Distances to Lattice Points in Rational Polyhedra

Aled Williams

2022

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

School of Mathematics
Ysgol Mathemateg

Summary

Let $\boldsymbol{a} \in \mathbb{Z}^n_{>0}$, $n \geq 2$, $\gcd(\boldsymbol{a}) := \gcd(a_1, \ldots, a_n) = 1$, $b \in \mathbb{Z}_{\geq 0}$ and denote by $\|\cdot\|_\infty$ the ℓ_∞-norm. Consider the knapsack polytope

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : \boldsymbol{a}^T \boldsymbol{x} = b\right\}$$

and assume that $P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset$ holds. The main result of this thesis states that for any vertex $\boldsymbol{x}^*$ of the knapsack polytope $P(\boldsymbol{a}, b)$ there exists a feasible integer point $\boldsymbol{z}^* \in P(\boldsymbol{a}, b)$ such that, denoting by s the size of the support of $\boldsymbol{z}^*$, i.e. the number of nonzero components in $\boldsymbol{z}^*$ and upon assuming $s > 0$, the inequality

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^{s-1}}{s} < \|\boldsymbol{a}\|_\infty$$

holds. This inequality may be viewed as a transference result which allows strengthening the best known distance (proximity) bounds if integer points are not sparse and, vice versa, strengthening the best known sparsity bounds if feasible integer points are sufficiently far from a vertex of the knapsack polytope. In particular, this bound provides an exponential in s improvement on the previously best known distance bounds in the knapsack scenario. Further, when considering general integer linear programs, we show that a resembling inequality holds for vertices of Gomory's corner polyhedra [49, 96]. In addition, we provide several refinements of the known distance and support bounds under additional assumptions.

Acknowledgements

Throughout the time I have been working on this thesis it really does amaze me to think about how many parts of my life have changed. Because of this it would be impossible to thank everyone who supported me during this time, however, I would like to acknowledge a small subset of those people here.

First and foremost, I would like to express my profound gratitude to my supervisor, Dr Iskander Aliev. I am incredibly thankful for all your support, encouragement and guidance. I feel very fortunate to have had the opportunity to work with you during this time. I would also like to thank my second supervisor, Dr Timm Oertel, for your willingness to evaluate my work and for the encouragement you have offered throughout.

I would like to thank both Dr Marcel Celaya and Dr Martin Henk who, in addition to Dr Iskander Aliev, their collaboration on [5] provided some of the results appearing in Chapter 4.

I would like to next thank my family who, despite a number of changes which happened throughout the period, supported me wholeheartedly. I am also so grateful for all the love and kindness you have given me in my life. I have no doubt in saying that I would not have completed this thesis without all your love, patience and support over the years.

As for my wife, Isabelle - I do not think anything I could write here would really provide an insight into just how thankful I am for all which you have done. I can confidently say that it is meeting you throughout the writing of this thesis for which I am most thankful for. You have provided the continuous love and support in order for me to not only finish the thesis, but to personally develop through some really testing times.

Finally, I would like to take the opportunity to thank my good friend and peer Abdullah Alasmari for both your support and friendship. Despite coronavirus locking down the world, I am very thankful for each of our remote meetings during the pandemic and I am grateful that I met such a good friend during this time.

Contents

— Chapter 1 —

Introduction

1.1 Motivation

We consider two directions of research in the theory of integer linear programming. In order to introduce these directions we let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ and $\boldsymbol{b} \in \mathbb{Z}^m$ and assume that the polyhedron $P(A, \boldsymbol{b}) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\right\}$ contains integer points. The first direction of interest asks to bound the distance from a vertex of $P(A, \boldsymbol{b})$ to its nearest integer point in $P(A, \boldsymbol{b})$. Such results are useful when one wants to know how far a solution of an integer linear program (IP) is from the corresponding relaxed linear program (LP) [28, 41, 72]. The second direction of research asks to bound the minimum size of support (sparsity) of integer points in $P(A, \boldsymbol{b})$. Note (see e.g. [7]) that polynomial support bounds for integer programming [39, 6] have been utilised successfully in areas including logic and complexity [71, 68], in fixed parameter tractability [58, 59], as a component in the solution of the cutting-stock problem with a fixed number of item types [47] and in compressed sensing [42, 44].

The aforementioned research directions have classically been treated distinctly, however, in the knapsack scenario (i.e. $m = 1$), we prove (Theorem 4.1.6 in Chapter 4) the following result. Let $\boldsymbol{a} \in \mathbb{Z}^n_{>0}$, $n \geq 2$, $\gcd(\boldsymbol{a}) := \gcd(a_1, \ldots, a_n) = 1$, $b \in \mathbb{Z}_{\geq 0}$ and denote by $\|\cdot\|_\infty$ the ℓ_∞-norm. Consider the knapsack polytope

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : \boldsymbol{a}^T\boldsymbol{x} = b\right\}$$

and assume that $P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset$. Then for any vertex $\boldsymbol{x}^*$ of $P(\boldsymbol{a}, b)$ there exists a feasible integer point $\boldsymbol{z}^*$ such that, denoting by s the size of the support of $\boldsymbol{z}^*$, i.e. the number of nonzero components in $\boldsymbol{z}^*$ and upon assuming $s > 0$, the inequality

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^{s-1}}{s} < \|\boldsymbol{a}\|_\infty \tag{1.1}$$

holds. The inequality (1.1) can be viewed as a transference result which allows strengthening of the previously best known distance (proximity) bounds if feasible integer points are not sparse and, vice versa, strengthening the best known sparsity bounds if feasible integer points are sufficiently far from a vertex of the knapsack polytope. In particular, (1.1) provides an exponential in s improvement on the previously best known distance bounds in the knapsack scenario. In light of (1.1) it would be natural to consider if the two research directions are interrelated in the general setting. It should be noted for completeness that Lee et al. [72] follow a similar direction as they apply new sparsity-type bounds in a recent work to refine the distance bounds in certain scenarios.

1.2 Outline

Following reviewing some preliminary material in Chapter 2, we then provide an overview of notable distance and sparsity bounds in Chapter 3. Chapter 4 considers the connection between the distances to and sparsity of solutions to an IP. In particular, we prove that a transference bound for vertices of corner polyhedra holds. Furthermore, we present a resembling result that connects the minimum absolute nonzero entry of an optimal integer solution with the size of its support. Chapter 5 details several refinements of the best known distance and sparsity bounds in special cases.

— Chapter 2 —

Preliminaries

2.1 Basics and Notation

The symbols $\mathbb{R}$, $\mathbb{Q}$ and $\mathbb{Z}$ will be used to denote the real, rational and integer numbers, respectively. The cardinality of a given set U will be denoted by $|U|$. In order to denote that a set U is a subset of a set V we will write $U \subset V$. It should be emphasised that the symbol $\subset$ may include the scenario where the sets U and V coincide, however, this will be clarified given the context.

The greatest common divisor of the n integers $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ with $a_i \neq 0$ for at least one i is the largest common factor of $a_1, a_2, \ldots, a_n$. This will be denoted by $\gcd(a_1, a_2, \ldots, a_n)$ throughout the thesis and can be computed using the Euclidean algorithm (see e.g. [30, Chapter 31]) or for example Lehmer's algorithm [74].

2.2 Norms

A function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is called a *norm* if it satisfies:

(i) $\|\boldsymbol{x}\| \geq 0$ for any vector $\boldsymbol{x} \in \mathbb{R}^n$, and $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$,

(ii) $\|\alpha \boldsymbol{x}\| = |\alpha| \, \|\boldsymbol{x}\|$ for any vector $\boldsymbol{x} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$, and

(iii) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for any vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

The last property is called the *triangle inequality*. The most commonly used norms belong to the family of *p-norms*, or *ℓ_p-norms*, which are defined with $\boldsymbol{x} \in \mathbb{R}^n$ as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p},$$

where $p \geq 1$ is a real number. It should be noted that there is an ℓ_0-"norm" which counts the cardinality of the support. The ℓ_0-"norm" will be discussed in more detail in Section 2.7, however, it is not actually a norm since it does not satisfy property (ii).

The following ℓ_p-norms are of particular interest:

- $p = 1$: the ℓ_1-norm, which is also referred to as both the Taxicab or the Manhattan norm, is
$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|,$$
- $p = 2$: the ℓ_2-norm, which is also referred to as the *Euclidean norm*, is
$$\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}},$$
where $\boldsymbol{x}^T$ denotes the transpose of the vector $\boldsymbol{x}$, and
- $p = \infty$: the ℓ_∞-norm, which is also referred to as the maximum norm, is
$$\|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i|.$$

2.3 Convexity

2.3.1 Convex Sets, Bodies, Combinations and the Convex Hull

A set $C \subset \mathbb{R}^n$ is called *convex* if the line segment between any two points in C lies in C. In other words, a set C is convex if for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ and any $\alpha \in \mathbb{R}$ with $0 \le \alpha \le 1$, we have
$$\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in C.$$

Further, a convex set $K \subset \mathbb{R}^n$ which is compact, i.e. closed and bounded, and has a nonempty interior will be called a *convex body*. Several examples of convex bodies are Euclidean balls, hypercubes and cross-polytopes.

Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ be a finite set of points from $\mathbb{R}^n$. We will call a point
$$\boldsymbol{x} = \sum_{i=1}^m \alpha_i \boldsymbol{x}_i, \quad \text{where } \sum_{i=1}^m \alpha_i = 1 \quad \text{and} \quad \alpha_i \ge 0 \quad \text{for } i = 1, \ldots, m$$
a *convex combination* of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$. In light of this definition, it follows that a set is convex if and only if it contains every convex combination of its points.

The *convex hull* of a set K, denoted by $\operatorname{conv}(K)$, is the intersection of all convex sets in $\mathbb{R}^n$ containing K. In particular, the convex hull of a set K is the smallest convex set with respect to inclusion that contains K (see e.g. [53, Chapter 3]).

2.3.2 The Affine Hull

If $K \subset \mathbb{R}^n$ is a set, then the *affine hull* of K is the smallest affine set in $\mathbb{R}^n$ which contains K. The dimension of K, denoted $\dim(K)$, is the dimension of the affine hull of K, where $\dim(\emptyset) = -1$. It should be noted that the affine hull of two distinct points in $\mathbb{R}^n$ is the unique line connecting them, while, the convex hull of those two points comprises of those two points and the line segment connecting them.

2.3.3 Half-spaces and Supporting Hyperplanes

Let $\boldsymbol{a} \in \mathbb{R}^n$ with $a_i \neq 0$ for at least one $i \in \{1, \dots, n\}$ and $b \in \mathbb{R}$. For an affine hyperplane $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} = b\}$ we associate two open (affine) half-spaces

$$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} > b\} \quad \text{and} \quad \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} < b\}$$

and two (affine) closed half-spaces

$$\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} \geq b\} \quad \text{and} \quad \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} \leq b\}.$$

A half-space will be called a *linear half-space* if $b = 0$. It is worth noting for completeness that an affine hyperplane can be equivalently defined as $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{n} \cdot \boldsymbol{x} = b\}$, where $\boldsymbol{n} \in \mathbb{R}^n$ denotes a (nonzero) normal vector and b may be related (via the normal vector $\boldsymbol{n}$) to the perpendicular distance with respect to the ℓ_2-norm from the origin to the hyperplane.

Let $C \subset \mathbb{R}^n$ be a closed convex set. A hyperplane $\mathscr{H} \subset \mathbb{R}^n$ is called a *supporting hyperplane* for the set C if $\mathscr{H} \cap C \neq \emptyset$ holds and the set C lies in one of the two closed half-spaces bounded by the hyperplane $\mathscr{H}$.

2.3.4 Convex Cones and Polyhedra

A set $C \subset \mathbb{R}^n$ is called a *cone* if for every $\boldsymbol{x} \in C$ and $\alpha \geq 0$ we have $\alpha \boldsymbol{x} \in C$. Further, if a cone C is additionally convex, then it is called a *convex cone*. In particular, a set C is a convex cone if for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ and $\alpha_1, \alpha_2 \geq 0$, we have

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 \in C.$$

Further, we will call a set an *affine convex cone* if the set can be established by applying an affine transformation to a convex cone. Throughout this thesis, we work only with those cones which are convex and, for this reason, we state several further definitions which build upon this notion.

A cone C is *polyhedral* if $C = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{0}\}$ for some matrix A, i.e. if the cone C is the intersection of finitely many linear half-spaces, where each row of the matrix A defines a half-space. It is worth emphasising that a polyhedral cone is necessarily a convex cone, however, the converse is not true in general. Any set which can be expressed as the intersection of a finite number of closed affine half-spaces is called a *polyhedron*. Further, we will call a bounded polyhedron a *polytope*.

2.3.5 Pointed Cones

Let C be a polyhedral cone in $\mathbb{R}^n$, then the *lineality space* of C is the linear subspace

$$\mathrm{Lin}(C) = \{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{x} + \lambda \boldsymbol{y} \in C \text{ for all } \boldsymbol{x} \in C, \lambda \in \mathbb{R}\}.$$

Further, a polyhedral cone C is said to be *pointed* if $\mathrm{Lin}(C) = \{\boldsymbol{0}\}$. It is worth noting that one could equivalently state that a convex cone C is pointed if and only if $C \cap (-C) = \{\boldsymbol{0}\}$ holds, where the notation $-C$ denotes the *negative cone* associated with C, which is the related convex cone $-C = \{-\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{y} \in C\}$.

In light of these definitions, if a pointed convex cone has a vertex, then that vertex must be unique and located at the origin $\boldsymbol{0}$. Furthermore, it follows that if C is a pointed convex cone in $\mathbb{R}^n$, then there exists a hyperplane $\mathscr{H}$ passing through the origin $\boldsymbol{0}$, which separates the convex cone C from its negative cone $-C$ and additionally supports both C and $-C$ simultaneously.

2.3.6 Vertices and Faces of a Polyhedron

Let $A \in \mathbb{Z}^{m \times n}$, $\boldsymbol{b} \in \mathbb{Z}^m$ and let $P = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}$ be a polyhedron. A subset $F \subset P$ is called a *face* of P if $F = P$ or if F is the intersection of P with a supporting hyperplane of P. In other words, F is a face of P if and only if F is nonempty and $F = \left\{\boldsymbol{x} \in P : A'\boldsymbol{x} = \boldsymbol{b}'\right\}$ for some subsystem $A'\boldsymbol{x} \leq \boldsymbol{b}'$ of $A\boldsymbol{x} \leq \boldsymbol{b}$.

A *facet* of P is a maximal face distinct from P relative to inclusion. It should be noted that the dimension of any facet of P is one less than the dimension of P. Further, each face of P, except for P itself, is the intersection of facets of P.

A *minimal face* of the polyhedron P is a face that does not contain any other face. In other words, a face F of P is a minimal face if and only if F is an affine subspace [89, Chapter 8]. Further, each minimal face of the polyhedron P is a translate of the lineality space of P. Note that all minimal faces of P have the same dimension, namely n minus the rank of A. Further, it follows that if the polyhedron P is pointed, then each minimal face consists of just one point. In such case, these points, i.e. the minimal faces of P, are called the *vertices* of the polyhedron P. It is worth emphasising that each vertex corresponds to a basic feasible solution to $A\boldsymbol{x} = \boldsymbol{b}$ and is hence determined by the linearly independent restrictions from the linear system [89, Chapter 8].

2.4 Lattice Preliminaries

2.4.1 Lattices

For linearly independent $\boldsymbol{b}_1, \dots, \boldsymbol{b}_k \in \mathbb{R}^n$, the set

$$\Lambda = \left\{ \sum_{i=1}^{k} x_i \, \boldsymbol{b}_i : x_i \in \mathbb{Z} \right\}$$

is a k-dimensional *lattice* generated with *basis* $\boldsymbol{b}_1, \dots, \boldsymbol{b}_k$. If $B = [\boldsymbol{b}_1, \dots, \boldsymbol{b}_k] \in \mathbb{R}^{n \times k}$ is a matrix with columns $\boldsymbol{b}_1, \dots, \boldsymbol{b}_k$, the lattice Λ can be expressed equivalently as $\Lambda = \left\{ B\boldsymbol{x} : \boldsymbol{x} \in \mathbb{Z}^k \right\}$. If $n = k$, then Λ is called a full rank lattice, as illustrated in Figure 2.1.

From the definition of a lattice given above, it is worth emphasising that the choice of basis is not unique as Figure 2.1 shows that different bases can generate the same lattice. It may seem natural at this point to question how one would go about transforming one basis into another basis which still generates the same lattice. In order to answer this we use the concept of unimodular basis transformation which allows us to easily find different bases that generate the same underlying lattice.

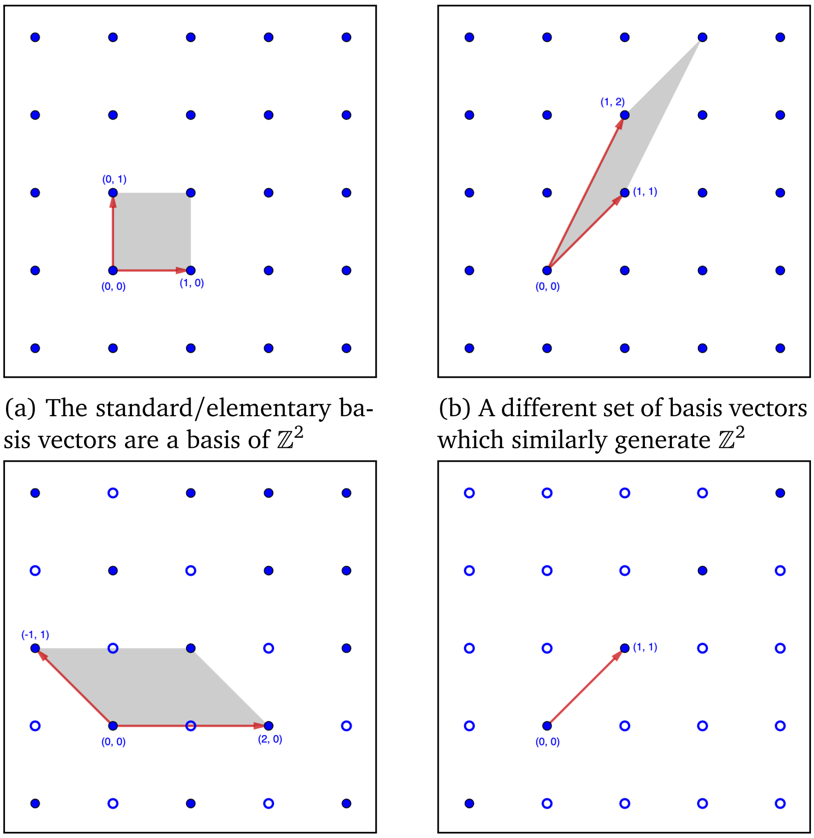
It turns out (see e.g. [89, Chapter 4]) that we can change the basis of a lattice Λ by simply multiplying the matrix B from the right by an appropriately sized *unimodular* matrix $U \in \mathbb{Z}^{k \times k}$, which is a square integral matrix whose determinant is ± 1. In particular, if the equality $B' = BU$ holds, then the columns from B and B' generate the same lattice Λ. Despite the fact that lattices are invariant under performing unimodular transformations, it is worth emphasising that the length of the basis vectors (with respect to any choice of norm) will generally vary under such operations, as shown in Figure 2.1.

2.4.2 The Determinant of a Lattice

In order to define the determinant of a lattice we firstly will introduce the *fundamental parallelepiped* of a given lattice. Recall that $\Lambda \subset \mathbb{R}^n$ denotes the lattice generated by the basis $B = [\boldsymbol{b}_1, \dots, \boldsymbol{b}_k] \in \mathbb{R}^{n \times k}$. The fundamental parallelepiped of the lattice Λ associated with basis B is

$$\mathscr{P}(B) = \{ B\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n, 0 \le x_i < 1 \}.$$

It is worth emphasising that because the column vectors in B provide a basis of Λ, it follows immediately that the fundamental parallelepiped $\mathscr{P}(B)$ does not contain any



(a) The standard/elementary basis vectors are a basis of $\mathbb{Z}^2$

(b) A different set of basis vectors which similarly generate $\mathbb{Z}^2$

(c) Two linearly independent vectors that are not a basis for $\mathbb{Z}^2$

(d) The single vector does not generate a full rank lattice

Figure 2.1: Several examples demonstrating how linearly independent sets of integral vectors (red) generate integral lattices (blue) and, in addition, how the basis vectors correspond to the fundamental parallelepipeds (grey). It should be noted that the linearly independent vectors in (c) do not provide a basis for $\mathbb{Z}^2$, however, they do indeed provide a basis for some integral sublattice of $\mathbb{Z}^2$.

lattice points other than the origin, i.e. $\mathscr{P}(B) \cap \Lambda = \{\mathbf{0}\}$. Further, the fundamental parallelepiped $\mathscr{P}(B)$ clearly depends on the basis B that generates the lattice Λ.

The *determinant* of the lattice Λ generated by basis B is the k-dimensional volume of the fundamental parallelepiped $\mathscr{P}(B)$ associated to B and is calculated [51] by

$$\det(\Lambda) = \mathrm{vol}_k(\mathscr{P}(B)) = \sqrt{\det(B^T B)},$$

where B^T denotes the transpose of B and $\mathrm{vol}_k(\cdot)$ denotes the k-dimensional volume or Lebesgue measure (see e.g. [14, Chapter 13]). In the special case that Λ is a full rank lattice (where $n = k$) then the matrix B is a square matrix and hence the determinant of the lattice is given by $\det(\Lambda) = |\det(B)|$, where $|\cdot|$ denotes the absolute value.

In order to show that the determinant is well defined assume that $B_1, B_2 \in \mathbb{R}^{n\times k}$ are two bases that generate the same lattice Λ. In light of this we know that the equality $B_2 = B_1 U$ holds for some unimodular matrix $U \in \mathbb{Z}^{k\times k}$. Considering the determinant $\det(\Lambda)$ associated with the bases B_1 and B_2 yields that

$$\det(\Lambda) = \sqrt{\det(B_2^T B_2)} = \sqrt{\det(U^T B_1^T B_1 U)} = \sqrt{\det(B_1^T B_1)} = \det(\Lambda)$$

holds and, in consequence, we conclude that the determinant of a lattice is well defined in the sense that its value is invariant on the basis of the lattice Λ.

2.5 Sparsity and the ℓ_0-“norm”

The *sparsity* of a vector provides a description of how many of the entries from the vector are zero. For $\boldsymbol{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, we will denote by

$$\operatorname{supp}(\boldsymbol{x}) = \{i \in \{1, \dots, n\} : x_i \neq 0\}$$

the *support* of $\boldsymbol{x}$. It is worth emphasising that the support of $\boldsymbol{x}$ is an index set associated with $\boldsymbol{x}$ which includes the positions/labels of all nonzero entries in the given vector. For example, given the vector $(-1, 7, 0, 0, 11)^T$, its support is $\operatorname{supp}\big((-1, 7, 0, 0, 11)^T\big) = \{1, 2, 5\}$. This allows us to state the ℓ_0-“norm” of the vector $\boldsymbol{x}$ is precisely the cardinality of its support, namely $\|\boldsymbol{x}\|_0 = |\operatorname{supp}(\boldsymbol{x})|$.

The ℓ_0-“norm” is a function widely used in the theory of compressed sensing (see e.g. [44, Chapter 2]). It is known [44] that finding a solution to a system of linear equations with minimal support even in the continuous case, i.e. solving the ℓ_0-minimisation problem, is $\mathcal{NP}$-hard [79].

2.6 Laplace’s Expansion Formula

The *Laplace expansion formula*, also known as the cofactor expansion, is an expression for the determinant of B, namely $\det(B)$, of an n-dimensional matrix B. The formula expresses $\det(B)$ as a weighted sum of the determinants of precisely n submatrices of B, each of which is a square matrix containing exactly $n-1$ rows and columns. The Laplace expansion formula is of interest as one of several ways to compute the determinant of a matrix, however, for large matrices the formula becomes inefficient in comparison to other methods because of its $\mathcal{O}(n!)$ worst-case time complexity. Let $\bar{B}_{ij}$ denote the submatrix of B which is formed by deleting the i-th row and j-th column from B.

Theorem 2.6.1 (Laplace's expansion formula). *Suppose $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ is a square n-dimensional matrix and fix any $i, j \in \{1,2,\ldots,n\}$. Then the determinant of B is given by*

$$\det(B) = \sum_{i=1}^{n}\Big((-1)^{i+j}\, b_{ij}\, \det(\bar{B}_{ij})\Big) = \sum_{j=1}^{n}\Big((-1)^{i+j}\, b_{ij}\, \det(\bar{B}_{ij})\Big),$$

which denotes expansion along the j-th column and i-th row of B, respectively.

2.7 Cramer's Rule

Given a square matrix $A \in \mathbb{Z}^{n\times n}$ with full row rank and a vector $\boldsymbol{b} \in \mathbb{Z}^n$, which both contain integral entries for simplicity, it is well-known that one can solve the corresponding matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ within polynomial time using various algorithms (see e.g. [30, Chapter 28]). One such result used throughout this thesis for solving suitable linear systems is the celebrated *Cramer's rule*. This result is named after Gabriel Cramer, who published the rule for an arbitrary number of unknowns in 1750 [32], however, despite its name, it should be noted that Maclaurin published special cases of the rule slightly earlier in 1748 [78]. Cramer's rule provides an explicit formula for the solution of a square system of linear equations, provided that the square constraint matrix is invertible. The formula depends upon particular determinants of the constraint matrix A and the right-hand side vector $\boldsymbol{b}$.

Lemma 2.7.1 (Cramer's Rule). *Let $A \in \mathbb{Z}^{n\times n}$ and $\boldsymbol{b} \in \mathbb{Z}^n$. The components of the solution to a linear system of the form $A\boldsymbol{x} = \boldsymbol{b}$ when $\det(A) \neq 0$ are given by*

$$x_k = \frac{\det(A_k(\boldsymbol{b}))}{\det(A)}, \quad k = 1,\ldots,n,$$

where $A_k(\boldsymbol{b})$ denotes the matrix obtained by replacing the entries in the k-th column of A by the entries in the column vector $\boldsymbol{b}$.

Cramer's rule is a result primarily of theoretical interest, however, it has recently been shown [54] that one can implement the method with time complexity $\mathcal{O}(n^3)$. This means that Cramer's rule is comparable to other methods commonly used for solving such systems of linear equations.

2.8 Linear and Integer Linear Programming

2.8.1 An Introduction to Linear and Integer Linear Programming

Linear programming concerns the problem of maximising or minimising a linear function over a polyhedron. The idea of linear programming is apparent in the work of Fourier

[45] in 1827, but the development of this discipline in addition to a widespread acknowledgment of its importance followed much later in consequence to the celebrated work of Dantzig [34, 36, 33], Hitchcock [56], Kantorovich [61, 62, 89], Koopmans [70] and von Neumann [82, 80, 81].

Let $A \in \mathbb{Z}^{m \times n}$, $\boldsymbol{b} \in \mathbb{Z}^m$ and $\boldsymbol{c} \in \mathbb{Q}^n$. It is well-known (see e.g. [89, Chapter 7]) that there are several equivalent ways to formulate the linear programming problem, however, throughout this thesis we will only work with linear programs (LPs) that are presented in *standard* form, namely

$$\max\left\{\boldsymbol{c}^T\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{R}^n_{\geq 0}\right\}. \tag{2.1}$$

This problem asks to find an optimal nonnegative solution $\boldsymbol{x} \in \mathbb{R}^n_{\geq 0}$ which satisfies the system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ (known as the *constraints* of the problem) and maximises the linear function $\boldsymbol{c}^T\boldsymbol{x}$ (known as the *objective* or cost function). It is worth emphasising that an optimal solution to (2.1) does not necessarily exist in general.

Khachiyan [66] first demonstrated in 1979 that a LP is solvable in polynomial time through the introduction of the *ellipsoid method*. Shortly after this breakthrough, Karmarkar [64, 86] introduced a novel *interior-point* method for solving LPs in polynomial time that was arguably of greater theoretical and practical importance. Further, the celebrated *simplex method*, which was developed around 1947 by Dantzig [34, 35], is currently used to solve large-scale LPs in various applications despite not yielding a polynomial worst-case time complexity.

Integer linear programming concerns the problem of maximising or minimising a linear function over a polyhedron, where the solution is additionally required to contain only integer entries. In a similar fashion there are several equivalent ways to formulate an integer linear programming problem, however, we only work with integer linear programs (IPs) that have been presented in standard form, namely

$$\max\left\{\boldsymbol{c}^T\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}\right\}. \tag{2.2}$$

The LP (2.1) is often called the linear programming relaxation of the IP (2.2) because one obtains the LP (2.1) by “relaxing” (or dropping) the integrality constraint in the IP (2.2). Furthermore, one noteworthy special case of the IP (2.2) and the LP (2.1) is known as the *knapsack problem*, which is the case when the problem features just one linear constraint (i.e. when $m = 1$). It should be noted that when we later consider the knapsack scenario we will make use of traditional vector notation in order to distinguish this setting from the general case by replacing A and $\boldsymbol{b}$ by $\boldsymbol{a} = (a_1, a_2, \dots, a_n)^T \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$, respectively.

Despite the seeming similarity to the LP (2.1), solving the IP (2.2) is in general a significantly more difficult task and, in particular, solving the decision version of the IP (2.2) in general is known to be $\mathscr{NP}$-complete [89, Chapter 18]. However, in 1983 it was shown by Lenstra [75] that any IP given in inequality form with a fixed number of variables n can be solved in polynomial time. In light of the aforementioned $\mathscr{NP}$-completeness, it is common within the domain of computational mathematics when given such a problem to obtain an approximate solution by solving some relaxed yet related problem. Throughout this thesis, we approximate solutions to (2.2) by solving the related linear programming relaxation (2.1). The research of Ford and Fulkerson [43] tells us that the optimal solution to the LP (2.1) and IP (2.2) coincide in certain scenarios, meaning that the IP (2.2) is sometimes polynomial solvable by making use of the aforementioned linear programming relaxation method. In particular, this occurs when considering various combinatorial optimisation problems involving networks including the shortest-path [1], the maximum-flow [43, 46], the assignment [3, 2], the minimum cost flow [84] and the transportation [69] problems.

2.8.2 Geometry of Linear and Integer Linear Programming

The constraints of the linear programming problem define the problem's *feasible region*, which is the space containing all *feasible solutions* to the LP. The feasible region defined by the linear constraints in (2.1) is a convex polyhedron and, therefore, the feasible region may be characterised by its facets or vertices. Being a little more precise, geometrically the set containing of all feasible solutions to (2.1) is the polyhedron

$$P(A, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\}\,. \tag{2.3}$$

Recall that an optimal solution to the LP (2.1) does not necessarily exist in general. In particular, if the constraints of the linear system $A\boldsymbol{x} = \boldsymbol{b}$ are inconsistent, then the polyhedron $P(A, \boldsymbol{b})$ is empty. Furthermore, if $P(A, \boldsymbol{b})$ is unbounded, then in certain situations the corresponding objective function value will be unbounded and, in consequence, no optimal solution to (2.1) exists.

In the scenario where the linear system appearing in the LP (2.1) is consistent, then it is well-known (see e.g. [89, Chapter 8]) that for any given rational objective function $\boldsymbol{c} \in \mathbb{Q}^n$, there either exists at least one optimal solution to the LP (2.1) positioned in at least one vertex of the polyhedron $P(A, \boldsymbol{b})$ or the maximum is infinite. Further, if the given instance of the problem yields more than one optimal solution, then since $P(A, \boldsymbol{b})$ is convex, it follows that each optimal solution lies upon the same facet of the feasible region, where one ignores linearly dependent facets and, in consequence, all feasible points along that facet are additionally optimal solutions to the LP (2.1).

In light of this, throughout this thesis, we will focus on those optimal solutions located at the vertices of the feasible region (2.3). In particular, we will call such an optimal solution an *optimal vertex solution* and will denote such a solution by $\boldsymbol{x}^*$.

The *integer hull* $P_I(A, \boldsymbol{b})$ of the polyhedron $P(A, \boldsymbol{b})$ is the convex hull of all integer points in $P(A, \boldsymbol{b})$, namely

$$P_I(A, \boldsymbol{b}) = \operatorname{conv}\left(\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\} \cap \mathbb{Z}^n\right) = \operatorname{conv}\left(P(A, \boldsymbol{b}) \cap \mathbb{Z}^n\right).$$

Note that each feasible solution to the IP (2.2) must be an element of the integer hull $P_I(A, \boldsymbol{b})$ because $P_I(A, \boldsymbol{b})$ is defined as the convex hull of all nonnegative integer solutions from $P(A, \boldsymbol{b})$. Furthermore, the vertices of $P_I(A, \boldsymbol{b})$ play an important role because analogously to the LP (2.1), by making use of standard linear programming theory, we know that the set of optimal solutions to IP (2.2) must contain at least one vertex of the integer hull $P_I(A, \boldsymbol{b})$ and, vice versa, each vertex of the integer hull is an optimal solution for some rational objective function $\boldsymbol{c}$. It for this reason that vertices of the integer hull are of special interest in the theory of integer optimisation.

2.9 Corner Polyhedra

Let $A \in \mathbb{Z}^{m \times n}$ where $m < n$, $\boldsymbol{b} \in \mathbb{Z}^m$ and $\boldsymbol{c} \in \mathbb{Q}^n$. We assume without loss of generality that the integral matrix A has full row rank m, i.e. that the dimension of the row space of A is m. It is well-known (see e.g. [91, Chapter 3]) that row rank is equal to column rank and as such we will simply state that A has full rank. Further, let $\tau = \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ with $i_1 < i_2 < \cdots < i_k$. We will use the notation A_τ for the $m \times k$ submatrix of A with columns indexed by τ. In the same manner, given $\boldsymbol{x} \in \mathbb{R}^n$, we will denote by $\boldsymbol{x}_\tau$ the vector $(x_{i_1}, \ldots, x_{i_k})^T$. The compliment of τ will be denoted by $\bar{\tau}$. We will say that τ is a *basis* of A if $|\tau| = m$ and the submatrix A_τ is nonsingular. Throughout this thesis, the notation γ will be used to denote when the indexed columns provide a basis of A.

Let A_γ be any nonsingular submatrix consisting of m columns from A. Upon reordering the columns of A if necessary, we can assume without loss of generality that A_γ consists of the first m columns of A, i.e. $\gamma = \{1, \ldots, m\}$ and, in consequence, the matrix A has the form

$$A = (A_\gamma, A_{\bar{\gamma}}).$$

This allows us to express the IP (2.2) in terms of A_γ and $A_{\bar{\gamma}}$ as

$$\max\left\{\boldsymbol{c}_\gamma^T \boldsymbol{x}_\gamma + \boldsymbol{c}_{\bar{\gamma}}^T \boldsymbol{x}_{\bar{\gamma}} : A_\gamma \boldsymbol{x}_\gamma + A_{\bar{\gamma}} \boldsymbol{x}_{\bar{\gamma}} = \boldsymbol{b},\, \boldsymbol{x} = (\boldsymbol{x}_\gamma, \boldsymbol{x}_{\bar{\gamma}})^T \in \mathbb{Z}^n_{\geq 0}\right\}, \tag{2.4}$$

where $\boldsymbol{x}_\gamma$ and $\boldsymbol{x}_{\bar{\gamma}}$ are called the basic and nonbasic variables, respectively (see e.g. [97, Chapter 2]).

