Universität des Saarlandes

Fachbereich 9 – Mathematik

Mathematischer Preprint

Computing Gröbner Bases and Syzygies Using Second Syzygies

Theo de Jong

Preprint No. 6

Saarbrücken 2000

Universität des Saarlandes

Fachbereich 9 – Mathematik

Computing Gröbner Bases and Syzygies Using Second Syzygies

Theo de Jong

Saarland University
Department of Mathematics
Postfach 15 11 50
D–66041 Saarbrücken
Germany
E-Mail: dejong@math.uni-sb.de

submitted: January 7, 2000

Preprint No. 6

Saarbrücken 2000

Edited by
FB 9 – Mathematik
Im Stadtwald
D–66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Computing Gröbner Bases and Syzygies Using Second Syzygies

Theo de Jong
FB 9 Mathematik
Universität Saarbrücken
e-mail:dejong@math.uni-sb.de

January 7, 2000

1 Introduction

Let K be a field, and $>$ be a well-ordering on the monomials of polynomial ring $R = K[x_1, \ldots, x_n]$ in n variables, which is compatible with the semigroup structure of the monomials. Let $F = \oplus_{i=1}^{t} R \cdot \eta_i$ be a finitely generated free R–module with a given basis $\eta_1, \ldots, \eta_t$. Consider a monomial order $>$ on F compatible with the monomial order on R. A set $G = \{f_1, \ldots, f_s\}$ of elements in F is called a Gröbner basis if the ideal generated by $L(f_i)$ is equal to the ideal $L(I)$. Here, $I = (f_1, \ldots, f_s)$, and $L(g)$ denotes the leading monomial of g with respect to the monomial ordering, and $L(I)$ is the monomial ideal generated by $L(g)$ for $g \neq 0$ and $g \in I$. Without loss of generality, and for simplicity, we will assume that the leading coefficient of f_i is one for all i. The well-known Buchberger criterion gives a characterization of Gröbner bases. To formulate this, consider the so–called S–polynomial of f and g:

$$S(f, g) := \frac{\mathrm{lcm}(L(f), L(g))}{L(f))} f - \frac{\mathrm{lcm}(L(f), L(g))}{L(g)} g.$$

This is only defined if $L(f) = x^{\alpha} \eta_i$ and $L(f) = x^{\beta} \eta_i$ with the *same* i. In this case we say that $L(f)$ and $L(g)$ lie in the same summand. If this is not the case, we formally put $S(f, g) = 0$.

Theorem 1.1 (Buchberger, [1]). *The set $G = \{f_1, \ldots, f_s\}$ is a Gröbner basis for $(f_1, \ldots, f_s)$ if and only if for all pairs $i \neq j$ the remainder on division, or normal form,* $\mathrm{NF}(S(f_i, f_j)|G)$ *of $S(f_i, f_j)$ by G is zero.*

Therefore, to show that a certain set $\{f_1, \ldots, f_s\}$ is a Gröbner basis we have to do, a priori, $\frac{s(s-1)}{2}$ checks consisting of divisions with remainder. Since these are computationally often very expensive, it is important, from a practical point of view, to reduces these number of checks. Therefore, it would be nice

to give criteria for reducing these number of checks. Known are the so-called product and chain criteria. In this paper, we propose a new criterion, which is a generalization of the chain criterion. The main idea is to use the second syzygies between the leading monomials. We use them to show that, under suitable assumptions, if $S(f_j, f_k)$ and $S(f_i, f_k)$ are reduced to zero then $S(f_i, f_j)$ reduces to zero, without actually doing the calculation.
This will be done in the second section.
In the third section we will consider syzygies. Schreyer [4] proved that if $f_1, \dots, f_s$ is a Gröbner basis, then we can also find a Gröbner basis of the syzygy module $\mathrm{syz}(f_1, \dots, f_s)$ if we use the so-called Schreyer ordering. We will give an algorithm for computing syzygies, while computing some second syzygies along with it. This will, hopefully, be particular useful for computing resolutions. We hope to report on implementations in the (near) future.

2 A criterion for S–polynomials to reduce to zero.

We fix a monomial well ordering on the free R–module F, where R is the polynomial ring in n–variables. For an element $f \in F$ we denote by $L(f)$ its leading monomial. We need two definitions.

