Gröbner Bases Computation Using Syzygies*

H. Michael Möller† Teo Mora‡ Carlo Traverso§

Abstract

For investigating multivariate polynomials over fields, Gröbner bases are an important tool. Therefore the main Computer Algebra Systems contain procedures for computing such ideal bases. These algorithms typically spend much time in reducing so called "S-polynomials" to 0. One of Buchberger's two criteria for avoiding superfluous reductions to 0 was interpreted in paper by Gebauer & Möller (1988) as a criterion for finding in a generating system of particular syzygies redundant ones. In the present paper this idea is extended. While the Gröbner basis is constructed, a set of syzygies is updated and an "S-polynomial" is not considered if its corresponding syzygy turns out to depend on the already known syzygies. Using this concept, we cover both of Buchberger's criteria, are able to include informations on polynomials (i.e. syzygies) either obtained from input or from intermediate calculations, and avoid more superfluous reductions. We illustrate this procedure by two examples.

1 Introduction

Among ideal bases, Gröbner bases are especially well suited for constructively solving problems in Commutative Algebra and some areas of applications, cf. for instance Buchberger (1985). The standard algorithm for computing these ideal bases is Buchberger's. It computes, whenever a (finite) ideal basis F of a polynomial ideal I is given, a Gröbner basis for I by calculating so called S-polynomials, by reducing them modulo F as long as possible, and by inserting these irreducible polynomials into F if they are nonzero. At termination, F is a Gröbner basis of I.

The most time consuming part of the algorithm is the reduction of the S-polynomials. If an S-polynomial reduces modulo F to the zero polynomial, its computation and reduction give no contribution to the final Gröbner basis. Such zero reduction can be avoided, if a criterion is available, which predicts that the S-polynomial reduces to 0. Two different kinds of criteria for avoiding many of these zero reductions are developped by Buchberger and resumed in Buchberger (1985). A practical implementation and an interpretation of one kind of Buchberger's criteria in terms of syzygies are given by Gebauer & Möller (1988). Although the criteria avoid many zero reductions, many still remain undetected (especially near to the termination of the algorithm).

The crucial idea for involving syzygies in the detection of superfluous zero reductions can be easily understood by using an observation by Lazard (1983). Since an ideal can be considered as an infinite dimensional vector space, an ideal possesses simultaneously ideal bases and Gröbner bases. Lazard pointed out the connection between Gröbner bases and linear bases of the same ideal. Buchberger's algorithm can be considered as a triangularisation of a linear basis by Gaussian elimination. The reduction of a polynomial to zero means a linear dependence of this polynomial on the polynomials employed in the reduction procedure. Since all polynomials are here of type $t_i f_i$, $f_i \in F$, t_i a power product, the linear dependence relation can be written as a syzygial relation

$$\sum_{f_i \in F} g_i f_i = 0,$$

where the g_i are linear combinations of some power products t_{ij}, i. e. they are polynomials. (In this connection the vector $(g_1, \ldots, g_s)$ is called a syzygy w. r. t. $(f_1, \ldots, f_s)$, if $F = \{f_1, \ldots, f_s\}$.)

In this paper, we describe the connection between linear bases, Gröbner bases, and syzygies. Interpreting every S-polynomial reduction to 0 as syzygy, we can predict every zero reduction provided we know a suitable

*This research was made in the preliminary part of the research financed with the CEC Basic Research ESPRIT contract n. 6846 POSSO. The authors acknowledge the support of this contract in the final preparation part.

†FernUniversität Hagen, Fachbereich Mathematik, Lützowstrasse 125, HAGEN, `ma105@dhafeu11.bitnet`

‡Università di Genova, Dipartimento di Matematica, Via L. B Alberti 4, I-16132 GENOVA, `theomora@igecuniv.bitnet`

§Università di Pisa, Dipartimento di Matematica, Via Buonarroti 2, I-56127 PISA, `traverso@dm.unipi.it`

basis of the module of syzygies. This gives a unifying interpretation of both kinds of Buchberger's criteria. In addition, it enables us to use informations as known syzygies among input polynomials for shortcutting Buchberger's algorithm. The algorithm we developped is first sketched in terms of linear bases. It calculates a linear basis for the ideal from which a Gröbner basis can be read off and a Gröbner basis for the module of syzygies. The second more detailed version of the algorithm calulates a Gröbner basis for the ideal and a Gröbner basis for the module of syzygies w.r.t. this Gröbner basis. We discuss variants in which only parts of the whole module of syzygies is computed and present two detailed examples.

2 Gröbner bases and syzygies

2.1 Let $\mathcal{P}$ denote the ring of polynomials in $X_1, \ldots, X_n$ over a field $I\!\!K$. The set T of terms (power products) in $X_1, \ldots, X_n$ is as usual ordered by a multiplicatively compatible ordering $<_T$ with $1 \leq_T t$ for all $t \in T$. For nonzero polynomials $f \in \mathcal{P}$ we denote the leading term by $lt(f)$, its corresponding coefficient by $lc(f)$, and its maximal form by $M_T(f) := lc(f)\, lt(f)$. For convenience $lt(0) := M_T(0) := 0$ and $0 <_T t$ for all $t \in T$.

2.2 For $F := \{f_1, \ldots, f_r\} \subset \mathcal{P}$ the reduction modulo F is defined recursively by

$$\begin{aligned} & f \underset{F}{\longrightarrow} g \iff \\ & \quad \exists f_i \in F \exists h \in \mathcal{P} : g = f - h f_i \ , \ lt(g) <_T lt(f) \\ & \quad \text{or } \Big((M_T(f) = M_T(g) \ , \ f - M_T(f) \underset{F}{\longrightarrow} g - M_T(g) \Big) . \end{aligned}$$

Denoting by $\underset{F}{\longrightarrow}^*$ the reflexive transitive closure of $\underset{F}{\longrightarrow}$, then $f \underset{F}{\longrightarrow}^* g$ implies that there are polynomials $g_i \in \mathcal{P}$ with $f = g + \sum g_i\, f_i$ and $lt(g_i)\, lt(f_i) \leq_T lt(f)$. Especially if $f \underset{F}{\longrightarrow}^* 0$, then f has a *T-representation* $f = \sum g_i\, f_i$ with $lt(g_i)\, lt(f_i) \leq_T lt(f)$.

