Affine Monoids, Hilbert Bases and Hilbert Functions

Dissertation
zur Erlangung des Doktorgrades

dem Fachbereich Mathematik/Informatik
der Universität Osnabrück

vorgelegt von
Robert Koch
aus Nürnberg

Contents

Preface

> The LORD is my shepherd;
> I shall not be in want.
> He makes me lie down in green pastures
> and leads me beside still waters.
>
> PSALTER 23: 1–2

The aim of this thesis is to introduce the reader to the fascinating world of affine monoids and, thereby, to present some results. The main parts are the description of an algorithm for the computation of the Hilbert basis and the Hilbert series of the integral closure of an affine monoid, as well as two theorems on the growth of the Hilbert function.

All the background information needed to understand and appreciate our considerations is developed in detail. This background material also forms a basis for studying other aspects of affine monoids, their algebras and Hilbert functions. Prerequisities for reading the thesis are modest: standard linear algebra, basic calculus, elementary point set topology and some higher algebra will suffice.

The author is grateful to the many people who have contributed to this work, be it directly or indirectly: several colleagues, in particular Morten Brun and Tim Römer, supplied valuable information; Antje Blomeyer cared for the final schedule of the thesis; and my family and friends backed me up in many different ways.

Special thanks go to my doctoral advisor, Professor Winfried Bruns: he offered me, coming from outwards, the opportunity to earn a doctorate in Osnabrück; he was prepared to help and to discuss at any time; he read through several drafts of the thesis and suggested many essential improvements; and, not least, he cared for a pleasant working atmosphere up to the very end.

Osnabrück, March 2003 ROBERT KOCH

1. Introduction

Assume we are given a finitely generated rational 'cone' in $\mathbb{R}^3$, that is, the set

$$C = \mathbb{R}_+(x_1,\dots,x_m) = \{a_1x_1+\cdots+a_mx_m : a_i \in \mathbb{R}_+\}$$

of all $\mathbb{R}_+$-linear combinations of the vectors $x_1,\dots,x_m \in \mathbb{Z}^3$. The central object of investigation now is the monoid

$$\bar{S} = C \cap \mathbb{Z}^3$$

of all lattice points in the cone; naturally, two batteries of questions arise:

(1) Is $\bar{S}$ finitely generated, that is, an 'affine monoid'?
(1′) If so, how to find a system of generators?
(2) Under which conditions does there exist a unique minimal system of generators?
(2′) And if it exists, how to find this 'Hilbert basis' of $\bar{S}$?

Among other things, we shall discuss these problems in the present work. As we shall see, question (1) can always be answered affirmatively, and as for question (2), we shall state a necessary and sufficient condition. Also, the more constructive questions (1′) and (2′) will be treated in detail.

Of course, the reader now asks himself what the bar in the name of $\bar{S}$ stands for. We immediately clarify this here: namely, we normally start with the affine monoid

$$S = \mathbb{N}(x_1,\dots,x_m) = \{a_1x_1+\cdots+a_mx_m : a_i \in \mathbb{N}\}$$

that is generated by $x_1,\dots,x_m$. Then C is the cone generated by S, and $\bar{S}$ is the 'integral closure' of S (in $\mathbb{Z}^3$).

We also mention yet another example that also gives a good impression of the environment in which this work is placed: imagine a 2-dimensional convex polygon whose vertices are integral; how to find all lattice points that are contained in it? Of course, there are bounds for the coordinates, and one could start a naive procedure which tests all the lattice points from the respective rectangle for being contained in the polygon. But as one can easily imagine, upon considering examples in higher dimension, a lot of needless points are tested, and each single test is rather time-consuming.

The computer program NORMALIZ proceeds essentially more effective. By embedding the polygon $P \subseteq \mathbb{R}^2$ in level 1 into the space next in size,

$$\tilde{P} := \{(y,1) \in \mathbb{R}^3 : y \in P\},$$

the task is reduced to problem (1′) above: namely, once a generating system of the monoid $\bar{S}$ of all lattice points in the cone

$$C = \mathbb{R}_+(x_1,\dots,x_m)$$

(whose generators $x_1,\dots,x_m \in \mathbb{Z}^3$ now are given via the vertices $y_1,\dots,y_m$ of P) is known, the lattice points $y \in P$ precisely appear in the form $x = (y,1)$ among the generators of $\bar{S}$. A central part of this thesis is dedicated to the algorithm that is used by NORMALIZ for computing the generators of $\bar{S}$. We avoid testing unnecessary points (that are not at all contained in C), and rather determine the generating lattice points directly.

In these introductory examples, we have already touched most of the central objects of the thesis: convex polytopes, finitely generated cones, affine monoids, their integral closure and the Hilbert basis. We shall now outline the arrangement of our studies.

1.1. Overview. The present work consists of five mathematical sections: the next three sections provide the foundations, and also contain already some more special statements; then the two main sections follow, containing our own considerations and results. Subsequently, we sketch the content of each section.

Section 2 contains a general introduction to convex geometry. After affine and convex sets and their faces, we consider the three central classes: convex polytopes, polyhedral sets and (finitely generated) cones. We shall study the faces of these objects in more detail and ask for existence and uniqueness of minimal representations. We shall also investigate the dual cone of a cone and describe the Fourier-Motzkin elimination for cones. We close the section with an analysis of positive cones and the connection to convex polytopes; not least, we discuss the construction of triangulations.

Section 3 then introduces the central objects of this thesis, namely, the affine monoids. Two constructions starting with an affine monoid are crucial: on the one hand, we associate a polyhedral cone and thereby have access to the theory of the previous section, which yields a geometric description of the normalization of an affine monoid; on the other hand, we associate an affine algebra and, thereby, open the gate to commutative algebra. At the end of the section, we investigate positive affine monoids and the standard embedding; this leads to Hilbert bases.

Section 4 introduces graded monoids and rings and the respective modules. This allows us to define the Hilbert function and the Hilbert series for these objects, so that we can state their properties in the homogeneous and positively graded case, respectively. This leads to Hilbert polynomials and Hilbert quasi-polynomials. We are particularly interested when the degrees of the partial polynomials of such a quasi-polynomial coincide; therefore, we shall state a sufficient condition for this to happen. We end the section by describing the Stanley decomposition of an affine algebra.

Section 5 now describes a proper application of the theory, namely, an algorithm for the computation of the normalization of affine monoids. The algorithm has been implemented, in the computer program NORMALIZ, and outputs the Hilbert basis and the Hilbert function of the normalization (if the monoid is positive). Note that there are lots of applications of NORMALIZ: it computes the lattice points in a lattice polytope, the integral closure of a monomial ideal, and it would be possible to extend it to solve Diophantine systems of linear inequalities.

Section 6 finally takes up again the notion of multigraded Hilbert functions and delves into it: we investigate the effect of the growth of the Hilbert function along arithmetic progressions (within the grading set) on global growth. This study is motivated by the case of a finitely generated module over a homogeneous ring: there, the Hilbert function grows with a degree which is well determined by the degree of the Hilbert polynomial (and the Krull dimension). Our study requires some particular resources which we shall develop over the section.

1.2. Notation and conventions. A *semigroup* $(S,+)$ is a set S, together with an associative operation $+$. S is an *Abelian* (or *commutative*) semigroup if the operation $+$ is commutative. S is called a *monoid* if there exists a neutral element $0 \in S$ with respect to the operation $+$. A *ring* $(R,+,\cdot)$ is always assumed to be commutative with unit element 1; more precisely, we require that

(1) $(R,+)$ is an Abelian group with neutral element $0 \in R$,
(2) $(R,\cdot)$ is an Abelian monoid with neutral element $1 \in R$, and
(3) the law of distributivity holds.

As usual, the *kernel* of a structure homomorphism φ is the set of those elements that are mapped to 0; we write $\ker\varphi$ for this set. Note that, in the case of monoids, the kernel does not characterize the injectivity of the homomorphism.

With $\mathbb{N}$ we denote the additive monoid

$$\mathbb{N} = \{0,1,2,\dots\}$$

of natural numbers. $\mathbb{Z}$ is the ring of integers, and $\mathbb{Q}$ and $\mathbb{R}$ denote the fields of rational and real numbers, respectively. We will sometimes need to refer to two important submonoids: $\mathbb{Q}_+$ and $\mathbb{R}_+$ denote the additive submonoids of non-negative rational and real numbers, respectively.

The symbol '$\subseteq$' will stand for 'is a subset of'; strict inclusions are denoted by the symbol '$\subsetneq$'. In order to indicate that a union of sets is disjoint, we use the symbol '$\uplus$'. The cardinality of a (possibly infinite) set X is denoted by $\operatorname{card} X$. For two subsets X,Y of an additive monoid V, we define the *(Minkowski) sum* to be

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

Note that V usually will be a finite dimensional vector space over $\mathbb{R}$.

Since they shall be used frequently, we also mention here two further symbols, introduced in the subsequent sections: C^* denotes the dual cone of a cone C, and S_0 denotes the group of units of a monoid S. Note that, by abuse of notation, M_0 also denotes the 0-th graded component of a graded module M; however, the respective meaning will always be clear from the context.

Finally, as for ordered lists, we shall stick to the following conventions:

(1) Arabic numerals are used for (independent) conditions, typically in a definition; also, they are sometimes used for indicating different steps, typically in a proof;
(a) Latin letters are used for (independent) statements, typically in a theorem;
(i) Roman numerals are used for equivalent statements, typically in a theorem.

Most of the terminology is explained in the text; for unexplained terminology, we refer the reader to the textbooks listed in the bibliography at the end of the thesis.

2. Convex geometry

This section provides the basic facts from the theory of convex geometry. After investigating general properties of affine and convex sets and their faces, we turn our attention to those three classes that are central for our work: the convex polytopes (with finite internal representation), the polyhedral sets (with finite external representation) and the cones (without finiteness requirements at first). As for these classes, we ask for existence and uniqueness of minimal representations, and we explore the connections: we shall see that the convex polytopes are precisely the bounded polyhedral sets, and that a cone is finitely generated if and only if it is polyhedral. Therefore, the class of polyhedral sets proves to be the one in which all relevant objects are contained. Consequently, we also study the facial structure of these sets in more detail.

We shall also investigate the dual cone of a cone and describe the Fourier-Motzkin elimination for cones. Finally, we establish a connection between positive polyhedral cones and convex polytopes, via a cross-section. Another important construction is the triangulation of a polyhedral cone: it plays a crucial role in practically computing the Hilbert basis and Hilbert series in Section 5.

We introduce most of the notions on our own and actually presume only basic knowledge from elementary set-theoretic topology and linear algebra. Throughout this section, the standard space $\mathbb{R}^n$ is equipped with the Euclidean scalar product (and the Euclidean norm); therefore, whenever talking about (right) angles and lengths, we refer to it.

Recommendable references for this section are the books of Brøndsted [Br] and McMullen and Shephard [MS].

2.1. Affine sets. An *affine subspace* of $\mathbb{R}^n$ is either the empty set $\emptyset$ or a translate of a linear subspace, that is, a subset $A = x + L$ where $x \in \mathbb{R}^n$ and L is a linear subspace of $\mathbb{R}^n$. By an *affine set* (or *affine space*) we mean an affine subspace of some $\mathbb{R}^n$.

A subset A of $\mathbb{R}^n$ is an affine subspace if and only if $a_1x_1 + a_2x_2 \in A$ for all $x_1, x_2 \in A$ and all $a_1, a_2 \in \mathbb{R}$ with $a_1 + a_2 = 1$. When x_1 and x_2 are distinct points from $\mathbb{R}^n$, then the set

$$\ell(x_1,x_2) := \{a_1x_1 + a_2x_2 : a_1,a_2 \in \mathbb{R},\ a_1 + a_2 = 1\} = \{(1-a)x_1 + ax_2 : a \in \mathbb{R}\}$$

is called the *line* through x_1 and x_2. With this notation, a set A is affine if and only if the line through any two points of A is contained in A.

An *affine combination* of points $x_1,\dots,x_m$ from $\mathbb{R}^n$ is an $\mathbb{R}$-linear combination $a_1x_1 + \cdots + a_mx_m$ where $a_1 + \cdots + a_m = 1$. It is clear that a set A is affine if and only if any affine combination of points from A is again in A.

Note that the intersection of any family of affine subspaces of $\mathbb{R}^n$ is again an affine subspace of $\mathbb{R}^n$. (Here it is important to accept $\emptyset$ as an affine subspace.) Therefore, for any subset M of $\mathbb{R}^n$ there is a smallest affine subspace containing M, namely, the intersection of all affine subspaces containing M. This affine subspace is called the *affine hull* of M, and is denoted by $\operatorname{aff} M$. A more explicit description of the affine hull is the following: since the set of all affine combinations of points from M is affinely closed, we have

$$\operatorname{aff} M = \left\{ \sum_{i=1}^{m} a_i x_i \,\middle|\, m \in \mathbb{N},\ a_i \in \mathbb{R},\ \sum_{i=1}^{m} a_i = 1,\ x_i \in M \right\}.$$

The vectors $x_1, \ldots, x_m$ from $\mathbb{R}^n$ are said to be *affinely independent* if, for coefficients $a_1, \ldots, a_m \in \mathbb{R}$ with $a_1 + \cdots + a_m = 0$, the equation $a_1 x_1 + \cdots + a_m x_m = 0$ is only possible for $a_1 = \cdots = a_m = 0$. This notion is motivated by the following fact: affine independence of points $x_1, \ldots, x_m$ is equivalent to the condition that any point $x \in \operatorname{aff}(x_1, \ldots, x_m)$ has a unique representation as an affine linear combination of $x_1, \ldots, x_m$. Note that this is also equivalent to saying that none of the points is an affine combination of the remaining points, or to saying that one (equivalently: all) of the $(m-1)$-tuples

$$x_1 - x_i,\ \ \ldots,\ \ x_{i-1} - x_i,\ \ x_{i+1} - x_i,\ \ \ldots,\ \ x_m - x_i$$

is linearly independent.

An *affine basis* of an affine space A is an affinely independent family $x_1, \ldots, x_m$ of points from A such that $A = \operatorname{aff}(x_1, \ldots, x_m)$. Note that an affine basis may contain the zero vector (namely, when A is a linear space). The *dimension* of a non-empty affine space A is the dimension of the linear subspace L in a representation $A = x + L$. (Since L is unique, $\dim A$ is well-defined.) We set $\dim \emptyset = -1$. Therefore, if $x_1, \ldots, x_m$ is an affine basis of A, then $\dim A = m - 1$. Yet in other words, we have $\dim A = d$ if and only if the maximal number of affinely independent vectors from A is $d + 1$.

The 0-dimensional affine spaces are the one-point sets. The 1-dimensional affine spaces have already been mentioned: these are precisely the lines $\ell(x_1, x_2)$. Finally, an $(n-1)$-dimensional affine subspace of an n-dimensional affine space A, where $n \geq 1$, is called a *hyperplane* in A. Most frequently, we will consider hyperplanes in $A = \mathbb{R}^n$. If A is an affine subspace of $\mathbb{R}^n$, then the hyperplanes in A are precisely the sets $A \cap H$, where H is a hyperplane in $\mathbb{R}^n$ such that $A \cap H$ is a non-empty proper subset of A.

For later purposes, we define the *dimension* of a subset M of $\mathbb{R}^n$ to be the dimension of the affine hull of M, that is, $\dim M := \dim(\operatorname{aff} M)$. It is convenient to call M a *d-set* in order to indicate that d is the dimension of M. We also introduce two notions which will frequently be used throughout the text: the *relative interior* of M, $\operatorname{relint} M$ for short, refers to the interior of M, when viewed within its affine hull $\operatorname{aff} M$; likewise, we define the *relative boundary*, $\operatorname{relbd} M$, of M.

Let A be an affine subspace of $\mathbb{R}^n$. A mapping $\varphi : A \to \mathbb{R}^m$ is called an *affine mapping* if it preserves affine combinations, that is,

$$\varphi\left(\sum_{i=1}^{n} a_i x_i \right) = \sum_{i=1}^{n} a_i \varphi(x_i)$$

whenever $\sum_i a_i = 1$. When φ is affine, then $\varphi(A)$ is an affine subspace in $\mathbb{R}^m$. When $A = x + L$, where L is a linear subspace of $\mathbb{R}^n$, then a mapping $\varphi : A \to \mathbb{R}^m$ is affine if and only if there exists a linear mapping $\Phi : L \to \mathbb{R}^m$ and a point $y \in \mathbb{R}^m$ such that $\varphi(x+z) = y + \Phi(z)$ for all $z \in L$. Note that affine mappings are continuous.

An affine mapping $\varphi : A \to \mathbb{R}$ is called an *affine form* on A. For each hyperplane H in A, there is a non-constant affine form φ on A such that $H = \varphi^{-1}(0)$. And conversely, $\varphi^{-1}(0)$ is a hyperplane in A for each non-constant affine form φ on A. In the most important case, namely, when $A = \mathbb{R}^n$, any affine form $\varphi : \mathbb{R}^n \to \mathbb{R}$ has a representation $\varphi = c + \Phi$, for some $c \in \mathbb{R}$ and a linear form Φ on $\mathbb{R}^n$. Note that c and Φ are uniquely determined by φ.

2.2. Convex sets. In this subsection we shall introduce the notion of a convex set, and we shall prove some basic facts about such sets.

A subset C of $\mathbb{R}^n$ is called a *convex set* if $a_1x_1 + a_2x_2$ belongs to C for all $x_1, x_2 \in C$ and all $a_1, a_2 \in \mathbb{R}_+$ with $a_1 + a_2 = 1$. When x_1 and x_2 are distinct points from $\mathbb{R}^n$, then the set

$$[x_1, x_2] := \{a_1x_1 + a_2x_2 : a_1, a_2 \in \mathbb{R}_+,\ a_1 + a_2 = 1\} = \{(1-a)x_1 + ax_2 : a \in [0,1]\}$$

is called the *closed segment* between x_1 and x_2. *Half-open segments* $]x_1, x_2]$, $[x_1, x_2[$ and *open segments* $]x_1, x_2[$ are defined analogously. With this notation, a set C is convex if and only if the closed segment between any two points of C is contained in C.

For example, the affine subspaces of $\mathbb{R}^n$, including $\emptyset$ and $\mathbb{R}^n$ itself, are convex; in particular, linear subspaces are convex. Likewise any (closed or open) halfspace is convex.

The image of a convex set under an affine mapping is again convex. In particular, translates of convex sets are again convex.

By a *convex combination* of points $x_1, \ldots, x_m$ from $\mathbb{R}^n$ we mean a linear combination $a_1x_1 + \cdots + a_mx_m$, where $a_1, \ldots, a_m \in \mathbb{R}_+$ and $a_1 + \cdots + a_m = 1$. Note that every convex combination is also an affine combination. Clearly, a subset C of $\mathbb{R}^n$ is convex if and only if any convex combination of points from C is again in C.

It is clear that the intersection of any family of convex sets in $\mathbb{R}^n$ is again convex. Therefore, for any subset M of $\mathbb{R}^n$ there is a smallest convex set containing M, namely, the intersection of all convex sets in $\mathbb{R}^n$ containing M. This convex set is called the *convex hull* of M, and is denoted by $\operatorname{conv} M$. It is well known that $\operatorname{conv} M$ also has a more explicit description, namely, it is the set of all convex combinations of points from M:

$$\operatorname{conv} M = \left\{ \sum_{i=1}^{m} a_i x_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{R}_+,\ \sum_{i=1}^{m} a_i = 1,\ x_i \in M \right\}.$$

In some sense, this is the construction of $\operatorname{conv} M$ 'from the bottom'.

It is clear that $\operatorname{conv}(x + M) = x + \operatorname{conv} M$ for any point x and any set M. More generally, we have $\operatorname{conv}(\varphi(M)) = \varphi(\operatorname{conv} M)$ when φ is an affine mapping.

Up to now there has been complete analogy with the concepts of linear and affine subspaces in $\mathbb{R}^n$, respectively. The concept of a basis of a linear or affine subspace, however, has no analogue for convex sets in general. Still, we have the following substitute.

Proposition 2.2.1: *For any subset M in $\mathbb{R}^n$, the convex hull of M is the set of all convex combinations $\sum_{i=1}^m a_i x_i$ such that $x_1, \ldots, x_m$ are affinely independent vectors from M.*

Proof: We shall prove that, if a point x is a convex combination of m points $x_1, \ldots, x_m$ from M which are affinely dependent, then x is already a convex combination of $m-1$ of the points $x_1, \ldots, x_m$. Repeating this argument if necessary, we eventually arrive at a representation of x as a convex combination of affinely independent vectors from M.

So, suppose that we have a convex combination

$$x = \sum_{i=1}^m a_i x_i\,,$$

where $x_1, \ldots, x_m \in M$ are affinely dependent. If some a_i is 0, then x is already a convex combination of $m-1$ of the points $x_1, \ldots, x_m$. Hence we may assume that all a_i are > 0. The affine dependence means that there are real numbers b_i, not all 0, such that

$$\sum_{i=1}^m b_i x_i = 0 \quad \text{and} \quad \sum_{i=1}^m b_i = 0\,.$$

Combining the equations, we see that for any real c we have

$$x = \sum_{i=1}^m (a_i - b_i c) x_i \quad \text{and} \quad \sum_{i=1}^m (a_i - b_i c) = 1\,.$$

We now simply seek a value of c (in fact, a positive value) such that $a_i - b_i c \geq 0$ for all i, with equality holding for at least one i. Clearly, we have $a_i - b_i c > 0$ for any $c > 0$ when $b_i \leq 0$. And when $b_i > 0$, we have $a_i - b_i c \geq 0$ provided that $c \leq a_i / b_i$, with $a_i - b_i c = 0$ if and only if $c = a_i / b_i$. Noting that we must have $b_i > 0$ for at least one i, we see that

$$c := \min\{a_i / b_i \mid b_i > 0\}$$

fulfils the requirements. □

The following corollary is known as *Carathéodory's Theorem.*

Corollary 2.2.2: *For any subset M of $\mathbb{R}^n$ with $\dim M = d$, the convex hull of M is the set of all convex combinations of at most $d+1$ vectors from M.*

Proof: It suffices to note that, by definition of the dimension, $d+1$ is the maximal number of affinely independent points from M. □

We conclude this subsection with an important application of Carathéodory's Theorem.

Proposition 2.2.3: *For any compact subset M of $\mathbb{R}^n$, the convex hull* $\operatorname{conv} M$ *is again compact.*

Proof: Let $d := \dim M$ and
$$N := \{(a_0, \dots, a_d) \in \mathbb{R}_+^{d+1} : a_0 + \cdots + a_d = 1\}.$$
Clearly N is compact. Now define a mapping $f : M^{d+1} \times N \to \mathbb{R}^n$ by
$$f((x_0, \dots, x_d), (a_0, \dots, a_d)) := \sum_{i=0}^{d} a_i x_i .$$
Clearly f is continuous, and by Carathéodory's Theorem, the image of f is precisely the convex hull $\operatorname{conv} M$ of M. But since M and N are compact, so too is $M^{d+1} \times N$. The assertion now follows from the fact that the image of a compact set under a continuous mapping is again compact. □

2.3. Supporting hyperplanes of closed convex sets. It is intuitively clear that, if x is a relative boundary point of a convex set C, then there is a hyperplane H passing through x such that C is contained in one of the (closed) halfspaces determined by H. This motivates the following notion.

Let C be a non-empty closed convex set in $\mathbb{R}^n$. By a *supporting hyperplane* of C we mean a hyperplane H in $\mathbb{R}^n$ such that $C \cap H$ is non-empty, and C is contained in one of the closed halfspaces determined by H. Occasionally, we may also say that H *supports* C.

Recall that H has the form $H = x_0 + L$, with a fixed point $x_0 \in H$ and a linear subspace L of dimension $n-1$. It is well known that L can be represented as the kernel of a linear form $\varphi \in (\mathbb{R}^n)^*$. But then we have
$$H = \{x \in \mathbb{R}^n : \varphi(x) = c\},$$
where we have defined $c := \varphi(x_0)$. Now the two (closed) halfspaces associated with H are
$$H^+ := \{x \in \mathbb{R}^n : \varphi(x) \geq c\} \quad \text{and} \quad H^- := \{x \in \mathbb{R}^n : \varphi(x) \leq c\}.$$
Without loss of generality, we may always assume that when H is a supporting hyperplane of C, then C is contained in H^+.

Our objective in this subsection is to show the existence of supporting hyperplanes which separate a point $x \in \mathbb{R}^n \setminus C$ from C. In order to prove this result (Lemma 2.3.3), we associate with x a 'nearest point' x_C from C.

Let $x \in \mathbb{R}^n \setminus C$ be a given point. Since C is non-empty, there exists a (closed) ball $B = B(x,r)$ with centre x and radius r such that $C' := C \cap B$ is non-empty. Since C is closed, C' is compact. Therefore the infimum
$$\operatorname{dist}(x,C) := \inf_{y \in C} \|x - y\| = \inf_{y \in C'} \|x - y\|$$

is attained, and is finite and strictly positive. It is called the *distance* of x from C. The infimum is attained for a unique point x_C from C; in fact, suppose it were attained at two distinct points x_C and x'_C, so that $\|x - x_C\| = \|x - x'_C\|$. Since C

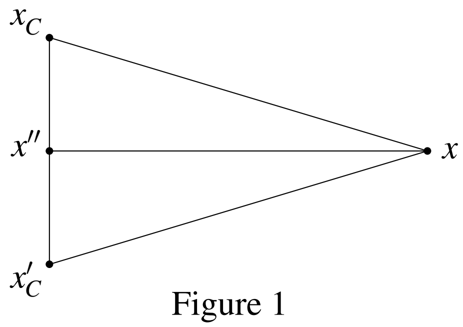

Figure 1

is convex, the mid-point $x'' = (x_C + x'_C)/2$ of the segment $[x_C, x'_C]$ belongs to C, and since the triangle with vertices x, x_C, x'' is right-angled at x'' (see Figure 1), we conclude that

$$\|x - x''\| < \|x - x_C\|,$$

which contradicts the choice of x_C. Here is a first summary:

Lemma 2.3.1: *Let C be a non-empty closed convex set in $\mathbb{R}^n$, and $x \in \mathbb{R}^n \setminus C$. Then the distance of x from C,*

$$\operatorname{dist}(x, C) = \inf_{y \in C} \|x - y\|,$$

is finite and strictly positive, and the infimum is attained at a unique point $x_C \in C$.

As a consequence of the lemma, the mapping $\pi_C : \mathbb{R}^n \to C$, given by

$$\pi_C(x) := \begin{cases} x & \text{for } x \in C, \\ x_C & \text{for } x \notin C, \end{cases}$$

is well-defined; it is called the *(nearest point) projection* on to C. In the following, we use the notation

$$\ell_+(x_1, x_2) := x_1 + \mathbb{R}_+(x_2 - x_1),$$

where x_1, x_2 are two distinct points from $\mathbb{R}^n$; it is the halfline starting at x_1 and passing through x_2. The following lemma shows that π_C is constant on the halflines $\ell_+(x_C, x)$. Note that $x_C \neq x$ for $x \notin C$.

Lemma 2.3.2: *Let $x \in \mathbb{R}^n \setminus C$ and $y \in \ell_+(x_C, x)$. Then we have $y_C = x_C$.*

Proof: For $y = x_C$ or $y = x$, the assertion is obvious. Therefore assume that $y \neq x_C$, $y \neq x$, and $y_C \neq x_C$. If $y \in\,]x_C, x[$, then we have

$$\begin{aligned} \|x - y_C\| &\leq \|x - y\| + \|y - y_C\| && \text{by the triangle inequality} \\ &< \|x - y\| + \|y - x_C\| && \text{by definition of } y_C \\ &= \|x - x_C\| && \text{since } y \in\,]x_C, x[\,; \end{aligned}$$

but this contradicts the definition of x_C.

It remains to do the case $y \notin [x_C, x]$. Then we must have $x \in \,]x_C, y[$. Let x' be the intersection of the segment $[x_C, y_C]$ with the line through x parallel to $\ell(y, y_C)$. (See

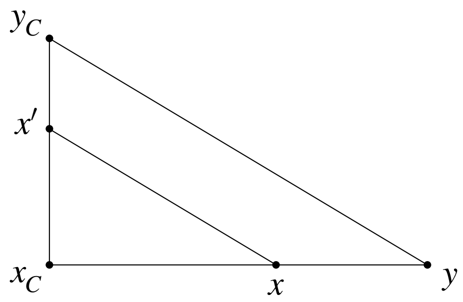

Figure 2

Figure 2; note that the figure is again close to reality, in the sense that all of the points lie in an affine space of dimension 2.) We have $\|y - y_C\| < \|y - x_C\|$, by our assumption. Since the triangles with vertices x_C, x', x and x_C, y_C, y, respectively, are similar by construction, it follows that $\|x - x'\| < \|x - x_C\|$. But this contradicts the definition of x_C, since x' lies in the segment $[x_C, y_C]$, hence in C by convexity. □

Now comes the first main result of this subsection; it is a special case of the Hahn-Banach Separation Theorem (where two disjoint closed convex sets are considered). It will be an important argument for the main result in the next subsection. Recall that we have equipped $\mathbb{R}^n$ with the standard scalar product, so that it is possible to measure angles.

Lemma 2.3.3: *Let C be a non-empty closed convex set in $\mathbb{R}^n$, and let $x \in \mathbb{R}^n \smallsetminus C$. Then the hyperplane H through x_C and perpendicular to $\ell(x_C, x)$ supports C.*

Proof: Since $x_C \in C \cap H$, $C \cap H$ is non-empty, and it remains to show that C is contained in one of the closed halfspaces determined by H. In fact, we will show that $C \subseteq H^+$ if $x \in H^-$.

So suppose to the contrary that we can find a point z in $C \cap (H^- \smallsetminus H)$, that is, a point z lying in the same open halfspace as x. Then the hyperplane through z, nor-

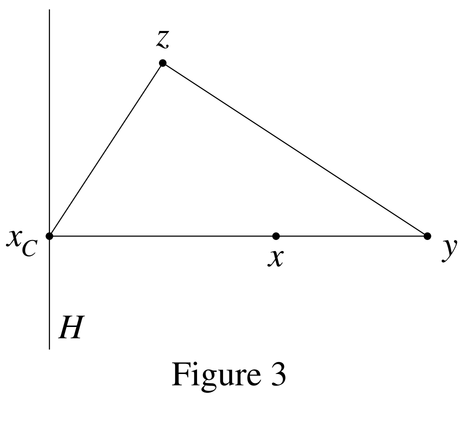

Figure 3

mal to $\ell(z,x_C)$, will meet $\ell_+(x_C,x)$ in a point y (see Figure 3). Now by Lemma 2.3.2, $x_C = y_C$, and since the triangle with vertices x_C, y, z is right-angled at z, we have

$$\|y-z\| < \|y-x_C\| = \|y-y_C\| .$$

This contradicts the definition of y_C. □

Lemma 2.3.4: (Busemann-Feller-Lemma) *Let C be a closed convex set in $\mathbb{R}^n$. Then the nearest point projection $\pi_C : \mathbb{R}^n \to C$ is Lipschitz with constant 1, that is,*

$$\|\pi_C(x) - \pi_C(y)\| \leq \|x-y\|$$

for all $x, y \in \mathbb{R}^n$. In particular, π_C is uniformly continuous.

Proof: We use the notation $x_C = \pi_C(x)$, $y_C = \pi_C(y)$. Clearly, we may assume that $x_C \neq y_C$. Let ℓ be the line through x_C and y_C. We denote by H_{x_C} and H_{y_C} the hyperplanes perpendicular to ℓ through x_C and y_C, respectively. We choose an orientation such that $x_C \in H^+_{y_C}$ and $y_C \in H^+_{x_C}$.

Now, for any point z in the interior of the stripe $H^+_{x_C} \cap H^+_{y_C}$, the foot $z_0 \in \,]x_C, y_C[$

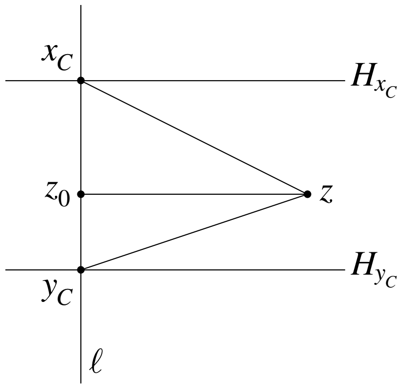

Figure 4

of the orthogonal projection of z on ℓ is contained in C; moreover,

$$\|z-z_0\| < \|z-x_C\| \quad \text{and} \quad \|z-z_0\| < \|z-y_C\| .$$

(See Figure 4.) This shows that x and y cannot lie in $\operatorname{int}(H^+_{x_C} \cap H^+_{y_C})$. Also, for any point $z \in H^-_{y_C}$, we have

$$\|z-y_C\| < \|z-x_C\| ,$$

again by elementary geometry. This shows that x cannot lie in $H^-_{y_C}$, and it follows that $x \in H^-_{x_C}$. Likewise, $y \in H^-_{y_C}$. But then

$$\|x-y\| \geq \operatorname{dist}(H_{x_C}, H_{y_C}) = \|x_C - y_C\| .$$ □

Proposition 2.3.5: *Let C be a closed convex set in $\mathbb{R}^d$, and let $x \in \operatorname{relbd} C$ be a relative boundary point. Then there exists a supporting hyperplane of C passing through x.*

Proof: It suffices to do the case where $\dim C = d$. We choose a sequence $B_n = B(x, r_n)$ of d-balls in $\mathbb{R}^d$, with centre x and decreasing radii $r_n = 1/n$. Since x is on the boundary of C, there exists a sequence of points $x_n \in B_n \setminus C$. Clearly, we have $x_n \to x$ as $n \to \infty$. And since the nearest point map π_C is continuous (Lemma 2.3.4), we also have $\pi_C(x_n) \to \pi_C(x) = x$.

Therefore it is possible to find a $(d-1)$-sphere S in $\mathbb{R}^d$ that encloses all points which have occurred up to now. For every $n \in \mathbb{N}$, let y_n be (the point in) the intersection of S with the halfline $\ell_+(\pi_C(x_n), x_n)$. Note that y_n is well-defined and $\pi_C(y_n) = \pi_C(x_n)$, by Lemma 2.3.2.

But since S is compact, there exists a convergent subsequence of (y_n), which converges to some point $y \in S$. We may assume that $y_n \to y$. Again by the continuity of π_C, we conclude that $\pi_C(y_n) \to \pi_C(y)$, and since limit points are unique, we see that $\pi_C(y) = \pi_C(x) = x$. Since we assume that $x \notin S$, we have $x \neq y$, and therefore $y \notin C$. Now it is easy: the hyperplane through x and perpendicular to $\ell(x, y)$ supports C, by Lemma 2.3.3. □

2.4. Faces of closed convex sets. In this subsection, we shall introduce certain 'extreme' convex subsets of a closed convex set C, called the 'faces' of C. As a main application, we will study the facial structure of a convex polytope in more detail in the next subsection.

In the following, let C be again a closed convex set in $\mathbb{R}^n$, and let d denote the dimension of C. (Recall that this means $d = \dim(\operatorname{aff} C)$.) Recall also that a hyperplane H in $\mathbb{R}^n$ is a supporting hyperplane of C if it intersects C in at least one point, and C is contained in one of the closed halfspaces determined by H.

Definition 2.4.1: Let H be a supporting hyperplane of the closed convex set C. Then the set $C \cap H$ is called a *face* of C.

For technical reasons, it is convenient to include among the faces of C two further faces, namely, the empty set $\emptyset$ and C itself. These are called the *improper faces* of C; all other faces are called *proper faces*. The set of all faces of C is denoted by $\mathscr{F}(C)$. It is a partially ordered set (with respect to inclusion), and is called the *face lattice* of C.

Note that there is a formal problem in connection with the definition of a proper face of C, namely, that it depends on the choice of the particular affine space containing C. If C is 'initially' lying in $\mathbb{R}^n$, we would like to be free to consider it as a subset of any affine subspace A of $\mathbb{R}^n$ containing $\operatorname{aff} C$. We can, however, easily get away with this difficulty, since the hyperplanes in A are precisely the non-empty intersections $A \cap H$, where H is a hyperplane in $\mathbb{R}^n$ not containing A.

Recall that if F is a face of C and $\dim F = j$, then F is also called a j-face of C. The 0-faces of C are called *vertices*, and the set of all vertices of C (regarded as points) is denoted by $\operatorname{vert} C$. The 1-faces of C are called *edges*, and the $(d-1)$-faces are called *facets* of C. Thus vertices and facets of C are proper faces of the least and greatest possible dimension, respectively.

Being the intersection of two closed convex sets, every face is closed and convex itself. Therefore it makes sense to talk about a 'face of a face' (of a closed convex set).

Lemma 2.4.2: *Let C be a closed convex set in $\mathbb{R}^n$. Let F and G be faces of C such that $F \subseteq G$. Then F is a face of G.*

Proof: There is nothing to show if F is an improper face of C. So we may assume that $F = C \cap H$, where H is a supporting hyperplane of C and, therefore, also of G. But then $F = G \cap H$, which proves the lemma. □

It is worth noting that the converse of Lemma 2.4.2 is false for closed convex sets in general. As Figure 5 illustrates, $F := \{x\}$ can be a face of $G := [x, y]$, and

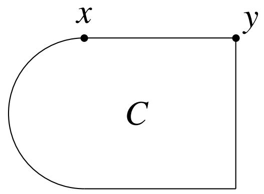

Figure 5

G a face of C, although F is not a face of C. However, for a more special class of closed convex sets, namely, the polyhedral sets, the converse of Lemma 2.4.2 is true (see Proposition 2.7.5).

Lemma 2.4.3: *Let C be a closed convex set in $\mathbb{R}^n$, and let $F_1, \ldots, F_m$ be faces of C. Then the intersection*

$$F := F_1 \cap \cdots \cap F_m$$

is a face of C, too.

Proof: It obviously suffices to show the lemma for $m = 2$ and $\dim C = n$. If F_1 or F_2 is an improper face of C, then the assertion is clearly valid. So we assume F_1 and F_2 to be proper faces of C, given as intersections $F_i = C \cap H_i$, $i = 1, 2$, where the H_i are hyperplanes in $\mathbb{R}^n$ such that $C \subseteq H_i^+$. Then we have $F = C \cap H_1 \cap H_2$, and if $F = \emptyset$, then there is nothing to show. (F is an improper face of C then.) So assume that there exists $x_0 \in F$. Without loss of generality, we may even assume that $x_0 = 0 \in F$. (A translation of the coordinate system affects neither the hypothesis nor the assertion.) Then there are linear forms φ_1, φ_2 on $\mathbb{R}^n$, both non-zero, such that

$$H_i = \{x \in \mathbb{R}^n : \varphi_i(x) = 0\} \quad \text{and} \quad H_i^+ = \{x \in \mathbb{R}^n : \varphi_i(x) \geq 0\}$$

for $i = 1, 2$.

