Proc. Int. Conf. – Commutative Algebra and Combinatorics No. 4, 2007, pp. 1–38.

Commutative Algebra Arising from the Anand–Dumir–Gupta Conjectures

Winfried Bruns

Universität Osnabrück, FB Mathematik/Informatik, 49069 Osnabrück, Germany
e-mail: wbruns@uos.de

Abstract. In 1966 the Indian mathematicians H. Anand, V. C. Dumir, and H. Gupta investigated a combinatorial distribution problem and formulated some conjectures on the number of solutions. These conjectures were solved and extended by R. Stanley. His solution, based on methods of commutative algebra, was one of the starting points of combinatorial commutative algebra. In this article we describe the conjectures and their proofs and introduce the notions of commutative algebra on which the proofs are based.

Meinem Freund und Lehrer Udo Vetter zum 65. Geburtstag

Introduction

In 1966 the Indian mathematicians H. Anand, V. C. Dumir, and H. Gupta [2] investigated a combinatorial distribution problem that they had learnt from a paper of K. Manjo: *suppose that n distinct objects, each available in r identical copies, are distributed among n persons in such a way that each person receives exactly r objects. What can be said about the number $H(n,r)$ of such distributions?* Based on their investigation of $H(3,r)$, for which they gave an explicit formula, Anand, Dumir, and Gupta formulated several conjectures:

(ADG-1) there exists a polynomial $P_n(r)$ of degree $(n-1)^2$ such that $H(n,r) = P_n(r)$ for all $r \gg 0$;
(ADG-2) $H(n,r) = P_r(n)$ for all $r > -n$; in particular $P_n(-r) = 0, r = 1, \dots, n-1$;
(ADG-3) $P_n(-r) = (-1)^{(n-1)^2} P_n(r-n)$ for all $r \in \mathbb{Z}$.

In the original formulation (ADG-1) is not present. We have split the original conjecture into two parts since (ADG-1) is substantially easier to prove than (ADG-2).

The combinatorial problem can be recast as follows. If we let a_{ij} denote the number of copies of object i that person j receives, then $A = (a_{ij}) \in \mathbb{Z}_+^{n\times n}$ is an $n \times n$ matrix with non-negative integral entries such that

$$\sum_{k=1}^{n} a_{ik} = \sum_{l=1}^{n} a_{lj} = r, \qquad i, j = 1, \dots, n. \tag{1}$$

8	1	6
3	5	7
4	9	2

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

Figure 1. Two famous magic squares

Clearly $H(n, r)$ is the number of such matrices A. The system of equations (1) is part of the definition of *magic squares*. For such squares one usually requires that the sums over their diagonal and anti-diagonal also have the same value r, and true magic $n \times n$ squares, like the ones in Figure 1, must have the entries $1, \dots, n^2$. The 3×3 square has been found in ancient Chinese sources and the 4×4 appears in Albrecht Dürer's engraving *Melancholia* (1514). It has remarkable symmetries and shows the year of its creation.

The system of equations (1) is satisfied by, from a magician's view point, rather trivial objects, like the zero matrix, the identity matrix, and the matrix $\mathbf{1}$ with all entries equal to 1. Although it is highly doubtful that such matrices possess any witchcraft, mathematicians have called them "magic squares", and we will stick to this terminology.

Generating functions

A very strong tool of enumerative combinatorics, endowed with real magic power, is the concept of generating function. We have to introduce it in order to describe Stanley's extension of the ADG-conjectures. Consider the formal power series

$$H_n(t) = \sum_{r=0}^{\infty} H(n, r)t^r.$$

Then, as we will see, (ADG-1) is equivalent to the fact that $H_n(t)$ is the power series expansion at 0 of a rational function, denoted by the same name,

$$H_n(t) = \frac{h_0 + h_1 t + \cdots + h_u t^u}{(1-t)^{(n-1)^2+1}}, \qquad h_0 = 1, h_1, \dots, h_u \in \mathbb{Z}.$$

(ADG-2) and (ADG-3) can be converted into properties of this function:

$$\text{(ADG-2)} \quad \Longleftrightarrow \quad \deg H_n(t) = u - ((n-1)^2 + 1) = -n,$$

and once (ADG-2) has been proved,

$$\text{(ADG-3)} \quad \Longleftrightarrow \quad h_i = h_{u-i}, \quad i = 0, \dots, u.$$

Stanley extended the conjectures by

(ADG-4) $h_0, \dots, h_u \geq 0$;
(ADG-5) $h_0 \leq h_1 \leq \cdots \leq h_{\lceil u/2 \rceil}$.

In [22], [23] he proved (ADG-1)–(ADG-4) via a translation of the problem into commutative algebra. See Stanley's book [27] for more information on the pre-ADG history of the problem and on its solution. The major steps of our notes follow the development in [27].

For results and references on explicit formulas for $H(n, r)$ and a computer approach to them we refer the reader to Ahmed, De Loera and Hemmecke [1].

The commutative algebra way

The basic step is to understand $H(n, r)$ as the Hilbert function of a graded K-algebra, where K is an arbitrary field.

Let $\mathcal{M}_{nr}$ denote the set of solutions to (1) in $\mathbb{Z}_+^{n\times n}$ and set

$$\mathcal{M}_n = \bigcup_{r=0}^{\infty} \mathcal{M}_{nr}$$

Since $A + B \in \mathcal{M}_{n,r+s}$ for $A \in \mathcal{M}_{nr}$, $B \in \mathcal{M}_{ns}$, the set $\mathcal{M}_n$ has an algebraic structure: it is a submonoid of $\mathbb{Z}_+^{n\times n}$ (under addition). By the theorem of Birkhoff-von Neumann the monoid $\mathcal{M}_n$ is generated by $\mathcal{M}_{n1}$, the set of $n \times n$ permutation matrices.

We form the monoid algebra $K[\mathcal{M}_n]$. For this algebra the magic sum r serves as the degree. As just stated, $K[\mathcal{M}_n]$ is finitely generated by the degree 1 elements representing the permutation matrices. Moreover it has Krull dimension $(n-1)^2 + 1$. After these observations (ADG-1) follows immediately from the theory of Hilbert functions. The equation $H(n, r) = P_n(r)$ holds for all $r > \deg H_n(t)$.

The crucial observation for the proof of (ADG-2)–(ADG-4), if approached from the commutative algebra view point, is the normality of the monoid $\mathcal{M}_n$. Then (ADG-4) follows from Hochster's theorem that normal affine monoid algebras are Cohen–Macaulay rings.

For (ADG-2) and (ADG-3) one has to compute the canonical module ω of $K[\mathcal{M}_n]$. By a theorem of Danilov and Stanley it is the ideal I in $K[\mathcal{M}_n]$ generated by all elements corresponding to strictly positive magic squares. This interior ideal I is easily seen to be a principal ideal, generated by the monomial representing the matrix $\mathbf{1}$; it has magic sum n. Since $-\deg H_r(t)$ is the lowest degree of a non-zero element in ω, it follows that $\deg H_r(t) = -n$. This proves (ADG-2).

The Hilbert series of a graded Cohen–Macaulay ring R and its canonical module ω are related by the functional equation $H_\omega(t) = (-1)^d H_R(t^{-1})$, with $d = \dim R$. Since $\omega \cong K[\mathcal{M}_n]$, more precisely $\omega \cong K[\mathcal{M}_n](-n)$ as graded modules, $K[\mathcal{M}_n]$ is a Gorenstein ring whose canonical module is generated by an element of positive degree, and in this case the functional equation appears in the strong version (ADG-3).

We will develop, or at least discuss, all these results, trying to restrict the prerequisites to a level covered by introductions to commutative algebra, for example Atiyah

and Macdonald [4] or Sharp [21]. Our main tool will be Noether normalizations and homogeneous systems of parameters, but for the proof of (ADG-2) and (ADG-3) more advanced techniques will be needed. Our general reference is Bruns and Herzog [8].

In 2003 conjecture (ADG-5) was proved by Athanasiadis [3] via a reduction to Stanley's g-theorem for simplicial polytopes. His result has now been extended significantly by Römer and the author [10]. Finding a commutative algebra proof of the g-theorem is still a real challenge.

Other developments

The ADG-conjectures were a strong motivation for Stanley's work, as is quite apparent from [27]. Nevertheless, and needless to say, the title of these notes is an exaggeration. At least three other closely related developments must be mentioned:

(i) Hochster [17] proved his theorem on the Cohen–Macaulay property of normal monoid algebras as a significant step on his way towards the Hochster–Roberts theorem by which invariant rings of linearly reductive groups acting on polynomial rings are Cohen–Macaulay. Normal affine monoid rings arise as invariant rings of torus actions.

(ii) A combinatorial problem generalizing the ADG-conjectures had been studied by E. Ehrhart [12]. Let P be an integral or, more generally, a rational polytope, and let $E(P, r)$ denote the number of integral points in the multiple rP, $r \in \mathbb{Z}_+$. If P is the polytope spanned by the $n \times n$-permutation matrices, then $E(P, r) = H(n, r)$. It is not difficult to interpret $E(P, r)$ for arbitrary P as the Hilbert function of a normal monoid algebra. We will discuss it at the end of Section 6.

(iii) Normal affine monoid rings are the coordinate rings of affine and projective toric varieties, and therefore extremely interesting objects for this branch of algebraic geometry. See Danilov [11], Kempf, Knudsen, Mumford, and Saint-Donat [19] or Fulton [14].

1. Affine monoids

Finite generation

Let us first prove that $\mathcal{M}_n$ is indeed generated by the permutation matrices.

Theorem 1.1 (Birkhoff-von Neumann). *The monoid $\mathcal{M}_n$ is generated by the permutation matrices $A_\pi \in \mathcal{M}_1$, $\pi \in S_n$.*

Proof. Let $B = (b_{ij}) \in \mathcal{M}_n$, $B \neq 0$. We must show that there exists a permutation $\pi \in S_n$ such that $B - A_\pi \in \mathcal{M}_n$. For each row $i = 1, \dots, n$ set

$$C_i = \{j : b_{ij} \neq 0\}.$$

Consider $I \subset \{1, \dots, n\}$. Then $\#\bigcup_{i\in I} C_i \geq \#I$, as the reader may verify. Therefore the marriage theorem implies that we can find a permutation $\pi \in S_n$ such that $b_{i\pi(i)} \neq 0$ for $i = 1, \dots, n$. Clearly $B - A_\pi \in \mathcal{M}_n$. □

Without an explicit reference to the magic sum r we could have defined $\mathcal{M}_n$ as the set of all matrices $A = (a_{ij}) \in \mathbb{Z}_+^{n\times n}$ satisfying the system

$$\sum_{k=1}^{n} a_{ik} = \sum_{l=1}^{n} a_{lj}, \qquad i, j = 1, \dots, n, \tag{2}$$

of linear homogeneous equations. More generally, if M is the set of non-negative solutions of a homogeneous diophantine linear system of equations in m variables, then M is certainly a submonoid of $\mathbb{Z}_+^m$, but it is not immediately clear that M, like $\mathcal{M}_n$, has a finite system of generators.

Definition 1.2. A submonoid $M \subset \mathbb{Z}^m$ is called *affine* if it is finitely generated. It is *normal* if it contains all elements $x \in \mathbb{Z}^m$ such that $cx \in M$ for some $c \in \mathbb{Z}$, $c > 0$.

Normal affine monoids are the discrete analogues of polyhedral cones, and some polyhedral geometry is necessary, or at least very useful, for their exploration.

Theorem 1.3. *Let $C \neq \emptyset$ be a subset of $\mathbb{R}^m$.*

(a) *Then the following are equivalent:*

  (i) *there exist finitely many elements $y_1, \dots, y_n \in \mathbb{R}^m$ such that $C = \mathbb{R}_+y_1 + \dots + \mathbb{R}_+y_n$;*
  (ii) *there exist finitely many linear forms $\lambda_1, \dots, \lambda_s$ such that C is the intersection of the half-spaces $H_i^+ = \{x : \lambda_i(x) \geq 0\}$.*

(b) *If C generates $\mathbb{R}^m$ as a vector space and the representation $C = H_1^+ \cap \dots \cap H_s^+$ is irredundant, then the halfspaces in* (ii) *above are uniquely determined (up to enumeration). Equivalently, the linear forms λ_i are unique up to positive scalar factors.*

(c) *The generating elements $y_1, \dots, y_n$ can be chosen in $\mathbb{Q}^m$ (or $\mathbb{Z}^m$) if and only if the λ_i can be chosen as linear forms with rational (or integral) coefficients.*

(d) *If $Y = \{y_1, \dots, y_n\} \subset \mathbb{Q}^m$, then $\mathbb{Q}^m \cap \mathbb{R}_+Y = \mathbb{Q}_+Y$.*

The reader is referred to Ziegler [29] or [6] for a proof. We use the implication (a)(ii)⇒(i) in order to prove *Gordan's lemma*. It shows that the finite generation of $\mathcal{M}_n$ is not an accident.

Theorem 1.4. *Let $\lambda_1, \dots, \lambda_p$ and $\mu_1, \dots, \mu_q$ be $\mathbb{Z}$-linear forms on $\mathbb{Z}^m$, and set*

$$M = \{x : \lambda_i(x) = 0,\ i = 1, \dots, p\} \cap \{x : \mu_j(x) \geq 0,\ j = 1, \dots, q\}.$$

Then M is a normal affine monoid.

Proof. We can replace every equation $\lambda_i(x) = 0$ by two inequalities so that we may assume $p = 0$. Obviously M is a normal monoid.

In $\mathbb{R}^m$ we consider the cone

$$C = \{x \in \mathbb{R}^m : \mu_j(x) \geq 0,\ j = 1, \dots, q\}$$

after the extension of the μ_j to $\mathbb{R}^m$. Then $M = C \cap \mathbb{Z}^m$. By Theorem 1.3 there exist finitely many elements $y_1, \dots, y_n \in M = C \cap \mathbb{Z}^m$ such that every element $x \in C$ can

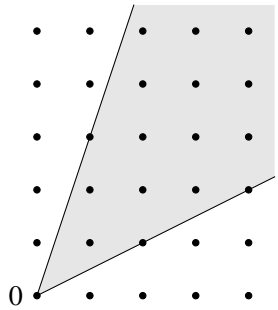

Figure 2. A normal monoid

be written as a linear combination $x = a_1y_1 + \cdots + a_ny_n$ with non-negative coefficients $a_i \in \mathbb{R}_+$. Then

$$x = x' + x'', \qquad x' = \lfloor a_1 \rfloor y_1 + \cdots + \lfloor a_n \rfloor y_n.$$

Both x', x'' belong to M, and $x'' = b_1y_1 + \cdots + b_ny_n$ with $b_j \in [0, 1)$, $j = 1, \ldots, n$. Thus x'' belongs to the intersection of $\mathbb{Z}^m$ with a bounded set B, and $y_1, \ldots, y_n$ together with the elements of the finite set $M \cap B$ generate M as a monoid. □

Normality and purity

The set M of non-negative solutions of a homogeneous linear diophantine system of equations in m variables is a special type of submonoid of $\mathbb{Z}_+^m$: let $G = \mathrm{gp}(M)$ be the subgroup of $\mathbb{Z}^m$ generated by M; then $M = G \cap \mathbb{Z}_+^m$. One says that M is a *pure* submonoid of $\mathbb{Z}_+^m$.