$f \in \mathcal{P}$ is called irreducible modulo F if $f \underset{F}{\longrightarrow}^* g$ implies $f = g$.

2.3 For a fixed polynomial r-tuple $\mathbf{f} := (f_1, \ldots, f_r) \in \mathcal{P}^r$, an r-tuple $(g_1, \ldots, g_r) \in \mathcal{P}^r$ is called a *syzygy w.r.t.* $\mathbf{f}$, if $\sum g_i f_i = 0$. The set of all syzygies w.r.t. $\mathbf{f}$ is a module and abbreviated by $Syz(f_1, \ldots, f_r)$. The polynomials $f_1, \ldots, f_r$ will be fixed in the following. For their leading terms, we write briefly $\tau_i := lt(f_i)$, $i = 1, \ldots, r$, and define their least common multiple $\tau_{ij} := l.c.m.\{\tau_i, \tau_j\}$ for all pairs i, j, $1 \leq i, j \leq r$.

2.4 As in Möller & Mora (1986) we use a decomposition of $\mathcal{P}^r$ into one dimensional $I\!\!K$-vector spaces $\{c\; t\; e_i \mid c \in I\!\!K\}$ with e_i the canonical i-th unit vector in $\mathcal{P}^r$, $t \in T$. Their generators $t\; e_i$, $t \in T$, $i = 1, \ldots, r$, will be called *module-terms* and ordered by

$$t_i\, e_i <_G t_j\, e_j :\iff t_i\, \tau_i <_T t_j\, \tau_j \text{ or } (t_i\, \tau_i = t_j\, \tau_j,\; i < j).$$

Since every $\mathbf{g} := (g_1, \ldots, g_r) \in \mathcal{P}^r$, $0 \neq \mathbf{g}$, has exactly one representation as a linear combination of module-terms, the maximal module-term (w.r.t. $<_G$) in this representation is denoted by $lt(\mathbf{g})$. There is an index $k \in \{1, \ldots, r\}$ with $lt(\mathbf{g}) = lt(g_k) e_k$. This k will be denoted as the *index* of $\mathbf{g}$, $ind(\mathbf{g})$, and $lc(\mathbf{g}) := lc(g_k)$. For $0 \in \mathcal{P}^r$, we define $lt(0) := 0$ and $0 <_G te_i$ for all module-terms te_i.

2.5 **Definition.** A set $\{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$ is called a Gröbner basis (for short: G-basis) of a module $M \subseteq \mathcal{P}^r$, if it is contained in M and if the module generated by all $\{lt(\mathbf{m}) \mid \mathbf{m} \in M\}$ has $\{lt(\mathbf{g}_1), \ldots, lt(\mathbf{g}_m)\}$ as basis.

2.6 **Proposition.** (Möller & Mora, 1986, Lemma 7.8) $Syz(M_T(f_1), \ldots, M_T(f_r))$ has a G-basis $\{\mathbf{s}_{ij} \mid 1 \leq i < j \leq r\}$, with

$$\mathbf{s}_{ij} := \frac{\tau_{ij}}{M_T(f_i)} \cdot e_i - \frac{\tau_{ij}}{M_T(f_j)} \cdot e_j .$$

2.7 **Theorem.** Let A be the ideal with basis $F := \{f_1, \ldots, f_r\}$ and let B be a G-basis of $Syz(M_T(f_1), \ldots, M_T(f_r))$. Then the following conditions are equivalent.

i) $f_1, \ldots, f_r$ is a Gröbner basis.

ii) $\forall f \in A : f \underset{F}{\longrightarrow}^* 0$.

iii) $\forall\; \mathbf{b} := (b_1, \ldots, b_r) \in B : \sum b_i\; f_i \underset{F}{\longrightarrow}^* 0$.

iv) $\exists S \subset Syz(f_1, \ldots, f_r)$, S finite:
$\forall \mathbf{g} = (g_1, \ldots, g_r) \in \mathcal{P}^r : (\exists \mathbf{s} \in S : lt(\mathbf{s}) \mid lt(\mathbf{g}))$
or
$lt(\mathbf{g}) = \tau e_k,\ \tau\tau_k = lt(\sum g_i f_i)$.

Proof. i) $\iff$ ii) is standard, ii) $\implies$ iii) trivial since $\sum b_i\; f_i \in A$.
iii) $\implies$ iv): Using the T-representation of $\sum b_i\; f_i$ we obtain for every $\mathbf{b} \in B$ an $\mathbf{s} \in Syz(f_1, \ldots, f_r)$ with $lt(\mathbf{b}) = lt(\mathbf{s})$. Let S be the (finite) set of all such $\mathbf{s}$. Consider $\mathbf{g} \in \mathcal{P}^r$. If $u := max\; lt(g_i)\tau_i$ equals $lt(\sum g_i f_i)$, then $u = lt(g_k)\tau_k$ for $k = ind(\mathbf{g})$ and $lt(\mathbf{g}) = \frac{u}{\tau_k} e_k$. Else the maximal leading terms in $\sum g_i f_i$ cancel,

$$\begin{aligned} \mathbf{g}^* := \ & (\varepsilon_1 M_T(g_1), \ldots, \varepsilon_r M_T(g_r)) \\ & \in Syz(M_T(f_1), \ldots, M_T(f_r)), \end{aligned}$$

where $\varepsilon_j := 1$ if $lt(g_j)\tau_j = u$ and $\varepsilon_j := 0$ else. Hence $lt(\mathbf{g}) = lt(\mathbf{g}^*)$ is a multiple of an $lt(\mathbf{b}), \mathbf{b} \in B$, i.e. of an

$lt(\mathbf{s}), \mathbf{s} \in S$.
iv) $\Longrightarrow$ i): For every $f = \sum g_i f_i$ either $lt(f)$ is a multiple of an $lt(f_k)$ or $\mathbf{s} = (s_1, \ldots, s_r) \in S, \tau \in T$ and $c \in I\!\!K$ exist, such that $lt(g_1 - c\tau s_1, \ldots, g_r - c\tau s_r) <_G lt(g_1, \ldots, g_r)$ and $f = \sum (g_i - c\tau s_i) f_i$ allowing an inductive argument. $\square$

2.8 Remark. The set $S \subset Syz(f_1, \ldots, f_r)$ of condition iv) is already a G-basis as it can be seen by considering only those $\mathbf{g}$ with $\sum g_i f_i = 0$, i.e. $\mathbf{g} \in Syz(f_1, \ldots, f_r)$.

