UNIVERSITATIS IAGELLONICAE ACTA MATHEMATICA, FASCICULUS XXXIX
2001

COMPUTING THE INTEGRAL CLOSURE OF AN AFFINE SEMIGROUP

BY WINFRIED BRUNS AND ROBERT KOCH

1. Introduction. An *affine semigroup* S is a finitely generated subsemigroup of a finitely generated free abelian group (or lattice) $\mathbb{Z}^n$. (We use the term 'semigroup' as a synonym for 'monoid'; so all our semigroups have a neutral element 0.) Let L be a sublattice of $\mathbb{Z}^n$ containing S. Then the *integral closure* of S in L is the set

$$\bar{S}_L = \{x \in L : \ mx \in S \text{ for some } m \in \mathbb{N}, \ m > 0\}.$$

In the special case where L coincides with the group $\operatorname{gp}(S)$ of differences of S, one calls $\bar{S} = \bar{S}_L$ the *normalization* of S. Obviously $\bar{S}_L$ is a subsemigroup of L.

The integral closure can be described geometrically. Let $C(S)$ be the cone generated by S in the vector space $\mathbb{R}^n$, i.e., the set of all linear combinations of elements of S with non-negative real coefficients. It is an elementary fact that $\bar{S}_L = C(S) \cap L$. Note that $C(S)$ is finitely generated by rational vectors since S is so. It follows (and is in fact equivalent) that $C(S)$ is the intersection of finitely many rational vector halfspaces H_i, $i = 1, \dots, r$. Moreover, $\bar{S}_L$ is itself finitely generated by Gordan's lemma.

We call S *positive* if $x, -x \in S$ is possible only for $x = 0$. It is not hard to show that a positive affine semigroup can be embedded into a positive orthant $\mathbb{Z}^s_+$ for some s (actually the smallest possible value $s = \operatorname{rank} \operatorname{gp}(S)$ suffices). It then follows that every element of S can be written as the sum of irreducible elements, and since S is finitely generated, it can have only finitely many irreducibles. The finite set of irreducibles is the unique minimal generating set of S. We call this set the *Hilbert basis*, $\operatorname{Hilb}(S)$, of S.

The authors have developed the computer program normaliz [**5**] for the computation of $\mathrm{Hilb}(\bar{S}_L)$ (evidently $\bar{S}_L$ is positive if S is). Normaliz has already proved very useful in various investigations; see Villarreal [**10**]. In particular it has played a crucial role in finding a counterexample to the unimodular covering conjecture and the discrete Carathéodory property of normal affine semigroups (Bruns and Gubeladze [**2**], Bruns et al. [**3**]). It is the purpose of this article to explain the algorithm used by normaliz. Most likely, all the ideas involved have appeared elsewhere, and we do not claim originality for them.

Because of the embedding $S \to \mathbb{Z}_+^s$, each positive affine semigroup can be graded, i.e., there exists a semigroup homomorphism $\deg : S \to \mathbb{N}$ such that $\deg(x) = 0$ if and only if $x = 0$. Then $K[S]$ is a positively graded K-algebra (under the canonical extension of the degree function, where K is an arbitrary field). If deg is the restriction of a $\mathbb{Z}$-linear form on L (and it always is after multiplication by a positive integer), then S is a graded subsemigroup of $\bar{S}_L$. If, in addition, S is generated by all $x \in S$ with $\deg(x) = 1$, then we say that S is *homogeneous with respect to* L (and simply *homogeneous* if $L = \mathrm{gp}(S)$). In this case normaliz can compute the *Hilbert function* of $\bar{S}_L$ given by $H(\bar{S}_L, i) = \mathrm{card}\{x \in \bar{S}_L \ \deg(x) = i\}$. Since $\bar{S}_L$ is a finite module over a homogeneous semigroup. we call it *almost homogeneous*.

Note that our nomenclature is consistent with its use in commutative algebra. Let K be a field. Upon the choice of a basis $e_1, \dots, e_n$ we can identify the group algebra $K[L]$ with the Laurent polynomial ring $K[X_1^{\pm 1}, \dots, X_n^{\pm 1}]$, and the semigroup ring $K[S]$ with a monomial subalgebra. Then $K[\bar{S}_L]$ is the integral closure of $K[S]$ in $K[L]$ (or its field of fractions). In particular, $K[\bar{S}]$ is the normalization of $K[S]$. The Hilbert function of a graded semigroup S coincides with the Hilbert function of the semigroup algebra $K[S]$.

