An extension of Buchberger’s criteria for Gröbner basis decision

John Perry

Abstract

Two fundamental questions in the theory of Gröbner bases are decision (‘Is a basis \( G \) of a polynomial ideal a Gröbner basis?’) and transformation (‘If it is not, how do we transform it into a Gröbner basis?’) This paper considers the first question. It is well known that \( G \) is a Gröbner basis if and only if a certain set of polynomials (the \( S \)-polynomials) satisfy a certain property. In general there are \( m(m-1)/2 \) of these, where \( m \) is the number of polynomials in \( G \), but criteria due to Buchberger and others often allow one to consider a smaller number. This paper presents two original results. The first is a new characterization theorem for Gröbner bases that makes use of a new criterion that extends Buchberger’s criteria. The second is the identification of a class of polynomial systems \( G \) for which the new criterion has dramatic impact, reducing the worst-case scenario from \( m(m-1)/2 \) \( S \)-polynomials to \( m-1 \).

1. Introduction

Gröbner bases ease significantly the investigation of many important questions in commutative algebra and algebraic geometry. Fundamental questions in the theory of Gröbner bases include (1) the decision problem, Is a basis \( G \) of a polynomial ideal a Gröbner basis? and (2) the transformation problem, If it is not, how do we transform it into one? This paper considers question (1).

Buchberger [4] showed that \( G \) is a Gröbner basis if and only if the \( S \)-polynomials of every pair of the polynomials in \( G \) satisfy a certain property. Ordinarily, if \( G \) contains \( m \) polynomials, one has to examine \( m(m-1)/2 \) \( S \)-polynomials. Buchberger and others [2, 4, 6, 8, 12, 13, 15, 18] have found criteria on the leading terms of \( G \) that often detect the property before building the \( S \)-polynomial, reducing significantly the number of \( S \)-polynomials that require inspection.

In § 2 we show that one of Buchberger’s two fundamental criteria can be extended in a new and non-trivial way. We will see that this extended criterion specializes to the criterion of [13]. The main theorem uses the extended criterion to formulate a new characterization theorem for Gröbner bases. In § 3 we prove the main theorem. In § 4 we identify a class of polynomial systems where Buchberger’s criteria have no effect, whereas the extended criterion reduces the maximum number of \( S \)-polynomials required to answer question (1) from \( m(m-1)/2 \) to \( m-1 \).

2. The extended criterion

We begin with a review of the essential notation and background material. Standard references in the theory of Gröbner bases are [1, 3, 10].

Fix a commutative ring \( \mathcal{R} \) of polynomials in \( x_1, x_2, \ldots, x_n \) over a field, and an admissible term ordering \( \prec \) over the terms of \( \mathcal{R} \). (In this paper, a term is a monomial whose coefficient is \( 1 \).) For any non-zero \( p \in \mathcal{R} \), we denote the leading term of \( p \) with respect to \( \prec \) by \( \text{lt}_{\prec}(p) \), and the leading coefficient by \( \text{lc}_{\prec}(p) \).
DEFINITION 1 (Gröbner basis). We say that \( G \in \mathbb{R}^m \) is a Gröbner basis with respect to \( \prec \) if, for every polynomial \( p \) in the ideal \( I \) generated by \( G \), there exists some \( g \in G \) such that \( \text{lcm}(\prec(g), \prec(p)) \).

Gröbner bases provide an elegant framework that allows one to decide easily many otherwise difficult problems in commutative algebra and algebraic geometry \([3, 5, 10, 11, 16]\). From an algorithmic perspective, however, Definition 1 is not useful; after all, \( p \) ranges over the infinite set \( I \), so it is impossible to decide whether \( G \) is a Gröbner basis by inspecting every \( p \in I \). Buchberger launched the theory of Gröbner bases by developing a characterization that requires finitely many inspections.

Before stating Buchberger’s characterization, we need a little more notation. For any \( f, g \in \mathbb{R} \), write
\[
\sigma_{f,g} = \frac{\text{lcm}(\prec(f), \prec(g))}{\text{lt}(f)},
\]
and define the \( S \)-polynomial of \( f \) and \( g \) as
\[
S_\prec(f, g) = \frac{\text{lt}(g)}{\text{lt}(f)} \sigma_{f,g} f - \frac{\text{lt}(f)}{\text{lt}(g)} \sigma_{g,f} g.
\]