Definition 2.1. For a subset G of the polynomial ring $K[x_1, \dots, x_n]$ we define

$$V(G) = \langle x_i : x_i \mid L(f) \text{ for some } f \in G \rangle.$$

Definition 2.2. Let f_i, f_j, f_j be nonzero elements of F, whose leading monomial lie in the same summand of F. We define

$$\alpha(i, j; k) := \frac{\mathrm{lcm}(L(f_i), L(f_j), L(f_k))}{\mathrm{lcm}(L(f_i, L(f_k))}.$$

Thus $\alpha(i, j; k) \in R$. We denote by B the set of elements of type (i, j) with $1 \leq i < j \leq s$ such that $L(f_i)$ and $L(f_j)$ lie in the same summand.

Theorem 2.3. *Consider a set $G = \{f_1, \dots, f_s\}$ of elements of F. Let D be a subset of B such that for all $(i, j) \in D$ we have $\mathrm{NF}(S(f_i, f_j)|G) = 0$. Suppose that for all $(i, j) \in B \setminus D$ there exists a $k \in \{1, \dots, s\}$ and a subset $G(i, j; k)$ of G such that*

1. $\mathrm{NF}(S(f_p, f_q)|G(i, j; k)) = 0$, *for all $f_p, f_q \in G(i, j; k)$, that is, $G(i, j; k)$ is a Gröbner basis of $\langle G(i, j; k) \rangle$.*
2. $\mathrm{NF}(S(f_i, f_k)|G(i, j; k)) = \mathrm{NF}(S(f_j, f_k)|G(i, j; k)) = 0$,
3. $\alpha(i, j; k) \notin V(G(i, j; k))$.

Then $\mathrm{NF}(S(f_i, f_j)|G(i, j; k)) = 0$. In particular G is a Gröbner basis.

Proof. We have the following identity between the S–polynomials

$$\alpha(i,j;k)S(f_i,f_j) = \alpha(i,k;j)S(f_i,f_k) - \alpha(j,k;i)S(f_j,f_k).$$

Using the assumptions, this shows that

$$\mathrm{NF}((\alpha(i,j;k)S(f_j,f_k)|G(i,j;k)) = 0,$$

By assumption $\alpha(i,j;k)$ is a non-zerodivisor of $\langle L(G(i,j;k)\rangle$, and $G(i,j;k)$ is a Gröbner basis. So it follows that $\mathrm{NF}(S(f_i,f_j)|G(i,j;k)) = 0$. $\square$

Remarks 2.4.

1. Our proof of the theorem in fact has a little bonus. Suppose that for a certain S–polynomial (f_i,f_j) there exist a k with (i,k) and (j,k) in D, but our criterion does not apply because the third condition is violated. Then we have to start reducing the S–polynomial $S(f_i,f_j)$. Take an element in f in $G(i,j;k)$, and say it occurs in $G(i,k)$. At a certain point in the calculation of the reduction of $S(f_i,f_k)$ we have done all the reductions we had to do with f. ¿From this order on, we can drop f from the set $G(i,k)$, and therefore the set $G(i,j;k)$, and hence $V(G(i,j;k))$ might get smaller. Then our criterion might hold from that point on, and we do not have to reduce the $S(f_i,f_j)$ until the end.

2. Suppose F is graded and we have a graded monomomial order on F. Then our criterion works degree for degree, that is we only have to know that the g_i form a standard basis up to the degree of the S–polynomial.

3. In the same vein as the previous remark, we only need that $\alpha(i,j;k) \notin V(G')$, where $G' = \{g \in G(i,j;k) : L(g) \leq L(S(f_i,f_j))\}$. This makes the criterion somewhat stronger.

4. The chain criterion criterion corresponds to the case that $\alpha(i,j;k) = 1$. In this case the first condition can be dropped.