Affine semigroup rings are the coordinate rings of (not necessarily normal) toric varieties, and homogeneous such rings are the homogeneous coordinate rings of projective toric varieties. Therefore normaliz has applications in commutative algebra and algebraic geometry.

In its present version normaliz requires the generators of S as input and allows only the choices $L = \mathrm{gp}(S)$ or $L = \mathbb{Z}^n$. These choices for L cover almost all potential applications.

It is the aim of this note to explain the algorithm used by normaliz. Many facts which we will use without proof belong to the classical theory of convex polyhedral cones. See Gale [**7**] and Gerstenhaber [**8**].

We do not attempt to describe the potential applications of normaliz. See the documentation of normaliz, the book [**10**] of Villarreal, and Bruns, Gubeladze, and Trung [**4**] for more information.

2. Computing the Hilbert basis. *Finiteness of the integral closure.* We start by showing that the integral closure of an affine semigroup $S \subset \mathbb{Z}^n$, in a sublattice L of $\mathbb{Z}^n$ which contains S, is finitely generated, and give a geometric description of the integral closure. The subcone of $\mathbb{R}^n$ generated by S is denoted by $C(S)$.

PROPOSITION 2.1. (a) (Gordan's lemma) *Let $C \subset \mathbb{R}^n$ be a finitely generated rational cone (i.e., generated by finitely many vectors from $\mathbb{Q}^n$). Then $\mathbb{Z}^n \cap C$ is an affine semigroup and integrally closed in $\mathbb{Z}^n$.*

(b) *Let S be an affine subsemigroup of the lattice $L \subset \mathbb{Z}^n$. Then*
   (i) *$\bar{S}_L = L \cap C(S)$;*
   (ii) *there exist $z_1, \dots, z_u \in \bar{S}_L$ such that $\bar{S}_L = \bigcup_{i=1}^{u}(z_i + S)$;*
   (iii) *$\bar{S}_L$ is an affine semigroup.*

PROOF. (a) Note that C is generated by finitely many elements $x_1, \dots, x_m \in \mathbb{Z}^n$. Let $x \in \mathbb{Z}^n \cap C$. Then $x = a_1x_1 + \dots + a_mx_m$ with non-negative rational a_i. Set $b_i = \lfloor a_i \rfloor$. Then

$$(*) \qquad x = (b_1x_1 + \dots + b_mx_m) + (r_1x_1 + \dots + r_mx_m), \qquad 0 \le r_i < 1.$$

The second summand lies in the intersection of $\mathbb{Z}^n$ with a bounded subset of C. Thus there are only finitely many choices for it. These elements together with $x_1, \dots, x_m$ generate $\mathbb{Z}^n \cap C$. That $\mathbb{Z}^n \cap C$ is integrally closed in $\mathbb{Z}^n$ is evident.

(b) Set $C = C(S)$, and choose a system $x_1, \dots, x_m$ of generators of S. Then every $x \in L \cap C$ has a representation $(*)$. Multiplication by a common denominator of $r_1, \dots, r_m$ shows that $x \in \bar{S}_L$. On the other hand, $L \cap C$ is integrally closed in L, and so $\bar{S}_L = L \cap C$.

The elements $z_1, \dots, z_u$ can now be chosen as those vectors $r_1x_1 + \dots + r_mx_m$ that appear in $(*)$ and belong to L. Their number is finite since they are all integral and contained in a bounded subset of $\mathbb{R}^n$. Together with $x_1, \dots, x_m$ they certainly generate $\bar{S}_L$ as a semigroup. $\square$

In principle Proposition 2.1 tells us how to determine the generators of $\bar{S}_L$: we only need to search these elements in a bounded subset of $\mathbb{Z}^n$. However, it is difficult to generate the candidates in an effective way without some preparations.

Reduction to a full rank embedding. In the first step we reduce the problem to computing the integral closure of S in a lattice L' such that $\operatorname{rank} L' = \operatorname{rank} \operatorname{gp}(S)$. (We will write $\operatorname{rank} S$ for $\operatorname{rank} \operatorname{gp}(S)$ in the following.)

Applying the elementary divisor algorithm one finds a basis $e_1, \dots, e_n$ of $\mathbb{Z}^n$ and integers $\alpha_1, \dots, \alpha_n$ such that $f_i = \alpha_i e_i$, $i = 1, \dots \operatorname{rank} L$, is a basis of L. After a linear transformation we can assume that $L = \mathbb{Z}^n$, and that we have to compute the integral closure of S in $\mathbb{Z}^n$. Henceforth the index L will be dropped.

