Canonical Subalgebra Bases in Non-commutative Polynomial Rings

Patrik Nordbeck

Department of Mathematics
Lund Institute of Technology
Box 118, S-221 00 Lund, Sweden
nordbeck@maths.lth.se
www.maths.lth.se

Abstract

Canonical bases, also called SAGBI bases, for subalgebras of the non-commutative polynomial ring are investigated. The process of subalgebra reduction is defined. Methods, including generalizations of the standard Gröbner bases techniques, are developed for the test whether bases are canonical, and for the completion procedure of constructing canonical bases. The special case of homogeneous subalgebras is discussed.

In this paper we study canonical bases, analogs of Gröbner bases for ideals, for subalgebras of $K\langle X\rangle$, the free associative algebra over the field K in the set of indeterminates X. Following the paper [8] by Robbiano and Sweedler, we will call these bases *SAGBI bases*, where SAGBI is an abbreviation for **S**ubalgebra **A**nalog to **G**röbner **B**ases for **I**deals.

SAGBI bases theory was introduced, for commutative polynomial rings, by Kapur and Madlener (see [3]), and independently by Robbiano and Sweedler ([8]). This paper follows the approach of Robbiano and Sweedler. The subject has been studied further by e.g. Ollivier in [6], where also an implementation in a computer algebra system is discussed. See also [10] for additional references.

The SAGBI theory is strongly influenced by the theory of Gröbner bases, introduced by Bruno Buchberger in his thesis [2]. The reader familiar with Gröbner bases techniques will recognize many of Buchberger's basic ideas.

Some problems concerning subalgebras, e.g. subalgebra membership, can be reduced to Gröbner basis problems. This has been performed for commutative polynomial rings by Shannon and Sweedler ([9]), and the easy generalization to the non-commutative case can be found in [5]. We will later refer to some of these techniques. We then of course need the theory of non-commutative Gröbner bases; the standard reference here is the paper [4] by Mora.

In the commutative case, SAGBI bases differ from Gröbner bases at one essential point; they may be infinite even for finitely generated subalgebras. This implies that the SAGBI approach, in contrast to the method of Shannon and Sweedler mentioned above, is not necessarily algorithmic[1]. However, it is easy to provide examples where the SAGBI basis computation is almost free whereas it is impossible, from the practical point of view, to apply the method of Shannon and Sweedler (due to the high complexity of Gröbner bases).

Passing to non-commutative polynomial rings, Gröbner bases are no longer in general finite, and we can easily find examples where SAGBI theory solves problems that would not be algorithmic using the approach of Shannon and Sweedler.

Non-commutative SAGBI bases have also shown to be applicable for solving systems of non-commutative polynomial equations, in an approach under development by Victor Ufnarovski. (What is really needed is canonical bases for subalgebras of quotients of the polynomial ring. A SAGBI theory for such factor algebras (both commutative and non-commutative) is now developed by the author, and will appear in a later paper.)

The first section of this paper contains some definition we will use, including the definition of a SAGBI basis.

In section 2 we define the process of subalgebra reduction, and investigate the consequences thereof.

Section 3 contains the main body of the text. This section concerns the test whether a given set is a SAGBI basis. In contrast to the commutative case we do not succeed in general to make the SAGBI test for a finite set of algebra generators algorithmic, so we leave this algorithmicity as an open problem. We also show how the Hilbert series can be used in SAGBI theory for homogeneous algebras.

In section 4 we indicate how the theory of codes and monoids can be used to verify SAGBI bases in some cases.

The last section concerns the construction of a SAGBI basis from a given set of algebra generators. As in Gröbner bases theory, this completion procedure is based on the theory developed for the test in section 3. Since our SAGBI test is not algorithmic, the SAGBI construction is certainly not. However, for homogeneous subalgebras we show how the SAGBI construction method implies an algorithmic solution for the subalgebra membership problem.

The author expresses his thanks to Victor Ufnarovski for valuable suggestions and inspiring discussions.

[1] By an algorithmic problem we will here mean a problem that can be solved by a finite procedure.

1 Basic definitions and notations

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite alphabet, and let $K\langle X\rangle$ denote the free associative algebra over the arbitrary field K. Denote by W the set of all words in X, including the empty word $\mathbf{1}$ (i.e. $W = X^*$, the free monoid generated by X).

If H is a subset of $K\langle X\rangle$, then the subalgebra S of $K\langle X\rangle$ generated by H is usually denoted $K\langle H\rangle$. This notion is natural since the elements of S are precisely the polynomials in the set of formal (non-commuting) variables H, viewed as elements of $K\langle X\rangle$. The elements of $K \subset S$ correspond to the constant polynomials.

We will in this paper use $K\langle H\rangle$ to denote the set of polynomials in the H-variables. We are particularly interested in the *monomials* in $K\langle H\rangle$, i.e. the monomials in the formal variables H. Contrary to common practice, we will always assume that the coefficient of such a monomial is 1. Note that our monomials are not (in general) words, i.e. monomials in $K\langle X\rangle$.

Example 1. Consider the subset $H = \{h_1 = yx^2 + \mathbf{1}, h_2 = 2xy - y, h_3 = yx\}$ of $K\langle x, y\rangle$. An example of a monomial $m \in K\langle H\rangle$ is $m(H) = h_1h_2 = (yx^2 + \mathbf{1})(2xy - y) = 2yx^3y - yx^2y + 2xy - y$.

Remark. In commutative SAGBI theory, the role of monomials is played by exponent functions; for a monomial $h_1^{\alpha_1}h_2^{\alpha_2}\cdots h_l^{\alpha_l}$, $\alpha_i \in \mathbb{N}$, we then speak of the exponent vector $\alpha = (\alpha_1, \ldots, \alpha_l)$. Since this approach is not possible in the non-commutative case, we here use monomials.