It is clear that the sum $\varphi := \varphi_1 + \varphi_2$ is a non-trivial linear form on $\mathbb{R}^n$. ($\varphi_1 + \varphi_2 = 0$ implies $H_2^+ = H_1^-$, and then C is contained in $H_1^+ \cap H_1^- = H_1$.) We find

$$\varphi(x) = \varphi_1(x) + \varphi_2(x) \geq 0$$

for all $x \in C$. Therefore $H := \{x \in \mathbb{R}^n : \varphi(x) = 0\}$ is a supporting hyperplane of C in $\mathbb{R}^n$. (Note that $0 \in C \cap H$.) Moreover, for $x \in C$, we have $\varphi(x) = 0$ if and only if $\varphi_1(x) = \varphi_2(x) = 0$; this implies that

$$C \cap H = C \cap H_1 \cap H_2 = F\,,$$

and so F is a face of C. □

The following lemma shows that proper faces are always contained in the (relative) boundary of C; intuitively, this is clear.

Lemma 2.4.4: *Let C be a closed convex set in $\mathbb{R}^n$.*

(a) *Let F be a proper face of C. Then F is contained in* $\operatorname{relbd} C$*, the relative boundary of C.*

(b) *Conversely, any point $x \in$* $\operatorname{relbd} C$ *is contained in some proper face of C.*

Proof: It suffices to prove the lemma in the case where $\dim C = n$.

(a) There is a supporting hyperplane H of C, such that $F = C \cap H$ (and C is contained in H^+). Now let $x \in F$. Then x lies in C, and therefore either in its interior, or on the boundary. Assume that, in contrast to the assertion, $x \in \operatorname{int} C$. Then there exists a (closed) n-ball $B = B(x, r)$ in $\mathbb{R}^n$, with centre x and radius $r > 0$, which is contained in C, hence also in H^+. But this is impossible, since the centre x lies in H.

(b) is just a reformulation of Proposition 2.3.5. □

Note that, in the proof, neither closedness nor convexity of C have been used. Faces of a closed convex set have yet another nice property: their relative interior (that is, relative to their affine hull) is never empty.

Lemma 2.4.5: *Let C be a closed convex set in $\mathbb{R}^n$, and let F be a non-empty face of C. Then* $\operatorname{relint} F$ *is also non-empty.*

Proof: We have already mentioned that F is (closed and) convex again. Let d be the dimension of F, and let $x_0, \ldots, x_d$ be affinely independent points in F. Now define $x := 1/(d+1) \cdot (x_0 + \cdots + x_d)$. We have $x \in F$, by convexity. It is easy to see that there is an n-ball $B = B(x, r)$ such that

$$B \cap \operatorname{aff} F \subseteq F\,.$$

Since the left-hand side is a d-ball with centre x in $\operatorname{aff} F$, this shows that $x \in \operatorname{relint} F$. □

Lemma 2.4.6: *Let C be an n-dimensional closed convex set in $\mathbb{R}^n$. Let F be a face of C, and let $x \in$* $\operatorname{relint} F$*. Then F is the smallest face of C containing x; in other words, if G is a face of C with $x \in G$, then $F \subseteq G$.*

Proof: G is non-empty, since $x \in G$; and for $G = C$, the assertion is true. Therefore, we may assume that G is a proper face, and as such, it has a representation $G = C \cap H$ with a supporting hyperplane H of C. Choose an orientation such that C is contained in the halfspace H^+. Let $d = \dim F$. Then, since $x \in \operatorname{relint} F$, there is a d-ball B with centre x which is contained in F. But then B is also contained in C, hence in the halfspace H^+. Since the centre x of B lies in the bounding hyperplane H, this is only possible if the whole ball B is contained in H. (Note that this implies $d < n$, hence $F \neq C$.) It follows that $\operatorname{aff} B = \operatorname{aff} F$ is contained in H, and so $F \subseteq H$. Since $F \subseteq C$ anyway, we have shown that $F \subseteq G$. □

2.5. Convex polytopes. In this subsection, we investigate a special class of closed convex sets, namely the polytopes. The main result will be that a polytope is the convex hull of its vertices.

By a *(convex) polytope* we mean the convex hull of a non-empty finite set of points $x_1, \ldots, x_m$ in $\mathbb{R}^n$. Clearly, if P is a polytope, then any translate $x + P$ of P is also a polytope; this follows from the fact that $x + \operatorname{conv} M = \operatorname{conv}(x + M)$. More generally, the image of a polytope under an affine mapping is again a polytope; this follows from the fact that $\varphi(\operatorname{conv} M) = \operatorname{conv} \varphi(M)$ whenever φ is an affine mapping. A polytope P which can be presented as the convex hull of affinely independent points $x_1, \ldots, x_m$ is called a *simplex*.

Since finite sets are compact, we obtain as an immediate consequence of Proposition 2.2.3 that polytopes are always compact.

Corollary 2.5.1: *Any convex polytope P in $\mathbb{R}^n$ is a compact set.*

As a first approach to our main objective, we prove that the face lattice $\mathscr{F}(P)$ of P is finite.

Proposition 2.5.2: *Let P be a convex polytope, given as the convex hull of the points $x_1, \ldots, x_m$ in $\mathbb{R}^n$. Then any face F of P is the convex hull $F = \operatorname{conv} M$ of a subset $M \subseteq \{x_1, \ldots, x_m\}$. In particular, each face of P is a convex polytope itself, and P has only a finite number of faces.*

Proof: Let H be an arbitrary supporting hyperplane of P, given by the equation $\varphi(x) = c$, for some linear form φ on $\mathbb{R}^n$ and some $c \in \mathbb{R}$. Without loss of generality, we may assume that $x_1, \ldots, x_r$ lie in H, and that $x_{r+1}, \ldots, x_m$ lie in the open halfspace

$$\operatorname{int} H^+ = \{x \in \mathbb{R}^n : \varphi(x) > c\}.$$

Then we have $\varphi(x_i) = c$ for $i = 1, \ldots, r$, and $\varphi(x_i) = c + d_i$ for $i = r+1, \ldots, m$, for some $d_i > 0$. Now let x be an arbitrary point of P; then x is a convex combination

$$x = \sum_{i=1}^{m} a_i x_i$$

of the points $x_1,\dots,x_m$, with $a_i \in \mathbb{R}_+$ and $\sum_{i=1}^m a_i = 1$, of course. Then we have

$$\varphi(x) = \sum_{i=1}^m a_i \varphi(x_i) = \sum_{i=1}^m a_i c + \sum_{i=r+1}^m a_i d_i = c + \sum_{i=r+1}^m a_i d_i \,.$$

The point x lies in H if and only if $\varphi(x) = c$, that is, if and only if $\sum_{i=r+1}^m a_i d_i = 0$; since $d_i > 0$ for all $i \geq r+1$, this means that $a_i = 0$ for $i = r+1,\dots,m$, or, in other words, that x is a convex combination of $x_1,\dots,x_r$ only. We conclude that

$$H \cap P = \operatorname{conv}(x_1,\dots,x_r)\,,$$

and so the face $H \cap P$ is the convex hull of the set $M = \{x_1,\dots,x_r\}$. □

We are now in a position to state the main result of this subsection. It shows that a polytope is 'spanned' by its extremal points (which is intuitively clear), and is known as *Minkowski's Theorem.*

Proposition 2.5.3: *Any convex polytope P is the convex hull of its vertices, that is,*

$$P = \operatorname{conv}(\operatorname{vert} P)\,.$$

Proof: Let P be the convex hull of the points $x_1,\dots,x_m$ in $\mathbb{R}^n$. We may assume that this is a minimal choice of points, in the sense that the polytopes

$$P_i := \operatorname{conv}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_m)$$

do not coincide with P for $i = 1,\dots,m$. (This is due to the fact that $\operatorname{conv} M = \operatorname{conv}(M \cup \{x\})$ if $x \in \operatorname{conv} M$.) The remainder of the proof consists in showing that

$$\operatorname{vert} P = \{x_1,\dots,x_m\}\,.$$

By Proposition 2.5.2, any 0-dimensional face of P has the form $\operatorname{conv}(x_i) = \{x_i\}$; therefore, the inclusion $\operatorname{vert} P \subseteq \{x_1,\dots,x_m\}$ is clear.

In order to prove the other inclusion, consider any point x_i, and the polytope P_i. Then $x_i \in \mathbb{R}^n \setminus P_i$, and so there is a nearest point $x_i^0 \in P_i$ with respect to x_i. By

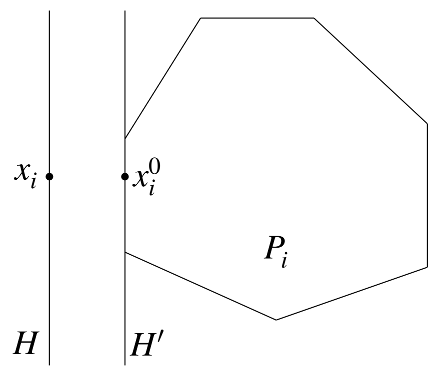

Figure 6

Lemma 2.3.3, the hyperplane H' through x_i^0 and normal to $\ell(x_i^0, x_i)$ supports P_i, and therefore, all the points $x_1,\dots,x_{i-1},x_{i+1},\dots,x_m$ lie in the closed halfspace bounded by H' which does not contain x_i (see Figure 6). Let H be the hyperplane through

x_i parallel to H'. Then $x_1, \ldots, x_m$ all lie in one of the closed halfspaces determined by H, say H^+, and so $P \subseteq H^+$. Since $x_i \in H \cap P$, we conclude that H supports P. Further, it is clear that no other of the points $x_1, \ldots, x_m$ lies in H, and so by Proposition 2.5.2,

$$H \cap P = \mathrm{conv}(x_i) = \{x_i\}.$$

Hence the point x_i is a vertex of P. □

The set X spanning a given polytope P is of course not unique (except when P is a 1-point set); in fact, one may always add any points from P. However, there is a unique *minimal* set X, namely, the set of vertices of P, already found in the preceding proposition.

Lemma 2.5.4: *Let P be a convex polytope and X a subset of P. Then $P = \mathrm{conv}\, X$ if and only if $\mathrm{vert}\, P \subseteq X$.*

Proof: The 'if' has been shown in Proposition 2.5.3. For the 'only if', suppose that there is a vertex v of P which is not in X. Then X is a subset of $P \setminus \{v\}$, and since $P \setminus \{v\}$ remains convex, it follows that $\mathrm{conv}\, X$ is also a subset of $P \setminus \{v\}$. □

2.6. Polyhedral sets. In the previous subsection we have introduced the geometric objects which have a finite 'internal' representation, namely, the convex polytopes. In this subsection we shall study the objects which have a finite 'external' representation, that is, which are intersections of a finite number of closed halfspaces.

A subset Q of $\mathbb{R}^n$ is called a *polyhedral set* (or *polyhedron*) if Q is the intersection of a finite number of closed halfspaces, the empty intersection being interpreted as the polyhedral set $Q = \mathbb{R}^n$.

Every hyperplane H in $\mathbb{R}^n$ is the intersection of the two closed halfspaces determined by H, and every proper affine subspace A of $\mathbb{R}^n$ is the intersection of a finite number of hyperplanes in $\mathbb{R}^n$. Therefore every affine subspace in $\mathbb{R}^n$ is polyhedral.

Clearly, the intersection of a finite number of polyhedral sets is again polyhedral; in particular, the faces of a polyhedral set are again polyhedral. Likewise, any translate of a polyhedron is again a polyhedron. Note that a polyhedron is closed and convex, but may be unbounded.

The main basic fact in polytope theory now is that the polytopes are precisely the bounded polyhedrons. This confirms, in fact, one's intuitive picture of a convex polytope. We start by showing the following proposition.

Proposition 2.6.1: *Let P be a convex polytope in $\mathbb{R}^n$. Then P is a bounded polyhedral set.*

Proof: First, it is evident that P is bounded, and it remains to show that P can be represented as a polyhedral set. Suppose, without loss of generality, that P is an n-polytope in $\mathbb{R}^n$. Let $\mathrm{vert}\, P = \{x_1, \ldots, x_m\}$, and let $F_1, \ldots, F_s$ denote the facets of P. (According to Proposition 2.5.2, there are only finitely many vertices and facets.) Then associated with each facet F_i is a supporting hyperplane H_i such that $H_i \cap P = F_i$ and $P \subseteq H_i^+$. Therefore, it is clear that

$$P \subseteq H_1^+ \cap \cdots \cap H_s^+ .$$

The proof of the proposition depends upon proving that, in fact,

$$P = H_1^+ \cap \cdots \cap H_s^+ .$$

So suppose that there exists a point $x \in H_1^+ \cap \cdots \cap H_s^+$, with $x \notin P$. Let D be the union of the affine subspaces of $\mathbb{R}^n$ spanned by at most n points, one of which is x, and the remaining $n-1$ are vertices of P. Then D is a finite union of affine subspaces of dimension at most $n-1$, and because $\dim P = n$, this implies that $\operatorname{int} P \nsubseteq D$. Hence we can find a point $y \in \operatorname{int} P$ with $y \notin D$.

Since $y \in \operatorname{int} P$, and $x \notin P$, there exists a unique point $z \in \operatorname{bd} P$ such that

$$[x,y] \cap \operatorname{bd} P = \{z\} ,$$

by the convexity of P. By Lemma 2.4.4, z is contained in some proper face of P. We shall show that z belongs to a facet of P, since it does not belong to any face of lower dimension. For suppose that z belonged to a j-face of P (with $0 \leq j \leq n-2$). Then, by Carathéodory's Theorem 2.2.2, z lies in some simplex of dimension $r \leq j$ whose vertices $w_0, \ldots, w_r$ are vertices of P. Thus

$$z \in \operatorname{conv}(w_0, \ldots, w_r) ,$$

and so, because $r < n-1$, we deduce that z is contained in one of the affine subspaces from D. As x is contained in each of these subspaces, this implies that $y \in D$, which is a contradiction.

So, let the facet to which z belongs be F_i. Then $z \in H_i$, and since $y \in \operatorname{int} P \subseteq \operatorname{int} H_i^+$, we deduce that $x \notin H_i^+$. But this contradicts our initial assumption that $x \in H_1^+ \cap \cdots \cap H_s^+$, whence it follows that

$$P = H_1^+ \cap \cdots \cap H_s^+ ,$$

as desired. □

Our next goal is to show the converse of Proposition 2.6.1. To this end, let $H_1, \ldots, H_s$ be hyperplanes in $\mathbb{R}^n$, and let Q be an n-polyhedron in $\mathbb{R}^n$, given as the intersection

$$Q = H_1^+ \cap \cdots \cap H_s^+ \tag{1}$$

of the halfspaces H_i^+.

Clearly, we may suppose that none of the halfspaces H_i^+ is redundant, in the sense that the polyhedral sets

$$Q_i := \bigcap_{j \neq i} H_j^+$$

do not coincide with Q for any $i = 1, \ldots, s$. In this case, the representation (1) of Q is called *irreducible*. Note that $Q \subset Q_i$, whence the polyhedron Q_i has dimension n, too.

As a first step towards the converse of Proposition 2.6.1, we show that the halfspaces in an irreducible representation of Q are uniquely determined by Q. So, consider the sets $F_i := H_i \cap Q = H_i \cap Q_i$ for $i = 1, \ldots, s$. It is easy to see that F_i is non-empty for all i; for if $H_i \cap Q_i = \emptyset$, then, since the polyhedral set Q_i is convex and therefore connected, it is contained either in H_i^+, or in H_i^-. In the former case, we conclude that $Q = Q_i$, which contradicts the irreducibility of the representation (1). In the latter, we find that $Q \subseteq H_i$, which is impossible due to the dimension. Therefore, we have shown that $Q \cap H_i$ is non-empty; since Q is contained in H_i^+, we conclude that H_i supports Q, and that F_i is a face of Q.

In the following proposition, we show even more, namely, that F_i is a facet of Q; the proposition can be taken as an analogue of Minkowski's Theorem 2.5.3, which dealt with representations of convex polytopes.

Proposition 2.6.2: *Let $H_1, \ldots, H_s$ be hyperplanes in $\mathbb{R}^n$. Let*

$$Q = H_1^+ \cap \cdots \cap H_s^+$$

be an irreducible representation of an n-polyhedron Q in $\mathbb{R}^n$. Then the following statements hold:

(a) *The hyperplane sections $F_i = Q \cap H_i$, $i = 1, \ldots, s$, are facets of Q.*

(b) *Every facet F of Q coincides with one of the facets F_i, $i = 1, \ldots, s$.*

In particular, Q has only a finite number of facets, and an irreducible representation of Q is uniquely determined by Q, via its facets.

Proof: (a) We have already shown above that F_i is a face of Q; it remains to show that it has dimension $n-1$. For this, we consider the interior of F_i, relative to H_i. This is precisely the set $f_i := H_i \cap \operatorname{int} Q_i$. We claim that this set is non-empty; for if $H_i \cap \operatorname{int} Q_i = \emptyset$, then, since $\operatorname{int} Q_i$ is convex and therefore connected, it is contained either in H_i^+, or in H_i^-. But then also Q_i is contained either in H_i^+ or in H_i^-, which leads to the same contradictions as above. (Above, we have started with Q_i instead of $\operatorname{int} Q_i$.)

Now, since there is a point $x \in f_i$, there is also an n-ball $B = B(x, r)$ with centre x and radius r which is contained in Q_i. But then the $(n-1)$-ball $B' := B \cap H_i$ (with centre x) is contained in F_i, and so $\dim F_i = n-1$.

(b) It follows immediately from the representation of Q that

$$\operatorname{int} Q = \operatorname{int} H_1^+ \cap \cdots \cap \operatorname{int} H_s^+ .$$

Since Q is closed, this implies that

$$\operatorname{bd} Q = Q \setminus \operatorname{int} Q = Q \cap \bigcup_{i=1}^{s} H_i = \bigcup_{i=1}^{s} F_i .$$

Now let F be a facet of Q, given as the intersection $F = H \cap Q$ of Q with a supporting hyperplane H in $\mathbb{R}^n$ (where $Q \subseteq H^+$). Since F is non-empty, so is $\operatorname{relint} F$ (by Lemma 2.4.5). So choose some point x in $\operatorname{relint} F$. Then x lies in the boundary of Q, since $F \subseteq \operatorname{bd} Q$ by Lemma 2.4.4. Thus $x \in F_i$ for some i. But then F is contained in F_i, since, by Lemma 2.4.6, F is the minimal face of Q which contains x. This implies an inclusion $H \subseteq H_i$ of the affine hulls, which must be an equality in view of their dimensions. Therefore, it turns out that $F = F_i$. □

In the proof of part (b), the following lemma has already been shown.

Lemma 2.6.3: *Let Q be a polyhedral set in $\mathbb{R}^n$. Then* $\operatorname{relbd} Q$*, the relative boundary of Q, is precisely the union of all facets of Q.*

The following proposition is another main result of this subsection; namely, it is the converse of Proposition 2.6.1.

Proposition 2.6.4: *Let Q be a bounded polyhedral set in $\mathbb{R}^n$. Then Q is a convex polytope.*

Proof: Let $H_1, \ldots, H_s$ be hyperplanes in $\mathbb{R}^n$, such that

$$Q = H_1^+ \cap \cdots \cap H_s^+$$

is the intersection of the corresponding halfspaces. Without loss of generality, we may suppose that this representation is irreducible, and that Q has dimension n.

Then the sets

$$F_i := H_i \cap Q = \bigcap_{j \neq i} (H_i \cap H_j^+)$$

are precisely the facets of Q, as shown in Proposition 2.6.2. Also, since (for $i \neq j$) $H_i \cap H_j^+$ is either the whole of H_i, or a closed halfspace in H_i, we see that F_i is itself a bounded polyhedral set in H_i.

We now use induction on the dimension n. The proposition clearly holds if $n = 1$, so assume that it holds in any dimension $\leq n-1$. Since $\dim F_i = n-1$, it follows that F_i is a convex polytope, and so by Minkowski's Theorem 2.5.3,

$$F_i = \operatorname{conv}(\operatorname{vert} F_i)\,.$$

Now let V be the finite set of all vertices of the facets F_i,

$$V = \bigcup_{i=1}^{s} \operatorname{vert} F_i\,.$$

Since V is contained in Q, and Q is convex, we conclude that $\operatorname{conv} V$ is also contained in Q.

We now show that $Q = \operatorname{conv} V$; hence it follows that Q is a convex polytope. In fact, for any point $x \in Q$, there are two possibilities:

(1) $x \in \operatorname{bd} Q$. Then, by Lemma 2.6.3, $x \in F_i$ for some i, and since the vertices of F_i all belong to V, it follows that $x \in \operatorname{conv} V$.

(2) $x \in \operatorname{int} Q$. Let ℓ be any line through x. Then ℓ meets $\operatorname{bd} Q$ in precisely two points (since Q is convex), say x_0 and x_1. Now $x_0 \in \operatorname{bd} Q$, and so, as in (1), $x_0 \in \operatorname{conv} V$. Likewise $x_1 \in \operatorname{conv} V$, and since $x \in [x_0, x_1]$, we deduce that $x \in \operatorname{conv} V$ again. □

We summarize Propositions 2.6.1 and 2.6.4.

Corollary 2.6.5: *A subset P of $\mathbb{R}^n$ is a convex polytope if and only if it is a bounded polyhedral set.*

We end this subsection with another separation lemma. We have already seen in Lemma 2.3.3 that a point x which lies outside a closed convex set C can be separated from C by a hyperplane. If we are dealing with a whole sequence of points and a polyhedral set, we can expect even more:

Lemma 2.6.6: *Let Q be a polyhedral set in an $\mathbb{R}$-vector space V, and $(x_n)_{n\in\mathbb{N}}$ a sequence of points in $V \smallsetminus Q$ which converges to some point $x \in Q$. Then there exists, for all $n \gg 0$, a supporting hyperplane H_n of Q such that $x_n \in \operatorname{int} H_n^-$ and $x \in H_n$ (and $Q \subseteq H_n^+$).*

Proof: We start by choosing hyperplanes $h_1, \dots, h_s$ in V such that

$$Q = h_1^+ \cap \cdots \cap h_s^+$$

is the intersection of the respective halfspaces. Without loss of generality we may assume that $x \in h_i$ for $i = 1, \dots, r$ and $x \in \operatorname{int} h_i^+$ for $i = r+1, \dots, s$. Since $x_n \to x$ as $n \to \infty$, we also have $x_n \in \operatorname{int} h_i^+$ for $i > r$ and n large enough, say $n \geq n_0$. But by hypothesis, the sequence (x_n) runs outside Q, and it follows that ($r \geq 1$ and) $x_n \notin h_1^+ \cap \cdots \cap h_r^+$ for $n \geq n_0$. So if $x_n \notin h_i^+$ (for some $1 \leq i \leq r$), we may choose $H_n := h_i$. □

As the proof shows, it is also possible to choose a fixed hyperplane which separates infinitely many points x_n from Q. In particular, we obtain the following statement.

Corollary 2.6.7: *Let Q be a polyhedral set in V, and $y, z \in V$ with $y \notin Q$, $z \in Q$. Let x be that point in Q where the halfline $\ell_+(y,z)$ (in its natural parametrization) enters Q. Then there exists a supporting hyperplane H of Q such that $x \in H$ and $y \in \operatorname{int} H^-$.*

Proof: Define the sequence

$$x_n := \frac{1}{n} \cdot y + \left(1 - \frac{1}{n}\right) \cdot x \in [y, x[$$

for $n > 0$; then $x_n \notin Q$, and the sequence converges to x. Now apply the lemma (or the above remark) and use the fact that, when H is a supporting hyperplane of Q, then $x_n \in \operatorname{int} H^-$ implies that $y \in \operatorname{int} H^-$ as well. □

2.7. Faces of polyhedral sets. We now investigate the facial structure of polyhedral sets more closely. We have already seen that a convex polytope has only a finite number of faces (Proposition 2.5.2). So one natural question which arises is: does the respective statement also hold for the more general class of polyhedral sets? Here is the answer:

Lemma 2.7.1: *Any polyhedral set Q has only a finite number of faces, and each proper face is a face of some facet of Q.*

Proof: We have seen in Proposition 2.6.2 that Q has only finitely many facets. So assume that $F \neq \emptyset$ is a face of Q of dimension $\leq d-2$, where d denotes the dimension of Q. Now choose $x \in \operatorname{relint} F$ (which is possible by Lemma 2.4.5). Since F is a proper face of Q, it is contained in $\operatorname{relbd} Q$ (by Lemma 2.4.4), and hence, so is x. But then x lies in some facet G of Q (Lemma 2.6.3), and it follows that the whole face F is contained in G (Lemma 2.4.6), and is therefore also a face of G (Lemma 2.4.2). Now the lemma follows easily, by induction on d. □

Corollary 2.7.2: *Let Q be a polyhedral set in $\mathbb{R}^n$, and $x \in \mathbb{R}^n$. Then there exists an n-ball $B = B_\varepsilon(x)$ with centre x such that B intersects Q only in faces G for which $x \in G$. Furthermore, if x lies in the relative interior of some face F of Q, then B intersects Q only in faces G for which $F \subseteq G$.*

Proof: Due to the finite number of faces, it suffices to show: if G is a face of Q with $x \notin G$, then the distance of x from G is strictly positive; but this has been observed in Lemma 2.3.1. The additional statement is an immediate consequence of Lemma 2.4.6. □

Imagine a polyhedron in $\mathbb{R}^2$: each vertex is the intersection of two edges; likewise, for a 3-polyhedron in $\mathbb{R}^3$, each edge is the intersection of two facets. This observation can be generalized:

Lemma 2.7.3: *Let Q be a polyhedral set in $\mathbb{R}^n$, let G be a facet of Q, and let F be a facet of G. Then F is the intersection of two facets of Q.*

Proof: Suppose that $\dim Q = n$, and that the representation

$$Q = H_1^+ \cap \cdots \cap H_s^+$$

of Q as an intersection of halfspaces (in $\mathbb{R}^n$) is irreducible. Then, by Proposition 2.6.2, the facets F_i of Q are precisely the intersections of Q with the bounding hyperplanes $H_i = \operatorname{relbd} H_i^+$. In particular, we have

$$G = Q \cap H_i = \bigcap_{j \neq i} (H_i \cap H_j^+)$$

for some i. Clearly, $H_i \cap H_j^+$ is either the whole of H_i or a halfspace in H_i: the empty set is excluded, since the representation of Q was assumed to be irreducible. (Also, G would be empty then.) By leaving out superfluous sets, we obtain an irreducible representation of the polyhedral set G as an intersection of halfspaces in H_i. (If

$G = H_i$, then there is nothing to show.) Therefore, again by Proposition 2.6.2, F is the intersection of G with a relative bounding hyperplane $\operatorname{relbd}(H_i \cap H_j^+)$ (for a suitable j):

$$F = G \cap \operatorname{relbd}(H_i \cap H_j^+) = G \cap H_i \cap H_j = Q \cap H_i \cap H_j = F_i \cap F_j\,.$$

This proves the lemma. □

Recall from Lemma 2.4.3 that the intersection of faces is again a face of Q. As one expects, the dimension decreases:

Lemma 2.7.4: *Let Q be a polyhedral set of dimension n in $\mathbb{R}^n$, and F_1, F_2, F_3 facets of Q. Then the following hold:*
(a) *If $F_1 \neq F_2$, then* $\dim(F_1 \cap F_2) \leq n-2$.
(b) *If F_1, F_2, F_3 are pairwise distinct, then* $\dim(F_1 \cap F_2 \cap F_3) \leq n-3$.

Proof: (a) This is an immediate consequence of the fact that the intersection of two different hyperplanes has dimension at most $n-2$. (In fact, the dimension equals $n-2$, unless the intersection is empty.)

(b) Unfortunately, the intersection of three different hyperplanes may have dimension $n-2$; therefore, we have to reason more precisely here, using the supporting property of the hyperplanes. First, we may assume that $F_1 \cap F_2 \cap F_3$ is non-empty, and by a translation of the coordinate system, we achieve that $0 \in F_1 \cap F_2 \cap F_3$. Then the respective supporting hyperplanes $H_i = \operatorname{aff} F_i$ are given by the vanishing of a linear form φ_i, for $i = 1,2,3$. By (a), we have

$$\dim(H_1 \cap H_2 \cap H_3) \leq \dim(H_1 \cap H_2) = n-2\,.$$

Now assume that, in contrast to the assertion, $\dim(H_1 \cap H_2 \cap H_3) = n-2$. Then the equation of H_3, $\varphi_3(x) = 0$, does not affect the linear system $H_1 \cap H_2$ (of rank 2), so that there is a dependency

$$\varphi_3 = a \cdot \varphi_1 + b \cdot \varphi_2 \tag{2}$$

with $a, b \in \mathbb{R}$. Using Lemma 2.4.5 and Corollary 2.7.2, we choose, for $i = 1,2,3$, a point $x_i \in \operatorname{relint} F_i$ and find that

$$\varphi_i(x_i) = 0 \quad \text{and} \quad \varphi_j(x_i) > 0$$

for $j \neq i$. Evaluating (2) on x_1, x_2, x_3, we successively obtain $b > 0$, $a > 0$ and, finally, a contradiction. □

The following proposition shows that the relation of being a face is transitive on the set $\mathscr{F}(Q)$ of all faces of a polyhedron Q.

Proposition 2.7.5: *Let Q be a polyhedral set in $\mathbb{R}^n$, let G be a face of Q, and F a face of G. Then F is a face of Q, too.*

Proof: We may suppose that both F and G are proper faces. We claim that F is the intersection of (finitely many) facets of Q; by Lemma 2.4.3, this proves the proposition. We prove our claim by induction on $d := \dim Q$. For $d = -1, 0, 1$, the claim is obvious. So assume that $2 \leq d \leq n$, and that the claim holds for all polyhedral sets of dimension $d-1$. Since G is assumed to be proper, we have $0 \leq \dim G < d$. Choose $x \in \operatorname{relint} G$ (which is possible by Lemma 2.4.5). Then x lies in the relative boundary of Q (Lemma 2.4.4), hence in some facet F_i of Q (Lemma 2.6.3). But then, by Lemma 2.4.6, we find that G is contained in F_i, and by Lemma 2.4.2, G is even a face of F_i. Now the induction hypothesis, applied to the polyhedral set F_i of dimension $d-1$, tells us that F is the intersection of finitely many facets of F_i. But then, by Lemma 2.7.3, F is also a finite intersection of facets of Q. This proves our claim. □

Corollary 2.7.6: *Every polyhedral set Q is the disjoint union*

$$Q = \biguplus_{F \in \mathscr{F}(Q)} \operatorname{relint} F$$

of the relative interiors of all its faces.

Proof: The assertion amounts to saying that every point $x \in Q$ is contained in the relative interior of precisely one face F of Q.

We show first the existence of such a face F. If $x \in \operatorname{relint} Q$, then we choose $F = Q$, of course. So assume that $x \in \operatorname{relbd} Q$. Then, by Lemma 2.6.3, x is contained in a facet G of Q. Since this is a polyhedral set of lower dimension, we can show by induction that there exists a face F of G such that $x \in \operatorname{relint} F$. But in fact, by Proposition 2.7.5, F is also a face of Q.

The face F is uniquely determined by x; namely, it is the smallest face of Q (with respect to inclusion) such that $x \in \operatorname{relint} F$, cf. Lemma 2.4.6. □

The rest of this subsection is dedicated to the following observation (and its consequences): for a 3-polyhedron in $\mathbb{R}^3$, each vertex is contained in an edge, and each edge is contained in a facet. More generally, the following lemma holds.

Lemma 2.7.7: *Let Q be a polyhedral set, and let $F_j \subseteq F_k$ be two faces of Q, of dimensions j and k, respectively. Then there are faces F_i, $j < i < k$, of Q such that*

$$F_j \subseteq F_{j+1} \subseteq \cdots \subseteq F_{k-1} \subseteq F_k$$

and $\dim F_i = i$ *for all i.*

Proof: By Lemma 2.4.2, F_j is a face of F_k. Suppose that it is a proper face; then, by Lemma 2.7.1, F_j is a face of some facet F_{k-1} of F_k. Now the lemma follows at once, by induction on k. □

Note that the lemma also applies when $F_j = \emptyset$ and $F_k = Q$; it then yields a maximal chain of faces of Q. The lemma allows the following conclusion; it is a characterization of the 'subfacets' of Q, which will play a crucial role for the investigations in Section 5.

Corollary 2.7.8: *Let Q be a polyhedral set of dimension n, and F a face of Q. Then F has dimension $n-2$ if and only if it is contained in precisely two facets of Q (and is the intersection of these).*

Proof: Suppose first that F has dimension $n-2$. Then there exists a facet G of Q which contains F (Lemma 2.7.7). Therefore, F is a facet of G (Lemma 2.4.2), so that Lemma 2.7.3 applies and yields that F is the intersection of two (different) facets of Q. Clearly, F is contained in these two facets. For reasons of dimension, F cannot be contained in yet another facet. (The intersection of three different facets has dimension at most $n-3$, by Lemma 2.7.4.)

Conversely, suppose that F is the intersection of the facets G, G' of Q, and is not contained in any other facet of Q. Since G and G' are different facets, the dimension of F is at most $n-2$. So there exists an $(n-2)$-face F' of Q such that $F \subseteq F' \subseteq G$ (Lemma 2.7.7). Therefore, F' is contained in (and the intersection of) precisely two facets of Q (by the already proven part of the corollary). But these facets must be G and G', since these are the only facets in which F is contained. This implies that $F = G \cap G' = F'$, and so $\dim F = n-2$. □

Note that the corollary confirms one's intuitive picture of a 3-dimensional polyhedron (or polytope) P in $\mathbb{R}^3$: each edge is in fact contained in precisely two facets (which in turn uniquely determine the edge, via intersection). However, a similar criterion for the lower-dimensional faces does not hold: a vertex of P may be contained in an arbitrary number of facets (and edges).

2.8. Cones. By a *cone* in a real vector space $V \cong \mathbb{R}^n$ we mean a subset $C \neq \emptyset$ which is closed under linear combinations with non-negative real coefficients. C is *finitely generated* if there exist $x_1, \dots, x_m \in C$ such that

$$C = \mathbb{R}_+(x_1, \dots, x_m) := \left\{ \sum_{i=1}^{m} a_i x_i \;\middle|\; a_i \in \mathbb{R}_+ \right\}.$$

The set $\mathbb{R}_+(x_1, \dots, x_m)$ is the cone *generated by* $x_1, \dots, x_m$; it is the smallest cone which contains $x_1, \dots, x_m$.

Note that a cone C is a convex set; it always contains the origin 0. If C is finitely generated, it is even a closed convex set (see Corollary 2.9.3), and the whole theory of Subsections 2.3 and 2.4 applies.

If the cone C is generated by vectors with rational or, equivalently, integral coordinates, then C is called *rational*. Note that C need not be finitely generated for this.

A cone $C \subseteq V$ generates an $\mathbb{R}$-vector subspace

$$\mathbb{R}C = \left\{ \sum_{i=1}^{m} a_i x_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{R},\ x_i \in C \right\}$$

of V; since 0 is contained in any cone, the affine hull aff$C$ and the linear hull $\mathbb{R}C$ of C coincide. Therefore there is also equality of the dimensions:

$$\dim C = \dim(\operatorname{aff} C) = \dim \mathbb{R}C\,.$$

Note that any supporting hyperplane H of a cone C must contain 0, that is, H is a linear hyperplane in V. In fact, let H be given by the equation $\varphi(x) = c$, for some linear form φ on V and some $c \in \mathbb{R}$; we have to show that $c = 0$. Now $C \cap H \neq \emptyset$ means that there is a point $x_0 \in C$ such that $\varphi(x_0) = c$; and $C \subseteq H^+$ means that $\varphi(x) \geq c$ for all $x \in C$. Since $0 \in C$, we already see that $c \leq 0$. But since $x_0 \in C$, we also have $2x_0 \in C$; therefore

$$c \leq \varphi(2x_0) = 2\varphi(x_0) = 2c\,,$$

and it follows that $c \geq 0$ as well. This observation has an immediate consequence for the face lattice $\mathscr{F}(C)$ of C.

Lemma 2.8.1: *Let C be a cone in $\mathbb{R}^n$. Then the following statements hold:*
(a) *Any supporting hyperplane H of C is a linear subspace of $\mathbb{R}^n$.*
(b) *C has at most one vertex, namely, the origin* 0.

Proof: (a) has been shown above. For (b), it suffices to note that if a supporting hyperplane H shares a point $x \neq 0$ with C, then H and C already share the halfline $\ell_+(0,x)$. □

2.9. Dual cones. For a cone C in a real vector space $V \cong \mathbb{R}^n$, one defines the *dual cone* by

$$C^* = \{\varphi \in V^* : \varphi(x) \geq 0 \text{ for all } x \in C\}\,.$$

Note that C^* is in fact closed under addition and multiplication with scalars from $\mathbb{R}_+$. A linear form $\varphi \in C^*$ is called a *dual form* of C. Evidently, each dual form defines a face of C, by intersecting it with the respective hyperplane. Those dual forms which define a facet of C are uniquely determined up to a positive real number, and are called the *support forms* of C; we will refine this notion previous to Lemma 2.9.7.

It is an important fact that C^* is finitely generated if C is so. We will now give a constructive proof of this fact; the algorithm is known as the *Fourier-Motzkin elimination* for cones [Zi, Section 1.3].

So let $x_1, \ldots, x_m \in \mathbb{R}^n$. We want to find the dual cone of $C = \mathbb{R}_+(x_1, \ldots, x_m)$. We can assume that $\mathbb{R}^n$ is generated by $x_1, \ldots, x_m$ as a vector space (that is, $\dim C = n$).

So we first search for n vectors among $x_1, \ldots, x_m$ that form a basis of $\mathbb{R}^n$, say $x_1, \ldots, x_n$. For each $i = 1, \ldots, n$, we compute a linear form φ_i such that $\varphi_i(x_i) > 0$, $\varphi_i(x_j) = 0$ for $j \neq i$. Clearly, φ_i is uniquely determined up to a positive factor. One also checks immediately that $\varphi_1, \ldots, \varphi_n$ are a basis of $(\mathbb{R}^n)^*$, and generate the dual cone of $C_0 = \mathbb{R}_+(x_1, \ldots, x_n)$. (Obviously, one can choose $\varphi_1, \ldots, \varphi_n$ as the dual basis of $x_1, \ldots, x_n$.) We now describe how the dual cone changes if we enlarge C by another generator.