Given a basis γ, the corresponding basic variables can be uniquely determined by

$$\boldsymbol{x}_\gamma = A_\gamma^{-1}\left(\boldsymbol{b} - A_{\bar{\gamma}}\boldsymbol{x}_{\bar{\gamma}}\right)$$

and, in particular, if $\boldsymbol{x} = (\boldsymbol{x}_\gamma, \boldsymbol{x}_{\bar{\gamma}})^T$ is to satisfy $A\boldsymbol{x} = \boldsymbol{b}$ such that $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$, a nonnegative integer vector $\boldsymbol{x}_{\bar{\gamma}}$ must be chosen such that the corresponding basic variables $\boldsymbol{x}_\gamma$ are both integral and nonnegative.

In the seminal paper of Gomory [49], the nonnegativity condition placed on the basic variables $\boldsymbol{x}_\gamma$ is dropped and, in particular, Gomory considers the relaxation

$$A_\gamma \boldsymbol{x}_\gamma + A_{\bar{\gamma}}\boldsymbol{x}_{\bar{\gamma}} = \boldsymbol{b}\,,\ \boldsymbol{x}_{\bar{\gamma}} \geq \boldsymbol{0},\ (\boldsymbol{x}_\gamma, \boldsymbol{x}_{\bar{\gamma}})^T \in \mathbb{Z}^n. \tag{2.5}$$

This relaxation allows us to define the *corner polyhedron* as the convex hull of the integer solutions to (2.5). This geometric object is sometimes called the corner relaxation associated with the basis γ [22, 15]. We can equivalently define corner polyhedron $\mathscr{C}_\gamma(A, \boldsymbol{b})$ associated with the basis γ of the matrix A as

$$\mathscr{C}_\gamma(A,\boldsymbol{b}) = \operatorname{conv}\left(\left\{\boldsymbol{x} = (\boldsymbol{x}_\gamma, \boldsymbol{x}_{\bar{\gamma}})^T \in \mathbb{Z}^n : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x}_{\bar{\gamma}} \geq \boldsymbol{0}\right\}\right).$$

It should be noted that we introduce this relaxation more precisely later by utilising the fact that A_γ is nonsingular, meaning that a suitable projection from $\mathbb{R}^n$ to $\mathbb{R}^{n-m}$ onto the nonbasic variables and, vice versa, is bijective under certain assumptions. Further, in Chapter 4 we define a corner polyhedron in a more general way, where the nonnegativity restriction placed upon the nonbasic variables $\boldsymbol{x}_{\bar{\gamma}}$ is extended to also restrict some of the basic variables $\boldsymbol{x}_\gamma$ under appropriate assumptions, however, for simplicity, we discuss this in much more detail later.

This geometric object $\mathscr{C}_\gamma(A, \boldsymbol{b})$ has been studied in the integer linear programming literature as a framework for deriving and evaluating practical cutting planes [50, 25], mixed-integer cutting planes [31, 26, 37, 17, 18], in addition to offering a substitute to LP relaxations in branch-and-bound [11] and branch-and-cut [12] techniques. For completeness, note that other famous cuts include the Chvátal-Gomory cuts [48, 23, 76, 77, 38], implied bounds [57] and the more general split cuts [10, 29, 13].

— Chapter 3 —

Previously Known Distance and Sparsity Bounds

In this chapter we provide an overview of the best known distance and sparsity bounds.

3.1 Previously Known Distance (Proximity) Bounds

Given $A = (a_{ij}) \in \mathbb{Z}^{m\times n}$, where a_{ij} denotes the entry in the i-th row and j-th column of A with full row rank, $\boldsymbol{b} \in \mathbb{Z}^m$ and $\boldsymbol{c} \in \mathbb{Q}^n$, consider the integer linear program (IP)

$$\max\left\{\boldsymbol{c}^T\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}\right\}. \tag{3.1}$$

We obtain the linear programming relaxation of the IP (3.1) by "relaxing" (or dropping) the integrality constraint in (3.1), which yields the linear program (LP)

$$\max\left\{\boldsymbol{c}^T\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{R}^n_{\geq 0}\right\}. \tag{3.2}$$

The optimal values of the IP (3.1) and the LP (3.2) will be denoted by $IP(\boldsymbol{c},A,\boldsymbol{b})$ and $LP(\boldsymbol{c},A,\boldsymbol{b})$, respectively. It is worth emphasising that $IP(\boldsymbol{c},A,\boldsymbol{b})$ and $LP(\boldsymbol{c},A,\boldsymbol{b})$ are the maximum objective function values associated with (3.1) and (3.2), respectively.

Recall from Chapter 2 that $LP(\boldsymbol{c},A,\boldsymbol{b})$ can be found in polynomial time [67, 64], however, in general solving the decision version of (3.1) is $\mathcal{NP}$-complete [89, Chapter 18]. It seems rather natural to therefore compare $IP(\boldsymbol{c},A,\boldsymbol{b})$ and $LP(\boldsymbol{c},A,\boldsymbol{b})$. In particular, our focus will be first on bounding the distance between the solutions in $\mathbb{R}^n$ which yield the optimal values $IP(\boldsymbol{c},A,\boldsymbol{b})$ and $LP(\boldsymbol{c},A,\boldsymbol{b})$ upon substitution into the given linear objective function $\boldsymbol{c}^T\boldsymbol{x}$.

In order to state the previous proximity bounds we assume that the IP (3.1) is feasible and bounded. Let $\boldsymbol{x}^*$ be an optimal vertex solution corresponding to the optimal value $LP(\boldsymbol{c},A,\boldsymbol{b})$ and let $\boldsymbol{z}^*$ be an optimal integral solution which corresponds to $IP(\boldsymbol{c},A,\boldsymbol{b})$. Let $\Delta_k(A)$ with $k \in \{1,\ldots,m\}$ denote the maximum absolute value of all k-dimensional subdeterminants of the matrix A, namely

$$\Delta_k(A) = \max\{|\det(B)| : B \text{ is a } k \times k \text{ submatrix of } A\}.$$

In particular, note that $\Delta_1(A) = \|A\|_\infty = \max_{i,j} |a_{ij}|$ is the maximum absolute entry of the matrix A. Further, in order to simplify the notation we let $\Delta(A) = \Delta_m(A)$. The notation $\log(\cdot)$ will denote the logarithm with base two.

Blair and Jeroslow [19, 20] demonstrated that the distance between the optimal solutions $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ with respect to any norm can be bounded independently of $\boldsymbol{b}$. This was improved by the classical sensitivity theorem of Cook et al. [28] which states that

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq n \cdot \Delta(A)$$

holds. It should be noted that the bound of Cook et al. [28] holds for general polyhedra and depends upon $\max\{\Delta_k : k \in \{1,\dots,m\}\}$ instead of $\Delta(A)$, however, simply using $\Delta(A)$ is sufficient [72]. Furthermore, Lee et al. [72] tell us that a closer analysis of the proof given by Cook et al. [28] yields an upper bound with respect to the ℓ_1-norm, namely

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 \leq (m+1)\, n \cdot \Delta(A).$$

Recently, Eisenbrand and Weismantel [41] improved upon the aforementioned results of Cook et al. [28] through a novel use of Steinitz Lemma [93] in order to show that the distance between $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ with respect to the ℓ_1-norm satisfies

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 \leq m\,(2m\|A\|_\infty + 1)^m. \tag{3.3}$$

In particular, notice that the bound (3.3) is independent of the dimension n. Furthermore, Lee et al. [72] noted that simply applying a different norm when making use of Steinitz Lemma allow us to restate (3.3) as

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 \leq m\,(2m+1)^m \cdot \Delta(A).$$

Lee et al. [72] recently demonstrate that

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < 3m^2 \log\left(2\sqrt{m} \cdot \Delta(A)^{1/m}\right) \cdot \Delta(A)$$

holds by making use of bounds on the sparsity of $\boldsymbol{z}^*$. It should be emphasised that this upper bound demonstrates that $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1$ is bounded by a polynomial which depends on m and $\Delta(A)$. Further, Lee et al. [73] very recently show that

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 \leq m(m+1)^2 \cdot \Delta^3(A) + (m+1) \cdot \Delta(A)$$

holds in a slightly more general setting.

Furthermore, in the *knapsack scenario*, i.e. when $m = 1$, upon following traditional vector notation, Aliev et al. [8] show that the distance from $\boldsymbol{x}^*$ to a nearby feasible integral solution $\boldsymbol{z}^*$ with respect to the ℓ_∞-norm is bounded by

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$$

and that this bound is optimal.

It is worth noting that providing worst-case upper bounds on the distance between $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ is not the only direction of research. In particular, the research domain entitled parametric integer programming studies how the distance varies on the average as one varies the vector $\boldsymbol{b}$ (see e.g. [83, 8, 85, 40]). Further, there has been research into placing lower bounds on the aforementioned worst-case upper bounds between $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ (see e.g. [8, 72, 16]).

3.2 Previously Known Sparsity Bounds

Recall that the support of $\boldsymbol{x} \in \mathbb{R}^n$ is $\mathrm{supp}(\boldsymbol{x}) = \{i \in \{1, \ldots, n\} : x_i \neq 0\}$ and the ℓ_0-"norm" of $\boldsymbol{x}$ is the cardinality of its support, namely $\|\boldsymbol{x}\|_0 = |\,\mathrm{supp}(\boldsymbol{x})|$.

It turns out that sparsity is a topic of interest in several areas of optimisation. In the theory of compressed sensing, given a certain underdetermined system of linear equations, a central problem is the ℓ_0-minimisation problem (over the reals), which asks to find the most sparse solution $\boldsymbol{x}$ that is consistent with $A\boldsymbol{x} = \boldsymbol{b}$. The ℓ_0-minimisation problem is precisely the problem

$$\min\{\|\boldsymbol{x}\|_0 : A\boldsymbol{x} = \boldsymbol{b}\}. \tag{3.4}$$

This problem (3.4) is in general unfortunately $\mathcal{NP}$-hard [79] and, therefore, in order to obtain a solution efficiently some method of approximation is required. For completeness, note that despite the apparent complexity of (3.4), one very popular provably effective method of solution is known as basis pursuit or ℓ_1-minimisation (see e.g. [44, Chapter 4]) and, under certain conditions (see e.g. [42, Chapter 1]), the optimal solutions to the basis pursuit and ℓ_0-minimisation (3.4) problems are unique and coincide.

In the continuous setting the celebrated classical theorem of Carathéodory implies that each optimal vertex (i.e. basic feasible) solution $\boldsymbol{x}^*$ to the LP (3.2) satisfies

$$\|\boldsymbol{x}^*\|_0 \leq m\,.$$

In particular, this implies that the optimal solution to (3.4) is bounded by m provided the linear system is consistent. It is worth noting that this classical result is well-known in the theory of linear programming because each optimal vertex solution $\boldsymbol{x}^*$ corresponds to a vertex of the polyhedron $P(A,\boldsymbol{b}) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\right\}$, where each vertex of $P(A,\boldsymbol{b})$ is known to be determined by a basis of the matrix A.

It would seem natural here to question how much larger can the minimum support of an integer solution to the linear system $A\boldsymbol{x} = \boldsymbol{b}$ be in comparison to the LP (3.2). In particular, we are interested in the integral ℓ_0-minimisation problem

$$\min\left\{\|\boldsymbol{x}\|_0 : A\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}\right\}. \tag{3.5}$$

Notice that (3.5) is $\mathcal{NP}$-hard since deciding the feasibility of $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ [89, Chapter 18] or solving the ℓ_0-minimisation problem (3.4) is $\mathcal{NP}$-hard [79].

Despite the relative ease in which one obtained an upper bound for (3.4) using the aforementioned classical result of Carathéodory, the story is unfortunately significantly more complicated when considering the related problem (3.5). There has however been a great deal of research into bounding the optimal value of (3.5) and it turns out that the upper bound is not significantly larger than the bound of m attained for (3.4). In order to state the previously known sparsity bounds we assume $\{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}\} \cap \mathbb{Z}^n_{\geq 0} \neq \emptyset$.

Eisenbrand and Shmonin [39] show that there exists some $\boldsymbol{x} \in \mathbb{Z}^n_{\geq 0}$ such that the minimal value of the integral ℓ_0-minimisation problem (3.5) is bounded by

$$\|\boldsymbol{x}\|_0 \leq 2m\log(4m\|A\|_\infty). \tag{3.6}$$

It should be noted that earlier bounds on the minimal value of (3.5) date back to Cook, Fonlupt and Schrijver [27] and Sebö [90] who additionally assume that the cone generated by the columns of A is pointed and those columns of A form a Hilbert basis for the cone. This bound (3.6) was strengthened by Aliev et al. [6, 7] who show that the optimal value of (3.5) is bounded by

$$\|\boldsymbol{x}\|_0 \leq m + \left\lfloor \log\left(\frac{\sqrt{\det(AA^\top)}}{\gcd(A)}\right)\right\rfloor \leq 2m\log(2\sqrt{m}\|A\|_\infty), \tag{3.7}$$

where $\gcd(A)$ denotes the greatest common divisor of all m-dimensional subdeterminants of A provided that $\Delta(A)$ is positive. Lee et al. [72] recently show that the optimal value of (3.5) additionally satisfies

$$\|\boldsymbol{x}\|_0 < 2m\log\left(\sqrt{2m}\cdot\Delta(A)^{1/m}\right).$$

In the scenario when the columns of the matrix A positively span $\mathbb{R}^m$, the aforementioned bounds were recently strengthened by Aliev et al. [4] who show that one can construct an integral solution satisfying

$$\|\boldsymbol{x}\|_0 \leq 2m + \Omega_m\left(\frac{|\det(A_\tau)|}{\gcd(A)}\right) \tag{3.8}$$

within polynomial time, where $\Omega_m(\cdot)$ denotes the truncated prime Ω-function (i.e. the Ω-function from number theory [55, Chapter 22] with a lower threshold of m) and τ denotes a selection of m columns from A such that the square matrix A_τ is nonsingular. It should be noted that the value of the function $\Omega_m(z)$ is bounded from above by the logarithmic function $\log(z)$ for every $z \in \mathbb{Z}_{>0}$ and, further, is much smaller than $\log(z)$ on the average since as $z \to \infty$, the average values of $\frac{1}{z}(\Omega(1) + \cdots + \Omega(z))$ are known [55, Chapter 22] to have order $\log\log z$. Furthermore, the Cauchy-Binet formula (see e.g. [91, Chapter 2]) implies that $|\det(A_\tau)| \leq \sqrt{\det(AA^\top)}$ and, in consequence, when m is fixed, the bound (3.8) is generally tighter than the upper bound (3.7).

In the knapsack scenario, (3.7) demonstrates that there exists an integral solution $\boldsymbol{x}$ satisfying

$$\|\boldsymbol{x}\|_0 \leq 1 + \left\lfloor \log\left(\frac{\|\boldsymbol{a}\|_2}{\gcd(\boldsymbol{a})}\right)\right\rfloor. \tag{3.9}$$

It was shown by Aliev et al. [6] that if all the components of $\boldsymbol{a}$ have the same sign, then the upper bound (3.9) can be strengthened by replacing the ℓ_2-norm of the vector $\boldsymbol{a}$ with the ℓ_∞-norm. Furthermore, Aliev et al. [4] recently show that under the same assumptions there exists an integral solution satisfying

$$\|\boldsymbol{x}\|_0 \leq 1 + \left\lfloor \log\left(\frac{\min\{a_1,\ldots,a_n\}}{\gcd(\boldsymbol{a})}\right)\right\rfloor.$$

Finally, if instead $\boldsymbol{a}$ is a vector containing both positive and negative entries, then it follows from (3.8) that one can construct within polynomial time an integral solution $\boldsymbol{x}$ satisfying

$$\|\boldsymbol{x}\|_0 \leq 2 + \min\left\{\omega\left(\frac{|a_i|}{\gcd(\boldsymbol{a})}\right) : i \in \{1,\ldots,n\}\right\}, \tag{3.10}$$

where $\omega(\cdot)$ denotes the prime-ω-function from number theory [55, Chapter 22].

— Chapter 4 —

Distance-Sparsity Transference for Vertices of Corner Polyhedra

4.1 Introduction and Statement of Results

The main contribution of this chapter shows that a relation between two well-established areas of research holds, namely the proximity and sparsity of solutions to integer linear programs (IPs), when considering Gomory's corner polyhedra [49]. This chapter is based on the collaboration with Iskander Aliev, Marcel Celaya and Martin Henk [5].

In order to state the main results of this chapter, we need the notation that was introduced in Chapter 2. Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$, and let $\tau = \{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ with $i_1 < i_2 < \cdots < i_k$. Recall that we use the notation A_τ for the $m \times k$ submatrix of A with columns indexed by τ. In the same manner, given $\boldsymbol{x} \in \mathbb{R}^n$, we will denote by $\boldsymbol{x}_\tau$ the vector $(x_{i_1}, \dots, x_{i_k})^T$. The complement of τ in $\{1, \dots, n\}$ will be denoted by $\bar{\tau} = \{1, \dots, n\}\backslash\tau$. We will say that τ is a *basis* of A if $|\tau| = m$ and the submatrix A_τ is nonsingular, i.e. $\det(A_\tau) \neq 0$. When τ is a basis, we will replace τ by γ. Recall that $\Delta(A)$ denotes the maximum absolute m-dimensional subdeterminant of A, namely

$$\Delta(A) = \max\{|\det(A_\tau)| : \tau \subset \{1, \dots, n\} \text{ with } |\tau| = m\}.$$

When $\Delta(A)$ is positive, the notation $\gcd(A)$ will denote the greatest common divisor of all m-dimensional subdeterminants of A. The notation $\log(\cdot)$ will be used throughout this chapter to denote the logarithm with base two.

Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ and $\boldsymbol{b} \in \mathbb{Z}^m$. Without loss of generality, we will assume that the matrix A has full rank m. Consider the polyhedron

$$P(A, \boldsymbol{b}) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\right\}$$

and, assuming $P(A, \boldsymbol{b})$ is not empty, take any vertex $\boldsymbol{x}^*$ of $P(A, \boldsymbol{b})$. Since the matrix A has rank m by assumption, there is a basis γ of A such that

$$\boldsymbol{x}^*_\gamma = A_\gamma^{-1}\boldsymbol{b} \text{ and } \boldsymbol{x}^*_{\bar{\gamma}} = \boldsymbol{0}. \tag{4.1}$$

Observe that in general, for a given vertex $\boldsymbol{x}^*$ of $P(A, \boldsymbol{b})$, there can be many choices for the basis γ in (4.1). However, if one assumes that $\boldsymbol{x}^*$ is nondegenerate; that is, if the size of the support of $\boldsymbol{x}^*$ is m, then there is a unique choice for the basis γ, namely when $\gamma = \operatorname{supp}(\boldsymbol{x}^*)$.

Recall that for a set $S \subset \mathbb{R}^n$, we denote by $\operatorname{conv}(S)$ the *convex hull* of S. Let τ be a subset of $\{1, \ldots, n\}$. Following Gomory [49] and Thomas [96], we define the *corner polyhedron* $C_\tau(A, \boldsymbol{b})$ associated with τ as

$$\mathscr{C}_\tau(A, \boldsymbol{b}) = \operatorname{conv}(\{\boldsymbol{x} \in \mathbb{Z}^n : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x}_{\bar{\tau}} \geq 0\}).$$

It is worth emphasising that the classical corner polyhedron $\mathscr{C}_\gamma(A, \boldsymbol{b})$ associated with the basis γ was introduced by Gomory [49], where only the nonbasic entries of $\boldsymbol{x}$ need to be nonnegative. IPs of the form

$$A\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x}_{\bar{\tau}} \geq \mathbf{0}, \quad \boldsymbol{x} \in \mathbb{Z}^n,$$

where τ is the support of a vertex of $P(A, \boldsymbol{b})$ are referred to as the *Gomory relaxation with respect to* τ [96]. The Gomory relaxation is indeed a "relaxation" since only those entries of $\boldsymbol{x}$ corresponding to $\bar{\tau} = \{1, \ldots, n\} \backslash \tau$ need to be nonnegative.

Theorem 4.1.1. *Let $A \in \mathbb{Z}^{m \times n}$ with $m < n$ be a matrix of full rank m and $\boldsymbol{b} \in \mathbb{Z}^m$. Let $\boldsymbol{x}^*$ be a vertex of the polyhedron $P(A, \boldsymbol{b})$ given by a basis γ as in (4.1) and let $\mathscr{C}_\gamma(A, \boldsymbol{b}) \neq \emptyset$. Let $\boldsymbol{z}^*$ be an integral vertex of $\mathscr{C}_\gamma(A, \boldsymbol{b})$ and let $r = \|\boldsymbol{z}^*_{\bar{\gamma}}\|_0$. Then*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0, \tag{4.2}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq \frac{\Delta(A)}{\gcd(A)} - 1 \text{ if } r = 1, \text{ and} \tag{4.3}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^r}{r} \leq \frac{\Delta(A)}{\gcd(A)} \text{ if } r \geq 2. \tag{4.4}$$

The equality (4.2) is sharp by inspection. Further, the bounds (4.3) and (4.4) are optimal in certain scenarios. Being more specific, the upper bound (4.3) is attained already in the knapsack scenario (with the choice of parameters given in (4.20), initially presented in [8]). The following example shows that the bound (4.4), in its turn, is attained for $r = 2$.

Example 4.1.2. *Given the data*

$$A = \begin{pmatrix} 2 & 0 & 5 & 5 \\ 0 & 4 & 2 & -1 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 20 \\ 3 \end{pmatrix}$$

the point $\boldsymbol{x}^* = (10, 3/4, 0, 0)^T$ *is a vertex of* $P(A, \boldsymbol{b})$*. In this case, observe that* $\Delta(A) = 20$*,* $\gcd(A) = 1$ *and that* $\boldsymbol{z}^* = (0, 1, 1, 3)^T$ *is the unique integral vertex of the corner polyhedron* $\mathscr{C}_{\{1,2\}}(A, \boldsymbol{b})$*. In particular, notice that* $r = 2$ *and the inequality* $10 \cdot 2^2/2 = 20 \leq 20/1 = 20$ *holds, which demonstrates that the inequality (4.4) from Theorem 4.1.1 is sharp.*

We remark that this example was obtained by closely analysing the tight cases of inequality (4.34) in Lemma 4.2.3, which takes a special form when $r = 2$.

Theorem 4.1.1 shows that for the classical Gomory corner polyhedron $\mathscr{C}_\gamma(A, \boldsymbol{b})$ associated with the vertex $\boldsymbol{x}^*$ of $P(A, \boldsymbol{b})$ a strong *proximity-sparsity transference* holds. In particular, the distance (proximity) from any vertex $\boldsymbol{x}^*$ of $P(A, \boldsymbol{b})$ to any integral vertex $\boldsymbol{z}^*$ of the corner polyhedron *exponentially* decreases when the size of support of $\boldsymbol{z}^*$ grows, and vice versa, the size of support of $\boldsymbol{z}^*$ reduces with the growth of its distance to the vertex $\boldsymbol{x}^*$. It should be emphasised that $\boldsymbol{z}^*$ is a vertex of $\mathscr{C}_\gamma(A, \boldsymbol{b})$ and, hence, in general we can only guarantee that $n - m$ entries of $\boldsymbol{z}^*$ are nonnegative.

Furthermore, in light of Theorem 4.1.1, we can state the following corollary which provides a resembling proximity (distance) bound with respect to the ℓ_1-norm. If additionally one assumes that $\boldsymbol{z}^*$ is a feasible integral solution to the IP then the following corollary would provide a refinement over the best known distance bounds [41, 72].

Corollary 4.1.3. *Let* $A \in \mathbb{Z}^{m \times n}$ *with* $m < n$ *be a matrix of full rank* m *and* $\boldsymbol{b} \in \mathbb{Z}^m$*. Let* $\boldsymbol{x}^*$ *be a vertex of* $P(A, \boldsymbol{b})$ *given by a basis* γ *as in (4.1) and let* $\mathscr{C}_\gamma(A, \boldsymbol{b}) \neq \emptyset$*. Let* $\boldsymbol{z}^*$ *be an integral vertex of the corner polyhedron* $\mathscr{C}_\gamma(A, \boldsymbol{b})$ *and let* $r = \|\boldsymbol{z}^*_{\bar{\gamma}}\|_0$*. Then*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0, \tag{4.5}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < (m+1)\left(\frac{\Delta(A)}{\gcd(A)} - 1\right) \text{ if } r = 1, \text{ and} \tag{4.6}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < \frac{r(r+m)}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right) < \left(\frac{9}{8} + \frac{m}{2}\right)\left(\frac{\Delta(A)}{\gcd(A)}\right) \text{ if } r \geq 2\,. \tag{4.7}$$

Suppose next that the polyhedron $P(A, \boldsymbol{b})$ is integer feasible and consider its *integer hull* $P_I(A, \boldsymbol{b}) = \operatorname{conv}(P(A, \boldsymbol{b}) \cap \mathbb{Z}^n)$. A natural direction for a further research would be to derive a distance-sparsity transference bound for the vertices of $P_I(A, \boldsymbol{b})$. Notice that the set $P(A, \boldsymbol{b}) \cap \mathbb{Z}^n$ is obtained from $\mathscr{C}_\gamma(A, \boldsymbol{b}) \cap \mathbb{Z}^n$ by enforcing back the nonnegativity constraints $\boldsymbol{x}_\gamma \geq 0$ and this may potentially result in cutting off all vertices of the corner polyhedron. In Subsection 4.1.1 we show that in the knapsack scenario at least one vertex of $\mathscr{C}_\gamma(A, \boldsymbol{b})$ avoids the cut and Theorem 4.1.1 implies an optimal distance-sparsity transference bound for lattice points in the knapsack polytope. We expect that a certain transference bound holds for vertices of $P_I(A, \boldsymbol{b})$ in the general setting, however, further research will have to be carried out in order for one to obtain such a transference result.

Although it remains an open problem to extend Theorem 4.1.1 to vertices of the integral hull $P_I(A, \boldsymbol{b})$ in the general setting, the next result of this chapter provides additional information in the case when the vertex $\boldsymbol{x}^*$ is degenerate; that is, $\tau = \operatorname{supp}(\boldsymbol{x}^*)$ has size strictly less than m, i.e. when $|\tau| < m$. In particular, the result applies to a tighter relaxation of the IP, where we enforce back the nonnegativity constraints $\boldsymbol{x}_\gamma \geq 0$ that are tight at the vertex $\boldsymbol{x}^*$ at the cost of a slightly weaker bound. Theorems 4.1.1 and 4.1.4 do however coincide when $\boldsymbol{x}^*$ is nondegenerate.

Recall that when the vertex $\boldsymbol{x}^*$ is degenerate, the choice of basis γ in (4.1) is typically not unique. However, what we show is that there exists at least one basis γ for which the conclusions of Theorem 4.1.1 remain valid for this polyhedron, up to a factor which depends on the number of zero coordinates of $\boldsymbol{x}^*_\tau$.

Theorem 4.1.4. *Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ be a matrix of full rank m and let $\boldsymbol{b} \in \mathbb{Z}^m$. Let $\boldsymbol{x}^*$ be a vertex of the polyhedron $P(A, \boldsymbol{b})$ with $\tau = \operatorname{supp}(\boldsymbol{x}^*)$. Let $\mathscr{C}_\tau(A, \boldsymbol{b}) \neq \emptyset$ and let $\boldsymbol{z}^*$ be an integral vertex of $\mathscr{C}_\tau(A, \boldsymbol{b})$. Then there exists a basis γ of A with $\tau \subset \gamma$ such that, letting $r = \|\boldsymbol{z}^*_{\bar{\gamma}}\|_0$ and $d = m - |\tau|$, we have*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0, \tag{4.8}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq \frac{\Delta(A)}{\gcd(A)} - 1 \text{ if } r = 1, \text{ and} \tag{4.9}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^r}{r^{d+1}} \leq \frac{\Delta(A)}{\gcd(A)} \text{ if } r \geq 2. \tag{4.10}$$

It is worth emphasising in contrast to Theorem 4.1.1, $\boldsymbol{z}^*$ is a vertex of the corner polyhedron $\mathscr{C}_\tau(A, \boldsymbol{b})$ and, hence, in general we can guarantee that $n - |\tau|$ entries of $\boldsymbol{z}^*$ are nonnegative, where $0 \leq |\tau| \leq m$. However, for each additional nonnegativity constraint that is enforced back, the bound (4.10) slightly weakens.

Although the statement of Theorem 4.1.4 is rather similar to the statement of Theorem 4.1.1, the proofs of the two results are quite different. This is because Theorem 4.1.1 provides a transference bound for classical Gomory corner polyhedra, where the underlying cone is simplicial, allowing the result to be proven using standard integer optimisation tools. Theorem 4.1.4 in contrast applies to generalised corner polyhedra, introduced by Thomas [96], where the affine cone $A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x}_{\bar{\tau}} \geq \boldsymbol{0}$ can have considerably more complicated geometry when τ has cardinality strictly smaller than m. In particular, the cone need not be a simplicial cone, whose orthogonal projection onto the $\bar{\gamma}$ coordinates of $\mathbb{R}^n$ is simply the nonnegative orthant. It is for this reason that the proof of Theorem 4.1.4 is carried out using more sophisticated convex geometric arguments in Section 4.5, in addition to the lattice-based arguments appearing in the proof of Theorem 4.1.1 in Section 4.3.

We can similarly state a corollary which provides a resembling proximity (distance) bound with respect to the ℓ_1-norm.

Corollary 4.1.5. *Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ be a matrix of full rank m and $\boldsymbol{b} \in \mathbb{Z}^m$. Let $\boldsymbol{x}^*$ be a vertex of $P(A, \boldsymbol{b})$ with $\tau = \operatorname{supp}(\boldsymbol{x}^*)$. Let $\mathscr{C}_\tau(A, \boldsymbol{b}) \neq \emptyset$ and let $\boldsymbol{z}^*$ be an integral vertex of $\mathscr{C}_\tau(A, \boldsymbol{b})$. Then there exists a basis γ of A with $\tau \subset \gamma$ such that, letting $r = \|\boldsymbol{z}^*_{\bar\gamma}\|_0$ and $d = m - |\tau|$, we have*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0\,, \tag{4.11}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < (|\tau| + 1)\left(\frac{\Delta(A)}{\gcd(A)} - 1\right) \text{ if } r = 1\,, \text{ and} \tag{4.12}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < \frac{r^{d+1}(r + |\tau|)}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right) \text{ if } r \geq 2\,. \tag{4.13}$$

4.1.1 Distance-sparsity transference for knapsacks

We now consider the case $A \in \mathbb{Z}^{1\times n}_{>0}$, known as *knapsack scenario*. Following traditional vector notation, we replace the matrix A and the right-hand side $\boldsymbol{b}$ with a positive integer vector $\boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{Z}^n_{>0}$ and integer $b \in \mathbb{Z}$. In this setting, the polyhedron $P(A, \boldsymbol{b})$ is referred to as the *knapsack polytope*, which is precisely

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : \boldsymbol{a}^T\boldsymbol{x} = b\right\}.$$

Note that if $P(\boldsymbol{a}, b) \neq \emptyset$, then clearly $b \geq 0$. Further, if $b > 0$, then the knapsack polytope $P(\boldsymbol{a}, b)$ is an $(n-1)$-dimensional simplex in $\mathbb{R}^n$ with vertices

$$\left(\frac{b}{a_1}, 0, \ldots, 0\right)^T, \left(0, \frac{b}{a_2}, 0 \ldots, 0\right)^T, \ldots, \left(0, \ldots, 0, \frac{b}{a_n}\right)^T.$$

The vertices can be equivalently expressed as $(b/a_1)\boldsymbol{e}_1, \ldots, (b/a_n)\boldsymbol{e}_n$, where $\boldsymbol{e}_i$ denotes the i-th standard/elementary basis vector. Note that each vertex of the knapsack polytope $P(\boldsymbol{a}, b)$ has at most one nonzero entry.

Given a vertex $\boldsymbol{x}^*$ of $P(\boldsymbol{a}, b)$ with $\gamma = \operatorname{supp}(\boldsymbol{x}^*)$, the corner polyhedron associated with $\boldsymbol{x}^*$ can be written as

$$\mathscr{C}_\gamma(\boldsymbol{a}, b) = \operatorname{conv}\left(\{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{a}^T\boldsymbol{x} = b, \boldsymbol{x}_{\bar\gamma} \geq 0\}\right).$$

In what follows, we will exclude the trivial case $n = 1$ and assume that the dimension $n \geq 2$. Furthermore, we assume without loss of generality that $\boldsymbol{a}$ is a primitive integer vector with nonzero entries. In other words, we will assume the following conditions:

$$\begin{aligned} &(i)\ \boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{Z}^n_{>0},\ n \geq 2 \text{ and} \\ &(ii)\ \gcd(\boldsymbol{a}) := \gcd(a_1, \ldots, a_n) = 1. \end{aligned} \tag{4.14}$$

In addition, we assume that $P(\boldsymbol{a}, b)$ contains integer points, i.e. that $P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset$. This second assumption is equivalent to assumption that the integer b belongs to the *semigroup*

$$Sg(\boldsymbol{a}) = \left\{ \boldsymbol{a}^\top \boldsymbol{z} : \boldsymbol{z} \in \mathbb{Z}^n_{\geq 0} \right\}$$

generated by the entries of the vector $\boldsymbol{a}$. Notice that the semigroup $Sg(\boldsymbol{a})$ contains all nonnegative integral combinations of $a_1, \ldots, a_n$. It should be emphasised that any element of the semigroup $Sg(\boldsymbol{a})$ must be divisible by the $\gcd(\boldsymbol{a})$. In light of this, if $\gcd(\boldsymbol{a}) \neq 1$ and $P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset$, then b must also be divisible by $\gcd(\boldsymbol{a})$ and, in such case, we could simply replace $\boldsymbol{a}$, b by $\boldsymbol{a}/\gcd(\boldsymbol{a})$, $b/\gcd(\boldsymbol{a})$. In particular, the second assumed condition from (4.14) follows directly.

Recall that Aliev et al. [8] proved that for any vertex $\boldsymbol{x}^*$ of the knapsack polytope $P(\boldsymbol{a}, b)$, there exists an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b)$ such that

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1 \tag{4.15}$$

and, further, that the bound (4.15) is sharp in the following sense. For any positive integer k and any dimension n, there exists a positive integral vector $\boldsymbol{a}$ satisfying (4.14) with $\|\boldsymbol{a}\|_\infty = k$ and $b \in \mathbb{Z}$ such that the knapsack polytope $P(\boldsymbol{a}, b)$ contains exactly one integer point $\boldsymbol{z}$, where the equality in (4.15) is attained, namely such that the equality $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \|\boldsymbol{a}\|_\infty - 1$ holds.

Recall also that the best known sparsity-type estimate in the knapsack scenario, obtained by Aliev et al. [4], guarantees the existence of an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b)$ which satisfies the sparsity bound

$$\|\boldsymbol{z}\|_0 \leq 1 + \log(\min\{a_1, \ldots, a_n\}), \tag{4.16}$$

where recall that $\log(\cdot)$ denotes the logarithm with base two. The next result will combine and refine the bounds (4.15) and (4.16) as follows.