Let \( G \in \mathbb{R}^m \) and \( p \in \mathbb{R} \), with \( p \neq 0 \). We say that \( p \) reduces to zero with respect to \( G \) if \( p = 0 \) or there exist monomials \( q_1, q_2, \ldots, q_r \) and integers \( \nu_1, \nu_2, \ldots, \nu_r \in \{1, 2, \ldots, m\} \) such that:
- \( p = q_1 g_{\nu_1} + q_2 g_{\nu_2} + \ldots + q_r g_{\nu_r} \);
- \( \text{lt}(q_i) \text{lt}(g_{\nu_i}) \) is a term of \( p \); and
- for \( i > 1 \), each \( \text{lt}(q_i) \text{lt}(g_{\nu_i}) \) is a term of \( p = q_1 g_{\nu_1} - q_2 g_{\nu_2} - \ldots - q_{i-1} g_{\nu_{i-1}} \).

If \( p \neq 0 \) and no \( \text{lt}(q_j) \) divides a term of \( p \), then \( p \) does not reduce to zero with respect to \( G \).

The notions of \( S \)-polynomials and reduction to zero allowed Buchberger to formulate the following \([4]\).

THEOREM 2 (Buchberger’s characterization). Let \( G \in \mathbb{R}^m \). The following are equivalent.

(A) \( G \) is a Gröbner basis with respect to \( \prec \).

(B) For every \( i, j \) such that \( 1 \leq i < j \leq m \), \( S_\prec(g_i, g_j) \) reduces to zero with respect to \( G \).

Unlike \( p \) in Definition 1, \( i \) and \( j \) in (B) range over finitely many integers. Moreover, deciding whether a polynomial reduces to zero with respect to \( G \) requires a finite number of steps. This gives Buchberger’s characterization a decided computational advantage over Definition 1.

Nevertheless, it is usually burdensome to check all the \( S \)-polynomials. Buchberger developed two criteria \([4, 15]\) that modify condition (B) of Buchberger’s characterization.

THEOREM 3. Let \( G \in \mathbb{R}^m \). The following are equivalent.

(A) \( G \) is a Gröbner basis with respect to \( \prec \).

(B) For every \( i, j \) such that \( 1 \leq i < j \leq m \), one of the following holds.

(B0) \( S_\prec(g_i, g_j) \) reduces to zero with respect to \( G \).

(B1) \( \text{lt}(g_i) \) and \( \text{lt}(g_j) \) are relatively prime.

(B2) There exist \( k_1, \ldots, k_n \) such that \( i = k_1 \), \( j = k_n \), each of the \( \text{lt}(g_{k_i}) \) divides \( \text{lcm}(\text{lt}(g_i), \text{lt}(g_j)) \), and each \( S_\prec(g_{k_i}, g_{k_{i+1}}) \) reduces to zero with respect to \( G \).

These criteria, along with adaptations of them, are widely used in both decision and transformation \([2, 7, 8, 12, 18]\). On this account, we make the following definition.

DEFINITION 4 (Buchberger’s criteria). Let \( t_1, t_2, \) and \( t_3 \) be terms of \( \mathbb{R} \). If \( t_1 \) and \( t_2 \) are relatively prime, we say that \( (t_1, t_2) \) satisfies Buchberger’s \text{gcd} criterion. If \( t_2 | \text{lcm}(t_1, t_3) \), we say that \( (t_1, t_2, t_3) \) satisfies Buchberger’s \text{lcm} criterion.

REMARK. The criteria of Definition 4 appear under different names in the literature. Some authors refer to them as Buchberger’s \text{first} and \text{second criteria} \([3]\); others refer to them merely
as optimizations of the singular Buchberger’s criterion, by which they mean criterion (B) of Theorem 2 [10, 16]. The convention we have chosen is based on the content of the criterion, and should avoid confusion.

The lcm criterion can be viewed as a ‘chain condition’ due to the requirement that the chain \( S_2(g_{k_1}, g_{k_{t+1}}) \) reduces to zero, linking \( g_i \) to \( g_j \). The \( \text{gcd} \) criterion can also be viewed as a chain condition, where the ‘chains’ contain zero \( S \)-polynomials. It is not obvious that the \( \text{gcd} \) criterion can, or even should, be extended to longer chains.

A number of researchers have studied how to apply Buchberger’s criteria as efficiently as possible [8, 12]. The algorithm described by Gebauer and Möller [12] is considered a standard benchmark algorithm for approaches to question (2) posed in the introduction.

The main contribution of this paper is to introduce the following criterion, which addresses question (1) by means of a new characterization theorem (the main theorem) as well as the identification of a class of polynomial systems for which the criterion gives a dramatic reduction in the number of \( S \)-polynomials required to answer the question (§ 4).