2.9 Remark. If B is the basis described in 2.6, then $\sum b_i \ f_i \underset{F}{\longrightarrow}^* 0$ of iii) is Buchberger's

$$Sp(f_i, f_j) := \frac{\tau_{ij}}{M_T(f_i)} \ f_i - \frac{\tau_{ij}}{M_T(f_j)} \ f_j \underset{F}{\longrightarrow}^* 0$$

for all $1 \leq i < j \leq r$.

2.10 Usually, Buchberger's algorithm is roughly described as follows. Take every non-superfluous S-polynomial $Sp(f_i, f_j)$ and reduce it modulo $F = \{f_1, \ldots, f_r\}$ to an irreducible polynomial. If this is not the zero polynomial, enlarge F by it and pair it with every previous f_i for the forthcoming S-polynomial reductions. Using the correspondence between S-polynomials and basis elements $\mathbf{s}_{ij}$, S-polynomials are superfluous which correspond to $\mathbf{s}_{ij}$'s being redundant in the basis given in proposition 2.6. This gives a criterion for superfluousness of S-polynomials, see Gebauer & Möller (1988). But these authors also showed that this criterion is essentially one of Buchberger's two criteria for detecting superfluous S-polynomials.

In contrast to Gebauer & Möller (1988), who used condition iii) of theorem 2.7, the method presented here is based on condition iv). The computation of the element in $Syz(f_1, \ldots, f_r)$ with maximal form $\mathbf{s}_{ij}$ requires the reduction of the S-polynomial $Sp(f_i, f_j)$. This is time and space consuming. In order to avoid as much S-polynomial reductions as we can, we employ as many elements of $Syz(f_1, \ldots, f_r)$ as we know a priori, like so called *principal* syzygies $(0, \ldots, f_j, \ldots, -f_i, \ldots, 0)$, where the two only nonzero entries are at position i and $j > i$. Also syzygies may be used, which relate the input polynomials and are eventually known by their construction. Since we employ more syzygies than Gebauer & Möller (1988), it is to be expected that in this way more S-polynomial redutions to 0 will be avoided. In fact, this will be shown in the forthcoming.

3 Linear bases and Gröbner bases

3.1 Every polynomial ideal $A \subseteq \mathcal{P}$ can be considered as a linear $I\!\!K$-vector space. Hence it possesses (ideal) bases and linear bases. In this section we give an interpretation of our syzygial approach in terms of linear algebra. Therefore we will distiguish carefully between the different notions of bases. We will say simply "basis" only for ideal bases and use always "linear basis" for linear vector space bases.

3.2 We denote by $\mathcal{P}_t$ the linear vector space of all $f \in \mathcal{P}$ with $lt(f) \leq_T t$. Then $\mathcal{P}$ is the union of all subspaces $\mathcal{P}_t, t \in T$, insofar that every $f \in \mathcal{P}$ belongs to one $\mathcal{P}_t$. An immediate consequence of the notion of (ideal) bases is that an ideal $A \subseteq \mathcal{P}$ has $F := \{f_1, \ldots, f_r\}$ as basis, if and only if the set $\langle F \rangle := \{t \ f_i \mid 1 \leq i \leq r, t \in T\}$ is a linear generator of A considered as $I\!\!K$-vector space. The definition of Gröbner bases for ideals gives even more.

3.3 __**Theorem.**__ Let $F := \{f_1, \ldots, f_r\} \subset \mathcal{P} \setminus \{0\}$ and $A \subseteq \mathcal{P}$ be an ideal. Let $\langle F \rangle_t := \langle F \rangle \cap \mathcal{P}_t$ and $A_t := A \cap \mathcal{P}_t$ for all $t \in T$. Then the following conditions are equivalent.

i) F is a Gröbner basis of A.

ii) $\forall \ f \in A \ \exists \ f_i \in F \ \exists \ t \in T : lt(f) = t \ lt(f_i)$.

iii) $\langle F \rangle_t$ generates linearly A_t for every $t \in T$.

Proof. Let F be a Gröbner basis of A and $f \in A$. Then $lt(f) \in (lt(f_1), \ldots, lt(f_r))$ i.e. ii) holds. If ii) holds, then $\langle F \rangle_t \subset A_t$ for all $t \in T$. For showing that $\langle F \rangle_t$ generates A_t, consider an $f \in A_t$, $t := lt(f)$. Then $M_T(f) = c \ u \ M_T(f_i)$ for suitable $u \in T$, $c \in I\!\!K$, $f_i \in F$ and $lt(f - u \ t \ f_i) <_T lt(f)$ allowing an inductive argument. Hence iii) holds. Conversely, let $\langle F \rangle_t$ generate A_t for all $t \in T$. If for $f \in A$ with $lt(f) = t$ there is no $uf_i \in \langle F \rangle_t$ with $lt(u) lt(f_i) = lt(f)$, then $\langle F \rangle_t$ generates a proper subspace of A_t, a contradiction to iii). Hence $M_T(f) \in (M_T(f_1), \ldots, M_T(f_r))$. $\square$

If F is a Gröbner basis of A with more than one element, then $\langle F \rangle$ being an (ideal) basis generates the $I\!\!K$-vector space A, but it is no linear basis. This follows from the fact that there are $u_i f_i, u_j f_j \in \langle F \rangle$ with $f_i \neq f_j$ and $u_i lt(f_i) = u_j lt(f_j)$. Therefore $\exists c_i, c_j \in I\!\!K \setminus \{0\}$ with $c_i u_i f_i - c_j u_j f_j \in \langle F \rangle_t$, $t <_T u_i lt(f_i)$ allowing to express $u_i f_i$ as linear combination of $u_j f_j$ and of elements from $\langle F \rangle_t$ with $t <_T u_i lt(f_i)$.