Theorem 4.1.6. *Let $\boldsymbol{a}$ satisfy (4.14), $b \in Sg(\boldsymbol{a})$ and let $\boldsymbol{x}^*$ be a vertex of the knapsack polytope $P(\boldsymbol{a}, b)$ with basis $\gamma = \mathrm{supp}(\boldsymbol{x}^*)$. Then $P(\boldsymbol{a}, b)$ contains an integral vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ associated with $\boldsymbol{x}^*$ such that, letting $r = \|\boldsymbol{z}^*_{\bar{\gamma}}\|_0$,*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0, \tag{4.17}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1 \text{ if } r = 1, \text{ and} \tag{4.18}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^r}{r} < \|\boldsymbol{a}\|_\infty \text{ if } r \geq 2. \tag{4.19}$$

It is worth emphasising in contrast to Theorems 4.1.1 and 4.1.4, Theorem 4.1.6 is rather special because it is the only case where we can guarantee that at least one integral vertex $\boldsymbol{z}^*$ of $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ is feasible in the sense that it belongs to the knapsack polytope $P(\boldsymbol{a}, b)$. Further, we believe the result is of particular interest for researchers working on knapsack and subset sum problems (see e.g. [65, Chapter 4]). Theorem 4.1.6 can therefore be viewed as a transference result that allows strengthening the distance bound (4.15) if integer points in the knapsack polytope are not sparse and, vice versa, strengthening the sparsity bound (4.16) if feasible integer points are sufficiently far from a vertex of the knapsack polytope. Example 5.2.3 and Figure 5.2 in Chapter 5 demonstrate that feasibility cannot be ensured in the scenario when all entries of the vector $\boldsymbol{a}$ are not strictly positive. Furthermore, Figure 4.1 below demonstrates that integral vertices of corner polyhedra may be infeasible when the polyhedron $P(A, \boldsymbol{b})$ is defined by more than one linear constraint.

It follows from the proof of Theorem 1 (ii) in [8] that the bound (4.18) (and hence (4.3) for $m = 1$) corresponding to the case $r = 1$ is optimal. For completeness, we recall that Aliev et al. [8] show that it is sufficient to choose a positive integer k and set

$$\boldsymbol{a} = (k, \ldots, k, 1)^T, \; b = k - 1 \text{ and } \boldsymbol{x}^* = \frac{k-1}{k}\boldsymbol{e}_1 . \tag{4.20}$$

In this scenario, the knapsack polytope $P(\boldsymbol{a}, b)$ contains precisely one integer point, namely the integer point $\boldsymbol{z}^* = (k-1)\boldsymbol{e}_n$ and consequently we notice that $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty = k - 1 = \|\boldsymbol{a}\|_\infty - 1$ as required.

In light of Theorem 4.1.6, we can state the following corollary which improves the best known proximity (distance) bounds [8, 41, 72] with respect to the ℓ_1-norm for the knapsack scenario without further assumptions since at least one vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ is feasible in the sense that $\boldsymbol{z}^* \in P(\boldsymbol{a}, b)$.

Corollary 4.1.7. *Let $\boldsymbol{a}$ satisfy (4.14), $b \in Sg(\boldsymbol{a})$ and let $\boldsymbol{x}^*$ be a vertex of the knapsack polytope $P(\boldsymbol{a}, b)$ with basis $\gamma = \operatorname{supp}(\boldsymbol{x}^*)$. Then $P(\boldsymbol{a}, b)$ contains an integral vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ associated with $\boldsymbol{x}^*$ such that, letting $r = \|\boldsymbol{z}^*_{\bar\gamma}\|_0$,*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \text{ if } r = 0, \tag{4.21}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < 2(\|\boldsymbol{a}\|_\infty - 1) \text{ if } r = 1, \text{ and} \tag{4.22}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 < \frac{r(r+1)}{2^r}\|\boldsymbol{a}\|_\infty \le \frac{3}{2}\|\boldsymbol{a}\|_\infty \text{ if } r \ge 2 . \tag{4.23}$$

Furthermore, we can state a corollary to Theorem 4.1.6 which provides an upper bound for the *(additive) integrality gap*, which is a quantity introduced below. Given a

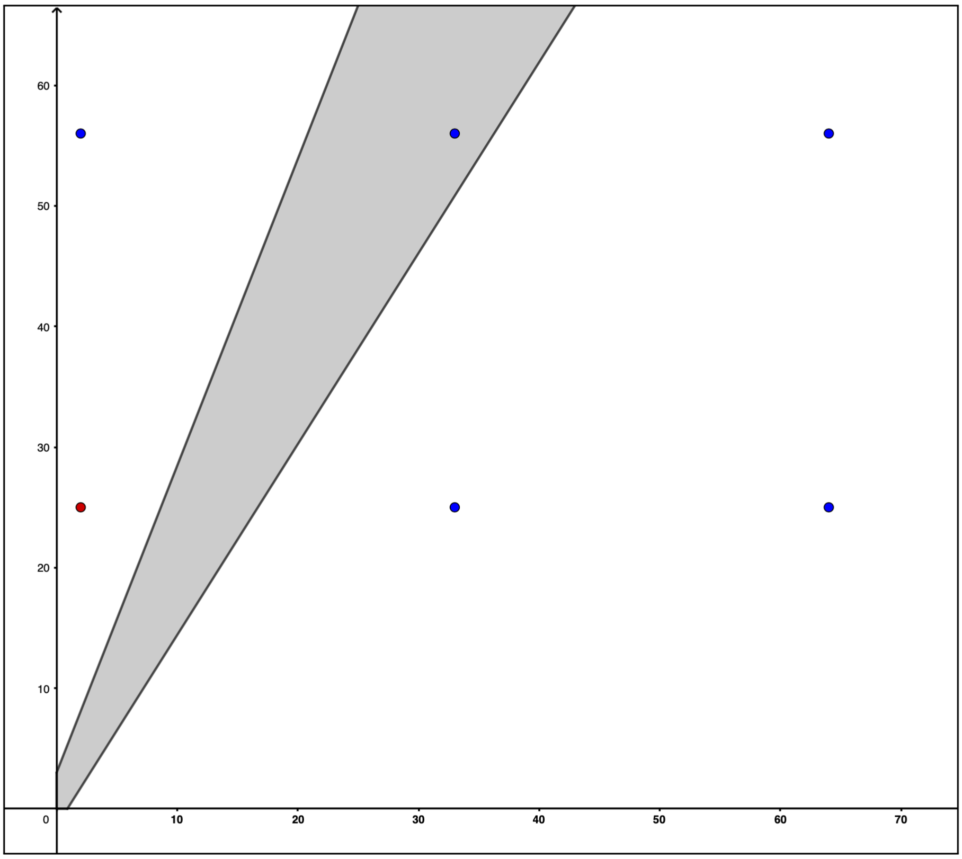


Figure 4.1: This figure illustrates that unique integral vertex (red) of the corner polyhedron corresponds to an infeasible point since the point lies outside of the projected feasible region (grey). It should be emphasised that the polyhedron $P(A, \boldsymbol{b})$ is defined here with two distinct linear constraints.

cost vector $\boldsymbol{c} \in \mathbb{Z}^n$, we will now consider the *integer knapsack problem*

$$\min\left\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n\right\}, \tag{4.24}$$

where (4.24) is feasible since $b \in Sg(\boldsymbol{a})$ by assumption.

Let $IP(\boldsymbol{c}, \boldsymbol{a}, b)$ and $LP(\boldsymbol{c}, \boldsymbol{a}, b)$ denote the optimal values of (4.24) and its linear programming relaxation

$$\min\left\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in P(\boldsymbol{a}, b)\right\}, \tag{4.25}$$

respectively. The *(additive) integrality gap* $IG(\boldsymbol{c}, \boldsymbol{a}, b)$ of the integer knapsack problem (4.24) is defined as

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) = IP(\boldsymbol{c}, \boldsymbol{a}, b) - LP(\boldsymbol{c}, \boldsymbol{a}, b).$$

In the knapsack scenario, the integrality gap $IG(\boldsymbol{c}, \boldsymbol{a}, b)$ can be bounded from above by

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) \le d(\boldsymbol{a}, b)\, \|\boldsymbol{c}\|_1,$$

where $d(\boldsymbol{a}, b)$ denotes the (maximum) vertex distance, which is defined as

$$d(\boldsymbol{a}, b) = \begin{cases} \max_{\boldsymbol{x}^*} \min_{\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n} \|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty, & \text{if } P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset, \\ & -\infty, \text{ otherwise.} \end{cases}$$

The following provides an upper bound for the integrality gap $IG(\boldsymbol{c}, \boldsymbol{a}, b)$.

Corollary 4.1.8. *Let $\boldsymbol{a}$ satisfy (4.14), $b \in Sg(\boldsymbol{a})$ and $\boldsymbol{c} \in \mathbb{Z}^n$. Let $\boldsymbol{x}^*$ be an optimal vertex solution to (4.25) with basis γ. Let further $\boldsymbol{z}^*$ be any integal vertex of $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ such that $\boldsymbol{z}^* \in P(\boldsymbol{a}, b)$. Then, letting $r = \|\boldsymbol{z}^*_{\bar{\gamma}}\|_0$, we have*

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) = 0 \textit{ if } r = 0, \tag{4.26}$$

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) \leq 2\,(\|\boldsymbol{a}\|_\infty - 1)\|\boldsymbol{c}\|_\infty \textit{ if } r = 1\,, \textit{ and} \tag{4.27}$$

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) < \frac{r(r+1)}{2^r}\|\boldsymbol{a}\|_\infty \|\boldsymbol{c}\|_\infty \textit{ if } r \geq 2\,. \tag{4.28}$$

The next result presented in this chapter shows that the bounds that appear in Theorems 4.1.1, 4.1.4 and 4.1.6 are optimal in the knapsack scenario for $r \geq 2$.

Theorem 4.1.9. *Fix an integer $n \geq 3$. For any $\epsilon > 0$ there exists an integer vector $\boldsymbol{a} \in \mathbb{Z}^n$ satisfying (4.14) and $b \in Sg(\boldsymbol{a})$ such that for a vertex $\boldsymbol{x}^*$ of the knapsack polytope $P(\boldsymbol{a}, b)$ with $\gamma = \operatorname{supp}(\boldsymbol{x}^*)$ and an integral vertex $\boldsymbol{z}^*$ of $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ with $\|\boldsymbol{z}^*_{\bar{\gamma}}\|_0 = n - 1$*

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^{n-1}}{n-1} > (1-\epsilon)\|\boldsymbol{a}\|_\infty\,. \tag{4.29}$$

This result intuitively tells us that one can create examples where the left-hand side of the strict inequality (4.19) becomes arbitrarily close to the right-hand side, however, critically an equality is not attained in the knapsack scenario as is found in the more general setting of both Theorems 4.1.1 and 4.1.4.

4.1.2 A refined sparsity-type bound for solutions to integer programs

The next result presented in this chapter aims to refine the general sparsity-type bound obtained by Aliev et al. [7]. Let

$$\rho(\boldsymbol{x}) = \min\{|x_i| : i \in \operatorname{supp}(\boldsymbol{x})\}$$

denote the minimum absolute nonzero entry of a given vector $\boldsymbol{x}$. Let $\boldsymbol{c} \in \mathbb{Z}^n$. We will consider the general integer linear programming problem that is given in standard form, namely the IP

$$\max\left\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{x} \in P(A, \boldsymbol{b}) \cap \mathbb{Z}^n\right\}. \tag{4.30}$$

We assume that $P(A, \boldsymbol{b})$ contains integer points, i.e. $P(A, \boldsymbol{b}) \cap \mathbb{Z}^n \neq \emptyset$, such that the general IP problem (4.30) is feasible.

It was shown in [7] that there exists an optimal integral solution $\boldsymbol{z}^*$ to the IP for (4.30) which satisfies the sparsity bound

$$\|\boldsymbol{z}^*\|_0 \leq m + \log\left(\frac{\sqrt{\det(AA^T)}}{\gcd(A)}\right). \tag{4.31}$$

Recall that any vertex solution to the linear programming relaxation of (4.30) has the size of support $\leq m$. Any non-vertex solution $\boldsymbol{z}^*$, in its turn, belongs to the interior of the face

$$\mathscr{F} = P(A, \boldsymbol{b}) \cap \{\boldsymbol{x} \in \mathbb{R}^n : x_i = 0 \text{ for } i \notin \operatorname{supp}(\boldsymbol{z}^*)\}$$

of the polyhedron $P(A, \boldsymbol{b})$. In light of this, the minimum absolute nonzero entry $\rho(\boldsymbol{z}^*)$ is precisely the ℓ_∞-distance from $\boldsymbol{z}^*$ to the boundary of $\mathscr{F}$. To obtain a refinement of the sparsity bound (4.31), we will link the minimum absolute nonzero entry and the size of support of solutions to the IP (4.30).

Theorem 4.1.10. *Let $A \in \mathbb{Z}^{m \times n}$ with $m < n$ be a matrix of full rank m, $\boldsymbol{b} \in \mathbb{Z}^m$, $\boldsymbol{c} \in \mathbb{Z}^n$ and suppose that the IP (4.30) is feasible. Then there is an optimal integral solution $\boldsymbol{z}^*$ to the IP (4.30) such that, letting $s = \|\boldsymbol{z}^*\|_0$,*

$$(\rho(\boldsymbol{z}^*) + 1)^{s-m} \leq \frac{\sqrt{\det(AA^T)}}{\gcd(A)}. \tag{4.32}$$

4.2 Preliminary and Auxiliary Results Regarding Lattices and Corner Polyhedra

Recall from Chapter 2 that for linearly independent $\boldsymbol{b}_1, \dots, \boldsymbol{b}_l \in \mathbb{R}^d$, the set

$$\Lambda = \left\{\sum_{i=1}^{l} x_i \boldsymbol{b}_i : x_i \in \mathbb{Z}\right\}$$

is an l-dimensional lattice with basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_l$ and determinant

$$\det(\Lambda) = \left(\det(\boldsymbol{b}_i \cdot \boldsymbol{b}_j)_{1 \leq i,j \leq l}\right)^{1/2},$$

where $\boldsymbol{b}_i \cdot \boldsymbol{b}_j$ is the standard inner product of the basis vectors $\boldsymbol{b}_i$ and $\boldsymbol{b}_j$. The *Minkowski sum* $X + Y$ of the sets $X, Y \subset \mathbb{R}^d$ consists of all points $\boldsymbol{x} + \boldsymbol{y}$ with $\boldsymbol{x} \in X$ and $\boldsymbol{y} \in Y$. The *difference set* $X - X$ is the Minkowski sum of X and $-X$. For a lattice $\Lambda \subset \mathbb{R}^d$ and $\boldsymbol{y} \in \mathbb{R}^d$, the set $\boldsymbol{y} + \Lambda$ is an *affine lattice* with determinant $\det(\Lambda)$.

Let $\Lambda \subset \mathbb{Z}^d$ be a d-dimensional integer lattice. The point $\boldsymbol{x} \in \mathbb{Z}^d_{\geq 0}$ is called *irreducible (with respect to the lattice* Λ*)* if for any two points $\boldsymbol{y}, \boldsymbol{y}' \in \mathbb{Z}^d_{\geq 0}$ with $0 \leq y_i \leq x_i$, $0 \leq y'_i \leq x_i$ where $i \in \{1, \ldots, d\}$, the inclusion $\boldsymbol{y} - \boldsymbol{y}' \in \Lambda$ implies $\boldsymbol{y} = \boldsymbol{y}'$.

Lemma 4.2.1 (Theorem 1 in [49])**.** *If* $\boldsymbol{x} \in \mathbb{Z}^d_{\geq 0}$ *is irreducible with respect to the lattice* Λ *then*

$$\prod_{i=1}^{d} (x_i + 1) \leq \det(\Lambda)\,. \tag{4.33}$$

Proof. The lattice Λ can be viewed as a subgroup of the additive group $\mathbb{Z}^d$. The number of points $\boldsymbol{y} \in \mathbb{Z}^d_{\geq 0}$ with $0 \leq y_i \leq x_i$, $i \in \{1, \ldots, d\}$ is equal to $\prod_{i=1}^{d}(x_i + 1)$. Since the point $\boldsymbol{x}$ is irreducible by assumption, each such $\boldsymbol{y}$ corresponds to a unique coset (affine lattice) $\boldsymbol{y} + \Lambda$ of Λ. Finally, notice that there are only $\det(\Lambda)$ different cosets yields that (4.33) holds as required. $\square$

Recall that $\operatorname{conv}(S)$ denotes the *convex hull* of $S \subset \mathbb{R}^d$. Let $\boldsymbol{p} \in \mathbb{Z}^d$ and consider the affine lattice $\Gamma = \boldsymbol{p} + \Lambda$. We will call the set $\mathscr{E}(\Gamma) = \operatorname{conv}(\Gamma \cap \mathbb{R}^d_{\geq 0})$ the *sail* associated with Γ, as illustrated in Figure 4.2.

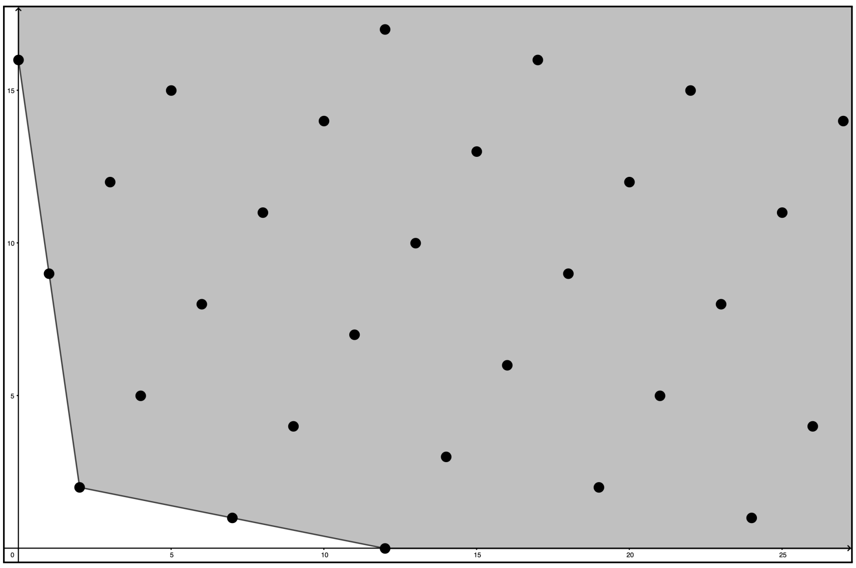


Figure 4.2: This figure illustrates, for an affine integer lattice Γ (black dots), the sail $\mathscr{E}(\Gamma) = \operatorname{conv}(\Gamma \cap \mathbb{R}^2_{\geq 0})$ associated with that affine lattice Γ (shaded grey).

Lemma 4.2.2. *Every vertex of the sail* $\mathscr{E}(\Gamma)$ *is irreducible.*

Proof. Let $\boldsymbol{x}$ be a vertex of the sail $\mathscr{E}(\Gamma)$. Suppose, to derive a contradiction, that $\boldsymbol{x}$ is reducible. Then there are distinct points $\boldsymbol{y}, \boldsymbol{y}' \in \mathbb{Z}^d_{\geq 0}$ with $0 \leq y_i \leq x_i$, $0 \leq y'_i \leq x_i$ where $i \in \{1, \ldots, d\}$ such that $\boldsymbol{y} - \boldsymbol{y}' \in \Lambda$.

Since $\boldsymbol{x}-\boldsymbol{y} \in \mathbb{Z}^d_{\geq 0}$ and $\boldsymbol{x}-\boldsymbol{y}' \in \mathbb{Z}^d_{\geq 0}$, the vectors $\boldsymbol{v}_1 = \boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}'$ and $\boldsymbol{v}_2 = \boldsymbol{x}-\boldsymbol{y}'+\boldsymbol{y}$ have nonnegative integer entries. Furthermore, $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Gamma$ and $\boldsymbol{x} = (\boldsymbol{v}_1+\boldsymbol{v}_2)/2$. Therefore, we conclude that $\boldsymbol{x}$ is not a vertex of the sail $\mathscr{E}(\Gamma)$, which yields the contradiction as required. □

Lemma 4.2.3. *For $r \geq 2$ and $x_1, \ldots, x_r \geq 1$ the inequality*

$$x_1 + \cdots + x_r \leq \frac{r(x_1+1)\cdots(x_r+1)}{2^r} \tag{4.34}$$

holds.

Proof. Suppose that (4.34) is satisfied for $x_1 = y_1, \ldots, x_r = y_r$. We will first show that for any $\epsilon > 0$ and any $i \in \{1, \ldots, r\}$ the inequality (4.34) is satisfied for $x_1 = y_1, \ldots, x_{i-1} = y_{i-1}, x_i = y_i + \epsilon, x_{i+1} = y_{i+1}, \ldots, x_r = y_r$. After possible renumbering, it is sufficient to consider the case $i = 1$. We have

$$\begin{aligned}(y_1+\epsilon)+y_2+\cdots+y_r &\leq \frac{r(y_1+1)(y_2+1)\cdots(y_r+1)}{2^r}+\epsilon \\ &\leq \frac{r(y_1+1)(y_2+1)\cdots(y_r+1)}{2^r}+\epsilon\frac{r(y_2+1)\cdots(y_r+1)}{2^r},\end{aligned}$$

which follows in light of the lower bound

$$\frac{r(y_2+1)\cdots(y_r+1)}{2^r} \geq \frac{r\,\overbrace{(1+1)\cdots(1+1)}^{(r-1)\text{-terms}}}{2^r} = \frac{r\cdot 2^{r-1}}{2^r} = \frac{r}{2} \geq 1,$$

where this holds by noting that $y_2, \ldots, y_r \geq 1$ and $r \geq 2$. Furthermore, we notice that

$$\begin{aligned}\frac{r(y_1+1)(y_2+1)\cdots(y_r+1)}{2^r}+\epsilon\frac{r(y_2+1)\cdots(y_r+1)}{2^r}& \\ = \frac{r(y_1+1+\epsilon)(y_2+1)\cdots(y_r+1)}{2^r}&.\end{aligned}$$

To complete the proof, it is sufficient to observe that the inequality (4.34) holds for the case $y_1 = \cdots = y_r = 1$. □

Given $A \in \mathbb{Z}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{Z}^m$, we will denote by $\Gamma(A, \boldsymbol{b})$ the set of integer points in the affine subspace

$$H(A, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}\},$$

that is

$$\Gamma(A, \boldsymbol{b}) = H(A, \boldsymbol{b}) \cap \mathbb{Z}^n.$$

The set $\Gamma(A, \boldsymbol{b})$ is either empty or an affine lattice of the form $\Gamma(A, \boldsymbol{b}) = \boldsymbol{p} + \Gamma(A)$, where $\boldsymbol{p}$ is any integer vector such that $A\boldsymbol{p} = \boldsymbol{b}$ and $\Gamma(A) = \Gamma(A, \boldsymbol{0})$ denotes the lattice formed by all integer points in the kernel of the matrix A.

Fix a basis γ of A and let π_γ denote the projection map from $\mathbb{R}^n$ to $\mathbb{R}^{n-m}$ which forgets about the coordinates indexed by γ, that is $\pi_\gamma : \boldsymbol{u} \mapsto \boldsymbol{u}_{\bar{\gamma}}$. Recall that the submatrix A_γ is nonsingular. In light of this, it follows that the restricted map $\pi_\gamma|_{H(A,\boldsymbol{b})} : H(A,\boldsymbol{b}) \to \mathbb{R}^{n-m}$ is bijective. Specifically, any projected point $\boldsymbol{u}_{\bar{\gamma}} \in \mathbb{R}^{n-m}$ is uniquely mapped by $\pi_\gamma^{-1}|_{H(A,\boldsymbol{b})}$ to a point $\boldsymbol{u} \in H(A,\boldsymbol{b})$ with

$$\boldsymbol{u}_\gamma = A_\gamma^{-1}\left(\boldsymbol{b} - A_{\bar{\gamma}}\boldsymbol{u}_{\bar{\gamma}}\right). \tag{4.35}$$

For technical reasons, it is convenient for us to consider the projected affine lattice $\Lambda_\gamma(A,\boldsymbol{b}) = \pi_\gamma(\Gamma(A,\boldsymbol{b}))$ and the projected lattice $\Lambda_\gamma(A) = \pi_\gamma(\Gamma(A))$.

Lemma 4.2.4. *Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ be a matrix with a basis γ. Then*

$$\det\left(\Lambda_\gamma(A)\right) = \frac{|\det(A_\gamma)|}{\gcd(A)}. \tag{4.36}$$

Proof. Upon reordering the columns of A if necessary, we may assume without loss of generality that $\gamma = \{1,2,\ldots,m\}$. Let $\boldsymbol{g}_1,\ldots,\boldsymbol{g}_{n-m}$ be a basis of $\Gamma(A)$. Since the map $\pi_\gamma|_{H(A,\boldsymbol{0})}$ is bijective, the vectors $\boldsymbol{b}_1 = \pi_\gamma(\boldsymbol{g}_1),\ldots,\boldsymbol{b}_{n-m} = \pi_\gamma(\boldsymbol{g}_{n-m})$ form a basis of the projected lattice $\Lambda_\gamma(A)$. Let $G \in \mathbb{Z}^{n\times(n-m)}$ be the matrix with columns $\boldsymbol{g}_1,\ldots,\boldsymbol{g}_{n-m}$. We will denote by F the $(n-m)\times(n-m)$-submatrix of G which consists of the last $n-m$ rows of G. Hence, the columns of F are $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_{n-m}$. Then $\det(\Lambda_\gamma(A)) = |\det(F)|$. The rows of the matrix A span the m-dimensional rational subspace of $\mathbb{R}^n$ orthogonal to the $(n-m)$-dimensional rational subspace spanned by the columns of G. Therefore, by Lemma 5G and Corollary 5I in [88], we have $|\det(F)| = |\det(A_\gamma)|/\gcd(A)$ and, consequently, (4.36) holds as required, which concludes the proof of Lemma 4.2.4. □

Theorem 4.2.5. *Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ be a matrix with a basis γ and $\boldsymbol{b} \in \mathbb{Z}^m$. For any vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(A,\boldsymbol{b})$, the bound*

$$\prod_{j\in\bar{\gamma}}\left(z_j^* + 1\right) \le \frac{|\det(A_\gamma)|}{\gcd(A)}$$

holds.

Proof. Because the restricted map $\pi_\gamma|_{H(A,\boldsymbol{b})}$ is bijective, the point $\boldsymbol{y}^* = \pi_\gamma(\boldsymbol{z}^*)$ is a vertex of the sail $\mathscr{E}(\Lambda_\gamma(A,\boldsymbol{b}))$. The result now follows directly in light of Lemma 4.2.2, Lemma 4.2.1 and the equality (4.36). □

4.3 Proof of Theorem 4.1.1

Theorem 4.1.1 is an immediate consequence of Theorem 4.2.5 and the following lemma:

Lemma 4.3.1. *Let $A \in \mathbb{Z}^{m\times n}$ with $m < n$ be a matrix of rank m and $\boldsymbol{b} \in \mathbb{Z}^m$. Let $\boldsymbol{x}^*$ be a vertex of $P(A, \boldsymbol{b})$, and let γ be any basis of A containing* $\operatorname{supp}(\boldsymbol{x}^*)$*. Let $\boldsymbol{z}^*$ be an integral vector satisfying $A\boldsymbol{z}^* = \boldsymbol{b}$ with $\boldsymbol{z}^*_{\bar\gamma} \geq \boldsymbol{0}$, and let $r = \|\boldsymbol{z}^*_{\bar\gamma}\|_0$. Then*

$$\boldsymbol{x}^* = \boldsymbol{z}^* \textit{ if } r = 0, \tag{4.37}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \leq \frac{\Delta(A)}{|\det(A_\gamma)|} \prod_{j\in\bar\gamma} \left(z^*_j + 1\right) - 1 \textit{ if } r = 1, \textit{ and} \tag{4.38}$$

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^r}{r} \leq \frac{\Delta(A)}{|\det(A_\gamma)|} \prod_{j\in\bar\gamma} \left(z^*_j + 1\right) \textit{ if } r \geq 2. \tag{4.39}$$

Proof. If $r = \|\boldsymbol{z}^*_{\bar\gamma}\|_0 = 0$ the vector $\boldsymbol{z}^*_\gamma$ is the unique solution to the system $A_\gamma \boldsymbol{x}_\gamma = \boldsymbol{b}$ and, in consequence, the equality (4.37) holds.

In the rest of the proof we assume that $r \geq 1$. Reordering the columns of A if necessary, without loss of generality we may additionally assume that $\gamma = \{1, \ldots, m\}$. We will set $\delta = \|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty$ and consider two cases. Firstly, suppose that there exists an index $j \in \bar\gamma$ such that $\delta = |x^*_j - z^*_j| = z^*_j$. Observe that r of the numbers $z^*_{m+1}, \ldots, z^*_n$ are nonzero. Hence, the product of the nonbasic entries of $\boldsymbol{z}^*$ can be bounded by

$$(\delta + 1)\, 2^{r-1} \leq \prod_{j\in\bar\gamma} \left(z^*_j + 1\right), \tag{4.40}$$

and upon rearranging and multiplying through by $2/r$, we can write (4.40) equivalently as

$$\delta \frac{2^r}{r} \leq \frac{2}{r} \prod_{j\in\bar\gamma} \left(z^*_j + 1\right) - \frac{2^r}{r}. \tag{4.41}$$

Since $\Delta(A) \geq |\det(A_\gamma)|$, the inequality (4.41) justifies both (4.38) and (4.39).

Now suppose that $\delta = x^*_j - z^*_j$ for $j \in \gamma$. We can write

$$A_\gamma \boldsymbol{z}^*_\gamma + A_{\bar\gamma} \boldsymbol{z}^*_{\bar\gamma} = \boldsymbol{b} \text{ and } A_\gamma \boldsymbol{x}^*_\gamma = \boldsymbol{b}.$$

Therefore, combining the two equations and manipulating yields

$$A_\gamma(\boldsymbol{x}^*_\gamma - \boldsymbol{z}^*_\gamma) = A_{\bar\gamma} \boldsymbol{z}^*_{\bar\gamma}. \tag{4.42}$$

Given a vector $\boldsymbol{v} \in \mathbb{R}^m$, we will denote by $A^j_\gamma(\boldsymbol{v})$ the matrix obtained from A_γ by replacing its j-th column with $\boldsymbol{v}$. Let $A_1, \ldots, A_n$ be the columns of the matrix A. Solving (4.42) using Cramer's rule (Lemma 2.7.1) yields that

$$\begin{aligned} \delta &= x^*_j - z^*_j = \frac{\det\left(A^j_\gamma(A_{\bar\gamma} \boldsymbol{z}^*_{\bar\gamma})\right)}{\det\left(A_\gamma\right)} \\ &= \frac{1}{\det(A_\gamma)} \left(z^*_{m+1} \det\left(A^j_\gamma(A_{m+1})\right) + \cdots + z^*_n \det\left(A^j_\gamma(A_n)\right)\right). \end{aligned} \tag{4.43}$$

For the case $r = 1$, we consider the following two cases. Firstly, suppose that $|\det(A^j_\gamma(A_i))| < |\det(A_\gamma)| \le \Delta(A)$, then equation (4.43) yields

$$\delta = \frac{z_i^* \det\left(A^j_\gamma(A_i)\right)}{\det(A_\gamma)} < z_i^* = \prod_{j\in\bar{\gamma}}\left(z_j^*+1\right)-1 \le \frac{\Delta(A)}{|\det(A_\gamma)|}\prod_{j\in\bar{\gamma}}\left(z_j^*+1\right)-1 \tag{4.44}$$

as required.

Now suppose that $|\det(A_\gamma)| \le |\det(A^j_\gamma(A_i))| \le \Delta(A)$. Then, in a similar fashion, for some $i \in \bar{\gamma}$ we can write (4.43) as

$$\begin{aligned}\delta = \frac{z_i^* \det\left(A^j_\gamma(A_i)\right)}{\det(A_\gamma)} &= \left(z_i^*+1\right)\frac{\det\left(A^j_\gamma(A_i)\right)}{\det(A_\gamma)} - \frac{\det\left(A^j_\gamma(A_i)\right)}{\det(A_\gamma)} \\ &\le \frac{\Delta(A)}{|\det(A_\gamma)|}\prod_{j\in\bar{\gamma}}\left(z_j^*+1\right)-1,\end{aligned} \tag{4.45}$$

where (4.44) and (4.45) imply that the inequality (4.38) holds as required.

To settle the case $r \ge 2$, observe that (4.43) implies

$$\delta \le \left(z^*_{m+1}+\cdots+z^*_n\right)\frac{\Delta(A)}{|\det(A_\gamma)|}. \tag{4.46}$$

Without loss of generality, assume that $z_i^* \ne 0$ for $i \in \{m+1,\dots,m+r\}$ and $z_i^* = 0$ for $m+r < i \le n$. Then, by (4.46) and Lemma 4.2.3, we have

$$\delta \le \frac{r\left(z^*_{m+1}+1\right)\cdots\left(z^*_{m+r}+1\right)\Delta(A)}{2^r|\det(A_\gamma)|},$$

which establishes that the inequality (4.39) holds as required, which concludes the proof of Lemma 4.3.1. □

4.4 Proof of Corollary 4.1.3

Recall that $\boldsymbol{z}^*$ is an integral vertex of $\mathscr{C}_\gamma(A,\boldsymbol{b})$ associated with the vertex $\boldsymbol{x}^*$ of the polyhedron $P(A,\boldsymbol{b})$. In light of Theorem 4.1.1, we can bound $\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1$ as follows. If $r=0$, we have $\boldsymbol{x}^*=\boldsymbol{z}^*$, which justifies (4.5).

If $r = 1$, then after reordering if necessary, we can assume without loss of generality that the vertex $\boldsymbol{x}^*$ and the vertex of the corner polyhedron $\boldsymbol{z}^*$ have the form $\boldsymbol{x}^* = (x_1^*,\dots,x_m^*,0,\dots,0)^T$ and $\boldsymbol{z}^* = (z_1^*,\dots,z_m^*,z_{m+1}^*,0,\dots,0)^T$, respectively. In consequence, $\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1$ can be bounded by

$$\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1 = \left|x_1^*-z_1^*\right|+\cdots+\left|x_m^*-z_m^*\right|+\left|z_{m+1}^*\right| < (m+1)\left(\frac{\Delta(A)}{\gcd(A)}-1\right),$$

where the final inequality follows directly from (4.3) and by noting that there must exist at least one $i \in \{1, \ldots, m\}$ such that $\left|x_i^* - z_i^*\right| \notin \mathbb{Z}$ holds. This yields the bound (4.6) as required.

If $r \geq 2$, then after reordering if necessary, we can assume without loss of generality that the form of vertex $\boldsymbol{x}^*$ and the vertex of the corner polyhedron $\boldsymbol{z}^*$ have the form $\boldsymbol{x}^* = (x_1^*, \ldots, x_m^*, 0, \ldots, 0)^T$ and $\boldsymbol{z}^* = (z_1^*, \ldots, z_m^*, z_{m+1}^*, \ldots, z_{m+r}^*, 0, \ldots, 0)^T$ respectively. In consequence, $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1$ can be bounded by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 &= \left|x_1^* - z_1^*\right| + \cdots + \left|x_m^* - z_m^*\right| + \left|z_{m+1}^*\right| + \cdots + \left|z_{m+r}^*\right| \\ &< \underbrace{\frac{r}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right) + \cdots + \frac{r}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right)}_{(r+m)\text{-terms}} = \frac{r(r+m)}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right),\end{aligned}$$

where the strict inequality follows in light of (4.4) and by similarly noting that there must exist at least one $i \in \{1, \ldots, m\}$ such that $\left|x_i^* - z_i^*\right| \notin \mathbb{Z}$ holds when $r \geq 2$. Furthermore, in the case when $r \geq 2$, we notice that

$$\frac{r(r+m)}{2^r} = \frac{r^2}{2^r} + \frac{rm}{2^r} \leq \frac{9}{8} + \frac{m}{2} \tag{4.47}$$

holds. Observe $\frac{r^2}{2^r} = \frac{9}{8}$ only if $r = 3$ and $\frac{rm}{2^r} = \frac{m}{2}$ only if $r = 2$. In particular, because the two equalities are attained for different values of r, we conclude that the final equality in (4.47) can be tightened to state that

$$\frac{r(r+m)}{2^r} = \frac{r^2}{2^r} + \frac{rm}{2^r} < \frac{9}{8} + \frac{m}{2}$$

holds. This strict upper bound further yields the bound (4.7) as required and concludes the proof of Corollary 4.1.3.