Proposition 2.9.1: *Let $x_1, \ldots, x_m, y \in \mathbb{R}^n$ be such that $x_1, \ldots, x_m$ generate $\mathbb{R}^n$ as a vector space. Suppose that $\varphi_1, \ldots, \varphi_t$ generate the dual cone of $C = \mathbb{R}_+(x_1, \ldots, x_m)$. For each pair (i, j), $i, j = 1, \ldots, t$, with $\varphi_i(y) > 0$ and $\varphi_j(y) < 0$, we set*

$$\psi_{ij} := \varphi_i(y) \cdot \varphi_j - \varphi_j(y) \cdot \varphi_i .$$

Then the dual cone of $\tilde{C} = C + \mathbb{R}_+ y$ is generated by the ψ_{ij} and those φ_k with $\varphi_k(y) \geq 0$.

Proof: We have to show that

$$\tilde{C}^* = \mathbb{R}_+\{\psi_{ij} : \varphi_i(y) > 0,\ \varphi_j(y) < 0\} + \mathbb{R}_+\{\varphi_k : \varphi_k(y) \geq 0\}.$$

The inclusion '$\supseteq$' is clear: $\tilde{C}$ is generated by the vectors $x_1, \ldots, x_m, y$, and

$$\psi_{ij}(x_r) = \underbrace{\varphi_i(y)}_{>0} \cdot \underbrace{\varphi_j(x_r)}_{\geq 0} - \underbrace{\varphi_j(y)}_{<0} \cdot \underbrace{\varphi_i(x_r)}_{\geq 0} \geq 0$$

for $r = 1, \ldots, m$. Furthermore, $\psi_{ij}(y) = 0$ by definition, and for the φ_k, there is in fact nothing to show.

Now consider the converse inclusion, and let $\varphi \in \tilde{C}^*$. Then $\varphi(y) \geq 0$ and $\varphi \in C^*$, whence

$$\varphi = a_1\varphi_1 + \cdots + a_t\varphi_t$$

for some $a_1, \ldots, a_t \in \mathbb{R}_+$. We may assume that $a_l > 0$ for all l. If $\varphi_l(y) \geq 0$ for $l = 1, \ldots, t$, then we are done. Therefore assume that $\varphi_l(y) < 0$ for some l, say $\varphi_t(y) < 0$. We will replace the summand $a_t\varphi_t$ in the above representation of φ by a suitable linear combination of the ψ_{ij} and the φ_k (with non-negative real coefficients). In order to enable an inductive argument, we assume right away that

$$\varphi = a_1\varphi_1 + \cdots + a_t\varphi_t + \sum b_{ij}\psi_{ij}$$

with $a_l > 0$ and $b_{ij} \geq 0$. Since $\varphi(y) \geq 0$ and $\varphi_t(y) < 0$, there must exist some φ_l such that $\varphi_l(y) > 0$. Let exactly $\varphi_1, \ldots, \varphi_s$ satisfy this inequality, and consider the linear forms

$$\psi_{it} = \varphi_i(y) \cdot \varphi_t - \varphi_t(y) \cdot \varphi_i$$

for $i = 1, \ldots, s$. We may substitute

$$\varphi_t = \frac{1}{\varphi_i(y)} \cdot (\psi_{it} + \varphi_t(y) \cdot \varphi_i)$$

in the representation of φ.

But since it is not clear which i to choose for the substitution, we do it slightly more generally: consider $c_1, \ldots, c_s \in \mathbb{R}_+$ such that $\sum_i c_i = 1$; then

$$\varphi_t = \sum_{i=1}^{s} c_i\varphi_t = \sum_{i=1}^{s} \frac{c_i}{\varphi_i(y)} \cdot (\psi_{it} + \varphi_t(y) \cdot \varphi_i).$$

Substituting this in the representation of φ yields new coefficients

$$b'_{it} = b_{it} + \frac{c_i a_t}{\varphi_i(y)} \geq b_{it}$$

for the ψ_{it}, and

$$a_i' = a_i + \frac{c_i a_t \varphi_t(y)}{\varphi_i(y)}$$

for the φ_i, $i = 1, \ldots, s$. We have $a_i' \geq 0$ if and only if

$$c_i \leq -\frac{a_i \varphi_i(y)}{a_t \varphi_t(y)};$$

hence the c_i can be chosen suitably if and only if

$$-\sum_{i=1}^{s} \frac{a_i \varphi_i(y)}{a_t \varphi_t(y)} \geq 1.$$

But this inequality follows immediately from

$$0 \leq \varphi(y) = a_1 \varphi_1(y) + \cdots + a_t \varphi_t(y) \leq a_1 \varphi_1(y) + \cdots + a_s \varphi_s(y) + a_t \varphi_t(y). \qquad \square$$

Note that some of the generators of the dual cone may be superfluous. However, for our current purposes, it is not necessary to find a minimal system of generators; we are content with a finite one. For a discussion about the minimal system, we refer the reader to Subsection 5.1 and, in particular, Lemma 5.1.2. Also, a geometric explanation of the algorithm can be found there. (Of course, also some generators of the original cone may be superfluous; but the algorithm above evidently detects such a generator y.)

The following proposition now justifies calling C^* the 'dual' cone.

Proposition 2.9.2: *Let C be a cone in the real vector space V.*

(a) *The bidual cone C^{**} is the (standard) topological closure of C in V (if we identify V with its bidual V^{**} via the natural isomorphism).*

(b) *If C is finitely generated, then so is C^*; moreover, $C^{**} = C$. If, in addition, C is also rational, then so is C^*.*

Proof: (a) Since linear forms are continuous, the standard topological closure $\bar{C}$ is contained in $C^{**} = \{x \in V : \varphi(x) \geq 0 \text{ for all } \varphi \in C^*\}$.

For the converse inclusion, consider $x \in C^{**}$. Evidently, $\bar{C}$ is again a cone; in particular, it is a convex set. If $x \notin \bar{C}$, then the Hahn-Banach Separation Theorem (Lemma 2.3.3) yields a linear form φ such that $\varphi(x) < 0$ and $\varphi(y) \geq 0$ for all $y \in \bar{C}$. (Note that supporting hyperplanes of cones are linear subspaces, by Lemma 2.8.1.) But then $\varphi \in C^*$, and we obtain a contradiction.

(b) If C is finitely generated, then C is closed, and therefore $C^{**} = C$ by (a). The remaining statements are immediate consequences of Proposition 2.9.1 (which provides an inductive argument). $\square$

Part (b) of the proposition can be stated as follows: let C be a finitely generated cone, then C^* is also finitely generated, say

$$C^* = \mathbb{R}_+(\varphi_1, \ldots, \varphi_s).$$

It follows that

$$C = C^{**} = \{x \in V : \varphi_i(x) \geq 0 \text{ for all } i = 1, \ldots, s\}\,.$$

Summing up, we have just seen that finitely generated cones are polyhedral sets; even more: they can be represented as an intersection of linear halfspaces. (A halfspace $H^+ \subseteq V$ is called *linear* if $0 \in H$.) Such cones will be called *linearly polyhedral cones* for a moment. As one expects, there is no need to distinguish between (ordinary) polyhedral cones and linearly polyhedral cones, and we will use this notion only in the following corollary.

Corollary 2.9.3: *For a cone C in V, the following statements are equivalent:*
(i) *C is finitely generated.*
(ii) *C is linearly polyhedral.*
(iii) *C is polyhedral.*
(iv) *C is closed, and C^* is finitely generated.*
In this case, C is rational if and only if C^* is rational.

Proof: The implications (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iv) have been shown above, and the implication (ii) $\Rightarrow$ (iii) is clear.

For (iii) $\Rightarrow$ (ii), assume that

$$C = H_1^+ \cap \cdots \cap H_s^+$$

is an irreducible representation of C, with halfspaces $H_i^+ \subseteq V$. Since $0 \in C$, $\operatorname{aff} C$ is a linear subspace in V, and we may assume that $\dim C = \dim V$ is maximal. But then the H_i are supporting hyperplanes, by Proposition 2.6.2, and by Lemma 2.8.1, they are linear subspaces in V.

Finally, we prove the implications (ii) $\Rightarrow$ (iv) $\Rightarrow$ (i). By hypothesis, there is a (not necessarily irreducible) representation

$$C = H_1^+ \cap \cdots \cap H_s^+\,,$$

with linear halfspaces $H_i^+ = \{x \in V : \varphi_i(x) \geq 0\}$, defined via linear forms $\varphi_i \in V^*$. Now consider the subcone $C' := \mathbb{R}_+(\varphi_1, \ldots, \varphi_s)$ in V^*; its dual cone is

$$\begin{aligned}(C')^* &= \{x \in V : \varphi(x) \geq 0 \text{ for all } \varphi \in C'\}\\ &= \{x \in V : \varphi_i(x) \geq 0 \text{ for all } i = 1, \ldots, s\} = C\,,\end{aligned}$$

the original cone C. Dualizing once more yields $(C')^{**} = C^*$. Since C' is finitely generated by definition, we find that $C' = (C')^{**} = C^*$, by Proposition 2.9.2 (b). Hence C^* is finitely generated, and by the same proposition, C^{**} is also finitely generated. But $C^{**} = C$, by part (a) of this proposition. (Note that C is closed, being a polyhedral set.)

The additional statement on rationality is clear. □

The corollary indicates why finitely generated cones appear in linear optimization, where constraints are given by linear inequalities.

For us, the dual representation of finitely generated cones also has many advantages. For instance, it allows the following immediate conclusions.

Corollary 2.9.4: *Let C and C' be finitely generated cones in V. Then the sum $C+C'$ and the intersection $C\cap C'$ are again finitely generated cones.*

Proof: The statement on the sum follows directly from the definition, and the statement on the intersection follows from Corollary 2.9.3. □

Furthermore, the dual representation offers an efficient method to prove the following lemmas. The first states that halflines running in a 'bad' direction will eventually leave a (possibly shifted) cone.

Lemma 2.9.5: *Let C be a finitely generated cone in V, and let $a,b\in V$, $b\neq 0$. Then the halfline $\ell_+(a,a+b)$ starting in a and running through $a+b$ intersects the cone C*

(a) *in a bounded (possibly empty) segment $[x_1,x_2]$ if $b\notin C$;*
(b) *not at all if there exists a supporting hyperplane H of C such that $a\in \operatorname{int} H^-$ and $b\in C\cap H$ (and $C\subseteq H^+$).*

Proof: (a) The intersection of the two closed convex sets $\ell_+(a,a+b)$ and C is closed and convex itself. Since it is contained in the halfline, it therefore suffices to prove boundness. By Corollary 2.9.3, C is linearly polyhedral, so that there exist linear forms $\varphi_1,\dots,\varphi_m\in V^*$ such that

$$C=\{x\in V:\varphi_i(x)\geq 0 \text{ for } i=1,\dots,m\}.$$

Since $b\notin C$, there is (at least) one linear form among the above, say φ_i, such that $\varphi_i(b)<0$. The points x on the halfline $\ell_+(a,a+b)$ have a representation $x=a+t\cdot b$ with $t\in\mathbb{R}_+$. But then

$$\varphi_i(x)=\varphi_i(a)+t\cdot\varphi_i(b)\longrightarrow -\infty$$

as $t\longrightarrow\infty$, so that the halfline eventually must leave the cone (if it has ever met it).

(b) By Lemma 2.8.1, the hyperplane

$$H=\{x\in V:\varphi(x)=0\}$$

is the kernel of a suitable linear form φ on V. Now everything is easy: by hypothesis, we have $\varphi(a)<0$, $\varphi(b)=0$ and $\varphi(C)\subseteq\mathbb{R}_+$. For any point $x=a+t\cdot b$ on the halfline, this yields

$$\varphi(x)=\varphi(a)+t\cdot\varphi(b)=\varphi(a)<0,$$

so that it is impossible for x to lie in the cone C. □

Lemma 2.9.6: *Let C be a polyhedral cone in V, given by dual forms $\varphi_1,\dots,\varphi_t$. Let $y\in V$, $z\in C$ and $\varphi\in C^*$ be such that $\varphi(y)<0$ and $\varphi(z)=0$. Then also*

$$\varphi_i(y)<0 \quad \text{and} \quad \varphi_i(z)=0$$

for some i.

Proof: This is easy: since C^* is generated by $\varphi_1, \ldots, \varphi_t$, we may, without loss of generality, assume that

$$\varphi = a_1\varphi_1 + \cdots + a_u\varphi_u$$

with $a_i > 0$ for $i = 1, \ldots, u$. Now $\varphi(z) = 0$ implies that $\varphi_i(z) = 0$ for all $i = 1, \ldots, u$, and $\varphi(y) < 0$ implies that $\varphi_i(y) < 0$ for some of these i. □

For the rest of this subsection, let C be a rational polyhedral cone in $V = \mathbb{R}^n$. Assume in addition that C has full dimension (that is, $\dim C = n$), and that the representation

$$C = H_1^+ \cap \cdots \cap H_s^+$$

of C as an intersection of linear halfspaces $H_i^+ = \{x \in V : \varphi_i(x) \geq 0\}$ is irreducible. As mentioned earlier, the (linear) hyperplanes $H_i = \{x \in V : \varphi_i(x) = 0\}$ are uniquely determined by C then: they correspond to the facets of C (Proposition 2.6.2). Moreover, the linear forms φ_i are uniquely determined up to a positive real number, and since they are rational, there is a unique representation with coprime integers. With such a representation, the φ_i are called the *support forms* of C. These considerations lead to the following

Lemma 2.9.7: *Let C be an n-dimensional rational polyhedral cone in $\mathbb{R}^n$, and let $\sigma_1, \ldots, \sigma_s$ be its support forms. Then*

$$C = \bigcap_{i=1}^{s} \{x \in \mathbb{R}^n : \sigma_i(x) \geq 0\},$$

and this representation is irreducible.

2.10. Positive cones. Let C be a cone in the finite dimensional real vector space V. We call C *positive* if the origin 0 is the only invertible element in C. In other words, this means that C does not contain a whole line (through the origin). Yet in other words, positivity of C is characterized by the fact that the *group of units*, that is,

$$C_0 := \{x \in C : -x \in C\},$$

is the trivial group. Note that C_0 is the unique maximal $\mathbb{R}$-vector subspace of C.

Now let C be an n-dimensional rational polyhedral cone in $\mathbb{R}^n$, given as an intersection

$$C = \{x \in V : \sigma_i(x) \geq 0 \text{ for all } i = 1, \ldots, m\},$$

where $\sigma_1, \ldots, \sigma_m \in C^*$ are the support forms of C. They define a linear map $\sigma = (\sigma_1, \ldots, \sigma_m) : \mathbb{R}^n \to \mathbb{R}^m$ whose kernel is

$$\ker \sigma = \{x \in \mathbb{R}^n : \sigma_i(x) = 0 \text{ for all } i\} = C_0,$$

the group of invertible elements in C.

This proves already part of the following lemma.

Lemma 2.10.1: *Let C be an n-dimensional rational polyhedral cone in $\mathbb{R}^n$. Then the following statements are equivalent:*

(i) *C is positive.*
(ii) $C_0 = \{0\}$.
(iii) *σ is injective.*
(iv) *C^* has dimension n.*
(v) *0 is a vertex of C.*

Proof: It only remains to connect (iv) and (v) with the other statements.

First we show (ii) $\Leftrightarrow$ (iv). Let $x \in V$. It is clear that both x and $-x$ lie in C if and only if $\varphi(x) = 0$ for all $\varphi \in C^*$. (Recall that $C^{**} = C$, by Proposition 2.9.2.) Now such an element $x \neq 0$ exists if and only if $\dim C^* < n$.

Now we show (iii) $\Rightarrow$ (v). Let H be the hyperplane in $\mathbb{R}^n$, given by (the vanishing of) the linear form $\varphi := \sigma_1 + \cdots + \sigma_m$. Then $C \subseteq H^+$ and

$$C \cap H = \ker \sigma = \{0\}\,.$$

Finally, let us prove (v) $\Rightarrow$ (ii). Let H be a supporting hyperplane of C, given by a linear form φ, such that $C \subseteq H^+ = \{x : \varphi(x) \geq 0\}$ and $C \cap H = \{0\}$. Now let $x \in \mathbb{R}^n$ be such that both x and $-x$ lies in C. Then $\varphi(x) = 0$ and $x \in H$, and so $x = 0$. □

Condition (v) explains why positive cones are also called ‘pointed cones’, or ‘cones with apex 0’, by some authors. As an immediate consequence of the lemma, we note that, for a polyhedral cone C of maximal dimension, the dual cone C^* is always positive (since $C^{**} = C$).

For a positive rational polyhedral cone C (of full dimension), the mapping $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ arising from the support forms of C defines an embedding $C \hookrightarrow \mathbb{R}^m_+$; hence σ (or more precisely, its restriction to C) is called the *standard embedding* of C. To a certain degree, this observation also explains the notion of positivity: via this embedding, C can be viewed as a subcone of the positive orthant $\mathbb{R}^m_+$.

Our next goal in this subsection is to associate a polytope Q with a positive polyhedral cone $C \subseteq V$. (V denotes a real vector space of finite dimension n.) For this purpose, we may again assume that C has maximal dimension, that is, $\dim C = n$. The construction of Q is established via the linear form $\varphi = \sigma_1 + \cdots + \sigma_m$ which has already appeared in the above proof. Note that $\varphi \neq 0$; otherwise we would have $\sigma_i(x) = 0$ for all i and all $x \in C$, which would imply $\sigma_i = 0$ for all i, since $\dim C = n$.

This time we consider the hyperplane $H := \{x \in V : \varphi(x) = 1\}$. (In the above proof, we have taken the hyperplane parallel to H, through the origin 0.) It is clear that the intersection

$$Q := C \cap H$$

is a polyhedral set. But even more: for any $x \in Q$, we have $\sigma_i(x) \geq 0$ for all i and $\varphi(x) = 1$, which implies that $0 \leq \sigma_i(x) \leq 1$ for all i. It follows that the monomorphism σ sends Q into the unit cube $[0,1]^m$ (in $\mathbb{R}^m$). This can only happen if Q is bounded.

Let us summarize our considerations.

Lemma 2.10.2: *Let C be a positive polyhedral cone in V. Then there exists a hyperplane H in V, not containing the origin* 0*, and such that the intersection $Q = C \cap H$ is a (non-empty) convex polytope. If, in addition, C is rational, then H can be chosen rational, so that Q becomes a rational convex polytope.*

In the situation of the lemma, the hyperplane H is called *transversal* to C, and the polytope Q is called a *cross-section* of C. Cross-sections are introduced because their facial structure is closely related to the facial structure of C (as one can imagine).

We start with a preparatory lemma.

Lemma 2.10.3: (a) *There is a bijection between the set*

$$\mathscr{H} = \{\mathbb{R}_+ x : x \in C,\ x \neq 0\}$$

of halflines in C and the polytope Q, given by the assignment $\mathbb{R}_+ x \longmapsto Q \cap \mathbb{R}_+ x$.

(b) *For any subset $X \subseteq Q$, we have*

$$\mathbb{R}_+ X = \mathbb{R}_+(\operatorname{conv} X) = \bigcup_{x \in \operatorname{conv} X} \mathbb{R}_+ x\,.$$

(c) $C = \bigcup_{x \in Q} \mathbb{R}_+ x$.

Proof: (a) For any $x \in C, x \neq 0$, we must have $\sigma_i(x) > 0$ for at least one i. Therefore there is precisely one value of $t \in \mathbb{R}_+$ such that $\varphi(tx) = 1$. So the mapping

$$r : \mathscr{H} \to Q\,, \quad \mathbb{R}_+ x \longmapsto Q \cap \mathbb{R}_+ x\,,$$

is well-defined, and the mapping

$$e : Q \to \mathscr{H}\,, \quad x \longmapsto \mathbb{R}_+ x\,,$$

obviously behaves inversely to r.

(b) The first equality is clear, by definition of the two sets. It is also clear that the third set is contained in $\mathbb{R}_+(\operatorname{conv} X)$. In order to prove the remaining inclusion, choose a point

$$x = a_1 x_1 + \cdots + a_m x_m$$

in $\mathbb{R}_+ X$ (with $a_i \in \mathbb{R}_+$ and $x_i \in X$). If $a := \sum a_i = 0$, then $x = 0$; otherwise, we find that $x_0 := x/a$ is contained in $\operatorname{conv} X$, and $x = a x_0$ is contained in the halfline $\mathbb{R}_+ x_0$.

(c) is just a consequence of (a) and (b). □

Proposition 2.10.4: *Let C be a positive polyhedral cone in V, and let Q be a cross-section of C. Then the mappings*

$$r : \mathscr{F}(C) \smallsetminus \{\emptyset\} \longrightarrow \mathscr{F}(Q), \quad r(F) := F \cap Q,$$
$$e : \mathscr{F}(Q) \longrightarrow \mathscr{F}(C) \smallsetminus \{\emptyset\}, \quad e(f) := \mathbb{R}_+(f),$$

are mutually inverse isomorphisms of partially ordered sets. Moreover, $\dim r(F) = \dim F - 1$ *for all faces $F \neq \emptyset$ of C.*

Proof: First we show that the 'restriction map' r is well-defined, that is, $F \cap Q$ is a face of Q for any face F of C. This is easy: let H be a hyperplane in V such that $C \subseteq H^+$ and $C \cap H = F$. Then $h := H \cap \operatorname{aff} Q$ is a hyperplane in $\operatorname{aff} Q$ (since $\operatorname{aff} Q$ is a hyperplane in V not containing 0), and it follows from $C \subseteq H^+$ that $Q \subseteq h^+$, by intersecting with $\operatorname{aff} Q$. Furthermore, the intersection

$$Q \cap h = Q \cap H = Q \cap H \cap C = Q \cap F$$

leads to $r(F)$, whence h supports Q at the face $r(F)$.

The next point is the well-definedness of e, the 'extension map': choose a face f of Q, and a hyperplane h in $\operatorname{aff} Q$ such that $Q \subseteq h^+$ and $Q \cap h = f$. Then, clearly, 0 is not contained in h, and the affine hull

$$H := \operatorname{aff}(h \cup \{0\}) = \mathbb{R}(h \cup \{0\}) = \mathbb{R}(h)$$

is a linear hyperplane in V; in particular, $\dim H = \dim h + 1$. Note also that the intersection $H \cap Q = h$ leads back to h. Then, since $Q \subseteq h^+ \subseteq H^+$ (we *define* H to induce this orientation), it is an immediate consequence of Lemma 2.10.3 that C is contained in the halfspace H^+. Moreover, $\mathbb{R}_+(f)$ is contained in $C \cap H$, since f is contained there. For the other inclusion, choose $x \in C \cap H$; without loss of generality, we may in addition assume that $x \in Q$. But then $x \in f$. Summing up, we have shown that H supports C at the face $e(f) = \mathbb{R}_+(f)$.

The remaining statements of the proposition are clear: using Lemma 2.10.3, we see that e and r are mutually inverse, and as for the statement on the dimensions, this amounts to showing that

$$\dim(\mathbb{R}f) = \dim(\operatorname{aff} f) + 1$$

for all faces f of Q. But this follows from the fact that $\operatorname{aff} f \subseteq \operatorname{aff} Q$ does not contain the origin. □

We finish this subsection by discussing minimal 'internal' and 'external' representations of a positive polyhedral cone: are they uniquely determined and how do we find them? To some extent, the following proposition (and its proof) is a summary of the most important statements of this section. As one expects, it refers to the one-dimensional and the one-codimensional faces of C (called 'edges' and 'facets', respectively). Recall that C has only a finite number of faces (Lemma 2.7.1).

Proposition 2.10.5: *Let C be a polyhedral cone in V.*

(a) *Suppose that C is positive. Then a subset X of C is a minimal generating set of C if and only if X contains exactly one vector $x_E \neq 0$ from each edge E of C (and no other vectors).*

(b) (Weyl's Theorem) *Suppose that* $\dim C = \dim V$. *Then there is exactly one irreducible representation of C as an intersection of halfspaces, namely the one where each halfspace H_F^+ is associated with a facet F of C (by requiring that the bounding hyperplane H_F of H_F^+ is the affine hull of F).*

Proof: (a) It is clear that any element of a generating set may be replaced by a positive real multiple. Therefore, we may choose the generators in a cross-section Q of C. So let X be a subset of Q. We now characterize when X generates the whole cone C: by Lemma 2.10.3, this happens if and only if

$$C = \bigcup_{x \in Q} \mathbb{R}_+ x = \bigcup_{x \in \operatorname{conv} X} \mathbb{R}_+ x,$$

and this in turn is equivalent to the condition that $Q = \operatorname{conv} X$. Now the theory of convex polytopes comes into play: by Lemma 2.5.4, this equality holds if and only if $\operatorname{vert} Q \subseteq X$. Clearly, the unique minimal choice for this is $X = \operatorname{vert} Q$. Recall now that the vertices of Q correspond to the edges of C, by Proposition 2.10.4. This proves part (a).

(b) is just a repetition of Proposition 2.6.2. □

2.11. Simplicial cones and triangulations. In this subsection we introduce the notion of a triangulation of a finitely generated cone C. Its basic building blocks are the simplicial cones. Triangulations play a crucial role for the computation of the normalization of an affine monoid; see Section 5.

A cone δ in an $\mathbb{R}$-vector space V is called *simplicial* if it is generated by linearly independent vectors. As an immediate consequence, we see that simplicial cones are always positive.

Moreover, the faces of a simplicial cone δ are extremely easy to determine: suppose that δ is generated by n linearly independent vectors $x_1, \dots, x_n$ in $\mathbb{R}^n$. As outlined previous to Proposition 2.9.1, the dual basis $\varphi_1, \dots, \varphi_n$ of the basis $x_1, \dots, x_n$ generates the dual cone δ^* of δ. In fact, each dual form φ_i leads to a facet

$$\delta_i = \{x \in \delta : \varphi_i(x) = 0\} = \mathbb{R}_+(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

of δ, which is again a simplicial cone. Proceeding by induction, one obtains the following lemma.

Lemma 2.11.1: *Let δ be a simplicial cone, generated by linearly independent vectors $x_1, \dots, x_n$. Then the faces of δ are precisely the simplicial cones $\mathbb{R}_+(E)$, E running through the subsets of $X := \{x_1, \dots, x_n\}$. In particular, there is a bijection between the face lattice $\mathscr{F}(\delta)$ of δ and the power set $\mathscr{P}(X)$ of X.*

Now we introduce the main notion of this subsection: a *triangulation* of a cone C in V is a family Δ of finitely many simplicial subcones such that
(1) $C = \bigcup_{\delta \in \Delta} \delta$,
(2) the faces of each $\delta \in \Delta$ are themselves members of Δ, and
(3) the intersection of each pair $\delta, \varepsilon \in \Delta$ is a face of both δ and ε.
Due to (2), it evidently suffices to know the maximal subcones of Δ. The reader may recall that a family Δ of finitely many simplices which satisfies only (2) and (3) is known as a *simplicial complex* in algebraic topology.

Before clarifying the existence of triangulations for certain cones (namely, the finitely generated ones), we discuss some fundamental properties of triangulations. For instance, we obtain a disjoint decomposition of the cone, which will be an important fact when computing the Hilbert function of affine monoids in Subsection 5.2.

Lemma 2.11.2: *Let Δ be a triangulation of the cone C. Then the following hold:*
(a) *For each $\delta \in \Delta$, we have*

$$\delta = \biguplus_{\substack{\varepsilon \in \Delta \\ \varepsilon \subseteq \delta}} \operatorname{relint} \varepsilon \, .$$

(b) *C is the disjoint union*

$$C = \biguplus_{\delta \in \Delta} \operatorname{relint} \delta \, .$$

Proof: (a) By Corollary 2.7.6, δ is the disjoint union of all its faces $\varepsilon \in \mathscr{F}(\delta)$, and by the definition of a triangulation, we have

$$\mathscr{F}(\delta) = \{\varepsilon \in \Delta : \varepsilon \subseteq \delta\} \, .$$

(For ε in the right-hand set, consider the intersection with δ.)

(b) It follows from (a) that C is the union of the relative interiors of all $\delta \in \Delta$; it remains to show that this union is disjoint. So let x be a common point of $\operatorname{relint} \delta$ and $\operatorname{relint} \varepsilon$ for some cones δ, ε in Δ. Then x is contained in $\delta \cap \varepsilon$, and by (3), this intersection is a face of δ (and ε). Since δ is the smallest face of δ which contains x (Lemma 2.4.6), we find that δ is contained in $\delta \cap \varepsilon$; therefore $\delta \cap \varepsilon = \delta$. Analogously, one finds that $\delta \cap \varepsilon = \varepsilon$. □

Another nice property of a triangulation of a cone C is that it also induces a triangulation of the faces of C:

Lemma 2.11.3: *Let Δ be a triangulation of the cone C, and F a face of C. Then*

$$F = \bigcup_{\substack{\delta \in \Delta \\ \delta \subseteq F}} \delta .$$

Proof: We may assume that F is a proper face of C; so write $F = C \cap H$ with a supporting hyperplane H of C. Using property (1) of the triangulation, this leads to

$$F = \bigcup_{\delta \in \Delta} (\delta \cap H) .$$

Now, if $\delta \cap H$ is non-empty, then H supports δ, and $\delta \cap H$ is in Δ, being a face of δ. From this, the assertion follows. □

We now turn our attention to the main result of this subsection. Let C be a finitely generated cone in $\mathbb{R}^n$. Proposition 2.9.1 is a constructive algorithm for computing the dual forms of C, that is, generators of the dual cone C^*. We shall develop an analogous algorithm for computing a triangulation of C. (In fact, this relies on the Dual Cone Algorithm.)

We start just as previous to the proposition: we may assume that $\dim C = n$, and can then find n linearly independent vectors $x_1, \ldots, x_n$ among the generators of C. These vectors generate a simplicial subcone C_0 of C and form a basis of $\mathbb{R}^n$; the dual basis $\varphi_1, \ldots, \varphi_n$ generates the dual cone of C_0 and, therefore, leads to an external representation

$$C_0 = \{x \in \mathbb{R}^n : \varphi_i(x) \geq 0 \text{ for } i = 1, \ldots, n\}$$

of C_0. Moreover, a triangulation of C_0 is obviously given by $\Delta_0 := \mathscr{F}(C_0)$, the face lattice of C_0. We now describe how to extend the triangulation if we enlarge a cone C by another generator.

Proposition 2.11.4: *Let C be an n-dimensional cone in $\mathbb{R}^n$, given internally by generators $x_1, \ldots, x_m$ and externally by dual forms $\varphi_1, \ldots, \varphi_t$. Let Δ be a triangulation of C. Then, for $y \in \mathbb{R}^n$, a triangulation of the cone $\tilde{C} = C + \mathbb{R}_+ y$ is given by*

$$\tilde{\Delta} := \Delta \cup \left\{ \delta + \mathbb{R}_+ y : \delta \in \Delta \textit{ and } \varphi_i(\delta) = \{0\},\, \varphi_i(y) < 0 \textit{ for some } i \right\} .$$

We postpone the proof of the proposition; instead, we first draw the desired conclusions and give a geometric explanation of the algorithm.

Corollary 2.11.5: *For each finitely generated cone C, there is a triangulation Δ such that each simplicial subcone $\delta \in \Delta$ is generated by some of the generators of C.*

Note that, if y is already contained in C (so that $\tilde{C} = C$), then the triangulation Δ remains unchanged. We may therefore suppose that $y \notin C$; then the following notion makes sense: a subset X of C is *visible* from y if, for each $x \in X$, the line segment $[y,x]$ from y to x intersects C exactly in x. One now obtains a triangulation of $\tilde{C}$ by joining Δ with the set of all cones $\delta + \mathbb{R}_+ y$ where $\delta \in \Delta$ is visible from y; exactly these cones δ are selected in Proposition 2.11.4.

Using the notion of visibility, we also give a brief sketch of part (1) of the proof (which deals with the covering property of the triangulation): given a point $x \in \tilde{C} \setminus C$, the halfline from y through x meets the cone C in a point z which is visible from y; the visible area of C is triangulated by (a subset of) Δ, and so z is contained in a visible cone $\delta \in \Delta$, whence x is contained in the cone $\delta + \mathbb{R}_+ y$, being a convex combination of y and z. This idea is formalized in the following proof.

Proof of Proposition 2.11.4: First of all, note that $\tilde{\Delta}$ is finite, since Δ is so. Moreover, $\tilde{\Delta}$ again consists of simplicial cones, since each 'new' cone $\delta + \mathbb{R}_+ y$ (if any) involves a simplicial subcone δ which is contained in some hyperplane $H_i = \{x : \varphi_i(x) = 0\}$ (while y is not). Subsequently, we check the three properties from the definition of a triangulation; note that, in any case, we may, and do, assume that $y \notin C$.

(1) By definition, it is clear that each $\tilde{\delta} \in \tilde{\Delta}$ is contained in $\tilde{C}$. So let $x \in \tilde{C}$; we have to find a simplicial subcone $\tilde{\delta} \in \tilde{\Delta}$ which contains x. For this, we may assume that $x \notin C$. Then there is a representation

$$x = a_1 x_1 + \cdots + a_m x_m + by$$

with $a_i \geq 0$ and $b > 0$. Upon defining $z := 2(a_1 x_1 + \cdots + a_m x_m) \in C$ and replacing y with $y/2b$, we have reached a situation where

$$x = \frac{1}{2} y + \frac{1}{2} z$$

is the mid-point of the segment $[y, z]$. Now, since $y \notin C$, $z \in C$ and C is convex, the halfline $\ell_+(y, z)$ starting in y and passing through z enters C in a point z'. Note that x is also a convex combination of y and z' now; since this is the only property of x which will be important at the end, we may (and do) assume that $z = z'$.

By Corollary 2.6.7, there exists a supporting hyperplane H of C which contains z and separates y from C. Let φ be the respective dual form of C; then, with respect to its behaviour on the points y and z, φ may be replaced with one of the dual forms $\varphi_1, \ldots, \varphi_t$, that is, $\varphi_i(y) < 0$ and $\varphi_i(z) = 0$ for some i (Lemma 2.9.6). Let F denote the respective face of C, that is,

$$F = \{v \in C : \varphi_i(v) = 0\}.$$

We then have $z \in F$, and by Lemma 2.11.3, we find some $\delta \in \Delta$, $\delta \subseteq F$, which contains z. Therefore, $\varphi_i(\delta) = \{0\}$ and

$$x \in \operatorname{conv}(y, z) \subseteq \delta + \mathbb{R}_+ y.$$

This shows that $\tilde{C}$ is covered by the simplicial cones in $\tilde{\Delta}$.

(2) Now we show that the faces of a simplicial cone $\tilde{\delta} \in \tilde{\Delta}$ are again in $\tilde{\Delta}$. For $\tilde{\delta} \in \Delta$, this is clear. So suppose that $\tilde{\delta} = \delta + \mathbb{R}_+ y$ for some $\delta \in \Delta$ with $\varphi_i(\delta) = \{0\}$ and $\varphi_i(y) < 0$ (for some i). Now let X be a set of linearly independent generators of δ. Then the faces of $\tilde{\delta}$ are precisely the cones $\mathbb{R}_+(E)$ with a subset E of $X \cup \{y\}$,

by Lemma 2.11.1. But this yields cones of the form δ' or $\delta' + \mathbb{R}_+ y$, where δ' is a face of δ. Thus δ' is in Δ, and each face of $\tilde{\delta}$ is in $\tilde{\Delta}$.

(3) We finally show that, for each pair $\tilde{\delta}, \tilde{\varepsilon} \in \tilde{\Delta}$, the intersection is a face of both $\tilde{\delta}$ and $\tilde{\varepsilon}$. If $\tilde{\delta}, \tilde{\varepsilon} \in \Delta$, then this is the respective property of Δ. So let $\tilde{\delta} = \delta + \mathbb{R}_+ y$ and $\tilde{\varepsilon} = \varepsilon \in \Delta$ (with $\varphi_i(\delta) = \{0\}$ and $\varphi_i(y) < 0$ for some i). Since $\tilde{\delta}$ is simplicial and $\varphi_i(\varepsilon) \subseteq \mathbb{R}_+$, it follows that

$$\tilde{\delta} \cap \tilde{\varepsilon} = \delta \cap \varepsilon \,,$$

and this is a common face of δ and ε, hence also of $\tilde{\delta}$ and $\tilde{\varepsilon}$.

Finally, let $\tilde{\delta} = \delta + \mathbb{R}_+ y$ and $\tilde{\varepsilon} = \varepsilon + \mathbb{R}_+ y$ (with $\delta, \varepsilon \in \Delta$, $\varphi_i(\delta) = \{0\}$, $\varphi_i(y) < 0$, $\varphi_j(\varepsilon) = \{0\}$ and $\varphi_j(y) < 0$ for some i, j). We claim that

$$\tilde{\delta} \cap \tilde{\varepsilon} = (\delta \cap \varepsilon) + \mathbb{R}_+ y \,,$$

by which everything is proven. The inclusion '$\supseteq$' is clear, the main point being the reverse inclusion. So choose $x \in \tilde{\delta} \cap \tilde{\varepsilon}$, say

$$x = x_\delta + ay = x_\varepsilon + by$$

with $x_\delta \in \delta$, $x_\varepsilon \in \varepsilon$ and $a, b \in \mathbb{R}_+$. Evaluating the linear forms φ_i and φ_j on x, we obtain the equations

$$a \cdot \varphi_i(y) = \varphi_i(x_\varepsilon) + b \cdot \varphi_i(y) \quad \text{and} \quad \varphi_j(x_\delta) + a \cdot \varphi_j(y) = b \cdot \varphi_j(y) \,.$$

But now, dividing by the negative numbers $\varphi_i(y)$ and $\varphi_j(y)$, respectively, yields that

$$a = b + \frac{\varphi_i(x_\varepsilon)}{\varphi_i(y)} \leq b \quad \text{and} \quad b = a + \frac{\varphi_j(x_\delta)}{\varphi_j(y)} \leq a \,;$$

hence $a = b$ and $x_\delta = x_\varepsilon \in \delta \cap \varepsilon$. □

3. Affine monoids and their algebras

In this section we finally introduce the central objects of our work, namely, the affine monoids. They are the discrete analogues of finitely generated cones. In fact, we shall also establish the connection between these two classes (Gordan's Lemma). But at least as important is the bridge between affine monoids and their algebras; it opens the gate to the essentially richer theory of rings, and from there, we shall transfer a number of results to the theory of monoids.

We end the section by investigating positive affine monoids. We shall work out the correspondence with positive polyhedral cones, and via the standard embedding, we shall (affirmatively) answer the question about existence and uniqueness of a minimal generating system, the Hilbert basis.

As an alternative source for the topics of this section, we recommend the books by Bruns and Herzog [BH] and Ewald [Ew].

3.1. Affine monoids. We start by introducing some notions. Finitely generated free Abelian groups are called *lattices* in the following. Of course, these are exactly the ones which are isomorphic to $\mathbb{Z}^n$ for some $n \in \mathbb{N}$ (by the structure theorem for finitely generated Abelian groups). An *affine monoid* is a finitely generated submonoid S of some lattice L; *finitely generated* means that there exist finitely many elements $v_1, \dots, v_m \in S$ such that the *monoid generated by* $v_1, \dots, v_m$, which we denote by

$$\mathbb{N}(v_1, \dots, v_m) := \mathbb{N}v_1 + \cdots + \mathbb{N}v_m\,,$$

coincides with S. We emphasize that the adjective 'affine', when used in conjunction with monoids and submonoids, always includes the property of being finitely generated.