Remark 2.5. It is now quite obvious how to use our criterion in an algorithm. We take the usual Buchberger algorithm, and remembers which elements $G(i,j)$ were used to reduce $S(f_i,f_j)$ to zero. If we take a pair (i,j), we search for a k such that $S(f_i,f_k)$ and $S(f_j,f_k)$ have been reduced to zero. Then one looks whether $G(i,k) \cup G(j,k)$ is a subset of a certain set $G(i,j;k)$ such that all $S(f_p,f_q)$ with $f_p,f_q \in G(i,j;k)$ are reduced to zero using only elements in $G(i,j;k)$. If this is the case then one checks the variable criterion $\alpha(i,j;k) \in V(G(i,j;k))$ holds. If this is the case, one can reduce $S(f_i,f_j)$ to zero by using elements of the set $G(i,j;k)$. If the criterion does not apply we have to reduce $S(f_i,f_j)$ in the classical way. It is quite obvious that one needs good strategies for choosing the pairs. Also, in an implementation, one has the cost of book-keeping, as one has to remember which elements have been used to reduce S–polynomials to zero.

Example 2.6. We take the cone over the rational normal curve of degree four, given by the elements

$$\begin{aligned} f_1 &= x_1^2 - x_0x_2 \\ f_2 &= x_2^2 - x_1x_3 \\ f_3 &= x_3^2 - x_2x_4 \\ f_4 &= x_1x_2 - x_0x_3 \\ f_5 &= x_2x_3 - x_1x_4 \\ f_6 &= x_1x_3 - x_0x_4. \end{aligned}$$

We wish to use our criterion with respect to the graded reverse lexicographical ordering. The product criterion says that we do not have to consider the pairs $(1,2)$, $(1,3)$, $(2,3)$, $(1,5)$, $(2,6)$ and $(3,4)$. We calculate and find that the S–polynomials for the pairs (f_1, f_4) and (f_1, f_6) reduce to zero modulo $G(4,6;1) = \{f_2, f_5, f_3\}$, and we get $V(G(4,6;1)) = (x_2, x_3)$. Our criterion asks whether $x_1 \notin (x_2, x_3)$, so it applies and we do not have to reduce the S–polynomial of (f_4, f_6). Similarly by considering (f_2, f_5) and (f_2, f_4) we see that we do not have to calculate (f_4, f_5). The chain criterion gives that we do not have to calculate for the pair (f_5, f_6). The remaining pairs (f_3, f_5) and (f_3, f_6) remain to be calculated, and in fact we already did that in the proof of the fact that for example $G(4,6;1) = \{f_2, f_5, f_3\}$ is a Gröbner basis. So we gain the non-computation of the reduction of two S–polynomials.

Remarks 2.7.

1. It is quite obvious that our algorithm works well if we can assure that $V(G(i,j;k))$ is small. The chance for this to be the case is probably the biggest for the lexicographical ordering.
2. If our criterion applies, we proved that a certain $S(i,j)$ reduces to zero modulo a certain set $G(i,j;k)$. The actual set might in fact, and probably will, in many cases, be smaller. Thus it might be better not to use our criterion in the "beginning" of the algorithm, so that we can collect more information. Experiments have to show what the best strategy is.

3 Computing Syzygies

We will use similar ideas to compute syzygies, that is, we will use second syzygies to compute first syzygies. More precisely, we try, whenever possible, to compute a first and second syzygy simultaneously. For simplicity, we will first assume that we have a standard basis $f_1, \ldots, f_s$ of an R–module M, where R is the polynomial ring $k[x_1, \ldots, x_n]$ for k a field. Consider the map

$$\begin{aligned} \varphi : \oplus_{i=1}^s R \cdot e_i &\longrightarrow M \\ e_i &\mapsto f_i. \end{aligned}$$

The kernel of φ is called the syzygy module between the $f_1, \ldots, f_s$, which we denote by $\operatorname{syz}(f_1, \ldots, f_s)$. On $\oplus_{i=1}^s R \cdot e_i$ we define a monomial order, called the

Schreyer ordering, see [4]. It is defined as follows:

$$m \cdot e_i > n \cdot e_j \text{ if either } L(mf_i) > L(nf_i) \text{ or } L(mf_i) = L(nf_i) \text{ and } i < j.$$

We recall Schreyer's Theorem for computing syzygies. Let $(i, j) \in B$ and consider the S–polynomial

$$S(f_i, f_j) := \frac{\operatorname{lcm}(L(f_i), L(f_j))}{L(f_j)} f_j - \frac{\operatorname{lcm}(L(f_i), L(f_j))}{L(f_i)} f_i.$$