As in Gröbner bases theory, a central point is the notion of order. We will always in what follows assume that W is given an *admissible order*, i.e. a well-order preserving multiplication: $f < g$ implies $hfk < hgk$ for all $f, g, h, k \in W$, such that the smallest word always is the unity $\mathbf{1}$. Note that an admissible order by definition is well-founded, i.e. every infinite sequence $u_1 \geq u_2 \geq \ldots u_i \geq \ldots$ in W stabilizes.

In the examples below we will use the following admissible order called *deglex* (degree lexicographical): If $|u|$ denotes the length of $u \in W$, we let $u > v$ if either $|u| > |v|$ or $|u| = |v|$ but u is larger than v lexicographically.

When we have chosen an admissible order we can, if terms with identical words are collected together using the operations over K, with every non-zero element $s \in K\langle X\rangle$ associate its *leading word* $\mathrm{LW}s$, i.e. the word in s that is larger (relative the order) than every other word occurring in s. Moreover, the *leading term* $\mathrm{LT}s$ of s is the leading word times its coefficient.

When we speak of the leading word (term) of an element in $K\langle H\rangle$, we will always mean the leading word (term) of the element viewed as an element of $K\langle X\rangle$, relative the order in $K\langle X\rangle$. Thus in Example 1 above we have, using deglex, $\mathrm{LW}m(H) = yx^3y$ and $\mathrm{LT}\, m(H) = 2yx^3y$. We also define, for a subset H of $K\langle X\rangle$, $\mathrm{LW}(H) = \{\mathrm{LW}h \mid h \in H\}$.

Following Robbiano and Sweedler we introduce the notions of height and breadth.

Definition 1 *Let $H \subset K\langle X\rangle$, and let $\sum_{i=1}^t k_im_i(H)$, $k_i \in K$, $m_i \in K\langle H\rangle$, be a sum of monomials. The* height *of the sum is $\max\{\mathrm{LW}m_i(H) | 1 \leq i \leq t\}$, where maximum is taken relative the order in $K\langle X\rangle$. The* breadth *of the sum is the number of i's such that $\mathrm{LW}m_i(H)$ is equal to the height.*

Note that the leading word of $s = \sum k_im_i(H)$ can be smaller than the height of $\sum k_im_i(H)$. This is the case if (and only if) some words larger than $\mathrm{LW}s$ cancel in the sum, and the breadth of the sum is then necessarily at least two.

Now to our main definition.

Definition 2 *Let S be a subalgebra of $K\langle X\rangle$. A subset $H \subset S$ is called a* SAGBI basis *for S if for every $s \in S$, $s \neq 0$, there exists a monomial $m \in K\langle H\rangle$ such that $\mathrm{LW}s = \mathrm{LW}m(H)$.*

Remark. Since orders are preserved after multiplication we have, if $m(H) = h_{i_1}h_{i_2}\cdots h_{i_t}$, $h_{i_j} \in H$, $\mathrm{LW}m(H) = \mathrm{LW}h_{i_1}\mathrm{LW}h_{i_2}\cdots\mathrm{LW}h_{i_t}$. Thus an equivalent formulation of the definition is that H is a SAGBI basis if the leading word of every non-zero element in S can be written as a product of leading words of elements in H.

Clearly, every subalgebra is a SAGBI basis for itself, so every subalgebra has a SAGBI basis. Another trivial example is the following.

Example 2. If H consists only of words (or terms), then H is a SAGBI basis for the subalgebra S generated by H; every word of $s = p(H) \in S$, $p \in K\langle H\rangle$, must be a product of words in $H = \mathrm{LW}(H)$, in particular the leading word of s.

The SAGBI property may depend on which order we use.

Example 3. Let $H = \{h_1 = yz, h_2 = zy, h_3 = x - y\} \in K\langle x, y, z\rangle$, and let S be the subalgebra generated by H. If we use deglex with $x > y > z$ we have $\mathrm{LW}(H) = \{yz, zy, x\}$, and H is easily seen to be a SAGBI basis by Proposition 4 below.

If we let $z > y > x$ we have $\mathrm{LW}(H) = \{yz, zy, y\}$. But $p(H) = h_1h_3 - h_3h_2 = yzx - xzy \in S$, and $\mathrm{LW}p(H) = yzx$ can not be written as a product of words in $\mathrm{LW}(H)$. We conclude that H is not SAGBI in this case.

2 Reduction

Next we define the process of *reduction*.

Reduction *The reduction of $s \in K\langle X\rangle$ over a subset $H \subset K\langle X\rangle$ is performed as follows:*

1. $s_0 = s$.
2. *If $s_i = 0$, or if there is no monomial $m \in K\langle H\rangle$ with $\mathrm{LW}s_i = \mathrm{LW}m(H)$, then terminate. In case of termination, this s_i will be referred as the* result *of the reduction.*
3. *Find a monomial $m_i \in K\langle H\rangle$ and $k_i \in K$ such that $\mathrm{LT}\, s_i = \mathrm{LT}\, k_im_i(H)$. (This is possible since we have not terminated in step 2.) Now let $s_{i+1} = s_i - k_im_i(H)$.*
4. *Go to step 2 ($i + 1 \mapsto i$).*

We note that when step 3 has been performed, the leading word of s_{i+1} is strictly smaller than the leading word of s_i (by the choice of m_i and k_i). Since the order is well-founded it then follows that the reduction always terminates after a finite number of steps.