4.5 Proof of Theorem 4.1.4

As in the proof of Theorem 4.1.1, we have that Theorem 4.1.4 is an immediate consequence of Lemma 4.3.1 and the generalization of Theorem 4.2.5 given below. Recall that τ denotes the support of $\boldsymbol{x}^*$, and that $\mathscr{C}_\tau(A, \boldsymbol{b})$ denotes the polyhedron defined as

$$\mathscr{C}_\tau(A, \boldsymbol{b}) = \operatorname{conv}\left(\{\boldsymbol{x} \in \mathbb{Z}^n : A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x}_{\bar{\tau}} \geq \boldsymbol{0}\}\right).$$

Theorem 4.5.1. *Let $\boldsymbol{z}^*$ be a vertex of $\mathscr{C}_\tau(A, \boldsymbol{b})$. Then there exists a basis γ of A containing τ such that*

$$\prod_{j \in \bar{\gamma}} \left(z_j^* + 1\right) \leq r^d \, \frac{\left|\det(A_\gamma)\right|}{\gcd(A)},$$

where $r = \|\boldsymbol{z}_{\bar{\gamma}}^*\|_0$ and $d = m - |\tau|$.

Theorem 4.5.1 is proved over the remainder of this section, by constructing a convex set P such that

$$\prod_{j\in\bar{\gamma}}\left(z_j^*+1\right)\leq \mathrm{vol}_{n-m}(P)\leq r^d\,\frac{\left|\det(A_\gamma)\right|}{\gcd(A)}\,. \tag{4.48}$$

The notation $\mathrm{vol}_k(S)$ denotes the k-dimensional volume, or Lebesgue measure [14, Chapter 13], of $S\subset\mathbb{R}^d$. It is worth noting that if S is a k-dimensional convex body, then the two straightforward measurable properties of S are the k-dimensional volume/content and the $(k-1)$-dimensional boundary content.

The next subsection (entitled *convex geometry lemmas*) collects some facts from convex geometry that are used in this proof. The subsequent subsection (entitled *a special case of Theorem 4.5.1*) establishes that the inequalities (4.48) hold, and hence Theorem 4.5.1 holds, in the special case when $\tau=\mathrm{supp}(\boldsymbol{x}^*)$ and $\mathrm{supp}(\boldsymbol{z}^*)$ together cover $\{1,\dots,n\}$, namely that $\tau\cup\mathrm{supp}(\boldsymbol{z}^*)=\{1,\dots,n\}$. Finally, the concluding subsection (entitled *proof of Theorem 4.5.1*) utilises this special case in order to establish the general case of Theorem 4.5.1.

4.5.1 Convex geometry lemmas

Lemma 4.5.2 (Blichfeldt's lemma [24, Chapter 3, Theorem I])**.** *Let $K\subset\mathbb{R}^d$ be a bounded, nonempty and Lebesgue measurable set and let Λ be a full-dimensional lattice in $\mathbb{R}^d$. Suppose that the difference set $K-K$ contains no nonzero lattice points from Λ. Then* $\mathrm{vol}_d(K)\leq\det(\Lambda)$.

Theorem 4.5.3 (Brunn's concavity principle [9, Theorem 1.2.1])**.** *Let K be a convex body, and let F be a k-dimensional subspace of $\mathbb{R}^d$. Then the function $g: F^\perp\to\mathbb{R}$ defined by*

$$g(x)=\mathrm{vol}_k(K\cap(F+x))^{1/k}$$

is concave on its support.

By a *slab* in $\mathbb{R}^d$ we mean the closed region bounded by two distinct parallel hyperplanes. Let $\boldsymbol{q}\in\mathbb{R}^d$ be nonzero. The *width* of a set $K\subset\mathbb{R}^d$ along $\boldsymbol{q}$ is defined to be

$$w_{\boldsymbol{q}}(K):=\left(\sup_{\boldsymbol{x}\in K}\boldsymbol{q}^T\boldsymbol{x}\right)-\left(\inf_{\boldsymbol{x}\in K}\boldsymbol{q}^T\boldsymbol{x}\right).$$

Proposition 4.5.4. *Let $K\subset\mathbb{R}^d$ be a centrally symmetric convex body with centre $\mathbf{c}\in\mathbb{R}^d$. Let S be a slab containing $\mathbf{c}$, such that $\mathbf{c}$ is equidistant from the two facets defining S with respect to the Euclidean norm. Let $\boldsymbol{q}\in\mathbb{R}^d$ be a normal vector to either of the hyperplanes* *bounding S. If S does not contain K, then*

$$\mathrm{vol}_d(K \cap S) \geq \frac{w_{\boldsymbol{q}}(S)}{w_{\boldsymbol{q}}(K)} \cdot \mathrm{vol}_d(K).$$

Proof. Note that distance is invariant under translation and, in consequence, without loss of generality, we may assume $\boldsymbol{c}$ is the origin. For $\lambda \in [-1, 1]$, define the affine hyperplane

$$L_\lambda := \left\{ \boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{q}^\top \boldsymbol{x} = \lambda \cdot w_{\boldsymbol{q}}(K)/2 \right\}.$$

Let $K_\lambda := K \cap L_\lambda$, and define the cross-sectional volume

$$f(\lambda) := \mathrm{vol}_{d-1}(K_\lambda).$$

It should be emphasised that the function $f(\lambda)$ is defined with respect to the dimension of the affine hyperplane L_λ, namely $d-1$, rather than the dimension d since any linear subspace (or affine subspace) of $\mathbb{R}^d$ with dimension strictly less than d has zero d-dimensional standard Lebesgue measure [14, Chapter 11].

Recall that K is a centrally symmetric convex body and therefore

$$f(\lambda) = \mathrm{vol}_{d-1}(K_\lambda) = \mathrm{vol}_{d-1}(-K_{-\lambda}) = \mathrm{vol}_{d-1}(K_{-\lambda}) = f(-\lambda) \tag{4.49}$$

holds, where the penultimate equality holds because multiplication by -1 is volume preserving. In light of (4.49), we that note $f(\lambda)$ is an even function on $[-1, 1]$. Further to this, $g(\lambda) := (f(\lambda))^{1/(d-1)}$ is consequently an even function since

$$g(\lambda) = (f(\lambda))^{1/(d-1)} = (f(-\lambda))^{1/(d-1)} = g(-\lambda)$$

holds. Because $g(\lambda)$ is a concave function on $[-1, 1]$ in light of Brunn's concavity principle (Theorem 4.5.3), we have that $g(\lambda)$, and therefore $f(\lambda)$, are monotonically decreasing (non-increasing) functions on $[0, 1]$.

For convenience, let $\delta := w_{\boldsymbol{q}}(S)/w_{\boldsymbol{q}}(K)$. Then by Fubini's theorem (see e.g. [98, Chapter 4]), symmetry and monotonicity on $[0, 1]$, we conclude that

$$\mathrm{vol}_d(K \cap S) = \int_{-\delta}^{\delta} f(\lambda) d\lambda = 2 \int_0^{\delta} f(\lambda) d\lambda \geq 2\delta \int_0^1 f(\lambda) d\lambda = \frac{w_{\boldsymbol{q}}(S)}{w_{\boldsymbol{q}}(K)} \cdot \mathrm{vol}_d(K)$$

holds as required, which concludes the proof of Proposition 4.5.4. □

The notion of irreducibility from Lemma 4.2.1 can be mildly generalized by making use of pointed cones. Recall (from Chapter 2) that a convex cone $C \subset \mathbb{R}^d$ is pointed if and only if $C \cap (-C) = \{\boldsymbol{0}\}$ holds, where $-C$ denotes the negative cone of C.

Let C be a pointed cone and $\Lambda \subset \mathbb{Z}^d$ be a d-dimensional integer lattice. The integral point $\boldsymbol{x} \in C \cap \mathbb{Z}^d$ is called *irreducible (with respect to* Λ *and* C*)* if

$$(-\boldsymbol{x} + C) \cap (\boldsymbol{x} - C) \cap \Lambda = \{\boldsymbol{0}\}$$

holds. It should be emphasised that the previous notion of irreducibly (from Lemma 4.2.1) is obtained from this notion of irreducibly by letting $C = \mathbb{R}^d_{\geq 0}$, where the d-dimensional positive orthant $\mathbb{R}^d_{\geq 0}$ is simply an example of a pointed cone since $\mathbb{R}^d_{\geq 0}$ contains no subspace of $\mathbb{R}^d$ other than the origin $\{\boldsymbol{0}\}$.

Let $\boldsymbol{p} \in \mathbb{Z}^d$ and consider the affine lattice $\Gamma = \boldsymbol{p} + \Lambda$. We will call the set $\mathscr{E}(\Gamma, C) = \operatorname{conv}(\Gamma \cap C)$ the *sail* associated with the affine lattice Γ and the pointed cone C. Figure 4.3 provides an exemplification of the sail associated with an affine lattice Γ and a pointed cone C.

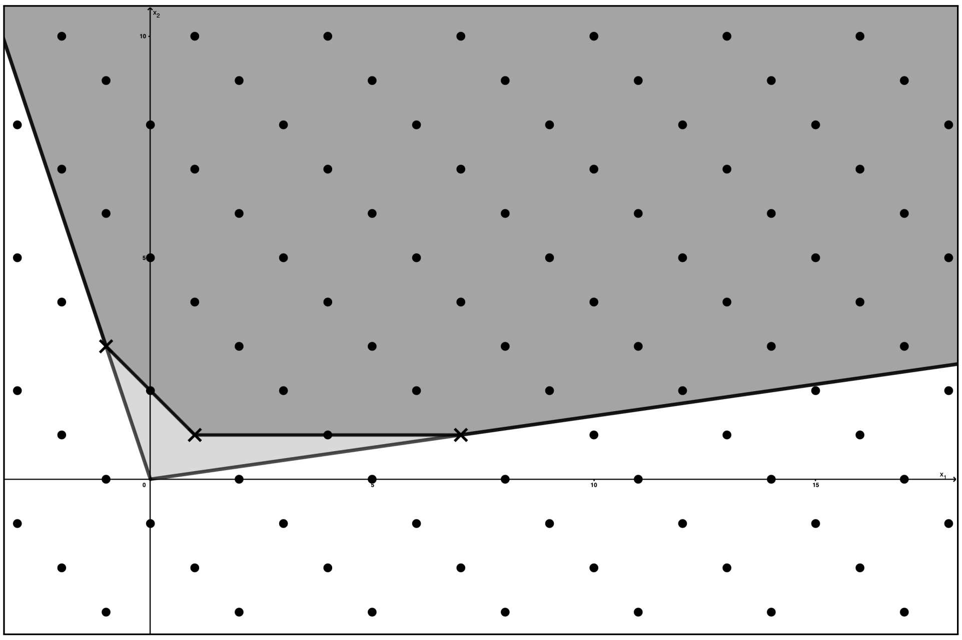


Figure 4.3: This figure illustrates, for an affine integer lattice Γ (black dots) and a pointed cone C (grey), the sail $\mathscr{E}(\Gamma, C) = \operatorname{conv}(\Gamma \cap C)$ (dark grey) associated with the affine lattice Γ and pointed cone C, where the three vertices of the sail $\mathscr{E}(\Gamma, C)$ are those points shown using black crosses.

Lemma 4.5.5. *Every vertex of the sail* $\mathscr{E}(\Gamma, C)$ *is irreducible (with respect to* Λ *and* C*)*.

Proof. Let $\boldsymbol{x}$ be a vertex of the sail $\mathscr{E}(\Gamma, C)$ associated with the affine lattice Γ and the pointed cone C. Suppose, to derive a contradiction, that the vertex $\boldsymbol{x}$ is reducible. Then in light of the notion of irreducibility given above, there exists a nonzero $\boldsymbol{\lambda} \in \Lambda \backslash \{\boldsymbol{0}\}$ and vectors $\boldsymbol{y}, \boldsymbol{y}' \in C$ such that $\boldsymbol{\lambda} = -\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} - \boldsymbol{y}'$. Notice that $\boldsymbol{x}$ is a vertex of $\mathscr{E}(\Gamma, C)$ implies that $\boldsymbol{x} \in \Gamma$ and, in consequence, both $\boldsymbol{y} = \boldsymbol{\lambda} + \boldsymbol{x}$ and $\boldsymbol{y}' = -\boldsymbol{\lambda} + \boldsymbol{x}$ are integral points contained in $\Gamma \cap C$, and further in $\mathscr{E}(\Gamma, C)$. It should be noted that the conclusion that $\boldsymbol{y}, \boldsymbol{y}' \in \Gamma \cap C$ follows because $\boldsymbol{x} \in \Gamma$, $\boldsymbol{\lambda} \in \Lambda \backslash \{\boldsymbol{0}\}$ and $\boldsymbol{y}, \boldsymbol{y}' \in C$. Since $\boldsymbol{\lambda} \in \Lambda \backslash \{\boldsymbol{0}\}$ is nonzero, we conclude that the vertex $\boldsymbol{x}$ can be written as $\boldsymbol{x} = (\boldsymbol{y} + \boldsymbol{y}')/2$ and is therefore not a vertex of $\mathscr{E}(\Gamma, C)$. This contradiction completes the proof of Lemma 4.5.5. □

4.5.2 A special case of Theorem 4.5.1

We next choose the basis γ of Theorem 4.5.1, define the convex set P in terms of this γ, and establish the lower and upper bounds of (4.48) in the special case when τ and $\mathrm{supp}(\boldsymbol{z}^*)$ together cover $\{1, \ldots, n\}$, namely that $\tau \cup \mathrm{supp}(\boldsymbol{z}^*) = \{1, \ldots, n\}$. In other words, throughout this subsection we assume that the condition $\bar{\tau} \subset \mathrm{supp}(\boldsymbol{z}^*)$ holds.

Proposition 4.5.6. *If $\bar{\tau} \subset \mathrm{supp}(\boldsymbol{z}^*)$, then there exists a basis γ of A containing τ which satisfies*

$$z_i^* + 1 \geq \frac{1}{r} \sum_{j \in \bar{\gamma}} \left| \left(A_\gamma^{-1} A_{\bar{\gamma}} \right)_{i,j} \left(z_j^* + 1 \right) \right| \tag{4.50}$$

for each $i \in \gamma \backslash \tau$, where $r = \|\boldsymbol{z}_{\bar{\gamma}}^*\|_0$.

Proof. Among all bases of A containing $\tau = \mathrm{supp}(\boldsymbol{x}^*)$, we choose a basis γ such that the quantity $\left|\det\left(A_\gamma\right)\right| \cdot \prod_{i \in \gamma} \left(z_i^* + 1\right)$ is as large as possible. It should be noted that this choice of γ implies that for any basis γ' of A containing τ, then the inequality

$$\left|\det\left(A_\gamma\right)\right| \prod_{i \in \gamma} \left(z_i^* + 1\right) \geq \left|\det\left(A_{\gamma'}\right)\right| \prod_{i' \in \gamma'} \left(z_{i'}^* + 1\right)$$

holds. In light of $\gamma = \tau \cup (\gamma \backslash \tau)$ and $\gamma' = \tau \cup (\gamma' \backslash \tau)$, we can express this inequality equivalently as

$$\left|\det\left(A_\gamma\right)\right| \prod_{i_1 \in \tau} \left(z_{i_1}^* + 1\right) \prod_{i_2 \in \gamma \backslash \tau} \left(z_{i_2}^* + 1\right) \geq \left|\det\left(A_{\gamma'}\right)\right| \prod_{i_1' \in \tau} \left(z_{i_1'}^* + 1\right) \prod_{i_2' \in \gamma' \backslash \tau} \left(z_{i_2'}^* + 1\right),$$

where noting that $\prod_{i_1 \in \tau} \left(z_{i_1}^* + 1\right) = \prod_{i_1' \in \tau} (z_{i_1'}^* + 1) \neq 0$ further yields that

$$\prod_{i_2 \in \gamma \backslash \tau} \left(z_{i_2}^* + 1\right) \geq \frac{\left|\det\left(A_{\gamma'}\right)\right|}{\left|\det\left(A_\gamma\right)\right|} \prod_{i_2' \in \gamma' \backslash \tau} \left(z_{i_2'}^* + 1\right).$$

The sets $\gamma \backslash \tau$ and $\gamma' \backslash \tau$ are not in general disjoint and, in consequence, one could similarly divide through by those brackets corresponding to those elements from $(\gamma \backslash \tau) \cap (\gamma' \backslash \tau)$. This simplification is the intuitive idea which we make use of below to conclude the proof of this proposition.

If $i \in \gamma$ and $j \in \bar{\gamma}$, then making use of Cramer's rule (Lemma 2.7.1), we find that

$$\left(A_\gamma^{-1} A_{\bar{\gamma}}\right)_{i,j} = \frac{\det\left(A_\gamma^i(A_j)\right)}{\det\left(A_\gamma\right)}, \tag{4.51}$$

where $A_\gamma^i(A_j)$ denotes the matrix obtained by replacing the i-th column of A_γ with the j-th column of A. It is worth noting for completeness that one could have equivalently obtained the equality (4.51) from a close inspection of $A_\gamma^{-1} A_{\bar{\gamma}} = {}^1/_{\det(A_\gamma)} \operatorname{adj}\left(A_\gamma\right) A_{\bar{\gamma}}$, where $\operatorname{adj}\left(A_\gamma\right)$ denotes the classical adjugate of the matrix A_γ (see e.g. [52, Chapter 4]).

The choice of basis γ implies that if $i \in \gamma \backslash \tau$ and $j \in \bar{\gamma}$, then

$$z_i^* + 1 \geq \left| \frac{\det\left(A_\gamma^i(A_j)\right)}{\det\left(A_\gamma\right)} \left(z_j^* + 1\right) \right| = \left| \left(A_\gamma^{-1} A_{\bar{\gamma}}\right)_{i,j} \left(z_j^* + 1\right) \right|.$$

The assumed condition $\bar{\tau} \subset \operatorname{supp}(\boldsymbol{z}^*)$ implies that $r = |\bar{\gamma}|$ and, in consequence, for all $i \in \gamma \backslash \tau$, we have that

$$z_i^* + 1 \geq \frac{1}{r} \sum_{j \in \bar{\gamma}} \left| \left(A_\gamma^{-1} A_{\bar{\gamma}}\right)_{i,j} \left(z_j^* + 1\right) \right|$$

holds as required, which concludes the proof of Proposition 4.5.6. □

We now fix a basis γ of A containing τ satisfying the hypotheses of Proposition 4.5.6 so that in particular $\bar{\tau} \subset \operatorname{supp}(\boldsymbol{z}^*)$ holds and let $r = \|\boldsymbol{z}_{\bar{\gamma}}^*\|_0$. Reordering the columns of A if necessary, we may assume without loss of generality that $\gamma = \{1, 2, \ldots, m\}$ and further assume $\gamma \backslash \tau = \{1, 2, \ldots, d\}$. We denote the rows of the matrix $-A_\gamma^{-1} A_{\bar{\gamma}}$ by $\boldsymbol{q}_1^T, \boldsymbol{q}_2^T, \ldots, \boldsymbol{q}_m^T$. Observe that the equalities $A\boldsymbol{z}^* = \boldsymbol{b}$ and $A_\gamma^{-1}\boldsymbol{b} = \boldsymbol{x}^*$ imply in light of (4.35) that $\boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* = z_i^* - x_i^*$ for all $i \in \gamma$.

For convenience, observe that making use of the notation $\boldsymbol{q}_1^T, \boldsymbol{q}_2^T, \ldots, \boldsymbol{q}_m^T$ allows us to rewrite the inequality (4.50) from Proposition 4.5.6 as

$$r\left(z_i^* + 1\right) \geq \sum_{j \in \bar{\gamma}} \left| \left(A_\gamma^{-1} A_{\bar{\gamma}}\right)_{i,j} \left(z_j^* + 1\right) \right| = \sum_{j \in \bar{\gamma}} \left| q_{i,j} \left(z_j^* + 1\right) \right|, \tag{4.52}$$

where $q_{i,j}$ denotes the j-th entry of the row vector $\boldsymbol{q}_i^T$.

Let $\mathbf{1}_{n-m} \in \mathbb{R}^{n-m}$ be the $(n-m)$-dimensional vector where each entry is equal to one, and define, for each $i \in \gamma \backslash \tau$,

$$S_i := \left\{ \boldsymbol{x} \in \mathbb{R}^{n-m} : -\frac{1}{2} < \boldsymbol{q}_i^T \boldsymbol{x} < \boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* + \frac{1}{2} \right\}.$$

Further to this, define

$$B := \left\{ \boldsymbol{x} \in \mathbb{R}^{n-m} : -\frac{1}{2}\mathbf{1}_{n-m} < \boldsymbol{x} < \boldsymbol{z}_{\bar{\gamma}}^* + \frac{1}{2}\mathbf{1}_{n-m} \right\}$$

and for each $i \in \gamma\backslash\tau$, let $P_i := P_{i-1} \cap S_i$ with $P_0 = B$. Recall that $\gamma\backslash\tau = \{1, 2, \ldots, d\}$ by assumption and hence, for convenience, we let $P := P_d$. It is worth emphasising that P is by construction a centrally symmetric convex set and that $P \subset P_i \subset B$ for each $i \in \gamma\backslash\tau$.

Lemma 4.5.7. *If* $\bar{\tau} \subset \operatorname{supp}(\boldsymbol{z}^*)$, *then*

$$\operatorname{vol}_{n-m}(P) \geq \frac{1}{r^d} \prod_{j \in \bar{\gamma}} \left(z_j^* + 1\right).$$

Proof. Suppose that $i \in \gamma\backslash\tau$. If S_i contains P_{i-1}, then by the construction of P_i, we notice that $P_{i-1} = P_i$ and therefore

$$\operatorname{vol}_{n-m}(P_i) = \operatorname{vol}_{n-m}(P_{i-1}).$$

Otherwise, we define

$$\lambda_i := \frac{w_{\boldsymbol{q}_i}(S_i)}{w_{\boldsymbol{q}_i}(P_{i-1})}.$$

Recall that by assumption $\bar{\tau} \subset \operatorname{supp}(\boldsymbol{z}^*)$ and, in consequence, we notice that $z_i^* \geq 1$ for all $i \in \gamma\backslash\tau$. Using the definitions of S_i and B allows us to bound λ_i from below by

$$\lambda_i = \frac{w_{\boldsymbol{q}_i}(S_i)}{w_{\boldsymbol{q}_i}(P_{i-1})} \geq \frac{w_{\boldsymbol{q}_i}(S_i)}{w_{\boldsymbol{q}_i}(B)} = \frac{\boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* + 1/2 - (-1/2)}{\sum_{j \in \bar{\gamma}} \left| q_{i,j} \left(z_j^* + 1\right)\right|} = \frac{\boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* + 1}{\sum_{j \in \bar{\gamma}} \left| q_{i,j} \left(z_j^* + 1\right)\right|}.$$

Further, upon noting that $\boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* + 1 = z_i^* + 1$ holds for $i \in \gamma\backslash\tau$ and that Proposition 4.5.6 applies, we can bound λ_i from below by

$$\lambda_i \geq \frac{\boldsymbol{q}_i^T \boldsymbol{z}_{\bar{\gamma}}^* + 1}{\sum_{j \in \bar{\gamma}} \left| q_{i,j} \left(z_j^* + 1\right)\right|} = \frac{z_i^* + 1}{\sum_{j \in \bar{\gamma}} \left| q_{i,j} \left(z_j^* + 1\right)\right|} \geq \frac{z_i^* + 1}{r\left(z_i^* + 1\right)} = \frac{1}{r}, \tag{4.53}$$

where the final inequality follows from (4.52). When S_i does not contain P_{i-1} by assumption, we can apply Proposition 4.5.4 in addition to (4.53) to find

$$\operatorname{vol}_{n-m}(P_i) \geq \frac{w_{\boldsymbol{q}_i}(S_i)}{w_{\boldsymbol{q}_i}(P_{i-1})} \operatorname{vol}_{n-m}(P_{i-1}) = \lambda_i \operatorname{vol}_{n-m}(P_{i-1}) \geq \frac{1}{r} \operatorname{vol}_{n-m}(P_{i-1}). \tag{4.54}$$

Observe that (4.54) additionally holds when S_i contains P_i. We can in consequence apply induction to the sequence of polytopes $P = P_d, \ldots, P_1, P_0 = B$, which upon making use of (4.54) yields that

$$\begin{aligned}\operatorname{vol}_{n-m}(P) = \operatorname{vol}_{n-m}(P_d) &\geq \frac{1}{r} \operatorname{vol}_{n-m}(P_{d-1}) \geq \frac{1}{r^2} \operatorname{vol}_{n-m}(P_{d-2}) \geq \cdots \\ &\geq \frac{1}{r^d} \operatorname{vol}_{n-m}(P_0) = \frac{1}{r^d} \operatorname{vol}_{n-m}(B) = \frac{1}{r^d} \prod_{j \in \bar{\gamma}} \left(z_j^* + 1\right)\end{aligned}$$

holds as required, which concludes the proof of Lemma 4.5.7. □

Lemma 4.5.8. *If* $\bar{\tau} \subset \operatorname{supp}(\boldsymbol{z}^*)$, *then*

$$\operatorname{vol}_{n-m}(P) \leq \frac{\left|\det\left(A_\gamma\right)\right|}{\gcd(A)}. \tag{4.55}$$

Proof. Recall that we defined the lattice $\Lambda_\gamma(A) = \pi_\gamma\left(\ker(A) \cap \mathbb{Z}^n\right)$, whose determinant is given by $\left|\det\left(A_\gamma\right)\right| / \gcd(A)$ in light of (4.36). We show that $(P - P) \cap \Lambda_\gamma(A) = \{\mathbf{0}\}$, where $P - P$ denotes the difference body associated to P. The conclusion then follows immediately by Blichfeldt's lemma (Lemma 4.5.2).

Suppose that $\boldsymbol{u}, \boldsymbol{v} \in P$ and $\boldsymbol{u} - \boldsymbol{v} \in \Lambda_\gamma(A)$. Since P is by construction centrally symmetric, then the difference body $P - P$ is in consequence precisely the origin-symmetric translate of $2P$. In light of this, making use of the definitions of P, S_i and B yields

$$\begin{aligned} -\boldsymbol{z}^*_{\bar{\gamma}} - \mathbf{1}_{n-m} < \boldsymbol{u} - \boldsymbol{v} < \boldsymbol{z}^*_{\bar{\gamma}} + \mathbf{1}_{n-m} \\ -\boldsymbol{q}_i^T \boldsymbol{z}^*_{\bar{\gamma}} - 1 < \boldsymbol{q}_i^T(\boldsymbol{u} - \boldsymbol{v}) < \boldsymbol{q}_i^T \boldsymbol{z}^*_{\bar{\gamma}} + 1 \text{ for all } i \in \gamma \backslash \tau. \end{aligned} \tag{4.56}$$

Observe that the lattice $\Lambda_\gamma(A)$ can be equivalently characterized as the set of points $\boldsymbol{x} \in \mathbb{Z}^{n-m}$ such that $\boldsymbol{q}_i^T \boldsymbol{x} \in \mathbb{Z}$ for each $i \in \gamma$. Hence, the integrality of $\boldsymbol{u} - \boldsymbol{v}$ and the inequalities from (4.56) imply

$$\begin{aligned} -\boldsymbol{z}^*_{\bar{\gamma}} \leq \boldsymbol{u} - \boldsymbol{v} \leq \boldsymbol{z}^*_{\bar{\gamma}} \\ -\boldsymbol{q}_i^T \boldsymbol{z}^*_{\bar{\gamma}} \leq \boldsymbol{q}_i^T(\boldsymbol{u} - \boldsymbol{v}) \leq \boldsymbol{q}_i^T \boldsymbol{z}^*_{\bar{\gamma}} \text{ for all } i \in \gamma \backslash \tau. \end{aligned}$$

In particular, $\boldsymbol{u} - \boldsymbol{v}$ lies in the polyhedron $(-\boldsymbol{z}^*_{\bar{\gamma}} + C) \cap (\boldsymbol{z}^*_{\bar{\gamma}} - C)$, where

$$C := \left\{\boldsymbol{x} \in \mathbb{R}^{n-m} : \boldsymbol{x} \geq \mathbf{0}, \boldsymbol{q}_i^T \boldsymbol{x} \geq 0 \text{ for all } i \in \gamma \backslash \tau\right\}.$$

Observe that C is a pointed cone. Recall that $\boldsymbol{z}^*_{\bar{\gamma}}$ is by assumption a vertex of the sail $\mathscr{E}(\Lambda_\gamma(A, \boldsymbol{b}), C)$. In consequence, $\boldsymbol{z}^*_{\bar{\gamma}}$ is irreducible (with respect to the lattice $\Lambda_\gamma(A)$ and the pointed cone C) by Lemma 4.5.5 and therefore $\boldsymbol{u} = \boldsymbol{v}$. In other words, we notice that $(P - P) \cap \Lambda_\gamma(A) = \{\mathbf{0}\}$ holds and, in consequence to Lemma 4.5.2 (Blichfeldt's lemma), we obtain that the inequality (4.55) holds as required, which concludes the proof of Lemma 4.5.8. □

In order to summarize this subsection, notice that we have proven the following special case of Theorem 4.5.1.

Corollary 4.5.9. *Suppose that* τ *and* $\operatorname{supp}(\boldsymbol{z}^*)$ *together cover* $\{1, \ldots, n\}$. *Then there exists a basis* γ *of* A *containing* τ *such that*

$$\prod_{j \in \bar{\gamma}} \left(z_j^* + 1\right) \leq r^d \, \frac{\left|\det\left(A_\gamma\right)\right|}{\gcd(A)}.$$

4.5.3 Proof of Theorem 4.5.1

Recall that $\tau = \mathrm{supp}(\boldsymbol{x}^*)$. In order to complete the proof of Theorem 4.5.1, it remains to deal with the case when $\bar{\tau}$ is not necessarily contained in $\mathrm{supp}(\boldsymbol{z}^*)$. Fix a vertex $\boldsymbol{z}^*$ of $\mathscr{C}_\tau(A, \boldsymbol{b})$ and let $\mu = \tau \cup \mathrm{supp}(\boldsymbol{z}^*)$. Upon reordering if necessary, we may assume without loss of generality that $\mu = \{1, 2, \ldots, |\mu|\}$, i.e. that the matrix A is written as $A = (A_\mu, A_{\bar{\mu}})$, and let $\bar{A}_\mu$ be any integer matrix with full row rank which has the same row space as A_μ. Notice that $\boldsymbol{x}^*_\mu$ is a basic feasible solution of the system

$$\bar{A}_\mu \boldsymbol{x}_\mu = \bar{\boldsymbol{b}}, \boldsymbol{x}_\mu \geq \boldsymbol{0}, \tag{4.57}$$

where $\bar{\boldsymbol{b}} := \bar{A}_\mu \boldsymbol{x}^*_\mu$. Moreover, let

$$\mathscr{C}_\tau(\bar{A}_\mu, \bar{\boldsymbol{b}}) := \mathrm{conv}\left(\{\boldsymbol{x}_\mu \in \mathbb{Z}^{|\mu|} : \bar{A}_\mu \boldsymbol{x}_\mu = \bar{\boldsymbol{b}}, \boldsymbol{x}_{\mu \setminus \tau} \geq \boldsymbol{0}\}\right).$$

We have that $\mathscr{C}_\tau(\bar{A}_\mu, \bar{\boldsymbol{b}}) \times \{\boldsymbol{0}\}^{|\bar{\mu}|}$ is the face of the polyhedron $\mathscr{C}_\tau(A, \boldsymbol{b})$ for which the constraints $\boldsymbol{x}_{\bar{\mu}} \geq \boldsymbol{0}$ are tight and, moreover, this face contains $\boldsymbol{z}^* = (\boldsymbol{z}^*_\mu, \boldsymbol{0})$. For completeness, this conclusion follows since $\mathscr{C}_\tau(\bar{A}_\mu, \bar{\boldsymbol{b}}) \times \{\boldsymbol{0}\}^{|\bar{\mu}|} \subset \mathscr{C}_\tau(A, \boldsymbol{b})$. Hence, $\boldsymbol{z}^*_\mu$ is a vertex of $\mathscr{C}_\tau(\bar{A}_\mu, \bar{\boldsymbol{b}})$ such that $\mu \setminus \tau \subset \mathrm{supp}(\boldsymbol{z}^*_\mu)$ holds. In light of this, we can therefore apply Corollary 4.5.9 to both $\boldsymbol{x}^*_\mu$ and $\boldsymbol{z}^*_\mu$ in addition to the linear system (4.57) in order to obtain a basis $\sigma \subset \mu$ of $\bar{A}_\mu$ containing τ which satisfies

$$\prod_{j \in \mu \setminus \sigma} \left(z^*_j + 1\right) \leq r^{|\sigma| - |\tau|} \frac{\left|\det\left(\bar{A}_\sigma\right)\right|}{\gcd(\bar{A}_\mu)}. \tag{4.58}$$

Now let γ be a basis of A containing σ. Upon noting that

$$\mu \cup \gamma = \underbrace{\sigma \cup (\mu \setminus \sigma)}_{\mu} \cup \underbrace{\sigma \cup (\gamma \setminus \sigma)}_{\gamma} = \sigma \cup (\mu \setminus \sigma) \cup (\gamma \setminus \sigma) = \mu \cup (\gamma \setminus \sigma)$$

holds, we notice that μ and $\gamma \setminus \sigma$ partition $\mu \cup \gamma$. In consequence, up to invertible row operations, we can write

$$A_{\mu \cup \gamma} = \begin{pmatrix} \bar{A}_\mu & \bar{A}_{\gamma \setminus \sigma} \\ \boldsymbol{0} & \bar{\bar{A}}_{\gamma \setminus \sigma} \end{pmatrix} = \begin{pmatrix} \bar{A}_{\mu \setminus \sigma} & \bar{A}_\sigma & \bar{A}_{\gamma \setminus \sigma} \\ \boldsymbol{0} & \boldsymbol{0} & \bar{\bar{A}}_{\gamma \setminus \sigma} \end{pmatrix}, \tag{4.59}$$

where both $\bar{A}_\sigma$ and $\bar{\bar{A}}_{\gamma \setminus \sigma}$ are both invertible. This follows upon recalling that γ is a basis of A and σ is a basis of $\bar{A}_\mu$, which implies that $\det(A_\gamma) = \det(A_\sigma, A_{\gamma \setminus \sigma}) \neq 0$, $\det(\bar{A}_\sigma) \neq 0$ and, in consequence, that $\det(\bar{\bar{A}}_{\gamma \setminus \sigma}) \neq 0$. Now, every nonzero maximal subdeterminant of $A_{\mu \cup \gamma}$ is the product of $\det(\bar{\bar{A}}_{\gamma \setminus \sigma})$ with a maximal subdeterminant of $\bar{A}_\mu$. It follows in consequence that

$$\gcd(A_{\mu \cup \gamma}) = \left|\det\left(\bar{\bar{A}}_{\gamma \setminus \sigma}\right)\right| \cdot \gcd(\bar{A}_\mu)$$

holds. It should be noted that this could have been equivalently deduced in light of (4.59) and results on the Smith normal form [92].

