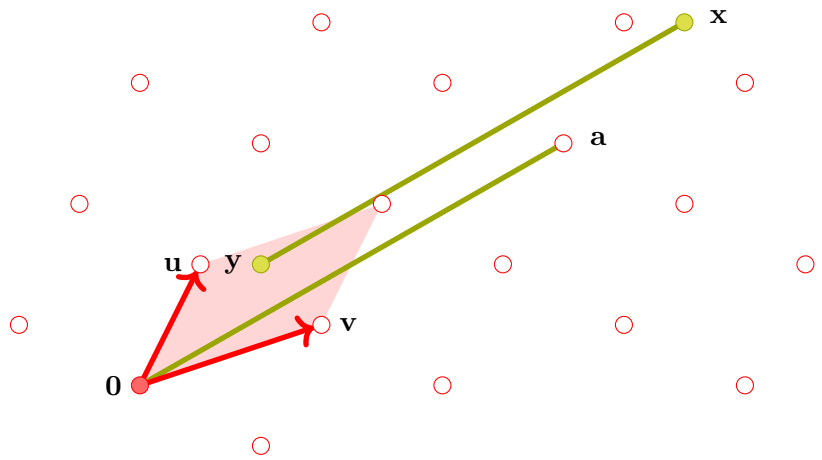


Figure 2.3.: A fundamental parallelepiped Π. Note that the points $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u}+\mathbf{v}$ are not contained in Π.

See also Figure 2.3. Then $\mathbf{y} \in \Pi(\mathrm{B})$, $\mathbf{a} \in \Lambda_\mathrm{B}$, and $\mathbf{x} = \mathbf{a} + \mathbf{y}$.

Now assume that there is a second decomposition $\mathbf{x} = \mathbf{a}' + \mathbf{y}'$ with $\mathbf{a} \neq \mathbf{a}'$ and, consequently, also $\mathbf{y} \neq \mathbf{y}'$. We can write $\mathbf{y}$ and $\mathbf{y}'$ as

$$\mathbf{y} := \sum_{i=1}^{d} \alpha_i \mathbf{b}_i \qquad \text{and} \qquad \mathbf{y}' := \sum_{i=1}^{d} \alpha_i' \mathbf{b}_i$$

for some $0 \leq \alpha_i, \alpha_i' < 1$ and $1 \leq i \leq d$. Then

$$\mathbf{a}' - \mathbf{a} = \mathbf{y} - \mathbf{y}' = \sum_{i=1}^{d} (\alpha_i - \alpha_i') \mathbf{b}_i$$

and $\mathbf{a}' - \mathbf{a} \in \Lambda_\mathrm{B}$ implies that $\alpha_i - \alpha_i' \in \mathbb{Z}$ for $1 \leq i \leq d$.

Now $|\alpha_i - \alpha_i'| < 1$ for all i, so $\alpha_i - \alpha_i' = 0$. This implies that $\mathbf{y} = \mathbf{y}'$ and $\mathbf{a} = \mathbf{a}'$. □

Proposition 2.7. *Let $\mathrm{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\} \subseteq \mathbb{R}^d$ be linearly independent. Then Λ_B is a lattice.*

Proof. Let $\mathbf{z} \in \Pi(\mathrm{B}) \cap \mathbb{R}^d$ be any interior point of $\Pi(\mathrm{B})$. Then there is $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon(\mathbf{z}) \subseteq \Pi(\mathrm{B})$. We claim that $\mathcal{B}_\varepsilon(\mathbf{x}) \cap \Lambda = \{\mathbf{x}\}$ for all $\mathbf{x} \in \Lambda_\mathrm{B}$.

Indeed, if $\mathbf{y} \in \mathcal{B}_\varepsilon(\mathbf{x}) \cap \Lambda_\mathrm{B}$ and $\mathbf{y} \neq \mathbf{x}$, then $\mathbf{x}' := \mathbf{x} - \mathbf{y} \in \Lambda_\mathrm{B} \setminus \{0\}$ and $\mathbf{x}' + \mathbf{z} \in \Pi(\mathrm{B})$. This is a contradiction to Lemma 2.6. □

So any linearly independent set of vectors generates a lattice. While Problem 2.4 shows that this is generally not true for linear dependent sets, we will see later that any finite set of *rational* vectors generates a lattice.

It follows directly from Lemma 2.6 that the parallelepipeds of a lattice with basis B tile the space. The full proof is left to the reader in Problem 2.7.

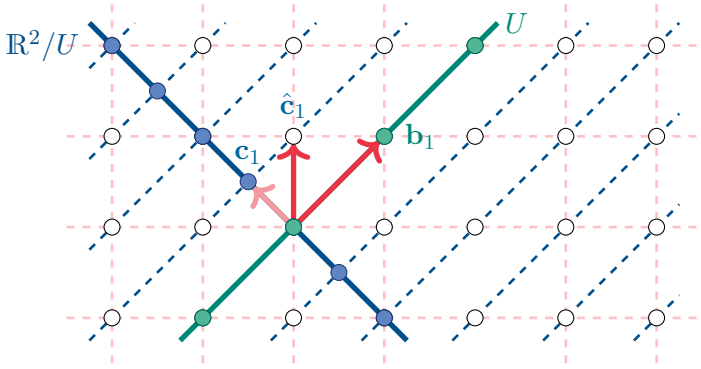


Figure 2.4.: A one-dimensional Λ-rational subspace U with lattice basis $\{\mathbf{b}_1\}$ of $\Lambda \cap U$, the projection $\mathbb{R}^2/U$ with $\{\mathbf{c}_1 + U\}$ as lattice basis of $\pi(\Lambda)$ and the basis obtained by the basis $\{\mathbf{b}_1\}$ of U with the pull-back $\{\hat{\mathbf{c}}_1\}$ of the basis in $\mathbb{R}^2/U$

Corollary 2.8. *Let Λ be a lattice in $\mathbb{R}^d$ with basis B. Then V is the disjoint union of all translates of $\Pi(\mathrm{B})$ by vectors in Λ.* □

Problem 2.7

Thus, sets of the form Λ_B for a basis B of linearly independent vectors are indeed lattices. We now aim for the opposite direction and want to show that any lattice can be obtained from some basis.

We will do this by induction and construct a basis first in a subspace and then extend it to a basis of the full lattice. Although we have already seen that for a k-dimensional subspace U of $\mathbb{R}^d$ the intersection $\Lambda' := \Lambda \cap U$ of a lattice Λ of rank d with U is again a lattice, it is not true that the rank of Λ' must coincide with the dimension of U. A simple example is the integer lattice $\mathbb{Z}^2 \subseteq \mathbb{R}^2$ together with the subspace U spanned by the vector $\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$. Hence, we make the following definition.

Definition 2.9. A subspace $U \subseteq \mathbb{R}^d$ is *Λ-rational* if it is generated by elements of Λ.

The following proposition then shows that this condition is sufficient to make our inductive approach to construct a basis feasible.

Proposition 2.10. *Let $\Lambda \subset \mathbb{R}^d$ be a lattice with a Λ-rational subspace $U \subseteq \mathbb{R}^d$ and the quotient map $\pi\colon \mathbb{R}^d \to \mathbb{R}^d/U$.*

(i) $\pi(\Lambda) \subset \mathbb{R}^d/U$ is a lattice.

(ii) If $\Lambda \cap U$ has a basis $\mathbf{b}_1, \dots, \mathbf{b}_r$, and $\pi(\Lambda)$ has a basis $\mathbf{c}_1, \dots, \mathbf{c}_s$, then any choice of preimages $\hat{\mathbf{c}}_i \in \Lambda$ of the $\mathbf{c}_i$ for $1 \le i \le s$ yields a Λ-basis $\mathbf{b}_1, \dots, \mathbf{b}_r, \hat{\mathbf{c}}_1, \dots, \hat{\mathbf{c}}_s$.

In the situation of the proposition, one often writes Λ/U for $\pi(\Lambda)$.

Proof. (i) $\pi(\Lambda)$ is the image of a group under a homomorphism. Hence, it is a subgroup of $\mathbb{R}^d/U$. The hard part of the proposition is to prove that $\pi(\Lambda)$ is discrete in $\mathbb{R}^d/U$.

The space U is Λ-rational. So we can choose a vector space basis

$$\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \subseteq \Lambda \cap U$$

of U. We can extend this basis to a vector space basis $\mathrm{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_d\} \subset \Lambda$ of $\operatorname{lin}\Lambda$. These bases yield maximum norms

$$\left\| \sum_{i=1}^{d} \lambda_i \mathbf{v}_i \right\|_{\Lambda} := \max\left(\{|\lambda_i| \,:\, i = 1, \dots, d\}\right)$$

on $\operatorname{lin}\Lambda$ and

$$\left\| \left(\sum_{i=1}^{d} \lambda_i \mathbf{v}_i \right) + U \right\|_{\Lambda/U} := \max\left(\{|\lambda_i| \,:\, i = r+1, \dots, d\}\right)$$

on $\operatorname{lin}\Lambda/U$. Denote the unit ball of $\operatorname{lin}\Lambda$ by W. By Problem 2.3, the set $W \cap \Lambda$ is finite. Set

$$\varepsilon := \min\left(\{1\} \cup \{\|\mathbf{v} + U\|_{\Lambda/U} \,:\, \mathbf{v} \in W \cap \Lambda \setminus U\}\right).$$

This minimum over a finite set of positive numbers is positive. Now suppose

$$\mathbf{v} = \sum_{i=1}^{d} \lambda_i \mathbf{v}_i \in \Lambda$$

with $\|\mathbf{v} + U\|_{\Lambda/U} < \varepsilon$. Then

$$\mathbf{v}' := \sum_{i=1}^{r} (\lambda_i - \lfloor \lambda_i \rfloor)\mathbf{v}_i + \sum_{i=r+1}^{d} \lfloor \lambda_i \rfloor \mathbf{v}_i \in \Lambda$$

represents the same coset: $\mathbf{v} + U = \mathbf{v}' + U$, and $\mathbf{v}' \in W \cap \Lambda$. We conclude $\mathbf{v}' \in U$ and thus $\mathbf{v}' + U = \mathbf{0} \in \mathbb{R}^d/U$.

(ii) Let $\mathbf{b}_1, \dots, \mathbf{b}_r, \hat{\mathbf{c}}_1, \dots, \hat{\mathbf{c}}_s$ be as in the proposition, and let $\mathbf{v} \in \Lambda$. Because the $\mathbf{c}_j$ form a lattice basis of $\pi(\Lambda)$, there are integers $\lambda_1, \dots, \lambda_s$ so that

$$\pi(\mathbf{v}) = \sum_{j=1}^{s} \lambda_j \mathbf{c}_j .$$

Thus,

$$\mathbf{v} - \sum_{j=1}^{s} \lambda_j \hat{\mathbf{c}}_j \in \ker \pi = U .$$

Because the $\mathbf{b}_i$ form a lattice basis of $\Lambda \cap U$, there are integers $\mu_1, \dots, \mu_r$ so that

$$\mathbf{v} - \sum_{j=1}^{s} \lambda_j \hat{\mathbf{c}}_j \;=\; \sum_{i=1}^{r} \mu_i \mathbf{b}_i \,.$$

So $\mathbf{b}_1, \dots, \mathbf{b}_r, \hat{\mathbf{c}}_1, \dots, \hat{\mathbf{c}}_s$ generate Λ. They must be linearly independent for dimension reasons. □

Definition 2.11. A set $\mathbf{b}_1, \dots, \mathbf{b}_k \in \Lambda$ of linearly independent lattice vectors is *primitive* if their integral span coincides with the intersection of of the lattice with their linear span, *i.e.*

$$\operatorname{lin}\{\mathbf{b}_1, \dots, \mathbf{b}_k\} \cap \Lambda \;=\; \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{b}_i \;:\; \lambda_i \in \mathbb{Z} \right\}.$$

In particular, a non-zero lattice vector $\mathbf{v} \in \Lambda$ is primitive if it is not a positive multiple of another lattice vector, *i.e.* $\mathsf{conv}(\mathbf{0}, \mathbf{v}) \cap \Lambda = \{\mathbf{0}, \mathbf{v}\}$.

Any lattice basis is also a linear basis of its linear span, so it follows from linear algebra, that any subset of a lattice basis is a primitive set.

Theorem 2.12. *Every lattice has a basis.*

Proof. We proceed by induction on $r := \operatorname{rank} \Lambda$. For $r = 0$, the empty set is a basis for Λ. For $r = 1$, a primitive vector yields a basis.

Assume $r \geq 2$. Let $\mathbf{b} \in \Lambda$ be primitive, and set $U := \operatorname{lin}\{\mathbf{b}\}$. Then $\{\mathbf{b}\}$ is a basis for $U \cap \Lambda$, and Λ/U is a lattice by the first statement of Proposition 2.10. Because $\operatorname{rank} \Lambda/U = r-1$, it has a basis by induction. By the second statement of Proposition 2.10, we can lift to a basis of Λ. □

Problem 2.8

If B is a basis of a lattice, then no further lattice points except 0 and B can be inside the simplex $\mathsf{conv}(\{0\} \cup \mathrm{B})$, which is a subset of $\Pi(\mathrm{B}) \cup \mathrm{B}$. However, with Problem 2.9 you will prove that this is not sufficient to characterize a basis, we need to consider the full fundamental parallelepiped.

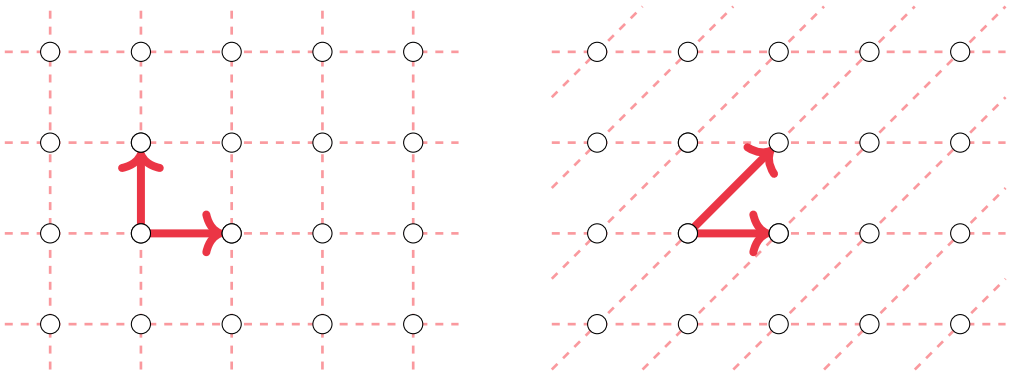
Problem 2.9
Problem 2.10
Problem 2.11

The proof of Proposition 2.10 shows that we can extend any primitive set of lattice vectors to a lattice basis of the the whole lattice. The proof of the next proposition is left as an exercise. Here we set $\gcd(0, 0) := 0$.

Proposition 2.13. *Let $\mathbf{b}_1, \dots, \mathbf{b}_d \subseteq \Lambda$ be a lattice basis of Λ, $\mathbf{a} := \sum_{i=1}^{d} \lambda_i \mathbf{b}_i$ for $\lambda_i \in \mathbb{Z}$, and let $1 \leq j \leq d$. Then the set $\mathcal{A} := \{\mathbf{b}_1 \dots, \mathbf{b}_{j-1}, \mathbf{a}\}$ is primitive if and only if $\gcd(\lambda_j, \dots, \lambda_d) = 1$.* □

Proof. See Problem 2.12. □

Problem 2.12



(a) The lattice $\mathbb{Z}^2$.

(b) The lattice sheared with $\mathbf{e}_1 \mapsto \mathbf{e}_1$ and $\mathbf{e}_2 \mapsto \mathbf{e}_1 + \mathbf{e}_2$.

Figure 2.5.: Shearings are unimodular transformations

2.2. The Hermite normal form

Definition 2.14. Let Λ and Λ' be lattices. A linear map

$$\mathsf{T} : \lin \Lambda \longrightarrow \lin \Lambda'$$

inducing a bijection $\Lambda \to \Lambda'$ is called *unimodular* or a *lattice transformation*. T is a *lattice isomorphism* if $\Lambda = \Lambda'$.

See Figure 2.5 for an example of a unimodular transformation.

Lemma 2.15. *Let* B *and* B' *be bases of lattices* Λ *and* Λ' *of full rank respectively.*

A linear map $\mathsf{T}\colon \lin \Lambda \to \lin \Lambda'$ *is unimodular if and only if the matrix representation* A *of* T *with respect to the bases* B *and* B' *is integral and satisfies* $|\det A| = 1$.

Proof. Let d be the rank of Λ and Λ', and $A = (a_{ij})_{1 \le i,j \le d}$. If A has only integral entries, then $\mathsf{T}(\Lambda) \subseteq \Lambda'$. Conversely, let $\mathbf{b}_j \in \mathrm{B}$ be the j-th basis vector. Then

$$\mathbf{x} := T(\mathbf{b}_j) = \sum_{i=1}^{d} a_{ij}\mathbf{b}'_i \in \Lambda'.$$

As the vectors in B' are linearly independent the representation of $\mathbf{x}$ in this basis is unique, and thus all a_{ij}, $1 \le i \le d$ must be integral (here we have used that the lattices have full rank).

Further, if T is unimodular, then the inverse transformation exists, and its matrix A^{-1} also has integral entries. Thus, $\det A$ and $\det A^{-1}$ are integers with product 1.

Conversely, if A is integral with $|\det A| = 1$, then, by Cramer's rule A^{-1} exists and is integral. □

Corollary 2.16. *An integral matrix $A \in \mathbb{Z}^{d\times d}$ is the matrix representation of a unimodular transformation of a lattice if and only if $|\det A| = 1$.* □

The set of matrices corresponding to unimodular transformations is denoted by $\mathrm{Gl}(d, \mathbb{Z})$.

Corollary 2.17. *Let Λ be a lattice with basis $\mathbf{b}_1, \ldots, \mathbf{b}_d \in \Lambda$. Then $\mathbf{c}_1, \ldots, \mathbf{c}_d \in \Lambda$ is another basis of Λ if and only if there is a unimodular transformation $\mathsf{T} :\in \Lambda \to \operatorname{lin}\Lambda$ such that $\mathsf{T}(\mathbf{b}_i) = \mathbf{c}_i$ for $1 \le i \le d$.* □

Problem 2.13
Problem 2.14

Let $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ be a generating set of a lattice Λ in $\mathbb{R}^d$. We do not require that $\mathcal{A}$ is a basis. Recall, however, that not all sets of vectors generate a lattice (Problem 2.4).

Given any basis C in V (not necessarily a lattice basis), we can represent the vectors of $\mathcal{A}$ in this basis. Writing the vectors in $\mathcal{A}$ as coefficient vectors in C we obtain an $(d \times m)$-matrix A. We assume in the following that $A \in \mathbb{Q}^{d\times m}$, which also ensures that $\mathcal{A}$ spans a lattice.

We want to find linear combinations of the vectors in $\mathcal{A}$ that yield a particularly nice basis of Λ (in the representation as coordinate vectors of C).

For this, we introduce the *Hermite normal form* H of a matrix A and show that we can find a unimodular transformation $U \in \mathrm{Gl}(m, \mathbb{Z})$ that maps coefficient vectors *w.r.t.* A into coefficient vectors *w.r.t.* H. Further, we will see that the Hermite normal form is unique (while the transformation U is unique only if $\mathcal{A}$ was already a basis).

Definition 2.18. Let $A = (a_{ij}) \in \mathbb{Q}^{d\times m}$ with $m \ge d$ be of full row-rank. The matrix A is in *Hermite normal form* if

- $a_{ij} = 0$ for $j > i$ and
- $a_{jj} > a_{ij} \ge 0$ for $i > j$.

So a matrix in Hermite normal form is an lower triangular matrix, and the largest entry in each row is on the diagonal.

We can view the columns of U as coefficients of linear combinations on the columns of A that produce the new basis vectors (and possibly some representations of 0, if A has more than d columns). Alternatively, if $\lambda \in \mathbb{Z}^m$ is the vector of coefficients of a point $\mathbf{x} \in \Lambda$ *w.r.t.* A, then $U\lambda$ give the coefficients for $\mathbf{x}$ *w.r.t.* the columns of H.

Example 2.19. A matrix in Hermite normal form.

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 1 & 4 & 0 \end{bmatrix}$$

Depending on the context we sometimes use the *transposed* matrix, i.e. we claim that a matrix is in Hermite normal form if it has at least as many rows as columns, it is upper

triangular, and the largest entry in each column is on the diagonal (and if the matrix is square we can also consider *lower* triangular matrices).

Theorem 2.20. *Let $A \in \mathbb{Q}^{d\times m}$ of full row-rank. Then there is a unimodular matrix $U \in \mathbb{Z}^{m\times m}$ such that AU is in Hermite normal form.*

The matrix H is unique.

Proof. This was proved in the class *Discrete Optimization*. Alternatively you can find a proof Schrijver's book.[3]

You can prove this yourself with Problem 2.15. □

Problem 2.15

Example 2.21. The matrix of Example 2.19 is the Hermite normal form of

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 1 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 10 & -5 & 0 & 15 \\ 2 & -1 & -2 & 1 \\ 10 & -7 & -1 & 12 \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 & -1 & 2 \\ 0 & 1 & -2 & 1 \\ -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{bmatrix}$$

The transformation $U \in \mathbb{Z}^{m\times m}$ is unimodular, so $U^{-1} \in \mathbb{Z}^{m\times m}$ is also integral. So any integral linear combination $\mathbf{x} := A\lambda$ of the columns of A for some $\lambda \in \mathbb{Z}^m$ corresponds to an integral linear combination $\mathbf{x} = H\mu$ for integral coefficients $\mu := U^{-1}\lambda$ and vice versa. This implies that the columns of A and H span the same lattice. The non-zero columns of H are linearly independent, as H is a lower triangular matrix. Thus, if we delete the zero columns from H we obtain a basis of the lattice spanned by the columns of H.

If the columns of A are linearly independent, then we obtain a new basis of the same lattice. The linear map between the two lattices as in Corollary 2.17 is given in coordinates by the matrix U^{-1}.

Example 2.22.

$$A := \begin{bmatrix} 10 & 5 & 15 \\ 6 & 3 & 7 \\ 12 & 8 & 15 \end{bmatrix} \quad \text{and} \quad H := \begin{bmatrix} 5 & 0 & 0 \\ 1 & 2 & 0 \\ 3 & 1 & 4 \end{bmatrix}$$

H is the Hermite normal form of A with unimodular transformation

$$U := \begin{bmatrix} -2 & 2 & -1 \\ 0 & -1 & 2 \\ 1 & -1 & 0 \end{bmatrix} \quad \text{and} \quad U^{-1} := \begin{bmatrix} 2 & 1 & 3 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

[3] Alexander Schrijver, *Theory of linear and integer programming.* Thm. 4.1 and 4.2.

The columns of A and H span the same lattice, and the coefficient vectors

$$\lambda := \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \qquad \textit{w.r.t.} \text{ the basis given by } A \text{ and}$$

$$\mu := U^{-1}\lambda = \begin{bmatrix} 9 \\ 6 \\ 2 \end{bmatrix} \qquad \textit{w.r.t.} \text{ the basis given by } H$$

both correspond to the lattice vector

$$\mathbf{x} := \begin{bmatrix} 45 \\ 21 \\ 41 \end{bmatrix}.$$

We want to use the Hermite normal form later in algorithms, so we want to compute it in polynomial time. The next proposition shows that this is indeed possible. Bounding the number of steps with a polynomial in the input size relies on the fact that we can run the *extended Euclidean algorithm* on two integers a and b in time $\mathcal{O}(\log(a) \cdot \log(b))$. This algorithm computes the greatest common divisor g of a and b together with integers x and y that linearly combine g in a and b, *i.e.*

$$g = \gcd(a, b) = x \cdot a + y \cdot b$$

via a succession of divisions with remainder. A bound on the size of the entries of H follows from the observation that the product $h_{11} \cdot h_{22} \cdots h_{dd}$ of the diagonal entries of H is the greatest common divisor D of $(d \times d)$-subdeterminants of A. Finally, we can bound the size of all intermediate matrices with the observation, that adding integer multiples of D to entries of the matrix does not change the final result. The full proof is left as Problem 2.16. It can also be found in the book of Schrijver.[4]

Proposition 2.23. *The Hermite normal form of a rational matrix A can be computed in polynomial time in the size of the input matrix A. In particular, the size of H is polynomially bounded in the size of A.* □

Problem 2.16

We can interpret the fact that we can add multiples of D to entries of intermediate results also geometrically. We will explain this in the next section, once we have introduced sublattices.

Problem 2.17

Problem 2.17 shows a nice application of Hermite normal forms to linear Diophantine equations.

While the Hermite normal form certainly is of high theoretical *and* computational value in the theory of lattices, the Hermite normal form can have very bad geometric properties. For example, there are bases $\mathbf{b}_1, \ldots, \mathbf{b}_d$ of a lattice, such that $\mathbf{b}_i \in \{0, \pm 1\}^d$

[4] Alexander Schrijver, *Theory of linear and integer programming.* Sec. 5.2 and 5.3.

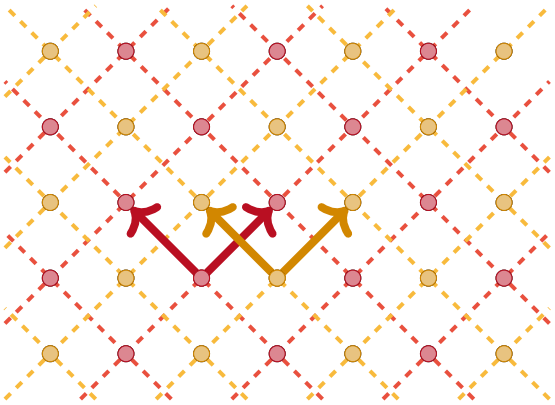


Figure 2.6.: The root lattice D_2 is a sublattice of index 2 in $\mathbb{Z}^2$.

for $1 \leq i \leq d$, but *all* basis vectors in the Hermite normal form $H = (\mathbf{h}_i)_{1 \leq i \leq d}$ have length

$$\|\mathbf{h}_i\| \;\geq\; 2^{\Omega(d)} \;\gg\; \sqrt{d} \;\geq\; \|vb_j\|$$

for all $1 \leq i, j \leq d$. You will construct such a basis in

2.3. Sublattices

Implicitly we have met already *sublattices*, which are subsets of lattices respecting the group structure, when we considered lattices Λ generated by a basis of integral vectors. Such vectors are contained in the integer lattice $\mathbb{Z}^d$. With the following definition we want to formalize this and explore relations between a lattice and its sublattices.

Definition 2.24. Let $\Lambda \subset \mathbb{R}^d$ be a lattice. Any lattice $\Gamma \subseteq \Lambda$ is a *sublattice* of Λ.

Sets of the form $\mathbf{a} + \Gamma := \{\mathbf{a} + \mathbf{x} \mid \mathbf{x} \in \Gamma\}$ for some $\mathbf{a} \in \Lambda$ are the *cosets* of Γ in Λ, and the set of all cosets is Λ/Γ.

The size $[\Gamma : \Lambda] := |\Lambda/\Gamma|$ is the *index* of Γ in Λ.

See Figure 2.6 for an example of a sublattice of index 2 in $\mathbb{Z}^2$.

Definition 2.25. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice of full rank with basis B. Then

$$\det \Lambda \;:=\; |\det B|$$

is the *determinant* of Λ.

This definition assumes that we have chosen a basis of $\mathbb{R}^d$ and written the basis of the lattice in these coordinates to obtain a matrix of column vectors. We choose the usual basis of unit vectors in $\mathbb{R}^d$ in all of the following. With another choice some results

below may need an additional factor $\det T$ for a transformation mapping the standard basis into the chosen one.

Problem 2.19
Problem 2.20

Observe, that the fundamental parallelepiped of Λ depends on the chosen lattice basis. However, by Lemma 2.15 and Corollary 2.17 the determinant is independent of the particular choice. The determinant $\det \Lambda$ also coincides with the usual Euclidean *volume* of the fundamental parallelepiped. Hence, also the volume of $\Pi(\mathrm{B})$ is also independent of the chosen basis B.

Problem 2.21

The following proposition connects the index and the determinant with the number of lattice points in the fundamental parallelepiped of the lattice in some basis.

Proposition 2.26. *Let $\Lambda \subseteq \mathbb{Z}^d$ be a sublattice of rank d, and let $\mathrm{B} := \{\mathbf{b}_1, \ldots, \mathbf{b}_d\}$ be a basis of Λ with fundamental parallelepiped $\Pi(\mathrm{B})$. Then*

$$|\mathbb{Z}^d/\Lambda| \;=\; |\Pi(\mathrm{B}) \cap \mathbb{Z}^d| \;=\; \det \Lambda\,.$$

See Figure 2.7 for an illustration.

Proof. Let $\mathbf{v}, \mathbf{w} \in \Pi(\mathrm{B}) \cap \mathbb{Z}^d$ with $\mathbf{v} \neq \mathbf{w}$. Then $\mathbf{v} \notin \mathbf{w} + \Lambda$ by Lemma 2.6, so cosets for different points in $\Pi(\mathrm{B}) \cap \mathbb{Z}^d$ are disjoint, and we have at least $|\Pi(\mathrm{B}) \cap \mathbb{Z}^d|$ cosets. On the other hand, the union of the cosets for points in $\Pi(\mathrm{B}) \cap \mathbb{Z}^d$ covers $\mathbb{Z}^d$. This implies the first equality.

For the second we note that the k-th scaling of $\Pi(\mathrm{B})$ is the disjoint union of translates of the fundamental parallelepiped $\Pi(\mathrm{B})$,

$$k \cdot \Pi(\mathrm{B}) \;=\; \dot{\bigcup_{\substack{0 \le m_i < k \\ 1 \le i \le d}}} (m_1\mathbf{b}_1 + \cdots + m_d\mathbf{b}_d) + \Pi(\mathrm{B})\,.$$

This follows with the same arguments as for Corollary 2.8. Further, we clearly have

$$|\,(\mathbf{v} + \Pi(\mathrm{B})) \cap \mathbb{Z}^d| \;=\; |\Pi(\mathrm{B}) \cap \mathbb{Z}^d|$$

for any $\mathbf{v} \in \mathbb{Z}^d$. Hence,

$$|\,(k \cdot \Pi(\mathrm{B})) \cap \mathbb{Z}^d| \;=\; k^d |\Pi(\mathrm{B}) \cap \mathbb{Z}^d|$$

We compute the volume of the fundamental parallelepiped by covering it with a collection of small cubes as in Proposition A.4. This implies

$$\begin{aligned}
\mathrm{vol}(\Pi(\mathrm{B})) \;&:=\; \lim_{k\to\infty} \frac{1}{k^d} |\Pi(\mathrm{B}) \cap \left(\tfrac{1}{k}\mathbb{Z}\right)^d| \\
&= \lim_{k\to\infty} \frac{1}{k^d} |k \cdot \Pi(\mathrm{B}) \cap \mathbb{Z}^d| \\
&= \lim_{k\to\infty} \frac{1}{k^d} k^d |\Pi(\mathrm{B}) \cap \mathbb{Z}^d| \;= |\Pi(\mathrm{B}) \cap \mathbb{Z}^d|\,.
\end{aligned}$$

Finally, by definition, $\mathrm{vol}\ \Pi(\mathrm{B}) \;=\; \det(\mathrm{B}) \;=\; \det \Lambda$. This proves the result. □

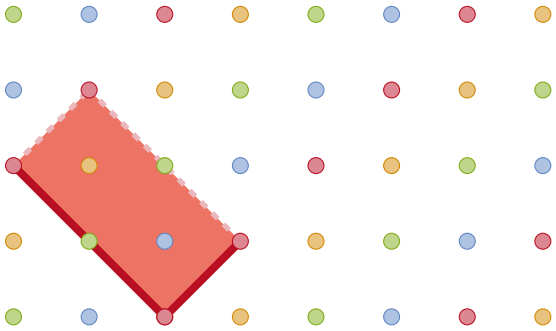


Figure 2.7.: The lattice spanned by $\left[\begin{smallmatrix}-2\\2\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ has index 4 in Z^2, determinant 4, 4 points in the fundamental parallelepiped and 4 cosets (drawn in different colors). Each coset has a unique representative in the fundamental parallelepiped.

More generally we can prove the following with the same arguments as in the proof of Proposition 2.26.

Corollary 2.27. *Let $\Gamma \subseteq \Lambda$ be a sublattice of rank d in a lattice Λ, and let the columns of B be a basis of Γ with fundamental parallelepiped $\Pi(\mathrm{B})$. Then*

$$|\Lambda/\Gamma| \;=\; |\Pi(\mathrm{B}) \cap \Lambda| \;=\; \frac{\det \Gamma}{\det\, \Lambda}\,.$$

Proof. Let B_0 be a basis of Λ. We can apply the transformation T given in matrix form by B_0^{-1} to both lattices. Then $T(\Gamma)$ is still a sublattice with the same index, and $\mathbf{x} \in \Pi(\mathrm{B}) \cap \Lambda$ if and only if $T\mathbf{x} \in \Pi(T\,\mathrm{B}) \cap T\Lambda$. Further,

$$\frac{\det\, T\Gamma}{\det\, T\Lambda} \;=\; \frac{\det\, T\,\mathrm{B}}{\det\, T\,\mathrm{B}_0} \;=\; \frac{\det\, T \cdot \det\, \mathrm{B}}{\det\, T \cdot \det\, \mathrm{B}_0} \;=\; \frac{\det\, \mathrm{B}}{\det\, \mathrm{B}_0}.$$

T maps Λ into the standard lattice, so the claim follows from the previous Proposition 2.26. □

A useful fact about a sublattice Γ of a lattice Λ is the observation that a suitable scaling of the Λ is in turn a sublattice of Γ. You will prove the following fact in Problem 2.22. In the next section we will see that we can even find bases of both lattices in such a way that the i-th basis vector of Γ is an integral multiple of the i-th basis vector of Λ. This much stronger observation will of course also imply the next proposition.

Problem 2.22

Proposition 2.28. *Let Γ be a sublattice of Λ of index $D := |\Gamma/\Lambda|$. Then*

$$D\Lambda \;\subseteq\; \Gamma \;\subseteq\; \Lambda\,,$$

i.e. $D\Lambda$ is a sublattice of Γ.

This gives a *geometric* interpretation of the fact used in the previous section in the proof that the entries of all intermediate matrices in the computation of the Hermite normal

form can be bounded by the greatest common divisor D of all maximal subdeterminants of A. Namely, D is the index of the lattice Λ_A spanned by the columns of A as a sublattice of $\mathbb{Z}^d$, and so $D\mathbb{Z}^d$ is a sublattice of Λ. Hence, adding the columns $D\mathbf{e}_i$ for $1 \leq i \leq d$ of A does not change the lattice Λ_A, and we can use these lattice generators to reduce the entries of intermediate results.

2.4. The Smith normal form

While the Hermite normal form is computationally important, the geometrically more important canonical form is a variation of this, the *Smith normal form*.

Theorem 2.29. *Let $\Gamma \subseteq \Lambda$ be lattices with $\lin \Lambda = \lin \Gamma$.*

Then there is a basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ of Λ and integers $k_1, \ldots, k_r \in \mathbb{Z}_{>0}$ with $k_i | k_{i+1}$ for $1 \leq i \leq d-1$ such that $k_1\mathbf{b}_1, \ldots, k_r\mathbf{b}_r$ is a basis of Γ.

Proof. We proceed by induction on $r := \operatorname{rank}\Lambda = \operatorname{rank}\Gamma$. For $r = 1$, a Λ-primitive vector has a positive integral multiple which is Γ-primitive.

Assume $r \geq 2$. By assumption, Λ and Γ span the same linear spaces, so for every $\mathbf{v} \in \Lambda$ there is a positive integer k so that $k\mathbf{v} \in \Gamma$. Among all Λ-primitive vectors we choose $\mathbf{b}_1 \in \Lambda$ and $k_1 \in \mathbb{Z}_{>0}$ with $k_1\mathbf{b}_1 \in \Gamma$ so that k_1 is minimal with this property.

Set $U := \lin \mathbf{b}_1$. Then $\mathbf{b}_1$ is a basis for $U \cap \Lambda$, and $k_1\mathbf{b}_1$ is a basis for $U \cap \Gamma$. By Proposition 2.10, $\Gamma/U \subseteq \Lambda/U$ are lattices of rank $r-1$. By induction, there is a basis $\overline{\mathbf{b}}_2, \ldots, \overline{\mathbf{b}}_r$ of Λ/U together with positive integers $k_2, \ldots, k_r$ so that $k_2\overline{\mathbf{b}}_2, \ldots, k_r\overline{\mathbf{b}}_r$ is a basis for Γ/U, and $k_i | k_{i+1}$ for $2 \leq i \leq r-1$.

Let $\mathbf{b}_i \in \Lambda$ be representatives of $\overline{\mathbf{b}}_i$ and $\mathbf{c}_i \in \Gamma$ of $k_i\overline{\mathbf{b}}_i$ for $i = 2, \ldots, r$. By Proposition 2.10,

$$\begin{aligned} \mathbf{b}_1, \ldots, \mathbf{b}_r & \qquad \text{is a basis for } \Lambda, \text{ and} \\ k_1\mathbf{b}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r & \qquad \text{is a basis for } \Gamma\,. \end{aligned}$$

By adding a suitable multiple of $k_1\mathbf{b}_1 \in \Gamma$ to the $\mathbf{c}_i$ we may assume that

$$\mathbf{c}_i \;=\; k_i\mathbf{b}_i + l_i\mathbf{b}_1$$

for $0 \leq l_i < k_1$ and for all $i = 2, \ldots, r$. We can write $\mathbf{c}_i$ similarly as a positive integral multiple of some Λ-primitive vector in the form

$$\mathbf{c}_i \;=\; m_i\mathbf{a}_i\,.$$

The two expressions for $\mathbf{c}_i$ together imply either $l_i = 0$ or $m_i \leq l_i < k_r$. However, the latter is in contradiction to the minimality of k_r, so $l_i = 0$.

Now assume that there is i such that $k_1 \not| k_i$, and let g be the greatest common divisor

with factors p_1, p_i such that $k_1 = gp_1$ and $k_i = gp_i$. Then

$$\mathbf{c} \;=\; \frac{1}{g}(k_1\mathbf{b}_1 \;+\; k_i\mathbf{b}_i) \;=\; p_1\mathbf{b}_1 \;+\; p_i\mathbf{b}_i$$

is a primitive lattice vector in Λ and $g\mathbf{c} \in \Gamma$. But $0 < g < k_1$, which contradicts the choice of k_1. □

The theorem can also be proved using coordinates in a similar way as for the Hermite normal form. You may try this with Problem 2.23. We give the resulting matrix form with the following theorem.

Theorem 2.30 (Smith normal form)**.** *Let $A \in \mathbb{Z}^{d\times m}$ be a matrix of full row rank. Then there are unimodular matrices $L \in \mathbb{Z}^{d\times d}$ and $R \in \mathbb{Z}^{m\times m}$ such that $S = (s_{ij})_{1\le i\le d, 1\le j\le m} := LAR$ satisfies*

(i) $s_{ij} = 0$ *for* $i \neq j$,

(ii) $s_{ii} > 0$ *for* $1 \le i \le d$, *and*

(iii) $s_{i-1,i-1}$ *divides* s_{ii} *for* $2 \le i \le d$.

The matrix S is unique, the companion matrices L and R are not.

Problem 2.23

The last statement about the non-uniqueness of L and R follows from the observation that there are unimodular matrices that commute with S. When actually computing smith normal forms with their companions, this fact can be used for an attempt to keep entries in L and R small. This is in fact crucial if we want to compute the Smith normal form in polynomial time. The proof of this is essentially the same as for the Hermite normal form.

The standard form of an integral matrix obtained with the Smith normal form implies Theorem 2.29, as you will prove in Problem 2.24.

Problem 2.24

2.5. The Dual

Definition 2.31. Let $\Lambda \subset \mathbb{R}^d$ be a lattice of full rank. The set

$$\Lambda^\star \;:=\; \{\alpha \in (\mathbb{R}^d)^\star \mid \alpha(\mathbf{a}) \in \mathbb{Z} \text{ for all } \mathbf{a} \in \Lambda\}$$

is the *dual lattice* to Λ.

In the usual identification of $(\mathbb{R}^d)^\star$ with $\mathbb{R}^d$ via the scalar product we can write

$$\alpha(\mathbf{a}) \;=\; \langle\, \alpha, \mathbf{a} \,\rangle\,.$$

Given a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ of Λ we can define a corresponding dual basis $\alpha_1, \ldots, \alpha_d$ via

$$\alpha_i(\mathbf{b}_j) \;=\; \begin{cases} 1 & \text{if } i = j, \text{ and} \\ 0 & \text{otherwise}\,. \end{cases}$$

Problem 2.25

This is a linear basis of the dual vector space $(\mathbb{R}^d)^\star$ by Problem 2.25.

These vectors also span $\Lambda^\star$ as a lattice, which you prove in Problem 2.26. Hence, the dual lattice is indeed a lattice. Further, dualizing twice gives us back the original lattice, $\Lambda^{\star\star} = \Lambda$, as $\mathbf{b}_1, \dots, \mathbf{b}_d$ is a dual basis to $\alpha_1, \dots, \alpha_d$.

Problem 2.26

The following proposition shows that determinant of the dual lattice $\Lambda^\star$ is determined by that of Λ.

Proposition 2.32. $\det(\Lambda)\det(\Lambda^\star) = 1$.

Proof. See Problem 2.27. □

Problem 2.27

For a Λ-rational subspace L we can define the quotient space $M := \mathbb{R}^d/L$ with projection $\pi : \mathbb{R}^d \to M$ and the orthogonal complement $M_0 \subseteq \mathbb{R}^d$ of L *w.r.t.* the standard scalar product. We can naturally identify M and M_0, as any element of M has a unique representative in M_0. This defines a volume on M, so we may define the determinant of the lattice $\pi(\Lambda)$ as the volume of a fundamental parallelepiped in M.

Further, let $L^\perp \subseteq (\mathbb{R}^d)^\star$ be the annihilator, *i.e.* the space of all functionals that vanish on L. We obtain a lattice in $L^\perp$ as $\Lambda^\star \cap L^\perp$. You will prove the following relations with Problem 2.29.

Problem 2.29

Proposition 2.33. *(i) The projection $\Gamma := \pi(\Lambda)$ of Λ onto M is a lattice in M of rank $d-k$ with*

$$\det \Gamma \cdot \det(\Lambda \cap L) = \det \Lambda .$$

(ii) For the dual lattice we have

$$\Gamma^\star = \Lambda^\star \cap L^\perp .$$

and

$$\det(\Lambda \cap L) = \det \Lambda \cdot \det \Gamma^\star .$$

Proof. See Problem 2.29 . □

2.6. Problems

2.1. By definition, a subset $\Lambda \subseteq \mathbb{R}^d$ is discrete if for all $\mathbf{x} \in \Lambda$ there is $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon(\mathbf{x}) \cap \Lambda = \{\mathbf{x}\}$. Show that if Λ is a lattice, then we can choose the same ε for all $\mathbf{x} \in \Lambda$.

2.2. Show that a discrete additive subgroup in $\mathbb{R}^d$ is closed (as a subset in the usual topology induced by some norm $\|.\|$).

2.3. Let Λ be a discrete closed subset of $\mathbb{R}^d$ (for instance, a discrete additive subgroup as in Problem 2.2) and B a bounded subset of $\mathbb{R}^d$.

Show that $\Lambda \cap B$ is a finite set. Give an example that shows that this is not correct for arbitrary discrete subsets.

2.4. Find an example of a finite, but linearly dependent, set $\mathcal{A}$ of vectors such that $\Lambda_{\mathcal{A}}$ is not a lattice.

2.5. Show that the collection $\mathcal{A}$ of the vectors $\mathbf{e}_i - \mathbf{e}_{i+1}$ for $1 \leq i \leq d-1$ is a basis of the lattice A and that $\mathcal{A} \cup \{ \mathbf{e}_{d-1} + \mathbf{e}_d \}$ is a basis of D.

2.6. In Example 2.2 we have shown that the root systems A and D are lattices. Here we want to consider the remaining root systems.

The root system B spans the lattice $\mathbb{Z}^d$, and that of the root system C coincides with the one of D. The root system E_7 spans the set

$$\{ \mathbf{x} \in \mathbb{R}^7 \} \; 2\mathbf{x} \in \mathbb{Z}^7 \text{ and } \sum_{i=1}^{d} x_i \text{ is even} .$$

Show that this is a lattice. For even d we define

$$\mathrm{D}_d^{1/2} \; := \; \mathrm{D}_d + \left(\frac{1}{2}\mathbf{1} + \mathrm{D}_d \right)$$

Show that this is a lattice. For $d = 8$ this is the root system E_8. E_7 is the sublattice of E_8 of all vectors perpendicular to one of the generators of E_8.

There are 3 more root systems E_6, F and G.

2.7. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice with fundamental parallelepiped Π. Show that the lattice translates of Π cover $\mathbb{R}^d$ without overlap, i.e.

$$\bigcup_{\mathbf{x} \in \Lambda} (\mathbf{x} + \Pi) \; = \; \mathbb{R}^d$$

and $(\mathbf{x} + \Lambda) \cap (\mathbf{y} + \Lambda) = \varnothing$ for $x, y \in \Lambda$, $x \neq y$.

2.8. Show that any full-dimensional Λ-rational cone contains a lattice basis.

Hint: Use induction over the dimension.

2.9. Let $\Lambda \subseteq \mathbb{R}^2$ be a 2-dimensional lattice. Show that a pair of vectors $\mathbf{b}_1, \mathbf{b}_2$ is a basis of Λ if and only if

$$\mathsf{conv}(0, \mathbf{b}_1, \mathbf{b}_2) \; \cap \; \Lambda \; = \; \{ 0, \mathbf{b}_1, \mathbf{b}_2 \} .$$

Can you extend this to higher dimensions?

2.10. A *lattice triangle* is a triangle $\Delta \subseteq \mathbb{R}^2$ whose vertices are in $\mathbb{Z}^2$.

Prove that any lattice triangle with only three lattice points has area $1/2$.

This is the *Theorem of Pick*.

2.11. Let P be a lattice polygon, *i.e.* a convex polygon in $\mathbb{R}^2$ whose vertices are in $\mathbb{Z}^2$. Let a be its volume, i the number of integral points in the interior of P and b the number of integral points on the boundary (including the vertices). Prove that

$$a \; = \; i \; + \; \frac{b}{2} \; - \; 1$$

This is *Pick's Formula*.

(i) Conclude that the number $|k \cdot P \cap \mathbb{Z}^2|$ of integral points in the k-th multiple of P is given by a quadratic polynomial in k.
(ii) Show that the same is true for the number of *interior* integral points in $k \cdot P$.
(iii) Is there a similar formula for polytopes in dimension 3?

2.12. Prove Proposition 2.13.

Hint: For one direction consider $\sum_{i=j}^{d} \lambda_i \mathbf{b}_i$ and use that you can divide the coefficients by the greatest common divisor to obtain a point in the linear span of $\mathcal{A}$.

For the other direction use the fact that lattice points in the linear span of $\mathcal{A}$ have two representations and compare coefficients.

2.13. Let $\Lambda = \mathbb{Z}^2$ be the standard lattice.

Can you construct an equilateral lattice triangle, *i.e.* an equilateral triangle whose vertices are points in $\mathbb{Z}^2$? Can you do it, if you are allowed to choose a different lattice?

2.14. Let Δ be a lattice triangle, *i.e.* a triangle whose vertices are lattice points, with one lattice point in the interior and no other lattice points on the boundary except the vertices.

Show that the interior lattice point is the centroid of the triangle.

Is a similar statement also true in higher dimensions?

2.15. Let $A \in \mathbb{Q}^{d \times m}$ of full row-rank.

- ▷ Show that there is a unimodular matrix $U \in \mathbb{Z}^{m \times m}$ such that AU is in Hermite normal form.
- ▷ Show that H is unique.

2.16. Show that the Hermite normal form can be computed in polynomial time.

2.17. Let $A \in \mathbb{Z}^{d \times m}$ and $\mathbf{b} \in \mathbb{Z}^d$ and consider the system $A\mathbf{x} = \mathbf{b}$ of linear Diophantine equations.

- ▷ Show that the system $A\mathbf{x} = \mathbf{b}$ has an integral solution if and only if $\mathbf{y}^t\mathbf{b}$ is an integer for all $\mathbf{y}^t$ such that $\mathbf{y}^t A$ is integral.
- ▷ Show that, if the system $A\mathbf{x} = \mathbf{b}$ has an integral solution $\mathbf{x}_0$, then there are linearly independent $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{Z}^d$ for $k = m - \operatorname{rank} A$ such that

$$\left\{ \mathbf{x} \in \mathbb{Z}^d \;:\; A\mathbf{x} = \mathbf{b} \right\} \;=\; \left\{ \mathbf{x}_0 + \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \;:\; \lambda_i \in \mathbb{Z} \text{ for } 1 \le i \le k \right\}. \tag{2.2}$$

- ▷ Show that the representation of (2.2) can be found in polynomial time.

2.18. Show that there exist lattices bases $\mathbf{b}_1, \ldots, \mathbf{b}_d$ of a lattice Λ, such that $\mathbf{b}_i \in \{0, \pm 1\}^d$ for $1 \le i \le d$, but but *all* basis vectors in the Hermite normal form $H = (\mathbf{h}_i)_{1 \le i \le d}$ have length

$$\|\mathbf{h}_i\| \;\ge\; 2^{\Omega(d)} \;\gg\; \sqrt{d} \;\ge\; \|vb_j\|$$

Additionally, you may think about the following two questions.

- ▷ Can you make the Hermite normal form even longer?
- ▷ Can you also find a basis with $\mathbf{b}_j \in \{0, 1\}^d$ with this property?

Hint: You may start by finding a basis $\mathbf{b}_j$ such that at least one of the $\mathbf{h}_i$ has norm at least $2^{\Omega(d)}$.

2.19. Let Λ be a lattice of full rank and B linearly independent lattice vectors that minimize $\det \mathrm{B}$. Then B is a lattice basis.

2.20. Let $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \left(\mathbb{Z}^d\right)^\star$ and $m_1, \ldots, m_k \in \mathbb{Z}$ for some $k > 0$. We define

$$\Lambda \;:=\; \left\{ \mathbf{z} \in Z^d \;:\; \mathbf{a}_i(\mathbf{z}) \equiv 0 \pmod{m_i} \text{ for } 1 \le i \le k \right\}.$$

Show that Λ is a lattice and that $\det \Lambda < \prod_{i=1}^{k} m_i$.

2.21. Let Λ be a lattice with basis B. Show that for any $\varepsilon > 0$ there is a radius r depending on ε, d, and B, so that

$$(1-\varepsilon) \cdot \operatorname{vol}(\mathcal{B}_r(0)) \;\le\; |\,\mathcal{B}_r(0) \cap \Lambda|\, \det \Lambda \;\le (1+\varepsilon) \cdot \operatorname{vol}(\mathcal{B}_r(0))$$

2.22. Let Γ be a sublattice of Λ of index $D := |\Gamma/\Lambda|$. Then

$$D\Lambda \subseteq \Gamma \subseteq \Lambda,$$

i.e. $D\Lambda$ is a sublattice of Γ.

2.23. Prove Theorem 2.30

2.24. Prove Theorem 2.29 using the Smith normal form from Theorem 2.30.

2.25. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be a basis of $\mathbb{R}^d$. For $x = \sum_{i=1}^{d} \lambda_i \mathbf{b}_i \in \mathbb{R}^d$ we define functionals $\mathbf{b}_i^*(\mathbf{x}) = \lambda_i$ for $1 \leq i \leq d$.
Show that $\mathbf{b}_1^*, \ldots, \mathbf{b}_d^*$ is a basis of $(\mathbb{R}^d)^*$.

2.26. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be a lattice basis of the lattice Λ in $\mathbb{R}^d$. Then $\mathbf{b}_1^*, \ldots, \mathbf{b}_d^*$ is a lattice basis of the lattice Λ^* in $(\mathbb{R}^d)^*$.

2.27. Prove Proposition 2.32.

2.28. Show that for arbitrary subspaces $L \subseteq \mathbb{R}^d$ the projection of a lattice in $\mathbb{R}^d$ onto L need not be a lattice.

Hint: $d = 2$ suffices for an example.

2.29. Prove Proposition 2.33

3. Geometry of Numbers

Geometry of numbers deals with the relation between convex bodies and lattices. The basic question of this area asks for connections between the number of lattice points of a convex body and its area.

Research on this started with the work of Hermann Minkowski, who used convex geometric methods, in particular his fundamental theorem, which is now known as *Minkowski's Theorem* (see Corollary 3.3), in order to bound class numbers in algebraic number theory. In the 20*th* century this new research field, the *Geometry of Numbers*, has grown into an established field of research with connections into many branches of mathematics.

The theory deals with general convex bodies. Our applications in later chapters for the results that we discuss in this chapter are mostly to polyhedra. So we will narrow our focus to polyhedra in the the presentation whenever it makes the statements or proofs easier.

3.1. Minkowski's Theorems

Minkowski's two theorems are the basis of this whole branch of discrete mathematics. Both essentially tell us something about generators of the lattice and prove that we can find such generators with a bounded Euclidean length. Constructing such bases, or at least finding one short direction in a lattice is the key ingredient to solve integer programming in polynomial time in fixed dimension.

With the theory developed in this chapter we will obtain quite powerful bounds on the lengths of short lattice vectors. However, most of the proofs are not constructive. Thus, we cannot immediately use them in the context of algorithms. We will have to reconsider some of the results in later chapters to come up with a construction of short vectors. This will come at the price of much weaker bounds, which, nevertheless, will still be sufficient to prove polynomiality.

Throughout this chapter $\Lambda \subset \mathbb{R}^d$ is a lattice of rank d (the reader may think of $\mathbb{Z}^d$).

3.1.1. Minkowski's First Theorem

Definition 3.1. A subset $K \subseteq \mathbb{R}^d$ of $\mathbb{R}^d$ is *centrally symmetric* if $-\mathbf{x} \in K$ for all $\mathbf{x} \in K$. K is a *convex body* if it is bounded and convex.

The set of convex bodies in $\mathbb{R}^d$ is denoted by $\mathcal{C}$, and the subset of *centrally symmetric* convex bodies is $\mathcal{C}_0$.

Note that the definition of the term *convex body* slightly varies in the literature.

With the following theorem we state a first fundamental correspondence between lattice points in a centrally symmetric convex body and its volume.

Theorem 3.2 (van der Corput, 1935)**.** *Let $K \subset \mathbb{R}^d$ be a centrally symmetric convex set. Then*

$$\operatorname{vol}(K) \leq 2^d \, |K \cap \Lambda| \det \Lambda \, .$$

The inequality is strict for compact K.

Minkowski's First Theorem, that he proved almost forty years earlier, is now a direct corollary of this. This result is the fundamental theorem in this area and it is considered to be the starting point of the theory. We state it before we give a proof of van der Corput's Theorem (Theorem 3.2).

Corollary 3.3 (Minkowski's First Theorem, 1898)**.** *Let $K \subseteq \mathbb{R}^d$ be convex and centrally-symmetric with $\operatorname{vol} K > 2^d \det \Lambda$.*

Then there exists $\mathbf{a} \neq 0$ in $K \cap \Lambda$. If K is also compact, then it suffices to assume $\operatorname{vol} K \geq 2^d \det \Lambda$. □

Equivalently we can also state that any centrally symmetric convex body K such that $\operatorname{int}(K) \cap \Lambda = \{0\}$ has volume bounded by $2^d \det \Lambda$. So, intuitively, if we have a convex body of volume exactly $2^d \det \Lambda$, then all nonzero lattice points in K must lie on the boundary. You will make this and similar observations more precise with Problem 3.1, Problem 3.2, and Problem 3.3.

Problem 3.1
Problem 3.2
Problem 3.3

For the proof of van der Corput's Theorem (Theorem 3.2) we will use the following lemma. Its proof uses a nice pigeonhole-style argument to show that the intersection of a sufficiently large set with some affine translate of the lattice is large.

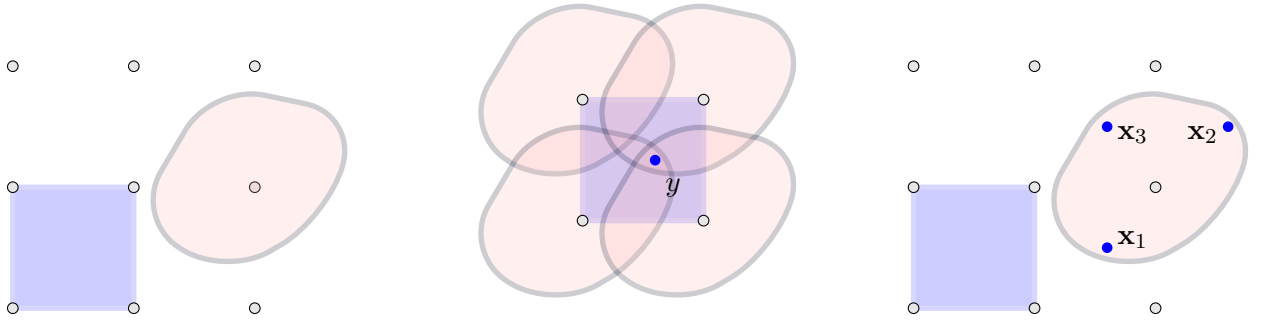
Lemma 3.4 (Generalized Blichfeldt's Theorem, 1914)**.** *Let $\mathcal{S} \subseteq \mathbb{R}^d$ be a (Jordan measurable) set with $\operatorname{vol}(\mathcal{S}) > m \det(\Lambda)$ for a positive integer m.*

Then there exist $m+1$ pairwise distinct points $\mathbf{p}_1, \ldots, \mathbf{p}_{m+1} \in \mathcal{S}$ such that $\mathbf{p}_i - \mathbf{p}_j \in \Lambda$ for all i, j.

Proof. By considering a sufficiently large subset, we may assume that $\mathcal{S}$ is bounded. Choose a *closed* fundamental parallelepiped (see Definition 2.5) $\overline{\Pi} := \overline{\Pi(\Lambda)}$ of Λ. Note that still $\det \Lambda = \operatorname{vol} \overline{\Pi}$. For any $\mathbf{x} \in \Lambda$ let

$$\mathcal{S}_{\mathbf{x}} := \{\mathbf{y} \in \overline{\Pi} \mid \mathbf{x} + \mathbf{y} \in \mathcal{S}\} = \overline{\Pi} \cap (\mathcal{S} - \mathbf{x})$$

Note that $\mathcal{S}_{\mathbf{x}} \neq \emptyset$ if and only if $\mathbf{x} \in (\mathcal{S} - \overline{\Pi}) \cap \Lambda$. For an illustration of this and the following argument see also Figure 3.1.



(a) The set $\mathcal{S}$ and the fundamental parallelepiped $\overline{\Pi}$

(b) And the shifted intersections with the fundamental parallelepiped

(c) Translating the point y gives the desired points $\mathbf{x}_i$

Figure 3.1.: Illustrating the proof of Generalized Blichfeldt's Theorem (Lemma 3.4)

As $\mathcal{S} - \overline{\Pi}$ is bounded, Problem 2.3 implies that there are only finitely many $\mathbf{x} \in \Lambda$ with $\mathcal{S}_\mathbf{x} \neq \emptyset$. This implies that the function

$$f \;:=\; \sum_{x \in \Lambda} \mathrm{id}_\mathbf{x}\,,$$

where $\mathrm{id}_\mathbf{x}$ is the indicator function on $\mathcal{S}_\mathbf{x}$ (i.e., it evaluates to 1 on $\mathcal{S}_\mathbf{x}$ and 0 elsewhere), is well-defined. Using that fundamental parallelepiped tile the space (Corollary 2.8) we compute

$$\begin{aligned}
\int_{\overline{\Pi}} f\,dx \;&=\; \sum_{x\in\Lambda} \int_{\overline{\Pi}} \mathrm{id}_\mathbf{x}\;dx \;=\; \sum_{x\in\Lambda} \mathrm{vol}(\mathcal{S}_\mathbf{x}) \\
&=\; \sum_{x\in\Lambda} \mathrm{vol}(\overline{\Pi} \cap (\mathcal{S} - \mathbf{x})) \;=\; \sum_{x\in\Lambda} \mathrm{vol}(\mathcal{S} \cap (\mathbf{x} + \overline{\Pi})) = \mathrm{vol}(\mathcal{S}) \\
&>\; m \det\,\Lambda \;=\; \int_{\overline{\Pi}} m\;dx
\end{aligned}$$

Hence, there is $\mathbf{y} \in \overline{\Pi}$ with $f(\mathbf{y}) > m$. Since f only evaluates to integers, we get $f(\mathbf{y}) \geq m+1$. In particular, there exist $\mathbf{x}_1, \ldots, \mathbf{x}_{m+1} \in \Lambda$ such that $\mathbf{y} \in \mathcal{S}_{\mathbf{x}_1} \cap \cdots \cap \mathcal{S}_{\mathbf{x}_{m+1}}$. Therefore, defining

$$\mathbf{p}_i \;:=\; \mathbf{y} \;+\; \mathbf{x}_i \;\in\; \mathcal{S} \qquad\qquad \text{for} \qquad\qquad i \;= 1, \ldots, m+1$$

yields $m+1$ points with the desired properties. □

With this preparation we can prove van der Corput's Theorem (Theorem 3.2) and Minkowski's First Theorem (Corollary 3.3).

Proof of van der Corput's Theorem (Theorem 3.2). We will give an indirect proof. Let us assume that

$$\mathrm{vol}(K) \;>\; m2^d \det(\Lambda)$$

for a positive integer m. Our goal is to show that there exist m distinct *non-zero* lattice

points $\mathbf{x}_1, \ldots, \mathbf{x}_m$ in K. Together with the origin this will give $m+1$ lattice points in K.

Let $\boldsymbol{T} := \frac{1}{2}K$. Then $\operatorname{vol} \boldsymbol{T} = \frac{\operatorname{vol} K}{2^d} > m \det \Lambda$. Hence, by the Generalized Blichfeldt's Theorem (Lemma 3.4), there are $m+1$ distinct points $\mathbf{p}_1, \ldots, \mathbf{p}_{m+1} \in \boldsymbol{T}$ such that $\mathbf{p}_i - \mathbf{p}_j \in \Lambda$ for all i, j. Choose $\mathbf{x}_i := \mathbf{p}_i - \mathbf{p}_{m+1}$ for $i = 1, \ldots, m$ as the desired lattice points. Note that $\mathbf{x}_i = \mathbf{p}_i + (-\mathbf{p}_{m+1}) \in \boldsymbol{T} + \boldsymbol{T} = K$. This proves the main part of the theorem.

For the second claim assume that K is compact and $\operatorname{vol} K = 2^d \det \Lambda$. Now, since K is compact, we can find $0 < \epsilon_{\mathbf{x}} < 1$ for each $\mathbf{x} \in 2K \backslash K$ such that $\mathbf{x} \notin (1+\epsilon_{\mathbf{x}})K$. Boundedness of $2K$ implies that $2K \cap \Lambda$ is finite (Problem 2.3).

Let ϵ be the minimum over $\epsilon_{\mathbf{x}}$ for all $\mathbf{x} \in (2K \backslash K) \cap \Lambda$. This choice ensures that $(1+\epsilon)K$ and K have the same set of lattice points (note that $\alpha K \subseteq \alpha' K$ for $0 < \alpha < \alpha'$ as K is centrally-symmetric and convex). Since $\operatorname{vol}((1+\epsilon)K) > 2^d \det \Lambda$, the result follows. □

Problem 3.4

Centrally-symmetric convex bodies with the origin as their only interior lattice point which have maximal volume $2^d \det(\Lambda)$ are also called *extremal bodies*.

Minkowski's theorem does not tell us how to find the integral point, it just tells us it exists. However, finding a short lattice vector is a very important problem in integer optimization and in cryptography. Good surveys are, *e.g.* in the book of Grötschel et. al[1] and the two books of Schrijver.[2] There are in fact polynomial time algorithms to explicitly find such a point (if the dimension is fixed), but only for a much larger volume bound. We will address this problem in the next two chapters.

Although we cannot easily compute a shortest vector of a lattice, Minkowski's Theorem at least allows us to estimate the length of such a vector.

Proposition 3.5. *Let $\Lambda \subset \mathbb{R}^d$ be a lattice. Then there is a vector $\mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}$ such that*

$$\|\mathbf{v}\| \;\leq\; \sqrt{d}(\det \Lambda)^{1/d}\,.$$

Note that the right hand side clearly has the correct exponent. If we scale the lattice Λ by a factor of α, then also the length of a shortest vector scales with α, but the volume scales with α^d.

Proof. Let $\mathrm{V_d}$ be the volume of the d-dimensional unit ball $\mathcal{B}_\mathrm{d}$ and choose

$$\alpha \;:=\; 2\left(\frac{\det \Lambda}{\mathrm{V_d}}\right)^{1/d}. \tag{3.1}$$

Then

$$\operatorname{vol}(\alpha\, \mathcal{B}_\mathrm{d}) \;=\; \alpha^d\, \mathrm{V_d} \;\geq\; 2^d \det \Lambda\,.$$

[1] Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*.
[2] Alexander Schrijver, *Combinatorial optimization. Polyhedra and efficiency (3 volumes)*. Alexander Schrijver, *Theory of linear and integer programming*.

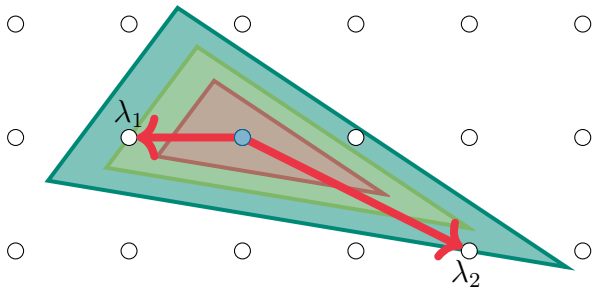


Figure 3.2.: A triangle in the plane together with two scaled copies with scaling factors λ_1 and λ_2.

By Minkowski's First Theorem (Corollary 3.3) there is a non-zero lattice point $\mathbf{v}$ in $\alpha\,\mathcal{B}_\mathrm{d}$, hence, of length at most α. We need to estimate the size of α.

The volume of the unit ball is

$$\mathrm{V_d} \;:=\; \frac{\pi^{\lfloor d/2\rfloor}2^{\lceil d/2\rceil}}{\prod_{0\le i<d/2}(d-2i)} \;\approx\; \left(\frac{2\pi e}{d}\right)^{d/2} \;\ge\; \left(\frac{4}{d}\right)^{d/2},$$

see Problem 3.5. The first approximation follows from Stirling's formula $d! \approx \sqrt{2\pi d}\frac{d}{e}^d$ and the second from $2\pi e \ge 4$. Inserting this into (3.1) proves the result. □

Problem 3.5

Using a more careful analysis one can improve the bound to

$$(1+o(1))\sqrt{2d/e\pi}\,(\det\,\Lambda)^{1/d}\,.$$

On the other hand, we know that there are lattices which essentially realize this bound, *i.e.* there are lattices with shortest vector of length $\Omega(\sqrt{d}(\det\Lambda)^{1/d})$.

The bound given by Proposition 3.5 can be arbitrarily bad already in dimension 2. This can already be seen from the simple basis $\mathbf{e}_1$ and $M\mathbf{e}_2$ for some $M \in \mathbb{Z}_{>0}$. The lattice contains a vector of length 1, while the determinant is M.

3.1.2. Successive Minima

Definition 3.6. Let $K \in \mathcal{C}_0$. For $1 \le k \le d$ we define the *k-th successive minimum of K* to be the number

$$\lambda_k \;:=\; \lambda_k(K) \;:=\; \inf_{\lambda>0}\{\dim \mathrm{lin}(\lambda K \cap \Lambda) \;\ge\; k\}.$$

For $K = \mathcal{B}_\mathrm{d}(0)$ we call $\lambda_k := \lambda_k(\mathcal{B}_\mathrm{d})$ the *k-th successive minimum of the lattice Λ*.

Figure 3.2 shows an example for λ_1 and λ_2 in dimension 2. The successive minima satisfy

$$0 \;<\; \lambda_1 \;\le\; \lambda_2 \;\le\; \cdots \;\le\; \lambda_d\,,$$

where the first inequality follows as Λ is discrete. If $\mathbf{v} \in \Lambda \setminus \{0\}$ is a shortest *non-zero* lattice vector in the norm defined by K, then

$$\lambda_1(K) \;=\; \|\mathbf{v}\|_K \;.$$

There is no similar simple relation for the higher successive minima. In the standard Euclidean norm we obtain a bound for $\lambda_1 = \lambda_1(\mathcal{B}_{\mathrm{d}}(0))$ from Proposition 3.5.

Corollary 3.7. *Let $\Lambda \subset \mathbb{R}^d$ be a lattice. Then*

$$\lambda_1 \;\leq \sqrt{d}(\det \Lambda)^{1/d} \,.$$

The following corollary is equivalent to Minkowski's First Theorem (Corollary 3.3).

Corollary 3.8. *Let $K \in \mathcal{C}_0$. Then*

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

All theorems above deal with *centrally symmetric* convex bodies, and the results are wrong for more general sets. One reason for this is the connection of compact centrally symmetric convex sets 3to metric geometry, which allows us to obtain bounds on the norm. More specifically, any *compact centrally symmetric* convex body defines a *norm* $\|.\|_K$ on $\mathbb{R}^d$ via

$$\|\mathbf{x}\|_K \;:=\; \max_{\mu} \{\, \mu\mathbf{x} \in K \,\} \;,$$

and, conversely, any norm $\|.\|$ is of this form, with

$$K \;:=\; \{\, \mathbf{x} \;:\; \|\mathbf{x}\| \leq 1 \,\} \;.$$

Proposition 3.9. *Let $K \in \mathcal{C}_0$ be compact and $\Lambda \subseteq \mathbb{R}^d$ be a lattice with successive minima $\lambda_1, \ldots, \lambda_d$ with respect to K. Then there is a (vector space) basis $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \Lambda$ such that $\|\mathbf{v}_i\|_K = \lambda_i$ for $1 \leq i \leq d$.*

Proof. Pick some index $1 \leq j \leq d$. By definition of λ_j there is a sequence $(\mathbf{w}_i)_{i\geq 1} \subseteq \Lambda$ of lattice vectors such that $\lim_{i\to\infty} \|\mathbf{w}_i\| = \lambda_j$. For sufficiently large i we have $\mathbf{w}_i \in 2K$. K is compact, so we can find a convergent sub-sequence $\mathbf{w}_{i_k}$, converging to some vector $\mathbf{w}$. We need to prove that $\mathbf{w} \in \Lambda$. By definition, $\lim_{k\to\infty} \|\mathbf{w} - \mathbf{w}_{i_k}\|_K = 0$, so for sufficiently large k

$$\|\mathbf{w} - \mathbf{w}_{i_k}\|_K \;<\; \lambda_1/2 \,.$$

The triangle inequality then implies for sufficiently large k, l

$$\|\mathbf{w}_{i_k} - \mathbf{w}_{i_l}\|_K \;\leq\; \|\mathbf{w} - \mathbf{w}_{i_k}\|_K + \|\mathbf{w} - \mathbf{w}_{i_l}\|_K < \lambda_1 \,.$$

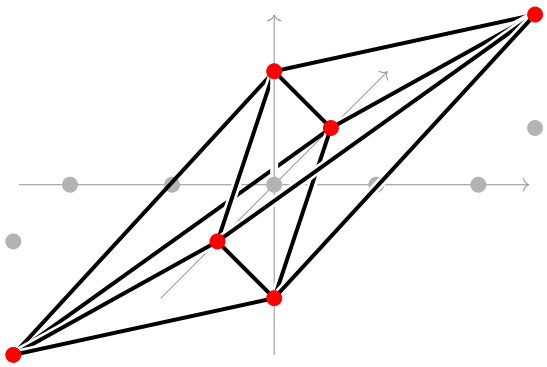


Figure 3.3.: The lattice points determining the successive minima need not be a lattice basis. In the polytope $P := \mathsf{conv}\,(\pm\mathbf{e}_2, \pm\mathbf{e}_3, \pm(\mathbf{e}_2 + \mathbf{e}_3 + 2\mathbf{e}_1))$ the vertices in the positive orthant determine the first linearly independent set of lattice points, but they do not span $\mathbb{Z}^3$.

But $\mathbf{w}_{i_l} - \mathbf{w}_{i_l}$ is a lattice vector, so $\mathbf{w}_{i_k} = \mathbf{w}_{i_l}$ for sufficiently large k, l. Hence, $\mathbf{w}_{i_k} = \mathbf{w}$ for sufficiently large k, and $\mathbf{w}$ is a lattice vector. □

The vectors found in the previous proposition need not be a basis of the lattice Λ. For an example, the lattice polytope

$$P \;:=\; \mathsf{conv}\,(\pm\mathbf{e}_2, \pm\mathbf{e}_3, \pm(\mathbf{e}_2 + \mathbf{e}_3 + 2\mathbf{e}_1))$$

in the lattice $\mathbb{Z}^3$ is centrally symmetric and its lattice points are the vertices and the origin. See Figure 3.3.

Hence, the successive minima are $\lambda_1 = \lambda_2 = \lambda_3 = 1$, but no subset of the vertices is a lattice basis of $\mathbb{Z}^3$. You will see in Problem 3.6 that also a set of lattice vectors of length $\lambda_1, \ldots, \lambda_d$ is not a lattice basis in general. However, Problem 3.8 shows that this is true in dimensions up to 4.

Problem 3.6

Problem 3.7
Problem 3.8

We can, however, achieve the following much weaker result on the connection of successive minima and a lattice basis. You will prove this in Problem 3.9.

Proposition 3.10. *There is a basis* $\mathrm{B} = \{\, \mathbf{b}_1, \ldots, \mathbf{b}_d \,\}$ *of* Λ *such that*

$$\Lambda \cap \lambda_i \operatorname{int} K \;\subseteq \operatorname{lin}\,(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}) \cap \Lambda\,.$$

Problem 3.9

The following result is a cornerstone of the theory of successive minima. We will not prove this much stronger theorem here.

Theorem 3.11 (Minkowski's Second Theorem, 1896)**.** *Let* $K \in \mathcal{C}_0$. *Then*

$$\frac{1}{d!} \cdot 2^d \det \Lambda \;\leq\; \lambda_1 \cdots \lambda_d \operatorname{vol} K \;\leq\; 2^d \det\, \Lambda.$$

The proof of the lower bound is simple, and you will do this in Problem 3.10, where you also show that the bound is tight. A proof of the upper bound, which is much more involved, can be found in the paper of Henk.[3]

Problem 3.10

Corollary 3.7 bounds the first successive minimum by d and the determinant. We can use Minkowski's Second Theorem (Theorem 3.11) to show that the same bound also holds for the larger geometric average of all successive minima.

Corollary 3.12. *Let $\Lambda \subset \mathbb{R}^d$ be a lattice. Then*

$$\left(\prod_{i=1}^{d} \lambda_i\right)^{1/d} \leq \sqrt{d}(\det \Lambda)^{1/d}.$$

Proof. Recall that $\lambda_i = \lambda_i(\mathcal{B}_\mathrm{d}(0))$, so we want to use Minkowski's Second Theorem (Theorem 3.11) for $K = \mathcal{B}_\mathrm{d}(0)$. We use the same lower bound for the volume of $\mathcal{B}_\mathrm{d}(0)$ as in Proposition 3.5,

$$\mathrm{V_d} \;\geq\; \left(\frac{4}{d}\right)^{d/2} \;=\; 2^d \frac{1}{\sqrt{d}^{\,d}}.$$

Thus,

$$\lambda_1 \cdots \lambda_d \cdot 2^d \frac{1}{\sqrt{d}^{\,d}} \;\leq\; \lambda_1 \cdots \lambda_d \,\mathrm{V_d} \;\leq\; 2^d \det \Lambda.$$

Rearranging the inequality and taking the d-th root gives the claim. □

Historically, the length of a shortest lattice vector in the Euclidean length was studied first in the context of quadratic forms. Hermite[4] studied the quotient

$$\gamma(\Lambda) \;:=\; \left(\frac{\lambda_1(\Lambda)}{(\det \Lambda)^{1/d}}\right)^2,$$

which is now known as the *Hermite factor* of the lattice. The *Hermite constant* is the supremum

$$\gamma_d \;:=\; \sup_{\Lambda} \gamma(\Lambda).$$

It follows from Proposition 3.5 that

$$\gamma_d \;\leq\; d.$$

The precise value is only known in a few cases for small d. See Table 3.1. $\gamma_2 = \frac{\sqrt{3}}{2}$ is realized by the hexagonal lattice, or, as historically first descovered, by the lattice of

[3] Henk, "Successive minima and lattice points".

[4] Hermite, "Extraits de lettres de M. Ch. Hermite à M. Jacobi sur différents objects de la théorie des nombres".

d	1	2	3	4	5	6	7	8	24
γ_d^d	1	$\frac{4}{3}$	2	4	8	$\frac{64}{3}$	64	2^8	4^{24}

Table 3.1.: Known values of the Hermite constant. Note that the table gives γ_d^d instead of γ_d.

Eisenstein integers.

Problem 3.11

* 3.1.3. The Lattice Point Enumerator

Betke, Henk, and Wills[5] conjectured a similar result as Minkowski's Second Theorem also for the lattice point count. As far as we know this conjecture is still open.