The elementary divisor algorithm is applied again in order to find a basis $f_1 \dots, f_n$ of $L\mathbb{Z}^n$ and integers $\beta_1, \dots .\beta_r$, $r = \operatorname{rank} S$, such that $\beta_1 f_1, \dots, \beta_r f_r$ is a basis of $\operatorname{gp}(S)$. The integral closure of S in $\mathbb{Z}^n$ evidently coincides with the integral closure of S in $L' = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_r$. Consequently we may further assume that $\operatorname{rank} S = n$.

It is clear that at the end of all the computations the linear transformations inverse to those above have to be applied in order to rewrite the output in the coordinates of the input.

The program normaliz allows for L only $\mathbb{Z}^n$ or $\operatorname{gp}(S)$. Therefore one application of the elementary divisor algorithm is sufficient: if $L = \mathbb{Z}^n$, then one chooses $L' = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_r$, and if $L = \operatorname{gp}(S)$ one has to take $L' = \mathbb{Z}\beta_1 f_1 + \cdots + \mathbb{Z}\beta_r f_r$.

Extreme rays, faces and facets. Let $C \subset \mathbb{R}^n$ be a finitely generated cone. We can assume that $\dim C = n$, replacing $\mathbb{R}^n$ by the vector subspace generated by C if necessary. (Above we have computed such an embedding in the discrete case.) Let H be a vector subspace of dimension $n-1$. It determines two halfspaces. If C is contained in one of these (closed) halfspaces, then $F = C \cap H$ is called a *face* of C. The *dimension* of a face is the dimension of the vector subspace that it generates. Faces of dimension $n-1$ are called *facets*. It is often useful to count C as a face of C.

From the algebraic and also from the computational point of view a subspace H is the kernel of a linear form ϕ, uniquely determined up to a nonzero constant factor. Replacing ϕ by $-\phi$ if necessary, we can always assume that C is contained in the positive halfspace associated with ϕ (or H).

An *extreme ray* is a halfline starting in 0 that is contained in a 1-dimensional face. Evidently each 1-dimensional face is either an extreme ray or the union of two extreme rays. In particular, if C does not contain a full line (and this will be the case later on), then we can identify 1-dimensional faces and extreme rays.

PROPOSITION 2.2. *Let C be a finitely generated cone.*

(a) *A subset X of C is a minimal generating set if and only if $0 \notin X$ and X contains exactly one element from each extreme ray.*
(b) *There is exactly one irredundant representation of C as the intersection of vector halfspaces, namely that by the positive halfspaces associated with the facets of C.*

For each facet F the linear form σ_F defining the half-space in (b) is unique up to a positive factor, and we may speak of σ_F as the *support form* associated with the facet F. If C is rational, then there is a natural choice for σ_F, and we always use it in the rational case: σ_F is rational (since its kernel is generated by rational vectors), and it has a unique multiple with coprime integral coefficients. (Normaliz always clears denominators and removes common divisors during its computations.) Next we discuss how to compute the σ_F.

The dual cone algorithm. For a cone $C \subset \mathbb{R}^n$ one defines the *dual* (or *polar*) *cone* by

$$C^* = \{\phi \in (\mathbb{R}^n)^* : \phi(x) \geq 0 \text{ for all } x \in C\}.$$

The following proposition justifies this term.

PROPOSITION 2.3. *Let C be a cone in $\mathbb{R}^n$.*

(a) *The bidual cone C^{**} is the topological closure of C in $\mathbb{R}^n$ (which we identify with its bidual $(\mathbb{R}^n)^{**}$ via the natural isomorphism).*

(b) *If C is finitely generated (rational), then C^* is finitely generated (rational). Moreover, $C^{**} = C$.*

PROOF. (a) Since linear forms are continuous, the topological closure $\hat{C}$ is contained in $C^{**} = \{x \in \mathbb{R}^n : \phi(x) \geq 0 \text{ for all } \phi \in C^*\}$.

For the converse inclusion consider $x \in C^{**}$. Evidently $\hat{C}$ is convex. If $x \notin \hat{C}$, then the Hahn-Banach separation theorem yields a linear form ϕ and a real number α such that $\phi(x) < \alpha$ and $\phi(y) > \alpha$ for all $y \in \hat{C}$. This is impossible if $\phi(z) < 0$ for some $z \in \hat{C}$, since $\beta z \in \hat{C}$ for all $\beta \geq 0$. Thus $\phi \in C^*$, and we obtain a contradiction.