In consequence, we obtain

$$\frac{\left|\det\left(\bar{A}_{\sigma}\right)\right|}{\gcd(\bar{A}_{\mu})}=\frac{\left|\det\left(\bar{A}_{\sigma}\right)\right|\cdot\left|\det\left(\bar{\bar{A}}_{\gamma\setminus\sigma}\right)\right|}{\gcd(A_{\mu\cup\gamma})}=\frac{\left|\det\left(A_{\gamma}\right)\right|}{\gcd(A_{\mu\cup\gamma})}, \tag{4.60}$$

where the final equality follows because γ is a basis of the matrix A containing σ. Upon noting that $|\bar{\gamma}|\leq|\bar{\sigma}|$, $\gcd(A)\leq\gcd(A_{\mu\cup\gamma})$ and $|\sigma|-|\tau|\leq m-|\tau|=d$, then in consequence to (4.58) and (4.60), we conclude that

$$\prod_{j\in\bar{\gamma}}\left(z_j^*+1\right)\leq\prod_{j\in\mu\setminus\sigma}\left(z_j^*+1\right)\leq r^{|\sigma|-|\tau|}\,\frac{\left|\det\left(A_{\gamma}\right)\right|}{\gcd(A_{\mu\cup\gamma})}\leq r^d\,\frac{\left|\det\left(A_{\gamma}\right)\right|}{\gcd(A)}$$

holds as required, which concludes the proof of Theorem 4.5.1.

4.6 Proof of Corollary 4.1.5

Recall that $\boldsymbol{z}^*$ is an integral vertex of $\mathscr{C}_{\tau}(A,\boldsymbol{b})$ associated with the vertex $\boldsymbol{x}^*$ of the polyhedron $P(A,\boldsymbol{b})$, where $\tau=\operatorname{supp}(\boldsymbol{x}^*)$. In light of Theorem 4.1.4, we can bound the $\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1$ as follows. If $r=0$, we have $\boldsymbol{x}^*=\boldsymbol{z}^*$ by (4.8) which justifies (4.11).

If $r=1$, then after reordering if necessary, we can assume without loss of generality that the vertex $\boldsymbol{x}^*$ and the vertex of the corner polyhedron $\boldsymbol{z}^*$ have the form $\boldsymbol{x}^*=(x_1^*,\ldots,x_{|\tau|}^*,0,\ldots,0)^T$ and $\boldsymbol{z}^*=(z_1^*,\ldots,z_{|\tau|}^*,z_{|\tau|+1}^*,0,\ldots,0)^T$, respectively. In consequence, $\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1$ can be bounded by

$$\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1=\left|x_1^*-z_1^*\right|+\cdots+\left|x_{|\tau|}^*-z_{|\tau|}^*\right|+\left|z_{|\tau|+1}^*\right|<(|\tau|+1)\left(\frac{\Delta(A)}{\gcd(A)}-1\right),$$

where the final inequality follows directly from (4.9) and by noting that there must exist at least one $i\in\{1,\ldots,|\tau|\}$ such that $\left|x_i^*-z_i^*\right|\notin\mathbb{Z}$ holds. This yields the bound (4.12) as required.

If $r\geq 2$, then after reordering if necessary, we can assume without loss of generality that the form of vertex $\boldsymbol{x}^*$ and the vertex of the corner polyhedron $\boldsymbol{z}^*$ have the form $\boldsymbol{x}^*=(x_1^*,\ldots,x_{|\tau|}^*,0,\ldots,0)^T$ and $\boldsymbol{z}^*=(z_1^*,\ldots,z_{|\tau|}^*,z_{|\tau|+1}^*,\ldots,z_{|\tau|+r}^*,0,\ldots,0)^T$, respectively. In consequence, $\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1$ can be bounded with $d=m-|\tau|$ by

$$\begin{aligned}\|\boldsymbol{x}^*-\boldsymbol{z}^*\|_1&=\left|x_1^*-z_1^*\right|+\cdots+\left|x_{|\tau|}^*-z_{|\tau|}^*\right|+\left|z_{|\tau|+1}^*\right|+\cdots+\left|z_{|\tau|+r}^*\right|\\&<\underbrace{\frac{r^{d+1}}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right)+\cdots+\frac{r^{d+1}}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right)}_{(r+|\tau|)\text{-terms}}=\frac{r^{d+1}(r+|\tau|)}{2^r}\left(\frac{\Delta(A)}{\gcd(A)}\right),\end{aligned}$$

where the strict inequality follows in light of (4.10) and by similarly noting that there must exist at least one $i \in \{1, \ldots, |\tau|\}$ such that $\left|x_i^* - z_i^*\right| \notin \mathbb{Z}$ holds when $r \geq 2$. This strict upper bound further yields the bound (4.13) as required and concludes the proof of Corollary 4.1.5.

4.7 Proof of Theorem 4.1.6

Without loss of generality, we assume in this proof that $\gamma = \{1\}$. Hence, we assume that the vertex $\boldsymbol{x}^*$ has the form

$$\boldsymbol{x}^* = \frac{b}{a_1} \boldsymbol{e}_1 \,.$$

The corner polyhedron associated with the vertex $\boldsymbol{x}^*$ can consequently be written as

$$\mathscr{C}_\gamma(\boldsymbol{a}, b) = \operatorname{conv}\left(\{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{a}^T \boldsymbol{x} = b, x_2 \geq 0, \ldots, x_n \geq 0\}\right).$$

First we will show that the knapsack polytope $P(\boldsymbol{a}, b)$ contains a vertex of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$. Let $\boldsymbol{z}^*$ be a vertex of $\mathscr{C}_\gamma(\boldsymbol{a}, b)$ that gives an optimal solution to the linear program

$$\max\left\{x_1 : \boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathscr{C}_\gamma(\boldsymbol{a}, b)\right\}.$$

By the definition of $\mathscr{C}_\gamma(\boldsymbol{a}, b)$, the vertex $\boldsymbol{z}^*$ is in $P(\boldsymbol{a}, b)$ if and only if $z_1^* \geq 0$. Since $P_I(\boldsymbol{a}, b) \subset \mathscr{C}_\gamma(\boldsymbol{a}, b)$, where $P_I(\boldsymbol{a}, b)$ denotes the integral hull of $P(\boldsymbol{a}, b)$, it is sufficient to choose any integer point $\boldsymbol{z} = (z_1, \ldots, z_n)^T \in P(\boldsymbol{a}, b)$ and observe that $z_1^* \geq z_1 \geq 0$.

Applying Theorem 4.1.1 with the vertex $\boldsymbol{z}^* \in P(\boldsymbol{a}, b)$ we immediately obtain (4.17) and (4.18). Further, the bound (4.4) implies for $r \geq 2$ the non-strict inequality

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \frac{2^r}{r} \leq \|\boldsymbol{a}\|_\infty \,. \tag{4.61}$$

In order to show that (4.61) is strict (and hence that (4.19) holds), it is sufficient to prove that the bound (4.46) in the proof of Theorem 4.1.1 is strict in the knapsack scenario. Specifically, we need to prove that for the vertex $\boldsymbol{z}^*$

$$\delta = \left|x_1^* - z_1^*\right| < \frac{\left(z_2^* + \cdots + z_n^*\right) \|\boldsymbol{a}\|_\infty}{a_1} \,.$$

Set $A = (a_1, \ldots, a_n) \in \mathbb{Z}^{1 \times n}$ and consider the affine lattice $\Lambda(\boldsymbol{a}, b) := \Lambda_\gamma(A, b)$. We can write the affine lattice as

$$\Lambda(\boldsymbol{a}, b) = \left\{(\lambda_2, \ldots, \lambda_n)^T \in \mathbb{Z}^{n-1} : \lambda_2 a_2 + \cdots + \lambda_n a_n \equiv b \,(\operatorname{mod} a_1)\right\}. \tag{4.62}$$

Following (4.35), the map $\pi_\gamma|_{H(A,b)}$ is a bijection. It follows that the point $\boldsymbol{y}^* = \pi_\gamma(\boldsymbol{z}^*)$ is a vertex of the sail $\mathscr{E}(\Lambda(\boldsymbol{a}, b))$.

Suppose, to derive a contradiction, that the equality

$$\delta = \frac{\left(z_2^* + \cdots + z_n^*\right) \|\boldsymbol{a}\|_\infty}{a_1} \tag{4.63}$$

holds. By (4.43) we have

$$\delta = \frac{z_2^* a_2 + \cdots + z_n^* a_n}{a_1}$$

and, consequently, (4.63) implies $a_2 = \cdots = a_n = \|\boldsymbol{a}\|_\infty$. Therefore, using (4.62), the affine lattice $\Lambda(\boldsymbol{a}, b)$ contains the points

$$(z_2^* + \cdots + z_n^*)\boldsymbol{e}_j, \; j \in \{2, \ldots, n\}. \tag{4.64}$$

The point $\boldsymbol{y}^* = (z_2^*, \ldots, z_n^*)^T$, in its turn, belongs to the simplex with vertices (4.64) and has $\|\boldsymbol{y}\|_0 = r \geq 2$. Therefore $\boldsymbol{y}^*$ cannot be a vertex of the sail $\mathscr{E}(\Lambda(\boldsymbol{a}, b))$. The derived contradiction completes the proof of Theorem 4.1.6.

4.8 Proof of Corollary 4.1.7

By Theorem 4.1.6 we know that the knapsack polytope $P(\boldsymbol{a}, b)$ contains an integral vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$. In consequence, we can bound $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1$ as follows. If $r = 0$, we clearly have $\boldsymbol{x}^* = \boldsymbol{z}^*$, which justifies (4.21).

If $r = 1$, then after reordering if necessary, we can assume without loss of generality that the vertex of the corner polyhedron $\boldsymbol{z}^*$ has the form $\boldsymbol{z}^* = (z_1^*, z_2^*, 0, \ldots, 0)^T$ and, in consequence, the vertex distance $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1$ can be bounded by

$$\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 = \left|x_1^* - z_1^*\right| + \left|z_2^*\right| < 2(\|\boldsymbol{a}\|_\infty - 1),$$

where the final inequality follows directly from (4.18) and by noting that $\left|x_1^* - z_1^*\right| \notin \mathbb{Z}$. This yields (4.22) as required.

If $r \geq 2$, then after reordering if necessary, we can assume without loss of generality that the vertex of the corner polyhedron $\boldsymbol{z}^*$ has the form $\boldsymbol{z}^* = (z_1^*, z_2^*, \ldots, z_{r+1}^*, 0, \ldots, 0)^T$ and, in consequence, $\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1$ can be bounded by

$$\begin{aligned} \|\boldsymbol{x}^* - \boldsymbol{z}^*\|_1 &= \left|x_1^* - z_1^*\right| + \left|z_2^*\right| + \cdots + \left|z_{r+1}^*\right| \\ &< \underbrace{\frac{r}{2^r}\|\boldsymbol{a}\|_\infty + \cdots + \frac{r}{2^r}\|\boldsymbol{a}\|_\infty}_{(r+1)\text{-terms}} = \frac{r(r+1)}{2^r}\|\boldsymbol{a}\|_\infty, \end{aligned}$$

where the strict inequality follows in light of (4.19) and through the observation that $\left|x_1^* - z_1^*\right| \notin \mathbb{Z}$ when $r \geq 2$. Furthermore, in the case when $r \geq 2$, we notice that

$$\frac{r(r+1)}{2^r} \leq \frac{3}{2}$$

holds with equality only if $r \in \{2,3\}$. This upper bound further yields the bound (4.23) as required and concludes the proof of Corollary 4.1.7.

4.9 Proof of Corollary 4.1.8

By Theorem 4.1.6 we know that the knapsack polytope $P(\boldsymbol{a}, b)$ contains an integral vertex $\boldsymbol{z}^*$ of the corner polyhedron $\mathscr{C}_\gamma(\boldsymbol{a}, b)$. Therefore, the integrality gap $IG(\boldsymbol{c}, \boldsymbol{a}, b)$ can be bounded by

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) \le \|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \sum_{i \in \operatorname{supp}(\boldsymbol{x}^* - \boldsymbol{z}^*)} |c_i| \,. \tag{4.65}$$

If $r = 0$, we clearly have $\boldsymbol{x}^* = \boldsymbol{z}^*$, that justifies (4.26). Furthermore, noting that $\boldsymbol{x}^*$ has at most nonzero entry, the inequality (4.65) implies that

$$IG(\boldsymbol{c}, \boldsymbol{a}, b) \le (r+1)\|\boldsymbol{x}^* - \boldsymbol{z}^*\|_\infty \|\boldsymbol{c}\|_\infty,$$

where performing a simple rearrangement immediately yields the bounds (4.27) and (4.28) as required which concludes the proof of Corollary 4.1.8.

4.10 Proof of Theorem 4.1.9

For $n \ge 2$, we set

$$\boldsymbol{a}^{(n)} = (2^{n-1}, 2^{n-2}, \dots, 1)^T \text{ and } b^{(n)} = \mathbf{1}_n^T \boldsymbol{a}^{(n)} = 2^n - 1\,,$$

where $\mathbf{1}_n = (1, \dots, 1)^T \in \mathbb{Z}^n$ denotes the n-dimensional integral vector where each entry is equal to one. Let $P_I\left(\boldsymbol{a}^{(n)}, b^{(n)}\right) = \operatorname{conv}\left(P\left(\boldsymbol{a}^{(n)}, b^{(n)}\right) \cap \mathbb{Z}^n\right)$ be the integer hull of the knapsack polytope $P\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$.

We will make use of the following two observations.

Lemma 4.10.1. *The point $\mathbf{1}_n$ is a vertex of the polytope $P_I\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$.*

Proof. We will use induction on n. The basis step $n = 2$ holds as there are only two integer points $\mathbf{1}_2$ and $(0,3)^T$ in the polytope $P\left(\boldsymbol{a}^{(2)}, b^{(2)}\right)$. In order to verify the inductive step, suppose that the result does not hold for some $n \ge 3$. Observe that any integer point $\boldsymbol{z} = (z_1, \dots, z_n)^T \in P\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$ has $z_1 \le 1$. Consequently, the point $\mathbf{1}_n$ belongs to the face $P_I\left(\boldsymbol{a}^{(n)}, b^{(n)}\right) \cap \{\boldsymbol{x} \in \mathbb{R}^n : x_1 = 1\}$ of the polyhedron $P_I\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$. Hence, the point $\mathbf{1}_n$ is a convex combination of some integer points in $P\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$ that have the first entry 1. Therefore, removing the first entry we obtain a convex combination of integer points from $P\left(\boldsymbol{a}^{(n-1)}, b^{(n-1)}\right)$ equal to $\mathbf{1}_{n-1}$. The obtained contradiction completes the proof of Lemma 4.10.1. $\square$

In light of the assumption regarding the size of the dimension n in the statement of Theorem 4.1.9, we assume during the rest of the proof of this result that $n \geq 3$.

Lemma 4.10.2. *The point $\mathbf{1}_{n-1}$ is a vertex of the sail $\mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)\right)$.*

Proof. Using (4.62), the affine lattice $\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$ can be written as

$$\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right) = \left\{\boldsymbol{x} \in \mathbb{Z}^{n-1} : 2^{n-2}x_2 + \cdots + x_n \equiv -1 \, (\bmod 2^{n-1})\right\},$$

which follows since $2^n - 1 = 2^{n-1} \cdot 2 - 1 \equiv -1 (\bmod 2^{n-1})$. Therefore,

$$\mathscr{H} = \left\{\boldsymbol{x} \in \mathbb{R}^{n-1} : 2^{n-2}x_2 + \cdots + x_n = 2^{n-1} - 1\right\}$$

is a supporting hyperplane of $\mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)\right)$. Consequently,

$$P_I\left(\boldsymbol{a}^{(n-1)}, b^{(n-1)}\right) = \mathscr{H} \cap \mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)\right)$$

is a face of $\mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)\right)$. The result now follows in light of Lemma 4.10.1. □

For a positive integer t, set

$$\boldsymbol{a}^{(n)}(t) = \left(a_1^{(n)}(t), \ldots, a_n^{(n)}(t)\right)^T = \left(2^{n-1}, 2^{n-2} + t2^{n-1}, \ldots, 1 + t2^{n-1}\right)^T$$

and $b^{(n)}(t) = \mathbf{1}_n^T \boldsymbol{a}^{(n)}(t) = 2^n + (n-1)t2^{n-1} - 1$. Consider the vertex

$$\boldsymbol{v}^{(n)}(t) = \left(\frac{b^{(n)}(t)}{a_1^{(n)}(t)}\right)\boldsymbol{e}_1$$

of the knapsack polytope $P\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)$.

In view of (4.62), we have $\Lambda\left(\boldsymbol{a}^{(n)}, b^{(n)}\right) = \Lambda\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)$. Therefore, by Lemma 4.10.2, the point $\mathbf{1}_{n-1}$ is a vertex of the sail $\mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)\right)$. Observe that the sail $\mathscr{E}\left(\Lambda\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)\right)$ is the image of the corner polyhedron $\mathscr{C}_{\{1\}}\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$ under the bijective linear map $\pi_{\{1\}}|_{H\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)}(\cdot)$. Using (4.35), the n-dimensional point

$$\mathbf{1}_n = \pi_{\{1\}}^{-1}|_{H\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)}(\mathbf{1}_{n-1})$$

is a feasible vertex of $\mathscr{C}_{\{1\}}\left(\boldsymbol{a}^{(n)}, b^{(n)}\right)$. Note also that $\mathbf{1}_n \in P\left(\boldsymbol{a}^{(n)}(t), b^{(n)}(t)\right)$.

It is now sufficient to show that for any $\epsilon > 0$,

$$\|\boldsymbol{v}^{(n)}(t) - \mathbf{1}_n\|_\infty \frac{2^{n-1}}{n-1} > (1-\epsilon)\|\boldsymbol{a}^{(n)}(t)\|_\infty \tag{4.66}$$

holds for sufficiently large t. Upon recalling the form of $\boldsymbol{v}^{(n)}(t)$, we have

$$\begin{aligned}\|\boldsymbol{v}^{(n)}(t)-\mathbf{1}_n\|_\infty &= \frac{b^{(n)}(t)}{a_1^{(n)}(t)}-1=\frac{2^n+(n-1)t2^{n-1}-1}{2^{n-1}}-1\\ &= \frac{2^{n-1}\cdot 2+(n-1)t2^{n-1}-1}{2^{n-1}}-1=(n-1)t+1-\frac{1}{2^{n-1}}\,.\end{aligned}$$

Finally, observe that

$$\frac{\|\boldsymbol{v}^{(n)}(t)-\mathbf{1}_n\|_\infty}{\|\boldsymbol{a}^{(n)}(t)\|_\infty}=\frac{(n-1)t+1-2^{-(n-1)}}{2^{n-2}+t2^{n-1}}\longrightarrow\frac{n-1}{2^{n-1}}$$

as $t\to\infty$, which implies that (4.66) and (4.29) hold as required, concluding the proof of Theorem 4.1.9.

4.11 Proof of Theorem 4.1.10

We will apply the following result by Bombieri and Vaaler [21] and, in consequence, the proof of Theorem 4.1.10 follows in a rather straightforward manner. Recall that for a given matrix $A\in\mathbb{Z}^{m\times n}$, we denote by $\Gamma(A)=\Gamma(A,\mathbf{0})$ the lattice formed by all integer points in the kernel of the matrix A, i.e. $\Gamma(A)=\{\boldsymbol{x}\in\mathbb{R}^n : A\boldsymbol{x}=\mathbf{0}\}\cap\mathbb{Z}^n$.

Theorem 4.11.1 ([21, Theorem 2]). *Let $A\in\mathbb{Z}^{m\times n}$ with $m<n$ be a matrix of full rank m. There exist $n-m$ linearly independent integral vectors $\boldsymbol{y}_1,\dots,\boldsymbol{y}_{n-m}\in\Gamma(A)$ satisfying*

$$\prod_{i=1}^{n-m}\|\boldsymbol{y}_i\|_\infty\le\frac{\sqrt{\det(AA^T)}}{\gcd(A)}$$

This result intuitively tells us that that if the coefficients in the system $A\boldsymbol{y}=\mathbf{0}$ are small integers, then there will exist $n-m$ nontrivial linearly independent integral solutions to the homogeneous linear system $A\boldsymbol{y}=\mathbf{0}$ which each involve only small integers.

Let $\boldsymbol{z}^*$ be an integral vertex of the integer hull $P_I(A,\boldsymbol{b})$ that gives an optimal solution to the IP (4.30). We will show that $\boldsymbol{z}^*$ satisfies (4.32) and firstly we argue that it suffices to consider the case $\|\boldsymbol{z}^*\|_0=n$. Suppose that $\|\boldsymbol{z}^*\|_0<n$. For $\tau=\operatorname{supp}(\boldsymbol{z}^*)$, we set $\bar{A}=A_\tau$, $\bar{\boldsymbol{b}}=\boldsymbol{b}$, $\bar{\boldsymbol{c}}=\boldsymbol{c}_\tau$, and $\bar{\boldsymbol{z}}^*=\boldsymbol{z}^*_\tau$. By removing linearly dependent rows if necessary, we may assume that $\bar{A}$ has full row rank. Let $\bar{m}=\operatorname{rank}(\bar{A})\le m$. Observe that $\bar{\boldsymbol{z}}^*$ is an optimal solution for the corresponding problem (4.30) with minimal support. Furthermore, $\bar{\boldsymbol{z}}^*$ has full support. Now, if (4.32) holds true for $\bar{\boldsymbol{z}}^*$, then

$$(\rho(\boldsymbol{z}^*)+1)^{s-m}\le(\rho(\bar{\boldsymbol{z}}^*)+1)^{s-\bar{m}}\le\frac{\sqrt{\det\left(\bar{A}\bar{A}^T\right)}}{\gcd(\bar{A})}\,.\qquad(4.67)$$

Further, using [6, Lemma 2.3], we have

$$\frac{\sqrt{\det\left(\bar{A}\bar{A}^T\right)}}{\gcd(\bar{A})} \leq \frac{\sqrt{\det\left(AA^T\right)}}{\gcd(A)}, \tag{4.68}$$

where combining (4.67) and (4.68) yields (4.32) as required, which completes the proof for the scenario when $\|\boldsymbol{z}^*\|_0 < n$.

From now on, we assume that $\|\boldsymbol{z}^*\|_0 = n$. Suppose, to derive a contradiction, that (4.32) does not hold, that is

$$(\rho(\boldsymbol{z}^*)+1)^{n-m} > \frac{\sqrt{\det\left(AA^T\right)}}{\gcd(A)}$$

By Theorem 4.11.1, there exists a vector $\boldsymbol{y} \in \mathbb{Z}^n \setminus \{\boldsymbol{0}\}$ such that

$$A\boldsymbol{y} = \boldsymbol{0} \text{ and } \|\boldsymbol{y}\|_\infty \leq \left(\frac{\sqrt{\det\left(AA^T\right)}}{\gcd(A)}\right)^{\frac{1}{n-m}} < \rho(\boldsymbol{z}^*)+1.$$

It follows that both $\boldsymbol{z}^*+\boldsymbol{y}$ and $\boldsymbol{z}^*-\boldsymbol{y}$ are integral points in the polyhedron $P(A,\boldsymbol{b})$ and, in consequence, $\boldsymbol{z}^*$ is not a vertex of the integral hull $P_I(A,\boldsymbol{b})$. The obtained contradiction completes the proof of Theorem 4.1.10.

— Chapter 5 —

Refinements for Special Cases

In this chapter we collect several results that in special cases provide refinements of the known distance and sparsity bounds under additional assumptions which motivate future research.

5.1 An Optimal Bound for the $A \in \mathbb{Z}^{m\times(m+1)}$ Scenario

5.1.1 Introduction

We assume that a matrix $A \in \mathbb{Z}^{m\times n}$ has full rank m and let $\boldsymbol{b} \in \mathbb{Z}^m$ be an m-dimensional integer vector. Consider the polyhedron

$$P = P(A, \boldsymbol{b}) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : A\boldsymbol{x} = \boldsymbol{b}\right\}$$

and, assuming P is nonempty, take any vertex $\boldsymbol{x}^*$ of the polyhedron P. Because the matrix A has full rank m by assumption, it again follows that there exists a basis γ of A such that

$$\boldsymbol{x}^*_\gamma = A_\gamma^{-1}\boldsymbol{b} \text{ and } \boldsymbol{x}^*_{\bar{\gamma}} = 0\,. \tag{5.1}$$

Recall that $\Lambda = \Lambda(A, \boldsymbol{b})$ denotes the affine lattice in $\mathbb{R}^n$ formed by taking integer points in the (affine) flat that is described by the linear system $A\boldsymbol{x} = \boldsymbol{b}$, namely

$$\Lambda = \Lambda(A, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{Z}^n : A\boldsymbol{x} = \boldsymbol{b}\}\,.$$

In order to simplify notation slightly, throughout this section we let $\pi(\cdot) : \mathbb{R}^{m+1} \to \mathbb{R}$ denote the projection onto the final coordinate, i.e. the projection that forgets about the first m coordinates. Furthermore, recall that $\Delta(A)$ and $\gcd(A)$ denote the greatest absolute valued m-dimensional subdeterminant and the greatest common divisor of all m-dimensional subdeterminants of A, respectively.

The first result of this chapter provides an optimal upper bound in a special case for the ℓ_∞-distance from any vertex $\boldsymbol{x}^*$ of the polyhedron P which is given by a basis γ as in (5.1) to the set of its lattice points $\boldsymbol{z} \in P \cap \mathbb{Z}^n$. It should be emphasised that by optimal we mean that one can find an example, namely Example 5.1.2, where the equality appearing in (5.4) is attained. Further, it is worth noting that $\boldsymbol{z}$ is used in order to denote that the lattice point $\boldsymbol{z}$ need not be an optimal integral solution to a given IP nor a vertex of a corner polyhedron.

For this purpose, we recall that the *(maximum) vertex distance* $d(A, \boldsymbol{b})$ is

$$d(A, \boldsymbol{b}) = \begin{cases} \max_{\boldsymbol{x}^*} \min_{\boldsymbol{z} \in P \cap \mathbb{Z}^n} \|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty\,, & \text{if } P \cap \mathbb{Z}^n \neq \emptyset, \\ & -\infty\,, \text{ otherwise,} \end{cases}$$

where the maximum is taken over all vertices $\boldsymbol{x}^*$ of the polyhedron P. The (maximum) vertex distance is the largest ℓ_∞-distance from any vertex $\boldsymbol{x}^*$ of P to a nearby feasible integer point $\boldsymbol{z} \in P \cap \mathbb{Z}^n$.

Recall (from Chapter 3) that provided $P \cap \Lambda \neq \emptyset$, then the results of Eisenbrand and Weismantel [41] show that from any vertex $\boldsymbol{x}^*$ of P there exists a lattice point $\boldsymbol{z} \in P \cap \Lambda$ satisfying

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 \leq m\,(2m\|A\|_\infty + 1)^m. \tag{5.2}$$

This upper bound has recently been strengthened by Lee et al. [72, 73]. Further, recall (from Chapter 3) that in the *knapsack scenario*, i.e. when $m = 1$, Aliev et al. [8] show that the (maximum) vertex distance $d(\boldsymbol{a}, b)$ is optimally bounded by

$$d(\boldsymbol{a}, b) \leq \|\boldsymbol{a}\|_\infty - 1\,, \tag{5.3}$$

where we follow traditional vector notation, where we replacing A and $\boldsymbol{b}$ by $\boldsymbol{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$, respectively.

In particular, the following theorem provides an optimal upper bound for the (maximum) vertex distance for the $m \times (m+1)$ dimensional scenario. It is worth noting that the upper bound depends only on subdeterminants of the matrix A. Furthermore, we provide an example for the case $m = 2$ that demonstrates the optimality of this upper bound. It is worth emphasising that optimal here means that we can construct an example where the equality appearing in (5.4) is attained.

Theorem 5.1.1. *Let $A \in \mathbb{Z}^{m\times(m+1)}$ with full rank m and $\boldsymbol{b} \in \mathbb{Z}^m$. If $P \cap \mathbb{Z}^{m+1} \neq \emptyset$, then*

$$d(A, \boldsymbol{b}) \leq \frac{\Delta(A)}{\gcd(A)} - 1\,. \tag{5.4}$$

The following example demonstrates the optimality of Theorem 5.1.1 in the case when $m = 2$ and, additionally, one can visualise this example using Figure 5.1.

Example 5.1.2. *Consider*

$$A = \begin{pmatrix} 3 & 11 & 7 \\ -5 & 7 & 3 \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 154 \\ 58 \end{pmatrix}.$$

In this case, the 2*-dimensional subdeterminants of A are*

$$\det\begin{pmatrix} 3 & 11 \\ -5 & 7 \end{pmatrix} = 76\,, \det\begin{pmatrix} 3 & 7 \\ -5 & 3 \end{pmatrix} = 44 \text{ and } \det\begin{pmatrix} 11 & 7 \\ 7 & 3 \end{pmatrix} = -16\,,$$

respectively. In particular, note that $\Delta(A) = 76$ and $\gcd(A) = 4$. In this case the polyhedron P is precisely a line segment (Figure 5.1) connecting $\boldsymbol{x}_1^* = (110/19, 236/19, 0)^T$ and $\boldsymbol{x}_2^* = (14/11, 0, 236/11)^T$ and, in addition, the only feasible integer point is $\boldsymbol{z} = (2, 2, 18)^T$. We consider the maximum vertex distance $d(A, \boldsymbol{b})$ from the two vertices to the feasible integral point $\boldsymbol{z}$. The corresponding ℓ_∞-norm distances from the vertices $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ to $\boldsymbol{z}$ are

$$\|\boldsymbol{x}_1^* - \boldsymbol{z}\|_\infty = \max\left(\left|\frac{110}{19} - 2\right|, \left|\frac{236}{19} - 2\right|, 18\right) = 18$$

and

$$\|\boldsymbol{x}_2^* - \boldsymbol{z}\|_\infty = \max\left(\left|\frac{14}{11} - 2\right|, 2, \left|\frac{236}{11} - 18\right|\right) = \frac{38}{11}\,,$$

respectively. The maximum vertex distance is consequently $d(A, \boldsymbol{b}) = 18$ and noting that

$$\frac{\Delta(A)}{\gcd(A)} - 1 = \frac{76}{4} - 1 = 19 - 1 = 18$$

holds demonstrates that the inequality (5.4) from Theorem 5.1.1 is sharp.

In light of the optimal upper bound (5.3) of Aliev et al. [8] and since Theorem 5.1.1 provides an optimal worst-case upper bound on the distance from any vertex $\boldsymbol{x}^*$ of the polyhedron P to a nearby feasible integer point $\boldsymbol{z}$ with respect to the ℓ_∞-norm that is independent of both n and m in the $m \times (m+1)$-dimensional scenario, we propose for the general setting the following.

Conjecture 1. *Let $A \in \mathbb{Z}^{m\times n}$ be a matrix of full rank m and $\boldsymbol{b} \in \mathbb{Z}^m$. If $P \cap \mathbb{Z}^n \neq \emptyset$, then*

$$d(A, \boldsymbol{b}) \le \frac{\Delta(A)}{\gcd(A)} - 1\,.$$

5.1.2 Proof of Theorem 5.1.1

Reordering the columns of the matrix A if necessary, we may assume without loss of generality that $\gamma = \{1, 2, \ldots, m\}$, i.e. that A is written in the form $A = (A_\gamma, A_{\bar\gamma})$, where

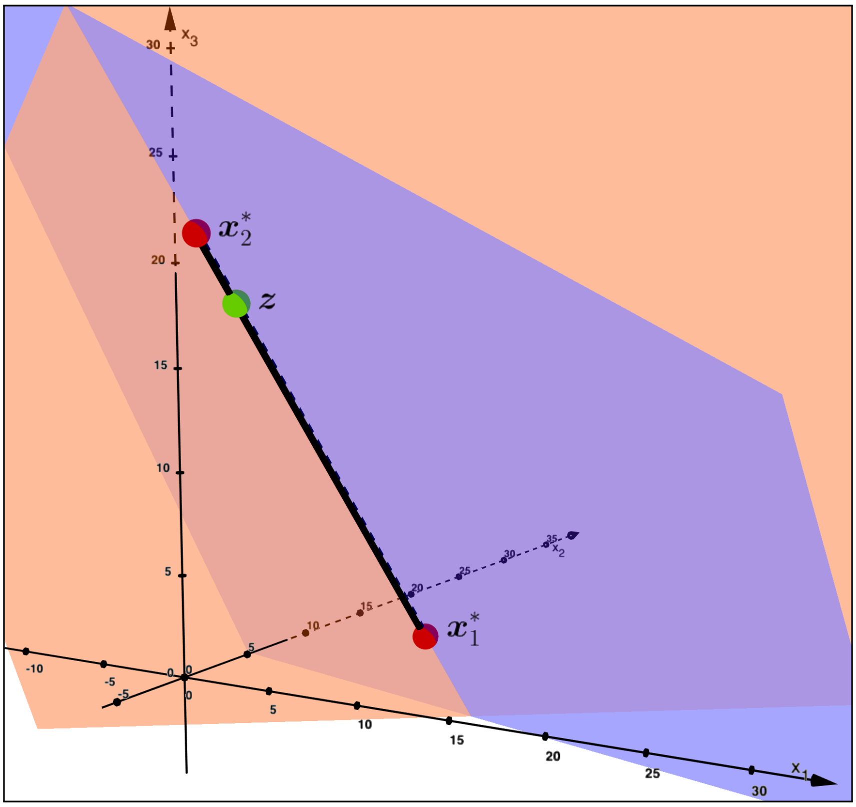


Figure 5.1: This figure provides a visualisation of Example 5.1.2. The blue and orange planes are $3x_1 + 11x_2 + 7x_3 = 154$ and $-5x_1 + 7x_2 + 3x_3 = 58$, respectively. The polyhedron is precisely the intersection of the two hyperplanes restricted to the positive orthant. It is precisely the black line segment connecting the two red vertices $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$, which additionally contains the green unique feasible integer point $\boldsymbol{z}$.

$\det(A_\gamma) \neq 0$. In this case $A_{\bar{\gamma}}$ is an m-dimensional column vector. The polyhedron P can be consequently written as

$$P = \left\{ \boldsymbol{x} \in \mathbb{R}^{m+1}_{\geq 0} : A_\gamma \boldsymbol{x}_\gamma + A_{\bar{\gamma}} x_{\bar{\gamma}} = \boldsymbol{b} \right\},$$

where $\boldsymbol{x} = (\boldsymbol{x}_\gamma, x_{\bar{\gamma}})^T$, i.e. $\boldsymbol{x}_\gamma \in \mathbb{R}^m$ and $x_{\bar{\gamma}} \in \mathbb{R}$ contain the entries of the vector $\boldsymbol{x}$ corresponding to A_γ and $A_{\bar{\gamma}}$, respectively. In particular, it is worth emphasising that since $A = (A_\gamma, A_{\bar{\gamma}})$, then $x_{\bar{\gamma}} = x_{m+1}$. The condition (5.1) on $\boldsymbol{x}^*$ can be equivalently expressed as

$$\boldsymbol{x}_\gamma^* = A_\gamma^{-1} \boldsymbol{b} \text{ and } x_{\bar{\gamma}}^* = x_{m+1}^* = 0.$$

Furthermore, using Cramer's Rule (Lemma 2.7.1), the vertex $\boldsymbol{x}^*$ has the form

$$\boldsymbol{x}^* = \left(\frac{\det\left(A_\gamma^1(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, \frac{\det\left(A_\gamma^2(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, \ldots, \frac{\det\left(A_\gamma^m(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, 0 \right)^T, \tag{5.5}$$

where $A^i_\gamma(\boldsymbol{b})$ denotes the submatrix A_γ whose i-th column has been replaced by the integral column vector $\boldsymbol{b}$. Observe that when the polyhedron P is bounded, then P is a line segment in $\mathbb{R}^{m+1}$ connecting its two vertices. If instead the polyhedron P is unbounded, then P is a ray in $\mathbb{R}^{m+1}$ where $\boldsymbol{x}^*$ is the only vertex of P.