It is easy to see that submonoids of $\mathbb{Z}$ are always affine.

Lemma 3.1.1: *Let S be a submonoid of $\mathbb{Z}$. Then S is a union of finitely many arithmetic progressions, that is, there are finitely many $a_i, b_i \in S$, $i = 1, \dots, r$, such that*

$$S = \bigcup_{i=1}^{r} (a_i + \mathbb{N}b_i)\,.$$

In particular, S is affine.

Proof: We may assume that $S \neq 0$.

(1) If S contains (strictly) positive and negative integers, then S is a subgroup of $\mathbb{Z}$: choose $a = \min\{n > 0 : n \in S\}$ and $b = \max\{n < 0 : n \in S\}$ and consider $a + b$; it turns out that $a = -b$. So $S = \mathbb{Z}a$ in this case.

(2) Now assume that S is contained in $\mathbb{N}$. In this case, we fix some $d \in S$, $d > 0$. In every residue class $a + \mathbb{Z}d$, $a = 0, \dots, d-1$, one then chooses $d_a =$

$\min S \cap (a + \mathbb{Z}d)$. This results in a representation

$$S = \bigcup_{a=0}^{d-1} (d_a + \mathbb{N}d)$$

(where the union is even disjoint).

(3) If $S \subseteq \mathbb{Z}_-$, then $-S \subseteq \mathbb{N}$, which leads back to the discussion in step (2). $\square$

Note that submonoids of $\mathbb{Z}^2$ need no longer be finitely generated: consider, for example, the subset

$$S := \{(0,0)\} \cup \{(x,y) \in \mathbb{N}^2 : y > 0\};$$

clearly, S is a submonoid, but is not finitely generated.

Now let S be an affine submonoid in the lattice L. We write $\mathbb{Z}S$ for the *group generated by* S (within L):

$$\mathbb{Z}S = \left\{ \sum_{i=1}^{m} a_i v_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{Z},\ v_i \in S \right\}.$$

Note that $\mathbb{Z}S$ fulfils a universal property which characterizes it up to isomorphism: every monoid homomorphism from S to a group can be extended to $\mathbb{Z}S$ in a unique way. From this point of view, $\mathbb{Z}S$ is the group of differences of S. Any element $x \in \mathbb{Z}S$ can be represented as $x = s - t$ for some $s,t \in S$. If S is also contained in a lattice L', then, by the universal property, the groups $\mathbb{Z}S$ and $(\mathbb{Z}S)'$ generated by S in L and L', respectively, are naturally isomorphic. Furthermore, since L is finitely generated free Abelian, we have $\mathbb{Z}S \cong \mathbb{Z}^r$ for some r (uniquely determined by S). r is called the *rank* of S:

$$\operatorname{rk} S := \operatorname{rk} \mathbb{Z}S.$$

We introduce abbreviations for the $\mathbb{Q}$- and the $\mathbb{R}$-vector spaces generated by S (within $\mathbb{R}^n$ if $S \subseteq \mathbb{Z}^n$):

$$\mathbb{Q}S := \left\{ \sum_{i=1}^{m} a_i v_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{Q},\ v_i \in S \right\} \cong \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}S,$$

$$\mathbb{R}S := \left\{ \sum_{i=1}^{m} a_i v_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{R},\ v_i \in S \right\} \cong \mathbb{R} \otimes_{\mathbb{Z}} \mathbb{Z}S.$$

As above, we prefer this ‘explicit’ point of view. The construction is, of course, independent of the chosen lattice L (in which S is contained); in fact, $\mathbb{Q}S$ and $\mathbb{R}S$ again satisfy universal properties which determine them up to isomorphism. By construction, we have natural inclusions

$$S = \mathbb{N}S \subseteq \mathbb{Z}S \subseteq \mathbb{Q}S \subseteq \mathbb{R}S$$

and equalities

$$r = \operatorname{rk} S = \operatorname{rk} \mathbb{Z}S = \dim_{\mathbb{Q}} \mathbb{Q}S = \dim_{\mathbb{R}} \mathbb{R}S.$$

So S is a submonoid of $\mathbb{Z}S$, $\mathbb{Z}S$ is a subgroup of $\mathbb{Q}S$, and $\mathbb{Q}S$ is a $\mathbb{Q}$-vector subspace of $\mathbb{R}S$.

Let S further on be an affine submonoid of the lattice L. An element $x \in L$ is called *integral* over S if $c \cdot x \in S$ for some $c \in \mathbb{N}$, $c > 0$. The set of all such x is called the *integral closure* $\bar{S}_L$ of S in L. Clearly $S \subseteq \bar{S}_L$. In the case where equality $S = \bar{S}_L$ holds, we say that S is *integrally closed* in L. Obviously, $\bar{S}_L$ is again a submonoid of L. As we shall see in Lemma 3.1.2, it is even finitely generated (since S is so), and can be described geometrically.

We introduce special terminology for the lattice $L = \mathbb{Z}S$: in this case, the integral closure $\bar{S} = \bar{S}_{\mathbb{Z}S}$ is called the *normalization* of S, and S is *normal* if $S = \bar{S}$. Note that there is no dependency on the lattice in which S is originally contained.

Let S be a submonoid of the lattice L. Then the *associated cone*, that is, the cone

$$C(S) := \mathbb{R}_+ S = \left\{ \sum_{i=1}^{m} a_i v_i \;\middle|\; m \in \mathbb{N},\ a_i \in \mathbb{R}_+,\ v_i \in S \right\}$$

generated by S in the real vector space $V = \mathbb{R}L$, is always rational. It is finitely generated if S is so. Note that the converse is also true; see Corollary 3.3.7.

Lemma 3.1.2: *Let $L = \mathbb{Z}^n$ be the standard lattice.*

(a) (Gordan's Lemma) *Let $C \subseteq \mathbb{R}^n$ be a finitely generated rational cone. Then the* associated semigroup $S(C) := L \cap C$ *is an affine monoid, and is integrally closed in L.*

(b) *Conversely, let S be an affine submonoid of L. Then $\bar{S}_L = L \cap C(S)$, and $\bar{S}_L$ is again an affine monoid.*

Proof: (a) Set $S = S(C)$ for short. By assumption, C is generated by finitely many vectors $v_1, \dots, v_m \in L$. Let $v \in S$. Then $v = a_1 v_1 + \cdots + a_m v_m$ with $a_i \in \mathbb{R}_+$. Therefore we can write

$$v = \sum_{i=1}^{m} a_i' v_i + \sum_{i=1}^{m} a_i'' v_i\,,$$

where $a_i' \in \mathbb{N}$ and $a_i'' \in \mathbb{R}$, $0 \le a_i'' < 1$. We have $v'' := \sum a_i'' v_i \in C \cap L = S$. But v'' also lies in the bounded set $B = \{\sum b_i v_i : 0 \le b_i < 1\}$. Therefore $L \cap B$ is finite, and S is generated by the finite set

$$(L \cap B) \cup \{v_1, \dots, v_m\}\,.$$

That $S = \bar{S}_L$ is quite evident: let $x \in \bar{S}_L$, that is, $x \in L$ and $c \cdot x \in S$ for some $c > 0$. But then $cx \in C$, hence $x \in C \cap L = S$.

(b) Since S is finitely generated, $C(S)$ is a finitely generated rational cone. Therefore (a) applies and yields that the monoid $S' := L \cap C(S)$ is finitely generated and integrally closed in L. But $S \subseteq S'$, hence $\bar{S}_L \subseteq S'$. The converse inclusion follows from the following lemma. In fact, let $S = \mathbb{N}(v_1, \dots, v_m)$ with $v_i \in L$, and let $v \in S'$. Then $v \in \mathbb{Q}_+(v_1, \dots, v_m)$ by the lemma, and multiplying by a common denominator $c > 0$ yields $cv \in S$. □

Lemma 3.1.3: *Let* $v_1, \ldots, v_m \in \mathbb{Z}^n$. *Then*

$$\mathbb{Z}^n \cap \mathbb{R}_+(v_1, \ldots, v_m) \subseteq \mathbb{Q}_+(v_1, \ldots, v_m).$$

Proof: Let $v = a_1 v_1 + \cdots + a_m v_m \in \mathbb{Z}^n$ with $a_i \in \mathbb{R}_+$. Then the (Diophantine) system of linear equations

$$x_1 v_1 + \cdots + x_m v_m = v$$

(with n equations and m indeterminates $x_1, \ldots, x_m$) is soluble over $\mathbb{R}$, hence also soluble over $\mathbb{Q}$ (by Gaussian elimination). Since $\mathbb{Q}$ is dense in $\mathbb{R}$, the rational solution space $L_{\mathbb{Q}}$ is dense in $L_{\mathbb{R}}$, the set of solutions over $\mathbb{R}$. Therefore it is even possible to find a solution $x \in \mathbb{Q}_+^m$. □

Note that, upon replacing $\mathbb{Z}^n$ with $\mathbb{Q}^n$ in Lemma 3.1.3, there is equality

$$\mathbb{Q}^n \cap \mathbb{R}_+(v_1, \ldots, v_m) = \mathbb{Q}_+(v_1, \ldots, v_m).$$

In fact, this is an immediate consequence of Lemma 3.1.3: simply replace the v_i and a vector $v \in \mathbb{Q}^n \cap \mathbb{R}_+(v_1, \ldots, v_m)$ by suitable positive multiples $c_i v_i$ and cv in $\mathbb{Z}^n$.

We end this subsection with a final remark: the monoids $S(C) = L \cap C$ associated with finitely generated rational cones C are always normal (independent of the chosen lattice L, provided that $\mathbb{R}L \subseteq \mathbb{R}C$). Conversely, every normal affine monoid S has such a representation, since $S = S(C(S))$ (within the lattice $L = \mathbb{Z}S$).

3.2. Faces of affine monoids. The aim of this subsection is to extend the notion of a face from polyhedral cones (and closed convex sets in general) to affine monoids.

So let S be an affine monoid, and $C = C(S)$ the cone of S. Clearly, C is rational and polyhedral and has only a finite number of faces (Lemma 2.7.1). The intersection of a face F of C with the monoid S, $F_S := F \cap S$, will henceforth be referred to as a *face* of S. The faces of S are collected in the finite set

$$\mathscr{F}(S) := \{F_S : F \in \mathscr{F}(C)\}.$$

Note that each face of S is a discrete set by definition, and that the *improper faces* $\emptyset$ and S are again included here.

Recall that the faces of a convex polytope are again convex polytopes (Proposition 2.5.2). For polyhedral sets, this is clear by definition, as well as it is for cones and monoids. But what about the property of being finitely generated? The following lemma states that these properties are as well inherited by the faces:

Lemma 3.2.1: *Let S be an affine monoid, generated by the vectors* $v_1, \dots, v_s$*, and C the cone of S. Let F be a non-empty face of C and* F_S *the respective face of S. Assume that, among the generators of S, precisely* $v_1, \dots, v_r$ *are contained in F. Then*

$$F = \mathbb{R}_+(v_1, \dots, v_r) \quad \text{and} \quad F_S = \mathbb{N}(v_1, \dots, v_r).$$

In particular,

(a) *F is a finitely generated rational subcone of C,*
(b) F_S *is an affine submonoid of S,*
(c) *F is the cone of* F_S *and*
(d) *there is a bijection between the faces of C and the faces of S.*

Proof: If F is an improper face of C, then there is nothing to show. So write $F = C \cap H$ with a supporting hyperplane H of C, and let H be given by the vanishing of a linear form φ (Lemma 2.8.1). Now, if an arbitrary point $x = \sum_{i=1}^{s} a_i v_i$, $a_i \in \mathbb{R}_+$, from C is also contained in H, then all the coefficients $a_{r+1}, \dots, a_s$ must be zero.

Literally the same argument holds for S (and F_S). □

We should mention that there is, of course, an equivalent way of introducing the faces of an affine monoid S, namely, along the same lines as for closed convex sets (and cones, in particular): a subset of S is a face of S if and only if it is the intersection of S with a supporting hyperplane of S (where supporting hyperplanes are defined in the usual way). This characterization relies on the fact that supporting hyperplanes of S are linear subspaces (which can be shown just as for cones, see Lemma 2.8.1); therefore, a hyperplane H supports S if and only if it supports $C(S)$.

The natural next step is to define the relative interior of a face F_S of S. We simply set

$$\operatorname{relint} F_S := S \cap \operatorname{relint} F = F_S \cap \operatorname{relint} F$$

(where $F = C(F_S)$ denotes the respective face of the cone $C(S)$, as described in Lemma 3.2.1). Recall that, if F is non-empty, then $\operatorname{relint} F$ is also non-empty (Lemma 2.4.5). The same holds true for F_S:

Lemma 3.2.2: *Let* F_S *be a non-empty face of the affine monoid S. Then the relative interior of* F_S *is also non-empty.*

Proof: Fix the notation from Lemma 3.2.1, so that

$$F = \mathbb{R}_+(v_1, \dots, v_r) \quad \text{and} \quad F_S = \mathbb{N}(v_1, \dots, v_r).$$

It is clear then that the sum $v_1 + \cdots + v_r$ of the generators is a relative interior point of F, and hence also of F_S. □

3.3. Affine monoid algebras. Now let K be a field, and S again a submonoid of a lattice L. (S need not yet be finitely generated at this point.) We write $K[S]$ for the monoid algebra corresponding to S, that is, the K-vector space with basis X^v, $v \in S$. These elements of $K[S]$ are called *monomials*. A general element of $K[S]$ is of the form

$$a_1 X^{v_1} + \cdots + a_m X^{v_m},$$

where $m \in \mathbb{N}$, $a_i \in K$, and $v_i \in S$. The additive structure of $K[S]$ is clear, and the multiplicative structure arises in a natural way from the law for monomials: $X^v \cdot X^{v'} = X^{v+v'}$ (where '$+$' denotes the monoid operation, as usual). For example, $K[\mathbb{Z}^n]$ is isomorphic to the Laurent polynomial ring $K[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]$ if we let X_i correspond to the i-th element e_i of the canonical basis of $\mathbb{Z}^n$; similarly, $K[\mathbb{N}^n]$ is isomorphic to the ordinary polynomial ring $K[X_1, \ldots, X_n]$.

An inclusion $S \subseteq \mathbb{Z}^n$ of monoids induces an inclusion $K[S] \subseteq K[\mathbb{Z}^n]$ of the corresponding K-algebras. Therefore, $K[S]$ is an integral domain. Analogously, the inclusion $S \subseteq \mathbb{Z}S$ yields an inclusion $K[S] \subseteq K[\mathbb{Z}S]$; but then $K[S]$ and $K[\mathbb{Z}S]$ clearly have the same field $\mathscr{K}$ of fractions. It follows that the Krull dimension of $K[S]$ coincides with the rank of S (see [Ma] or [BH] for the notion of Krull dimension):

Lemma 3.3.1: *For a submonoid S of a lattice L, we have* $\dim K[S] = \operatorname{rk} S$.

Proof: This is clear by the above remarks, for if $\mathbb{Z}S \cong \mathbb{Z}^r$, then

$$\dim K[S] = \operatorname{tr.deg}_K \mathscr{K} = \dim K[\mathbb{Z}S] = \dim K[\mathbb{Z}^r] = r. \qquad \square$$

But there are much more connections between S and $K[S]$. We next investigate the property of being finitely generated.

So let $S = \mathbb{N}(v_1, \ldots, v_r)$ be a finitely generated (affine) monoid. Then the corresponding monoid ring is

$$K[S] = \left\{ \sum_{i=1}^{m} a_i X^{b_i} \;\middle|\; m \in \mathbb{N},\ a_i \in K,\ b_i \in S \right\}.$$

We claim that $K[S]$ is a finitely generated K-algebra, namely, we have

$$K[S] = K[X^{v_1}, \ldots, X^{v_r}].$$

The inclusion '$\supseteq$' is clear; for the other inclusion, it suffices to show that $X^v \in K[X^{v_1}, \ldots, X^{v_r}]$ for $v \in S$. But write $v = n_1 v_1 + \cdots + n_r v_r$ with $n_i \in \mathbb{N}$, then

$$X^v = (X^{v_1})^{n_1} \cdots (X^{v_r})^{n_r} \in K[X^{v_1}, \ldots, X^{v_r}].$$

Therefore the property of being finitely generated is transferred from monoids to the corresponding monoid rings. The following proposition says that the converse is also true.

Proposition 3.3.2: *Let S be a submonoid of a lattice L, and $K[S]$ the corresponding monoid ring (where K is a field). Then the following are equivalent:*

(i) *S is a finitely generated monoid.*

(ii) *$K[S]$ is a finitely generated K-algebra.*

Proof: The implication (i) $\Rightarrow$ (ii) has been shown above. So let $K[S]$ be generated as a K-algebra by finitely many (finite) sums $\sum a_i X^{b_i}$. Then $K[S]$ clearly is as well generated by all monomials $X^{v_1},\dots,X^{v_r}$ occurring in all these sums. We claim that $S=\mathbb{N}(v_1,\dots,v_r)$, hence S is finitely generated. In order to see this, choose $v\in S$. Then there exists a polynomial $f\in K[X_1,\dots,X_r]$ such that

$$X^v = f(X^{v_1},\dots,X^{v_r}).$$

Since all the monomials X^w, $w\in S$, are K-linearly independent (by construction of $K[S]$), the right-hand side of this equation must be reducible to a single summand

$$f(X^{v_1},\dots,X^{v_r}) = a\cdot (X^{v_1})^{n_1}\cdots (X^{v_r})^{n_r}$$

(with $a\in K$, $n_i\in\mathbb{N}$). But then $a=1$ and $v=n_1v_1+\dots+n_rv_r\in\mathbb{N}(v_1,\dots,v_r)$. □

The algebras $K[S]$, where K is a field and S is an affine monoid, are called *affine monoid algebras* (or *affine monoid rings*). Proposition 3.3.2 states, in particular, that these algebras are always finitely generated.

Corollary 3.3.3: *Affine monoid algebras are Noetherian integral domains.*

For further properties of finitely generated K-algebras, see Subsection 4.1 (in particular, Proposition 4.1.3).

Our next goal is to establish the link between the normality of an affine monoid S and the normality of the corresponding affine monoid algebra $K[S]$. As we shall see (Proposition 3.3.4), these properties are equivalent, too. We start with a reminder of the respective notions from the theory of rings.

Let $R\subseteq R'$ be two integral domains. As is generally known, an element $x\in R'$ is called *integral over R* if there exists a monic polynomial $f\in R[X]$ such that $f(x)=0$. Those elements in R' which are integral over R form a subring $\bar{R}_{R'}$; it is called the *integral closure* of R in R'. In the particular case when $R'=\operatorname{Frac}R$ is the field of fractions of R, the integral closure is referred to as the *normalization* $\bar{R}$ of R; accordingly, R is called *normal* if it coincides with its normalization.

We return to affine monoid algebras and so consider the ring $R=K[S]$, where K is a field and S an affine monoid. It is not hard to see that normality of R entails normality of S: let $v\in\mathbb{Z}S$ and $c\cdot v\in S$ for some $c>0$. Then $X^v\in\operatorname{Frac}R$ and $(X^v)^c\in R$. Now, if R is normal, it follows already that $X^v\in R$, so that $v\in S$.

Proposition 3.3.4: *Let S be an affine monoid and $K[S]$ the corresponding monoid algebra (with a field K). Then the following are equivalent:*

(i) *S is normal.*
(ii) *$K[S]$ is normal.*

Proof: The implication (ii) $\Rightarrow$ (i) has just been observed. So we now assume that S is normal. Without loss of generality, we may further assume that $\mathbb{Z}S=\mathbb{Z}^n$. By Lemma 3.1.2, we then have

$$S = C(S)\cap\mathbb{Z}^n,$$

where $C(S)$ denotes the associated cone of S. This is a finitely generated rational cone, and within $\mathbb{R}^n$, it has full dimension. Let $\sigma_1, \ldots, \sigma_m$ be the support forms of $C(S)$, so that

$$C(S) = \{x \in \mathbb{R}^n : \sigma_i(x) \geq 0 \text{ for all } i = 1, \ldots, m\},$$

by Lemma 2.9.7. Note that each support form σ_i is rational (even integral). Now let

$$H_i^+ := \{x \in \mathbb{R}^n : \sigma_i(x) \geq 0\}$$

be the rational halfspace defined by σ_i, and H_i the bounding rational hyperplane.

We consider the sublattice $L_i := H_i \cap \mathbb{Z}^n$. H_i has $\mathbb{Q}$-dimension $n-1$, thus L_i has rank $n-1$. Then the factor group $\mathbb{Z}^n/L_i$ has rank 1, and since this group is torsion-free (if $v \in \mathbb{Z}^n$ is not in L_i, then neither is any positive multiple of v), it is an infinite cyclic group. Therefore, there is a vector $v \in \mathbb{Z}^n$ such that

$$\mathbb{Z}^n = L_i \oplus \mathbb{Z} \cdot v;$$

we may assume that $v \in H_i^+$. Therewith, we find that the (normal) affine monoid

$$S_i := H_i^+ \cap \mathbb{Z}^n$$

has a direct sum decomposition $S_i = L_i \oplus \mathbb{N}v$. We conclude that the affine monoid algebra

$$K[S_i] = K[L_i \oplus \mathbb{N}v] \cong K[\mathbb{Z}^{n-1} \oplus \mathbb{N}]$$

is factorial and, hence, also normal. But since $S = S_1 \cap \cdots \cap S_m$, the algebra

$$K[S] = K[S_1] \cap \cdots \cap K[S_m]$$

is normal as well. This proves the proposition. □

We note an immediate conclusion.

Corollary 3.3.5: *$K[\bar{S}]$ is the normalization of $K[S]$, and is a finitely generated $K[S]$-module.*

Proof: By Lemma 3.1.2 and Proposition 3.3.2, $K[\bar{S}]$ is a finitely generated K-algebra, so more than ever a finitely generated $K[S]$-algebra. One can choose the generators monomially; obviously, they are integral over $K[S]$ then. Therefore, $K[\bar{S}]$ is a finitely generated $K[S]$-module. □

This proof already gives a first impression how fruitful the interplay between the categories of affine monoids on the one hand and affine monoid algebras on the other hand can be. Also, the following proposition has a nice combinatorial analogue; originally, it was stated and proven by Artin and Tate [AT, Theorem 1] from whom we quote.

Proposition 3.3.6: *Let R_0 be a Noetherian ring, R a finitely generated R_0-algebra and A an intermediate algebra such that R is a finitely generated A-module. Then A is a finitely generated R_0-algebra.*

Proof: Write $R = R_0[x_1,\ldots,x_n] = Ay_1 + \cdots + Ay_m$. Then there exist expressions of the form

$$x_i = \sum_{j=1}^{m} a_{ij} y_j\,, \qquad i = 1,\ldots,n\,, \qquad a_{ij} \in A\,, \tag{3}$$

$$y_i y_j = \sum_{k=1}^{m} b_{ijk} y_k\,, \qquad i,j = 1,\ldots,m\,, \qquad b_{ijk} \in A\,. \tag{4}$$

Let A_0 be the R_0-algebra generated by the a_{ij} and the b_{ijk}. Clearly, A_0 is Noetherian and contained in A.

An element of R is a polynomial in the x_i with coefficients in R_0. Substituting (3) and making repeated use of (4) shows that

$$R = A_0 y_1 + \cdots + A_0 y_m\,,$$

so that R is a finitely generated A_0-module. Hence A is also a finitely generated A_0-module, and, therefore, is a finitely generated R_0-algebra. □

The following corollary is a nice application of some of the preceding propositions.

Corollary 3.3.7: *Let S be a submonoid of a lattice L. Then the following statements are equivalent:*
(i) *S is finitely generated.*
(ii) *$\bar{S}$ is finitely generated.*
(iii) *$C(S)$ is finitely generated.*

Proof: The implication (i) ⇒ (iii) is clear, while (iii) ⇒ (ii) immediately follows from Lemma 3.1.2. It remains to prove the implication (ii) ⇒ (i). So suppose that $\bar{S}$ is finitely generated, and choose a field K. Then $K[\bar{S}]$ is a finitely generated K-algebra (Proposition 3.3.2), and according to Proposition 3.3.6 (and Corollary 3.3.5), so too is $K[S]$. But then S is finitely generated, again by Proposition 3.3.2. □

The next lemma continues the series of interaction of affine monoids and their algebras. It states that each affine monoid is ‘almost normal’.

Lemma 3.3.8: *Let S be an affine monoid. Then there exists some $s_0 \in S$ such that*

$$s_0 + \bar{S} \subseteq S\,.$$

In other words, in the region of the ‘shifted’ cone $s_0 + C(S)$, S contains all lattice points (from $\mathbb{Z}S$).

Proof: We consider the corresponding affine monoid algebras $K[S]$ and $K[\bar{S}]$. By Corollary 3.3.5, $K[\bar{S}]$ is a finitely generated $K[S]$-module. Using the abbreviation $Q := K[\bar{S}]/K[S]$ for the factor module, we have a short exact sequence

$$0 \longrightarrow K[S] \xrightarrow{\ i\ } K[\bar{S}] \xrightarrow{\ p\ } Q \longrightarrow 0$$

of finitely generated $K[S]$-modules, where i is the inclusion map and p the canonical projection. Obviously, the subset

$$\mathscr{M} := \mathscr{M}(K[S]) := \{X^s : s \in S\}$$

of all monomials in $K[S]$ is a multiplicative submonoid, and since localization is exact, we obtain a short exact sequence

$$0 \longrightarrow K[S]_{\mathscr{M}} \xrightarrow{i_{\mathscr{M}}} K[\bar{S}]_{\mathscr{M}} \xrightarrow{p_{\mathscr{M}}} Q_{\mathscr{M}} \longrightarrow 0$$

of $K[S]_{\mathscr{M}}$-modules. But clearly, we have

$$K[S]_{\mathscr{M}} = K[\mathbb{Z}S] = K[\bar{S}]_{\mathscr{M}},$$

so that the inclusion map $i_{\mathscr{M}}$ has turned to identity. But this means that $Q_{\mathscr{M}} = 0$, and since Q is finitely generated, it is annihilated by a single monomial $X^{s_0} \in \mathscr{M}$. This in turn means that

$$X^{s_0} \cdot K[\bar{S}] \subseteq K[S].$$

This is the ring-theoretic view of the assertion; the assertion follows immediately from it. (In fact, both statements are equivalent.) □

We consider the following example: let S be the affine submonoid

$$S = \mathbb{N}\big((2,0), (2,1), (2,2), (3,1)\big)$$

of $\mathbb{N}^2$. In Figure 7, the four generators are highlighted, and the 'first' elements of S

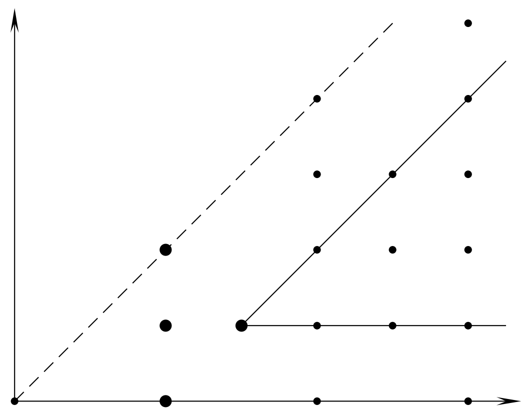

Figure 7

(with small values of the coordinates) are plotted. One finds that

$$\mathbb{Z}S = \mathbb{Z}^2, \quad C(S) = \mathbb{R}_+((1,0),(1,1)) \quad \text{and} \quad \bar{S} = \mathbb{N}((1,0),(1,1)).$$

S is not normal, but for $s_0 := (3,1)$, we have $s_0 + \bar{S} \subseteq S$.

The reader should notice that, in this example, $\bar{S}$ is covered by S and three more hyperplanes (which are lines here, of course). This holds true in any case:

Corollary 3.3.9: *Let S be an affine monoid in $\mathbb{Z}^n$. Then there exist finitely many hyperplanes $H_1,\dots,H_u$ in $\mathbb{R}^n$ such that $\bar{S}$ is covered by S and these hyperplanes:*

$$\bar{S} \subseteq S \cup H_1 \cup \cdots \cup H_u .$$

In other words, S and $\bar{S}$ differ only in $H_1,\dots,H_u$.

Proof: It is clear that we may assume that S has rank n. Then the cone $C(S)$ is an n-dimensional rational polyhedral cone in $\mathbb{R}^n$, and as such, it has a well-defined set of support forms $\sigma_1,\dots,\sigma_s$, for which

$$C(S) = \{x \in \mathbb{R}^n : \sigma_i(x) \geq 0 \text{ for } i = 1,\dots,s\}$$

(Lemma 2.9.7). Now choose $s_0 \in S$ according to Lemma 3.3.8, and suppose we are given a point $x \in \bar{S}$ which is not contained in $s_0 + C(S)$. Since

$$s_0 + C(S) = \{x \in \mathbb{R}^n : \sigma_i(x) \geq \sigma_i(s_0) \text{ for } i = 1,\dots,s\},$$

and since each σ_i maps $\bar{S}$ to $\mathbb{N}$, we obtain an integral double inequality

$$0 \leq \sigma_i(x) < \sigma_i(s_0)$$

for some i. Therefore, $\bar{S}$ is covered by $s_0 + \bar{S}$ (which in turn is covered by S) and the hyperplanes

$$H_{ij} := \{x \in \mathbb{R}^n : \sigma_i(x) = j\}$$

for $i = 1,\dots,s$ and $j = 0,\dots,\sigma_i(s_0) - 1$. □

Using the same ideas as in this proof, we even obtain a more general result:

Corollary 3.3.10: *Let S be an affine monoid in $\mathbb{Z}^n$, and $t \in \mathbb{Z}^n$ a point. Then any two of the following sets differ in only finitely many hyperplanes:*

$$S, \quad t+S, \quad \bar{S}, \quad t+\bar{S}.$$

Proof: Corollary 3.3.9 proves the assertion for the two sets S and $\bar{S}$, and then also for $t+S$ and $t+\bar{S}$; in order to prove the assertion for $\bar{S}$ and $t+\bar{S}$, one argues along the same lines as in the proof of Corollary 3.3.9 (the role of s_0 is then taken by t). □

We continue with general theory. In analogy with the respective notion in ring theory, we define: a subset I of an (affine) monoid S is called an *ideal* (in S) if $S + I \subseteq I$. (The sum denotes the Minkowski sum of the sets.) Note that S always contains two trivial ideals, namely, the empty set and S itself. In Lemma 3.3.8, we have seen that the set

$$\operatorname{cond}_S \bar{S} := \{s \in S : s + \bar{S} \subseteq S\}$$

is non-empty; clearly, $\operatorname{cond}_S \bar{S}$ is an ideal in S (and also in $\bar{S}$). It is called the *conductor ideal* of $\bar{S}$ in S. The analogy between ideals in S and monomial ideals in $K[S]$ is established as follows: let I be an arbitrary subset of S and $K(I)$ the K-vector subspace of $K[S]$ which is generated by the monomials X^a, $a \in I$. Then I is an ideal in S if and only if $K(I)$ is an ideal in $K[S]$.

Likewise, we also introduce the more general notion of a module over a monoid at this point: let S be an affine monoid in a lattice L. Then a subset $T \subseteq L$ is called an *S-module* if $S+T \subseteq T$. Of course, every ideal I of S can also be viewed as an S-module. In particular, the empty set and S itself are S-modules again. The link between S-modules and $K[S]$-modules is the following: a subset $T \subseteq L$ is an S-module if and only if the K-vector subspace $K(T)$, which is generated by the respective monomials X^a, $a \in T$, in $K[L]$, is a $K[S]$-module.

Besides the two trivial S-modules and the conductor ideal, we know yet another S-module, namely, the normalization $\bar{S}$ of S. We have seen that $K[\bar{S}]$ is a finitely generated $K[S]$-module (Corollary 3.3.5). But then the generators can even be chosen monomially, that is, there are monomials $X^{a_1},\dots,X^{a_n}$ (with $a_i \in \bar{S}$) such that

$$K[\bar{S}] = X^{a_1} \cdot K[S] + \cdots + X^{a_n} \cdot K[S]\,.$$

It follows that

$$\bar{S} = (a_1 + S) \cup \cdots \cup (a_n + S)\,.$$

We introduce a notion for this situation: an S-module T is *finitely generated* if it is the union of finitely many *cyclic* S-modules $a+S$, $a \in T$. As we have seen (in the case of $T = \bar{S}$), this property is transferred from the S-module T to the $K[S]$-module $K(T)$, and vice versa.

Lemma 3.3.11: *For an affine monoid S, the normalization $\bar{S}$ is a finitely generated S-module.*

3.4. Positive monoids. Let S be a monoid. Then the invertible elements of S form a subgroup S_0. It is the largest group contained in S. In analogy with the corresponding notion for cones, we say that S is *positive* if $S_0 = 0$. In the affine case (where S is contained in some lattice), S_0 has a geometric description as

$$S_0 = \{v \in S : -v \in S\}\,,$$

since the inverse element $-v$ can be found in the group $\mathbb{Z}S$ of differences of S. Here, S_0 is a subgroup of $\mathbb{Z}S$, and $\mathbb{Z}S/S_0$ carries a group structure again.

We consider an example, namely, the submonoid

$$S = \mathbb{N}\big(\pm(2,0)\,,\,(1,1)\,,\,(0,1)\big)$$

in $\mathbb{Z}^2$. Its group of differences is $\mathbb{Z}S = \mathbb{Z}^2$, and its group of units is $S_0 = \mathbb{Z}\cdot(2,0)$; therefore, the group $\mathbb{Z}S/S_0 = \mathbb{Z}/2 \times \mathbb{Z}$ is not torsion-free. Note that S is not normal. As we shall see now, there is a connection between these two properties.

To this end, let S be a normal affine monoid. We choose some $v \in \mathbb{Z}S$ such that $\bar{v}$ is a torsion element in the group $\mathbb{Z}S/S_0$. This means that $c \cdot \bar{v} = \bar{0}$ for some $c \in \mathbb{N}$, $c > 0$. Therefore $cv \in S_0$, so that both $cv \in S$ and $-cv \in S$. But since S is normal, it follows that $v \in S$ and $-v \in S$, whence $\bar{v} = \bar{0}$. Therefore $\mathbb{Z}S/S_0$ is torsion-free, and we have already shown part of the following lemma.

Lemma 3.4.1: *Let S be a normal affine monoid of rank r. Then the following statements hold:*

(a) *The factor group $\mathbb{Z}S/S_0$ is torsion-free.*
(b) $S_0 \cong \mathbb{Z}^s$ *for some* $s \in \mathbb{N}$, $s \leq r$.
(c) $\mathbb{Z}S/S_0 \cong \mathbb{Z}^{r-s}$.

Proof: (a) has been shown above. (b) is clear from the inclusion $S_0 \subseteq \mathbb{Z}S \cong \mathbb{Z}^r$.

(c) We apply the structure theorem for finitely generated Abelian groups to $\mathbb{Z}S/S_0$: by (a), there is no torsion part, and the exact sequence

$$0 \longrightarrow S_0 \longrightarrow \mathbb{Z}S \longrightarrow \mathbb{Z}S/S_0 \longrightarrow 0$$

shows that $\mathbb{Z}S/S_0$ has rank $r-s$. □

Now let S be a monoid, and let S_1, S_2 be submonoids. Then we write

$$S = S_1 \oplus S_2$$

and say that S is the *(inner) direct sum* of S_1 and S_2 if, for every $v \in S$, there exist *unique* elements $v_1 \in S_1$ and $v_2 \in S_2$ such that $v = v_1 + v_2$. Note that this implies $S = S_1 + S_2$ and $S_1 \cap S_2 = \{0\}$, whereas the converse implication is not true for monoids in general. (For groups, both conditions are equivalent, of course.) If S is normal, then S splits into a direct sum of S_0 and a positive normal submonoid.

Proposition 3.4.2: *Let S be a normal affine monoid. Then $S = S_0 \oplus S'$ with a positive normal submonoid S' of S.*

Proof: Let $r = \operatorname{rk} S$ and $s = \operatorname{rk} S_0$ be as in Lemma 3.4.1. By the Elementary Divisors Theorem, there exists a basis $v_1, \dots, v_r$ of $\mathbb{Z}S$, and $a_1, \dots, a_s \in \mathbb{N}$, such that $a_1 v_1, \dots, a_s v_s$ is a basis of S_0. Then

$$\mathbb{Z}S/S_0 \cong \mathbb{Z}^{r-s} \times \mathbb{Z}/a_1 \times \cdots \times \mathbb{Z}/a_s,$$

and Lemma 3.4.1 implies that $a_1 = \cdots = a_s = 1$. Therefore $\mathbb{Z}S = S_0 \oplus U$ for the group $U := \mathbb{Z}(v_{s+1}, \dots, v_r)$.

Now we show that this decomposition of $\mathbb{Z}S$ induces a corresponding one of S. For this, we consider the submonoid $S' := S \cap U$ of S and claim that

(1) S' is positive,
(2) S' is normal, and
(3) $S = S_0 \oplus S'$.

For (1), let $x \in S'$ such that $-x \in S'$, too. But then both $x \in S$ and $-x \in S$, hence $x \in S_0 \cap U = \{0\}$.

For (2), choose $x \in \mathbb{Z}S'$ such that $cx \in S'$ for some $c > 0$. Since $\mathbb{Z}S' \subseteq U \subseteq \mathbb{Z}S$, and S is normal, this immediately implies that $x \in S \cap U = S'$.

For (3), let $x \in S$. The main point is to show the *existence* of a representation $x = v + w$ with $v \in S_0$ and $w \in S'$; *uniqueness* is inherited from $\mathbb{Z}S$. But since $S \subseteq \mathbb{Z}S = S_0 \oplus U$, we find $v \in S_0$ and $w \in U$ such that $x = v + w$. Then $-v \in S$ and $w = x - v \in S$, hence $w \in S'$, as desired. □

We now leave the area of direct sum decompositions of normal affine monoids, and turn to the question how one can check the positivity of a (not necessarily normal) affine monoid S. The following lemma gives an answer: as one expects, positivity of S is determined by the associated cone $C(S)$. The lemma can be seen as an analogue of Lemma 2.10.1 which characterizes positive cones.

Lemma 3.4.3: *For an affine monoid S, the following statements are equivalent:*
(i) *S is positive.*
(ii) *$\bar{S}$ is positive.*
(iii) *$C(S)$ is positive.*

Proof: The implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) are clear, due to inclusions.