(b) If C is finitely generated, then C is closed, and so $C^{**} = C$ by (a). It remains to show that C^* is finitely generated if C is. The dual cone algorithm, outlined below, will show this. If C is rational, then it finds rational generators for C^*. $\square$

COROLLARY 2.4. *Let the cone C be the intersection of finitely many (rational) halfspaces. Then it is finitely generated (and rational).*

In fact, C is closed, and we can apply Proposition 2.3. The corollary indicates why finitely generated cones appear in linear optimization where constraints are given by linear inequalities.

Let $x_1, \dots, x_m \in \mathbb{R}^n$. We want to find the dual cone of $C = \mathbb{R}_+ x_1 + \dots + \mathbb{R}_+ x_m$. We can assume that $\mathbb{R}^n$ is generated by $x_1, \dots, x_m$ as a vector space. (For the data for which we want to use the algorithm this assumption is satisfied after we have passed to a full rank embedding.)

We first search for n vectors among $x_1 \dots, x_m$ that form a basis of $\mathbb{R}^n$, say $x_1 \dots, x_n$. For each $i = 1, \dots, n$ we compute a linear form ϕ_i such that $\phi_i(x_i) > 0$, $\phi_i(x_j) = 0$ for $j \neq i$. Clearly ϕ_i is uniquely determined up to a positive factor. One also checks immediately that $\phi_1, \dots, \phi_n$ is a basis of $(\mathbb{R}^n)^*$ and generate the dual cone of $C_0 = \mathbb{R}_+x_1 + \dots + \mathbb{R}_+x_n$.

This initialization is useful because it simultaneously starts a triangulation of the cone C. This triangulation will be needed for the computation of the Hilbert basis. We now describe how the dual cone changes if we enlarge C by another generator.

PROPOSITION 2.5. *Let $x_1, \dots, x_m, y \in \mathbb{R}^n$ be such that $x_1, \dots, x_m$ generate $\mathbb{R}^n$ as a vector space. Suppose that $\phi_1, \dots, \phi_t$ generate the dual cone of $C = \mathbb{R}_+x_1 + \dots + \mathbb{R}_+x_m$. For each pair (i,j), $i,j = 1, \dots, t$, with $\phi_i(y) > 0$ and $\phi_j(y) < 0$ we set*

$$\psi_{ij} = \phi_i(y)\phi_j - \phi_j(y)\phi_i.$$

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

For the proof see Burger [**6**]. A geometric explanation follows after the next proposition.

The generating set of C^* specified by Proposition 2.5 contains a minimal system of generators, and because of Proposition 2.3 it consists exactly of a set of support forms σ_F, F running through the facets of C. It is not difficult to find this minimal generating set:

PROPOSITION 2.6. *With the notation of Proposition 2.5, suppose that $\phi_1, \dots. \phi_t$ are the support forms of C, and denote by H_i the hyperplane given by the vanishing of ϕ_i. Then the support forms of $\tilde{C}$ are given by the ϕ_i with $\phi_i(y) \geq 0$ and those ψ_{ij} such that $H_i \cap H_j \cap C$ is not contained in one of the hyperplanes H_k, $k \neq i,j$.*

Let us say that a subset X of C is *visible* from y if for each $x \in X$ the line segment from y to x intersects C exactly in x. It is geometrically evident that one finds the facets of $\tilde{C}$ by taking first those facets of C that do not separate y from C, and second those hyperplanes that pass through y and the $(n-2)$-dimensional faces of C that bound the part of C that is visible from y. Exactly these hyperplanes are specified by the two propositions above: an $(n-2)$-dimensional face is contained in exactly two facets, and it bounds the visible area if exactly one of these facets is visible from C.

Once the dual cone (equivalently, the support forms) of C have been computed, we can decide whether C is *positive*, i.e., 0 is the only element $x \in C$ such that $-x \in C$, too.

PROPOSITION 2.7. *Let $C \subset \mathbb{R}^n$ be an n-dimensional cone, and $S \subset \mathbb{Z}^n$ an affine semigroup.*

(a) *C is positive if and only if C^* has dimension n.*

(b) *S is positive if and only if $C(S)$ is positive.*
(c) *If S is positive, then it has a unique minimal generating set, given by its finitely many irreducible elements.*

PROOF. (a) It is obvious that $x, -x \in C$ if and only if $\phi(x) = 0$ for all $\phi \in C^*$, and such an element $x \neq 0$ exists if and only if $\dim(C^*) < n$.

(b) This is trivial.