As $f_1, \ldots, f_s$ is a Gröbner basis, the $S(f_i, f_j)$ reduce to zero modulo $f_1, \ldots, f_s$. Hence we can find $a_k \in R$ (depending on i and j) such that

$$S(f_i, f_j) + \sum_k a_k f_k = 0$$

with $L(a_k f_k) < L(S(f_i, f_j))$ for all k. Thus we get the syzygy

$$R(i, j) := S(i, j) + \sum_k a_k e_k$$

where $S(i, j)$ by definition is

$$S(i, j) := \frac{\operatorname{lcm}(L(f_i), L(f_j))}{L(f_j)} e_j - \frac{\operatorname{lcm}(L(f_i), L(f_j))}{L(f_i)} e_i.$$

Theorem 3.1 (Schreyer). *The $R(i, j)$ for $(i, j) \in B$ form a Gröbner basis for the syzygy module* $\operatorname{syz}(f_1, \ldots, f_s)$ *with respect to the Schreyer order.*

¿From now on we will always use the convention that $R(i, j) = -R(j, i)$. We consider the map

$$\begin{aligned} \psi : \oplus_{(i,j)\in G} R \cdot \varepsilon_{ij} &\longrightarrow \oplus_{i=1}^{s} R \cdot e_i \\ \varepsilon_{ij} &\mapsto R(i, j). \end{aligned}$$

The kernel of ψ is called the second syzygy module. We now present our algorithm for computing syzygies.

Algorithm 3.2.

Input: A Gröbner basis $\{f_1, \ldots, f_s\}$ of M.
Output: The syzygies $R(i, j)$ for $(i, j) \in B$ together with some of the second syzygies $S(i, j; k)$.
Initialization

$B := \{\{i, j\} : L(f_i) \text{ and } L(f_j) \text{ are in the same summand of } F,\ 1 \leq i < j \leq s\}$

$R(i, j) := S(i, j)$. $D := \emptyset$

Iteration

WHILE $B \neq \emptyset$ DO

IF there does not exists a k with $\{i,k\}$ and $\{j,k\}$ in D and $L(f_i)$, $L(f_j)$ and $L(f_k)$ lie in the same summand of F
THEN
  COMPUTE $R(i,j)$.
ELSE
  Take k with $\{i,k\}$ and $\{j,k\}$ in D, and $L(f_i)$, $L(f_j)$ and $L(f_k)$ lie in the same summand.
  $\widetilde{R}(i,j) := -\alpha(i,k;j)R(i,k) + \alpha(j,k;i)R(j,k) + \alpha(i,j;k)R(i,j)$
  $S(i,j;k) := \alpha(i,j;k)\varepsilon_{ij} - \alpha(i,k;j)\varepsilon_{ik} + \alpha(j,k;i)\varepsilon_{jk}$
  WHILE $\widetilde{R}(i,j) \neq 0$ DO
  Let m be the leading term of $\widetilde{R}(i,j)$.
  IF $\alpha(i,j;k) \mid m$
  THEN
    $\widetilde{R}(i,j) := \widetilde{R}(i,j) - m$.
    $R(i,j) := R(i,j) - \frac{m}{\alpha(i,j;k)}$
  ELSE
    Find (α,β) and γ with $m = x^\gamma L(R(\alpha,\beta))$.
    IF $\{\alpha,\beta\} \in B$ THEN COMPUTE $R(\alpha,\beta)$ FI
    IF $\{\alpha,\beta\} \neq \{i,j\}$
    THEN
      $\widetilde{R}(i,j) := \widetilde{R}(i,j) - x^\gamma R(\alpha,\beta)$
      $S(i,j;k) := S(i,j;k) - x^\gamma \varepsilon_{\alpha\beta}$
    ELSE
      $\widetilde{R}(i,j) := 0$
$B := B \setminus \{\{i,j\}\}$
$D := D \cup \{\{i,j\}\}$

Here by COMPUTE $R(i,j)$ we mean computing the syzygy in the standard classical way, as described above.