If H is a finite set, then it is a constructive matter to determine whether a given word is a product of elements in $\mathrm{LW}(H)$, and the reduction is therefore algorithmic. This is also the case if H is infinite but sorted by e.g. the length of the leading words.

If the result (i.e. the last s_i) of a reduction of s over H is denoted $\overline{s^H}$, and if the reduction terminated after t iterations of step 3 above, then we have, with the notations above,

$$s = \sum_{i=0}^{t-1} k_i m_i(H) + \overline{s^H}. \tag{1}$$

If $t = 0$ the right hand side of (1) is of course just $\overline{s^H}$. But if step 3 is performed at least once we have, by above,

$$\mathrm{LW}s = \mathrm{LW}m_0(H) > \mathrm{LW}m_1(H) > \ldots > \mathrm{LW}m_{t-1}(H), \tag{2}$$

so the sum in the right hand side of (1) is clearly of breadth one and height equal to $\mathrm{LW}s$. *We will use these facts several times in the sequel* (in particular when $\overline{s^H} = 0$).

There may of course be several different possibilities to choose the m_i's in step 3, so the result of the reduction depends (in general) on how we choose these monomials.

We will be particularly interested in the case when $\overline{s^H} = 0$ above. We say that *s reduces to zero weakly over H* if there exists one reduction (i.e. one choice of the m_i's) with $\overline{s^H} = 0$, and that *s reduces to zero strongly over H* if every reduction (every choice) yields $\overline{s^H} = 0$. However, in most cases it does not matter which formulation we use, and we will then simply say that *s reduces to zero over H*. Note that, by definition, $s = 0$ reduces to zero.

As a first application of SAGBI bases we mention the *Subalgebra Membership Problem*, i.e. the decision whether an element $s \in K\langle X\rangle$ is in a given subalgebra.

Proposition 1 *Let $H \subset K\langle X\rangle$ be a SAGBI basis for the subalgebra S, and let $s \in K\langle X\rangle$. Then $s \in S$ if and only if s reduces to zero over H.*

Proof. First assume $s \in S$. By (1) we have $\overline{s^H} = s - \sum_{i=1}^{t-1} k_i m_i(H)$. Since every $m_i(H) \in S$ it follows that $\overline{s^H} \in S$. The only possibility of termination in step 2 is $\overline{s^H} = 0$. Otherwise, if $\overline{s^H} \neq 0$ there is, by Definition 2, a monomial $m \in K\langle H\rangle$ with $\mathrm{LW}\overline{s^H} = \mathrm{LW}m(H)$, and this contradicts the termination.

Conversely, it follows directly from (1) that $\overline{s^H} = 0$ implies $s \in S$. $\square$

Corollary 1 *If H is a SAGBI basis for the subalgebra S, then H generates S.*

Proof. Clear from Proposition 1 and (1). $\square$

In view of Corollary 1 we may (and will) simply say that $H \subset K\langle X\rangle$ is a SAGBI basis, meaning that H is a SAGBI basis for the subalgebra of $K\langle X\rangle$ generated by H.

We conclude this section with a few remarks. Our definition of reduction is sufficient for our purposes; Proposition 1 is what we will need. However, an arbitrary element of $K\langle X\rangle$ has not in general a unique result of reduction over a SAGBI basis H. The reader familiar with Gröbner bases techniques can easily do the changes necessary for uniqueness: what is needed is to modify step 2 above. Instead of terminating when the leading term no longer can be written as a product of $\mathrm{LW}(H)$, we move this term to some kind of remainder (generalizing $\overline{s^H}$) and continue with the other terms.

This modification also allows us to do computations concerning equivalence classes modulo subalgebras of $K\langle X\rangle$; we then have $f \equiv g$ ($f - g \in S$) if and only if (with the notations above) $\overline{f^H} = \overline{g^H}$.

3 SAGBI basis test

The following proposition gives a first method to test for the SAGBI basis property.

Proposition 2 *$H \subset S$ is a SAGBI basis for the subalgebra S if and only if every $s \in S$ reduces to zero over H.*

Proof. If H is a SAGBI basis, then every $s \in S$ reduces to zero by Proposition 1.

Conversely, let $s \in S$, $s \neq 0$, be arbitrary. Since s reduces to zero, reduction yields $s = \sum k_i m_i(H)$, where at least $k_0 \neq 0$. By (2) we have $\mathrm{LW}s = \mathrm{LW}m_0(H)$, so H is a SAGBI basis by Definition 2. $\square$

The following definition and proposition generalize one of the cornerstones in Buchberger theory. It allows us to reduce the number of elements we need to consider for the SAGBI basis test.

Definition 3 *Let H be a subset of $K\langle X\rangle$. A* critical pair *(m, m') of H is a pair of monomials $m, m' \in K\langle H\rangle$ with $\mathrm{LW}m(H) = \mathrm{LW}m'(H)$. If $k \in K$ is such that $\mathrm{LT}\,m(H) = \mathrm{LT}\,km'(H)$ we define the T-*polynomial[2] *of (m, m') as $T(m, m') = m(H) - km'(H)$.*

Remark. The constant is chosen such that the leading words cancel in $T(m, m')$, and thus we have $\mathrm{LW}T(m, m') < \mathrm{LW}m(H) = \mathrm{LW}m'(H)$. It is also clear that $T(m, m') = k'T(m', m)$ for some $k' \in K$. It follows that $T(m, m')$ reduces to zero weakly (alternatively strongly) if and only if $T(m', m)$ does.

Proposition 3 *H is a SAGBI basis if and only if the T-polynomials of all critical pairs of H reduce to zero over H.*

Proof. Let S be the subalgebra generated by H.