The implication (i) $\Rightarrow$ (ii) is easy to see: choose $x \in \bar{S}$ such that $-x \in \bar{S}$, too. Then both cx and $-cx$ are in S for some common $c > 0$. This implies that $cx = 0$ (since S is assumed to be positive), hence $x = 0$.

Finally, let us prove the implication (ii) $\Rightarrow$ (iii). We consider all objects within the $\mathbb{R}$-vector space $\mathbb{R}S$ generated by S. There, the rational polyhedral cone $C := C(S)$ has full dimension, and via its support forms $\sigma_1, \dots, \sigma_m$, there is a representation

$$C = \{x \in \mathbb{R}S : \sigma_i(x) \geq 0 \text{ for all } i = 1, \dots, m\}$$

(see Lemma 2.9.7). The group of units C_0 then is

$$C_0 = \{x \in \mathbb{R}S : \sigma_i(x) = 0 \text{ for all } i = 1, \dots, m\},$$

that is, the kernel of the rational (even integral) linear map $\sigma = (\sigma_1, \dots, \sigma_m) : \mathbb{R}S \to \mathbb{R}^m$. It follows that $C_0 = \mathbb{R}(v_1, \dots, v_t)$ is an $\mathbb{R}$-vector subspace of $\mathbb{R}S$, generated by rational basis vectors $v_j \in \mathbb{Q}S$. Clearly, we may even assume that $v_j \in \mathbb{Z}S$. But then the vectors

$$\pm v_j \in C_0 \cap \mathbb{Z}S \subseteq C \cap \mathbb{Z}S = \bar{S}$$

are both contained in the normalization $\bar{S}$, and since this was assumed to be positive, it follows that $v_j = 0$ for each j. Therefore $C_0 = 0$. □

A short analysis of the proof shows that we have actually established a stronger result.

Lemma 3.4.4: *For the groups of units of an affine monoid S and the associated cone $C = C(S)$, we have $C_0 = \mathbb{R}(S_0)$.*

Now let S be a positive affine monoid. As we have just seen, the cone $C(S)$ then is positive, too. Anyway, $C(S)$ is rational and polyhedral, and if we confine ourselves to the vector space $\mathbb{R}S$, then it is also full-dimensional. In this case, there is a unique set $\sigma_1, \dots, \sigma_m$ of (integral) support forms, and the respective $\mathbb{R}$-linear map $\sigma = (\sigma_1, \dots, \sigma_m) : \mathbb{R}S \to \mathbb{R}^m$ is injective (Lemma 2.10.1). In Subsection 2.10, we have already defined the standard embedding $\sigma|_C : C \hookrightarrow \mathbb{R}^m_+$ of C into $\mathbb{R}^m_+$. Since σ is even integral (that is, σ can be restricted to an injective homomorphism

$\sigma|_{\mathbb{Z}S} : \mathbb{Z}S \to \mathbb{Z}^m$ of groups), it also induces an embedding $\sigma|_S : S \hookrightarrow \mathbb{N}^m$; this is called the *standard embedding* of the positive affine monoid S into $\mathbb{N}^m$.

We take the liberty of anticipating the following section: due to the embedding $S \hookrightarrow \mathbb{N}^m$ and the 'total degree' homomorphism $\mathbb{N}^m \to \mathbb{N}$ (which assigns the sum of the coordinates to a vector), we arrive at a homomorphism $\deg : S \to \mathbb{N}$ of positive affine monoids. We have $\deg(x) = 0$ if and only if $x = 0$ (which, of course, does not mean that deg is injective). Furthermore, each fibre

$$S_n := \deg^{-1}(n) = \{x \in S : \deg(x) = n\}$$

(with $n \in \mathbb{N}$) is finite, and we have $S_m + S_n \subseteq S_{m+n}$. In this case, the monoid $S = \biguplus_{n\in\mathbb{N}} S_n$ is *positively graded* (via deg), and S_n is the n-th *graded component* of S. We will introduce the more general notion of a G-graded monoid in Subsection 4.1 (where G denotes an affine monoid). We record:

Lemma 3.4.5: *Every positive affine monoid S has a positive grading*

$$S = \biguplus_{n\in\mathbb{N}} S_n$$

with $S_0 = \{0\}$ and finite components S_n, $n \in \mathbb{N}$.

At the end of this section, we ask about minimal internal representations of a positive affine monoid S: does there exist a unique minimal set of generators of S? As we shall see in a moment, this is in fact true; in doing so, the standard embedding of S will be very helpful.

But first, we introduce another notion: an element $v \in S$, $v \neq 0$, is called *irreducible* (in S) if an equation $v = v_1 + v_2$ with elements $v_1, v_2 \in S$ is only possible for $v_1 = 0$ or $v_2 = 0$. It is clear that this notion makes sense for positive monoids only. Of course, an element $v \in S$, $v \neq 0$, is irreducible in S if and only if the respective monomial X^v is irreducible in $K[S]$. (K denotes an arbitrary field.)

Now comes the promised statement as for minimal generation of S. It is an analogue of the respective proposition for cones (Proposition 2.10.5). Here, the irreducible elements of S play a crucial role.

Proposition 3.4.6: *For a positive affine monoid S, the following statements hold:*

(a) *S has only finitely many irreducible elements.*

(b) *S has a unique minimal generating set, namely, its irreducible elements.*

Proof: We use the degree homomorphism $\deg : S \to \mathbb{N}$ introduced above. It follows easily that each element v of S is the sum of irreducible elements, by induction on the degree of v. Therefore, S is generated by its irreducible elements, and since S has a finite system of generators, and every system of generators must contain the irreducibles (by definition of irreducibility), their number is finite. □

We end this subsection with the following definition: the unique minimal generating set of a positive affine monoid S is called the *Hilbert basis* of S, and is denoted by $\operatorname{Hilb} S$.

4. Hilbert functions

In this section we initiate the notion of grading for rings and modules; the objects introduced in the previous section, that is, affine monoids (and their modules), are well-suited as grading sets. For graded objects, the Hilbert function and Hilbert series are defined. With these notions at hand, we can state the well-known results on the Hilbert function of homogeneous and, more generally, positively graded rings (and their modules). This naturally leads to Hilbert polynomials and Hilbert quasi-polynomials.

We are particularly interested in the degrees of the partial polynomials of such a quasi-polynomial. In the cases that are relevant to us, these degrees will be equal; this is an essential basis for our studies in Section 6. We end the current section with another preparatory work for Section 6: we reduce the Hilbert function of an affine algebra to that of a direct sum of shifted polynomial rings; this procedure is known as the Stanley decomposition of an affine algebra.

Recommendable references for this section are the textbooks by Bruns and Herzog [BH], Matsumura [Ma] and Eisenbud [Ei].

4.1. Graded monoids, rings and modules. In this subsection we shall investigate monoids, rings and modules which, like a polynomial ring, admit a decomposition of their elements into homogeneous components. Throughout this subsection, G denotes an affine monoid.

Definition 4.1.1: A *G-graded ring* is a ring R, together with a decomposition $R = \bigoplus_{g\in G} R_g$ (as an Abelian group) such that $R_g \cdot R_{g'} \subseteq R_{g+g'}$ for all $g, g' \in G$.

A *G-graded R-module* is an R-module M, together with a decomposition $M = \bigoplus_{g\in G} M_g$ (as an Abelian group) such that $R_g \cdot M_h \subseteq M_{g+h}$ for all $g, h \in G$. One calls M_g the g-th *graded* (or *homogeneous*) *component* of M.

The elements $x \in M_g$ are called *homogeneous (of degree g)*. According to this definition, the zero element is homogeneous of arbitrary degree. The degree of x is denoted by $\deg x$. An arbitrary element $x \in M$ has a unique representation $x = \sum_{g\in G} x_g$ as a finite sum of homogeneous elements $x_g \in M_g$. The summands x_g are uniquely determined by x, and are called the *homogeneous components* of x.

In order to get used to working with gradings, we show that the unit element $1 \in R$ is homogeneous of degree 0: write $1 = x_1 + \cdots + x_n$ with homogeneous elements $x_i \neq 0$ of different degrees $d_i \in G$. Multiplying by an arbitrary homogeneous element $x \neq 0$ of degree d yields

$$x = xx_1 + \cdots + xx_n\,.$$

Since, on the left-hand side, there is only one summand (of degree d), the same must happen on the right-hand side; but there, the degrees of the summands are $d + d_1, \ldots, d + d_n$, pairwise distinct. Therefore, all but one of these summands vanish, and the remaining summand, say xx_1, equals x and has degree d. It follows that $d_1 = 0$ and $d_i \neq 0$ for $i \neq 1$. This in turn enforces the equation

$$x = xx_1$$

for *all* homogeneous elements $x \in R$, whence it even holds for arbitrary $x \in R$ (since each element is a finite sum of its homogeneous components). Thus $x_1 = 1$.

Summing up, we have shown that R_0 is a subring of R. Clearly, each graded piece M_g of a G-graded R-module M is an R_0-module, and $M = \bigoplus M_g$ is a direct sum decomposition of M as an R_0-module.

We mention the most common example of a graded ring, namely, the polynomial ring $R = K[X_1, \ldots, X_n]$ over a field K. It is usually made into an $\mathbb{N}$-graded ring by defining the degree of a monomial $X_1^{a_1} \cdots X_n^{a_n}$ to be the total degree $a_1 + \cdots + a_n$; however, R has other useful gradings. For example, R has a natural $\mathbb{N}^n$-grading in which $X_1^{a_1} \cdots X_n^{a_n}$ has degree $(a_1, \ldots, a_n)$; in this case, $\dim_K R_g = 1$ for all $g \in \mathbb{N}^n$, and R is isomorphic with the monoid ring $K[\mathbb{N}^n]$ (which originally constitutes the construction of R). Alternatively, giving each of the X_i some suitable weight $d_i \in \mathbb{N}$ and letting the monomial $X_1^{a_1} \cdots X_n^{a_n}$ have weight $a_1 d_1 + \cdots + a_n d_n$ defines another $\mathbb{N}$-grading on R.

Definition 4.1.2: Let R be a G-graded ring, and let M and N be G-graded R-modules. An R-linear map $\varphi : M \to N$ is called *homogeneous of degree $d \in G$* if $\varphi(M_g) \subseteq N_{d+g}$ for all $g \in G$. It simply is *homogeneous* if it is homogeneous of degree 0.

Let M be a G-graded R-module, and N a submodule of M. N is a *G-graded submodule* if it is a G-graded module such that the inclusion map $N \hookrightarrow M$ is homogeneous. This is equivalent to saying that $N_g = N \cap M_g$ for all $g \in G$. Yet in other words, N is a graded submodule of M if and only if N is generated by those homogeneous elements of M which belong to N. In particular, if $x \in N$, then all homogeneous components of x belong to N. Furthermore, $M/N = \bigoplus M_g/N_g$ carries a natural G-grading. If φ is a homogeneous map between graded modules, then the kernel and the image of φ are graded submodules.

We extend the notion of grading to the category of monoids.

Let S be an (additive) monoid. If there exists a disjoint decomposition $S = \biguplus_{g \in G} S_g$ such that $S_g + S_{g'} \subseteq S_{g+g'}$ for all $g, g' \in G$, then we call S a *G-graded monoid*, and S_g the g-th *graded* (or *homogeneous*) *component* of S. For example, S is trivially G-graded by defining $S_0 := S$ and $S_g := \emptyset$ for all $g \neq 0$. Indicating the other extreme case, G itself can be G-graded via $G_g := \{g\}$ for $g \in G$. Recall the notion of a *positively graded* monoid, already introduced in Subsection 3.4: this means that $G = \mathbb{N}$, $S_0 = \{0\}$ and that each component S_n is finite.

We also recall some terminology from Subsection 3.3: if S is a submonoid of G, then a subset $T \subseteq G$ is called an *S-module* if $S + T \subseteq T$ (including the case $T = \emptyset$). T is *finitely generated* if there exist $t_1, \ldots, t_n \in T$ such that $T = \bigcup_i (t_i + S)$. If S is finitely generated and T is a finitely generated S-module, then every S-submodule $T' \subseteq T$ is also finitely generated. For example, this follows immediately by 'linearization' with coefficients in a field K: the K-vector space $M := K(T) \subseteq K[G]$ is a finitely generated module over the Noetherian ring $R := K[S]$, and so all its submodules are finitely generated over R. For $K(T')$, this implies the finite generation of T' over S.

First we note a result on the finite generation of certain subalgebras of G-graded algebras and submodules of G-graded modules.

Proposition 4.1.3: [BG, Theorem 4.4.1] *Let G be an affine monoid, S an affine submonoid of G, and $T \subseteq G$ a finitely generated S-module. Furthermore, let R be a Noetherian G-graded ring and M a finitely generated G-graded R-module. Then the following statements hold:*

(a) *R_0 is a Noetherian ring, and each graded component M_g, $g \in G$, of M is a finitely generated R_0-module.*
(b) *$A := \bigoplus_{s\in S} R_s$ is a finitely generated R_0-algebra.*
(c) *$N := \bigoplus_{t\in T} M_t$ is a finitely generated A-module.*

In particular, R is a finitely generated R_0-algebra.

Proof: (a) In order to see that R_0 is Noetherian, choose an ascending chain

$$I_0 \subseteq I_1 \subseteq \ldots \subseteq I_n \subseteq \ldots$$

of ideals in R_0, and extend it to an ascending chain

$$J_0 \subseteq J_1 \subseteq \ldots \subseteq J_n \subseteq \ldots$$

of ideals in R (by setting $J_i := R \cdot I_i$). Since R is Noetherian by assumption, the latter chain is stationary. We now show that $I = RI \cap R_0$ for any ideal I of R_0; then the former chain of ideals of R_0 must be stationary as well. But this equality is easily checked: the inclusion '$\subseteq$' is clear, and for the reverse inclusion, one chooses an arbitrary element $x = a_1 r_1 + \cdots + a_k r_k$ in RI, with G-homogeneous elements $a_i \in I \subseteq R_0$ and $r_i \in R_{d_i}$; requiring x to be additionally contained in R_0, we may assume that all of the r_i have degree 0, that is, each r_i is contained in R_0. But then $x \in I$, since I is an ideal in R_0.

Similarly, one checks that $RM' \cap M_g = M'$ for each R_0-submodule M' of M_g. Therefore, ascending chains of such submodules M' of M_g are stationary. This shows that the R_0-module M_g is Noetherian and finitely generated.

(b) For the rest of the proof, we may, and do, assume that G is a group, say $G = \mathbb{Z}^m$. (Replace G with its group of differences, and fill in the new components of R and M trivially, if any.)

First we do the case in which S is integrally closed in $\mathbb{Z}^m$. Let $\varphi : \mathbb{Z}^m \to \mathbb{Z}$ be a non-zero linear form. It induces a $\mathbb{Z}$-grading on R whose components are given by

$$R'_n := \bigoplus_{g\in\varphi^{-1}(n)} R_g .$$

Let R' denote R with this $\mathbb{Z}$-grading. Now set $R'_+ := \bigoplus_{n\geq 0} R'_n$ and define R'_- analogously. By [BH, Theorem 1.5.5], R'_0 is Noetherian, and both R'_+ and R'_- are finitely generated R'_0-algebras. On the other hand, R'_0 is a $(\ker\varphi)$-graded ring in a natural way, and by induction on m, we conclude that R'_0 is a finitely generated R_0-algebra. (Note that $\ker\varphi \cong \mathbb{Z}^{m-1}$.)

At this point, it follows in particular that R is a finitely generated R_0-algebra; note that this is the assertion if $S = \mathbb{Z}^m$. Otherwise $S = \mathbb{Z}^m \cap C(S)$ (by Lemma 3.1.2), and the rational polyhedral cone $C(S)$ has at least one supporting hyperplane. Then

$$S = \{x \in \mathbb{Z}^m : \varphi_i(x) \geq 0 \text{ for } i = 1, \dots, u\}$$

with $u \geq 1$ and (non-zero) integral linear forms $\varphi_i : \mathbb{R}^m \to \mathbb{R}$ (Lemma 2.9.7). In order to see that A is a finitely generated R_0-algebra, we use induction on u, and the induction hypothesis applies to $R' := \bigoplus_{s\in S'} R_s$, where

$$S' := \{x \in \mathbb{Z}^m : \varphi_i(x) \geq 0 \text{ for } i = 1, \dots, u-1\}.$$

In particular, R' is Noetherian, since so is R_0 by (a). Applying the above argument with $\varphi = \varphi_u$, one concludes that

$$R'_+ = \bigoplus_{\substack{s\in S' \\ \varphi(s)\geq 0}} R_s = A$$

is a finitely generated R_0-algebra.

In the general case for S we denote by S' the integral closure of S in $\mathbb{Z}^m$. Then $A' := \bigoplus_{s\in S'} R_s$ is a finitely generated algebra over the Noetherian ring R_0, as we have already shown. In particular, A' is also a finitely generated A-algebra. It is not hard to check that, in addition, A' is integral over A: in fact, let $x \in R_s$ for some $s \in S'$; then $cs \in S$ for some $c > 0$, and therefore, x^c is contained in R_{cs} (and A). It follows that A' is a finitely generated A-module. Now Proposition 3.3.6 implies that A is a finitely generated R_0-algebra.

(c) By hypothesis, T is the union of finitely many translates $t + S$. Therefore, we may assume that $T = t + S$. Passing to the shifted module $M(-t)$, given by $M(-t)_g = M_{g-t}$, we may even assume that $T = S$. Now the proof follows the same pattern as that of (b). In order to deal with an integrally closed submonoid of $G = \mathbb{Z}^m$, one notes that M_+ is a finitely generated module over R_+, where '$+$' denotes the positive part with respect to a $\mathbb{Z}$-grading (induced by a linear form $\varphi : \mathbb{Z}^m \to \mathbb{Z}$). This is shown as follows: the extended module RM_+ is finitely generated over R, and any of its generating systems $E \subseteq M_+$, together with finitely many components M_n, $n \geq 0$, generates M_+ over R_+; furthermore, the M_n are finitely generated over R_0 by (a).

For the general situation, we consider N' defined analogously as A'. It is a finitely generated A'-module, by the previous argument. Since A' is a finitely generated A-module, N' is finitely generated over A, and thus so is its submodule N. □

The proposition has a purely combinatorial consequence which will be very important in the following.

Corollary 4.1.4: *In the lattice L, let S and S′ be affine submonoids, T a finitely generated S-module and T′ a finitely generated S′-module. Then the following hold:*
(a) *$S \cap S'$ is an affine monoid, and*
(b) *$T \cap T'$ is a finitely generated $(S \cap S')$-module.*

Proof: We choose a field K of coefficients and set $R = K[S']$, $M = K(T')$. Then the hypotheses of Proposition 4.1.3 are satisfied, which therefore implies the finite generation of

$$A = \bigoplus_{s \in S} R_s = K[S \cap S'] \quad \text{and} \quad N = \bigoplus_{t \in T} M_t = K(T \cap T')$$

as a K-algebra and an A-module, respectively. (Note that $R_0 = K$.) However, finite generation of the 'linearized' objects is equivalent to that of the combinatorial ones (Proposition 3.3.2). □

4.2. Hilbert functions and Hilbert series. In this subsection L denotes a lattice. Let $R = \bigoplus R_\ell$ be a Noetherian L-graded ring, and $M = \bigoplus M_\ell$ a finitely generated L-graded R-module. Then, by Proposition 4.1.3, each M_ℓ is a finitely generated R_0-module. Suppose further that R_0 is a finitely generated K-algebra, for some field K. Then each M_ℓ is a K-vector space, and we can define the Hilbert function. Note that M_ℓ may have infinite K-dimension in general; but in the particular case when R_0 is a finitely generated K-module, the dimension will be finite.

Definition 4.2.1: The numerical function

$$H(M,\cdot) : L \to \mathbb{N} \cup \{\infty\}, \quad H(M,\ell) := \dim_K M_\ell \,,$$

is called the *(L-graded) Hilbert function* of M. In addition, if the Hilbert function is finite, then we define the *(L-graded) Hilbert series* of M to be the formal series

$$H_M(X) := \sum_{\ell \in L} H(M,\ell) \cdot X^\ell \in \mathbb{N}[\![L]\!] \,.$$

For the Hilbert function and Hilbert series, there are lots of classical statements; we shall present some of them in Subsections 4.3 and 4.4. For a more complete collection, we refer the reader to the textbook [BH, Chap. 4].

But before investigating the Hilbert function of certain rings and modules, we turn our attention to affine monoids again, and introduce the notion of Hilbert function also for this class. As we shall see, the Hilbert function of an affine monoid coincides with the Hilbert function of the respective affine monoid algebra; therefore, the Hilbert functions share each of their properties.

So let A be an affine monoid and $\varphi : A \to L$ a homomorphism. Then φ induces an L-grading on A:

$$A = \biguplus_{\ell \in L} A_\ell \quad \text{with} \quad A_\ell := \varphi^{-1}(\ell) = \{a \in A : \varphi(a) = \ell\} \,.$$

The *(L-graded) Hilbert function* of A (with respect to φ) now is defined to be the numerical function

$$H(A,\cdot) : L \to \mathbb{N} \cup \{\infty\}, \quad H(A,\ell) := \operatorname{card} A_\ell .$$

Upon using the field K and considering the corresponding affine monoid algebra $R := K[A]$, we can relate the Hilbert functions of A and R: namely, R is L-graded in a natural way,

$$R = \bigoplus_{\ell \in L} R_\ell \quad \text{with} \quad R_\ell := K(A_\ell) = \bigoplus_{a \in A_\ell} K \cdot X^a ;$$

it is obvious now that the L-graded Hilbert function of R coincides with the L-graded Hilbert function of A.

Next we fix some $t \in L$ and consider a shifted copy $A(-t)$ of A: its ℓ-th graded component $A(-t)_\ell := A_{\ell-t}$ corresponds to the ℓ-th graded component

$$R(-t)_\ell = R_{\ell-t} = \bigoplus_{a \in A_{\ell-t}} K \cdot X^a$$

of $R(-t)$, so that the respective Hilbert functions coincide again. Note that $A(-t)$ is a cyclic A-module (with generator 0 in degree t), just as $R(-t)$ is a (free) cyclic R-module (with generator 1 in degree t). Upon defining the *shifted homomorphism* $\psi : A \to L$, $\psi(x) := t + \varphi(x)$, we can express the Hilbert function of $A(-t)$ as

$$H(A(-t),\ell) = \operatorname{card} A_{\ell-t} = \operatorname{card}\{a \in A : \varphi(a) = \ell - t\} = \operatorname{card} \psi^{-1}(\ell) .$$

This suggests calling it also the Hilbert function of ψ; in fact, we shall use this speech in Section 6, and write $\psi = t + \varphi$ for short.

Finally we note: if we are given several affine monoid homomorphisms $\varphi_i : A_i \to L$ and shifts $t_i \in L$, $i = 1, \ldots, r$, then the Hilbert function of $\biguplus_{i=1}^r A_i(-t_i)$ (which is defined in an obvious way) precisely corresponds to the Hilbert function of $\bigoplus_{i=1}^r R_i(-t_i)$.

4.3. Homogeneous rings and the Hilbert polynomial. Throughout this subsection, we will assume that $R = K[x_1, \ldots, x_m]$ is a finitely generated $\mathbb{Z}$-graded K-algebra, with homogeneous generators x_i of degree 1. (Note that $R_0 = K$ now.) In this case, the $\mathbb{N}$-graded ring R is called *homogeneous*.

We say that a numerical function $f : \mathbb{Z} \to \mathbb{Z}$ is *of polynomial type (of degree d)* if there exists a polynomial $p \in \mathbb{Q}[X]$ (of degree d) such that $f(n) = p(n)$ for all $n \in \mathbb{Z}$, $n \gg 0$. With this terminology at hand, we can now quote the main result of this subsection, which dates back to Hilbert.

Proposition 4.3.1: *Let $R = \bigoplus_{n \in \mathbb{N}} R_n$ be a homogeneous K-algebra (with $R_0 = K$ a field), and $M = \bigoplus_{n \in \mathbb{Z}} M_n$ a finitely generated $\mathbb{Z}$-graded R-module of Krull dimension d. Then the Hilbert function $H(M,\cdot)$ of M is of polynomial type of degree $d-1$.*

Proof: See [BH, Theorem 4.1.3]. □

It is clear that the polynomial $P_M \in \mathbb{Q}[X]$ (of degree $d-1$) for which $H(M,n) = P_M(n)$ for all $n \gg 0$ is uniquely determined by M. It is called the *Hilbert polynomial* of M. A note on the degree of the zero polynomial: in the proposition, this degree is defined to be -1; however, we shall frequently set the degree to $-\infty$.

We mention the most prominent example, namely, the polynomial ring $R = K[X_1,\dots,X_m]$, $\mathbb{N}$-graded by the total degree, where K is a field. The number of monomials of total degree n is $\binom{m+n-1}{m-1}$: this follows from the bijection

$$\{(a_1,\dots,a_m) \in \mathbb{N}^m : a_1+\cdots+a_m = n\} \\ \longleftrightarrow \quad \{(b_1,\dots,b_{m-1}) \in \mathbb{N}^{m-1} : 0 \le b_1 < b_2 < \cdots < b_{m-1} \le n+m-2\},$$

given by

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{m-1} \end{pmatrix} := \begin{pmatrix} a_1 \\ a_1+a_2+1 \\ a_1+a_2+a_3+2 \\ \vdots \\ a_1+a_2+\cdots+a_{m-1}+m-2 \end{pmatrix}.$$

Therefore, we have

$$H(R,n) = \dim_K R_n = \binom{m+n-1}{m-1}$$

for all $n \ge 0$, and the right-hand side is $P_R(n)$. Thus

$$P_R(X) = \binom{X+m-1}{m-1} = \frac{(X+m-1)(X+m-2)\cdots(X+1)}{(m-1)!} \\ = \frac{1}{(m-1)!}\cdot X^{m-1} + (\text{lower terms}).$$

4.4. Positively graded rings and Hilbert quasi-polynomials. Let $R = \bigoplus R_n$ be an $\mathbb{N}$-graded ring such that $R_0 = K$ is a field, and $R = K[x_1,\dots,x_m]$ is a finitely generated K-algebra, generated by homogeneous elements x_i of degree $d_i > 0$. Such K-algebras will be called *positively graded*. Note that R is Noetherian, by Hilbert's Basis Theorem.

Now let $d > 0$ be a common multiple of the degrees $d_1,\dots,d_m$, and define $S := K[R_d]$. Clearly, S is a $d\mathbb{N}$-graded K-subalgebra of R; however, note that $S \neq \bigoplus_{n\in d\mathbb{N}} R_n$ in general, that is, the graded components of S are note necessarily those of R (but proper subsets). The equality

$$S = K[x_1^{a_1}\cdots x_m^{a_m} : a_1d_1+\cdots+a_md_m = d]$$

shows that S is a finitely generated K-algebra, hence Noetherian (again by Hilbert's Basis Theorem).

So we already have shown part of the following lemma.

Lemma 4.4.1: *Let $R = \bigoplus_{n\in\mathbb{N}} R_n = K[x_1,\ldots,x_m]$ be a positively graded K-algebra (with $R_0 = K$ a field), and set $S = K[R_d]$, where d is a common multiple of the degrees of the generators x_i. Then*

(a) *S is a Noetherian $d\mathbb{N}$-graded K-algebra, finitely generated by elements of degree d, and*
(b) *R is a finitely generated S-module.*

Proof: (a) has been shown above, and for (b), one checks that the finite set

$$\{x_1^{a_1}\cdots x_m^{a_m} : 0 \le a_i < \frac{d}{d_i} \text{ for } i = 1,\ldots,m\}$$

generates R as an S-module: clearly, x_i^{d/d_i} is homogeneous of degree d, whence it is contained in S for all i. □

Now let $M = \bigoplus M_n$ be a finitely generated $\mathbb{Z}$-graded R-module. Then, by Lemma 4.4.1 (b), M is also finitely generated as an S-module. Next we define some S-submodules of M, namely,

$$N_j := \bigoplus_{i\equiv j\ (d)} M_i = \bigoplus_{n\in\mathbb{Z}} M_{j+dn}$$

for $j = 0,\ldots,d-1$. Being a finitely generated module over the Noetherian ring S, M is Noetherian, and each of the N_j is a finitely generated S-module. Recalling that S is $d\mathbb{N}$-graded (and finitely generated by elements of degree d over $S_0 = K$) and observing that N_j is $(j+d\mathbb{Z})$-graded, we can regrade S and N_j in such a way that S becomes $\mathbb{N}$-graded homogeneous (all generators now have degree 1) and N_j turns into a finitely generated $\mathbb{Z}$-graded S-module. (Effectively, we have replaced d with 1.)

We have therefore arrived at a situation where Hilbert's Proposition 4.3.1 applies and yields that the (original) Hilbert function $H(N_j,\cdot)$ of N_j is of polynomial type (on the residue class $j+d\mathbb{Z}$; it vanishes elsewhere, of course). Let $P_j \in \mathbb{Q}[X]$ be the Hilbert polynomial of N_j on $j+d\mathbb{Z}$. Then, since $M = \bigoplus_{j=0}^{d-1} N_j$, we find that, for all $n \gg 0$,

$$H(M,n) = \begin{cases} P_0(n) & \text{if } n \in d\mathbb{Z}, \\ P_1(n) & \text{if } n \in 1+d\mathbb{Z}, \\ \quad\vdots & \qquad\vdots \\ P_{d-1}(n) & \text{if } n \in d-1+d\mathbb{Z}. \end{cases}$$

For obvious reasons, the function $Q : \mathbb{Z} \to \mathbb{Q}$, defined via the right-hand side for all $n \in \mathbb{Z}$, is called a *quasi-polynomial (of period d)*; likewise, in our situation, the Hilbert function $H(M,\cdot)$ of M is said to be *of quasi-polynomial type (with period d)*.

We summarize our considerations.

Proposition 4.4.2: *Let $R = \bigoplus_{n\in\mathbb{N}} R_n = K[x_1,\ldots,x_m]$ be a positively graded K-algebra (with $R_0 = K$ a field), d a common multiple of the degrees of the x_i and $M = \bigoplus_{n\in\mathbb{Z}} M_n$ a finitely generated $\mathbb{Z}$-graded R-module. Then the Hilbert function $H(M,\cdot)$ of M is of quasi-polynomial type with period d.*

Clearly, this proposition generalizes Hilbert's Proposition 4.3.1, and since the quasi-polynomial Q satisfying $H(M,n) = Q(n)$ for all $n \gg 0$ is uniquely determined by M, it is called the *Hilbert quasi-polynomial* Q_M of M. We will sometimes write $Q_M = (P_0,\ldots,P_{d-1})$ in order to specify the partial polynomials P_j of the quasi-polynomial Q_M.

After these more theoretical considerations, it is time to present an example: let us determine the Hilbert function and quasi-polynomial of the polynomial ring $R = K[X_1,X_2]$, where $\deg X_1 = 1$ and $\deg X_2 = 2$. Clearly, for $n \in \mathbb{N}$ even, we have

$$H(R,n) = \dim_K R_n = \operatorname{card}\{(a_1,a_2) \in \mathbb{N}^2 : a_1 + 2a_2 = n\} = \frac{n}{2} + 1\,.$$

(This is the number of possible choices for a_2.) Similarly, for $n \in \mathbb{N}$ odd, one finds $H(R,n) = (n-1)/2 + 1$. Consequently, the Hilbert quasi-polynomial

$$Q_R(n) = \begin{cases} \dfrac{n+2}{2} & \text{for } n \equiv 0 \pmod 2\,, \\ \dfrac{n+1}{2} & \text{for } n \equiv 1 \pmod 2\,, \end{cases}$$

coincides with the Hilbert function $H(R,n)$ for all $n \geq 0$.

Having this example in mind, we return to general theory: when dealing with Hilbert quasi-polynomials, it is natural to ask what can be said about the degrees of the polynomials being involved. Since the degree of a Hilbert polynomial is directly determined by the Krull dimension of the respective module (namely, $\deg P_M = \dim M - 1$; see Proposition 4.3.1), we first quote a result on the invariance of the dimension under certain ring extensions.

Lemma 4.4.3: *Let $S \subseteq R$ be an extension of Noetherian rings such that R is a finitely generated S-module. Then, for any ideal I of R,* $\dim R/I = \dim S/(I\cap S)$.

Proof: See [BH, Corollary A.8]. (Note at this point that R is a finitely generated S-module if and only if R is integral over S and finitely generated as an S-algebra.) □

Now let R, d, S, M, N_j and P_j (for $j = 0,\ldots,d-1$) be as earlier in this subsection. As an immediate consequence of the lemma, we see that $\dim S = \dim R$ (choose $I = 0$), and that the R-module M, which can also be viewed as a module over S, has a well-defined Krull dimension, independent of the base ring: choosing $I = \operatorname{Ann}_R M$ yields

$$\dim_{(R)} M = \dim R/I = \dim S/(I\cap S) = \dim(S/\operatorname{Ann}_S M) = \dim_{(S)} M\,.$$

At this point, the modules N_j come into play: each of them is a finitely generated S-module, and since $M = \bigoplus_j N_j$, we have $\mathrm{Ann}_S M = \bigcap_j \mathrm{Ann}_S N_j$, which further yields that

$$\begin{aligned}\dim M &= \dim(S/\bigcap_{j=0}^{d-1} \mathrm{Ann}_S N_j) \\ &= \max\{\dim(S/\mathrm{Ann}_S N_j) : 0 \le j < d\} = \max\{\dim N_j : 0 \le j < d\}\,.\end{aligned}$$

The second equality results from Lemma 4.4.5 below, by induction on d. But before stating and proving the lemma, we recall that $\dim N_j = \deg P_j + 1$ for each j and sum up the above considerations.

Proposition 4.4.4: *Let the assumptions and the notation be as above. Then*

$$\max\{\deg P_j : 0 \le j < d\} = \dim M - 1\,.$$

For an ideal I in a ring R, we define $V(I)$ to be the *variety* of I, that is, the set of all prime ideals of R that contain I. Due to the correspondence of the (prime) ideals in R/I and the (prime) ideals in R that contain I, the variety of I is useful for the computation of $\dim R/I$. Namely, we obtain the following lemma.

Lemma 4.4.5: *Let R be a commutative ring, and I, J ideals in R. Then $V(I \cap J) = V(I) \cup V(J)$; in particular,*

$$\dim R/(I \cap J) = \max\{\dim R/I\,, \dim R/J\}\,.$$

Proof: Clearly, any prime ideal $\mathfrak{p}$ of R containing I or J also contains $I \cap J$; and conversely, if $\mathfrak{p}$ contains $I \cap J$, then $\mathfrak{p}$ also contains the product ideal $I \cdot J$, whence I or J is contained in $\mathfrak{p}$ (since $\mathfrak{p}$ is prime). □

Recall the above example: we had $\dim M = 2$ and $\deg P_0 = \deg P_1 = 1$. So yet another question arises, concerning the degrees of the P_j: when does $\deg P_i = \deg P_j$ for all i, j? In this case, we refer to the common degree of the polynomials $P_0, \ldots, P_{d-1}$ as the *degree* of the quasi-polynomial Q_M. Before stating a sufficient condition for this to happen, we remind the reader of yet another notion: for a module M over a ring R, an element $a \in R$ is called *M-regular* if it acts as a non-zerodivisor on M, that is, the equation $ax = 0$ (with $x \in M$) is possible only for $x = 0$.

Proposition 4.4.6: *Let the assumptions and the notation be as in Proposition* 4.4.2, *and assume in addition that there exists a homogeneous M-regular element $a \in R$ of degree* 1. *Then all the polynomials P_j have the same degree and the same leading coefficient. In particular, the Hilbert quasi-polynomial Q_M of M has degree* $\dim M - 1$.

Proof: First of all, the multiplication map $\mu_a : M \to M$, $\mu_a(x) := ax$, is K-linear and injective: linearity is evident, and $\ker \mu_a = \{x \in M : ax = 0\} = \{0\}$, due to the M-regularity of a. Furthermore, since a is homogeneous of degree 1, so too is μ_a. Therefore, all the restrictions $M_n \xrightarrow{a\cdot} M_{n+1}$, $n \in \mathbb{Z}$, of μ_a are K-monomorphisms. This in turn implies that $\dim_K M_n \leq \dim_K M_{n+1}$ for all $n \in \mathbb{Z}$, which means that the Hilbert function $H(M,\cdot)$ is monotonic increasing:

$$H(M,n) \leq H(M,n+1) \leq \cdots \leq H(M,n+d-1) \leq H(M,n+d).$$

For $n \gg 0$ and $n \equiv 0 \pmod{d}$, this results in a chain

$$P_0(n) \leq P_1(n+1) \leq \cdots \leq P_{d-1}(n+d-1) \leq P_0(n+d)$$

of inequalities for the partial polynomials of the Hilbert quasi-polynomial.

Now assume that P_j has leading term $a_j X^{m_j}$, that is, $P_j = a_j X^{m_j} + (\text{lower terms})$ for each j. Then the above chain yields the following chain for the leading terms with respect to n:

$$a_0 n^{m_0} \preccurlyeq a_1 n^{m_1} \preccurlyeq \cdots \preccurlyeq a_{d-1} n^{m_{d-1}} \preccurlyeq a_0 n^{m_0},$$

which is valid for all $n \gg 0$, $n \equiv 0 \pmod{d}$. This is possible only if $m_i = m_j$ and $a_i = a_j$ for all i, j. (Divide by $a_0 n^{m_0}$ and send $n \to \infty$.) So all the polynomials P_j have the same leading term. The additional statement on the degree of Q_M is clear, by Proposition 4.4.4. □

Note that the requirement for the degree of a can be weakened. A simple analysis of the proof shows that the proposition remains true if a has a degree which has no common divisor with d; then the multiplication map μ_a, when applied to the component M_n, reaches components M_i in every residue class modulo d.

In our example, where $M = R = K[X_1, X_2]$, one can choose $a = X_1$; this is a homogeneous non-zerodivisor of degree 1 in R.

4.5. Initial ideals and the Hilbert function. In this subsection we consider a G-graded K-algebra $A = K[x_1, \ldots, x_n]$, generated by homogeneous elements x_i of degrees $d_i \in G$. (G denotes an affine monoid and K a field.) Clearly, A has a representation

$$A \cong R/I$$

as a residue class ring of the polynomial ring $R = K[X_1, \ldots, X_n]$, for some ideal I of R. By assigning X_i the degree d_i, R turns into a G-graded ring, I is a homogeneous ideal and the above isomorphism has degree 0. Therefore, it induces an isomorphism of K-vector spaces in each G-degree, and the respective Hilbert functions of A and R/I coincide.