We want to show that every normal affine monoid M can be realized as a pure submonoid of $\mathbb{Z}_+^s$ for suitable s, provided M is *positive*: $x, -x \in M \Rightarrow x = 0$. Positivity is an evident necessary condition for the embeddability into $\mathbb{Z}_+^s$.

For brevity we let "affine" include "positive" in the following since we are mainly interested in normal affine monoids M, and every such monoid decomposes into the direct sum

$$M = U(M) \oplus M' \tag{3}$$

where $U(M) = \{x \in M : -x \in M\}$ and M' is normal, affine and positive. In fact, the normality of M implies that $\mathbb{Z}^m/U(M)$ is torsionfree. Hence $U(M)$ is a free direct summand of $\mathbb{Z}^m$, and we can choose M' as the image of M under the natural projection $\mathbb{Z}^m \mapsto \mathbb{Z}^m/U(M)$. In our context, direct summands that are groups are almost irrelevant.

Proposition 1.5. *Let $M \subset \mathbb{Z}^m$ be a affine monoid (positive, as understood). Then the following are equivalent:*

(a) *M is normal;*
(b) *$M = \mathbb{R}_+M \cap \mathrm{gp}(M)$;*
(c) *M is isomorphic to a pure submonoid of $\mathbb{Z}_+^s$ for suitable s.*

Proof. Suppose that $M \subset \mathbb{Z}^m$ is normal, and let $x \in \mathbb{R}_+M \cap \mathrm{gp}(M)$. The cone $\mathbb{R}_+M$ is generated by finitely many elements $y_1, \dots, y_n \in M$, and $x = a_1y_1 + \cdots + a_ny_n$ with $a_1, \dots, a_n \in \mathbb{Q}_+$ by Theorem 1.3. Let $a \in \mathbb{Z}$, $a > 0$, be a common denominator for $a_1, \dots, a_n$. Then $ax \in M$, and $x \in M$ by hypothesis. This shows (a) $\Rightarrow$ (b).

For (b) $\Rightarrow$ (c) we identify the subgroup $\mathrm{gp}(M)$ of $\mathbb{Z}^m$ with $\mathbb{Z}^r$, $r = \operatorname{rank} \mathrm{gp}(M)$. The number r is called the *rank* of M. We can replace $\mathbb{Z}^m$ by $\mathbb{Z}^r$. Let $C = \mathbb{R}_+M$ be the cone generated by M in $\mathbb{R}^r$. By Theorem 1.4 there exist finitely many linear forms $\sigma_1, \dots, \sigma_s$ on $\mathbb{R}^r$ with integral coefficients such that:

(i) C is the intersection of the half-spaces $H_i^+ = \{x \in \mathbb{R}^r : \sigma_i(x) \geq 0\}, i = 1, \dots, s$;
(ii) the representation $C = H_1^+ \cap \cdots \cap H_s^+$ is irredundant;
(iii) $\sigma_i(\mathbb{Z}^r) = \mathbb{Z}$, $i = 1, \dots, s$.

Up to enumeration, the forms $\sigma_1, \dots, \sigma_s$ are uniquely determined. (Property (iii) can be achieved by extracting the greatest common divisor of the coefficients.) We call them the *support forms*.

Consider the $\mathbb{Z}$-linear map $\sigma : \mathbb{Z}^r \to \mathbb{Z}^s$, $\sigma(x) = (\sigma_1(x), \dots, \sigma_s(x))$. Since M is positive, σ is injective. It maps M isomorphically on a submonoid M' of $\mathbb{Z}_+^s$. Every element x of $\mathrm{gp}(M')$ belongs to $\sigma(\mathrm{gp}(M))$, and $x \in M'$ if and only if $\sigma(x) \in \mathbb{Z}_+^s$.

The remaining implication (c) $\Rightarrow$ (a) is trivial. □

Monoid algebras

Let K be a field and M a commutative monoid. Then the monoid algebra $K[M]$ is a vector space with a basis indexed by the elements of M for which the addition in M serves as the multiplication table. If $M \subset \mathbb{Z}^m$ is an affine monoid, then $K[M]$ is a subalgebra of $K[\mathbb{Z}^m]$ since the monoid embedding $M \hookrightarrow \mathbb{Z}^m$ extends to an algebra embedding $K[M] \hookrightarrow K[\mathbb{Z}^m]$. The monoid algebra $K[\mathbb{Z}^m]$ is isomorphic to the Laurent polynomial ring $K[X_1^{\pm 1}, \dots, X_m^{\pm 1}]$ via the assignment

$$a = (a_1, \dots, a_m) \mapsto X^a = X^{a_1} \cdots X^{a_m}.$$

It follows that the monoid algebras $K[M]$, M affine, can be identified with subalgebras of the Laurent polynomial rings $K[X_1^{\pm 1}, \dots, X_m^{\pm 1}]$ that are generated by finitely many monomials. By Hilbert's basis theorem such algebras are Noetherian, and so we can apply the rich structure theory of Noetherian rings to the study of affine monoids.

The terms "normal" and "pure" have been chosen judiciously:

Theorem 1.6. *Let M be an affine monoid.*

(a) *Then M is normal if and only if $K[M]$ is a normal ring, i. e. integrally closed in its field of fractions.*
(b) *$M \subset \mathbb{Z}_+^s$ is a pure submonoid if and only if $K[M]$ is a pure subalgebra of $K[\mathbb{Z}_+^s] = K[Y_1, \dots, Y_s]$, i. e. $K[\mathbb{Z}_+^s]$ splits into the direct sum $K[M] \oplus T$ of $K[M]$ and a $K[M]$-submodule T of $K[Y_1, \dots, Y_s]$.*

Proof.

(a) Let $x \in \mathrm{gp}(M)$ such that $mx \in M$ for some $m \in \mathbb{Z}$, $m > 0$. For the algebra $K[M]$ and the monomial $\tilde{x} \in K[\mathrm{gp}(M)] \subset \mathrm{QF}(K[M])$ with exponent vector x this implies $\tilde{x}^m \in K[M]$. If $K[M]$ is normal, then $\tilde{x} \in K[M]$ or, equivalently, $x \in M$.

Conversely, let M be normal, and set $C = \mathbb{R}_+M$. We identify $\mathrm{gp}(M)$ with $\mathbb{Z}^r$. Then $M = \mathbb{Z}^r \cap C$ and $C = H_1^+ \cap \cdots \cap H_s^+$ where each H_i^+ is a closed halfspace defined by a linear form with integral coefficients. Thus

$$M = N_1 \cap \cdots \cap N_s, \qquad N_i = \mathbb{Z}^r \cap H_i^+.$$

Consequently

$$K[M] = K[N_1] \cap \cdots \cap K[N_s].$$

Since an intersection of normal domains is normal it is sufficient that $K[N]$ is normal for the intersection N of $\mathbb{Z}^r$ with a half-space H^+ defined by an integral linear form. Now

$$N \cong \mathbb{Z}^{r-1} \oplus \mathbb{Z}_+,$$

as follows from the splitting (3) of N: for $x \in \mathbb{Z}^r$ one has $x, -x \in N$ if and only if x lies in the hyperplane bounding the halfspace H^+. Thus $K[N]$ is isomorphic to a partial Laurent polynomial ring $K[Y_1^{\pm 1}, \ldots, Y_{r-1}^{\pm 1}, Z]$. It is even a factorial ring, which is certainly normal.

(b) Let $R = K[Y_1, \ldots, Y_s]$ and $S = K[M]$. Set $W = \mathbb{Z}_+^s \setminus M$ and let U be the vector subspace of $R = K[\mathbb{Z}_+^s]$ generated by the monomials with exponent vector in W. The purity of M translates into the condition $M + W \subset W$ or, equivalently, $SU \subset U$. Since $R = S \oplus U$ as a K-vector space, it is of course sufficient for the splitting of R that U is an S-submodule of $K[Y_1, \ldots, Y_S]$.

Conversely, if $R = S \oplus T$ for some S-submodule T, then one certainly has the equation

$$I = IR \cap S$$

for all ideals I of S since the ideal IR of R decomposes into the direct sum $I \oplus IT$ (the reader should verify this claim). Take I as the principal ideal generated by the monomial corresponding to $x \in M$. If $x + W$ contains an element $y \in M$, then IR contains the monomial corresponding to y, an element of S which, on the other hand, is not in I. □

In a more general context, one uses a weaker property to define the purity of a ring extension; see [8, p. 293].

In order to avoid cumbersome formulations we will sometimes consider the monoids as subsets of their algebras, identifying each monoid element with the monomial of which it is the exponent vector.

The class group

We have used the support forms associated with the facets of the cone $\mathbb{R}_+M$ in order to realize $R = K[M]$ as a pure subalgebra of a polynomial ring. Let us observe that they also define prime ideals in R that determine the ideal theory of R to a great extent. Let σ_i be a support form and set

$$\mathfrak{p}_i = \sum_{x\in M,\ \sigma_i(x)\geq 1} Kx.$$

Then $\mathfrak{p}_i$ is an ideal of R, and even prime:

$$K[M]/\mathfrak{p}_i \cong K[M \cap F_i]$$

where F_i is the facet of $\mathbb{R}_+M$ belonging to σ_i. The *symbolic powers* of $\mathfrak{p}_i$ are defined as follows:

$$\mathfrak{p}_i^{(c)} = \sum_{x\in M,\ \sigma_i(x)\geq c} Kx.$$

One ideal in $K[M]$ will receive our special attention, namely

$$\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_s.$$

It is the "interior" ideal generated by all monomials in the interior of the cone $\mathbb{R}_+M$.

More generally, let R be a Noetherian normal domain (or just a Krull domain). A subset J of the field $\mathrm{QF}(R)$ of fractions of R is a *fractional ideal* if there exists $x \in R$, $x \neq 0$, such that xJ is an ideal of R. A fractional ideal is divisorial if $(J^{-1})^{-1} = J$; here $J^{-1} = \{x \in K : xJ \subset R\}$ can be identified with $\mathrm{Hom}_R(J, R)$. From the view point of linear algebra, the divisorial fractional ideals are exactly the reflexive ones: $(J^{-1})^{-1} = J$ if and only if the natural homomorphism $J \to \mathrm{Hom}_R(\mathrm{Hom}(J, R), R)$ is an isomorphism.

The isomorphism classes of divisorial ideals are parameterized by a classical invariant, the *class group*; see Fossum [13]. (In the number-theoretic situation R is a Dedekind domain, and over such domain every fractional ideal is divisorial.) For a normal affine monoid ring $R = K[M]$ the class group can be calculated as follows: let x be a monomial in the interior of $\mathbb{R}_+M$. Then $R[x^{-1}]$ contains all monomials corresponding to elements in $\mathbb{Z}^d$, and therefore is the Laurent polynomial ring $L = K[\mathbb{Z}^d]$, a factorial ring. Since x is a monomial, the ideal Rx has a vector space of monomials, and by the normality of M the quotient y/x belongs to M if and only if $\sigma_i(y) \geq \sigma_i(x)$; therefore

$$Rx = \mathfrak{p}_1^{(v_1)} \cap \cdots \cap \mathfrak{p}_s^{(v_s)}, \qquad v_i = \sigma_i(x).$$

Nagata's theorem [13] implies that the class group is generated by the classes of the divisorial prime ideals containing x. These are exactly the monomial prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ corresponding to the facets, as defined above. This implies: for each divisorial ideal J there exist integers $c_1, \ldots, c_s \in \mathbb{Z}$ such that

$$J \cong \mathfrak{p}_1^{(c_1)} \cap \cdots \cap \mathfrak{p}_s^{(c_s)}. \tag{4}$$

Figure 3 illustrates this construction. See Bruns and Gubeladze [7] for a detailed study of the class group of a normal monoid algebra.

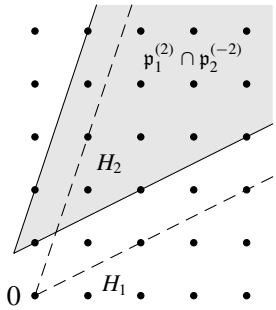

Figure 3. A divisorial fractional ideal

2. Graded rings and modules

The monoid algebra $R = K[\mathcal{M}_n]$ has a natural decomposition as a vector space: since $\mathcal{M}_n = \bigcup_{r=0}^{\infty} \mathcal{M}_{nr}$ we have $R = \bigoplus_{r=0}^{\infty} R_r$ where R_r is the vector subspace spanned by the monomials corresponding to the $n \times n$ magic squares of magic sum r. Since $\mathcal{M}_{nr} + \mathcal{M}_{ns} \subset \mathcal{M}_{n,r+s}$, this decomposition is compatible with the multiplicative structure: $R_r R_s \subset R_{r+s}$. Thus R is a graded ring:

Definition 2.1. A *grading* of a ring R is a decomposition $R = \bigoplus_{i\in\mathbb{Z}} R_i$ of R as an abelian group such that $R_i R_j \subset R_{i+j}$ for all i, j. A *graded ring* is a ring together with a grading.

A *grading* of a module M over a graded ring R is a decomposition $M = \bigoplus_{i\in\mathbb{Z}} M_i$ of R as an abelian group such that $R_i M_j \subset M_{i+j}$ for all i, j. A *graded R-module* is an R-module together with a grading.

The elements of M_i are *homogeneous* of degree i.

The main example of a graded ring is the polynomial ring $S = K[X_1, \dots, X_n]$ with the grading in which the graded component of degree i is the vector space spanned by the monomials of total degree i. More generally we can define a grading on S by assigning degrees $g_i \in \mathbb{Z}$ to the indeterminates X_i and setting

$$\deg X_1^{a_1} \cdots X_n^{a_n} = \sum_{j=1}^{n} a_j g_j.$$

The degree i component is again spanned by the monomials of degree i.