Conjecture 3.13 (Betke, Henk, Wills, 1993)**.** *Let $K \in \mathcal{C}_0$. Then*

$$|K \cap \mathbb{Z}^n| \;\leq\; \prod_{i=1}^{d} \left\lfloor \frac{2}{\lambda_i} + 1 \right\rfloor .$$

Some evidence for the correctness of the conjecture is given by the observation that the corresponding analogue for Minkowski's First Theorem is true.

Theorem 3.14 (Betke, Henk, Wills, 1993[6])**.** *Let $K \in \mathcal{C}_0$. Then*

$$|K \cap \mathbb{Z}^n| \;\leq\; \left\lfloor \frac{2}{\lambda_1} + 1 \right\rfloor^d$$

Proof. We follow the original proof. Let $m := \left\lfloor \frac{2}{\lambda_1} + 1 \right\rfloor$. Suppose there are $\mathbf{x}, \mathbf{y} \in K$, $\mathbf{x} \neq \mathbf{y}$ such that

$$x_i \;\equiv\; y_i \pmod{m} \qquad \text{for} \qquad 1 \leq i \leq d\,.$$

We consider the point

$$\mathbf{z} \;:=\; \frac{1}{2}\left(\frac{2}{m}\mathbf{x}\right) \;+\; \frac{1}{2}\left(\frac{2}{m}\mathbf{y}\right) \;=\; \frac{1}{m}(\mathbf{x}-\mathbf{y})\,.$$

By assumption, each entry of $\mathbf{x} - \mathbf{y}$ is divisible by m, so $\mathbf{z}$ is a lattice point. By choice of m we have $\frac{2}{m} < \lambda_1$, so $\mathbf{z} \in \lambda_1 \operatorname{int} K \cap \mathbb{Z}^d$ and non-zero. This contradicts the definition of λ_1, so there are no two points in K whose difference is m times a lattice vector.

The pigeonhole principle implies that there can be at most m^d different points in K that are not congruent modulo m. This implies the theorem. □

[5] Betke, Henk, and Wills, "Successive-minima-type inequalities".
[6] Betke, Henk, and Wills, "Successive-minima-type inequalities".

Later, Henk[7] verified the conjecture with a weaker bound using Proposition 3.10

Theorem 3.15 (Henk 2007)**.**

$$|K \cap \mathbb{Z}^n| \;\leq\; 2^{d-1} \prod_{i=1}^{d} \left\lfloor \frac{2}{\lambda_i} + 1 \right\rfloor .$$

Proof. Set

$$\alpha_i \;:=\; \left\lfloor \frac{2}{\lambda_i} + 1 \right\rfloor .$$

Then α_i is the smallest integer such that

$$\frac{2}{\alpha_i} \;>\; \lambda_i \,,$$

and $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_d$. Choose a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ as in Proposition 3.10 and choose numbers $k_1 \geq \ldots \geq k_{d-1} \leq k_d := 1$ such that $\eta_i := 2^{k_i}\alpha_d$ satisfies

$$\eta_d \;:=\; \alpha_d \qquad\qquad \text{and } \alpha_i \qquad\qquad \leq\; \eta_i \;\leq\; 2\alpha_i$$

for $1 \leq i \leq d-1$. The η_i satisfy

$$\eta_k | \eta_j \text{ for } k \geq j \,.$$

Now consider the lattice

$$\Lambda' \;:=\; \bigoplus \mathbb{Z}\eta_i \mathbf{b}_i \,.$$

We want to show that $\Lambda' \cap 2K = \{\,\mathbf{0}\,\}$. Assume not, and let $\mathbf{u} \in \Lambda'$ with $\|\mathbf{u}\|_K \leq 2$. We can write $\mathbf{u}$ as

$$\mathbf{u} \;:=\; \sum_{i=1}^{d} \mu_i \eta_i \mathbf{b}_i$$

for some integers $\mu_1, \ldots, \mu_d$. Let i be the largest index with $\mu_i \neq 0$. As η_k divides $\eta_1, \ldots, \eta_{i-1}$ we have $\frac{1}{\eta_i}\mathbf{u} \in \Lambda'$, and so

$$\left\| \frac{1}{\eta_i}\mathbf{u} \right\| \;\leq\; \frac{2}{\eta_i} \;\leq\; \frac{2}{\alpha_i} \;\leq\; \lambda_i \,.$$

Hence, $\frac{1}{\eta_i}\mathbf{u} \in \operatorname{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$, which is a contradiction to the choice of the basis.

Thus, the projection $\pi : \mathbb{R}^d \longrightarrow \mathbb{R}^d/\Lambda'$ is injective on K, so

$$|K \cap \mathbb{Z}^d| \;=\; |\pi(K \cap \mathbb{Z}^d)|$$

[7] Henk, "Successive minima and lattice points", Thm 1.5.

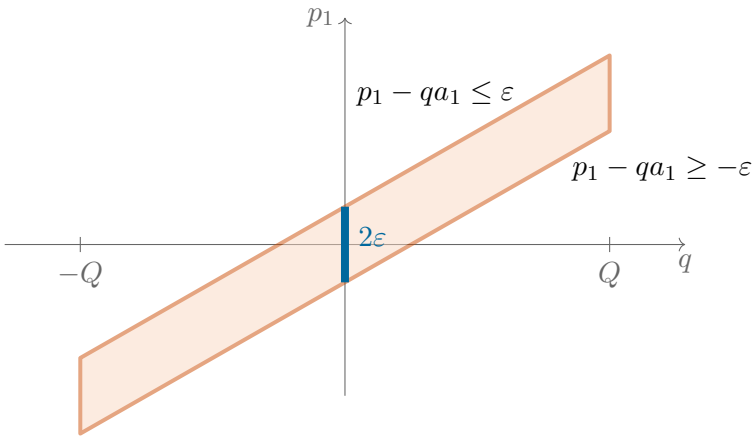


Figure 3.4.: The centrally symmetric convex body K. This is a cut off tube of width 2ε around $\mathbf{p} - q \cdot \mathbf{a} = 0$.

But $\pi(K \cap \mathbb{Z}^d) \subseteq \Lambda/\Lambda'$, so

$$|K \cap \mathbb{Z}^d| \;\leq\; |\Lambda/\Lambda'| \;=\; \prod \eta_i \;\leq\; 2^{d-1} \prod \alpha_i \,. \qquad \square$$

3.1.4. Dirichlet's Theorem

We want to discuss one application of Minkowski's First Theorem (Corollary 3.3) in this section. You can find another one in Problem 3.13.

Assume we are given a vector $\mathbf{a} \in [0,1]^d$ with entries $a_i \in \mathbb{R}$. We want to approximate $\mathbf{a}$ with a *rational* vector $\overline{\mathbf{a}} \in \mathbb{Q}^d$ with a denominator bounded by some given Q. This is a common task in many applications, as computers cannot deal efficiently with non-rational numbers (although we may represent some exactly using, *e.g.* field extensions). On possible choice would be

$$\overline{\mathbf{a}} \;:=\; \begin{bmatrix} \frac{\lceil a_1 Q \rfloor}{Q} \\ \vdots \\ \frac{\lceil a_d Q \rfloor}{Q} \end{bmatrix} .$$

where $\lceil x \rfloor$ denotes rounding to the nearest integer. The rounding error with this choice is at most $\frac{1}{2Q}$. Can we do better? We may observe that we have not used the assumption that the denominator should be *at most* Q. This, together with Minkowski's First Theorem (Corollary 3.3) sometimes allows a better approximation.

Theorem 3.16 (Dirichlet). *For any $\mathbf{a} \in (0,1]^d$ and $Q \in \mathbb{Z}_{>0}$ there are $p_1, \ldots, p_d \in \mathbb{Z}_{\geq 0}$ and $q \in \{1, \ldots, Q\}$ such that*

$$\left|\frac{p_i}{q} - a_i\right| \;\leq\; \frac{1}{Q^{1/d} q} \qquad \textit{for} \qquad 1 \leq i \leq d \,.$$

Proof. Let $\varepsilon : -\frac{1}{Q^{1/d}}$ and define

$$K \;:=\; \left\{ (p_1, \ldots, p_d, q)^t \in \mathbb{R}^{d+1} \;:\; |p_i \;-\; q \cdot a_i| \;\leq\; \varepsilon \text{ for } 1 \leq i \leq d \text{ and } |q| \leq Q \right\}.$$

This set is defined by a finite set of linear inequalities, so it is a polyhedron. It is also bounded and centrally symmetric, so a centrally symmetric convex body. See Figure 3.4 for an example. The volume is

$$\text{vol } K \;=\; 2Q \cdot (2\varepsilon)^d \;=\; 2Q \left(\frac{2}{Q^{1/d}} \right)^d \;=\; 2^{d+1}.$$

It follows from Minkowski's First Theorem (Corollary 3.3) that $K \cap \mathbb{Z}^{d+1} \setminus \{0\}$ is not empty. Pick any integral point $(p_1, \ldots, p_d, q)$ in this intersection. By symmetry, we may assume that $q \geq 0$. If $q = 0$, then $|p_i| \leq \varepsilon < 1$. This would imply $p_i = 0$, which is not possible. So $q > 0$ and

$$|p_i \;-\; q \cdot a_i| \;\leq\; \varepsilon\,.$$

Dividing both sides by q gives the desired approximation. □

Problem 3.12
Problem 3.13

3.2. Coverings and Packings

For $r > 0$ and $\mathbf{z} \in \mathbb{R}^d$ let

$$\mathcal{B}_r^\circ(\mathbf{z}) \;:=\; \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x} - \mathbf{z}\| \;<\; r\}$$

be the *open ball* of radius r around $\mathbf{z}$. In this section we consider the configuration of all translates of such a ball to all lattice points. We want to determine for which radii these translates are pairwise disjoint or cover the whole space, and relations between these two.

We start with the first and introduce the *packing radius* of a lattice, which, in plain words, is the largest radius of a ball such that any two translates to a lattice point either coincide or are disjoint.

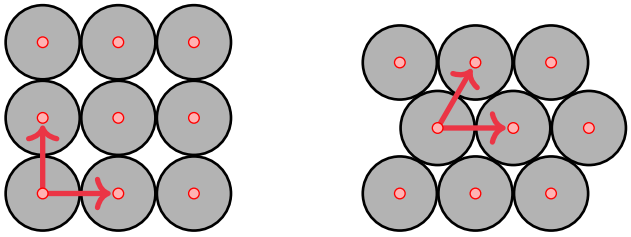
Definition 3.17. Let Λ be a lattice in $\mathbb{R}^d$. The *packing radius* is

$$\varrho(\Lambda) \;:=\; \sup_{r>0} \big(\mathcal{B}_r^\circ(\mathbf{x}) \cap \mathcal{B}_r^\circ(\mathbf{y}) \;=\; \varnothing \text{ for all } \mathbf{x}, \mathbf{y} \in \Lambda \big)\,,$$

i.e. the largest $r > 0$ such that the open balls of radius r around any two distinct lattice points do not intersect.

Problem 3.14

See Figure 3.5 for some examples. You will prove the following connection between the packing radius and the first successive minimum in Problem 3.14.



(a) The standard lattice $\mathbb{Z}^2$ (b) The hexagonal lattice

Figure 3.5.: Examples of Packings in the square and hexagonal lattice

Proposition 3.18. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. Then $\varrho(\Lambda) = \frac{1}{2}\lambda_1$.*

Problem 3.16

Recall that in Definition 2.31 we have defined the dual of a lattice Λ to be the set of all linear functionals that map lattice points to integers. This is itself a lattice $\Lambda^\star$ in $(\mathbb{R}^d)^\star$. As with many other places where we have dual objects, transformations like scalings induce the inverse operation on the dual, which connects invariants on the primal and dual side. This is also true for the packing radius, as we prove with the next proposition.

Proposition 3.19. *Let Λ be a lattice in $\mathbb{R}^d$ with dual lattice $\Lambda^\star$. Then*

$$\varrho(\Lambda) \cdot \varrho(\Lambda^\star) \leq {}^{d}\!/_{4}\,.$$

Proof. By Proposition 3.18 the packing radius is half the length of a shortest non-zero lattice vector, and by Proposition 3.5 we can bound this length with

$$\varrho(\Lambda) \leq \frac{1}{2}\sqrt{d}(\det \Lambda)^{1/d} \qquad \varrho(\Lambda^\star) \leq \frac{1}{2}\sqrt{d}(\det \Lambda^\star)^{1/d}$$

both for Λ and its dual. The proposition now follows as $\det \Lambda \cdot \det \Lambda^\star = 1$. □

Problem 3.17

The following statement is immediate from the previous two propositions.

Corollary 3.20. *For any Λ we have $\lambda_1 \cdot \lambda_1^* \leq d$.* □

Banaszczyk has proved the much stronger result that we can replace one λ_1 by λ_n.[8] It follows from this corollary that if λ_1 is large, say $\lambda_1 \gg N$ for some $N \geq d$, then the corresponding value for the dual must be small, *i.e.* $\lambda_1^\star \leq {}^{d}\!/_{N}$. However, the converse is not necessarily true, *i.e.* both λ_1 and $\lambda_1^\star$ can be small, see Problem 3.18. Similarly, one can show that λ_n and $\lambda_n^\star$ cannot both be small at the same time, but they can both be large, see Problem 3.19.

Problem 3.18
Problem 3.19

[8] W. Banaszczyk, "Inequalities for convex bodies and polar reciprocal lattices in $\mathbf{R}^n$. II. Application of K-convexity".

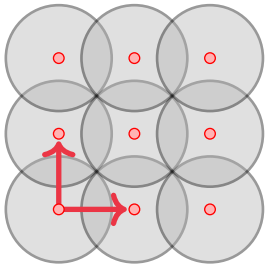


Figure 3.6.: The covering radius of the standard lattice is $\sqrt{2}/2$.

Now we switch the view and want to find out how large we need to make the radius of our balls so that the translates cover the whole space. This is captured with the next definition.

Definition 3.21 (Covering Radius)**.** Let Λ be a lattice in $\mathbb{R}^d$ of full rank. The *covering radius* is

$$\mu(\Lambda) \;:=\; \max_{\mathbf{x}\in\mathbb{R}^d} \mathrm{d}(\mathbf{x},\Lambda)\,,$$

i.e. the largest possible distance between any point in $\mathbb{R}^d$ and its nearest lattice point.

See Figure 3.6 for an example. You should convince yourself that the covering radius is indeed well-defined and finite. In particular, the maximum is attained for some point $\mathbf{x} \in \mathbb{R}^d$ (by a standard compactness argument).

Problem 3.20
Problem 3.21

Lemma 3.22. *Let Λ be a lattice with successive minima $\lambda_1, \ldots, \lambda_d$ and linearly independent lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_d$ such that $\lambda_i = \|\mathbf{v}_i\|$ for $1 \le i \le d$. Then*

$$\mu(\Lambda) \;\ge\; \frac{1}{2}\,\|\mathbf{v}_i\| \qquad\qquad \text{for } 1 \le i \le d\,.$$

Proof. Let $\mathbf{u} = \frac{1}{2}\mathbf{v}_d$. Assume there is $\mathbf{w} \in \Lambda$ such that $\mathrm{d}(\mathbf{u},\mathbf{w}) < \frac{1}{2}\,\|\mathbf{v}_d\|$. Then

$$\|\mathbf{w}\| \;\le\; \|\mathbf{u}\| \;+\; \mathrm{d}(\mathbf{u},\mathbf{w}) \;<\; \|\mathbf{v}_d\|\,,$$

so $\mathbf{w}$ cannot be linearly independent of $\mathbf{v}_1, \ldots, \mathbf{v}_{d-1}$ by the choice of $\mathbf{v}_d$. Hence, $\mathbf{w}$ is in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_{d-1}$. But then $2\mathbf{w} - \mathbf{v}_d$ is linearly independent, and

$$\|2\mathbf{w} - \mathbf{v}_d\| \;=\; \|2(\mathbf{w} - \mathbf{u})\| \;<\; \|\mathbf{v}_d\|$$

again contradicting the choice of $\mathbf{v}_d$. Hence, $\mathrm{d}(\mathbf{u},\Lambda) = \mathrm{d}(\mathbf{u},\mathbf{0}) = \frac{1}{2}\,\|\mathbf{v}_d\|$. This implies that

$$\mu(\Lambda) \;\ge\; \frac{1}{2}\,\|\mathbf{v}_d\| \;\ge\; \frac{1}{2}\,\|\mathbf{v}_i\|$$

for $1 \le i \le d$, where the latter follows from $\|\mathbf{v}_d\| \ge \|\mathbf{v}_i\|$ for all i. □

Corollary 3.23. *The covering radius is bounded from below by half the last successive minimum,*

$$\mu(\Lambda) \;\geq\; \frac{1}{2}\lambda_d\,.$$

Problem 3.22
Problem 3.23
Problem 3.24

Lemma 3.24. *Let Λ be a lattice in $\mathbb{R}^d$ with dual lattice $\Lambda^\star$. Then*

$$4\,\mu(\Lambda)\cdot\varrho(\Lambda^\star) \;\geq\; 1$$

Proof. Let Λ be a lattice with successive minima $\lambda_1,\dots,\lambda_d$ and linearly independent vectors $\mathbf{v}_1,\dots,\mathbf{v}_d\in\Lambda$ such that $\lambda_i=\|\mathbf{v}_d\|$ for $1\leq i\leq d$. Let $\mathbf{u}$ be a shortest non-zero lattice vector in $\Lambda^\star$. Proposition 3.18 and Lemma 3.22 imply for any $1\leq i\leq d$

$$4\,\mu(\Lambda)\cdot\varrho(\Lambda^\star) \;=\; 2\,\mu(\Lambda)\cdot\|\mathbf{u}\| \;\geq\; \|\mathbf{v}_i\|\cdot\|\mathbf{u}\|\;. \tag{3.2}$$

The vectors $\mathbf{v}_1,\dots,\mathbf{v}_d$ are a basis, so for at least one i we have $|\mathbf{v}_i(\mathbf{u})|\geq 1$. Hence, for that i

$$\|\mathbf{v}_i\|\cdot\|\mathbf{u}\| \;\geq\; 1\,,$$

which, together with (3.2) implies the claim. □

It follows from this lemma, that the product $\lambda_1\cdot\lambda_d^\star\geq 1$, see Problem 3.25. More generally, one can also show that for any $1\leq k\leq d$

$$\lambda_k\cdot\lambda^\star_{d-k+1} \;\geq\; 1\,,$$

see Problem 3.26.

Problem 3.25
Problem 3.26

So if λ_d (or the covering radius $\mu\geq\lambda_d/2$) is small, then the dual minimum length $\lambda_1^\star\geq 1/\lambda_d$ must necessarily be large. Here, also the converse is true, as already noted above, below Corollary 3.20. However, this result needs new methods for the proof.

The following theorem is the key ingredient for the flatness theorem that we will prove in Section 3.3.

Theorem 3.25. *Let Λ be a lattice in $\mathbb{R}^d$ with dual lattice $\Lambda^\star$. Then*

$$1 \;\leq\; 4\,\mu(\Lambda)\cdot\varrho(\Lambda^\star) \;\leq\; d^{3/2}\,.$$

Our proof of this theorem is based on an argument by Schnorr, Lagarias, and Lenstra.[9]

[9] J. C. Lagarias, H. W. Lenstra Jr., and C.-P. Schnorr, "Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice".

Proof. The left inequality is Lemma 3.24.

We have to prove the right inequality. This is done by induction over d. For $d = 1$ we have, for some $\lambda > 0$,

$$\Lambda \;=\; \lambda\mathbb{Z} \qquad\qquad \text{and} \qquad\qquad \Lambda^\star \;=\; \lambda^{-1}\mathbb{Z}^\star \;\cong\; \lambda^{-1}\mathbb{Z}\,.$$

Thus, $\mu(\Lambda) = \frac{1}{2}\lambda$ and $\varrho(\Lambda^\star) = \frac{1}{2}\lambda^{-1}$, so that

$$\mu(\Lambda)\,\varrho(\Lambda^\star) \;=\; \tfrac{1}{4} \;\leq\; \tfrac{1}{4}\cdot 1^{3/2}\,.$$

Now let $d > 1$. We choose a shortest non-zero lattice vector $\mathbf{v} \in \Lambda$. Then $\|\mathbf{v}\| = 2\,\varrho(\Lambda)$. Let L be the orthogonal complement of $\mathbf{v}$ with projection $\pi : \mathbb{R}^d \to L$ and $\Gamma := \pi(\Lambda)$. Then Γ is a lattice in L and $\Gamma^\star \subseteq \Lambda^\star$ by Problem 2.29. Hence

$$\varrho(\Gamma^\star) \;\geq \varrho(\Lambda^\star)\,. \tag{3.3}$$

We now want to bound $\mu(\Lambda)$. For this, let $\mathbf{x} \in \mathbb{R}^d$, $\mathbf{y} = \pi(\mathbf{x})$ and $\mathbf{u} \in \Gamma$ a closest point to $\mathbf{y}$ in L. Then

$$\|\mathbf{u} - \mathbf{y}\| \;\leq\; \mu(\Gamma)\,.$$

Consider the line $\pi^{-1}(\mathbf{u})$. Any two neighboring lattice points of Λ on this line have distance $\|\mathbf{v}\|$. Hence, we can pick a point $\mathbf{w} \in \Lambda \cap \pi^{-1}(\mathbf{u})$ such that

$$\mathrm{d}(\mathbf{x}, \mathbf{w} + (\mathbf{y} - \mathbf{u})) \;=\; \|\mathbf{x} - (\mathbf{w} + (\mathbf{y} - \mathbf{u}))\| \;\leq\; \frac{1}{2}\,\|\mathbf{v}\|\,.$$

Using the right angled triangle $\mathbf{x}, \mathbf{w}, \mathbf{w} + (\mathbf{y} - \mathbf{u})$ we compute

$$\|\mathbf{x} - \mathbf{w}\|^2 \;\leq\; \|\mathbf{x} - (\mathbf{w} + (\mathbf{y} - \mathbf{u}))\|^2 + \|\mathbf{y} - \mathbf{u}\|^2\,.$$

Now $\mathbf{x}$ was chosen arbitrary, so we can assume it is a point with maximum distance to the lattice and we can estimate (note that $\mathbf{w}$ need not be a lattice point closest to $\mathbf{x}$)

$$\mu(\Lambda)^2 \;\leq\; \|\mathbf{x} - \mathbf{w}\|^2 \;\leq\; \mu(\Gamma)^2 + \frac{1}{4}\,\|\mathbf{v}\|^2 \;=\; \mu(\Gamma)^2 + \varrho(\Lambda)^2\,.$$

Hence, we obtain

$$\begin{aligned}
\mu(\Lambda)^2 \cdot \varrho(\Lambda^\star)^2 \;&\leq\; \mu(\Gamma)^2 \cdot \varrho(\Lambda^\star)^2 + \varrho(\Lambda)^2 \cdot \varrho(\Lambda^\star)^2 \\
&\leq \mu(\Gamma)^2 \cdot \varrho(\Gamma^\star)^2 + \varrho(\Lambda)^2 \cdot \varrho(\Lambda^\star)^2 \\
&\leq\; \frac{1}{16}(d-1)^3 + \frac{1}{16}d^2 \\
&\leq\; \frac{1}{16}d^3\,,
\end{aligned}$$

where the second inequality follows from (3.3), the third from Proposition 3.19 and induction. This proves the theorem. □

Remark 3.26. We can refine the notion of the covering minima similarly to the sequence

$\lambda_1, \ldots, \lambda_d$ of successive minima and define the covering minimal μ_j of a convex body K to be

$$\mu_j(\Lambda, K) \ := \ \inf_{\mu>0} \left((\mu K + \Lambda) \cap T \neq \varnothing \text{ for any } j\text{-dimensional subspace } T\right) .$$

Then $\mu(\Lambda) = \mu_d(\Lambda, \mathcal{B}_\mathrm{d})$. For any convex body K we get a centrally symmetric convex body $K_0 := K - K$. let $\lambda_1^\star$ be the first successive minimum of $\Lambda^\star$ *w.r.t.* to the norm defined by $(K-K)^\star$. Kannan and Lovasz[10] show that

$$\mu_1(\Lambda, K) \cdot \lambda_1^\star \ = \ 1 .$$

3.3. Flatness Theorem

Playing around with two-dimensional convex sets the reader may get the impression that a convex body without interior lattice points cannot be arbitrarily wide. Indeed this is a fundamental fact in the geometry of numbers. The following considerations are based on an argument given in Barvinok's book.[11]

Definition 3.27. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice with dual lattice $\Lambda^\star$. Let $K \subset \mathbb{R}^d$ be a full-dimensional convex body. The *width of K with respect to a non-zero lattice vector $a \in \Lambda^\star$* is defined as

$$\mathrm{width}(K; a) := \max_{x \in K} a(x) - \min_{x \in K} a(x).$$

We define the *width* of K with respect to Λ as

$$\mathrm{width}_\Lambda(K) := \inf(\mathrm{width}(K; a) \ : \ a \in \Lambda^\star \setminus \{0\}).$$

You will show in Problem 3.28 that for full-dimensional convex bodies the infimum is actually a minimum, and in Problem 3.27 that the width of convex bodies with dimension less than the ambient dimension is actually 0.

Problem 3.27
Problem 3.28
Problem 3.29

Recall that an *ellipsoid* is the image of a ball (in some norm) under an affine linear map. See Definition A.7 for a full definition and for some properties that we need in the following. Our approach to bound the lattice width of empty convex bodies will proceed in three steps. We first prove it for balls, then extend to ellipsoids and finally use Lemma A.11 to approximate an arbitrary convex body with ellipsoids from the interior and the exterior. The following lemma does the first two steps.

Lemma 3.28. *Let Λ be a lattice, $\mathbf{v} \in \Lambda^\star$ a shortest non-zero lattice vector and E an ellipsoid such that $E \cap \Lambda = \varnothing$. Then* $\mathrm{width}_\mathbf{v}(E) \leq d^{3/2}$.

[10] Ravi Kannan and László Lovász, "Covering minima and lattice-point-free convex bodies", Lemma 2.3.
[11] A. Barvinok, *Integer points in polyhedra*.

Proof. We prove this first for the case that E is a ball. In this case we know by Proposition 3.18 that $\|\mathbf{v}\| = 2\,\varrho(\Lambda^\star)$. Let r be the radius of the ball. Then $r \leq \mu(\Lambda)$. Now

$$\operatorname{width}(E;\mathbf{v}) \;=\; r\,\|\mathbf{v}\| \;\leq\; 2\,\varrho(\Lambda^\star)\,\mu(\Lambda)\,,$$

and the latter is at most $d^{3/2}$ by Theorem 3.25.

For the extension to ellipsoids we use that the bound $d^{3/2}$ obtained is independent of the lattice. Further, any ellipsoid is a linear image of a ball and the image of a lattice Λ for a non-singular linear map T is a lattice.

More precisely, let $x \mapsto Tx + \mathbf{t}$ be the affine map such that $T(E) = \mathcal{B}$ is a ball, and let $\Lambda' := T(\Lambda)$. Then Λ' is a lattice in $\mathbb{R}^d$ and $\mathcal{B} \cap \Lambda' = \varnothing$. Hence, for a shortest non-zero vector $\mathbf{v}' \in (\Lambda')^\star$, its preimage $T\mathbf{v}$ is a shortest non-zero vector $\mathbf{w} \in \Lambda^\star$ we have

$$\operatorname{width}(E;\mathbf{w}) \;\leq\; \operatorname{width}(E;\mathbf{v}') \;=\; \operatorname{width}(\mathcal{B};\mathbf{v}') \;\leq\; d^{3/2}\,. \qquad \square$$

We can extend our bound for the width of a convex body from balls and ellipsoids to general convex bodies with empty interior, albeit only with a weaker right hand side. The key observation for this is Lemma A.11, which tells us the we can estimate any convex body from the interior and exterior with a suitably chosen ellipsoid.

Theorem 3.29 (Flatness Theorem)**.** *Let $K \subset \mathbb{R}^d$ be a convex body with $K \cap \Lambda = \emptyset$. Then*

$$\operatorname{width}_\Lambda(K) \;\leq\; d^{\frac{5}{2}}\,.$$

Note that also in this general theorem the upper bound only depends on the dimension and not on the given lattice.

Proof. Let E be a maximum volume ellipsoid (see Definition A.9) in K with center $\mathbf{z}$. Then also $E \cap \Lambda = \varnothing$. Let $\mathbf{v}$ be a shortest non-zero lattice vector in Λ such that $\operatorname{width}(E;\mathbf{v}) \leq d^{3/2}$ by the previous Lemma 3.28.

Clearly, the width of K is translation invariant, so we can assume that $\mathbf{z}$ is the origin. By Lemma A.11 we deduce $K \subseteq dE$, and thus

$$\operatorname{width}(K;\mathbf{v}) \;\leq\; d\operatorname{width}(E;\mathbf{v}) \leq d \cdot d^{3/2} \;=\; d^{5/2}\,. \qquad \square$$

The bound from the previous theorem is not optimal. If we look for a function $f(d)$ such that

$$\operatorname{width}_\Lambda(K) \;\leq\; f(d)\,.$$

for all convex K with $K \cap \Lambda = \emptyset$, then we can ask how small we can make $f(d)$. Consider the polytope defined by the inequalities

$$x_i \;\geq\; \mu \qquad\qquad x_1 \;+\; \cdots \;+\; x_d \;\leq\; d - \mu$$

for some small $\mu \geq 0$. This is the unit simplex scaled by $d(1-\mu) - \frac{\mu}{d}$ and then shifted by $\mu\mathbf{1}$. It is lattice free, and its width approaches d for $\mu \to 0$. So $f(d) \geq d$.

On the other hand, it follows from Remark 3.26 and Problem 3.29 that $f(d) = \mathcal{O}(d^2)$.

Problem 3.30

In fact, the bound of the previous theorem can be strengthened to be of order $d^{3/2}$, so that

$$\operatorname{width}_\Lambda(K) \;\leq\; \mathcal{O}\left(d^{\frac{3}{2}}\right).$$

This is a result of Banaszczyk et al.[12] The current best upper bound is $\mathcal{O}(d^{4/3}\log^a d)$ for some integer $a > 0$.[13] It is unknown and an active subject of current research, whether the sharp bound is actually of the form $\mathcal{O}(d)$.

If K is centrally symmetric, then Banaszczyk proved that one can choose $f(d) = \mathcal{O}(d \log d)$.[14] We get the same bound for simplices.[15] For ellipsoids we have $f(d) = \mathcal{O}(d)$.[16]

From the algorithmic point of view we should also ask whether we can actually compute such a vector $\mathbf{c}$ that realizes the lattice width of a given convex body K. It turns out that computing $\mathbf{c}$, though possible, cannot be done in polynomial time. We can, however, compute an approximation in polynomial time, albeit with much weaker bounds on the width they realize. Here, the currently known best bound $\mathcal{O}(2^d)$ that we can achieve has been given by Grötschel et al.[17]

3.4. Lower Bounds on Flatness

There has been less work on lower bounds so far. The unit simplex shows that

$$\operatorname{width}_\Lambda(K) \geq d\,.$$

Sebő constructed empty simplices in dimension d of width $d - 2$.[18] Here, a polytope P is *empty* if the vertices are the only lattice points in P. For 3-dimensional empty lattice polytopes (not just simplices) the width is 1. The maximal widths of empty simplices and polytopes differ in higher dimensions.

The maximum width also differs for convex bodies and lattice polytopes (polytopes whose vertices are lattice points) for dimensions $d \geq 2$ (In dimension 1 both are 1). In dimension 2 we have a maximum width of 2 for a lattice polytope (obtained by twice the unit simplex) and $2 + 2/\sqrt{3}$ obtained by a slightly rotated simplex.

In dimension 3 the maximum on lattice polytopes is 3,[19] while the precise value is

[12] Wojciech Banaszczyk, Litvak, Pajor, and Szarek, "The flatness theorem for nonsymmetric convex bodies via the local theory of Banach spaces", Thm. 2.4.
[13] Rudelson, "Distances between non-symmetric convex bodies and the MM^*-estimate", Cor. 2.
[14] W. Banaszczyk, "Inequalities for convex bodies and polar reciprocal lattices in $\mathbf{R}^n$. II. Application of K-convexity", Eq. (8).
[15] Wojciech Banaszczyk, Litvak, Pajor, and Szarek, "The flatness theorem for nonsymmetric convex bodies via the local theory of Banach spaces".
[16] W. Banaszczyk, "Inequalities for convex bodies and polar reciprocal lattices in $\mathbf{R}^n$. II. Application of K-convexity", Eq. (9).
[17] Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*.
[18] Sebő, "An introduction to empty lattice simplices".
[19] Averkov, Krümpelmann, and Weltge, *Notions of maximality for integral lattice-free polyhedra: the case of*

not known for general convex bodies. We have a lower bound of $2+\sqrt{2}$, which was proved by Codenotti and Santos.[20] They conjecture that this is the true value. They also show that the width of a lattice polytope without interior lattice points can be larger than the dimension. More precisely, they construct a lattice polytope with no interior lattice points of width 15 in dimension 14 and a simplex of with 408 in dimension 404.

3.5. Problems

3.1. Let $K \subseteq \mathbb{R}^d$ be a centrally symmetric convex body with $\operatorname{int}(K) \cap \mathbb{Z}^d = \{0\}$.

- ▷ Assume K is a polytope. Show that at most $2(2^d-1)$ facets contain a lattice point in their relative interior.
- ▷ Assume that $\operatorname{vol}(K) = 2^d \det \Lambda$. Show that K is a polytope and each facet of K contains at least one lattice point in its relative interior.

Hint: For the second part: For Any lattice point $\mathbf{x}$ choose a half space $H_{\mathbf{x}}$ containing both $\mathbf{x}$ and K. Let $S_{\mathbf{x}} := H_{\mathbf{x}} \cap -H_{\mathbf{x}}$. Now consider the intersection of all $S_{\mathbf{x}}$ and prove that this satisfies the assumptions of Minkowski's First Theorem (Corollary 3.3).

3.2. Let $K \subseteq \mathbb{R}^d$ be a centrally symmetric polytope with $\operatorname{int}(K) \cap \mathbb{Z}^d = \{0\}$. Show that K contains at most 3^d lattice points.

This is a result of Minkowski from 1910.

Hint: Choose lattice points in the interior of facets. Consider their coordinates module 3 (*i.e.*, their image under $\mathbb{Z}^d \to (\mathbb{Z}/3\mathbb{Z})^d$). Look at the difference of two points having the same image and use the pigeon-hole principle.

3.3. Let $K \subseteq \mathbb{R}^d$ be a centrally symmetric polytope with $\operatorname{int}(K) \cap \mathbb{Z}^d = \{0\}$. Assume that no lattice point in the boundary is in the convex hull of some other lattice points in K.

Show that K has at most $2^{d+1}-1$ lattice points

3.4. Let K be a centrally symmetric convex body with $\operatorname{vol} K > k \cdot 2^d \det \Lambda$

Show that K contains at least $2k+1$ lattice points.

Can you get k linearly independent points?

Hint: You may start with finding just $k+1$ points.

3.5. Prove that the unit ball in dimension d has volume

$$\mathrm{V_d} := \frac{\pi^{\lfloor d/2 \rfloor} 2^{\lceil d/2 \rceil}}{\prod_{0 \le 2i \le d}(d-2i)}.$$

3.6. Let Λ be a lattice with successive minima $\lambda_1 \le \ldots \le \lambda_d$ and $\mathbf{b}_1, \ldots, \mathbf{b}_d$ linearly independent lattice vectors with $\lambda_i = \|\mathbf{b}_i\|$ for $1 \le i \le d$. Show that these vectors are a lattice basis for $d \le 2$, but not necessarily for $d \ge 5$.

Hint: Consider the lattice $\Lambda := \{\, \mathbf{x} \,:\, x_1 \equiv x_2 \equiv \cdots \equiv x_d \mod 2 \,\}$.

3.7. Let Λ be a lattice generated by $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and $\mu > 0$ such that

$$\|\mathbf{b}_i\| \le \mu \qquad \text{for all } 1 \le i \le d,$$

Then we have, for any $\mathbf{v} \in \mathbb{R}^d$

$$\min_{\mathbf{u} \in \Lambda} \|\mathbf{v} - \mathbf{u}\| \le \frac{\sqrt{d}}{2}\mu$$

dimension three.

[20] Codenotti and Santos, *Hollow polytopes of large width*.

with equality if and only if $\mathbf{b}_1, \ldots, \mathbf{b}_d$ are pairwise orthogonal, $\|\mathbf{b}_i\| = \mu$ for $1 \le i \le d$ and

$$\mathbf{v} \;=\; \sum_{i=1}^{d} \left(\eta_i + \frac{1}{2}\right) \mathbf{b}_i \qquad\qquad \text{for some} \eta_i \in \mathbb{Z}\,, 1 \le i \le d\,.$$

3.8. Show that for $d \le 4$ any linearly independent set of lattice vectors $\mathbf{b}_i$, $1 \le i \le d$ with $\|\mathbf{b}_i\| = \lambda_i$ is a lattice basis.

Hint: Use induction and Problem 3.7.

3.9. Prove Proposition 3.10

3.10. Let Λ be a lattice with successive minima $\lambda_1, \ldots, \lambda_d$ and $\mathbf{v}_1, \ldots, \mathbf{v}_d$ linearly independent lattice vectors such that $\|\mathbf{v}_i\| = \lambda_i$ for $1 \le i \le d$ (This is possible by Proposition 3.9).
Let Γ be the lattice spanned by these vectors and consider the polyhedron

$$C \;:=\; \left\{ \mathbf{x} = \sum \mu_i \mathbf{v}_i \;:\; \sum \left|\frac{\mu_i}{\lambda_i}\right| \;\le\; 1 \right\}$$

Compute the volume of this polyhedron. Use this to prove the lower bound in Minkowski's Second Theorem (Theorem 3.11) and show that it cannot be improved.

3.11. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice such that $\gamma(\Lambda) = \gamma_d$. Show that the successive minima are all equal, *i.e.* $\lambda_1 = \lambda_2 = \cdots = \lambda_d$.

Hint: Use Corollary 3.12.

3.12. Let p be a prime with $p \equiv 1 \mod 4$. Show that there are integers $a, b \in \mathbb{Z}$ such that $p = a^2 + b^2$.

Hint: You may want to use Euler's criteron, which states that for a prime p and coprime a there is q such that $q^2 \equiv a \mod p$ if $a^{\frac{p-1}{2}} \equiv 1 \mod p$.

3.13. Show that every $x \in \mathbb{Z}_{\ge 0}$ can be written as a sum of four squares, *i.e.* for any such x there are $a, b, c, d \in \mathbb{Z}_{\ge 0}$ such that

$$x \;=\; a^2 \;+\; b^2 \;+\; c^2 \;+\; d^2\,.$$

This is the *Theorem of Lagrange*.

Hint: Reduce first to prime x by showing that the product of two sums of four squares can be written as a sum of four squares.

Now show that you can find $\alpha, \beta \in \mathbb{Z}$ such that $\alpha^2 + \beta^2 \equiv -1 \mod x$. For odd x consider the sets $S_\alpha := \{\alpha^2 \mod x \;:\; 0 \le \alpha < \frac{x}{2}\}$ and $S_\beta := \{-1 - \beta^2 \mod c \;:\; 0 \le \beta < \frac{x}{2}\}$ and use the pigeon hole principle to show that their intersection is not empty.

Now consider

$$\Lambda \;:=\; \{\mathbf{a} \in \mathbb{Z}^4 \;:\; a_1 \equiv \alpha \cdot a_3 + \beta \cdot a_4 \mod x \text{ and } a_2 \equiv \beta \cdot a_3 - \alpha \cdot a_4 \mod x\}\,.$$

Show that this is a lattice of index x^2 in $\mathbb{Z}^4$ and apply Minkowski's Theorem to the open ball of radius $\sqrt{2x}$.

3.14. Show that the packing radius is finite and equals $\frac{1}{2}\lambda_1$, which is half of the length of a shortest non-zero lattice vector.

3.15. Show that for $r > 0$

$$|\Lambda \;\cap\; r \cdot \mathcal{B}_\mathrm{d}\,| \;\le\; \left(\frac{2r}{\lambda_1} + 1\right)^d\,.$$

3.16. Let Λ be a lattice in $\mathbb{R}^d$. Show that there is a non-zero $\mathbf{x} \in \Lambda$ such that

$$\|\mathbf{x}\|_\infty \;\le\; (\det\,\Lambda)^{1/d}\,,.$$

3.17. Let $\Lambda_0 \subseteq \Lambda \subseteq \mathbb{R}^d$ be lattices. Show that

$$\varrho(\Lambda) \;\leq\; \varrho(\Lambda_0) \;\leq\; \left|\Lambda/\Lambda_0\right| \varrho(\Lambda)\,.$$

3.18. Show that for $d \geq 2$ and any $\varepsilon > 0$ there is a lattice Λ of rank d such that $\lambda_1 = \lambda_1^\star \leq \varepsilon$.

3.19. Show that for $d \geq 2$ and any $c > 0$ there is a lattice Λ of rank d such that $\lambda_n = \lambda_n^\star \geq c$.

3.20. Show that the covering radius $\mu(\Lambda)$ is finite and attained for some $\mathbf{x} \in \mathbb{R}^d$.

3.21. Prove

(i) $\mu(\mathbb{Z}^d) \;=\; \sqrt{d}/2$
(ii) $\mu(D_3) \;=\; 1$
(iii) $\mu(D_d) \;=\; \sqrt{n}/2$ for $d \geq 4$

3.22. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. Show that

$$\mu(\Lambda) \;\leq\; \frac{\sqrt{d}}{2}\lambda_d\,.$$

Show that this bound is tight.

Hint: Choose appropriate linearly independent vectors and modify the fundamental parallelepiped to a cube of appropiate side lengths (*i.e.* choose an orthogonal basis).

3.23. Show that there is $\mathbf{v} \in \mathbb{R}^d$ with $\|\mathbf{v} - \mathbf{u}\| \geq \frac{1}{2}\lambda_d$ for any $\mathbf{u} \in \Lambda$.

3.24. Show that there are lattices Λ such that $\mu(\Lambda) \leq 2\lambda_1(\Lambda)$. Deduce that the bound in Proposition 3.5 is essentially tight.

Hint: Assume you have a lattice with $\mu(\Lambda) \leq 2\lambda_1(\Lambda)$. Show that there is a lattice point $\mathbf{v} \in \Lambda$ so that $\frac{1}{2}\mathbf{v}$ has distance at least $\lambda_1(\Lambda)$ from lattice. Consider the new lattice $\Lambda \cup \left(\frac{1}{2}\mathbf{v} + \Lambda\right)$.

For the second claim note that the volume of a ball used in a covering must be at least the volume of a parallelepiped.

3.25. Show that $\lambda_1 \cdot \lambda_d^* \geq 1$.

3.26. Show that $\lambda_k \cdot \lambda_{d-k+1}^* \geq 1$.

3.27. Show that the lattice width of a low dimensional convex body is 0.

3.28. Show that in the definition of lattice width we can replace the infimum with a minimum for a full-dimensional convex body K. Thus, the width is strictly positive.

3.29. Any centrally symmetric convex body C defines a norm on $\mathbb{R}^d$ via $\|\mathbf{x}\|_C := r$ if $\frac{1}{r}\mathbf{x} \in \partial C$ for any $\mathbf{x} \neq 0$.

Let K be a convex body. Then $K - K := \{\, \mathbf{x} - \mathbf{y} \;:\; \mathbf{x}, \mathbf{y} \in K \,\}$ is a centrally symmetric convex body, and so is its polar $(K-K)^\star$. We can define the successive minima $\lambda_i(\Lambda, K - K^\star)$ *w.r.t.* to this norm.

Show that $\operatorname{width}_\Lambda(K) \;=\; \lambda_1(\Lambda, K - K^\star)$.

3.30. Let $C_1, \ldots, C_m$ be convex bodies in $\mathbb{R}^{d_i}$ with lattices Λ_i containing the origin and $\lambda_1, \ldots, \lambda_m > 0$. Let

$$C \;:=\; \bigoplus_{i=1}^{m} \lambda_i C_i \;=\; \lambda_1 C_1 \oplus \cdots \oplus \lambda_m C_m\,.$$

be the scaled free sum. Show that C is a lattice polytope if all C_i are lattice polytopes and that

$$\operatorname{width}_\Lambda(C) \;= \min_i \left(\lambda_i \operatorname{width}_{\Lambda_i}(C_i)\right)\,.$$

Further, if $\operatorname{int} C_i \cap \Lambda = \varnothing$ for all i and $\sum \frac{1}{\lambda_i} \geq 1$, then also $\operatorname{int} C \cap \Lambda = \varnothing$.

4. The Shortest Vector Problem

We have seen in the previous chapter that, as a consequence of Minkowski's First Theorem (Corollary 3.3), any lattice Λ of rank d contains a vector $\mathbf{v} \in \Lambda$ with length bounded by

$$\|\mathbf{v}\|_2 \;\leq\; \sqrt{d}\,(\det\,\Lambda)^{1/d}\,,$$

see Proposition 3.5. Note that this bound is only true for the *Euclidean norm*, and in this and the following chapter we will always use this norm and just write $\|.\|$ for $\|.\|_2$.

We have also seen that the proof of Minkowski's Theorem is not constructive and so far we do not have an algorithm to construct a short vector. However, constructing such a vector, or at least an approximation, is one of the fundamental algorithmic problems not only in integer optimization, but also in cryptography, number theory and other fields. It has therefore gained a lot of attraction over recent years.[1]

4.1. Motivation

Here is the precise formulation of the task to find a short vector in a lattice as an optimization problem.

(SVP). Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. The *Shortest Vector Problem (SVP)* is the task to find a non-zero vector $\mathbf{u} \in \Lambda \setminus \{\mathbf{0}\}$ of shortest possible length.

Observe that a shortest non-zero lattice vector has length $\lambda_1(\Lambda)$, the first successive minimum introduced in Definition 3.6. An important relaxation of the shortest vector problem is the *approximate shortest Vector problem (SVP)*$_\gamma$.

(SVP)$_\gamma$**.** Let $\gamma \geq 1$ and $\Lambda \subseteq \mathbb{R}^d$ be a lattice. The *Approximate Shortest Vector Problem (SVP)*$_\gamma$ is the task to find a non-zero vector $\mathbf{u} \in \Lambda \setminus \{\mathbf{0}\}$ of length bounded by

$$\|\mathbf{u}\| \;\leq\; \gamma\lambda_1\,.$$

Both the exact problem and its approximate version are NP-hard in general.[2] If we want to compute something in polynomial time we need to relax the problem further, fix some parameters, allow γ to depend on the dimension, or restrict to special cases.

[1] Aardal, *Lattice basis reduction and Integer programming*; Hanrot, Pujol, and Stehlé, "Algorithms for the shortest and closest lattice vector problems"; Nguyen and Vallée, *The LLL Algorithm*; Tateiwa, Shinano, Yamamura, Yoshida, Kaji, Yasuda, and Fujisawa, *CMAP-LAP: Configurable Massively Parallel Solverfor Lattice Problems*; Wübben, Seethaler, Jaldén, and Matz, "Lattice Reduction".

[2] Khot, "Hardness of approximating the shortest vector problem in lattices".

The algorithm for **(SVP)** that we will discuss below is based on a solution for the approximate problem using a *reduced basis* for the lattice. In such a basis the coefficients of a shortest vector are bounded by a constant depending on the dimension only. Hence, we can solve the shortest vector problem by enumerating over all possible coefficients, at least in fixed dimension.

In this chapter we will introduce the *LLL-reduced bases* of Arjen Lenstra, Hendrik Lenstra and László Lovász[3] and look at the consequences for **(SVP)**. We will construct such bases in the next chapter and show how to compute them in polynomial time, at least for sublattices of $\mathbb{Z}^d$.

Let us first discuss why the computation of a short vector, or a basis with short vectors is more involved for lattice bases than for vector space bases. We know from Proposition 3.5 that there is a vector of length at most $\sqrt{d}\,(\det \Lambda)^{1/d}$, but the proof was non-constructive. This is not just an artefact of the proof, as in an arbitrary lattice basis the coefficients of a shortest vector in this basis can be arbitrarily large. To see this, consider *e.g.* the basis

$$\mathbf{b}_1 := \begin{pmatrix} n \\ 1 \end{pmatrix} \qquad \mathbf{b}_2 := \begin{pmatrix} n^2 - 1 \\ n \end{pmatrix}$$

of $\mathbb{Z}^2$ for any integer $n \in \mathbb{Z}$. Then

$$\pm\mathbf{e}_1 = \pm n\mathbf{b}_1 \pm \mathbf{b}_2 \qquad \pm\mathbf{e}_2 = \pm(1-n^2)\mathbf{b}_1 \pm n\mathbf{b}_2\,.$$

are representations of the shortest vectors in the lattice $\mathbb{Z}^2$. Hence, brute force enumeration of coefficients is not sufficient to solve the problem without more information on the basis. Note that one of the basis vectors is much longer than the other, and they are far from orthogonal. We will later see that this is necessarily so for *bad* bases.

On the other hand, assume that we have a lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ of pairwise orthogonal vectors, and assume $\|\mathbf{b}_1\| \leq \|\mathbf{b}_j\|$ for $1 \leq j \leq d$. Any non-zero shortest vector $\mathbf{x} \in \Lambda$ has a representation in this basis with integral coefficients $\lambda_1, \ldots, \lambda_d \in \mathbb{Z}$, and not all can be 0. Let j be an index with $\lambda_j > 0$. Then

$$\|\mathbf{x}\|^2 = \left\|\sum \lambda_i \mathbf{b}_i\right\|^2 = \sum |\lambda_i|^2 \|\mathbf{b}_i\|^2 \geq \|\mathbf{b}_j\|^2 \geq \|\mathbf{b}_1\|^2\,.$$

Hence, the shortest of the lattice basis vectors is already a non-zero lattice vector of shortest length.

So it might be desirable to find a lattice basis with orthogonal vectors. From plain linear algebra we know how to find such a basis using the *Gram-Schmidt orthogonalization*. This is a simple and efficient algorithm that runs in polynomial time. However, already simple examples in dimension 2 show that it does not respect the lattice structure.

Furthermore, the hexagonal lattice of Figure 4.1 shows in fact that, in general, a basis with pairwise orthogonal vectors does not exist. We have also seen bases with vectors of necessarily different lengths.

If we do not want to discard the idea of orthogonal bases completely we should check whether we can at least get *close* to an orthogonal basis (in a sense we still have to define), and whether such an almost orthogonal basis is still good enough to solve

[3] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".

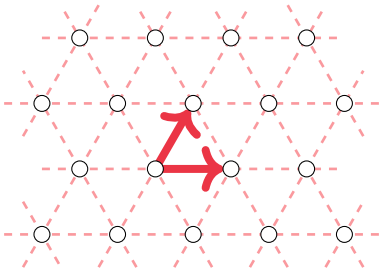


Figure 4.1.: The hexagonal lattice

(SVP). We also want a polynomial time algorithm for this, so we may also have to discuss the relation between a better approximation and one that we can still compute efficiently.

4.2. Reduced Bases

So in the following we want to revisit the Gram-Schmidt orthogonalization and see what we can learn from it for lattice bases. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice with a basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d \in \mathbb{R}^d$. The order of the basis vectors is relevant in the orthogonalization process, and we will sometimes speak of *ordered bases* to emphasize this.

We consider the increasing chain of Λ-rational subspaces

$$V_0 := \{\mathbf{0}\} \qquad V_k := \operatorname{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_k) \qquad \text{for } 1 \leq k \leq d\,. \tag{4.1}$$

together with the induced lattices $\Lambda_k := \Lambda \cap V_k$ on these spaces. For $0 \leq k \leq d$ let $\pi_k : \mathbb{R}^d \to V_k$ be the orthogonal projection onto the subspace V_k. The Gram-Schmidt-orthogonalization of the basis are the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_d$ defined via

$$\begin{aligned} \mathbf{w}_k &:= \mathbf{b}_k - \pi_{k-1}(\mathbf{b}_k) \\ &= \mathbf{b}_k - \sum_{j=1}^{k-1} \lambda_{jk} \mathbf{w}_j \qquad \text{with} \quad \lambda_{jk} := \frac{\langle\, \mathbf{b}_k, \mathbf{w}_j \,\rangle}{\|\mathbf{w}_j\|^2}\,. \end{aligned}$$

So in particular we have

$$\langle\, \mathbf{w}_i, \mathbf{w}_j \,\rangle = 0 \quad \text{for } 1 \leq i < j \leq d \qquad \text{and} \qquad \mathrm{d}(\mathbf{b}_k, V_{k-1}) = \|\mathbf{w}_k\|\,. \tag{4.2}$$

Example 4.1. Let $\mathbf{b}_1 := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{b}_2 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. This is a basis of $\mathbb{R}^2$ with Gram-Schmidt-orthogonalization

$$\mathbf{w}_1 = \mathbf{b}_1 \qquad \text{and} \qquad \mathbf{w}_2 = \begin{pmatrix} -3/5 \\ 6/5 \end{pmatrix} = -4/5\mathbf{b}_1 + \mathbf{b}_2\,,$$

see Figure 4.2.

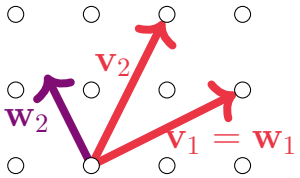


Figure 4.2.: The lattice basis of Example 4.1

We can write the original lattice basis in terms of the Gram-Schmidt basis as

$$\mathbf{b}_k = \mathbf{w}_k + \sum_{j=1}^{k-1} \lambda_{jk} \mathbf{w}_j \quad \text{for } 1 \leq k \leq d\,. \tag{4.3}$$

The vectors $\mathbf{w}_1, \ldots, \mathbf{w}_d$ are pairwise orthogonal, so

$$\begin{aligned} \|\mathbf{w}_k\| \;&\leq \|\mathbf{b}_k\|\,, \\ \det \Lambda = \prod_{j=1}^{d} \|\mathbf{w}_j\|\,. \qquad &\text{and} \qquad \det \Lambda_k = \prod_{j=1}^{k} \|\mathbf{w}_j\|\,. \end{aligned} \tag{4.4}$$

We can use these equations to measure how far our given lattice basis is from orthogonality. We introduce the following invariant for this.

Definition 4.2. The *orthogonality defect of* the lattice basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d$ is

$$M(\mathbf{b}_1, \ldots, \mathbf{b}_d) \;:=\; \frac{1}{\det \Lambda} \prod_{j=1}^{d} \|\mathbf{b}_j\|\,.$$

Observe that the basis vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d$ are pairwise orthogonal if and only if $M(\mathbf{b}_1, \ldots, \mathbf{b}_d) = 1$. In all other cases M_Λ is strictly larger than 1.

It will follow from our considerations that there is in fact a universal bound M_d depending on the dimension only such that any lattice Λ has a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ with

$$M(\mathbf{b}_1, \ldots, \mathbf{b}_d) \;\leq\; M_d\,.$$

We say that a lattice basis is *reduced* if it satisfies this bound.

The Gram-Schmidt vectors give a lower bound for the length of a shortest vector of the lattice.

Theorem 4.3. *Let Λ be a d-dimensional lattice with basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d$ and Gram-Schmidt-orthogonalization $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_d$. Then*

$$\|\mathbf{u}\| \;\geq\; \min(\|\mathbf{w}_1\|, \ldots, \|\mathbf{w}_d\|)$$

for all $\mathbf{u} \in \Lambda \setminus \{0\}$.

Using the first successive minimum λ_1 we can rewrite this as

$$\lambda_1 \;\geq\; \min(\|\mathbf{w}_1\|, \ldots, \|\mathbf{w}_d\|)$$

Problem 4.2
Problem 4.3

Proof. We can write $\mathbf{u}$ as a linear combination

$$\mathbf{u} \;=\; \sum_{j=1}^{d} \eta_j \mathbf{w}_j \,.$$

Let k be the highest index such that $\eta_k \neq 0$. We can rewrite

$$\mathbf{u} \;=\; \eta_k \mathbf{w}_k + \sum_{j=1}^{k-1} \eta_j \mathbf{w}_j$$

By orthogonality, all lattice points lie in lattice hyperplanes parallel to V_{k-1} at distance $r\,\|\mathbf{w}_k\|$ apart, for some $r \in \mathbb{Z}$. Hence, $|\eta_k| \geq 1$, as $\eta_k \neq 0$. So

$$\|\mathbf{u}\| \;\geq\; |\eta_k|\,\|\mathbf{w}_k\| \;\geq\; \|\mathbf{w}_k\| \,. \qquad \square$$

Problem 4.4

In Corollary 3.12 we deduced that

$$\left(\prod_{i=1}^{d} \lambda_i\right)^{1/d} \leq \sqrt{d}(\det \Lambda)^{1/d} \,.$$

for the successive minima *w.r.t.* the Euclidean norm from Theorem 3.11 by specializing the norm to the Euclidean norm. With Problem 4.5 you can use the Gram-Schmidt orthogonalization and the inscribed ellipsoids that we already used for the Flatness Theorem (Theorem 3.29) to give an independent proof of this bound.

Problem 4.5

The previous theorem suggests that a comparison of a lattice basis to its Gram-Schmidt orthogonalization might be useful to obtain information on the length of short vectors. We define a special version of a *reduced bases*.

In the following, let $\mathbf{b}_1, \ldots, \mathbf{b}_d \subseteq \mathbb{R}^d$ be any ordered lattice basis of our lattice Λ, with Gram-Schmidt orthogonalization $\mathbf{w}_1, \ldots, \mathbf{w}_d \in \mathbb{R}^d$ and coefficients λ_{jk} for $1 \leq k \leq d$, $1 \leq j \leq k-1$ as in (4.3).

Definition 4.4. The basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ is *LLL-reduced* if

$$|\lambda_{jk}| \;\leq\; \tfrac{1}{2} \qquad \text{for all} \qquad 1 \;\leq\; j \;<\; k \;\leq d \qquad (4.5)$$

$$\tfrac{3}{4}\,\|\mathbf{w}_k\|^2 \;\leq\; \|\lambda_{k,k+1}\mathbf{w}_k + \mathbf{w}_{k+1}\|^2 \qquad \text{for all} \qquad 1 \;\leq\; k \;\leq\; d-1\,. \qquad (4.6)$$

We say that a coefficient λ_{jk} is *weakly reduced* if it satisfies condition (4.5) of the definition. A lattice basis is *weakly reduced* if all coefficients are weakly reduced.

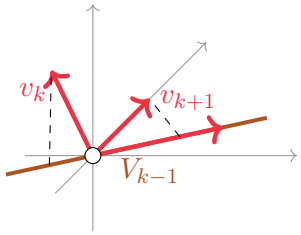


Figure 4.3.: The second condition for LLL

Using (4.2) we can rewrite the second condition equivalently form.

$$\frac{3}{4}\,\mathrm{d}(\mathbf{b}_k, V_{k-1})^2 \;\leq\; \mathrm{d}(\mathbf{b}_{k+1}, V_{k-1})^2\,,$$

for the subspaces V_j defined in (4.1). Geometrically, a basis is LLL-reduced if the vector $\mathbf{b}_{k+1}$ is not much closer to the subspace spanned by the first $k-1$ basis vectors than the vector $\mathbf{b}_k$, see Figure 4.3.

More generally, we can define *δ-reduced* bases for some $1/4 < \delta < 1$ by replacing the second condition with

$$\delta\,\|\mathbf{w}_k\|^2 \;\leq\; \|\lambda_{k,k+1}\mathbf{w}_k + \mathbf{w}_{k+1}\|^2\,.$$

The construction of the basis, that we will do below, works also for this more general definition, but the application to shortest vectors requires the specialization to $\delta = \frac{3}{4}$, so we do this right from the beginning.

Example 4.5. Here is an example of an LLL-reduced basis and one that is not. Both are illustrated in Figure 4.4. Consider the three vectors

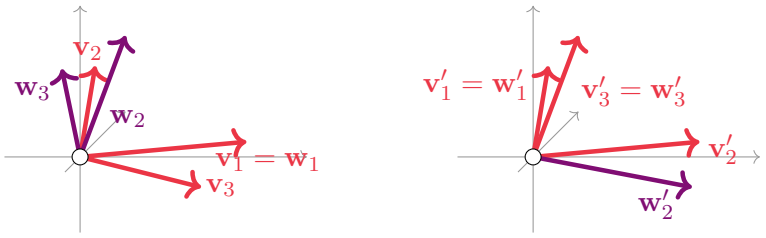
$$\mathbf{v}_1 := \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2 := \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 := \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$$

Their Gram-Schmidt-orthogonalization is

$$\mathbf{w}_1 := \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{w}_2 := \begin{bmatrix} -2/5 \\ 4/5 \\ 1 \end{bmatrix} = \mathbf{v}_2 - \frac{2}{5}\mathbf{w}_1$$

$$\mathbf{w}_3 := \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \mathbf{v}_3 - \frac{1}{5}\mathbf{w}_1 - \mathbf{w}_2$$

Then λ_{13} violates condition (4.5) of the definition, and $\mathbf{v}_1$ and $\mathbf{v}_2$ violate condition (4.6), as $\mathrm{d}(\mathbf{v}_1, V_0)^2 = 5$ and $\mathrm{d}(\mathbf{v}_1, V_0)^2 = 9/5$ but $3/4 \cdot 5 > 9/5$.



(a) Not a reduced basis (b) A reduced basis

Figure 4.4.: A lattice basis that is not $3/4$-reduced and one for the same lattice that is reduced.

However, the basis

$$\mathbf{v}'_1 := \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}'_2 := \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}'_3 := \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix},$$

that spans the same lattice, is LLL-reduced. It has Gram-Schmidt-orthogonalization

$$\mathbf{w}'_1 := \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{w}'_2 := \begin{bmatrix} 2 \\ 1/2 \\ -1/2 \end{bmatrix} = \mathbf{v}'_2 - \frac{1}{2}\mathbf{w}'_1 \quad \mathbf{w}'_3 := \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \mathbf{v}'_3 .$$

Using $\langle \mathbf{w}_k, \mathbf{w}_{k+1} \rangle = 0$ we can expand the right hand side of (4.6) and rewrite this condition as

$$\|\mathbf{w}_{k+1}\|^2 \geq \left(\frac{3}{4} - \lambda_{k,k+1}^2\right) \|\mathbf{w}_k\|^2 . \tag{4.7}$$

This implies that [4]

$$\|\mathbf{w}_k\|^2 \leq 2\,\|\mathbf{w}_{k+1}\|^2 \qquad \text{for all} \qquad 1 \leq k \leq d-1 . \tag{4.8}$$

So, while it is allowed in the definition that the length of the $\mathbf{w}_k$ may decrease with increasing k, it may not drop by too much.

Proposition 4.6. *Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be an LLL-reduced basis of Λ with Gram-Schmidt orthogonalization $\mathbf{w}_1, \ldots, \mathbf{w}_d$. Then*

$$\|\mathbf{b}_1\| \leq 2^{(d-1)/2} \lambda_1 \tag{4.9}$$
$$\|\mathbf{b}_1\| \leq 2^{(d-1)/4} (\det \Lambda)^{1/d} \tag{4.10}$$

Proof. By (4.8) we know $\|\mathbf{w}_j\|^2 \leq 2\,\|\mathbf{w}_{j+1}\|^2$, so by induction

$$\|\mathbf{w}_j\|^2 \leq 2^{k-j}\,\|\mathbf{w}_k\|^2 \qquad \text{for } 1 \leq j < k \leq d . \tag{4.11}$$

[4] We could also use this as the defining condition. It implies δ-reduced for for $\delta = \frac{1}{2}$.

Hence, we obtain for all $1 \le j \le d$

$$\|\mathbf{b}_1\|^2 \;=\; \|\mathbf{w}_1\|^2 \;\le 2^{j-1}\,\|\mathbf{w}_j\|^2 \;\le\; 2^{d-1}\,\|\mathbf{w}_j\|^2\,, \tag{4.12}$$

so

$$\|\mathbf{b}_1\|^2 \;\le\; 2^{d-1}\min_j \|\mathbf{w}_j\|^2$$

and by Theorem 4.3

$$\|\mathbf{b}_1\| \;\le\; 2^{\frac{d-1}{2}}\lambda_1\,.$$

This proves the first bound. Taking the product of (4.12) for all j gives

$$\begin{aligned}\|\mathbf{b}_1\|^{2d} \;&\le\; \prod_{i=1}^{d} 2^{i-1}\,\|\mathbf{w}_i\|^2 \\ &=\; 2^{\frac{d(d-1)}{2}}\,\|\mathbf{w}_1\|^2\cdots\|\mathbf{w}_d\|^2 \;=\; 2^{\frac{d(d-1)}{2}}\,(\det\Lambda)^2\,.\end{aligned}$$

□

Problem 4.7

Corollary 4.7. *Let $\Lambda \subset \mathbb{R}^d$ be a lattice with LLL-reduced basis $\mathbf{b}_1,\dots,\mathbf{b}_d$. Then*

$$\det\,\Lambda \;\le\; \prod_{i=1}^{d}\|\mathbf{b}_i\| \;\le\; 2^{\frac{1}{2}\binom{d}{2}}\det\Lambda\,.$$

The lower bound in this corollary is also known as the *Hadamard Inequality*.

Proof. We have seen the lower bound already in the definition of the orthogonality defect. For the upper we compute

$$\|\mathbf{b}_j\|^2 \;\le\; \|\mathbf{w}_j\|^2 + \frac{1}{4}\sum_{k=1}^{j-1}\|\mathbf{w}_k\|^2 \;\le\; \|\mathbf{w}_j\|^2\left(1+\frac{1}{4}\sum_{k=1}^{j-1}2^{j-k}\right) \;\le\; 2^{j-1}\,\|\mathbf{w}_j\|^2$$

and

$$\prod_{j=1}^{d}\|\mathbf{b}_j\| \;\le\; \prod_{j=1}^{d}2^{\frac{1}{2}(j-1)}\,\|\mathbf{w}_j\| \;=\; 2^{\frac{1}{2}\binom{d}{2}}\prod_{j=1}^{d}\|\mathbf{w}_j\| \;\le\; 2^{\frac{1}{2}\binom{d}{2}}\det\Lambda\,.$$

□

Corollary 4.8. *The orthogonality defect M of an LLL-reduced basis is at most $2^{\frac{1}{2}\binom{d}{2}}$.* □

Problem 4.8
Problem 4.9

4.3. Short Vectors

The first successive minimum in the previous proposition is precisely the length of any shortest nonzero vector $\mathbf{x}$ in the lattice. Hence, (4.9) in Proposition 4.6 shows that

any LLL-reduced basis solves the approximate shortest vector problem **(SVP)**$_\gamma$ for the constant $\gamma = 2^{(d-1)/2}$. We just take the first basis vector in the reduced basis as our approximate shortest vector.

Corollary 4.9. *Let Λ be a lattice and λ_1 be the length of a shortest nonzero vector in Λ The first basis vector $\mathbf{b}_1$ in an ordered LLL-reduced lattice basis satisfies*

$$\|\mathbf{b}_1\| \;\leq\; \gamma\lambda_1$$

for $\gamma = 2^{(d-1)/2}$. □

We have started our discussion of the shortest vector problem with the bound obtained from Minkowski's First Theorem (Corollary 3.3) in Proposition 3.5

$$\|\mathbf{v}\|_2 \;\leq\; \sqrt{d}\,(\det\,\Lambda)^{1/d}\,,$$

The solution of the approximate shortest vector problem **(SVP)**$_\gamma$ in Corollary 4.9 bounds the norm of the first basis vector in terms of the shortest vector directly. So apparently we have not used the bound via the determinant.

However, this is just hidden by our special choice of a reduced basis. Let $\mathbf{b}_k$ be the shortest vector in an LLL-reduced basis. Then

$$\|\mathbf{b}_k\|^d \;\leq \prod_{i=1}^{d}\|\mathbf{b}_i\| \;=\; \det\,\Lambda\cdot\frac{1}{\det\,\Lambda}\prod_{i=1}^{d}\|\mathbf{b}_i\| \;=\; \det\,\Lambda\;\cdot\;M \;\leq\; 2^{\frac{1}{2}\binom{d}{2}}\det\,\Lambda\,, \quad (4.13)$$

where we have used the bound of Corollary 4.8. Given any approximation of the bound from Minkowski's Theorem in the form of a vector $\mathbf{v}$ that satisfies

$$\|\mathbf{v}\| \;\leq\; f(d)\,(\det\,\Lambda)^{1/d} \quad (4.14)$$

for some constant $f(d)$ depending only on the dimension we prove with the next theorem that we can use this to solve **(SVP)**$_\gamma$.

Note that a lattice point $\mathbf{v}$ bounded as in (4.14) will in general not be sufficient, as the bound depends on the determinant, which may be large. The main idea in the following proof will be that we can apply the same bound in the *dual* lattice. As the product of the determinants is 1, at least one of the bounds must be small, and we can exploit this to construct a short vector.

Problem 4.10

Theorem 4.10. *Let $f : \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0}$ be a nondecreasing function and assume that we have a polynomial time algorithm that finds $\mathbf{v} \in \Lambda \setminus \{0\}$ with*

$$\|\mathbf{v}\| \;\leq\; f(d)\,(\det\,\Lambda)^{1/d}\,.$$

Then there is a polynomial time algorithm for **(SVP)**$_\gamma$ *with $\gamma = f(d)^2$.*

Proof. Let $\overline{\mathbf{v}} \in \Lambda \setminus \{0\}$ be a shortest vector in the lattice. With the algorithm given by

assumption we can also find a lattice vector $\mathbf{a} \neq 0$ in the dual lattice $\Lambda^\star$ such that

$$\|\mathbf{a}\| \;\leq\; f(d)\,(\det\,\Lambda^\star)^{1/d}\,. \tag{4.15}$$

By definition of the dual lattice we know that $\langle\,\mathbf{a},\mathbf{x}\,\rangle \in \mathbb{Z}$ for all $\mathbf{x} \in \Lambda$, and thus any two hyperplanes

$$H_k \;:=\; \{\,\mathbf{x} \;:\; \langle\,\mathbf{a},\mathbf{x}\,\rangle \;=\; k\,\}$$

for different k are at least $\frac{1}{\|\mathbf{a}\|}$ apart.

We distinguish two cases, depending on the unknown vector $\overline{\mathbf{v}}$ and show, that in both cases we obtain a nonzero lattice point whose norm is bounded by $\gamma\lambda_1$.

Assume first that $\langle\,\mathbf{a},\overline{\mathbf{v}}\,\rangle = k \neq 0$, We can give a lower bound on the norm of $\overline{\mathbf{v}}$ via

$$\|\overline{\mathbf{v}}\| \;\geq\; \frac{1}{\|\mathbf{a}\|} \;\geq\; \frac{1}{f(d)}\,(\det\,\Lambda^\star)^{-1/d} \;\Longrightarrow\; f(d)\,(\det\,\Lambda^\star)^{1/d}\,\|\overline{\mathbf{v}}\| \geq 1\,.$$

Now we use the assumed algorithm to find a lattice point $\mathbf{v} \in \Lambda \setminus \{0\}$ bounded as in (4.14). Together with the lower bound on $\|\overline{\mathbf{v}}\|$ we compute

$$\|\mathbf{v}\| \;\leq\; f(d)\,(\det\,\Lambda)^{1/d} \;\leq\; f(d)^2\,(\det\,\Lambda)^{1/d}\,(\det\,\Lambda^\star)^{1/d}\,\|\overline{\mathbf{v}}\| \;=\; f(d)^2\,\|\overline{\mathbf{v}}\|\,,$$

which proves the desired bound.

If $\langle\,\mathbf{a},\overline{\mathbf{v}}\,\rangle = 0$, then $\overline{\mathbf{v}}$ is in the sublattice

$$\Gamma \;:=\; \{\,\mathbf{x} \in \Lambda \;:\; \langle\,\mathbf{a},\mathbf{x}\,\rangle \;=\; 0\,\}\,.$$

This lattice has rank $d-1$. We restrict to this lattice and use induction. Note that by Problem 4.9 we can find a lattice basis of Γ in polynomial time. However, the determinant of Γ may be much larger than the one of Λ.

Further, if $d = 1$, then it is easy to find a short lattice vector (any basis vector is a shortest vector).

By induction we can find $\mathbf{v} \in \Gamma \setminus \{0\}$ such that

$$\|\mathbf{v}\| \;\leq\; f(d-1)^2\lambda_1(\Gamma) \;\leq\; f(d)^2\lambda_1(\Lambda)\,,$$

where the last inequality follows from $f(d-1) \leq f(d)$, as we assumed that f is monotonously increasing. Hence, also in this case we have found a short vector.

As we don't know $\overline{\mathbf{v}}$ we don't know in which case we are. So in our algorithm we just compute the shortest vector in both cases and return the shorter one. □

Note that by (4.13) we can use

$$f(d) \;:=\; \left(2^{\frac{1}{2}\binom{d}{2}}\right)^{1/d} \;=\; 2^{\frac{1}{2}\frac{d-1}{2}}$$

which leads to the same $\gamma = 2^{(d-1)/2}$ as before.

We can also use reduced bases to solve the exact shortest vector problem **(SVP)** in polynomial time in fixed dimension. Namely, with the following theorem we will prove

that we can bound the size of the coefficients of a short lattice vector in the representation in a reduced basis. This then allows us to enumerate all possible candidates and pick the shortest.

Theorem 4.11. *Let Λ be a lattice with reduced basis $\mathbf{b}_1, \dots, \mathbf{b}_d$ and let $\mathbf{u} \in \Lambda \backslash \{0\}$ be a shortest non-zero lattice vector. Then*

$$\mathbf{u} \;=\; \sum_{j=1}^{d} \lambda_j \mathbf{b}_j \qquad \textit{with} \qquad |\lambda_j| \;\leq\; \sqrt{d} M \quad \textit{for } 1 \leq j \leq d.$$

Proof. Let $\mathbf{b}_1$ be the shortest vector among $\mathbf{b}_1, \dots, \mathbf{b}_d$, and let $\mathrm{B} = (\mathbf{b}_1, \dots, \mathbf{b}_d)$. Then $\mathbf{u} = \mathrm{B}\,\lambda$ for $\lambda = (\lambda_1, \dots, \lambda_d)$. Hence, $\lambda = \mathrm{B}^{-1}\,\mathbf{u}$.

By Cramer's rule all entries of B^{-1} are determinants of $(d-1) \times (d-1)$-minors of B, divided by $\det \mathrm{B}$. So each entry of B^{-1} is bounded by

$$\|\mathbf{b}_2\| \cdot \ldots \cdot \|\mathbf{b}_d\| \cdot \frac{1}{\det \mathrm{B}} \;\leq\; \frac{M}{\|\mathbf{b}_1\|}.$$

So

$$|\lambda_j| \leq \sum |u_i| \frac{M}{\|\mathbf{b}_1\|} \;\leq\; \sqrt{d}\,\|\mathbf{u}\|\,\frac{M}{\|\mathbf{b}_1\|} \;\leq\; \sqrt{d} M\,,$$

where the second inequality uses the Cauchy-Schwartz inequality for the scalar product $\langle\, \mathbf{1}, |\mathbf{u}| \,\rangle$, where $|\mathbf{u}|$ is the vector of absolute values of the entries of $\mathbf{u}$, and the last inequality uses $\|\mathbf{u}\| \leq \|\mathbf{b}_1\|$. □

This provides a simple enumeration algorithm to compute the shortest vector, once we have a reduced basis. The next chapter will show that we can compute one in polynomial time in the input size of the original basis an the dimension. This gives the following corollary.

Corollary 4.12. *Let $\Lambda \subseteq \mathbb{R}^d$ be a d-dimensional lattice. Then we can solve **(SVP)** in time $2^{\mathcal{O}(d^3)}$.* □

However, this is a rather crude bound. With an improved enumeration scheme we can do better. For this we use the fact that each subset $\mathbf{b}_1, \dots, \mathbf{b}_k$ for some $1 \leq k \leq d$ is a reduced basis in Λ_k with corresponding Gram-Schmidt basis $\mathbf{w}_1, \dots, \mathbf{w}_k$.

Now consider the sets

$$S_k(\mathbf{c}, r) \;:=\; \mathcal{B}_r(\mathbf{c}) \cap \Lambda_k \;\subseteq\; \Lambda_k$$

of lattice points inside the ball of radius r around $\mathbf{c}$. We want to show that we can enumerate $S_{k+1}(\mathbf{c}, r)$ for a given lattice point $\mathbf{c}$ and a radius r if we know how to enumerate the sets $S_k(\mathbf{c}', r')$ for parameters $\mathbf{c}'$ and r'.

Lemma 4.13. *Let $2 \le k \le d$, $\mathbf{c} \in V_k$, and $r > 0$.*

Let $f_k(r)$ be an upper bound on the number of steps needed to enumerate a set $S_{k-1}(\mathbf{c}', r)$ for a given $\mathbf{c}' \in V_{k-1}$.

Then we can enumerate $S_k(\mathbf{c}, r)$ with at most $\frac{2r}{\|\mathbf{w}_k\|} \cdot f_k(r)$ steps.

Proof. We will actually enumerate a set $\tilde{S}_k(\mathbf{c}, r) \supseteq S_k(\mathbf{c}, r)$. Let $\mathbf{v} \in \mathcal{B}_r(\mathbf{c}) \cap \Lambda_k$. Then $\mathbf{v}$ is a lattice point and $\|\mathbf{v} - \mathbf{c}\| \le r$. We can write

$$\mathbf{v} - \mathbf{c} \;:=\; \sum_{i=1}^{k} \eta_i \mathbf{w}_i \qquad \text{and} \qquad \mathbf{v} \;:=\; \mu \mathbf{b}_k + \mathbf{v}' \tag{4.16}$$

for some $\eta_i \in \mathbb{R}$, $\mathbf{v}' \in \Lambda_{k-1}$ and $\mu \in \mathbb{Z}$. We want to determine how large $|\mu|$ can be.