3.4 __**Corollary.**__ Let $A \subseteq \mathcal{P}$ be a $I\!\!K$-vector space and let $B \subset A \setminus \{0\}$. If for every $0 \neq f \in A$ there exists exactly one $b \in B$ such that $lt(f) = lt(b)$, then $B_t := B \cap \mathcal{P}_t$ is a linear basis of $A_t = A \cap \mathcal{P}_t$ for all $t \in T$.

Proof. B generates linearly A. This follows as iii) from ii) in the preceding theorem. The $f_i \in B$ are linearly independent, because considering the leading terms of both sides of $0 := \sum_{i=1}^r c_i f_i$, we get $0 = \max lt(c_i f_i)$, i.e. $c_i = 0$ for $i = 1, \ldots, r$. $\square$

3.5 **Definition.** Let A be a $I\!K$-vector subspace of $\mathcal{P}$ and $B \subset A \setminus \{0\}$. If

$$f \in A \Longrightarrow \exists b \in B : lt(f) = lt(b)$$

then B is called a *staggered generator of* A. If in the last equation the $b \in B$ is uniquely determined, then B is called a *staggered linear basis of* A.

The notion "stagger", introduced first in Gebauer & Möller (1986), refers to the fact that a staggered linear basis B contains for every $t \in T$ a linear basis $B_t = B \cap \mathcal{P}_t$ and $B_t \subseteq B_u$ for $t <_T u$. So writing down every basis B_t for increasing t, we get a staggered structure.

3.6 If $F := \{f_1, \ldots, f_r\}$ is a Gröbner basis of an ideal A, then $\langle F \rangle$ is a staggered generator but contains for $card(F) > 1$ elements with the same leading term. By introducing a linear order in $\langle F \rangle$ and by cancelling all polynomials of higher order but same leading term in $\langle F \rangle$, it becomes a staggered linear basis of A.

The converse procedure, the construction of Gröbner bases by means of staggered generators is easy as well, since all staggered generator elements may be cancelled having a leading term which is a proper multiple of another leading term.

3.7 **Proposition.** Let B be a staggered generator of an ideal A. Then the set B' of all $b' \in B$ with

$$\begin{array}{l} \forall b \in B : lt(b) \mid lt(b') \Longrightarrow lt(b) = lt(b') \\ \text{and} \\ \forall b \in B' : lt(b) = lt(b') \Longrightarrow b = b' \end{array}$$

is a Gröbner basis of A. If B is a staggered linear basis, then no proper subset of B' is a Gröbner basis of A.

Proof. The ideal generated by the leading terms of all $f \in A$ is generated by all $lt(b)$, $b \in B$, since B is a staggered generator of A. Therefore $\{lt(b) \mid b \in B'\}$ generates this monomial ideal as well. B' is finite and contained in $B \subset A$, hence a Gröbner basis of A. If B is a staggered linear basis, then no two elements of B and hence no two of B' have the same leading term. By construction, no leading term of an $b \in B'$ divides properly the leading term of an other $b' \in B'$. Therefore, the Gröbner basis B' is irreducible. □

4 The algorithm for staggered linear bases.

4.1 In this section we present an algorithm for computing a staggered linear basis of an ideal A. By proposition 3.7, a Gröbner basis of A can be obtained by cancelling all elements of this linear basis having a leading term which is a proper multiple of the leading term of another basis element. The algorithm will start with an arbitrary basis $F = \{f_1, \ldots, f_r\}$, $0 \notin F$, of the ideal A and with the linear generator $B := \{t_i f_i \mid t_i \in T, 1 \le i \le r\}$. If there are in B elements with the same leading term, all apart of one have to be removed. The modified B has still to generate the linear $I\!K$-vector space A. Therefore it has to be checked, whether the element to be removed depends linearly on the remaining ones and, if not, an element has to be inserted into B with different leading term and allowing to express the removed element by elements of the new B.

4.2 If an element of such B depends linearly on some other elements of B, then there is a dependence relation $\sum_i (\sum_j c_{ij} t_j) f_i = 0$. With $g_i := \sum_j c_{ij} t_j \in \mathcal{P}$, this can be written as $\sum_i g_i f_i = 0$. Hence $\mathbf{g} := (g_1, \ldots, g_r)$ is a nonzero syzygy w.r.t. $(f_1, \ldots, f_r)$. Conversely, every nonzero syzygy $\mathbf{g}$ w.r.t. $(f_1, \ldots, f_r)$ can be interpreted as dependence relation $\sum_i (\sum_j c_{ij} t_j) f_i = 0$, such that every $t_j f_i$ with $c_{ij} \neq 0$ depends linearly on the other polynomials appearing in the sum. In the following we take for linearly dependent only that $t_j f_i$ with nonzero coefficient c_{ij}, which corresponds to the maximal module-term of $\mathbf{g}$, i.e. $lt(\mathbf{g}) = t_j e_i$ or in other words $i = ind(\mathbf{g})$ and with $\mathbf{g} =: (g_1, \ldots, g_r)$ we have $t_j = lt(g_i)$ and $c_{ij} = lc(g_i)$. In that case we say that $t_j f_i$ is a linear combination of lower order polynomials $t f_k$ displaying the fact that for the corresponding module-terms $t e_k <_G t_j e_i$ holds. Observe that with $\mathbf{g}$ also all $t\,\mathbf{g}$ with $t \in T$ are nonzero syzygies with $lt(t\,\mathbf{g}) = t\; lt(\mathbf{g})$, such that $t\; t_j f_i$ is a linear combination of lower order polynomials if $t_j f_i$ does.

4.3 **Proposition.** Let $B := \{t f_i \mid t \in T, 1 \le i \le r\}$ generate the ideal A considered as $I\!K$-vector space. Let S be a finite set of nonzero syzygies w.r.t. $(f_1, \ldots, f_r)$. Then

$$B' := \{t f_i \mid t \in T_i, 1 \le i \le r\}$$

with

$$T_i := T \setminus \{t\tau \mid t \in T, \exists g \in S : lt(g) = \tau e_i\}$$

also generates A.