Let M be an affine monoid. Then a $\mathbb{Z}$-linear map $\gamma : \mathrm{gp}(M) \to \mathbb{Z}$ is called a grading on M. It induces a grading on $K[M]$ in which the degree i component is spanned by the elements $x \in M$ with $\gamma(x) = i$. We say that γ is *positive* if $\gamma(x) > 0$ for all $x \in M$, $x \neq 0$. It follows from Proposition 1.5 that every (normal) affine monoid admits a positive grading, as does every submonoid of $\mathbb{Z}_+^s$.

A submodule U of M is graded if $U = \bigoplus_{i\in\mathbb{Z}} U \cap M_i$. In this case the residue class module M/U is naturally graded, too, since $M/U = \bigoplus_{i\in\mathbb{Z}} M_i/(U \cap M_i)$ by a natural isomorphism.

The classical example of a graded submodule is an ideal I in the polynomial ring S as above that is generated by homogeneous polynomials. Then S/I is the homogeneous coordinate ring of the projective variety defined by I.

A *homogeneous* R-linear map $\varphi : M \to N$ of graded R-modules preserves the graded structure: one has $\varphi(M_i) \subset N_i$. The kernel and the cokernel of a graded homomorphism are graded R-modules.

In order to make certain natural linear maps homogeneous we must shift the degree of one (or both) of the modules involved: the module $M(s)$ is identical to M as an R-module, but

$$M(s)_i = M_{s+i}.$$

In other words, degree j elements in M have degree $j - s$ in $M(s)$. Most often one needs shifts that increase the degrees of elements. Therefore they are usually written with a negative sign.

Let K be a field. We say that a graded K-algebra R is *positively graded* if $R = K[x_1, \ldots, x_n]$ with homogeneous elements x_i of positive degree. Then $R_i = 0$ for $i < 0$ and $R_0 = K$.

Graded Ext functors

At a certain point in Section 6 it seems inevitable to use derived functors, and we need that $\operatorname{Ext}^i_R(M, N)$ carries a natural grading for finitely generated graded modules M, N over a Noetherian graded ring R. First we endow $\operatorname{Hom}_R(M, N)$ with a natural grading: its homogeneous component of degree i consists of all R-linear maps φ such that $\varphi(M_j) \subset \varphi(N_{i+j})$ for all j. We leave it to the reader to show that $\operatorname{Hom}_R(M, N)$ is indeed the sum of these abelian subgroups.

Next, one constructs a graded free resolution of M,

$$\mathcal{F} : \cdots \to F_n \xrightarrow{\varphi_n} F_{n-1} \to \cdots \to F_1 \xrightarrow{\varphi_1} F_0 \to 0, \qquad M = \operatorname{Coker} \varphi_1,$$

in which the F_i are finitely generated graded free modules and the morphisms $F_{i+1} \to F_i$ are homogeneous. It follows that $\operatorname{Ext}^i_R(M, N) = H^i(\operatorname{Hom}_R(\mathcal{F}, N))$ is certainly a graded R-module, and the grading does not depend on the choice of the homogeneous free (or projective) resolution. Evidently such resolutions exist: if M is graded, generated by elements $x_1, \ldots, x_m$ of degrees $g_1, \ldots, g_m$, then

$$F_0 = \bigoplus_{i=1}^{m} R(-g_i) \to M, \qquad e_i \mapsto x_i,$$

is a homogeneous R-linear map from a graded free R-module onto M. Its kernel is again a finitely generated graded R-module, onto which we can map a graded free R-module etc.

For numerical applications it is usually best to group the summands $R(-g_i)$ according to their shifts:

$$\bigoplus_{i=1}^{m} R(-g_i) = \bigoplus_{j \in \mathbb{Z}} R(-j)^{\beta_j}, \qquad \beta_j = \#\{i : g_i = j\}.$$

The notion of grading can be generalized by allowing an arbitrary group or even an arbitrary commutative monoid as the set of degrees for R and M.

3. Krull dimension and Noether normalization

Let R be a Noetherian ring. The *Krull dimension* $\dim R$ of R is the supremum of all integers n such that there exists a strictly ascending chain

$$\mathfrak{p}_0 \subset \mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n$$

of prime ideals. A priori, the value ∞ is not excluded, but it cannot occur for finitely generated algebras R over a field K. The Krull dimension of a positively graded K-algebra can be described as follows:

Theorem 3.1. *Let $R = \bigoplus_{k=0}^{\infty} R_k$ be a positively graded, finitely generated K-algebra.*

(a) *Then* $\dim R$ *is the smallest integer d for which there exist homogeneous elements $x_1, \ldots, x_d \in R$ such that $\mathfrak{m} = \bigoplus_{k=1}^{\infty} R_k$ is the radical of the ideal $(x_1, \ldots, x_d)$.*
(b) *Such elements $x_1, \ldots, x_d$ are algebraically independent over K, and R is a finitely generated module over $K[x_1, \ldots, x_d]$.*
(c) *If K is infinite and $R = K[R_1]$, then $x_1, \ldots, x_d$ can be chosen of degree* 1.

Elements $x_1, \ldots, x_d$ as in Theorem 3.1 are said to form a *homogeneous system of parameters* for R. More generally, if M is an R-module, then a *homogeneous system of parameters for M* is just a homogeneous system of parameters for $R/\operatorname{Ann} M$. (By $\operatorname{Ann} M$ we denote the annihilator ideal $\{x \in R : xM = 0\}$.) The subalgebra $K[x_1, \ldots, x_d]$ is called a *graded Noether normalization* of R. It is isomorphic to a polynomial ring over K, and so Theorem 3.1 shows that every finitely generated graded module over R, in particular R itself, can be understood as such a module over the polynomial ring $K[X_1, \ldots, X_d]$. We will intensively use this argument in the next sections.

We cannot include a full proof of Theorem 3.1. The first step is to show that the homogeneous elements in a prime ideal $\mathfrak{p}$ of a graded ring generate again a prime ideal (see [8, 1.5.6]). It follows that all minimal prime ideals of a graded ring are graded. Starting from this basic fact one constructs a chain of graded prime ideals

$$\mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_d = \mathfrak{m}$$

of length $d = \dim R$. (Since every homogeneous ideal of R, except R itself, is contained in $\mathfrak{m}$, all such chains must end in $\mathfrak{m}$.) Now Krull's principal ideal theorem shows that $\mathfrak{m}$ cannot be the radical of an ideal generated by fewer than d elements.

The existence of $x_1, \ldots, x_d$ is shown by induction on d, based on the following lemma:

Lemma 3.2. *Let I be an ideal in a graded ring, generated by elements of positive degree, and let $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ be prime ideals such that $I \not\subset \mathfrak{p}_i$ for $i = 1, \ldots, m$.*

(a) *Then I contains a homogeneous element $x \notin \mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_m$.*
(b) *If R is a graded algebra over an infinite field K and I is generated by elements $x_1, \ldots, x_n$ of the same degree g, then there exist $a_1, \ldots, a_n \in K$ such that $x = a_1x_1 + \cdots + a_nx_n \notin \mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_m$.*

For the combinatorial application we are free to choose the field K (and often one can extend K anyway). Therefore it may be enough to prove (b): by hypothesis the intersection $\mathfrak{p}_i \cap V$ is a proper subspace of $V = Kx_1 + \cdots + Kx_n$, and a vector space over an infinite field cannot be the union of finitely many proper subspaces.

For the proof of Theorem 3.1 suppose first that $d = 0$. Then $\mathfrak{m}$ is the only prime ideal in R, and every element in $\mathfrak{m}$ is nilpotent, since the nilradical of R is the intersection of its prime ideals. Therefore $\mathfrak{m}$ is the radical of the ideal generated by the empty set.

Now let $d \geq 1$. To conclude the proof of Theorem 3.1(a) we apply the lemma to $I = \mathfrak{m}$ and the minimal prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ of R, choosing $x_1 = x$. Then $R/(x_1)$ has Krull dimension $d-1$, and we are done by induction, lifting a homogeneous system of parameters of $R/(x_1)$ to R.

This argument shows that we can find a homogeneous system of parameters consisting of degree 1 elements if R is generated over K in degree 1 and K is infinite.

It remains to prove (c). Let I be the ideal in R generated by $x_1, \ldots, x_d$ and $y_1, \ldots, y_m$ be a homogeneous system of generators of the K-algebra R. Since $y_i^e \in I$ for $e \gg 0$, there exists $f \in \mathbb{Z}$ such that $R_i \subset I$ for $i \geq f$. Then a vector space basis of $R_0 \oplus \cdots \oplus R_{f-1}$ generates R as an $K[x_1, \ldots, x_d]$-module. The reader may show this as an exercise or look up [8, p. 38].

For the algebraic independence of $x_1, \ldots, x_d$ it is sufficient that $\dim K[x_1, \ldots, x_d] = \dim R = d$. This follows from the going up theorem for module-finite or, more generally, integral ring extensions.

Krull dimension and transcendence degree

Suppose that R is an integral domain, $y \in R$, and let S be a graded Noether normalization. Since R is a finitely generated S-module, and submodules of finitely generated modules over Noetherian rings are again finitely generated, there exists an exponent q such that y^q is an S-linear combination of the powers $y^0 = 1, y_1, \ldots, y^{q-1}$. It follows that the quotient field $\mathrm{QF}(R)$ is an algebraic extension of $\mathrm{QF}(S)$, and the latter has transcendence degree $d = \dim R$ over K. Therefore d is the transcendence degree of $\mathrm{QF}(R)$ over K.

Let M be an affine monoid. Then $\mathrm{QF}(K[M]) = \mathrm{QF}(K[\mathrm{gp}(M)])$, and the transcendence degree of the latter equals $\mathrm{rank}\, M$. So $\dim K[M] = \mathrm{rank}\, M$.

We use this fact to show that $\dim K[\mathcal{M}_n] = (n-1)^2 + 1$. Let $G = \mathrm{gp}(\mathcal{M}_n)$. Let $\mathbf{1}$ be the magic $n \times n$-square with all entries equal to 1, and $A \in \mathbb{Z}^{n \times n}$ a solution of the homogeneous system of equations (2) defining $\mathcal{M}_n$, without sign restriction. Then $A + k\mathbf{1} \in \mathcal{M}_n$ for $k \gg 0$, and it follows that $\mathrm{gp}(\mathcal{M}_n)$ is the group of all solutions to the system (2). The system (2) has rank $2n-2$, so $\mathrm{rank}\, \mathcal{M}_n = n^2 - (2n-2) = (n-1)^2 + 1$.

4. Hilbert functions of graded modules

Suppose that K is a field, considered as a graded ring in which all elements have degree 0, and let $V = \bigoplus_{i \in \mathbb{Z}} V_i$ be a graded K-vector space. Then the *Hilbert function*

$H(M,\_)$ of V is defined as follows:

$$H(M,\_) : \mathbb{Z} \to \mathbb{Z} \cup \{\infty\}, \qquad H(M,i) = \dim_K M_i.$$

A fundamental property of the Hilbert function is its additivity along exact sequences of homogeneous linear maps: If $0 \to V^m \to \cdots \to V^0 \to 0$ is an exact and homogeneous sequence of graded K-vector spaces V^i, then $\sum(-1)^i H(V^i,\_) = 0$. This follows immediately from the additivity of K-dimension, applied to the homogeneous components.

Let R be a K-algebra. If R is graded, then the graded components R_i are automatically K-vector spaces since $K \subset R_0$, and for the same reason the graded components M_i of a graded R-module M are K-vector spaces. If R is a positively graded, finitely generated K-algebra, then the Hilbert function of a finitely generated graded R-module M takes only finite values and moreover $H(M,i) = 0$ for $i \ll 0$. Therefore the *Hilbert* (or *Poincaré*) *series*

$$H_M(t) = \sum_{i\in\mathbb{Z}} H(M,i)t^i$$

is a Laurent series with a finite principal part. It belongs to the field $\mathbb{C}[\![t]\!][t^{-1}]$. Note that $\mathbb{C}[\![t]\!][t^{-1}]$ contains the polynomial ring $\mathbb{C}[t]$ and therefore the field $\mathbb{C}(t)$ of rational functions. The embedding $\mathbb{C}(t) \hookrightarrow \mathbb{C}[\![t]\!][t^{-1}]$ identifies a rational function with its Laurent expansion at 0. In the following we will use the same notation for a rational function and its Laurent expansion at 0.

The fundamental structural result about $H_M(t)$ is the following theorem.

Theorem 4.1 (Hilbert–Serre). *Let R be a graded K-algebra generated by homogeneous elements $x_1, \dots, x_n$ of degrees $g_1, \dots, g_n > 0$. Furthermore let M be a non-zero, finitely generated graded R-module. Then there exists a Laurent polynomial $Q \in \mathbb{Z}[t, t^{-1}]$ such that*

$$H_M(t) = \frac{Q(t)}{\prod_{i=1}^n (1 - t^{g_i})}.$$

Proof. We use induction on n. In the case $n = 0$ we can simply take $Q = H_M$. Let $n > 0$. Then multiplication on M by x_n induces an exact sequence

$$0 \to (0 :_M x_n)(-g_n) \to M(-g_n) \xrightarrow{x_n} M \to M/x_nM \to 0.$$

Both the kernel $U = (0 :_M x_n)$ of the multiplication map and $N = M/x_nM$ are finite modules over the graded algebra $R/(x_n)$ (or over $K[x_1, \dots, x_{n-1}]$). Since the Hilbert function is additive along exact sequences, we obtain

$$(1 - t^{g_n})H_M = H_N - t^{g_n} H_U.$$

It only remains to apply the induction hypothesis. We are allowed to divide by $1 - t^{g_n}$ in the field $\mathbb{C}[\![t]\!][t^{-1}]$, to which H_M, H_U, H_N belong. (Here it is used that $g_i > 0$ for all i.) □

In particular, if $R = K[X_1, \dots, X_n]$ is the polynomial ring over the field K, graded in such a way that $\deg X_i = g_i$, $i = 1, \dots, n$, then

$$H_R(t) = \frac{1}{\prod_{i=1}^{n}(1 - t^{g_i})}. \tag{5}$$

The formula follows by the same induction with $x_i = X_i$. In this case the kernel U is always 0, since X_n is a non-zero-divisor on R.