We compute

$$\begin{aligned} \eta_k \|\mathbf{w}_k\| \;&=\; \frac{1}{\|\mathbf{w}_k\|} \langle\, \mathbf{v} - \mathbf{c}, \mathbf{w}_k \,\rangle \;=\; \frac{1}{\|\mathbf{w}_k\|} \langle\, \mu \mathbf{b}_k + \mathbf{v}' - \mathbf{c}, \mathbf{w}_k \,\rangle \\ &= \frac{1}{\|\mathbf{w}_k\|} \langle\, \mu \mathbf{b}_k - \mathbf{c}, \mathbf{w}_k \,\rangle \\ &= \mu \|\mathbf{w}_k\| \;+\; \frac{1}{\|\mathbf{w}_k\|} \langle\, \mathbf{c}, \mathbf{w}_k \,\rangle \,. \end{aligned}$$

Hence

$$\begin{aligned} \left(\mu \|\mathbf{w}_k\| + \frac{\langle\, \mathbf{c}, \mathbf{w}_k \,\rangle}{\|\mathbf{w}_k\|} \right)^2 \;&=\; \eta_k^2 \|\mathbf{w}_k\|^2 \\ &\le \sum_{i=1}^{d} \eta_i^2 \|\mathbf{w}_i\|^2 \;=\; \|\mathbf{v} - \mathbf{c}\|^2 \qquad (4.17) \\ &\le\; r^2 \,, \end{aligned}$$

where the equation in (4.17) follows as the $\mathbf{w}_j$ are pairwise orthogonal. Hence, we need to consider μ for

$$-\frac{r}{\|\mathbf{w}_k\|} \;-\; \frac{\langle\, \mathbf{c}, \mathbf{w}_k \,\rangle}{\|\mathbf{w}_k\|^2} \;\le\; \mu \;\le\; \frac{r}{\|\mathbf{w}_k\|} \;-\; \frac{\langle\, \mathbf{c}, \mathbf{w}_k \,\rangle}{\|\mathbf{w}_k\|^2} \tag{4.18}$$

There are at most $\frac{2r}{\|\mathbf{w}_k\|}$ values for μ.

If we now show that $\mathbf{v}' \in S_{k-1}(\pi_{k-1}(\mathbf{c} - \mu \mathbf{b}_k), r)$, then we obtain $\tilde{S}_k(\mathbf{c}, r)$ by enumerating $\mathbf{v} = \mathbf{v}' + \mu \mathbf{b}_k$ for $\mathbf{v}' \in S_{k-1}(\pi_{k-1}(\mathbf{c} - \mu \mathbf{b}_k), r)$ and μ in the range (4.18).

We set

$$\mathbf{v}' - \pi_{k-1}(\mathbf{c} - \mu \mathbf{b}_k) \;=\; \sum_{i=1}^{k-1} \zeta_i \mathbf{w}_i \,.$$

Then

$$\begin{aligned}\langle\, \mathbf{v}' - \pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k), \mathbf{w}_i \,\rangle &= \langle\, \mathbf{v} - \mu\mathbf{b}_k, \mathbf{w}_i \,\rangle - \langle\, \pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k), \mathbf{w}_i \,\rangle \\ &= \langle\, \mathbf{v} - \mu\mathbf{b}_k, \mathbf{w}_i \,\rangle - \langle\, \mathbf{c} - \mu\mathbf{b}_k, \mathbf{w}_i \,\rangle \\ &= \langle\, \mathbf{v} - \mathbf{c}, \mathbf{w}_i \,\rangle\,,\end{aligned}$$

so we must have $\zeta_i = \eta_i$ for the coefficients η_i ,$1 \leq i \leq k-1$, defined in (4.16). Hence,

$$\left\|\mathbf{v}' - \pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k)\right\|^2 \;=\; \sum_{i=1}^{k-1} \zeta_i^2 \left\|\mathbf{w}_i\right\|^2 \;=\; \sum_{i=1}^{k} \eta_i^2 \left\|\mathbf{w}_i\right\|^2 \;=\; \left\|\mathbf{v} - \mathbf{c}\right\|^2 \;\leq\; r^2\,.$$

This implies $\mathbf{v}' \in \mathcal{B}_r(\pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k))$, so $\mathbf{v}' \in S_{k-1}(\pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k), r)$.

By assumption, we need $f_k(r)$ steps to enumerate $S_{k-1}(\pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k), r)$, and we need to do this for all μ that satisfy (4.18), which are at most $\frac{2r}{\|\mathbf{w}_k\|}$. This proves the bound on the number of steps. □

We can now recursively apply this lemma to compute the $S_{k-1}(\pi_{k-1}(\mathbf{c} - \mu\mathbf{b}_k), r)$. We will actually compute supersets, but this does not affect the total bound. We obtain the following corollary.

Corollary 4.14. *Given an LLL-reduced basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ we can solve **(SVP)** in time $\mathcal{O}(2^{d^2})$.*

Proof. We use the notation from the statement and proof of Lemma 4.13. So in particular we consider the sets

$$S_k(\mathbf{c}, r) \;:=\; \mathcal{B}_r(\mathbf{c}) \cap \Lambda_k \;\subseteq\; \Lambda_k$$

and the upper bound $f_k(r)$ for the enumeration of a superset $\tilde{S}_k(\mathbf{c}, r) \supseteq S_k(\mathbf{c}, r)$ of $S_k(\mathbf{c}, r)$ for some $\mathbf{x} \in \mathbb{R}^d$ (which need not be a lattice point) and some radius $r > 0$.

Clearly we can enumerate the set $S_1(\mathbf{c}, r)$ with at most

$$\frac{2r}{\|\mathbf{b}_1\|} \;=\; \frac{2r}{\|\mathbf{w}_1\|}$$

steps. So, recursively applying Lemma 4.13, we obtain $f_k(r) = \prod_{i=1}^{k-1} \frac{2r}{\|\mathbf{w}_i\|}$. Doing this up to $k = d$ we get

$$\prod_{i=1}^{d} \frac{2r}{\|\mathbf{w}_i\|} \;=\; \frac{2^d r^d}{\det\,\Lambda}$$

Now we plug in $r = \|\mathbf{b}_1\|$, which is an upper bound for the length of a shortest vector. So by enumerating all points in $S_d(0, \|\mathbf{b}_1\|)$ we necessarily also find the shortest nonzero vector.

From (4.10) of Proposition 4.6 this is bounded by $2^{(d-1)/4}\,(\det\,\Lambda)^{1/d}$, which implies the claim. □

This bound is still not optimal. Kannan[5] has shown that one can solve **(SVP)** in time $2^{d \log d}$, also using an enumeration over the coefficients in a basis, but using a different basis reduction algorithm that yields a more suitable basis for this. Ajtai, Kumar and Sivakumar provided a randomized sieve method, that computes the shortest vector in time $2^{\mathcal{O}(d)}$.[6] In 2015, Aggarwal, Dadush, Regev, and Stephens-Davidowitz gave an algorithm, that runs in time 2^d.[7]

While Kannan's algorithm uses only polynomial space, all others require exponential space during computation. It is open whether one can bring this down to polynomial space.

4.4. Problems

4.1. Let Λ be a lattice and λ_1 the first successive minimum (the length of the shortest vector). Show that for any $k \in \mathbb{Z}_{\geq 0}$

$$|\Lambda \cap \mathcal{B}_{k\lambda_1}(0)| \leq (2k+1)^d .$$

4.2. Show that

(1) $\|\mathbf{w}_j\|^2 \leq \|\mathbf{v}_j\|^2 \leq \|\mathbf{w}_j\|^2 + \frac{1}{4}\sum_{i=1}^{j-1} \|\mathbf{w}_i\|^2$ for $1 \leq j \leq d$

(2) $\|\mathbf{w}_j\| \leq \|\mathbf{b}_j\| \leq \left(\frac{1}{2} + 2^{i-2}\right) \|\mathbf{w}_j\|$ for $1 \leq j \leq d$

(3) $\|\mathbf{b}_k\| \leq 2^{(j-1)/2} \|\mathbf{w}_j\|$ for $1 \leq k \leq j \leq d$.

4.3. Let $\mathbf{v}_1, \ldots, \mathbf{v}_d$ be lattice vectors such that $\|\mathbf{v}_i\| = \lambda_i$, where $\lambda_1, \ldots, \lambda_d$ are the successive minima of Λ. We can write them in the lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ as

$$\mathbf{v}_k = \sum_{i=1}^{d} \mu_{ki} \mathbf{b}_i .$$

For each $1 \leq k \leq d$ let j_k be the largest index such that $\mu_{k,j_k} \neq 0$. Show that

(1) $\|\mathbf{v}_j\| \geq \|\mathbf{w}_{j_k}\|$

(2) $\|\mathbf{b}_j\| \leq 2^{(d-1)/2} \cdot \lambda_i$ for $1 \leq j \leq i \leq d$

(3) $2^{(1-i)/2}\lambda_i \leq \|\mathbf{b}_i\| \leq 2^{(d-1)/2}\lambda_i$

4.4. Let Λ be a lattice of full rank with basis B, and $\mathbf{b}_1, \ldots, \mathbf{b}_d \in \Lambda$ linearly independent. Let $\mathbf{w}_d$ be the last vector in the Gram-Schmidt orthogonalization process. Then

$$\|\mathbf{w}_d\| \leq \max(\|\mathbf{b}_1\|, \ldots, \|\mathbf{b}_d\|)$$

Hint: This should follow from the proof of Theorem 4.3.

4.5. Give a proof of Corollary 3.12 that is independent of Minkowski's Second Theorem (Theorem 3.11).

[5] Ravi Kannan, "Improved Algorithms for Integer Programming and Related Lattice Problems".
[6] Ajtai, Kumar, and Sivakumar, "A sieve algorithm for the shortest lattice vector problem".
[7] Aggarwal, Dadush, and Stephens-Davidowitz, "Solving the Closest Vector Problem in 2^n Time— The Discrete Gaussian Strikes Again!"

4.6. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be a lattice basis of a lattice Λ with Gram-Schmidt vectors $\mathbf{w}_1, \ldots, \mathbf{w}_d$. Show that the covering radius satisfies

$$\mu(\Lambda) \;\leq\; \frac{1}{2}\sqrt{\sum_{i=1}^{d} \|\mathbf{w}_i\|^2}\,.$$

Find a lattice basis so that we have equality.

4.7. Check, which of the following bases represent an LLL-reduced basis for the lattice they span.

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \qquad \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \qquad \begin{bmatrix} 10 & 0 \\ 0 & 9 \end{bmatrix}$$

4.8. Λ with basis B. Then there is a basis B' spanning a sublattice Λ' of Λ with orthogonality defect bounded by $n^{O(n)}$.

Hint: Use Corollary 3.12.

There is also a basis of Λ with this bound, the *Khorkine-Zolotarav-basis*. This is (much) harder.

4.9. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice of full rank with lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_d \in \Lambda$, and $\mathbf{a} \in \Lambda^\star$. Let

$$L \;:=\; \{\, \mathbf{v} \;:\; \langle \mathbf{a}, \mathbf{v} \rangle \;=\; 0 \,\}$$

be the Λ-rational subspace defined by $\mathbf{a}$.

Show that one can compute a lattice basis of L in polynomial time.

4.10. Let $\mathbf{b}_1, \ldots \mathbf{b}_k \subseteq \mathbb{R}^d$ be linearly independent, $k \leq d$. Let Z be the zonotope of dimension k spanned by these vectors and B the matrix, whose columns are the $\mathbf{b}_1$. Then

$$\text{vol } Z \;=\; \sqrt{\det(B^t B)}\,,$$

where we use the relative volume in the subspace defined by $\mathbf{b}_1, \ldots, \mathbf{b}_k$.

5. Reduced Bases

We have seen in the previous section that reduced bases allow the computation of a shortest vector in polynomial time. Yet, so far we don't even know that such bases exist for our lattice Λ. We address this question in this chapter with a polynomial time algorithm to compute an LLL-reduced lattice basis. This method was first described by Arjen Lenstra, Hendrik Lenstra and László Lovász in 1982.[1]

More background and many more applications may be found in the books of Cohen,[2] von zur Gathen and Gerhard,[3] Grötschel, Lovász and Schrijver[4] and the proceedings by Nguyen and Vallée.[5] Also the original paper by Lenstra, Lenstra, and Lovász[1] from 1982 is a good source.

There is a similar notion of a reduced basis in dimension 2 due to Lagrange and Gauss, and nowadays often referred to as *Gauss' Algorithm*.[6] As a good preparation for the rest of this chapter you can study this with Problem 5.1.

Problem 5.1

We will show that LLL-reduced bases for a lattice Λ always exist for sublattices of $\mathbb{Z}^d$. We will present the algorithm for any lattice, but we restrict to sublattices of $\mathbb{Z}^d$ for the proof that it can actually computed in polynomial time (in the dimension and the input size). This extends to the general case, but the analysis is much more involved. You can find a proof in the original paper.[1]

Here is the main theorem that we want to prove. The construction of the algorithm claimed here will cover the rest of the chapter.

Theorem 5.1 (Lenstra, Lenstra, Lovász, 1982). *$\Lambda \subset \mathbb{R}^d$ a lattice with an integral basis $\mathbf{b}'_1, \dots, \mathbf{b}'_d$. Then there is a LLL-reduced basis $\mathbf{b}_1, \dots, \mathbf{b}_d$ of Λ such that*

$$\prod_{i=1}^{d} \|\mathbf{b}_i\| \;\leq\; 2^{\frac{1}{2}\binom{d}{2}} \det \Lambda$$

with orthogonality defect bounded by $2^{\frac{1}{2}\binom{d}{2}}$. We can compute this basis in time polynomial in d and $\log \|\mathbf{b}'_i\|_2$.

This will fill the missing piece in the computation of a shortest lattice vector from the previous chapter.

[1] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".
[2] Cohen, *A course in computational algebraic number theory*.
[3] Gathen and Gerhard, *Modern computer algebra*.
[4] Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*.
[5] Nguyen and Vallée, *The LLL algorithm*.
[6] Nguyen and Stehlé, "Low-dimensional lattice basis reduction revisited".

Note that we have restricted to sublattices of the integer lattice in the formulation of the theorem. We will only prove this restricted case here, but the theorem actually holds without this restriction, using the same method for the construction of a basis. We will in fact also present the construction in full generality, and only use the restriction to sublattices of $\mathbb{Z}^d$ in the proof that that the algorithm runs with polynomially many steps. A proof of the general case, which is more involved, but does not contain any new ideas, can be found in the original paper of Lenstra, Lenstra, and Lovász.[7]

Also, in the same way as already in the definition of reduced bases, we have restricted the formulation to the case $\delta = \frac{3}{4}$. The theorem is true for all δ in the range $\frac{1}{4} < \delta < 1$. Also here, the proof essentially remains the same, but some computations get more involved. It follows from the construction that we obtain better bases for larger δ, but the running time increases. The algorithm itself also works for $\delta = 1$, but then the bound on the running time becomes exponential in this case.

The construction of the basis will have two parts. In the first part we will discuss how we can obtain a *weakly reduced* basis. In the second part, we will show how we can move closer to a *reduced* basis if some entries in the basis violate the condition. The resulting basis may not be weakly reduced anymore after this step. We will, however, prove that we can reapply the first step to make this basis weakly reduced again without loosing the improvement made in the second step.

Finally, we will show that we do not need too many iterations of this process, *i.e.* at most polynomially many. Together with a proof that we can do each step with polynomially many operations this will give the result.

5.1. Weakly reduced bases

We need some preparations. Recall our definition in (4.1) of Λ-rational subspaces

$$V_0 := \{\mathbf{0}\} \qquad \text{and} \qquad V_k := \operatorname{lin}(\mathbf{b}_1, \dots, \mathbf{b}_k) \qquad \text{for } 1 \le k \le d\,.$$

together with orthogonal projections $\pi_k : \mathbb{R}^d \to V_k$ and induced lattices $\Lambda_k := \Lambda \cap V_k$. We are given some basis $\mathbf{b}_1, \dots, \mathbf{b}_d$ of the lattice, and compute the Gram-Schmidt orthogonalization $\mathbf{w}_1, \dots, \mathbf{w}_d$ of it.

As seen in (4.3) we have a representation of our given basis in the form

$$\mathbf{w}_k := \mathbf{b}_k - \pi_{k-1}(\mathbf{b}_k) \tag{5.1}$$

$$= \mathbf{b}_k - \sum_{j=1}^{k-1} \lambda_{jk} \mathbf{w}_j \qquad \text{with} \quad \lambda_{jk} := \frac{\langle\, \mathbf{b}_k, \mathbf{w}_j \,\rangle}{\|\mathbf{w}_j\|^2}\,. \tag{5.2}$$

Assume that for some pair of indices $i < k$ the absolute value of the coefficient λ_{ik} is larger than $\frac{1}{2}$. Then there are unique $\mu_{ik} \in \mathbb{R}$ and $a_{ik} \in \mathbb{Z}$ such that

$$|\mu_{ik}| \le \tfrac{1}{2} \qquad\qquad \lambda_{ik} = a_{ik} + \mu_{ik}\,.$$

[7] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".

We set

$$\mathbf{b}'_k \;:=\; \mathbf{b}_k - a_{ik}\mathbf{b}_i \qquad \text{and} \qquad \mathbf{b}'_j \;:=\; \mathbf{b}_j \quad \text{for } j \neq k$$

and define new subspaces

$$V'_0 := \{\mathbf{0}\} \qquad V'_k := \operatorname{lin}(\mathbf{b}'_1, \dots, \mathbf{b}'_k) \qquad \text{for } 1 \leq k \leq d\,.$$

By construction,

$$V_j \;=\; V'_j \qquad \text{for} \qquad 1 \;\leq\; j \;\leq\; d \tag{5.3}$$

and $\mathbf{b}'_1, \dots, \mathbf{b}'_j$ is still a basis of the lattice Λ_j for $1 \leq j \leq d$. It follows from (5.1) that (5.3) implies that the Gram-Schmidt orthogonalization $\mathbf{w}_1, \dots, \mathbf{w}_d$ does not change.

We want to compute the coefficients λ'_{jk} for the new basis in the representation (5.2). For this, we consider a fixed k between 1 and d. As the Gram-Schmidt vectors do not change, the formula in (5.2) shows that only coefficients that involve $\mathbf{b}'_k$ can change. So

$$\lambda'_{jl} \;=\; \lambda_{jl} \qquad \text{for} \qquad l \;\neq\; k \quad \text{and} \quad j \;<\; l\,.$$

To compute the new coefficients for $\mathbf{b}'_k$ we consider

$$\begin{aligned}
\mathbf{b}'_k \;=\; \mathbf{b}_k - a_{ik}\mathbf{b}_i \;&=\; \mathbf{w}_k + \sum_{j=1}^{k-1} \lambda_{jk}\mathbf{w}_j - a_{ik}\mathbf{b}_i \\
&=\; \mathbf{w}_k + \sum_{j=1}^{k-1} \lambda_{jk}\mathbf{w}_j - a_{ik}\sum_{j=1}^{i} \eta_j \mathbf{w}_j \\
&=\; \mathbf{w}_k + \sum_{j=1}^{i} (\lambda_{jk} - a_{ik}\eta_j)\mathbf{w}_j + \sum_{j=i+1}^{k-1} \lambda_{jk}\mathbf{w}_j\,,
\end{aligned}$$

for some $\eta_j \in \mathbb{R}$, $1 \leq j \leq i$. The second equation follows as $\mathbf{b}_i \in V_i$, which is spanned by $\mathbf{w}_1, \dots, \mathbf{w}_i$. Furthermore, $\mathbf{b}_i - \mathbf{w}_i \in V_{i-1}$ implies $\eta_i = 1$.

We can now bound the size of $\lambda'_{ik} := \lambda_{ik} - \eta_i a_{ik}$ with

$$|\lambda_{ik} - \eta_i a_{ik}| \;=\; |\lambda_{ik} - a_{ik}| \;=\; |\mu_{ik}| \;\leq\; \tfrac{1}{2}\,.$$

So λ'_{ik} is weakly reduced, and the new coefficients for $\mathbf{b}'_k$ are

$$\begin{aligned}
\lambda'_{jk} \;&:=\; \lambda_{jk} - a_{ik}\eta_j \qquad && \text{for} \qquad j \;\leq\; i \\
\lambda'_{jk} \;&:=\; \lambda_{jk} && \text{for} \qquad j \;>\; i
\end{aligned}$$

For fixed k, making λ_{ik} weakly reduced only affects λ_{jk} for $j < i$. Hence, we can make λ_{ik} weakly reduced for all $1 \leq i < k$ if we apply the reduction starting from $j = k-1$ in decreasing order. Doing this for all $1 \leq k \leq d$ gives us a weakly reduced basis.

There are $\binom{d-1}{2}$ coefficients, and reducing λ_{ik} affects at most $i - 1 \leq d$ other coefficients, so this process terminates after at most $\mathcal{O}(d^3)$ steps.

For an LLL-reduced basis we also need to satisfy the condition (4.6),

$$\frac{3}{4}\,\|\mathbf{w}_k\|^2 \;\le\; \|\lambda_{k,k+1}\mathbf{w}_k + \mathbf{w}_{k+1}\|^2 \qquad \text{for} \qquad 1 \;\le\; k \;\le\; d-1\,.$$

Now assume that this condition is violated for some j, *i.e.*

$$\frac{3}{4}\,\|\mathbf{w}_j\|^2 \;>\; \|\lambda_{j,j+1}\mathbf{w}_j + \mathbf{w}_{j+1}\|^2\,.$$

Let $\mathbf{b}'_1,\ldots,\mathbf{b}'_d$ be the basis obtained by exchanging $\mathbf{b}_j$ and $\mathbf{b}_{j+1}$, i.e.

$$\mathbf{b}'_{j+1} := \mathbf{b}_j \qquad \mathbf{b}'_j := \mathbf{b}_{j+1} \qquad \mathbf{b}'_i := \mathbf{b}_i \qquad \text{for } i \neq j, j+1\,.$$

Let V'_i, Λ'_i be the new subspaces and lattices, for $1 \le i \le d$. Then

$$V'_i \;=\; V_i\,, \qquad \Lambda'_i \;=\; \Lambda_i\,, \qquad \text{and} \qquad \det\,\Lambda'_i \;=\; \det\,\Lambda_i \qquad \text{for } i \neq j\,,$$

while

$$V'_j := \mathrm{lin}(\mathbf{b}'_1,\ldots,\mathbf{b}'_{j-1},\mathbf{b}'_j) = \mathrm{lin}(\mathbf{b}_1,\ldots,\mathbf{b}_{j-1},\mathbf{b}_{j+1})\,.$$

We want to compute $\det\,\Lambda'_j$. we know that

$$\det\,\Lambda_j \;=\; \prod_{i=1}^{j}\|\mathbf{w}_i\| \qquad \text{and} \qquad \det\,\Lambda'_j \;=\; \prod_{i=1}^{j}\|\mathbf{w}'_i\|\,,$$

so we need the Gram-Schmidt vectors $\mathbf{w}'_1,\ldots,\mathbf{w}'_j$.

We conclude from (5.1) that

$$\mathbf{w}'_i \;=\; \mathbf{w}_i \qquad \text{for} \qquad i \;<\; j \text{ and } i \;>\; j+1\,,$$

as neither $\mathbf{b}_i$ nor π_{i-1} change for these i.

Again using (5.1) we compute

$$\begin{aligned}\mathbf{w}'_j \;&:=\; \mathbf{b}_{j+1} \;-\; \pi_{j-1}(\mathbf{b}_{j+1})\\ &=\; \mathbf{b}_{j+1} \;-\; \sum_{i=1}^{j-1}\lambda_{i,j+1}\mathbf{w}_i\\ &=\; \mathbf{w}_{j+1} \;+\; \lambda_{j,j+1}\mathbf{w}_j\end{aligned}$$

So

$$\det\Lambda'_j \;=\; \prod_{i=1}^{j}\|\mathbf{w}'_i\| \;=\; \left(\prod_{i=1}^{j-1}\|\mathbf{w}'_i\|\right)\cdot\|\lambda_{j,j+1}\mathbf{w}_j + \mathbf{w}_{j+1}\|$$

$$\leq \left(\prod_{i=1}^{j-1} \|\mathbf{w}_i'\|\right) \cdot \sqrt{\frac{3}{4}}\,\|\mathbf{w}_j\|$$
$$= \sqrt{\frac{3}{4}}\det\,\Lambda_j\,,$$

i.e. the determinant of the j-th lattice drops by a constant factor, while all other determinants do not change (Note that also $\mathbf{w}_{j+1}$ may change, but this effect must be cancelled in the computation of the determinant of Λ_k' for $k \geq j+1$ by the change in $\mathbf{w}_j$, as the lattice is spanned by the same set $\mathbf{b}_1, \ldots, \mathbf{b}_k$.).

In the new order of the basis vectors the basis may not be weakly reduced anymore. We can fix this by repeating the above method. However, we see from (5.2) that only λ_{ik} for $i, k \in \{\, j, j+1 \,\}$ may change when we swap $\mathbf{b}_j$ and $\mathbf{b}_{j+1}$. Making them weakly reduced again may require us to change λ_{ik} for $k \in \{\, j, j+1 \,\}$ and all $i < k$. In total, we change at most $2d$ of the coefficients, and for each we may need to change at most d other coefficients. Hence, we need at most $\mathcal{O}(d^2)$ steps.

Note, that in this process that lattices Λ_i do not change, so also their determinants do not change. Hence, if there are more pairs of basis vectors that violate (4.6), we can continue to exchange such pairs of vectors and make the basis weakly reduced again. With each swap, the determinant of one of the lattices Λ_j drops by a factor of at least $\sqrt{\frac{3}{4}}$, while all others do not change.

This is more conveniently captured with the following notion of a *potential* of a lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ defined as

$$D(\mathbf{b}_1, \ldots, \mathbf{b}_d) \;:=\; \prod_{j=1}^{d} \det \Lambda_j \;=\; \prod_{j=1}^{d}\prod_{k=1}^{j} \|\mathbf{w}_k\|\,. \tag{5.4}$$

Note that $D(\mathbf{b}_1, \ldots, \mathbf{b}_d)$ depends on the order of the basis vectors. This gives the following lemma.

Lemma 5.2. *In each iteration of the algorithm the potential drops by at least a factor of* $\sqrt{\frac{3}{4}}$. □

With this we can estimate the number of steps our method needs to produce an LLL-reduced basis and prove that this is in fact polynomial in the input size, at least in the case that all entries in the original basis where integral. So here is the one place where we need to restrict to sublattices of $\mathbb{Z}^d$ in our argument. The crucial fact following from this is the observation that in this case

$$D(\mathbf{b}_1, \ldots, \mathbf{b}_d) \;\geq\; 1\,. \tag{5.5}$$

Proposition 5.3. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice with a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d \in \mathbb{Z}^d$ of integral vectors*

(i.e. Λ is a sublattice of $\mathbb{Z}^d$). Let B be a bound on the Euclidean length of the $\mathbf{b}_i$, i.e.

$$B \;\geq\; \max\left(\|\mathbf{b}_1\|_2^2, \ldots, \|\mathbf{b}_d\|_2^2\right).$$

Then

$$1 \;\leq\; D(\mathbf{b}_1, \ldots, \mathbf{b}_d) \;\leq\; B^{d(d-1)/2}.$$

Proof. The norm of the Gram-Schmidt vectors is bounded by that of the corresponding basis vector, so

$$\|\mathbf{w}_i\|_2 \;\leq\; \|\mathbf{b}_i\|_2 \;\leq\; B.$$

The upper bound now follows from (5.4).

For the lower bound we need that $\det\,\Lambda_i \in \mathbb{Z}$ for all i. Let B_i be the $(d \times i)$-matrix of the first i basis vectors. It follows from Problem 4.10 that

$$\det\,\Lambda_i \;=\; \mathrm{vol}\,(\Pi(\mathbf{b}_1, \ldots, \mathbf{b}_i)) \;=\; \sqrt{\det\,B_i^t B_i}\,.$$

The latter is integral if all $\mathbf{b}_i$ are integral, which proves the lower bound. □

By Lemma 5.2, the value of D drops by at least a factor of $\sqrt{\frac{3}{4}}$ each time we swap a pair of basis vectors that violate condition (4.5), so, using the bounds of the previous proposition, we need at most

$$k \;=\; \left\lceil \frac{2}{\log\left(\frac{4}{3}\right)} \cdot \frac{d(d-1)}{2} \log\,B \right\rceil \tag{5.6}$$

such swaps. We summarize this with the next corollary.

Corollary 5.4. *Let $\Lambda \subseteq \mathbb{R}^d$ be a sublattice of $\mathbb{Z}^d$ with a basis $\mathrm{B} := \{\,\mathbf{b}_1, \ldots, \mathbf{b}_d\,\} \in \mathbb{Z}^d$ of integral vectors. Then the LLL-algorithm applied to this basis uses at most*

$$\mathcal{O}(d^2 \cdot \log\,B)$$

iterations. □

After each swap we have to restore the property that our lattice basis is weakly reduced. We have seen above that this takes at most $\mathcal{O}(d^2)$ steps. So overall, the lattice basis reduction needs at most

$$\mathcal{O}\left(d^4 \log\,B\right)$$

steps. Initially we have to compute a Gram-Schmidt orthogonalization once. From the description we can easily see that this requires at most $\mathcal{O}(d^3)$ steps. This is subsumed in the number of steps we need to do for the LLL algorithm.

Clearly, if $\Lambda \subset \mathbb{Z}^d$ is a sublattice of the integer lattice, then LLL can be implemented using exact arithmetic over the rationals, and hence exact integer arithmetic. But to show that LLL actually runs in polynomial time we also have to show that all numbers computed in intermediate steps of the algorithm have a binary encoding size bounded by a polynomial in I.

Bounding the size of the numbers during the algorithm consists of several parts. Given the basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ with Euclidean norm bounded by B as above, we need to show that also the Gram-Schmidt basis is bounded by this constant. This is true, and can either be derived directly from the algorithm, or you can consult text books in linear algebra. The size of intermediate results is bounded by $d \cdot B$.

We also have to bound the intermediate and final values of the basis vectors $\mathbf{b}_i$ and the coefficients λ_{ik}. For this we refer to the original paper of Lenstra, Lenstra and Lovász[8], where they prove the following lemma.

Lemma 5.5. *For a lattice $\Lambda \subseteq \mathbb{Z}^d$ with a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ whose Euclidean length is bounded by B the LLL algorithm requires arithmetic operations on integers of size $\mathcal{O}(d \log B)$.* □

In the form we have discussed the LLL algorithm needs at most $\mathcal{O}\left(d^4 \log B\right)$ steps with arithmetic operations on numbers of size $\mathcal{O}(d \cdot \log B)$. Using naïve arithmetic for the the elementary operations we thus arrive at a total running time of

$$\mathcal{O}\left(d^6 \log^3 B\right) . \tag{5.7}$$

This corresponds to the original form of the algorithm given by Lenstra, Lenstra and Lovász.[8] There have been found several improvements that run significantly faster. Here are some references. Schnorr[9] has improved the bound to $\mathcal{O}(d^3 \log B)$ on integers of size at most $\mathcal{O}(d \cdot \log B)$. Storjohann[10] has another algorithm for a reduced basis in time $\mathcal{O}(d^3(\log B))$ on integers of size at most $\mathcal{O}(d \cdot \log B)$. Using fast matrix multiplication and an improvement of the LLL-algorithm by Schönhage will lead to a running time of $\mathcal{O}(d^{2.381}(\log B))$ with arithmetic operations on integers of the same size.

Problem 5.2
Problem 5.3
Problem 5.4

5.3. Further Notes

It is important to note that the notion of reduced bases and the algorithm we have seen to compute these bases has many more important applications apart from the computation of a shortest vector.[11] Here are some further applications.

- ▷ In the next chapter we use the approximation of the shortest vector for a polynomial algorithm to solve an integer linear program with a fixed number of variables.

[8] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".
[9] Koy and C. P. Schnorr, "Segment LLL-reduction of lattice bases"; Claus-P. Schnorr, "A hierarchy of polynomial time lattice basis reduction algorithms".
[10] Storjohann, *Faster Algorithms for Integer Lattice Basis Reduction*.
[11] Nguyen and Vallée, *The LLL algorithm*.

- ▷ We give an application to a knapsack problem arising from cryptography in Chapter 7.
- ▷ We will use the LLL-algorithm in Chapter 8 for an approximation of the closest vector problem.
- ▷ Lovász gave another possible attach on knapsack cryptosystems using lattice reduction and Diophantine approximation.[12]
- ▷ We can use LLL to find the minimal polynomial of an algebraic number given by a sufficiently good approximation, *e.g.* return $x^2 - 2 = 0$ on the input 1.41421356.
- ▷ In the original paper of Lenstra, Lenstra, and Lovász[13] the authors consider an application to factoring polynomials, *e.g.* find the two factors of $x^2 - 1 = (x+1)(x-1)$. For this see also the work of Klüners.[14]
- ▷ Find integer relations on a set of numbers $x_1, \ldots, x_k \in \mathbb{R}^d$, *i.e.* find $a_1, \ldots, a_k \in \mathbb{Z}$ such that

$$a_1 x_1 \ + \ \cdots \ + \ a_k x_k \ = \ 0$$

and not all a_i are zero. An example of such an integer relation is Machin's formula,

$$\frac{\pi}{4} \ = \ 4 \arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) .$$

5.4. Problems

5.1. Let $\Lambda \subseteq \mathbb{R}^2$ be a lattice. An ordered basis $\mathbf{b}_1, \mathbf{b}_2 \subseteq \mathbb{R}^2$ of Λ is *Lagrange-reduced* if

$$\|\mathbf{b}_1\|_2 \ \leq \ \|\mathbf{b}_2\|_2 \ \leq \|\mathbf{b}_2 + a\mathbf{b}_1\|$$

for all $a \in \mathbb{Z}$.

(i) Let λ_1 and λ_2 be the successive minima of Λ. Show that a Lagrange-reduced basis of Λ satisfies $\|\mathbf{b}_1\|_2 = \lambda_1$ and $\|\mathbf{b}_2\|_2 = \lambda_2$.

(ii) Show that the function

$$a \ \longmapsto \ \|\mathbf{b}_2 + a\mathbf{b}_1\|^2 \ = \ \|\mathbf{b}_2\|^2 \ + \ 2a\langle \mathbf{b}_1, \mathbf{b}_2 \rangle \ + \ a^2 \|\mathbf{b}_1\|^2$$

takes its minimum at $a_0 := \frac{\langle \mathbf{b}_1, \mathbf{b}_2 \rangle}{\|\mathbf{b}_1\|^2}$.

(iii) Show that an ordered basis is Lagrange reduced if and only if

$$\|\mathbf{b}_1\|_2 \ \leq \ \|\mathbf{b}_2\|_2 \ \leq \|\mathbf{b}_2 \pm \mathbf{b}_1\|$$

(iv) Show that the following algorithm transforms any lattice basis $\mathbf{b}_1, \mathbf{b}_2$ of Λ into a Lagrange reduced one.

a) replace $\mathbf{b}_2$ by $\mathbf{b}_1 - \lceil a_0 \rfloor \mathbf{b}_2$
b) if $\|\mathbf{b}_2\| \ < \ \|\mathbf{b}_1\|$, then swap $\mathbf{b}_1$ and $\mathbf{b}_2$ and repeat.
c) otherwise return $\mathbf{b}_1, \mathbf{b}_2$.

[12] J. C. Lagarias, "Knapsack public key cryptosystems and Diophantine approximation (extended abstract)".
[13] A. K. Lenstra, H. W. Lenstra Jr., and L. Lovász, "Factoring polynomials with rational coefficients".
[14] Klüners, "The van Hoeij Algorithm for Factoring Polynomials".

This algorithm applied to a basis $\mathbf{b}_1, \mathbf{b}_2 \subseteq \mathbb{Z}^2$ runs in fact in polynomial time

$$\mathcal{O}(\log^3(\max(\|\mathbf{b}_1\|_2, \|\mathbf{b}_2\|_2)).$$

You may attempt to prove this, but this is more difficult.

5.2. Assume that

$$\|\mathbf{b}_1\| \leq \|\mathbf{w}_i\| \qquad \text{for all} \qquad 1 \leq i \leq d.$$

Show that $\|\mathbf{b}_1\|$ is a shortest non-zero vector.

5.3. Let $\mathbf{b}_1, \ldots \mathbf{b}_d$ be LLL-reduced and choose a permutation σ on $\{1, \ldots, d\}$ such that

$$\left\|\mathbf{b}_{\sigma(i)}\right\| \leq \left\|\mathbf{b}_{\sigma(i+1)}\right\| \qquad \text{for} \qquad 1 \leq i \leq d-1.$$

(i) Give an example with $\|\mathbf{b}_1\| \neq \left\|\mathbf{b}_{\sigma(1)}\right\|$

(ii) Show that

$$\left\|\mathbf{b}_{\sigma(i)}\right\| \leq 2^{\frac{1}{4}d(d-1)} \cdot (\det \Lambda)^{1/(n+1-i)}.$$

5.4. Compute an LLL-reduced basis for some examples in dimensions 2 and 3.

6. Integer Programming

The integer programming problem is the task to find an integral point in a polyhedron defined by linear constraints that maximizes a linear functional, *i.e.* find, for $A \in \mathbb{Z}^{m\times d}$, $\mathbf{b} \in \mathbb{Z}^m$ and $\mathbf{c} \in \mathbb{Z}^d$,

$$\begin{aligned} &\max \langle \mathbf{c}, \mathbf{x} \rangle \\ &\text{subject to} \quad A\mathbf{x} \leq \mathbf{b} \\ &\qquad\qquad\quad \mathbf{x} \in \mathbb{Z}^d. \end{aligned}$$

Different from the linear programming problem, where we drop the requirement that $\mathbf{x} \in \mathbb{Z}^d$, this problem is known to be in the class NP, but not known to be in P.[1]

However, integer programming appears at the core of many optimization problems, and having efficient tools for its solution is vitally important in many cases. This has lead to rich theory of methods to solve integer programs despite its theoretical complexity, often in special cases where we know more about the algebraic or geometric structure of the problem. In a second direction there have also been developed many ways to obtain approximations of the exact solution. You may have met many such tools in the class on *Discrete Optimization*.

Standard approaches to obtain exact solutions are the Branch&Bound or Branch&Cut methods. Both methods start with the linear relaxation of the problem and are, broadly speaking, recursive methods that lay out *systematic* ways either to *split* the feasible set into smaller pieces and look at those separately, or to shrink the feasible region by *cutting off* pieces that cannot contain an integer solution. In both cases we need methods to *measure* or *estimate* which splits or cuts are most promising and should be chosen next. You can find good introductions into this theory in various textbooks, *e.g.* the one of Schrijver,[2] or Nemhauser and Wolsey,[3] or Bertsimas and Weismantel.[4]

We will look at one such approach in Chapter * 11, where we will look at cuts and lattice free polytopes.

Many special cases allow polynomial solutions or, if not, at least more efficient superpolynomial solutions. The additional structure may come from more information on the origin of the problem, *e.g.* in combinatorial optimization, where one often deals with problems whose underlying structure is a graph, or where the solution space

[1] Recall that the usually applied simplex algorithm for the computation of a solution to a linear program is not known to be in P, although it is practically usually fast and efficient. The proof of polynomiality is via the *ellipsoid method* of Khinchine. This is a theoretically polynomial, but practically not (yet) efficient algorithm, so that many software tools implement variations of the simplex method

[2] Alexander Schrijver, *Theory of linear and integer programming.*

[3] Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*.

[4] Bertsimas and Weismantel, *Optimization over integers*.

is restricted to vectors from $\{0,1\}^d$. This may allow to solve (subproblems of) the problem with methods from other fields, *e.g.* graph algorithms.[5]. We may also have more information on the constraint matrix or the whole system of constraints, *e.g.* the matrix may be totally unimodular or the system may be TDI.[6]

Prominent examples for methods the retreat to approximations of the exact solution are the Lagrange or Benders' decomposition, that either split off some constraints or some variables from the problem in the hope to obtain subproblems that are easier to solve (preferably because we can then find additional structure that makes the problem polynomially solvable). To solve the original problem one then has to find a way to insert the neglected constraints or variables back into the problem without loosing (all of) the advantage of the simpler problem structure.[7]

In all of the above we consider both the size of the constraint matrix and the number of variables (*i.e.* the ambient dimension of the polyhedron defined by the constraints) as input size. One may ask in which way these two different input parameters contribute to the complexity of the problem, and whether fixing or restricting one of them may lead to polynomial or, at least, more efficient, algorithms. Bounding the complexity of the constraint matrix may indeed lead to more efficient algorithms, and fixing some parameters even gives polynomial time algorithms, see *e.g.* the resent survey of Eisenbrand at al.[8]

In this chapter we will discuss the integer programming problem for fixed dimension or number of variables, and show that this problem is polynomial. The algorithm we present is due to Hendrik Lenstra. It is based on flatness and its proof via ellipsoids, together with the computation of reduced bases with the LLL-algorithm. The version of flatness that we need here, is more general than the version we have seen in the Flatness Theorem (Theorem 3.29), but the method of proof is just an extension of what we have seen before.

In the second part of this chapter on integer programs we will look at the minimum infeasible set problem. You may have discussed the corresponding linear version already in a course on linear programming. If we are given an infeasible linear program

$$\max\left(\langle \mathbf{c}, \mathbf{x} \rangle \ : \ A\mathbf{x} \leq \mathbf{b}\right),$$

i.e. a program where the feasible region $P := \{\, \mathbf{x} \ : \ A\mathbf{x} \leq \mathbf{b} \,\}$ is empty, then we can ask which, and how many of the constraints we need such that already the polyhedron defined by these inequalities is empty. In linear programming it can be shown that a subset of at most $d+1$ of the constraints is sufficient.

This theorem has no direct extension to integer programming, as it is not anymore true that infeasibility of the problem implies that the linear constraints have no common solution. The polyhedron P may well be a proper full-dimensional set, but $P \cap \Lambda$ may be empty. We will see in Section 6.3 that we nevertheless can recover a similar statement, but the bound of $d+1$ necessary inequalities increases significantly.

[5] Korte and Vygen, *Combinatorial Optimization*.
[6] Alexander Schrijver, *Theory of linear and integer programming.*
[7] Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*.
[8] Eisenbrand, Hunkenschröder, K.-M. Klein, Koutecký, Levin, and Onn, *An Algorithmic Theory of Integer Programming*.

6.1. Flatness Revisited

Let us look at the flatness theorem of Section 3.3 again. As before, our considerations will have three steps, and we first consider balls, almost trivially extend the result to ellipsoids, and then we use the John ellipsoids again to squeeze a polytope P between two ellipsoids and use the method for ellipsoids to find a feasible integer point in P or assert that there is none.

The refined version of flatness is necessary to not only consider convex bodies without interior lattice points, but to slice a polytope with a polynomial number of integral translates of a lattice hyperplane and descend in dimension if we cannot yet decide whether P has an integral point.

Let us make this precise with the following theorem. We follow ideas from Grötschel at al.[9] in this proof.

Theorem 6.1. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice of full rank and $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\} \subseteq \mathbb{R}^d$ with $A \in \mathbb{R}^{d\times m}$ and $\mathbf{b} \in \mathbb{R}^m$ a polyhedron. Then we can find, in polynomial time in d, $\log \|A\|_\infty$ and $\log \|\mathbf{b}\|_\infty$,*

(i) either a point $\mathbf{u} \in P \cap \Lambda$

(ii) or a direction $\mathbf{c} \in \Lambda^\star$ with

$$\operatorname{width}(P;\mathbf{c}) \;\leq\; f(d)$$

for some function $f(d) \;\leq\; 2^{\mathcal{O}(d^2 \log^2 d)}$.

We will discuss the complexity a bite more later, but we will not work this out precisely. This is technical, and we will not discuss all ingredients in enough detail to obtain a precise bound.

Observe the large gap in the constants in the above theorem, compared to the Flatness Theorem (Theorem 3.29). In the latter, we have a constant of $d^{5/2}$ that implies that P is lattice point free, while in the above we can only construct a direction with a flatness constant of $2^{d^2 \log d}$. However, reducing this to polynomial would imply a polynomial algorithm for integer programming, so this may be too much to ask for.

As explained above, the proof has three steps, for balls, for ellipsoids, and finally the extension to a proof of the above theorem. In this process, only the first step, proving the result for balls, will need substantial effort. Here is the proposition that solves flatness for balls.

Proposition 6.2. *Let Λ be a lattice with basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ and $\mathcal{B}_\mathrm{d} := \mathcal{B}_d(\mathbf{z})$ be the unit ball with center $\mathbf{z}$. Then we can find, in polynomial time,*

(i) either a lattice point $\mathbf{u} \in \mathcal{B}_\mathrm{d} \cap \Lambda$, or

[9] Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*.

(ii) a lattice functional $\mathbf{c} \in \Lambda^\star$ with $\|\mathbf{c}\| \leq 2^{\mathcal{O}(d^2)}$ such that $\mathcal{B}_\mathrm{d} \cap \Lambda$ is covered by at most $2^{\mathcal{O}(d^2 \log d)}$ hyperplanes of the form

$$\{\mathbf{x} \,:\, \langle\, \mathbf{c}, \mathbf{x} \,\rangle \;=\; \beta\} \;.$$

for some $\beta \in \mathbb{Z}$.

Proof. Assume that $\mathbf{b}_1, \ldots, \mathbf{b}_d$ is a LLL-reduced basis with orthogonality defect bounded by

$$M \;\leq\; \prod_{j=1}^{d} \frac{\|\mathbf{b}_i\|}{\|\mathbf{w}_i\|} \;\leq\; 2^{\frac{1}{2}\binom{d}{2}} \,,$$

where $\mathbf{w}_1, \ldots, \mathbf{w}_d$ is the associated Gram-Schmidt basis. Reorder the basis such that

$$\|\mathbf{b}_1\| \;\leq\; \|\mathbf{b}_2\| \;\leq\; \ldots \;\leq\; \|\mathbf{b}_d\| \;.$$

This may destroy reducedness of the basis, but does not affect the orthogonality defect. We distinguish two cases.

If $\|\mathbf{b}_d\| \leq \frac{1}{d}$, then we consider the representation

$$\mathbf{z} \;=\; \sum_{i=1}^{d} \lambda_i \mathbf{b}_i$$

of the center in our basis. The point

$$\mathbf{u} \;=\; \sum_{i=1}^{d} \lfloor \lambda_i \rfloor \mathbf{b}_i$$

is a lattice point and

$$\|\mathbf{u} - \mathbf{z}\| \;\leq\; \sum_{i=1}^{d} \|\mathbf{b}_i\| \;\leq\; d\,\|\mathbf{b}_d\| \;\leq\; 1 \,.$$

Hence $\mathbf{u} \in \mathcal{B}_\mathrm{d} \cap \Lambda$.

Otherwise we have $\|\mathbf{b}_d\| > \frac{1}{d}$. Let

$$L \;:=\; \mathrm{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}).$$

Then $L + \mathbf{b}_d = L + \mathbf{w}_d$, so

$$\Lambda \;\subseteq\; \bigcup_{\beta \in \mathbb{Z}} L + \beta \mathbf{b}_d \;=\; \bigcup_{\beta \in \mathbb{Z}} L + \beta \mathbf{w}_d \,.$$

Now

$$\frac{\|\mathbf{b}_d\|}{\|\mathbf{w}_d\|} \;\leq\; M \;\leq\; 2^{\frac{1}{2}\binom{d}{2}}\,,$$

so

$$\|\mathbf{w}_d\| \;\geq\; 2^{-\frac{1}{2}\binom{d}{2}}\,\|\mathbf{b}_d\| \;\geq\; \frac{1}{d}2^{-\frac{1}{2}\binom{d}{2}}\,.$$

Hence, translates of L are at least $\frac{1}{d}2^{-\frac{1}{2}\binom{d}{2}}$ apart. On the other hand, we can bound the distance in which we need to search trivially by

$$\max_{\beta\in\mathbb{Z}}\left(L+\beta\mathbf{b}_d\cap\mathcal{B}_\mathrm{d}\neq\varnothing\right) \;-\; \max_{\beta\in\mathbb{Z}}\left(L+\beta\mathbf{b}_d\cap\mathcal{B}_\mathrm{d}\neq\varnothing\right) \;\leq 2\,.$$

Combining this with the minimum distance of two translates of L we conclude that $L+\beta\mathbf{b}_d\cap\Lambda\neq\varnothing$ for at most $2d2^{\frac{1}{2}\binom{d}{2}}$ different β's.

Let $\mathbf{c}:=\frac{\mathbf{w}_d}{\|\mathbf{w}_d\|^2}$. Then $L+\beta\mathbf{w}_d=\{\,\mathbf{x}\;:\;\langle\,\mathbf{c},\mathbf{x}\,\rangle=\beta\,\}$, and any $\mathbf{u}\in\Lambda$ can be written as $\mathbf{u}=\mu\mathbf{w}_d+\mathbf{h}$ for $\mathbf{h}\in L$ and $\mu\in\mathbb{Z}$. This implies

$$\langle\,\mathbf{c},\mathbf{u}\,\rangle \;=\; \mu \;\in\; \mathbb{Z}\,.$$

so that $\mathbf{c}\in\Lambda^\star$ and $\|\mathbf{c}\|\leq\frac{1}{\|\mathbf{w}_d\|}\leq 2^{\mathcal{O}(d^2)}$. □

Geometrically, this proposition tells us that either we can find a reduced basis with only short vectors, or there must be a short dual vector, albeit with a very different bound. Namely, all basis vectors of our LLL-reduced basis have length at most $\frac{1}{d}$, or there is a dual vector of length at most $\frac{1}{\|\mathbf{w}_d\|}\leq 2^{\mathcal{O}(d^2)}$.

Problem 6.1

Corollary 6.3. *The proposition is also true if we replace the ball by an ellipsoid $E=T\,\mathcal{B}_\mathrm{d}+\mathbf{a}$ for an invertible linear transformation T.*

Proof. This follows in the same way as for the previous flatness theorem. Just pull back the ellipsoids to a ball with the linear map T^{-1} and consider the lattice $T^{-1}\Lambda$. □

By Lemma A.11 we know that for a convex body K there is an ellipsoid E centered at the origin such that

$$\mathbf{a}\;+\;E\;\subseteq\;K\;\subseteq\;\mathbf{a}\;+\;d\cdot E\,.$$

Unfortunately, the proof of this theorem is not constructive, and so far, also no polynomial time algorithm is known that allows to compute E. However, the precise scaling factor is not really important, as long as it depends polynomially on d, and we may use the approximation of Theorem A.34 instead. It guarantees the computation of an ellipsoid E with center $\mathbf{a}$ such that

$$\mathbf{a}\;+\;E\;\subseteq P\;\subseteq\;\mathbf{a}\;+\;d^{3/2}\cdot E\,. \tag{6.1}$$

The algorithm is essentially the same as for the ellipsoid method of linear programming, but needs more general cuts than the central cuts used in this method. However, the ideas used and the proofs are otherwise the same.

With this result we can finally prove Theorem 6.1.

Proof of Theorem 6.1. Let us first assume that P is bounded. Let E an ellipsoid with center $\mathbf{a}$ that satisfies (6.1). After translation we may assume that

$$E \subseteq P \subseteq d^{3/2}(E)\,.$$

Then Corollary 6.3 implies that we can either find an integral point in E, and hence in P, or there is a direction $\mathbf{c} \in \Lambda^\star$ such that at most $2^{\mathcal{O}(d^2 \log d)}$ hyperplanes defined by $\mathbf{c}$ intersect E. So at most $d^{3/2}2^{\mathcal{O}(d^2 \log d)} = 2^{\mathcal{O}(d^2 log^2 d)}$ hyperplanes can intersect P.

Finally, we want to reduce the unbounded to the bounded case. Let φ be an upper bound on the size of an inequality in the description of P. We know from Proposition A.30 that the size ν of the generators in a representation

$$P = \mathsf{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n) + \mathrm{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_m)$$

for some points $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and rays $\mathbf{r}_1, \ldots, \mathbf{r}_m$ is bounded by $\nu \leq 4d^2\varphi$.

Let $C > 0$ be the constant hidden in the $2^{\mathcal{O}(d^2 \log d)}$ above. We add one additional inequality

$$\|\mathbf{x}\| \leq C \cdot 2^{\nu+k+1} d^{3/2} 2^{d^2} \tag{6.2}$$

to the system. Then we have a bounded problem and we can apply the above result. We are done if we are in the first option and find a lattice point. This is also a lattice point in P.

Otherwise, we obtain a vector $\mathbf{c}$ such that there is δ with

$$\delta \leq \langle\, \mathbf{c}, \mathbf{x} \,\rangle \leq \delta + C \cdot d^{3/2} 2^{d^2}$$

for all $\mathbf{x} \in P$ that satisfy (6.2). Every $\mathbf{x} \in P$ can be written as

$$\mathbf{x} = \sum \lambda_i \mathbf{v}_i + \sum \mu_j \mathbf{r}_j$$

for $\lambda_i, \mu_j \geq 0$, $\sum \lambda_i = 1$ and the points and rays defined above of size at most ν. Then $\|\mathbf{v}_i\| \leq 2^\nu$, so

$$\mathbf{y} := \sum \lambda_i \mathbf{v}_i$$

is in P and satisfies (6.2). Hence, we are done if we can show that $\langle\, \mathbf{c}, \mathbf{r}_j \,\rangle = 0$.

Suppose not, then $|\langle\, \mathbf{c}, \mathbf{r}_j \,\rangle| > 2^{-\nu}$, and, changing the sign if necessary, we may assume that $\langle\, \mathbf{c}, \mathbf{r}_j \,\rangle > 2^{-\nu}$. Consider

$$\mathbf{z} := \mathbf{y} + 2^\nu d^{3/2} \cdot C \cdot 2^{d^2} \mathbf{r}_j\,.$$

Then

$$\langle \mathbf{c}, \mathbf{z} \rangle \;\geq\; \delta \;+\; C \cdot d^{3/2} 2^{d^2} \,.$$

This is a contradiction, as $\mathbf{z}$ is in P and satisfies (6.2).

So any lattice point in P is already on one of the hyperplanes returned by the above method if we add (6.2). □

It follows from work of Babai[10] and Lenstra[11] that we can reduce the number of hyperplanes to c^d for some constant c.

6.2. Integer Programming in Fixed Dimension

With these preparations we finally arrive at Lenstra's algorithm[11] for integer programming with a fixed number of variables.

Theorem 6.4 (Lenstra). *For a polytope $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ we can decide, in time $2^{\mathcal{O}(d^3 \log d)}$ times a polynomial in the encoding length of A and $\mathbf{b}$, whether P contains an integer point.*

Proof. We use Theorem 6.1. We are done, if we find an integral point. Otherwise, we obtain the short direction $\mathbf{c}$. We then consider the problem in each of the slices

$$P \cap \{\mathbf{x} : \langle \mathbf{c}, \mathbf{x} \rangle = \beta\} \tag{6.3}$$

for all $\beta \in \mathbb{Z}$ where this intersection is not empty. For this, we have to compute a lattice basis in the subspace $\{\mathbf{x} : \langle \mathbf{c}, \mathbf{x} \rangle = 0\}$, extend this to a lattice basis of the whole lattice, and use the transformation from the original basis (usually the $\mathbb{Z}^d$-basis) to transform the problem (6.3), which is a $(d-1)$-dimensional problem in d-dimensional space, into one in $(d-1)$-dimensional space. This can be done, *e.g.* with the Hermite normal form algorithm.

It remains to check the running time. For this, let $T(d)$ denote the number of recursive calls in dimension d. Then the total running time is $T(d)$ times a polynomial in d and the encoding length of A and $\mathbf{b}$.

By the Theorem 6.1 we know that

$$T(d) \;\leq\; T(d-1) \cdot 2^{\mathcal{O}(d^2 \log^2 d)} \,,$$

This implies

$$T(d) \;\leq\; \prod_{k=1}^{d} 2^{\mathcal{O}(d^2 \log^2 d)} \;\leq\; 2^{\mathcal{O}(d^3 \log^2 d)} \,. \qquad \square$$

[10] Babai, "On Lovász' lattice reduction and the nearest lattice point problem".
[11] Hendrik W. Lenstra, "Integer Programming with a fixed number of variables".

Note that we omitted one detail in the proof. In the transformation of the $(d-1)$-dimensional problems in $\mathbb{R}^d$ into problems in $\mathbb{R}^{d-1}$ we have to check that the size of the new constraints is still bounded in the input size. This is, however, not hard to check, not very instructive, and rather technical. So we refrain from doing this.

The above theorem only solves the feasibility problem. However, using binary search, we can easily solve the *optimization problem* in polynomial time.

Corollary 6.5. *For any fixed $d \geq 1$ we can solve the integer programming problem*

$$\max\Big(\langle \mathbf{c}, \mathbf{x} \rangle \;:\; A\mathbf{x} \leq \mathbf{b}\,,\; \mathbf{x} \in \mathbb{Z}^d\Big)$$

in polynomial time. □

The currently best known algorithm for integer programming in fixed dimension is due to Kannan [12]. It takes time $d^{\mathcal{O}(d)}$ times a polynomial in the size of A and $\mathbf{b}$. The main new idea here is that he applies the LLL-algorithm again in each iteration to improve the lattice basis. This brings down the number of parallel planes needed to check to a polynomial in d. It is still open whether we can bring the running time down to single exponential, *i.e.* whether we can replace $d^{\mathcal{O}(d)}$ to $2^{\mathcal{O}(d)}$.

6.3. Minimal infeasible subsets

We consider the feasible region P of a linear program in the form

$$P := \{\mathbf{x} \;:\; A\mathbf{x} \leq \mathbf{b}\}$$

for some $A \in \mathbb{Q}^{m\times n}$ and $\mathbf{b} \in \mathbb{Q}^d$. A well known theorem in linear optimization states that , if the program is infeasible, *i.e.* if $P = \varnothing$, then we can find a subset $A'\mathbf{x} \leq \mathbf{b}'$ containing at most $d+1$ of the original constraints, such that already

$$\{\mathbf{x} \;:\; A'\mathbf{x} \leq \mathbf{b}'\} = \varnothing\,.$$

In other words, if a system of linear inequalities has no solution, then there is a subset of at most $d+1$ of the inequalities that is already infeasible. yet in other words, we know that there is a short proof of infeasibility. The proof of this result is not difficult, you may attempt this with Problem 6.2.

Problem 6.2

We may ask if the theorem is still true if we ask for *integer* solutions instead of any solution. It is, however, not hard to see that the claim already fails in dimension 2. To see this we can look at the polytope

$$P := \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;:\; \pm\left(x_1 - \frac{1}{2}\right) \pm \left(x_2 - \frac{1}{2}\right) \leq \frac{3}{4} \right\}$$

This is a square with four facets, and $P \cap \mathbb{Z}^2 = \varnothing$. See also Figure 6.1. However, if we

[12] Ravi Kannan, "Improved Algorithms for Integer Programming and Related Lattice Problems".

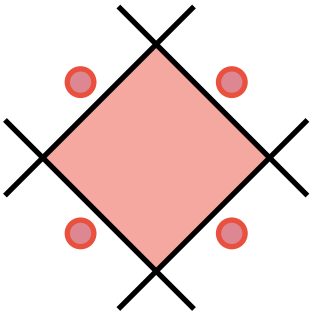


Figure 6.1.: A set of $4 = 2^2$ inequalities that are integer infeasible, but each subset is feasible.

remove any facet, then we obtain an unbounded polyhedron that contains points from $\mathbb{Z}^2$.

So a direct translation does not work. If you experiment further in dimension 2 you may get the impression that with more inequalities you can always discard one. Hence, we may ask if the claim holds if we replace $d+1$ by a larger number. Indeed, this is possible, and this is a theorem of Doignon[13] from 1973.

Theorem 6.6. *Let $A\mathbf{x} \leq \mathbf{b}$ be a linear system $A\mathbf{x} \leq \mathbf{b}$ of inequalities that is* integer infeasible*, i.e. $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\} \cap \mathbb{Z}^d = \varnothing$. Then there is a subsystem $A'\mathbf{x} \leq \mathbf{b}$ of at most 2^n rows, such that already $A'\mathbf{x} \leq \mathbf{b}'$ is integer infeasible.*

This has also been proved by Bell[14] and by Scarf.[15]

Problem 6.3

Proof. Let $\mathbf{a}_i^t\mathbf{x} \leq \beta_i$ for $\leq i \leq m$ be a set of linear inequalities in R^d that has no integer solution, but any subset has an integral solution. This implies that we can find, for each i, an integer point $\mathbf{x}_i$ such that

$$\mathbf{a}_i^t\mathbf{x}_i > \beta_i \qquad \text{and} \qquad \mathbf{a}_j^t\mathbf{x}_i \leq \beta_j \quad \text{for } i \neq j\,.$$

In particular, all $\mathbf{x}_i$ are distinct. We set

$$S := \mathsf{conv}(\mathbf{x}_1 \ldots, \mathbf{x}_m) \cap Z^d\,.$$

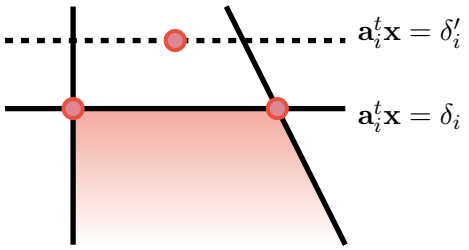
Now look at the set D of all tuples $(\delta_1, \ldots, \delta_m)$ such that

$$\delta_i \geq \min(\mathbf{a}_i^t\mathbf{x} : \mathbf{a}_i^t\mathbf{x} > \beta_i \text{ and } \mathbf{x} \in S) \tag{6.4}$$

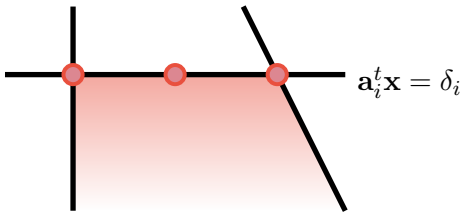
while the system

$$\mathbf{a}_j^t\mathbf{x} < \delta_j \qquad \text{for} \qquad 1 \leq j \leq m \tag{6.5}$$

[13] Doignon, "Convexity in cristallographical lattices".
[14] Bell, "A theorem concerning the integer lattice".
[15] Scarf, "An observation on the structure of production sets with indivisibilities".



(a) If all lattice points are on the boundary of the facet for $\mathbf{a}_i$, then we can increase δ_i to δ_i'



(b) If we cannot push out, then there is a lattice point in the relative interior.

Figure 6.2.: Finding lattice points in the relative interior.

has no solution in S. The latter condition means that the polyhedron

$$\left\{ \mathbf{x} \,:\, \mathbf{a}_j^t\mathbf{x} \leq \delta_j \text{ for } 1 \leq j \leq m \right\}$$

contains no point of S in its interior. We know that we have at least one such tuple $(\delta_1, \ldots, \delta_m)$ as we can certainly choose

$$\delta_i \;=\; \min(\mathbf{a}_i^t\mathbf{x} \;:\; \mathbf{a}_i^t\mathbf{x} \;>\; \beta_i \text{ and } \mathbf{x} \in S) \qquad \text{for all} \qquad i\,.$$

So D is not empty. It is also bounded, as $\delta_i > \mathbf{a}_j^t\mathbf{x}_j$ would imply that $\mathbf{x}_j$ is a solution of (6.5). The set of all δ_i that satisfy (6.4) and the condition that there is $\mathbf{s} \in S$ such that

$$\mathbf{a}_j^t\mathbf{s} \;<\; \beta_j \qquad \text{for} \qquad 1 \;\leq\; j \;\leq\; m$$

is an open set, so D is also closed. This implies that the sum $\delta_1 + \cdots + \delta_m$ assumes its maximum in D. Let $\delta_1, \cdots, \delta_m$ be such a choice of parameters.

We claim that there are $\mathbf{y}_i \in S$ such that

$$\mathbf{a}_i^t\mathbf{y}_i \;=\; \delta_i \qquad \text{and} \qquad \mathbf{a}_j^t\mathbf{y}_i \;<\; \delta_j \quad \text{for } i \;\neq\; j\,.$$

Assume that this is not the case. Then there is some i such that all $\mathbf{y} \in S$ with $\mathbf{a}_i^t\mathbf{y} = \delta_i$ satisfy at least one other inequality of our system with equality. So all point of S that are on the hyperplane

$$\left\{ \mathbf{x} \,:\, \mathbf{a}_i^t\mathbf{y} \;=\; \delta_i \right\}$$

are on the boundary of the facet defined by this inequality. Hence, we can increase δ_i slightly to some $\delta_i' > \delta_i$ without violating (6.5), see Figure 6.2. Hence, our choice was not maximal, a contradiction.

If now $m > 2^d$, then there is a pair $\mathbf{y}_j, \mathbf{y}_k$ for $j \neq k$ such that

$$\mathbf{y}_j \;\equiv\; \mathbf{y}_k \mod 2\,.$$

Hence, $\mathbf{y} := \frac{1}{2}(\mathbf{y}_j + \mathbf{y}_k) \in S$ and satisfies (6.5). This is a contradiction, so $m \leq 2^d$. □

Problem 6.4

6.4. Problems

6.1. Show that for any lattice $\lambda_d(\Lambda)\lambda_1(\Lambda^\star) \leq 2^{\mathcal{O}(d^2)}$.

6.2. Let $A\mathbf{x} \leq \mathbf{b}$ be a system of linear inequalities in $\mathbb{R}^d$. Show that, if this system is infeasible, then there is a subsystem $A'\mathbf{x} \leq \mathbf{b}'$ of at most $d+1$ rows that is already infeasible.

6.3. Prove that the bound of Doignon's Theorem is tight by constructing a polytope with 2^d facets that contains no integral point, but removing any facet produces a polyhedron with interior lattice points.

6.4. Let $K_1, \ldots, K_m \subseteq \mathbb{R}^d$ be convex and bounded with

$$\left(\bigcap_{i=1}^{m} K_i\right) \cap \mathbb{Z}^d = \varnothing\,.$$

Show that there is a subset $I \subseteq \{1, \ldots, m\}$ of size at most 2^d such that

$$\left(\bigcap_{i\in I} K_i\right) \cap \mathbb{Z}^d = \varnothing\,.$$

7. The Subset Sum Problem

This chapter introduces another application of the LLL algorithm. We will consider the *subset sum* or *knapsack* problem. The general case of this problem is hard (NP-complete) to solve. However, we will see that a special case can be solved in polynomial time using a reduced basis. This has also a nice application in cryptography, which we will explain in the next section, before we discuss an algorithm using LLL to break certain encryption schemes based on knapsacks. .

Let us first formalize the general problem.

(SubsetSum). Let $\mathbf{a} := (a_1, \ldots, a_d)$ be positive integer weights, and $s \in \mathbb{Z}_{>0}$ such that there is $\mathbf{x} = (x_1, \ldots, x_d) \in \{0,1\}^d$ with $s = \sum x_i a_i$.
The *Subset Sum Problem* is the task to find $\mathbf{x}$ given $\mathbf{a}$ and s.

In this general formulation the problem is NP-complete. But this applies to the worst case. It might still be that some or most instances can be solved efficiently. We will see below that some *structured* instances can be solved easily, and if the bit length of all a_i is large compared to n, then we can use the LLL-algorithm to find $\mathbf{x}$ with high probability.

7.1. A knapsack cryptosystem

Many NP-hard problems attract researches in the field of cryptography, as they may serve as a basis for a strong public key encryption schemes (as long as P$\neq$ NP).

In our case we want to look at knapsack problems, and the scheme would consist of a choice of weights $\mathbf{a} := (a_1, \ldots, a_d)$ somehow chosen from a specified distribution. Given some data $\mathbf{x} := (x_1, \ldots, x_d) \in \{0,1\}^d$, the encryption computes the sum

$$s := \sum_{i=1}^{d} a_i x_i$$

This process of encryption is pretty simple, which is a major advantage over other public key encryption schemes, which use modular exponentiation or elliptic curves. As recovering the encrypted text requires us to solve the subset sum instance with input data $\mathbf{a}$ and s, which is a NP-complete problem, it is also secure.

However, in this form it is not yet useful. For the true receiver of the message, who generated the set of weights $\mathbf{a}$, it should be possible to recover the original message without solving an NP-hard problem. To make this feasible, we have to find an instance of a subset sum problem, which is easy to solve with some additional information on the system, but which is hard to solve without. In particular, it should not be possible to deduce the required additional information from the publicly known weights.

Hence, we need an instance of the subset sum problem which is easy to solve, together with a method to hide this. Here is one option of an easily solvable subset sum instance. We say that a sequence $a_1, a_2, \ldots, a_d$ of weights is *superincreasing* if

$$a_j > \sum_{i=1}^{j-1} a_i \qquad \text{for} \qquad 1 \leq j \leq n\,.$$

Given such a sequence we can easily solve any subset sum instance. Namely, if we have the total weight s, then we check if $s \geq a_d$. If it is, then necessarily $x_d = 1$, and otherwise $x_d = 0$. We can now replace s with $s - x_d a_d$. The remaining weights are still superincreasing, so we can repeat the same process with a_{n-1} to find x_{n-1}. We can continue in this way until we arrive at a_1. At this point we must have recovered the original message $\mathbf{x}$.

Now we have an easily solvable instance. However, if we now publish these weights, then also any attacker can solve the problem. We must somehow disguise this special property of our weights. The following method to do this was proposed by Merkle and Hellman.[1]

(i) Choose a superincreasing sequence $b_1, \ldots, b_d$.

(ii) Choose a number $N > \sum b_i$ and a nonzero number $m \in \mathbb{Z}_N$.

(iii) Choose a permutation π of $(1, \ldots, n)$.

Now we set

$$a_i := m \cdot b_{\pi(i)} \mod N \qquad \text{for} \qquad 1 \leq j \leq n\,.$$

Then our public key is $\mathbf{a} = (a_1, \ldots, a_d)$ and the required data to obtain a simple solution is (N, m, π).

Given a message $\mathbf{x} = (x_1, \ldots, x_d)$ we receive

$$s := \sum x_i a_i = m \cdot \sum x_i b_{\pi(i)}\,.$$

To decrypt this message using our additional data we first compute $t := s/m$. By the choice of N we know that $t < N$, so this the true sum for the weights $b_1, \ldots, b_d$. Now we can solve the problem using this permuted superincreasing sequence.

At first sight this might be a good encryption scheme, if we choose our superincreasing sequence suitably. In particular, it should certainly not follow any pattern, like setting $b_i = 2^i$ or similar. However, we will see in the next section that this scheme is in fact inherently insecure.

7.2. Solving Sparse Knapsack Instances in Polynomial Time

We now want to show that in certain cases we can use the computation of a short vector to solve the knapsack instance **(SubsetSum)** in polynomial time. The systems we want to look at are the *sparse* knapsack systems, where we call a system *sparse* if its *density*,

[1] Merkle and Hellman, "Hiding information and signatures in trapdoor knapsacks".

which is defined as

$$\frac{d}{\max_{1\le i\le d}(\log\, a_i)}$$

is bounded by $\mathcal{O}(\frac{1}{d})$ and we choose the a_i uniformly at random from

$$a_i \;\in\; \{\,1,\ldots,N\,\}\;.$$

for $N := 2^{d^2(\frac{1}{2}+\varepsilon)}$ and small $\varepsilon > 0$.

Let $\mathbf{x} \in \{\,0,1\,\}^d$ be the plain message, and $s := \sum a_i x_i$ be the knapsack sum. We may assume that $s > \frac{1}{2}\sum_{i=1}^{d} a_i$, otherwise we replace s by $\sum_{i=1}^{d} a_i - s$ and flip zeros and ones in the solution that we compute. This implies in particular, that $\mathbf{x} \neq 0$. For such systems, Lagarias and Odlydzko,[2] with later simplifications by Frieze[3] have shown that we can solve them efficiently with high probability.

Theorem 7.1. *There is a polynomial time algorithm that, with high probability, on input* $(\mathbf{a}, s)$ *recovers* $\mathbf{x}$.

The key idea for the proof is to define a lattice Λ of rank $d+1$ in which the desired solution $\mathbf{x}$ is a shortest nonzero lattice vector and, moreover, all lattice vectors not parallel to this $\mathbf{x}$ have length at least $2^{d/2}\,\|\mathbf{x}\|$. Now we can efficiently solve **(SVP)**$_\gamma$ for $\gamma = 2^{d/2}$ (see Corollary 4.9 and recall that our lattice will have rank $d+1$). Hence, by construction, the approximation we compute with this algorithm must actually be $\mathbf{x}$.

Proof. Let $M := \lceil \sqrt{d}\cdot 2^{d/2} + 1 \rceil$ and define the lattice Λ as the integral span of the columns of

$$\mathrm{B} \;:=\; \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ -Ma_1 & -Ma_2 & \cdots & -Ma_d & Ms \end{bmatrix},$$

where all entries in the matrix that are not specified are 0.

For the coefficient vector $\mathbf{u} := \left[\begin{smallmatrix}\mathbf{x}\\ 1\end{smallmatrix}\right]$, where $\mathbf{x} \in \{0,1\}^d$ is the solution we search for, we have $\mathrm{B}\,\mathbf{u} = \left[\begin{smallmatrix}\mathbf{x}\\ 0\end{smallmatrix}\right] \neq 0$ (recall that $s = \sum a_i x_i$) and this vector has norm at most $\sqrt{d}$.

Now we use the LLL-algorithm to compute an approximate shortest vector $\mathbf{v}$ as in Corollary 4.9. This vector satisfies

$$\|\mathbf{v}\| \;\le\; 2^{d/2}\lambda_1(\Lambda)\,, \tag{7.1}$$

We want to show that the last coordinate v_{d+1} of such a vector is necessarily 0. Assume this is not the case. By our choice of a basis the last coordinate w_{d+1} of any lattice vector

[2] J. C. Lagarias and Odlyzko, "Solving low-density subset sum problems".
[3] Frieze, "On the Lagarias-Odlyzko algorithm for the subset sum problem".

$\mathbf{w} \in \Lambda$ is divisible by M. Thus, if $v_{d+1} \neq 0$, then

$$\|\mathbf{v}\| \;\geq\; M \;>\; 2^{d/2}\,\|\mathbf{x}\| \;\geq\; 2^{d/2}\lambda_1(\Lambda)\,,$$

which is a contradiction to (7.1). Hence, we know that $v_{d+1} = 0$.

We will now show that with high probability the vector $\mathbf{v}$ must be a multiple of $\mathbf{x}' := \left[\begin{smallmatrix}\mathbf{x}\\0\end{smallmatrix}\right]$, but let us first discuss why this will prove the claim. For this we first recall that the construction of the approximation of a shortest vector in $\textbf{(SVP)}_\gamma$ applies the LLL-algorithm and then returns the first vector of the computed basis. So the vector $\mathbf{v}$ that we have computed is part of a basis. As such, $\mathbf{v}$ must be primitive, *i.e.* $\mathbf{v}$ cannot be a multiple $\lambda\mathbf{w}$ of a lattice vector for some $\lambda \in \mathbb{Z}$ with $|\lambda| \geq 2$. This implies

$$\mathbf{v} \;=\; \pm\left[\begin{smallmatrix}\mathbf{x}\\0\end{smallmatrix}\right]\,,$$

so we know $\mathbf{x}$ up to sign. But $\mathbf{x} \in \{0,1\}^d$, so we just multiply by -1 if $\mathbf{x} \leq 0$.

Now to prove that $\mathbf{v}$ is a multiple of $\mathbf{x}'$, we show that with high probability, multiples of $\left[\begin{smallmatrix}\mathbf{x}\\0\end{smallmatrix}\right]$ are the only lattice vectors that have a length bounded by $2^{d/2}\sqrt{d} < M$.

So let $\left[\begin{smallmatrix}\mathbf{z}\\0\end{smallmatrix}\right] \in \mathbb{Z}^{d+1}$ be any nonzero lattice vector with $\|\mathbf{z}\| < 2^{d/2}\sqrt{d}$ that is *not* a multiple of $\left[\begin{smallmatrix}\mathbf{x}\\0\end{smallmatrix}\right]$. We have to bound the probability that

$$\begin{bmatrix}\mathbf{z}\\0\end{bmatrix} \;=\; \mathrm{B}\begin{bmatrix}\mu\\\eta\end{bmatrix}$$

is a vector in Λ for some $\mu \in \mathbb{Z}^d$ and $\eta \in \mathbb{Z}$. Note that, by our choice of the basis, the first d coefficients must coincide with the entries of the vector, *i.e.*

$$\mu_i \;=\; z_i \qquad \text{for} \qquad 1 \;\leq\; i \;\leq\; d \qquad \text{and} \qquad \|\mathbf{z}\| \;=\; \|\mu\|\,. \tag{7.2}$$

In this case we have

$$s \cdot |\eta| \;=\; |s \cdot \eta| \;=\; |\sum_{i=1}^{d} \mu_i a_i| \;\leq\; \|\mu\| \sum_{i=1}^{d} a_i\,.$$

By assumption, $s > \frac{1}{2}\sum_{i=1}^{d} a_i$, so we get

$$|\eta| \;\leq\; 2\,\|\mu\|\,. \tag{7.3}$$

If $\left[\begin{smallmatrix}\mathbf{z}\\0\end{smallmatrix}\right]$ is in Λ, then

$$\sum_{i=1}^{d} \mu_i a_i \;=\; \eta \cdot s \;=\; \eta \sum_{i=1}^{d} x_i a_i\,,$$

so, for $\xi := \mu_i - \eta x_i$

$$\sum_{i=1}^{d} \xi_i a_i \;=\; 0\,.$$

By assumption, $\left[\begin{smallmatrix}\mathbf{z}\\0\end{smallmatrix}\right]$ is not a multiple of $\left[\begin{smallmatrix}\mathbf{x}\\0\end{smallmatrix}\right]$, so not all ξ_i can be zero. Thus, up to reordering we may assume that $\xi_1 \neq 0$. This implies

$$a_1 \;=\; -\frac{1}{\xi_1}\sum_{i=2}^{d}\xi_i a_i\,.$$

Hence, for any fixed (μ, η) satisfying the constraints, the probability that $\left[\begin{smallmatrix}\mathbf{z}\\0\end{smallmatrix}\right]$ is in Λ is bounded by

$$\Pr_{a_i\in\{\,1,\ldots\,N\,\}}\left[\sum_{i=1}^{d}\xi_i a_i \;=\; 0\right] \;=\; \Pr_{a_1\in\{\,1,\ldots\,N\,\}}\left[a_1 \;=\; -\frac{1}{\xi_1}\sum_{i=2}^{d}\xi_i a_i\right] \;\leq\; \frac{1}{N}\,,$$

as the a_i are chosen uniformly at random in $\{\,1,\ldots,N\,\}$, for $N = 2^{d^2(\frac{1}{2}+o(1))}$.