Proof. Consider an arbitrary $0 \neq f \in A$. Since B generates A, f has a representation

$$f = \sum_i \sum_j c_{ij} t_j f_i, \quad c_{ij} \in I\!K, \; t_j f_i \in B.$$

Let $U := \{t_j f_i \mid c_{ij} \neq 0\}$ be the support of this representation. U is finite and non void. Hence there exists a maximal element in

$$\{t f_k \in U \mid \exists\, \mathbf{g} = (g_1, \ldots, g_r) \in S : ind(\mathbf{g}) = k, \; lt(g_k) \mid t\}.$$

As shown in 4.2, this maximal $t f_k$ can be expressed as linear combination of lower order $t_i f_i \in B$. Hence it can be replaced in the representation of f. For this new representation of f this procedure is repeated, a new maximal element (lesser than the previous one) is found and replaced etc. until all $t_i f_i$ in the support satisfy $t_i \in T_i$. □

4.4 The proposition 4.3 can also be used when some $t_i f_i \in B$ have the same leading term τ. Let $K := \{i \mid t_i\tau_i = \tau\}$ and $i := min(K)$. Then for every $i \neq j \in K$, τ is a multiple of τ_{ij} and

$$t_j f_j - \frac{lc(f_j)}{lc(f_i)} t_i f_i = -\frac{\tau}{\tau_{ij}}\ lc(f_j)\ Sp(f_i, f_j).$$

Reduce $Sp(f_i, f_j)$ modulo $F := \{f_1, \ldots, f_r\}$ to an $h = h(i, j, F)$, $Sp(f_i, f_j) \underset{F}{\longrightarrow}^* h$. Then $Sp(f_i, f_j)$ can be reduced to 0 either modulo F, if $h = 0$, or modulo $F' := F \cup \{h\}$. This means that there is in any case a T-representation

$$Sp(f_i, f_j) = \sum_{k=1}^{r} g_k f_k + h,$$

i.e. a syzygy w.r.t. $(f_1, \ldots, f_r)$, if $h = 0$ or w.r.t. $(f_1, \ldots, f_r, h)$ otherwise. This syzygy is

$$\mathbf{S}_{ij} := (0, \ldots, 0, \frac{\tau_{ij}}{\tau_i}, 0, \ldots, 0, -\frac{\tau_{ij}}{\tau_j}, 0, \ldots, 0) - (g_1, \ldots, g_k)$$

with $lt(\mathbf{S}_{ij}) = \frac{\tau_{ij}}{\tau_j}\ e_j$, and either $k = r$ if $h = 0$, or $k = r + 1$, $g_{r+1} := 1$ otherwise.

Hence if $h \neq 0$, we enlarge $B = \{tf_i \mid t \in T,\ 1 \leq i \leq r\}$ to $\{tf_i \mid t \in T,\ 1 \leq i \leq r + 1\}$ with $f_{r+1} := h$, and modify S to

$$\{(g_1, \ldots, g_r, 0) \mid (g_1, \ldots, g_r) \in S\},$$

which we will denote again by B and S resp. Now we may enlarge S to $S \cup \{\mathbf{S}_{ij}\}$, giving in any case a reduced B' by proposition 4.3 which generates the same space as B does. This is A, because we enlarged B at most by elements $\{th \mid t \in T\} \subset A$, since $h \in A$. But the application of proposition 4.3 gives now, that for the specific i and j no $t_i f_i \in B'$ and $t_j f_j \in B'$ have a common leading term.

4.5 The algorithm can now be described as follows. It is initialized with a set $F := \{f_1, \ldots, f_r\} \subset \mathcal{P}^r$, $0 \notin F$, generating an ideal A, and a finite subset S of $Syz(f_1, \ldots, f_r)$, $0 \notin S$, which contains the principal syzygies $f_j e_i - f_i e_j, 1 \leq i < j \leq r$. Then for $i = 1, \ldots, r$

$$T_i := T \setminus \{t\ lt(g_i) \mid \mathbf{g} \in S, ind(\mathbf{g}) = i\}$$

and $B := \{tf_i \mid t \in T_i, 1 \leq i \leq r\}$. As long as there are $t_i \in T_i, t_j \in T_j$ with $t_i\tau_i = t_j\tau_j$ take such a pair $(t_i, t_j) \in T_i \times T_j$ with minimal (w.r.t. $<_T$) $t := t_i\tau_i = t_j\tau_j$. Reduce $Sp(f_i, f_j)$ modulo F to an irreducible polynomial h. Compute simultaneously $\mathbf{S}_{ij}$. If $h \neq 0$, then modify B and G as in 4.4, enlarge F by $f_{r+1} := h$ and r by 1, modify G by inserting principal syzygies and eventually other known syzygies. At termination, B is a staggered linear basis and S a Gröbner basis for $Syz(f_1, \ldots, f_r)$.

In the algorithm, S can be kept autoreduced. This means, that whenever new elements are inserted into S, such that there are $\mathbf{g}_1, \mathbf{g}_2 \in S$ with $lt(\mathbf{g}_2) = \tau lt(\mathbf{g}_1)$, $\tau \in T$, then $\mathbf{g}_2$ will be replaced in S by $\mathbf{g}_2 - \frac{lc(\mathbf{g}_2)}{lc(\mathbf{g}_1)}\tau\mathbf{g}_1$.

4.6 Proof of correctness. Every actual set B is a linear generator of A by proposition 4.3. The required minimality of $t \in T$ with $t = t_i\tau_i = t_j\tau_j$, $t_i \in T_i, t_j \in T_j$ implies $t = t_{ij}$. Therefore after a loop, all multiples of f_j are removed from B, which have the same leading term as a polynomial $tf_i, t \in T_i$. At termination the leading terms of all $tf_i \in B$ are different. Thus at the end, B is a staggered linear basis of A. Since S contains all $\mathbf{S}_{ij}$ it is a Gröbner basis by proposition 2.6. □

5 The algorithm using syzygies

5.1 The algorithm computes a Gröbner basis together with a Gröbner basis of syzygies; it begins with $(f_1, \ldots, f_r)$ such that τ_i does not divide τ_j if $i < j$, and completes it to a Gröbner basis. In the algorithm module-terms τe_i, $\tau \in T, e_i$ canonical i-th unit vector, are considered in increasing order $<_G$.