Algebras generated in degree 1

Note that the Hilbert function of a graded module M depends only on the decomposition of M as a K-vector space. We are therefore free to replace R by any other graded K-algebra as long as the grading of M is not changed. This allows us to refine Theorem 4.1. First we consider the classical case in which $g_1, \dots, g_n = 1$ and the Hilbert function is of polynomial type:

Theorem 4.2. *Under the hypotheses of Theorem 4.1 suppose that* $g_1 = \cdots = g_n = 1$ *and let* $d = \dim M$. *Then there exists a Laurent polynomial* $Q \in \mathbb{Z}[t, t^{-1}]$ *such that*

$$H_M(t) = \frac{Q(t)}{(1 - t)^d}.$$

Moreover:

(a) *There exists a polynomial* $P_M \in \mathbb{Q}[X]$ *such that*

$$H(M, i) = P_M(i), \qquad i > \deg H_M,$$
$$H(M, i) \neq P_M(i), \qquad i = \deg H_M.$$

(b) $e(M) = Q(1) > 0$, *and if* $d \geq 1$, *then*

$$P_M = \frac{e(M)}{(d-1)!}X^{d-1} + \textit{terms of lower degree}.$$

Proof. We may assume that K is infinite. Otherwise we choose an infinite extension L of K, replace R by $R \otimes_K L$, and M by $M \otimes_K L$. Next we substitute $R/\operatorname{Ann} M$ for R, and can assume that $d = \dim R$. Then there exists a Noether normalization S of R generated by (necessarily) d homogeneous elements of degree 1. Since R is a finitely generated S-module, so is M, and direct specialization of Theorem 4.1 yields

$$H_M(t) = \frac{Q(t)}{(1 - t)^d}.$$

We expand into a Laurent series:

$$\frac{t^j}{(1-t)^d} = \sum_{k=0}^{\infty} \binom{k+d-1}{d-1} t^{k+j} = \sum_{i \in \mathbb{Z}} c_j(i) t^i$$

with

$$c_j(i) = \binom{i+d-1-j}{d-1}, \quad i \geq j, \qquad \text{and} \qquad c_j(i) = 0, \quad i < j.$$

Set

$$p_j(X) = \binom{X+d-1-j}{d-1} = \frac{(X+d-1-j)\cdots(X+1-j)}{(d-1)!}.$$

Then $p_j \in \mathbb{Q}[X]$, and we have

$$c_j(i) = p_j(i), \quad i > j-d, \qquad \text{and} \qquad c_j(i) \neq p_j(i), \quad i = j-d.$$

With $Q(t) = h_a t^a + \cdots + h_b t^b$, $a, b \in \mathbb{Z}$, $a \leq b$, $h_a, h_b \neq 0$, the desired polynomial is

$$P_M = \sum_{j=a}^{b} h_j p_j.$$

That $Q(1) > 0$ will be shown in Proposition 4.5. Then the assertion on the leading coefficient of P_M follows since each of the polynomials p_j has degree $d-1$ (if $d \geq 1$) and leading coefficient $1/(d-1)!$. Therefore P_M has leading coefficient

$$\sum_{j=a}^{b} \frac{h_j}{(d-1)!} = \frac{Q(1)}{(d-1)!}.$$

□

Definition 4.3. The polynomial P_M is called the *Hilbert polynomial* of M, and $e(M)$ is the *multiplicity* of M.

Positively graded algebras

If M cannot be written as a finitely generated module over a K-algebra generated in degree 1, then its Hilbert function need not be of polynomial type. It is however of *quasi-polynomial type*: let $P_0, \ldots, P_{m-1}$ be polynomials. Then the function P,

$$P(km+\ell) = P_\ell(k), \qquad k \in \mathbb{Z},\ \ell = 0, \ldots, m-1,$$

is called a *quasi-polynomial of period* m.

Theorem 4.4. *Under the hypotheses of Theorem 4.1 suppose that* $d = \dim M$. *Then there exist a Laurent polynomial* $Q \in \mathbb{Z}[t, t^{-1}]$ *and integers* $e_1, \ldots, e_d > 0$ *such that*

$$H_M(t) = \frac{Q(t)}{\prod_{i=1}^{d}(1-t^{e_i})}, \qquad Q(1) > 0.$$

Moreover there exists a quasi-polynomial P_M, *whose period divides* $\operatorname{lcm}(e_1, \ldots, e_d)$, *such that*

$$H(M,i) = P_M(i), \qquad i > \deg H_M,$$
$$H(M,i) \neq P_M(i), \qquad i = \deg H_M.$$

Proof. As above we can assume that K is infinite and replace R by a graded Noether normalization of $R/\operatorname{Ann} M$. Then $e_1, \dots, e_d$ can be chosen as the degrees of the elements in a homogeneous system of parameters. The formula for H_M follows again by direct specialization of Theorem 4.1.

With $m = \operatorname{lcm}(e_1, \dots, e_d)$ we choose $S = R^{(m)} = \bigoplus_{k=0}^{\infty} R_{km}$, the m-th Veronese subalgebra of R. Dividing by m, we normalize the degrees in S. While S need not be generated in degree 1, it is a finite module over $S' = K[S_1] = K[R_m]$, as the reader may check. As an S'-module M splits into the direct sum of its submodules

$$N_\ell = \bigoplus_{k \equiv \ell \ (m)} M_k, \qquad \ell = 0, \dots, m-1.$$

Now we can compose P from the Hilbert polynomials of the N_ℓ. The reader should check that the assertion on the equality of $P(i)$ and $H(M, i)$ follows from the corresponding statement in Theorem 4.2.

Again it remains to show $Q(1) > 0$, but this will be done in the next proposition. □

The theorem contains the statement that the Krull dimension of M is the pole order of H_M at $t = 1$. The proof is not yet complete, but we know at least that the pole order is bounded above by the Krull dimension. We use this fact in the proof of

Proposition 4.5. *Under the hypothesis of Theorem 4.4 suppose that $e_1, \dots, e_d$ are the degrees of the elements in a homogeneous system of parameters $x_1, \dots, x_d$ for M. Let $S = K[x_1, \dots, x_d]$. Then $Q(1) = \operatorname{rank}_S M > 0$.*

Proof. The annihilator of M as a module over S is 0. We can replace R by S, but must keep in mind that M has annihilator 0. The polynomial ring R over K is an integral domain and every R-module has a well-defined rank, given by the vector space dimension of $M \otimes \mathrm{QF}(R)$ over $\mathrm{QF}(R)$. There exist homogeneous elements $y_1, \dots, y_r \in M$ that form a basis of $M \otimes \mathrm{QF}(R)$ over $\mathrm{QF}(R)$. Let F be the R-submodule of M generated by $y_1, \dots, y_r$. Then $y_1, \dots, y_r$ are linearly independent, and there is an exact sequence of graded modules

$$0 \to F \to M \to N \to 0, \qquad N = M/F.$$

Since N has non-zero annihilator, $\dim N < d$. With self-explaining notation we have

$$H_N(t) = \frac{Q_M - Q_F}{\prod_{i=1}^{d}(1 - t^{e_i})}.$$

Since $\dim N < d$, the difference $Q_M - Q_F$ must be divisible by $1 - t$ since the pole order of $H_N(t)$ at $t = 1$ is smaller than d. Therefore $Q_M(1) = Q_F(1)$, and that $Q_F(1) = \operatorname{rank} F = \operatorname{rank} M$ is clear:

$$H_F(t) = \frac{f_a t^a + \cdots + f_b t^b}{\prod_{i=1}^{d}(1 - t^{e_i})}$$

where f_i counts the number of degree i elements in a homogeneous basis of F. □

Remark 4.6.

(a) The module-theoretic argument we have used in the proof of Theorem 4.4 can be replaced by a formal reasoning. It amounts to writing $H_M(t)$ as a rational function with denominator $(1-U)^d$ where $U = T^m$ and to split the numerator according to the residue classes modulo m of the exponents of the T^j. It follows that the assertion on the period of P does not depend on the fact that we can really find a Noether normalization generated in degrees $e_1, \dots, e_d$.

Similarly $Q(1) > 0$ whenever $e_1, \dots, e_d$ admit a representation of $H_M(t)$ with denominator $\prod_{i=1}^d (1-t^{e_i})$. In fact two such denominators differ only by polynomial factors $1 + t + \cdots + t^f$, having positive value at 1.

(b) Suppose that there exists a homogeneous element y in R that is a non-zero-divisor of M and whose degree is coprime to m. Then the Hilbert polynomials of the modules N_ℓ have the same degrees and the same leading coefficients. This holds since there exist $u_{ij} \in \mathbb{Z}_+$ for $i, j = 0, \dots, m-1$ such that $y^{u_{ij}} N_i \subset N_j$. This result can be substantially extended: let I be the ideal of R generated by those homogeneous elements that are coprime to m; if $\dim M/IM \le k$, then the coefficients of the degree j terms in the Hilbert polynomials coincide for all $j \ge k$. See [9].

Hilbert functions and homological invariants

The Hilbert series can be calculated from a graded finite free resolution. In the next theorem we use the standard convention of writing the shifts with a negative sign and grouping the terms in a free graded module $R(-s_1) \oplus \cdots \oplus R(-s_m)$ according to their shifts.

Theorem 4.7. *Let M be a finite graded R-module of finite projective dimension, and let*

$$0 \to \bigoplus_j R(-j)^{\beta_{pj}} \to \cdots \to \bigoplus_j R(-j)^{\beta_{0j}} \to M \to 0$$

be a graded free resolution of M. Then

$$H_M(t) = S_M(t) H_R(t)$$

where $S_M(t) = \sum_{i,j} (-1)^i \beta_{ij} t^j$. In particular, if $R = K[X_1, \dots, X_n]$ is the polynomial ring over the field K, graded in such a way that $\deg X_i = g_i$, $i = 1, \dots, n$, *then*

$$H_M(t) = \frac{S_M(t)}{\prod_{i=1}^n (1-t^{g_i})}.$$

Proof. This follows immediately from the additivity of Hilbert functions along exact sequences. The formula for the Hilbert function of the polynomial ring has already been proved above. □

If $R = K[X_1, \dots, X_n]$, then every finitely generated module over R has a finite free resolution by *Hilbert's syzygy theorem* (for example, see [8, 2.214]). Under the hypotheses of Theorem 4.1 we can choose $S = K[X_1, \dots, X_n]$ with $\deg X_i = g_i$ and

map it onto $R = K[x_1, \dots, x_n]$ via the substitution $X_i \mapsto x_i$. In conjunction with the syzygy theorem we therefore obtain a new (but more difficult) proof of Theorem 4.1. This is Hilbert's original approach [16] to his theorem that the Hilbert function is of polynomial type (in the classical case $\deg X_i = 1$).

The difference between the Hilbert function and the Hilbert polynomial can be expressed in terms of local cohomology. It would take us too far astray to explain this concept here. See [8].

Theorem 4.8 (Serre). *Under the hypotheses of Theorem 4.4 let* $\mathfrak{m}$ *denote the maximal ideal* $(x_1, \dots, x_n)$ *and* $d = \dim M$. *Then*

$$H(M, i) - P_M(i) = \sum_{j=0}^{d} (-1)^j \dim_K H^j_{\mathfrak{m}}(M)_i$$

for all $i \in \mathbb{Z}$.

Here $H^j_{\mathfrak{m}}(M)$ denotes the jth local cohomology of M with support in $\mathfrak{m}$. The proof follows the same pattern as that of Theorem 4.1, but this time one must work in the field $\mathbb{C}[[t^{-1}]][t]$ after one has shown that the generating function of the right hand belongs to it – for the left hand this has been proved above. We refer the reader to [8, 4.4.3] for the details.

5. Cohen–Macaulay rings

Before giving a definition we examine a simple, but very instructive case. Also in the next section it will guide us to a general theorem.

Simplicial normal monoids

An affine monoid M is *simplicial* if the cone $\mathbb{R}_+M$ is generated by $d = \operatorname{rank} M$ vectors $x_1, \dots, x_d$. Suppose that M is normal, embedded into $\mathbb{Z}^d = \operatorname{gp}(M)$, and let $y \in M$. Then y has a unique representation $y = \sum_{i=1}^d a_i x_i$ with $a_i \in \mathbb{Q}_+$. It follows that

$$y = \sum_{i=1}^{d} b_i x_i + \sum_{i=1}^{d} q_i x_i, \qquad b_1, \dots, b_n \in \mathbb{Z}_+, \quad q_1, \dots, q_n \in [0, 1) \cap \mathbb{Q}.$$

In other words, $y = y' + y''$ where y' belongs to the free submonoid $N = \sum \mathbb{Z}_+ x_i$ and y'' is an integral point in the semi-open parallelotope

$$\operatorname{par}(x_1, \dots, x_d) = \left\{ \sum_{i=1}^{d} q_i x_i : q_i \in [0, 1) \right\}.$$

The vectors y', y'' are uniquely determined by y. Since M is normal, $y'' \in M$. Set

$$B = \operatorname{par}(x_1, \dots, x_d) \cap \mathbb{Z}^d.$$

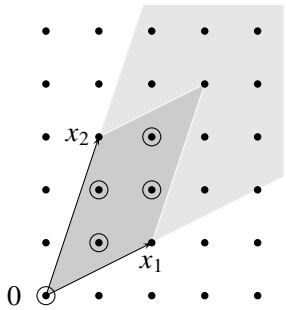

Figure 4. The semi-open parallelotope

If M is an affine monoid with a positive grading, then we can assign a Hilbert function to M directly, namely

$$H(M,k) = \#\{y \in M : \deg y = k\}.$$

Clearly $H(M,k) = H(K[M],k)$ for the induced grading on $K[M]$. Similarly we use the notion of Hilbert series for M.

Proposition 5.1.

(a) *M is the disjoint union of the sets $z + N$, $z \in B$;*
(b) *Let* $\deg : M \to \mathbb{Z}_+$ *be a positive grading on M. Then*

$$H_M(t) = \frac{\sum_{i=0}^{u} h_i t^i}{\prod_{j=1}^{d}(1-t^{g_j})}, \qquad h_i = \#\{z \in B : \deg z = i\} \geq 0, \quad i = 0, \dots, u.$$

(c) $\deg H_M < 0$, *and* $P_M(i) = H_M(i)$ *for all* $i \geq 0$.

Proof. (a) has been proved above. For (b) it is now enough that $H_N(t) = 1/\prod_{j=1}^{d}(1-t^{g_j})$. (This follows from equation (5), but can also be proved directly.) For (c) we note that every element in B has degree $< \sum g_j$ so that $u < \sum g_j$. The last assertion follows from Theorem 4.4. □

Proposition 5.1 contains statements of type (ADG-2) and (ADG-4), but it cannot be applied directly since the monoid of magic squares is simplicial only for $n = 1, 2$.