If H is a SAGBI basis, then, since every $T(m, m') \in S$, $T(m, m')$ reduces to zero by Proposition 1

Conversely, take arbitrary non-zero $s \in S$. It is sufficient to show that s can be written as a sum $\sum k_i m_i(H)$ ($k_i \in K$, $m_i \in K\langle H\rangle$ monomials) of height $\mathrm{LW}s$; then there exists (at least) one monomial m_j in the sum with $\mathrm{LW}m_j(H) = \mathrm{LW}s$.

Assume contrarily that no such sum exists. Then every representation $s = \sum_{i=1}^{t} k_i m_i(H)$ (we have at least one since $s \in S$) has height higher than s, and the breadth must be at least two (since some leading terms must cancel). Now choose, among all such representations, one with primarily minimal height, and secondarily minimal breadth for that height.

By renumbering we can assume that $\mathrm{LW}m_1(H) = \mathrm{LW}m_2(H) =$ the height of the sum. If $T(m_1, m_2) =$

[2] The Gröbner bases term is *S-polynomial*. We here use the T since the counterpart of a critical pair in Robbiano and Sweedler is called a *tête-a-tête*.

$m_1(H) - km_2(H)$ we can write

$$s = \sum_{i=1}^{t} k_i m_i(H) =$$

$$= k_1(m_1(H) - km_2(H)) + (k_2 + k_1 k)m_2(H) + \sum_{i=3}^{t} k_i m_i(H) =$$

$$= k_1 T(m_1, m_2) + (k_2 + k_1 k)m_2(H) + \sum_{i=3}^{t} k_i m_i(H). \quad (3)$$

Since $T(m_1, m_2)$ reduces to zero, $T(m_1, m_2)$ is either zero or can be written as a sum of monomials of height $\mathrm{LW}T(m_1, m_2) < \mathrm{LW}m_1(H) = \mathrm{LW}m_2(H)$. If $k_2 + k_1 k = 0$ and the breadth of our original sum was two, then the right hand side of (3) is a representation of lower height than before. Otherwise the height is preserved, but we have decreased the breadth by (at least) one. In both cases, some of our two minimality conditions is contradicted. □

We can now prove the following proposition which gives a sufficient condition for a set to be a SAGBI basis. The simple proof gives a first indication of what problems concerning leading words we are dealing with.

Proposition 4 *Let $H \subset K\langle X\rangle$ be such that $\mathrm{LW}h_i \neq \mathrm{LW}h_j$ if $h_i \neq h_j$, $h_i, h_j \in H$. If either no word in $\mathrm{LW}(H)$ is a prefix (proper left factor) of some other word in $\mathrm{LW}(H)$, or no word in $\mathrm{LW}(H)$ is a suffix (proper right factor) of some other, then H is a SAGBI basis.*

Proof. Assume first that no word is a prefix of some other, and let (m, m') be an arbitrary critical pair of H. By the remark after Definition 2 we can write $\mathrm{LW}m(H) = u_1 u_2 \ldots u_s$, $\mathrm{LW}m'(H) = u'_1 u'_2 \ldots u'_t$ with $u_i, u'_j \in \mathrm{LW}(H)$, and (since (m, m') is a critical pair) $u_1 u_2 \ldots u_s = u'_1 u'_2 \ldots u'_t$. We here have $u_1 = u'_1$, for otherwise one of them must be a prefix of the other. By the same argument we have $u_2 = u'_2$ and, by induction, $s = t$ and $u_s = u'_t$. We conclude that $\mathrm{LW}m(H)$ and $\mathrm{LW}m'(H)$ are "built up" by the same leading words, and since there are no different elements in H with the same leading word, $m(H)$ and $m'(H)$ are built up by the same elements of H, i.e. $m(H) = km'(H)$ for some $k \in K$. Thus $T(m, m') = 0$, and H is a SAGBI basis by Proposition 3.

The case when no word is a suffix of some other is treated similarly. The only difference is that we now have to start comparing leading words of H at the end of $\mathrm{LW}m(H)$ and $\mathrm{LW}m'(H)$. □

In order to further reduce the elements to be considered for the SAGBI basis test, we need another version of Proposition 3. In the proof of the sufficiency we only used that every $T(m, m')$ either was zero, or could be written as a sum of monomials of height less than $\mathrm{LW}m(H) = \mathrm{LW}m'(H)$. We will say, sometimes without mentioning the corresponding m and m', that such a T-polynomial *admits a low representation.* Since this property of the $T(m, m')$'s is clear if H is a SAGBI basis (by reduction), we have the following proposition.

Proposition 5 *H is a SAGBI basis if and only if every T-polynomial of H admits a low representation.* □

In view of Proposition 5, the following proposition reduces the number of critical pairs to be considered.

Proposition 6 *Let (m, m') be a critical pair of H, and assume there are factorizations $m = m_1 m_2, m' = m'_1 m'_2$, m_1, m_2, m'_1, m'_2 monomials in $K\langle H\rangle$ with $\mathrm{LW}m_i(H) = \mathrm{LW}m'_i(H)$, $i = 1, 2$. If $T(m_i, m'_i)$ admits a low representation, $i = 1, 2$, then also $T(m, m')$ admits a low representation.*

Proof. If $T(m, m') = m - km'$ $(= m(H) - km'(H))$ and $T(m_i, m'_i) = m_i - k_i m'_i$, $i = 1, 2$, then, by comparing leading coefficients, it is easily seen that $k = k_1 k_2$. We can thus write

$$\begin{aligned} T(m, m') &= m_1 m_2 - k_1 m'_1 k_2 m'_2 = \\ &= (m_1 - k_1 m'_1)m_2 + k_1 m'_1(m_2 - k_2 m'_2) = \\ &= T(m_1, m'_1)m_2 + k_1 m'_1 T(m_2, m'_2). \quad (4) \end{aligned}$$