This motivates the study of the ring R/I. Since the monomials in R are assumed to be G-homogeneous, we may at first equip R with the finest possible grading, namely, the natural $\mathbb{N}^n$-grading. Due to the monoid homomorphism $\varphi : \mathbb{N}^n \to G$, given by the assignment $e_i \mapsto d_i$, every G-graded component of R is a direct sum of $\mathbb{N}^n$-graded components:

$$R = \bigoplus_{g \in G} R_g \quad \text{with} \quad R_g = \bigoplus_{\substack{a \in \mathbb{N}^n \\ \varphi(a) = g}} K \cdot X^a .$$

Here we use the following abbreviation: if $a = (a_1, \ldots, a_n) \in \mathbb{N}^n$, then X^a denotes the monomial $X_1^{a_1} \cdots X_n^{a_n}$.

We start with two lemmas which deal with monomial and initial ideals. They will prove very helpful in the reduction process of the main theorem in Section 6. An ideal J in R is called a *monomial ideal* if it is generated by monomials in the variables $X_1, \ldots, X_n$; since R is Noetherian, finitely many monomials always suffice to generate J.

Lemma 4.5.1: *Let $R = K[X_1, \ldots, X_n]$ be a polynomial ring over the field K, and let $J = (X^{a_1}, \ldots, X^{a_m})$ be a monomial ideal in R (with $a_i \in \mathbb{N}^n$). Then, for any $f \in R$, the following are equivalent:*

(i) $f \in J$.

(ii) *Every monomial of f is a multiple of one of the monomials X^{a_i}.*

Proof: The implication (ii) $\Rightarrow$ (i) is clear. For the other implication, write $f = f_1 \cdot X^{a_1} + \cdots + f_m \cdot X^{a_m}$ with $f_i \in R$. Now it suffices to observe that, in every polynomial $f_i \cdot X^{a_i}$, every monomial is divisible by X^{a_i}. □

Corollary 4.5.2: (a) *The set $\mathscr{M}(J)$ of all monomials in J is a K-basis of J.*

(b) *The set $\mathscr{M}(R) \smallsetminus \mathscr{M}(J)$ of all monomials not contained in J forms a K-basis of R/J. (Actually, of course, one has to take the residue classes of the monomials modulo J.)*

The next lemma deals with monomial orders and initial ideals. A total order $>$ on the set $\mathscr{M}(R)$ of all monomials in R is called a *monomial order* on R if, for all monomials $m_1, m_2, n \in \mathscr{M}(R)$ such that $m_1 > m_2$ and $n \neq 1$, we have $m_1 n > m_2 n > m_2$. It is worth noting that, due to the latter double inequality, any monomial order is Artinian, that is, descending chains of monomials eventually are stationary; this enables us to use the induction principle in the context of monomials.

Three important monomial orders in practice are the lexicographic ('lex'), the homogeneous lexicographic ('hlex') and the reverse lexicographic order ('rlex'), respectively. For instance, the *homogeneous lexicographic order* $>_{\text{hlex}}$ is given by

$$\begin{aligned} X^a >_{\text{hlex}} X^b \quad &:\Longleftrightarrow \quad a >_{\text{hlex}} b \\ &:\Longleftrightarrow \quad |a| > |b| \quad \text{or} \quad \begin{cases} |a| = |b| \text{ and } a_i > b_i \text{ for the} \\ \text{first index } i \text{ such that } a_i \neq b_i . \end{cases} \end{aligned}$$

(Here $a, b \in \mathbb{N}^n$, and $|a| = a_1 + \cdots + a_n$ denotes the total degree of a and X^a.)

If $>$ is a monomial order, then, for any polynomial $f \in R$, we define the *initial term* $\text{in}(f)$ of f to be the greatest term of f with respect to the order $>$, that is, the one where the greatest monomial is involved. For an ideal I of R, we write $\operatorname{in}(I)$ for the (monomial) ideal generated by the elements $\operatorname{in}(f)$, $f \in I$. This ideal is called the *initial ideal* of I. Note that it is finitely generated. As the following lemma shows, the transition from an arbitrary ideal to its initial ideal is of fundamental importance: initial ideals are very useful for the computation of the Hilbert function.

Lemma 4.5.3: *Let $R = K[X_1, \ldots, X_n]$ be a polynomial ring over the field K, and I any ideal in R. Furthermore, let $>$ be a monomial order on $\mathscr{M}(R)$, and define $\mathscr{M} := \mathscr{M}(R) \smallsetminus \operatorname{in}(I)$ to be the set of all monomials not contained in the initial ideal of I. Then the following statements hold:*

(a) *$\mathscr{M}$ forms, modulo $\operatorname{in}(I)$, a K-basis of $R/\operatorname{in}(I)$.*

(b) (Macaulay) *$\mathscr{M}$ forms, modulo I, a K-basis of R/I.*

Proof: (a) is an immediate consequence of Corollary 4.5.2 (b).

(b) First, we show that $\mathscr{M}$ is K-linearly independent. Let $m_1, \ldots, m_r \in \mathscr{M}$, and assume that there is a dependency $f := a_1 m_1 + \cdots + a_r m_r \in I$ with $a_i \in K^*$. This implies that $\operatorname{in}(f) \in \operatorname{in}(I)$. Since $\operatorname{in}(f) = a_i m_i$ for some i, this is a contradiction.

Now we show that $\mathscr{M}$ generates R/I (as a K-vector space). This in turn means to prove the inclusion $R \subseteq K \cdot \mathscr{M} + I$ (where $K \cdot \mathscr{M}$ stands for the K-vector space generated by $\mathscr{M}$). So choose $f \in R$ arbitrarily. Since $>$ is a total order on $\mathscr{M}(R)$, we may use induction on the position of $\operatorname{in}(f)$ in the chain of all monomials of R. For $f = \operatorname{in}(f) = 0$ there is nothing to show, so suppose that $f \neq 0$ and $\operatorname{in}(f) = a \cdot m$ (for some $a \in K$ and $m \in \mathscr{M}(R)$).

If $m \notin \operatorname{in}(I)$, then $am \in K \cdot \mathscr{M}$, and it suffices to deal with $\tilde{f} := f - a \cdot m$; since $\operatorname{in}(\tilde{f}) < \operatorname{in}(f)$, it follows that $\tilde{f} \in K \cdot \mathscr{M} + I$ (by the induction hypothesis), whence $f \in K \cdot \mathscr{M} + I$ also.

Now suppose that $m \in \operatorname{in}(I)$. Then there exists a polynomial $g \in I$ such that $m = \operatorname{in}(g)$. Therefore $\tilde{f} := f - a \cdot g$ has an initial term less than $\operatorname{in}(f)$. This implies that $\tilde{f} \in K \cdot \mathscr{M} + I$, and then $f \in K \cdot \mathscr{M} + I$ again. □

The lemma states, in particular, that the K-vector spaces R/I and $R/\operatorname{in}(I)$ are isomorphic, and that their Hilbert functions (with respect to K-dimension) coincide. The next lemma finally clarifies how to determine the Hilbert function of $R/\operatorname{in}(I)$.

Lemma 4.5.4: *Let K be a field, $R = K[X_1, \ldots, X_n]$ a polynomial ring over K, $\mathbb{Z}^n$-graded in the usual way, and $J = (X^{a_1}, \ldots, X^{a_m})$ a monomial ideal.*

Then there exists an isomorphism

$$R/J \cong \bigoplus_i (R/\mathfrak{p}_i)(-d_i)$$

of $\mathbb{Z}^n$-graded K-vector spaces, with finitely many prime ideals $\mathfrak{p}_i$ in R, each generated by some of the indeterminates, and suitable shifts $d_i \in \mathbb{Z}^n$.

Proof: R carries the natural $\mathbb{Z}^n$-grading, with respect to which J is a graded ideal. Therefore the K-vector space $V := R/J$ can also be viewed as an $\mathbb{Z}^n$-graded R-module. Suppose that $V \neq 0$. Then there exists a prime ideal $\mathfrak{p} \in \operatorname{Ass} V$ (since R is Noetherian). By a generalization of [BH, Lemma 1.5.6], $\mathfrak{p}$ is even $\mathbb{Z}^n$-graded. (This lemma remains true for G-graded rings and modules whenever the monoid G can be totally ordered, with an order compatible with the operation on G; clearly, such an order on G corresponds to a monomial order on the monoid ring $K[G]$.) Therefore $\mathfrak{p}$ is generated by monomials, and since it is prime, it is even generated by indeterminates.

The lemma cited above also tells us that, in a representation $\mathfrak{p} = \operatorname{Ann}_R \bar{x}$ (with an element $x \in R$ and $\bar{x} := x + J$), we may assume x to be homogeneous with respect to the $\mathbb{Z}^n$-grading. So let $d = \deg x \in \mathbb{Z}^n$ be the degree of the monomial (or term) x. Then the mapping

$$R(-d) \to V\,, \quad a \mapsto a \cdot \bar{x}\,,$$

is R-linear of degree 0, hence induces an isomorphism $(R/\mathfrak{p})(-d) \to R\bar{x}$ of $\mathbb{Z}^n$-graded R-modules. Note that this in turn induces an isomorphism of K-vector spaces in each $\mathbb{Z}^n$-degree.

Now define the ideal $J' := J + (x)$ of R and the K-vector space $V' := R/J'$. Since J' is generated by monomials, both J' and V' are $\mathbb{Z}^n$-graded R-modules. Choosing suitable sets of monomials in R, we find an isomorphism

$$V \cong V' \oplus J'/J$$

of K-vector spaces which retains this property in each $\mathbb{Z}^n$-degree (Corollary 4.5.2). Since there is an isomorphism

$$J'/J = R \cdot \bar{x} \cong (R/\mathfrak{p})(-d)$$

(of $\mathbb{Z}^n$-graded R-modules), the lemma follows: if $V' = 0$, we are done anyway; and if $V' \neq 0$, we may apply an inductive argument: the sequence $V, V', V'', \ldots$ corresponds to a chain

$$J \subseteq J' \subseteq J'' \subseteq \ldots$$

of ideals in the Noetherian ring R, hence must become stationary. □

The lemma remains true if we replace $\mathbb{Z}^n$ with an arbitrary lattice L; the only requirement on L is that all monomials are homogeneous in the L-grading. In fact, the lemma and the following corollary are equivalent, and are known as the *Stanley decomposition* of R/J.

Corollary 4.5.5: *Let L be a lattice, K a field, $R = K[X_1, \ldots, X_n]$ an L-graded polynomial ring such that all monomials are homogeneous, and $J = (X^{a_1}, \ldots, X^{a_m})$ a monomial ideal.*

Then there exists an isomorphism

$$R/J \cong \bigoplus_i (R/\mathfrak{p}_i)(-d_i)$$

of L-graded K-vector spaces, where the $\mathfrak{p}_i$ are finitely many prime ideals in R, each generated by some of the indeterminates, and the d_i are suitable shifts in L.

Proof: By hypothesis, the L-grading of R can be refined to the standard $\mathbb{Z}^n$-grading, so that there is a homomorphism $\deg : \mathbb{Z}^n \to L$ of groups. Applying Lemma 4.5.4, we obtain an isomorphism of $\mathbb{Z}^n$-graded K-vector spaces (and shifts in $\mathbb{Z}^n$), which then induces an isomorphism of L-graded K-vector spaces (and shifts in L), via the degree homomorphism. □

5. Computing the normalization of an affine monoid

Let L be a sublattice of $\mathbb{Z}^n$ for some n, and S an affine monoid in L. Recall that the integral closure of S in L is given by

$$\bar{S}_L = \{x \in L : mx \in S \text{ for some } m \in \mathbb{N},\ m > 0\},$$

including the normalization $\bar{S}$ of S, for the choice $L = \mathbb{Z}S$. We have also seen that the integral closure can be described geometrically: we have $\bar{S}_L = L \cap C$, where $C = C(S)$ denotes the cone generated by S in the vector space $\mathbb{R}^n$ (Lemma 3.1.2). Note that, by the same proposition, $\bar{S}_L$ is itself finitely generated. Moreover, it is positive if S is so (Lemma 3.4.3). In this case, $\bar{S}_L$ has a unique minimal generating set, namely, its finitely many irreducible elements (Proposition 3.4.6); this set is referred to as the *Hilbert basis*, $\operatorname{Hilb}\bar{S}_L$, of $\bar{S}_L$.

For the computation of $\operatorname{Hilb}\bar{S}_L$, the computer program NORMALIZ [BK] has been developed. It has already proved very useful in various investigations; see Villarreal [Vi]. In particular, it has played a crucial role in finding a counterexample to the Unimodular Covering Conjecture and the discrete Carathéodory property of normal affine monoids (Bruns and Gubeladze [BG2], Bruns et al. [BGHMW]). It is the purpose of this section to explain the algorithm used by NORMALIZ; the text is based on the article [BK2]. Most likely, all the ideas involved have appeared elsewhere, and we do not claim originality for them.

5.1. Computing the Hilbert basis. We start with an overview of this subsection. After having mentioned some applications of NORMALIZ, we analyze in more detail which of the lattice points in the cone C suffice to generate the integral closure of S in L. Then a reduction step follows: we shall see that, after performing a base change, we have to compute the integral closure of the affine monoid S of rank n in $\mathbb{Z}^n$. Since the best way to test a lattice point $x \in \mathbb{Z}^n$ for being contained in C is via the dual representation, we then discuss how to determine the support forms of C. In fact, NORMALIZ computes a triangulation of C along with the support forms. Therefore, the final topic is how to find those lattice points in each simplicial subcone of the triangulation that are necessary for the generation of the integral closure.

Applications of the program. In its present version, NORMALIZ requires the generators of S as input and allows only the choices $L = \mathbb{Z}S$ or $L = \mathbb{Z}^n$. These choices for L cover almost all potential applications; we mention two of them. Note that we have predefined special working modes for these purposes; see the documentation [BK].

Suppose first that we are given a *lattice polytope*, that is, a polytope P in $\mathbb{R}^n$ which is generated by lattice points, say $x_1, \dots, x_m \in \mathbb{Z}^n$. We wish to compute all lattice points in P. This turns out to be a task that is connected with the integral closure of an affine monoid: let S be the monoid generated by the vectors $(x_1, 1), \dots, (x_m, 1)$ in $\mathbb{Z}^{n+1}$; then, for a lattice point x in P, the vector $(x, 1)$ obviously belongs to $\operatorname{Hilb}\bar{S}_{\mathbb{Z}^{n+1}}$, the Hilbert basis of the integral closure of S in $\mathbb{Z}^{n+1}$.

Now consider a monomial ideal I in the polynomial ring $R = K[X_1,\dots,X_n]$ over a field K, say $I = (X^{a_1},\dots,X^{a_m})$. The task is to compute the integral closures of I and its Rees algebra. For this, one starts as in the case of a lattice polytope and embeds the exponent vectors $a_1,\dots,a_m$ into $\mathbb{N}^{n+1}$ (by adding a final coordinate 1); this amounts to multiplying the generators of I by an additional indeterminate. Furthermore, one adds the vectors $e_1,\dots,e_n$ of the canonical basis of $\mathbb{Z}^{n+1}$; these vectors represent the indeterminates $X_1,\dots,X_n$. At this point, the task is again performed by computing the integral closure of an affine monoid (in $\mathbb{Z}^{n+1}$).

For further applications, see Villarreal's book [Vi], and Bruns, Gubeladze and Trung [BGT].

Finiteness of the integral closure. We have already mentioned the importance of Lemma 3.1.2. We extend it to the following lemma, which is also a more precise rendering of Lemma 3.3.11.

Lemma 5.1.1: *Let S be an affine submonoid of the lattice $L \subseteq \mathbb{Z}^n$. Then $\bar{S}_L$ is a finitely generated S-module; in fact, if $x_1,\dots,x_m$ are generators of S, then $\bar{S}_L = \bigcup_z (z+S)$, where z runs through the finite set*

$$\operatorname{par}_L(x_1,\dots,x_m) := L \cap \{a_1x_1 + \dots + a_mx_m : a_i \in [0,1[\,\}.$$

In particular, this set, together with the vectors $x_1,\dots,x_m$, generates $\bar{S}_L$ (as a monoid).

Proof: Actually, this has already been shown in the proof of Lemma 3.1.2: we have

$$\bar{S}_L = L \cap C(S) = L \cap \mathbb{R}_+(x_1,\dots,x_m) = \mathbb{N}(x_1,\dots,x_m) + \operatorname{par}_L(x_1,\dots,x_m),$$

where the sum is the Minkowski sum of the two sets. □

The notation 'par' (introduced by Sebö [Se]) is suggested by the fact that the elements of $\operatorname{par}_L(x_1,\dots,x_m)$ are exactly the lattice points in the semi-open parallelepiped spanned by $x_1,\dots,x_m$. One should note that the coefficients a_i may as well be chosen in $\mathbb{Q} \cap [0,1[$; see Lemma 3.1.3.

In principle, Lemma 5.1.1 tells us how to determine the generators of $\bar{S}_L$: we only need to search these elements in a bounded subset of $\mathbb{Z}^n$. However, it is difficult to generate the candidates in an effective way without some preparations.

Reduction to a full rank embedding. In the first step we reduce the problem to computing the integral closure of S in a lattice L' such that $\operatorname{rk} L' = \operatorname{rk} S$. (Recall that $\operatorname{rk} S = \operatorname{rk} \mathbb{Z}S$, by definition.)

Applying the Elementary Divisors Algorithm, one finds a basis $e_1,\dots,e_n$ of $\mathbb{Z}^n$ and integers $\alpha_1,\dots,\alpha_n$ such that $f_i = \alpha_i e_i$, $i = 1,\dots,\operatorname{rk} L$, is a basis of L. After a linear transformation we can assume that $L = \mathbb{Z}^n$, and that we have to compute the integral closure of S in $\mathbb{Z}^n$. Henceforth the index L will therefore be dropped.

The Elementary Divisors Algorithm is applied again in order to find a basis $f_1,\dots,f_n$ of $\mathbb{Z}^n$ and integers $\beta_1,\dots\beta_r$, $r = \operatorname{rk} S$, such that $\beta_1 f_1,\dots,\beta_r f_r$ is a basis of $\mathbb{Z}S$. The integral closure of S in $\mathbb{Z}^n$ evidently coincides with the integral closure of S in $L' = \mathbb{Z}(f_1,\dots,f_r)$. Consequently, we may further assume that $\operatorname{rk} S = n$.

It is clear that, at the end of all the computations, the linear transformations inverse to those above have to be applied in order to rewrite the output in the coordinates of the input.

The program NORMALIZ allows for L only $\mathbb{Z}^n$ or $\mathbb{Z}S$. Therefore, one application of the Elementary Divisors Algorithm is sufficient: if $L = \mathbb{Z}^n$, then one chooses $L' = \mathbb{Z}(f_1,\dots,f_r)$, and if $L = \mathbb{Z}S$, one has to take $L' = \mathbb{Z}(\beta_1 f_1,\dots,\beta_r f_r)$.

To sum up, we are now in a position where we have to compute the integral closure of the affine monoid $S \subseteq \mathbb{Z}^n$ of rank n in the lattice $L = \mathbb{Z}^n$. As observed several times already, this amounts to computing the lattice points in the cone $C = C(S)$. Unfortunately, testing a lattice point $x \in \mathbb{Z}^n$ for being contained in C is a hard task if only the generators of C are known; the test is much easier if one has a dual representation of the form

$$C = \{x \in \mathbb{R}^n : \sigma_i(x) \geq 0 \text{ for } i = 1,\dots,s\} \tag{5}$$

at hand.

The Dual Cone Algorithm. Fortunately, the cone C of S is a finitely generated rational cone of dimension n in $V = \mathbb{R}^n$, and the bulk of the theory in Section 2 applies: by Corollary 2.9.3, C is in fact polyhedral, and due to rationality, there is a unique irreducible representation (5) of C, where $\sigma_1,\dots,\sigma_s$ are the support forms of C (Lemma 2.9.7, Proposition 2.10.5). A brief reminder: a support form σ_F of C is characterized by the facts that the respective hyperplane H (given by the vanishing of σ_F) leads to a facet F of C, and that σ_F has a representation with coprime integers. Next we discuss how to compute the σ_F.

Namely, the preparations for this purpose have been made: since C is finitely generated, so too is the dual cone

$$C^* = \{\varphi \in V^* : \varphi(x) \geq 0 \text{ for all } x \in C\},$$

and C coincides with its bidual C^{**} (upon identifying V with its bidual V^{**}; see Proposition 2.9.2). In the language of dual objects, a (finite) system of generators of C^* is equivalent to a polyhedral representation of C; and a *minimal* system of generators then corresponds to an *irreducible* polyhedral representation. (Note that C^* is positive since C has full dimension.)

The algorithm for the computation of the dual cone starts just as previous to Proposition 2.9.1: we first search for n linearly independent vectors $x_1,\dots,x_n$ among the generators of C (or S, respectively); these constitute a basis of V then. We compute the respective dual basis $\varphi_1,\dots,\varphi_n$ of V^* and therewith have also found a generating system of the dual cone of $C_0 = \mathbb{R}_+(x_1,\dots,x_n)$. This initialization is useful because it simultaneously starts a triangulation of the cone C. This triangulation will be needed for the computation of the Hilbert basis.

Proposition 2.9.1 now offers a constructive algorithm for finding a generating system of the dual cone if one adds another generator y of C (or S): suppose that

$\varphi_1, \ldots, \varphi_t$ generate the dual cone of $D = \mathbb{R}_+(x_1, \ldots, x_m)$; for each pair $1 \le i, j \le t$ with $\varphi_i(y) > 0$ and $\varphi_j(y) < 0$, set

$$\psi_{ij} := \varphi_i(y) \cdot \varphi_j - \varphi_j(y) \cdot \varphi_i\,;$$

then the dual cone of $\tilde{D} = \mathbb{R}_+(x_1, \ldots, x_m, y)$ is generated by the ψ_{ij} and those φ_k such that $\varphi_k(y) \ge 0$. Clearly, this enables an inductive proceeding, and after finitely many steps we have integrated all generators of C (or S, respectively); at this point, a generating system of C^* has been found.

One should note that all computations can be performed with integers: since $S \subseteq \mathbb{Z}^n$, the generators of C can be chosen to be lattice points, and therefore, it is possible to find an integral representation of all linear forms in each inductive step. This is precisely what NORMALIZ does, and even more: it always clears denominators and removes common divisors during its computations, therewith producing the most effective (and unique) representation of each linear form.

The second, and final, remark refers to selecting a *minimal* generating system of C^*: of course, we wish to omit superfluous linear forms in order to prevent unnecessary computations when testing lattice points for being contained in C later on. In the initial step the generators $\varphi_1, \ldots, \varphi_n$ of $(C_0)^*$ are minimally chosen already: obviously, each φ_i belongs to a facet of C_0. The following lemma clarifies which of the dual forms (found in an inductive step) belong to a facet.

Lemma 5.1.2: *Let D be an n-dimensional cone in $\mathbb{R}^n$, generated by the vectors $x_1, \ldots, x_m$ from $\mathbb{Z}^n$. Let $\varphi_1, \ldots, \varphi_t$ be the support forms of D, corresponding to its facets $F_1, \ldots, F_t$. Suppose further that the ψ_{ij} have been computed as in Proposition* 2.9.1*, and have (been replaced by) a representation with coprime integers. Then, for $y \in \mathbb{Z}^n$, the support forms of $\tilde{D} = D + \mathbb{R}_+ y$ are given by*

(1) *those ψ_{ij} such that $F_i \cap F_j \not\subseteq F_k$ for $k \neq i, j$, and*

(2) *those φ_k such that $\varphi_k(y) \ge 0$.*

Proof: We have to show that, in (1) and (2), precisely those dual forms which belong to a facet of $\tilde{D}$ are chosen from the generating system of $\tilde{D}^*$. So let H_{ij} and H_k denote the respective hyperplanes in $\mathbb{R}^n$ (given by the vanishing of ψ_{ij} and φ_k, respectively), and G_{ij} and G_k the respective faces of $\tilde{D}$. (Note that we only consider indices k with $\varphi_k(y) \ge 0$ here.)

Our task now is to clarify that (or when) these faces contain $n-1$ linearly independent vectors from $\tilde{D}$; of course, it suffices to search for these vectors among the generators $x_1, \ldots, x_m, y$ (Lemma 3.2.1). As for the faces G_k, this is immediate: one already finds $n-1$ linearly independent vectors among $x_1, \ldots, x_m$, since φ_k is a support form of D.

Finally, let us consider the faces G_{ij}. Obviously, we have $\psi_{ij}(y) = 0$, so that y is contained in G_{ij}. In order to decide which of the other generators $x_1, \ldots, x_m$ of $\tilde{D}$

are also contained in G_{ij}, the intersection of G_{ij} and D suffices:

$$\begin{aligned} D\cap G_{ij} &= \{x\in D : \psi_{ij}(x)=0\} \\ &= \{x\in D : \varphi_i(x)=\varphi_j(x)=0\} = D\cap H_i\cap H_j = F_i\cap F_j\,. \end{aligned}$$

Since y is not contained in H_i, whereas the intersection of H_{ij} and D is (as we have just seen), we find that

$$\dim G_{ij} = \dim(F_i\cap F_j)+1\,.$$

Consequently, G_{ij} is a facet of $\tilde{D}$ if and only if $\dim(F_i\cap F_j)=n-2$. By Corollary 2.7.8, this happens if and only if the condition in (1) is satisfied. □

Before continuing, we offer a geometric interpretation of the procedure. To this end, we remind the reader of a notion introduced in Subsection 2.11: we say that a subset X of D is *visible* from y if, for each $x\in X$, the line segment $[y,x]$ from y to x intersects D exactly in x. It is geometrically evident that one finds the facets of $\tilde{D}$ by taking first those facets of D that do not separate y from D, and second those hyperplanes that pass through y and the $(n-2)$-dimensional faces of D that bound the part of D that is visible from y. Exactly these hyperplanes are specified by Proposition 2.9.1 and Lemma 5.1.2: an $(n-2)$-dimensional face is contained in exactly two facets, and it bounds the visible area if exactly one of these facets is visible from y.

However, the question remains how to efficiently check the condition in item (1) in the lemma. The best way to do this is via an 'incidence matrix' which looks like the following:

	x_1	$\cdots$	x_m
φ_1			
$\vdots$	$(=0)$	or	(>0)
φ_t			

The entries are obtained by evaluating each support form of D on all the generators; note that this yields non-negative integers in our setup. The face $F_i\cap F_j$ of D is generated by those vectors x_k for which $\varphi_i(x_k)=\varphi_j(x_k)=0$, and NORMALIZ selects the respective columns (submatrix) of the matrix. Now, condition (1) is checked by searching for another, third, row of zeros in this submatrix.

Once the dual cone, equivalently, the support forms, of C have been computed, we can decide whether the original monoid S is positive: by Lemma 3.4.3, this happens if and only if C is positive, and by Lemma 2.10.1, this is equivalent to the fact that C^* has dimension n.

Thus, at this point NORMALIZ tests whether S is positive. If not, it stops. It would not be difficult to extend the program in such a way that it covers the general case. Let U be the kernel of the homomorphism $\sigma=(\sigma_1,\dots,\sigma_s):\mathbb{Z}^n\to\mathbb{Z}^s$. Then the image T of S in $\mathbb{Z}^n/U$ is a positive affine monoid. It is enough to lift the Hilbert basis of $\bar{T}$ back to $\mathbb{Z}^n$, because $\bar{S}$ is generated by U and preimages of the Hilbert basis of $\bar{T}$. (We are still assuming that $S\subseteq\mathbb{Z}^n$ has rank n.)

At this point we can also discuss how to find the irreducible elements of $\bar{S}$, and thus $\operatorname{Hilb}\bar{S}$, once a system of pairwise distinct generators $x_1,\dots,x_m \neq 0$ of $\bar{S}$ is known: x_i is irreducible if $x_i - x_j \notin \bar{S}$ for all $j \neq i$. This criterion holds for arbitrary S; however, the condition $x_i - x_j \notin S$ is difficult to verify in general. For $\bar{S}$, it is easy: we simply test whether the condition $\sigma_F(x_i - x_j) \geq 0$ is violated for at least one facet F of C.

To sum up, we are now in a position where we can efficiently test a lattice point $x \in \mathbb{Z}^n$ for being contained in $\bar{S}$ (or C); however, we still do not know how to produce a system of generators of $\bar{S}$ (apart from the information given in Lemma 5.1.1). We would still be forced to start a naive search for the lattice points in C (or a bounded part of it); in order to avoid this, we decompose C into simplicial subcones.

Computing the triangulation. The notion of a triangulation Δ of C has been introduced in Subsection 2.11. The most important property that we need here is the covering property: C is the union of all cones $\delta \in \Delta$. Therefore, the search for lattice points in C is reduced to the search in each simplicial subcone δ. Recall also that there is a triangulation Δ such that each $\delta \in \Delta$ is generated by some of the generators of C (or S, originally); see Corollary 2.11.5.

We mention a computational aspect: since a triangulation Δ is uniquely determined by its maximal members (with respect to inclusion), it suffices to save those $\delta \in \Delta$ such that $\dim \delta = n$. Therefore, the triangulation can be described by a list of n-tuples of vectors in $\mathbb{Z}^n$, where each n-tuple contains the generators of an n-dimensional simplicial cone.

We briefly recall the construction of Δ: let $x_1,\dots,x_m \in \mathbb{Z}^n$ be the generators of C. In the computation of the dual cone C^* we have started with a simplicial subcone C_0 generated by a linearly independent subset of $\{x_1,\dots,x_m\}$. It has a trivial triangulation by its faces (including C_0 itself). From then on, an inductive procedure starts, as described in Proposition 2.11.4; this procedure runs hand in hand with the Dual Cone Algorithm. (Of course, NORMALIZ uses the support forms, rather than arbitrary dual forms, when applying Proposition 2.11.4.)

With the notation of Lemma 5.1.2, the 'new' n-dimensional simplicial subcones found in an inductive step are generated by y and the $(n-1)$-dimensional visible cones δ' in the triangulation Δ' of D. Such a subcone δ' is visible if and only if it lies in a facet of D that is visible from y, and the visible facets F are characterized by the fact that $\varphi_F(y) < 0$. This makes it easy for NORMALIZ to find the new n-dimensional members of the triangulation, since the support forms of D are known. (However, note that a facet of D itself is not simplicial in general, and even if it is: Δ' may subdivide it, if the given generating set of D is not minimal; see Figure 8.)

For purposes of efficiency, especially when the Hilbert series is to be computed (see Subsection 5.2), we subdivide the triangulation into *blocks* and *subblocks*; in fact, this subdivision plays a crucial role for the performance of NORMALIZ, but is not of fundamental importance for the algorithm. Every block consists of exactly those simplicial cones found in an inductive step, as described above. That is, whenever a new generator y has to be integrated, a new block starts and collects all simplicial cones which are constructed due to y (and which, in particular, contain y). The vector y is then called the *characteristic vector* of the block. A similar, but slightly more difficult operation leads to subblocks within the blocks.

In the example in Figure 8 (imagine the picture as a cross-section of the cones), the triangulation of the cone D generated by $x_1,\dots,x_5$ consists of three blocks already, each of them containing only one maximal simplicial cone. Now, upon

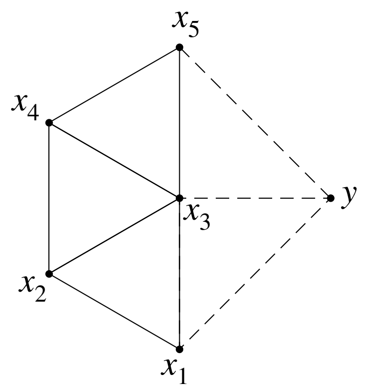

Figure 8

adding a further generator y, a fourth block starts, containing precisely two maximal simplicial cones δ_1 and δ_2, namely, those generated by (x_1,x_3,y) and (x_3,x_5,y). (The other members of this block are those faces of δ_1 and δ_2 which contain y.) Here, y is the characteristic vector of the fourth block.

Let us sum up again: suppose that we have found a system of generators for the monoid $\delta\cap\mathbb{Z}^n$ for each δ in a triangulation Δ of C. Then the union of all these systems obviously generates $\bar{S}$. We have already constructed a triangulation Δ, and each $\delta\in\Delta$ is specified by a set of integral, linearly independent vectors generating δ. It only remains to find the generators of $\bar{S}$ if S is simplicial.

Simplicial cones. Let $x_1,\dots,x_n$ be linearly independent elements of $\mathbb{Z}^n$, let C be the cone spanned by them and S the affine monoid they generate. Then, according to Lemma 5.1.1, $\bar{S}=C\cap\mathbb{Z}^n$ is generated by $x_1,\dots,x_n$ themselves and the set

$$\operatorname{par}(x_1,\dots,x_n)=\mathbb{Z}^n\cap\{q_1x_1+\dots+q_nx_n : q_i\in[0,1[\}\,.$$

Lemma 5.1.3: *The set* $\operatorname{par}(x_1,\dots,x_n)$ *contains exactly one representative from each residue class of* $\mathbb{Z}^n$ *modulo* $U:=\mathbb{Z}(x_1,\dots,x_n)$. *Therefore,*

$$\operatorname{card}\operatorname{par}(x_1,\dots,x_n)=\operatorname{card}(\mathbb{Z}^n/U)=|\det(x_1,\dots,x_n)|\,.$$

Moreover, $\bar{S}$ *is the disjoint union of the sets* $z+S$, $z\in\operatorname{par}(x_1,\dots,x_n)$.

Proof: The first statement is evident, and it implies the first equation. The second equation results from the Elementary Divisors Theorem. That $\bar{S}$ is the union of the sets $z+S$ has been shown in Lemma 5.1.1, and that the union is disjoint follows immediately from the fact that the $z \in \operatorname{par}(x_1,\dots,x_n)$ represent different residue classes. □

In the language of commutative algebra: let K be a field; then $\operatorname{par}(x_1,\dots,x_n)$ is a basis of the free $K[S]$-module $K[\bar{S}]$, and $K[S]$ is actually a polynomial ring over K.

Together with $x_1,\dots,x_n$, the set $\operatorname{par}(x_1,\dots,x_n)$ generates $\bar{S}$. Therefore, it is enough to find an efficient method for producing $\operatorname{par}(x_1,\dots,x_n)$ from $x_1,\dots,x_n$.

First one applies the Elementary Divisors Algorithm to find a basis $u_1,\dots,u_n$ of $\mathbb{Z}^n$, and positive integers $\lambda_1,\dots,\lambda_n$ such that $\lambda_1 u_1,\dots,\lambda_n u_n$ is a basis of $\mathbb{Z}S$. Clearly $d := |\det(x_1,\dots,x_n)| = \lambda_1\cdots\lambda_n$, and $d\mathbb{Z}^n \subseteq \mathbb{Z}S$, since $\mathbb{Z}^n/\mathbb{Z}S$ is a direct sum of n cyclic groups of orders $\lambda_1,\dots,\lambda_n$.

The residue classes of $\mathbb{Z}^n$ modulo $\mathbb{Z}S$ are represented by the vectors

$$u = b_1u_1+\cdots+b_nu_n, \qquad b_i = 0,\dots,\lambda_i - 1, \quad i = 1,\dots,n.$$

Each such vector u has a representation $u = a_1x_1+\cdots+a_nx_n$ with *rational* coefficients a_i. Now we set $q_i := a_i - \lfloor a_i \rfloor$, so that $u' := q_1x_1+\cdots+q_nx_n$ represents the residue class of u and belongs to $\operatorname{par}(x_1,\dots,x_n)$. Note that d is a suitable common denominator for the a_i, since $d\mathbb{Z}^n \subseteq \mathbb{Z}S$. Therefore one can keep all the coefficients integral by first passing to du and dividing by d at the end.

5.2. Computing the Hilbert series. Because of the embedding $\sigma : S \to \mathbb{N}^s$ via its support forms, each positive affine monoid S can be positively graded, that is, there exists a monoid homomorphism $\deg : S \to \mathbb{N}$ such that $\deg(x) = 0$ if and only if $x = 0$. Then $K[S]$ is a positively graded K-algebra (where K is an arbitrary field). If deg is the restriction of a $\mathbb{Z}$-linear form on L (and it always is after multiplication by a positive integer), then S is a graded submonoid of $\bar{S}_L$. If, in addition, S is generated by all $x \in S$ with $\deg(x) = 1$, then we say that S is *homogeneous with respect to L* (and simply *homogeneous* if $L = \mathbb{Z}S$). In this case NORMALIZ can compute the Hilbert function of $\bar{S}_L$, given by

$$H(\bar{S}_L,k) = \operatorname{card}\{x \in \bar{S}_L : \deg(x) = k\}.$$

We start by recalling some facts from Subsection 4.2 and, thereby, connecting this notion with its use in commutative algebra.

The Hilbert series of an affine monoid. Suppose that A is a positively graded affine monoid; one has fixed a homomorphism $\deg : \mathbb{Z}A \to \mathbb{Z}$ such that $\deg(A) \subseteq \mathbb{N}$ and 0 is the only element of A having degree 0. Then the set $A_k = \{x \in A : \deg(x) = k\}$ is finite for each $k \in \mathbb{N}$ (Lemma 3.4.5), and we can consider the Hilbert function, $H(A,k) = \operatorname{card} A_k$, and the Hilbert series,

$$H_A(X) = \sum_{k=0}^{\infty} (\operatorname{card} A_k)\cdot X^k,$$

of A. If K is a field, then the monoid algebra $K[A]$ inherits the grading, and the Hilbert series of A is just the Hilbert series of $K[A]$. Therefore, $H_A(X)$ has all the properties that are known for Hilbert series of positively graded K-algebras (see [BH, Chap. 4]). In particular, $H_A(X)$ represents a rational function,

$$H_A(X) = \frac{Q(X)}{(1-X^{d_1})\cdots(1-X^{d_n})},$$

where $Q(X)$ is a polynomial, $n = \operatorname{rk} A$, and $d_1,\dots,d_n$ are positive integers.

The situation further simplifies if A is *almost homogeneous*. This means that there exists an affine submonoid $A_0 \subseteq A$ which is generated by elements of degree 1 and over which A is a finitely generated module, that is, there exist $x_1,\dots,x_m \in A$ such that $A = \bigcup_{i=1}^m (x_i + A_0)$. Then $K[A]$ is a finitely generated module over the homogeneous ring $K[A_0]$, and therefore, $H_A(X)$ can be represented in the form

$$H_A(X) = \frac{Q(X)}{(1-X)^n}.$$

For $k \gg 0$, the Hilbert function $H(A,k)$ is given by the Hilbert polynomial $P_A(k)$; see Subsection 4.3. It is a polynomial of degree $n-1$ with leading coefficient $e(A)/(n-1)!$, where $e(A) = Q(1)$ is the *multiplicity* of A.

In NORMALIZ the role of A_0 is played by the given affine monoid S, and that of A is played by the integral closure $\bar{S}$. (We assume, as before, that $S \subseteq \mathbb{Z}^n$, $\operatorname{rk} S = n$, and the integral closure is taken with respect to $\mathbb{Z}^n$.) We want to compute the Hilbert series of $\bar{S}$. In principle this would be possible for the general case, but so far it has only been implemented in the almost homogeneous case (simply called 'homogeneous' in the NORMALIZ documentation), and from now on we restrict ourselves to this case.