We describe the algorithm in a way that several versions are possible, redefining suitable subalgorithms; one of these versions coincides with Buchberger algorithm, in the installation of Gebauer-Möller (1988), keeping trace of the computation to derive syzygies, and with a marginal manipulation of them to obtain at the end a Gröbner basis of syzygies.

The algorithm deals with three sets: the basis B, the syzygies S and the module-term set P. In the beginning B is empty, S is empty, and P is the set of the first r unit vectors $e_i \in \mathcal{P}^m$, m sufficiently large. The algorithm consists of a loop that stops when P is empty. Then B is a Gröbner basis of the ideal and S a Gröbner basis of the module of syzygies.

$B := \emptyset$;
$S := \emptyset$;
$P := \{e_i \mid 1 \leq i \leq r\}$;
REPEAT until $P = \emptyset$
  CHOOSE a module-term $\tau e_i \in P$
    (*Selection Strategy*)
  INTERREDUCE S (*)
  UNLESS exists $Y \in S$ such that
    $lt(Y)$ divides τe_i
    OR

```
  B-criterion(τe_i) (**)
f := τf_i
Y := −τe_i
REPEAT while exists f_j ∈ B such that
    (τ_j|lt(f) AND (i ≠ j OR Y ≠ τe_i))
  CHOOSE one such f_j (Simplification strategy)
  f := f − M_T(f)/M_T(f_j)f_j
  Y := Y + M_T(f)/M_T(f_j)e_j
IF f ≠ 0 THEN
  ADD f to B, as f_k, where k − 1
    is the number of elements of B
  Y := Y − 1e_k
  ADD to S further syzygies that we
    might know (Syzygy addition) (***)
  ADD to P all τe_k such that exists i ≠ k,
    ττ_k = τ_ik and τ minimal
ADD Y to S (Syzygy S_ij )
```

The points (*), (**) and (***) have to be specified to obtain different versions of the algorithm (moreover, all the selections of elements have to be precised, but this is the same in Buchberger's original algorithm). The point (**) (B-criterion) has to be specified in such a way that if it is true, then a syzygy Y such that $lt(Y)$ divides τe_i is computed and added to S as a side-effect; the point (*) (INTERREDUCE S) has to perform manipulations on S in such a way that the module generated by S remains unchanged, and the module generated by $lt(S)$ is increased or unchanged; the point (***) has to add syzygies to S. Under these conditions, at the end of the procedure, B is a Gröbner basis of the ideal, and S a Gröbner basis of the syzygies.

5.2 The correctness of the algorithm is based on condition iv) of theorem 2.7. In addition we need the following.

Lemma. Let $(g_1, \ldots, g_r)$ be a syzygy w.r.t $(f_1, \ldots, f_r)$, and let τe_i be its leading module-term; then there exists $j < i$ such that τ_{ij} divides $\tau\tau_i$.

Proof. The second module-term of the syzygy is a σe_j, and necessarily $\sigma\tau_j = \tau\tau_i$ (otherwise no cancellation can occur in the syzygy). □

The proof that the algorithm is correct follows from the assertion "For every $f = \sum g_i f_i$ then either exists τ_j dividing $lt(f)$ or exists an Y of S such that $lt(Y)$ divides $lt(g_1, \ldots, g_r)$ or exists a τe_i in P dividing $lt(g_1, \ldots, g_r)$" (module-terms are a "promise of a syzygy"). The assertion is trivially true at the beginning, and remains true at every turn of the outermost loop.

5.3 The distinctive points of the algorithm are the addition of known syzygies and the syzygy interreduction; without them the algorithm is just a slightly different version of the Buchberger algorithm; the main difference is that instead of Buchberger's pairs of elements we store τe_j that substantially corresponds to an equivalence class of pairs: all (i, j), $i < j$, such that $\tau_{ij} = \tau\tau_j$. In Buchberger algorithm only one of these pairs is computed, and in our algorithm the choice corresponds to the choice of the first simplification of τf_j: the result is $\mathbf{S}_{ij}$. This different form of the pairs has been suggested by A. Galligo and D. Lazard, and corresponds to Hironaka division. We favour its use mainly as a matter of taste, since some descriptions are easier: it is anyway not difficult to convert one form into the other.

5.4 To obtain precisely the Gebauer-Möller installation of the Buchberger algorithm:
(*) has to do nothing,
(**) `B-criterion` (for module-term τe_i) has to be defined as follows:
`IF` exists $Y_1, Y_2 \in S$ such $Y_1 = \mathbf{S}_{ik}$ and $Y_2 = \mathbf{S}_{jk}$, $j < i < k$, and τ_k divides τ_{ij}, $\tau_{ik} \neq \tau_{ij} \neq \tau_{ik}$, $\tau\tau_i$ multiple of τ_{ij},
`THEN` add $\frac{\tau\tau_i}{\tau_{ik}}\mathbf{S}_{ik} - \frac{\tau\tau_i}{\tau_{jk}}\mathbf{S}_{jk}$ (it is a syzygy with leading module-term τe_i) to S and return true.
`ELSE` false.
(***) (Syzygy addition) has to be defined as follows:
Add to S all $f_j e_k - f_k e_j$ for all j such that τ_j and τ_k are coprime (this corresponds to criterion 2 of Buchberger (1985)).

5.5 Different algorithms are obtained using more syzygies and reductions; for example, an extreme version is obtained defining:
(*) perform all possible interreductions of S
(**) `IF` exists $k > i$, $Y_1 \neq Y_2 \in S$ such that $lt(Y_l) = \sigma_l\ e_k$, and if the leading term of $Sp(Y_1, Y_2) := \frac{\sigma}{\sigma_1 lc(Y_1)} Y_1 - \frac{\sigma}{\sigma_2 lc(Y_2)} Y_2$ with $\sigma := l.c.m.\{\sigma_1, \sigma_2\}$ divides $\tau\ e_i$
`THEN` add $Sp(Y_1, Y_2)$ to S and return true
`ELSE` false.
(***) Add to S all $f_j e_k - f_k e_j$ for all j, (the principal syzygies), and any other syzygy that we might know (either syzygies that we independently know, for example see 5.7, or syzygies derived from computed ones, for example obtained from a known syzygy dividing by a common factor of all coefficients).