Since $T(m_1, m'_1)$ and $T(m_2, m'_2)$ both admit low representations, we see from the right hand side of (4) that the same is true for $T(m, m')$. □

By induction, the result in the proposition is valid for any number of factors, i.e. for any factorizations $m = m_1 \cdots m_t$, $m' = m'_1 \cdots m'_t$ with $\mathrm{LW}m_i(H) = \mathrm{LW}m'_i(H)$ for all i. The only critical pairs necessary for the SAGBI test are therefore those which can not be factored is the sense above. This is the case only for the critical pairs of form (h, h'), $h, h' \in H$ with $\mathrm{LW}h = \mathrm{LW}h'$, and pairs (m, m'), $m = h_{i_1} \cdots h_{i_s}$, $m' = h'_{i_1} \cdots h'_{i_t}$, s or (and) $t > 1$, where the leading words of the factors overlap, i.e. the pairs with leading words of form

$$\begin{array}{lcl} \mathrm{LW}m & = & \overbrace{\qquad}^{\mathrm{LW}h_{i_1}}\ \overbrace{\qquad\qquad}^{\mathrm{LW}h_{i_2}} - \cdots - \overbrace{\qquad}^{\mathrm{LW}h_{i_s}} \\ \mathrm{LW}m' & = & \underbrace{\qquad\qquad}_{\mathrm{LW}h'_{i_1}}\ \underbrace{\qquad}_{\mathrm{LW}h'_{i_2}} \ \text{——} \cdots \text{——}\ \underbrace{\qquad}_{\mathrm{LW}h'_{i_t}} \end{array} . \quad (5)$$

We will call a critical pair with leading words of form (5) an *overlapping pair.* To find all such pairs we can proceed as follows. The reader is advised to keep (5) in mind.

An overlapping pair must begin with two elements of H where one of the leading words is a prefix of the other. If these words are u_1 and $u_2 = u_1 v$ we get the *overlap* v. We must then find all possibilities to continue on v. This could be with a leading word equal to v (in which case we have obtained a critical pair), with a prefix of v, or finally with a word of which v is a prefix. In the two latter cases we get new overlaps that might be continued. If we get an overlap we have obtained before (starting with u_1, u_2 or a previous pair) we need of course not continue with this; we have already examined how it can be continued. As mentioned before, we get critical pairs whenever a leading word fits on an overlap.

If H is a finite set, then the process above is algorithmic. This rests on the fact that there only can be a finite number of different overlaps shorter than a given length. However, there may be an infinite number of overlapping pairs. This is the case if (and only if) the pair of leading words of some overlapping pair contains a segment beginning and ending with the same overlap. We will call such a segment a *loop.*

We illustrate with a few examples.

Example 4. In Example 1 we have, using deglex, $\mathrm{LW}(H) = \{yx^2, xy, yx\}$. The only possible beginning is $\dfrac{\overline{yx}}{\underline{yxx}}$, and we can continue only with $\dfrac{\overline{yx}\,\overline{xy}}{\underline{yxx}}$. (We here use the bars to separate leading words.) Now we have two possibilities. The first is $\frac{\overline{yx}\,\overline{xy}}{\underline{yxx}\,\underline{yxx}}$ which can not be continued further, and the other is $\frac{\overline{yx}\,\overline{xy}}{\underline{yxx}\,\underline{yx}}$. But we have already checked the overlap x, and the only way to continue is the loop $\frac{\overline{xy}}{\underline{yx}}$ which can not be ended to give a pair of form (5). (We can of course also consider $\frac{\overline{yx}}{\underline{xy}}$ as a loop.) Thus there are no overlapping pairs, and we conclude that H is a SAGBI basis. (This was of course clear by the suffix condition in Proposition 4.)

Example 5. We now add one element to H in Example 1 to get $H = \{h_1 = yx^2 + \mathbf{1}, h_2 = 2xy - x, h_3 = yx, h_4 = y\}$, which yields $\mathrm{LW}(H) = \{yx^2, xy, yx, y\}$. We then have $\frac{\overline{yx}\,\overline{xy}}{\underline{yxx}\,\underline{y}}$, and also $\frac{\overline{yx}\cdots\overline{xy}}{\underline{yxx}\cdots\underline{y}}$, where we can insert an arbitrary number of the loop $\frac{\overline{xy}}{\underline{yx}}$. We must now also check $\frac{\overline{y}}{\underline{yxx}}$ (harmless) and $\frac{\overline{y}}{\underline{yx}}$, which gives, since we already checked the overlap x, $\frac{\overline{y}\cdots\overline{xy}}{\underline{yx}\cdots\underline{y}}$ (with the same loop as above). It seems that we have to consider an infinite number of critical pairs. But, as we will see below, it is sufficient to calculate $T(h_3h_2, h_1h_4) = -h_1 - 2h_4 + 1$ and $T(h_4h_2, h_3h_4) = -h_3$. These are representations as in Proposition 5 and 6, and it follows that H is a SAGBI basis.