Computing the Hilbert series. NORMALIZ has computed a triangulation Δ of the cone C generated by S such that each simplicial cone $\delta \in \Delta$ is generated by elements of S; by assumption, these elements have degree 1 in $\bar{S}$. The triangulation defines a disjoint decomposition

$$C = \biguplus_{\delta\in\Delta} \operatorname{relint}\delta;$$

see Lemma 2.11.2. We set

$$\omega_\delta := \mathbb{Z}^n \cap \operatorname{relint}\delta.$$

It then follows that

$$H_{\bar{S}}(X) = \sum_{\delta\in\Delta} H_{\omega_\delta}(X),$$

where the terms on the right-hand side are defined in an obvious way.

Therefore, in order to compute $H_{\bar{S}}(X)$, two tasks have to be carried out, namely,

(1) the decomposition of C into the cones $\delta \in \Delta$, and

(2) the computation of $H_{\omega_\delta}(X)$ for each $\delta \in \Delta$.

The most time consuming part of NORMALIZ is step (1) because of the extreme combinatorial complexity of triangulations in general: recall that NORMALIZ only saves the maximal (that is, n-dimensional) simplicial cones in Δ; the problem that arises when running through these maximal members δ of Δ is to decide whether or not a face of δ has already been considered before. (Note that δ has 2^n faces, by Lemma 2.11.1.)

Of course, for the computation of the Hilbert basis, it is enough to consider the maximal simplicial cones in Δ, and it would be foolish to insist on a disjoint decomposition. (Some vectors are tested more than once for being members of the Hilbert basis if they belong to two or more maximal simplicial subcones; but this effect is negligible.)

However, as for the Hilbert series, one cannot avoid the disjoint decomposition. To this end, we have introduced the subdivision of the list of maximal simplicial cones into blocks and subblocks; see the example in Figure 8. Assume that we have to process a maximal simplicial cone δ from block B. Then, by construction, we only have to investigate those faces ε of δ which contain the characteristic vector y of B; moreover, if ε has been accounted before, then this must have happened within B. In particular, when δ happens to be the first member of B, then all these faces ε have to be accounted. (There are 2^{n-1} such faces.) Thus, using the concept of blocks and characteristic vectors reduces the investigation to the current block and speeds up the computation considerably.

As for step (2), we denote by $x_1,\dots,x_r$ the linearly independent degree 1 elements of S that generate the cone $\delta \in \Delta$. The monoid S_δ they generate is free, and so $H_{S_\delta}(X) = 1/(1-X)^r$. Furthermore, in analogy with Lemma 5.1.3, one has a disjoint decomposition

$$\omega_\delta = \biguplus_{x \in \operatorname{par}'(x_1,\dots,x_r)} (x + S_\delta)\,,$$

where

$$\operatorname{par}'(x_1,\dots,x_r) := \mathbb{Z}^n \cap \{q_1x_1 + \cdots + q_rx_r : q_i \in \,]0,1]\}\,.$$

Therefore

$$H_{\omega_\delta}(X) = \frac{\sum_{k=1}^{r} (\operatorname{card} B_k)\cdot X^k}{(1-X)^r}\,,$$

where $B_k := \{x \in \operatorname{par}'(x_1,\dots,x_r) : \deg(x) = k\}$.

Final remarks. (a) Often one is only interested in a single numerical invariant, namely the *multiplicity* $e(\bar{S})$. (It coincides with the multiplicity of S if $\mathbb{Z}S = \mathbb{Z}^n$; in general one has $e(\bar{S}) = e(S)\cdot \operatorname{card}(\mathbb{Z}^n/\mathbb{Z}S)$.) The different maximal simplicial cones in the triangulation intersect each other only in lower-dimensional cones, so that the leading coefficient of the Hilbert polynomial can be calculated without taking care of the lower-dimensional cones. Then

$$e(\bar{S}) = \sum_{\substack{\delta \in \Delta \\ \dim \delta = n}} e(\bar{S}_\delta)\,,$$

where $\bar{S}_\delta = \delta \cap \mathbb{Z}^n$. Let $x_1,\dots,x_n$ be the linearly independent degree 1 generators of δ. Then $S_\delta = \mathbb{N}(x_1,\dots,x_n)$ is a free affine monoid and therefore of multiplicity 1. The integral closure $\bar{S}_\delta$ is a free S_δ-module (Lemma 5.1.3), and therefore, its multiplicity coincides with the number of elements in its basis $\operatorname{par}(x_1,\dots,x_n)$. To sum up,

$$e(\bar{S}_\delta) = \operatorname{card}(\operatorname{par}(x_1,\dots,x_n)) = |\det(x_1,\dots,x_n)|\,.$$

Therefore it is not necessary to compute the Hilbert basis in order to find the multiplicity. We offer the option '`-v`' for NORMALIZ. It restricts all computations to multiplicities and those data which determine the triangulation.

The letter '`v`' has been chosen since the multiplicity of S can be interpreted as the normalized volume of the polytope spanned by the generators of S in the hyperplane of degree 1 elements. Thus, '`normaliz -v`' can be used for the computation of volumes of lattice polytopes.

(b) It is not necessary to compute $\operatorname{par}'(x_1,\dots,x_r)$ separately. In fact, we have $y \in \operatorname{par}'(x_1,\dots,x_r)$ if and only if

$$(x_1+\dots+x_r) - y \in \operatorname{par}(x_1,\dots,x_r)\,.$$

We use this observation as follows.

The n-dimensional simplicial cones $\delta \in \Delta$ are scanned for the computation of the Hilbert basis. Suppose that δ is spanned by the degree 1 elements $x_1,\dots,x_n$. Then the elements $x \in \operatorname{par}(x_1,\dots,x_n)$ are computed since they are candidates for the Hilbert basis. We can write $x = q_1 x_{i_1} + \dots + q_r x_{i_r}$ with $0 < q_i < 1$. For each subset $J \subseteq \{1,\dots,n\}$ with $\{i_1,\dots,i_r\} \subseteq J$, the vector $\sum_{j\in J} x_j - x$ belongs to $\operatorname{par}'(x_j : j \in J)$, and all the vectors necessary for the computation of the Hilbert series are produced by this method.

(c) We have discussed above how to compute the numerator polynomial

$$Q(X) = h_0 + h_1 X + \dots + h_{n-1} X^{n-1}$$

in the Hilbert series $H_{\bar{S}}(X) = Q(X)/(1-X)^n$ of $\bar{S}$. (In fact, in order to see that $Q(X)$ has no higher terms, one can write the Hilbert series of $\bar{S}$ as an alternating sum of Hilbert series $H_{\bar{S}_\delta}$, where each S_δ is a free monoid, as above; by Lemma 5.1.3, the claim is clear for each summand $H_{\bar{S}_\delta}$.) The coefficients occurring here constitute the *h-vector* $(h_0,\dots,h_{n-1})$ of $\bar{S}$. It is clear that, with the h-vector at hand, one can determine the Hilbert polynomial $P_{\bar{S}}$ of $\bar{S}$, and in fact, the h-vector even provides information about from when on the Hilbert function and the Hilbert polynomial coincide: in our situation, we always have $H(\bar{S},k) = P_{\bar{S}}(k)$ for $k \geq 0$.

6. Growth of Hilbert functions

We now turn our attention to the growth of the Hilbert function of certain modules. This study is motivated by the following observations. Let R be a homogeneous ($\mathbb{N}$-graded) K-algebra and M a finitely generated $\mathbb{Z}$-graded R-module of Krull dimension $d > 0$. Then, according to Proposition 4.3.1, the Hilbert function of M is of polynomial type of degree $d-1$. In particular, there is a constant $c > 0$ such that

$$H(M,n) \geq c \cdot n^{d-1}$$

for all $n \in \mathbb{Z}$, $n \gg 0$. Loosely speaking, we could say that the Hilbert function grows 'with degree $d-1$'.

Now, more generally, suppose that the K-algebra R is $\mathbb{Z}^m$-graded, and that we are given a finitely generated $\mathbb{Z}^m$-graded R-module M. The question is: are there again constants $c > 0$ and $d \geq 0$ such that

$$H(M,\ell) \geq c \cdot \|\ell\|^d$$

for all $\ell \in \mathbb{Z}^m$, $\|\ell\| \gg 0$ (for which $M_\ell \neq 0$)? In any case, it is easy to see that one can no longer take $d = \dim M - m$; namely, consider the example $R = M = K[X,Y,Z]$, $\mathbb{Z}^2$-graded by setting $\deg X := (1,0)$ and $\deg Y := \deg Z := (0,1)$. Then, for $n \in \mathbb{N}$, we have

$$M_{(n,0)} = K \cdot X^n \quad \text{and} \quad M_{(0,n)} = K(Y^i Z^j : i + j = n),$$

so that $H(M,(n,0)) = 1$ and $H(M,(0,n)) = n+1$. Consequently, we arrive at estimates

$$H(M,\ell) \geq 1 \cdot \|\ell\|^0 \quad \text{and} \quad H(M,\ell) \geq 1 \cdot \|\ell\|^1$$

for $\ell \in \mathbb{N} \times \{0\}$ and $\ell \in \{0\} \times \mathbb{N}$, respectively. Note that the exponents here are best, although $\dim M = 3$. In addition, we observe another effect: the degree of growth may depend on the direction in which the Hilbert function is examined.

The aim of this section is to show a result that can, loosely speaking, be stated as follows: if the Hilbert function grows with a certain (fixed) degree d along each arithmetic progression in the grading monoid, then it globally grows with degree d. Originally, this study was motivated by a result of Bruns and Gubeladze [BG, Theorem 4.4.3]; it concerns the growth of Hilbert functions of certain multigraded modules and algebras. Roughly speaking, it says that the Hilbert function takes values $\leq C$ only at finitely many graded components, provided that this holds along each arithmetic progression in the grading monoid. This theorem can be applied to the problem of finding the minimal number of generators of divisorial ideals (representing the full divisor class group $\mathrm{Cl}(R)$ of an affine monoid algebra $R = K[S]$): this problem is transformed into a statement on a multigraded Hilbert function which can be controlled along arithmetic progressions. Our intention now is to generalize [BG, Theorem 4.4.3].

6.1. The main theorems. Let K be a field and $L = \mathbb{Z}^m$ a lattice. Let R be a Noetherian L-graded K-algebra for which R_0 is a finitely generated K-algebra, and M

a finitely generated L-graded R-module. Let H be the L-graded Hilbert function of M, that is, $H(M,\ell) := \dim_K M_\ell$ for all $\ell \in L$.

Now choose $a, b \in L$, $b \neq 0$, and consider the ring $R_{(b)}$ and its module $M_{(a,b)}$, defined as

$$R_{(b)} := \bigoplus_{g \in \mathbb{N}b} R_g \quad \text{and} \quad M_{(a,b)} := \bigoplus_{g \in a+\mathbb{N}b} M_g\,,$$

respectively. By Proposition 4.1.3, $R_{(b)}$ is a (finitely generated) positively graded K-algebra, and $M_{(a,b)}$ is a finitely generated $R_{(b)}$-module. Hence Proposition 4.4.2 applies and yields that the Hilbert function of $M_{(a,b)}$ is of quasi-polynomial type. As we shall see in Subsection 6.3 (Lemma 6.3.3), the respective quasi-polynomial has a well-defined degree d, namely, $d = \dim M_{(a,b)} - 1$. Note that we are in a position to determine this degree, since the Krull dimension is computable in principle. Since the Hilbert functions of M and $M_{(a,b)}$ coincide on $a + \mathbb{N}b$, we conclude that

$$H(M, a+nb) \geq c(a,b) \cdot n^d$$

for all $n \in \mathbb{N}$, $n \gg 0$, with a positive constant $c(a,b)$ depending only on a and b. Conversely, note that if this inequality holds for some $d \in \mathbb{N}$, then we must have $\dim M_{(a,b)} \geq d+1$.

We introduce one more notion: the *support* of R in L, $\operatorname{Supp}_L R$, is the set of those degrees $\ell \in L$ in which R is occupied, that is, such that $R_\ell \neq \{0\}$. Likewise, we define the *support* $\operatorname{Supp}_L M$ of M in L. The next theorem is our main result on the growth of Hilbert functions. It generalizes [BG, Theorem 4.4.3] and answers the question to what extent local growth (along arithmetic progressions) determines global growth. Note that we do not assume that $R_0 = K$; the graded components of R and M may have arbitrary (possibly infinite) K-dimension.

Theorem 6.1.1: *Let K be a field and $L = \mathbb{Z}^m$ a lattice. Let R be a Noetherian L-graded K-algebra for which R_0 is a finitely generated K-algebra, and M a finitely generated L-graded R-module. Let S be an affine submonoid of L which contains* $\operatorname{Supp}_L R$, *and T a finitely generated S-submodule of L which contains* $\operatorname{Supp}_L M$. *Finally, let H be the L-graded Hilbert function of M, that is, $H(M,\ell) := \dim_K M_\ell$ for all $\ell \in L$.*

Now suppose that there exists an integer $d \geq 0$ such that

$$H(M, a+nb) \geq c(a,b) \cdot n^d$$

for all $a \in T$, $b \in S$, $b \neq 0$, and $n \gg 0$, where $c(a,b) > 0$ is a constant depending on a and b only. Then there exists a constant $c > 0$ such that

$$H(M,t) \geq c \cdot \|t\|^d$$

for all $t \in T$, $\|t\| \gg 0$, where $\|\cdot\|$ represents an arbitrary norm on $\mathbb{R}^m$.

Note that the hypothesis on the growth of H along arithmetic progressions can be stated equivalently as follows: for each $t \in a + \mathbb{N}b$, $\|t\| \gg 0$, we have

$$H(M,t) \geq c(a,b) \cdot \|t\|^d\,;$$

in fact, $\|a+nb\| = n \cdot \|a/n+b\|$, and the latter factor converges to the positive value $\|b\|$ for $n \to \infty$. The reader may find the following corollary useful; since we restrict ourselves to as few objects as possible, it may offer a better understanding of the context at first.

Corollary 6.1.2: *Let K be a field. Let R be a Noetherian $\mathbb{Z}^m$-graded K-algebra with $R_0 = K$, and M a finitely generated $\mathbb{Z}^m$-graded R-module. Finally, let H be the $\mathbb{Z}^m$-graded Hilbert function of M, that is, $H(M,\ell) := \dim_K M_\ell$ for all $\ell \in \mathbb{Z}^m$. Now suppose that there exists an integer $d \geq 0$ such that*

$$H(M, a+nb) \geq c(a,b) \cdot n^d$$

for all $a,b \in \mathbb{Z}^m$, $b \neq 0$, and $n \gg 0$. Then there exists a constant $c > 0$ such that

$$H(M,\ell) \geq c \cdot \|\ell\|^d$$

for all $\ell \in \mathbb{Z}^m$, $\|\ell\| \gg 0$.

Proof of Theorem 6.1.1: We use induction on m. In the case when $m = 1$, we use Lemma 3.1.1: S is a finite union of arithmetic progressions, hence so is T, say $T = \bigcup_i (a_i + \mathbb{N}b_i)$. Therefore, in this case, the assertion is covered by the hypothesis of the theorem: the constant c can be chosen as the minimum of the finitely many constants $c(a_i, b_i)$. Note here that, in principle, the general case is handled in the same way.

As a first step, we want to extend the hypothesis on the Hilbert function from a '1-dimensional' condition to a '1-codimensional' condition by an application of the induction hypothesis.

So let L' be a sublattice of L with $\operatorname{rk} L' < m$, pick some $\ell_0 \in L$ and define

$$S' := S \cap L', \quad T' := T \cap (\ell_0 + L'), \quad R' := \bigoplus_{g \in L'} R_g \quad \text{and} \quad M' := \bigoplus_{g \in \ell_0 + L'} M_g .$$

By Corollary 4.1.4, S' is an affine monoid, and T' a finitely generated S'-module. Moreover, Proposition 4.1.3 implies that R' is a finitely generated R_0-algebra and M' is a finitely generated R'-module. Clearly, the support of R' is contained in S', and the support of M' is contained in T'. (Note that M' can be viewed as an L'-graded module.) By definition of M', the Hilbert functions of M and M' coincide on $\ell_0 + L'$. Therefore,

$$H(M', a+nb) \geq c(a,b) \cdot n^d$$

for all $a \in T'$, $b \in S'$, $b \neq 0$, and $n \gg 0$, with a positive constant $c(a,b)$. The induction hypothesis now implies that

$$H(M,t) = H(M',t) \geq c(\ell_0, L') \cdot \|t\|^d \tag{6}$$

for all $t \in T'$, $\|t\| \gg 0$, with a positive constant $c(\ell_0, L')$ depending on ℓ_0 and L'. We shall fall back on condition (6) later on. At first, we show that R may be replaced with a polynomial ring, and M with a direct sum of shifted polynomial rings; this procedure often is referred to as the *Stanley decomposition* in the literature (see, for instance, [Va, Sec. 1.4]).

By Proposition 4.1.3 (b), R is a finitely generated R_0-algebra (since L itself is a finitely generated monoid), and thus a finitely generated K-algebra. We represent R as the residue class ring of an L-graded polynomial ring $P = K[X_1,\dots,X_r]$ in a natural way. Note that, in particular, the monomials in P are homogeneous in the L-grading.

Now we show that P_0 inherits the property of being a finitely generated K-algebra (by hypothesis, R_0 has this property). Both the set $\mathscr{M}(P)$ of all monomials in P and the set $\mathscr{M}(\hat{P})$ of all monomials in the Laurent polynomial ring $\hat{P} = K[X_1^{\pm 1},\dots,X_r^{\pm 1}]$ is an affine monoid (the former is isomorphic to $\mathbb{N}^r$, the latter to $\mathbb{Z}^r$). Now $\hat{P}$ also carries a natural L-grading, and the kernel of the degree homomorphism $\deg : \mathscr{M}(\hat{P}) \to L$ is (a subgroup and) again an affine monoid. Being the intersection of the two affine monoids $\mathscr{M}(P)$ and $\ker(\deg)$, the set $\mathscr{M}(P_0)$ of all monomials in P_0 is a finitely generated monoid, by Corollary 4.1.4. This shows that P_0 is a finitely generated K-algebra. Thus we may assume that R itself is generated by finitely many algebraically independent elements as a K-algebra.

Since M is finitely generated, there is a filtration

$$0 = M_0 \subseteq M_1 \subseteq \ldots \subseteq M_t = M$$

where each successive quotient M_{i+1}/M_i is a cyclic L-graded R-module, that is,

$$M_{i+1}/M_i \cong (R/I_i)(-s_i)$$

with an L-graded (annihilator) ideal I_i in R and a shift $s_i \in L$. This isomorphism of L-graded R-modules induces an isomorphism of K-vector spaces in each L-degree. As far as the Hilbert function is concerned, we can therefore replace M by the direct sum of these cyclic modules. After the introduction of a monomial order, we can even replace R/I_i by $R/\operatorname{in}(I_i)$, where $\operatorname{in}(I_i)$ is the initial ideal (see Lemma 4.5.3 and recall that the monomials are homogeneous in the L-grading). By Corollary 4.5.5, we may finally replace $R/\operatorname{in}(I_i)$ by a direct sum of modules $(R/\mathfrak{p}_{ij})(-d_{ij})$, where the $\mathfrak{p}_{ij}$ are prime ideals in R, each generated by some of the indeterminates, and the d_{ij} are shifts in L.

Altogether this reduces the problem to the case in which the K-vector space M is isomorphic to a direct sum of vector spaces $R_i(-t_i)$, where R_i is a polynomial ring generated by indeterminates with degrees in L, and t_i a shift in L. Furthermore, we can use that the Hilbert function of M satisfies condition (6). The Hilbert function now counts the total number of monomials in each degree. Replacing the monomials by their exponent vectors, we can deduce the theorem from the next one. □

However, it may be useful to recall the notion of Hilbert function of an affine monoid, as introduced in Subsection 4.2. So let $\varphi : A \to S$ be a homomorphism of affine monoids, and S a submonoid of the lattice L; then φ induces an L-grading on A and $R := K[A]$, such that the L-graded Hilbert function of A,

$$H(A,\ell) = \operatorname{card} A_\ell = \operatorname{card} \varphi^{-1}(\ell)\,,$$

coincides with the L-graded Hilbert function of R. Likewise, for a shift $t \in L$, the Hilbert functions of $A(-t)$ and $R(-t)$ coincide, and are given by

$$H(A(-t),\ell) = \operatorname{card} A_{\ell-t} = \operatorname{card} \psi^{-1}(\ell),$$

where $\psi : A \to L$ is the shifted homomorphism $\psi := t + \varphi$. This justifies calling it also the Hilbert function of ψ. Finally, if we are given several affine monoid homomorphisms $\varphi_i : A_i \to S$ and shifts $t_i \in L$, $i = 1, \dots, r$, then the Hilbert functions of $\biguplus_{i=1}^r A_i(-t_i)$ and $\bigoplus_{i=1}^r R_i(-t_i)$ coincide. With these considerations in mind, we can state the combinatorial analogue of Theorem 6.1.1, and a generalization of [BG, Theorem 4.4.4]:

Theorem 6.1.3: *In the lattice $L = \mathbb{Z}^m$, let S be an affine submonoid and T a finitely generated S-submodule. Furthermore, for $i = 1, \dots, r$, let A_i be an affine monoid and $\varphi_i : A_i \to S$ a homomorphism. Let $t_1, \dots, t_r \in L$ and $\psi_i : A_i \to L$, $\psi_i(x) := t_i + \varphi_i(x)$. Finally, define the amalgamation $\psi := \bigoplus \psi_i : \biguplus A_i \to L$ and its Hilbert function $H(y) := \operatorname{card} \psi^{-1}(y)$ for $y \in L$.*

Now suppose that there exists an integer $d \geq 0$ such that

$$H(t) \geq c(w,W) \cdot \|t\|^d$$

for all $w \in L$, all subgroups W in L with $\operatorname{rk} W < m$, and all $t \in T \cap (w+W)$ with $\|t\| \gg 0$ (where $c(w,W) > 0$ is a constant depending on w and W, and $\|\cdot\|$ represents an arbitrary norm on $\mathbb{R}^m$). Then there exists a constant $c > 0$ such that

$$H(t) \geq c \cdot \|t\|^d$$

for all $t \in T$, $\|t\| \gg 0$.

The proof of Theorem 6.1.3 will be prepared in the following subsections; it can finally be found in Subsection 6.6. In order to point out the structure of the proof and to indicate the connections with the subsequent subsections, we include a short summary of the proof at this point.

In a first step one shows that it suffices to prove the theorem in a special case: one may assume, essentially, that the groups of differences of S and the φ_i-images S_i coincide with the ambient lattice L. Among other things, we must clarify that a certain subset of a sublattice U of L generates the whole lattice U again. We will investigate this in Subsection 6.2. Additionally, we may assume that the affine monoid S coincides with T and is positive; therefore, we are interested in covering S by objects for which we can control the degree of growth of the Hilbert function.

In the second step we then show some covering properties of the normalization $\bar{S}$ of S. In fact, $\bar{S}$ is covered by the normalizations $\bar{S}_i$ of the images of the mappings φ_i (what can be seen without further resources). But instead, we argue more precisely in order to cover $\bar{S}$ by some hyperplanes (which do not require further attention, due to the hypothesis) and the normalizations of certain submonoids $S_i^{(n)}$ (for which it is possible to control the degree of growth of the Hilbert function). Subsection 6.3 is dedicated to the analysis of the degree of growth, and Subsection 6.4 then explicitly describes the construction of the submonoids $S_i^{(n)}$.

In the third and last step, we compile all our considerations: it only remains to check the effect (on the growth of the Hilbert function) of the shifts t_i of the homomorphisms φ_i, as well as the shifts between S and the cyclic components of T. The respective statements are presented in Subsection 6.5 (and also already in Subsection 6.2).

6.2. Some linear algebra. At some points in our work, we will need the following facts from elementary linear algebra.

Lemma 6.2.1: *Let $\mathscr{S}$ be a system of rational equalities and inequalities, given in matrix form as*

$$A \cdot x = b\,, \quad C \cdot x \geq d\,,$$

where $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$, $C \in \mathbb{Q}^{k\times n}$, and $d \in \mathbb{Q}^k$. (The vector inequality is to be read componentwise.) Let $x \in \mathbb{R}^n$ be a solution of $\mathscr{S}$. Then there exist rational solutions, arbitrarily close to x; that is, for any $\varepsilon > 0$, there exists a solution $x' \in \mathbb{Q}^n$ of $\mathscr{S}$ such that $\|x - x'\| < \varepsilon$.

Proof: We may assume that x solves the inequalities strictly, that is, $Cx > d$. Recall that solvability of $Ax = b$ over $\mathbb{Q}$ and over $\mathbb{R}$ is equivalent. Therefore, let the rational solution space of $Ax = b$ be given by rational parameters λ_i. Then x is obtained by a certain choice $\lambda_i = r_i$ of real values $r_i \in \mathbb{R}$. Next we observe that the solution space of $Cx > d$ is an open subset in $\mathbb{Q}^n$ (or $\mathbb{R}^n$, respectively; the adjective 'open' refers to the standard topology). Thus choosing the parameters λ_i to be rational values $q_i \in \mathbb{Q}$, closely enough to r_i, will lead to a solution $x' \in \mathbb{Q}^n$ of both $Ax = b$ and $Cx > d$, as desired. □

Lemma 6.2.2: *Let L be a lattice of rank m, $S \subseteq L$ an affine submonoid and $U \subseteq \mathbb{Z}S$ a subgroup of rank m. Then $\mathbb{Z}(S \cap U) = U$.*

Proof: The inclusion '$\subseteq$' is clear (since U is a group). For the other inclusion, we first note that we may assume that $\mathbb{Z}S = L = \mathbb{Z}^m$ (since S and U are contained in $\mathbb{Z}S$). Now choose a basis $u_1, \ldots, u_m$ of U, and choose $s_0 \in S$ such that $s_0 + \bar{S} \subseteq S$ (according to Lemma 3.3.8). Next define the cone $C := C(S)$ of S and the (closed) parallelepiped

$$P := \{a_1 u_1 + \cdots + a_m u_m : a_i \in [0,1]\}$$

that is spanned by $u_1, \ldots, u_m$. The point now is that there exists some $s \in \mathbb{R}^m$ such that the shifted parallelepiped $s + P$ is contained in the intersection

$$I := (s_0 + C) \cap (s_0 - u_1 + C) \cap \cdots \cap (s_0 - u_m + C)$$

of shifts of the cone C. (Note that P is bounded and that I contains a shifted copy of C, since $\dim C = m$.)

By definition of P, the set $s + P$ intersects the lattice U in (at least) a point u_0. It follows that u_0 is also contained in S:

$$u_0 \in I \cap U \subseteq (s_0 + C) \cap \mathbb{Z}S = s_0 + (C \cap \mathbb{Z}S) = s_0 + \bar{S} \subseteq S\,.$$

Likewise, we see that $u_0 \in s_0 - u_i + C$ for $i = 1, \dots, m$, so that

$$u_0 + u_i \in (s_0 + C) \cap U \subseteq S.$$

Since this sum is contained in U anyway, we conclude that each $u_i = (u_0 + u_i) - u_0$ is contained in the group of differences of $S \cap U$. □

The following two lemmas deal with the equivalence of functions on a lattice L. Let $f, g : L \to \mathbb{R}_+$ be two functions on L; then f, g are *equivalent (on L)* if there exist positive real numbers c, c' such that

$$c \cdot f(x) \leq g(x) \leq c' \cdot f(x)$$

for all $x \in L$. If this double inequality only holds for almost all $x \in L$ (that is, only for $\|x\| \gg 0$), then f and g are *equivalent almost everywhere*.

Lemma 6.2.3: *Let $U \subseteq L$ be two lattices of equal ranks. Let $\|\cdot\|, \|\cdot\|'$ be norms on $\mathbb{R}L$ and $\mathbb{R}U$, respectively. Then $\|\cdot\|$ and $\|\cdot\|'$ are equivalent on U.*

Proof: This is easy: the inclusion $U \subseteq L$ of lattices implies an inclusion $\mathbb{R}U \subseteq \mathbb{R}L$ of $\mathbb{R}$-vector spaces, and the equality $\operatorname{rk} U = \operatorname{rk} L$ of ranks implies an equality $\dim_{\mathbb{R}} \mathbb{R}U = \dim_{\mathbb{R}} \mathbb{R}L$ of dimensions; therefore $\mathbb{R}U = \mathbb{R}L$. Now the assertion is a well-known fact from linear algebra. □

Lemma 6.2.4: *Let L be a lattice, $u \in L$ and $\|\cdot\|$ a norm on $\mathbb{R}L$. Then the functions $\|\cdot\|$ and $\|\cdot - u\|$ are equivalent almost everywhere, that is, there are constants $c, c' > 0$ such that*

$$c \cdot \|x\| \leq \|x - u\| \leq c' \cdot \|x\|$$

for all $x \in L$, $\|x\| \gg 0$. In fact, one can choose $c = 1/2$ and $c' = 2$.

Proof: For $\|x\| \geq 2\|u\|$, we have

$$\|x\| \leq 2\|x\| - 2\|u\| = 2(\|x - u + u\| - \|u\|) \leq 2\|x - u\|.$$

And if $\|x\| \geq \|u\|$, then $\|x - u\| \leq \|x\| + \|u\| \leq 2\|x\|$. □

6.3. A single homomorphism. In this subsection, we consider two affine monoids A and S, connected via a *surjective* homomorphism $\varphi : A \to S$; in addition, we assume that S is positive. (Note that assuming φ to be surjective is no restriction, upon replacing S with $\operatorname{im} \varphi$.) Recall now that the S-graded Hilbert function of A is given by

$$H : S \to \mathbb{N} \cup \{\infty\}, \quad H(s) := \operatorname{card} \varphi^{-1}(s).$$

We start by investigating when the Hilbert function is finite; it turns out that this happens precisely when $\ker \varphi = \{0\}$. Note that this does not imply the injectivity of φ since we are dealing with monoids.

Lemma 6.3.1: (a) *If the kernel of φ is trivial, then $H(s)$ is finite for all $s \in S$; otherwise, $H(s) = \infty$ for all $s \in S$.*

(b) *For any two elements $s, s' \in S$, there is an inequality $H(s) \leq H(s + s')$.*

Proof: (a) Suppose first that the kernel of φ is trivial. Since S is positive, there exists a positive grading $S = \biguplus_{n\in\mathbb{N}} S_n$ such that all components S_n are finite, and $S_0 = \{0\}$ (Lemma 3.4.5). Let A be generated by $a_1, \ldots, a_n$. Then S is generated by the images $s_i := \varphi(a_i)$. We may assume that $a_i \neq 0$ for all i. Therefore $\deg s_i > 0$ for all i. Now let $s \in S$ and $a \in \varphi^{-1}(s)$, say $a = \sum \lambda_i a_i$ with $\lambda_i \in \mathbb{N}$. This implies that $s = \sum \lambda_i s_i$ and $\deg s = \sum \lambda_i \deg s_i$. But this equation admits only a finite number of values for the λ_i.

Now suppose that $\varphi(a) = 0$ for some $a \in A$, $a \neq 0$. Then the kernel of φ contains the infinite set $\mathbb{N}a$, and so $H(0) = \infty$. The assertion now follows from (b).

(b) It suffices to note here that a fixed preimage $a' \in A$ of s' (which exists, since φ is surjective) induces an injective mapping $\varphi^{-1}(s) \longrightarrow \varphi^{-1}(s+s')$, given by the assignment $a \mapsto a + a'$. □

We remind the reader of some notions and results, introduced in Subsection 4.4: a function $\tilde{Q} : \mathbb{Z} \to \mathbb{Q}$ is *of quasi-polynomial type (of degree d)* if it coincides with a quasi-polynomial $Q(n)$ (of degree d) for all $n \in \mathbb{Z}$, $n \gg 0$; we have seen in Proposition 4.4.2 that the Hilbert function of a finitely generated module over a positively graded K-algebra is of quasi-polynomial type.

We continue our investigations by considering special sequences σ within the positive affine monoid S.

Definition 6.3.2: Let $a, b \in S$, $b \neq 0$, and let $\sigma(a,b)$ be the arithmetic progression $\sigma(a,b) := a + \mathbb{N}b$ in S. We say that $\sigma(a,b)$ has *degree* $d \in \mathbb{N}$ if the Hilbert function $H(a+nb)$ is of quasi-polynomial type of degree d (with respect to the variable n).

We note an immediate consequence which links the results of this subsection with the main theorem (Theorem 6.1.3): the arithmetic progression $\sigma(a,b)$ has degree $\geq d$ if and only if there exists a constant $c > 0$ such that

$$H(a+nb) \geq c \cdot n^d$$

for all $n \gg 0$. Moreover, since φ is supposed to be surjective, we could even assume that this inequality holds for all $n \geq 0$.

The next lemma is crucial for our investigations; it shows that the progressions $\sigma(a,b)$ do in fact have a certain degree.

Lemma 6.3.3: *Let φ and H be as above, and suppose that H is finite. Let $a, b \in S$, $b \neq 0$. Then the following statements hold:*

(a) *$A' := \varphi^{-1}(\mathbb{N}b)$ is an affine submonoid of A.*

(b) *$\sigma(0,b)$ has degree $d := \operatorname{rk} A' - 1$.*

(c) *$\sigma(a,b)$ too has degree d.*

Proof: (a) Note first that the finiteness of the Hilbert function implies that the kernel of φ is trivial (Lemma 6.3.1). Now fix a field K and recall that the affine monoid algebra $R := K[A]$ is S-graded in a natural way,

$$R = \bigoplus_{s \in S} R_s \quad \text{with} \quad R_s := \bigoplus_{\substack{\alpha \in A \\ \varphi(\alpha)=s}} K \cdot X^{\alpha},$$

such that the S-graded Hilbert functions of A and R coincide. With respect to this grading, we have $R_0 = K$, since $\ker\varphi = \{0\}$. Now, by Proposition 4.1.3, the K-algebra

$$R' := \bigoplus_{s \in \mathbb{N}b} R_s = \bigoplus_{n \in \mathbb{N}} R_{nb}$$

is finitely generated, and is therefore positively graded (in the sense of Subsection 4.4). Since $R' = K[A']$ by definition, the property of being finitely generated is inherited by A' (Proposition 3.3.2).

(b) This is an immediate consequence of Proposition 4.4.6 and Lemma 3.3.1 now (note that $R_b \neq \{0\}$ since φ is surjective):

$$\deg \sigma(0,b) = \dim R' - 1 = \operatorname{rk} A' - 1 .$$

(c) Here we consider the R'-module

$$M' := \bigoplus_{s \in a+\mathbb{N}b} R_s = \bigoplus_{n \in \mathbb{N}} R_{a+nb}$$

for which the same arguments as in (a) and (b) apply: it is finitely generated over R' (Proposition 4.1.3), and therefore

$$\deg \sigma(a,b) = \dim M' - 1 ,$$

by Proposition 4.4.6. Finally, it is clear that $\dim M' = \dim R'$, since $\operatorname{Ann}_{R'} M' = \{0\}$. ($R$ is an integral domain.) □

We note an immediate, but very helpful consequence.

Corollary 6.3.4: *Let $b \in S$, $b \neq 0$, and $d \in \mathbb{N}$. Then the following are equivalent:*

(i) *$\sigma(0,b)$ has degree d.*

(ii) *$\sigma(a,b)$ has degree d, for some $a \in S$.*

(iii) *$\sigma(a,b)$ has degree d, for every $a \in S$.*

This motivates, and establishes, the following definition which extends the notion of degree from arithmetic progressions to elements of S.

Definition 6.3.5: An element $b \in S$, $b \neq 0$, has *degree d* if one of the (equivalent) conditions (i)–(iii) in Corollary 6.3.4 is satisfied, that is, we define

$$\deg b := \deg \sigma(0,b)$$

if the Hilbert function is finite; otherwise, we set $\deg b := \infty$.

We have just seen that affine shifts $a \in S$ do not affect the degree of growth of the Hilbert function in direction of $b \in S$. We will now compare the degree of growth with respect to different directions $b, b' \in S$.

Lemma 6.3.6: *Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids with S positive, and $H : S \to \mathbb{N}$ the S-graded Hilbert function of A. Let F be a face of S and b, b' relative interior points of F. Then the following hold:*
(a) $\deg b = \deg mb$ *for any integer* $m > 0$.
(b) $\deg b = \deg b'$.

Proof: (a) Obviously, $A'' := \varphi^{-1}(\mathbb{N}mb)$ is a submonoid of $A' := \varphi^{-1}(\mathbb{N}b)$; therefore $\operatorname{rk} A'' \le \operatorname{rk} A'$ and $\deg mb \le \deg b$, by Lemma 6.3.3. Equality holds because, on the other hand, $H(b) \le H(mb)$, by Lemma 6.3.1 (b).

(b) Since F is itself an affine monoid (Lemma 3.2.1), we may assume that $F = S$. (Otherwise, consider the restriction $\varphi' : A' \to F$ of φ to the monoid $A' := \varphi^{-1}(F)$; although irrelevant, A' can be shown to be affine, see Lemma 6.5.1.) According to Lemma 3.3.8, there exists an element $s_0 \in S$ such that the 'shifted' normalization $s_0 + \bar{S}$ is contained in S.

Now let C denote the cone of S. By hypothesis, b' is contained in the relative interior of C, and in view of (a), we may replace b' by an arbitrary multiple mb', $m > 0$. We may therefore assume that

$$b' - b - s_0 \in C \cap \mathbb{Z}S = \bar{S}$$

(imagine a dual representation of C and note that each dual form takes strictly positive values on b'). It follows that $b'' := b' - b$ is contained in S, and so

$$H(b') = H(b + b'') \ge H(b),$$

by Lemma 6.3.1. This implies that $\deg b \le \deg b'$, and by symmetry, the assertion follows. □

Part (b) of the lemma enables us to define the degree of a face F of S. Recall that, if F is non-empty, then $\operatorname{relint} F$ is also non-empty (Lemma 3.2.2).

Definition 6.3.7: The *degree* of a non-empty face $F \ne \{0\}$ of the positive affine monoid S is defined to be the degree of any of its relative interior points. The *degree* of a non-empty face $F \ne \{0\}$ of the cone $C(S)$ is defined to be the degree of the respective face F_S of S.

So far we have only considered points $b, b' \in S$ for which the degree has shown to be equal. In the following situation, one will possibly have inequality.

Lemma 6.3.8: *Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids with S positive, and $H : S \to \mathbb{N}$ the S-graded Hilbert function of A. Let F, G be faces of S such that $F \subseteq G$. Then* $\deg F \le \deg G$.