(c) Let $\sigma_1, \dots, \sigma_s$ be the support forms of $C(S)$. Then we consider the homomorphism $\sigma : \mathbb{R}^n \to \mathbb{R}^s$, $\sigma(x) = (\sigma_1(x), \dots, \sigma_s(x))$. Because of (a) and (b) σ is injective, and it maps S isomorphically onto a subsemigroup of $\mathbb{Z}_+^s$. We can assume that $S \subset \mathbb{Z}_+^s$. Then it follows easily that each element x of S is the sum of irreducible elements, for example by induction on the sum of the components of x. Therefore S is generated by its irreducible elements, and since S has a finite system of generators and every system of generators must contain the irreducibles, their number is finite. $\square$

At this point normaliz tests whether S is positive. If not, it stops. It would not be difficult to extend the program in such a way that it covers the general case. Let U be the kernel of the homomorphism $\sigma : \mathbb{Z}^n \to \mathbb{Z}^s$. Then the image T of S in $\mathbb{Z}^n/U$ is a positive affine semigroup. It is enough to lift the Hilbert basis of $\bar{T}$ back to $\mathbb{Z}^n$, because $\bar{S}$ is generated by U and preimages of the Hilbert basis of $\bar{T}$. (We are still assuming that $S \subset \mathbb{Z}^n$ has rank n.)

Computing the triangulation. A cone is simplicial if it is generated by a linearly independent set of vectors. By a *triangulation* of a cone C we mean a decomposition into a family Δ of finitely many simplicial subcones such that the intersection of $\delta, \epsilon \in \Delta$ is a face of both δ and ϵ. A triangulation Δ is uniquely determined by those $\delta \in \Delta$ such that $\dim \delta = \dim C$. Therefore the triangulation can be described by a list of n-tuples of vectors in $\mathbb{R}^n$, where each n-tuple contains the generators of an n-dimensional simplicial cone.

Let $C \subset \mathbb{R}^n$ be a cone of dimension n, given by generators $x_1, \dots, x_m$. In the computation of the dual cone C^* we have started with a simplicial subcone C_0 generated by a linearly independent subset of $\{x_1, \dots, x_m\}$. It has a trivial triangulation by its faces (including C_0 itself). Therefore it is enough to describe how to pass from a triangulation of C to a triangulation of $\tilde{C}$ generated by $x_1, \dots, x_m, y$.

If y is contained in C, then we can (and do) simply keep the triangulation δ. So suppose that $y \notin C$. Let Δ be a triangulation of C. Then we obtain a triangulation of $\tilde{C}$ by joining Δ with the set of all cones $\delta + \mathbb{R}_+ y$ where $\delta \in \Delta$ is visible from y.

The new n-dimensional simplicial subcones are generated by y and the $(n-1)$-dimensional visible cones δ' of Δ. Such a subcone is visible if and only if it lies in a facet F of C that is visible from y, and the visible facets F are characterized by the fact that $\sigma_F(y) < 0$. This makes it easy for normaliz to find the new n-dimensional members of the triangulation, since the support forms of C are known. (However, note that a facet of C itself is not simplicial in general, and even if it is: Δ may subdivide it, if the given generating set of C is not minimal.)

At this point we can also discuss how to find the irreducible elements of $\bar{S}$, and thus $\operatorname{Hilb}(\bar{S})$, once a system of generators $x_1, \dots, x_m \neq 0$ of $\bar{S}$ is known: x_i is irreducible if $x_i - x_j \notin \bar{S}$ for all $j \neq i$. This criterion holds for arbitrary S. However, the condition $x_i - x_j \notin S$ is difficult to verify in general. For $\bar{S}$ it is easy: we simply test whether the condition $\sigma_F(x_i - x_j) \geq 0$ is violated for at least one facet F.

Suppose further that we have found a system of generators for the semigroup $\delta \cap \mathbb{Z}^n$ for each δ in a triangulation of $C(S)$. Then the union of all these systems obviously generates $\bar{S}$. We have already constructed a triangulation Δ, and each $\delta \in \Delta$ is specified by a set of integral, linearly independent vectors generating δ. It only remains to find the generators of $\bar{S}$ if S is simplicial.