5.6 This version considers by far too many syzygies (many of them reduce to 0 at once). Other versions were experimented, that introduce only a subset of principal syzygies: in one version, a principal syzygy is added only if it is not immediately reducible by an existing syzygy; in another version, a principal syzygy of a pair is added only immediately before considering the corresponding Buchberger pair for simplification.

A tuning of the algorithm is necessary to find a good version; moreover, if one is not interested in a basis for syzygies, but only in discovering efficiently some extra set of useless pair, some partial information on syzygies might be sufficient. This will be discussed in the following section.

5.7 The introduction of "other known syzygies" may be used in many cases in which these syzygies are known from the construction of the f_i. Here are two such cases.

Assume that the f_i are determinants of minors of a matrix, the Plücker relations are known: if for example the matrix is a $n \times n+1$ matrix, repeating a row on it we obtain a matrix with determinant 0, and expanding the matrix with respect to the added row with Laplace expansion we have a syzygy between the minors.

Another possibility of getting syzygies is to apply pseudo-reduction. If it is possible to reduce an appropriate multiple $gf_i, g \in \mathcal{P}$, to 0 by subtracting scalar multiples of finitely many tf_j, $t \in T$, $te_j <_G lt(g)e_i$, then a syzygy with leading module-term $lt(g)e_i$ is found. The obvious choice $g = f_j$ leads to the principal syzygy $f_ie_j - f_je_i$. Less obvious choices may give new syzygies.

6 Storing of syzygies

6.1 If we need to compute the syzygies, then the algorithm should be a good one: performing the interreductions of syzygies at an early stage can allow savings since the computation should be done anyway to reduce the basis. This may be questionable in general, since the syzygy managing can be hard both in space and time.

Keeping track of the module-terms processed is necessary anyway, at least in the original description, and this can be seen as a first rough approximation to syzygies; it corresponds to keeping the leading module-term of the syzygies, and allows to discard the module-terms since for this only the leading module-term is used. But it does not allow to draw the conclusions obtainable from the syzygy interreduction, since any information on the rest is lost.

6.2 A good compromise could be the following: for each syzygy keep a few initial module-terms. When interreducing syzygies the first terms are cancelled, but it may happen that it is still possible to find the leading module-term of a new syzygy, and this might allow to discard some terms. Whenever all the information kept is used, we do not longer know the leading term of the resulting syzygy, hence all the information is lost and the syzygy should be discarded.

6.3 Another possibility is to store the syzygies modulo a prime p; we have to distinguish between monomial that do not exist and monomials with coefficient 0 mod p; a cancellation explicitly performed can delete monomials, but accidental cancellation produces coefficients 0; if these are leading coefficients the whole syzygy should be discarded. Of course, a probabilistic modular approach is also possible, but the approach sketched above is deterministic: no uncertain information is ever used.

A combination of both suggestions is of course possible.

7 Examples

We compute two examples in full; we will be sketchy, you can easily complete the computations yourselves. The version that we use introduces a principal syzygy only when the corresponding pair is computed. In the computation display, exponents are not raised (`x2y3` means x^2y^3), `xy.3` is xye_3, syzygies are represented as linear combinations of module-terms, `PS(2,3)` means the principal syzygy $f_2.3 - f_3.2$, `->` means derivation through either evaluation or interreduction.

Here is the first example, using a graduated reverse lexicographical ordering:

$$\begin{aligned} f_0 &= x^2y - z^2 \\ f_1 &= xz^2 - y^2 \\ f_2 &= yz^3 - x^2 \end{aligned}$$

```
P = ((1.0),(1.1),(1.2))
(1.0) ->  x2y-z2 = f0
    P = ((1.1)(1.2))
(1.1) ->  xz2-y2 = f1
    P = ((1.2)(xy.1))
(1.2) ->  yz3-x2 = f2
    P=((x.2)(xy.1))
(x.2)->  y3z-x3 = f3
    S0 = x.2-yz.1-1.3
    PS(1,2) = xz2.2-yz3.1-y2.2+x2.1 ->
                    -> z2.3-y2.2+x2.1 = S1
    P = ((xy.1)(z2.3)(xz.3)(x2.3))
(xy.1)-> -xy3+z4 = f4
    S2 = xy.1-z2.0-1.4
    PS(0,1) = x2y.1-xz2.0+y2.0-z2.1 ->
                     -> x.4+y2.0-z2.1 = S3
    P = ((z.4)(x.4)(z2.3)(xz.3)(x2.3))
(z.4) -> z5-x4 = f5
    S4 = z.4+x.3-1.5
    PS(3,4) = y3z.4+xy3.3-z4.3-x3.4 ->
                -> y3.5-y2z2.2+x2y2.0 ->
                ->  (dividing by y2)
                -> y.5-z2.2+x2.0       = S5
    P = ((x.4)(y.5)(x.5)(z2.3)(xz.3)(x2.3))
(x.4) discarded by S3
(y.5) discarded by S5
(x.5) -> -x5+z2  = f6
    S6 = x.5-z3.1-1.6-y.2-1.0
    PS(1,5) = xz2.5-z5.1+x4.1-y2.5 ->
                 -> z2.6+x4.1+x2y.0+z2.0 = S7
```

```
    P = ((z2.3)(xz.3)(x2.3)(y.6)(z2.6))
(z2.3) discarded by S1
(xz.3) -> y5-x4z = f7
    S8 = xz.3-y3.1-1.7
    PS(1,3) = xz2.3-y3z.1-y2.3+x3.1 ->
                    ->   z.7-y2.3+x3.1 = S9
    P = ((x2.3)(z.7)(x.7)(y.6)(z2.6))
(x2.3): B-criterion applies with S3 and S4;
          considering z.S4-x.S3, we obtain
    S10 = -x2.3+y2z.0+1.6+y.2+1.0
    PS(0,3) = x2y.3-y3z.0+x3.0-z2.3 ->
                ->  -y.6-x3.0-x2.1-y.0 = S11
(z.7) discarded by S9
(x.7) -> 0
    S12 = x.7+y2.4-z.6-yz.2-z.0
    PS(4,7) = -xy3.7-y5.4+x4z.4+z4.7 -> 0
(y.6) discarded by S11
(z2.6) discarded by S7
```

The algorithm has ended. Remark that only one useless pair, (x.7), has remained. In the best form of Buchberger algorithm there were 7 useless pairs (and 5 useful ones).