The fact that saves us from an infinite procedure in Example 5 is that every loop contains the overlap y, and $y \in \mathrm{LW}(H)$. We can therefore use induction on the next proposition. This proposition deals with critical pairs (m, m') where the leading words contain an overlap which is equal to $\mathrm{LW}\tilde{m}$ for some monomial $\tilde{m} \in K\langle H\rangle$ ($\tilde{m}(H) = h_4$ in Example 5), i.e. when we have

$$\begin{array}{rcl} \mathrm{LW}m & = & \overbrace{\hphantom{xxxxxx}\underbrace{\hphantom{xx}}_{\mathrm{LW}\tilde{m}}}^{\mathrm{LW}m_1}\overbrace{\hphantom{xxxxx}}^{\mathrm{LW}m_2} \\ \mathrm{LW}m' & = & \underbrace{\hphantom{xxxxxx}}_{\mathrm{LW}m'_1}\underbrace{\overbrace{\hphantom{xx}}^{\mathrm{LW}\tilde{m}}\hphantom{xxxxx}}_{\mathrm{LW}m'_2} \end{array} . \qquad (6)$$

Proposition 7 *Let (m, m') be a critical pair of H, and assume there are factorizations such as in (6). If $T(m_1, m'_1\tilde{m})$ and $T(\tilde{m}m_2, m'_2)$ both admit low representations, then also $T(m, m')$ admits a low representation.*

Proof. If $T(m, m') = m - km'$, $T(m_1, m'_1\tilde{m}) = m_1 - k_1m'_1\tilde{m}$, $T(\tilde{m}m_2, m'_2) = h_im_2 - k_2m'_2$, then, again by comparing leading coefficients, $k = k_1k_2$, and thus

$$\begin{aligned} T(m, m') &= m_1m_2 - k_1m'_1k_2m'_2 = \\ &= (m_1 - k_1m'_1\tilde{m})m_2 + k_1m'_1(\tilde{m}m_2 - k_2m'_2) = \\ &= T(m_1, m'_1\tilde{m})m_2 + k_1m'_1T(\tilde{m}m_2, m'_2). \end{aligned} \qquad (7)$$

The result now follows as in Proposition 6. □

By Proposition 7 we need not, in the search for overlapping pairs above, on an overlap continue the "chain" of leading words if we can find a leading word, or a product of leading words, that fits on the overlap; in view of the factorization in (6) we will still get all the overlapping pairs necessary for the SAGBI test.

If H is a finite set, and if all loops, just as in Example 5, can be factored in the sense of (6), then the number of necessary critical pairs is finite. Together with the reduction algorithm we then get an algorithmic SAGBI test (see Theorem 1 below).

But if we conclude that there is some loop to which we not can apply Proposition 7, then it seems that we have consider the infinite number of critical pairs this loop causes. It would be desirable to have results available to handle also these loops, but in lack of that the following remains unsolved here.

Problem *For a finite set H, is the SAGBI basis test in general algorithmic?*

Anyhow, we summarize:

Theorem 1 (Test) *A subset $H \subset K\langle X\rangle$ is a SAGBI basis if and only if the T-polynomials of all the necessary critical pairs described above reduce to zero over H.*

Proof. The sufficiency follows from Proposition 5, 6 and 7 since reduction yields the monomial representations in these propositions.

Conversely, since all T-polynomials are in the subalgebra generated by H, they reduce to zero by Proposition 1. □

If all elements of H are homogeneous, then the homogeneous subalgebra S generated by H breaks up into the direct sum $S = \bigoplus_{i=0}^{\infty} S_i$, where S_j is the homogeneous component of degree j. We can then use the *Hilbert series* $h_S(t) = \sum_{i=0}^{\infty}(\dim_K S_i)\, t^i$ to test for the SAGBI property.

It is well known (and easy to prove) that S (homogeneous) and the K-span of $\mathrm{LW}(S)$ (always homogeneous) have equal Hilbert series. An equivalent formulation of the SAGBI basis definition is that $H \subset S$ is a SAGBI basis for S if $\mathrm{LW}(S)$ is equal to $\mathrm{LW}(H)^*$ (the monoid generated by $\mathrm{LW}(H)$), and this is clearly equivalent to the fact that the algebra generated by $\mathrm{LW}(H)$ is equal to the K-span of $\mathrm{LW}(S)$. Conversely, if these two algebras have the same Hilbert series, then the algebras must be equal (since the first is included in latter). Summarizing we get:

Proposition 8 *H is a SAGBI basis if and only if the subalgebras generated by H and $\mathrm{LW}(H)$ have equal Hilbert series.* □

Corollary 2 *If a homogeneous subalgebra has a finite SAGBI basis (in some order), then its Hilbert series is rational.*

Proof. Let H be a finite SAGBI basis. The subalgebra generated by $\mathrm{LW}(H)$ has the same Hilbert series as the regular language $\mathrm{LW}(H)^*$. This series is well-known to be rational, and it can be calculated (algorithmically) using an automaton (compare [11]). □

Assume now that H is a finite homogeneous set generating the algebra S. By Gröbner bases techniques described in [5] we can find (theoretically) an isomorphic factor algebra $K\langle Y\rangle/I$ (I two-sided ideal), and the Hilbert series of this algebra can be calculated by various software packages, e.g. ANICK ([7]) by Podoplelov and Ufnarovski.

Example 6. Let us consider the subalgebra S of $K\langle x_1, x_2\rangle$ generated by $H = \{x_2 + x_1, x_1x_2 + x_1^2, x_2^2 + x_2x_1\}$. S is isomorphic to $K\langle y_1, y_2, y_3\rangle/\langle y_3 + y_2 - y_1^2\rangle$, where we let y_1 have degree one, y_2 and y_3 degree two. The algebra generated by $\mathrm{LW}(H) = \{x_2, x_1x_2, x_2^2\}$ ($x_1 < x_2$) is isomorphic to $K\langle y_1, y_2, y_3\rangle/\langle y_3 - y_1^2\rangle$ (with the same grading on the y_i's). Both these factor algebras have the Hilbert series $h(t) = (1 - t - 2t^2 + t^3 + t^4)^{-1}$, so we conclude that H is a SAGBI basis.