Simplicial cones. Let $x_1, \dots, x_n$ be linearly independent elements of $\mathbb{Z}^n$ and let C be the cone spanned by them and S the affine semigroup they generate. Then each $y \in \bar{S} = C \cap \mathbb{Z}^n$ has a representation

$$y = (a_1 x_1 + \cdots + a_n x_n) + (q_1 x_1 + \cdots + q_n x_n), \qquad a_i \in \mathbb{Z}_+,\ q_i \in \mathbb{Q},\ 0 \leq q_i < 1.$$

We collect the second summands in the set

$$\operatorname{par}(x_1, \dots, x_n) = \mathbb{Z}^n \cap \{q_1 x_1 + \cdots + q_n x_n : q_i \in \mathbb{Q},\ 0 \leq q_i < 1\}.$$

The notation par (introduced by Sebö [**9**]) is suggested by the fact that the elements of $\operatorname{par}(x_1, \dots, x_n)$ are exactly the lattice points in the semi-open parallelepiped spanned by $x_1, \dots, x_n$.

LEMMA 2.8. *The set* $\operatorname{par}(x_1, \dots, x_n)$ *contains exactly one representative from each residue class of* $\mathbb{Z}^n$ *modulo* $U = \mathbb{Z}x_1 + \cdots + \mathbb{Z}x_n$. *Therefore*

$$\operatorname{card} \operatorname{par}(x_1, \dots, x_n) = \operatorname{card}(\mathbb{Z}^n / U) = |\det(x_1, \dots, x_n)|.$$

Moreover, $\bar{S}$ *is the disjoint union of the sets* $z + S$, $z \in \operatorname{par}(x_1, \dots, x_n)$.

PROOF. The first statement is evident, and it implies the first equation. The second equation results from the elementary divisor theorem. That $\bar{S}$ is the union of the sets $z + S$ has been shown in Proposition 2.1, and that the union is disjoint follows immediately from the fact that the $z \in \operatorname{par}(x_1, \dots, x_n)$ represent different residue classes. $\square$

In the language of commutative algebra: let K be a field; then $\mathrm{par}(x_1,\dots, x_n)$ is a basis of the free $K[S]$-module $K[\bar{S}]$, and $K[S]$ is actually a polynomial ring over K.

Together with $x_1,\dots,x_n$ the set $\mathrm{par}(x_1,\dots,x_n)$ certainly generates $\bar{S}$. Therefore it is enough to find an efficient method for producing $\mathrm{par}(x_1,\dots,x_n)$ from $x_1,\dots,x_n$.

First one applies the elementary divisor algorithm to find a basis $u_1\dots,u_n$ of $\mathbb{Z}^n$, and positive integers $\lambda_1,\dots,\lambda_n$ such that $\lambda_1u_1,\dots,\lambda_nu_n$ is a basis of $\mathrm{gp}(S)$. Clearly $d = \det(x_1,\dots,x_n) = \lambda_1\cdots\lambda_n$, and $d\mathbb{Z}^n \subset \mathrm{gp}(S)$, since $\mathbb{Z}^n/\mathrm{gp}(S)$ is a direct sum of n cyclic groups of orders $\lambda_1,\dots,\lambda_n$.

The residue classes of $\mathbb{Z}^n$ modulo $\mathrm{gp}(S)$ are represented by the vectors

$$e = b_1u_1 + \dots + b_nu_n, \qquad b_i = 0,\dots,\lambda_i - 1, \quad i = 1,\dots,n.$$

Each such vector e has a representation $e = a_1x_1 + \dots + a_nx_n$ with *rational* coefficients a_i. Now we set $q_i = a_i - \lfloor a_i \rfloor$, so that $e' = q_1x_1 + \dots + q_nx_n$ represents the residue class of e and belongs to $\mathrm{par}(x_1,\dots,x_n)$. Note that d is a suitable common denominator for the a_i, since $d\mathbb{Z}^n \subset \mathrm{gp}(S)$. Therefore one can keep all the coefficients integral by first passing to de and dividing by d at the end.

3. Computing the Hilbert series. Suppose that A is a positively graded affine semigroup; one has fixed a homomorphism $\deg : \mathrm{gp}(A) \to \mathbb{Z}$ such that $deg(A) \subset \mathbb{Z}_+$ and 0 is the only element of A having degree 0. Then the set $A_k = \{x \in A : \deg x = k\}$ is finite for each $k \in \mathbb{Z}_+$, and we can define the *Hilbert function*

$$H(A,k) = \operatorname{card} A_k, \qquad k \in \mathbb{Z}_+,$$

of A. The *Hilbert series* of A is the formal power series

$$H_A(T) = \sum_{k=0}^{\infty} (\operatorname{card} A_k)T^k.$$

If K is a field, then the semigroup algebra $K[A]$ inherits the grading, and the Hilbert series of A is just the Hilbert series of $K[A]$. Therefore $H_A(T)$ has all the properties that are known for Hilbert series of positively graded K-algebras (see [**1**, Chap. 4]). In particular, $H_A(T)$ represents a rational function,

$$H_A(T) = \frac{Q(T)}{(1-T^{d_1})\cdots(1-T^{d_n})}$$

where $Q(T)$ is a polynomial, $n = \operatorname{rank} A$, and $d_1,\dots,d_n$ are positive integers.