Recall that we are upper bounding the distance with respect to the ℓ_∞-norm from the vertex $\boldsymbol{x}^*$ to some nearby feasible integral point. We denote such an integral point by $\boldsymbol{z} = (z_1, \ldots, z_{m+1})^T \in P \cap \mathbb{Z}^{m+1}$. We assume that $\boldsymbol{z}$ is the feasible integral point with minimal final coordinate. It is worth emphasising that this assumption means that $\boldsymbol{z}$ is the closest integral point from $\boldsymbol{x}^*$ in the x_{m+1}-th coordinate direction with respect to the ℓ_∞-norm.

Note that the form of $\boldsymbol{x}^*$ (5.5) implies that its projection is $\pi(\boldsymbol{x}^*) = 0$. In particular, we now consider $\|\pi(\boldsymbol{x}^*) - \pi(\boldsymbol{z})\|_\infty = z_{m+1}$. Recall that Λ denotes the affine lattice in $\mathbb{R}^{m+1}$ containing all integer points in the (affine) flat described by the linear system $A\boldsymbol{x} = \boldsymbol{b}$. In particular, its projection $\pi(\Lambda)$ is a one-dimensional affine lattice which, in light of Lemma 4.2.4 (from Chapter 4), has determinant $\det(\Lambda) = |\det(A_\gamma)|/\gcd(A)$. It follows that all projected affine lattice points from $\pi(\Lambda)$ belong to the same congruence class. Further, upon noting that the least residue in this congruence class is one of the integers $\{0, 1, \ldots, \det(\Lambda) - 1\}$, it follows since $\boldsymbol{z}$ is the closest integral point from $\boldsymbol{x}^*$ in the x_{m+1}-th coordinate direction with respect to the ℓ_∞-norm by assumption, that $\pi(\boldsymbol{z})$ satisfies

$$\pi(\boldsymbol{z}) = z_{m+1} \leq \det(\Lambda) - 1 = |\det(A_\gamma)|/\gcd(A) - 1\,.$$

Observe that if one fixes the value of z_{m+1} then the corresponding m-dimensional integer solution $\boldsymbol{z}_\gamma$ is uniquely determined by $\boldsymbol{z}_\gamma = A_\gamma^{-1}\left(\boldsymbol{b} - A_{\bar{\gamma}} z_{m+1}\right)$ because the matrix A_γ is nonsingular by assumption. In order to simplify the subsequent notation, we let $\tilde{\boldsymbol{b}} = \boldsymbol{b} - A_{\bar{\gamma}} z_{m+1}$. In particular, using Cramer's rule (Lemma 2.7.1), the m-dimensional integer solution $\boldsymbol{z}_B$ is given by

$$\boldsymbol{z}_B = (z_1, z_2, \ldots, z_m)^T = \left(\frac{\det\left(A^1_\gamma(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, \frac{\det\left(A^2_\gamma(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, \ldots, \frac{\det\left(A^m_\gamma(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)} \right)^T. \tag{5.6}$$

Further, the m-dimensional solution (5.6) to the linear system $B\boldsymbol{x}_B = \tilde{\boldsymbol{b}}$ can be "lifted" to yield a solution in the $(m+1)$-dimensional space by simply appending the fixed z_{m+1} to the $(m+1)$-th entry of the solution. In other words, we can write

$$\boldsymbol{z} = (z_1, z_2, \ldots, z_m, z_{m+1})^T = \left(\frac{\det\left(A^1_\gamma(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, \frac{\det\left(A^2_\gamma(\tilde{\boldsymbol{b}})\right)}{\det(A_\gamma)}, \ldots, \frac{\det\left(A^m_\gamma(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, z_{m+1} \right)^T.$$

Recalling the assumed form (5.5) of the vertex $\boldsymbol{x}^*$, the vertex distance $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty$ is precisely given by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \Bigg\| &\left(\frac{\det\left(A_\gamma^1(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, \frac{\det\left(A_\gamma^2(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, \ldots, \frac{\det\left(A_\gamma^m(\boldsymbol{b})\right)}{\det\left(A_\gamma\right)}, 0 \right)^T \\ &- \left(\frac{\det(A_\gamma^1(\tilde{\boldsymbol{b}}))}{\det\left(A_\gamma\right)}, \frac{\det\left(A_\gamma^2(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, \ldots, \frac{\det\left(A_\gamma^m(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)}, z_{m+1} \right)^T \Bigg\|_\infty .\end{aligned}$$

This can be expressed equivalently using the definition of the ℓ_∞-norm as

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \max\Bigg(&\left| \frac{\det\left(A_\gamma^1(\boldsymbol{b})\right) - \det\left(A_\gamma^1(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)} \right|, \left| \frac{\det\left(A_\gamma^2(\boldsymbol{b})\right) - \det\left(A_\gamma^2(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)} \right|, \\ &\ldots, \left| \frac{\det\left(A_\gamma^m(\boldsymbol{b})\right) - \det\left(A_\gamma^m(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)} \right|, z_{m+1} \Bigg).\end{aligned} \tag{5.7}$$

Recall we have shown that $z_{m+1} \leq |\det\left(A_\gamma\right)|/\gcd(A) - 1$ holds and, furthermore, notice that $z_{m+1} \leq \Delta(A)/\gcd(A) - 1$ holds. In particular, this observation implies that (5.4) holds for the final difference appearing in (5.7), namely $|x_{m+1}^* - z_{m+1}| = z_{m+1} \leq \Delta(A)/\gcd(A) - 1$.

It remains sufficient to consider the first m differences appearing in (5.7). Take the j-th difference for $j \in \{1, \ldots, m\}$ which appears in (5.7), namely

$$\left| \frac{\det\left(A_\gamma^j(\boldsymbol{b})\right) - \det\left(A_\gamma^j(\tilde{\boldsymbol{b}})\right)}{\det\left(A_\gamma\right)} \right| . \tag{5.8}$$

In order to enable us to work with the determinants appearing in the numerator in (5.8), we make use of Laplace's expansion formula (Theorem 2.6.1). For this purpose, denote by $M_{i,j}$ and $\tilde{M}_{i,j}$ the $(m-1)$-dimensional square submatrices formed by deleting the i-th row and j-th column of matrices $A_\gamma^j(\boldsymbol{b})$ and $A_\gamma^j(\tilde{\boldsymbol{b}})$, respectively. Upon expanding both determinants from the numerator of (5.8) along the j-th column, we may equivalently express (5.8) as

$$\begin{aligned}&\left| \frac{\sum_{i=1}^m \left((-1)^{i+j} \cdot b_i \cdot \det\left(M_{i,j}\right)\right) - \sum_{i=1}^m \left((-1)^{i+j} \cdot \tilde{b}_i \cdot \det\left(\tilde{M}_{i,j}\right)\right)}{\det\left(A_\gamma\right)} \right| \\ &\qquad = \left| \frac{\sum_{i=1}^m \left((-1)^{i+j} \cdot (b_i - \tilde{b}_i) \cdot \det\left(M_{i,j}\right)\right)}{\det\left(A_\gamma\right)} \right| ,\end{aligned} \tag{5.9}$$

where the equality follows since $M_{i,j} = \tilde{M}_{i,j}$ holds because $A^j_\gamma(\boldsymbol{b})$ and $A^j_\gamma(\tilde{\boldsymbol{b}})$ vary only in their j-th column, which is the column that has been forgotten about by the expansion leading to the submatrices $M_{i,j}$ and $\tilde{M}_{i,j}$.

Recall that $\tilde{\boldsymbol{b}} = \boldsymbol{b} - A_{\bar{\gamma}} z_{m+1}$ for fixed z_{m+1} and hence (5.9) can be equivalently expressed after simple algebraic manipulation as

$$\begin{aligned}
&\left| \frac{\sum_{i=1}^{m} \left((-1)^{i+j} \cdot (b_i - (b_i - a_{i(m+1)} \cdot z_{m+1})) \cdot \det\left(M_{i,j}\right) \right)}{\det\left(A_\gamma\right)} \right| \\
&= \left| \frac{z_{m+1} \cdot \sum_{i=1}^{m} \left((-1)^{i+j} \cdot a_{i(m+1)} \cdot \det\left(M_{i,j}\right) \right)}{\det\left(A_\gamma\right)} \right| = \left| \frac{z_{m+1} \cdot \left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right|}{\left|\det\left(A_\gamma\right)\right|} \right|,
\end{aligned} \tag{5.10}$$

where the final equality follows by making use of Laplace's expansion formula (Theorem 2.6.1) in order to write the summations equivalently as the determinant of a square m-dimensional matrix. We could alternatively deduce the final expression appearing in (5.10) directly from (5.8) using several fundamental properties of determinants.

Recall that $z_{m+1} \le \left|\det\left(A_\gamma\right)\right| / \gcd(A) - 1$ holds and hence (5.10) is further bounded from above by

$$\begin{aligned}
\left| \frac{z_{m+1} \cdot \left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right|}{\left|\det\left(A_\gamma\right)\right|} \right| &\le \left| \frac{\left(\left|\det\left(A_\gamma\right)\right| / \gcd(A) - 1 \right) \cdot \left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right|}{\left|\det\left(A_\gamma\right)\right|} \right| \\
&= \left| \left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right| \cdot \left(\frac{1}{\gcd(A)} - \frac{1}{\left|\det(A_\gamma)\right|} \right) \right|.
\end{aligned} \tag{5.11}$$

Notice that the matrix $A^j_\gamma(A_{\bar{\gamma}})$ contains only columns from A and, in consequence, $\left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right| \le \Delta(A)$ holds. Finally, since $\gcd(A) \le |\det\left(A_\gamma\right)| \le \Delta(A)$ holds it follows that (5.11) is bounded by

$$\left| \left|\det\left(A^j_\gamma(A_{\bar{\gamma}})\right)\right| \cdot \left(\frac{1}{\gcd(A)} - \frac{1}{\left|\det(A_\gamma)\right|} \right) \right| \le \left| \Delta(A) \cdot \left(\frac{1}{\gcd(A)} - \frac{1}{\Delta(A)} \right) \right| = \frac{\Delta(A)}{\gcd(A)} - 1\,,$$

which implies that (5.4) holds as required and concludes the proof of Theorem 5.1.1.

5.2 A Refined ℓ_1 Proximity Bound for Knapsacks

In this section, provided that a knapsack polyhedron is integer feasible, we provide a refined upper bound for the distance from any vertex of a knapsack polyhedron to its nearest feasible lattice point with respect to the ℓ_1-norm. Furthermore, we provide several extremal examples which demonstrate the tightness of this bound. Further, this upper bound provides insight into the sharpness of the seminal proximity bound (5.2) provided by Eisenbrand and Weismantel [41].

5.2.1 Introduction

Given $\boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$, then $P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : \boldsymbol{a}^T\boldsymbol{x} = b\right\}$ is a *knapsack polyhedron*. In the scenario when $P(\boldsymbol{a}, b)$ is bounded, we will once more refer to $P(\boldsymbol{a}, b)$ as a *knapsack polytope*.

In what follows, we exclude the trivial case $n = 1$, where the ℓ_1-norm only takes value 0 or is undefined. Furthermore, we assume without loss of generality that $\boldsymbol{a}$ satisfies the following conditions introduced in Chapter 4:

$$\begin{aligned} &(i)\ \boldsymbol{a} = (a_1, \ldots, a_n)^T \in \mathbb{Z}^n,\ n \geq 2,\ a_i \neq 0,\ i = 1, \ldots, n \text{ and} \\ &(ii)\ \gcd(\boldsymbol{a}) := \gcd(a_1, \ldots, a_n) = 1\,. \end{aligned} \tag{5.12}$$

In addition, we assume that the integer b belongs to the semigroup

$$Sg(\boldsymbol{a}) = \left\{\boldsymbol{a}^\top \boldsymbol{z} : \boldsymbol{z} \in \mathbb{Z}^n_{\geq 0}\right\}$$

that is generated by the entries of the vector $\boldsymbol{a}$, which ensures $P(\boldsymbol{a}, b) \cap \mathbb{Z}^n \neq \emptyset$.

The following result provides the aforementioned refined upper bound for the distance from a vertex of a knapsack polyhedron to the nearest feasible lattice point with respect to the ℓ_1-norm, which depends only on the greatest absolute valued entry of the vector $\boldsymbol{a}$ and is independent of both the dimension n and the integer b. In the following $\|\cdot\|_1$ and $\|\cdot\|_\infty$ will be used to denote the ℓ_1-norm and ℓ_∞-norms, respectively.

Theorem 5.2.1. *Let $\boldsymbol{a} \in \mathbb{Z}^n$ satisfy (5.12) and $b \in Sg(\boldsymbol{a})$. Suppose that at least one of the conditions $P(\pi_n(\boldsymbol{a}), b)$ is unbounded, $\gcd(a_1, \ldots, a_{n-1}) = 1$, $\pi_n(\boldsymbol{x}^*) = \boldsymbol{0}$ and (5.20) does not hold. Then from any suitable vertex $\boldsymbol{x}^*$ of the knapsack polyhedron $P(\boldsymbol{a}, b)$ there exists an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ satisfying*

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2 \ \textit{if}\ \|\boldsymbol{a}\|_\infty \neq 1,$$

while, $\boldsymbol{x}^* = \boldsymbol{z}$ holds, otherwise.

It remains an open problem to prove (or disprove) that the inequality appearing in Theorem 5.2.1 holds in the scenario that the aforementioned conditions all hold. In the knapsack scenario the seminal proximity bound (5.2) of Eisenbrand and Weismantel [41] yields

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 \leq 2\,\|\boldsymbol{a}\|_\infty + 1. \tag{5.13}$$

In order to refine the upper bound (5.13), we combine the recent result (5.3) of Aliev et al. [8] that bounds the maximum vertex distance with properties of a geometric object called the *sail* associated with an underlying lattice, that was previously defined in Chapter 4. Let $\boldsymbol{p} \in \mathbb{Z}^d$ and consider the affine lattice $\Gamma = \boldsymbol{p} + \Lambda$ associated with the lattice $\Lambda \subset \mathbb{Z}^d$. Recall the set $\mathscr{E}(\Gamma) = \operatorname{conv}(\Gamma \cap \mathbb{R}^d_{\geq 0})$ the sail associated with Γ, as illustrated in Figure 4.2 (in Chapter 4).

5.2.2 Proof of Theorem 5.2.1

Let $\pi_n(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n-1}$ denote the projection map which forgets about the n-th coordinate. For technical reasons, it is convenient as in the paper of Aliev et al. [8] to consider firstly two special cases, namely Lemma 5.2.2 and Lemma 5.2.4, which prove Theorem 5.2.1 for positive $\boldsymbol{a}$, corresponding to the case where $P(\boldsymbol{a}, b)$ is bounded, and for tuples $(\boldsymbol{a}, b)$ that correspond to bounded or empty polyhedra $P(\pi_n(\boldsymbol{a}), b)$, respectively. It will then remain to consider the scenario where the polyhedra $P(\pi_n(\boldsymbol{a}), b)$ is unbounded, namely Lemma 5.2.5.

During the proof of Theorem 5.2.1 we one more will make use of the following notation. Let $\Lambda(\boldsymbol{a}, b)$ denote the affine lattice formed by integer points in the affine hyperplane $\boldsymbol{a}^T\boldsymbol{x} = b$, that is

$$\Lambda(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{a}^T\boldsymbol{x} = b\right\}.$$

For convenience, we set $Q(\boldsymbol{a}, b) = \pi_n(P(\boldsymbol{a}, b))$ and $L(\boldsymbol{a}, b) = \pi_n(\Lambda(\boldsymbol{a}, b))$. Observe that the projected affine lattice $L(\boldsymbol{a}, b)$ is an affine $(n-1)$-dimensional lattice which can be written in the form

$$L(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{Z}^{n-1} : a_1x_1 + a_2x_2 + \cdots + a_{n-1}x_{n-1} \equiv b \bmod |a_n|\right\}. \tag{5.14}$$

Further, note that $L(\boldsymbol{a}, 0)$ is a lattice of determinant $\det(L(\boldsymbol{a}, 0)) = |a_n|$ and $L(\boldsymbol{a}, b) = L(\boldsymbol{a}, 0) + \boldsymbol{y}$ is an affine $(n-1)$-dimensional integral lattice for some $\boldsymbol{y} \in \mathbb{Z}^{n-1}$.

As discussed above, we start with two special cases. Firstly, suppose that all entries of $\boldsymbol{a}$ are positive, i.e. that $\boldsymbol{a} \in \mathbb{Z}_{>0}$. In such case, the knapsack polyhedron $P(\boldsymbol{a}, b)$ is bounded and is called the knapsack polytope (as in Chapter 4).

Lemma 5.2.2. *Let $\boldsymbol{a} \in \mathbb{Z}^n_{>0}$ satisfy (5.12) and $b \in Sg(\boldsymbol{a})$. Then from any vertex $\boldsymbol{x}^*$ of the knapsack polytope $P(\boldsymbol{a}, b)$ there exists an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ satisfying*

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2 \text{ if } \|\boldsymbol{a}\|_\infty \neq 1, \tag{5.15}$$

while, $\boldsymbol{x}^* = \boldsymbol{z}$ holds, otherwise.

Proof. Because $b \in Sg(\boldsymbol{a})$ by assumption then $b \geq 0$. Further, if $b = 0$, then $P(\boldsymbol{a}, b) = \{\mathbf{0}\}$, where $\mathbf{0}$ denotes the zero vector. It is sufficient to assume that b is a positive integer and then the polytope $P(\boldsymbol{a}, b)$ is a simplex with vertices

$$\left(\frac{b}{a_1}, 0, \ldots, 0\right)^T, \left(0, \frac{b}{a_2}, 0 \ldots, 0\right)^T, \ldots, \left(0, \ldots, 0, \frac{b}{a_n}\right)^T.$$

Let $\boldsymbol{x}^*$ be any vertex of the knapsack polytope $P(\boldsymbol{a}, b)$. Upon reordering the entries of the vector $\boldsymbol{a}$ if necessary, we may assume without loss of generality that $\boldsymbol{x}^*$ has the form $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$.

It should be observed that if $\|\boldsymbol{a}\|_\infty = 1$, then the conditions (5.12) and the positivity of $\boldsymbol{a}$ implies that $\boldsymbol{a} = (1, 1, \ldots, 1)^T$. This is precisely the totally unimodular case, where each vertex of $P(\boldsymbol{a}, b)$ is integral since $b \in \mathbb{Z}_{>0}$. In this case, the distance with respect to the ℓ_1-norm from any vertex to a nearby integral point is zero as required. In light of this it is sufficient throughout the rest of the proof to assume that the integral vector $\boldsymbol{a}$ is not unimodular, i.e. that $\|\boldsymbol{a}\|_\infty \neq 1$.

It was shown by Aliev et al. [8] through making use of covering arguments and the projected affine lattice $L(\boldsymbol{a}, b)$ (5.14) that a feasible lattice point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ exists which satisfies $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$. In particular, for technical reasons we assume that $\boldsymbol{z}$ is an integral vertex of the corner polyhedron $\mathscr{C}_n(\boldsymbol{a}, b)$ associated with $\boldsymbol{x}^* = (b/a_n)\boldsymbol{e}_n$ whose existence is guaranteed in light of Theorem 4.1.6 (in Chapter 4) and is an integral point satisfying $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$. Recall from Chapter 4 that $\mathscr{C}_n(\boldsymbol{a}, b)$ is defined as

$$\mathscr{C}_n(\boldsymbol{a}, b) = \operatorname{conv}\bigl(\{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{a}^T\boldsymbol{x} = b, x_1 \geq 0, \ldots, x_{n-1} \geq 0\}\bigr).$$

Note that the definition of the projected affine lattice $L(\boldsymbol{a}, b)$ (5.14) implies that the lattice point $\boldsymbol{z}$ has the form

$$\boldsymbol{z} = (z_1, z_2, \ldots, z_n)^T = \left(z_1, z_2, \ldots, z_{n-1}, \frac{b}{a_n} - \frac{a_1 z_1 + a_2 z_2 + \ldots + a_{n-1} z_{n-1}}{a_n}\right)^T.$$

Recalling that the vertex $\boldsymbol{x}^*$ has by assumption the form $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$, we notice that the distance between $\boldsymbol{x}^*$ and $\boldsymbol{z}$ with respect to the ℓ_∞-norm is

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \left\| \left(0, \ldots, 0, \frac{b}{a_n}\right)^T - \left(z_1, z_2, \ldots, z_{n-1}, \frac{b}{a_n} - \frac{a_1 z_1 + a_2 z_2 + \ldots + a_{n-1} z_{n-1}}{a_n}\right)^T \right\|_\infty .$$

This can be expressed equivalently using the definition of the ℓ_∞-norm as

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \max\left(z_1, z_2, \ldots, z_{n-1}, \frac{a_1 z_1 + a_2 z_2 + \ldots + a_{n-1} z_{n-1}}{a_n}\right). \tag{5.16}$$

Recall that $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \le \|\boldsymbol{a}\|_\infty - 1$ holds. In particular, notice that the n-th entry appearing in (5.16) satisfies $|x_n^* - z_n| = \|\boldsymbol{a}\|_\infty - 1$ only if $x_n^* \in \mathbb{Z}$. Furthermore, this equality holds only if

$$0 = |x_n^* - z_n| = \|\boldsymbol{a}\|_\infty - 1 = 0 ,$$

which implies that equality holds only if $\boldsymbol{a}$ is unimodular, i.e. if $\|\boldsymbol{a}\|_\infty = 1$. Recall that we have assumed that $\|\boldsymbol{a}\|_\infty \neq 1$ and, in consequence, the upper bound for $|x_n^* - z_n|$ can be strengthened to the strict inequality $|x_n^* - z_n| < \|\boldsymbol{a}\|_\infty - 1$. This strict upper bound will now be utilised in our consideration of $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1$.

In light of the assumptions on the lattice point $\boldsymbol{z}$, the form of the vertex $\boldsymbol{x}^*$ and by the definition of the ℓ_1-norm, we obtain

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 = z_1 + z_2 + \ldots + z_{n-1} + \left|\frac{b}{a_n} - z_n\right| < z_1 + z_2 + \ldots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1.$$

Let $\boldsymbol{w} = \pi_n(\boldsymbol{z}) = (z_1, z_2, \ldots, z_{n-1})^T$. This notation allows us to equivalently express the inequality above as

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < z_1 + z_2 + \ldots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1 = \|\boldsymbol{w}\|_1 + \|\boldsymbol{a}\|_\infty - 1. \tag{5.17}$$

In light of this, it is sufficient to upper bound the ℓ_1-norm of $\boldsymbol{w}$ in order to upper bound the distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_1-norm as required.

We may assume without loss of generality that precisely k of the components from $\boldsymbol{w}$ are nonzero, where $0 \le k \le n-1$. Further, upon rearranging and relabelling the $n-1$ entries of $\boldsymbol{w}$ if necessary, we can assume that only the first k entries of $\boldsymbol{w}$ are nonzero. In order to complete the proof we now consider three cases, namely when $k = 0$, $k = 1$ and $k \ge 2$, respectively.

If $k = 0$, then $\boldsymbol{w}$ has no nonzero components and hence $\|\boldsymbol{w}\|_1 = 0 < \|\boldsymbol{a}\|_\infty - 1$ holds because $\|\boldsymbol{a}\|_\infty \neq 1$ by assumption. In particular, simply inspecting (5.17) implies that (5.15) holds as required.

If $k=1$, then (5.17) becomes $\|\boldsymbol{x}^*-\boldsymbol{z}\|_1 < z_1+\|\boldsymbol{a}\|_\infty - 1$. Since the integral point $\boldsymbol{z}$ was chosen such that $\|\boldsymbol{x}^*-\boldsymbol{z}\|_\infty \le \|\boldsymbol{a}\|_\infty - 1$ holds, it follows that (5.15) similarly holds as required.

Suppose now that $k \ge 2$. In this case, $\boldsymbol{w}$ has at least two nonzero components and upon making use of Lemma 4.2.3 (from Chapter 4) we obtain

$$\|\boldsymbol{w}\|_1 = z_1+z_2+\cdots+z_k \le \frac{k(z_1+1)(z_2+1)\cdots(z_k+1)}{2^k}. \tag{5.18}$$

Recall that $\boldsymbol{z}$ is a feasible integral vertex of the corner polyhedron $\mathscr{C}_n(\boldsymbol{a},b)$ associated with $\boldsymbol{x}^*=(b/a_n)\boldsymbol{e}_n$ which satisfies $\|\boldsymbol{x}^*-\boldsymbol{z}\|_\infty \le \|\boldsymbol{a}\|_\infty - 1$. In light of this and since the projection $\pi_n(\cdot)$ is bijective, we deduce that the projected integral point $\boldsymbol{w}$ is a vertex of the sail $\mathscr{E}(L(\boldsymbol{a},b))$ associated the projected affine lattice $L(\boldsymbol{a},b)$ and, in consequence, $\boldsymbol{w}$ is therefore irreducible with respect to the projected lattice $L(\boldsymbol{a},0)$ by Lemma 4.2.2 (from Chapter 4). Recall that the sail $\mathscr{E}(L(\boldsymbol{a},b))$ of the projected affine lattice $L(\boldsymbol{a},b)$ is $\mathscr{E}(L(\boldsymbol{a},b)) = \operatorname{conv}(L(\boldsymbol{a},b)\cap\mathbb{R}^{n-1}_{\ge 0})$.

Because the projected lattice point $\boldsymbol{w}$ is irreducible with respect to $L(\boldsymbol{a},0)$, we can apply Lemma 4.2.1 (from Chapter 4), which implies

$$\prod_{i=1}^{n-1}\bigl(1+w_i\bigr) = \prod_{i=1}^{k}\bigl(1+w_i\bigr) \le \det\bigl(L(\boldsymbol{a},0)\bigr) = a_n\,. \tag{5.19}$$

In particular, (5.19) implies that $(z_1+1)\cdots(z_k+1) \le a_n$ holds and, in consequence, (5.18) becomes

$$\|\boldsymbol{w}\|_1 \le \frac{k\cdot a_n}{2^k} < a_n\,,$$

where the final inequality follows since $k<2^k$ holds with $k\ge 2$. Finally, the integrality here implies that $\|\boldsymbol{w}\|_1 \le \|\boldsymbol{a}\|_\infty - 1$ holds, which upon similarly inspecting (5.17) yields that (5.15) holds as required, which concludes the proof of Lemma 5.2.2. □

The sail based argument used above for the case where the polyhedron $P(\boldsymbol{a},b)$ is bounded does not immediately apply in the general case. In particular, the following example demonstrates that under projection all vertices of the sail associated with the projected affine lattice $L(\boldsymbol{a},b)$ need not be feasible in the sense that they do not lie within $Q(\boldsymbol{a},b)$, which tells us that the vertices of the sail may not correspond to feasible integer points from the knapsack polyhedron $P(\boldsymbol{a},b)$ as required. Figure 5.2 illustrates this example.

Example 5.2.3. *Consider*

$$\boldsymbol{a} = (-5,49,10)^T \in \mathbb{Z}^3 \textit{ and } b = 13\,.$$

In this case, the knapsack polyhedron $P(\boldsymbol{a}, b)$ is unbounded because $\boldsymbol{a}$ contains both positive and negative entries and, further, the knapsack polyhedron $P(\boldsymbol{a}, b)$ features two vertices, namely the points $\boldsymbol{x}_1^* = (0, 13/49, 0)^T$ and $\boldsymbol{x}_2^* = (0, 0, 13/10)^T$. Under projection we yield the projected affine lattice $L(\boldsymbol{a}, b)$ and projected knapsack polyhedron

$$Q(\boldsymbol{a}, b) = \{\boldsymbol{x} \in \mathbb{R}^2_{\geq 0} : -5x_1 + 49x_2 \leq 13\}.$$

In such case, all vertices of the sail $\mathscr{E}(L(\boldsymbol{a}, b)) = \operatorname{conv}(L(\boldsymbol{a}, b) \cap \mathbb{R}^2_{\geq 0})$ associated with the projected affine lattice $L(\boldsymbol{a}, b)$ are infeasible, implying that the argument used in the proof of Lemma 5.2.2, does not immediately apply in general. Figure 5.2 illustrates this example.

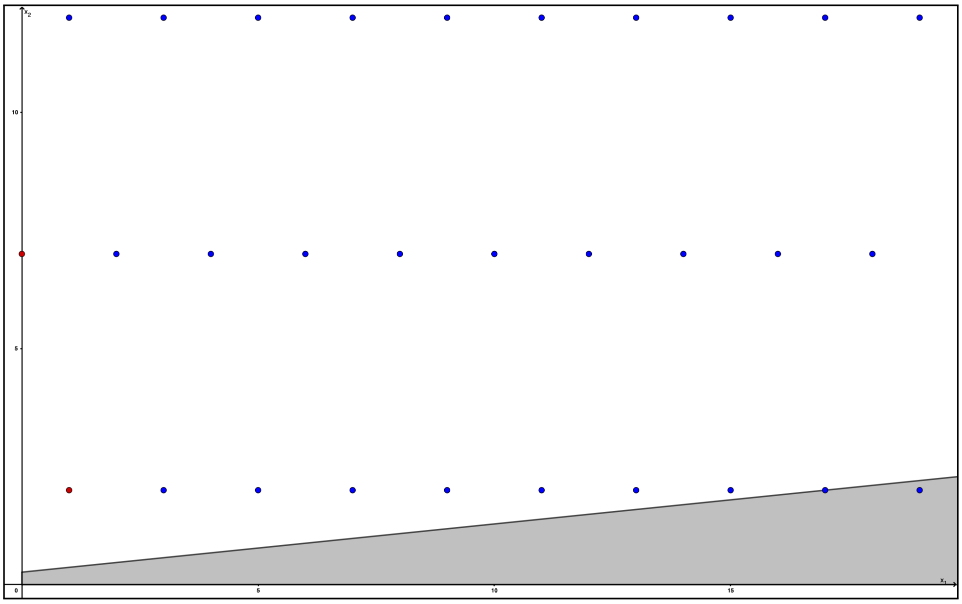


Figure 5.2: This figure provides a visualisation of Example 5.2.3. In this case, the grey area is the projected polyhedron $Q(\boldsymbol{a}, b) = \{\boldsymbol{x} \in \mathbb{R}^2_{\geq 0} : -5x_1 + 49x_2 \leq 13\}$ and the dots are nonnegative points from the projected affine lattice $L(\boldsymbol{a}, b)$. In particular, the two red dots are the only vertices of the sail $\mathscr{E}(L(\boldsymbol{a}, b))$ associated with the affine lattice $L(\boldsymbol{a}, b)$, however, these vertices lie outside of the projected feasible region $Q(\boldsymbol{a}, b)$ and consequently they do not correspond to integer points from $P(\boldsymbol{a}, b)$.

Despite this, in the case when the knapsack polyhedron $P(\boldsymbol{a}, b)$ is unbounded but the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is either bounded or empty, it turns out that one can perform a suitable translation in order to make use of the auxiliary results Lemmas 4.2.2 and 4.2.1 (from Chapter 4) in order to once more bound the distance from any vertex $\boldsymbol{x}^*$ of the knapsack polyhedron $P(\boldsymbol{a}, b)$ to a nearby feasible integer point with respect to the ℓ_1-norm. In light of this, we now consider the case when at least one of the entries of $\boldsymbol{a}$

is negative, the entries of $\boldsymbol{a}$ satisfy the condition

$$a_n = \min_i |a_i| < \|\boldsymbol{a}\|_\infty \tag{5.20}$$

and the polyhedron $P(\pi_n(\boldsymbol{a}), b) = \left\{\boldsymbol{x} \in \mathbb{R}^{n-1}_{\geq 0} : \pi_n(\boldsymbol{a})^T \boldsymbol{x} = b\right\}$ is bounded or empty. It is worth emphasising that $P(\pi_n(\boldsymbol{a}), b) \neq Q(\boldsymbol{a}, b)$ and, in particular, $P(\pi_n(\boldsymbol{a}), b)$ is a face of the projected polyhedron $Q(\boldsymbol{a}, b)$.

It should be observed that provided we consider the case $\min_i |a_i| = \|\boldsymbol{a}\|_\infty$ separately and since we can rearrange the entries of $\boldsymbol{a}$ and replace $\boldsymbol{a}$, b by $-\boldsymbol{a}$, $-b$ if necessary, we may assume that the condition (5.20) holds without loss of generality. In particular, notice that if the conditions (5.12) and $\min_i |a_i| = \|\boldsymbol{a}\|_\infty$ hold, then the vector $\boldsymbol{a}$ has the form $\boldsymbol{a} = (\pm 1, \ldots, \pm 1, 1)^T$ and, in consequence, the knapsack polyhedron $P(\boldsymbol{a}, b)$ is an integral polyhedron, i.e. each vertex $\boldsymbol{x}^*$ of $P(\boldsymbol{a}, b)$ is integral which implies that the equality $\boldsymbol{x}^* = \boldsymbol{z}$ from Theorem 5.2.1 holds.

Lemma 5.2.4. *Let $\boldsymbol{a} \in \mathbb{Z}^n$ satisfy (5.12) and $b \in Sg(\boldsymbol{a})$. If $\boldsymbol{a}$ has at least one negative entry, (5.20) holds and the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is bounded or empty, then from any vertex $\boldsymbol{x}^*$ of the knapsack polyhedron $P(\boldsymbol{a}, b)$ there exists an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ satisfying*

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2\,. \tag{5.21}$$

Proof. Because (5.20) holds and $\boldsymbol{a}$ has at least one negative entry by assumption it follows that the knapsack polyhedron $P(\boldsymbol{a}, b)$ is unbounded. Further, it follows that the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is bounded or empty only when all the entries of $\pi_n(\boldsymbol{a}) = (a_1, \ldots, a_{n-1})^T$ are negative. Throughout the proof we will consider two cases, namely when $b > 0$ and $b \leq 0$, which are the cases corresponding to when the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is empty and bounded, respectively.

Suppose firstly that $b > 0$. In this case, the knapsack polyhedron $P(\boldsymbol{a}, b)$ has the single vertex $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$. Furthermore, the positivity of b implies that the polyhedron $P\left(\pi_n(\boldsymbol{a}), b\right)$ is empty and the projection of $P(\boldsymbol{a}, b)$ is $Q(\boldsymbol{a}, b) = \mathbb{R}^{n-1}_{\geq 0}$. Recall that it was shown by Aliev et al. [8] that a feasible lattice point exists $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ which satisfies $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$. For technical reasons we assume that $\boldsymbol{z}$ is an integral vertex of the corner polyhedron $\mathscr{C}_n(\boldsymbol{a}, b)$ associated with $\boldsymbol{x}^* = (b/a_n)\,\boldsymbol{e}_n$ which is a point satisfying $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$, whose existence is guaranteed by noting that enforcing back the nonnegativity constraint $x_n \geq 0$ will not cut off any vertices of $\mathscr{C}_n(\boldsymbol{a}, b)$ because $Q(\boldsymbol{a}, b) = \mathbb{R}^{n-1}_{\geq 0}$.