In commutative algebra language 5.1(a) reads as follows: the monoid algebra $S = K[N]$ is a Noether normalization of $K[M]$, and $R = K[M]$ *is a free module over its Noether normalization S*. In fact, S is generated by d elements, $d = \dim R$. Furthermore the finitely many monomials $z \in B$ generate R as a S-module. Their linear independence follows easily from the disjointness of the decomposition in 5.1(a).

Cohen–Macaulay rings

Graded algebras or modules that are free over a Noether normalization deserve (and have) a special name. Its classical and usual definition uses the notion of regular sequence:

Definition 5.2. Let R be a commutative ring and M an R-module. We say that $x_1, \dots, x_m \in R$ form an *M-sequence* or are *M-regular* if

(i) $M/(x_1, \dots, x_m)M \neq 0$;
(ii) x_i is a non-zero-divisor on $M/(x_1, \dots, x_{i-1})M$ for $i = 1, \dots, m$.

Condition (ii) is equivalent to the injectivity of the multiplication by x_i on $M/(x_1, \dots, x_{i-1})M$. Therefore the property of being an M-sequence behaves well under flat ring extensions, like passage to a ring of fractions or to $R \otimes_K L$ if R contains a field K and L is an extension field of K.

The next proposition connects the notion of M-sequence with the freeness of M over a Noether normalization. Moreover, it shows that this property does not depend on the choice of normalization:

Proposition 5.3. *Let R be a positively graded finitely generated K-algebra and M a finitely generated graded R-module. Furthermore, let $x_1, \dots, x_d$ and $y_1, \dots, y_d$ be homogeneous systems of parameters for M. Then the following are equivalent:*

(a) *M is a free module over $S = K[x_1, \dots, x_d]$;*
(b) *$x_1, \dots, x_d$ is an M-sequence;*
(c) *$y_1, \dots, y_d$ is an M-sequence.*

Proof. Both (a) and (b) concern only the structure of M as an S-module. In proving their equivalence we may therefore assume that $R = S = K[X_1, \dots, X_d]$. The implication (a) $\Rightarrow$ (b) is then obvious: M is a direct sum of copies of R, and (b) is trivial for $M = R$.

For the converse we choose a minimal homogeneous system $z_1, \dots, z_m$ of generators of M. "Minimal" means that none of $z_1, \dots, z_m$ can be omitted. We want to show that $z_1, \dots, z_m$ are linearly independent. Choose a presentation

$$F \xrightarrow{\varphi} M \to 0$$

mapping the i-th basis element e_i of the graded free module F to z_i. Our claim amounts to $\operatorname{Ker} \varphi = 0$.

On the contrary, let us assume that $z_1, \dots, z_m$ are linearly dependent. Then there exist $a_1, \dots, a_m \in R$, not all 0, such that $a_1 z_1 + \cdots + a_m z_m = 0$. Since $z_1, \dots, z_m$ are homogeneous, we can split this equation into its homogeneous components. Therefore we may assume that the a_i are homogeneous and that we have chosen a non-trivial homogeneous relation of the smallest possible degree.

None of the $a_i \neq 0$ can be in K. Otherwise we could write z_i as a linear combination of the z_j, $j \neq i$. In other words, $\operatorname{Ker} \varphi \subset \mathfrak{m}F$, where $\mathfrak{m} = (X_1, \dots, X_d)$. This implies that the residue classes of $z_1, \dots, z_m$ form a basis of the K-vector space $M/\mathfrak{m}M$.

For $d = 0$, $R = K$, there is nothing to prove. So let $d > 0$. Set $\bar{R} = R/(X_1)$, $\bar{M} = M/X_1 M$. Taking intermediate residue classes with respect to (X_1) does not change the fact that the residue classes of $\bar{z}_1, \dots, \bar{z}_m \in \bar{M}$ form a basis of $M/\mathfrak{m}M$. Moreover $X_2, \dots, X_d$ form an $\bar{M}$-sequence. By the induction hypothesis $\bar{z}_1, \dots, \bar{z}_m$ are linearly independent over $\bar{R}$.

For our relation $a_1 z_1 + \cdots + a_m z_m = 0$ this implies $a_i \in (X_1)$, $a_i = b_i X_1$, $i = 1, \dots, m$. Since X_1 is a non-zero-divisor on M, we can factor out X_1 from the equation

$a_1z_1 + \cdots + a_mz_m = 0$, and obtain a contradiction to the minimality of $\deg a_i$. This completes the proof of (b) $\Rightarrow$ (a).

We postpone the somewhat technical proof of the equivalence of (b) and (c) to the end of this section. □

Definition 5.4. Let R be an affine, positively graded K-algebra, and M a finitely generated graded R-module. Then M is *Cohen–Macaulay* if it admits an M-regular homogeneous system of parameters. If R itself is a Cohen–Macaulay module, then it is called a *Cohen–Macaulay ring*.

As stated in Proposition 5.3 (but not yet proved) it follows that every homogeneous system of parameters of a graded Cohen–Macaulay module M is M-regular.

The main combinatorial consequence of the Cohen–Macaulay property is the non-negativity (or even positivity) of the h-vector.

Theorem 5.5. *Let R be a positively graded affine K-algebra, M a finitely generated graded Cohen–Macaulay R-module, and let $g_1, \dots, g_d$ be the degrees of the elements in a homogeneous system $x_1, \dots, x_d$ of parameters for M. Let*

$$H_M(t) = \frac{\sum_{i=a}^{b} h_i t^i}{\prod_{i=1}^{d}(1 - t^{g_i})}$$

be the Hilbert series of M. Then $h_i \geq 0$ for all i.

If R is a Cohen–Macaulay ring generated in degree 1 *and*

$$H_R(t) = \frac{\sum_{i=0}^{u} h_i t^i}{(1-t)^d}, \qquad h_u \neq 0,$$

then $h_i > 0$ for $i = 0, \dots, u$.

Proof. The first assertion follows as in the special case of a simplicial normal monoid algebra considered above: h_i counts the number of degree i basis elements of M over the Noether normalization.

Alternatively we could argue as follows: $\sum_{i=a}^{b} h_i t^i$ is the Hilbert series of $M/(x_1, \dots, x_d)M$. This follows immediately by induction on d (compare the proof of Theorem 4.1).

For the second assertion we first extend the field in order to obtain a degree 1 homogeneous system $x_1, \dots, x_d$ of parameters. The K-algebra $\bar{R} = R/(x_1, \dots, x_d)$ is generated in degree 1 and $\sum_{i=0}^{u} h_i t^i$ is its Hilbert series. If $\bar{R}_i = 0$, then $\bar{R}_j = 0$ for all $i \geq j$. □

Remark 5.6. In the second part of Theorem 5.5 we have used the fact that $\bar{R} = K[\bar{R}_1]$ only in a very weak form. Which h-vectors $(h_0, \dots, h_u)$ can occur for such algebras is determined by *Macaulay's theorem*; see [8, 4.2.15].

In order to apply Theorem 5.5 to $\mathcal{M}_n$ or other normal affine monoids we must prove

Theorem 5.7 (Hochster). *Let M be a normal affine monoid (with a positive grading). Then $K[M]$ is Cohen–Macaulay for every field K.*

The Hochster–Roberts theorem

We want to derive Hochster's theorem from a more general result:

Theorem 5.8 (Hochster–Roberts). *Let K be a field and let S be a an affine graded K-subalgebra of a polynomial ring $R = K[X_1, \dots, X_n]$. If there exists an S-submodule T of R such that $R = S \oplus T$, then S is Cohen–Macaulay.*

Theorem 5.8, in conjunction with Proposition 1.5, immediately implies Hochster's theorem. The theorem of Hochster–Roberts is proved by reduction to characteristic p. However, in general the hypothesis $R = S \oplus T$ in the theorem would not survive the reduction. Therefore one has to prove a more general statement.

Theorem 5.9. *Let K be a field, and let $f_1, \dots, f_s$ be algebraically independent homogeneous elements of positive degree in $R = K[X_1, \dots, X_n]$. Suppose that S is a module-finite graded $K[f_1, \dots, f_s]$-algebra such that there exists a homogeneous homomorphism $\psi : S \to R$ of $K[f_1, \dots, f_s]$-algebras. If $g_{r+1} f_{r+1} = g_1 f_1 + \cdots + g_r f_r$ with $g_1, \dots, g_{r+1} \in S$ for some r, $0 \le r \le s-1$, then $\psi(g_{r+1}) \in (f_1, \dots, f_r)R$.*

For the derivation of the Hochster–Roberts theorem we choose $f_1, \dots, f_s$ as a homogeneous system of parameters for S and $\psi : S \to R$ as the given embedding. We want to show that $f_1, \dots, f_s$ is S-regular. If it should fail, then we have an equation $g_{r+1} f_{r+1} = g_1 f_1 + \cdots + g_r f_r$ with $g_{r+1} \notin (f_1, \dots, f_r)S$ for some r. On the other hand, Theorem 5.9 shows that $g_{r+1} \in (f_1, \dots, f_r)R$, and this is a contradiction: since $R = S \oplus T$, one has $IR \cap S = I$ for every ideal I of S.

Proof of Theorem 5.9. It needs two steps:

(i) the case of positive characteristic;
(ii) the reduction from characteristic 0 to positive characteristic.

Under the general hypothesis of the theorem step (ii) is technically demanding (see [8, p. 294]). Therefore we restrict ourselves to step (i), and do step (ii) only for algebras that, roughly speaking, can be defined over the ring $\mathbb{Z}$ of integers; see Proposition 5.11 below.

So let K be a field of characteristic $p > 0$. We can replace K by its algebraic closure L. The hypothesis certainly survives the base field extension to L, and so does the conclusion since $IL[X_1, \dots, X_n] \cap K[X_1, \dots, X_n] = I$ for every ideal I of $K[X_1, \dots, X_n]$. We may therefore assume that $K = K^p$.

There exist a finitely generated free $K[f_1, \dots, f_s]$-submodule $F \subset S$ and a non-zero element $c \in K[f_1, \dots, f_s]$ such that $cS \subset F$ (just because $K[f_1, \dots, f_s]$ is a domain and S is finitely generated).

Let $q = p^e$ be a power of the characteristic p, take the q-th power of the equation $g_{r+1} f_{r+1} = g_1 f_1 + \cdots + g_r f_r$ and multiply by c to obtain

$$(cg_{r+1}^q) f_{r+1}^q = \sum_{i=1}^{r} (cg_i^q) f_i^q.$$

The elements cg_i^q, $i = 1, \dots, r+1$, are in the free $K[f_1, \dots, f_s]$-module F, and in $K[f_1, \dots, f_s]$ the elements $f_1, \dots, f_s$ behave like indeterminates. Thus an elementary

argument yields the existence of $h_{iq} \in F$ with $cg_i^q = h_{iq} f_{r+1}^q$ for $i = 1, \dots, r$. By substituting these expressions into the previous equation and applying $\psi : S \to K[X_1, \dots, X_n]$ one has

$$cf_{r+1}^q \psi(g_{r+1})^q = \sum_{i=1}^r f_i^q f_{r+1}^q \psi(h_{iq}), \quad \text{hence} \quad c\psi(g_{r+1})^q = \sum_{i=1}^r f_i^q \psi(h_{iq}).$$

Let M be the set of monomials $\mu = X_1^{\mu_1} \cdots X_1^{\mu_n}$ with $\mu_i < q$ for $i = 1, \dots, n$. Taking q-th powers in the algebraically closed field K is bijective. Therefore every element $h \in K[X_1, \dots, X_n]$ has a necessarily *unique* representation $h = \sum_{\mu \in M} (h_\mu)^q \mu$; in particular,

$$\psi(h_{iq}) = \sum_{\mu \in M} (h_{iq\mu})^q \mu.$$

Thus

$$\sum_{i=1}^r f_i^q \psi(h_{iq}) = \sum_{\mu \in M} \left(\sum_{i=1}^r h_{iq\mu} f_i \right)^q \mu = \sum_{\mu \in M} (h_{q\mu})^q \mu$$

with $h_{q\mu} \in (f_1, \dots, f_r)R$.

The crucial point is that c does not depend on q. We choose q so large that $c = \sum_{\mu \in M} c'_\mu \mu$ with $c'_\mu \in K$. Let $c'_\mu = (c_\mu)^q$. Then

$$\sum_{\mu \in M} (c_\mu \psi(g_{r+1}))^q \mu = \sum_{\mu \in M} (h_{q\mu})^q \mu.$$

Since $c \neq 0$ there exists μ with $c_\mu \neq 0$, and so

$$\psi(g_{r+1}) = \frac{1}{c_\mu} h_{q\mu} \in (f_1, \dots, f_r)R.$$

□

Remark 5.10.

(a) This proof of the Hochster–Roberts theorem is due to F. Knop. It is a condensation of a tight closure proof by Hochster and Huneke (see [8, 10.1.14]).

(b) The Hochster–Roberts theorem can be strengthened. In characteristic 0 direct summands R of polynomial rings (or the affine varieties defined by them) have rational singularities by a theorem of Boutot [5]; also see Gurjar [15]. In characteristic p Hochster and Huneke showed that R is F-regular; see [8, 10.1.3]. These properties include that R is Cohen–Macaulay.

(c) A general Cohen–Macaulay criterion for affine monoid rings was established by Trung and Hoa [28].

Now Stanley's conjecture (ADG-4) has been completely proved. In fact, for the combinatorial application we are free to choose the coefficient field K, and so we take an infinite field of arbitrary positive characteristic. Moreover, we have shown that every homogeneous system of parameters in a direct summand of a polynomial ring over such a field, for example in $K[\mathcal{M}_n]$, is a regular sequence.

Free $\mathbb{Z}$-algebras

Let M be an affine monoid. Then we can define the monoid algebra $\mathbb{Z}[M]$ in the same way as the monoid algebra $K[M]$ where K is a field: we choose a free $\mathbb{Z}$-module with a basis corresponding to the elements of M, and use the monoid operation for the multiplication table. Then $K[M] = \mathbb{Z}[M] \otimes_{\mathbb{Z}} K$ for every field K.

Proposition 5.11. *Let R be a positively graded, finitely generated and torsionfree $\mathbb{Z}$-algebra. If $R \otimes \mathbb{Z}/(p)$ is Cohen–Macaulay for at least one prime number p, then $R \otimes K$ is Cohen–Macaulay for every field K of characteristic* 0.

Proof. Note that $\bar{R} = R \otimes \mathbb{Z}/(p)$ is a positively graded affine algebra over the field $\mathbb{Z}/(p)$. We choose a homogeneous system of parameters $\bar{x}_1, \ldots, \bar{x}_d$ and homogeneous representatives $x_1, \ldots, x_d$ in R. Moreover we choose a finite homogeneous system of generators $\bar{y}_1, \ldots, \bar{y}_m$ (of positive degree) of $\bar{R}$ over $\mathbb{Z}/(p)[\bar{x}_1, \ldots, \bar{x}_d]$ and homogeneous representatives $y_1, \ldots, y_m \in R$.