**Definition 5 (The extended criterion).** Let \( t_1, \ldots, t_m \) be terms of \( \mathcal{R} \). We say that \( (t_1, \ldots, t_m) \) satisfies the extended criterion \((EC)\) if it satisfies \((EDiv)\) and \((EVar)\), where:

- \((EDiv)\) for every \( k \) such that \( 1 \leq k \leq m \), \( \gcd(t_1, t_m) \) divides \( t_k \); and
- \((EVar)\) for every variable \( x \), \( \deg_x \gcd(t_1, t_m) = 0 \) or \( \{\deg_x t_k\}_{k=1}^m = 1 \) is a monotonic sequence.

Observe that \( (t_1, t_2, \ldots, t_m) \) satisfies the extended criterion if and only if its reversal \( (t_m, t_{m-1}, \ldots, t_1) \) does. This is because \((EVar)\) tests for ‘monotonic’ without reference to a direction.

**Example 6.** The list \( T_1 = (x_0x_1, x_0x_2, \ldots, x_0x_m) \) satisfies \((EC)\). Why? \((EDiv)\) is satisfied because \( x_0 \) divides \( t_k \) for \( k = 1, \ldots, m \), and \((EVar)\) is satisfied because \( \{\deg_{x_0} t_k\}_{k=1}^m = (1, 1, \ldots, 1) \) and \( \deg_{x_0} \gcd(t_1, t_m) = 0 \) for \( i = 1, \ldots, m \). Observe that no pair or triplet of terms in \( T \) satisfies either of Buchberger’s criteria.

Similarly, the list \( T_2 = (x_0x_1, x_0^2x_2, x_0^3x_3, x_0^4x_4) \) satisfies \((EC)\) without satisfying Buchberger’s criteria, as illustrated by Figure 1: \( \gcd(t_1, t_4) = x_0 \) divides both \( t_2 \) and \( t_3 \), and \( \{\deg_{x_0} t_k\}_{k=1}^4 = (1, 2, 2, 3) \) is monotonic.

On the other hand, the list \( T_3 = (x_0x_1, x_0^2x_2, x_0^3x_3, x_0^4x_4) \) does not satisfy \((EC)\), because \((EVar)\) is violated: \( \{\deg_{x_0} t_k\}_{k=1}^4 = (1, 2, 3, 2) \) is not monotonic. This is illustrated by Figure 2. A permutation of \( T_3 \), \( (x_0x_1, x_0^2x_2, x_0^3x_3, x_0^4x_4) \), would satisfy \((EC)\), but such permutations are not always possible if \( t_1 \) and \( t_m \) share more than one variable; consider \( x_1y^z, x_2y^2z, x_3y^z, x_4y^3z^2, x_5yz \).

We can use the extended criterion to generalize Buchberger’s characterization theorem.

**Main Theorem.** Let \( G \in \mathcal{R}^m \). The following are equivalent.

(A) \( G \) is a Gröbner basis with respect to \( \prec \).

(B) For every \( i, j \) such that \( 1 \leq i < j \leq m \), one of the following holds.

- (B0) \( S_\prec(g_i, g_j) \) reduces to zero with respect to \( G \).
- (B1) \( \lt_\prec(g_i) \) and \( \lt_\prec(g_j) \) are relatively prime.
- (B2) There exist \( k_1, \ldots, k_n \) such that \( i = k_1, j = k_n \), each of the \( \lt_\prec(g_{k_i}) \) divides \( \text{lcm}(\lt_\prec(g_i), \lt_\prec(g_j)) \), and each \( S_\prec(g_{k_i}, g_{k_{i+1}}) \) reduces to zero with respect to \( G \).
- (B3) There exist \( k_1, \ldots, k_n \) such that \( i = k_1, j = k_n \), the list of leading terms of \( g_{k_1}, \ldots, g_{k_n} \) satisfies \((EC)\), and each \( S_\prec(g_{k_i}, g_{k_{i+1}}) \) reduces to zero with respect to \( G' = (g_{k_1}, \ldots, g_{k_n}) \).
It is essential that in (B3), the reductions to zero are with respect to \( G' \) and not to \( G \). If we use \( G \) instead of \( G' \), then we may not have a Gröbner basis; see Example 8. This also makes it a bad idea to try to combine (B3) and (B2) into one disjunction.