The definition of the projected affine lattice $L(\boldsymbol{a}, b)$ (5.14) implies that the lattice point $\boldsymbol{z}$ has the form

$$\boldsymbol{z} = \left(z_1, z_2, \dots, z_{n-1}, \frac{b}{a_n} - \frac{a_1 z_1 + a_2 z_2 + \dots + a_{n-1} z_{n-1}}{a_n} \right)^T$$

and, in consequence, the distance between $\boldsymbol{x}^*$ and $\boldsymbol{z}$ with respect to the ℓ_∞-norm is

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \max \left(z_1, z_2, \dots, z_{n-1}, \frac{a_1 z_1 + a_2 z_2 + \dots + a_{n-1} z_{n-1}}{a_n} \right). \tag{5.22}$$

Since $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$ holds, similarly to as in the proof Lemma 5.2.2 it follows that the n-th entry appearing in (5.22) satisfies $|x_n^* - z_n| < \|\boldsymbol{a}\|_\infty - 1$ since $\|\boldsymbol{a}\|_\infty \neq 1$ in light of the assumptions (5.12) and (5.20). It is this strict upper bound which will be used in our consideration of $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1$. It follows by the definition of the ℓ_1-norm that

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 = z_1 + z_2 + \dots + z_{n-1} + \left| \frac{b}{a_n} - z_n \right| < z_1 + z_2 + \dots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1.$$

Let $\boldsymbol{w} = \pi_n(\boldsymbol{z}) = (z_1, z_2, \dots, z_{n-1})^T$. This notation allows us to express this inequality as

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < z_1 + z_2 + \dots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1 = \|\boldsymbol{w}\|_1 + \|\boldsymbol{a}\|_\infty - 1. \tag{5.23}$$

In light of this, it is sufficient for the case $b > 0$ to upper bound the ℓ_1-norm of $\boldsymbol{w}$ in order to upper bound the distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_1-norm as required.

We may assume without loss of generality that precisely k of the components from $\boldsymbol{w}$ are nonzero, where $0 \leq k \leq n-1$. Further, upon rearranging and relabelling the $n-1$ entries of $\boldsymbol{w}$ if necessary, we can assume that only the first k entries of $\boldsymbol{w}$ are nonzero. In order to complete the proof we now consider three cases, namely when $k = 0$, $k = 1$ and $k \geq 2$, respectively.

If $k = 0$, then $\boldsymbol{w}$ has no nonzero components and hence $\|\boldsymbol{w}\|_1 = 0 < \|\boldsymbol{a}\|_\infty - 1$ holds in light of (5.12) and (5.20). In particular, simply inspecting (5.23) implies that (5.21) holds as required.

If $k = 1$, then (5.23) becomes $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < z_1 + \|\boldsymbol{a}\|_\infty - 1$. Because the integral point $\boldsymbol{z}$ was chosen such that $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$ holds, it follows that (5.21) similarly holds as required.

Suppose now that $k \geq 2$. In this case, $\boldsymbol{w}$ has at least two nonzero components and making use of Lemma 4.2.3 (from Chapter 4) implies that (5.18) holds. Recall that $\boldsymbol{z}$ is a feasible integral vertex of $\mathscr{C}_n(\boldsymbol{a}, b)$ associated with $\boldsymbol{x}^*$ by assumption which satisfies $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$. In light of this and since the projection map $\pi_n(\cdot)$ is bijective it once more follows that $\boldsymbol{w}$ is a vertex of the sail $\mathscr{E}(L(\boldsymbol{a}, b))$ associated the projected affine lattice $L(\boldsymbol{a}, b)$ and, in consequence, $\boldsymbol{w}$ is irreducible with respect to the projected lattice $L(\boldsymbol{a},0)$ by Lemma 4.2.2 (from Chapter 4). Because $\boldsymbol{w}$ is irreducible with respect to the projected lattice $L(\boldsymbol{a},0)$, we can once more apply Lemma 4.2.1 (from Chapter 4) in order to deduce that the inequality (5.19) holds. Further, since $k \geq 2$, combining (5.19) and (5.18) implies that $\|\boldsymbol{w}\|_1 < a_n \leq \|\boldsymbol{a}\|_\infty - 1$ holds, where the final inequality follows by (5.20) and integrality of $\boldsymbol{a}$. Finally, upon simply inspecting (5.23) it follows that (5.21) holds as required, which concludes the case $b > 0$.

Suppose now that $b \leq 0$ and let $\boldsymbol{x}^*$ denote any vertex of the knapsack polyhedron $P(\boldsymbol{a}, b)$. In this case, it follows that the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is bounded and the condition (5.20) implies that any vertex of $P(\boldsymbol{a}, b)$ has the form form

$$\boldsymbol{x}^* = \left(0, \ldots, 0, \frac{b}{a_j}, 0, \ldots, 0\right)^T,$$

for some $1 \leq j < n$. It should be emphasised that in contrast to when $b > 0$, the projected vertex $\pi_n(\boldsymbol{x}^*)$ is not in general the zero vector, which means that the argument utilised in both the proof of Lemma 5.2.2 and for the case $b > 0$ will not work immediately here. Despite this, it will become apparent that after performing a suitable translation that a very similar argument will indeed apply.

Let $\boldsymbol{z}$ be some feasible lattice point satisfying $z_j \geq b/a_j$ and $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$, whose existence is guaranteed due to the results of Aliev et al. [8] and since $P(\boldsymbol{a}, b)$ is unbounded. Note that later in the proof we will make a further assumption about $\boldsymbol{z}$ after performing a suitable projection and translation and, for this reason, one should not think of $\boldsymbol{z}$ as being fixed at this point in the proof.

It follows upon recalling the form of the vertex $\boldsymbol{x}^* = (b/a_j)\boldsymbol{e}_j$ for $1 \leq j < n$ that the distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_∞-norm is

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty = \max\left(z_1, z_2, \ldots, z_{j-1}, \left|\frac{b}{a_j} - z_j\right|, z_{j+1}, \ldots, z_n\right). \tag{5.24}$$

Observe that because $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \leq \|\boldsymbol{a}\|_\infty - 1$ holds, the n-th entry from (5.24) clearly satisfies $z_n \leq \|\boldsymbol{a}\|_\infty - 1$, however, in this case one cannot strengthen this upper bound in contrast to previously considered cases since $z_n \in \mathbb{Z}$. It follows by the definition of the ℓ_1-norm that

$$\begin{aligned}
\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &= z_1 + z_2 + \cdots + z_{j-1} + \left|\frac{b}{a_j} - z_j\right| + z_{j+1} + \cdots + z_n \\
&\leq z_1 + z_2 + \cdots + z_{j-1} + \left|\frac{b}{a_j} - z_j\right| + z_{j+1} + \cdots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1.
\end{aligned}$$

Notice that this can be also upper bounded by

$$\begin{aligned}
\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 \leq{}& z_1 + z_2 + \cdots + z_{j-1} + \left|\lceil b/a_j \rceil - z_j\right| \\
&+ \left|\lceil b/a_j \rceil - b/a_j\right| + z_{j+1} + \cdots + z_{n-1} + \|\boldsymbol{a}\|_\infty - 1,
\end{aligned} \tag{5.25}$$

where $\lceil \cdot \rceil$ denotes the ceiling function.

Recall that $\boldsymbol{w} = \pi_n(\boldsymbol{z})$. In order to upper bound (5.25), it is sufficient for us to consider $\|\pi_n(\boldsymbol{x}^*) - \boldsymbol{w}\|_1$, which satisfies

$$\|\pi_n(\boldsymbol{x}^*) - \boldsymbol{w}\|_1 \le \left\| \lceil \pi_n(\boldsymbol{x}^*) \rceil - \boldsymbol{w} \right\|_1 + \left\| \lceil \pi_n(\boldsymbol{x}^*) \rceil - \pi_n(\boldsymbol{x}^*) \right\|_\infty . \tag{5.26}$$

Because $\pi_n(\boldsymbol{x}^*)$ is a vertex of the bounded polyhedron $P(\pi_n(\boldsymbol{a}), b)$ and $P(\boldsymbol{a}, b)$ is unbounded it follows that $\lceil \pi_n(\boldsymbol{x}^*) \rceil \in Q(\boldsymbol{a}, b)$.

For convenience, within the projected space we perform a translation along the j-th coordinate axis, which has the effect of translating $\lceil \pi_n(\boldsymbol{x}^*) \rceil$ to the origin. Being a little more precise, we perform a translation which maps every point $\boldsymbol{y} \in \mathbb{R}^{n-1}$ to $\boldsymbol{y}' \in \mathbb{R}^{n-1}$ such that

$$\boldsymbol{y} = (y_1, \dots, y_{n-1})^T \mapsto \boldsymbol{y}' = \left(y_1, \dots, y_{j-1}, y_j - \lceil b/a_j \rceil, y_{j+1}, \dots, y_{n-1}\right)^T .$$

In particular, the translations of the projected lattice point $\boldsymbol{w}$, the projected affine lattice $L(\boldsymbol{a}, b)$ and $Q(\boldsymbol{a}, b)$ will be denoted by $\boldsymbol{w}'$, $L'(\boldsymbol{a}, b)$ and $Q'(\boldsymbol{a}, b)$, respectively. Since $\lceil b/a_j \rceil \in \mathbb{Z}$ it follows that $L'(\boldsymbol{a}, b) \subset \mathbb{Z}^{n-1}$. Further, because distance is invariant under translation we note that $\| \lceil \pi_n(\boldsymbol{x}^*) \rceil - \boldsymbol{w}\|_1 = \|\boldsymbol{w}'\|_1$. In order to upper bound (5.26) we now focus on upper bounding the ℓ_1-norm of $\boldsymbol{w}'$.

Let $Q'_+(\boldsymbol{a}, b) = Q'(\boldsymbol{a}, b) \cap \mathbb{R}^{n-1}_{\ge 0}$ and observe $Q'_+(\boldsymbol{a}, b) = \mathbb{R}^{n-1}_{\ge 0}$. In particular, because the integral point $\boldsymbol{z}$ satisfies $z_j \ge b/a_j$ it follows that $\boldsymbol{w}' \in Q'_+(\boldsymbol{a}, b)$. Since $\boldsymbol{z}$ satisfies $\|\boldsymbol{x}^* - \boldsymbol{z}\|_\infty \le \|\boldsymbol{a}\|_\infty - 1$ it is reasonable in consequence to assume that $\boldsymbol{w}'$ is an integral vertex of the sail $\mathscr{E}(L'(\boldsymbol{a}, b))$ associated with the affine lattice $L'(\boldsymbol{a}, b)$. In other words, since the composite mapping is bijective we assume that $\boldsymbol{z}$ is the preimage of some integral vertex $\boldsymbol{w}'$ of the sail $\mathscr{E}(L'(\boldsymbol{a}, b))$. Since $\boldsymbol{w}'$ is a vertex of the sail $\mathscr{E}(L'(\boldsymbol{a}, b))$ it follows that $\boldsymbol{w}'$ is irreducible with respect to $L'(\boldsymbol{a}, 0)$ by Lemma 4.2.2 (from Chapter 4).

Without loss of generality, we similarly assume that precisely k of the components from $\boldsymbol{w}'$ are nonzero, where $0 \le k \le n-1$. Further, upon rearranging and relabelling the $n-1$ entries of $\boldsymbol{w}'$ if necessary, we assume that only the first k entries of $\boldsymbol{w}'$ are nonzero. In particular, it is once more useful for us to now consider three cases, namely when $k = 0$, $k = 1$ and $k \ge 2$, respectively.

If $k = 0$, then $\boldsymbol{w}'$ has no nonzero components and hence $\|\boldsymbol{w}'\|_1 = 0 \le a_n - 1$ holds in light of (5.12) and (5.20). If $k = 1$, then since $\boldsymbol{w}'$ is irreducible with respect to $L'(\boldsymbol{a}, 0)$ it follows using 4.2.1 (from Chapter 4) that

$$\prod_{i=1}^{n-1} \left(1 + w'_i\right) = 1 + w'_1 \le \det\left(L'(\boldsymbol{a}, 0)\right) = \det\left(L(\boldsymbol{a}, 0)\right) = a_n ,$$

which follows since the determinant of a lattice is invariant under translation. In particular, we deduce that $\|\boldsymbol{w}'\|_1 = w'_1 \leq a_n - 1$ holds. Suppose now that $k \geq 2$. Following (5.18) and (5.19) from the proof of Lemma 5.2.2 we can again make use of Lemmas 4.2.3 and 4.2.1 (from Chapter 4) in order to deduce that $\|\boldsymbol{w}'\|_1 < a_n$ holds. The integrality further implies that $\|\boldsymbol{w}'\|_1 \leq a_n - 1$ holds.

Recall that $\|\boldsymbol{w}'\|_1 = \|\lceil \pi_n(\boldsymbol{x}^*)\rceil - \boldsymbol{w}\|_1$ and, in consequence, we deduce that

$$\|\lceil \pi_n(\boldsymbol{x}^*)\rceil - \boldsymbol{w}\|_1 \leq a_n - 1$$

holds. Using (5.26) yields

$$\|\pi_n(\boldsymbol{x}^*) - \boldsymbol{w}\|_1 \leq a_n - 1 + \left\|\lceil \pi_n(\boldsymbol{x}^*)\rceil - \pi_n(\boldsymbol{x}^*)\right\|_\infty < a_n - 1 + 1 = a_n \leq \|\boldsymbol{a}\|_\infty - 1,$$

which follows by (5.20) and since $\|\lceil \pi_n(\boldsymbol{x}^*)\rceil - \pi_n(\boldsymbol{x}^*)\|_\infty < 1$.

Finally, recall that $z_n \leq \|\boldsymbol{a}\|_\infty - 1$ holds which implies that the ℓ_1-distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ is bounded

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 = \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{w}\|_1 + z_n < \|\boldsymbol{a}\|_\infty - 1 + \|\boldsymbol{a}\|_\infty - 1 = 2\,\|\boldsymbol{a}\|_\infty - 2,$$

which implies that the inequality (5.21) holds when $b \leq 0$ as required, which concludes the proof of Lemma 5.2.4. □

In order to complete the proof of Theorem 5.2.1 it remains to consider the case when the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is unbounded, namely the following.

Lemma 5.2.5. *Let $\boldsymbol{a} \in \mathbb{Z}^n$ satisfy (5.12) and $b \in Sg(\boldsymbol{a})$. Suppose $P(\pi_n(\boldsymbol{a}), b)$ is unbounded and (5.20) holds. Unless $\pi_n(\boldsymbol{x}^*) = \mathbf{0}$ when $\gcd(a_1, \ldots, a_{n-1}) = 1$, then from any suitable vertex $\boldsymbol{x}^*$ of the knapsack polyhedron $P(\boldsymbol{a}, b)$ there exists an integer point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ satisfying*

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2. \tag{5.27}$$

Proof. In order to prove Theorem 5.2.1 in the general case, we mirror the proof presented by Aliev et al. [8] and make use of an induction argument on n, where the base case corresponds to when the dimension is $n = 2$.

Notice that the base step where $n = 2$ is immediately settled in light of Lemma 5.2.2 and Lemma 5.2.4. In particular, we have that the distance from each vertex $\boldsymbol{x}^*$ to a nearby integer point $\boldsymbol{z}$ provided $\|\boldsymbol{a}\|_\infty \neq 1$ is bounded by $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2$. In consequence to this observation, we suppose now that $n \geq 3$, that the distance from any vertex $\boldsymbol{x}^*$ to $\boldsymbol{z}$ is bounded by $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\|\boldsymbol{a}\|_\infty - 2$ and that this upper bound holds in all dimensions $2 \leq k < n$. In light of Lemma 5.2.2 and Lemma 5.2.4, we may assume that at least one of the entries of $\boldsymbol{a}$ is negative where, since the condition (5.20) is assumed to hold, the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is unbounded. For completeness, recall that we can assume (5.20) holds without loss of generality since we can rearrange the entries of $\boldsymbol{a}$ and replace $\boldsymbol{a}$, b by $-\boldsymbol{a}$, $-b$ if necessary and since the case $\min_i |a_i| = \|\boldsymbol{a}\|_\infty$ was considered separately.

Let $\boldsymbol{x}^*$ be any vertex of the knapsack polyhedron $P(\boldsymbol{a}, b)$. Recall that $\pi_n(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n-1}$ denotes a projection which forgets about the n-th coordinate. Observe that the vertex $\boldsymbol{x}^*$ has at most one nonzero entry $x_j^* = b/a_j$ for some $1 \le j \le n$ and, further, that $\pi_n(\boldsymbol{x}^*)$ is a vertex of the projected polyhedron $Q(\boldsymbol{a}, b)$. In light of this, we consider two cases, namely the cases when $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$ and $\pi_n(\boldsymbol{x}^*) = \boldsymbol{0}$, respectively.

Firstly, suppose that $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$. Upon rearranging the first $n-1$ entries if necessary, we may assume without loss of generality that the projected vertex $\pi_n(\boldsymbol{x}^*)$ has the form $\pi_n(\boldsymbol{x}^*) = (0, \dots, 0, b/a_{n-1})^T$. In this case, observe that $\pi_n(\boldsymbol{x}^*)$ is a vertex of the polyhedron $P(\pi_n(\boldsymbol{a}), b)$. It should be emphasised that $P(\pi_n(\boldsymbol{a}), b)$ is an unbounded facet of the projected polyhedron $Q(\boldsymbol{a}, b)$ and that either $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) \neq \emptyset$ or $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) = \emptyset$. In light of this, we consider these two distinct cases in turn.

Firstly, suppose that $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) \neq \emptyset$. Denote the greatest common divisor of the first $n-1$ entries of $\boldsymbol{a}$ by g, namely let $g = \gcd(\pi_n(\boldsymbol{a}))$. By the inductive hypothesis, applied with $\pi_n(\boldsymbol{a})/g$ and b/g, there exists an $(n-1)$-dimensional integer point $\boldsymbol{y} = (y_1, y_2, \dots, y_{n-1})^T \in P(\pi_n(\boldsymbol{a}), b)$ satisfying

$$\|\pi_n(\boldsymbol{x}^*) - \boldsymbol{y}\|_1 < \frac{2\,\|\pi_n(\boldsymbol{a})\|_\infty}{g} - 2,$$

where the strict inequality follows since $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$. Simply observing that the n-dimensional integer point $\boldsymbol{z} = (y_1, y_2, \dots, y_{n-1}, 0)^T \in P(\boldsymbol{a}, b)$ is feasible implies that the distance from the vertex $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_1-norm is strictly bounded by $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - 2$, which implies that the strict inequality (5.27) holds as required.

Suppose alternatively that $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) = \emptyset$. Upon recalling the form of the affine projected lattice $L(\boldsymbol{a}, b)$ (5.14), we have

$$P(\pi_n(\boldsymbol{a}), t) \cap L(\boldsymbol{a}, t) = P(\pi_n(\boldsymbol{a}), t) \cap \mathbb{Z}^{n-1} \quad \text{for any } t \in \mathbb{Z}. \tag{5.28}$$

Since $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) = \emptyset$ and the polyhedron $P(\pi_n(\boldsymbol{a}), b)$ is unbounded by assumption, we deduce that $g \ge 2$. It is worth emphasising that when $g = 1$, then since $\pi_n(\boldsymbol{a})$ contains both positive and negative entries in light of $P(\pi_n(\boldsymbol{a}), b)$ being unbounded, there necessarily exists a nonnegative integer linear combination of the entries of $\pi_n(\boldsymbol{a})$ corresponding to an integral point, meaning that $P(\pi_n(\boldsymbol{a}), b) \cap L(\boldsymbol{a}, b) \neq \emptyset$ holds, which is in contradiction to the current assumption. In light of the conditions

(5.12), notice that $\gcd(g, a_n) = 1$ and because the projection of the knapsack polyhedron $Q(\boldsymbol{a}, b)$ has the form $Q(\boldsymbol{a}, b) = \left\{ \boldsymbol{x} \in \mathbb{R}^{n-1}_{\geq 0} : \pi_n(\boldsymbol{a})^T \boldsymbol{x} \leq b \right\}$, there exists an integer t such that

(a) t is in the interval $[b - g \cdot a_n + 1, b)$,

(b) $P(\pi_n(\boldsymbol{a}), t) \cap L(\boldsymbol{a}, b) \neq \emptyset$, i.e. that $t \equiv 0 (\operatorname{mod} g)$ and $t \equiv b (\operatorname{mod} a_n)$, and

(c) $P(\pi_n(\boldsymbol{a}), t) \subset Q(\boldsymbol{a}, b)$.

It is worth noting that (a) implies (c).

Let us choose a vertex $\boldsymbol{p}$ of the polyhedron $P(\pi_n(\boldsymbol{a}), t)$ in the following way. If the $(n-1)$-dimensional point $\boldsymbol{p}' = (0, 0, \ldots, 0, t/a_{n-1})^T$ is a vertex of $P(\pi_n(\boldsymbol{a}), t)$, then set $\boldsymbol{p} = \boldsymbol{p}'$. If not, we select $\boldsymbol{p}$ as an arbitrary vertex of $P(\pi_n(\boldsymbol{a}), t)$. By the inductive hypothesis, applied with $\pi_n(\boldsymbol{a})/g$ and t/g, there exists an $(n-1)$-dimensional integer point $\boldsymbol{y} = (y_1, y_2, \ldots, y_{n-1})^T \in P(\pi_n(\boldsymbol{a}), t)$ such that

$$\|\boldsymbol{p} - \boldsymbol{y}\|_1 \leq \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2. \tag{5.29}$$

Because $\boldsymbol{y} \in P(\pi_n(\boldsymbol{a}), t)$, we note that $a_1 y_1 + \cdots + a_{n-1} y_{n-1} = t$ holds. In light of (5.28), we notice that $\boldsymbol{y} \in L(\boldsymbol{a}, b)$ and, because $\boldsymbol{y}$ is feasible in the sense that $\boldsymbol{y} \in Q(\boldsymbol{a}, b)$ by (c), then there exists a corresponding n-dimensional feasible integer point $\boldsymbol{z} = (y_1, y_2, \ldots, y_{n-1}, z)^T \in P(\boldsymbol{a}, b)$. In consequence, observe that $a_1 y_1 + \cdots + a_{n-1} y_{n-1} + a_n z = b$ holds and hence, using (a), combining the two equations yields

$$z = \frac{b - t}{a_n} \leq \frac{b - (b - g \cdot a_n + 1)}{a_n} = \frac{g \cdot a_n - 1}{a_n} = g - \frac{1}{a_n} < g \leq \|\boldsymbol{a}\|_\infty .$$

In particular, note that since we have shown that $z < g$ holds, where $z, g \in \mathbb{Z}$, then it follows immediately that $z \leq g - 1$ holds.

Recall that $\boldsymbol{p} = \left(0, \ldots, 0, t/a_j, 0, \ldots, 0\right)^T$ for some $1 \leq j \leq n-1$ and the projected vertex $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$ has the form $\pi_n(\boldsymbol{x}^*) = (0, \ldots, 0, b/a_{n-1})^T$. If $j = n-1$, then by (a) and in light of (5.29), we yield

$$\begin{aligned} \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{y}\|_1 = \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{p}\|_1 + \|\boldsymbol{p} - \boldsymbol{y}\|_1 &\leq \left| \frac{b - t}{a_{n-1}} \right| + \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2 \\ &\leq \frac{g \cdot a_n - 1}{a_{n-1}} + \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2 . \end{aligned}$$

The distance from the vertex $\boldsymbol{x}^*$ to the n-dimensional integer point $\boldsymbol{z}$ with respect to the ℓ_1-norm is consequently bounded using $z \leq g - 1$ by

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 \leq \frac{g \cdot a_n - 1}{a_{n-1}} + \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2 + g - 1 = \frac{g \cdot a_n - 1}{a_{n-1}} + \frac{2\|\boldsymbol{a}\|_\infty}{g} + g - 3 . \tag{5.30}$$

Firstly, suppose that $\|\boldsymbol{a}\|_\infty = g$. In consequence, $|a_1| = |a_2| = \cdots = |a_{n-1}| = g$, and making use of the assumption (5.20) implies that $a_n \leq g-1$. Therefore, the distance from the vertex $\boldsymbol{x}^*$ to an integral point $\boldsymbol{z}$ with respect to the ℓ_1-norm (5.30) can be bounded from above by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &\leq \frac{g\cdot a_n - 1}{a_{n-1}} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + g - 3 \leq \frac{g\cdot(g-1)-1}{g} + \frac{2g}{g} + g - 3 \\ &= g - 1 - \frac{1}{g} + 2 + g - 3 = 2g - \frac{1}{g} - 2 < 2g - 2 = 2\,\|\boldsymbol{a}\|_\infty - 2\,,\end{aligned}$$

which implies that the inequality (5.27) holds in the case $\|\boldsymbol{a}\|_\infty = g$ as required.

It is sufficient to assume now that $\|\boldsymbol{a}\|_\infty \geq 2g$. Then, using the assumption (5.20) and since $g \geq 2$, we can bound the distance from the vertex $\boldsymbol{x}^*$ to an integral point $\boldsymbol{z}$ with respect to the ℓ_1-norm (5.30) by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &\leq \frac{g\cdot a_n - 1}{a_{n-1}} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + g - 3 < \frac{g\cdot a_n}{a_n} - \frac{1}{a_{n-1}} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + g - 3 \\ &= 2g - \frac{1}{a_n} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 3 \leq \|\boldsymbol{a}\|_\infty - \frac{1}{a_n} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 3 \\ &\leq \|\boldsymbol{a}\|_\infty - \frac{1}{a_n} + \frac{2\,\|\boldsymbol{a}\|_\infty}{2} - 3 = 2\,\|\boldsymbol{a}\|_\infty - \frac{1}{a_n} - 3 < 2\,\|\boldsymbol{a}\|_\infty - 2\,,\end{aligned}$$

which implies that the inequality (5.27) holds as required.

In order to complete the proof for the case $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$, it remains to we now consider when $j \neq n-1$. In such case, since $\boldsymbol{p}$ has the form $\boldsymbol{p} = \big(0,\ldots,0,t/a_j,0,\ldots,0\big)^T$ for some $1 \leq j < n-1$ and since the projected vertex $\pi_n(\boldsymbol{x}^*) \neq 0$ has the form $\pi_n(\boldsymbol{x}^*) = (0,\ldots,0,b/a_{n-1})^T$, it follows that the distance from the vertex $\boldsymbol{x}^*$ to an integral point $\boldsymbol{z}$ with respect to the ℓ_1-norm is bounded by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &= \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{y}\|_1 + z = \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{p}\|_1 + \|\boldsymbol{p} - \boldsymbol{y}\|_1 + z \\ &\leq \left|\frac{b}{a_{n-1}}\right| + \left|\frac{t}{a_j}\right| + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + g - 1\,,\end{aligned} \tag{5.31}$$

where the final inequality follows since $z \leq g-1$. Observe that in light of the construction of the point $\boldsymbol{p}$, if $j \neq n-1$ then $t/a_{n-1} < 0$ and $bt < 0$ hold and, in consequence to (a), $|b| + |t| \leq g\cdot a_n - 1$.

Firstly, suppose that $\|\boldsymbol{a}\|_\infty = g$ such that $|a_1| = |a_2| = \cdots = |a_{n-1}| = g$ and hence $a_n \leq g-1$ holds by (5.20). The distance from $\boldsymbol{x}^*$ to an integral point $\boldsymbol{z}$ with respect to the ℓ_1-norm (5.31) in this case can be bounded by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &\leq \left|\frac{b}{a_{n-1}}\right| + \left|\frac{t}{a_j}\right| + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + g - 3 = \left|\frac{b}{g}\right| + \left|\frac{t}{g}\right| + \frac{2g}{g} + g - 3 \\ &= \frac{|b| + |t|}{g} + g - 1 \leq \frac{g\cdot a_n - 1}{g} + g - 1 = a_n - \frac{1}{g} + g - 1\,.\end{aligned}$$

Recalling that $a_n \leq g-1$ and $g = \|\boldsymbol{a}\|_\infty$ hold, yields that $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1$ is further bounded by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 \leq \frac{g \cdot a_n - 1}{g} + g - 1 = a_n - \frac{1}{g} + g - 1 \leq g - 1 - \frac{1}{g} + g - 1 \\ = 2g - 2 - \frac{1}{g} < 2g - 2 = 2\,\|\boldsymbol{a}\|_\infty - 2\,,\end{aligned}$$

which implies that the inequality (5.27) holds as required.

It is sufficient to suppose that $\|\boldsymbol{a}\| \geq 2g$ holds, i.e. that $g \leq \|\boldsymbol{a}\|/2$. In such case, the distance from $\boldsymbol{x}^*$ to an integral point $\boldsymbol{z}$ with respect to the ℓ_1-norm (5.31) can be bounded from above by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &\leq \left|\frac{b}{a_{n-1}}\right| + \left|\frac{t}{a_j}\right| + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + g - 1 \\ &\leq \frac{|b| + |t|}{\min\{a_{n-1}, a_j\}} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + g - 3 \\ &\leq \frac{g \cdot a_n - 1}{\min\{a_{n-1}, a_j\}} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} + \frac{\|\boldsymbol{a}\|_\infty}{2} - 3\,.\end{aligned}$$

Noting that $g \geq 2$ and that $a_n < \min\{a_{n-1}, a_j\}$ hold (5.20) allows us to further bound $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1$ by

$$\begin{aligned}\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &< \frac{g \cdot a_n - 1}{a_n} + \frac{2\,\|\boldsymbol{a}\|_\infty}{2} + \frac{\|\boldsymbol{a}\|_\infty}{2} - 3 = g - \frac{1}{a_n} + \|\boldsymbol{a}\|_\infty + \frac{\|\boldsymbol{a}\|_\infty}{2} - 3 \\ &\leq \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{a_n} + \|\boldsymbol{a}\|_\infty + \frac{\|\boldsymbol{a}\|_\infty}{2} - 3 = 2\,\|\boldsymbol{a}\|_\infty - 3 - \frac{1}{a_n} < 2\,\|\boldsymbol{a}\|_\infty - 2\,,\end{aligned}$$

which implies that the inequality (5.27) holds as required and completes the proof of Lemma 5.2.5 in the case when $\pi_n(\boldsymbol{x}^*) \neq \boldsymbol{0}$.

Suppose now that $\pi_n(\boldsymbol{x}^*) = \boldsymbol{0}$, so that $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$. It is worth emphasising that one can assume that $b/a_n \notin \mathbb{Z}$ since, if this was not the case the vertex $\boldsymbol{x}^*$ would be integral and $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 = 0 < 2\,\|\boldsymbol{a}\|_\infty - 2$ holds as required. In this setting, we need to consider separately the case $\|\boldsymbol{a}\|_\infty = g = \gcd(\pi_n(\boldsymbol{a}))$. In such case, there exists an index $1 \leq i < n$ such that $\pi_i(\boldsymbol{a})$ has at least one negative entry and, in consequence, the polyhedron $P(\pi_i(\boldsymbol{a}), b)$ is unbounded since $a_n > 0$ by (5.20). Further, because $|a_1| = |a_2| = \cdots = |a_{n-1}| = g$ and $a_n/g \notin \mathbb{Z}$ by (5.12), it follows that $\gcd(\pi_i(\boldsymbol{a})) = 1$.

Note that $\pi_i(\boldsymbol{x}^*)$ is a vertex of $P(\pi_i(\boldsymbol{a}), b)$ and, since $\gcd(\pi_i(\boldsymbol{a})) = 1$, we have $P(\pi_i(\boldsymbol{a}), b) \cap \mathbb{Z}^{n-1} \neq \emptyset$. Then, by the inductive hypothesis, there exists an integer point $\boldsymbol{y} \in P(\pi_i(\boldsymbol{a}), b)$ satisfying

$$\|\pi_i(\boldsymbol{x}^*) - \boldsymbol{y}\|_1 < 2\,\|\pi_i(\boldsymbol{a})\|_\infty - 2\,,$$

where the strictness follows since $\pi_i(\boldsymbol{x}^*) \neq \mathbf{0}$. Therefore, observing that the integral point $\boldsymbol{z} = (y_1, y_2, \ldots, y_{i-1}, 0, y_{i+1}, \ldots, y_{n-1})^T \in P(\boldsymbol{a}, b)$ and recalling $\boldsymbol{x}^*$ has the form $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$ implies the distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_1-norm is strictly bounded by $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\|\boldsymbol{a}\|_\infty - 2$ implying that the inequality (5.27) holds as required.

It is once more sufficient throughout the remainder of the proof to assume that $\|\boldsymbol{a}\|_\infty \geq 2g$. Noting that $\mathbf{0} = \pi_n(\boldsymbol{x}^*) \in Q(\boldsymbol{a}, b)$, there exists an integer t such that

(a′) t is in the interval $[-g \cdot a_n + 1, 0]$,

(b′) $P(\pi_n(\boldsymbol{a}), t) \cap L(\boldsymbol{a}, b) \neq \emptyset$, and

(c′) $P(\pi_n(\boldsymbol{a}), t) \subset Q(\boldsymbol{a}, b)$.

Let $\boldsymbol{p}$ be a vertex of the polyhedron $P(\pi_n(\boldsymbol{a}), t)$. By the inductive hypothesis, applied with $\pi_n(\boldsymbol{a})/g$ and t/g, there exists an integer point $\boldsymbol{y} = (y_1, y_2, \ldots, y_{n-1})^T \in P(\pi_n(\boldsymbol{a}), t)$ such that

$$\|\boldsymbol{p} - \boldsymbol{y}\|_1 \leq \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2. \tag{5.32}$$

In particular, we assume $\boldsymbol{y} \in P(\pi_n(\boldsymbol{a}), t) \cap \mathbb{Z}^{n-1}$ and, in light of (5.28), it follows that $\boldsymbol{y} \in L(\boldsymbol{a}, t)$. Therefore, since $\boldsymbol{y}$ is feasible in the sense $\boldsymbol{y} \in Q(\boldsymbol{a}, b)$ by (c′), there exists a corresponding n-dimensional feasible integer point $\boldsymbol{z} = (y_1, y_2, \ldots, y_{n-1}, z)^T \in P(\boldsymbol{a}, b)$. In consequence, $a_1 y_1 + \cdots + a_{n-1} y_{n-1} + a_n z = b$ and hence, making use of (a′), we upper bound $|z - b/a_n|$ by

$$\left| z - \frac{b}{a_n} \right| = \left| \frac{b-t}{a_n} - \frac{b}{a_n} \right| = \left| -\frac{t}{a_n} \right| \leq \left| -\frac{(-g \cdot a_n + 1)}{a_n} \right| = g - \frac{1}{a_n}. \tag{5.33}$$

Recall that $\boldsymbol{p} = \left(0, \ldots, 0, t/a_j, 0, \ldots, 0\right)^T$ for some $1 \leq j \leq n-1$. Since the vertex $\boldsymbol{x}^*$ has the form $\boldsymbol{x}^* = (0, \ldots, 0, b/a_n)^T$, we bound the distance from $\boldsymbol{x}^*$ to $\boldsymbol{z}$ with respect to the ℓ_1-norm using the inequality (5.33) by

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 = \|\pi_n(\boldsymbol{x}^*) - \boldsymbol{p}\|_1 + \|\boldsymbol{p} - \boldsymbol{y}\|_1 + \left| z - \frac{b}{a_n} \right| \leq \left| \frac{t}{a_j} \right| + \frac{2\|\boldsymbol{a}\|_\infty}{g} - 2 + g - \frac{1}{a_n}, \tag{5.34}$$

where the final inequality follows in light of the upper bounds (5.32) and (5.33).