Now there are at most

$$(2M+1)^d\cdot(4M+1) \;\leq\; (5M)^{d+1} \;\leq\; 2^{d^2(\frac{1}{2}+o(1))}$$

possible choices for (μ,η). This rather crude bound follows from $\|\mathbf{z}\| \leq M$, so $|z_i| \leq M$, and the observations in (7.2) and (7.3).

This shows that also the total probability that there exists any $\left[\begin{smallmatrix}\mathbf{z}\\0\end{smallmatrix}\right] \in \Lambda$ satisfying the above constraints is at most $2^{-\Omega(d^2)}$, which is small. □

8. The Closest Vector Problem

In Chapter 4 we discussed how we may compute the short vector that is guaranteed by Minkowski's First Theorem (Corollary 3.3). We showed that we can approximate the vector up to some constant with a reduced basis, and then find the true shortest vector with a simple enumeration.

In this chapter we want to consider a seemingly similar problem, the *Closest Vector Problem (CVP)* and some variants. Here, we are given a lattice $\Lambda \subseteq \mathbb{R}^d$ and some point $\mathbf{t} \in \mathbb{R}^d$ and we want to find a lattice point $\mathbf{v} \in \Lambda$ that minimizes $\|\mathbf{v} - \mathbf{t}\|$.

Yet, despite its similarity, we need a new technique. Fortunately, our algorithms will still use the same tools that we developed in Chapter 3 and Chapter 5. In particular, we will uses reduced bases and their properties again.

Let us first give a formal definition of the problem.

(CVP). Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice and $\mathbf{t} \in \mathbb{R}^d$. The *Closest Vector Problem (CVP)* is the task to find a vector $\mathbf{u} \in \Lambda$ that minimizes $\|\mathbf{u} - \mathbf{t}\|$.

As for the shortest vector problem it proves to be much harder to find the exact solution compared to an approximate solution. We may consider two variants for the approximation. Either, we want to find a vector $\mathbf{u}$ whose distance from $\mathbf{t}$ is at most some factor γ longer than the optimal distance, or we provide a bound $r > 0$ and ask for a point $\mathbf{u}$ at distance at most r from $\mathbf{t}$. Let us make this precise. Here is version to find a relative approximation.

(CVP)$_\gamma$. Let $\gamma \geq 1$, $\Lambda \subseteq \mathbb{R}^d$ a lattice and $\mathbf{t} \in \mathbb{R}^d$. The *(Relative) Approximate Closest Vector Problem (CVP)$_\gamma$* is the task to find $\mathbf{u} \in \Lambda \setminus \{\mathbf{0}\}$ at distance at most

$$\|\mathbf{u} - \mathbf{t}\| \;\leq\; \gamma \cdot \min_{\mathbf{v} \in \Lambda} \|\mathbf{t} - \mathbf{v}\| \;.$$

In some problems it may be more interesting to have an absolute bound and consider the following problem

(AbsCVP)$_r$. Let $r > 0$, $\Lambda \subseteq \mathbb{R}^d$ a lattice and $\mathbf{t} \in \mathbb{R}^d$. The *Approximate Absolute Closest Vector Problem (AbsCVP)$_r$* is the task to find $\mathbf{u} \in \Lambda \setminus \{\mathbf{0}\}$ at distance at most

$$\|\mathbf{u} - \mathbf{t}\| \;\leq\; r \,.$$

Clearly, if r is too small, then this problem will not have a solution.

We will discuss two methods below. The first one, the nearest plain algorithm by

Babai[1] only solves **(CVP)**$_\gamma$, with $\gamma = 2^{d/2}$. However, it runs in polynomial time.

The second algorithm expands the idea of this algorithm and refines the method where in Babai's algorithm we may miss the exact solution. Thus, the algorithm solves the exact problem **(CVP)**. The main drawback is, that we need a similar technique as in Section 6.2 and use recursion in lower dimension on a bounded number of hyperplanes. Here, we need to search $\mathcal{O}(2^d)$ hyperplanes, which makes the algorithm only polynomial if the dimension is not part of the input. Its total running time is $\mathcal{O}(2^{d^2})$.

Micciancio and Voulgaris[2] have shown that one can solve the problem in $\mathcal{O}(2^d)$ using Voronoi cells around lattice points, which is, as far as I know, the fastest known algorithm for this problem.

There a more methods known, but nothing that runs in polynomial time if the dimension is part of the input. In fact, the exact complexity class is not known. In particular it is not known whether the closest vector problem is NP-hard. On the other hand we do know that it is not more difficult than the shortest vector problem.[3]

8.1. Babai's Nearest Plane Algorithm

Recall the discussion at the beginning of Chapter 4. There we observed that solving **(SVP)** is simple if all basis vectors of the lattice are pairwise orthogonal. The same is true for **(CVP)**.

Assume that $\mathbf{b}_1, \dots, \mathbf{b}_d$ is a lattice basis of pairwise orthogonal vectors, and we want to determine the closest lattice vector to some $\mathbf{t} \in \mathbb{R}^d$. We can write

$$\mathbf{t} = \sum_{i=1}^{d} \mu_i \mathbf{b}_i$$

for some $\mu_i \in \mathbb{R}^d$. If $\mathbf{u}$ is any lattice vector, then we can write $\mathbf{u}$ as

$$\mathbf{u} = \sum_{i=0}^{d} \lambda_i \mathbf{b}_i$$

for $\lambda_i \in \mathbb{Z}$. Thus,

$$\|\mathbf{u} - \mathbf{t}\|^2 = \sum_{i=1}^{d} (\mu_i - \lambda_i)^2 \|\mathbf{b}_i\|^2 . \tag{8.1}$$

This sum is minimal, when each of the $(\mu_i - \lambda_i)^2$ is minimal. As $\lambda_i \in \mathbb{Z}$, this is the case if $\lambda_i = \lceil \mu_i \rfloor$, *i.e.* the rounding of μ_i to the nearest integer.

We want to apply the same idea, rounding coefficients to the nearest integer, also for general lattice bases. These are usually not orthogonal, not even if we apply the

[1] Babai, "On Lovász' lattice reduction and the nearest lattice point problem".
[2] Daniele Micciancio and Voulgaris, "A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations [extended abstract]".
[3] Goldreich, D. Micciancio, Safra, and Seifert, "Approximating shortest lattice vectors is not harder than approximating closest lattice vectors".

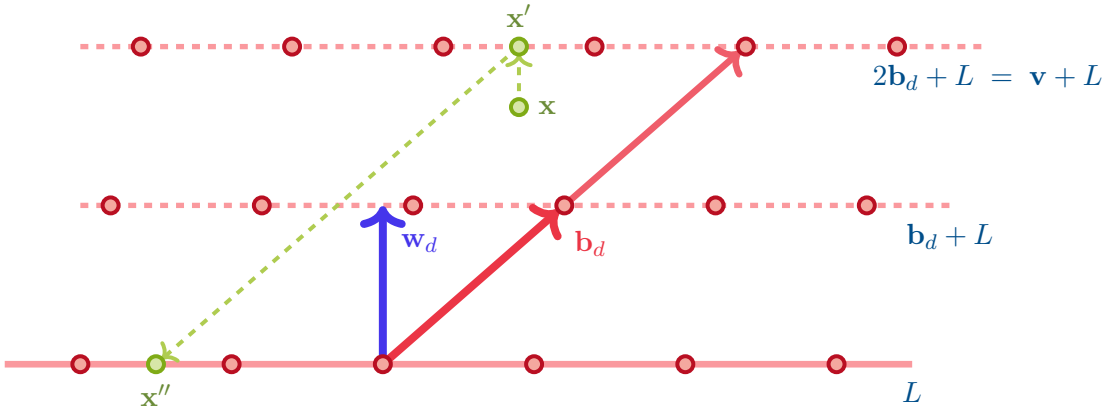


Figure 8.1.: The nearest plane algorithm

LLL-algorithm. We cannot expect to find the exact closest vector with such a method, but we will find a $2^{d/2}$-approximation to a vector $\mathbf{x}$, *i.e.* a vector $\mathbf{v} \in \Lambda$ such that

$$\|\mathbf{x} - \mathbf{v}\| \;\leq\; 2^{d/2} \cdot \|\mathbf{x} - \mathbf{u}\|$$

where $\mathbf{u} \in \Lambda$ is a closest vector to $\mathbf{x}$ in Λ. This goes back to work of Babai.[4]

So we want to round the coefficients of some representation of $\mathbf{x}$ successively to their nearest integer. In this process, being orthogonal will turn out more important than being a lattice basis, so we will not do this for the coefficients in the lattice basis, but for the coefficients in the associated Gram-Schmidt basis.

Let us first consider this geometrically, before we compute our approximation algebraically. We are given a lattice Λ with basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ and some point $\mathbf{x} \in \mathbb{R}^d$. The method will use recursion over the dimension. For this, we define

$$L \;:=\; \operatorname{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}) \qquad \text{and} \qquad \Lambda' \;:=\; L \cap \Lambda\,.$$

The lattice Λ' is spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}$. As before, we denote by $\mathbf{w}_1, \ldots, \mathbf{w}_d$ the Gram-Schmidt orthogonalization of the lattice basis. Let $\mu_1, \ldots, \mu_d$ be the coefficients of $\mathbf{x}$ in the basis $\mathbf{w}_1, \ldots, \mathbf{w}_d$, *i.e.*

$$\mathbf{x} \;:=\; \sum_{i=1}^{d} \mu_i \mathbf{w}_i\,.$$

All lattice points of Λ are contained in one of the hyperplanes

$$k \cdot \mathbf{b}_d \;+\; L \;=\; k \cdot \mathbf{w}_d \;+\; L$$

for some $k \in \mathbb{Z}$. The idea of the algorithm is to find k such that the distance from $\mathbf{x}$ to $k\mathbf{b}_d + L$ is minimal, see Figure 8.1. Let $\mathbf{x}'$ be the orthogonal projection of $\mathbf{x}$ onto

[4] Babai, "On Lovász' lattice reduction and the nearest lattice point problem".

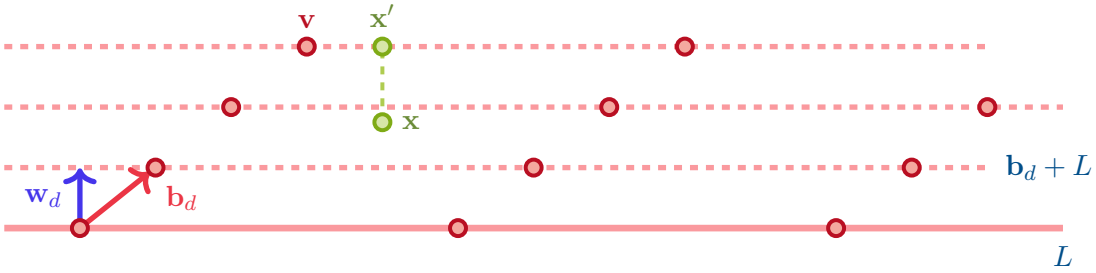


Figure 8.2.: The closest lattice point need not be on the closest lattice hyperplane

$k\mathbf{b} + L$, and set

$$\mathbf{x}'' := \mathbf{x}' - k\mathbf{b}_d\,.$$

Now we recursively find an approximate solution $\mathbf{v}''$ to the closest vector problem for the point $\mathbf{x}''$ in the lower dimensional space L. Then

$$\mathbf{v} := \mathbf{v}'' + k\mathbf{b}_d$$

is the solution to the original problem. Note that $k = \lfloor \mu_d \rceil$, so we use the same rounding for the coefficients as we needed in (8.1).

However, we cannot expect to obtain the true closest vector with this procedure. Figure 8.2 shows an example where the closest vector $\mathbf{v}$ to $\mathbf{x}$ is not on the nearest hyperplane. As we are recursing to find the closest lattice point to a projection on the nearest plane we will never find a closest point in this example.

The orthogonal projection of $\mathbf{x}$ onto $k\mathbf{b}_d + L$ is

$$\mathbf{x}' := \sum_{i=1}^{d-1} \mu_i \mathbf{w}_i + \lfloor \mu_d \rceil \mathbf{w}_d\,.$$

By construction, this point has minimal distance from $\mathbf{x}$ among all points in $k\mathbf{b} + L$. Recall from (4.3) that

$$\mathbf{b}_d = \mathbf{w}_d + \sum_{i=1}^{d-1} \lambda_{id} \mathbf{w}_i \tag{8.2}$$

We compute the representation of $\mathbf{x}''$,

$$\begin{aligned}
\mathbf{x}'' &= \mathbf{x}' - k\mathbf{b}_d \\
&= \sum_{i=1}^{d-1} \mu_i \mathbf{w}_i + \lfloor \mu_d \rceil \mathbf{w}_d - \lfloor \mu_d \rceil \left(\mathbf{w}_d + \sum_{i=1}^{d-1} \lambda_{id} \mathbf{w}_i \right) \\
&= \sum_{i=1}^{d-1} (\mu_i - \lfloor \mu_d \rceil \lambda_{id})\, \mathbf{w}_i\,.
\end{aligned} \tag{8.3}$$

We now have a full description of the algorithm. With (8.3) we have a representation of $\mathbf{x}''$ in the Gram-Schmidt orthogonalization of the basis $\mathbf{b}_1, \dots, \mathbf{b}_{d-1}$ of the lattice Λ'. We continue by rounding the last coefficient $\mu_i - \lfloor \mu_d \rceil \lambda_{id}$ to the nearest integer. We discuss below that we can simplify this slightly. There we show that we omit the projection of $\mathbf{x}$ onto $k\mathbf{b}_d + L$ and use $\mathbf{x} - k\mathbf{b}_d$ instead of $\mathbf{x}''$.

We have not yet discussed how close we get to an exact closest vector. For this we have to assume that our basis $\mathbf{b}_1, \dots, \mathbf{b}_d$ is LLL-reduced. The following lemma bounds the approximation by the lengths of the Gram-Schmidt vectors. We use this in the next theorem to compute the approximation factor of the algorithm.

Lemma 8.1. *Let $\mathbf{b}_1, \dots, \mathbf{b}_d$ be an LLL-reduced lattice basis of Λ with Gram-Schmidt vectors $\mathbf{w}_1, \dots, \mathbf{w}_d$ and $\mathbf{x} \in \mathbb{R}^d$. Let further $\mathbf{v}$ be the lattice point returned by Babai's algorithm for the input $\mathbf{x}$. Then*

$$\|\mathbf{x} - \mathbf{v}\|^2 \leq \left(2^{d-2} - \frac{1}{4}\right) \|\mathbf{w}_d\|^2$$

Proof. We use induction. For $d = 1$ the closest vector to $\mathbf{x} = \mu \mathbf{w}_1$ is $\mathbf{v} = \lfloor \mu \rceil \mathbf{b}_1$, as $\mathbf{w}_1 = \mathbf{b}_1$. The claim follows, as $(\mu - \lfloor \mu \rceil)^2 \leq \frac{1}{4}$.

Let $\mathbf{v} = \mathbf{v}' + \lfloor \mu_d \rceil \mathbf{b}_d$ be the result of the algorithm, where $\mathbf{v}'$ is the closest vector computed in the lattice Λ'. By induction

$$\|\mathbf{v}' - \mathbf{x}''\|^2 \leq \left(2^{d-3} - \frac{1}{4}\right) \|\mathbf{w}_{d-1}\|^2 .$$

Hence,

$$\begin{aligned}
\|\mathbf{x} - (\mathbf{v}' + \lfloor \mu_d \rceil \mathbf{b}_d)\|^2 &= \|\mathbf{x} - \mathbf{x}' + \mathbf{x}' - (\mathbf{v}' + \lfloor \mu_d \rceil \mathbf{b}_d)\|^2 \\
&= \|\mathbf{x} - \mathbf{x}'\|^2 + \|\mathbf{x}'' - \mathbf{v}'\|^2 \\
&\leq \frac{1}{4} \|\mathbf{w}_d\|^2 + \left(2^{d-3} - \frac{1}{4}\right) \|\mathbf{w}_{d-1}\|^2 \\
&= \left(\frac{1}{4} + 2 \cdot \left(2^{d-3} - \frac{1}{4}\right)\right) \|\mathbf{w}_d\|^2 \\
&= \left(2^{d-2} - \frac{1}{4}\right) \|\mathbf{w}_d\|^2
\end{aligned}$$

where we have used (4.8) in the second but last equation. □

Problem 8.1
Problem 8.2

Theorem 8.2. *Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be an LLL-reduced basis of Λ and $\mathbf{x} \in \mathbb{R}^d$. Let further $\mathbf{v}$ be the lattice point returned by Babai's algorithm. Then*

$$\|\mathbf{x} - \mathbf{v}\| \;\leq\; 2^{d/2} \|\mathbf{x} - \mathbf{u}\| \qquad \textit{for any} \qquad \mathbf{u} \in \Lambda\,.$$

Proof. We use induction again. For $d = 1$ the nearest plane algorithm returns the exact closest vector, so the bound holds.

Let $d \geq 2$ and let $\mathbf{u}$ be a closest vector to $\mathbf{x}$. Recall from Figure 8.2 that $\mathbf{u}$ need not be on the hyperplane $k\mathbf{b}_d + L$ that we choose in the algorithm. We discuss the case $\mathbf{u} \in k\mathbf{b}_d + L$ and $\mathbf{u} \notin k\mathbf{b}_d + L$ separately. Let $\mathbf{x}'$ be the orthogonal projection of $\mathbf{x}$ onto $k\mathbf{b}_d + L$.

In the first case the vector $\mathbf{u} - k\mathbf{b}_d$ is a closest vector to $\mathbf{x}'' = \mathbf{x}' - k\mathbf{b}_d$, and we know by induction that

$$\|\mathbf{x}'' - \mathbf{v}'\| \;\leq\; 2^{(d-1)/2} \|(\mathbf{u} - k\mathbf{b}_d) - \mathbf{x}''\|\,,$$

where $\mathbf{v}'$ is the vector returned by the algorithm applied to $\mathbf{x}''$ in L. Using

$$\mathbf{x}'' - \mathbf{v}' \;=\; (\mathbf{x}' - k\mathbf{b}_d) - \mathbf{v}' \;=\; \mathbf{x}' - (\mathbf{v}' + k\mathbf{b}_d) \;=\; \mathbf{x}' - \mathbf{v}$$

and

$$(\mathbf{u} - k\mathbf{b}_d) - \mathbf{x}'' \;=\; \mathbf{u} - (\mathbf{x}'' + k\mathbf{b}_d) \;=\; \mathbf{u} - \mathbf{x}'$$

we obtain

$$\|\mathbf{x}' - \mathbf{v}\| \;\leq\; 2^{(d-1)/2} \|\mathbf{u} - \mathbf{x}'\|\,.$$

The Theorem of Pythagoras now implies

$$\begin{aligned}
\|\mathbf{x} - \mathbf{v}\|^2 \;&=\; \|\mathbf{x} - \mathbf{x}'\|^2 \;+\; \|\mathbf{x}' - \mathbf{v}\|^2 \\
&\leq\; \|\mathbf{x} - \mathbf{x}'\|^2 \;+\; 2^{d-1} \|\mathbf{u} - \mathbf{x}'\|^2 \\
&\leq\; \|\mathbf{x} - \mathbf{u}\|^2 \;+\; 2^{d-1} \|\mathbf{u} - \mathbf{x}\|^2 \\
&\leq\; 2^d \|\mathbf{x} - \mathbf{u}\|^2\,,
\end{aligned}$$

where the second inequality follows from the triangle inequality.

Now assume that $\mathbf{u} \notin k\mathbf{b}_d + L$. The hyperplanes $\mu\mathbf{b}_d + L$ for different $\mu \in \mathbb{Z}$ are at least $\frac{1}{2}\|\mathbf{w}_d\|$ apart, so

$$\|\mathbf{x} - \mathbf{u}\|^2 \;\geq\; \frac{1}{4} \|\mathbf{w}_d\|^2\,.$$

The previous Lemma 8.1 implies

$$\|\mathbf{x} - \mathbf{v}\|^2 \;\leq\; \frac{1}{4}(2^d - 1) \|\mathbf{w}_d\|^2 \;\leq\; (2^d - 1) \|\mathbf{x} - \mathbf{u}\|^2 \;\leq 2^d \|\mathbf{x} - \mathbf{u}\|^2\,. \qquad \square$$

This theorem proves that Babai's nearest plane algorithms returns a $2^{d/2}$-approximation of the closest vector. In other words, it solves **(CVP)**$_\gamma$ for $\gamma = 2^{d/2}$.

Computing the running time of this algorithm is rather simple. Let $\Lambda \subseteq \mathbb{Z}^d$ be a lattice with an LLL-reduced basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$, whose Euclidean length is bounded by B. This is the same setting as we used for the LLL-algorithm in Chapter 5 for our running time estimates, *e.g.* in Proposition 5.3, Corollary 5.4, and (5.7). You will prove in Problem 8.4 that the nearest plane algorithm on this input runs in time

$$\mathcal{O}\left(d^5 \log^2 B\right) .$$

Problem 8.4

We know from (5.7) that the LLL-algorithms requires time

$$\mathcal{O}\left(d^6 \log^3 B\right) .$$

This dominates the running time of the nearest plane algorithm on a reduced basis.

We can simplify the computations in the algorithm slightly. For this we observe that $\mathbf{x} - \mathbf{x}' = \mu \mathbf{w}_d$ for some $\mu \in \mathbb{R}$. Hence, any lattice point in L that is a closest lattice point to $\mathbf{x}'' = \mathbf{x}' - k\mathbf{b}_d \in L$ is also a closest lattice point to $\mathbf{x} - k\mathbf{b}_d$. This point does not lie in L anymore, but this does not affect the algorithm. Recall that from the representation of $\mathbf{x}$ in the Gram-Schmidt basis as

$$\mathbf{x} \;=\; \sum_{i=1}^{d} \mu_i \mathbf{w}_i$$

we obtain k as the nearest integer $\lfloor \mu_d \rceil$ of μ_d.

Instead of computing $\mathbf{x}''$, and in particular the coefficients of $\mathbf{x}''$ in the Gram-Schmidt basis, we can set $\mathbf{y} := \mathbf{x} - \lfloor \mu_d \rceil \mathbf{b}_d$ and compute the closest vector $\mathbf{v}'$ to $\mathbf{y}$ in L and return $\mathbf{v}' + \lfloor \mu_d \rceil \mathbf{b}_d$.

Now in the k-th iteration we need the highest coefficient of the current $\mathbf{y}$ in the Gram-Schmidt basis to compute $\mathbf{y} - \lfloor \mu_k \rceil \mathbf{b}_k$ and then round this to the nearest integer. We can obtain this via

$$\mu_k \;=\; \frac{\langle \mathbf{y}, \mathbf{w}_k \rangle}{\|\mathbf{w}_k\|^2} .$$

To show that this indeed is the same coefficient as used in (8.3) we compute, using (8.2),

$$\frac{\langle \mathbf{x} - \lfloor \mu_d \rceil \mathbf{b}_d, \mathbf{w}_{d-1} \rangle}{\|\mathbf{w}_{d-1}\|^2} \;=\; \frac{(\mu_{d-1} - \lfloor \mu_d \rceil \lambda_{d-1,d}) \, \|\mathbf{w}_{d-1}\|^2}{\|\mathbf{w}_{d-1}\|^2} \;=\; \mu_{d-1} - \lfloor \mu_d \rceil \lambda_{d-1,d} .$$

To summarize, on an LLL-reduced basis we initiate the result $\mathbf{v}$ with $\mathbf{0}$ and compute, starting with $j = d$ the coefficient

$$k \;:=\; \left\lfloor \frac{\langle \mathbf{x}, \mathbf{w}_j \rangle}{\|\mathbf{w}_j\|^2} \right\rceil ,$$

replace $\mathbf{v}$ with $\mathbf{v} + k\mathbf{b}_j$, $\mathbf{x}$ with $\mathbf{x} - k\mathbf{b}_d$, and j with $j-1$, and continue until $j = 0$. The approximate closest vector is now the vector $\mathbf{v}$.

Some variants or improvements to this algorithm are known, but none is significantly better. The example in Figure 8.2 suggests that the choice of the order of the basis may influence the result. Klein[5] used randomization in this choice to solve the exact **(CVP)** if the input vector is close to a lattice point.

With these arguments we can also solve the absolute closest vector approximation problem **(AbsCVP)**$_r$ for some $\gamma \geq \frac{1}{2}\sqrt{\sum_{i=1}^{d} \|\mathbf{w}_i\|^2}$. You will prove this in Problem 8.5.

Problem 8.5

* 8.2. Fundamental Domains revisited

The bound of Lemma 8.1 has a nice geometric interpretation, that we want to explore in this section.

We define a new fundamental parallelepiped for the lattice. This will use the Gram-Schmidt vectors instead of the lattice basis, so it will in fact be a fundamental rectangle.

Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice in $\mathbb{R}^d$ with basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ and Gram-Schmidt orthogonalization $\mathbf{w}_1, \ldots, \mathbf{w}_d$. We define

$$\mathrm{F}_{\mathrm{GS}}(\mathbf{b}_1, \ldots, \mathbf{b}_d) := \left\{ \sum_{i=1}^{d} \mu_i \mathbf{w}_i \ : \ 0 \leq \mu_i < 1 \text{ for } 1 \leq i \leq d \right\} .$$

More generally, we say that a set F is a fundamental domain for a lattice Λ, if

$$\mathbf{u} + \mathrm{F} \cap \mathbf{u}' + \mathrm{F} = \varnothing \quad \text{for} \quad \mathbf{u} \neq \mathbf{u}' \quad \text{and} \quad \mathbb{R}^d = \bigcup_{\mathbf{u} \in \Lambda} \mathbf{u} + \mathrm{F} .$$

It follows from (4.4) that

$$\mathrm{vol}\, \Pi(\mathbf{b}_1, \ldots, \mathbf{b}_d) = \det \Lambda = \prod_{j=1}^{d} \|\mathbf{w}_j\| = \mathrm{vol}\ \mathrm{F}_{\mathrm{GS}}(\mathbf{b}_1, \ldots, \mathbf{b}_d),,$$

so intuitively, if translates of $\mathrm{F}_{\mathrm{GS}}(\mathbf{b}_1, \ldots, \mathbf{b}_d)$ by different lattice points are disjoint, then these translates should cover $\mathbb{R}^d$. This is in fact true, as we have the following proposition.

Proposition 8.3. *The set* $\mathrm{F}(\mathbf{b}_1, \ldots, \mathbf{b}_d)$ *is a fundamental domain.*

The proof requires a lemma, which you will prove in Problem 8.6.

[5] P. Klein, "Finding the closest lattice vector when it's unusually close".

Lemma 8.4. *Let Λ be a lattice with basis $\mathbf{b}_1, \dots, \mathbf{b}_d$, L the orthogonal complement of $\mathbf{b}_1$ with projection $\pi : \mathbb{R}^d \to L$ and $\Gamma := \pi(\Lambda)$.*

If F_Γ is a fundamental domain of Γ in L, then

$$\mathrm{F} := \{ \mathbf{x} + \mu \mathbf{b}_1 \,:\, \mathbf{x} \in \mathrm{F}_\Gamma \text{ and } 0 \le \mu < 1 \} = \mathrm{F}_\Gamma + \{ \mu \mathbf{b}_1 \,:\, 0 \le \mu < 1 \}$$

is a fundamental domain for Λ.

Problem 8.6

The proof of Proposition 8.3 follows almost immediately from this lemma.

Proof of Proposition 8.3. We prove the claim by induction. For $d = 1$ we have $\mathbf{b}_1 = \mathbf{w}_1$, so the claim follows.

Now assume that $d > 1$. We consider the orthogonal complement L of $\mathbf{b}_1$ with the projection $\pi : \mathbb{R}^d \to L$ and the projected lattice $\Gamma := \pi(\Lambda)$.

The vectors $\mathbf{b}'_j := \pi(\mathbf{b}_j)$ for $2 \le j \le d$ are a basis for Γ, with Gram-Schmidt orthogonalization $\mathbf{w}_2, \dots, \mathbf{w}_d$.[6]

By assumption, $\mathrm{F}_\Gamma := \mathrm{F}(\mathbf{b}'_2, \dots, \mathbf{b}'_d)$ is a fundamental domain of Γ, and thus, by the previous lemma, $\mathrm{F}_\Gamma + \{ \mu \mathbf{b}_1 \,:\, 0 \le \mu < 1 \}$ is one of Λ. As $\mathbf{b}_1 = \mathbf{w}_1$ the claim follows. □

A simple consideration now shows that $\mathbf{v} + \mathrm{F}(\mathbf{b}_1, \dots, \mathbf{b}_d)$ for $\mathbf{v} \in \Lambda$ contains a single lattice point (this is also true for $\mathbf{x} + \mathrm{F}(\mathbf{b}_1, \dots, \mathbf{b}_d)$ for any $\mathbf{x} \in \mathbb{R}^d$, and the lattice point will be in the interior if $\mathbf{x} \notin \Lambda$). We can use this fundamental domain to give a new proof of Theorem 4.3, see Problem 8.7.

Problem 8.7

8.3. A $\mathcal{O}(2^{d^2})$-algorithm for the closest vector problem

We now want to address **(CVP)**. So let $\mathbf{t} \in \mathbb{R}^d$, and we want to find $\mathbf{u} \in \Lambda$ that minimizes $\|\mathbf{t} - \mathbf{v}\|$ over all $\mathbf{v} \in \Lambda$.

Recall the sublattice $\Lambda_{d-1} := \left\{ \sum_{i=1}^{d-1} \lambda_i \mathbf{b}_i \,:\, \lambda_i \in \mathbb{Z} \right\}$ with linear subspace $V_{d-1} = \operatorname{lin} \Lambda_{d-1}$. The translates

$$\mu \mathbf{w}_d + V_{d-1} \qquad \text{for} \qquad \mu \in \mathbb{Z}$$

cover the lattice Λ with *layers* $L_\mu := \mu \mathbf{w}_d + V_{d-1}$. As discussed in the previous section, we can project $\mathbf{t}$ onto one of the layers L_μ and use recursion to find $\mathbf{u} \in L_\mu$ to solve the exact problem, if we would know which of the layers L_μ contains the closest lattice point $\mathbf{u}$.

However, we don't know how to find this layer. The example of Figure 8.2 shows that it needs not be the layer closest to $\mathbf{x}$. Using this nonetheless leads to Babai's approximation algorithm.

[6] here we slightly abuse the notation, as the $\mathbf{w}_j$ for $j \ge 2$ are in $\mathbb{R}^d$, but as they are orthogonal to $\mathbf{b}_1$, they are in the subspace L, and $\mathbf{w}_j = \pi(\mathbf{w}_j)$.

We aim for the exact solution, so our strategy will be to find a subset of the layers that can possibly contain $\mathbf{u}$, compute the closest vector to the projection in each of these layers, and finally choose the one that is closest to $\mathbf{x}$. So we need to determine at which layers we have to look at.

Here we can use Lemma 8.1 which bounds the distance between input and result in Babai's algorithm in terms of the last Gram-Schmidt vector. This also bounds the distance to the true closest vector.

Corollary 8.5. *Let Λ be a lattice with an LLL-reduced basis $\mathbf{b}_1, \dots, \mathbf{b}_d$ and Gram-Schmidt vectors $\mathbf{w}_1, \dots, \mathbf{w}_d$. Let $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{u} \in \Lambda$ a closest vector to $\mathbf{x}$. Then*

$$\|\mathbf{u} - \mathbf{x}\| \;\leq\; 2^{(d-2)/2} \|\mathbf{w}_d\| \,.$$

□

Hence, we need to search at most the $2 \cdot 2^{(d-2)/2} = 2^{d/2}$ closest layers around $\mathbf{t}$ to find the right layer. So if $f(k)$ is the number of iterations to solve **(CVP)** in dimension k, then

$$f(k+1) \;=\; 2^{(k+1)/2} f(k) \,.$$

Hence, $f(d) \in \mathcal{O}(2^{d^2})$. In each step the number of computations is bounded by a polynomial in $\max \log \|\mathbf{b}_i\|_\infty$. To summarize, we obtain the following theorem.

Theorem 8.6. ***(CVP)** can be solved in time $\mathcal{O}(2^{d^2})$ times a polynomial in $\max \log \|\mathbf{b}_i\|_\infty$.* □

Looking back at this algorithm we may realize that in each iteration we solve $\mathcal{O}(2^k)$ closest vector problems in the *same* lattice independently. So if we find a way to reuse information in these computations we may be able to gain in the running time. That this is possible is the key idea in the algorithm of Micciancio and Voulgaris[7]. The show that we can speed up significantly if we, in each iteration, compute a *Voronoi* cell of the lattice and use this in each of the $\mathcal{O}(2^k)$ closest vector problems.

A *Voronoi* cell $V_\mathbf{v}$ around a lattice vector $\mathbf{v}$ of the lattice is the set of all points that are closer to $\mathbf{v}$ than to any other lattice point,

$$V_\mathbf{v} \;:=\; \Big\{ \mathbf{x} \in \mathbb{R}^d \;:\; \|\mathbf{x} - \mathbf{v}\| \;<\; \|\mathbf{x} - \mathbf{u}\| \text{ for all } \mathbf{u} \in \Lambda \setminus \{\mathbf{v}\} \Big\} \,.$$

Clearly, Voronoi cells to different lattice points are just translates of each other. In the definition we have not included the boundary, which is the set of points that have the same shortest distance to more than one lattice point. One can show that the closure of a Voronoi cell is a polytope with at most 2^{d+1} facets.[8]

[7] Daniele Micciancio and Voulgaris, "A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations [extended abstract]".

[8] In other applications one defines more generally the Voronoi diagram of a discrete set X of points as

Once the Voronoi cell $V_{\mathbf{0}}$ around 0 is known, one just needs to find the translate $\mathbf{v} + V_{\mathbf{0}}$ that contains the target $\mathbf{x}$. Then $\mathbf{v}$ is the closest vector. While the idea of the algorithm is pretty simple, working out how to compute the Voronoi cell efficiently and then analyzing the running time requires some work. If done right, the algorithm will solve **(CVP)** in time $\mathcal{O}(2^d)$.

8.4. Problems

8.1. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be an LLL-reduced lattice basis of a lattice Λ. Let $\mathbf{x} \in \mathbb{R}^d$ and assume that there is $\mathbf{v} \in \Lambda$ such that

$$\|\mathbf{x} - \mathbf{v}\| \leq \frac{1}{2} \|\mathbf{w}_i\|$$

for all i. Show that Babai's nearest plane algorithm returns $\mathbf{v}$.

8.2. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be a lattice basis of a lattice Λ. Let $\mathbf{x} \in \mathbb{R}^d$. If Babai's nearest plane algorithm returns $\mathbf{v}$, then

$$\|\mathbf{x} - \mathbf{v}\|^2 \leq \sum_{i=1}^{d} \frac{1}{4} \|\mathbf{w}_i\|^2 .$$

8.3. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice with basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ and Gram-Schmidt orthogonalization $\mathbf{w}_1, \ldots, \mathbf{w}_d$. Let $\mathbf{v}$ be the vector returned by the Nearest-Plane-Algorithm for $\mathbf{t}$. Define

$$F := \left\{ \mathbf{x} + \sum_{i=1}^{d} \mu_i \mathbf{w}_i \ : \ \frac{1}{2} \leq \mu_i \leq \frac{1}{2} \right\} .$$

Show that $\mathbf{v} \in F$. If $\mathbf{t}$ is not a lattice point, then $\mathbf{v}$ is the only lattice point in F.

8.4. Let $\Lambda \subseteq \mathbb{Z}^d$ be a lattice with an LLL-reduced basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$, whose Euclidean length is bounded by B. Prove that the nearest plane algorithm runs in time

$$\mathcal{O}\left(d^5 \log^2 B)\right) .$$

8.5. Show that Babai's nearest plane algorithm solves **(AbsCVP)**$_\gamma$ for any $\gamma \geq \frac{1}{2}\sqrt{\sum_{i=1}^{d} \|\mathbf{w}_i\|^2}$.

8.6. Prove Lemma 8.4.

8.7. Prove Theorem 4.3 using the fundamental domain $\mathrm{F}(\mathbf{b}_1, \ldots, \mathbf{b}_d)$.

the polyhedral complex of all cells around points in the set. If X is not a lattice, then the cells around different points may differ. Computing such cells, or the dual *Delaunay triangulations* are an important problem in many applications.

9. Counting Lattice Points

In this chapter we want to take a different approach to integer optimization and see how we can *encode* and *count* all integer points in a polyhedron efficiently. This requires us to solve two separate tasks.

If we want to encode *all* integer points efficiently, then we cannot just enumerate the points, as a polyhedron may contain an exponential number *w.r.t.* to the input size, even if it is bounded. If the polyhedron is not bounded, then enumerating the points is not even feasible. So we need a better way to write down all integer points.

Further, if we want to count the points efficiently, then we must also be able to compute our encoding of the points efficiently, *i.e.* in polynomial time, and we must also be able to obtain the size of the set in polynomial time.

We address both questions in the next sections. We start with the first and discuss a method to encode all integer points efficiently. We motivate the idea in the next section, before we work out the details.

9.1. Motivation

Let P be a polyhedron and S the set of integer points in P, *i.e.* the set $S := P \cap \mathbb{Z}^d$. All of the following also works for general lattices, but as only one lattice is involved and we can reduce to this case by picking a basis of the lattice and writing all our coordinates in this basis, we will stick to $\mathbb{Z}^d$.

To count the integer points in a polyhedron P, *i.e.* to compute the size of S, we have to find a way to distinguish integer points in a polyhedron P from all others, i.e. a way to encode them, preferably in an *efficient* and *explicit* way, that we can easily write down in a *short* and *concise* form. It should be simple from our notation to decide whether a point is in our list or not.

In principle we already know at least two options to decide whether an integer point is in the polyhedron or not. Namely, an integer point $\mathbf{x}$ is in our polyhedron P if it is either a convex combination of the vertices of P or satisfies all defining inequalities of P. Both methods are certainly fine if we need to check a particular point. But they do not tell us much about the whole set of points, nor about the structure of the set.

For this, we need to find a way to make our description of the set S more explicit. In a first, rather naïve approach, we could now be tempted to explicitly list all lattice points in our polytope (this clearly only works well for bounded objects). To make a simple example, look at the polytope $P_3 := [0, 3]$. This is the simple segment shown in

Figure 9.1.: The polytope P_3.

Figure 9.1. The naïve approach gives us the list

$$0,\ 1,\ 2,\ 3\,.$$

This works well in this small example, but consider the structurally similar example $P_{10002} := [0, 10002]$. Here, our list,

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, \dots\,,$$

if written out completely, easily exceeds the line, and also this page, and if we replace 10002 by 10002^2, then also the length of these lecture notes. To get a compact encoding of the points we need a better idea.

Here is the key idea. This approach may look really strange at first, but will prove to be very powerful. We can replace each point $k \in P_3$ with its monomial t^k. With this we define a polynomial that contains precisely the monomials corresponding to points in our polytope and write

$$1 + t + t^2 + t^3 \ = \sum_{i=0}^{3} t^i\,.$$

The option to write our polynomial as a sum already shows a quite compact way to encode the lattice points. Observe, that the representation is not really more complicated for P_{10002}. However, it is pretty obvious that this particular compact notation as a sum is only possible in very special cases, so we need to look further.

If you look at the polynomial you may recall from your calculus class that there is another option to write this sum in a more condensed form using the geometric series

$$G_{P_3}(t) \ := \ \frac{1-t^4}{1-t}\,.$$

whose expansion is again our polynomial. Doing the same for P_{10002} gives

$$G_{[0,10002]}(t) \ := \ \frac{1-t^{10003}}{1-t}\,,$$

so it does not really make this notation more complicated.

We will see that this idea of using a geometric series to specify the lattice points in a polyhedron is both sufficiently flexible to work for all polyhedra, and efficient enough that we can use it to really study the structure of the set of lattice points.

Here comes another surprising and powerful property of our last observation. If we try to do write down the lattice points of the unbounded polyhedron $P_\infty := [0, \infty)$, then our first two approaches obviously become infeasible. However, the third works

and turns out to be even shorter and more appealing![1] As a geometric series we can concisely describe all lattice points in P_∞ via the monomials in

$$G_{[0,\infty)}(t) \;:=\; \frac{1}{1-t}\,.$$

As this extended example suggests, the generating function we use here to encode the lattice points will indeed provide a powerful bookkeeping tool for counting and enumerating lattice points in polytopes.

It will soon become apparent that it is indeed quite useful and natural to encode lattice points not only in polytopes, but more generally in any bounded or unbounded subset of $\mathbb{R}^d$, as in the last example of a ray in $\mathbb{R}^1$. You should keep this in mind for the following considerations.

9.2. Generating Functions

In the previous section we have seen that we can use rational functions in one variable t to describe the infinite series of all monomials corresponding to the lattice points for the one-dimensional cone

$$C \;:=\; \{\, x \;:\; x \geq 0 \,\} \;\subseteq\; \mathbb{R}\,.$$

In this section we want to formalize this idea, and directly generalize it to arbitrary dimensions.

Let $\Bbbk$ be some ground field (you can just think of $\Bbbk = \mathbb{C}$, if you like). We assign the monomial

$$\mathbf{t}^{\mathbf{a}} := t_1^{a_1} t_2^{a_2} \cdots t_d^{a_d}$$

in d variables to an integral point $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{Z}^d$. In the initial example of the nonnegative axis all integral points were non-negative and thus lead to ordinary monomials as you know them from multidimensional calculus or algebra. In general, the coordinates of a are allowed to be negative, so this is a *Laurent polynomial* living in the *Laurent polyonmial ring*

$$\mathsf{L} \;=\; \Bbbk[t_1^{\pm 1}, \ldots, t_d^{\pm 1}]\,.$$

Moreover, note that the sum of monomials for the cone $x \geq 0$ is infinite. Since we do not care about convergence, we will actually consider our sums not as Laurent polynomials, but as series in a subset of the L-module

$$\widehat{\mathsf{L}} \;:=\; \Bbbk[\![t_1^{\pm 1}, \ldots, t_d^{\pm 1}]\!]$$

of *formal Laurent series*. We associate a series in this module to any convex set in $\mathbb{R}^d$.

[1] You may wander about what happens at $t = 1$. However, this is not an issue for us, as we will not deal with any analytic convergence issues and usually will not evaluate at certain values. We will do this once later for one specific value, and postpone the discussion until then.

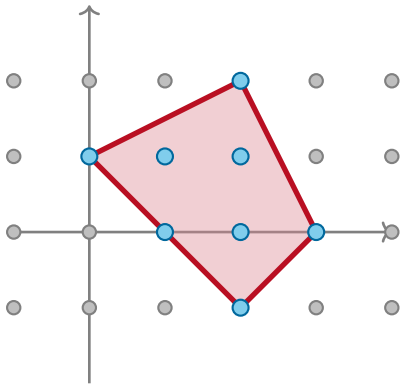


Figure 9.2.: The polygon of Example 9.2.

Definition 9.1. For $S \subset \mathbb{R}^d$ the *integer point series* $\hat{G}_S$ is the formal Laurent series

$$\hat{G}_S(\mathbf{t}) := \sum_{\mathbf{a} \in S \cap \mathbb{Z}^d} \mathbf{t}^{\mathbf{a}} \in \widehat{\mathsf{L}} \,.$$

Here is an example of such a series.

Example 9.2. Let P be the polygon

$$P := \mathsf{conv} \begin{bmatrix} 0 & 2 & 2 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix}$$

(see Figure 9.2). Recall that the convex hull of a matrix is defined to be the convex hull of the column vectors of the matrix. We list the lattice points as monomials in the Laurent polynomial

$$\begin{array}{ccccccc} & & & & t_1^2t_2^2 & & \\ + & t_2 & + & t_1t_2 & + & t_1^2t_2 & \\ & & + & t_1 & + & t_1^2 & + \; t_1^3 \\ & & & & + & t_1^2t_2^{-1} & \end{array} \,.$$

Translating a set $S \subseteq \mathbb{R}^d$ by some integral vector $\mathbf{a} \in \mathbb{Z}^d$ amounts to multiplication of its generating series with $\mathbf{t}^{\mathbf{a}}$,

$$\hat{G}_{\mathbf{a}+S}(\mathbf{t}) = \mathbf{t}^{\mathbf{a}} \hat{G}_S(\mathbf{t}) \,.$$

You will realize (or you have seen this already in some other course) that dealing with formal power series is more subtle than computations with polynomials. Seemingly simple operations may not work as expected, as the following example with formal

Laurent series shows.

$$\begin{aligned}\hat{G}_{\mathbb{R}}(t)\cdot(1-t) &= (\cdots+t^{-2}+t^{-1}+1+t+t^2+\cdots)\cdot(1-t)\\ &= (\cdots+t^{-2}+t^{-1}+1+t+t^2+\cdots)\\ &\quad -(\cdots+\ t^{-1}+1+t+t^2+t^3\cdots)\\ &= 0\,.\end{aligned}$$

We might be tempted to divide by $(1-t)$ and deduce

$$\hat{G}_{\mathbb{R}}(t) \;=\; \cdots+t^{-2}+t^{-1}+1+t+t^2+\cdots \;=\; \frac{0}{1-t} \;=\; 0\,. \tag{9.1}$$

However, the left side of the equation is definitely not zero as a Laurent series, and the operation is not allowed, at least not in this simple form. We will meet one option to fix this later.

Actually, not all Laurent series appear as a generating series for lattice points in polyhedra. The ones we will encounter have a nice additional structure that we will work out with the next definitions and theorems.

Definition 9.3. A Laurent series $\hat{G}\in\widehat{\mathsf{L}}$ is *summable* if there is a Laurent polynomial $g\in\mathsf{L}$ such that the series $g\hat{G}$ is a Laurent polynomial.

Clearly all Laurent polynomials are summable. On the other hand, the series

$$1+t^2+t^3+t^5+t^7+t^{11}+t^{13}+t^{17}+\ldots \;=\; 1+\sum_{k\text{ prime}} t^k$$

cannot be summable. We will denote the set of all summable Laurent series by $\mathsf{L}^{\mathrm{sum}}$. We leave the proof of the following proposition to the reader as Problem 9.1.

Proposition 9.4. $\mathsf{L}^{\mathrm{sum}}$ *is a* L*-submodule of* $\widehat{\mathsf{L}}$. □

Problem 9.1

In fact, the summable series we are interested in will be constructed from summable series of cones. We will derive those in the next paragraphs, starting with some simple cones, until we obtain the series of a general polyhedral cone spanned by a finite set of generators in Proposition 9.7.

We begin with the polyhedron $P_\infty=[0,\infty)$ that we introduced in the previous section. The integer point series is

$$\hat{G}_{P_\infty}(t) \;=\; \sum_{a\in\mathbb{Z}_{\geq 0}} t^a = 1+t+t^2+t^3+\cdots\,.$$

Using the polynomial $g(t):=(1-t)$ we obtain $g(t)G_{P_\infty}(t)=1$, so $\hat{G}_{P_\infty}(t)$ is a summable series.

Now let us consider the 2-dimensional cone $C := \mathrm{cone}(\mathbf{e}_1, \mathbf{e}_2)$, where $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^2$ are the standard unit vectors. Then

$$\begin{aligned}\hat{G}_C(t,s) &= \sum_{a,b\in\mathbb{Z}_{\geq 0}} t^a s^b \\ &= \left(\sum_a t^a\right)\left(\sum_b s^b\right) = 1 + t + s + t^2 + s^2 + ts + t^3 + \cdots .\end{aligned}$$

Here we can use the polynomial

$$g(t,s) := (1-t)(1-s)$$

to obtain

$$g(t,s)\cdot\hat{G}_C(t,s) = (1-t)\cdot\left(\sum_{a\in\mathbb{Z}_{\geq 0}} t^a\right)\cdot(1-s)\cdot\left(\sum_{b\in\mathbb{Z}_{\geq 0}} s^b\right) = 1.$$

Hence, $\hat{G}_C(t,s)$ is a summable series. Having seen these two cases you probably already see the pattern. And indeed, if consider the cone $C := \mathrm{cone}(\mathbf{e}_1, \ldots, \mathbf{e}_d)$ spanned by the d unit vectors in $\mathbb{R}^d$, then

$$\begin{aligned}\prod_{i=1}^d (1-\mathbf{t}^{\mathbf{e}_i})\cdot\hat{G}_C(\mathbf{t}) &= \prod_{i=1}^d (1-\mathbf{t}^{\mathbf{e}_i})\cdot\sum_{\mathbf{a}\in\mathbb{Z}^d_{\geq 0}} \mathbf{t}^{\mathbf{a}} \\ &= \prod_{i=1}^d (1-\mathbf{t}^{\mathbf{e}_i})\cdot\sum_{\mathbf{a}\in\mathbb{Z}^d_{\geq 0}} (\mathbf{t}^{\mathbf{e}_1})^{a_1}\cdots(\mathbf{t}^{\mathbf{e}_d})^{a_d} = 1\end{aligned} \tag{9.2}$$

Hence, $\hat{G}_C(\mathbf{t})$ is summable. In all three cases it is tempting to divide both sides by the function

$$g(\mathbf{t}) = \prod_{i=1}^d (1-\mathbf{t}^{\mathbf{e}_i})$$

to solve the equation for $\hat{G}_C(\mathbf{t})$. However, it is not immediately clear that this is an allowed operation. The next proposition solves this for us. We leave the proof for Problem 9.2.

Proposition 9.5. *There is a natural homomorphism*

$$\Phi : \mathsf{L}^{\mathrm{sum}} \longrightarrow \mathsf{R} := \Bbbk(t_1, \ldots, t_d),$$

from summable series to rational functions that maps $\hat{G}$ to f/g if $g\hat{G} = f$ in $\widehat{\mathsf{L}}$.

Problem 9.2

With this definition we can give a correct version of the computation in (9.1). If we

replace the series with its image under Φ, then we obtain

$$\begin{aligned}\Phi(\hat{G}_{\mathbb{R}}(t)) &= \Phi(\cdots + t^{-2} + t^{-1} + 1 + t + t^2 + \cdots) \\ &= \frac{0}{1-t} = 0.\end{aligned}$$

In other words, not the Laurent series is zero but only its associated rational function. While it is often very convenient identify summable Laurent series with rational function in equations instead of using a cumbersome and non-standard notation such as Φ, one cannot stress enough that one must be aware that such an equality only holds on the level of rational functions and not on the level of Laurent series.

In particular, we see from the previous example that Φ is not an injective map. However, it clearly is for Laurent *polynomials* (you should check this). In other words, L is a submodule of $\mathsf{L}^{\mathrm{sum}}$, and $\Phi|_{\mathsf{L}}$ is the identity map. A more general criterion on injectivity is proven in Problem 9.3.

Problem 9.3

Definition 9.6. Let $S \subseteq \mathbb{R}^d$ and assume that $\hat{G}_S(\mathbf{t})$ is summable. The *integer point generating function* of S is

$$G_S(\mathbf{t}) := \Phi(\hat{G}_S(\mathbf{t}))\,.$$

If $S \subseteq \mathbb{R}^d$ is bounded, then we are allowed to identify (see also Problem 9.3)

$$G_S(\mathbf{t}) = \sum_{\mathbf{a} \in S \cap \mathbb{Z}^d} \mathbf{t}^{\mathbf{a}}\,.$$

In the one-dimensional example $P_\infty = [0, \infty)$ we have already computed the image of the generating series in R,

$$G_{P_\infty}(t) = \frac{1}{1-t}\,.$$

We can generalize this observation to rational simplicial cones. Let C be a simplicial rational cone in $\mathbb{R}^d$ with primitive ray generators $R := \{\mathbf{r}_1, \ldots, \mathbf{r}_d\}$. We recall the *fundamental parallelepiped* of V from Definition 2.5

$$\Pi(R) := \left\{ \sum_{i=1}^{d} \mu_i \mathbf{r}_i \;:\; \mu_i \in [0,1) \text{ for } 1 \leq i \leq d \right\}$$

We know from Corollary 2.8 that the fundamental parallelepipeds tile the space without overlap (strictly, there we talked about lattice bases, however, the same argument works for the generating set R). We define

$$F := \Pi(R) \cap \mathbb{Z}^d\,,$$

the set of all integral points in the fundamental parallelepiped of R. Note that F is a

finite set, so its integer point generating function $G_{\Pi(R)}(\mathbf{t})$ is a polynomial. We write

$$\sigma_C(\mathbf{t}) \;:=\; G_{\Pi(R)}(\mathbf{t})$$

for this.

Proposition 9.7. *In this notation, $\hat{G}_C(\mathbf{t})$ is summable with*

$$G_C(\mathbf{t}) \;=\; \frac{\sigma_C(\mathbf{t})}{\prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i})}$$

Proof. Let S be the set of all nonnegative integer linear combinations of R,

$$S \;:=\; \left\{ \mathbf{z} \;=\; \sum \lambda_i \mathbf{r}_i \;:\; \lambda_i \in \mathbb{Z}_{\geq 0} \right\}.$$

By replacing the Laurent monomial $t^{\mathbf{e}_i}$ by $t^{\mathbf{r}_i}$ in (9.2), we get

$$\prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i}) \cdot \sum_{\mathbf{z}\in S} \mathbf{t}^{\mathbf{z}} \;=\; \prod_{i=1}^{d}(1-t^{\mathbf{r}_i}) \cdot \sum_{\mathbf{a}\in\mathbb{Z}^d_{\geq 0}} (\mathbf{t}^{\mathbf{r}_1})^{a_1} \cdots (\mathbf{t}^{\mathbf{r}_d})^{a_d} \;=\; 1.$$

By Corollary 2.8 we get

$$\begin{aligned}\prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i}) \cdot \sum_{\mathbf{x}\in C\cap\mathbb{Z}^d} \mathbf{t}^{\mathbf{x}} \;&=\; \prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i}) \cdot \sum_{\mathbf{y}\in F}\sum_{\mathbf{z}\in S} \mathbf{t}^{\mathbf{y}+\mathbf{z}} \\ &=\; \prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i}) \cdot \sum_{\mathbf{z}\in S} \mathbf{t}^{\mathbf{z}} \cdot \sum_{\mathbf{y}\in F} \mathbf{t}^{\mathbf{y}} \\ &=\; \sum_{\mathbf{y}\in F} \mathbf{t}^{\mathbf{y}} \qquad\qquad\qquad\qquad = \;\sigma_C(\mathbf{t}).\end{aligned}$$

Dividing by $\prod_{i=1}^{d}(1-\mathbf{t}^{\mathbf{r}_i})$ gives the claim. □

Our integer point generating series $\hat{G}_S(\mathbf{t})$ contains a monomial for every integral point in a set S. Now assume we have a second set S'. If we want to obtain the series for the union $S \cup S'$, then we have to list each monomial of a lattice point in the union precisely once. Hence, we cannot just add the two series if the intersection of the sets is not empty.

However, we can obtain the series by first adding the two series of S and S', and then *subtracting* the series of $S \cap S'$. This principle clearly extends to the union of any finite number of sets. We can compute the generating series from the generating series of the sets and all partial intersections if we keep track of the multiplicities a partial intersection appears in the total sum. This is called the *principle of inclusion-exclusion*. You will study this in more detail in Problem 9.4.

Problem 9.4

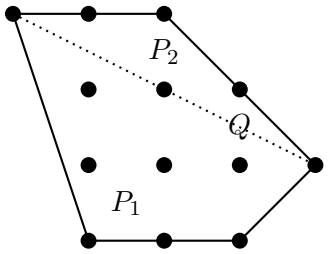


Figure 9.3.: Let Q be the dotted chord of the polygon and P_1, P_2 the two polygons obtained by cutting P along Q. Then $|P \cap \mathbb{Z}^2| = |P_1 \cap \mathbb{Z}^2| + |P_2 \cap \mathbb{Z}^2| - |Q \cap \mathbb{Z}^2|$.

Now let us come back to the special case of integer points in polyhedra. Triangulating a rational polyhedral cone into rational simplicial cones (see Section A.8)) and using inclusion-exclusion (see also Figure 9.3) yields the following general result.

Corollary 9.8. *The integer point generating series of a rational polyhedral cone is summable.*

9.3. The Theorem of Brion

With the last Corollary 9.8 in the previous section we have found a way to express the lattice points in a polyhedral cone as a rational function. However, we set out for a way to encode and count the integer points in a polytope. So in order to make the results we have obtained so far useful we need a method to express the integer point generating function of a polytope via the integer point generating functions of some cones.

We will show in this section that this is indeed possible. With Brion's Theorem (Theorem 9.14) we prove that the integer point generating function of a polytope can be computed from the integer point generating functions of all vertex cones of P.

We need some preparations for this. Let P be a rational d-dimensional polytope and F a face of P. Recall the tangent cone of F in P from Definition A.42, which is defined by

$$\mathsf{T}_F P := \left\{ \mathbf{v} \in \mathbb{R}^d \ : \ \text{there is } \mathbf{w} \in F,\ \varepsilon > 0 \text{ with } \mathbf{w} + \varepsilon(\mathbf{v} - \mathbf{w}) \in P \right\}.$$

We know from Problem A.14 that the tangent cone is the common intersection of all supporting halfspaces at F, and we know from Proposition A.44 that the shifted cone $\mathsf{T}_F P - \mathbf{v}$ for some $\mathbf{v} \in F$ is dual to the normal cone of F.

The following theorem makes a connection between the generating series of the tangent cones of *all* faces of the polytope and the generating series of the polytope. This is the first step towards our final result. To obtain the generating functions we afterwards apply our map Φ on both sides and use linearity of this map.

In the final step we will then argue that the tangent cones of faces of dimension $d \geq 1$ doe not contribute anymore to the sum after we have applied Φ. So it will suffice to consider only the vertices to obtain the generating functions.

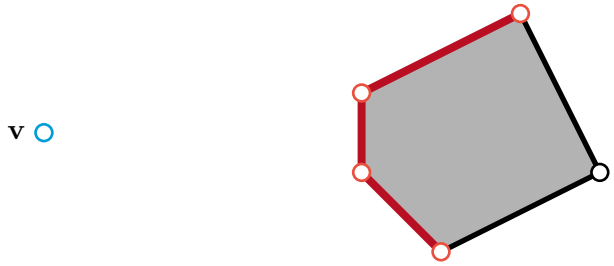


Figure 9.4.: $\mathbf{v} \notin P$. The complex $G \in \mathrm{visible}_P(\mathbf{v})$ is drawn in red.

Theorem 9.9 (Theorem of Brianchon-Gram)**.** *Let P be a rational polytope. Then*

$$\hat{G}_P(\mathbf{t}) = \sum_{F \preceq P} (-1)^{\dim F} \hat{G}_{\mathsf{T}_F P}(\mathbf{t}) \,,$$

where the sum is over all non-empty faces of P. This equation is the Brianchon-Gram identity.

The Brianchon-Gram identity is already valid on the level of indicator functions for lattice points. For its proof we introduce the complex of *visible faces*.

Definition 9.10. Let $\mathbf{v} \notin P$. A face F of a polytope P is *visible* from $\mathbf{v}$ if for some (equivalently every) $\mathbf{w} \in \mathrm{relint}\, F$ the segment $\mathsf{conv}(\mathbf{v}, \mathbf{w})$ intersects P only in $\mathbf{w}$.

The collection $\mathrm{visible}_P(\mathbf{v})$ of all visible faces of P *w.r.t.* $\mathbf{v}$ is the *complex of visible faces*.

We prove some some properties for visible faces, before we turn to the proof of the Theorem of Brianchon-Gram (Theorem 9.9). See Figure 9.4 for an illustration. The next lemma characterizes the visible faces in terms of the tangent cones.

Lemma 9.11. *F is visible from $\mathbf{v}$ if and only if $\mathbf{v} \notin \mathsf{T}_F P$.*

Proof. Let $\mathbf{w}$ be a point in $\mathrm{relint}\ F$. As $\mathsf{T}_F P$ is the intersection of all valid half spaces of P we know that $\mathrm{aff}\ F$ is the minimal face of $\mathsf{T}_F P$. In particular,

$$\langle \mathbf{a}, \mathbf{w} \rangle \ \leq \ \langle \mathbf{a}, \mathbf{x} \rangle$$

for all valid functionals $\mathbf{a}$ and $\mathbf{x} \in \mathsf{T}_F P$.

If $\mathbf{v} \notin \mathsf{T}_F P$, there is a linear functional $\mathbf{a} \in \left(\mathbb{R}^d\right)^\star$ that separates $\mathbf{v}$ from the tangent cone and the polyhedron, so

$$\langle \mathbf{a}, \mathbf{v} \rangle \ < \ \min \left(\langle \mathbf{a}, \mathbf{x} \rangle \ : \ \mathbf{x} \in \mathsf{T}_F P \right) .$$

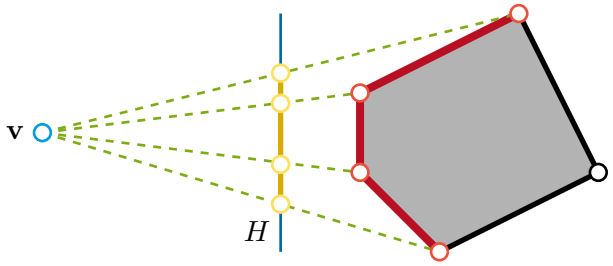


Figure 9.5.: The hyperplane H is drawn in blue, and the polytopal subdivision in H is drawn in yellow.

We compute, for $0 \leq \lambda < 1$,

$$\begin{aligned}\langle \mathbf{a}, (1-\lambda)\mathbf{v} + \lambda\mathbf{w} \rangle &= (1-\lambda)\langle \mathbf{a}, \mathbf{v} \rangle \ + \ \lambda\langle \mathbf{a}, \mathbf{w} \rangle \\ &< (1-\lambda)\langle \mathbf{a}, \mathbf{w} \rangle \ + \ \lambda\langle \mathbf{a}, \mathbf{w} \rangle \\ &= \langle \mathbf{a}, \mathbf{w} \rangle \\ &\leq \min\left(\langle \mathbf{a}, \mathbf{x} \rangle \ : \ \mathbf{x} \in \mathsf{T}_F P\right) \\ &\leq \min\left(\langle \mathbf{a}, \mathbf{x} \rangle \ : \ \mathbf{x} \in P\right) .\end{aligned}$$

Hence, $(1-\lambda)\mathbf{v} + \lambda\mathbf{w} \notin P$ for all $0 \leq \lambda < 1$. In other words, F is visible from $\mathbf{v}$.

If, conversely, $\mathbf{v} \in \mathsf{T}_F P$, then we know that there are $\mathbf{w}' \in F$ and an $\varepsilon_0 > 0$ so that

$$\mathbf{v}' \ := \ \mathbf{w}' + \varepsilon(\mathbf{v} - \mathbf{w}') \ \in \ P$$

for all $0 < \varepsilon \leq \varepsilon_0$. As $\mathbf{w} \in \operatorname{relint} F$ we can choose $0 < \varepsilon \leq \varepsilon_0$ such that

$$\mathbf{w}'' \ := \ \mathbf{w} + \varepsilon(\mathbf{w} - \mathbf{w}') \ \in \ F .$$

But then

$$\begin{aligned}(1-\varepsilon)\mathbf{w}'' \ + \ \varepsilon\mathbf{v}' &= (1-\varepsilon)\mathbf{w} \ + \ (1-\varepsilon)\varepsilon(\mathbf{w} - \mathbf{w}') \ + \ \varepsilon\mathbf{w}' \ + \ \varepsilon^2(\mathbf{v} - \mathbf{w}') \\ &= \big((1-\varepsilon)(1+\varepsilon)\mathbf{w} \ - \ (1-\varepsilon)\varepsilon\mathbf{w}'\big) \ + \ \big(\varepsilon(1-\varepsilon)\mathbf{w}' \ + \ \varepsilon^2\mathbf{v}\big) \\ &= (1-\varepsilon^2)\mathbf{w} \ + \ \varepsilon^2\mathbf{v}\end{aligned}$$

belongs to both $\mathsf{conv}(\mathbf{v}, \mathbf{w})$ and to P, and F is not visible from $\mathbf{v}$. □

Corollary 9.12. *Let P be a polytope and $\mathbf{v} \notin P$.*
Then $\operatorname{visible}_P(\mathbf{v})$ is a polyhedral complex that is isomorphic to subdivision of a polytope.

Proof. The definition of visibility implies $G \in \operatorname{visible}_P(\mathbf{v})$ if $G \preceq F \in \operatorname{visible}_P(\mathbf{v})$. So $\operatorname{visible}_P(\mathbf{v})$ is a subcomplex of the boundary of P.

Let H be a hyperplane separating $\mathbf{v}$ from P, and consider the polytope $Q := H \cap \mathsf{conv}(P \cup \{\mathbf{v}\})$, see Figure 9.5. Then

$$\left\{H \cap \mathsf{conv}(F \cup \{\mathbf{v}\}) \ : \ F \in \operatorname{visible}_P(\mathbf{v})\right\}$$

is a subdivision of Q which is combinatorially isomorphic to $\mathrm{visible}_P(\mathbf{v})$. □

Now we have all ingredients to prove the Theorem of Brianchon-Gram (Theorem 9.9).

Proof of the Theorem of Brianchon-Gram (Theorem 9.9). We think of the Laurent polynomial on the left hand side as an infinite Laurent series that contains all possible monomials, but most coefficients are 0.

We compare coefficients of an arbitrary monomial $\mathbf{t}^{\mathbf{u}}$ for $\mathbf{u} \in \Lambda$ on both sides to prove the identity. For this, we deal with the two cases $\mathbf{u} \in P$ and $\mathbf{u} \notin P$ separately.

If $\mathbf{u} \in P$, then $\mathbf{u} \in \mathsf{T}_F P$ for every non-empty face F of P. Hence, the coefficient of $\mathbf{t}^{\mathbf{u}}$ on the right hand side is

$$\sum_{\emptyset \neq F \preceq P} (-1)^{\dim F} \;=\; \mathsf{f}_0 - \mathsf{f}_1 + \cdots + (-1)^d \mathsf{f}_d \,.$$

This sum is the Euler characteristic of the polytope, which is always 1 by Euler's relation (see (A.13) and Proposition A.46). Note that we have included the polytope itself in the sum.

On the other hand, if $\mathbf{u} \notin P$, then we obtain from Lemma 9.11 that the coefficient of $\mathbf{t}^{\mathbf{u}}$ on the right hand side is

$$\sum_{\emptyset \neq F \preceq P} (-1)^{\dim F} - \sum_{\emptyset \neq F \in \mathrm{visible}_P(\mathbf{u})} (-1)^{\dim F} \;=\; 1 - 1 \;=\; 0\,,$$

using the fact that the Euler characteristic of $\mathrm{visible}_P(\mathbf{u})$ is 1 by Corollary 9.12 and Proposition A.46. See also Figure 9.5. □

In the previous section we have applied $\Phi : \widehat{\mathsf{L}} \longrightarrow \mathsf{R}$ only to *pointed* polyhedral cones. We now want to study this map also in the case of cones that have a nontrivial lineality space, which is defined as

$$\mathrm{lineal}(C) \;:=\; C \cap (-C)\,.$$

It is the maximal linear subspace contained in C.

Let us first look at an example for what we want to prove. Consider the sets

$$C^+ \;:=\; [0,\infty) \;\subseteq\; \mathbb{R} \qquad \text{and} \qquad C^- \;:=\; 3 - C^+ = (-\infty, 3] \;\subseteq\; \mathbb{R}\,.$$

C^+ is a one-dimensional cone, C^- is an affine cone with apex in 3. Let P be the intersection of the cones, *i.e.*

$$P \;:=\; C^+ \cap C^- \;=\; [0,3].$$

Then P is a one-dimensional polytope. We compute the integer point generating function

and the image under Φ for C^+ and C^-. The series are

$$\begin{aligned}\hat{G}_{C^+}(t) &= \sum_{k\geq 0} t^k \\ \hat{G}_{C^-}(t) &= \sum_{k\leq 3} t^k \;=\; t^3 \sum_{k\leq 0} t^k \;=\; t^3 \sum_{k\geq 0} t^{-k}\,,\end{aligned}$$

so we obtain the functions

$$\begin{aligned}G_{C^+}(t) &= \Phi(\hat{G}_{C^+}(t)) \;=\; \frac{1}{1-t} \\ G_{C^-}(t) &= \Phi(\hat{G}_{C^-}(t)) \;=\; t^3 \frac{1}{1-\frac{1}{t}} \;=\; \frac{-t^4}{1-t}\end{aligned}$$

The integer point generating function of P is the finite geometric series

$$\hat{G}_P(t) \;=\; G_P(t) \;=\; \frac{1-t^4}{1-t} \;=\; 1+t+t^2+t^3\,.$$

We observe that

$$G_P(t) \;=\; G_{C^+}(t) + G_{C^-}(t)\,.$$

Using the map Φ we can make the following symbolic calculation

$$\begin{aligned}G_P(t) &= \Phi(\hat{G}_{C^+}(t)) + \Phi(\hat{G}_{C^-}(t)) &&= \Phi(\hat{G}_{C^+}(t) + \hat{G}_{C^-}(t)) \\ &= \Phi(\hat{G}_{\mathbb{R}+P}(t)) &&= \Phi(\hat{G}_{\mathbb{R}}(t)) + \Phi(\hat{G}_P(t))\end{aligned}$$

This can only hold if $\Phi(\hat{G}_{\mathbb{R}}t) = 0$, i.e. if Φ maps the infinite series $\sum_{k\in\mathbb{Z}} t^k$ to 0. The following proposition shows that this indeed holds in general for cones with nontrivial lineality space.

Proposition 9.13. *Let $C \subseteq \mathbb{R}^d$ be a polyhedral cone with $\operatorname{lineal} C \neq \{0\}$.*
Then $\Phi(\hat{G}_C(\mathbf{t})) = 0$.

Proof. Let $\mathbf{v} \in \operatorname{lineal}(C) \setminus \{0\}$ and $L := \{\, \lambda\mathbf{v} \;:\; \lambda \in \mathbb{R} \,\}$. Then $L \subseteq C$, so that

$$\mathbf{t}^{\mathbf{v}} \hat{G}_C(\mathbf{t}) \;=\; \hat{G}_C(\mathbf{t})\,.$$

Applying the map Φ gives

$$\mathbf{t}^{\mathbf{v}} \Phi(\hat{G}_C(\mathbf{t})) \;=\; \Phi(\hat{G}_C(\mathbf{t})) \qquad\Longleftrightarrow\qquad (1-\mathbf{t}^{\mathbf{v}})\Phi(\hat{G}_C(\mathbf{t})) = 0\,.$$

$\mathbf{v} \neq \mathbf{0}$ implies $\Phi(\hat{G}_C(\mathbf{t})) = 0$. □

Now we can finally deduce a connection between the generating function of cones and the generating function of the polytope.

Theorem 9.14 (Brion's Theorem)**.** *Let P be a polytope. Then*

$$G_P(\mathbf{t}) \;=\; \sum_{v \text{ vertex of } P} G_{\mathsf{T}_v P}(\mathbf{t})\,.$$

Proof. Apply the map Φ to both sides of the equation in the Theorem of Brianchon-Gram (Theorem 9.9). The only non-pointed tangent cones are those originating from a vertex of P, so by Proposition 9.13 only the contributions of the vertices are non-zero on the right hand side. □

Problem 9.5

Here is one example for the equation of this theorem.

Example 9.15. Let P be the unit square in $\mathbb{R}^2$, *i.e.* the convex hull of $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_1 + \mathbf{e}_2$. Then

$$\begin{aligned} G_P(x,y) &= \frac{1}{(1-x)(1-y)} + \frac{x}{(1-\frac{1}{x})(1-y)} \\ &\qquad + \frac{y}{(1-x)(1-\frac{1}{y})} + \frac{xy}{(1-\frac{1}{x})(1-\frac{1}{y})} \\ &= \frac{1}{(1-x)(1-y)} + \frac{-x^2}{(1-x)(1-y)} \\ &\qquad + \frac{-y^2}{(1-x)(1-y)} + \frac{x^2y^2}{(1-x)(1-y)} \\ &= \frac{(1-x^2)(1-y^2)}{(1-x)(1-y)} \\ &= 1 + x + y + xy \end{aligned}$$

So $G_P(1,1) = 1+1+1+1 = 4$.

The theorem provides us with a general method to compute the function $G_P(\mathbf{t})$ for polytopes by computing the generating functions for the tangent cones at vertices (which are dual to the normal cones at vertices) and adding them up.

However, recall from the previous section that we need a triangulation of the cones for this, in each cone we have to enumerate the lattice points in the fundamental domain, and then use inclusion-exclusion with all lower dimensional faces of the cone to compute the generating function of the cone. This is still an expensive operation. We will see in the next section how we can avoid this and compute generating functions for cones efficiently.

9.4. Barvinok's Algorithm

In Section 9.2 we have seen that we can express the lattice points in a cone as a rational function. The proof is somehow constructive, as it tells us the exact from of the numerator of the rational function for simplicial cones, see Corollary 9.8 and the remarks above this

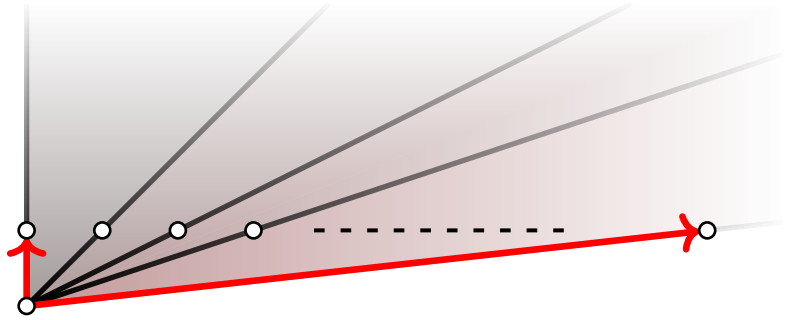


Figure 9.6.: The cone spanned by $\mathbf{e}_2$ and $k\mathbf{e}_1 + \mathbf{e}_2$ needs k unimodular cones in its decomposition

corollary. We have to enumerate the lattice points in the fundamental parallelepiped and combine the corresponding monomials into a polynomial. The denominator always has a canonical form depending only on the generators of the cone, with a factor of $(1 - \mathbf{t}^{\mathbf{r}})$ for each generator $\mathbf{r}$. For more general cones we use the fact that we can triangulate cones into simplicial ones, and that counting lattice points in such subdivisions can be done with inclusion-exclusion on the intersections.

Yet, as we will see below, this is in general not a task that can be done in polynomial time in the size of input and dimension. There are two related problems with this approach. In general, it is difficult to enumerate the lattice points in the fundamental parallelepiped (and this is also the task we actually want to solve with the generating functions that we develop) of a cone. On the other hand, this is easy, if the generators of the cone are a lattice basis. Such cones are called *unimodular*. As we know that any lattice point can be written as the sum of a lattice point in the fundamental parallelepiped and an integral linear combination of the cone generators we conclude in this case that 0 is the only lattice point. So in subdividing a cone into simplicial ones we should aim for a subdivision in which each such cone is generated by a lattice basis. However, as we will see, this leads to an exponential number of cones in the subdivision (relative to the size of the input, which is the number of generators of the input cone).

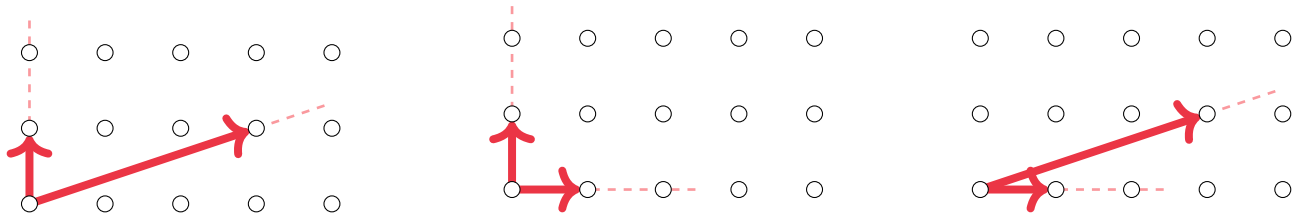
Here is an example that illustrates this problem. Consider the cone

$$C \;:=\; \operatorname{cone}(\mathbf{e}_2, k \cdot \mathbf{e}_1 + \mathbf{e}_2) \tag{9.3}$$

for some $k \in \mathbb{Z}_{>0}$. We need k monomials in the numerator or k cones in a unimodular triangulation. This is not polynomial in the input, as the input is given just by the two generators of the cone. See also Figure 9.6

If we want to use the simplicity of unimodular cones to compute the generating function of some cone efficiently, then we need a new idea. In this section we will develop such an approach using *signed decompositions* of a cone to compute the multivariate rational generating function $G(\mathbf{t})$ of a simplicial cone in polynomial time, at least for fixed dimension.