The bridge from characteristic p to characteristic 0 is the ring of fractions $R_{(p)} = R \otimes \mathbb{Z}_{(p)}$ where $\mathbb{Z}_{(p)}$ is the ring of all rational numbers whose denominators are not divisible by p – it is just the localization of $\mathbb{Z}$ with respect to the prime ideal (p). We claim that $R_{(p)}$ is generated over its subalgebra $\mathbb{Z}_{(p)}[x_1, \ldots, x_d]$ by $y_1, \ldots, y_m$. In fact, consider a single graded component $(R_i) \otimes \mathbb{Z}_{(p)}$. It is a finitely generated $\mathbb{Z}_{(p)}$-module, since R_i is finitely generated over $\mathbb{Z}$. Let N be the submodule of $(R_{(p)})_i$ generated by the elements of type μy_j where μ is a monomial in $x_1, \ldots, x_d$. By the choice of the x_i and y_j we have $(R_{(p)})_i = N + p(R_{(p)})_i$, since $(R_{(p)})_i / p(R_{(p)})_i = R_i / pR_i$. By the elementary divisor theorem, applied to the free module $(R_{(p)})_i$ and its submodule N over the Euclidean domain $\mathbb{Z}_{(p)}$ (or Nakayama's lemma), it follows that $(R_{(p)})_i = N$, as desired.

Before we discuss the Cohen–Macaulay property of $R \otimes K$, let us observe that $x_1, \ldots, x_d$ is a homogeneous system of parameters for $R \otimes \mathbb{Q}$. The inversion of p leads us from $R_{(p)}$ to $R_{(0)} = R \otimes \mathbb{Q}$, and therefore $R \otimes \mathbb{Q}$ is a finitely generated module over $\mathbb{Q}[x_1, \ldots, x_d]$. But $R \otimes \mathbb{Q}$ has the same Krull dimension as $R \otimes \mathbb{Z}/(p)$. For example, this can be concluded from the coincidence of their Hilbert functions:

$$H(R \otimes \mathbb{Z}/(p), i) = \operatorname{rank}_{\mathbb{Z}} R_i = H(R \otimes \mathbb{Q}, i).$$

By hypothesis $\bar{x}_1, \ldots, \bar{x}_d$ is an $\bar{R}$-sequence. Moreover p is a non-zero-divisor on R since R is torsionfree. Therefore $p, x_1, \ldots, x_d$ is an R-sequence. This property is not lost if we pass to the ring of fractions $R_{(p)} = R \otimes \mathbb{Z}_{(p)}$. Now we can conclude that

(i) $x_1, \ldots, x_d$ are algebraically independent over $\mathbb{Z}_{(p)}$ and
(ii) $R_{(p)}$ is a free module over $\mathbb{Z}_{(p)}[x_1, \ldots, x_d]$.

Claim (i) holds since $x_1, \ldots, x_d$ are algebraically independent over $\mathbb{Q}$, and claim (ii) can be proved by the same induction as the implication (b) $\Rightarrow$ (a) of 5.3. For the case $d = 0$ one uses that $\mathbb{Z}_{(p)}$ is a discrete valuation domain over which a finitely generated module is free if p is a non-zero-divisor on it.

Now it is clear that we can permute $p, x_1, \ldots, x_d$ without loosing the property that it is an $R_{(p)}$-sequence. By inversion of p it follows that $x_1, \ldots, x_d$ is an $R \otimes \mathbb{Q}$-sequence.

The passage from $\mathbb{Q}$ to an arbitrary field of characteristic 0 is harmless, as observed immediately after Definition 5.2. □

Homogeneous systems of parameters in Cohen–Macaulay rings

It remains to show the equivalence of (b) and (c) in Proposition 5.3. We can assume that $\operatorname{Ann} M = 0$. There is nothing to show in dimension $d = 0$, but $d = 1$ is the critical case. Let x_1 and y_1 be homogeneous system of parameters for M. If x_1 is a non-zero-divisor, then all zero-divisors of M must be contained in minimal prime ideals of R. On the other hand, y_1 cannot be contained in such a prime ideal, and therefore is a non-zero-divisor.

The equivalence 5.3 (a) $\Longleftrightarrow$ (b) shows that we can permute a homogeneous system of parameters without destroying the property of being an M-sequence. This will now be used.

For $d > 1$ we argue by induction. Suppose that $x_1, \dots, x_d$ is a homogeneous system of parameters and an M-sequence. We can replace the x_i by suitable powers to obtain elements of the same degree: $K[x_1, \dots, x_d]$ is a free module over $S = K[x_1^{e_1}, \dots, x_d^{e_d}]$, and so M is free over S. We can assume that K is infinite.

The residue class ring $\bar{R} = R/(y_1, \dots, y_{d-1})$ has dimension 1. It is not possible that the residue classes of $x_1, \dots, x_d$ are all contained in a single minimal prime ideal of $\bar{R}$. By the prime avoidance lemma 3.2 we therefore find a linear combination $x' = a_1x_1 + \cdots + a_dx_d$ whose residue class is not contained in such a prime ideal. Equivalently, $y_1, \dots, y_{d-1}, x'$ is a homogeneous system of parameters.

Permuting the x_i if necessary we can assume that $a_1 \neq 0$. As follows from 5.3 (a) $\Rightarrow$ (b), also $x', x_2, \dots, x_d$ is an M-sequence. So $x_2, \dots, x_d$ is an $(M/x'M)$-sequence. By induction $y_1, \dots, y_{d-1}$ is an $(M/x'M)$-sequence, too. Thus $x', y_1, \dots, y_{d-1}$ is an M-sequence. Finally, a similar argument allows us to replace x' by y_d (here it is used that $d > 1$).

6. The canonical module and Gorenstein rings

A functional equation for the Hilbert series

Let $R = K[\mathcal{M}_n]$ and $d = \dim R = (n-1)^2 + 1$. The equation $P_n(-k) = (-1)^{d-1}P_n(k-n)$ in (ADG-3) (together with the not yet proved fact that $H(R, r) = P_n(r)$ for all $r \geq 0$) encodes a functional equation between the rational functions $H_R(t)$ and $H_R(t^{-1})$ where the latter is obtained by the substitution of t^{-1} for t. The substitution $t \mapsto t^{-1}$ transforms the Laurent expansion of $H_R(t)$ at 0 into the Laurent expansion of $H_R(t^{-1})$ at ∞, but what we really want to compare are the Laurent expansions at 0:

Lemma 6.1. *Let $P : \mathbb{Z} \to \mathbb{C}$ be a quasi-polynomial. Set $H(t) = \sum_{k=0}^{\infty} P(k)t^k$ and $G(t) = -\sum_{k=1}^{\infty} P(-k)t^k$. Then H and G are rational functions. Moreover*

$$H(t) = G(t^{-1}).$$

Proof. We have to exercise some care and need precise notation for expansions. If R is a rational function, then we let $\mathcal{L}_0(R)$ and $\mathcal{L}_\infty(R)$ denote the Laurent expansions at 0 and ∞ resp. One obtains $\mathcal{L}_\infty(R)$ as follows: in the Laurent expansion of $R(t^{-1})$ at 0 one replaces t by t^{-1}.

Both $\mathcal{L}_0(R)$ and $\mathcal{L}_\infty(R)$ are elements of the abelian group $\mathbb{C}[\![t, t^{-1}]\!]$ of formal Laurent series $\sum_{k\in\mathbb{Z}} a_k t^k$. Note that the multiplication of elements of $\mathbb{C}[\![t, t^{-1}]\!]$ by Laurent polynomials is well-defined. It makes $\mathbb{C}[\![t, t^{-1}]\!]$ a module over $\mathbb{C}[t, t^{-1}]$. Moreover, the maps $\mathcal{L}_0$ and $\mathcal{L}_\infty$ from the field of rational functions to $\mathbb{C}[\![t, t^{-1}]\!]$ are both injective and $\mathbb{C}[t, t^{-1}]$-linear.

Set $J = G(t^{-1})$, $Q = (1 - t^e)^{d+1}$ and $F = \mathcal{L}_0(H) - \mathcal{L}_\infty(J)$. Since $\mathcal{L}_\infty(J) = -\sum_{k=1}^{\infty} P(-k)t^{-k}$, we have $F = \sum_{k\in\mathbb{Z}} P(k)t^k$. Let d be the degree of P and e its period. Then

$$QF = (1 - t^e)^{d+1} F = 0.$$

In fact, $(1 - t^e)F = \sum_{k\in\mathbb{Z}} \tilde{P}(k)t^k$ with the quasi-polynomial $\tilde{P}$ whose components $\tilde{P}_i$, $i = 1, \ldots, e$, are given by $\tilde{P}_i(j) = P_i(j) - P_i(j-1)$, $j \in \mathbb{Z}$. Thus $\tilde{P}$ has lower degree than P, and the claim follows by induction on d.

By $\mathbb{C}[t, t^{-1}]$-linearity one has

$$Q\mathcal{L}_0(H) - Q\mathcal{L}_\infty(J) = Q(\mathcal{L}_0(H) - \mathcal{L}_\infty(J)) = QF = 0.$$

The k-th coefficient of $Q\mathcal{L}_0(H)$ vanishes for $k < 0$, and the ℓ-th coefficient of $Q\mathcal{L}_\infty(J)$ vanishes for $\ell > (d+1)e$. So the equation can only hold if $Q\mathcal{L}_0(H)$ and $Q\mathcal{L}_\infty(J)$ are the same Laurent polynomial K.

By the injectivity and $\mathbb{C}[t, t^{-1}]$-linearity of the assignment $R \mapsto \mathcal{L}_0(R)$ we conclude that $H = K/Q$, and similarly $G(t^{-1}) = J = K/Q$. □

Remark 6.2. We have not really used the hypothesis that P is a quasi-polynomial. It is enough that F is annihilated by a non-zero polynomial or, equivalently, that the values $P(k)$ satisfy a linear difference equation with constant coefficients.

We know from Theorem 4.2 that (ADG-2) follows if the Hilbert series of $K[\mathcal{M}_n]$ has degree $-n$. If this should indeed be true, then (ADG-3) is equivalent to the validity of a functional equation for the Hilbert series:

Corollary 6.3. Let R be a positively graded, finitely generated K-algebra of Krull dimension d, and let P be its Hilbert quasi-polynomial. Suppose that the Hilbert series $H_R(t)$ has degree $g < 0$. Then the following are equivalent:

(a) $P(-k) = (-1)^{d-1} P(k+g)$ for all $k \in \mathbb{Z}$;
(b) $(-1)^d H_R(t^{-1}) = t^{-g} H_R(t)$.

This follows from Lemma 6.1 by comparison of coefficients if one takes into account that $H(R, k) = P(k)$ for all $k \geq 0$ and $P(-k) = 0$ for $k = 1, \ldots, g-1$ (see Theorem 4.4).

The canonical module and Gorenstein rings

The approach to the proof of (ADG-2) and (ADG-3) that we will take consists of two major steps:

(i) the construction of a graded R-module $\omega = \omega_R$ for a *Cohen–Macaulay* positively graded K-algebra R such that $H_\omega(t) = (-1)^d H_R(t^{-1})$, and
(ii) the exact computation of ω_R for normal affine monoid algebras, and in particular for $K[\mathcal{M}_n]$.

Step (i) is not difficult. Let us first consider the case in which $R = K[X_1, \dots, X_d]$ is the polynomial ring over K with grading given by $\deg X_i = g_i$. Set $e = -\sum_{i=1}^d g_i$ and $\omega = R(e)$. Then

$$H_\omega(t) = t^{-e} H_R(t) = \frac{\prod_{i=1}^d t^{g_i}}{\prod_{i=1}^d (1 - t^{g_i})} = (-1)^d H_R(t^{-1}). \tag{6}$$

If R is a Cohen–Macaulay positively graded K-algebra with graded Noether normalization $S = K[x_1, \dots, x_d] \cong K[X_1, \dots, X_d]$, we set

$$\omega_R = \operatorname{Hom}_S(R, \omega_S).$$

The module ω_R (whose construction so far may depend on the choice of S) has indeed the structure of an R-module defined by $(a\varphi)(b) = \varphi(ab)$ for all $a, b \in R$.

For the computation of its Hilbert series we only need the S-structure. Every direct summand $S(-j)$ of the free S-module R accounts for a summand $\operatorname{Hom}_S(S(-j), S(e)) \cong S(e + j)$ of ω_R. Let

$$H_R(t) = \frac{\sum_{j=0}^u h_j t^j}{\prod_{i=1}^d (1 - t^{g_i})}.$$

Then $R = \bigoplus_j S(-j)^{h_j}$, and so

$$H_{\omega_R}(t) = \frac{\sum_{j=0}^u h_j t^{-e-j}}{\prod_{i=1}^d (1 - t^{g_i})} = \frac{t^{-e}}{\prod_{i=1}^d (1 - t^{g_i})} \sum_{j=0}^u h_j t^{-j} = (-1)^d H_R(t^{-1}). \tag{7}$$

We have reached our goal: ω_R has the right Hilbert series, independently of the choice of the Noether normalization used in its construction. We supplement the combinatorial information on ω_R by

Proposition 6.4. *Let R be a positively graded Cohen–Macaulay K-algebra of Krull dimension d with Hilbert quasi-polynomial P. Then*

(a) $\deg H_R(t) = -\min\{i : (\omega_R)_i \neq 0\}$;
(b) $k \mapsto (-1)^{d-1} P(-k)$ *is the Hilbert quasi-polynomial of ω_R, and*

$$H(\omega_R, k) = (-1)^{d-1} P(-k) \qquad \text{for all } k \geq 1.$$

Part (a) follows from equation (7) since the lowest degree appearing in the numerator polynomial of $H_\omega(t)$ is $-\deg H_R(t)$. For part (b) one writes

$$H_R(t) = \sum_{k=0}^{\infty} P(k) t^k + \sum_{k=0}^{\deg H_R(t)} (H(R, k) - P(k)) t^k \tag{8}$$

and applies Lemma 6.1.

As our notation ω_R, in which the Noether normalization does not appear, suggests, this module is uniquely determined by R (up to homogeneous isomorphism).

Definition 6.5. The module ω_R is called the *canonical module* of R. One says that R is a *Gorenstein ring* if ω_R is isomorphic to $R(h)$ for some $h \in \mathbb{Z}$.