Suppose that $g \geq 2$ and recall $\|\boldsymbol{a}\|_\infty \geq 2g$. In light of (a′) and the assumption (5.20), we further bound $\|\boldsymbol{x}^* - \boldsymbol{z}\|_1$ (5.34) by

$$\begin{aligned}
\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &\leq \left|\frac{t}{a_j}\right| + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + g - \frac{1}{a_n} \\
&\leq \frac{g \cdot a_n - 1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{a_n} \\
&< \frac{g \cdot a_n}{a_n} - \frac{1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{a_n}\,.
\end{aligned}$$

This upper bound can be written equivalently and further bounded upon noting that $\|\boldsymbol{a}\|_\infty \geq 2g$ by

$$\begin{aligned}
\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 &< g - \frac{1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{a_n} \\
&\leq \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 + \frac{\|\boldsymbol{a}\|_\infty}{2} - \frac{1}{a_n} \\
&= \|\boldsymbol{a}\|_\infty - \frac{1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{g} - 2 - \frac{1}{a_n} \leq \|\boldsymbol{a}\|_\infty - \frac{1}{|a_j|} + \frac{2\,\|\boldsymbol{a}\|_\infty}{2} - 2 - \frac{1}{a_n},
\end{aligned}$$

where the final inequality follows since $g \geq 2$. Firstly, this upper bound can be written equivalently and strictly bounded by (5.12) and (5.20) as

$$\|\boldsymbol{x}^* - \boldsymbol{z}\|_1 < 2\,\|\boldsymbol{a}\|_\infty - \frac{1}{|a_j|} - 2 - \frac{1}{a_n} < 2\,\|\boldsymbol{a}\|_\infty - 2\,,$$

which implies that (5.27) holds, which concludes the proof of Lemma 5.2.5. □

5.3 The Existence of Sparse Solutions to the Unbounded Knapsack Problem

Recall that given $\boldsymbol{a} \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$, then $P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : \boldsymbol{a}^T\boldsymbol{x} = b\right\}$ is a knapsack polyhedron. In this section, we consider the case where the knapsack polyhedron $P(\boldsymbol{a}, b)$ is unbounded, namely when $\boldsymbol{a}$ contains both positive and negative entries.

In this setting we can similarly assume without loss of generality that $\boldsymbol{a}$ is a primitive integer vector with nonzero entries, i.e. that $\boldsymbol{a}$ satisfies (5.12). Further, we once more assume that the knapsack polyhedron $P(\boldsymbol{a}, b)$ is integer feasible, namely that the integer b belongs to the semigroup $Sg(\boldsymbol{a})$ that is generated by the entries of the vector $\boldsymbol{a}$.

The result of this section demonstrates that under further assumptions there exists an integral solution whose ℓ_0-"norm" is bounded by three. It is worth noting that Aliev et al. [4] show for the unbounded case that if $b \in Sg(\boldsymbol{a})$, then there exists an integer solution to the knapsack problem satisfying (3.10) from Chapter 3. It is worth emphasising that the method we use will be based on properties of the Frobenius number. Further, our proof is constructive and yields a polynomial time algorithm which finds an integral solution satisfying the conditions of Theorem 5.3.1. It would be interesting to know what conditions on the vector $\boldsymbol{a}$ and integer b imply that the size of support is bounded by a constant in general. The following result provides a partial answer to this question.

Theorem 5.3.1. *Let $\boldsymbol{a} \in \mathbb{Z}^n$ satisfy (5.12) and $b \in Sg(\boldsymbol{a})$. If $\boldsymbol{a}$ contains both positive and negative entries and there exists two coprime entries of $\boldsymbol{a}$ with the same sign, then there exists a feasible integral solution $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ with $\|\boldsymbol{z}\|_0 \leq 3$. Under the above assumptions, for a given $\boldsymbol{a}$ and b, such a sparse integer solution $\boldsymbol{z}$ can be computed within polynomial time.*

During the proof of Theorem 5.3.1 we use the following notation. For $q \in \mathbb{Q}$, the notation $\lceil q \rceil_0$ denotes $\lceil q \rceil_0 = \max\{\lceil q \rceil, 0\}$, namely the standard ceiling function with a lower threshold of zero.

5.3.1 The Frobenius Number

Given a set of positive integers $p_1, p_2, \ldots, p_k \in \mathbb{Z}_{>0}$ with $\gcd(p_1, p_2, \ldots, p_k) = 1$ the *Frobenius number* of those integers will be denoted by $F(p_1, p_2, \ldots, p_k)$. The Frobenius number is precisely the largest integer that cannot be represented as a nonnegative integral combination of $p_1, p_2, \ldots, p_k$. Provided that the greatest common divisor of the set of integers is equal to one, then the Frobenius number always exists and takes a finite value [63]. For completeness, note that if the greatest common divisor of a set of positive integers is not equal to one, then each integer in that set is a multiple of the greatest common divisor and, in consequence, only multiples of the greatest common divisor can be represented through integer linear combinations of the given set of integers.

In 1996, Ramirez-Alfonsin [87] showed that computing the Frobenius number in general is $\mathcal{NP}$-hard through a reduction to the integer knapsack problem. Despite this, Kannan [60] developed a polynomial time algorithm for computing the Frobenius number in 1992 provided that the number of coprime integers is fixed, however, this algorithm is known to be hard to implement. For two integers p_i and p_j with $\gcd(p_i, p_j) = 1$ Sylvester [94, 95] tells us that the Frobenius number can be found can be found using the formula

$$F\left(p_i, p_j\right) = p_i p_j - p_i - p_j .$$

In the proof of Theorem 5.3.1 we only need to compute the Frobenius number for the two coprime entries with the same sign and, in particular, we avoid the $\mathcal{NP}$-hardness by essentially fixing the dimension of the problem to two.

5.3.2 Proof of Theorem 5.3.1

Recall that the vector $\boldsymbol{a}$ features both positive and negative entries by assumption. Since we can replace $\boldsymbol{a}$, b by $-\boldsymbol{a}$, $-b$ if necessary, we can assume without loss of generality that $b \geq 0$. Further, rearranging the entries of $\boldsymbol{a}$ if necessary, we can assume without loss of generality that $\boldsymbol{a}$ is written such that only its first k entries are positive, where $1 \leq k \leq n-1$. In other words, we assume that $\boldsymbol{a}$ is written in the form

$$\boldsymbol{a} = (\,|a_1|, |a_2|, \ldots, |a_k|, -|a_{k+1}|, \ldots, -|a_n|\,)^T\,, \tag{5.35}$$

where $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ and $a_i \neq 0$ for all $i \in \{1, 2, \ldots, n\}$ in light of the conditions (5.12). The knapsack polyhedron $P(\boldsymbol{a}, b)$ can be expressed in light of (5.35) as

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : |a_1|x_1 + |a_2|x_2 + \cdots + |a_k|x_k - (\,|a_{k+1}|x_{k+1} + \cdots + |a_n|x_n\,) = b\right\}\,,$$

or, after rearranging, as

$$\begin{aligned} P(\boldsymbol{a}, b) = \big\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : |a_1|x_1 + |a_2|x_2 + \cdots + |a_k|x_k& \\ = b + (\,|a_{k+1}|x_{k+1} + \cdots + |a_n|x_n\,)\big\}&\,. \end{aligned} \tag{5.36}$$

Recall that there exists at least two entries from $\boldsymbol{a}$ with the same sign whose greatest common divisor equals to one by assumption. Observe that this equivalently tells us that by assumption there exists two coprime entries on either the left-hand or right-hand side of the equality appearing in (5.36). Intuitively the approach we take is to say that since we can create a suitable integral linear combination for any integer greater than the Frobenius number of the coprime pair using them, then provided the other side of the equality appearing in (5.36), namely the opposite side to where the aforementioned coprime pair are found, is “sufficiently large”, then the equality will necessarily hold. In particular, since we can make either the left-hand or the right-hand side “sufficiently large” as required, it is sufficient for simplicity to assume that the two coprime knapsack entries are negative, meaning they are found on the right-hand side of the equality appearing in (5.36). Further, upon rearranging the negative entries of the knapsack if required, we assume that these coprime entries have indices $k+1$ and $k+2$, respectively. In other words, we assume $\gcd(a_{k+1}, a_{k+2}) = 1$. It is worth adding that if instead the coprime pair were instead found on the left-hand side of the equality appearing in (5.36) a problem with feasibility may arise in our approach if $b > F + 1$, however, in such case we may simply make use of the standard ceiling function with a lower threshold of zero in order to ensure nonnegativity is achieved.

Let
$$y = \frac{|a_{k+1}|x_{k+1} + \cdots + |a_n|x_n}{g},$$
where $g = \gcd(a_{k+1}, \ldots, a_n)$. This notation allows us to express $P(\boldsymbol{a}, b)$ (5.36) equivalently as

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : |a_1|x_1 + |a_2|x_2 + \cdots + |a_k|x_k = b + gy\right\}.$$

Recall that $\gcd(a_{k+1}, a_{k+2}) = 1$ and, in consequence, $g = 1$, meaning that $P(\boldsymbol{a}, b)$ can be expressed equivalently as

$$P(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^n_{\geq 0} : |a_1|x_1 + |a_2|x_2 + \cdots + |a_k|x_k = b + y\right\}. \tag{5.37}$$

Recall that $F(p_1, \ldots, p_n)$ denotes the Frobenius number of the set of positive integers $p_1, \ldots, p_n \in \mathbb{Z}_{>0}$ with $\gcd(p_1, \ldots, p_n) = 1$. Let $F = F(a_{k+1}, a_{k+2})$ denote the Frobenius number for the coprime pair a_{k+1} and a_{k+2}. Because any integer greater than the Frobenius number F can be represented as a positive integral linear combination of the pair a_{k+1} and a_{k+2} and the knapsack polyhedron $P(\boldsymbol{a}, b)$ is unbounded, it is sufficient to consider only the case when $y > F$. In other words, it is sufficient to consider only the case $y > F$ since we can always create a suitable positive combination of the pair a_{k+1} and a_{k+2} that is "sufficiently large" in the sense that the combination is strictly larger than the Frobenius number. Moreover, since the Frobenius number is integral by definition, then clearly we consider only the scenario when $y \geq F + 1$.

Let $\pi_{[k]}(\cdot) : \mathbb{R}^n \to \mathbb{R}^k$ denote the projection onto the first k coordinates and, in particular, let $Q(\boldsymbol{a}, b) = \pi_{[k]}(P(\boldsymbol{a}, b))$. Because we consider only the case when $y \geq F + 1$, we will show the existence of a projected integral point from

$$Q'(\boldsymbol{a}, b) = \left\{\boldsymbol{x} \in \mathbb{R}^k_{\geq 0} : |a_1|x_1 + |a_2|x_2 + \cdots + |a_k|x_k \geq b + F + 1\right\},$$

where $Q'(\boldsymbol{a}, b) \subset Q(\boldsymbol{a}, b)$. In light of the definition of the Frobenius number, observe that every integer point in $Q'(\boldsymbol{a}, b) \cap \mathbb{Z}^k$ corresponds to the projection of at least one n-dimensional integer point in the unbounded knapsack polyhedron $P(\boldsymbol{a}, b)$. It is worth adding that this follows since for all integers $y \geq F + 1$, there exists at least one nonnegative integral linear combination of x_{k+1} and x_{k+2} satisfying the equality appearing in (5.37). It should be emphasised that the method of projection does not correspond to a bijective mapping and, in consequence, we cannot guarantee that the projection of a given n-dimensional lattice point is unique and further, we cannot perform a standard method of "lifting" in order to write a given k-dimensional projected integer point as some n-dimensional feasible lattice point.

Because all integral points in $Q'(\boldsymbol{a}, b)$ are feasible in the sense that they correspond to at least one feasible n-dimensional lattice point, we can therefore select an integral point in $Q'(\boldsymbol{a}, b)$ with minimal ℓ_0-"norm" before subsequently constructing an appropriate positive integral combination satisfying the equality appearing in (5.37). In particular, we choose a projected integer point $\pi_{[k]}(\boldsymbol{z}) \in Q'(\boldsymbol{a}, b) \cap \mathbb{Z}^k$ with at most one nonzero entry. Note we require in the worst-case one nonzero entry since $\pi_{[k]}(\boldsymbol{0}) \notin Q'(\boldsymbol{a}, b)$ in general.

More specifically, we choose a k-dimensional integral point with the form

$$\pi_{[k]}(\boldsymbol{z}) = \left(0, \ldots, 0, \left\lceil \frac{b'}{|a_i|} \right\rceil_0, 0, \ldots, 0\right)^T$$

for some $1 \leq i \leq k$, where $b' = b + F + 1$. It is worth emphasising that the lower threshold of zero for the ceiling function is required in order to ensure that the k-dimensional integer point $\pi_{[k]}(\boldsymbol{z})$ is feasible in the sense that $\pi_{[k]}(\boldsymbol{z}) \in Q'(\boldsymbol{a}, b)$.

Because $\pi_{[k]}(\boldsymbol{z}) \in Q'(\boldsymbol{a}, b)$ it follows that this k-dimensional integral point corresponds to at least one n-dimensional lattice point $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$, which has the form

$$\boldsymbol{z} = \left(0, \ldots, 0, \left\lceil \frac{b'}{|a_i|} \right\rceil_0, 0, \ldots, 0, z_{k+1}, z_{k+2}, \ldots, z_n\right)^T,$$

for some $z_{k+1}, z_{k+2}, \ldots, z_n \in \mathbb{Z}_{\geq 0}$. In light of the equality appearing in (5.36), the integral point $\boldsymbol{z}$ corresponds to the integral linear combination

$$|a_i| \left\lceil \frac{b'}{|a_i|} \right\rceil_0 = b + |a_{k+1}| z_{k+1} + |a_{k+2}| z_{k+2} + \cdots + |a_n| z_n .$$

Furthermore, since $b' = b + F + 1$, this equality can be expressed equivalently as

$$|a_i| \left\lceil \frac{b + F + 1}{|a_i|} \right\rceil_0 = b + |a_{k+1}| z_{k+1} + |a_{k+2}| z_{k+2} + \cdots + |a_n| z_n .$$

Recall that F denotes the Frobenius number of the coprime pair a_{k+1} and a_{k+2} and, by definition, every positive integer greater than F can be expressed as a positive integral linear combination of a_{k+1} and a_{k+2}. In particular, because $y \geq F + 1$, we deduce that the equality

$$|a_i| \left\lceil \frac{b + F + 1}{|a_i|} \right\rceil_0 = b + |a_{k+1}| z_{k+1} + |a_{k+2}| z_{k+2}$$

holds for some suitable z_{k+1}, z_{k+2}. In particular, this shows the existence of some integral solution $\boldsymbol{z} \in P(\boldsymbol{a}, b) \cap \mathbb{Z}^n$ to the knapsack problem with the form

$$\boldsymbol{z} = \left(0, \ldots, 0, \left\lceil \frac{b + F + 1}{|a_i|} \right\rceil_0, 0, \ldots, 0, z_{k+1}, z_{k+2}, 0, \ldots, 0\right)^T,$$

which is an integral point satisfying $\|\boldsymbol{z}\|_0 \leq 3$ as required. Finally, observe that the proof boils down to computing the greatest common divisor, which can be done using methods including the Euclidean algorithm (see e.g. [30, Chapter 31]), which concludes the proof of Theorem 5.3.1.

— Chapter 6 —

Conclusions and Future Work

During this final chapter, we provide a brief overview of the results presented in this thesis and, in addition, discuss future research directions.

6.1 Conclusions

In Chapter 4, we show that a surprising relation links proximity and sparsity of solutions to IPs when considering vertices of Gomory's corner polyhedra [49]. In particular, Theorems 4.1.1 - 4.1.6 demonstrate that a transference result holds which allows the best known distance bounds to be strengthened when vertices of the corresponding corner polyhedron are not sparse and, vice versa, strengthening the sparsity bounds if the vertices of the corner polyhedron are sufficiently far from a given vertex of the polyhedron $P(A, \boldsymbol{b})$. Recall that the corresponding vertices of the corner polyhedron need not be feasible for the original IP problem in general, however, in the knapsack scenario, we demonstrate in Theorem 4.1.6 that at least one vertex is feasible. In addition, the aforementioned transference results yield resembling ℓ_1-distance bounds, namely Corollaries 4.1.3 - 4.1.8. Furthermore, Theorem 4.1.10 presents a resembling result for optimal integer solutions to IPs which similarly connects proximity with sparsity.

In Chapter 5, we present three refinements of the known distance and sparsity bounds in special cases. In particular, Theorems 5.1.1 and 5.2.1 provide improvements for the estimates for the worst-case distance from any vertex of the polyhedron $P(A, \boldsymbol{b})$ to a nearby feasible integral point provided that $P(A, \boldsymbol{b})$ is integer feasible. Further, we present Theorem 5.3.1, which demonstrates that under certain assumptions one can always generate an integral solution to a knapsack problem whose ℓ_0-"norm" is bounded by three in polynomial time.

6.2 Future Work

It still remains an open question to provide an optimal worst-case upper bound on the distance from any vertex of the polyhedron $P(A, \boldsymbol{b})$ to a nearby feasible integer point in general. In particular, if it was proven that Conjecture 1 (from Chapter 5) holds, then the upper bound would be tight in light of the optimality of Theorem 5.1.1. Further, it remains an open problem in the general setting to prove (or disprove) that a certain transference result holds linking the proximity and sparsity when connecting vertices of the polyhedron $P(A, \boldsymbol{b})$ with vertices of the integral hull $\operatorname{conv}(P(A, \boldsymbol{b}) \cap \mathbb{Z}^n)$ as discussed in Chapter 4.

Bibliography

[1] Ravindra K Ahuja, Kurt Mehlhorn, James Orlin, and Robert E Tarjan. “Faster algorithms for the shortest path problem”. In: *Journal of the ACM (JACM)* 37.2 (1990), pp. 213–223.

[2] Ravindra K Ahyja, James B Orlin, and Thomas L Magnanti. *Network flows: theory, algorithms, and applications*. Prentice-Hall, 1993.

[3] Carlos A Alfaro, Sergio L Perez, Carlos E Valencia, and Marcos C Vargas. “The equivalence between two classic algorithms for the assignment problem”. In: *arXiv preprint arXiv:1810.03562* (2018).

[4] Iskander Aliev, Gennadiy Averkov, Jesús A. De Loera, and Timm Oertel. “Sparse representation of vectors in lattices and semigroups”. In: *Mathematical Programming* (2021). DOI: 10.1007/s10107-021-01657-8. URL: https://doi.org/10.1007/s10107-021-01657-8.

[5] Iskander Aliev, Marcel Celaya, Martin Henk, and Aled Williams. “Distance-sparsity transference for vertices of corner polyhedra”. In: *SIAM Journal on Optimization* 31.1 (2021), pp. 200–216.

[6] Iskander Aliev, Jesús A De Loera, Timm Oertel, and Christopher O’Neill. “Sparse solutions of linear Diophantine equations”. In: *SIAM Journal on Applied Algebra and Geometry* 1.1 (2017), pp. 239–253.

[7] Iskander Aliev, Jesus De Loera, Fritz Eisenbrand, Timm Oertel, and Robert Weismantel. “The Support of Integer Optimal Solutions”. In: *SIAM Journal on Optimization* 28.3 (2018), pp. 2152–2157.

[8] Iskander Aliev, Martin Henk, and Timm Oertel. “Distances to lattice points in knapsack polyhedra”. In: *Mathematical Programming* (2019). ISSN: 1436-4646. DOI: 10.1007/s10107-019-01392-1. URL: https://doi.org/10.1007/s10107-019-01392-1.

[9] Shiri Artstein-Avidan, Apostolos A Giannopoulos, and Vitali D. Milman. *Asymptotic Geometric Analysis, Part I*. Mathematical Surveys and Monographs. American Mathematical Society, 2015.

[10] Egon Balas. “Disjunctive programming”. In: *Annals of discrete mathematics*. Vol. 5. Elsevier, 1979, pp. 3–51.

[11] Egon Balas, Sebastián Ceria, and Gérard Cornuéjols. “A lift-and-project cutting plane algorithm for mixed 0–1 programs”. In: *Mathematical programming* 58.1-3 (1993), pp. 295–324.

[12] Egon Balas, Sebastian Ceria, Gérard Cornuéjols, and N Natraj. “Gomory cuts revisited”. In: *Operations Research Letters* 19.1 (1996), pp. 1–9.

[13] Egon Balas and Michael Perregaard. “Lift-and-project for mixed 0–1 programming: recent progress”. In: *Discrete Applied Mathematics* 123.1-3 (2002), pp. 129–154.

[14] Robert G Bartle and Robert G Bartle. *The elements of integration and Lebesgue measure*. Vol. 27. Wiley Online Library, 1995.

[15] David E Bell and Marshall L Fisher. “Improved integer programming bounds using intersections of corner polyhedra”. In: *Mathematical Programming* 8.1 (1975), pp. 345–368.

[16] Sebastian Berndt, Klaus Jansen, and Alexandra Lassota. “Tightness of Sensitivity and Proximity Bounds for Integer Linear Programs”. In: *International Conference on Current Trends in Theory and Practice of Informatics*. Springer. 2021, pp. 349–360.

[17] Robert Bixby, Mary Fenelon, Zonghao Gu, Ed Rothberg, and Roland Wunderling. “MIP: Theory and practice Closing the Gap”. In: *IFIP Conference on System Modeling and Optimization*. Springer. 1999, pp. 19–49.

[18] Robert Bixby and Edward Rothberg. “Progress in computational mixed integer programminga look back from the other side of the tipping point”. In: *Annals of Operations Research* 149.1 (2007), pp. 37–41.

[19] Charles E Blair and Robert G Jeroslow. “The value function of a mixed integer program: I”. In: *Discrete Mathematics* 19.2 (1977), pp. 121–138.

[20] Charles E Blair and Robert G Jeroslow. “The value function of a mixed integer program: II”. In: *Discrete Mathematics* 25.1 (1979), pp. 7–19.

[21] Enrico Bombieri and Jeffrey Vaaler. “On Siegel’s lemma”. In: *Inventiones mathematicae* 73.1 (1983), pp. 11–32.

[22] Manoel Campêlo and Gérard Cornuéjols. “Stable sets, corner polyhedra and the Chvátal closure”. In: *Operations Research Letters* 37.6 (2009), pp. 375–378.

[23] Alberto Caprara and Matteo Fischetti. “{0, 1/2}-Chvátal-Gomory cuts”. In: *Mathematical Programming* 74.3 (1996), pp. 221–235.

[24] John W S Cassels. *An Introduction to the Geometry of Numbers*. Classics in Mathematics. Springer Berlin Heidelberg, 1996.

[25] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. “Corner polyhedron and intersection cuts”. In: *Surveys in Operations Research and Management Science* 16.2 (2011), pp. 105–120.

[26] William Cook, Sanjeeb Dash, Ricardo Fukasawa, and Marcos Goycoolea. “Numerically safe Gomory mixed-integer cuts”. In: *INFORMS Journal on Computing* 21.4 (2009), pp. 641–649.

[27] William Cook, Jean Fonlupt, and Alexander Schrijver. “An integer analogue of Caratheodory’s theorem”. In: *Journal of Combinatorial Theory, Series B* 40.1 (1986), pp. 63–70.

[28] William Cook, Albertus M H Gerards, Alexander Schrijver, and Éva Tardos. “Sensitivity Theorems in Integer Linear Programming”. In: *Mathematical Programming* 34.3 (1986), pp. 251–264.

[29] William Cook, Ravindran Kannan, and Alexander Schrijver. “Chvátal closures for mixed integer programming problems”. In: *Mathematical Programming* 47.1-3 (1990), pp. 155–174.

[30] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.

[31] Gérard Cornuéjols, Yanjun Li, and Dieter Vandenbussche. “K-cuts: A variation of Gomory mixed integer cuts from the LP tableau”. In: *INFORMS Journal on Computing* 15.4 (2003), pp. 385–396.

[32] Gabriel Cramer. *Introduction a l’analyse des lignes courbes algebriques*. chez les freres Cramer & Cl. Philibert, 1750.

[33] George B Dantzig. *Linear programming and extensions*. Vol. 48. Princeton university press, 1998.

[34] George B Dantzig. “Maximization of a linear function of variables subject to linear inequalities”. In: *Activity analysis of production and allocation* 13 (1951), pp. 339–347.

[35] George B Dantzig. "Origins of the simplex method". In: *A history of scientific computing*. 1990, pp. 141–151.

[36] George B Dantzig. "Reminiscences about the origins of linear programming". In: *Mathematical Programming The State of the Art*. Springer, 1983, pp. 78–86.

[37] Sanjeeb Dash and Oktay Günlük. "On the strength of Gomory mixed-integer cuts as group cuts". In: *Mathematical Programming* 115.2 (2008), pp. 387–407.

[38] Santanu S Dey, Jean-Philippe P Richard, Yanjun Li, and Lisa A Miller. "On the extreme inequalities of infinite group problems". In: *Mathematical Programming* 121.1 (2010), pp. 145–170.

[39] Friedrich Eisenbrand and Gennady Shmonin. "Carathéodory bounds for integer cones". In: *Operations Research Letters* 34.5 (2006), pp. 564–568.

[40] Friedrich Eisenbrand and Gennady Shmonin. "Parametric integer programming in fixed dimension". In: *Mathematics of Operations Research* 33.4 (2008), pp. 839–850.

[41] Friedrich Eisenbrand and Robert Weismantel. "Proximity results and faster algorithms for integer programming using the Steinitz lemma". In: *ACM Transactions on Algorithms (TALG)* 16.1 (2019), pp. 1–14.

[42] Yonina C Eldar and Gitta Kutyniok. *Compressed sensing: theory and applications*. Cambridge university press, 2012.

[43] Lester R Ford Jr and Delbert R Fulkerson. *Flows in networks*. Princeton university press, 1962.

[44] Simon Foucart and Holger Rauhut. "A mathematical introduction to compressive sensing". In: *Bull. Am. Math* 54 (2017), pp. 151–165.

[45] Joseph Fourier. *Analyse des travaux de l'Academie Royale des Sciences, pendant l'année 1827. Partie mathématique*. 1827.

[46] Saul I Gass and Arjang A Assad. *An annotated timeline of operations research: An informal history*. Vol. 75. Springer Science & Business Media, 2005.

[47] Michel X Goemans and Thomas Rothvoss. "Polynomiality for bin packing with a constant number of item types". In: *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*. SIAM. 2014, pp. 830–839.

[48] Ralph E Gomory. "Outline of an algorithm for integer solutions to linear programs and an algorithm for the mixed integer problem". In: *50 Years of Integer Programming 1958-2008*. Springer, 2010, pp. 77–103.

[49] Ralph E Gomory. “Some Polyhedra Related to Combinatorial Problems”. In: *Linear algebra and its applications* 2.4 (1969), pp. 451–558.

[50] Ralph E Gomory, Ellis L Johnson, and Lisa Evans. “Corner polyhedra and their connection with cutting planes”. In: *Mathematical Programming* 96.2 (2003), pp. 321–339.

[51] Eugene Gover and Nishan Krikorian. “Determinants and the volumes of parallelotopes and zonotopes”. In: *Linear Algebra and its Applications* 433.1 (2010), pp. 28–40.

[52] Werner H Greub. *Linear algebra*. Vol. 23. Springer Science & Business Media, 1975.

[53] Peter M Gruber. *Convex and Discrete Geometry*. Vol. 336. Springer Science & Business Media, 2007.

[54] Ken Habgood and Itamar Arel. “A condensation-based application of Cramers rule for solving large-scale linear systems”. In: *Journal of Discrete Algorithms* 10 (2012), pp. 98–109.

[55] Godfrey H Hardy and Edward M Wright. *An introduction to the theory of numbers*. Oxford university press, 1979.

[56] Frank L Hitchcock. “The distribution of a product from several sources to numerous localities”. In: *Journal of mathematics and physics* 20.1-4 (1941), pp. 224–230.

[57] Karla L Hoffman and Manfred Padberg. “Improving LP-representations of zero-one linear programs for branch-and-cut”. In: *ORSA Journal on Computing* 3.2 (1991), pp. 121–134.

[58] Klaus Jansen. “An EPTAS for scheduling jobs on uniform processors: using an MILP relaxation with a constant number of integral variables”. In: *SIAM Journal on Discrete Mathematics* 24.2 (2010), pp. 457–485.

[59] Klaus Jansen, Kim-Manuel Klein, and José Verschae. “Closing the gap for makespan scheduling via sparsification techniques”. In: *Mathematics of Operations Research* 45.4 (2020), pp. 1371–1392.

[60] Ravi Kannan. “Lattice translates of a polytope and the Frobenius problem”. In: *Combinatorica* 12.2 (1992), pp. 161–177.

[61] Leonid V Kantorovich. “A new method of solving of some classes of extremal problems”. In: *Dokl. Akad. Nauk SSSR*. Vol. 28. 1940, pp. 211–214.

[62] Leonid V Kantorovich. "Mathematical methods of organizing and planning production". In: *Management science* 6.4 (1960), pp. 366–422.

[63] Jui-Yi Kao, Jeffrey Shallit, and Zhi Xu. "The Frobenius problem in a free monoid". In: *arXiv preprint arXiv:0708.3224* (2007).

[64] Narendra Karmarkar. "A new polynomial-time algorithm for linear programming". In: *Proceedings of the sixteenth annual ACM symposium on Theory of computing*. 1984, pp. 302–311.

[65] Hans Kellerer, Ulrich Pferschy, and David Pisinger. *Knapsack Probelms*. Springer-Verlag New York, 2004.

[66] Leonid G Khachiyan. "A polynomial algorithm in linear programming". In: *Doklady Akademii Nauk*. Vol. 244. 5. Russian Academy of Sciences. 1979, pp. 1093–1096.

[67] Leonid G Khachiyan. "Polynomial algorithms in linear programming". In: *USSR Computational Mathematics and Mathematical Physics* 20.1 (1980), pp. 53–72.

[68] Emanuel Kieronski, Jakub Michaliszyn, Ian Pratt-Hartmann, and Lidia Tendera. "Two-variable first-order logic with equivalence closure". In: *SIAM Journal on Computing* 43.3 (2014), pp. 1012–1063.

[69] Morton Klein. "A primal method for minimal cost flows with applications to the assignment and transportation problems". In: *Management Science* 14.3 (1967), pp. 205–220.

[70] Tjalling C Koopmans. "Activity analysis of production and allocation". In: *Activity Analysis of Production and Allocation* (1951).

[71] Viktor Kuncak and Martin Rinard. "Towards efficient satisfiability checking for Boolean Algebra with Presburger Arithmetic". In: *International Conference on Automated Deduction*. Springer. 2007, pp. 215–230.

[72] Jon Lee, Joseph Paat, Ingo Stallknecht, and Luze Xu. "Improving Proximity Bounds Using Sparsity". In: *Combinatorial Optimization*. Ed. by Mourad Baïou, Bernard Gendron, Oktay Günlük, and A. Ridha Mahjoub. Cham: Springer International Publishing, 2020, pp. 115–127.

[73] Jon Lee, Joseph Paat, Ingo Stallknecht, and Luze Xu. "Polynomial upper bounds on the number of differing columns of an integer program". In: *arXiv preprint arXiv:2105.08160* (2021).

[74] Derrick H Lehmer. "Euclid's algorithm for large numbers". In: *The American Mathematical Monthly* 45.4 (1938), pp. 227–233.

[75] Hendrik W Lenstra Jr. "Integer programming with a fixed number of variables". In: *Mathematics of operations research* 8.4 (1983), pp. 538–548.

[76] Adam N Letchford and Andrea Lodi. "Strengthening Chvátal–Gomory cuts and Gomory fractional cuts". In: *Operations Research Letters* 30.2 (2002), pp. 74–82.

[77] Andrea Lodi, Gilles Pesant, and Louis-Martin Rousseau. "On counting lattice points and Chvátal-Gomory cutting planes". In: *International Conference on AI and OR Techniques in Constriant Programming for Combinatorial Optimization Problems*. Springer. 2011, pp. 131–136.

[78] Colin MacLaurin. "1748, A Treatise of Algebra in Three Parts, A". In: *Millar, London* (1748).

[79] Balas K Natarajan. "Sparse approximate solutions to linear systems". In: *SIAM journal on computing* 24.2 (1995), pp. 227–234.

[80] John von Neumann. "A model of general economic equilibrium". In: *Readings in the Theory of Growth*. Springer, 1971, pp. 1–9.

[81] John von Neumann. "Discussion of a maximum problem". In: *John von Neumann Collected Works* 6 (1947), pp. 89–95.

[82] John von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1953.

[83] Timm Oertel, Joseph Paat, and Robert Weismantel. "The distributions of functions related to parametric integer optimization". In: *SIAM Journal on Applied Algebra and Geometry* 4.3 (2020), pp. 422–440.

[84] James B Orlin. "A polynomial time primal network simplex algorithm for minimum cost flows". In: *Mathematical Programming* 78.2 (1997), pp. 109–129.

[85] Christos H Papadimitriou. "On the complexity of integer programming". In: *Journal of the ACM (JACM)* 28.4 (1981), pp. 765–768.

[86] Florian A Potra and Stephen J Wright. "Interior-point methods". In: *Journal of Computational and Applied Mathematics* 124.1-2 (2000), pp. 281–302.

[87] Jorge L Ramirez-Alfonsin. "Complexity of the Frobenius problem". In: *Combinatorica* 16.1 (1996), pp. 143–147.

[88] Wolfgang M Schmidt. *Diophantine approximations and Diophantine equations*. Springer, 2006.

[89] Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1998.

[90] András Sebö. "Hilbert bases, Caratheodory's theorem and combinatorial optimization". In: *Proceedings of the 1st Integer Programming and Combinatorial Optimization Conference*. 1990, pp. 431–455.

[91] Igor R Shafarevich and Alexey O Remizov. *Linear algebra and geometry*. Springer Science & Business Media, 2012.

[92] Henry J S Smith. "I. On systems of linear indeterminate equations and congruences". In: *Proceedings of the Royal Society of London* 11 (1862), pp. 86–89.

[93] Ernst Steinitz. "Bedingt konvergente Reihen und konvexe Systeme." In: *Journal für die reine und angewandte Mathematik* 143 (1913), pp. 128–176.

[94] James J Sylvester. "On subvariants, ie semi-invariants to binary quantics of an unlimited order". In: *American Journal of Mathematics* 5.1 (1882), pp. 79–136.

[95] James J Sylvester. "Problem 7382". In: *Educational Times* 37.26 (1884).

[96] Rekha R Thomas. "The structure of group relaxations". In: *Handbooks in Operations Research and Management Science* 12 (2005), pp. 123–170.

[97] Robert J Vanderbei. *Linear programming: foundations and extensions*. Vol. 285. Springer Nature, 2015.

[98] Alan J Weir. *Lebesgue integration and measure*. Vol. 1. Cambridge University Press, 1973.