The situation further simplifies if A is *almost homogeneous*. This means that there exists an affine subsemigroup $A_0 \subset A$ which is generated by elements of degree 1 and over which A is a finite module, i.e., there exist $x_1,\dots,x_m \in A$

such that $A = \bigcup_{i=1}^m (A_0 + x_i)$. Then $K[A]$ is a finitely generated module over $K[A_0]$, and therefore $H_A(T)$ can be represented in the form

$$H_A(T) = \frac{Q(T)}{(1-T)^n}.$$

For $k \gg 0$ the Hilbert function $H(A,k)$ is given by a polynomial, the *Hilbert polynomial* $P_A(k)$. It is a polynomial of degree $n-1$ with leading coefficient $e(A)/(n-1)!$, where $e(A) = Q(1)$ is the *multiplicity* of A.

In normaliz the role of A_0 is played by the given affine semigroup S and that of A is played by the integral closure $\bar{S}$. (We assume, as before, that $S \subset \mathbb{Z}^n$, $\operatorname{rank} S = n$, and the integral closure is taken with respect to $\mathbb{Z}^n$.) We want to compute the Hilbert series of $\bar{S}$. In principle this would be possible for the general case, but so far it has only been implemented in the almost homogeneous case (simply called "homogeneous" in the normaliz documentation), and from now on we restrict ourselves to this case.

Normaliz has computed a triangulation Δ of the cone C generated by S such that each simplicial cone $\delta \in \Delta$ is generated by elements of S, and by assumption these have degree 1 in $\bar{S}$. The triangulation defines a disjoint decomposition

$$C = \bigcup_{\delta \in \Delta} \operatorname{relint}(\delta),$$

where $\operatorname{relint}(\delta)$ is the interior of δ relative to the vector subspace of $\mathbb{R}^n$ generated by δ. We set

$$\omega_\delta = \operatorname{relint}(\delta) \cap \mathbb{Z}^n.$$

It follows that

$$H_{\bar{S}}(T) = \sum_{\delta \in \Delta} H_{\omega_\delta}(T),$$

where the terms on the right hand side are defined in an obvious way.

Therefore, in order to compute $H_{\bar{S}}(T)$, two tasks have to be carried out, namely

1. the decomposition of C into the cones $\delta \in \Delta$, and
2. the computation of $H_{\omega_\delta}(T)$ for each δ.

The most time consuming part of normaliz is step 1 above because of the extreme combinatorial complexity of triangulations in general. For the computation of the Hilbert basis it is enough to consider the maximal simplicial cones in Δ, and it would be foolish to insist on a disjoint decomposition. (Some vectors are tested more than once for being members of the Hilbert basis if they belong to two or more maximal simplicial subcones. But this effect is negligible.) However, for the Hilbert series one cannot avoid the disjoint decomposition. We have recently implemented an essential improvement of the decomposition algorithm.

For step 2 we denote by $x_1,\dots,x_r$ the linearly independent degree 1 elements of S that generate δ. The semigroup S_δ they generate is free, and so $H_{S_\delta}(T)=1/(1-T)^r$. Furthermore one has a disjoint decomposition

$$\omega_\delta=\bigcup_{x\in\operatorname{par}'(x_1,\dots,x_r)}(x+S_\delta)$$

where

$$\operatorname{par}'(x_1,\dots,x_r)=\mathbb{Z}^n\cap\{q_1x_1+\cdots+q_rx_r : q_i\in\mathbb{Q},\ 0<q_i\le 1\}.$$

Therefore

$$H_{\omega_\delta}(T)=\frac{\sum_{k=1}^r\operatorname{card}(B_k)T^k}{(1-T)^r},$$

where $B_k=\{x\in\operatorname{par}'(x_1,\dots,x_r):\deg x=k\}$.