Before we indicate why ω_R is uniquely determined, we describe the combinatorial consequences of the Gorenstein property.

Theorem 6.6 (Stanley). *Let R be a positively graded, finitely generated K-algebra of Krull dimension d, let $H_R(t) = (h_0 + \cdots + h_u t^u)/\prod_{i=1}^{d}(1 - t^{g_i})$ with $h_u \neq 0$ be the Hilbert series of R, and P the Hilbert quasi-polynomial. Suppose that R is a Gorenstein ring. Then*

(a) $\omega_R \cong R(g)$, $g = \deg H_R(t)$;
(b) $h_i = h_{u-i}$ *for* $i = 0, \ldots, u$: *the h-vector is palindromic;*
(c) $H_R(t^{-1}) = (-1)^d t^{-g} H_R(t)$;
(d) $P(-k) = (-1)^{d-1} P(k + g)$ *for all* $k \in \mathbb{Z}$.

Conversely, if R is a Cohen–Macaulay integral domain such that

$$H_R(t^{-1}) = (-1)^d t^{-h} H_R(t)$$

for some $h \in \mathbb{Z}$, then R is Gorenstein.

Proof. Let R be a Gorenstein ring, say $\omega_R \cong R(h)$. Then we have $H_{\omega_R}(t) = t^{-h} H_R(t)$. Thus

$$(-1)^d H_R(t^{-1}) = t^{-h} H_R(t).$$

In particular the numerator polynomials on both sides must be equal:

$$t^{-g} \sum_{j=0}^{u} h_j t^{u-j} = t^{-h} \sum_{j=0}^{u} h_j t^j$$

This is only possible if $h = g$ and the h-vector is palindromic. This proves (a), (b), and (c). The equation in (d) follows from Proposition 6.4 by comparison of coefficients for $k \gg 0$. (However, if $g \geq 0$, (d) does not imply (c).)

For the converse we choose a non-zero element x of degree $-h$ in ω_R. The linear map $\xi : a \mapsto ax$ from R to ω_R is injective, since ω_R is torsionfree and R is an integral domain. Comparing the Hilbert series of the image with that of ω_R, we conclude that ξ is surjective, too. □

The proof that ω_R is unique up to homogeneous isomorphism requires the use of derived functors. First one shows that

$$\operatorname{Ext}_R^j(K, \omega_R) = \begin{cases} K, & j = d, \\ 0, & j \neq d. \end{cases} \tag{9}$$

This is to be understood as a relation between graded modules: $K \cong R/\mathfrak{m}$ (with $\mathfrak{m} = \bigoplus_{k=1}^{\infty} R_k$) lives in degree 0. Equation (9) implies that the localization $(\omega_R)_{\mathfrak{m}}$

is a canonical module of the local ring $R_{\mathfrak{m}}$. As such it is uniquely determined up to $R_{\mathfrak{m}}$-isomorphism. Finitely generated graded R-modules that become isomorphic after the passage to $R_{\mathfrak{m}}$ are isomorphic up to a shift. Thus graded canonical modules ω_R and ω'_R are isomorphic up to a shift. A non-zero shift is ruled out, since it would show up in $\operatorname{Ext}^d_R(K, \omega_R)$ (or in the Hilbert series). See [8, 3.6.9] for more details.

Equation (9) is proved by induction on d. First let $d = 0$. Then K itself is the Noether normalization, and $\omega_R = R^\vee = \operatorname{Hom}_K(R, K)$. Thus (9) follows from

(i) the exactness of the R-linear functor $V \mapsto V^\vee$ assigning each finitely generated graded R-module M its graded K-dual, and
(ii) the natural isomorphism $M = (M^\vee)^\vee$ if $\dim_K M < \infty$.

The functor $M \mapsto M^\vee$ is defined on the category of graded K-vector spaces: $\bigoplus M_i \mapsto \bigoplus N_i$ with $N_i = \operatorname{Hom}_K(M_{-i}, K)$, and (i) and (ii) follow from the general properties of graded K-duality. For (ii) it is enough that $\dim_K M_i < \infty$ for all i.

The general case is reduced inductively to the case $d = 0$. Let $x_1, \dots, x_d$ be a homogeneous system of parameters as usual, and $S = K[x_1, \dots, x_d]$. We set $\bar{R} = R/x_dR$ and $\bar{S} = S/x_dS$. By construction $\omega_{\bar{S}} = (\omega_S/x_d\omega_S)(\deg x_d)$, and it follows that

$$\omega_{\bar{R}} = (\omega_R/x_d\omega_R)(\deg x_d).$$

Since x_d annihilates K and is a non-zero-divisor on ω_R, the graded Rees lemma (see [8, 3.1.16 and 4.4.20]) yields

$$\operatorname{Ext}^i_{\bar{R}}(K, \omega_R/x_d\omega_R) = \operatorname{Ext}^{i+1}_R(K, \omega_R)(-\deg x_d) \qquad i \in \mathbb{Z},$$

and the two opposite shifts cancel each other in the passage from R to $\bar{R}$.

Remark 6.7.

(a) If we evaluate equation (8) beyond Proposition 6.4(b), then we obtain

$$H(\omega_R, k) = (-1)^d(H(R, -k) - P(-k)), \qquad k \in \mathbb{Z}.$$

The right hand side has appeared before, namely in Theorem 4.8. Since R is Cohen–Macaulay, the local cohomology modules $H^i_{\mathfrak{m}}(R)$ vanish for $i < d$, and so $H(\omega_R, k) = H(H^d_{\mathfrak{m}}(R), -k)$ for all k. The equality of the Hilbert functions shows that ω_R is the graded K-dual of $H^d_{\mathfrak{m}}(R)$, at least as a K-vector space.

This statement can be either viewed as part of the graded local duality theorem [8, 3.6.19] or used for an intrinsic definition of ω_R.

(b) Once it is known that ω_R is the graded K-dual of $H^d_{\mathfrak{m}}(R)$, equation (8) follows immediately from Theorem 4.8.

Simplicial monoids revisited

Now it has become clear that (ADG-3) is equivalent to the Gorenstein property of $K[\mathcal{M}_n]$. In order to prove it, we have to determine the canonical module. We will solve this task for every normal affine monoid ring. As in the previous section, the simplicial normal monoids lead us to the general theorem.

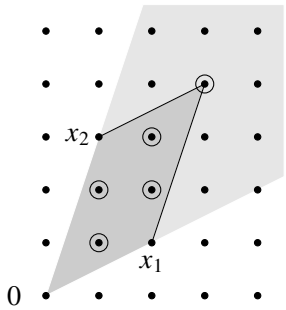

Figure 5. The basis of the interior ideal

Proposition 6.8. *Let M be a simplicial normal affine monoid, and let I be the ideal generated by the monomials in the interior of $\mathbb{R}_+M$. Then I is the graded canonical module of R.*

Proof. We choose $x_i \in M$, $i = 1, \dots, d$, as in the proof of Proposition 5.1, namely as a linearly independent system of generators of $\mathbb{R}_+M$. They form a homogeneous system of parameters, generating the Noether normalization S of R. As a realization of the module $\omega_S = S(-\sum \deg y_i)$ we choose the S-submodule Sx, $x = x_1 \cdots x_d$, of R.

Let $z_1, \dots, z_m$ be the monomial basis of the S-module R given by the monomials corresponding to the lattice points in $\operatorname{par}(x_1, \dots, x_d)$. Then $z_i' = x/z_i$ is a monomial in the interior of $\mathbb{R}_+M$, and, as is easily checked, the z_i' form a basis of I as an S-module. Now we identify $\operatorname{Hom}_S(R, \omega_S)$ with the S-module I as follows: $z_i'(z_j) = x$ if $i = j$, and $= 0$ otherwise.

One must of course check that $\operatorname{Hom}_S(R, \omega_S)$ and I are isomorphic as R-modules. This is left to the reader. □

At this point we should note that Proposition 6.8 opens a purely combinatorial doorway to counting the elements in a normal affine monoid. For the elementary notions of combinatorial topology used in the proof of the next proposition we refer the reader to [6] or [29]. By $\operatorname{relint}(F)$ we denote the relative interior of F, that is the interior of F with respect to the topology on its affine hull.

Proposition 6.9. *Let $M \subset \mathbb{Z}^d$ be a normal simplicial affine monoid with a positive grading* deg. *Set $M_k = \{x \in M : \deg x = k\}$. Then there exists a quasi-polynomial P such that*

$$\#M_k = P(k) \qquad \text{for all } k \geq 0.$$

Proof. We triangulate the cone $\mathbb{R}_+M$ into a family of simplicial subcones whose rays pass through elements of M. (The extreme rays of $\mathbb{R}_+M$ suffice.) This induces a disjoint decomposition of $\mathbb{R}_+M \setminus \{0\}$ into the relative interiors of the cones $F \in \mathcal{F}$, $F \neq \{0\}$. Thus, for all $k > 0$,

$$\#M_k = \sum_{F \in \mathcal{F}} \#(\operatorname{relint}(F) \cap M_k).$$

By Proposition 5.1 the number of degree k lattice points in F is given by a quasi-polynomial P_F for all $k \geq 0$; in particular $P_F(0) = 1$. Proposition 6.8 shows that

$$\#(\text{relint}(F) \cap M_k) = (-1)^{\dim F - 1} P_F(-k)$$

for all $k \geq 1$.

Let $H_M(t) = \sum_{k=0}^{\infty} \#M_k t^k$ be the generating function. We have shown that

$$H_M(t) = 1 + \sum_{k=1}^{\infty} \sum_{F \in \mathcal{F}} (-1)^{\dim F - 1} P_F(-k) t^k.$$

With $P(k) = \sum_{F \in \mathcal{F}} (-1)^{\dim F - 1} P_F(-k)$ it remains to verify that $P(0) = 1$.

We have not yet used the convexity of $\mathbb{R}_+ M$! Rewriting the sum $\sum (-1)^{\dim F - 1} P_F(0)$ according to the dimension of the faces we obtain

$$\sum_{F \in \mathcal{F}} (-1)^{\dim F - 1} P_F(-k) = \sum_{i=0}^{d-1} (-1)^i f_i,$$

where f_i counts the number of faces in $\mathcal{F}$ that have dimension $i + 1$. Consider a polytopal cross-section Π of $\mathbb{R}_+ M$. An $(i+1)$-dimensional face $F \in \mathcal{F}$ intersects Π in a simplex of dimension i. Therefore $P(0)$ is the Euler characteristic of Π. But Π is a convex polytope, homotopic to a point, and so $P(0) = 1$. □

For the combinatorial application the algebraic properties of $K[\mathcal{M}_n]$ and its interior ideal are irrelevant. It is enough to show that $H_I(t) = (-1)^d H_R(t^{-1})$. Using more advanced arguments of combinatorics, it is indeed possible to derive this equation from the simplicial case; see [26, 4.6].

The canonical module of a normal affine monoid algebra

Let us return to our main goal, the computation of the canonical module of a normal affine monoid algebra: Proposition 6.8 is a very strong sign that it should always be given by the interior ideal I. In fact, if this turns out to be true, then (ADG-2) and (ADG-3) have been proved simultaneously. A magic square x in the interior of $\mathbb{R}_+ \mathcal{M}_n$ is not contained in any of the coordinate hyperplanes, and we can subtract the square $\mathbf{1}$ with all entries 1 from it: $x - \mathbf{1} \in \mathcal{M}_n$. It follows that the monomial corresponding to $\mathbf{1}$ generates I. It has degree n. So $I \cong K[\mathcal{M}_n](-n)$, the ring $K[\mathcal{M}_n]$ is indeed Gorenstein, and its Hilbert series has degree $-n$.

At this point it seems inevitable to use the monomial structure of $K[M]$ in a deeper way than only through positive gradings. We need a definition of canonical module that respects this structure. So far the grading group of a graded ring has always been the group $\mathbb{Z}$. However, all the notions introduced in Section 2 can be transferred to more general grading groups. For us it is enough to allow the groups (isomorphic to) $\mathbb{Z}^n$ for $n \geq 1$. An affine monoid algebra $R = K[M]$ has a decomposition $\bigoplus_{\mu \in \text{gp}(M)} R_\mu$ by its very definition: $R_\mu = K\mu$ for all $\mu \in M$ and $R_\mu = 0$ for all monomials $\mu \in \text{gp}(M) \setminus M$.

Definition 6.10. Let R be a $\mathbb{Z}^n$-graded Cohen–Macaulay ring such that the homogeneous non-units generate a proper ideal $\mathfrak{p}$ of R. Then $\mathfrak{p}$ is necessarily a prime ideal; we set $d = \dim R_{\mathfrak{p}}$. One says that ω is a *$\mathbb{Z}^n$-graded canonical module* of R if

$$\operatorname{Ext}_R^j(R/\mathfrak{p}, \omega) = \begin{cases} R/\mathfrak{p}, & j = d, \\ 0, & j \neq d. \end{cases} \tag{10}$$

Our discussion of equation (9) and the uniqueness of the $\mathbb{Z}$-graded canonical module shows that the definitions 6.5 and 6.10 are compatible. Moreover, a generalization of the arguments of [8, 3.6.9] (to which we have referred already) shows that the $\mathbb{Z}^n$-graded canonical module is again uniquely determined. In the following we will simply speak of the multigraded canonical module.

In $K[M]$, with its grading by $\operatorname{gp}(M)$, the non-units even generate a maximal ideal so that it is a candidate for the use of Definition 6.10. As we have shown above, $K[M]$ has a multigraded canonical module if M is simplicial. (That equation (10) is satisfied can be shown by varying the positive $\mathbb{Z}$-grading.) However, in the non-simplicial case it is not clear whether the canonical module of R, constructed via a non-monomial homogeneous system of parameters, has a multigraded structure at all. There are various ways to find it; we use a divisorial approach.

The canonical module (with respect to a positive $\mathbb{Z}$-grading) is a Cohen–Macaulay module over R, and therefore a reflexive module: the natural homomorphism

$$\omega \to \operatorname{Hom}_R(\operatorname{Hom}(\omega, R), R)$$

is an isomorphism. Moreover ω has rank 1 as an R-module (since it has the same rank as R over a Noether normalization). Therefore it is isomorphic to a (fractional) divisorial ideal of R. According to equation (4) there exist integers $c_1, \dots, c_s$ such that

$$\omega_R \cong \mathfrak{p}_1^{(c_1)} \cap \cdots \cap \mathfrak{p}_s^{(c_s)}.$$

The prime ideals $\mathfrak{p}_i$ are defined by the support forms of the monoid M; see Section 1. We would like to show that $c_1 = \cdots = c_s = 1$ is a possible choice since the "interior ideal" I is exactly $\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_s$.