In conclusion, the Gröbner basis is

$$\begin{aligned}
f_0 &= x^2y - z^2\\
f_1 &= xz^2 - y^2\\
f_2 &= yz^3 - x^2\\
f_3 &= y^3z - x^3\\
f_4 &= xy^3 - z^4\\
f_5 &= z^5 - x^4\\
f_6 &= x^5 - z^2\\
f_7 &= y^5 - x^4z
\end{aligned}$$

and the Gröbner basis of syzygies is

$$\begin{aligned}
S_0 &= x.2 - yz.1 - 1.3\\
S_1 &= z^2.3 - y^2.2 + x^2.1\\
S_2 &= xy.1 - z^2.0 - 1.4\\
S_3 &= x.4 + y^2.0 - z^2.1\\
S_4 &= z.4 + x.3 - 1.5\\
S_5 &= y.5 - z^2.2 + x^2.0\\
S_6 &= x.5 - z^3.1 - 1.6 - y.2 - 1.0\\
S_7 &= z^2.6 + x^4.1 + x^2y.0 + z^2.0\\
S_8 &= xz.3 - y^3.1 - 1.7\\
S_9 &= z.7 - y^2.3 + x^3.1\\
S_{10} &= -x^2.3 + y^2z.0 + 1.6 + y.2 + 1.0\\
S_{11} &= -y.6 - x^3.0 - x^2.1 - y.0\\
S_{12} &= x.7 + y^2.4 - z.6 - yz.2 - z.0
\end{aligned}$$

The second example is an ideal of minors of the matrix $\begin{pmatrix} x & xz & yz \\ y & xz & xy \end{pmatrix}$. The ordering is as before. We use two syzygies coming from known relations between the minors. These allow to detect two useless pairs.

```
f0 = x2z-xyz
f1 = x2y-y2z
f2 = x2yz-xyz2
    S0 = x.2-xz.1+yz.0
    S1 = -xz.1+xy.0+y.2
    P = ((1.2),(z.1))
(1.2) -> xy2z-xyz2 = f3
    S2 = 1.2-y.0-1.3
    S0 ->  x.3-y.2-yz.0
    P = ((z.1)(x.3))
(z.1) -> xyz2-y2z2 = f4
    S3 = z.1-y.0-1.3-1.4
    S1 -> x.4 - yz.0
    P = ((y.4)(x.4)(x.3))
(y.4) -> -y3z2+y2z3
    S4 = y.4-z.3-1.5-z.4
    P = ((x.4)(x.3)(x.5))
(x.4) discarded by S1
(x.3) discarded by S0
(x.5) -> 0
    S5 = x.5+yz.3
```

In conclusion, the (redundant) Gröbner basis is

$$\begin{aligned}
f_0 &= x^2z - xyz\\
f_1 &= x^2y - y^2z\\
f_2 &= x^2yz - xyz^2\\
f_3 &= xy^2z - xyz^2\\
f_4 &= xyz^2 - y^2z^2
\end{aligned}$$

and the Gröbner basis of syzygies is

$$\begin{aligned}
S_0 &= x.3 - y.2 - yz.0\\
S_1 &= x.4 - yz.0\\
S_2 &= 1.2 - y.0 - 1.3\\
S_3 &= z.1 - y.0 - 1.3 - 1.4\\
S_4 &= y.4 - z.3 - 1.5 - z.4\\
S_5 &= x.5 + yz.3
\end{aligned}$$

8 Experimental implementation

A first rough implementation has been made modifying the system AlP*I*, see Traverso-Donati (1989). The first results show an ambiguous behaviour: many useless pairs are discovered, but this involves a lot of extra computation, so the execution time is increased. This might in part depend from several awkward choices that were made in order to have an implementation in a very short time, but in large part depends from a lot of computations that lead to syzygies that are never used, and eventually disappear. This could be fixed either developping good heuristics, or discovering syzygies between the syzygies that allow to avoid useless syzygy computations; this could lead to computing a resolution of the ideal in one single algorithm: we did not have the time to develop this possibility.

We report here, for some of the standard examples in Boege-Gebauer-Kredel (1986), the number of useless pairs in Buchberger algorithm and in one of the implemented versions of our algorithm (we introduce a principal syzygy only if it corresponds to a pair that is computed). This rough version misses to discover some useless pairs that are discovered by some other versions (in the example computed in the preceding section it has to compute 3 useless pairs instead of 1 as we have shown).

Katsura3	5	1
Katsura4	18	7
Trinks1	16	6
Trinks2	10	6
Raksanyi	1	0
Hairer1	10	4
Gonnet	26	19

A better implementation and careful experimenting is needed to be able to draw conclusions on the usefulness of the algorithm.

Acknowledgement. The authors gratefully acknowledge the comments of J. Backelin on a previous version of section 3.

REFERENCES.

- W. Boege, R. Gebauer, H. Kredel, Some examples for solving systems of algebraic equations by calculating Gröbner bases, J. of Symb. Comp., 2, 1986, 83–98
- B. Buchberger, Gröbner bases : An algorithmic method in polynomial ideal theory. Multidimensional Systems Theory (ed.: N.K.Bose), D. Reidel Publ. Comp. 1985, 184 –232.
- R. Gebauer & H.M. Möller, Buchberger's algorithm and staggered linear bases. Proceedings of SYMSAC 1986 (Waterloo/Ontario), pp. 218-221.
- R. Gebauer & H.M. Möller, On an installation of Buchberger's algorithm, J. of Symb. Comp. 6, 1988, 275 – 286.
- D. Lazard, Gröbner bases, Gaussian elimination and resolution of systems of algebraic equations, EUROCAL 1983, London, pp.146-156.
- H. M. Möller & T. Mora, New constructive methods in classical ideal theory, J. Algebra, Vol. 100, 1986, 138 – 178.
- C. Traverso, L. Donati, Experimenting the Gröbner basis algorithm with the AlP*I* system, ISSAC 89 proceedings, A. C. M. 1989