Finally we mention that if, for some set of generators H, the two Hilbert series above differ (i.e. H is not a SAGBI basis), then the series can be used to detect in which homogeneous component we have critical pairs not reducing to zero. This can be used to avoid unnecessary work in the SAGBI basis construction described in section 5.

4 Connections with codes and monoids

Recall that $H \subset S$ is a SAGBI basis for S if (and only if) $\mathrm{LW}(H)$ generates $\mathrm{LW}(S)$ as a monoid.

In the commutative case we can use the theory of monoids e.g. to develop an algorithmic SAGBI test, and to obtain certain finiteness conditions (see Robbiano and Sweedler). Non-commutative monoids do not allow such a rich theory. However, there are results concerning the existence of overlapping pairs.

A set V of words is called a *code* if for all s, t and $v_1, \ldots, v_s, v_1', \ldots, v_t' \in V$, the condition $v_1v_2\cdots v_s = v_1'v_2'\cdots v_t'$ implies $s = t$ and $v_i = v_i'$ for all i. The fact that V is a code is equivalent to the property that the monoid V^* generated by V is free.

The condition in the definition of a code is exactly what we deduced in the proof of Proposition 4, so if V satisfies the prefix or suffix condition in this proposition, then V is a code.

It is clear that the property of $\mathrm{LW}(H)$ being a code is equivalent to the absence of overlapping pairs. We thus get the following version of the Test theorem.

Proposition 9 *Let H be a subset of $K\langle X\rangle$. If $\mathrm{LW}(H)$ is a code ($\mathrm{LW}(H)^*$ is a free monoid), and if $T(h_i, h_j)$ reduces to zero for all $h_i, h_j \in H$ with $\mathrm{LW}h_i = \mathrm{LW}h_j$, then H is a SAGBI basis.* $\square$

The method of finding overlapping pairs in the previous section will of course determine whether $\mathrm{LW}(H)$ is a code ($\mathrm{LW}(H)^*$ is a free monoid) (cf. Example 4 and 5). We now present a few results that may simplify this matter. For the proofs, and for more results, we refer to [1].

Proposition 10 *If a set V of words (over our alphabet X) is a code, then any morphism $\varphi : Y^* \to X^*$ which induces a bijection of some alphabet Y onto V is injective. Conversely, if there exists an injective morphism $\varphi : Y^* \to X^*$ such that $V = \varphi(Y)$, then V is a code.* $\square$

To decide whether $\mathrm{LW}(H)$ is a code it therefore suffices to calculate the kernel of the morphism defined by $\varphi(y_i) = \mathrm{LW}h_i$ for each $h_i \in H$. This can be done with Gröbner bases methods as shown in [5].

Example 7. With $\mathrm{LW}(H)$ in Example 4 we get $\ker\varphi = \{0\}$, so we conclude (again) that $\mathrm{LW}(H)$ is a code.

In Example 5 we have (after renaming the variables) $\mathrm{LW}(H) = \{\mathrm{LW}h_1 = x_2x_1^2, \mathrm{LW}h_2 = x_1x_2, \mathrm{LW}h_3 = x_2x_1, \mathrm{LW}h_4 = x_2\}$. A (reduced) Gröbner basis for $\ker(\varphi)$, φ defined as above, is $\{y_3y_2 - y_1y_4, y_4y_2 - y_3y_4\}$, and these elements correspond to the overlapping pairs $\begin{array}{l}\overline{x_2x_1}\,\overline{x_1x_2}\\ \overline{x_2x_1x_1}\,\overline{x_2}\end{array}$ resp. $\begin{array}{l}\overline{x_2}\,\overline{x_1x_2}\\ \overline{x_2x_1}\,\overline{x_2}\end{array}$.

Remark. To find all overlapping pairs we would like to get, in the latter case above, $\{y_3y_2^sy_2 - y_1y_3^sy_4, y_4y_2^ty_2 - y_3y_3^ty_4 | s, t = 0, 1, 2, \ldots\}$, the powers corresponding to the loops. In Example 7 we used an order with $y_4 > y_3 > y_2 > y_1$, so e.g the leading word y_3y_2 is a subword of the leading word $y_3y_2y_2$, and the latter will therefore not appear in the Gröbner basis. However, if we let $y_1 > y_2 > y_3 > y_4$ we get the basis $\{y_1y_3^sy_4 - y_3y_2^sy_2, y_3y_4 - y_4y_2 | s = 0, 1, 2, \ldots\}$, and here we found one of the loops. Anyhow we conclude that the Gröbner basis method is not sufficient for finding all critical pairs.

A case when Proposition 7 always is applicable is when the monoid $\mathrm{LW}(H)^*$ is *stable*, i.e. when $u, v, uw, wv \in \mathrm{LW}(H)^*$ implies $w \in \mathrm{LW}(H)^*$; every overlap must then clearly be of form (6). However, in these cases we need not check any overlapping pairs thanks to the following.

Proposition 11 *A monoid is stable if and only if it is free.* $\square$

Finally we mention that the test for the code property is algorithmic for an infinite but recognizable set.

5 SAGBI basis construction

As in Buchberger theory, the Test theorem is the key to the construction of SAGBI bases.

If H fails to be a SAGBI basis, then for at least one of the T-polynomials $T(m, m')$ in the Test theorem, a result $\overline{T(m, m')}$ of reduction (over H) is not zero. If we to H add all $\overline{T(m, m')}$, $\overline{T(m, m')} \neq 0$, then all the T-polynomials clearly reduce to zero (weakly) over this new set H'. But H' may contain new critical pairs, and then there are new T-polynomials to check. This suggests the following algorithm.