This extends to all cones by using a triangulation of the input cone without new vertices. See Theorem A.48 for the construction of such triangulations in the case of polytopes. The same construction works for cones. Note that this construction of a triangulation is polynomial in the size of the input.



(a) A cone

(b) The positive summand in the decomposition

(c) The negative summand in the decomposition

Figure 9.7.: A cone and its signed decomposition. The cone on the left is the difference of the cones on the right.

We may count lattice points in a polytope via the generating functions of the vertex cones specialized at $\mathbf{x} = \mathbf{1}$, using Brion's Theorem (Theorem 9.14). However, these values are (removable) poles of the generating functions. We will discuss a possible approach to evaluate the functions efficiently in Section 9.4.2.

9.4.1. Computing a Generating Function

The initial idea for the algorithm presented in this section is due to Barvinok[2] and Barvinok and Pommersheim.[3] For fixed dimension d the algorithm is polynomial in the input size.

As we cannot influence the number of monomials from lattice points in the fundamental domain in a given cone we need a better way to do the subdivision, if we want to arrive at a polynomial time algorithm.

The key idea for this is to use *signed* decompositions, which are decompositions where we may take new rays outside of the original cone and use addition and subtraction of rational generating functions to obtain the desired rational generating function of the original cone.

Here is an example that should explain the idea and also suggest why this may make the computation simpler. Consider again the cone of (9.3) for $k = 3$ shown in Figure 9.7(a). Its fundamental domain contains three lattice points (of which one is the origin). However, up to lattice points in the intersection, we can obtain all lattice points in this cone from the lattice points in the cone in Figure 9.7(b), and then subtract the lattice points in Figure 9.7(c). Both cones have a fundamental parallelepiped which contains only the origin. Note that this means that the numerator of the integer point generating series of these cones is 1. We can detect whether we have such a cone by computing the determinant. It is 1 if and only if the origin is the only lattice point. Equivalently, the generators of the cone are a lattice basis. We fix a name for such cones with the next definition.

[2] A. I. Barvinok, "Computing the Ehrhart polynomial of a convex lattice polytope".
[3] A. Barvinok and Pommersheim, "An algorithmic theory of lattice points in polyhedra".

Definition 9.16. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. A cone C spanned by d rays $\mathbf{r}_1, \ldots, \mathbf{r}_d \in \Lambda$ is *unimodular* if its generators are a lattice basis.

You will prove with Problem 9.6 that any two unimodular cones differ by a unimodular transformation.

Problem 9.6

It was the achievement of Barvinok in the paper mentioned above to show that with this method you can get away with a polynomial number of (even unimodular) cones.

To make this precise, let C be a d-dimensional cone spanned by primitive rays $\mathbf{r}_1, \ldots, \mathbf{r}_d$. We associate an *index* with such a cone.

Definition 9.17. The *index* of C is

$$\operatorname{index}(C) \;:=\; \left|\Pi(\mathbf{v}_1, \ldots, \mathbf{v}_d) \cap \mathbb{Z}^d\right|$$

Note that equivalently

$$\operatorname{index}(C) \;=\; |\det(\mathbf{v}_1, \ldots, \mathbf{v}_d)| \;=\; \operatorname{vol}\Pi(\mathbf{v}_1, \ldots, \mathbf{v}_d)$$

By our definitions the cone C is unimodular if and only if $\operatorname{index}(C) = 1$.

In a triangulation $\mathcal{T}$ of our cone C we will record the index of each cone in $\mathcal{T}$ (note that, for the generating series we also need to take lower dimensional cones into account to account for overcounting in intersections via inclusion-exclusion). This collection of indices is both a measure of how far we are still from a unimodular triangulation and it gives an indicator whether we are already done or which cones we need to subdivide further.

Observe that the index of a face F of a cone C is bounded by the index of C, *i.e.*

$$\operatorname{index}(F) \;\leq\; \operatorname{index}(C)\,,$$

so we actually only need to track indices of maximal cones and subdivide those if the index is still larger than one. This is, of course, only useful if we can provide a method that subdivides a cone into cones of smaller index. Here the idea of signed decompositions is needed to make this efficiently.

The basic tool in the construction is Minkowski's First Theorem (Corollary 3.3), which tells us that for any compact, convex, and centrally symmetric $K \subseteq \mathbb{R}^d$ with $\operatorname{vol} K \geq 2^d$ there exists $\mathbf{a} \neq 0$ in $K \cap \mathbb{Z}^d$, which provides us with a short nonzero vector in the lattice. We need the algorithm developed in Chapter 4 to compute this shortest vector using the LLL-algorithm, as Minkowski's First Theorem (Corollary 3.3) is not constructive.

We use this theorem (or, more precisely, the algorithm for **(SVP)**) in the following way. If, for a cone C, the index $\operatorname{index}(C)$ is still larger than 1, then

$$K \;:=\; \left\{ \frac{1}{(\operatorname{index} C)^{1/d}} \sum \lambda_i \mathbf{v}_i \;:\; -1 \leq \lambda_i \leq 1 \right\}$$

is a compact, convex, and centrally symmetric body with volume

$$\mathrm{vol}(K) \;=\; 2^d\,.$$

Hence, we can conclude that there is $\mathbf{w} \in K \cap \mathbb{Z}^d$ different from $\mathbf{0}$. We can write $\mathbf{w}$ as a linear combination of the cone generators and, as $\mathbf{w}$ is contained in K, we obtain a bound on the size of the coefficients, that is, we know that

$$\mathbf{w} \;=\; \lambda_1\mathbf{r}_1 \;+\; \lambda_2\mathbf{r}_2 \;+\; \cdots \;+\; \lambda_d\mathbf{r}_d \qquad \text{for} \qquad 0 \le |\lambda_i| \;\le\; (\mathrm{index}(C))^{-1/d}\,. \tag{9.4}$$

Instead of the exact solution of **(SVP)** we may use an approximation, as this is easier to compute.

Consider the vector $\mathbf{w}$ obtained in (9.4). By replacing $\mathbf{w}$ with $-\mathbf{w}$ if necessary we can assume that $\mathbf{w}, \mathbf{r}_1, \ldots, \mathbf{r}_d$ lie in a common half-space. Additionally we may clearly assume that $\mathbf{w}$ is primitive. By construction, $|\lambda_i| \le (\mathrm{index}\, C)^{-\frac{1}{d}}$. We define d new cones

$$C_j \;:=\; \mathrm{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_{j-1}, \mathbf{w}, \mathbf{r}_{j+1}, \ldots, \mathbf{r}_d) \qquad \text{for} \qquad 1 \le j \le d$$

by replacing $\mathbf{r}_j$ by $\mathbf{w}$ in the j-th new cone. This is the *star subdivision* of the cone with apex $\mathbf{w}$.

We compute the index of these new cones.

$$\begin{aligned}
\mathrm{index}\, C_j \;&=\; |\det(\mathbf{r}_1, \ldots, \mathbf{r}_{j-1}, \mathbf{w}, \mathbf{r}_{j+1}, \ldots, \mathbf{r}_d)| \\
&=\; \sum_{k=1}^{d} |\lambda_k| \cdot |\det(\mathbf{r}_1, \ldots, \mathbf{r}_{j-1}, \mathbf{r}_k, \mathbf{r}_{j+1}, \ldots, \mathbf{r}_d)| \\
&=\; |\lambda_j| \cdot |\det(\mathbf{r}_1, \ldots, \mathbf{r}_d)| \\
&=\; |\lambda_j| (\mathrm{index}\, C) \\
&\le\; (\mathrm{index}\, C)^{-\frac{1}{d}} (\mathrm{index}\, C) \\
&=\; (\mathrm{index}\, C)^{\frac{d-1}{d}}\,.
\end{aligned}$$

The last term strictly less than $\mathrm{index}\, C$ if $\mathrm{index}\, C \ge 2$. As the index is an integral number we see that the index actually drops by at least one. In particular, if it is less than 2, then it must already be 1.

We define a corresponding sign function to make a signed subdivision of C with the cones C_j. For $1 \le j \le d$ let

$$\varepsilon_j := \begin{cases} 0 & \text{if } \dim C_j \;<\; d \\ 1 & \text{if } \det(\mathbf{v}_1, \ldots, \mathbf{v}_d) \cdot \det(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}, \mathbf{w}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_d) \;>\; 0 \\ -1 & \text{otherwise.} \end{cases}$$

With this decomposition and the corresponding sign function the integer point generat-

ing series takes the following form.

$$\hat{G}_C(\mathbf{t}) \;=\; \sum_{j=1}^{d} \varepsilon_j \hat{G}_{C_j}(\mathbf{t}) \;+\; \text{contributions from lower dimensional cones.}$$

This decomposition of a single cone creates at most d new d-dimensional cones in our list. Note that the number may be smaller than d, as the new generator may lie in a facet of the original cone. Also note that the collection of new cones obtained by subdividing all cones doe not form a subdivision in the classical sense as defined in Definition A.45. This is only the case if all ε_j for all cones are nonnegative. Otherwise, cones obtained from subdividing different cones may overlap.

As we have to use inclusion-exclusion over all lower dimensional faces to get rid of overcounting in the common boundaries of the cones it is, however, not enough to keep track of the number of d-dimensional cones. We need to count the number of *all* cones in the signed subdivision. But subdividing a fulldimensional cone into d new cones produces at most $2^d d$ cones of any dimension.

We repeat this decomposition for each cone of index ≥ 2 in our triangulation successively until there is no cone of index greater than one left. If we want a polynomial time algorithm we need to control the total number of cones we produce. So let us obtain an upper bound for this.

After n decomposition steps, a cone D in the decomposition has index at most

$$\operatorname{index} D \;\leq\; (\operatorname{index} C)^{\left(\frac{d-1}{d}\right)^n}. \tag{9.5}$$

The algorithm stops if this this number drops below 2 (recall that the index is integral, so it must be 1). To obtain a bound we take the logarithm twice in (9.5) to solve this expression for n.

$$\begin{aligned}
\log\left(\log\left((\operatorname{index} C)^{\left(\frac{d-1}{d}\right)^n}\right)\right) \;&=\; \log\left(\left(\frac{d-1}{d}\right)^n \log(\operatorname{index} C)\right) \\
&= n\log\left(\frac{d-1}{d}\right) \;+\; \log\log(\operatorname{index} C) \qquad (9.6)\\
&= -n\log\left(\frac{d}{d-1}\right) \;+\; \log\log(\operatorname{index} C)\,.
\end{aligned}$$

Hence, for

$$n \;>\; \frac{\log\log(\operatorname{index} C)}{\log\left(\frac{d}{d-1}\right)} \;=\; \mathcal{O}(d\log\log\operatorname{index} C)$$

the last term of (9.6) is negative, so that

$$\operatorname{index} D \;\leq\; \left\lfloor (\operatorname{index} C)^{\left(\frac{d-1}{d}\right)^n} \right\rfloor \;\leq\; (\operatorname{index} C)^{\left(\frac{d-1}{d}\right)^n} \;<\; 2\,.$$

So after at most n steps all indices are 1. This shows that the number of iterations until we reach unimodular cones is indeed polynomial.

However, we also need to check that the number of cones we produce is polynomial. In n steps we produce at most

$$\begin{aligned}(d2^d)^n &= 2^{nd\log\, d} \leq 2^{Md^2\log\, d\log\log \operatorname{index} C}\\ &= (\log \operatorname{index} C)^{Md^2\log\, d}\\ &= (\log \operatorname{index} C)^{\mathcal{O}(d^2\log d)}.\end{aligned}$$

cones. Hence, we conclude that with this approach indeed, in fixed dimension, the number of cones is bounded by a polynomial in $\log \operatorname{index} C$. This is in the order of the input size of our cone in binary encoding. We summarize the algorithm the following theorem.

Theorem 9.18. *Let $d \in \mathbb{Z}_{>0}$ be fixed. Then there is a polynomial time algorithm that computes the integer point generating function $G(t)$ in the form*

$$G_C(\mathbf{t}) := \sum_{i\in I} \varepsilon_i \frac{\mathbf{t}^{\mathbf{a}_i}}{(1-\mathbf{t}^{\mathbf{v}_{i1}})\cdots(1-\mathbf{t}^{\mathbf{v}_{is_i}})},$$

where $\varepsilon_i \in \{-1,1\}$, $\mathbf{a} \in \mathbb{Z}^d$, $\mathbf{v}_{ij} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ for all i,j and $s_i \leq d$, for any d-dimensional polyhedral cone C given in its exterior description. □

Note that, in this theorem, the index set I runs over all cones, including lower dimensional ones, in the subdivision.

Remark 9.19. You can use $G_P(\mathbf{t})$ also to solve linear programs. If you want to maximize over a functional $\mathbf{c} \in \mathbb{Z}^d$, then you can just substitute $\mathbf{t} = (z^{c_1}, z^{c_2}, \ldots, z^{c_d})$. The highest degree of a monomial in the result is the optimal solution.

9.4.2. Polynomial Time Evaluation

We can use Theorem 9.18 and Brion's Theorem (Theorem 9.14) to count lattice points in a polytope P by computing the rational generating functions of all vertex cones. However, this requires us to evaluate the generating function at $\mathbf{t} = \mathbf{1}$. Although we know from the theory that these are regular points of the rational functions, they are poles in the representation we obtain. We use tools from analysis to evaluate them (Note again, that expansion into a Taylor series, though theoretically a method to remove the pole, is not an option if we want evaluate this in polynomial time.).

One possible approach is to define a curve $\gamma(s)$ for a real parameter $s \geq 0$ such that $\gamma(0)$ is $\mathbf{1}$ and the only pole of our generating function on that curve occurs for $s = 0$. Then we take the limit $s \to 0$. We need the following lemma.

Lemma 9.20. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^d$. Then there is $\mathbf{m} \in \mathbb{R}^d$ such that $\langle \mathbf{m}, \mathbf{v}_j \rangle \neq 0$ for $1 \leq j \leq k$.*

Proof. We use the moment curve

$$\mathbf{m}(\lambda) \;:=\; \left(1, \lambda, \lambda^2, \ldots, \lambda^{d-1}\right) .$$

The map

$$\lambda \longmapsto \prod_{j=1}^{d} \langle\, \mathbf{m}(\lambda), \mathbf{v}_j \,\rangle$$

is a nonzero polynomial of degree $(d-1)k$. Hence, it has at most $(d-1)k$ zeros and we can try a polynomial number of values to find a λ such that $\mathbf{m}(\lambda)$ gives the claim. □

Using this lemma we can find $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$ such that, in the notation of Theorem 9.18,

$$\langle\, \mathbf{m}, \mathbf{a}_i \,\rangle \;\neq\; 0 \qquad\qquad \langle\, \mathbf{m}, \mathbf{v}_{ij} \,\rangle \;\neq\; 0 \qquad\qquad \text{for } i \in I \text{ and } 1 \leq j \leq s_i \,.$$

Now consider

$$\gamma(r) \;:=\; (e^{rm_1}, \ldots, e^{rm_d}) \,.$$

We get the desired evaluation as

$$\lim_{r \to 0} G_P(\gamma(r)) \,.$$

Let

$$\alpha_i \;:=\; \langle\, \mathbf{m}, \mathbf{a}_i \,\rangle \qquad\qquad \nu_{ij} \;:=\; \langle\, \mathbf{m}, \mathbf{v}_{ij} \,\rangle \,.$$

Then

$$G_P(\gamma(r)) \;=\; \sum_{i \in I} \varepsilon_i \frac{e^{\alpha_i r}}{\prod_{j=1}^{s_i} (1 - e^{\nu_{ij} r})} \,,$$

and the summands are all rational functions in one variable r which are defined for all $r > 0$. We want to compute the constant term of the Laurent expansion of all summands at $r = 0$. Now consider a single such fraction. We can transform it to obtain

$$\frac{e^{\alpha_i r}}{\prod_{j=1}^{s_i} (1 - e^{\nu_{ij} r})} \;=\; \frac{1}{r^{s_i}} e^{\alpha_i r} \prod_{j=1}^{s_i} \frac{r}{1 - e^{\nu_{ij} r}} \,. \tag{9.7}$$

Now each factor

$$\frac{r}{1 - e^{\mathbf{u}_{ij} r}}$$

is defined for all r and we can compute the Laurent expansion up to degree $s_i + 1$:

$$\frac{r}{(1 - e^{\nu_{ij} r})} \;=\; T_{ij}(r) \;+\; R_{ij}(r^{s_i+1}) \,,$$

and similarly we get

$$e^{\alpha_i r} \;=\; S_i(r) \;+\; R'_{ij}(r^{s_i+1})\,.$$

We compute the product up to degree $s_i + 1$:

$$P_i(r) \;:=\; S_i(r)\prod_{j=1}^{s_i} T_{ij}(r) \;+\; R''_i(r^{s_i+1})\,.$$

Let c_i be the coefficient of r^{s_i} (note that (9.7) has an additional factor of $\frac{1}{r^{s_i}}$, so for this product c_i is the constant coefficient). We sum them up with the given signs to obtain

$$c \;:=\; \sum_{i\in I} \varepsilon_i c_i\,.$$

This is the desired limit and thus the evaluation at $\mathbf{1}$.

Remark 9.21. Using the *Todd-polynomials* $\mathrm{td}_m(\xi_1,\ldots,\xi_d)$ defined by

$$\prod_{i=1}^{k}\frac{x\xi_i}{1-e^{-x\xi_i}} \;=\; \sum_{m=0}^{\infty}\mathrm{td}_m(\xi_1,\ldots,\xi_d)x^m$$

one may obtain a closed formula for the evaluation of $G_C(\mathbf{1})$.[4] However, this requires us to evaluate the Todd polynomials.

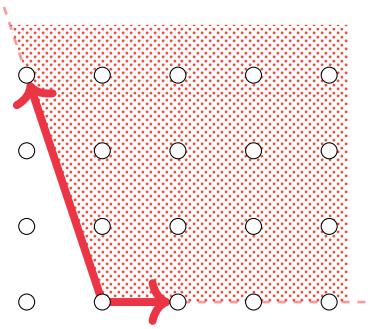
9.5. Half-open Decompositions

In computing the generating function of a cone we have to keep track of all contributions of lower dimensional cones to adjust overcounting from the intersection of cones, together with the multiplicity of overcounting. This is a huge computational effort, as we have seen above that this increases the number of newly created cones in each step of the algorithm potentially by $2^d \cdot d$ instead of just d. While we will not improve asymptotically it should be clear that we can reduce the computational effort considerably if we find a way to avoid all or most of the computations for lower dimensional cones and do not have to store the generating functions for those.

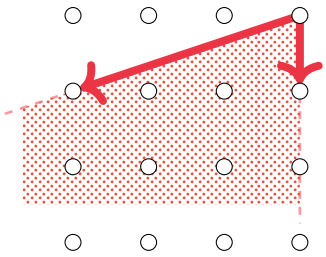
This is indeed possible, and various ways have been proposed to do so. The first idea employs the observation that cones with nontrivial lineality are mapped to 0 by Φ and that the dual of a low dimensional cone has such a nontrivial lineality space. The second modifies the subdivision in a way that each lattice point is contained in a unique cone of the subdivision, so no overcounting occurs in intersections. We briefly sketch the first approach and then turn to the second, which has the benefit that we can reuse the results in the next chapter.

For the first approach we observe that dualizing the (signed) decomposition of a cone is a (signed) decomposition of the original cone. Dualizing translates unimodular cones

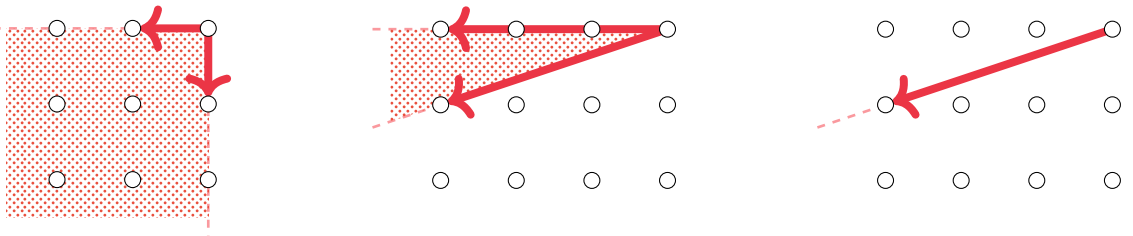
[4] Thm. 7.2.1 De Loera, Hemmecke, and Köppe, *Algebraic and geometric ideas in the theory of discrete optimization*.



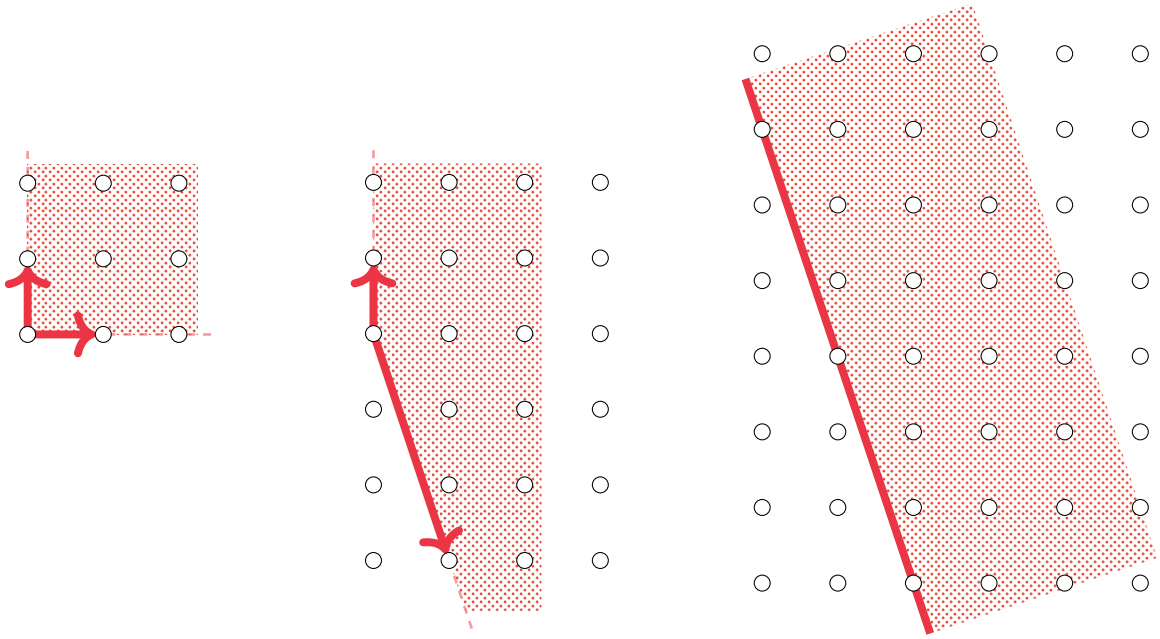
(a) A cone



(b) Its dual



(c) A signed decomposition of the dual cone. The left and right cone have positive sign, the middle one negative sign.



(d) The corresponding signed decomposition of the primal cone. Again, the left and right are taken with a positive sign, and the middle one with negative sign. The right one does not contribute to the rational function, as it is not pointed.

Figure 9.8.: A cone and a signed decomposition obtained from a signed decomposition of the dual cone.

into unimodular cones, and cones of dimension less than d into cones with a nontrivial lineality space. We can thus obtain a decomposition of a cone C in the following way. We first dualize the cone to obtain a cone $D := C^\star$. Now we apply the signed decomposition of Section 9.4.1 to D to obtain a collection $\mathcal{D} := (D_1, \ldots, D_m)$ of cones in dimensions $0 \leq k \leq d$ and signs $(\varepsilon_1, \ldots, \varepsilon_m)$ (or signs with multiplicities, if we only keep one copy of each cone appearing). Then $\mathcal{C} := (D_1^\star, \ldots, D_m^\star)$ is a signed decomposition with the same signs (and multiplicies) as $\mathcal{D}$. See Figure 9.8 for an example.

However, we can shorten this considerably. As the dual of a cone D_j with $\dim\, D_j < d$ is not pointed, ${}^\star D_j$ is mapped to 0 by the map Φ. So in the decomposition of D we can disregard all lower dimensional cones and only store the full dimensional cones. Hence, in each subdivision step of the algorithm we obtain at most d new cones instead of all $2^d \cdot d$ coming from low dimensional intersections. Also, we do not have cones with a multiplicity other than 1. So in the example of Figure 9.8 we can disregard the right cone of Figure 9.8(c), and correspondingly the right cone of Figure 9.8(d). The rational generating function of

$$C \;:=\; \operatorname{cone}\left(\left[\begin{smallmatrix}-1\\3\end{smallmatrix}\right],\, \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right)$$

is therefore

$$\begin{aligned} G_C(s,t) \;&=\; \frac{1}{(1-s)(1-t)} \;-\; \frac{1}{(1-t)(1-st^{-3})} \\ &=\; \frac{1}{(1-s)(1-t)} \;+\; \frac{s^{-1}t^3}{(1-t)(1-s^{-1}t^3)} \\ &=\; \frac{1-t^3}{(1-t)(1-s)(1-s^{-1}t^3)} \\ &=\; \frac{1+t+t^2}{(1-s)(1-s^{-1}t^3)}\,, \end{aligned}$$

where for computations we would stop with the first equation. The transformation in the form of the last line is only done to show the equivalence of the rational function in the first line with the one obtained from a naïve application of Corollary 9.8.

Also observe that a signed decomposition of the original cone would use $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ to produce the decomposition

$$\operatorname{cone}\left(\left[\begin{smallmatrix}-1\\3\end{smallmatrix}\right],\, \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right) \;=\; \operatorname{cone}\left(\left[\begin{smallmatrix}-1\\3\end{smallmatrix}\right],\, \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\right) \;+\; \operatorname{cone}\left(\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right],\, \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right) \;-\; \operatorname{cone}\left(\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\right)$$

to obtain the generating function in the form

$$\begin{aligned} G_C(s,t) \;&=\; \frac{1}{(1-s^{-1}t^3)(1-t)} \;+\; \frac{1}{(1-t)(1-s)} \;-\; \frac{1}{1-t} \\ &=\; \frac{1-t^3}{(1-t)(1-s)(1-s^{-1}t^3)} \\ &=\; \frac{1+t+t^2}{(1-s)(1-s^{-1}t^3)}\,, \end{aligned}$$

where again we would stop with the first line for computations. The following lines show that also this generating function coincides with the ones above.

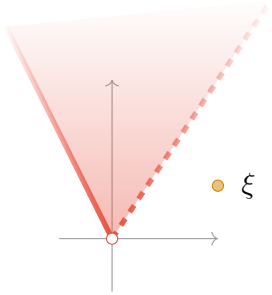


Figure 9.9.: Making a cone half open. The right face and the origin are not part of the half open cone.

Now we turn to the second approach to avoid overcounting in intersections. It has the advantage that is remains in primal space, so it avoids the expensive computation of dual cones. The general idea of the approach the following. For each lattice point $\mathbf{u}$ in the intersection of fulldimensional cones $C_1, \ldots, C_k$ we want to assign a unique cone among those that $\mathbf{u}$ should belong to, and we only count its contribution for that particular cone. There are two main ideas how we can do such a unique assignment. We discuss in the following the approach via *half-open decompositions*. I such a decomposition, we remove part of the boundary from some cells, so that each such boundary cell is contained in a unique fulldimensional one. We then argue that the whole process of computing generating functions also works for such half-open cones (with a modified fundamental domain).

The following definitions should make this precise.

Definition 9.22. Given a vector $\xi \in \mathbb{R}^d$, we define the *half-open cone* C^ξ with respect to $\xi \in \mathbb{R}^d$

$$C^\xi := \{\mathbf{y} \in C \ : \ \mathbf{y} + \varepsilon\xi \in C \text{ for all } \varepsilon > 0 \text{ small enough}\} .$$

We say $\xi \in \mathbb{R}^d$ is *generic* with respect to C (respectively, a triangulation $\mathcal{T}$ of C) if ξ is not in the linear hull of a $(d-1)$-dimensional face of C (respectively, any simplicial $(d-1)$-cone in $\mathcal{T}$).

C^ξ can also be described as precisely the set of elements in C that are not visible from ξ (Problem 9.7). See Figure 9.9 for an example.

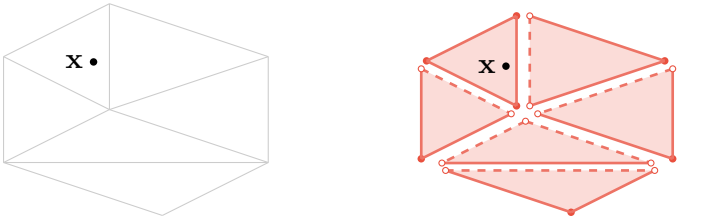
Problem 9.7

If $\xi \in C$, then $C^\xi = C$, and if ξ is generic with respect to C, then we can replace C with $\operatorname{int} C$ in the definition, so

$$C^\xi = \{\mathbf{y} \in C \ : \ \mathbf{y} + \varepsilon\xi \in \operatorname{int} C \text{ for all } \varepsilon > 0 \text{ small enough}\} ,$$

Problem 9.8

If $\xi \in C$ is generic with respect to a triangulation $\mathcal{T}$ of C, then it is also generic with respect to C. In this case, $\xi \in \operatorname{int} D \subset \operatorname{int} C$ for a unique $D \in \mathcal{T}[d]$. Our first goal is to show that *making cones half-open* is compatible with decompositions.



(a) A triangulation

(b) and a half open triangulation of it

Figure 9.10.: A triangulation and its half open decomposition.

Proposition 9.23. *Let $\mathcal{T}$ be a triangulation of the cone C, and let $\xi \in \mathbb{R}^d$ be generic with respect to $\mathcal{T}$. Then*

$$C^\xi \;=\; \bigsqcup_{D\in\mathcal{T}[d]} D^\xi\,.$$

where $\mathcal{T}[d]$ is the set of d-dimensional faces of the triangulation (see Definition A.35) is a disjoint union of half-open cones. If additionally $\xi \in C$, then

$$C \;=\; \bigsqcup_{D\in\mathcal{T}[d]} D^\xi\,.$$

A half-open decomposition in this way is illustrated in Figure 9.10, where this shows a slice through C containing ξ.

Proof. Let $\mathbf{y} \in D^\xi$. Then for any $\varepsilon > 0$ small enough $\mathbf{y} + \varepsilon\xi \in \operatorname{int}(D) \subset \operatorname{int}(C)$, so $\mathbf{y} \in C^\xi$. Conversely, let $\mathbf{y} \in C^\xi$, so $\mathbf{y} + \varepsilon\xi \in \operatorname{int}(C)$ for any $\varepsilon > 0$ small enough. This implies that there exists a unique $D \in \mathcal{T}[d]$ so that $\mathbf{y} + \varepsilon\xi \in \operatorname{int} D$ for small enough $\varepsilon > 0$. The uniqueness argument implies disjointness of the union on the right hand side. □

Problem 9.9

Let us now focus on simplicial d-cones $D \subset \mathbb{R}^d$. So let D be a simplicial d-cone in $\mathbb{R}^d$ and $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_d\} \subset \mathbb{R}^d$ the primitive ray generators. A point $\xi \in \mathbb{R}^d$ is generic with respect to D if and only if all coefficients λ_i in the unique representation

$$\xi \;=\; \sum_{i=1}^{d} \lambda_i \mathbf{v}_i$$

are non-zero.

Definition 9.24. In the setting above we define the sets

$$I_+(\xi) \;:=\; \{i \;:\; \lambda_i > 0\} \qquad \text{and} \qquad I_-(\xi) \;:=\; \{i \;:\; \lambda_i < 0\}\,.$$

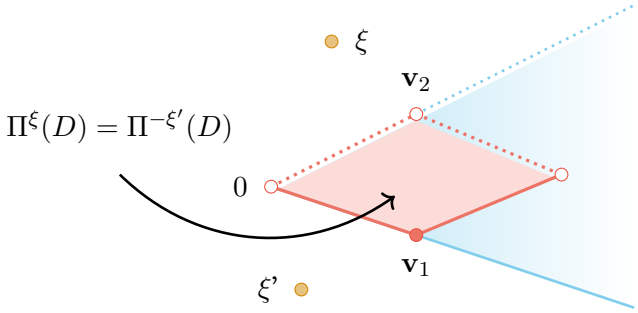


Figure 9.11.: Half open cone and fundamental parallelepiped for ξ and $-\xi'$. The dashed lines and the vertices with withe points are not part of the cone or fundamental parallelepiped.

With this notation we obtain an alternative description of a half-open simplicial cone. You will prove this result in Problem 9.10.

Lemma 9.25. *Let $\xi \in \mathbb{R}^d$ be generic with respect to a simplicial d-cone D with primitive ray generators $\mathbf{v}_1, \ldots, \mathbf{v}_d$. Then*

$$D^\xi = \left\{ \sum_{i=1}^{d} \mu_i \mathbf{v}_i \ : \ \begin{array}{l} \mu_i \geq 0 \text{ for } i \in I_+(\xi) \text{ and} \\ \mu_i > 0 \text{ for } i \in I_-(\xi) \end{array} \right\}.$$

Problem 9.10

We have seen in Corollary 2.8 that the lattice translates of the fundamental parallelepiped cover $\mathbb{R}^d$. We want to extend this to the half-open setting. With the next definition we propose the suitable generalization, and with the lemma below we will prove that the translates indeed cover $\mathbb{R}^d$.

Definition 9.26. Let D be a simplicial d-cone with primitive ray generators V and $\xi \in \mathbb{R}^d$ generic for D. We define the *half-open parallelepiped* $\Pi^\xi(D)$ with respect to ξ as

$$\Pi^\xi(D) := \left\{ \sum_{i-1}^{d} \mu_i \mathbf{v}_i \ : \ \begin{array}{l} \mu_i \in [0,1) \text{ for } i \in I_+(\xi) \text{ and} \\ \mu_i \in (0,1] \text{ for } i \in I_-(\xi) \end{array} \right\}$$

Note that $\Pi^\xi(D) \subset D^\xi$. See Figure 9.11 for an illustration.

For x strictly in the interior of D we recover the usual half-open fundamental parallelepiped of D with generating set V. The following result generalizes Corollary 2.8. For a proof see Problem 9.11.

Lemma 9.27. *Let $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_d\} \subset \mathbb{R}^d$ be linearly independent, and suppose $\xi \in \mathbb{R}^d$ is generic with respect to the simplicial cone $D := \operatorname{cone} V$. Denote by Λ the lattice generated by V.*

Then any point $\mathbf{w} \in \mathbb{R}^d$ has a unique representation $\mathbf{w} = \mathbf{y} + \mathbf{z}$ with $\mathbf{y} \in \Lambda$ and $\mathbf{z} \in \Pi^\xi(D)$.

Problem 9.11

We can further decompose each of the half-open simplicial cones into half-open boxes.

Proposition 9.28. *Let $D \subset \mathbb{R}^d$ be a simplicial cone with primitive ray generators $\mathbf{v}_1, \ldots, \mathbf{v}_d$ and $\xi \in \mathbb{R}^d$ be generic with respect to D. We define*

$$S := \left\{ \sum_{i=1}^{d} \lambda_i \mathbf{v}_i \ : \ \lambda_i \in \mathbb{Z}_{\geq 0} \right\},$$

Then we have the following disjoint union:

$$D^\xi = \bigsqcup_{\mathbf{w} \in S} \mathbf{w} + \Pi^\xi(D)$$

Proof. The fact that the translates by Λ-vectors are pairwise disjoint follows from the uniqueness in Lemma 9.27. From the existence part we see that $\mathbb{R}^d$ is covered by all Λ-translates of $\Pi^\xi(D)$. It remains to observe that for $\mathbf{w} \in \Lambda$

$$D^\xi \cap \Big(\mathbf{w} + \Pi^\xi(D)\Big) = \begin{cases} \mathbf{w} + \Pi^\xi(D) & \text{for } \mathbf{w} \in S \\ \emptyset & \text{else,} \end{cases}$$

We leave the verification of this identity to the reader (Problem 9.12). □

Problem 9.12
Problem 9.13
Problem 9.14

Finally, with all these preparations we can now compute the rational generating function of a simplicial half-open cone.

Corollary 9.29. *Let D be a simplicial cone with primitive ray generators $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\} \subseteq \mathbb{Z}^d$ and let $\xi \in \mathbb{R}^d$ be generic with respect to D.*

Then the integer point generating function of the half-open cone D^ξ is summable, and

$$G_{D^\xi}(\mathbf{t}) \ = \ \frac{G_{\Pi^\xi(D)}(\mathbf{t})}{(1-\mathbf{t}^{\mathbf{v}_1})(1-\mathbf{t}^{\mathbf{v}_2})\cdots(1-\mathbf{t}^{\mathbf{v}_d})}\,. \tag{9.8}$$

Using Proposition 9.28 the proof follows precisely along the lines of the proof of Proposition 9.7 (just replace Corollary 2.8 by Lemma 9.27). Together with Proposition 9.23 we get the following nice formula.

Corollary 9.30. *Let C be a rational cone in $\mathbb{R}^d$, let $\mathcal{T}$ be a triangulation of C into rational simplicial cones, and let $\xi \in C$ be generic. Then*

$$\hat{G}_C(\mathbf{t}) \ = \ \sum_{S \in \mathcal{T}[d]} \hat{G}_{S^\xi}(\mathbf{t}) \tag{9.9}$$

In particular the series is summable and (9.9) also holds on the level of rational functions.

Proof. Equation (9.9) is a translation of Proposition 9.23 into generating functions. By Corollary 9.29, all the summands are summable Laurent series. □

9.6. Problems

9.1. The goal of this exercise is to give a proof of Proposition 9.4. Show that the set $\mathsf{L}^{\mathrm{sum}}$ of summable Laurent series is an L-submodule of $\widehat{\mathsf{L}}$, i.e. show that for $f \in \mathsf{L}$ and $g, h \in \mathsf{L}^{\mathrm{sum}}$ also $f \cdot g$ and $g + h$ are summable.

9.2. Prove that there is a natural homomorphism from summable series to rational functions

$$\Phi : \widehat{\mathsf{L}} \longrightarrow \mathsf{R} := \Bbbk(x_1, \ldots, x_d)\,,$$

mapping $\hat{G}$ to f/g if $g\hat{G} = f$ in $\widehat{\mathsf{L}}$.

9.3. Let S, S' be subsets of the a (possibly translated) pointed cone in $\mathbb{R}^d$. Then $\hat{G}_S(\mathbf{t}) = \hat{G}_{S'}(\mathbf{t})$ implies $\hat{G}_S(\mathbf{t}) = \hat{G}_{S'}(\mathbf{t})$.

9.4. Let subsets $S_1, \ldots, S_m$ of $\mathbb{R}^d$ be given. Then

$$\hat{G}_{\bigcup_{i\in[m]} S_i} = \sum_{\emptyset \neq I \subseteq [m]} (-1)^{|I|+1} \hat{G}_{\bigcap_{i\in I} S_i}.$$

Remark: This is just the usual inclusion-exclusion formula for sets.

9.5. Apply Brion's identity to

$$P := \mathsf{conv}\begin{bmatrix} 0 & 2 & 2 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix}$$

and verify that both rational functions coincide (you may want to use a computer for this).

9.6. Let C be a unimodular cone spanned by $\mathbf{r}_1, \ldots, \mathbf{r}_d$. Show that $|\det(\mathbf{r}_1, \ldots, \mathbf{r}_d)|$ is 1 and that 0 is the only lattice point in the fundamental parallelepiped.

Show that any two unimodular cones are lattice equivalent, meaning that for any two unimodular cones C_1 and C_2 there is a linear map φ that maps $\mathbb{R}^d$ onto $\mathbb{R}^d$ and induced a bijection on the lattice such that $\varphi(C_1) = C_2$.

Those maps are the unimodular maps of Definition 2.14.

9.7. Let $\xi \in \mathbb{R}^d$ and C a d-cone. Then

$$C^{\xi} = \{\mathbf{y} \in C \,:\, \mathbf{y}_\varepsilon \in C \text{ for all } \varepsilon > 0 \text{ small enough}\}\,,$$

where $\mathbf{y}_\varepsilon := (1-\varepsilon)\mathbf{y} + \varepsilon\xi$.

9.8. Let C be a cone and ξ generic for C. Show that

$$C^{-\xi} = \operatorname{int} C\,.$$

9.9. Let $\mathcal{T}$ be a triangulation of the cone C, and let $\xi \in \mathbb{R}^d$ be generic with respect to $\mathcal{T}$. Recall the disjoint union

$$C^\xi \;=\; \bigsqcup_{D \in \mathcal{T}[d]} D^\xi$$

obtained in Proposition 9.23, where $\mathcal{T}[d]$ is the set of d-dimensional faces of the triangulation. In this problem we want to obtain a disjoint subdivision of the *interior* of C.

For this, show that we can use the negative of our generic ξ to obtain

$$\operatorname{int} C \;=\; \bigsqcup_{D \in \mathcal{T}[d]} D^{-\xi}\,.$$

Hint: You may want to solve Problem 9.8 first.

9.10. Prove Lemma 9.25.

9.11. Prove Lemma 9.27.

9.12. Check carefully and rigorously the last identity in the proof of Proposition 9.28.

9.13. Show directly that $C \backslash C[x]$ is a union of faces of C.

9.14. Let $\mathcal{T}$ be a triangulation of a full-dimensional cone C. Show that there is always a generic element $\xi \in \operatorname{int}(C)$.

9.15. Prove that, in the setting of Corollary 9.30

$$\hat{G}_{\operatorname{int} C}(\mathbf{t}) \;=\; \sum_{S \in \mathcal{T}[d]} \hat{G}_{S^{-\xi}}(\mathbf{t}).$$

Hint: You may want to solve Problem 9.9 first.

* 10. Counting Lattice Points in Dilates

In this chapter we want to look into the geometry of integral or lattice polytopes. We can use the generating functions developed in Chapter 9 to prove a classical result in the theory of lattice polytopes, the *Theorem of Ehrhart*. Its original version by Ehrhart[1] states that the number of lattice points in positive integral dilates of a d-dimensional lattice polytope are given by the evaluation of a polynomial, the *Ehrhart polynomial*, of degree d at integral values.

This theorem has since seen many generalizations, some already by Ehrhart. Many more appeared in later years, for example by MacDonald[2] on reciprocity, by Stanley[3] on the numerator of the generating function, the $h^\star$*-polynomial*, Hibi[4] on bounds of its coefficients, or the counting algorithm of Barvinok that we have seen in the previous Chapter 9. For surveys on the known results you can look at the books of Beck and Robins[5] or Barvinok.[6]

* 10.1. Some examples

We will start this chapter with a definition of the *Ehrhart counting function* $\operatorname{ehr}_P(k)$ and some simple examples of $\operatorname{ehr}_P(k)$ of a polytope P that we can compute directly. This will give some observations we will prove afterward.

Let $S \subseteq \mathbb{R}^d$, and $k \in \mathbb{Z}_{>0}$. The *k-th dilation of a set* of S is the set

$$k \cdot S \;:=\; \{\, k\mathbf{x} \;:\; \mathbf{x} \in S \,\} \,.$$

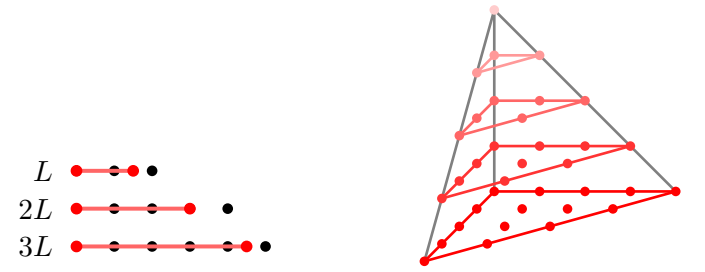
We introduce the following counting function.

Definition * 10.1. The *Ehrhart counting function* of a bounded subset $S \subseteq \mathbb{R}^d$ is the function

$$\begin{aligned} \operatorname{ehr}_S(k) \;:\; \mathbb{Z}_{>0} &\longrightarrow \mathbb{Z}_{>0} \\ k &\longmapsto \left| k \cdot S \cap \mathbb{Z}^d \right| . \end{aligned}$$

Problem * 10.1

[1] Ehrhart, *Polynômes arithmétiques et méthode des polyèdres en combinatoire*; Ehrhart, "Sur un problème de géométrie diophantienne linéaire. II. Systèmes diophantiens linéaires".
[2] Macdonald, "Polynomials associated with finite cell-complexes".
[3] Stanley, "Decompositions of rational convex polytopes".
[4] Hibi, "A lower bound theorem for Ehrhart polynomials of convex polytopes".
[5] Beck and Robins, *Computing the continuous discretely*.
[6] A. Barvinok, *A course in convexity*.



(a) Lattice points in an interval

(b) Lattice points in a standard simplex.

Figure * 10.1.: Some examples of Ehrhart polynomials

For $a, b \in \mathbb{R}$ we define the interval

$$L \;:=\; [a,b] \;\subseteq\; \mathbb{R}$$

on the real line. Counting lattice points in dilates is relatively simple here. The k-th dilate of L is $[ka, kb]$, and it contains $\lfloor kb \rfloor - \lceil ka \rceil + 1$ integral points, so

$$\mathrm{ehr}_L(k) \;=\; \lfloor kb \rfloor - \lceil ka \rceil + 1\,.$$

Figure * 10.1(a) shows the interval $I = [0, \frac{3}{2}]$ and its second and third dilation.

If the boundary points a and b are integral and $a \leq b$, then we can simplify the formula. In this case also all multiples of a and b are integral, and we can omit the floor and ceiling operations to obtain

$$\mathrm{ehr}_L(k) \;=\; k(b-a)+1\,.$$

We observe that this is a polynomial of degree 1 in k.

Let us consider some more examples in general dimension $d \geq 1$. The *standard simplex*

$$\Delta_d \;:=\; \mathsf{conv}(\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_d)\,,$$

see Figure * 10.1(b) for the lattice points in a multiple of this simplex. The following proposition gives the number of lattice points in its dilates.

Proposition * 10.2. *Let Δ_d be the d-dimensional standard simplex. Then*

$$\mathrm{ehr}_{\Delta_d}(k) \;=\; \binom{d+k}{d} \;=\; \frac{(d+k)\cdot(d+k-1)\cdot\ldots\cdot(k+1)}{d!}\,.$$

Observe that this is a polynomial in the variable k of degree d with leading coefficient $1/d!$.

Problem * 10.2

Proof. There is a bijection between the lattice points in $k\,\Delta_d$ and sequences of k dots and d bars: to each such sequence, assign the vector $\mathbf{x} \in \mathbb{R}^d$ whose i-th coordinate

equals the number of dots between the i-th bar and the $(i+1)$st bar for $1 \leq i \leq d-1$ (we don't write down the number of dots after the last bar, it is determined by the rest):

$$\cdot\,\cdot\,|\,\cdot\,\cdot\,\cdot\,||\,\cdot \qquad \longleftrightarrow \qquad \mathbf{x} \;=\; (2,3,0)$$

This yields a bijection between the sequences and lattice points with non-negative coordinates and with $\sum x_i \leq k$. □

Another simple, but very important example is the unit cube defined in Example A.19(i). The k-th dilate of the cube is $kC_d = k \cdot [0,1]^d = [0,k]^d$. Hence, the Ehrhart counting function is given by

$$\mathrm{ehr}_{C_d}(k) \;=\; (k+1)^d\,.$$

Note again that this is a polynomial in k of degree d.

Problem * 10.3

* 10.2. The Ehrhart Polynomial

Now we turn to the proof that the number of lattice points in dilates is given by a polynomial. We aim for the following theorem.

Theorem * 10.3 (Ehrhart's Theorem). *The Ehrhart counting function given by $k \mapsto \mathrm{ehr}_P(k)$ for $k \in \mathbb{Z}_{\geq 1}$ extends to a polynomial function $t \mapsto \mathrm{ehr}_P(t)$ of degree d.*

With the following definition we assign a name to the function in this theorem. The proof needs some more preparations.

Definition * 10.4 (Ehrhart *polynomial*). For a polytope P the polynomial $\mathrm{ehr}_P(t)$ as in the previous theorem is the *Ehrhart polynomial* of P.

Let us first argue why we should expect that the polynomial has degree d if it exists. For this we look at the *volume* of the polytope, or, more generally, any convex body. Computing the volume generally is a difficult task, and we will argue that counting lattice points in dilates is related.

Let $K \subseteq \mathbb{R}^d$ be a convex body. We can approximate the volume of K by counting the volume of little cubes with vertices at lattice points of the refined lattice $\left(\frac{1}{k}\mathbb{Z}\right)^d$. This approximates the volume for $k \to \infty$, see Proposition A.4. Figure * 10.2 shows an illustration of the approach. With this idea we can make the following computation.

$$\begin{aligned} \mathrm{vol}(K) \;=\; \int_P \mathrm{dx} \;&=\; \lim_{k\to\infty} \frac{1}{k^d} \left| K \cap \frac{1}{k}\mathbb{Z}^d \right| \\ &=\; \lim_{k\to\infty} \frac{1}{k^d} \left| k\cdot K \cap \mathbb{Z}^d \right| \\ &=\; \lim_{k\to\infty} \frac{1}{k^d}\, \mathrm{ehr}_K(k) \end{aligned} \qquad (* \ 10.1)$$

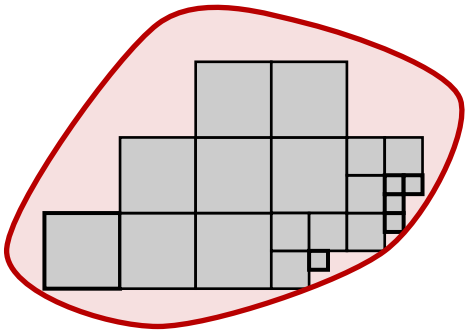


Figure * 10.2.: Approximating a convex body by smaller and smaller cubes

Hence, we can compute the limit if we can compute $\mathrm{ehr}_K(k)$ for all k. Further, if we know that $\mathrm{ehr}_K(k)$ is a polynomial function for some K, then it necessarily is a polynomial of degree d, as otherwise the limit would be 0 or ∞.

In particular, the Ehrhart polynomial would be determined by knowing $d+1$ many values of it. It can be shown (and you can deduce this by following the inclusion-exclusion for the computation of generating functions carefully) that the constant term will always be 1. Hence, in this case, $\mathrm{vol}(K)$ can be explicitly computed from $|k \cdot K \cap \mathbb{Z}^d|$ for $k = 1, \ldots, d$. This may remind you of Problem 2.11, where you proved that the integer points in dilates of a polygon are given by a polynomial of degree 2. You may want to extend this to dimension 3 with Problem * 10.4.

Problem * 10.4
Problem * 10.5
Problem * 10.6
Problem * 10.7

Let us now turn to the proof of Ehrhart's Theorem (Theorem * 10.3). For this, we homogenize our polytopes and work with the cone over P instead of P. Recall that we have defined $C(P)$ in (A.10) via

$$C(P) \;:=\; \mathrm{cone}\,(\{1\} \times P) \;\subseteq\; \mathbb{R}^{d+1}\,,$$

We usually write a vector $\mathbf{x} \in \mathbb{R}^{d+1}$ with indices starting from 0 and use x_0 for the special coordinate. The homogenization is convenient in our setting as we can recover all dilates of P from $C(P)$. More precisely, for any $k \geq 0$ we get the k-th dilate of P by intersecting $C(P)$ with the hyperplane $x_0 = k$, and the lattice points in kP by intersecting with $\{k\} \times \mathbb{Z}^d$.

Hence,

$$\hat{G}_{C(P)}(t, 1, \ldots, 1) \;=\; \sum_{k \geq 0} \Big| kP \cap \mathbb{Z}^d \Big|\; t^k \;=\; 1 + \sum_{k \geq 1} \mathrm{ehr}_P(t) t^k.$$

As substituting variables clearly keeps summability, by Corollary 9.8 the following definitions make sense.

Definition * 10.5. Let P be a lattice d-polytope. The *Ehrhart series* of P is the summable formal Laurent series

$$\widehat{\mathrm{Ehr}}_P(t) \;:=\; 1 + \sum_{k \geq 1} \mathrm{ehr}_P(t) t^k \;\in\; \Bbbk[\![t]\!]$$

in one variable t. The corresponding rational function will be denoted by

$$\mathrm{Ehr}_P(t) \;:=\; \Phi(\widehat{\mathrm{Ehr}}_P(t)) \;\in\; \Bbbk(t)\,.$$

To proceed we consider some well-known results on generating functions.

Lemma * 10.6. *For $j \in \mathbb{Z}_{\geq 0}$,*

$$\Phi\left(\sum_{\mathbb{Z}_{\geq 0}} \binom{k+d-j}{d} z^k\right) \;=\; \frac{z^j}{(1-z)^{d+1}}\,.$$

The proof will be given in Problem * 10.8.

Problem * 10.8

Proposition * 10.7. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be such that*

$$\Phi\left(\sum_{t=0}^{\infty} f(t)z^t\right) \;=\; \frac{g(z)}{(1-z)^{d+1}}\,.$$

Then $f(t)$ is a polynomial of degree at most d if and only if $g(z) = \sum_{k\in\mathbb{Z}_{\geq 0}} g_k z^k$ is a polynomial of degree at most d. In this case we can write f in terms of g via

$$f(t) \;=\; g_0\binom{t+d}{d} + g_1\binom{t+d-1}{d} + \ldots + g_d\binom{t}{d}\,.$$

and the leading coefficient of f is $\frac{g(1)}{d!}$. In particular, f has degree d if and only if $g(1) \neq 0$.

Problem * 10.9

Proof. We define the polynomials $f_j(t) := \binom{t+d-j}{d}$ for $0 \leq j \leq d$. By Problem * 10.9 the set $\{f_o, \ldots, f_d\}$ is a basis of $\mathbb{R}[t]_{\leq d}$.

Let f be a polynomial of degree at most d. Then there are $g_0, \ldots, g_d$ such that

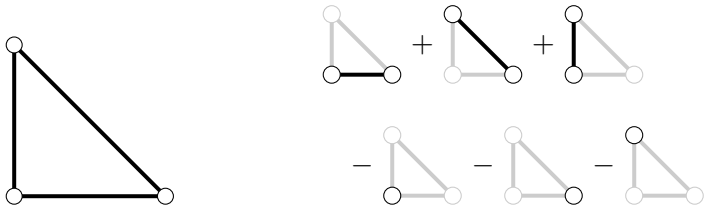
$$f(t) = \sum_{j=0}^{d} g_j f_j(t) = \sum_{j=0}^{d} g_j \binom{t+d-j}{d}\,.$$

The coefficient of t^d is $\frac{1}{d!}\sum g_j$. We compute

$$\sum_{t\geq 0}\sum_{j=0}^{d} g_j\binom{t+d-j}{d}z^k = \sum_{j=0}^{d} g_j \sum_{t\geq 0}\binom{t+d-j}{d}z^k\,.$$

Using Lemma * 10.6 we obtain

$$\Phi\left(\sum_{t\geq 0}\sum_{j=0}^{d} g_j\binom{t+d-j}{d}z^k\right) \;=\; \frac{\sum_{j=0}^{d} g_j z^j}{(1-z)^{d+1}} \;=\; \frac{g(z)}{(1-z)^{d+1}}$$



(a) The boundary complex of a triangle

(b) Inclusion-Exclusion on the boundary complex

Figure * 10.3.: Decomposing a triangle

For the converse direction, injectivity of Φ on polynomials implies $f(t) = \sum_{j=0}^{d} g_j \binom{t+d-j}{d}$. Again using Problem * 10.9, which tells us that the polynomials on the right are a basis, we obtain the claim. □

Now, let us compute the Ehrhart generating function for lattice simplices.

Proposition * 10.8. *Let S be a d-simplex. Then*

$$\mathrm{Ehr}_S(t) \;=\; \frac{h^\star(t)}{(1-t)^{d+1}}$$

where $h^\star$ is a polynomial of degree $\leq d$. Further, for $h^\star(t) = \sum_{k=0}^{d} h_k^\star t^k$, we have

$$h_k^\star \;=\; \Big|\Pi\left(C(S)\right) \,\cap\, \mathbb{Z}^{d+1} \,\cap\, \{\mathbf{x} \mid x_0 = k\}\Big| \;\in\; \mathbb{Z}_{\geq 0}.$$

In particular, $h_0^\star = 1$ and $h^\star(1) \neq 0$.

Proof. Let $\{\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_d\}$ be the vertex set of $\{1\} \times S$, with $\mathbf{a}_i = (1, \mathbf{v}_i)$ for $i = 0, \ldots, d$. Applying the substitution $(t_0, t_1, \ldots, t_d)$ by $(t_0, 1, \ldots, 1)$ to Proposition 9.7 we obtain that

$$\mathrm{Ehr}_S \;=\; \frac{h^\star(t_0)}{(1-t_0)^{d+1}}$$

for the polynomial $h^\star(t_0) = \sum_{(y_0,\mathbf{y}) \in \Pi(C(S)) \cap \mathbb{Z}^{d+1}} t_0^{y_0}$.

Let $(y_0, \mathbf{y}) = \sum_{i=0}^{d} \lambda_i (1, \mathbf{v}_i) \in \Pi(C(S)) \cap \mathbb{Z}^{d+1}$, so $0 \leq \lambda_i < 1$ for $i = 0, 1, \ldots, d$. In particular, $y_0 < d+1$, so $y_0 \leq d$. Moreover, $y_0 \geq 0$ with equality if and only if also $\mathbf{y} = (0, 0, \ldots, 0)$. □

We have now collected all necessary tools and definitions to prove Ehrhart's Theorem (Theorem * 10.3).

*Proof of Ehrhart's Theorem (Theorem * 10.3).* Combining Proposition * 10.8 with Proposition * 10.7 we get that the Ehrhart counting function of an n-dimensional lattice simplex in $\mathbb{R}^d$ uniquely extends to a polynomial function of degree at most n.

For a general polytope P we triangulate it into maximal-dimensional simplices F_i and consider the triangulation of $C(P)$ into the associated simplicial cones $C(F_i)$. Then we apply inclusion-exclusion using Problem 9.4. □

From this theorem and (* 10.1) we deduce the following corollary.

Problem * 10.10
Problem * 10.11

Corollary * 10.9. *The leading coefficient of* $\mathrm{ehr}_P(t)$ *is* $\mathrm{vol}(P)$. □

With this corollary we have determined the highest coefficient in the polynomial. So far, we do not know anything about the other coefficients. In particular, we do not know the constant coefficient. From our considerations so far you may come to the conclusion that

$$\mathrm{ehr}_P(0) \;=\; \left|0P \cap \mathbb{Z}^d\right| \;=\; 1\,.$$

This is indeed the case if P is a polytope, but it hides the true meaning of this coefficient, which has a different geometric interpretation.

Namely, instead of counting in a single polytope, we can count lattice points in dilations of complexes of lattice polytopes. The entire chain of arguments given carries over to this setting. We obtain a counting function which is the evaluation of a polynomial. Consider, for example, $\mathcal{C}$ to be the boundary of a standard triangle, see Figure * 10.3(a). Then our counting polynomial turns out to be $\mathrm{ehr}_{\mathcal{C}}(k) = 3k$. This polynomial has constant coefficient zero See Figure * 10.3(b).

You may now guess that the constant coefficient is the Euler characteristic of the complex, which we have already met in the proof of the Theorem of Brianchon-Gram (Theorem 9.9), see also (A.13). This is in fact true, and the 1 for polytopes comes from the fact that balls have Euler characteristic 1. If you recall how we computed the generating function of a cone via the principle of inclusion-exclusion, then you may deduce this fact by following the count of the origin in this. You will see that this amounts exactly to the computation of the Euler characteristic of a complex from its cells (or you know this already from a course on *Algebraic Topology*).

In Problem * 10.12 you will discover a geometric interpretation of the second highest coefficient of the Ehrhart polynomial.

Problem * 10.12
Problem * 10.13

* 10.3. Problems

* 10.1. Compute the Ehrhart generating function of $P = [0,1]^2$ using the Brion's Theorem (Theorem 9.14).

* 10.2. Show

$$\sum_{k=0}^{\infty} \binom{k+d}{d} x^k = \frac{1}{(1-x)^{d+1}}.$$

(Why is the equality sign justified here?)

* 10.3. Compute the Ehrhart counting function of the cross polytope.

* 10.4. Determine a formula for the volume of a 3-dimensional lattice polytope using the number of lattice points in the k-multiple for $k = 1, 2$ and 3.

* 10.5. Determine the Ehrhart polynomial of the Reeve simplices defined by

$$\mathsf{R}_d(m) \;:=\; \mathsf{conv}\,(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 \,+\, \mathbf{e}_2 + m\mathbf{e}_3) \qquad (* \ 10.2)$$

for $m \in \mathbb{Z}_{\geq 1}$.
What do you observe for $m = 20$?

* 10.6. Let P be a lattice polytope with Ehrhart polynomial $\mathrm{ehr}_P(t)$. Compute the Ehrhart polynomial of the bipyramid over P.

* 10.7. For integers p,q with $\gcd(p, q) = 1$ define the tetrahedron

$$\Delta_{pq} = \mathsf{conv} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & p \\ 0 & 0 & 0 & q \end{bmatrix} .$$

(i) Show that its vertices are its only lattice points.
(ii) Compute the Ehrhart polynomial of Δ_{pq}.
(iii) Determine for which parameters Δ_{pq} and $\Delta_{p'q'}$ are unimodularly equivalent.

White[7] proved a converse of the first claim. He showed that every lattice tetrahedron with only four lattice points is unimodularly equivalent to a Δ_{pq}.

* 10.8. Prove Lemma * 10.6.

Hint: do $j = 0$ first

* 10.9. Show that $\binom{t+d-j}{d}$ for $j = 0, \ldots, d$ is a basis of the polynomials of degree at most d.

* 10.10. Prove that the coefficients of the Ehrhart polynomial of a d-dimensional lattice polytope are in $\mathbb{Z}/d!$.

* 10.11. Show that $\sum_{i=0}^{d} k^2 = \binom{d+2}{3} + \binom{d+1}{3}$

Hint: You should try to do this via Ehrhart Theory. Consider a $(d-1)$-fold pyramid over a square.

* 10.12. Let P be a d-dimensional lattice polytope with Ehrhart polynomial $\sum_{k=0}^{d} c_k t^k$. Show that

$$c_{d-1} = \frac{1}{2}\,\mathrm{vol}(\partial P).$$

Here, $\mathrm{vol}(\partial P)$ denotes the surface area of P, namely,

$$\mathrm{vol}(\partial P) := \sum_{F \in \mathsf{F}(P)} \mathrm{vol}(F),$$

where $\mathsf{F}(P)$ is the set of facets of P and $\mathrm{vol}(F)$ denotes the (non-normalized) volume with respect to the lattice $\mathrm{aff}(F) \cap \mathbb{Z}^d$. For instance, note that $\mathrm{vol}(\mathsf{conv}((1,0),(0,1)))$ equals 1 and not $\sqrt{2}$. Hence,

$$\mathrm{vol}(\partial\, \mathsf{conv}((1,0),(0,1),(-1,0),(0,-1))) = 4\,.$$

* 10.13. A simplex which is unimodularly equivalent to the standard simplex is called unimodular. A triangulation is unimodular if all its simplices are.
(i) For a k-dimensional unimodular simplex Δ and $t \in \mathbb{Z}_{\geq 1}$ show that

$$|\mathbb{Z}^k \cap \mathrm{relint}(t\Delta)| = \binom{t-1}{k}.$$

[7] White, "Lattice tetrahedra".

(ii) Suppose P admits a unimodular triangulation $\mathcal{T}$ with $f_0(\mathcal{T})$ vertices, $f_1(\mathcal{T})$ edges, ..., $f_d(\mathcal{T})$ d-simplices. Show that

$$\mathrm{ehr}_P(t) = \sum_{k=0}^{d} f_k(\mathcal{T}) \binom{t-1}{k} .$$

(iii) Conclude that any two unimodular triangulations have the same f-vector $(f_0, \ldots, f_d)$.

* 11. Cuts and Lattice Free Polytopes

In this chapter we consider mixed integer linear programs (MILP) of the form

$$\begin{aligned} A\mathbf{x} &\leq \mathbf{b} \\ x_i &\in \mathbb{Z} \qquad \text{for} \quad i \in I \end{aligned} \tag{* 11.1}$$

for some $A \in \mathbb{R}^{m \times d}$, $\mathbf{b} \in \mathbb{R}^d$, and some index set $I \subseteq \{1, \ldots, d\}$. To simplify the notation we usually assume that $I = \{1, \ldots, k\}$ for some $0 \leq k \leq d$, so that we can replace the second line with

$$\mathbf{x} \in \mathbb{Z}^k \times \mathbb{R}^{d-k}.$$

We have an associated polyhedron

$$P := \left\{ \mathbf{x} \in \mathbb{R}^d : A\mathbf{x} \leq \mathbf{b} \right\}.$$

The set of feasible solutions is this polyhedron intersected with $\mathbb{Z}^k \times \mathbb{R}^{d-k}$. The *mixed integer hull* is the polyhedron

$$P_I := \mathsf{conv}\left(P \cap \mathbb{Z}^k \times \mathbb{R}^{d-k} \right).$$

We will only consider the task to decide feasibility of the mixed integer program. In general solving this problem is NP-hard, so there is no direct efficient algorithm known. As this is nevertheless a important problem for many real world applications, much effort has been put into methods to solve or approximate the solution efficiently for instances that appear in applications. We have discussed some methods in the beginning of Chapter 6.

Here we want to focus on the approach via cutting planes or *cuts*. Basically, a *cutting plane* is a hyperplane

$$H := \{ \mathbf{x} : \langle \mathbf{a}, \mathbf{x} \rangle = \delta \}$$

in $\mathbb{R}^d$ that induces a valid inequality for the feasible solutions of (* 11.1), *i.e.*

$$\langle \mathbf{a}, \mathbf{x} \rangle \leq \delta$$

for all $\mathbf{x}$ that satisfy $A\mathbf{x} \leq \mathbf{b}$ and $x_i \in \mathbb{Z}$ for $i \in I$, but which is not valid for the linear relaxation, where we drop all integrality constraints, *i.e.* there is $\mathbf{y} \in \mathbb{R}^d$ such that

$$\langle \mathbf{a}, \mathbf{y} \rangle > \delta.$$

Then necessarily $y_i \notin \mathbb{Z}$ for some $i \in I$.

Various ways have been devised to come up with cutting planes for a mixed integer program. Algorithmic methods usually start with a solution to the linear relaxation of the program. If this is not a feasible solution of the mixed integer program, then they search for a cutting plane that is violated by this solution.

On the more theoretical side one can ask for families C of pairs $(\mathbf{a}_i, \delta_i)$ defining cutting planes, such that

$$P_I \;=\; \{\, \mathbf{x} \;:\; \mathbf{x} \in P \text{ and } \langle\, \mathbf{a}_i, \mathbf{x} \,\rangle \leq \delta_i \text{ for all } i \,\}\;,$$

or for families of such hyperplanes that approximate P_I in the sense that

$$P \;\subsetneq\; Q \;\subseteq\; P_I \tag{* 11.2}$$

for the polyhedron Q obtained by adding all inequalities in the family. Most interesting in this contex are families of cuts that produce P_I after a finite number of steps.

An example for the first approach are the *Gomory (mixed) integer cuts* that you have discussed in *Discrete Optimization*. An example for the second are the *Chvátal cuts* for integer programs also discussed in *Discrete Optimization*, together with the associated *elementary closures*. The *elementary closure* $P^{(1)}$ of a polyhedron P is the intersection of P with all halfspaces of the form

$$\{\, \mathbf{x} \;:\; \langle\, \mathbf{a}, \mathbf{x} \,\rangle \leq \lfloor \delta \rfloor \,\} \tag{* 11.3}$$

where $\mathbf{a} \in \mathbb{Z}^d$, $\delta \in \mathbb{R}$, $\langle\, \mathbf{a}, \mathbf{x} \,\rangle \leq \delta$ is a valid inequality for P, but $\langle\, \mathbf{a}, \mathbf{x} \,\rangle \leq \lfloor \delta \rfloor$ is not. The polyhedron $P^{(1)}$ (which corresponds to Q in (* 11.2)) contains P_I, but they usually differ.

It can be shown that after a finite iteration of elementary closures

$$P^{(t)} \;:=\; \left(P^{(t-1)}\right)^{(1)}$$

we arrive at some t such that $P^{(t)} = P_I$. Any Gomory cut is a Chvátal cut, so that the former can be seen as an algorithmic version of the latter. Both can also be seen as *split cuts*, which arise in the following way. Given a polyhedron P, some functional $\mathbf{c} \in \left(\mathbb{R}^d\right)^\star$, and $\pi \in \mathbb{R}$ a *split cut* for P is any inequality $\langle\, \mathbf{a}, \mathbf{x} \,\rangle \leq \delta$ that is simultaneously valid for

$$\{\, \mathbf{x} \in P \;:\; \langle\, \mathbf{c}, \mathbf{x} \,\rangle \leq \lfloor \pi \rfloor \,\} \qquad \text{and} \qquad \{\, \mathbf{x} \in P \;:\; \langle\, \mathbf{c}, \mathbf{x} \,\rangle \geq \lfloor \pi \rfloor + 1 \,\} \tag{* 11.4}$$

In the integer case adding this inequality to the system does not change the feasible set. In the mixed integer case we take such an inequality in the first k variables and extend with a linear space in the remaining (nonintegral) ones.

The disjunction of (* 11.4) can be seen as intersecting P with the a polyhedron

$$Q \;:=\; S \times \mathbb{R}^{d-1} \;=\; \{\, \mathbf{x} \;:\; \lfloor \pi \rfloor \leq \langle\, \mathbf{c}, \mathbf{x} \,\rangle \leq \lfloor \pi \rfloor + 1 \,\}$$

for an interval

$$S := \{ \lambda \mathbf{r} : \lfloor \pi \rfloor \leq \langle \mathbf{c}, \lambda \mathbf{r} \rangle \leq \lfloor \pi \rfloor + 1 \}$$

and some $\mathbf{r}$ not in the nullspace of $\mathbf{c}$. So S is a onedimensional polyhedron whose facets are given by the two inequalities $\lfloor \pi \rfloor \leq \langle \mathbf{c}, \mathbf{x} \rangle$ and $\langle \mathbf{c}, \mathbf{x} \rangle \leq \lfloor \pi \rfloor + 1$ and that has one vertex in each of the two hyperplanes. Note that there is no integral point in the interior of S, but its two vertices may be integral. Such sets are called *lattice free*.

In the following we want to discuss a generalization of this, where we replace the interval S by any *lattice free* polyhedron, *i.e.* by a polyhedron that has no lattice points in its interior (but maybe some on the boundary).

In the next section we will first derive this idea from an algorithmic point of view that follows the approach for Gomory cuts, but uses more than one row of the simplex tableaux. We will see that we need lattice free polyhedra to describe the cuts obtained in this way, and the maximal cuts come from *maximal lattice free* polyhedra, which are those lattice free polyhedra, that are not contained in a strictly larger lattice free polyhedron.

In the second section we will give a characterization of maximal lattice free polyhedra due to Lovasz[1], with proofs by Basu et al.[2] and Averkov.[3] It can then be shown that it suffices to look at the integral maximal lattice free polyhedra (those with integral vertices),[4] to obtain a sequence of cuts that produce P_I after a finite number of iterations. Further, in each dimension we only have a finite number of maximal lattice free polyhedra.[5]

In the last section we give a brief outlook on further results on lattice free sets. In particular we sketch the proof that there are only finitely many lattice free polyhedra in each dimension, up to lattice equivalence. The polytopes have been classified up to dimension 3, and we conclude the chapter with the complete list of such polytopes.

* 11.1. Corner Polyhedra

In this section we want to give an algorithmic motivation for the study of lattice free polyhedra. For this we introduce the *corner polyhedron* of a mixed integer program

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq 0 \\ x_i &\in \mathbb{Z} \qquad \text{for} \qquad i \in I \end{aligned} \tag{* 11.5}$$

in in standard form for some $I \subseteq \{1, \ldots, k\}$ (and we will again mostly assume that $I = \{1 \ldots, k\}$ for some $0 \leq k \leq d$). This is the form used for the simplex algorithm in linear programming, and thus the usual form if one solves a relaxation of the mixed

[1] Lovász, "Geometry of numbers and integer programming".
[2] Basu, Conforti, Cornuéjols, and Zambelli, "Maximal lattice-free convex sets in linear subspaces".
[3] Averkov, *A proof of Lovász's theorem on maximal lattice-free sets*.
[4] Del Pia and Weismantel, "On convergence in mixed integer programming".
[5] Averkov, Wagner, and Weismantel, "Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three".

integer program and then tries to find new inequalities that separate a non-integral solution from the set of feasible solutions.

These corner polyhedra were already introduced by Gomory[6] as a generalization of cutting planes. and subsequently studied by Gomory and Johnson.[7] In our discussion we will follow an approach to generate new cuts introduced by Anderson et al.[8]. A good survey on this an many related results and extensions is given by Basu et al.[9]

We recall some notation for such a linear program. A basis B is a maximal minor of linearly independent columns of A, so $\det B \neq 0$. The associated basic solution is

$$\mathbf{x}_B := \begin{bmatrix} B^{-1}\mathbf{b} \\ 0 \end{bmatrix}$$

We collect the remaining columns not in B into a matrix N, the *nonbasic variables*. By slight abuse of notation we denote by $\mathbf{x}_B$ also the subset of the entries of $\mathbf{x}$ that correspond to the basis variables, and by $\mathbf{x}_N$ the remaining entries.

We can no give a formal definition of a corner polyhedron.

Definition * 11.1. For a mixed integer linear program as in (* 11.5) and a basis B the *corner polyhedron* associated to B is

$$\mathrm{Corner}(B) := \mathsf{conv}\left(\left\{ \mathbf{x} \in \mathbb{R}^d : \begin{array}{l} \mathbf{x}_B = B^{-1}\mathbf{b} - B^{-1}N\mathbf{x}_N \\ \mathbf{x}_N \geq 0 \\ x_i \in \mathbb{Z} \text{ for } i \in I \end{array} \right\} \right)$$

Note that in this formulation we have dropped the nonnegativity constraints on the basic variables. In the following we only look at one specific corner polyhedron.

If $(\mathbf{x}_B, \mathbf{x}_N) = (B^{-1}\mathbf{b}, 0)$ is a feasible solution then we have found the solution to the mixed integer program (assuming that the simplex tableaux corresponded to a final solution of the relaxation). Otherwise, note that $\mathbf{x}_N$ are the independent variables in the current formulation, and we aim for an inequality, that separates the solution $\mathbf{x}_N = 0$ obtained from the relaxation from the set of feasible solutions, which are all contained in $\mathrm{Corner}(B)$.

Let us introduce new variables to make the formulation simpler. We set

$$(\mathbf{c}, \mathbf{s}) := \mathbf{x}_N$$

where $\mathbf{c}$ collects the integer variables among the nonbasic variables and $\mathbf{s}$ the continuous ones. We can then split $B^{-1}N$ accordingly into

$$(C, R) := B^{-1}N\,,$$

set $\mathbf{f} := B^{-1}\mathbf{b}$ and replace $\mathbf{x}_B$ by a variable $\mathbf{y}$ to write the corner polyhedron as the

[6] Gomory, "Some polyhedra related to combinatorial problems".
[7] Gomory and Johnson, "Some continuous functions related to corner polyhedra"; Gomory and Johnson, "Some continuous functions related to corner polyhedra. II".
[8] Andersen, Louveaux, Weismantel, and L. A. Wolsey, "Inequalities from two rows of a simplex tableau".
[9] Basu, Conforti, and Summa, "A geometric approach to cut-generating functions".

convex hull of

$$\begin{aligned} \mathbf{y} &= \mathbf{f} - C\mathbf{c} - R\mathbf{s} \\ \mathbf{u}, \mathbf{v} &\geq 0 \\ \mathbf{u} &\in \mathbb{Z}^k \end{aligned} \qquad \text{(* 11.6)}$$

Note that $\mathbf{y}$ may still combine integer and continuous variables, as we have only split $\mathbf{x}_N$ in two parts. However, for the solutions to (* 11.6) the continuous variables among $\mathbf{y}$ are just given by the equation in the first line, and are not subject to any further constraint, so we can drop these variables in our considerations and assume that $\mathbf{y} \in \mathbb{Z}^d$.

Following Anderson et al.[10] we now also drop the integrality constraints on the nonbasic variables. So all nonbasic variables are in $\mathbf{s}$. Then (* 11.6) describes an affine cone, which we can write in the form

$$\begin{aligned} \mathbf{y} &= \mathbf{f} + \sum_{i=1}^{k} s_i \mathbf{r}_i \\ \mathbf{y} &\in \mathbb{Z}^m \\ s_i &\geq 0 \text{ for all } i\,. \end{aligned} \qquad \text{(* 11.7)}$$

where $\mathbf{r}_i$ are the rays of the cone given by the columns of R.

We denote the set of all solutions of (* 11.7) by $S_{\mathbf{f}}(\mathbf{r}_1, \ldots, \mathbf{r}_k)$, and the convex hull of this by

$$R_{\mathbf{f}}(\mathbf{r}_1, \ldots, \mathbf{r}_k) := \mathsf{conv}(S_{\mathbf{f}}(\mathbf{r}_1, \ldots, \mathbf{r}_k)\,.$$

We assume that $\mathbf{f}$ is a fractional solution, *i.e.* at least one f_i is not integral (otherwise its a feasible solution of the mixed integer program). We aim for inequalities that cut of the solution, *i.e.* the apex, from the cone.