When \( m = 3 \), EC is equivalent to the criterion of [13], which generalizes both of Buchberger’s criteria. For \( m > 3 \), this is not the case. Terms can satisfy Buchberger’s lcm criterion without satisfying EC and, as in Example 6, terms can satisfy EC without satisfying Buchberger’s lcm criterion.

However, if the terms \( t_1 \) and \( t_m \) are relatively prime, then \((t_1, \ldots, t_m)\) satisfies (EDiv) and (EVar) easily. Hence, pairs of leading terms that satisfy Buchberger’s gcd criterion also satisfy the extended criterion. It is not easy to condense (B1) and (B3) into one criterion, because (B3) requires a chain of \( S \)-polynomials that reduce to zero, while (B1) does not. We can therefore view EC as an extension of Buchberger’s gcd criterion to a chain condition.

The remainder of this section consists of examples:

- Example 7 provides a straightforward application of the main theorem;
- Example 8 shows an invalid application of the main theorem.

**Example 7.** Let \( G = (g_1, g_2, g_3, g_4) \), where

\[
g_1 = 4x_0x_1 + 2x_0x_2 + 3x_0x_4 - 8x_1 - 4x_2 - 6x_4,
g_2 = 3x_0^2x_2 + 2x_0^2x_4 - 6x_0x_2 - 4x_0x_1,
g_3 = 4x_0^2x_3 + 2x_0^2x_4 - 8x_0x_3 - 4x_0x_4,
g_4 = 2x_0^3x_4 - 2x_0^3x_3 - x_0^3x_4 + 4x_0x_3 - 6x_0x_4.
\]

Let \( \prec \) represent any term ordering such that \( \text{lt}_\prec(g_1) = x_0x_1 \), \( \text{lt}_\prec(g_2) = x_0^2x_2 \), \( \text{lt}_\prec(g_3) = x_0^2x_3 \), and \( \text{lt}_\prec(g_4) = x_0^3x_4 \). We pose this question: Is \( G \) a Gröbner basis with respect to \( \prec \)?
Routine computation verifies that the pairs (1, 2), (2, 3), and (3, 4) satisfy (B0) of Theorem 3 and of the main theorem; that is, \( S_<(1, g_2) \), \( S_<(g_2, g_3) \), and \( S_<(g_3, g_4) \) reduce to zero with respect to \( G \). We can say something more: in the process of reducing them, we discover that for \( i = 1, 2, 3 \) each \( S_<(g_i, g_{i+1}) \) reduces to zero with respect to \( \{g_i, g_{i+1}\} \). This will prove important in a moment.

As for the remaining pairs, they do not satisfy (B1) or (B2) of either theorem, because no permutation of the leading terms \( x_0 x_1, x_0^2 x_2, x_1^2 x_3, \) and \( x_0^3 x_4 \) satisfies Buchberger’s criteria. Thus, Theorem 3 does not help us answer the question posed.

However, the main theorem does. Observe that

\[
(\text{lt}_<(g_1), \text{lt}_<(g_2), \text{lt}_<(g_3), \text{lt}_<(g_4)) = T_2,
\]

where \( T_2 \) was defined in Example 6; the extended criterion applies to \( T_2 \). In addition, \( S_<(1, g_2) \), \( S_<(g_2, g_3) \), and \( S_<(g_3, g_4) \) reduce to zero with respect to \( G \). Hence, (1, 4) satisfies (B3) of the main theorem with \( G' = G \).

We are not quite done: to decide whether \( G \) is a Gröbner basis, we must resolve the pairs (1, 3) and (2, 4). The main theorem shows that these pairs also satisfy (B0).

- To show that \( S_<(g_1, g_3) \) reduces to zero, we claim that \( \{g_1, g_2, g_3\} \) is a Gröbner basis.
  * We know that the pairs (1, 2) and (2, 3) satisfy (B0) of the main theorem.
  * The extended criterion applies to \( (\text{lt}_<(g_1), \text{lt}_<(g_2), \text{lt}_<(g_3)) \).
  * Recalling that each \( S_<(g_i, g_{i+1}) \) reduces to zero with respect to \( \{g_i, g_{i+1}\} \), we infer that \( S_<(g_1, g_2) \) and \( S_<(g_2, g_3) \) reduce to zero with respect to \( G^{(1,2,3)} = (g_1, g_2, g_3) \). Thus, the pair (1, 3) satisfies (B3) of the main theorem.
  * This implies that \( G^{(1,2,3)} \) is a Gröbner basis, which implies that \( S_<(g_1, g_3) \) reduces to zero.