SAGBI basis construction *Let H be a subset of $K\langle X\rangle$.*

1. $H_0 = H$.
2. *Use the methods in section 3 to find the set M_i of all necessary critical pairs (m, m') of H_i.*
3. $H_{i+1} = H_i \bigcup \{\overline{T(m, m')} \mid (m, m') \in M_i, \overline{T(m, m')} \neq 0\}$. Here $\overline{T(m, m')}$ denotes a result of reduction over H_i.
4. *If $H_{i+1} \neq H_i$ then go to step 2 ($i + 1 \mapsto i$).*
5. $H_\infty = \bigcup H_i$.

Proposition 12 *H_∞ is a SAGBI basis for the subalgebra S generated by H.*

<u>*Proof.*</u> If (m, m') is a critical pair of $H = H_0$, then $\overline{T(m, m')} \in S$, so $H_1 \subset S$. By induction we have every $H_i \subset S$ and thus $H_\infty \subset S$, so H and H_∞ generate the same subalgebra.

If $H_i = H_{i+1}$ for some i, then H_i is a SAGBI basis (by the test theorem). Otherwise, let (m, m') be a critical pair of H_∞. Take j so large that every element of H_∞ occurring in m and m' is in H_j. It is then clear that $T(m, m')$ reduces to zero (weakly) over H_{j+1}, and thus of course also over H_∞. □

In general the algorithm will not stop, and we have seen in section 3 that step 2 may not be algorithmic even for a finite set H_i. We conclude that the construction algorithm is mostly of theoretical value.

However, if H is a set of homogeneous elements, then the situation is a bit more promising. It is then clear that all monomials $m(H)$ are homogeneous, and thus also all T-polynomials. Since reduction over H involves only homogeneous elements, every T-polynomial will reduce to a homogeneous element (or zero). It follows that H_1 must be homogeneous, and by induction every H_i (and H_∞).

If, moreover, H has only a finite number of element of degree (word length) less than or equal to $d \in \mathbb{N}$ (and we can find all these elements), then we can algorithmically obtain a "partial" SAGBI basis $H_{(d)} = \{h \in H_\infty | \deg h \leq d\}$ as follows.

Performing step 2 and 3 above we need only consider critical pairs of degree $\leq d$, because critical pairs of degree $> d$ can only reduce to new elements of degree $> d$ (or to zero). It is clear that the results $\overline{T(m, m')} \neq 0$ of reduction over H_i all have leading words not lying in $\mathrm{LW}(H_i)$ (since the reduction algorithm terminated). Since there are only a finite number of different words of length $\leq d$, we must sooner or later get $\overline{T(m, m')} = 0$ for all T-polynomials of degree $\leq d$ in step 3, and the current H_i is then our requested $H_{(d)}$.

To perform the reduction of an arbitrary element of $K\langle X \rangle$ over some subset, we need of course in each step of the reduction only know this subset up to a certain degree. We conclude that if H consists of homogeneous elements (with the finiteness condition above), then the Subalgebra Membership Problem (Proposition 1) is algorithmic.

References

[1] Berstel, J., and Perrin, D. *Theory of codes*, vol. 117 of *Pure and Applied Mathematics*. Academic Press Inc., Orlando, Fla., 1985.

[2] Buchberger, B. *On finding a vector space basis of the residue class ring modulo a zero dimensional polynomial ideal (German)*. PhD thesis, University of Innsbruck, Austria, 1965.

[3] Kapur, D., and Madlener, K. A completion procedure for computing a canonical basis for a k-subalgebra. In *Computers and mathematics (Cambridge, MA, 1989)*. Springer, New York, 1989, pp. 1–11.

[4] Mora, F. Groebner bases for noncommutative polynomial rings. In *Algebraic algorithms and error correcting codes (Grenoble, 1985)*, vol. 229 of *Lecture Notes in Comput. Sci.* Springer, Berlin, 1986, pp. 353–362.

[5] Nordbeck, P. On some basic applications of Gröbner bases in noncommutative polynomial rings. In *Gröbner bases and applications*, vol. 251 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, Cambridge, 1993, pp. 463–472.

[6] Ollivier, F. Canonical bases: relations with standard bases, finiteness conditions and application to tame automorphisms. In *Effective methods in algebraic geometry (Castiglioncello, 1990)*, vol. 94 of *Progr. Math.* Birkhäuser Boston, Boston, MA, 1991, pp. 379–400.

[7] Podoplelov, A., and Ufnarovski, V. ANICK, C code, 1997. Available by anonymous ftp from ftp.riscom.net in the directory: /pub/anick.

[8] Robbiano, L., and Sweedler, M. Subalgebra bases. In *Commutative algebra (Salvador, 1988)*, vol. 1430 of *Lecture Notes in Math.* Springer, Berlin, 1990, pp. 61–87.

[9] Shannon, D., and Sweedler, M. Using Gröbner bases to determine algebra membership, split surjective algebra homomorphisms determine birational equivalence. *J. Symbolic Comput. 6*, 2-3 (1988), 267–273. Computational aspects of commutative algebra.

[10] Sturmfels, B. *Gröbner bases and convex polytopes*, vol. 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.

[11] Ufnarovski, V. Combinatorial and asymptotic methods in algebra [MR 92h:16024]. In *Algebra, VI*, vol. 57 of *Encyclopaedia Math. Sci.* Springer, Berlin, 1995, pp. 1–196.