For this recall that any centrally symmetric convex body $K \subseteq \mathbb{R}^d$ defines a norm $\|.\|_K$ via

$$\|\mathbf{x}\|_K := \inf(t > 0 : \mathbf{x} \in t \cdot K)\,. \qquad \text{(* 11.8)}$$

We can obtain K from its norm via

$$K = \{\mathbf{x} : \|\mathbf{x}\|_K \leq 1\} \qquad \text{(* 11.9)}$$

and the norm is linear and satisfies the triangle inequality, *i.e.* for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$

$$\|\lambda\mathbf{x}\|_K = |\lambda|\,\|\mathbf{x}\|_K \qquad \|\mathbf{x}+\mathbf{y}\|_K \leq \|\mathbf{x}\|_K + \|\mathbf{y}\|_K\,. \qquad \text{(* 11.10)}$$

If K is not centrally symmetric, but still contains the origin in its strict interior, then K does not define a norm anymore, but the conditions (* 11.9) and (* 11.10) still hold. We will use σ_K instead of $\|.\|_K$ for the function defined in (* 11.8) in this case. This is the *Minkowski functional* or *gauge functional* defined by K.

[10] Andersen, Louveaux, Weismantel, and L. A. Wolsey, "Inequalities from two rows of a simplex tableau".

σ is also monotonic *w.r.t.* the set is defined for. If $K' \subseteq K$ are convex sets and K' contains the origin in its interior, then

$$\sigma_K(\mathbf{y}) \;\leq\; \sigma_{K'}(\mathbf{y})\,. \tag{* 11.11}$$

Problem * 11.1

You will give a proof for this in Problem * 11.1.

With this we can introduce a new inequality for the corner polyhedron.

Theorem * 11.2. *Let $\overline{K}$ be a closed convex set that contains $\mathbf{f}$ in its interior, but no lattice point in its interior, and let $K := \overline{K} - \mathbf{f}$. Then*

$$\sum_{i=1}^{k} \sigma_K(\mathbf{r}_i) s_i \;\geq\; 1$$

is a valid inequality for $R_{\mathbf{f}}(\mathbf{r}_1, \ldots, \mathbf{r}_k)$.

Proof. Let $\mathbf{s}$ satisfy (* 11.7) and $\overline{\mathbf{y}} := \mathbf{f} + \sum \mathbf{r}_i s_i$. Then $\overline{\mathbf{y}}$ is not in the interior of $\overline{K}$, as $\overline{\mathbf{y}}$ is integral. Sp $\overline{\mathbf{y}} - \mathbf{f}$ is not in the interior of K. So

$$\sigma_K(\overline{\mathbf{y}} - \mathbf{f}) \;\geq\; 1$$

by (* 11.9). This implies

$$\sum \sigma_K(\mathbf{r}_i) s_i \;\geq\; \sum \sigma_K(\mathbf{r}_i s_i) \;\geq\; \sigma_K\left(\sum \mathbf{r}_i s_i\right) \;\geq\; \sigma_K(\overline{\mathbf{y}} - \mathbf{f}) \;\geq\; 1\,.$$

where we have used the inequalities of (* 11.10). □

Now observe that, if for two inequalities of the form

$$\sum_{i=1}^{k} \lambda_i s_i \;\geq\; 1 \qquad\qquad \sum_{i=1}^{k} \lambda'_i s_i \;\geq\; 1$$

for $\mathbf{s} \in \mathbb{R}^k_{\geq 0}$ we have $\lambda_i \leq \lambda'_i$, then any solution $\mathbf{s}$ to the first also satisfies the second inequality. Hence, the second inequality is redundant. Now recall, that for Minkowski functionals we have (* 11.11). Hence, if, for the inequalities derived from convex sets K, K' with 0 in its relative interior as in Theorem * 11.2 we have $K' \subseteq K$, then the inequality from K' is redundant.

It follows that it suffices to look at those sets in Theorem * 11.2 that have $\mathbf{f}$ in its interior, no lattice point in its interior, and which are not properly contained in another set with the same properties. The following definition formalizes this.

Definition * 11.3. Let $K \subseteq \mathbb{R}^d$ be a convex set. Then K is *lattice free* if K contains no lattice point in its interior, K is *maximal lattice free* if it is not contained in another lattice free set.

Note that any affine $L \subseteq \mathbb{R}^d$ with $\dim L < d$ is lattice free, as $\operatorname{int} L = \varnothing$. It is, however, not necessarily maximally lattice free, as we will see in Lemma * 11.8.

With this definition, we obtain the strongest new inequalities from Theorem * 11.2 if $\overline{K}$ is a maximally lattice free convex set. Furthermore, it can be shown that all inequalities for the corner polyhedron are dominated by a cut from a lattice free set.[11] Hence, a characterization of maximal lattice free sets is sufficient to completely describe the corner polyhedron.

This suggests that we should understand these lattice free sets to apply the cuts efficiently. We will see in the next section that maximal lattice free sets are in fact polyhedra. This result by Lovász is also interesting in the theory of lattice polytopes and geometry of numbers. Averkov et al.[12] have shown that there is only a finite number of such polyhedra in each dimension (up to lattice isomorphism). Hence, we may also aim for a full classification of such, which is known so far only up to dimension 3.

* 11.2. Maximal lattice free sets

To use Theorem * 11.2 for new cuts, which, by the remarks at the end of the last section, would suffice to obtain all possible cuts, we need a classification of maximal lattice free convex sets. The following characterization of such sets is due to Lovasz.[13].

Theorem * 11.4. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. A set S is maximally lattice free if and only if one of the following two conditions is satisfied.*

(i) S is a polyhedron and can be written as

$$S = P + L$$

for a polytope P and a Λ-rational linear subspace L with

$$\dim S = \dim P + \dim L = d$$

such that S contains no point of Λ in its interior, but each facet of S contains a point of Λ in its relative interior.

(ii) S is an affine subspace of $\mathbb{R}^d$ of the form $S = v + L$ for a linear subspace L that is not *Λ-rational. In particular, $\dim S < d$.*

We follow ideas of Basu et al.[14] using Theorem 3.16 to find lattice points close to hyperplanes. A different proof by Averkov[15] uses Minkowski's First Theorem (Corollary 3.3) for this. We need some preparations before we can give the proof.

[11] Thm. 4.1 Basu, Conforti, and Summa, "A geometric approach to cut-generating functions".
[12] Averkov, Wagner, and Weismantel, "Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three".
[13] Lovász, "Geometry of numbers and integer programming".
[14] Basu, Conforti, Cornuéjols, and Zambelli, "Maximal lattice-free convex sets in linear subspaces".
[15] Averkov, *A proof of Lovász's theorem on maximal lattice-free sets*.

Lemma * 11.5. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice, $L \subseteq \mathbb{R}^d$ a linear subspace, and $\varepsilon > 0$. If L is not Λ-rational, then there is $y \in \Lambda \setminus L$ with $\mathrm{d}(\mathbf{y}, L) < \varepsilon$.*

Proof. We use induction over $k := \dim L$. Let $k = 1$ and $L := \operatorname{lin} \mathbf{r}$ for some $\mathbf{r} \in \mathbb{R}^d$. We know from Problem * 11.2 that we can find $\mathbf{v} \in \Lambda$ with distance at most ε from $\{ \lambda \mathbf{r} \,:\, \lambda \geq 0 \}$. If $\mathbf{v} \in L$, then L would be Λ-rational, so $\mathbf{v} \in \Lambda \setminus L$.

So consider $k \geq 2$. We assume first, that there is $\mathbf{u} \in L \cap \Lambda$. Let $\pi : \mathbb{R}^d \to U$ be the projection onto the orthogonal complement U of $\mathbf{u}$ and define

$$L' \;:=\; \pi(L) \qquad\qquad \Lambda' \;:=\; \pi(\Lambda)\,.$$

Then Λ' is a lattice in U. If L' is a Λ'-rational subspace of U, then there are

$$\mathbf{b}'_1, \ldots, \mathbf{b}'_k \;\in\; \Lambda'$$

that span L' (linearly). By construction, there are preimages $\mathbf{b}_i \in \pi^{-1}(\mathbf{b}'_i) \cap \Lambda$ for $1 \leq i \leq k$. But then, $r, \mathbf{b}_1, \ldots, \mathbf{b}_k$ would be a basis of L, so L is Λ-rational. This is a contradiction, so L' is not Λ'-rational.

By induction we can find $\mathbf{v}' \in \Lambda'$ at distance at most ε from L'. By construction, there is $\mathbf{v} \in \pi^{-1}(\mathbf{v}') \cap \Lambda$, and this point has distance at most ε from L. Because $\mathbf{v}' \notin L'$ also $\mathbf{v} \notin L$, so $\mathbf{v} \in \Lambda \setminus L$, as required.

If now $L \cap \Lambda = \{ 0 \}$, then we use Problem * 11.2 again to obtain $\mathbf{v} \in \Lambda \setminus \{ 0 \}$ at distance at most ε from L. Then also $\mathbf{v} \notin L$, so $\mathbf{v} \in \Lambda \setminus L$. □

Problem * 11.2

Proposition * 11.6. *Let $S \subseteq \mathbb{R}^d$ be a d-dimensional maximally lattice free set. Then*

$$\operatorname{rec} S \;=\; \operatorname{lineal} S\,.$$

Proof. Assume not, and let $C := \operatorname{rec} S$ be the recession cone. We will show that also

$$S - C \;:=\; \{ \mathbf{s} - \mathbf{c} \,:\, \mathbf{s} \in S \text{ and } \mathbf{c} \in C \}$$

is lattice free. As this contains S and strictly contains S if $C \neq \operatorname{lineal} S$ this would contradict maximality.

So let $\mathbf{u} \in \operatorname{int}(S - C) \cap \Lambda$. Then $\mathbf{u} \in \operatorname{int}(S) - C$ by Problem * 11.3. Let $\mathbf{r} \in C$. Then there is $\lambda \in \mathbb{R}_{\geq 0}$ such that $\mathbf{v} := \mathbf{u} + \lambda \mathbf{r} \in \operatorname{int} S$. As $\mathbf{u} \in \Lambda$ and $\operatorname{int} S \cap \Lambda = \emptyset$ we must have $\lambda > 0$.

Choose $\varepsilon > 0$ such that $\mathcal{B}_{\mathbf{v}}(\varepsilon) \subseteq S$. As $\mathbf{r} \in \mathbb{C} = \operatorname{rec} S$ we know that

$$\mathcal{B}_{\mathbf{v}}(\varepsilon) \;+\; \{ \mu \mathbf{r} \,:\, \mu \geq 0 \} \;\subseteq\; S\,. \tag{* 11.12}$$

We know from Problem * 11.2 that there is a lattice point $\mathbf{w}$ at distance at most ε from the affine ray $\{ \mathbf{u} + \mu \mathbf{r} \,:\, \mu \geq \lambda \}$. But then $\mathbf{w} \in \mathcal{B}_{\mathbf{v}}(\varepsilon) \;+\; \{ \mu \mathbf{r} \,:\, \mu \geq 0 \}$. By (* 11.12) we have $\mathbf{w} \in \operatorname{int} S$. This is a contradiction, so $S - C$ is lattice free. □

Proposition * 11.7. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice and S be a bounded maximally lattice free convex set with $\dim S = d$.*

Then S is a polytope and each facet of S contains a lattice point in its relative interior.

Proof. S is bounded, so we can find a cube $C := \left\{ \mathbf{x} \in \mathbb{R}^d \ : \ a_i \leq x_i \leq b_i \text{ for } 1 \leq i \leq d \right\}$ for some $a_i, b_i \in \mathbb{R}$ such that $S \subseteq C$.

Let $\mathbf{u} \in C \cap \Lambda$. As S is lattice free we can find a hyperplane

$$H_{\mathbf{u}} \ := \ \{ \mathbf{x} \ : \ \langle \mathbf{a_u}, \mathbf{x} \rangle = \delta_{\mathbf{u}} \}$$

for some $\mathbf{a_u} \in \mathbb{R}^{d\star}$ and $\delta_{\mathbf{u}} \in \mathbb{R}$ such that

$$S \ \subseteq \ H_{\mathbf{u}}^{\leq} \qquad \text{and} \qquad \langle \mathbf{a_u}, \mathbf{u} \rangle \ \geq \ 0 \,.$$

The set $C \cap \Lambda$ is finite, as C is bounded, so

$$P \ := \ \bigcap_{\mathbf{u} \in C \cap \Lambda} H_{\mathbf{u}}^{\leq}$$

is a polytope and contains S. It is also lattice free, so $S = P$ as S is maximal lattice free by assumption.

It remains to show that P contains a lattice point in the relative interior of each facet. However, if there is some facet F without a lattice point in the relative interior, then all lattice points on F lie on at least one other facet, and we can push out F slightly without picking up interior lattice points. This contradicts maximality, so F must contain a lattice point in its relative interior. □

Lemma * 11.8. *Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice, $L \subseteq \mathbb{R}^d$ be a linear subspace with $\dim L = d-1$ and $\mathbf{y} \in \mathbb{R}^d \setminus \Lambda$.*

Then $A := \mathbf{y} + L$ is maximal lattice free if and only if L is not a Λ-rational subspace.

Proof. Assume that L is Λ-rational, choose a lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}$ of $\Lambda \cap L$ and pick any $\mathbf{b}_d \in \Lambda$ such that $\mathbf{b}_1, \ldots, \mathbf{b}_d$ is a lattice basis of Λ.

Then $A = \left\{ \mu \mathbf{b}_d + \sum_{i=1}^{d-1} \lambda_i \mathbf{b}_i \ : \ \lambda_i \in \mathbb{R} \right\}$ for some $\mu \in \mathbb{R}$. But then

$$R \ := \ \left\{ \eta \mathbf{b}_d + \sum_{i=1}^{d-1} \lambda_i \mathbf{b}_i \ : \ \lambda_i \in \mathbb{R} \ : \ \lceil \mu - 1 \rceil \leq \eta \leq \lceil \mu \rceil \right\}$$

is lattice free and strictly contains S, so S is not maximal lattice free.

Conversely, assume that L is not Λ-rational. $\dim L < d$, so $\operatorname{int} L = \varnothing$. In particular $\operatorname{int}(A) \cap \Lambda = \varnothing$, so the space is lattice free. We need to show that it is maximal with this property.

Assume that there is maximal lattice free $S \subseteq \mathbb{R}^d$ with $\operatorname{int}(S) \cap \Lambda = \varnothing$ and $A \subsetneq S$. Let $\mathbf{u} \in S \setminus A$. As S is closed and convex we know that

$$\operatorname{conv}(\mathbf{y}, \mathbf{u}) \; + \; L \; \subseteq \; S \, .$$

Let $\delta := \mathrm{d}(\mathbf{y} + L, \mathbf{u} + L)$. By Lemma * 11.5 we can find $\mathbf{v} \in \Lambda$ at distance $\alpha < \delta$ from L. By taking $-\mathbf{v}$ if necessary we may assume that $\mathbf{v}$ is between L and $\mathbf{y} + L$. Let $\beta := \min(\mathrm{d}(0, \mathbf{y} + L), \mathrm{d}(0, \mathbf{u} + L))$ and set

$$\mathbf{w} \; := \; \left(\left\lfloor \frac{\beta}{\alpha} \right\rfloor + 1 \right) \mathbf{v} \, .$$

Then $\mathbf{w} \in \Lambda$ and strictly between $\mathbf{y} + L$ and $\mathbf{u} + L$, so $\mathbf{w} \in \operatorname{int}(K) \cap \Lambda$. This is a contradiction. □

We can now prove Theorem * 11.4.

*Proof of Theorem * 11.4.* Let us first show that in both cases S is maximal lattice free.

If S satisfies (ii) then S is a maximal lattice free set by the previous Lemma * 11.8.

If S satisfies (i), then by assumption S is lattice free, so we only have to show that S is maximal with this property.

Assume not, and let R be a maximal lattice free set strictly containing S. For any $\mathbf{y} \in R \setminus S$ there is a facet F of S such that the hyperplane $\operatorname{aff} F$ separates $\mathbf{y}$ from S. Further, by convexity, $\operatorname{conv}(S \cup \{ \mathbf{y} \}) \subseteq K$, and as $\dim S = d$ the relative interior of F is contained in the interior of K. But the relative interior of F contains a lattice point $\mathbf{u}$ by assumption, so $\mathbf{u} \operatorname{int} K$. This is a contradiction.

For the converse direction assume that S is maximal lattice free. If $\dim S < d$, then S is contained in some affine hyperplane H of $\mathbb{R}^d$. By maximality of S we know that $H = S$. So there is a linear space L of dimension $d-1$ and $\mathbf{u} \in \mathbb{R}^d$ such that $S = \mathbf{u} + L$. It follows from the previous Lemma * 11.8 that S satisfies (ii).

Now assume that $\dim S = d$. If S is bounded, then the claim follows from Proposition * 11.7. Hence, in the following we assume that S is unbounded. We know from Proposition * 11.6 that

$$L \; := \; \operatorname{rec} S \; = \; \operatorname{lineal} S \, .$$

We want to show that L is Λ-rational. Assume it is not. We pick any $\mathbf{y} \in \operatorname{int} S$ and some $\varepsilon > 0$ such that $\mathcal{B}_{\mathbf{y}}(\varepsilon) \subseteq S$. Then also

$$T \; := \; \mathcal{B}_{\mathbf{y}}(\varepsilon) + L \subseteq S \, .$$

Let L' be the maximal Λ-rational subspace of L and pick any $\mathbf{u} \in L \setminus L'$. By Problem * 11.2 we find $\mathbf{v} \in \Lambda$ at distance at most ε from the line spanned by $\mathbf{u}$. But then either $\mathbf{v} \in T$ or $-\mathbf{v} \in T$. Hence, $\operatorname{int} S \cap \Lambda \neq \varnothing$, which contradicts the assumption. So L is Λ-rational.

Hence, we may project onto the orthogonal complement of L. Let $\pi : \mathbb{R}^d \to L^\perp$ be

the projection and define

$$R := \pi(S) \qquad \text{and} \qquad \Lambda' := \pi(\Lambda)\,.$$

As L is Λ-rational the set Λ' is a lattice. Further,

$$S = R + L$$

is bounded,

$$\dim S = \dim R + \dim L \qquad \text{and} \qquad \dim R = d - \dim L\,.$$

and

$$\operatorname{int} S = \operatorname{relint} R + L\,,$$

so $\operatorname{relint} R = \varnothing$. If R is not maximal lattice free in the space $L^\perp$ with the lattice Λ', then R is strictly contained in a lattice free set R'. But then $S \subsetneq R' + L$ and the latter is lattice free in the lattice Λ. This contradicts maximality of S.

Now we use Proposition * 11.7 for R in the space $L^\perp$ with the lattice Λ' and obtain that P is a polytope with a lattice point of Λ' in the relative interior of each of its facets. We can lift the facets of R to facets of S and the lattice points in Λ' to lattice points in Λ. Hence, S satisfies (i). □

This theorem shows that the sets we need to generate cuts for the corner polyhedron are polyhedra with a simple structure. The situation is even nicer, as we will briefly summarize in the next section.

* 11.3. Convergence

In this section we want to sketch some further results on cuts from lattice free polytopes.

At the beginning of this chapter we have already discussed Chvátal cuts and elementary closures of rational polyhedra P. The elementary closure will in most cases differ from the integer hull $P_I := \mathsf{conv}(P \cap Z^d)$. However, it follows from results of Schrijver and Chvátal, that a finite number of iterations of this process produces P_I. The same is true for the split cuts also introduced above. From the point of view of linear programming, these results apply to the pure integer case of a linear program. In this setting, the results state that we arrive at the integer solution of the program after a *finite* number of cuts for a linear relaxation.

However, this is not true anymore in the mixed integer case. Cook et al.[16] showed that in this case a finite number of iterations of taking split closures may not suffice. Although it is still true that the split closure of a polyhedron P is strictly contained in P unless P is its own mixed integer hull, it may happen that the difference between successive steps in this process becomes arbitrarily small.

[16] Cook, R. Kannan, and A. Schrijver, "Chvátal closures for mixed integer programming problems".

It is a result of Del Pia and Weismantel[17] that in the mixed integer case successively applying cuts derived from maximal lattice free sets produces the mixed integer hull in a finite number of steps. For this, it even suffices to consider only those maximal lattice free sets, which are polyhedra by the result of the last section, that are also integral, *i.e.* whose minimal faces contain a lattice point.

Hence, to apply such cuts efficiently in applications we would like to understand the set of integral maximal lattice free polyhedra. The set of such polytopes is contained in the set $\mathcal{M}$ of all integral lattice free polyhedra, that are maximal within the class of integral polyhedra (whereas in the previous section we have taken the maximum over all rational polyhedra). These sets differ in dimensions $d \geq 4$.[18]

However, Averkov et al.[19] have shown that the set $\mathcal{M}$ is finite up to affine lattice transformations. Hence, also the set of lattice free polyhedra needed for the finite convergence of mixed integer programs is finite.

This result allows us to classify such polyhedra for small dimension d, and such a classification allows to obtain cuts for mixed integer programs for up to d rows of the simplex tableaux. The classification is known up to $d = 3$.[19] In dimension 1, there is, up to affine lattice transformations, just one such polytope, the segment $[0, 1]$ (note that any other interval $[, a + 1]$ for $a \in \mathbb{Z}$ is obtained by a lattice transformation from this interval). In dimension 2 there is also just one such polytope, the simplex

$$\Delta_2 \ := \ \mathsf{conv}(0, 2\mathbf{e}_1, 2\mathbf{e}_2) \, .$$

In dimension 3 there are 12 maximal lattice free integral polytopes. Those are the three simplices

$$\mathsf{conv}(0, a\mathbf{e}_1, b\mathbf{e}_2, c\mathbf{e}_3)$$

for

$$(a, b, c) \ \in \ \{ \, (2, 3, 6), (2, 4, 4), (3, 3, 3) \, \} \ ,$$

the four simplices

$$\mathsf{conv}(0, a\mathbf{e}_1, b\mathbf{e}_1 + c\mathbf{e}_2, d\mathbf{e}_1 + e\mathbf{e}_3)$$

for

$$(a, b, c, d, e) \ \in \ \{ \, (1, 2, 4, 3, 4), (1, 2, 5, 3, 5), (3, 1, 3, 2, 3), (4, 1, 2, 2, 4) \, \} \ ,$$

two pyramids

$$\begin{aligned} &\mathsf{conv} \, (\pm 2\mathbf{e}_1, \pm 2\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3) \\ &\mathsf{conv} \, (-\mathbf{e}_1, -\mathbf{e}_2, 2\mathbf{e}_1, 2\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + 3\mathbf{e}_3) \ , \end{aligned}$$

[17] Del Pia and Weismantel, "On convergence in mixed integer programming".
[18] Nill and Ziegler, "Projecting lattice polytopes without interior lattice points".
[19] Averkov, Wagner, and Weismantel, "Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three".

two prisms

$$\begin{aligned}&\mathsf{conv}\left(\mathbf{e}_1+\varepsilon\mathbf{u}, \mathbf{e}_2+\varepsilon\mathbf{u}, -\mathbf{e}_1-\mathbf{e}_2+\varepsilon\mathbf{u} \;:\; \varepsilon\in\{0,1\}, \mathbf{u}:=\mathbf{e}_1+2\mathbf{e}_2+3\mathbf{e}_3\right)\\ &\mathsf{conv}\left(\pm\mathbf{e}_1+\varepsilon\mathbf{u}, 2\mathbf{e}_2+\varepsilon\mathbf{u} \;:\; \varepsilon\in\{0,1\}, \mathbf{u}:=\mathbf{e}_1+2\mathbf{e}_3\right),\end{aligned}$$

and one parallelepiped

$$\mathsf{conv}\left(\sum_{i=1}^{3}\varepsilon_i\mathbf{u}_i \;:\; \varepsilon_i\in\{0,1\}, \mathbf{u}_1:=-\mathbf{e}_1+\mathbf{e}_2, \mathbf{u}_2:=\mathbf{e}_1+\mathbf{e}_2, \mathbf{u}_3:=\mathbf{u}_2+2\mathbf{e}_3\right).$$

The classification for $d\geq 4$ is still open. The proof of finiteness is based on the classical finiteness result about the volume of lattice polytopes with a fixed nonzero number of lattice points in the interior, which is due to Hensley[20] and Lagarias and Ziegler[21].

Theorem * 11.9. *Let $d, s, k\in\mathbb{Z}_{>0}$, $\Lambda:=s\mathbb{Z}^d$ and $\mathcal{P}_k$ the family of d-dimensional integer polytopes P (so with vertices in $\mathbb{Z}^d$) such that*

$$1 \;\leq\; |\operatorname{int} P\cap\Lambda| \;\leq\; k\,.$$

Then there is $V=V(d,s,k)>0$ such that

$$\operatorname{vol} P \;\leq\; V \quad \textit{for all} \quad P \textit{ in } \mathcal{P}_k\,.$$

As a consequence of this theorem, Lagarias and Ziegler[21] show that each equivalence class *w.r.t.* affine lattice transformations of an integer polytope that satisfies the conditions of the previous theorem has a representative that is contained in a cube with side lengths at most $d\cdot d!V$.

Theorem * 11.10. *Le $\mathcal{P}_k$ be as in the previous theorem. Then any $P\in\mathcal{P}_k$ is, up to an affine lattice isomorphism, contained in a cube with side lengths $d\cdot d!V$.*

In particular, the family $\mathcal{P}_k$ is finite, up to affine lattice isomorphisms.

The proof of this theorem is pretty similar to the proof of the Flatness Theorem (Theorem 3.29), but with simplices instead of ellipsoids. They first show that simplices in $\mathcal{P}_k$ are contained in a cube of sidelength $d!\cdot V$ for a suitable lattice transformation, and then show that any other polytope in $P\in\mathcal{P}_k$ contains a simplex S such that

$$\mathbf{a}\;+\;S\;\subseteq\;P\;\subseteq\;\mathbf{a}\;+\;d\cdot S\,.$$

The proof for simplices essentially follows by moving one vertex into the origin, considering the primitive edge directions of edges at this vertex as a lattice basis of a sublattice of $\mathbb{Z}^d$ and applying the Hermite normal form to obtain a normalized version of the

[20] Thm. 3.6 Hensley, "Lattice vertex polytopes with interior lattice points".
[21] Thm. 1 Jeffrey C. Lagarias and Ziegler, "Bounds for lattice polytopes containing a fixed number of interior points in a sublattice".

simplex. The bound on the side lengths of an enclosing cube follows by considering the entries of the Hermite normal form. Finally, the finiteness of the family $\mathcal{P}_k$ up to lattice isomorphisms follows, as a cube of fixed side length contains a finite number of lattice points.

Obtaining the finiteness of the family of lattice free integer polyhedra (which may be unbounded) up to lattice isomorphism from this requires some more steps. As a first simple observation we can reduce to polytopes, as for each polyhedron in this family the recession cone coincides with the lineality space, so it is the product of a polytope with a linear space.

In a second step one considers the *lattice diameter* of the polytope P, *i.e.* the largest number of lattice points on a line segment contained in P, and shows that this diameter must be bounded. This in turn shows that $|P \cap \Lambda|$, *i.e.* the number of lattice points on the boundary of P (as there is none in the interior by assumption), is bounded. Using the above results of Hensley and Lagarias and Ziegler now implies a bound on the volume, and subsequently the finiteness of the family of polyhedra. For the full proof see Section 3 in the paper of Averkov et al.[22]

There are many more results on cuts from lattice free sets and the classification of various subsets of such and variations thereof. A survey on cuts and cut generating functions is in the paper of Basu et al.[23] The lattice width of lattice free polyhedra is studied by Henk et al.[24] Averkov[25] studies the difference between maximizing integer lattice free polyhedra within the class of integer polyhedra and the class of general convex sets.

* 11.4. Problems

* 11.1. Prove the monotonicity of gauge functionals for convex sets $K' \subseteq K$ as in (* 11.11).

* 11.2. Let Λ be a lattice, $\mathbf{u} \in \Lambda$ and $\mathbf{r} \in \mathbb{R}^d$. We define the affine ray $R := \{\, \mathbf{u} + \lambda \mathbf{r} \;:\; \lambda \geq \lambda_0 \,\}$. Show that for any $\varepsilon > 0$ and $\lambda_0 \in \mathbb{R}$ there is $\mathbf{v} \in \Lambda$ with distance at most ε from R.

Hint: Note that this is easy if $\mathbf{r}$ is rational. For the other case you may want to use Dirichlet's Theorem (Theorem 3.16).

* 11.3. Let K be a d-dimensional convex body and $C := \operatorname{rec} K$. Show that

$$\operatorname{int}(K - C) \subseteq \operatorname{int}(K) - C\,.$$

[22] Averkov, Wagner, and Weismantel, "Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three".
[23] Basu, Conforti, and Summa, "A geometric approach to cut-generating functions".
[24] Henk, Kuhlmann, and Weismantel, *On lattice width of lattice-free polyhedra and height of Hilbert bases*.
[25] Averkov, *Difference between families of weakly and strongly maximal integral lattice-free polytopes*.

A. Convexity

In this chapter we will briefly collect some basic definitions and results from convex geometry.

A.1. Basics

We use $\mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ to denote the integer, rational, and real numbers, and we use $\mathbb{R}^d$ for the d-dimensional Euclidean space equipped with the standard inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle \;=\; \sum_{i=1}^{d} x_i y_i$$

for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. This scalar product can be restricted to one in $\mathbb{Q}^d$ and $\mathbb{Z}^d$. For $\mathbb{X} \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ we use

$$\mathbb{X}_{>0} \;:=\; \{x \in \mathbb{X} \mid x > 0\} \qquad\qquad \mathbb{X}_{\geq 0} \;:=\; \{x \in \mathbb{X} \mid x \geq 0\}$$

and similarly $\mathbb{X}_{<0}$ and $\mathbb{X}_{\leq 0}$.

In the following we restrict to $\mathbb{X} \in \{\mathbb{Q}, \mathbb{R}\}$. The *dual space* $(\mathbb{X}^d)^\star$ of $\mathbb{X}$ is the space of all linear functionals $\varphi : \mathbb{X}^d \to \mathbb{X}$. Given a scalar product we can write any such functional in the form

$$\begin{aligned} \varphi : \mathbb{X}^d &\longrightarrow \mathbb{X} \\ \mathbf{x} &\longmapsto \langle \mathbf{a}, \mathbf{x} \rangle \end{aligned}$$

for some $\mathbf{a} \in \mathbb{X}^d$, ans this gives a bijection between $\mathbb{X}^d$ and $(\mathbb{X}^d)^\star$.

We are mostly concerned with objects that can be defined from a, usually finite, subset $X \subseteq \mathbb{R}^d$. We can study spaces *generated* by such a set. The most commonly studied notion here is the linear span of X. Let $X \subseteq \mathbb{R}^d$. A *linear combination* of X is a sum

$$v \;:=\; \sum_{x \in X} \lambda_x x$$

where $\lambda_x = 0$ for all but finitely $x \in X$. The *linear hull* or *linear span* $\operatorname{lin}(X)$ of X is the set of all linear combinations of elements of X. The set X is a *linear space* if X equals its linear span. A linear combination is an *affine combination* if additionally the sum of the coefficients λ_x is 1. The *affine hull* $\operatorname{aff}(X)$ of X is the set of all affine combinations, and a space is *affine* if it coincides with its affine hull. We set $\operatorname{lin}(X) = \{0\}$ and $\operatorname{aff}(X) := \varnothing$ for $X = \varnothing$.

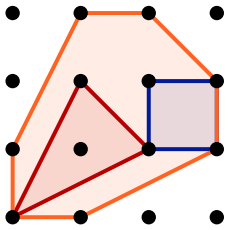


Figure A.1.: The orange polygon is the Minkowski sum of the red and blue polygons

The linear span of X is the smallest linear space containing Xor, equally, the common intersection of all linear spaces containing X. Similarly, the affine hull of X is the smallest affine space containing X or the intersection of all affine spaces containing X. For a matrix $A \in \mathbb{R}^{d\times n}$ with column vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ we also write

$$\operatorname{lin}(A) := \operatorname{lin}(\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}) \qquad \text{and} \qquad \operatorname{aff}(A) := \operatorname{aff}(\{\mathbf{a}_1, \ldots, \mathbf{a}_n\})$$

A set of points X is *linearly* or *affinely independent* if no point of X can be written as a linear or affine combination of the other points.

Linear spaces can always be spanned by a finite subset of X. All minimal such sets, the bases of $\operatorname{lin} X$, have the same size, which is the *dimension* of $\operatorname{lin} X$. The translation of a subset $Y \subseteq \mathbb{R}^d$ by a vector $\mathbf{t} \in \mathbb{R}^d$ is

$$Y - \mathbf{t} := \{y - \mathbf{t} : y \in Y\}$$

For any affine space $A = \operatorname{aff} X$ we can consider its translation by a vector $\mathbf{x} \in A$. This is a linear space. The dimension of A is the dimension of $A - \mathbf{x}$. Hence, any point in the affine hull of X can be written as an affine combination of at most $d+1$ points in X.

The *Minkowski sum* of two subsets $X, Y \subseteq \mathbb{X}^d$ is

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

See Figure A.1 for an example of the sum of two polygons. We similarly define the multiplication with a scalar $\lambda \in \mathbb{X}$ as

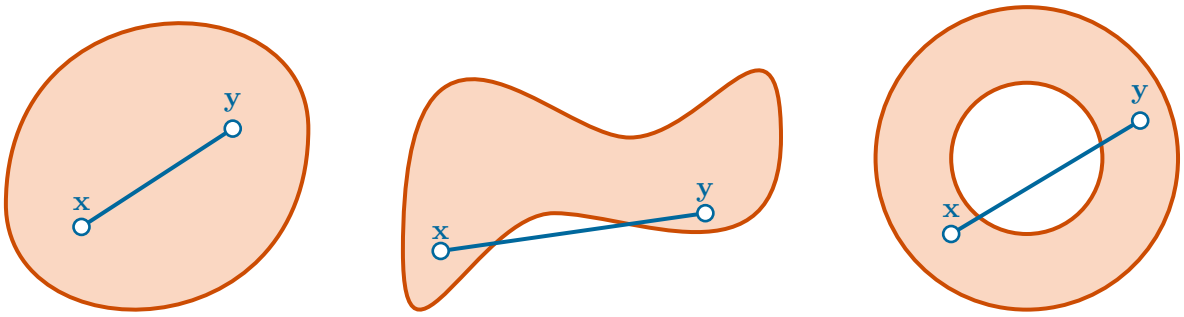
$$\lambda X := \{\lambda\mathbf{x} : \mathbf{x} \in X\}$$

and write $-X$ for $(-1)\cdot X$, $X - Y$ for $X + (-1)\cdot Y$ and $\mathbf{x} + Y$ as shorthand for $\{\mathbf{x}\} + Y$.

A.2. Convex Bodies

A subset $C \subseteq \mathbb{R}^d$ is *convex* if and only if for all $\mathbf{x}, \mathbf{y} \in C$ and all $0 \le \lambda \le 1$ also $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in C$. See Figure A.2 for some examples.

Definition A.1 (Convex Body)**.** A *convex body* is a compact convex set $K \subseteq \mathbb{R}^d$ such that $\operatorname{int} K \neq \varnothing$. It is *centrally symmetric* if for any $\mathbf{x} \in K$ also $-\mathbf{x} \in K$.



(a) A convex set (b) A nonconvex set (c) Another nonconvex set

Figure A.2.: A convex and two nonconvex sets

We denote the set of convex bodies in $\mathbb{R}^d$ by $\mathcal{C}$ and the subset of centrally symmetric convex bodies by $\mathcal{C}_0$.

See Figure A.2(a) for an example. Any d-dimensional centrally symmetric convex body K defines a norm in $\mathbb{R}^d$ via

$$\begin{aligned} \|.\|_K \ : \ \mathbb{R}^d &\longrightarrow \mathbb{R}_{\geq 0} \\ \mathbf{x} &\longmapsto \min_{\lambda \in \mathbb{R}_{\geq 0}} (\mathbf{x} \in \lambda K) \ . \end{aligned}$$

Conversely, any norm $\|.\|$ on $\mathbb{R}^d$ obviously defines a d-dimensional centrally symmetric body via

$$\mathcal{B}_1(0) \ := \ \left\{ \mathbf{x} \in \mathbb{R}^d \ : \ \|\mathbf{x}\| \leq 1 \right\} ,$$

the *unit ball* in this norm. This defines a bijection between norms and d-dimensional centrally symmetric convex bodies.

Definition A.2. For a linear functional given by some $\mathbf{a} \in \mathbb{R}^d$ and $\delta \in \mathbb{R}$ we define the *affine hyperplane*

$$H \ := \ H(\mathbf{a}, \delta) \ := \ \{ \mathbf{x} \ : \ \langle \mathbf{a}, \mathbf{x} \rangle = \delta \}$$

and the two *closed halfspaces*

$$H_{\leq} \ := \ H_{\leq}(\mathbf{a}, \delta) \ := \ \{ \mathbf{x} \ : \ \langle \mathbf{a}, \mathbf{x} \rangle \leq \delta \} \quad H_{\geq} \ := \ H_{\geq}(\mathbf{a}, \delta) \ := \ \{ \mathbf{x} \ : \ \langle \mathbf{a}, \mathbf{x} \rangle \geq \delta \} \ .$$

Their intersection is $H(\mathbf{a}, \delta)$.

We can use hyperplanes to separate disjoint convex bodies.

Theorem A.3. *Let K_1, K_2 be convex with $K_1 \cap K_2 = \varnothing$. Then there is a hyperplane $H := H(\mathbf{a}, \delta)$ such that $K_1 \subseteq H_{\leq}$ and $K_2 \subseteq H_{\geq}$.*

If K_1 is additionally compact, then we can find a strictly *separating hyperplane, i.e. a separating hyperplane H such that $K_1 \subsetneq H_{\leq}$ and $K_1 \cap H = K_2 \cap H = \varnothing$.*

We define the volume of a convex body K via the *Jordan measure*, *i.e.* we define the volume as the *Riemann integral*

$$\mathrm{vol}(K) \;:=\; \int_{\mathbb{R}^d} \chi_K(\mathbf{x}) \mathrm{d}\mathbf{x}$$

for the indicator function χ_K of K (*i.e.* the function that is 1 on K and 0 elsewhere). We say that a set is *Jordan measurable* if this integral exists.

We can compute the volume of K by approximating K with a collection S of small d-dimensional boxes $s_i := [a_1^i, b_1^i] \times \cdots \times [a_1^d, b_1^d]$ for $a_k^i, b_k^i \in R$, and

$$\mathrm{vol}(s_i) \;:=\; \prod_{k=1}^{d} (b_k^i - a_k^i)\,.$$

All sets of dimension $k < d$ thus have volume 0. A particularly useful approximation is given by the following proposition.

Proposition A.4. *Let K be Jordan measurable. Then*

$$\mathrm{vol}(K) \;:=\; \lim_{k\to\infty} \frac{1}{k^d} |K \cap \left(\tfrac{1}{k}\mathbb{Z}\right)^d|\,.$$

Definition A.5. Let $X \subseteq \mathbb{R}^d$ be a convex set. The *polar* or *dual* of X is the set

$$X^\vee \;:=\; \left\{ \mathbf{a} \in (\mathbb{R}^d)^\star \;:\; \langle\, \mathbf{a}, \mathbf{x} \,\rangle \;\leq\; 1 \text{ for all } \mathbf{x} \in X \right\}.$$

The polar body is a closed an convex set that contains the origin. It is bounded if and only if the origin is an interior point of X. The next proposition collects some more properties, whose proof is left as Problem A.1.

Proposition A.6. *Let $X \subseteq \mathbb{R}^d$ be a convex set.*

(i) Let M be an invertible linear transformation on $\mathbb{R}^d$. Then $(M \cdot X)^\vee = M^{-t} X^\vee$.

(ii) $(X^\vee)^\vee = X$ with equality if X is closed and $0 \in X$.

(iii) If X is a centrally symmetric convex body, then so is $X^\vee$.

Problem A.1
Problem A.2
Problem A.3
Problem A.4

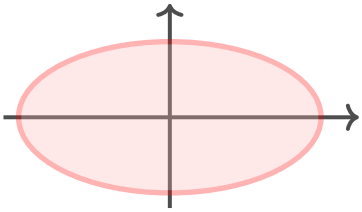


Figure A.3.: The ellipsoid for $T = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{t} = \mathbf{0}$

A.3. Ellipsoids

Let $\mathcal{B}_\mathrm{d} := \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\| = 1\}$ be the unit ball.

Definition A.7 (Ellipsoid)**.** Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be an invertible linear transformation and $\mathbf{t} \in \mathbb{R}^d$. The set

$$E := E(T, \mathbf{t}) := T(B) + \mathbf{t}$$

is the *ellipsoid* with center $\mathbf{t}$.

See Figure A.3 for an example. We can write the ellipsoid explicitly as

$$\begin{aligned} E &= \left\{\mathbf{x} \in \mathbb{R}^d \mid \langle\, T^{-1}(\mathbf{x} - \mathbf{t}), T^{-1}(\mathbf{x} - \mathbf{t})\,\rangle \leq 1\right\} \\ &= \left\{\mathbf{x} \in \mathbb{R}^d \mid \langle\, Q(\mathbf{x} - \mathbf{t}), \mathbf{x} - \mathbf{t}\,\rangle \leq 1\right\} \end{aligned}$$

for the positive semidefinite matrix $Q = (TT^t)^{-1}$. We can also assume that T is positive definite, as any linear transformation T decomposes into a product $T = US$ for an orthogonal matrix U and a positive definite matrix S, and $U(\mathcal{B}_\mathrm{d}) = \mathcal{B}_\mathrm{d}$.

In a basis of eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_d$ with eigenvalues $\lambda_1, \ldots, \lambda_d$ for Q this takes the form

$$E = \left\{\mathbf{x} \in \mathbb{R}^d \mid \lambda_1(x_1 - t_1)^2 + \cdots + \lambda_d(x_d - t_d) \leq 1\right\}.$$

The volume of the ellipsoid is

$$\operatorname{vol} E = |\det T| \operatorname{vol} \mathcal{B}_\mathrm{d} = \frac{\operatorname{vol} \mathcal{B}_\mathrm{d}}{\sqrt{\det Q}}.$$

Theorem A.8. *Let $K \subseteq \mathbb{R}^d$ be a convex body. Then there is a unique ellipsoid $E \subseteq K$ with center c such that*

$$E \subseteq K \subseteq c + d(E - c).$$

Definition A.9. The ellipsoid from the previous theorem is the *maximum volume, John-* or *Löwner-John*-Ellipsoid of K.

We prove this with the following three lemmas. With the first, we show that there is an ellipsoid in a convex body K that attains

$$\eta \;:=\; \sup(\text{vol } E \;:\; E \subseteq K \text{ ellipsoid})\,. \tag{A.1}$$

We then show that $K \subseteq c \;+\; d(E-c)$ for this ellipsoid and finally, that E is unique with this property.

Lemma A.10. *Let $K \subseteq \mathbb{R}^d$ be a convex body. Then the supremum of (A.1) is attained, i.e. there is an ellipsoid $E \subseteq K$ such that* $\text{vol } E = \eta$.

Proof. Let $\mathcal{B}_\mathrm{d}$ be the unit ball. We define the set

$$S \;:=\; \left\{\, (T,\mathbf{a}) \in \mathrm{Gl}(d) \times \mathbb{R}^d \;:\; T(\mathcal{B}_\mathrm{d}) + \mathbf{a} \;\subseteq\; K \,\right\}.$$

Any ellipsoid $E \subseteq K$ is of the form $R = T(\mathcal{B}_\mathrm{d}) + \mathbf{a}$ and

$$\mathrm{vol}(E) \;=\; |\det T| \cdot \mathrm{vol}(\mathcal{B}_\mathrm{d})\,. \tag{A.2}$$

K is compact, so there is $r > 0$ such that $\|\mathbf{x}\| \leq r$ for all $\mathbf{x} \in K$. Hence,

$$\|\mathbf{a}\| \;\leq\; r \qquad \text{and} \qquad \|T\| \;\leq 2r \qquad \text{for all } (T,\mathbf{a}) \;\in\; S\,. \tag{A.3}$$

Hence, S is a closed and bounded subset of $\mathrm{Gl}(d) \times \mathbb{R}^d$, and the map

$$(T,\mathbf{a}) \;\longmapsto\; |\det T|$$

attains its maximum at some $(T_0,\mathbf{a}_0) \in S$. As K is non-empty, we have $|\det T| > 0$ and $E_0 := T_0(\mathcal{B}_\mathrm{d}) + \mathbf{a}_0$ is an ellipsoid of maximum volume. □

We can use the maximum volume ellipsoid of a convex body K to approximate K up to a factor depending on the dimension alone. The inequality of the following lemma is the key result used in the flatness theorem (see Lemma 3.28).

Lemma A.11. *Let $K \subseteq \mathbb{R}^d$ be a convex body and E a maximum volume ellipsoid in K. If the center of E is the origin, then $K \subseteq d \cdot E$.*

Using a translation we can of course always assume the the center of E is the origin.

Proof. By definition there is an invertible linear transformation T such that $E = T(\mathcal{B}_\mathrm{d})$. We can apply T^{-1} to both K and E, so that in the following we can assume that $E = \mathcal{B}_\mathrm{d}$.

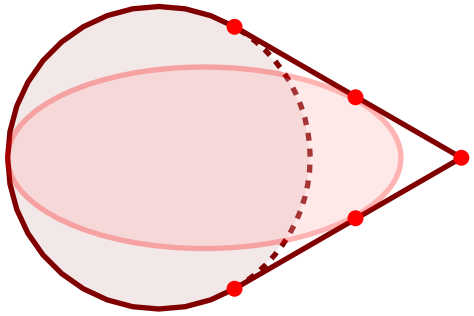


Figure A.4.: A sketch of the setting for the proof of Lemma A.11.

We then need to show that there is no point $\mathbf{z} \in K$ with $\|\mathbf{z}\| \geq d$. Assume on the contrary that there is such a point $\mathbf{z} \in K$ with $\|\mathbf{z}\| > d$ and let

$$L \;:=\; \mathsf{conv}(\mathcal{B}_\mathrm{d} \cup \{\mathbf{z}\}) \;\subseteq\; K\,.$$

We construct an ellipsoid inside L of volume larger than $\mathrm{vol}\ \mathcal{B}_\mathrm{d}$.

Using a linear transformation we can assume that $\mathbf{z} = m\mathbf{e}_1$. For parameters a, b and ε we consider the ellipsoid

$$F_d \;:=\; \left\{ \mathbf{x} \in \mathbb{R}^d \mid \frac{1}{a^2}(x_1 - \varepsilon)^2 + \frac{1}{b^2}\sum_{i=2}^{d} x_i^2 \;\leq\; 1 \right\}.$$

This is symmetric in the last $d-1$ coordinates. Hence, it suffices to consider the case $d = 2$, *i.e.*

$$F \;:=\; \left\{ \mathbf{x} \in \mathbb{R}^d \mid \frac{1}{a^2}(x_1 - \varepsilon)^2 + \frac{x_2^2}{b^2} \;\leq\; 1 \right\}.$$

See Figure A.4 for a sketch of the setting. Now clearly an ellipsoid of maximum volume in F will touch at the point $(-1, 0)$. Plugging this into the defining equation we obtain

$$a \;=\; \varepsilon + 1\,. \tag{A.4}$$

The tangent to F at a point (u, v) is given by the equation

$$\frac{u - \varepsilon}{a^2}(x_1 - \varepsilon) \;+\; \frac{v}{b^2}x_2 \;=\; 1\,. \tag{A.5}$$

and, as (u, v) is on the boundary of F,

$$1 \;=\; \frac{(u-\varepsilon)^2}{a^2} \;+\; \frac{v^2}{b^2} \tag{A.6}$$

Now we want to determine the particular tangent to the ellipsoid that also passes through $m\mathbf{e}_1$ and touches the unit ball, *i.e.* the boundary segment of L added in the convex hull of $\mathcal{B}_\mathrm{d}$ with $m\mathbf{e}_1$. This line touches the unit ball in a point (p, q) and can

thus also be written as

$$px_1 + qx_2 = 1\,. \tag{A.7}$$

It passes through $m\mathbf{e}_1$ and $p^2+q^2=1$, so

$$\frac{u-\varepsilon}{a^2} = \frac{1}{m-\varepsilon} \qquad p = \frac{1}{m} \qquad q = \frac{\sqrt{m^2-1}}{m}\,. \tag{A.8}$$

From (A.7) we deduce that the slope of the tangent is $-1/\sqrt{m^2-1}$. Computing the slope from (A.5) we obtain

$$-\frac{1}{\sqrt{m^2-1}} = -\frac{u-\varepsilon}{a^2}\frac{b^2}{v}\,.$$

Squaring and first using (A.6) and then the first equation of (A.8) we obtain

$$\begin{aligned}\frac{1}{m^2-1} &= \frac{(u-\varepsilon)^2}{a^4}b^2\left(1-\frac{(u-\varepsilon)^2}{a^2}\right)^{-1}\\ &= \frac{1}{(m-\varepsilon)^2}b^2\left(1-\frac{a^2}{(m-\varepsilon)^2}\right)^{-1}\end{aligned}$$

Using (A.4) and solving for b^2 gives

$$b^2 = \frac{(m-\varepsilon)^2 - (1+\varepsilon)^2}{m^2-1}\,.$$

Now let us return to the ellipsoid F_d in dimension d. Its volume is

$$\operatorname{vol} F_d = ab^{d-1}\operatorname{vol}\mathcal{B}_{\mathrm{d}}\,.$$

Now

$$ab^{d-1} = (1+\varepsilon)\left(\frac{(m-\varepsilon)^2-(1+\varepsilon)^2}{m^2-1}\right)^{(d-1)/2}$$

As a function in ε its derivative at 0 is

$$1-\frac{d-1}{2}\frac{2}{m-1} = \frac{m-d}{m-1}\,.$$

Hence, for $m>d$ and small ε the volume of F_d is larger than that of $\mathcal{B}_{\mathrm{d}}$. □

We have shown that an ellipsoid satisfying the inclusions of Theorem A.8 exists. To complete the proof of the theorem we need to show that E is unique.

Lemma A.12. *Let K be a convex body. Then K contains a unique ellipsoid of maximal volume.*

Proof. Any ellipsoid has the form $E = A\,\mathcal{B}_{\mathrm{d}}+t$ for a linear map A and a vector t. We

have seen above that we can assume that A is positive definite. Let P be the set of pairs (A, t) of a positive definite matrix A and a translation t.

We set $S := \{ (A, t) \in P \ : \ A\mathcal{B}_\mathrm{d} + t \subseteq K \}$. As K is convex also S is convex. Assume that we have two ellipsoids of maximal volume, corresponding to the pairs (A_1, t_1) and (A_2, t_2). The volume the ellipsoids is $\det A_1 \operatorname{vol} \mathcal{B}_\mathrm{d} = \det A_2 \operatorname{vol} \mathcal{B}_\mathrm{d}$, so $\det A_1 = \det A_2$, as both are maximal.

Let $A := 1/2(A_1 + A_2)$ and $t := 1/2(t_1 + t_2)$. Then $(A, t) \in S$ by convexity. By Minkowski's determinant inequality (Problem A.5) we obtain

$$\begin{aligned}(\det A)^{1/n} &= \frac{1}{2}(\det A_1 + \det A_2)^{1/n} \\ &\geq \frac{1}{2}\left((\det A_1)^{1/n} + (\det A_2)^{1/n}\right) = (\det A_1)^{1/n}.\end{aligned}$$

As $\det A_1$ is maximal, we must have equality throughout in this chain, and $\det A = \det A_1 = \det A_2$. Minkowski's determinant inequality states that equality implies that $A_1 = A_2$.

So the two ellipsoids are at most translations of each other. Hence, we can find an affine transformation T that maps both ellipsoids to unit ball $\mathcal{B}_{\mathrm{d}1}$ and $\mathcal{B}_{\mathrm{d}2}$ contained in the transformed convex body TK. We can further assume that the centers of the balls are at $-\mu\mathbf{e}_1$ and $\mu\mathbf{e}_1$ for some $\mu \geq 0$. A simple calculation then shows that $\mathsf{conv}(\mathcal{B}_{\mathrm{d}1} \cup \mathcal{B}_{\mathrm{d}2})$ contains the ellipsoid

$$E := \left\{ \mathbf{x} \ : \ \frac{x_1^2}{(\mu+1)^2} + x_2^2 + \cdots + x_d^2 \leq 1 \right\}$$

of volume $(\mu + 1) \operatorname{vol} \mathcal{B}_\mathrm{d}$. As $E \subseteq \mathsf{conv}(\mathcal{B}_{\mathrm{d}1} \cup \mathcal{B}_{\mathrm{d}2}) \subseteq TK$ we get, by maximality of the the volume of $\mathcal{B}_{\mathrm{d}1}$ inside K, that $\mu = 0$. Hence, $t_1 = t_2$ and the claim follows. □

Problem A.5
Problem A.6
Problem A.7

Clearly, uniqueness implies that the ellipsoid E has at least the symmetries of K. Hence, if K is centrally symmetric, then also E is centrally symmetric, and has its center at the origin. We can strengthen the bound in this case.

From the proof of Theorem A.13 it becomes pretty obvious that we can get a larger ellipsoid inside L if K is centrally symmetric and we use the knowledge that also $-\mathbf{z}$ is in K and we can thus replace L by $L := \mathsf{conv}(\mathcal{B}_\mathrm{d} \cup \{\pm\mathbf{z}\})$. With essentially the same proof this leads to the following approximation of a centrally symmetric convex body K by an ellipsoid, which you prove in Problem A.8.

Theorem A.13. *Let $K \subseteq \mathbb{R}^d$ be a* centrally symmetric *convex body and E a maximum volume ellipsoid in K. If the center of E is the origin, then $K \subseteq \sqrt{d} \cdot E$.* □

Problem A.8

A.4. Polyhedra

As before, we consider $\mathbb{X} \in \{ \mathbb{Q}, \mathbb{R} \}$. *Polyhedral cones* are the intersection of a finite set of linear half spaces in $\mathbb{X}$. Generalizing to intersections of affine half spaces leads to

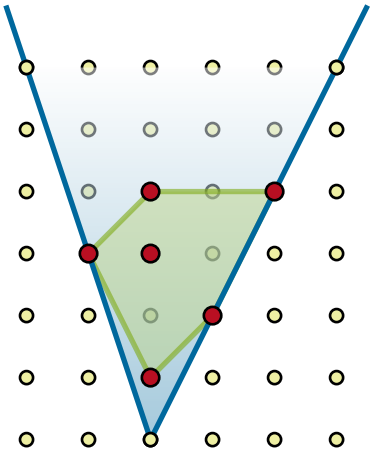


Figure A.5.: Cone (blue) and convex hull (green) of the red points.

polyhedra. We are mainly interested in the subset of bounded polyhedra, the *polytopes*.

Definition A.14. A linear combination is *conic* if all coefficients are nonnegative, and it is *convex* if it is conic and affine.

The set of all conic combinations of a set $X \subseteq \mathbb{Z}$ is the *cone* over X, denoted by $\operatorname{cone}(X)$. The set of all convex combinations of X is the *convex hull* $\mathsf{conv}(X)$. X is a *cone* if $X = \operatorname{cone}(X)$ and X is a *convex set* if $X = \mathsf{conv}(X)$.

The *dimension* of a cone is the dimension of its linear span. The *dimension* of a polytope is the dimension of the affine space it spans.

See Figure A.5 for an example. We sometimes write $\operatorname{cone}(A)$ and $\mathsf{conv}(A)$ for the conic or convex hull of the set of column vectors of a matrix $A \in \mathbb{X}^{d\times n}$. We are mostly interested in cones and convex sets defined by a finite set $X \subseteq \mathbb{X}^d$.

Definition A.15. A *finitely generated cone* C is the cone of a finite subset $X \subseteq \mathbb{X}^d$, and a *polytope* P is the convex hull of finitely many points in $\mathbb{X}^d$.

A cone or polytope is *full dimensional* if its dimension coincides with the dimension d of the space.

Polytopes of dimension 2 are *polygons*.

Definition A.16. Let K be a convex set. A point $\mathbf{x} \in K$ is an *interior point* of K if there is some $\varepsilon > 0$ such that $\mathcal{B}_{\mathbf{x}}(\varepsilon) \subseteq K$. Otherwise $\mathbf{x}$ is a *boundary point*.

$\mathbf{x} \in K$ is a *relative interior point* of K if it is an interior point of K if considered as a subset of $\operatorname{aff} K$. See also Figure A.7.

We need at least $d + 1$ affinely independent points in $\mathbb{R}^d$ to affinely span $\mathbb{R}^d$, so any full-dimensional polytope has at least $d + 1$ points in its defining set. Any polytope

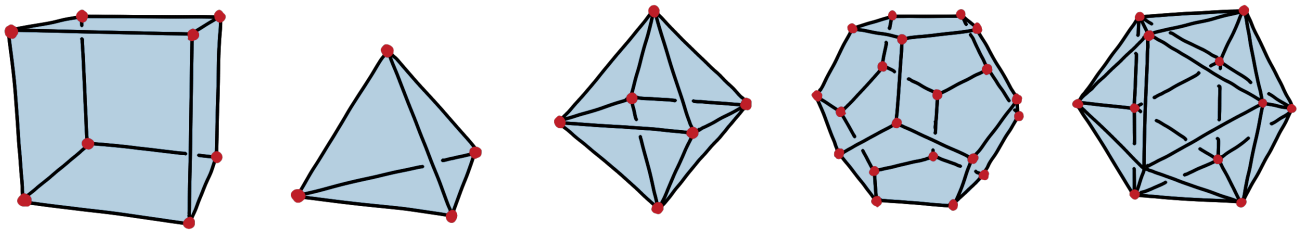


Figure A.6.: Simplex, Cube, Cross Polytope, Dodecahedron, Icosahedron

defined by precisely $d+1$ affinely independent points is called *simplex*. Any two simplices can be identified via a bijective affine map (if you translate both simplices such that one point is in the origin this a just a change of basis).

Clearly, if the dimension of a polytope is less than the dimension of the ambient space, then we can restrict to that affine space. Hence, we may assume that the dimension of our polytopes coincides with the dimension of the space (we will see later that it will be useful to also consider lower dimensional polytopes, though).

In dimension 3 there are the famous *regular* polytopes, which are the *cube*, the *tetrahedron*, the *octahedron*, the *dodecahedron*, and the *icosahedron*, see Figure A.6. Three of them can be generalized to higher dimensions. We have seen the simplex above, which is a tetrahedron in dimension 3. The *unit cube* C_d is the convex hull of the set $X := \{0,1\}^d$. The octahedron can be realized as the special case $d = 3$ of the polytope defined as the convex hull of $\pm\mathbf{e}_i$ for $1 \leq i \leq d$, where $\mathbf{e}_j$ is the j-th unit vector in $\mathbb{R}^d$. In general, those polytopes are called *cross polytopes*.

Let us also look at a slightly more complicated, but highly interesting group of polytopes, the *hypersimplices*. The hypersimplex $h(d,k) \subset \mathbb{R}^d$ for $1 \leq k \leq d-1$ is most easily defined as a polytope of one dimension less than its ambient space. It is the convex hull of all vertices of the unit cube whose coordinates sum up to k:

$$\begin{aligned} h(d,k) &:= \mathsf{conv}\left(\mathbf{x} \in \{0,1\}^d \;:\; \sum_{i=1}^{d} x_i = k\right) \\ &= C_d \cap \left\{\mathbf{x} \in \mathbb{R}^d \;:\; \sum_{i=1}^{d} x_i = k\right\}. \end{aligned}$$

For $k = 1$ and $k = d-1$ we obtain a $(d-1)$-dimensional simplex. You can of course extend the definition to $k = 0$ and $k = d$, but these are just single points in $\mathbb{R}^d$.

Similar to the linear and affine spaces above any point $\mathbf{x}$ in a cone can be written as the conic generation of at most d elements of X, and a point in the convex hull as the convex combination of at most $d+1$ elements of X. Differently from above, however, the choice of these points depends on $\mathbf{x}$. The following theorem makes this precise.

We leave the proof to the reader as Problem A.9.

Theorem A.17 (Carathéodory's Theorem). *Let $X \subseteq \mathbb{R}^d$, $C = \mathrm{cone}(X)$, and $\mathbf{y} \in C$.*

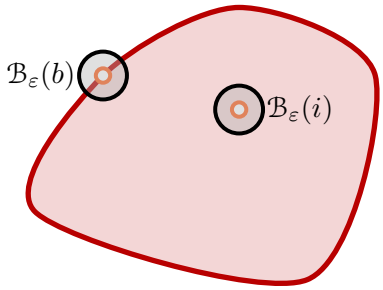


Figure A.7.: An interior point i and a boundary point b.

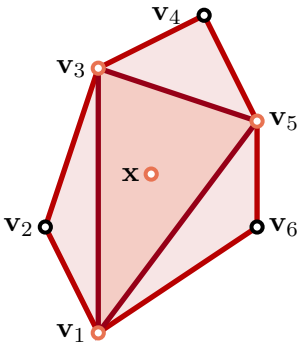


Figure A.8.: $\mathbf{x}$ can be written as a convex combination of $\mathbf{v}_1$, $\mathbf{v}_3$ and $\mathbf{v}_5$.

Then there are $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d \in X$ *and* $\lambda_1, \lambda_2, \ldots, \lambda_d \geq 0$ *such that*

$$\mathbf{y} \;=\; \sum_{i=1}^{d} \lambda_i \mathbf{x}_i \;.$$

Similarly, for $P \;=\; \mathsf{conv}(X)$ *and* $\mathbf{z} \;\in\; P$ *there are* $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_d \;\in\; X$ *and* $\lambda_0, \lambda_1, \ldots, \lambda_d \geq 0$ *such that*

$$\mathbf{z} \;=\; \sum_{i=0}^{d} \lambda_i \mathbf{x}_i \qquad \textit{and} \qquad \sum_{i=0}^{d} \lambda_i \;=\; 1 \;.$$

See also Figure A.8

Problem A.9

Recall the definition of an affine hyperplane from Definition A.2,

$$H \;:=\; H(\mathbf{a}, \delta) \;:=\; \{\, \mathbf{x} \;:\; \langle\, \mathbf{a}, \mathbf{x} \,\rangle = \delta \,\}$$

for some $\mathbf{a} \in \left(\mathbb{X}^d\right)^\star$ and $\delta \in \mathbb{X}$ together with the two associated halfspaces $H_\leq(\mathbf{a}, \delta)$ and $H_\geq(\mathbf{a}, \delta)$ of points $\mathbf{x}$ with $\langle\, \mathbf{a}, \mathbf{x} \,\rangle \leq \delta$ and $\langle\, \mathbf{a}, \mathbf{x} \,\rangle \geq \delta$ respectively. We say that a point $\mathbf{y} \in \mathbb{X}^d$ is *beneath* H if $\langle\, \mathbf{a}, \mathbf{y} \,\rangle < \delta$ and *beyond* H if $\langle\, \mathbf{a}, \mathbf{y} \,\rangle > \delta$. Note that $\lambda\mathbf{a}, \lambda\beta$ for any $\lambda \neq 0$ defines the same hyperplane as $\mathbf{a}, \beta$, and the same affine half space if $\lambda > 0$. Hence, the defining functional for a hyperplane or half space is unique only up to a non-zero and positive factor, respectively.

Definition A.18. A *polyhedron* P is the intersection of finitely many affine half spaces,

$$P \;=\; \bigcap \{\mathbf{x} \mid \langle\, \mathbf{a}_i, \mathbf{x} \,\rangle \leq \beta_i\} \;=\; \{\mathbf{x} \mid \langle\, \mathbf{a}_1, \mathbf{x} \,\rangle \leq \beta_1, \ldots, \langle\, \mathbf{a}_m, \mathbf{x} \,\rangle \leq \beta_m\}$$

for $\mathbf{a}_i \in \mathbb{X}^d$ and $\beta_i \in \mathbb{X}$ and $1 \leq i \leq k$. This is often written in the more concise form

$$P \;=\; \{\, \mathbf{x} \;\mid\; A\mathbf{x} \leq \mathbf{b} \,\}$$

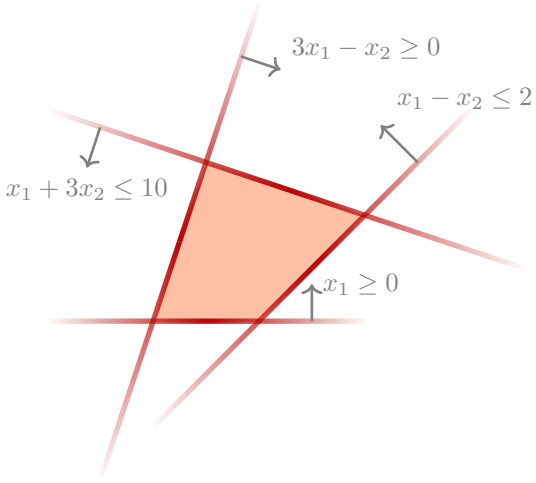


Figure A.9.: The polygon of Example A.19(ii)

where $A \in \mathbb{X}^{k\times d}$ whose rows are the functionals $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ and $\mathbf{b}$ is the vector with entries $\beta_1, \beta_2, \ldots, \beta_k$.

Example A.19. We look at some simple examples.

(i) The *unit cube* C_d is defined by the inequalities

$$\mathbf{x}_i \;\geq\; 0 \qquad\qquad \mathbf{x}_i \;\leq 1 \qquad\qquad \text{for } 1 \leq i \leq d\,.$$

(ii) The inequalities

$$x_1 \;\geq\; 0 \qquad x_1 - x_2 \;\leq\; 2 \qquad x_1 + 3x_2 \;\leq\; 10 \qquad 3x_1 - x_2 \;\geq\; 0$$

define a polygon with vertices

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 4 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

see Figure A.9.

Definition A.20. A polyhedron P is a *(polyhedral) cone* if all defining inequalities are linear, that is,

$$P \;=\; \bigcap H^-_{\mathbf{a}_i,0} \tag{A.9}$$

for some $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k \in (\mathbb{X}^d)^\star$.

We have already defined a *cone over a set* X as the set of all conic combinations in the previous section. We will see below that this and the newly defined notion of a

polyhedral cone coincide if X is a finite set, *i.e.* any polyhedral cone can equally be described as the cone over some suitably chosen finite set X, and any cone over a finite set is polyhedral.

We will not encounter non-polyhedral cones, that is, cones defined as the set of conic combinations over an infinite set X, in this book. Therefore, we will often omit the word polyhedral and just speak of cones in the text, and only stress this restriction in definitions and theorems.

The *dimension* of such a polyhedron defined by half spaces is again defined as the dimension of its affine hull. We sometimes use the notion *d-polytope* for a d-dimensional polyhedron. A polyhedron is *full dimensional* if $\dim P = d$.

Definition A.21. Let $P = \bigcap H^-_{\mathbf{a}_i,\beta_i}$ be a polyhedron. The *recession cone* and *lineality space* of P are

$$\operatorname{rec} P = \bigcap H^-_{\mathbf{a}_i,0} \qquad \text{and} \qquad \operatorname{lineal} P = \bigcap H_{\mathbf{a}_i,0}\,.$$

A polytope is *pointed* if $\operatorname{lineal} P = \varnothing$.

We can associate a cone to each polyhedron $P \subseteq \mathbb{X}^d$ that essentially has the same combinatorial and geometric properties. This is the *homogenization* of P or just the *cone over* P defined by

$$C(P) := \operatorname{cone}(\{1\} \times P) \subseteq \mathbb{X}^{d+1}\,, \tag{A.10}$$

so if $P := \{\mathbf{x} \mid \langle \mathbf{a}_1, \mathbf{x} \rangle \leq \beta_1, \ldots, \langle \mathbf{a}_m, \mathbf{x} \rangle \leq \beta_m\} \subseteq \mathbb{X}^d$ with $\mathbf{a}_i \in (\mathbb{X}^d)^\star$, $\beta_i \in \mathbb{X}$ for $i \in [m]$ then

$$C(P) = \{(x_0, \mathbf{x}) \mid -\beta_1 x_0 + \langle \mathbf{a}_1, \mathbf{x} \rangle \leq 0, \ldots, -\beta_m x_0 + \langle \mathbf{a}_m, \mathbf{x} \rangle \leq 0\}\,.$$

It is often convenient to look at the homogenization of the polyhedron instead of the polyhedron itself as it is defined by linear instead of affine inequalities. We can recover the polyhedron by intersecting the cone with the hyperplane $x_0 \equiv 1$ (and projecting).

A hyperplane $H := H(\mathbf{a}, \delta)$ for some $\mathbf{a} \in (\mathbb{R}^d)^\star$ and $\beta \in \mathbb{R}$ defines a *valid hyperplane* if P is contained in the negative half space $H_\leq(\mathbf{a}, \delta)$. A valid hyperplane is *supporting* if $P \cap H$ is non-empty.

Definition A.22. Let P be a polytope. A face F of P is either P itself or the intersection of P with a valid linear hyperplane. If $F \neq P$ then F is a *proper face*.

Observe that the empty set is also a face of P. For any face F we have

$$F \cap P = \operatorname{aff} F \cap P\,,$$

so faces of polyhedra are again polyhedra and a face of a face of the polyhedron is a face of the polyhedron. The *dimension of a face* of a polyhedron P is its dimension as a

polyhedron,

$$\dim F := \dim \operatorname{aff} F\,.$$

A *k-face* of P is a k-dimensional face of P. If

$$\beta \;:=\; \max\{\langle \mathbf{a}, \mathbf{x}\rangle \mid \mathbf{x} \in P\}$$

is finite then $H(\mathbf{a}, \beta)$ is a supporting hyperplane of P and $P \cap H$ is a face of P, the *face defined by* $\mathbf{a}$. The functionals defining a face are exactly those in the negative dual of the recession cone.

A 0-dimensional face is a *vertex* of P, a 1-dimensional face is an *edge* and a face of dimension $\dim\, P - 1$ is a *facet*. For *full-dimensional* polyhedra the facets have dimension $d-1$. The set of vertices of P is denoted by $\mathcal{V}(P)$.

Problem A.10
Problem A.11
Problem A.12

A set X is *finitely generated* if it can be written as a Minkowski sum of a polytope, a cone, and a linear space, that is, there are $\mathbf{v}_i \in \mathbb{X}^d$, $i = 1, \ldots, r$, $\mathbf{r}_j \in \mathbb{X}^d$, $j = 1, \ldots, s$, $vb_k \in \mathbb{X}^d$, $k = 1, \ldots, t$ such that

$$X \;=\; \left\{ \sum_{i=1}^{r} \lambda_i \mathbf{v}_i + \sum_{j=1}^{s} \mu_i \mathbf{r}_i + \sum_{k=1}^{r} \nu_i \mathbf{b}_i \;:\; \begin{array}{c} \lambda_i, \mu_j, \nu_k \in \mathbb{R}, \\ \lambda_i, \mu_j \geq 0, \sum_{i=1}^{r} \lambda_i = 1 \end{array} \right\}. \tag{A.11}$$

Theorem A.23 (Weyl-Minkowski Theorem)**.** *Let $P \subseteq \mathbb{X}^d$. Then P is a polyhedron if and only if it is finitely generated.*

A proof of this theorem can be found in the book of Schrijver [1] (and in many other books). Projections of polyhedra are again polyhedra, finitely generated by the projections of the generators.

Example A.24. In the notation of Weyl-Minkowski Theorem (Theorem A.23) we can define a polyhedron with

$$\mathbf{v}_1 := \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \mathbf{v}_2 := \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \mathbf{v}_3 := \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \mathbf{r}_1 := \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \mathbf{r}_2 := \begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

see Figure A.10. It is defined by the inequalities

$$3x_1 - x_2 \;\geq\; 3 \qquad x_1 \geq 1 \qquad x_1 + x_2 \;\geq\; 3 \qquad x_1 - 3x_2 \;\leq\; -1\,.$$

The Minkowski sum of a polytope with a cone C or a linear space L is unbounded if C or L have positive dimension. Hence, we can deduce the following duality for polytopes from the Weyl-Minkowski Theorem.

Corollary A.25 (Weyl-Minkowski-Duality)**.** *A bounded set $P \subseteq \mathbb{X}^d$ is a polytope if and only if it is the bounded intersection of a finite number of affine half spaces.* □

[1] Alexander Schrijver, *Theory of linear and integer programming.*

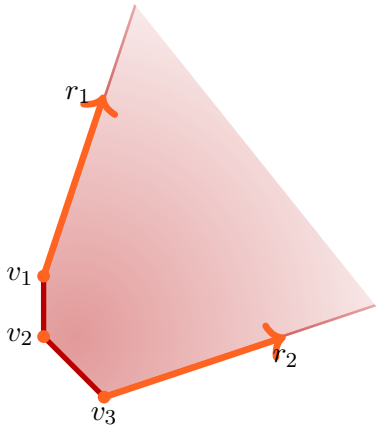


Figure A.10.: The polyhedron of example Example A.24

From this theorem we obtain two equivalent descriptions of a polytope:

(i) as the convex hull of a finite set of points in $\mathbb{X}^d$,

(ii) as the bounded intersection of a finite set of affine half spaces.

The first is called the *interior* or $\mathcal{V}$*-description*, The second is the *exterior* or $\mathcal{H}$*-description*. Both are important in polytope theory, as some things are easy to describe in one and may be difficult to define in the other.

A.5. Integer Hulls

Definition A.26. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice in $\mathbb{R}^d$. A polytope is a *lattice polytope* if all its vertices are in Λ.

In the case $\Lambda = \mathbb{Z}^d$ one often also calls them *integral polytopes*.

Definition A.27. Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice in $\mathbb{R}^d$ and $P \subseteq \mathbb{R}^d$ a polytope. The *integer hull* of P is the convex hull of all lattice points in P,

$$P_I := \mathsf{conv}(P \cap \Lambda)\,.$$

Computing the integer hull is an NP-complete problem (note that integer programming would be in P if we can compute the integer hull in polynomial time). Still, there are many methods to compute the inter hull, *e.g.* via Chvatal closures. See the book of Schrijver [2] for details.

[2] Alexander Schrijver, *Theory of linear and integer programming.*

A.6. Complexity of Polyhedra

In this section we want to collect some results on the complexity of representations of polyhedra. Let $P \in \mathbb{R}^d$ be any rational polyhedron. Recall that P can be written in the form

$$P := \{ \mathbf{x} \mid A\mathbf{x} \leq \mathbf{b} \} = \mathsf{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n) + \mathrm{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_m).$$

for some matrix A and right hand side $\mathbf{b}$, or vertices $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and ray generators $\mathbf{r}_1, \ldots, \mathbf{r}_m$.

Definition A.28. The *facet complexity* of P is the smallest $\varphi > d$ such that there exists a system of linear inequalities $A\mathbf{x} \leq \mathbf{b}$ such that $P = \{ \mathbf{x} : A\mathbf{x} \leq \mathbf{b} \}$ and each inequality in this system has size at most φ.

Definition A.29. The *vertex complexity* of P is the smallest $\nu > d$ such that we can find points $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and rays $\mathbf{r}_1, \ldots, \mathbf{r}_m$ of size at most ν such that

$$P = \mathsf{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n) + \mathrm{cone}(\mathbf{r}_1, \ldots, \mathbf{r}_m).$$

The following proposition bounds the facet complexity in terms of the vertex complexity and vice versa. You have seen a proof in *Discrete Optimization*. You can also find this as Theorem 10.2 in the book of Schrijver[3].

Proposition A.30. *Let P be a rational polyhedron with facet complexity φ and vertex complexity ν. Then*

$$\nu \leq 4d^2\varphi \qquad \text{and} \qquad \varphi \leq 4d^n u.$$

We can also bound the complexity of an integral point in the polyhedron.

Proposition A.31. *Let $P = \{ \mathbf{x} : A\mathbf{x} \leq \mathbf{b} \}$ be a rational polyhedron, and assume that the size of any inequality is bounded by φ.*

If $P_I \neq \varnothing$, then P contains an integral point of size at most $6n^3\varphi$.

A.7. Computing Ellipsoids

Unfortunately, the proof of Theorem A.8 is not constructive, and so far, also no polynomial time algorithm is known that allows to compute E. However, the precise scaling

[3] Alexander Schrijver, *Theory of linear and integer programming.*

factor is not really important in most applications, *e.g.* in Lenstra (Theorem 6.4), as long as it depends polynomially on d. So we may content ourselves with a worse bound, if this allows a polynomial computation of the ellipsoid E. Various authors [4] have shown that this is indeed possible, and we can even find an iterative method, that in the limit converges to the precise E.

We present a method that follows the book of Grötschel, Lovász and, Schrijver[5]. This is essentially the same approach that was proposed for the ellipsoid method of linear programming. There, it was used to solve the feasibility problem by iteratively constructing smaller enclosing ellipsoids. This was repeated until we found one ellipsoid whose center was contained in the polytope. At this point the algorithm stopped and returned the point.