Proof: By restriction arguments (as in the proof of Lemma 6.3.6), we may assume that $G = S$. So choose $b \in \operatorname{relint} F$ and $b' \in \operatorname{relint} S$. An easy check yields that we have already shown the desired inequality, $\deg b \le \deg b'$, again in the previous proof (under even weaker assumptions, namely, that $b \in S$ only). □

6.4. Affine monoids with least degree. Let $\varphi : A \to S$ continue to be a surjective homomorphism of affine monoids with S positive. As we have just seen in Lemma 6.3.8, an inclusion $F \subseteq G$ of faces of S leads to an inequality $\deg F \le \deg G$ of the corresponding degrees. We introduce a special notion if there is a lower bound.

Definition 6.4.1: Let $d \in \mathbb{N}$, $d \le \deg S$. An affine submonoid S' of S has *least degree d* (with respect to φ) if, for any face F' of S', one has $\deg F' \ge d$.

We illustrate this notion with an example, already mentioned in the introduction to this section: consider the homomorphism $\varphi : \mathbb{N}^3 \to \mathbb{N}^2$, given by

$$\varphi(e_1) = e_1\,, \quad \varphi(e_2) = \varphi(e_3) = e_2\,.$$

This assignment also defines the extension $\bar{\varphi} : \mathbb{Z}^3 \to \mathbb{Z}^2$ to the groups of differences. Hence we have $\ker\bar{\varphi} = \mathbb{Z}(e_2 - e_3)$. By Lemma 6.3.3, we have

$$\deg\sigma(0,b) = \begin{cases} 0 & \text{if } b \in \mathbb{N}e_1 \setminus \{0\}\,, \\ 1 & \text{if } b \in \mathbb{N}^2 \setminus \mathbb{N}e_1\,, \end{cases}$$

so that the two faces $F_x = \mathbb{N}e_1$ and $F_y = \mathbb{N}e_2$ of $\mathbb{N}^2$ have the degrees $\deg F_x = 0$ and $\deg F_y = 1$, respectively. Therefore, the whole monoid $S := \mathbb{N}^2$ has degree 1, but not all of its faces have this degree (with respect to φ).

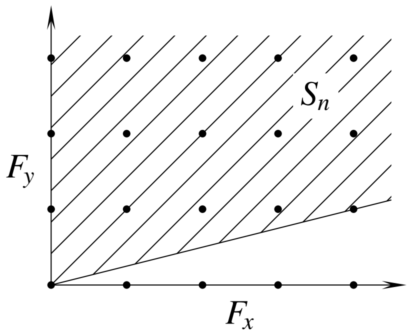

Figure 9

We now define a sequence of submonoids of S as follows: first, for all $n \in \mathbb{N}$, $n > 0$, let C_n be the cone generated by the vectors e_2 and $e_1 + 1/n \cdot e_2$ in $\mathbb{R}^2$; then, with C_n at hand, define the monoid $S_n := C_n \cap S$. Figure 9 visualizes these objects for the value $n = 4$. Note that each S_n is finitely generated (by Gordan's Lemma 3.1.2) and has least degree 1. Obviously, the S_n form an ascending chain $S_1 \subseteq S_2 \subseteq \cdots$ of submonoids of S, and any affine submonoid S' of S with least degree 1 must be contained in some S_n.

This example motivates the following construction. Let $\varphi : A \to S$ continue to be a surjective homomorphism of affine monoids with S positive. Note that we do not assume that S is normal. Let $d \in \mathbb{N}$, $d \le \deg S$. We will now construct a sequence $S_1 \subseteq S_2 \subseteq \cdots$ of submonoids S_n of S such that

(1) every S_n has least degree d, and

(2) every submonoid $S' \subseteq S$ with least degree d is contained in some S_n.

For this, consider the associated cone $C := C(S)$ of S. Since S is positive, so too is C (Lemma 3.4.3). Now let $Q \subseteq \mathbb{R}S$ be a fixed rational cross-section of C (which exists by Lemma 2.10.2). Recall that Q is a rational convex polytope and that there is a bijection between the faces of Q and the faces of C (Proposition 2.10.4). In particular, the vertices of Q correspond to the edges of C. Due to the further correspondence of $\mathscr{F}(C)$ and $\mathscr{F}(S)$ (Lemma 3.2.1), we may assign the degree of the respective face of S to each face F of Q; in particular, we define

$$\mathscr{F}_+(Q) := \{F \in \mathscr{F}(Q) : \deg F \geq d\}$$

to be the set of all 'good' faces of Q.

Now, for every $F \in \mathscr{F}_+(Q)$ (which is a rational convex polytope by itself), we do the following: first, we fix some rational point $z_F \in \operatorname{relint} F$. Note that z_F has degree $\geq d$, by definition. In a second step we define, for each vertex x_F of F, a sequence $(x_F^{(n)})_{n>0}$ of rational points

$$x_F^{(n)} := x_F + \frac{1}{n} \cdot (z_F - x_F) = \frac{1}{n} \cdot z_F + \left(1 - \frac{1}{n}\right) \cdot x_F \in [z_F, x_F[\,.$$

Note that $x_F^{(n)} \to x_F$ as $n \to \infty$, and that $x_F^{(n)}$ is contained in the relative interior of F. Therefore, $x_F^{(n)}$ has degree $\geq d$. Furthermore, we see that $x_F^{(n)}$ is the image of x_F under the dilation δ with centre z_F and factor $1 - 1/n$.

Next we define the rational convex polytope

$$Q_n := \operatorname{conv}\left(x_F^{(n)} : F \in \mathscr{F}_+(Q),\ x_F \in \operatorname{vert} F\right),$$

for each $n > 0$. Note that Q_n is a subpolytope of Q, and that all the points z_F are contained in Q_n. This can be seen as follows: since any polytope is the convex hull of its vertices (Proposition 2.5.3), we have

$$z_F \in F = \operatorname{conv}(x_F : x_F \in \operatorname{vert} F);$$

applying δ yields that

$$z_F = \delta(z_F) \in \operatorname{conv}(x_F^{(n)} : x_F \in \operatorname{vert} F).$$

Now let C_n denote the cone generated by Q_n in $\mathbb{R}S$, that is,

$$C_n := \mathbb{R}_+\left(x_F^{(n)} : F \in \mathscr{F}_+(Q),\ x_F \in \operatorname{vert} F\right)$$

for $n > 0$. We note that C_n is a finitely generated rational subcone of C. Finally, let $S_n := S \cap C_n$ be the corresponding submonoid of S.

Figure 10 illustrates the procedure: it shows the cross-section Q of a rational polyhedral cone C (with $\dim C = 3$), originally given by an affine monoid S. The vertices $v_1, \ldots, v_6$ of Q correspond to the edges $E_1, \ldots, E_6$ of C. We will identify corresponding faces of Q and C, respectively, and assume that only v_2 and v_6 have degree $\geq d$. This implies that $[v_1, v_2]$, $[v_2, v_3]$, $[v_1, v_6]$ and $[v_5, v_6]$ also have degree $\geq d$. As for the faces $[v_3, v_4]$ and $[v_4, v_5]$, we assume that these have degree less than d. Therefore, one chooses relative interior ('central') points

$$z_2 = v_2, \quad z_6 = v_6, \quad z_{12}, \quad z_{23}, \quad z_{56}, \quad z_{16} \quad \text{and} \quad z_Q,$$

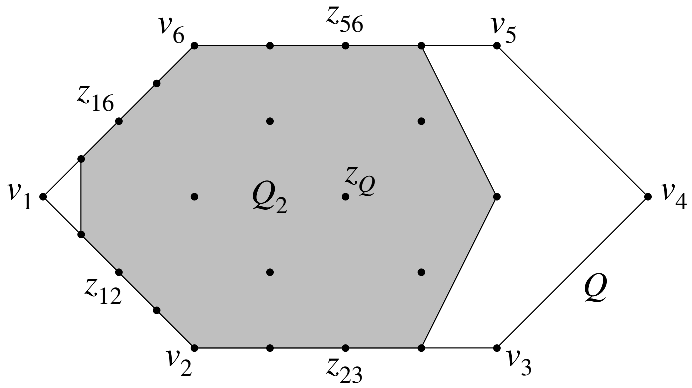

Figure 10

as indicated. Then, for $n = 2$, the cone C_2 is generated by the mid-points of the segments between these central points and the vertices; these mid-points are plotted without caption, in order to keep a clearly arranged figure. The intersection S_2 of S and C_2 then has least degree d. Note that, in this example, v_1 is replaced by three points, and v_4 is replaced by a single point in the interior of C. So, loosely speaking, the effect is as follows: we cut off all the 'bad' faces of C (but none of the 'good' ones).

The following proposition now states that the sequences (C_n) and (S_n) have the desired properties.

Proposition 6.4.2: *Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids with S positive, and let $d \in \mathbb{N}$, $d \le \deg_\varphi S$. Let $C = C(S)$ be the cone generated by S. Then the sequences constructed above have the following properties:*

(a) *$(C_n)_{n>0}$ is an ascending sequence of finitely generated rational subcones of C such that*

$$\bigcup_{n>0} C_n = \bigcup_{F \in \mathscr{F}_+(C)} \operatorname{relint} F\,.$$

($\mathscr{F}_+(C)$ denotes the set of those faces F of C for which $\deg F \ge d$.)

(b) *$(S_n)_{n>0}$ is an ascending sequence of affine submonoids of S such that each S_n has least degree d.*

Proof: Before proving (a) and (b), we show that each halfline $\mathbb{R}_+ x_F^{(n)}$ contains a non-trivial point from S: since $x_F^{(n)}$ is a rational point in C, we find that

$$c \cdot x_F^{(n)} \in C \cap \mathbb{Z}S = \bar{S}$$

for some $c > 0$, and then $c' \cdot x_F^{(n)} \in S$ for some $c' > 0$. By Lemma 2.5.4 and Proposition 2.10.4, the edges of C_n are among the halflines $\mathbb{R}_+ x_F^{(n)}$. Therefore, the edges of C_n have degree $\ge d$ (with respect to φ), and by Lemma 6.3.8, we conclude that each face of C_n has least degree d.

(a) By definition, each C_n is a finitely generated subcone of C. The sequence (C_n) is ascending since it follows from the definition that the point $x_F^{(n)}$ is contained in the segment $[z_F, x_F^{(n+1)}]$. (In fact, a detailed calculation yields that

$$x_F^{(n)} = \frac{1}{n^2} \cdot z_F + \frac{n^2-1}{n^2} \cdot x_F^{(n+1)} \text{.)}$$

But this segment clearly is contained in the convex cone C_{n+1}.

It remains to show the set-theoretic statement on the 'limit' of the sequence (C_n). The inclusion '$\subseteq$' is immediate: any point $x \in \bigcup C_n$ has degree at least d and is contained in C; but since C decomposes as

$$C = \biguplus_{F \in \mathscr{F}(C)} \operatorname{relint} F$$

into disjoint parts (Corollary 2.7.6), the inclusion follows.

Now let $x \in \operatorname{relint} F$ for some $F \in \mathscr{F}_+(C)$; since we have to detect x in a cone C_n, we may assume that $F \in \mathscr{F}_+(Q)$ (multiply x by a suitable positive real number). Now recall that $x_F^{(n)}$ is the image of x_F under the dilation δ with centre $z_F \in F$ and factor $1 - 1/n$. Therefore, the convex polytope

$$P_n := \operatorname{conv}(x_F^{(n)} : x_F \in \operatorname{vert} F)$$

is the image of F under δ. It is clear now that the ascending sequence (P_n) of convex polytopes will eventually cover every relative interior point of F: consider polyhedral representations of F and P_n (with linear inequalities), and use the fact that the relative interior point x satisfies the inequalities for F strictly.

(b) Being the intersection of the two monoids S and C_n, each S_n clearly is a submonoid of S. As for finite generation, we reason as follows: $\tilde{S} := \mathbb{Z}S \cap C_n$ is finitely generated (by Gordan's Lemma 3.1.2), and hence so is $S_n = S \cap \tilde{S}$ (by Corollary 4.1.4). It remains to see that S_n has least degree d. But in the preliminary remark, we have already shown that each face of C_n has least degree d. □

Corollary 6.4.3: *Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids with S positive, and $d \in \mathbb{N}$, $d \leq \deg_\varphi S$. Let C be the cone of S, and F a face of C with* $\deg F \geq d$. *Finally, let (C_n) be a sequence of subcones of C as in Proposition* 6.4.2, *and let x be a relative interior point of F.*

Then there exists a neighbourhood $U = U_\varepsilon(x)$ of x in C which is contained in C_n for sufficiently large n. In particular, C_n is a neighbourhood of x in C for $n \gg 0$.

Proof: By Corollary 2.7.2, there exists a ball B with centre x such that B intersects C only in faces G for which $F \subseteq G$. Note that $\deg G \geq d$ for these faces G, by Lemma 6.3.8. Due to Corollary 2.7.6, the intersection U of B and C can then be expressed as

$$U = B \cap C = \bigcup_{G \in \mathscr{F}_+(C)} (B \cap \operatorname{relint} G) \subseteq \bigcup_{G \in \mathscr{F}_+(C)} \operatorname{relint} G.$$

Clearly, U is a neighbourhood of x within C. In order to push it into C_n, it is convenient to argue via convexity: by re-choosing B (or the underlying metrics), we may assume that B is a convex polytope; then so too is U (Corollary 2.6.5), and it suffices to locate its finitely many vertices in some C_n. But since each vertex is contained in the relative interior of some face $G \in \mathscr{F}_+(C)$, we only have to apply Proposition 6.4.2 now. □

The following lemma states that, when the monoid S has least degree d, then the Hilbert function globally grows with degree d:

Lemma 6.4.4: *Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids such that $S \subseteq \mathbb{Z}^m$ has least degree d with respect to φ, and let H be the S-graded Hilbert function of A. Then there is a constant $c > 0$ such that*

$$H(x) \geq c \cdot \|x\|^d$$

for all $x \in S$.

Proof: By hypothesis, there exists, for each $b \in S$, $b \neq 0$, a constant $c(b) > 0$ such that

$$H(nb) \geq c(b) \cdot n^d$$

for all $n \in \mathbb{N}$. (We may assume that this holds for *all* n since φ is surjective.) Now S is finitely generated, say by $b_1, \ldots, b_r \neq 0$, and we shall show that c can be chosen via the constants $c(b_1), \ldots, c(b_r)$. In fact, let $x \in S$, say $x = n_1 b_1 + \cdots + n_r b_r$ with $n_i \in \mathbb{N}$. By Lemma 6.3.1, we have

$$H(x) \geq H(n_i b_i) \geq c(b_i) \cdot n_i^d$$

for all i; therefore, it remains to show that there is a constant $\gamma > 0$ such that $n_i \geq \gamma \cdot \|x\|$ for at least one i. But in fact, one can choose any $\gamma < 1/(\|b_1\| + \cdots + \|b_r\|)$; namely, by the triangle inequality, we have

$$\|x\| \leq n_1 \|b_1\| + \cdots + n_r \|b_r\| .$$ □

The situation becomes particularly interesting when not only one homomorphism φ is involved. We will investigate this in the proof of Theorem 6.1.3 in Subsection 6.6.

6.5. Shifted homomorphisms. Let $\varphi : A \to S$ be a surjective homomorphism of affine monoids, where S is a submonoid of a lattice L that has rank m.

Now consider an affine submonoid U of L. It defines a submonoid $S' := S \cap U$ in L. We may restrict φ via its image to S' and $A' := \varphi^{-1}(S')$. This leads to the mapping $\varphi' : A' \to S'$, another surjective homomorphism of affine monoids, as the following lemma shows. Note that, by definition, the L-graded Hilbert functions of A and A' coincide on S'.

Lemma 6.5.1: *Both S' and A' are finitely generated, that is, affine monoids.*

Proof: The assertion is clear for S', in view of Corollary 4.1.4. We may therefore write $S' = \sum \mathbb{N} s_i$, with finitely many $s_i \in S'$.

As for A', we consider the extension $\bar{\varphi} : \mathbb{Z}A \to \mathbb{Z}S$, and its kernel $\ker \bar{\varphi}$. Since $\mathbb{Z}A$ is a finitely generated free Abelian group, so too is $\ker \bar{\varphi}$, and thus $\ker \bar{\varphi}$ is also finitely generated as a monoid. Hence

$$\bar{\varphi}^{-1}(S') = \sum \mathbb{N}a_i + \ker \bar{\varphi}$$

is a finitely generated monoid (where we have chosen $a_i \in \varphi^{-1}(s_i)$ arbitrarily). This finally implies that this also holds for $A' = A \cap \bar{\varphi}^{-1}(S')$, again by Corollary 4.1.4. □

Now fix some shift $t \in L$, and define the mapping $\psi : A \to L$, $\psi(x) := t + \varphi(x)$. We will subsequently write $\psi = t + \varphi$ for short, and ψ will be called a *shifted homomorphism*. As shown in Subsection 4.2, the L-graded Hilbert function $H : L \to \mathbb{N}$ of $A(-t)$ is given by $H(y) = \operatorname{card} \psi^{-1}(y)$ for $y \in L$. Due to this connection and for simplicity, we will refer to H also as the *Hilbert function* of ψ. Furthermore, we fix some point $u \in L$ and consider the cyclic U-module $V := u + U$ in L.

In the next part of this subsection, we shall consider the following problem: find $t'_1, \ldots, t'_n \in V$ such that the amalgamation $\psi' := \psi'_1 \oplus \ldots \oplus \psi'_n$ of the shifted homomorphisms $\psi'_i = t'_i + \varphi'$ has a Hilbert function H' which is equivalent to the Hilbert function H of ψ on V. Note that the Hilbert function H' of ψ' is the sum of the Hilbert functions of the mappings ψ'_i; and recall from Subsection 6.2 that equivalence of H and H' on V means that there are positive constants c_1 and c_2 such that

$$c_1 \cdot H(y) \leq H'(y) \leq c_2 \cdot H(y)$$

for all $y \in V$.

In order to construct the t'_i as desired, it is recommendable to switch over to the corresponding objects in the category of modules over rings. So we fix an arbitrary field K, and define the monoid ring $R = K[A]$. Since A is finitely generated, R is Noetherian. It carries a natural L-grading, induced by the homomorphism φ. As for the module, we consider a shifted copy of R, namely $M = R(-t)$. It is a finitely generated L-graded R-module, and its L-graded Hilbert function coincides with the Hilbert function of ψ (as seen previous to Theorem 6.1.3).

Now comes the crucial step. Since U is a finitely generated submonoid of L, the R_0-algebra $R' := \bigoplus_{y \in U} R_y$ is a U-graded Noetherian ring (Proposition 4.1.3), and coincides with the monoid ring $K[A']$. And since V is a cyclic U-module by definition, the same proposition also implies that $M' := \bigoplus_{y \in V} M_y$ is a finitely generated (V-graded) R'-module. So, if $m'_1, \ldots, m'_n \in M'$ denotes a homogeneous system of generators of M' over R', of degrees $t'_1, \ldots, t'_n \in V$, we obtain a homogeneous and surjective homomorphism

$$\tilde{M} := \bigoplus_{i=1}^{n} R'(-t'_i) \longrightarrow M', \qquad e_i \longmapsto m'_i.$$

In order to see that the t'_i perform what we want, we note that there are inequalities

$$H(M',y) \leq H(\tilde{M},y) = \sum_{i=1}^{n} H(R'(-t'_i),y) \leq n \cdot H(M',y)$$

for all $y \in V$; the last inequality is due to the fact that the above homomorphism is injective on every direct summand $R'(-t'_i)$. Now the observation is complete because, for $y \in V$, we have $H(M',y) = H(M,y) = H(y)$, and $H(\tilde{M},y) = H'(y)$ is the Hilbert function of $\psi' = \psi'_1 \oplus \ldots \oplus \psi'_n$.

Summing up, we have shown the following proposition.

Proposition 6.5.2: *Let S be an affine submonoid of the lattice L, and let $\varphi : A \to S$ be a homomorphism of affine monoids. Furthermore, let $t,u \in L$ be elements and U an affine submonoid of L, and choose $\varphi' : A' \to S'$ to be the restriction of φ to $S' := S \cap U$. Then, with respect to the Hilbert function on $V = u + U$, the shifted homomorphism $\psi = t + \varphi$ is equivalent to an amalgamation ψ' of finitely many shifted homomorphisms $\psi'_i = t'_i + \varphi'$.*

In the final part of this subsection, we shall compare the Hilbert functions of a homomorphism φ and a corresponding shifted homomorphism $\psi = t + \varphi$. This study is necessary since, in the main theorem, both the assumption and the assertion refer to the Hilbert function of shifted homomorphisms, whereas in Subsections 6.3 and 6.4, we have considered homomorphisms only. Nevertheless, we shall be able to use these results; namely, we have the following link:

Lemma 6.5.3: *Let $\varphi : A \to S$ be a homomorphism of affine monoids and S a submonoid of the lattice L. Let $t \in L$ be a shift, and $\psi = t + \varphi$ a shifted homomorphism. Let H_φ and H_ψ denote the respective Hilbert functions, that is,*

$$H_\varphi(x) := \operatorname{card} \varphi^{-1}(x) \quad \textit{and} \quad H_\psi(x) := \operatorname{card} \psi^{-1}(x)$$

for all $x \in L$. Finally, let $d \in \mathbb{N}$ and X be a subset of L. Then the following are equivalent:

(i) *There is a constant $c > 0$ such that $H_\varphi(x) \geq c \cdot \|x\|^d$ for $x \in X$, $\|x\| \gg 0$.*

(ii) *There is a constant $c' > 0$ such that $H_\psi(x) \geq c' \cdot \|x\|^d$ for $x \in t + X$, $\|x\| \gg 0$.*

Proof: By definition of ψ, we have

$$H_\psi(x) = \operatorname{card} \psi^{-1}(x) = \operatorname{card} \varphi^{-1}(x - t) = H_\varphi(x - t)$$

for $x \in L$. The assertion now follows from Lemma 6.2.4: for instance, suppose (i) to be true, and choose $x \in t + X$ with $\|x\| \gg 0$; then $x = t + x'$ for some $x' \in X$ with $\|x'\| \gg 0$, whence

$$H_\psi(x) = H_\varphi(x') \geq c \cdot \|x'\|^d \geq c' \cdot \|x\|^d$$

for a suitable constant $c' > 0$. □

6.6. Proof of the main theorem. Finally, we are now in a position to prove the main theorem from Subsection 6.1.

Theorem 6.1.3: *In the lattice $L = \mathbb{Z}^m$, let S be an affine submonoid and T a finitely generated S-submodule. Furthermore, for $i = 1, \dots, r$, let A_i be an affine monoid and $\varphi_i : A_i \to S$ a homomorphism. Let $t_1, \dots, t_r \in L$ and $\psi_i : A_i \to L$, $\psi_i(x) := t_i + \varphi_i(x)$. Finally, define the amalgamation $\psi := \bigoplus \psi_i : \biguplus A_i \to L$ and its Hilbert function $H(y) := \operatorname{card} \psi^{-1}(y)$ for $y \in L$.*

Now suppose that there exists an integer $d \geq 0$ such that

$$H(t) \geq c(w, W) \cdot \|t\|^d$$

for all $w \in L$, all subgroups W in L with $\operatorname{rk} W < m$, and all $t \in T \cap (w + W)$ with $\|t\| \gg 0$ (where $c(w, W) > 0$ is a constant depending on w and W, and $\|\cdot\|$ represents an arbitrary norm on $\mathbb{R}^m$). Then there exists a constant $c > 0$ such that

$$H(t) \geq c \cdot \|t\|^d$$

for all $t \in T$, $\|t\| \gg 0$.

For convenience and to simplify matters, we introduce some abbreviations. We denote the images of the homomorphisms $\varphi_i : A_i \to S$ by S_i. Clearly, each S_i is an affine monoid. Another important tool for the proof will be the set

$$I := I_{\max} := \left\{ i \in \{1, \dots, r\} : \operatorname{rk} S_i = m \right\}$$

of those indices i such that S_i has maximal rank m. With this setup in mind, the proof of Theorem 6.1.3 can now be carried out in three steps.

Step 1: *Reduce to the case when $\mathbb{Z}S_i = L$ for $i \in I$, $T = S$ is positive and $m > 1$.*

Suppose that $\operatorname{rk} S_i = m$ for at least one i, so that the index set I is non-empty. Then the intersection

$$U := \bigcap_{i \in I} \mathbb{Z}S_i$$

defines a subgroup of $\mathbb{Z}S$ (and of L) of rank m. Since L consists of only finitely many residue classes modulo U, we may choose some $u \in L$ and restrict the search for the constant c to the residue class $V = u + U$, or, more precisely, to the region $T' := T \cap V$. We can now use Corollary 4.1.4 to see that $S' := S \cap U$ is an affine monoid and that T' is a finitely generated S'-module.

Next we consider the restrictions $\varphi_i' : A_i' \to S_i'$, as described in Subsection 6.5: their images are $S_i' := S_i \cap U$, and they are defined on the monoids $A_i' := \varphi_i^{-1}(S_i')$. Recall that φ_i' is a surjective homomorphism of affine monoids. By Proposition 6.5.2, there are, for every i, finitely many shifted homomorphisms $\psi_{ij}' = t_{ij}' + \varphi_i' : A_i' \to V$ such that the associated Hilbert functions H_i and $\sum_j H_{ij}'$ are equivalent on V. But then the Hilbert functions $H = \sum_i H_i$ and $H' = \sum_{i,j} H_{ij}'$, corresponding to the amalgamations of the ψ_i and the ψ_{ij}', respectively, are also equivalent on V. This shows that Theorem 6.1.3 is equivalent to its version for the objects (L, S', T', H') (instead of (L, S, T, H)).

Another shift by the common value $-u$ leads to mappings $\psi''_{ij} = \psi'_{ij} - u : A'_i \to U$. In view of Lemma 6.2.4, we obtain another equivalent version of Theorem 6.1.3, namely for the objects $(U, S', T'' := T' - u, H'')$. Note that, according to Lemma 6.2.2, we have $\mathbb{Z}S' = \mathbb{Z}S'_i = U$ for all $i \in I$. Therefore, we may restrict the proof of the theorem (in its original version) to the case where $\mathbb{Z}S_i = L$ for all $i \in I$.

Since T is finitely generated, it suffices to find a constant c for each cyclic part $t + S$ of T. Now, by Lemma 6.2.4, both the assumption and the assertion of the theorem are invariant under a shift of the coordinate system; therefore, we may assume that $T = S$.

Finally, note that L can be written as a union of finitely many positive affine submonoids U' (choose $U' = \mathbb{N}^m$ and so on). Therefore, it will clearly suffice to find a constant c for a submonoid $S' := S \cap U'$ of $S = T$. Now, S' is a positive affine submonoid of L (Corollary 4.1.4), and we may again consider the restricted homomorphisms $\varphi'_i : A'_i \to S'_i$ for $S'_i := S_i \cap U'$, as already done above (in a different situation in detail). By Lemma 6.5.1, each φ'_i is a surjective homomorphism of affine monoids, and by Proposition 6.5.2, each shifted homomorphism $\psi_i = t_i + \varphi_i$, when considered with respect to its Hilbert function on U', is equivalent to an amalgamation of finitely many shifted homomorphisms $\psi'_{ij} = t'_{ij} + \varphi'_i$. As above, this leads to an equivalence (on U') of the Hilbert functions of the amalgamations of the ψ_i and the ψ'_{ij}, respectively, and shows that it suffices to prove the theorem only in the particular case when $S = T$ is a positive affine submonoid of L.

Since the case $m = 1$ has already been treated in the proof of Theorem 6.1.1, we may assume that $m > 1$. □

We prepare the next step by introducing some further notation: for $i = 1, \ldots, r$, the cone generated by S_i will be denoted by C_i. Since S_i is affine, each C_i is a rational polyhedral subcone of $C := C(S)$, the cone generated by S. Note that $\dim C_i = m$ for $i \in I$, and $\dim C_i < m$ for $i \notin I$. For $i \notin I$, it will prove convenient to choose linear hyperplanes V_i in $\mathbb{R}^m$ which contain the respective cone C_i (and the monoid S_i). Recall that a bar is used to denote normalization, that is, integral closure within the group of differences. So, for example, we have

$$\bar{S} = \{x \in \mathbb{Z}S : ax \in S \text{ for some } a > 0\}.$$

By Lemma 3.1.2, we can express $\bar{S} = C \cap \mathbb{Z}S$ as the set of all ($\mathbb{Z}S$-)lattice points in C, and so, by Step 1, we have $\bar{S}_i = L \cap C_i$ for all $i \in I$.

We remind the reader of the notion of degree of the monoid S_i: it is defined to be the degree (with respect to φ_i) of any progression $a + \mathbb{N}b$ with $a \in S_i$ and $b \in \operatorname{relint} S_i$ (Corollary 6.3.4, Lemma 6.3.6). Moreover, the degree is monotonic increasing with respect to inclusions (Lemma 6.3.8). So, in view of Lemma 6.5.3, we may clearly omit those φ_i such that $\deg_i S_i < d$, and do therefore assume that $\deg_i S_i \geq d$ for all $i = 1, \ldots, r$. (We emphasize that, throughout this subsection, $\deg_i$ refers to the degree of growth with respect to the homomorphism φ_i, as discussed in Subsection 6.3.) Now consider, for every homomorphism φ_i, the sequences

$(C_i^{(n)})$ and $(S_i^{(n)})$ of finitely generated cones and monoids, respectively, as constructed in Subsection 6.4. Recall that, by Proposition 6.4.2, $C_i^{(n)}$ converges to $\bigcup_{F\in\mathscr{F}_+(C_i)} \operatorname{relint} F$ as $n \to \infty$, and that each $S_i^{(n)}$ has least degree d.

Step 2: *For $n \gg 0$, show that*

$$\bar{S} \subseteq \bigcup_{i\in I} \bar{S}_i^{(n)} \cup \bigcup_{i\notin I} V_i\,. \tag{7}$$

We argue via contradiction and, therefore, assume that the set

$$C \smallsetminus \bigcup_{i\in I} C_i^{(n)} \smallsetminus \bigcup_{i\notin I} V_i$$

is non-empty for all n. Since S is positive (Step 1), so too is C (Lemma 3.4.3); thus we may fix a rational cross-section Q of C (Lemma 2.10.2), and can then find a sequence (x_n) of rational points

$$x_n \in Q \smallsetminus \bigcup_{i\in I} C_i^{(n)} \smallsetminus \bigcup_{i\notin I} V_i\,.$$

Since Q is compact, there exists a limit point

$$x \in Q \smallsetminus \operatorname{relint} \bigcup_{i\in I} C_i \subseteq Q \smallsetminus \bigcup_{i\in I} \operatorname{relint} C_i\,.$$

By omitting points if necessary, we may assume that $x_n \to x$ as $n \to \infty$. For $i \in I$, we have $x \notin \operatorname{relint} C_i$, so that either $x \notin C_i$ or $x \in \operatorname{bd} C_i$. Let us re-choose x according to Lemma 6.2.1: this yields a rational point $x' \in Q \smallsetminus \bigcup_{i\in I} \operatorname{relint} C_i$ satisfying exactly the same coincidences as x (with respect to all cones C_i). Recall that we can even choose x' arbitrarily close to x. Using the rational points x_n and x', our strategy now is to construct a progression σ in S which violates the hypothesis for the degree of growth. For this, we preliminarily fix some $n \gg 0$ (in fact, we will specify the requirements for n later on). Since x_n and x' are rational points in C, we can find positive integers c and c' such that

$$a := c \cdot x_n \quad \text{and} \quad b := c' \cdot x'$$

lie in S. Now we consider the arithmetic progression

$$\sigma := a + \mathbb{N} \cdot b$$

in S. For this progression and each i, we shall show that

$$\deg_i(\sigma - t_i) < d\,. \tag{8}$$

Using Lemma 6.5.3, we see that the degree of growth of H_{ψ_i} along σ is less than d. This is clearly in contrast to the hypothesis of the theorem (choose $w = a$ and $W = \mathbb{Z}b$). In order to show inequality (8), we distinguish three cases.

Case 1: $i \in I$ and $x \notin C_i$. This is an easy case. Since x' satisfies the same coincidences as x, neither x' nor the directional vector b are contained in C_i. Consequently, $\sigma - t_i = a - t_i + \mathbb{N}b$ intersects the cone C_i only in a finite number of points (Lemma 2.9.5). More than ever, the intersection of $\sigma - t_i$ and S_i is finite.

Therefore, we have $\deg_i(\sigma - t_i) = -\infty$. By the way, note that the choice of the starting point a of σ is irrelevant in this case.

Case 2: $i \in I$ and $x \in \operatorname{bd} C_i$. By Corollary 2.7.6, there exists a (unique) face $F \neq \{0\}$ of C_i such that $x \in \operatorname{relint} F$. Note that also $b, x' \in \operatorname{relint} F$, and by Corollary 6.3.4, we have $\deg_i(\sigma - t_i) \leq \deg_i F$. (Actually, there is equality here, or the degree on the left equals $-\infty$.) So if $\deg_i F < d$, the desired contradiction is immediate, and we may (and do) suppose that $\deg_i F \geq d$. Now we recall Corollary 6.4.3: for n large enough, $C_i^{(n)}$ is a neighbourhood of x in C_i. But since x_n has been chosen outside $C_i^{(n)}$, we may also assume that x_n lies outside C_i.

At this point, Lemma 2.6.6 is useful: there exists a supporting hyperplane H of C_i such that $x_n \in \operatorname{int} H^-$ and $x \in H$ (and $b, x' \in H$). Therefore, upon choosing the factor c large enough, the point

$$a - t_i = c \cdot x_n - t_i$$

is contained in the interior of H^- as well. But then Lemma 2.9.5 helps again: the halfline $a - t_i + \mathbb{R}_+ b$ does not meet the cone C_i at all. Thus $\sigma - t_i$ and S_i do not intersect, and (8) follows.

Case 3: $i \notin I$. This case is easy again. By choice, x_n is not contained in V_i (neither in C_i, in particular). By varying c, we again achieve that

$$a - t_i = c \cdot x_n - t_i$$

neither is contained in V_i. Now elementary linear algebra yields that the halfline $a - t_i + \mathbb{R}_+ b$ meets the hyperplane V_i at most once. So once more, we have $\deg_i(\sigma - t_i) = -\infty$. □

Step 3: *Deduce the assertion of the theorem.*

This is easy now, after Step 2 has been performed. As mentioned earlier, each $S_i^{(n)}$ has least degree d with respect to φ_i (Proposition 6.4.2). Therefore, by Lemma 6.4.4, there is a constant $c_i > 0$ such that

$$H_{\varphi_i}(x) \geq c_i \cdot \|x\|^d$$

for all $x \in S_i^{(n)}$, $\|x\| \gg 0$. Then, in turn, Lemma 6.5.3 provides a constant $c_i' > 0$ such that

$$H_{\psi_i}(x) \geq c_i' \cdot \|x\|^d$$

for all $x \in t_i + S_i^{(n)}$, $\|x\| \gg 0$. But the sets $t_i + S_i^{(n)}$ and $\bar{S}_i^{(n)}$ differ in only finitely many hyperplanes (Corollary 3.3.10), and all these hyperplanes (found for $i \in I$), together with the hyperplanes V_i (for $i \notin I$), are covered by the hypothesis of the theorem. Summing up, we have covered the right-hand side in (7), and the proof of the theorem is complete. □

References

[AM] M. F. Atiyah, I. G. Macdonald: *Introduction to Commutative Algebra.* Addison-Wesley, Reading/Mass. 1969.

[Ar] E. Artin: *Collected Papers.* S. Lang, J. T. Tate (eds.), Addison-Wesley, Reading/Mass. 1965.

[AT] E. Artin, J. T. Tate: *A note on finite ring extensions.* J. Math. Soc. Japan **3** (1951), 74–77. Also included in [Ar].

[BG] W. Bruns, J. Gubeladze: *Semigroup algebras and discrete geometry.* In: L. Bonavero, M. Brion (eds.), *Geometry of toric varieties*, Séminaires et Congrès **6** (2002), 43–127.

[BG2] W. Bruns, J. Gubeladze: *Normality and covering properties of affine semigroups.* J. Reine Angew. Math. **510** (1999), 151–178.

[BGHMW] W. Bruns, J. Gubeladze, M. Henk, A. Martin, R. Weismantel: *A counterexample to an integer analogue of Carathéodory's theorem.* J. Reine Angew. Math. **510** (1999), 179–185.

[BGT] W. Bruns, J. Gubeladze, N. V. Trung: *Problems and algorithms for affine semigroups.* Semigroup Forum **64** (2002), 180–212.

[BH] W. Bruns, J. Herzog: *Cohen-Macaulay rings.* Cambridge University Press, Cambridge 1993.

[BH2] W. Bruns, J. Herzog: *Semigroup rings and simplicial complexes.* J. Pure Appl. Algebra **122** (1997), 185–208.

[BK] W. Bruns, R. Koch: NORMALIZ*, computing normalizations of affine semigroups.* Available via anonymous `ftp` from `ftp.math.uos.de/pub/osm/kommalg/software/`

[BK2] W. Bruns, R. Koch: *Computing the integral closure of an affine semigroup.* Univ. Iagell. Acta Math. **39** (2001), 59–70.

[Br] A. Brøndsted: *An Introduction to Convex Polytopes.* Springer, New York 1983.

[Bu] E. Burger: *Über homogene lineare Ungleichungssysteme.* Z. angew. Math. Mech. **36** (1956), 135–139.

[Ei] D. Eisenbud: *Commutative Algebra with a View Toward Algebraic Geometry.* Springer, New York 1995.

[Ew] G. Ewald: *Combinatorial Convexity and Algebraic Geometry.* Springer, New York 1996.

[Ga] D. Gale: *Convex polyhedral cones and linear inequalities.* In: T. C. Koopmans (ed.), *Activity analysis of production and allocation*, Chap. XVII, 287–297 (1951).

[Ge] M. Gerstenhaber: *Theory of convex polyhedral cones.* In: T. C. Koopmans (ed.), *Activity analysis of production and allocation*, Chap. XVIII, 298–316 (1951).

[Ma] H. Matsumura: *Commutative ring theory.* Cambridge University Press, Cambridge 1986.

[MS] P. McMullen, G. C. Shephard: *Convex Polytopes and the Upper Bound Conjecture.* Cambridge University Press, Cambridge 1971.

[Se] A. Sebö: *Hilbert bases, Carathéodory's theorem, and combinatorial optimization.* In: R. Kannan, W. R. Pulleyblank (eds.), *Integer Programming and Combinatorial Optimization*, 431–456 (1990).

[Sh] R. Y. Sharp: *Steps in Commutative Algebra.* Cambridge University Press, Cambridge 1990.

[Va] W. V. Vasconcelos: *Computational Methods in Commutative Algebra and Algebraic Geometry.* Springer, Berlin 1998.

[Vi] R. H. Villarreal: *Monomial algebras.* Dekker, New York 2001.

[Zi] G. M. Ziegler: *Lectures on Polytopes.* Springer, New York 1995.