The numbers $c_1, \dots, c_s$ are not uniquely determined. Therefore let us first make a choice that respects only the module structure and not yet the grading. Denote the chosen ideal by J. At least we have realized the R-module ω_R as a $\mathbb{Z}^d$-graded module J. It follows that $\operatorname{Ext}_R^d(K, J)$ is a $\mathbb{Z}^d$-graded module, too. On the other hand, it is just K: as a multigraded module we can identify it with Kv where v is a monomial in $\operatorname{gp}(M)$. We are free to replace J by $v^{-1}J$. Then $\operatorname{Ext}_R^d(K, v^{-1}J)$ is indeed K in multidegree 0: this is the right choice for the multigraded canonical module. Now $c_1, \dots, c_s$ are uniquely determined, and for the desired generalization of Proposition 6.8 it remains to show that $c_1 = \cdots = c_s = 1$.

Let F_i be the facet of $\mathbb{R}_+M$ that contains the set G of monomials not in $\mathfrak{p}_i$. We invert all of them and pass to the ring $S_i = R[G^{-1}]$. In this ring the non-units among the monomials generate the prime ideal $\mathfrak{P}_i = \mathfrak{p}_i S_i$. Then $S_i/\mathfrak{P}_i$ is the Laurent polynomial ring $K[U]$ where U is the group of monomials living in the support hyperplane of

$\mathbb{R}_+M$ through F_i. The multigraded canonical module ω_{S_i} is defined by the condition that $\operatorname{Ext}^j_{S_i}(S_i/\mathfrak{P}_i, \omega_{S_i}) = S_i/\mathfrak{P}_i$ for $j = 1$ (the Krull dimension of $(S_i)_{\mathfrak{P}_i}$) and $= 0$ otherwise. The ideal $\mathfrak{P}_i$ is a principal ideal generated by every monomial that has value 1 under the support form σ_i, and it is now easy to see that $\mathfrak{P}_i$ is the unique multigraded canonical module of S_i.

The final step is to show that the multigraded canonical module "localizes": $\omega_{S_i} = \omega_R \otimes S_i$. We omit this rather technical argument; it is indicated in Remark 6.12(a) and will be carried out in [6]. Since $\omega_R \otimes S_i = \mathfrak{P}_i^{c_i}$, we obtain $c_i = 1$ as desired:

Theorem 6.11 (Danilov, Stanley). *Let M be a normal affine monoid, and K a field. Then the ideal I generated by the monomials in the interior of $\mathbb{R}_+M$ is the multigraded canonical module of $K[M]$.*

Remark 6.12.

(a) The only case in which one has a multigraded Noether normalization, namely one generated by monomials, is that of a simplicial affine monoid. Consequently there is no simple way for the construction of the multigraded canonical module. A method that always works is to represent R as the residue class ring of a polynomial ring $P = K[X_1, \dots, X_n]$ over K. The grading of R can be pulled back to P by giving each indeterminate the degree of its residue class. The multigraded canonical module of P is $\omega_P = R(-\sum \deg X_i)$. As a P-module, R has a finite multigraded free resolution, and $\operatorname{Ext}^c_P(R, \omega_P)$, $c = \dim P - \dim R$, is a multigraded canonical module of R (see [8, 3.6.12] for the $\mathbb{Z}$-graded case).

This approach can be used for a proof of the localization of the multigraded canonical module, since it reduces the question to the polynomial ring P.

(b) One can also use local cohomology and graded local duality for the proof of Hochster's theorem and the computation of the canonical module, as in Stanley's original proof [25] or in the version presented in [8]. It uses the topological properties of the face lattice of the cone $\mathbb{R}_+M$.

Danilov [11] exploits the representation of the canonical module as the module of regular differential d-forms, $d = \dim R$. It can be determined by essentially the same method that we have applied for the divisorial approach.

Remark 6.13. Let us consider a general system of homogeneous linear diophantine equations $\mathcal{L}$, say in m variables. We can assume that there exists a strictly positive solution. Then the monoid M of non-negative solution is normal, all the results above apply, and for a suitable degree function on M (ADG-1)–(ADG-4) remain valid, apart from inevitable adjustments:

(i) in general the number $H(\mathcal{L}, k)$ of degree k solutions is given only by a quasi-polynomial Q for all $k \geq 0$;
(ii) one has $Q(-k) = 0$ for all $k = 1, \dots, s-1$, where s is the smallest degree of a strictly positive solution;
(iii) the general form of (ADG-3) is $Q(-r) = (-1)^{d-1}Q^+(k)$ where $Q^+(k)$ counts the number of the degree k strictly positive solutions and d is the rank of M;
(iv) if $(1, \dots, 1)$ is a solution to $\mathcal{L}$ (necessarily of degree s), then (ADG-3) remains fully valid: $Q(-k) = (-1)^{d-1}Q(k-s)$.

Stanley [25], [27] has extended the commutative algebra approach to the study of the solutions of the inhomogeneous systems associated with $\mathcal{L}$. The K-vector space D whose basis is given by the non-negative solutions to an associated inhomogeneous system is a rank 1 module over $K[M]$. If D is a Cohen–Macaulay module, then a reciprocity law holds similar to that in (iii).

Under certain conditions all these rank 1-modules are divisorial (and a Cohen–Macaulay rank 1 module is automatically divisorial). However, up to isomorphism over $K[M]$ there exist only finitely many Cohen–Macaulay divisorial ideals; see [7].

Ehrhart functions

Let $P \subset \mathbb{R}^n$ be a rational polytope, i. e. the convex hull of finitely many points $x_1, \ldots, x_m$. The study of the function

$$E(P,k) = \#(kP \cap \mathbb{Z}^n),$$

counting the lattice points in the multiples $kP = \{kx : x \in P\}, k \in \mathbb{Z}_+$, was pioneered by E. Ehrhart [12]. For the commutative algebra approach we let

$$C(P) = \mathbb{R}_+(P,1) = \mathbb{R}_+\{(y,1) \in \mathbb{R}^{n+1} : y \in P\}$$

be the cone over P. The cone is generated by finitely many rational vectors, and therefore $M = C(P) \cap \mathbb{Z}^{n+1}$ is a normal affine monoid. As the grading on M we choose the last coordinate of the elements $x \in M$. Let K be a field. Then clearly

$$E(P,k) = H(K[M],k),$$

and all the results on affine monoid algebras can be exploited for the study of $E(P,k)$. The Hilbert function of the canonical module of $K[M]$ is given by

$$E^+(P,k) = \#(\text{relint}(kP) \cap \mathbb{Z}^n).$$

The corresponding generating functions are

$$E_P(t) = \sum_{k=0}^{\infty} E(P,k)t^k \qquad \text{and} \qquad E_P^+(t) = \sum_{k=0}^{\infty} E^+(P,k)t^k.$$

Theorem 6.14 (Ehrhart). *Let $P \subset \mathbb{R}^n$ be a d-dimensional rational polytope, $d > 0$. Then*

(a) *$E_P(t)$ is a rational function, and there exist quasi-polynomials Q and Q^+ with $E(P,k) = Q(k)$ for all $k \geq 0$ and $E^+(P,k) = Q^+(k)$ for all $k \geq 1$;*
(b) *$E_P^+(t) = (-1)^{d+1}E_P(t^{-1})$, equivalently*

$$Q(k) = (-1)^d Q^+(-k) \qquad \textit{for all} \quad k \geq 1.$$

If P is even an integral polytope, i. e. its vertices belong to $\mathbb{Z}^n$, then $E(P,k)$ is given by a polynomial, since $K[M]$ is finite over its subalgebra S generated by the monomials $(x,1)$, x a vertex of P, and S is an algebra generated in degree 1. Therefore $K[M]$ has a well-defined multiplicity. In order to save us a discussion of volume functions, we restrict the corollary to full-dimensional polytopes:

Corollary 6.15. Let $P \subset \mathbb{R}^d$ be a d-dimensional integral polytope, and let $K[M]$ be the normal monoid algebra constructed above. Then

$$e(K[M]) = d!\ \mathrm{vol}\ P.$$

Proof. Elementary arguments of measure theory show that the volume of P is

$$\mathrm{vol}\ P = \lim_{k\to\infty} \frac{E(P,k)}{k^d}.$$

Being the Hilbert polynomial of the $(d+1)$-dimensional S-module, $E(P,k)$ has degree d. Thus its leading coefficient is given by vol P. On the other hand, it is also given by $e(k[M])/d!$. □

Unimodality of the h-vector

As we have seen, the h-vector $(h_0,\dots,h_u)$ of the Gorenstein ring $K[\mathcal{M}_n]$ is palindromic: $h_{u-i} = h_i$ for $i = 0,\dots,u$. Stanley's conjecture (ADG-5) asks whether it is *unimodal*, i.e. it increases until it reaches its maximum and decreases afterwards. This conjecture has been proved by Athanasiadis [3] and generalized to the Ehrhart functions of compressed Gorenstein polytopes – P is *Gorenstein* if $K[M]$ is Gorenstein, and P is compressed if all its so-called pulling triangulations are unimodular. Meanwhile it has been shown by Römer and the author [10] that unimodality holds for the Ehrhart functions of Gorenstein lattice polytopes, provided they have at least one regular unimodular triangulation. (See [6] for the notion of regular triangulation.)

The method of proof is a reduction to Stanley's g-theorem for simplicial polytopes, for which no access via commutative algebra has been found so far. The simplicial polytope is constructed from the unimodular triangulation and it seems difficult to further weaken the hypothesis on P.

The existence of a unimodular triangulation implies that $K[M]$ is generated by its degree 1 elements. If not even this condition is fulfilled, then the generators of higher degree induce "phenomena of higher period" in the h-vector. Unimodality can no longer be expected, and an example of a Gorenstein polytope P for which unimodality fails has been given by Musţată and Payne [20].

To the best of our knowledge it is still an open question whether there exist graded Gorenstein integral domains that are generated in degree 1, but whose h-vector is not unimodal.

Acknowledgement

These notes are an expanded version of my lectures at the International Conference on Commutative Algebra and Combinatorics organized by the Bhaskaracharya

Pratishthana, Pune and the Harish-Chandra Research Institute, Allahabad, December 2003. I am grateful to the DFG for a travel grant and to our Indian colleagues that they have made possible a very successful conference in the pleasant environment of the Harish Chandra Research Institute.[1]

References

[1] M. Ahmed, J. De Loera and R. Hemmecke, Polyhedral cones of magic cubes and squares, In: *New directions in computational geometry*, (eds.) B. Aronov et al., (Springer) (2003) 25–41
[2] H. Anand, V. C. Dumir and H. Gupta, A combinatorial distribution problem, *Duke Math. J.* **33** (1966) 757–769
[3] Ch. A. Athanasiadis, Ehrhart polynomials, simplicial polytopes, magic squares and a conjecture of Stanley, *J. Reine Angew. Math.* **583** (2005) 163–174
[4] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison–Wesley (1969)
[5] J.-F. Boutot, Singularités rationelles et quotients par les groupes réductifs, *Invent. Math.* **88** (1987) 65–68
[6] W. Bruns and J. Gubeladze, *Polytopes, rings and K-theory*, In preparation. Preliminary version at `http://www.math.uos.de/staff/brunsw`
[7] W. Bruns and J. Gubeladze, Divisorial linear algebra of normal semigroup rings, *Algebr. Represent. Theory* **6** (2003) 139–168
[8] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, (Rev. ed.) University Press, Cambridge (1998)
[9] W. Bruns and B. Ichim, On the coefficients of Hilbert quasipolynomials, *Proc. Amer. Math. Soc.*, to appear. `http://arxiv.org/format/math.AC/0512329`
[10] W. Bruns and T. Römer, h-vectors of Gorenstein polytopes, *J. Comb. Th. Ser. A*, in press
[11] V. I. Danilov, The geometry of toric varieties, *Russian Math. Surveys* **33** (1978) 97–154
[12] E. Ehrhart, *Polynômes arithméthiques et méthode des polyédres en combinatoire*, Birkhäuser (1977)
[13] R. Fossum, *The divisor class group of a Krull domain*, Springer (1973)
[14] W. Fulton, *Introduction to toric varieties*, Princeton University Press (1993)
[15] R. V. Gurjar, On a conjecture of C. T. C. Wall, *J. Math. Kyoto Univ.* **31** (1991) 1121–1124
[16] D. Hilbert, Über die Theorie der algebraischen Formen, *Math. Ann.* **36** (1890) 473–534
[17] M. Hochster, Rings of invariants of tori, Cohen–Macaulay rings generated by monomials, and polytopes, *Ann. of Math.* **96** (1972) 318–337
[18] M. Hochster and J. L. Roberts, Rings of invariants of reductive groups acting on regular rings are Cohen–Macaulay, *Adv. Math.* **13** (1974) 115–175
[19] G. Kempf, F. Knudsen, D. Mumford and B. Saint-Donat, *Toroidal embeddings I*, LNM, **339**, Springer (1973)
[20] M. Mustaţă and S. Payne, Ehrhart polynomials and stringy Betti numbers, *Math. Ann.* **333** (2005) 787–795
[21] R. Y. Sharp, *Steps in commutative algebra*, (2nd ed.) Cambridge University Press, (2001)
[22] R. P. Stanley, Linear homogeneous diophantine equations and magic labelings of graphs, *Duke Math. J.* **40** (1973) 607–632
[23] R. P. Stanley, Magic labelings of graphs, symmetric magic squares, systems of parameters, and Cohen–Macauly rings, *Duke Math. J.* **43** (1976) 511–531

[1] It was a special pleasure for the author to lecture about the ADG-conjectures in the presence of his old friend V. C. Dumir. In the academic year 1977/78 the Dumir and Bruns families were neighbors in the married student housing at Orchard Downs, Urbana, Illinois.

[24] R. P. Stanley, Hilbert functions of graded algebras, *Adv. Math.* **28** (1978) 57–83
[25] R. P. Stanley, Linear diophantine equations and local cohomology, *Invent. Math.* **68** (1982) 175–193
[26] R. P. Stanley, *Enumerative combinatorics*, Vol. I, Wadsworth & Brooks/Cole (1986)
[27] R. P. Stanley, *Combinatorics and commutative algebra*, (2nd ed.) Birkhäuser (1996)
[28] Ngô Viêt Trung and Lê Tuân Hoa, Affine semigroups and Cohen–Macaulay rings generated by monomials, *Trans. Amer. Math. Soc.* **298** (1986) 145–167
[29] G. M. Ziegler, *Lectures on polytopes*, Springer (1995)