We will need an ellipsoid that satisfies a weakened version of the bounds of Theorem A.8, so we will have to modify the process at this point and continue constructing new ellipsoids until our bounds are satisfied. Depending on assumptions and approximations, there are various similar bounds in the literature. In the following we want to show that, for a given polytope $P \subseteq \mathbb{R}^d$ in dimension d, we can find an ellipsoid E such that

$$E \;\subseteq\; P \;\subseteq\; (d+1)\sqrt{d+1}E\,.$$

As with the ellipsoid method, the initial step is the computation of an ellipsoid that contains the portion of a ball cut out by some affine halfspace intersecting the ball. More precisely, let $\mathcal{B}_\mathrm{d}$ be the unit ball in $\mathbb{R}^d$ (with origin 0). We consider halfspaces of the form $H_{\leq}(\mathbf{a}, \delta)$ as defined in Definition A.2. Using a linear transformation we may assume that $\mathbf{a} = \mathbf{e}_d$, the last unit vector, in the following. Then we look at

$$S \;:=\; B \;\cap\; H_{\leq}(\mathbf{e}_d, \delta)\,.$$

This is certainly only interesting if $|\delta| \leq 1$, but we have to restrict δ further in the theorem below. We now aim for ellipsoids that contains S, and want to find one of minimal volume. For $\delta = 1$ this is the unit ball itself, and it will follow from our consideration that this will still be true for all δ larger than a threshold $t > 0$. It will turn out that we need to make δ smaller than $\frac{1}{d}$ before the minimal volume ellipsoid is not anymore the original ball.

With the next theorem, we make this precise, and we also compute the volume of the minimal ellipsoid. Note that in our setting the problem is essentially 2-dimensional, as the resulting ellipsoid will be symmetric around the $\mathbf{e}_d$-axis. We will only give the main steps in the proof and leave all intermediate computations to the reader.

Theorem A.32. *Let $\mathcal{B}_\mathrm{d} \subseteq \mathbb{R}^d$ be the unit ball in $\mathbb{R}^d$ and $\mathbf{e}_d$ be the last unit vector. For a parameter $-1 \leq \delta \leq \frac{1}{d}$ we consider the set*

$$S \;:=\; B \;\cap\; H_{\leq}(\mathbf{e}_d, \delta)$$

[4] Goffin, "Variable metric relaxation methods. II. The ellipsoid method"; Todd and Yıldırım, "On Khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids".

[5] Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*.

for the half space $H_{\leq}(\mathbf{e}_d, \delta) := \{\, \mathbf{x} : \langle \mathbf{a}, \mathbf{x} \rangle \leq \delta \,\}$. The ellipsoid E of minimal volume containing S is given by

$$E := \left\{ (\mathbf{x}, y) \in \mathbb{R}^{d-1} \times \mathbb{R} : \frac{\|\mathbf{x}\|^2}{a^2} + \frac{(y-c)^2}{b^2} \leq 1 \right\}$$

for

$$a := \sqrt{\frac{d^2}{d^2-1}}\sqrt{1-\delta^2} \quad b := \frac{d}{d+1}(\delta+1) \quad c := 1-b = \frac{1}{d+1} - \frac{d}{d+1}\delta\,.$$

Its volume is

$$\operatorname{vol} E = f(d) \cdot \operatorname{vol} \mathcal{B}_{\mathrm{d}}\,.$$

for a function

$$f(d,\delta) := \left(\frac{d}{d+1}\right)^d \left(\frac{d+1}{d-1}\right)^{d-1/2} \sqrt{1-\delta^2}(\delta+1) < 1$$

only depending on d and δ.

Proof. We only sketch the relevant steps. All computations are left as Problem A.13.

Any ellipsoid E that has the $(d-2)$-sphere $S := \partial\,\mathcal{B}_{\mathrm{d}} \cap H(\mathbf{e}_d, \delta)$ and the point $-\mathbf{e}_d$ in its boundary is of the form

$$E := \left\{ (\mathbf{x}, y) \in \mathbb{R}^{d-1} \times \mathbb{R} : \frac{\|\mathbf{x}\|^2}{a^2} + \frac{(y-c)^2}{b^2} \leq 1 \right\}$$

with

$$a := \frac{\sqrt{1-\delta^2}\cdot b}{\sqrt{b^2-(\delta+1-b)^2}} \qquad\qquad c := b-1\,.$$

Now the volume of E is

$$\operatorname{vol} E = a^{d-1} b \cdot \operatorname{vol} \mathcal{B}_{\mathrm{d}} = \left(\frac{(1-\delta^2)\cdot b^2}{b^2-(\delta+1-b)^2}\right)^{d-1/2} \cdot b \cdot \operatorname{vol} \mathcal{B}_{\mathrm{d}}\,.$$

This takes its minimum for

$$a := \sqrt{\frac{d^2}{d^2-1}}\sqrt{1-\delta^2} \quad b := \frac{d}{d+1}(\delta+1) \quad c := 1-b = \frac{1}{d+1} - \frac{d}{d+1}\delta\,.$$

The corresponding volume is

$$\operatorname{vol} E = f(d,\delta) \cdot \operatorname{vol} \mathcal{B}_{\mathrm{d}}\,.$$

for

$$f(d,\delta) \;:=\; \left(\frac{d^2}{d^2-1}(1-\delta^2)\right)^{d-1/2} \frac{d}{d+1}(\delta+1)\,.$$

The factor is less than 1 and only depends on d and δ.

E may not contain $\mathcal{B}_\mathrm{d}$ (we only required it to contain the intersection with $H(\mathbf{e}_d,\delta)$ and the point $-\mathbf{e}_d$), but it certainly does if $a \geq 1$, which implies $\delta \leq \frac{1}{d}$. This explains the condition on δ in the theorem. □

Observe that $f(d,\frac{1}{d}) = 1$. For this $\delta = \frac{1}{d}$ the ellipsoid containing the sphere S and $-\mathbf{e}_d$ is the original unit ball, and for larger δ the ellipsoid E that we computed will no longer contain $\mathcal{B}_\mathrm{d}$. Hence, for lager δ we would nee to choose $E = \mathcal{B}_\mathrm{d}$, *i.e.* we would not obtain an enclosing ellipsoid of strictly smaller volume.

Problem A.13

We will in the following only use the theorem for

$$\delta \;:=\; d^{-3/2} \;<\; \frac{1}{d}\,,$$

so we define $f(d) := f(d,d^{-3/2})$. Note that this choice of δ is covered by the above theorem, so intersections of the ball with an halfspace whose defining right hand side is at most δ will lead to a decrease in the volume of the enclosing ellipsoid.

The following lemma is easy to prove and left as an exercise.

Lemma A.33. *Let $P := \{\mathbf{x} \;:\; A\mathbf{x} \leq \mathbf{b}\} \subseteq \mathbb{R}^d$ be a polytope contained in the unit ball $\mathcal{B}_\mathrm{d}$. If $b_i \geq \delta = d^{-3/2}$ for all elements b_i of the right hand side $\mathbf{b}$, then*

$$\delta \cdot \mathcal{B}_\mathrm{d} \;=\; d^{-3/2}\,\mathcal{B}_\mathrm{d} \;\subseteq\; P$$

The general idea is now the following. We first find a ball centered at 0 that contains our polytope. By scaling we can assume that the ball is the unit ball $\mathcal{B}_\mathrm{d}$. Then we search for an inequality that violates the condition in Lemma A.33, *i.e.* an inequality where $b_i < \delta$.

If we do not find such an inequality, then we have found our approximation. Otherwise, we apply Theorem A.32 and obtain an ellipsoid E of smaller volume that contains P. We normalize with an affine linear map so that E is again the unit ball with center at the origin, and repeat the process.

Note that affine linear maps preserve containment and relative scaling factors, *i.e.* if for the scaled and translated polytope P' we have

$$\lambda \cdot \mathcal{B}_\mathrm{d} \;\subseteq\; P' \;\subseteq\; \mathcal{B}_\mathrm{d}\,,$$

then if φ is an affine map that sends P' to P, we have

$$\lambda \cdot \varphi(\mathcal{B}_\mathrm{d}) \;\subseteq\; P \;\subseteq\; \varphi(\mathcal{B}_\mathrm{d})\,,$$

and $\varphi(\mathcal{B}_\mathrm{d})$ is an ellipsoid.

Now we need to estimate the size of P. We need to bound its volume from below and above, to obtain a ball that contains P, and to show that the iteration with smaller and smaller ellipsoids stops sufficiently fast.

By Proposition A.30, the facet complexity, which is part of our input, polynomially bounds the vertex complexity. Further, our polytope P is clearly contained in the ball with center at the origin and radius equal to the largest distance to a vertex. This is bounded by 2^ν, so we may choose the ball $\mathcal{B}_{2^\nu}(0)$ as the initial ellipsoid in our process.

Let $\mathbf{x}_0, \ldots, \mathbf{x}_d$ be any affinely independent points in P (which must exist, as P is full dimensional). Then

$$\begin{aligned} \operatorname{vol} P &\geq \operatorname{vol} \mathsf{conv}(\mathbf{x}_0, \ldots, \mathbf{x}_d) \\ &\geq \frac{1}{d!} \det \begin{bmatrix} 1 & \cdots & 1 \\ \mathbf{x}_0 & \cdots & \mathbf{x}_d \end{bmatrix} \\ &\geq \frac{1}{d^d} 2^{-d\nu} \geq 2^{-2d\nu}. \end{aligned}$$

where the lower bound on the determinant follows from the fact that each of $\mathbf{x}_0, \ldots, \mathbf{x}_d$ has size at most ν. Hence the determinant has denominator at most $2^{d\nu}$.

Recall that we only continue as long as we find inequalities whose right hand side (after normalization) is less than δ. We initially start with a ball of volume $2^{d\nu} \operatorname{vol} \mathcal{B}_\mathrm{d}$ and reduce the volume by a factor of $f(d)$ in each iteration.

Hence, if the process would not stop, then after some number N of iterations the volume of the enclosing ellipsoid would be smaller than the volume of the polytope it encloses. This is clearly a contradiction, so the process in fact stops. We can compute an upper bound on the number N of iterations via

$$f(d)^n 2^{d\nu} \operatorname{vol} \mathcal{B}_\mathrm{d} \leq 2^{-2d\nu} \quad \Longleftrightarrow \quad n \geq \frac{1}{|\log_2 f(d)|} \cdot (3d\nu - \log_2 \operatorname{vol} \mathcal{B}_\mathrm{d}).$$

Hence, if we set

$$N := \frac{1}{|\log_2 f(d)|} \cdot (3d\nu - \log_2 \operatorname{vol} \mathcal{B}_\mathrm{d})$$

then our process does at most N iterations before we have found the required ellipsoid. We summarize this in the following theorem.

Theorem A.34. *Let $P := \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$. Then we can find, in time polynomial in the size of A, $\mathbf{b}$ and the dimension d, an ellipsoid E with center $\mathbf{c}$ such that*

$$\mathbf{c} + d^{-3/2}(E - \mathbf{c}) \subseteq P \subseteq E.$$

By scaling E we can equally state that we can find an ellipsoid E, such that

$$E \subseteq P \subseteq \mathbf{c} + d^{3/2}(E - \mathbf{c}). \tag{A.12}$$

Note that our arguments above do not yet comprise a full proof of the theorem, and we also will not attempt to give one here. Besides the missing pieces in the derivation of the statement of the theorem above, there is a more serious issue that one needs to address if one wants to give a rigorous proof. We have seen in Theorem A.32 that the parameters for the new ellipsoid contain square roots. These will usually be not rational, so that we have to round them appropriately to represent them in a computer. But the ellipsoid E' with rounded data need not contain P anymore, even if the exact ellipsoid E does. We need to control the error. For this, we want to show that there is some ε such that

$$P \subseteq E \subseteq \left\{ \mathbf{x} : \exists \mathbf{y} \in E' \text{ with } \|\mathbf{x} - \mathbf{y}\| \leq \varepsilon \right\},$$

i.e. E is contained in E' together with a boundary of size ε around E. One can now show that we can bound the size of the necessary ε in terms of the smallest and largest eigenvalue of the defining matrix Q of E, and that we can bound the eigenvalues of E' in terms of those of E. You can find a proof of this in Section 13.2 of the book of Schrijver.[6] A proof of Theorem A.34 that takes this into account and also computes the precise bounds and the running time can be fond in Section 15.6 of this book.

A.8. Decompositions of Polyhedra

In this section we look at ways to subdivide a polyhedron into smaller pieces, or *cells*. All pieces should be polyhedra themselves, and the subdivision should be a partition, *i.e.* any two pieces intersect at most in their boundary, and the union of all pieces should cover the polyhedron. We will also require that the pieces intersect nicely, which should mean that any intersection is again a polyhedron (the empty polyhedron in many cases).

This will lead us to polyhedral complexes and to two special cases. First we will look at fans, which are polyhedral complexes in which all cells are cones. Secondly, we will consider complexes in which all cells are simplices. If the union of all cells is a polyhedron, polytope, or cone P, then the complex is a *triangulation* of P.

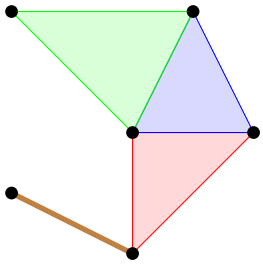
A.8.1. Polyhedral Complexes

Let us start with a definition of a polyhedral complex. For this recall the definition of a *face* F of a polyhedron P, which is the intersection of P with a valid hyperplane H, *i.e.* a hyperplane such that P is completely contained in one of the half spaces defined by H. So

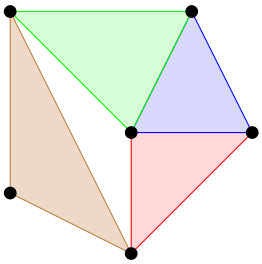
$$F := P \cap H .$$

We see that F is again a polyhedron (which is empty, unless H is a *supporting* hyperplane of P).

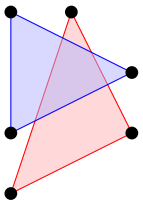
[6] Alexander Schrijver, *Theory of linear and integer programming.*



(a) A polyhedral complex



(b) A *pure* polyhedral complex



(c) Not a polyhedral complex

Figure A.11.: Examples of collections of polyhedra, of which two are polyhedral complexes, and one is not.

Definition A.35. A *polyhedral complex* $\mathcal{C}$ is a finite family of polyhedra (the *cells* of the complex) such that for all $P, Q \in \mathcal{C}$

(i) if $P \in \mathcal{C}$ and F is a face of P then $F \in \mathcal{C}$, and

(ii) $F := P \cap Q$ is a face of both P and Q.

A cell P is *maximal* if there is no $Q \in \mathcal{C}$ strictly containing it.

All cells are polyhedra itself and the *dimension* of a cell is its dimension as a polyhedron. $\mathcal{C}[k]$ is the set of all k-dimensional faces of $\mathcal{C}$.

The *dimension* of $\mathcal{C}$ is the maximal dimension of a cell of the complex. A complex is *pure* if all maximal cells have the same dimension. In this case the maximal cells are the *facets* of the complex.

A polyhedral complex $\mathcal{S}$ is a *subcomplex* of $\mathcal{C}$ if its cells are a subset of the cells of $\mathcal{C}$.

Example A.36. Figure A.11 shows two examples of polyhedral complexes, and one collection of two polyhedra that is not a polyhedral complex. Note that the union (as point sets) of all cells need not be a polyhedron. Figure A.11(b) shows an example.

See Figure A.11(a) for a non-pure polyhedral complex. It has three 2-dimensional maximal cells and one 1-dimensional maximal cell.

Any polytope or cone can be viewed as a polyhedral complex. This complex has one maximal cell, the cone or polytope itself. This is also called the *trivial subdivision* of the cone or polytope. In general, subdivisions are defined with the next definition below.

Also the *boundary complex* of a d-dimensional polytope naturally has the structure of a pure polyhedral complex. The maximal cells are the facets of the polytope, and its dimension is $d - 1$, the dimension of the facets of the polytope.

Definition A.37. The *face vector* of a pure d-dimensional polyhedral complex $\mathcal{C}$ is the vector

$$\mathsf{f}(\mathcal{C}) = (\mathsf{f}_0, \mathsf{f}_1, \ldots, \mathsf{f}_d)$$

where f_k counts the number of k-dimensional faces of $\mathcal{C}$.

We will see later that the entries of the face vector satisfy a linear relation, the Euler equation. The *Euler characteristic* of the complex $\mathcal{C}$ is

$$\chi(\mathcal{C}) \;:=\; \mathsf{f}_0 \,-\, \mathsf{f}_1 \,+\, \ldots \,+\, (-1)^d \mathsf{f}_d \,. \tag{A.13}$$

This satisfies some addition formula. Let $\mathcal{C}$ and $\mathcal{C}'$ be two polyhedral complexes such that $\mathcal{C} \cap \mathcal{C}'$ is a subcomplex of both. Then there union is also a polyhedral complex and

$$\chi(\mathcal{C}) + \chi(\mathcal{C}') = \chi(\mathcal{C} \cup \mathcal{C}') - \chi(\mathcal{C} \cap \mathcal{C}') \,. \tag{A.14}$$

A.8.2. Fans

Now we turn to our first special case of polyhedral complexes, the *fans*. Here is the definition.

Definition A.38. A *fan* is a pure connected polyhedral complex such that all cells of the complex are cones.

We can naturally associate a fan in dual space to each polyhedron. Its cones correspond to the faces of the polyhedron, where each cone collects all functionals $\mathbf{c}$ that are maximized on a particular face of the polyhedron.

Definition A.39. The *normal cone* $\mathsf{N}_P(F)$ of a face F of a polytope P is the set of linear functionals $\mathbf{c}$ such that there is some β with

$$\langle \mathbf{c}, F \rangle \;=\; \beta \qquad\qquad \text{and} \qquad\qquad \langle \mathbf{c}, P \rangle \;\leq\; \beta \,.$$

Proposition A.40. *The normal cone is a polyhedral cone whose ray generators are the facet normals defining the face F.* □

Normal cones of faces of a polyhedron are important in *Discrete Optimization*, where they appear in the construction of TDI systems. A system of affine inequalities $A\mathbf{x} \leq \mathbf{b}$ is *TDI (totally dual integral)* if, for any integral linear objective function $\mathbf{c}$ the dual optimal solution is integral. This is the case if and only if the subset of rows of A that are tight on a face F are a *Hilbert basis* of the cone spanned by these rows. Here, a finite set H of vectors is a *Hilbert basis* of a cone if any integral point in the cone can be written as a conic combination of vectors in H with integral coefficients.

Definition A.41. The *normal fan* of a polytope P is the collection of all normal cones of proper faces of P.

Fans naturally have the structure of a polyhedral complex. In this case all cells are cones. Note that we have an inclusion reversing bijection between the cells in the normal fan and the faces of the polyhedron. If R is the intersection of two faces F and F' of the polyhedron, then the normal cones of F and F' are faces of the normal cone of R, and vice versa.

Definition A.42. Let P be a d-polytope and F a face of P. The *tangent cone* $\mathsf{T}_F P$ of F is the cone

$$\mathsf{T}_F P := \{\mathbf{p} + \mathbf{v} \in \mathbb{R}^d \ : \ \mathbf{p} \in F,\, \mathbf{p} + \varepsilon\mathbf{v} \in P \text{ for some } \varepsilon > 0\}\,.$$

The tangent cone is the common intersection of all supporting half-spaces at F, see Problem A.14. Note that the tangent cones are not cones in the usual sense, as their apex is not in the origin. We call them *affine cone* if we want to emphasize this.

Definition A.43. Let C be a polyhedral cone. The *dual cone* $C^\star$ is the set of all functionals that are maximized at the minimal face of the cone (the apex if C is pointed), *i.e.*

$$C^\star \ := \ \left\{ \mathbf{a} \in \left(\mathbb{R}^d\right)^\star \ : \ \langle \mathbf{a}, \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in C \right\}\,.$$

We can use a point $\mathbf{w} \in F$ to shift the cone into the origin. You will prove the following proposition in Problem A.15.

Proposition A.44. *The shifted cone* $\mathsf{T}_F P - \mathbf{w}$ *is dual to the normal cone of* F.

Problem A.14
Problem A.15

A.8.3. Regular Subdivisions and Triangulations

Often it is useful to subdivide a polytope into smaller pieces and look at the pieces separately. It will turn out that the most useful subdivisions are those where all pieces are simplices. Such subdivisions are called *triangulations* of the polytope. Here is the formal definition.

Definition A.45. A *subdivision* of a polytope P is a pure polyhedral complex $\mathcal{S}$ such that $P = \bigcup_{C \in \mathcal{S}} C$.

A subdivision is a *triangulation* of P if all cells are simplices.

Recall the definition of the Euler characteristic from (A.13). For polytopes and subdivisions of polytopes we can compute this value. You can find a proof of this theorem in the book of Ziegler.[7]

Proposition A.46. *The Euler characteristic of a d-dimensional polytope P is $\chi(P) = 1-(-1)^d$, and the Euler characteristic of a subdivision $\mathcal{S}$ of a polytope is $\chi(\mathcal{S}) = 1$.*

A subdivision or triangulation is *without new vertices*, if $\mathcal{V}(\Delta_d) \subseteq \mathcal{V}(P)$ for any $\Delta_d \in \mathcal{T}$. We will use the basic fact that for every finite $V \subset \mathbb{R}^d$ the polytope $\mathsf{conv}\, V$ has a triangulation with vertex set V. Similarly, the cone $\operatorname{pos} V$ has a triangulation with rays $\{\mathbb{R}_{\geq 0}\mathbf{v} \,:\, \mathbf{v} \in V\}$.[8] We need a new definition for this.

Definition A.47. A subdivision $\mathcal{S}$ of a polytope with vertices $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ (of the subdivision) is regular if there is a weight vector $\mathbf{w}$ such that $\mathcal{S}$ is the projection of the lower hull of

$$\mathsf{conv}((w_i, \mathbf{v}_i) \mid 1 \leq i \leq m)\,,$$

where the lower hull is the polyhedral complex of those facets whose normal has negative first coordinate.

Given a set of points $V := \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ and a weight vector $\mathbf{w} \in \mathbb{R}^m$ we denote by $\mathcal{S}_{\mathbf{w}}(V)$ the regular subdivision obtained as the lower hull of

$$\operatorname{lift}(\mathbf{w}) \;:=\; \mathsf{conv}((w_i, \mathbf{v}_i) \mid 1 \leq i \leq m)\,.$$

A *regular triangulation* is a regular subdivision that is a triangulation.

This notion is important far beyond the following theorem, with many, also surprising consequences. You will show in Problem A.16 that all subdivisions of a polygon using only the vertices of the polygon are regular.

Problem A.16

The main fact about subdivisions that we need in these lecture notes is the following theorem, which guarantees that we can always find a triangulation of a polytope.

Theorem A.48. *Every d-polytope P has a regular triangulation using only the vertices of the polytope.*

Proof. Let $V := \mathcal{V}(P)$ be the vertices of the polytope. We can assume that P is full dimensional. We claim that any sufficiently generic vector $\mathbf{w}$ induces a regular triangulation.

The subdivision induced by $\mathbf{w}$ is a triangulation if and only if for each facet of the lower hull of $\operatorname{lift}(\mathbf{w})$ is a d-simplex, i.e. if at most $d+1$ of the points

$$(w_1, \mathbf{v}_1), \ldots, (w_d, \mathbf{v}_d)$$

[7] Sec. 8.2 Ziegler, *Lectures on polytopes*.
[8] De Loera, Rambau, and Santos, *Triangulations. Structures for algorithms and applications.*

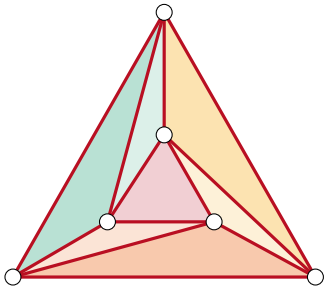


Figure A.12.: A non-regular subdivision

lie on a common hyperplane. For any $(d+2)$-tuple

$$(w_{i_1}, \mathbf{v}_{i_1}), \ldots, (w_{i_{d+2}}, \mathbf{v}_{i_{d+2}})$$

being on a common hyperplane means that the determinant

$$\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ w_{i_1} & w_{i_2} & \cdots & w_{i_{d+2}} \\ \mathbf{v}_{i_1} & \mathbf{v}_{i_2} & \cdots & \mathbf{v}_{i_{d+2}} \end{pmatrix}$$

vanishes. We can view this determinant as a linear functional in the entries of $\mathbf{w}$. There are $\binom{m}{d+2}$ different such functionals, hence, the complement Z^c of the union of the zero sets of these functionals is not empty. Choosing any $\mathbf{w} \in Z^c$ satisfies our requirements. □

It is important to realize that not all triangulations of a polytope are regular. See e.g. Figure A.12 for a simple example. You will prove that it is indeed not regular in Problem A.17

Problem A.17

Corollary A.49. *Every pointed cone C can be triangulated into simplicial cones without introducing new generators.*

Proof. If C is pointed, then there is a functional $\mathbf{u}$ such that

$$\mathbf{u}^t\mathbf{x} > 0 \qquad \text{for all } \mathbf{x} \in C\,.$$

Then $P := C \cap \{\mathbf{x} \mid \mathbf{u}^t\mathbf{x} = 1\}$ is a polytope, and C is the cone over P. By the previous Theorem A.48 P has a regular triangulation $\mathcal{T}$ without new vertices. The cones over the cells in this triangulation give a triangulation of the cone C without using new generators. □

Let $V \subseteq \mathbb{R}^d$ be a finite set of points. If the points are not in convex position, then not all points in V must be vertices of $\mathcal{S}_{\mathbf{w}}(V)$. For example, if $V = \{0, 1, 2\} \subseteq \mathbb{R}$ and we lift them to height 0, 1, and 0 in that order, then $\mathcal{S}_{(0,1,0)}(V)$ is the interval $[0, 2]$. In most settings one nevertheless distinguishes between this and the subdivision of $V' := \{0, 2\}$ obtained by lifting both points to height 0. Thus, we want to preserve the information

from which set V of points our subdivision was defined, even if not all points in V are a vertex of some cell in the subdivision.

It follows immediately from the definition that we can subdivide any finite set V of points into *some* regular subdivision by randomly choosing heights for the points. We have also seen that we can always find a regular *triangulation* of a set of points, if the points are the vertices of a polytope.

Theorem A.48 shows that we can also find a regular triangulation for an arbitrary finite set of points (in most applications we will consider $P \cap \Lambda$ for a polytope P and a lattice Λ) if we don't require that all points of the set are vertices of the triangulation. More generally, we obtain a triangulation if the heights are sufficiently generic. This is the case if no $d+1$ points

$$(w_{i_1}, \mathbf{v}_{i_1}), \ldots, (w_{i_{d+2}}, \mathbf{v}_{i_{d+2}})$$

lie on a common hyperplane. This is an open condition, so we can always choose weights that satisfy this. However, also this does not guarantee that all points are actually vertices of the subdivision.

We can fix this with the following operation. Assume that $\mathbf{w} := (w_i, \mathbf{v}_i)$ is not a vertex in the subdivision. This means that $\mathbf{w}$ is not a vertex of the lower hull (it may still be on a higher dimensional face of it). Hence, there is $w_i' \leq w_i$ such that $\mathbf{w}' := (w_i', \mathbf{v}_i)$ is on the lower hull, and contained in some facets $F_1, \ldots, F_k$ of it. We can decrease w_i' further, so that $\mathbf{w} = (w_i', \mathbf{v}_i)$ is beyond the facets $F_1, \ldots, F_k$, but still beneath all others. With this choice the set of heights remains generic, and $\mathbf{w}$ is now a vertex of the lower hull. So $\mathbf{v}_i$ is a vertex of the subdivision. We can repeat this process with any other point that is not yet a vertex in the subdivision. This proves that we can find a regular triangulation that uses all points.

Corollary A.50. *Let $V \subseteq \mathbb{R}^d$ be a finite set of points. Then there is a regular triangulation of V such that all points in V are vertices of the subdivision.* □

This idea also leads to a standard method to triangulate a finite set of points V, the *pulling triangulations*. This is obtained with the following procedure. We order the points in V arbitrarily as $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. We lift them all to the same height, say 0. Now we pick the first point and use the above process. So we lower the point slightly and obtain a subdivision. We remember the subdivision and put the point back to height 0. $\mathbf{v}_1$ is a vertex in this subdivision (among others). Now we pick the next point $\mathbf{v}_2$ and do the same separately *in each cell it is contained in*. It is not hard to see that this induces the same subdivision on the intersection of cells, so collecting all cells together gives a new subdivision of the original point set. Now also $\mathbf{v}_2$ is a vertex. We can repeat this process until $\mathbf{v}_k$. This gives a regular subdivision in which all points are vertices.

Pulling triangulations also prove that we can triangulate any finite point set in polynomial time, as each iteration above is a polynomial operation. You will show in Problem A.18 that one can construct a height function for this triangulation.

Problem A.18

We can use our process also for another common task. Suppose we are given regular subdivision $\mathcal{S} := \mathcal{S}_{\mathbf{w}}(V)$ of a point set $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ and some $\mathbf{u} \notin V$. Then $\mathcal{S}$ is

also a subdivision of $V \cup \{\mathbf{u}\}$. To see this, we just fix some height $w_{\mathbf{u}} > w_i$ for all $1 \leq i \leq k$ of $\mathbf{u}$. If we want $\mathbf{u}$ to be a vertex in this subdivision, then we lower the height as above. The new subdivision is called a *refinement* of $\mathcal{S}$. We can of course refine a subdivision also with a finite point set, by adding one point after the other.

Subdivisions of the same point set V are partially ordered by inclusion of the cells. Let $\mathcal{S}_1$, $\mathcal{S}_2$ be two subdivisions of V. We say that $\mathcal{S}_1 \preccurlyeq \mathcal{S}_2$ if for any cell $\sigma \in \mathcal{S}_1$ there is a cell $\tau \in \mathcal{S}_2$ such that $\sigma \subseteq \tau$.

Definition A.51. Let V be a finite point set and $\mathcal{S}_1$, $\mathcal{S}_2$ two subdivisions of V. The *common refinement* of $\mathcal{S}_1$ and $\mathcal{S}_2$ is the set

$$\{\, \sigma_i \cap \sigma_2 \;:\; \sigma_1 \in \mathcal{S}_1, \sigma_2 \in \mathcal{S}_2 \,\}\,.$$

You will prove in Problem A.19 that the common refinement is indeed a subdivision, and it is regular if both $\mathcal{S}_1$ and $\mathcal{S}_2$ are regular.

Problem A.19

A.9. Problems

A.1. Prove Proposition A.6.

A.2. Let $X_1 \subseteq X_2 \subseteq \mathbb{R}^d$.
Show that $X_2^\vee \subseteq X_1^\vee$.

A.3. Let $X_i \subseteq \mathbb{R}^d$ for $i \in I$.
Show that $\left(\bigcup_{i\in I} X_i\right)^\vee = \bigcap_{i\in I} X_i^\vee$.

A.4. Let $X \subseteq \mathbb{R}^d$ and $\mathbf{t} \in \mathbb{R}^d$ such that $0 \in \operatorname{int}(X) \cap \operatorname{int}(\mathbf{t} + X)$. Show that

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

A.5. Prove Minkowski's determinant inequality for positive definite matrices, *i.e.* prove that for any two positive definite square matrices A_1, A_2 with d rows and columns we have

$$(\det\, A_1 + A_2)^{1/n} \;\geq\; \left((\det\, A_1)^{1/n} \;+\; (\det\, A_2)^{1/n}\right)$$

with equality if and only if there is $\mu > 0$ such that $A_1 = \mu A_2$.
If also $\det\, A_1 = \det\, A_2$ in the equality case, then $A_1 = A_2$.

A.6. Let C_d be the cube defined by $|x_i| \leq 1$. Prove that the maximum volume ellipsoid is the unit ball.

A.7. Let $\Delta_d \subseteq \mathbb{R}^{d+1}$ be the d-dimensional simplex defined as the convex hull of the unit vectors. Prove that a maximum volume ellipsoid is the ball in the affine hull of Δ_d with center $1/(d+1)\mathbf{1}$.
Deduce that the bound of Theorem A.13 is best possible.

A.8. Prove Theorem A.13.

A.9. Prove Carathéodory's Theorem (Theorem A.17).

A.10. Show that the preimage of the projection of a face F is again a face (but not necessarily the original one).

A.11. If some $\mathbf{x} \in \mathbb{R}^d$ is in the relative interior of two faces of a convex set, then the two faces coincide.

A.12. Let $\pi : \mathbb{R}^d \to \mathbb{R}^m$ be a projection that maps a d-dimensional polytope P onto a m-dimensional polytope Q. Then, if x is a point in the interior of Q, $\operatorname{relint}(\pi^{-1}(x) \cap P) \subseteq \operatorname{int}(P)$.

A.13. Fill in the missing computations in the proof of Theorem A.32.

A.14. Prove that the tangent cone of a face of a polytope is precisely the intersection of the half spaces defining F.

A.15. Prove Proposition A.44.

A.16. Show that any subdivision $\mathcal{S}$ of a polygon P such that $\mathcal{V}(\mathcal{S}) = \mathcal{V}(P)$ is regular.

A.17. Prove that the subdivision in Figure A.12 is not regular.

A.18. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a finite set of points with pulling triangulation $\mathcal{T}$. Show how one can obtain a height function for this triangulation.

A.19. Show that the common refinement is a subdivision.
Show that the common refinement of two regular subdivisions is regular.

B. Solutions to some Exercises

B.1. Solutions for *The Geometry of Lattices* (*Chapter 2*)

2.1. We can move any element $\mathbf{x}$ into the origin by a translation of $\mathbb{R}^d$ that maps Λ into Λ. The inverse maps balls around the origin into balls around the $\mathbf{x}$. If a ball around the origin contains no other point of Λ, then so does its translate into $\mathbf{x}$.

2.2. Let $(\mathbf{x}_n)_{n\in\mathbb{Z}_{\geq 0}}$ be a convergent sequence in Λ and $\mathbf{x} = \lim_{n\to\infty} x_n$ its limit. We want to show that $\mathbf{x} \in \Lambda$. Let $\varepsilon > 0$ be such that $\mathcal{B}_{2\varepsilon}(\mathbf{x}_0) \cap \Lambda = \{\,\mathbf{x}\,\}$. By Problem 2.1 we know that we can use the same radius of a ball for any point in the lattice, *i.e.* $\mathcal{B}_{2\varepsilon}(\mathbf{x}_n) \cap \Lambda = \{\,\mathbf{x}\,\}$ for all $n \in \mathbb{Z}_{\geq 0}$.

As the sequence converges there is some $n_0 \in \mathbb{Z}_{\geq 0}$ such that for $n \geq n_0$

$$\|\mathbf{x}_n - \mathbf{x}_{n+1}\| \;\leq\; \varepsilon\,.$$

This implies $\mathbf{x}_{n+1} \in \mathcal{B}_{2\varepsilon}(\mathbf{x}_n)$, so $\mathbf{x}_{n+1} = \mathbf{x}_n$ for $n \geq n_0$. Hence, the sequence becomes stationary and $\mathbf{x} = \mathbf{x}_n$ for any $n \geq n_0$. So $\mathbf{x} \in \Lambda$.

2.3. For a closed subset this follows from the Theorem of Bolzano-Weierstrass. A counterexample among non-closed sets is, *e.g.*, the set $\{\frac{1}{n} \;:\; n \in \mathbb{Z}_{>0}\} \subseteq [0,1]$

2.4. Consider, *e.g.* the set $\mathcal{A} := \{1, \sqrt{2}\} \subseteq \mathbb{R}$. Then $\Lambda_{\mathcal{A}}$ is

$$\Lambda_{\mathcal{A}} \;=\; \left\{\, a + b\sqrt{2} \right\} a, b \in \mathbb{Z}$$

Now $0 < \sqrt{2}-1 < 1/2$, so for any $\varepsilon > 0$ there is n such that $(\sqrt{2}-1)^n < \varepsilon$. The binomial expansion of the left hand side shows that $(\sqrt{2}-1)^n$ is an element in $\Lambda_{\mathcal{A}}$, so any open neighborhood of 0 contains a nonzero element of $\Lambda_{\mathcal{A}}$.

This is, of course, also true for any other $\mathcal{A} := \{1, \sqrt{q}\}$ unless q is a prime, but the proof is slightly more involved. For this, consider $\lambda := \inf(x \mid x \in \Lambda_{\mathcal{A}})$ and show that $\lambda \in \Lambda_{\mathcal{A}}$ (otherwise, there are $y_1, y_2 \in \Lambda_{\mathcal{A}}$ such that $\lambda < y_1 < y_2 < 3/2\lambda$, but then $y_2 - y_1$ is a positive element of $\Lambda_{\mathcal{A}}$ that is smaller than λ) and that λ is a basis of $\Lambda_{\mathcal{A}}$ (for any $x \in \Lambda_{\mathcal{A}}$ let y be the largest integer less than x/λ, so that $0 < x - y\lambda < \lambda$ and thus $x = y\lambda$). This implies that $\sqrt{q}$ is a rational multiple of 1.

2.5. Direct computation.

2.6. Direct computation.

2.7. Use Lemma 2.6.

2.8. We use induction over the dimension d of the cone C. If $d = 1$, then C is either $\mathbb{R}_{\geq 0}$ or $\mathbb{R}_{\leq 0}$. We may assume $C = \mathbb{R}_{\geq 0}$. Let $\mathbf{v} \in \mathbb{R}$ be any generator of the lattice. Then also $-\mathbf{v}$ generates the lattice, so we may assume $\mathbf{v} > 0$ and thus $\mathbf{v} \in C$.

Now let $d \geq 2$. Choose any facet F of C defined by the inequality $\langle\,\mathbf{f}, \mathbf{x}\,\rangle \leq 0$ and consider the $(d-1)$-dimensional subspace $L := \operatorname{lin} F$. Let $\Lambda' := \Lambda \cap L$. The generators of the cone F are Λ-rational, so some multiple of the generators is in Λ. Hence, Λ' has rank $d-1$.

By our induction hypothesis we can find a basis $\mathrm{B}' := \{\,\mathbf{b}_1, \ldots, \mathbf{b}_{d-1}\,\}$ of Λ' such that $\mathbf{b}_i \in F$ for $1 \leq i \leq d-1$. Let $\mathbf{c} := \mathbf{b}_1 + \cdots \mathbf{b}_{d-1}$. Then $\mathbf{c}$ is a relative interior point of F.

Let $\mathbf{b}_d$ be any lattice vector that extends B' to a basis of Λ. Such a vector exists by Proposition 2.10. We may assume $\langle\,\mathbf{f}, \mathbf{b}_d\,\rangle < 0$ (otherwise take $-\mathbf{b}_d$). Clearly, also $\mathrm{B}' \cup \{\,\mathbf{b}_d + \lambda\mathbf{c}\,\}$ generates Λ for any $\lambda \in \mathbb{Z}$. Let $\mathbf{a}$ be any facet (outer) normal of C such that $]pro\mathbf{a}\mathbf{b}_d > 0$, *i.e.* $\mathbf{b}_d$ is beyond the

facet defined by $\mathbf{a}$. As $\mathbf{c} \in F$ also $\mathbf{c} \in C$, so $\langle \mathbf{a}, \mathbf{c} \rangle < 0$ (the inequality is strict as $\mathbf{c}$ cannot be in the facet defined by $\mathbf{a}$). Hence

$$\langle \mathbf{a}, \mathbf{b}_d + \lambda \mathbf{c} \rangle \ < \ 0 \qquad \text{for} \qquad \lambda \ > \ -\frac{\langle \mathbf{a}, \mathbf{b}_d \rangle}{\langle \mathbf{a}, \mathbf{c} \rangle} \, .$$

Observe that the right hand side of the second inequality is positive as $\langle \mathbf{a}, \mathbf{b}_d \rangle < 0$ by assumption. Hence, by replacing $\mathbf{b}_d$ by $\mathbf{b}_d + \left\lceil -\frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{c} \rangle} \right\rceil$ we obtain $\langle \mathbf{a}, \mathbf{b}_d \rangle < 0$. C has only finitely many facets, so we can replace $\mathbf{b}_d$ by a generator in the interior of C.

Note that the results is also true without the requirement that C is Λ-rational. For this we need to show that we can find a full-dimensional Λ-rational cone inside C.

2.9. The convex hull is

$$\mathsf{conv}(0, \mathbf{b}_1, v b_2) \ = \ \{ \, \mu_1 \mathbf{b}_1 + \mu_2 \mathbf{b}_2 \ : \ 0 \leq \mu_1, \mu_2 \leq 1 \text{ and } \mu_1 + \mu_2 \leq 1 \, \} \, .$$

Hence, $\mathsf{conv}(0, \mathbf{b}_1, v b_2) \setminus \{ \, \mathbf{b}_1, \mathbf{b}_2 \, \} \subseteq \Pi(\mathbf{b}_1, \mathbf{b}_2)$. If $\mathsf{conv}(0, \mathbf{b}_1, v b_2)$ contains a nonzero lattice point $\mathbf{a} \neq \mathbf{b}_1, \mathbf{b}_2$, then $\mathbf{a} \in \Pi(\mathbf{b}_1, \mathbf{b}_2)$, so $\mathbf{a} \notin \Lambda$.

Conversely, if $\mathbf{b}_1, \mathbf{b}_2$ are not a lattice basis, then there is $\mathbf{a} \in \Pi(\mathbf{b}_1, \mathbf{b}_2)$, so there are $0 \leq \mu_1, \mu_2 < 1$ such that

$$\mathbf{a} \ = \ \mu_1 \mathbf{b}_1 \ + \ \mu_2 \mathbf{b}_2 \, .$$

If $\mu_1 + \mu_2 \leq 1$, then $\mathbf{a} \in \mathsf{conv}(0, \mathbf{b}_1, \mathbf{b}_2)$. Otherwise,

$$\mathbf{a}' \ := \ \mathbf{b}_1 \ + \ \mathbf{b}_2 \ - \ \mathbf{a} \ = \ (1 - \mu_1) \mathbf{b}_1 \ + \ (1 - \mu_2) \mathbf{b}_2$$

is a lattice point with $0 < \mu_1, \mu_2 \leq 1$ and $\mu_1 + \mu_2 < 1$. Hence $\mathbf{a}' \in \mathsf{conv}(0, \mathbf{b}_1, \mathbf{b}_2)$.

The last argument fails in higher dimensions and indeed the whole claim fails, as we can see from the Reeve tetrahedra

$$\Delta \ := \ \mathsf{conv} \, (0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + m \mathbf{e}_3)$$

for some integral $m > 0$. We have

$$\Delta \ \cap \ \mathbb{Z}^3 \ = \ \{ \, 0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + m \mathbf{e}_3 \, \} \, .$$

but $\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + m \mathbf{e}_3$ is not a lattice basis for $m \geq 2$.

2.10. There are various options to prove this theorem. We give two of them, one using the previous exercise, and one independent one.

In both proofs we first translate the triangle with a lattice vector such that one vertex is the origin. Hence, we may assume that the vertices are

$$\mathbf{v}_0 \ := \ \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right] \qquad \mathbf{v}_1 \ := \ \left[\begin{smallmatrix} a \\ b \end{smallmatrix} \right] \qquad \mathbf{v}_2 \ := \ \left[\begin{smallmatrix} p' \\ q' \end{smallmatrix} \right] \, . \tag{B.1}$$

For the first proof observe that

$$\mathsf{conv}(0, \mathbf{v}_1, \mathbf{v}_2) \ \cap \ \mathbb{Z}^2 \ = \ \{ \, 0, \mathbf{v}_1, \mathbf{v}_2 \, \} \, .$$

By Problem 2.9 $\mathbf{v}_1$ and $\mathbf{v}_2$ are a basis of $\mathbb{Z}^2$, so $\det(\mathbf{v}_1, \mathbf{v}_2) = 1$. This is the volume of $\mathsf{conv}(0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2)$. The map $\varphi : \mathbf{x} \mapsto \mathbf{v}_1 + \mathbf{v}_2 - \mathbf{x}$ is a bijection from $\mathsf{conv}(0, \mathbf{v}_1, \mathbf{v}_2)$ to $\mathsf{conv}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2)$, as any

$$\mathbf{x} \ = \ \mu_0 \cdot 0 \ + \ \mu_1 \mathbf{v}_1 \ + \ \mu_2 \mathbf{v}_2 \qquad \text{with} \qquad 0 \leq \mu_0, \mu_1, \mu_2 \leq 1, \ \mu_0 + \mu_1 + \mu_2 = 1$$

is mapped to

$$\begin{aligned}\mathbf{v}_1+\mathbf{v}_2-\mathbf{x} &= (1-\mu_1)\mathbf{v}_1 + (1-\mu_2)\mathbf{v}_2 - \mu_0\cdot 0\\ &= \mu_2\mathbf{v}_1 + \mu_1\mathbf{v}_2 + (1-\mu_1-\mu_2)(\mathbf{v}_1+\mathbf{v}_2) = \mu_2\mathbf{v}_1 + \mu_1\mathbf{v}_2 + \mu_0(\mathbf{v}_1+\mathbf{v}_2)\,.\end{aligned}$$

So the area of $\Delta \cong \mathsf{conv}(0,\mathbf{v}_1,\mathbf{v}_2)$ is $\frac{1}{2}$.

We can also prove this without referring to lattices. For this, consider again the vertices as in (B.1). These are the only lattice points in the triangle, so we also know $\gcd(a,b)=\gcd(p',q')=1$. We can find $x,y\in Z$ with $1=xa+by$. We apply the unimodular transformation

$$A := \left[\begin{smallmatrix} x & y \\ -b & a\end{smallmatrix}\right]$$

that maps v_2 onto $\mathbf{e}_1$ and v_2 onto a point

$$\mathbf{v} := \left[\begin{smallmatrix} p \\ q\end{smallmatrix}\right].$$

Using a reflection if necessary we may assume $q>0$, and by applying

$$B := \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1\end{smallmatrix}\right]$$

or its inverse we can assume that $0\le p<q$. If $p=0$, then $q=1$, otherwise $\mathbf{e}_2$ is an additional lattice point. In this case Δ has area $\frac{1}{2}$.

So assume $p>0$. Then $q>1$, as $p<q$, and

$$\mathbf{w} := \left[\begin{smallmatrix} 1 \\ 1\end{smallmatrix}\right] = \alpha\left[\begin{smallmatrix} 1 \\ 0\end{smallmatrix}\right] + \beta\left[\begin{smallmatrix} p \\ q\end{smallmatrix}\right]$$

for

$$\alpha := \left(1-\frac{p}{q}\right) \qquad\qquad \beta := \frac{1}{q}\,.$$

But $0<\alpha,\beta<1$ and $0<\alpha+\beta\le 1$, so $\mathbf{w}$ is an additional lattice point. Hence, the case $p>0$ cannot occur.

2.11. We look at the following three cases.

(i) $b+i=3$

(ii) $b=3$ and $i>0$.

(iii) $b>3$.

In the first case the claim follows from Problem 2.10. In the other cases we use induction on $b+i$. In the second case we can split P into three triangles with less lattice points using an interior lattice point. In the third case we use two nonadjacent points on the boundary to split P into two pieces with less lattice points each.

The volume of P is in both cases the sum of the volumes of the pieces, and we can now verify the formula by direct computation using the induction assumption (express the number of interior and boundary points of P in the numbers of interior points and boundary points of the pieces. Observe that some boundary points of the pieces are interior points of P and my aappear in two different pieces).

2.12. If the greatest common divisor is 0, then $\lambda_j=\cdots=\lambda_d=0$ and $\mathcal{A}$ is linearly dependent. So assume $\gcd(\lambda_j,\ldots,\lambda_d)=g\ge 2$. Then $\mathcal{A}$ is linearly indepenendent. Let

$$\mathbf{v} := \sum_{i=j}^{d}\lambda_i\mathbf{b}_i = \mathbf{a} - \sum_{i=1}^{j-1}\lambda_i\mathbf{b}_i\,.$$

The coefficients of the second term are all divisible by g, so

$$\frac{1}{g}\mathbf{v} := \sum_{i=j}^{d}\frac{\lambda_i}{g}\mathbf{b}_i = \frac{1}{g}\mathbf{a} - \sum_{i=1}^{j-1}\frac{\lambda_i}{g}\mathbf{b}_i\,.$$

is a lattice vector. But the coefficient of $\mathbf{a}$ in the unique representation of $\frac{1}{g}\mathbf{v}$ in the set $\mathcal{A}$ is not integral, so $\mathcal{A}$ is not primitive.
Conversely, let $\mathbf{v} \in \operatorname{lin} \mathcal{A} \cap \Lambda$. Then $\mathbf{v}$ can be repesented in two ways as

$$\mathbf{v} \;=\; \sum_{i=1}^{j-1} \mu_i \mathbf{b}_i \;+\; \mu_j \mathbf{a} \;=\; \sum_{k=1}^{d} \nu_k \mathbf{b}_k$$

for $\mu_i \in \mathbb{R}$, $1 \le i \le j$ and $\nu_i \in \mathbb{Z}$ for $1 \le k \le d$. Using $\mathbf{a} := \sum_{i=1}^{d} \lambda_i \mathbf{b}_i$ we obtain

$$\mathbf{v} \;=\; \sum_{i=1}^{j-1} (\mu_i + \mu_j \lambda_i) \mathbf{b}_i \;+\; \sum_{k=j}^{d} \mu_j \lambda_k \mathbf{b}_k \;=\; \sum_{k=1}^{d} \nu_k \mathbf{b}_k \,.$$

Comparing coefficients we obtain $\mu_j \lambda_k = \nu_k$ for $1 \le k \le d$ and $\gcd(\lambda_j, \ldots, \lambda_d) = 1$ implies $\mu_j \in \mathbb{Z}$. Comparing the remaining coefficients gives $\mu_i \in \mathbb{Z}$ for $1 \le i \le j-1$.

2.13. ▷ The area A of a regular triangle with side length a is $\frac{\sqrt{3}}{2}a^2$, and a^2 is integral. So A is of the form $\frac{k}{2}\sqrt{3}$ for an integer k.
Yet, by Pick's Theorem, or, if we move one vertex into the origin, as the determinant of the two other vertices is integral, the area is integral.
▷ The hexagonal lattice allows a regular triangle.
Observe that lattice preserving maps are not angle preserving.

2.14. There are various options to prove this.
Here is a direct one. Let $\mathbf{v}$ be the interior lattice point, $\mathbf{v}_1, \mathbf{v}_2\, vv_3$ be the vertices and $\lambda_1, \lambda_2, \lambda_3 > 0$ such that

$$\mathbf{v} \;=\; \sum \lambda_i \mathbf{v}_i$$

Assume that $\lambda_3 \ge \lambda_2 \ge \lambda_3$, *i.e.* λ_3 is the largest of the three coefficients. We move $\mathbf{v}_3$ into the origin and look at the lattice spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. $\mathbf{v}$ is a point in the fundamental parallelepiped F and

$$\mathbf{v} \;=\; \lambda_1 \mathbf{v}_1 \;+\; \lambda_2 \mathbf{v}_2$$

By assumption $\lambda_1 \le \lambda_2 < \frac{1}{2}$, so also $2\mathbf{v} \in F$. Hence, also $\mathbf{v}' := (\mathbf{v}_1 + \mathbf{v}_2) - 2\mathbf{v} \in F$.
By assumption, $2\mathbf{v} \not\in \Delta$. But then $\mathbf{v}' \in \Delta$ and thus $\mathbf{v}' = \mathbf{v}$, so

$$(\mathbf{v}_1 + \mathbf{v}_2) - 2\mathbf{v} \;=\; \mathbf{v} \qquad \Longleftrightarrow \qquad \mathbf{v} \;=\; \frac{1}{3}(\mathbf{v}_1 + \mathbf{v}_2)\,.$$

Another proof uses the Theorem of Pick. Δ can be triangulated into three triangles of volume $\frac{1}{2}$. If we move $\mathbf{v}$ into the origin, then the volume of the triangles equals half the determinant of the two other vertices, *i.e.*

$$\det(\mathbf{v}_1, \mathbf{v}_2) \;=\; \det(\mathbf{v}_1, \mathbf{v}_3) \;=\; \det(\mathbf{v}_2, \mathbf{v}_3) \;=\; 1$$

Further, after a unimodular transformation we can assume that $\mathbf{v}_1 = \mathbf{e}_1$ and $\mathbf{v}_2 = \mathbf{e}_2$. It is easy to see that then $\mathbf{v}_3 = -(\mathbf{v}_1 + \mathbf{v}_2)$.
There is no similar statement in higher dimensions, as

$$P^k \;:=\; \operatorname{conv}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -k \end{bmatrix} \right)$$

for $k \in \{1, 2\}$ is a lattice simplex with exactly one interior point and only the vertices on the boundary.

2.15. Let g be the common denominator of all entries of A. Then gA is an integral matrix, and if H is in Hermite normal form with a unimodular transformation U such that $gA \;=\; HU$, then also $\frac{1}{g}H$

is in Hermite normal form and $A = \frac{1}{g}HU$. Hence, in the following we can replace A by gA and assume that A is an integral matrix.

Now observe that the following three transformations on the columns of a matrix A can be realized by a multiplication with suitably chosen unimodular matrix T from the right:

(i) Exchanging two columns, and

(ii) multiplying a column by -1, and

(iii) adding an integral multiple of one column to another column.

These operations are called *elementary transformations* for a matrix. Any succession of such operations is then realized by the product of the corresponding transformation matrices, which is again unimodular. In the following, we will show that we can transform A into its Hermite normal form using only such elementary transformations. The unimodular matrix U in the theorem is then given by the product of the corresponding transformation matrices.

We show that we can transform A into its Hermite normal form by induction on the rows of A. So assume that A already has the form

$$A = \begin{bmatrix} B & \mathbf{0} \\ M & C \end{bmatrix} \tag{B.2}$$

for matrices B, C, M where $B \in \mathbb{Z}^{k\times k}$ is in Hermite normal form and $k \geq 0$. Consider the first row $(c_{11}, \ldots, c_{1,m-k})$ of the matrix C. Using elementary column operations we can transform C such that

(i) $c_{11} \geq c_{12} \geq \ldots, c_{1,m-k} \geq 0$ and

(ii) $c := c_{11} + c_{12} + \cdots + c_{1,m-k}$ is as small as possible.

Then $c_{11} > 0$ as A has full row rank. Further, if $c_{12} \neq 0$, then we can subtract the second from the first column and reorder the columns if necessary to obtain a smaller total sum c. Hence, $c_{12} = c_{13} = \ldots = c_{1,m-k} = 0$. The column operations on C clearly extend to A without affecting B and M, so we can apply them to A to obtain a matrix

$$A = \begin{bmatrix} B & 0 & \mathbf{0} \\ m & c_{11} & 0 \\ M' & c'_1 & C' \end{bmatrix},$$

where m' is a row vector of length k, the first row of the matrix M. By adding or subtracting multiples of the $(k+1)$st column (the one containing c_{11}) to the first k columns of A we can assume that all entries of m are non-negative and smaller than c_{11}.

In this way we have again reached a matrix of the form (B.2), but this time B has size $(k+1)\times(k+1)$. After d steps A is in Hermite normal form using only elementary operations.

We still need to prove that the Hermite normal form is unique. Assume that there are two different Hermite normal forms $H_1 = (h^{(1)}_{ij})_{ij}$ and $H_2 = (h^{(2)}_{ij})_{ij}$ with unimodular transforms $A = H_1U_1 = H_2U_2$. The product of unimodular matrices is unimodular, so $H_1 = H_2U$ for some unimodular U.

We first look at the diagonal. Assume they differ and let i be the smallest index, where this happens. We may assume $h^{(1)}_{jj} < h^{(2)}_{jj}$, otherwise we swap H_1 and H_2 and use U^{-1}. Let $\mathbf{u} = (u_k)_k$ be the j-th column of U. Then $u_k = 0$ for $k < i$ and thus $h^{(2)}_{jj}u_i = h^{(1)}_{jj}$. So $0 < u_j < 1$, but U is integral.

Now assume $h^{(1)}_{ij} < h^{(2)}_{ij}$ and choose i minimal with this property. Then $j < i$. Let $\mathbf{u} = (u_k)_k$ be the j-th column of U. Then $u_k = 0$ for $k < j$, $u_j = 1$ and $u_k = 0$ for $j < k < i$. But $h^{(1)}_{ij} = h^{(2)}_{ij} + h^{(2)}_{ii}u_i$ is less than $h^{(2)}_{ij}$ and $0 \leq h^{(2)}_{ij} < h^{(2)}_{ii}$, so $-1 < u_i < 0$. However, U is integral.

2.16. The construction of the unimodular matrix in Problem 2.15 is not necessarily polynomial, as we

- ▷ cannot control the number of elementary operations necessary to turn the first row of C into one that has all zeros except for the first entry, and
- ▷ cannot control the size of the entries in the result and in all intermediate steps.

Hence, we need a couple of new observations to fix this. We first look for an improvement of the algorithm that lets us bound the number of steps necessary to obtain the Hermite normal form by a polynomial in the input size, before we look at the size of the result and finally the size of all intermediate steps.

Given any integers a and b with greatest common divisor g we can find, in polynomial time $\mathcal{O}(\log(a)\cdot\log(b)$ in the size of a and b, two integers x and y such that $g = x\cdot a + y\cdot b$. This can be done with the *extended Euclidean algorithm*.

In our computation we now use at most one elementary operation to make the first entry of the first row of C positive. Then we replace, for all j, the first column c_1 of C by $xc_1 + yc_j$ and the column c_j by $\frac{1}{g}(c_{1j}c_1 - c_{11}c_j)$. Note that in the second linear combination the coefficients $\frac{1}{g}c_{1j}$ and $\frac{1}{g}c_{11}$ are both integral. A simple consideration shows that the transformation matrix corresponding to the transformation used in the second step has determinant ± 1 and thus is unimodular.

We need at most d of these operations in each iteration of the original construction of H, which can all be done in polynomial time, so we obtain the Hermite normal form in polynomially many steps in the size of the input matrix A.

We also need to control the size of the output and all intermediate steps of this algorithm to prove that it runs in polynomial time. We look first at the size of the output, *i.e.* at the Hermite normal form H.

Observe that in each step we transform the first row of C so that only the first entry is nonzero. A simple consideration shows that this is the greatest common divisor of all the entries in the first row of C.

We obtain all transformations by swapping columns, multiplying a column with -1 and adding (a multiple of) a column to another. Linearity of the determinant in the columns of a matrix shows that these operations preserve the greatest common divisor of all subdeterminants of size d of A. This is clear for the first two operations.

For the last, let Q be a square submatrix of A with first column q, r any other column of A, Q' the matrix obtained from Q by replacing q with $q + \lambda r$ for some $\lambda \in \mathbb{Z}$ and R the matrix obtained from Q by replacing q with r. Observe that R is also a submatrix of A. By linearity we get

$$\det Q' \;=\; \det Q \;+\; \lambda \det R$$

$$\gcd(\det Q', \det R) \;=\; \gcd(\det Q + \lambda \det R, \det R) \;=\; \gcd(\det Q, \det R)\,.$$

Hence, the determinant of H is the greatest common divisor of all $(d\times d)$-subdeterminants of A, which bounds the size of the diagonal entries of H by a polynomial in the size of the entries of A. All entries outside the diagonal in H are at most the size of the diagonal entry in the same row, so those are bounded as well.

It remains to control the intermediate entries in the matrices in each step. For this we observe first the following.Let $H = AU$ be some matrix A with Hermite normal form H and a unimodular transformation U. Let v be an element in the integral span of the columns, *i.e.* there is an integral vector p such that $Ap = v$. Then

$$\begin{bmatrix} H & 0 \end{bmatrix} \;=\; \begin{bmatrix} A & v \end{bmatrix} \cdot \begin{bmatrix} U & p \\ 0 & 1 \end{bmatrix}\,.$$

So the Hermite normal form of $\begin{bmatrix} A & v \end{bmatrix}$ is $\begin{bmatrix} H & 0 \end{bmatrix}$. Let g be the greatest common divisor of all $(d\times d)$-submatrices of A as above. This is the product of the diagonal entries of H, so it follows from the representation $H = AU$ that the multiples $g\mathbf{e}_i$ of the unit vectors are in the integral span of the columns of A. We obtain the same Hermite normal form (up to some zero columns) if we replace A by $\begin{bmatrix} A & g\,\mathrm{I} \end{bmatrix}$. As above we can now use the columns of $g\,\mathrm{I}$ to reduce all entries in intermediate matrices that exceed g. This finally bounds the size of all intermediate results and proves the polynomiality of our algorithms for the Hermite normal form.

2.17. We start with the first claim. If the system has an integral solution $\mathbf{x} + 0$, then $\mathbf{y}^t\mathbf{b} = \mathbf{y}^tA\mathbf{x} + 0$ is an integer whenever $\mathbf{y}^tA$ is integral.

Now assume that $\mathbf{y}^t\mathbf{b}$ is an integer whenever $\mathbf{y}^tA$ is integral. Then $A\mathbf{x} = \mathbf{b}$ has a fractional solution, as otherwise we can find $\mathbf{y}^t$ with $\mathbf{y}^tA = 0$, but $\mathbf{y}^t\mathbf{b} \neq 0$. By discarding redundant rows we can assume that A has full row rank.

Let $H = AU$ be the Hermite normal form of A. As U^{-1} is integral the integral solutions $\mathbf{x}_0$ of $A\mathbf{x} = \mathbf{b}$ correspond to integral solutions $U^{-1}\mathbf{x}_0$ of $AU\mathbf{x} = \mathbf{b}$. Thus, we can assume that $A = [B, 0]$ is in Hermite normal form. Then $B^{-1}b$ is integral (the i-th coordinate of $B^{-1}b$ is $\mathbf{y}^t\mathbf{b}$ if we choose the i-th row of B^{-1} for $\mathbf{y}^t$). Hence, $\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ is an integral solution of $A\mathbf{x} = \mathbf{b}$.

For the second claim we may assume that A is in Hermite normal form $A = [B, 0]$ with full row rank. Then any solution of $A\mathbf{x} = \mathbf{b}$ is of the form $\mathbf{x}_0 + \mathbf{y}$, where $\mathbf{y}$ is an integral solution of $A\mathbf{x} = 0$.

But those are the vectors $\mathbf{y} \in \{0\}^d \times \mathbb{Z}^{m-d}$.
Finally, for the last claim we observe that by Problem 2.16 we can compute the Hermite normal form in polynomial time.

2.19. Assume it is not a lattice basis of Λ but spans a sublattice Γ. Then there is a lattice point $\mathbf{v} \in \Lambda \setminus \Gamma$, and $\mathrm{B} \cup \{\mathbf{v}\}$ spans a lattice Γ' with

$$\Gamma \subseteq \Gamma' \subseteq \Lambda .$$

and we can find integral matrices A and A' representing the basis of Γ in that of Γ', and that of Γ' in Λ. The determinant is independent of the chosen basis, so we obtain

$$\det B = \det A \cdot \det B' = \det A' \det C$$

for bases B' of Γ' and C of Λ. This contradicts minimality of $\det B$.

2.20. If $\mathbf{x}, \mathbf{y} \in \Lambda$, then $\langle \mathbf{a}_i, \mathbf{x} \rangle = k_{\mathbf{x}} m_i$ and $\langle \mathbf{a}_i, \mathbf{u} \rangle = k_{\mathbf{y}} m_i$ for some $k_{\mathbf{x}}, k_{\mathbf{y}} \in \mathbb{Z}$ so

$$\langle \mathbf{a}_i, \mathbf{x}+\mathbf{y} \rangle = \langle \mathbf{a}_i, \mathbf{x} \rangle + \langle \mathbf{a}_i, \mathbf{x} \rangle = k_{\mathbf{x}} m_i k_{\mathbf{y}} m_i = (k_{\mathbf{x}} + k_{\mathbf{y}}) m_i .$$

so $\mathbf{x}+\mathbf{y} \in \Lambda$. Similarly, also $-\mathbf{x} \in \Lambda$ and also $0 \in \Lambda$, so Λ is a lattice.
By construction, it is a sublattice of $\mathbb{Z}^d$. Further, $m_1 \mathbf{e}_i \in \Lambda$ for $1 \le i \le k$, so the lattice spanned by $\mathcal{A} := m_1 \mathbf{e}_1, \ldots, m_k \mathbf{e}_k$ is a sublattice of Λ. The latter has determinant $m_1 \cdots m_k$, so

$$1 \le \det \Lambda \le m_1 \cdots m_k .$$

2.21. We only prove this in the case $\Lambda = \mathbb{Z}^d$. In this case the fundamental parallelepiped F of Λ is the (half open) unit cube $[0,1)^d$. The longest vector in F has length $\sqrt{d}$ and $\mathrm{vol}\ F = 1$.
Let $B^{\pm} := \mathcal{B}_{r \pm \sqrt{d}}(0)$. For any $\mathbf{x}, \mathbf{y} \in \mathcal{B}_r(0) \cap \Lambda$ we have $\mathbf{x} + F \subseteq B^+$, and $\mathbf{x} + F \cap \mathbf{y} + F = \varnothing$ for $\mathbf{x} \neq \mathbf{y}$. Hence,

$$\begin{aligned} |\mathcal{B}_r(0) \cap \Lambda| &= \sum_{\mathbf{x} \in \mathcal{B}_r(0) \cap \Lambda} 1 = \sum_{\mathbf{x} \in \mathcal{B}_r(0) \cap \Lambda} \mathrm{vol}(F) \\ &= \sum_{\mathbf{x} \in \mathcal{B}_r(0) \cap \Lambda} \mathrm{vol}(\mathbf{x} + F) \le \mathrm{vol}\ B^+ . \end{aligned}$$

Similarly, B^- is completely covered by the translates $\mathbf{x} + F$ for $\mathbf{x} \in \mathcal{B}_r(0) \cap \Lambda$, so

$$\mathrm{vol}\ B^- \le |\mathcal{B}_r(0) \cap \Lambda| .$$

Now, for some chosen $\varepsilon > 0$,

$$\frac{\mathrm{vol}\, B^+}{|\mathcal{B}_r(0)|} = \frac{(r+\sqrt{d})^d}{r^d} \le (1+\varepsilon) \tag{B.3}$$

for r large enough. Similarly

$$\frac{\mathrm{vol}\, B^-}{|\mathcal{B}_r(0)|} = \frac{(r-\sqrt{d})^d}{r^d} \ge (1-\varepsilon) \tag{B.4}$$

for some r large enough. Note that this also depends on the basis (which we have chosen to be $\mathbb{Z}^d$).
For a general lattice the volume of F may change, and the length of a longest vector in F. The first is accounted for by the additional factor $\det \Lambda$ in the general formula, the latter changes the constant $\sqrt{d}$ that appears in the numerator of (B.3) and (B.4), and thus the r we need to choose.

2.22. Let B be a basis of Λ and B_0 a basis of Γ (given as matrices of column vectors). Then

$$|\Gamma/\Lambda| \;=\; \frac{\det\, B_0}{\det\, B}\,.$$

Further, there is an integral matrix T such that $B_0 = BT$, so $B = B_0T^{-1}$. Now $(\det\, T)T^{-1}$ is integral, $\det\, T$ is the index, and $(\det\, T)B$ is a basis of $(\det\, T)\Lambda$.

2.23. As in the proof of the Hermite normal form it suffices to show that we can transform A into its Smith normal form using elementary row and column operations. The existence of the companions then follows.

We again use induction. Suppose that after some elementary transformations A has the form

$$A \;=\; \begin{bmatrix} S & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix} \tag{B.5}$$

where S is a diagonal matrix with positive entries $s_{11}, \dots, s_{kk}$ for $k \geq 0$ on the diagonal such that $s_{j-1,j-1}$ divides s_{jj} for $2 \leq j \leq k$, and s_{kk} divides all entries of C.

Among all transformations of C that we can reach with elementary row and column operations we pick one such that

$$\min(|c_{ij}| \mid 1 \leq i \leq d, 1 \leq j \leq m \text{ and } c_{ij} \neq 0)$$

is minimal. We can also assume that this minimum is attained by c_{11}. Then clearly c_{11} is the only non-zero element in the first row and column, as otherwise we can obtain a smaller entry by a suitable row of column operation. Further, a similar consideration shows that c_{11} must divide all other entries of C. We have extended our induction form (B.5) from k to $k+1$.

Uniqueness of S follows from the observation that in each step the element c_{11} that we construct is the greatest common divisor of the elements in C.

2.24. We choose the any basis of Λ as a basis of the subspace $\operatorname{lin}\Lambda$. Then a basis of Λ' can be written in this basis with only integral coefficients. We apply Smith normal form to this matrix.

2.25. This is a classical fact from linear algebra. Let φ be any element of $(\mathbb{R}^d)^*$, *i.e.* any linear functional on $\mathbb{R}^d$ and define $\eta_i := \varphi(\mathbf{b}_i)$ for $1 \leq i \leq d$. For any $x = \sum_{i=1}^d \lambda_i \mathbf{b}_i$ we get

$$\varphi(\mathbf{x}) \;=\; \sum \lambda_i \varphi(\mathbf{b}_i) \;=\; \sum \lambda_i \eta_i \;=\; \sum \mathbf{b}_i^*(\mathbf{x})\eta_i \;=\; \left(\sum \eta_i \mathbf{b}_i^*\right)(\mathbf{x})$$

by linearity. Hence, any linear functional is in the span of the $\mathbf{b}_i^*$.

2.26. If φ is a lattice functional, then the η_i in Problem 2.25 are integral.

2.27. Let $\mathbf{b}_1, \dots, vb_d$ be a basis of Λ with dual basis $\mathbf{b}_1^*, \dots, \mathbf{b}_d^*$. If B is the matrix whose columns are the $\mathbf{b}_i$ and B^* the one with columns $\mathbf{b}_i^*$. Then

$$\det\, \Lambda \;=\; \det\, B \qquad\qquad \det\, \Lambda^\star \;=\; \det\, B^*$$

and

$$\det(\Lambda) \cdot \det(\Lambda^\star) \;=\; \det(B) \cdot \det(B^*) \;=\; \det(B^t) \cdot \det(B^*) \;=\; \det(B^t B^*) \;=\; \det(\mathrm{I}) \;=\; 1$$

as $\mathbf{b}_i^t \mathbf{b}_j$ is 1 if $i = j$ and 0 otherwise.

2.29.

B.2. Solutions for *Geometry of Numbers* (*Chapter 3*)

3.1. ▷ Let F be the set of a choice of a lattice point in the relative interior of a facet, if there is one in the facet. Note that we take at most one form each relative interior.

We assign to each the parity vector (*e.g.* the vector $(1,1,0)$ for the lattice point $(7,5,4)$). If we have more than 2^d vectors in F, then there are $\mathbf{x},\mathbf{y}\in F$ with $\mathbf{x}\neq\mathbf{y}$ that have the same parity. Then also $1/2(\mathbf{x}+\mathbf{y})$ is a lattice point, and in the interior of S.

This is a contradiction, so we have at most 2^d elements in F. In fact, as $\{0\}$ is in the interior of K there cannot be a lattice point in the boundary with only even coordinates, so $|F|\leq 2^d-1$.

Hence, we have at most $2(2^d-1)$ facets for K.

▷ For any lattice point $\mathbf{x}$ we choose a halfspace $H_\mathbf{x}$ containing $\mathbf{x}\cup K$. Let $S_\mathbf{x}:=H_\mathbf{x}\cap -H_\mathbf{x}$. Then $K\subseteq S_\mathbf{x}$. Let

$$S:=\bigcap_{\mathbf{x}\in\Lambda\setminus\{0\}} S_\mathbf{x}\,.$$

Then S is a centrally symmetric set of volume $\operatorname{nvol}_{\mathbb{Z}^d}(S)\geq\operatorname{nvol}_{\mathbb{Z}^d}(K)=2^d$. Also $\operatorname{int}S\cap\Lambda=\{0\}$, as all $\mathbf{x}\in\Lambda\setminus\{0\}$ are on one of the hyperplanes defining S.

If $S\neq K$, then there is $\mathbf{y}\in S\setminus K$, and thus also $-vy\in S\setminus K$. Let $K':=\mathsf{conv}(K\cup\{\pm\mathbf{y}\})$ So $K'\subseteq S$, $K'\cap\Lambda=\{0\}$ and $\operatorname{nvol}_{\mathbb{Z}^d}(K')>\operatorname{nvol}_{\mathbb{Z}^d}(K)=2^d$. Hence, by Minkowski's First Theorem (Corollary 3.3) K' contains a nonzero lattice point. This is a contradiction, so $S=K$.

We need to show that a finite number of the pairs of inequalities $H_\mathbf{x}$ and $-H_\mathbf{x}$ suffice. Let $F\subseteq\Lambda$ be the set of those $\mathbf{x}$ such that the hyperplane defined by $H_\mathbf{x}$ is irredundant for S. Let $H_\mathbf{y}$ be any hyperplane for $\mathbf{y}\in F$ such that $H_\mathbf{y}\cap K$ does not contain a lattice point in its relative interior. Then S is a strict subset of

$$S':=\bigcap_{\mathbf{x}\in F\setminus\{\mathbf{y}\}} S_\mathbf{x}$$

and $\operatorname{int}S'\cap\Lambda=\{0\}$. But $\operatorname{nvol}_{\mathbb{Z}^d}(\operatorname{int}S')>2^n\det\Lambda$, so $\operatorname{int}S'$ contains a nonzero lattice point by Minkowski's First Theorem (Corollary 3.3). This is a contradiction, so each of the hyperplanes given by $H_\mathbf{x}$ and $-H_\mathbf{x}$ for $\mathbf{x}\in F$ contain a lattice point in its relative interior.

The number of lattice points in the relative interiors is finite by the same argument as in the first part. So K is a polytope.

3.2. Consider the map $\pi:\mathbb{Z}^d\to(\mathbb{Z}/3\mathbb{Z})^d$ given by assigning each coordinate its congruence class modulo 3. This is a homomorphism, so that $\pi(\mathbf{x}\pm\mathbf{y})=\pi(\mathbf{x})\pm\pi(\mathbf{y})$.

If, for lattice points $\mathbf{x},\mathbf{y}\in K$ the images satisfy $\pi(\mathbf{x})=\pi(\mathbf{y})$, then $\pi(\mathbf{x}-\mathbf{y})=0$, so $\mathbf{z}:=\frac{1}{3}(\mathbf{x}-\mathbf{y})\in\mathbb{Z}^d$ and $\mathbf{z}$ is in the interior of the triangle spanned by 0, $\mathbf{x}$ and $-\mathbf{y}$. As K is centrally symmetric we know that $-\mathbf{y}\in K$, so $\mathbf{z}$ is contained in $\frac{2}{3}K$.

So, by assumption, $\mathbf{z}=0$ and $\mathbf{x}=\mathbf{y}$. Hence, π is injective and we have at most 3^d lattice points in K.

In fact, up to unimodular transformations the standard cube $[-1,1]^d$ is the only centrally-symmetric lattice polytope with $\operatorname{int}(K)\cap\Lambda=\{0\}$ and $|K\cap\Lambda|=3^d$.[1]

3.3. Consider the map $\pi:\mathbb{Z}^d\to(\mathbb{Z}/2\mathbb{Z})^d$. Assume that there is a boundary lattice point $\mathbf{x}$ in the boundary of K with $\pi(\mathbf{x})=0$. Then $\frac{1}{2}\mathbf{v}\in\mathbb{Z}^d$. Hence, we would have a nonzero lattice point in the interior, which is a contradiction. So 0 is not in the image of π.

Let $p\in(\mathbb{Z}/2\mathbb{Z})^d\setminus\{0\}$. Assume that there are lattice points $\mathbf{x},\mathbf{y}\in K$ with $\mathbf{y}\neq\pm\mathbf{x}$ (note that $-\mathbf{x}\in K$) such that $\pi(\mathbf{x})=\pi(\mathbf{y})=p$. Then $\pi(\mathbf{x}-\mathbf{y}))=$, so $\frac{1}{2}(\mathbf{x}-\mathbf{y})\in\mathbb{Z}^d$. As $-\mathbf{x}\in K$ the point $\mathbf{z}$ is in the interior of K, a contradiction, as $\pi(\mathbf{x})\neq 0$. So at most 2 points map to p, and the bound follows.

3.4. We may assume that $\Lambda=\mathbb{Z}^d$ and $\operatorname{vol}K>k\cdot 2^d$. By Generalized Blichfeldt's Theorem (Lemma 3.4) we can find $k+1$ pairwise different points $\mathbf{x}_0,\ldots,\mathbf{x}_k\in\frac{1}{2}K$ such that $\mathbf{x}_i-\mathbf{x}_j$ is a lattice point for all i,j.

Assume that $\|\mathbf{x}_0\|\leq\|\mathbf{x}_i\|$ for all i and let $\mathbf{y}_i:=\mathbf{x}_0-\mathbf{x}_i$ for $1\leq i\leq k$. Then $\mathbf{y}_i\neq\mathbf{y}_j$ for $i\neq j$ and

[1] Draisma, McAllister, and Nill, *Lattice width directions and Minkowski's 3^d-theorem*.

$\mathbf{y}_i \in K \cap \mathbb{Z}^d \setminus \{0\}$. By our choice of $\mathbf{x}_1$ we also know that

$$\langle \mathbf{x}_i, \mathbf{y}_i \rangle \;=\; \langle \mathbf{x}_i, \mathbf{x}_0 - \mathbf{x}_i \rangle \;=\; -\|\mathbf{x}_i\|^2 + \langle \mathbf{x}_i, \mathbf{x}_0 \rangle \;\leq\; -\|\mathbf{x}_i\|^2 + \|\mathbf{x}_i\|\,\|\mathbf{x}_0\| \;\leq\; 0$$

with equality in the second but last inequality if and only if $\mathbf{x}_i$ is a scalar multiple of $\mathbf{x}_0$. The last inequality follows from our choice of $\mathbf{x}_0$. If $\mathbf{x}_0$ and $\mathbf{x}_i$ are linearly independent, then they cannot have the same length, so in fact we can conclude

$$\langle \mathbf{x}_i, \mathbf{y}_i \rangle \;<\; 0\,.$$

Hence, all points $\pm\mathbf{y}_i$ are different, and together with 0 we obtain the required number of points.