REMARK 3.1. (a) Often one is only interested in a single numerical invariant, namely the *multiplicity* $e(\bar S)$. (It coincides with the multiplicity of S if $\operatorname{gp}(S)=\mathbb{Z}^n$; in general one has $e(\bar S)=e(S)\cdot\operatorname{card}(\mathbb{Z}^n/\operatorname{gp}(S))$.) The different maximal simplices in the triangulation intersect each other only in lower dimensional cones, so that the leading coefficient of the Hilbert polynomial can be calculated without taking care of the lower dimensional cones. Then

$$e(\bar S)=\sum_{\delta\in\Delta,\dim\delta=n}e(\bar S_\delta)$$

where $\bar S_\delta=\delta\cap\mathbb{Z}^n$. Let $x_1,\dots,x_n$ be the linearly independent degree 1 generators of δ. Then $S_\delta=\mathbb{Z}_+x_1+\cdots+\mathbb{Z}_+x_n$ is a free affine semigroup and therefore of multiplicity 1. The integral closure $\bar S_\delta$ is a free S_δ-module (as already observed), and therefore its multiplicity coincides with the number of elements in its basis $\operatorname{par}(x_1,\dots,x_n)$. To sum up,

$$e(\bar S_\delta)=\operatorname{card}(\operatorname{par}(x_1,\dots,x_n))=|\det(x_1,\dots,x_n)|.$$

Therefore it is not necessary to compute the Hilbert basis in order to find the multiplicity. We offer the option -v for normaliz. It restricts all computations to multiplicities and those data which determine the triangulation.

The letter v has been chosen since the multiplicity of S can be interpreted as the normalized volume of the polytope spanned by the generators of S in the hyperplane of degree 1 elements. Thus normaliz -v can be used for the computation of volumes of lattice polytopes.

(b) It is not necessary to compute $\operatorname{par}'(x_1,\dots,x_r)$ separately. In fact $y\in\operatorname{par}'(x_1,\dots,x_r)$ if and only if $(x_1+\cdots+x_r)-y\in\operatorname{par}(x_1,\dots,x_r)$. We use this observation as follows.

The n-dimensional simplicial cones $\delta\in\Delta$ are scanned for the computation of the Hilbert basis. Suppose that δ is spanned by the degree 1 elements

$x_1 \ldots, x_n$. Then the elements $x \in \operatorname{par}(\delta)$ are computed since they are candidates for the Hilbert basis. We can write $x = q_1 x_{i_1} + \cdots + q_r x_{i_r}$ with $0 < q_i < 1$. For each subset $J \subset \{1, \ldots, n\}$ with $\{i_1, \ldots, i_r\} \subset J$ the vector $\sum_{j \in J} x_j - x$ belongs to $\operatorname{par}'(x_j : j \in J)$, and all the vectors necessary for the computation of the Hilbert function are produced by this method.

References

1. Bruns W., Herzog J., *Cohen-Macaulay rings,* Cambridge University Press, Cambridge (1993).
2. Bruns W., Gubeladze J., *Normality and covering properties of affine semigroups,* J. Reine Angew. Math. **510** (1999), 151–178.
3. Bruns W., Gubeladze J., Henk M., Martin A. Weismantel R., *A counterexample to an integer analogue of Carathéodory's theorem* J. Reine Angew. Math. **510** (1999), 179–185.
4. Bruns W., Gubeladze J., Trung N.V., *Problems and algorithms for affine semigroups,* preprint.
5. Bruns W., Koch R., *Normaliz,* a program to compute normalizations of semigroups. Available via anonymous `ftp` from `ftp.mathematik.Uni-Osnabrueck.DE/pub/osm/kommalg/software/`.
6. Burger E., *Uber homogene lineare Ungleichungssysteme,* Z. angew. Math. Mech. **36** (1956), 135–139.
7. Gale D., *Convex polyhedral cones and linear inequalities,* Activity Analysis of Production and Allocation, Chap. XVII, (1951), 287–297.
8. Gerstenhaber M., *Theory of convex polyhedral cones,* Activity Analysis of Production and Allocation, Chap. XVIII, (1951), 298–316.
9. Sebö A., *Hilbert bases, Carathéodory's theorem, and combinatorial optimization,* Integer Programming and Combinatorial Optimization (R. Kannan, W. Pulleyblank, eds.), University of Waterloo Press, Waterloo (1990), 431–456.
10. Villarreal R., *Monomial algebras,* M. Dekker, New York (2001).

Received April 8, 2001

Universität Osnabrück,
FB Mathematik/Informatik, 49069 Osnabrück,
Germany
e-mail: `Winfried.Bruns@mathematik.uni-osnabrueck.de`
e-mail: `Robert.Koch@mathematik.uni-osnabrueck.de`