Proof of the correctness of the algorithm. The algorithm runs as in the classical case, except that we treat the case that we have a triple (i,j,k) whose leading terms lie in the same summand, and of which two syzygies, say $R(i,k)$ and $R(j,k)$ have been computed. Note the following identity in $\oplus_{i=1}^s R \cdot e_i$.

$$-\alpha(i,k;j)S(i,k) + \alpha(j,k;i)S(j,k) + \alpha(i,j;k)S(i,j) = 0$$

By assumption we calculated $R(i,k)$ and $R(j,k)$ which are syzygies whose largest terms are $S(i,k)$ and $S(j,k)$. It follows from our initialization that $\alpha(i,j;k)S(i,j) - \widetilde{R}(i,j)$ is a syzygy. As the $\{f_1,\ldots,f_s\}$ is a standard basis, we can lift the $R(i,j) = S(i,j)$ to $R'(i,j)$. Thus we can write $S(i,j) = R'(i,j) - A(i,j)$. Hence, $\alpha(i,j;k)R'(i,j) - \widetilde{R}(i,j) + \alpha(i,j;k)A(i,j)$ is a syzygy, and thus $-\widetilde{R}(i,j) + \alpha(i,j;k)A(i,j)$ is a syzygy. Thus we see that if the leading term of $\widetilde{R}(i,j)$ is not divisible by $\alpha(i,j;k)$, then it must be divisible by the leading term of $R(\alpha,\beta)$ for some (α,β). Note that in the first case $L(\frac{m}{\alpha(i,j;k)})$ is smaller that the terms in $S(i,j)$. In this case we subtract off m for $\widetilde{R}_{ij}$ and also the $R(i,j)$ and thus $A(i,j)$ are changed accordingly. In the second case we can reduce $\widetilde{R}(i,j)$ with $R(\alpha,\beta)$, thereby also computing a part of the *second syzygy.* What remains is still a syzygy. $\square$

Remarks 3.3. 1. A little generalization of the algorithm shows that we can drop the assumption that $\{f_1,\dots,f_s\}$ is a Gröbner basis, as we can compute a Gröbner along with it. Indeed, we just have to change the procedure COMPUTE $R(i,j)$. In this case then we compute the reduction modulo the set $\{f_1,\dots,f_s\}$. If nonzero, the set $\{f_1,\dots,f_s\}$ and the set of pairs B is extended.

2. Our algorithm also works when the ground ring is a quotient ring R/J of the polynomial ring, as soon as we proved that $\alpha(i,j;k)$ is a non-zerodivisor of R/J.

3. We can iteratively use our algorithm to compute a free resolution (of course truncated if not finite) in the obvious way. In the graded case we can do the computations degree by degree. Moreover, suppose that, in some way, we computed a third syzygy, and look at its components, which correspond to second syzygies. Suppose that all but one of those syzygies have been computed. Then we can use the third syzygy to compute a monomial times the second syzygy, and thus the second syzygy itself. Then the computed second syzygy can be used to eventually compute first syzygies. These ideas are very much related to the ideas of La Scala and Stillmann. Thus probably our algorithm can be combined with that of La Scala and Stillman [3]. They have the strategy of, if possible, computing higher order syzygies first. Other strategies for computing resolutions can be found in [2] and [5].

4. Of course we can make all kind of variants of this algorithms. Which of those will work best in empirical examinations have to show.

Example 3.4. We will consider the cone over the rational normal curve of degree four given by the following six elements in $K[x_0,\dots,x_4]$

$$\begin{aligned}
f_1 &= x_1^2 - x_0x_2\\
f_2 &= x_2^2 - x_1x_3\\
f_3 &= x_3^2 - x_2x_4\\
f_4 &= x_1x_2 - x_0x_3\\
f_5 &= x_2x_3 - x_1x_4\\
f_6 &= x_1x_3 - x_0x_4.
\end{aligned}$$

We already know that it is a standard basis with respect to the graded reverse lexicographical orderings. It is known that the syzygy module is generated by the eight elements. We do not have to consider the pairs $(1,2)$, $(1,3)$, $(2,3)$,