They need not be linearly independent. For this, consider *e.g.* the rectangle with vertices $(\pm 12, \pm 2k)$ with volume $4k = k \cdot 2^2$.

3.5. Let ω_d be the area of the $(d-1)$-dimensional unit sphere $S_{d-1} := \{\, \mathbf{x} \;:\; \|\mathbf{x}\| = 1 \,\}$, and $\mathrm{V_d}$ the volume of the unit ball.

Consider any 1-dimensional function $f : (0,\infty) \to \mathbb{R}$ defined on the positive axis. We consider this as a radially symmetric function on $\mathbb{R}^n$, and we want to integrate $f(|\mathbf{x}|)$ over $\mathbb{R}^d$. Writing this in spherical coordinates we obtain

$$\int_{\mathbb{R}^d} f(\|\mathbf{x}\|)\,\mathrm{d}\mathbf{x} \;=\; \omega_d \cdot \int_0^\infty f(r) r^{d-1} \mathrm{dr}\,, \tag{B.6}$$

where ω_d is the area of the unit sphere, as long as the integrals are well defined. Integrating the function f that is 1 for $r \in (0,1)$ and 0 for $r \geq 1$ we obtain

$$\mathrm{V_d} \;=\; d\omega_d$$

Hence, it would suffice to compute ω_d in order to obtain $\mathrm{V_d}$. We can obtain ω_d if we can find a function f for which we can compute both integrals in (B.6) explicitely.

A good candidate for this is certainly a function f for the computation of the right hand side is certainly one that splits into a product of functions that each depend on one coordinate only,

$$f(\|\mathbf{x}\|) \;=\; \prod_{i=1}^d \varphi_i(x_i)\,.$$

Radial symmetry of f implies that all φ_i are the same. A good choice for such a function is the Gaussian $\varphi(x) = e^{-x^2}$, because we know the Guassian integral

$$\int_{-\infty}^{\infty} e^{-x^2} \mathrm{dx} \;=\; \sqrt{\pi}\,.$$

This gives

$$\int_{\mathbb{R}^d} f(|\mathbf{x}|)\,\mathrm{d}\mathbf{x} \;=\; \int_{\mathbb{R}^d} e^{-\|\mathbf{x}\|^2}\,\mathrm{d}\mathbf{x} \;=\; \int_{\mathbb{R}^d} \prod_{i=1}^d e^{|x_i|^2}\,\mathrm{d}\mathbf{x} \;=\; \pi^{d/2}$$

Using the substritution $t = r^2$ we compute the right hand side of (B.6) as

$$\int_0^\infty e^{-r^2} r^{d-1} \mathrm{dr} \;=\; \frac{1}{2}\int_0^\infty t^{d-22} e^{-t} \mathrm{dt} \;=\; \frac{1}{2}\Gamma\left(\frac{d}{2}\right)\,,$$

where Γ is the Γ-function. From this we compute

$$\omega_d \;:=\; \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}\,.$$

and thus, using $\frac{d}{2}\cdot\Gamma\left(\frac{d}{2}\right) = \Gamma\left(\frac{d}{2}+1\right)$,

$$\mathrm{V_d} = \frac{\omega_d}{d} = \frac{2\pi^{d/2}}{d\cdot\Gamma\left(\frac{d}{2}\right)} = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$$

We need the following two relations for the Γ-function. For each integer n

$$\Gamma(n) = (n-1)! \qquad \Gamma\left(n+\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^n}\prod_{0\le 2i\le d}(d-2i)\,.$$

Using these we obtain

$$\mathrm{V_d} = \frac{\pi^{\lfloor d/2\rfloor}2^{\lceil d/2\rceil}}{\prod_{0\le 2i\le d}(d-2i)}\,.$$

3.6.

(i) This is clear in dimension 1.

(ii) In dimension 2 let $\mathbf{b}_1$, $\mathbf{b}_2$ be linearly independent lattice vectors that realize λ_1 and λ_2. They span a sublattice Γ of Λ. Assume they differ, then there is a lattice point $\mathbf{x}$ contained in the fundamental parallepiped spanned by $\mathbf{b}_1$ and $\mathbf{b}_2$, *i.e.* there are $0\le\mu_1,\mu_2<1$ with $\mathbf{x}=\mu_1\mathbf{b}_1+\mu_2\mathbf{b}_2$. If $\mu_1+\mu_2>1$, then

$$\mathbf{y} := \mathbf{b}_1+\mathbf{b}_2-\mathbf{x} = (1-\mu_1)\mathbf{b}_1+(1-\mu_2)\mathbf{b}_2$$

is another lattice point. We set $\eta_1 := 1-\mu_1$ and $\eta_2 := 1-\mu_2$. Then $0\le\eta_1,\eta_2$ and $\eta_1+\eta_2\le 1$. So either $\mathbf{x}$ or $\mathbf{y}$ is contained in the triangle spanned by 0 and $\mathbf{b}_1,\mathbf{b}_2$, say $\mathbf{x}$. Thus, $\mathbf{x}$ is shorter than at least one of $\mathbf{b}_1$ and $\mathbf{b}_2$. This must be $\mathbf{b}_2$ as $\mathbf{b}_1$ is a shortest lattice vector. But then $\|\mathbf{x}\|<\lambda_2$ and $\mathbf{b}_1$, $\mathbf{x}$ are linearly independent. This is a contradiction to the coice of $\mathbf{b}_2$.

(iii) Let $\mathbf{b}_i := 2\mathbf{e}_i$ for $1\le i\le d-1$ and $\mathbf{b}_d := \mathbf{e}_1+\cdots+\mathbf{e}_d$. Then $\mathbf{y}=2\mathbf{b}_d-\mathbf{b}_{d-1}-\cdots-\mathbf{b}_1=2\mathbf{e}_d$. For $d\ge 5$ the vector $\mathbf{b}_1$ is a shortest lattice vector (consider any integer linear combination of the $\mathbf{b}_i$ and show that it can have length at most 2 if and only if one of the coefficients of $\mathbf{b}_j$ for $j\le d-1$ is 1 and all others are 0). Hence, $\lambda_1=2$ and thus $\lambda_i\ge 2$ for $1\le i\le d$. But $\mathbf{b}_j$ for $j\le d-1$ and $\mathbf{y}$ have length 2 and are linearly independent, so $\lambda_i=2$ for all i.

But the lattice spanned by $\mathbf{b}_1,\ldots,\mathbf{b}_{d-1},\mathbf{y}$ is $(2\mathbb{Z})^d$ and Λ is a sublattice of index 2 in this lattice.

3.7. We prove this by induction. The claim is trivial for $d=1$, so we assume we know the result for dimensions up to $d-1$.

The $\mathbf{b}_i$ are a basis, so we can write $\mathbf{v}$ as

$$\mathbf{v} = \sum_{i=1}^{d}\xi_i\mathbf{b}_i$$

for some $\xi_i\in\mathbb{R}$. Choose integers a_i such that $a_i-\xi_i|\le 1/2$ for $1-\le i\le d$. Let $L := \mathrm{lin}(\mathbf{b}_1,\ldots,\mathbf{b}_{d-1})$ be the subspace spanned by the first $d-1$ basis vectors, Λ' the lattice spanned by $\mathbf{b}_1,\ldots,\mathbf{b}_{d-1}$ and let $\mathbf{v}'$, $\mathbf{b}_d'$ be the orthogonal projections of $\mathbf{v}$ and $\mathbf{b}_d$ onto L. Then

$$\mathbf{v}' = \sum_{i=1}^{d-1}\xi_i\mathbf{b}_i + \xi_d\mathbf{b}_d'$$

and $\mathbf{v}-\mathbf{v}'=\xi_d(\mathbf{b}_d-\mathbf{b}_d')$.

Note that $\mathbf{v}'-a_d\mathbf{b}_d'\in L$. By assumption, there is some $\mathbf{u}'\in\Lambda'$ such that

$$\left\|(\mathbf{v}'-a_d\mathbf{b}_d')-\mathbf{u}'\right\| - \min_{\mathbf{w}'\in\Lambda'}\left\|(\mathbf{v}'-a_d\mathbf{b}_d')-\mathbf{w}'\right\| \le \frac{\sqrt{d-1}}{2}\mu\,.$$

Now $\mathbf{b}'_d$ and $(\mathbf{b}_d - \mathbf{b}'_d)$ are orthogonal by construction, so

$$\left\|\mathbf{b}'_d\right\|^2 \;+\; \left\|\mathbf{b}_d - \mathbf{b}'_d\right\|^2 \;=\; \|\mathbf{b}_d\|^2$$

and thus

$$\left\|\mathbf{b}_d - \mathbf{b}'_d\right\|^2 \;=\; \|\mathbf{b}_d\|^2 \;-\; \left\|\mathbf{b}'_d\right\|^2 \;\le\; \|\mathbf{b}_d\|^2 \;\le\; \mu^2\,.$$

So

$$\left\|\mathbf{v} - \mathbf{v}' - a_d(\mathbf{b}_d - \mathbf{b}'_d)\right\| \;=\; |a_d - \xi_d|\left\|\mathbf{b}_d - \mathbf{b}'_d\right\| \;\le\; \frac{1}{2}\mu\,. \tag{B.7}$$

Further, $(\mathbf{v}' - a_d\mathbf{b}'_d) - \mathbf{u}' \in L$ and thus orthognal to $\mathbf{v} - \mathbf{v}' - a_d(\mathbf{b}_d - \mathbf{b}'_d)$, so

$$\begin{aligned}\left\|\mathbf{v} - a_d\mathbf{b}_d - \mathbf{u}'\right\|^2 \;&=\; \left\|\mathbf{v} - \mathbf{v}' - a_d(\mathbf{b}_d - \mathbf{b}'_d)\right\|^2 \;+\; \left\|(\mathbf{v}' - a_d\mathbf{b}'_d) - \mathbf{u}'\right\|^2 \\ &\le\; \frac{1}{4}\mu^2 \;+\; \frac{d-1}{4}\mu^2 \;=\; \frac{d}{4}\mu^2\,.\end{aligned} \tag{B.8}$$

Taking the square root gives the desired inequality for the lattice point $\bar{\mathbf{u}} := a_d\mathbf{b}_d + \mathbf{u}'$
In case of equality we must have $\|\bar{\mathbf{u}} - \mathbf{v}\| = \min_{\mathbf{u}} \in \Lambda \|\mathbf{u} - \mathbf{v}\|$ and equality in (B.8), hence also

$$\left\|\mathbf{v} - \mathbf{v}' - a_d(\mathbf{b}_d - \mathbf{b}'_d)\right\|^2 \;=\; \frac{1}{4}\mu \qquad\qquad \left\|(\mathbf{v}' - a_d\mathbf{b}'_d) - \mathbf{u}'\right\|^2 \;=\; \frac{d-1}{4}\mu\,.$$

By assumption the vectors $\mathbf{b}_1, \ldots, vb_{d-1}$ are pairwise orthogonal and $\|\mathbf{b}_i\| = \mu$ for $1 \le i \le d-1$. We conclude $\|\mathbf{b}_d\| = \mu$ and $\|\mathbf{b}'_d\| = 0$, so vb_d is orthogonal to all other $\mathbf{b}_i$. Again by assumption, the coefficients ξ_i for $1 \le i \le d-1$ are half-integers, and from the equality in (B.7) we conclude that also ξ_d is a half-integer.

Conversely, with these conditions we clearly have equality.

3.8. We use induction. This is clear for $d = 1$, so assume we know the result for dimension $d - 1 \ge 1$.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be linearly independent lattice vectors with $\|\mathbf{b}_i\| = \lambda_i$. Let $L := \mathrm{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_{d-1})$. This is a $(d-1)$-dimensional space and $\Lambda' := L \cap \Lambda$ is a lattice of rank $d-1$ in L (since it contains $\mathbf{b}_1, \ldots' \mathbf{b}_{d-1}$). By assumption, Λ' has a basis $\mathbf{c}_1, \ldots, \mathbf{c}_{d-1}$ of lattices vectors such that $\|\mathbf{c}_i\| = \lambda_i$. With $\mathbf{c}_d = \mathbf{b}_d$ we have another (vector space) basis $\mathbf{c}_1, \ldots, \mathbf{c}_d$ of lattice vectors. Let Γ be the lattice spanned by these vectors.

If the claim fails, then there is $\mathbf{v} \in \Lambda \setminus \Gamma$. By Problem 3.7 we can find $\mathbf{u} \in \Gamma$ such that

$$\|vv - \mathbf{u}\| \;\le\; \frac{\sqrt{d}}{2}\lambda_d\,.$$

Then $\mathbf{v} - \mathbf{u} \in \Lambda \setminus \Gamma$, and, as Λ' is contained in Γ, also $\mathbf{v} - \mathbf{u} \notin L$. This implies that $\mathbf{c}_1, \ldots, \mathbf{c}_{d-1}, \mathbf{v} - \mathbf{u}$ are linearly independent lattice vectors, and

$$\lambda_n \;\le\; \|\mathbf{v} - \mathbf{u}\| \;\le\; \frac{\sqrt{d}}{2}\,,$$

where the first inequality follows as otherwise $\mathbf{v} - \mathbf{u}$ would have been part of the original selection of vectors.

This implies $d \ge 4$, so for $d = 2, 3$ we have proved the claim. For $d = 4$ we have equality in Problem 3.7. Hence, the $\mathbf{c}_i$ are pairwise orthogonal, $\lambda_1 = \cdots = \lambda_d$, and any $\mathbf{v} \in \Lambda \setminus \Gamma$ has a representation

$$\mathbf{v} \;=\; \sum(a_i + \frac{1}{2})\mathbf{c}_i$$

for some integers $a_1, \ldots, a_d$. This shows that Λ is generated by $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \frac{1}{2}(\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 + \mathbf{c}_4)$ and the latter has length λ_4. This proves the claim.

3.9. Let $\mathbf{v}_i$, $1 \le i \le d$ be the vectors of Proposition 3.9. Then

$$\lambda_i \;=\; \min\left(\|\mathbf{v}\| \;:\; \mathbf{v} \in \Lambda \setminus \operatorname{lin}\left(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\right)\right) ,$$

with $\operatorname{lin}(\varnothing) = \{0\}$. We can thus take any lattice basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ with

$$\mathbf{v}_i \;\in\; \mathbb{Z}\mathbf{b}_1 \;+\; \cdots \;+\; \mathbb{Z}\mathbf{b}_i \,.$$

for $1 \le i \le d$.

3.10. C is a skew cross polytope with vertices

$$\pm\frac{1}{\lambda_1}\mathbf{v}_1, \ldots, \pm\frac{1}{\lambda_d}\mathbf{v}_d \,.$$

Consider the simplex

$$S \;:=\; \mathsf{conv}\left(\frac{1}{\lambda_1}\mathbf{v}_1, \ldots, \frac{1}{\lambda_d}\mathbf{v}_d\right) .$$

The volume of S is

$$\operatorname{vol} S \;=\; \frac{1}{d!\lambda_1 \cdots \lambda_d} \det \Gamma$$

Hence, the volume of C is

$$\operatorname{vol} C \;=\; \frac{2^d}{d!\lambda_1 \cdots \lambda_d} \det \Gamma \,.$$

Now consider the lower bound in Minkowski's Second Theorem (Theorem 3.11). By construction, $C \subseteq K$. The lattice Γ is a sublattice of Λ, so $\det \Lambda \le \det \Gamma$ and we obtain

$$\begin{aligned}
\lambda_1 \cdots \lambda_d \operatorname{vol} K \;&\ge\; \lambda_1 \cdots \lambda_d \operatorname{vol} C \\
&\ge\; \lambda_1 \cdots \lambda_d \cdot 2^d \frac{1}{d!} \frac{\det \Gamma}{\lambda_1 \cdots \lambda_d} \\
&\ge\; 2^d \frac{1}{d!} \det B \;\ge\; 2^d \frac{1}{d!} \det \Lambda \,.
\end{aligned}$$

The right and left hand side of this chain of inequalities are the lower bound in the theorem.

3.12. Choose $q < p$ such that $q^2 \equiv 1 \mod p$ using Euler's criterion for -1 and $(-1)^{\frac{1}{2}((4k+1)-1)} = (-1)^{2k} = 1 \equiv 1 \mod p$.
Now consider the lattice Λ spanned by

$$\mathbf{b}_1 \;:=\; \left[\begin{smallmatrix} 1 \\ q \end{smallmatrix}\right] \qquad\qquad \mathbf{b}_2 \;:=\; \left[\begin{smallmatrix} 0 \\ p \end{smallmatrix}\right] .$$

with determinant $\det \Lambda \;=\; p$.
Let $\mathcal{B}_{\mathrm{d}}\, r0$ be the open ball around the origin with radius $r = \sqrt{2}p$. Then K is centrally symmetric, convex, and

$$\operatorname{vol} K \;=\; \pi \cdot r^2 \;>\; 4 \cdot p \;=\; 2^2 \det \Lambda \,,$$

so by Minkowski's First Theorem (Corollary 3.3) there exists a lattice point

$$\mathbf{x} \;:=\; \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \;=\; \mu_1\mathbf{b}_1 \;+\; \mu_2\mathbf{b}_2 \;\neq\; 0 \,.$$

We compute

$$\begin{aligned} a^2 + b^2 &= \mu_1^2 + (\mu_1 q + \mu_2 p)^2 \\ &\equiv \mu^2 + \mu^2 q^2 \mod p \\ &\equiv \mu^2 (q^2 + 1) \mod p \\ &\equiv \mu^2 (11) \mod p \\ &\equiv 0 \mod p\,. \end{aligned}$$

Thus, p divides $a^2 + b^2$. Further, $\mathbf{x} \in K$ implies $a^2 + b^2 < 2p$, so the only choice left is $a^2 + b^2 = p$.

3.14. The open balls in the definition of the packing radius contain exactly one lattice point, their center $\mathbf{v}$, and they are centrally symmetric around $\mathbf{v}$. Translating $\mathbf{v}$ into the origin results in a centrally symmetric convex body containing only 0 in the interior. By Minkowski's First Theorem (Corollary 3.3) the volume is bounded. Hence, the radius must be finite.

Now let $\mathbf{v}$ be a shortest lattice vector of length λ_1. Then $\varrho(\Lambda) \leq {}^1\!/_2 \lambda_1$ as otherwise the balls around 0 and $\mathbf{v}$ intersect.

On the other hand, if $\varrho(\Lambda) < \lambda_1$, then there are $\mathbf{u}, \mathbf{v} \in \Lambda$ with $\mathbf{u} \neq \mathbf{v}$ such that the open balls of radius ${}^1\!/_2 \lambda_1$ around $\mathbf{u}$ and $\mathbf{v}$ intersect. We can assume that $\mathbf{u}$ is the origin, so $vv \in \Lambda \setminus \{0\}$. Then $\mathbf{v} \in \mathcal{B}^\circ_{\lambda_1}(0)$, so $\|\mathbf{v}\| < \lambda_1$. This is a contradiction.

3.16. The centrally symmetric cube C with sidelength $2(\det \Lambda)^{1/d}$ is compact and has volume $2^n \det \Lambda$. Minkowski's First Theorem (Corollary 3.3) implies that there is a nonzero lattice point $\mathbf{x}$ in C. All its coordinates are bounded by $(\det \Lambda)^{1/d}$.

3.17. The lattice points of Λ_0 are a subset of the lattice points of Λ. Hence, if balls of radius r around points of Λ do not intersect, then they also do not intersect if we only consider those around the points of Λ_0. This proves the first inequality.

For the second inequality observe that the packing radius of $\mu\Lambda$ is $\mu\, \varrho(\Lambda)$ for any lattice, and that $|\Lambda/\Lambda_0|\Lambda$ is a sublattice of Λ_0 by Problem 2.22. Hence, we can use the first inequality.

3.20. Let $\mathbf{b}_1, \ldots, \mathbf{b}_d$ be a lattice basis and $\mathbf{x} \in \mathbb{R}^n$. Then there are $\eta_1, \ldots, \eta_d$ such that

$$\mathbf{x} = \sum_{i=1}^{d} \eta_i \mathbf{b}_i$$

We set $\mathbf{y} := \sum_{i=1}^{d} \{\eta_i\} \mathbf{b}_i$. Then $\mathbf{x} - \mathbf{y}$ is a lattice point and $\mathrm{d}(\mathbf{x}, \Lambda) = \mathrm{d}(\mathbf{y}, \Lambda)$. The point $\mathbf{y}$ is contained in the fundamental parallelpiped. This is a bounded set, so there is a finite $\nu > 0$ such that $\|\mathbf{u}\| < \nu$ for any point in this set. So in particular $\mathrm{d}(\mathbf{x}, \Lambda) = \mathrm{d}(\mathbf{y}, \Lambda) < \nu$.

The argument also shows that it suffices to consider $\mathrm{d}(\mathbf{x}, \Lambda)$ for points in the closure of the fundamental parallelpiped. This is a compact set, so the maximum is attained.

3.23. The packing radius is at least $\frac{1}{2}\lambda_d$ by Lemma 3.22, so there must be a point in $\mathbb{R}^d$ that has at least distance $\frac{1}{2}\lambda_d$ from any lattice point.

3.25. We know $\lambda_1 = 2\varrho$ and $2\mu^* \geq \lambda_d *$, so $\lambda_1 \lambda_d^* \geq 4\varrho\mu^d \geq 1$.

3.28. As K is full dimensional there is some interior point $\mathbf{x} \in K$. Let $\varepsilon > 0$ so that $\mathcal{B}_\varepsilon(\mathbf{x}) \subseteq K$. Then the width of K *w.r.t.* any lattice vector $\mathbf{a} \in \Lambda^\star$ is at least $2\varepsilon \|\mathbf{a}\|$.

Choosing any nonzero $\mathbf{a} \in \Lambda^\star$ we obtain an upper bound W for the lattice width. Hence, any lattice vector such that $\mathrm{width}(K; \mathbf{a}) \leq W$ has length at most $\frac{W}{2\varepsilon}$. This is a compact set. So the infimum exists and is actually a minimum attained by some $\mathbf{a} \in \Lambda^\star$.

3.29. Let $\mathbf{v} \in \Lambda^\star$ be a vector of length λ_1 in the norm defined by $(K - K)^\star$. Then

$$\mathbf{v}(\mathbf{x}) \leq \lambda_1 \qquad \text{for all } \mathbf{x} \in K - K$$

Now

$$\mathrm{width}_\Lambda(K) = \min_{\mathbf{a} \in \Lambda^\star} \left(\max_{\mathbf{x} \in K} \mathbf{a}(\mathbf{x}) - \min_{\mathbf{y} \in K} \mathbf{a}(\mathbf{y}) \right)$$

Let $\mathbf{x}_0, \mathbf{y}_0 \in K$ be the vectors realising max and min for $\mathbf{v}$. Then $\mathbf{x}_0 - \mathbf{y}_0 \in K - K$ and

$$\lambda_1 \geq \mathbf{v}(\mathbf{x}_0 - \mathbf{y}_0) \ = \mathbf{v}(\mathbf{x}_0) \ - \ \mathbf{v}(\mathbf{y}_0) \ = \ \max_{\mathbf{x}\in K}(\mathbf{v}(\mathbf{x})) - \min_{\mathbf{y}\in K}(\mathbf{v}(\mathbf{y}))\,.$$

Hence, the width is bounded by λ_1. Conversely, if $\mathbf{w}$ produces a smaller width, then $\mathbf{w}$ is a vector in $\Lambda^\star$ with length smaller than λ_1 in the norm of $(K-K)^\star$.

3.30. The claim about lattice polytopes is obvious.

For the second claim observe that $\operatorname{width}_\lambda(()_i C_i) = \lambda_i \operatorname{width}_C(()_i)$, so we can assume $\lambda_i = 1$ for all i.

Let $\mathbf{v}_i \ \Lambda_i$ be a lattice direction for which $\operatorname{width}_C(()_i)$ is obtained. Then

$$\operatorname{width}_\Lambda(C) \ \leq \ \operatorname{width}_{(0,\ldots,0,\mathbf{v}_i,0,'ldots,0)}(C) \ = \ \operatorname{width}_{\mathbf{v}_i}(C_i) \ = \ \operatorname{width}_{\Lambda_i}(C_i)\,,$$

so

$$\operatorname{width}_{\Lambda_i}(C) \ \leq \ \min_i(\operatorname{width}_{\Lambda_i}(C_i))\,.$$

For the opposite direction we take and lattice functional

$$\mathbf{w} \ = \ (\mathbf{w}_1, \ldots, \mathbf{w}_m) \ \in \ \Lambda \setminus \{0\} \ = \ \bigoplus_i \Lambda_i \setminus \{0\}\,.$$

We want to show that $\operatorname{width}_{\mathbf{w}}(C) \ \geq \ \operatorname{width}_{\Lambda_i}(C_i)$ for some i. For this, let us choose any i with $\mathbf{w}_i \neq 0$. We compute

$$\begin{aligned}
\operatorname{width}_{\mathbf{w}}(C) \ &= \ \max_{\mathbf{x},\mathbf{y}\in C}(|\langle\, \mathbf{w}, \mathbf{x}\,\rangle \ \langle\, \mathbf{w}, \mathbf{y}\,\rangle|) \\
&\geq \ \max_{\mathbf{x}_i,\mathbf{y}_i\in C_i}(|\langle\, \mathbf{w}, (0,\ldots,0,\mathbf{x}_i,0,\ldots,0)\,\rangle \ \langle\, \mathbf{w}, (0,\ldots,0,\mathbf{y}_i,0,...,0)\,\rangle|) \\
&= \ \operatorname{width}_{\mathbf{w}_i}(C_i) \ \geq \ \operatorname{width}_\Lambda(C_i) \ \geq \ \min \operatorname{width}_\Lambda(C_i)\,.
\end{aligned}$$

For the last part assume that $\operatorname{int} C \cap \Lambda \neq \varnothing$ and let $\mathbf{x}$ be a lattice point in the interior of C. We can write $\mathbf{x}$ as

$$\mathbf{x} \ = \ (\mu_1\mathbf{x}_1, \ldots, \mu_m\mathbf{x}_m) \qquad\qquad \text{for } \mu_i \geq 0,\ \mathbf{x}_i \in \operatorname{int} C_i \text{ and } \sum_i \mu_i \ = \ 1\,.$$

As $\mathbf{x} \in \Lambda$ we have $\lambda_i\mu_i\mathbf{x}_i \in \Lambda_i$. As $\operatorname{int} C_i$ does not contain a lattice point we see that $\lambda_i\mu_i > 1$. But this implies

$$\sum_i \tfrac{1}{\lambda_i} \ > \ \sum_i \mu_i \ = \ 1\,.$$

which contradicts our assumption.

B.3. Solutions for *The Shortest Vector Problem* (*Chapter 4*)

4.2. By the definition of a weakly reduced basis we have

$$\mathbf{b}_k = \mathbf{w}_k + \sum_{k=1}^{j-1} \lambda_{jk}\mathbf{w}_j$$

for coefficients $|\lambda_{jk}| \leq 1/2$. The $\mathbf{w}_j$ are pairwise orthogonal, so taking the norm implies

$$\|\mathbf{b}_k\|^2 \;=\; \|\mathbf{w}_k\|^2 + \sum_{j=1}^{k-1} \lambda_{jk}^2 \|\mathbf{w}_j\|^2 \;\leq\; \|\mathbf{w}_k\|^2 + \frac{1}{4}\sum_{j=1}^{k-1} \|\mathbf{w}_j\|^2 \,.$$

This proves the first inequality.
For the second note that using (4.8) we get

$$\begin{aligned}\|\mathbf{w}_k\|^2 + \frac{1}{4}\sum_{j=1}^{k-1} \|\mathbf{w}_j\|^2 \;&\leq\; \|\mathbf{w}_k\| \left(1 + \frac{1}{4}\sum_{j=1}^{k-1} 2^{k-j}\right) \\ &=\; \|\mathbf{w}_k\|^2 \left(1 + \frac{1}{4}\cdot 2^{k-2}\right) \;=\; \|\mathbf{w}_k\|^2 \left(\frac{1}{2} + 2^{k-2}\right) .\end{aligned}$$

For the last we observe that $\left(\frac{1}{2} + 2^{k-2}\right) \leq 2^{k-1}$. So the previous inequality implies

$$\|\mathbf{b}_k\|^2 \;\leq 2^{k-1} \|\mathbf{w}_k\|^2 \,.$$

Again using (4.8) and taking square roots gives the result.

4.3. Let $\mathbf{x} = \sum_{i=1}^d \mu_i \mathbf{b}_i$ for integer coefficients $\mu_i \in \mathbb{Z}$. Let m be the largest index so that $\mu_i \neq 0$. By construction, $\langle\, \mathbf{b}_i, \mathbf{w}_m \,\rangle \;=\; 0$ for all $i < m$, so

$$\begin{aligned}\|\mathbf{w}_m\|^2 \;&\leq\; |\mu_m| \, \|\mathbf{w}_m\|^2 \;=\; |\mu_m| |\langle\, \mathbf{b}_m, \mathbf{w}_m \,\rangle| \;=\; |\langle \sum_{i=1}^d \mu_i \mathbf{b}_i, \mathbf{w}_m \,\rangle| \\ &\leq\; \left\| \sum_{i=1}^d \mu_i \mathbf{b}_i \right\| \cdot \|\mathbf{w}_m\| \;=\; \|\mathbf{x}\| \, \|\mathbf{w}_m\| \,.\end{aligned}$$

We can devide by $\|\mathbf{w}_k\|$ to get $\|\mathbf{w}_k\| \leq \|\mathbf{x}\|$. Using this for $\mathbf{x} = \mathbf{v}_j$ and $m = j_k$ gives the first claim.
For the second claim note that $k \leq j_k$, as otherwise $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \operatorname{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_{k-1})$, contrary to the assuption that these vectors are linearly independent.
Now we use Problem 4.2 to bound

$$\|\mathbf{b}_j\| \;\leq\; 2^{(j_k - 1)/2} \, \|\mathbf{w}_{j_k}\| \;\leq\; 2^{(d-1)/2} \, \|\mathbf{v}_j\| \;\leq\; 2^{(d-1)/2} \lambda_i \,.$$

The upper bound of the third claim follows from the second. For the lower bound observe that $\lambda_i \;\leq\; \max_{1 \leq j \leq i} \|\mathbf{b}_j\|$ as $\mathbf{b}_1, \ldots, \mathbf{b}_i$ are linearly independent. Further $\|\mathbf{b}_j\| \leq 2^{(i-1)/2} \, \|\mathbf{w}_i\|$ by Problem 4.2 and with $\|\mathbf{w}_i\| \leq \|\mathbf{b}_i\|$ we obtain the lower bound.

4.5. Recall again that an *ellipsoid* is the image of a ball under some linear map ψ. We can find an orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_d$ of eigenvectors of the matrix AA^T corresponding to ψ (in some basis) with eigenvectors $a_1^2, \ldots, a_d^2$, such that

$$\mathcal{E} \;:=\; \mathcal{E}(a_1, \ldots\, a_d) \;:=\; \psi\left(\mathcal{B}_1(0)\right) \;=\; \left\{ \mathbf{x} \in \mathbb{R}^d \;:\; \sum_{i=1}^d \frac{1}{a_i^2} \langle\, \mathbf{x}, \mathbf{u}_i \,\rangle^2 \;\leq\; 1 \right\} .$$

Here we have assumed that ψ is a bijection. The eigenvalues of AA^T are positive as the matrix is symmetric and positive semidefinite. If $\mathbf{v}_1, \ldots, \mathbf{v}_d$ are the eigenvetors of $A^T A$, then this has the same eigenvalues $a_1^2, \ldots, a_d^2$ and $\mathbf{u}_i = \frac{1}{a_i} A\mathbf{v}_i$. Hence, for $\mathbf{x} = \sum_{i=1}^d x_i \mathbf{v}_i$ with $x_i = \langle\, \mathbf{x}, \mathbf{v}_i \,\rangle$ we have $\mathbf{y} = A\mathbf{x} = \sum_{i=1}^d y_i a_i \mathbf{u}_i$ and

$$x_1^2 + \cdots + x_d^2 \;\leq 1 \qquad \text{iff} \qquad \frac{1}{a_1^2} y_1^2 + \cdots + \frac{1}{a_d^2} y_d^2 \;\leq\; 1 \,.$$

Its volume is

$$\text{vol } \mathcal{E} \;=\; \text{vol } \mathcal{B}_\text{d} \cdot \prod_{i=1}^{d} a_i$$

Let $\mathbf{b}_1, \ldots, \mathbf{b}_d \in \Lambda \setminus \{0\}$ be vectors such that $\lambda_i = \|\mathbf{b}_i\|$ and let $\mathbf{b}_1^*, \ldots, \mathbf{b}_d^*$ be the Gram-Schmidt orthgonalization. We keep the order of the vectors.
We consider the ellipsoid

$$\mathcal{E} \;:=\; \left\{ \mathbf{x} \in \mathbb{R}^d \;:\; \sum_{i=1}^{d} \frac{1}{\lambda_i^2} \frac{\langle \mathbf{x}, \mathbf{b}_i^* \rangle^2}{\|\mathbf{b}_i^*\|^2} \;\leq\; 1 \right\}.$$

We want to show that the interior of $\mathcal{E}$ contains only one lattice point, namely the origin. Let $\mathbf{x} \in \Lambda \setminus \{0\}$, and let j be such that

$$\lambda_j \;\leq\; \|\mathbf{x}\| \;\leq\; \lambda_{j+1}$$

if such a j exists, and $j = d$ otherwise. Then clearly $\mathbf{x} \in \text{lin}(\mathbf{b}_1, \ldots, \mathbf{b}_j) = \text{lin}(\mathbf{b}_1^*, \ldots, \mathbf{b}_j^*)$, as otherwise we would have chosen $\mathbf{x}$ instead of $\mathbf{b}_{j+1}$ if $j < d$. Now

$$\sum_{i=1}^{d} \frac{1}{\lambda_i^2} \frac{\langle \mathbf{x}, \mathbf{b}_i^* \rangle^2}{\|\mathbf{b}_i^*\|^2} \;\geq\; \frac{1}{\lambda_j^2} \sum_{i=1}^{j} \frac{\langle \mathbf{x}, \mathbf{b}_i^* \rangle^2}{\|\mathbf{b}_i^*\|^2} \;=\; \frac{1}{\lambda_j^2} \|\mathbf{x}\|^2 \;\geq\; 1\,.$$

Hence, $\mathbf{x} \notin \mathcal{E}$. Now $\mathcal{E}$ is a centrally symmetric convex body, so from Minkowski's First Theorem (Corollary 3.3) we conclude

$$2^d \det \Lambda \;\geq\; \text{vol } \mathcal{E} \;= \text{vol } \mathcal{B}_\text{d} \cdot \prod_{i=1}^{d} \lambda_i \;\geq\; \left(\frac{2}{\sqrt{d}}\right)^d \prod_{i=1}^{d} \lambda_i\,,$$

where we have approximated the volume of the ball with the cube $[-1/\sqrt{d}, 1/\sqrt{d}]^d$ contained in $\mathcal{B}_\text{d}$.

4.8. We can choose linearly indepenent lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \Lambda$ such that $\|\mathbf{v}_i\| = \lambda_i$ for the successive minima $\lambda_1, \ldots, \lambda_d$ of Λ. This may not be a lattice basis, but as the vectors are linearly independent they span a sublattice Γ of Λ. We clearly have that

$$\det \Lambda \;\leq\; \det \Gamma\,.$$

The orthogonality defect is defined via

$$M_\Gamma \;:=\; \frac{1}{\det \Gamma} \prod_{j=1}^{d} \|\mathbf{v}_j\|\,.$$

and Minkowski's Second Theorem (Theorem 3.11) states that

$$\lambda_1 \cdots \lambda_d \leq 2^d \det \Lambda$$

Combining this gives the desired result.

4.9. We may assume that $\mathbf{a}$ is primitive.
We consider the dual basis $\mathbf{b}_i^*$ in $\Lambda^\star$, set $L^* := \text{lin}\{\mathbf{a}\}$, and project the basis into a generating set $\overline{\mathbf{b}}_1^*, \ldots, \overline{\mathbf{b}}_d^*$ of the lattice $\Lambda^\star/L$ inside $U := \left(\mathbb{R}^d\right)^\star/L$. We can use the Hermite normal form algorithm to compute a basis $\overline{\mathbf{c}}_2, \ldots, \overline{\mathbf{c}}_d$ of U. Let $\mathbf{c}_2, \ldots, \ldots, \mathbf{c}_d$ be preimages of this basis, via

$$\mathbf{c}_i \;=\; \sum \mu_{ij} \mathbf{b}_i^* \qquad\qquad \text{if } \overline{\mathbf{c}}_i \qquad\qquad = \sum \mu_{ij} \overline{\mathbf{b}}_i^*$$

for some $\mu_{ij} \in \mathbb{Z}$. Then $\mathbf{a}, \mathbf{c}_2, \ldots, \mathbf{c}_d$ is a basis of $\Lambda^\star$, The dual basis $\mathbf{a}^*, \mathbf{c}_2^*, \ldots, \mathbf{c}_d^*$ is a lattice basis of Λ. By construction, $\mathbf{c}_2^*, \ldots, \mathbf{c}_d^*$ must span L.

4.10. Here is one option to prove this.

We can extend the set of vectors $\mathbf{b}_1, \ldots, vb_k$ to a basis of $\mathbb{R}^d$ with vectors $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_d$. We may assume that the latter are orthogonal to any $\mathbf{b}_i$ and have unit length. Let $\overline{Z}$ be the zonotope spanned by $\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_d$. It has Z as a k-dimensional face and

$$\text{vol } \overline{Z} \;=\; \text{vol } Z\,.$$

If $\overline{B}$ is the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_d$, then $\overline{B}^t\overline{B}$ has B^tB as a minor in the upper left corner, 1 on the remaining diagonal and 0 in all other entries. Hence

$$\det \overline{B}^t\overline{B} \;=\; \det B^tB\,.$$

But $det\, \overline{B}^t\overline{B} = (\det\, \overline{B})^2 = \text{vol } \overline{Z}$, and the claim follows.

B.4. Solutions for *Reduced Bases* (*Chapter 5*)

B.5. Solutions for *Integer Programming* (*Chapter 6*)

6.3. One example is the system of inequalities

$$\sum_{i\in I} x_i \;-\; \sum_{i\notin I} x_i \;\leq\; |I| \;-\; 1 \qquad\qquad \text{for} \qquad\qquad I \;\subseteq \{1,\ldots,d\}$$

of 2^d inequalities.

B.6. Solutions for *The Closest Vector Problem* (*Chapter 8*)

8.6. We show first that the translates of F cover the space. Let $\mathbf{v} \in \mathbb{R}^d$. We want to show that $\mathbf{v} - \mathbf{u} \in \text{F}$ for some $\mathbf{u} \in \Lambda$.

Let $\mathbf{v}' := \pi(\mathbf{v})$. As F_Γ is a fundamental domain, we have $\mathbf{u}' \in \Gamma$ and $\mathbf{c}' \in \text{F}_\Gamma$ such that

$$\mathbf{v}' \;=\; \mathbf{u}' \;+\; \mathbf{c}'$$

Pick any $\mathbf{u} \in \pi^{-1}(\mathbf{u}') \cap \Lambda$. Then

$$\mathbf{v}_1 \;:=\; \mathbf{v} - \mathbf{v}' \in \text{lin}(\mathbf{b}_1) \qquad\qquad \text{and} \qquad\qquad \mathbf{u}_1 \;:=\; \mathbf{u} - \mathbf{u}' \in \text{lin}(\mathbf{b}_1)\,,$$

so there are $\lambda \in \mathbb{Z}$ and $\mu \in [0,1)$ such that

$$\mathbf{v}_1 \;-\; \mathbf{u}_1 \;=\; (\lambda + \mu)\mathbf{b}_1\,.$$

We can write

$$\begin{aligned} \mathbf{v} \;&=\; \mathbf{v}' + (\mathbf{v} - \mathbf{v}') && =\; \mathbf{u}' + \mathbf{c}' + \mathbf{v}_1 \\ &=\; \mathbf{u} + (\mathbf{u}' - \mathbf{u}) + \mathbf{c}' + \mathbf{v}_1 && =\; \mathbf{u} - \mathbf{u}_1 + \mathbf{c}' + \mathbf{v}_1 \\ &=\; \mathbf{u} + \mathbf{c}' + (\lambda+\mu)\mathbf{b}_1 && =\; (\mathbf{u} + \lambda\mathbf{b}_1) + (\mathbf{c}' + \mu\mathbf{b}_1)\,. \end{aligned}$$

The left summand in the last term is in the lattice, and the right is in F.

We still need to prove that different translates of F do not intersect. So assume there is $\mathbf{u}, \mathbf{v} \in \Lambda$

with $\mathbf{u} \neq \mathbf{v}$ such that there is $\mathbf{x} \in \mathbf{u} + \mathrm{F} \cap \mathbf{v} + \mathrm{F}$. Then

$$\pi(\mathbf{x}) \ \in \ \pi(\mathbf{u}) \ + \ \mathrm{F}_\Gamma \ \cap \pi(\mathbf{v}) \ + \ \mathrm{F}_\Gamma \ .$$

As $\pi(\mathbf{u}), \pi(\mathbf{v}) \in \Gamma$ and F_Γ is a fundamental domain, we know that $\pi(\mathbf{u}) = \pi(\mathbf{v})$, so $\mathbf{v} - \mathbf{u} = \lambda \mathbf{b}_1$ for some $\lambda \in \mathbb{Z}$.
Further, there are $0 \leq \mu_1, \nu_i < 1$ for $1 \leq i \leq d$ such that

$$\mathbf{x} \ = \ \mathbf{u} + \sum \mu_i \mathbf{w}_i \ = \ \mathbf{v} + \sum \nu_i \mathbf{w}_i \, .$$

As $\mathbf{w}_2, \ldots, \mathbf{w}_d$ is the Gram-Schmidt basis in L we know that $\mu_i = \nu_i$ for $i \geq 2$, and thus

$$\lambda \mathbf{b}_1 \ = \ (\mu_1 - \nu_1)\mathbf{w}_1 \ = \ (\mu_1 - \nu_1)\mathbf{b}_1 \, .$$

But $-1 < \mu_1 - \nu_1 = \lambda < 1$ and $\lambda \in \mathbb{Z}$, so $\lambda = 0$, So $\mathbf{u} = \mathbf{v}$.

B.7. Solutions for *Cuts and Lattice Free Polytopes (Chapter * 11)*

* 11.3. Let $K = \{ \mathbf{x} \ : \ \langle \mathbf{a}_i, \mathbf{x} \rangle \leq \beta_i \text{ for } 1 \leq i \leq m \}$. We have $\mathrm{rec}\ K \subseteq \ rec\, K - C$.
Let $\mathbf{x} := \mathbf{y} - \mathbf{c} \in \mathrm{int}(K - C)$ for some $\mathbf{y} \in K$ and $\mathbf{c} \in C$, and choose $\varepsilon > 0$ such that $\mathcal{B}_{\mathbf{x}}(\varepsilon) \subseteq K - C$.
Let

$$\delta_i \ := \ \max \left(\langle \mathbf{a}_i, \mathbf{z} \rangle \ : \ \mathbf{z} \in \mathcal{B}_{\mathbf{x}}(\varepsilon) \right) \qquad \text{for} \qquad 1 \leq i \leq m \, .$$

Then $\delta_i > \infty$ as $\mathcal{B}_{\mathbf{x}}(\varepsilon)$ is finite, and for each i we can find k_i such that $\langle \mathbf{a}_i, k\mathbf{c} \rangle < \beta_i - \delta_i$, and let $k := \max(k_i)$. Then

$$\langle \mathbf{a}_i, \mathbf{z} + k\mathbf{c} \rangle \ = \ \langle \mathbf{a}_i, \mathbf{z} \rangle \ + \ \langle \mathbf{a}_i, k\mathbf{c} \rangle \ \leq \ \delta_i \ + \ \beta_i \ - \ \delta_i \ = \ \beta_i \, .$$

and hence, $\mathcal{B}_{\mathbf{x}+k\mathbf{c}}(\varepsilon) = \mathcal{B}_{\mathbf{x}}(\varepsilon) + k\mathbf{c} \subseteq K$. So $\mathbf{x} + k\mathbf{c} \in \mathrm{int}\ K$ and

$$\mathbf{x} \ = \ (\mathbf{x} + k\mathbf{c}) \ - \ (k+1)\mathbf{c} \ \in \mathrm{int}(K) \ - \ C \, .$$

B.8. Solutions for *Convexity (Appendix A)*

A.5. The matrix A_2 is positive definite, so there is M such that $A_2 = M^2$. Then

$$(\det(A_1 + A_2))^{1/d} \ = \ (\det(M + A_2))^{1/d} \ = \ (\det\ M)^{2/d} \left(\mathrm{I} + M^{-1} A_2 M^{-1} \right)^{1/d}$$

and

$$(\det\ A_1)^{1/d} \ + \ (\det\ A_2)^{1/d} \ = \ (\det\ M)^{2/d} \left(1 + \left(\det(M^{-1} A_2 M^{-1} \right)^{1/d} \right) \, .$$

Let $B := M^{-1} A_1 M^{-1}$. Then it suffices to show that

$$(\det(\mathrm{I} + B))^{1/d} \ \geq 1 \ + \ (\det\ B)^{1/d} \tag{B.9}$$

with equality if and only if B is a positive multiple of I. Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of B. Then (B.9) is equivalent to

$$\left(\prod_{i=1}^{d} (1 + \lambda_i) \right)^{1/d} \ \geq \ 1 \ + \ (\lambda_1 \cdots \lambda_d)^{1/d} \tag{B.10}$$

with equality if and only if $\lambda_i = \lambda$ for all i and some positive λ. Let

$$\sigma_k \;:=\; \sigma_k(\lambda_1,\ldots,\lambda_d) \;:=\; \sum_{1\le i_1<\cdots<i_k\le d} \lambda_{i_1}\cdots\lambda_{i_k}\,.$$

Then

$$\prod_{i=1}^{d}(1+\lambda_i) \;=\; 1 \,+\, \sigma_1 \,+\, \cdots \,+\, \sigma_d\,.$$

The inequality bewtween the arithmetic and geometric mean implies

$$\begin{aligned}
\sigma_k \;&=\; \binom{d}{k} \sum_{1\le i_1<\cdots<i_k\le d} \frac{1}{\binom{d}{k}} \lambda_{i_1}\cdots\lambda_{i_k} \\
&=\; \binom{d}{k} \prod_{1\le i_1<\cdots<i_k\le d} (\lambda_{i_1}\cdots\lambda_{i_k})^{1/\binom{d}{k}} \\
&=\; \binom{d}{k} (\lambda_{i_1}\cdots\lambda_{i_d})^{\binom{d-1}{k-1}/\binom{d}{k}} \\
&=\; \binom{d}{k} (\lambda_{i_1}\cdots\lambda_{i_d})^{k/d}
\end{aligned}$$

with equality if and only if

$$\lambda_{i_1}\cdots\lambda_{i_k} \;=\; \lambda_{j_1}\cdots\lambda_{j_k}$$

for all $1 \le i_1 < \cdots < i_k \le d$ and $1 \le j_1 < \cdots < j_k \le d$, so if and only if $\lambda_i = \lambda$ for some λ. This gives

$$\begin{aligned}
\left(1 \,+\, (\lambda_1\cdots\lambda_d)^{1/d}\right)^d \;&=\; \sum k = 0^d \binom{d}{k} (\lambda_{i_1}\cdots\lambda_{i_d})^{k/d} \\
&=\; 1 \,+\, \sigma_1 \,+\, \cdots \,+\, \sigma_d \;=\; \prod_{i=1}^{d}(1+\lambda_i)
\end{aligned}$$

with equality if and only if $\lambda_i = \lambda$ for all i and some $\lambda > 0$. This is equivalent to (B.10).

A.6. Let E be an ellipsoid in the cube C. We can use a linear transformation T to map it into the unit sphere. The map sends C into a parallelepiped enclosing the unit ball.

The map T preserves the volume ratio between the ellipsoid and the cube. Hence, instead of maximising the volume of E we can find parallelepipeds P enclosing the unit ball that minimize the volume.

We prove this by induction. In dimension 1 there is nothing to prove. SO assume we know this for dimensions up to $d-1$. Choose one of the facets F of P. it has an opposite parallel facet F' at distance 2. Let Q be the intersection of P with the hyperplane parallel to F through the origin. Then Q has distance 1 from F and F', and the volume of P is twice the volume of Q. Q encloses a unit ball in dimension $d-1$, but its facets may not touch. Yet, by assumption, only the cube $[-1,1]^{d-1}$ in $\operatorname{lin} Q$ minimizes the volume, so Q must be this cube. This, however, implies that all facets of P defining facets of Q must be tangent planes to the sphere at Q, so also P must be the cube $[-1,1]^d$.

A.7. We can map the simplex into R^d with some orthogonal map, and translate, so that the center of the maximal volume ellipsoid is the orign.

The maximum volume ellipsoid E is unique inside the simplex, so it must have (at least) the same symmetries as the simplex, which is the group S_d. Assume that E has a principal direction $\mathbf{a}$ of maximal length ℓ, and at least one, that has a smaller length.

Any orthogonal map maps principal directions onto such and preserves the length, so if there are

$k < d$ directions of the ℓ, then the symmetry group of E is a subgroup of S_k. This implies $k = d$ and all principal directions of E have the same length. Hence, E is a ball.

The other claims follow.

Index

Bibliography

Aardal, Karen. *Lattice basis reduction and Integer programming*. Tech. rep. CWI, 1999 (cit. on p. 57).

Aggarwal, Divesh, Daniel Dadush, and Noah Stephens-Davidowitz. "Solving the Closest Vector Problem in 2^n Time— The Discrete Gaussian Strikes Again!" In: *FOCS 2015* (Apr. 8, 2015). arXiv: 1504.01995 [cs.DS] (cit. on p. 70).

Ajtai, Miklós, Ravi Kumar, and D. Sivakumar. "A sieve algorithm for the shortest lattice vector problem". In: *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*. Hersonissos, Greece: ACM, New York, 2001, pp. 601–610. DOI: 10.1145/380752.380857. MR2120363 68W20 (68Q25 94B40) (cit. on p. 70).

Andersen, Kent, Quentin Louveaux, Robert Weismantel, and Laurence A. Wolsey. "Inequalities from two rows of a simplex tableau". In: *Integer programming and combinatorial optimization*. Ed. by Matteo Fischetti and David P. Williamson. Vol. 4513. Lecture Notes in Comput. Sci. Available electronically at http://www.springerlink.com/content/978-3-540-72791-0. Berlin: Springer, 2007, pp. 1–15. DOI: 10.1007/978-3-540-72792-7_1. MR2480507 (2011b:90082) 90C11 (52B12 90C05 90C49 90C57), rev. by Pablo Guerrero-Garcia. URL: https://doi.org/10.1007/978-3-540-72792-7 (cit. on pp. 156, 157).

Averkov, Gennadiy. *A proof of Lovász's theorem on maximal lattice-free sets*. Oct. 2011. eprint: 1110.1014 (cit. on pp. 155, 159).

– *Difference between families of weakly and strongly maximal integral lattice-free polytopes*. July 2018. eprint: 1807.06327 (cit. on p. 166).

Averkov, Gennadiy, Jan Krümpelmann, and Stefan Weltge. *Notions of maximality for integral lattice-free polyhedra: the case of dimension three*. Sept. 2015. eprint: 1509.05200 (cit. on p. 53).

Averkov, Gennadiy, Christian Wagner, and Robert Weismantel. "Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three". In: *Math. Oper. Res.* 36.4 (2011), pp. 721–742. DOI: 10.1287/moor.1110.0510. MR2855866 (2012j:90103) 90C11 (52B12 90C57). arXiv: 1010.1077 [math.CO] (cit. on pp. 155, 159, 164, 166).

Babai, L. "On Lovász' lattice reduction and the nearest lattice point problem". In: *Combinatorica* 6.1 (1986), pp. 1–13. DOI: 10.1007/BF02579403. MR856638 68Q25 (11H06 11Y16 12-04), rev. by A. K. Lenstra (cit. on pp. 89, 102, 103).

Banaszczyk, W. "Inequalities for convex bodies and polar reciprocal lattices in $\mathbf{R}^n$. II. Application of K-convexity". In: *Discrete Comput. Geom.* 16.3 (1996), pp. 305–311. DOI: 10.1007/BF02711514. MR1410163 11H60 (11H06 52C07), rev. by Martin Henk (cit. on pp. 47, 53).

Banaszczyk, Wojciech, Alexander E. Litvak, Alain Pajor, and Stanislaw J. Szarek. "The flatness theorem for nonsymmetric convex bodies via the local theory of Banach spaces". In: *Math. Oper. Res.* 24.3 (1999), pp. 728–750. DOI: 10.1287/moor.24.3.728. MR1854250 52C07 (46N10 90C10), rev. by Konrad J. Swanepoel (cit. on p. 53).

Barvinok, A. I. "Computing the Ehrhart polynomial of a convex lattice polytope". In: *Discrete Comput. Geom.* 12.1 (1994), pp. 35–48. DOI: 10.1007/BF02574364. MR1280575 (95e:52015) 52B05 (52B20 52B45 68Q25), rev. by Mark E. Hartmann. URL: http://dx.doi.org/10.1007/BF02574364 (cit. on p. 128).

Barvinok, Alexander. *A course in convexity*. Vol. 54. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2002, pp. x+366. MR1940576 (2003j:52001) 52-02 (49N15 52-01 90-02 90C05 90C22 90C25), rev. by P. McMullen (cit. on p. 143).

– *Integer points in polyhedra*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008, pp. viii+191. DOI: 10.4171/052. MR2455889 (2011a:52001) 52-02 (05-02 05A15 11H06 52B20 52B45 52B55 52C07 52C45), rev. by Martin Henk (cit. on p. 51).

Barvinok, Alexander and James E. Pommersheim. "An algorithmic theory of lattice points in polyhedra". In: *New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97)*. Ed. by Louis J. Billera, Anders Björner, Curtis Greene, Rodica E. Simion, and Richard P. Stanley. Vol. 38. Math. Sci. Res. Inst. Publ. Papers from

the MSRI Program on Combinatorics held in Berkeley, CA, 1996–1997. Cambridge: Cambridge Univ. Press, 1999, pp. 91–147. MR1731815 (2000k:52014) 52C07 (05A15) (cit. on pp. 11, 128).

Barvinok, Alexander and Kevin Woods. "Short rational generating functions for lattice point problems." English. In: *J. Am. Math. Soc.* 16.4 (2003), pp. 957–979. DOI: 10.1090/S0894-0347-03-00428-4 (cit. on p. 11).

Basu, Amitabh, Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. "Maximal lattice-free convex sets in linear subspaces". In: *Math. Oper. Res.* 35.3 (2010), pp. 704–720. DOI: 10.1287/moor.1100.0461. MR2724071 (2011i:90076) 90C11 (52A20 52B12 90C10), rev. by Ruriko Yoshida (cit. on pp. 155, 159).

Basu, Amitabh, Michele Conforti, and Marco Di Summa. "A geometric approach to cut-generating functions". In: *Mathematical Programming, vol.151(1), 2015, pp. 153-189* (Jan. 24, 2017). DOI: 10.1007/s10107-015-0890-5. arXiv: 1701.06692 [math.OC] (cit. on pp. 156, 159, 166).

Beck, Matthias and Sinai Robins. *Computing the continuous discretely*. Undergraduate Texts in Mathematics. Integer-point enumeration in polyhedra. New York: Springer, 2007, pp. xviii+226. MRMR2271992 (2007h:11119) 11P21 (05-02 05A15 11-02 11H06 52B05 52B20), rev. by Jesús A. de Loera (cit. on p. 143).

Bell, David E. "A theorem concerning the integer lattice". In: *Studies in Appl. Math.* 56.2 (1976), pp. 187–188. DOI: 10.1002/sapm1977562187. MR462617 90C99, rev. by George E. Andrews. URL: https://doi.org/10.1002/sapm1977562187 (cit. on p. 91).

Bertsimas, Dimitris and Robert Weismantel. *Optimization over integers*. Athena Scientific, 2005 (cit. on p. 83).

Betke, Ulrich, Martin Henk, and Jörg M. Wills. "Successive-minima-type inequalities". In: *Discrete Comput. Geom.* 9.2 (1993), pp. 165–175. DOI: 10.1007/BF02189316. MR1194034 52C07 (52C05 52C17), rev. by Peter M. Gruber. URL: http://dx.doi.org/10.1007/BF02189316 (cit. on p. 43).

Codenotti, Giulia and Francisco Santos. *Hollow polytopes of large width*. Dec. 2018. eprint: 1812.00916 (cit. on p. 54).

Cohen, Henri. *A course in computational algebraic number theory*. Vol. 138. Graduate Texts in Mathematics. Springer-Verlag, Berlin, 1993, pp. xii+534. DOI: 10.1007/978-3-662-02945-9. MR1228206 11Y40 (11Rxx 68Q40), rev. by Joe P. Buhler. URL: https://doi.org/10.1007/978-3-662-02945-9 (cit. on p. 73).

Cook, W., R. Kannan, and A. Schrijver. "Chvátal closures for mixed integer programming problems". In: *Math. Programming* 47.2, (Ser. A) (1990), pp. 155–174. DOI: 10.1007/BF01580858. MR1059391 (91c:90077) 90C11, rev. by Klaus Hofstedt (cit. on p. 163).

De Loera, Jesús A., Raymond Hemmecke, and Matthias Köppe. *Algebraic and geometric ideas in the theory of discrete optimization*. Vol. 14. MOS-SIAM Series on Optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2013, pp. xx+322. MR3024570 90-02 (13P10 52C07 90C10 90C27 90C52), rev. by Alexander I. Barvinok (cit. on p. 134).

De Loera, Jesús A., Jörg Rambau, and Francisco Santos. *Triangulations. Structures for algorithms and applications*. Vol. 25. Algorithms and Computation in Mathematics. Structures for algorithms and applications. Springer-Verlag, Berlin, 2010, pp. xiv+535. DOI: 10.1007/978-3-642-12971-1. MR2743368 (2011j:52037) 52B55 (05C10 52B05 57Q15 68U05). URL: http://dx.doi.org/10.1007/978-3-642-12971-1 (cit. on p. 192).

Del Pia, Alberto and Robert Weismantel. "On convergence in mixed integer programming". In: *Math. Program.* 135.1-2, Ser. A (2012), pp. 397–412. DOI: 10.1007/s10107-011-0476-9. MR2968262 90C11 (90C57), rev. by Mechthild Opperud. URL: http://dx.doi.org/10.1007/s10107-011-0476-9 (cit. on pp. 155, 164).

Doignon, Jean-Paul. "Convexity in cristallographical lattices". In: *J. Geom.* 3 (1973), pp. 71–85. DOI: 10.1007/BF01949705. MR387090 05B35 (52A35), rev. by G. L. Alexanderson (cit. on p. 91).

Draisma, Jan, Tyrrell B. McAllister, and Benjamin Nill. *Lattice width directions and Minkowski's 3^d-theorem*. Jan. 2009. eprint: 0901.1375 (cit. on p. 205).

Ehrhart, Eugène. *Polynômes arithmétiques et méthode des polyèdres en combinatoire*. International Series of Numerical Mathematics, Vol. 35. Basel: Birkhäuser, 1977, p. 165. MR0432556 (55 #5544) 10E99 (10B05 10E40), rev. by G. L. Alexanderson (cit. on p. 143).

– "Sur un problème de géométrie diophantienne linéaire. II. Systèmes diophantiens linéaires". In: *J. Reine Angew. Math.* 227 (1967), pp. 25–49. MR36 #105 10.10, rev. by C. G. Lekkerkerker (cit. on p. 143).

Eisenbrand, Friedrich, Christoph Hunkenschröder, Kim-Manuel Klein, Martin Koutecký, Asaf Levin, and Shmuel Onn. *An Algorithmic Theory of Integer Programming*. Apr. 2, 2019. arXiv: 1904.01361 [math.OC] (cit. on p. 84).

Frieze, A. M. “On the Lagarias-Odlyzko algorithm for the subset sum problem”. In: *SIAM J. Comput.* 15.2 (1986), pp. 536–539. DOI: 10.1137/0215038. MR837602 68Q25 (11B99 11Y16 65K05 90C10), rev. by J. C. Lagarias (cit. on p. 97).

Gathen, Joachim von zur and Jürgen Gerhard. *Modern computer algebra*. Third. Cambridge University Press, Cambridge, 2013, pp. xiv+795. DOI: 10.1017/CBO9781139856065. MR3087522 68W30 (11Y05 11Y11 13Pxx) (cit. on p. 73).

Goffin, Jean-Louis. “Variable metric relaxation methods. II. The ellipsoid method”. In: *Math. Programming* 30.2 (1984), pp. 147–162. DOI: 10.1007/BF02591882. MR758001 90C05 (52A40 65K10), rev. by F. Pickel (cit. on p. 184).

Goldreich, O., D. Micciancio, S. Safra, and J.-P. Seifert. “Approximating shortest lattice vectors is not harder than approximating closest lattice vectors”. In: *Information Processing Letters* 71.2 (1999), pp. 55–61. DOI: https://doi.org/10.1016/S0020-0190(99)00083-6. URL: https://www.sciencedirect.com/science/article/pii/S0020019099000836 (cit. on p. 102).

Gomory, Ralph E. “Some polyhedra related to combinatorial problems”. In: *Linear Algebra Appl.* 2 (1969), pp. 451–558. DOI: 10.1016/0024-3795(69)90017-2. MR256718 90.56, rev. by G. Berman. URL: https://doi.org/10.1016/0024-3795(69)90017-2 (cit. on p. 156).

Gomory, Ralph E. and Ellis L. Johnson. “Some continuous functions related to corner polyhedra”. In: *Math. Programming* 3 (1972), pp. 23–85. DOI: 10.1007/BF01584976. MR479415 90C99, rev. by Joachim Piehler. URL: https://doi.org/10.1007/BF01584976 (cit. on p. 156).

– “Some continuous functions related to corner polyhedra. II”. In: *Math. Programming* 3 (1972), pp. 359–389. DOI: 10.1007/BF01585008. MR479416 90C99, rev. by Joachim Piehler. URL: https://doi.org/10.1007/BF01585008 (cit. on p. 156).

Grötschel, Martin, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*. Second. Vol. 2. Algorithms and Combinatorics. Berlin: Springer-Verlag, 1993, pp. xii+362. MR1261419 (95e:90001) 90-02 (52C07 90C27), rev. by Ulrich Faigle (cit. on pp. 38, 53, 73, 85, 184).

Haase, Christian, Benjamin Nill, and Andreas Paffenholz. *Lattice Polytopes*. Lecture Notes. Dec. 2012. URL: http://www2.mathematik.tu-darmstadt.de/~paffenholz/daten/preprints/20210628_Lattice_Polytopes.pdf (cit. on p. 1).

Hanrot, Guillaume, Xavier Pujol, and Damien Stehlé. “Algorithms for the shortest and closest lattice vector problems”. In: *Coding and cryptology*. Ed. by Yeow Meng Chee, Zhenbo Guo, San Ling, Fengjing Shao, Yuansheng Tang, Huaxiong Wang, and Chaoping Xing. Vol. 6639. Lecture Notes in Comput. Sci. Springer, Heidelberg, 2011, pp. 159–190. DOI: 10.1007/978-3-642-20901-7_10. MR2834699 05B20. URL: https://doi.org/10.1007/978-3-642-20901-7 (cit. on p. 57).

Henk, Martin. “Successive minima and lattice points”. In: *Rend. Circ. Mat. Palermo (2) Suppl.* 70, part I (2002). IV International Conference in “Stochastic Geometry, Convex Bodies, Empirical Measures & Applications to Engineering Science”, Vol. I (Tropea, 2001), pp. 377–384. MR1962579 (2003m:11164) 11P21 (11H06 52C07), rev. by Matthias Beck. eprint: math.MG/0204158 (cit. on pp. 42, 44).

Henk, Martin, Stefan Kuhlmann, and Robert Weismantel. *On lattice width of lattice-free polyhedra and height of Hilbert bases*. Oct. 6, 2021. arXiv: 2110.02893 [math.CO] (cit. on p. 166).

Hensley, Douglas. “Lattice vertex polytopes with interior lattice points”. In: *Pacific J. Math.* 105.1 (1983), pp. 183–191. DOI: 10.2140/pjm.1983.105.183. MR688412 (84c:52016) 52A43 (10E05), rev. by J. M. Wills (cit. on p. 165).

Hermite, C. “Extraits de lettres de M. Ch. Hermite à M. Jacobi sur différents objects de la théorie des nombres”. In: *J. Reine Angew. Math.* 40 (1850), pp. 261–278. DOI: 10.1515/crll.1850.40.261. MR1578698 DML. URL: https://doi.org/10.1515/crll.1850.40.261 (cit. on p. 42).

Hibi, Takayuki. “A lower bound theorem for Ehrhart polynomials of convex polytopes”. In: *Adv. Math.* 105.2 (1994), pp. 162–165. DOI: 10.1006/aima.1994.1042. MR1275662 (95b:52018) 52B20, rev. by P. McMullen. URL: http://dx.doi.org/10.1006/aima.1994.1042 (cit. on p. 143).

Kannan, Ravi. “Improved Algorithms for Integer Programming and Related Lattice Problems”. In: *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*. STOC ’83. New York, NY, USA: Association for Computing Machinery, 1983, pp. 193–206. DOI: 10.1145/800061.808749 (cit. on pp. 70, 90).

– “Lattice translates of a polytope and the Frobenius problem”. In: *Combinatorica* 12.2 (1992), pp. 161–177. DOI: 10.1007/BF01204720. MR1179254 52B55 (52C07 90C60), rev. by Gerard Sierksma (cit. on p. 11).

– “Test sets for integer programs, ∀∃ sentences”. In: *Polyhedral combinatorics (Morristown, NJ, 1989)*. Ed. by William Cook and Paul D. Seymour. Vol. 1. DIMACS Ser. Discrete Math. Theoret. Comput. Sci.

Amer. Math. Soc., Providence, RI, 1990, pp. 39–47. DOI: 10.1090/dimacs/001. MR1105115 90C10 (03D15), rev. by Joachim Piehler. URL: https://doi.org/10.1090/dimacs/001 (cit. on p. 11).
Kannan, Ravi and László Lovász. "Covering minima and lattice-point-free convex bodies". In: *Ann. of Math. (2)* 128.3 (1988), pp. 577–602. DOI: 10.2307/1971436. MR970611 (89i:52020) 52A43 (11H06 11H31 52A45 90C10), rev. by J. M. Wills (cit. on p. 51).
Khot, Subhash. "Hardness of approximating the shortest vector problem in lattices". In: *J. ACM* 52.5 (2005), pp. 789–808. DOI: 10.1145/1089023.1089027. MR2176563 68Q17 (11H71 68Q15 94A60 94B35), rev. by Johan Håstad. URL: https://doi.org/10.1145/1089023.1089027 (cit. on p. 57).
Klein, Philip. "Finding the closest lattice vector when it's unusually close". In: *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (San Francisco, CA, 2000)*. Held in San Francisco, CA, January 9–11, 2000. ACM, New York, 2000, pp. 937–941. MR1755554 68Q25 (cit. on p. 108).
Klüners, Jürgen. "The van Hoeij Algorithm for Factoring Polynomials. Survey and Applications". In: *The LLL Algorithm: Survey and Applications*. Ed. by Phong Q. Nguyen and Brigitte Vallée. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 283–291. DOI: 10.1007/978-3-642-02295-1_8 (cit. on p. 80).
Korte, Bernhard and Jens Vygen. *Combinatorial Optimization*. Springer, 2008 (cit. on p. 84).
Koy, Henrik and Claus Peter Schnorr. "Segment LLL-reduction of lattice bases". In: *Cryptography and lattices (Providence, RI, 2001)*. Ed. by Joseph H. Silverman. Vol. 2146. Lecture Notes in Comput. Sci. Springer, Berlin, 2001, pp. 67–80. DOI: 10.1007/3-540-44670-2_7. MR1903888 11H55 (11H06 11Y16 68Q25), rev. by Martin Henk. URL: https://doi.org/10.1007/3-540-44670-2_7 (cit. on p. 79).
Lagarias, J. C. "Knapsack public key cryptosystems and Diophantine approximation (extended abstract)". In: *Advances in cryptology (Santa Barbara, Calif., 1983)*. Ed. by David Chaum. Plenum, New York, 1984, pp. 3–23. MR799717 94A60 (11T71) (cit. on p. 80).
Lagarias, J. C., H. W. Lenstra Jr., and C.-P. Schnorr. "Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice". In: *Combinatorica* 10.4 (1990), pp. 333–348. DOI: 10.1007/BF02128669. MR1099248 11H50 (11H55 11Y16), rev. by Simon N. Litsyn. URL: https://doi.org/10.1007/BF02128669 (cit. on p. 49).
Lagarias, J. C. and A. M. Odlyzko. "Solving low-density subset sum problems". In: *J. Assoc. Comput. Mach.* 32.1 (1985), pp. 229–246. DOI: 10.1145/2455.2461. MR832341 11Y16 (68Q25 90C09), rev. by A. K. Lenstra (cit. on p. 97).
Lagarias, Jeffrey C. and Günter M. Ziegler. "Bounds for lattice polytopes containing a fixed number of interior points in a sublattice". In: *Canad. J. Math.* 43.5 (1991), pp. 1022–1035. DOI: 10.4153/CJM-1991-058-4. MR1138580 (92k:52032) 52C07 (11H06), rev. by J. M. Wills (cit. on p. 165).
Lenstra, A. K., H. W. Lenstra Jr., and L. Lovász. "Factoring polynomials with rational coefficients". In: *Math. Ann.* 261.4 (1982), pp. 515–534. DOI: 10.1007/BF01457454. MR682664 12-04 (12A20 68C20 68C25), rev. by Daniel Lazard. URL: http://dx.doi.org/10.1007/BF01457454 (cit. on pp. 7, 58, 73, 74, 79, 80).
Lenstra, Hendrik W. "Integer Programming with a fixed number of variables". In: *Math. Oper. Res.* 8 (1983), pp. 538–548 (cit. on pp. 9, 89).
Lovász, László. "Geometry of numbers and integer programming". In: *Mathematical programming (Tokyo, 1988)*. Ed. by Tokyo SCIPRESS. Vol. 6. Math. Appl. (Japanese Ser.) SCIPRESS, Tokyo, 1989, pp. 177–201. MR1114315 90C10 (11H06 90C27), rev. by Michael A. Trick (cit. on pp. 12, 155, 159).
Macdonald, I.G. "Polynomials associated with finite cell-complexes". English. In: *J. Lond. Math. Soc., II. Ser.* 4 (1971), pp. 181–192 (cit. on p. 143).
Merkle, Ralph C. and Martin E. Hellman. "Hiding information and signatures in trapdoor knapsacks". In: *Secure communications and asymmetric cryptosystems*. Ed. by Gustavus J. Simmons. Vol. 69. AAAS Sel. Sympos. Ser. Westview, Boulder, CO, 1982, pp. 197–215. MR668725 94A05 (cit. on p. 96).
Micciancio, Daniele and Panagiotis Voulgaris. "A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations [extended abstract]". In: *STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*. Ed. by Daniel Bienstock and George Nemhauser. Vol. 3064. Lecture Notes in Computer Science. Berlin: ACM, New York, 2010, pp. 351–358. DOI: 10.1007/b97946. MR2743283 68Q25 (68P25). URL: http://dx.doi.org/10.1007/b97946 (cit. on pp. 102, 110).
Nemhauser, George and Laurence Wolsey. *Integer and Combinatorial Optimization*. John Wiley & Sons, Inc., June 1988. DOI: 10.1002/9781118627372 (cit. on pp. 9, 11, 83, 84).
Nguyen, Phong Q. and Damien Stehlé. "Low-dimensional lattice basis reduction revisited". In: *ACM Trans. Algorithms* 5.4 (2009), Art. 46, 48. DOI: 10.1145/1597036.1597050. MR2571909 68W40 (11A05 68Q25) (cit. on p. 73).

Nguyen, Phong Q. and Brigitte Vallée, eds. *The LLL Algorithm. Survey and Applications*. Springer Berlin, 2010. DOI: 10.1007/978-3-642-02295-1 (cit. on p. 57).
– eds. *The LLL algorithm*. Information Security and Cryptography. Survey and applications. Springer-Verlag, Berlin, 2010, pp. xiv+496. DOI: 10.1007/978-3-642-02295-1. MR2722178 11Y16 (11-06 11H06 94A60), rev. by Jr. Samuel S. Wagstaff. URL: http://dx.doi.org/10.1007/978-3-642-02295-1 (cit. on pp. 73, 79).
Nill, Benjamin and Günter M. Ziegler. "Projecting lattice polytopes without interior lattice points". In: *Math. Oper. Res.* 36.3 (2011), pp. 462–467. DOI: 10.1287/moor.1110.0503. MR2832401 (2012k:52034) 52B20 (11H06 52C07), rev. by Gennadiy Averkov. arXiv: 1101.4292 [math.CO] (cit. on p. 164).
Pikhurko, Oleg. "Lattice points in lattice polytopes". In: *Mathematika* 48.1-2 (2001), 15–24 (2003). DOI: 10.1112/S0025579300014339. MR1996360 (2004f:52009) 52B20 (52C07), rev. by Margaret M. Bayer. arXiv: math/0008028 [math.CO] (cit. on p. 10).
Rudelson, M. "Distances between non-symmetric convex bodies and the MM^*-estimate". In: *Positivity* 4.2 (2000), pp. 161–178. DOI: 10.1023/A:1009842406728. MR1755679 52A20 (46B20 52A40) (cit. on pp. 8, 53).
Scarf, Herbert E. "An observation on the structure of production sets with indivisibilities". In: *Proc. Nat. Acad. Sci. U.S.A.* 74.9 (1977), pp. 3637–3641. DOI: 10.1073/pnas.74.9.3637. MR452678 90C10, rev. by F. Giannessi. URL: https://doi.org/10.1073/pnas.74.9.3637 (cit. on p. 91).
Schnorr, Claus-P. "A hierarchy of polynomial time lattice basis reduction algorithms". In: *Theoret. Comput. Sci.* 53.2-3 (1987), pp. 201–224. DOI: 10.1016/0304-3975(87)90064-8. MR918090 11Y16 (11H55 68Q25), rev. by J. C. Lagarias. URL: http://dx.doi.org/10.1016/0304-3975(87)90064-8 (cit. on p. 79).
Schrijver, Alexander. *Combinatorial optimization. Polyhedra and efficiency (3 volumes)*. English. Algorithms and Combinatorics 24. Berlin: Springer., 2003 (cit. on p. 38).
– *Theory of linear and integer programming*. English. Wiley-Interscience Series in Discrete Mathematics. A Wiley-Interscience Publication. Chichester: John Wiley & Sons Ltd., 1986 (cit. on pp. 9, 23, 24, 38, 83, 84, 181–183, 188).
Scott, P. R. "On convex lattice polygons". In: *Bull. Austral. Math. Soc.* 15.3 (1976), pp. 395–399. MR0430960 (55 #3964) 52A10 (10E05), rev. by J. M. Wills (cit. on p. 10).
Sebő, András. "An introduction to empty lattice simplices". In: *Integer programming and combinatorial optimization (Graz, 1999)*. Ed. by Gérard Cornuéjols, Rainer E. Burkard, and Gerhard J. Woeginger. Vol. 1610. Lecture Notes in Comput. Sci. Berlin: Springer, 1999, pp. 400–414. DOI: 10.1007/3-540-48777-8_30. MR1709397 (2000e:90063) 90C27 (90C35). URL: http://dx.doi.org/10.1007/3-540-48777-8_30 (cit. on p. 53).
Stanley, Richard P. "Decompositions of rational convex polytopes". In: *Ann. Discrete Math.* 6 (1980). Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Comb. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978), pp. 333–342. DOI: 10.1016/S0167-5060(08)70717-9. MRMR593545 (82a:52007) 52A43, rev. by P. McMullen (cit. on p. 143).
Storjohann, Arne. *Faster Algorithms for Integer Lattice Basis Reduction*. 1996 (cit. on p. 79).
Tateiwa, Nariaki, Yuji Shinano, Keiichiro Yamamura, Akihiro Yoshida, Shizuo Kaji, Masaya Yasuda, and Katsuki Fujisawa. *CMAP-LAP: Configurable Massively Parallel Solverfor Lattice Problems*. Tech. rep. ZIB-Report 21-16. ZIB Berlin, July 2021 (cit. on p. 57).
Todd, Michael J. and E. Alper Yıldırım. "On Khachiyan's algorithm for the computation of minimum-volume enclosing ellipsoids". In: *Discrete Appl. Math.* 155.13 (2007), pp. 1731–1744. DOI: 10.1016/j.dam.2007.02.013. MR2348357 90C25 (90C05), rev. by Serge G. Kruk (cit. on p. 184).
White, G. K. "Lattice tetrahedra". In: *Canadian J. Math.* 16 (1964), pp. 389–396 (cit. on p. 150).
Wübben, Dirk, Dominik Seethaler, Joakim Jaldén, and Gerald Matz. "Lattice Reduction". In: *IEEE SIGNAL PROCESSING MAGAZINE* 70 (May 2011). DOI: 10.1109/MSP.2010.938758 (cit. on p. 57).
Ziegler, Günter M. *Lectures on polytopes*. Vol. 152. Graduate Texts in Mathematics. New York: Springer-Verlag, 1995, pp. x+370. DOI: 10.1007/978-1-4613-8431-1. MRMR1311028 (96a:52011) 52Bxx, rev. by Margaret M. Bayer (cit. on p. 192).