$(1,5)$, $(2,6)$, $(3,4)$ and $(4,5)$. We start off by computing $R(1,4)$ and $R(1,6)$. The result is

$$
\begin{aligned}
R(1,4) &= x_1 e_4 - x_2 e_1 - x_0 e_2 = S(1,4) - x_0 e_2 && (1)\\
R(1,6) &= x_1 e_6 - x_3 e_1 + x_0 e_5 = S(1,6) + x_0 e_5. && (2)
\end{aligned}
$$

Now we use the second syzygy between the leading monomials

$$S(1,4;6) = x_3 \varepsilon_{14} - x_2 \varepsilon_{16} + x_1 \varepsilon_{46}$$

so that

$$x_3 S(1,4) - x_2 S(1,6) + x_1 S(4,6) = 0.$$

Plugging in (1) and (2) we get the syzygy

$$x_0 x_3 e_2 - x_0 x_2 e_5 + x_1 S(4,6) \tag{3}$$

In particular $\alpha(4,6;1) = x_1$, and $\widetilde{R}(4,6) = x_0 x_3 e_2 - x_0 x_2 e_5$. The monomial x_1 does not divide the leading term of $\widetilde{R}(4,6)$. We see from $\widetilde{R}(4,6)$ that we now have to compute $R(2,5)$ and multiply it with x_0, so we do that:

$$R(2,5) = x_2 e_5 - x_3 e_2 - x_1 e_3.$$

We add $x_0 R(2,5)$ from (3) and get the syzygy

$$x_1 S(4,6) - x_0 x_1 e_3.$$

Now $\widetilde{R}(4,6) = -x_0 x_1 x_3$ is divisible by x_1. We thus get the syzygy

$$S(1,4) - x_0 e_3$$

which is equal to $R(1,4)$. Note that we also computed the second syzygy

$$x_3 \varepsilon_{14} - x_2 \varepsilon_{16} + x_1 \varepsilon_{46} + x_0 \varepsilon_{25}.$$

Similarly, we see that we can compute simultaneously the syzygy

$$R(5,6) = S(5,6) - x_4 e_1,$$

and the second syzygy

$$x_1 \varepsilon_{35} - x_2 \varepsilon_{36} + x_3 \varepsilon_{56} - x_4 \varepsilon_{24}.$$

Note that we now computed all eight syzygies, and two of the three second syzygies.

It is a curiosity that we can compute also syzygy by using a non-minimal second syzygy. For example, we compute $R(1,4)$, and we know $R(1,2)$ from the product criterion without doing any calculation. Thus we get the two syzygies

$$
\begin{aligned}
R(1,4) &= \quad S(1,4) - x_0 e_2\\
R(1,2) &= \quad S(1,2) - x_0 x_2 e_2 + x_1 x_3 e_1.
\end{aligned}
$$

We have the identity

$$S(1,2) + x_1 S(2,4) - x_2 S(1,4) = 0$$

so by plugging in we get the syzygy

$$x_1 S(2,4) - (-x_0 x_2 e_2 + x_1 x_3 e_1 + x_o x_2 e_2) = x_1 S(2,4) - x_1 x_3 e_1.$$

Dividing by x_1 we get that $R(2,4) = S(2,4) - x_3 e_1$. It is unclear whether computing syzygies in this way will speed up the computation of the syzygy module.

Acknowledgment. I thank Holger Cröni for asking me a question which started to make me think about these problems, and to Wolfram Decker for discussions. I am very grateful to Olaf Bachmann and in particular to Gerhard Pfister for convincing me that I was on the right track, at a time I thought all was lost.

References

[1] Buchberger, B. (1965) Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Dissertation Innsbruck (1965.)

[2] Capani, A., De Dominicis, G., Niesi, G., and Robbiano, L. (1997) Computing minimal free resolutions. Journal of Pure and Applied Algebra **117-118**, 105-117

[3] La Scala, R. and Stillman, M. (1998) Strategies for computing minimal free resolutions. To appear in Journal of Symbolic Computation.

[4] Schreyer, F.-O. (1980) Die Berechnung von Syzygien mit dem verallgemeinerten Weierstrass'schen Divisionssatz, Diplomarbeit, Hamburg, 1980.

[5] Siebert, T. (1999) Algorithms for Free Resolutions. In: Algorithmic Algebra and Number Theory, B.H. Matzat, G.-M. Greul, G. Hiss (Eds.) Springer Verlag, Berlin etc.