- To show that \( S_<(g_2, g_4) \) reduces to zero, we claim that \( \{g_2, g_3, g_4\} \) is a Gröbner basis.
  * We know that the pairs (2, 3) and (3, 4) satisfy (B0) of the main theorem.
  * The extended criterion applies to \( (\text{lt}_<(g_2), \text{lt}_<(g_3), \text{lt}_<(g_4)) \).
  * Recalling that each \( S_<(g_i, g_{i+1}) \) reduces to zero with respect to \( \{g_i, g_{i+1}\} \), we infer that \( S_<(g_2, g_3) \) and \( S_<(g_3, g_4) \) reduce to zero with respect to \( G^{(2,3,4)} = (g_2, g_3, g_4) \). Thus, the pair (2, 4) satisfies (B3) of the main theorem.
  * This implies that \( G^{(2,3,4)} \) is a Gröbner basis, which implies that \( S_<(g_2, g_4) \) reduces to zero.

Recall that (1, 4) satisfies (B3) of the main theorem with \( G' = G \). We now know that the other pairs satisfy (B0). It follows from the main theorem that \( G \) is indeed a Gröbner basis with respect to \( \prec \). We have answered the question posed by reducing only three of the six \( S \)-polynomials to zero.

To achieve this, we had to know not only that the \( S \)-polynomials reduced to zero, but also over which subsets of \( G \) they were reduced. Had those subsets been different, the extended criterion probably would not apply, as Example 8 shows below. Conversely, it is conceivable that one could apply the extended criterion but not realize it, because one has verified that the \( S \)-polynomials in question reduce to zero with respect to a different subset of \( G \) than the one needed.

The following example illustrates why (B3) of the main theorem requires \( G' \) and not \( G \).

**Example 8.** Let \( G = (g_1, g_2, g_3, g_4) \), where

\[
\begin{align*}
g_1 &= x^2 y + z, \\
g_2 &= x y z, \\
g_3 &= x y^2, \\
g_4 &= z^2.
\end{align*}
\]
Let $\prec$ be any ordering such that $x^2y \succ z$. Again we ask, is $G$ a Gröbner basis with respect to $\prec$?

It is easy to verify that the pairs $(1, 2), (1, 4), (2, 3), (2, 4), \text{ and } (3, 4)$ satisfy (B0) of the main theorem. The leading terms of $g_1, g_2,$ and $g_3$ satisfy the extended criterion, so set $G' = (g_1, g_2, g_3)$. A subquestion: Does (B3) of the main theorem imply that $G$ is a Gröbner basis? No, because the $S$-polynomials $S_<(g_1, g_2)$ and $S_<(g_2, g_3)$ reduce to zero with respect to $G$, but not with respect to $G'$. In fact, $S_<(g_1, g_3) = yz$ does not reduce to zero with respect to $G$ even though all the other $S$-polynomials do. Thus, $G$ is not a Gröbner basis with respect to $\prec$.

3. Proof of the main theorem

Before diving into details, we pause a moment to describe the fundamental goal of the proof. A previous example will serve us well. The polynomials of Example 7 factor as follows:

$$g_1 = (x_0 - 2)(4x_1 + 2x_2 + 3x_4)$$
$$g_2 = x_0(x_0 - 2)(3x_2 + 2x_4)$$
$$g_3 = 2x_0(x_0 - 2)(2x_3 + x_4),$$
$$g_4 = x_0(x_0 - 2)(2x_0x_4 + 3x_4 - 2x_3).$$

Any pair of the polynomials has a common divisor whose cofactors have relatively prime leading terms: for example, the common divisor of $g_1$ and $g_4$ is $x_0 - 2$, and the leading terms of the cofactors are $x_1$ and $x_0^2x_4$, respectively. From Theorem 3(B1), we know that the system of cofactors of the gcd is a Gröbner basis. Generating a new system whose polynomials are multiples of the cofactors does not alter this, provided that for each pair the multiple of the cofactors is common.

The fundamental goal of the proof is to generalize this observation. Theorem 18 accomplishes this. Lemma 11 is a technical lemma that fills in a crucial step of Lemma 16, which in its turn is a technical lemma that fills in a crucial step of Theorem 18. Lemmas 12 and 14 are also technical lemmas that help clarify some linear algebra necessary for the proof of Lemma 11.

Although Lemma 16 and Theorem 18 generalize similar lemmas and theorems in [13], the increased size of the list ($m > 3$) required the development of an entirely new lemma (Lemma 11), as well as substantial changes to the proof of Lemma 16. In addition, Theorem 18 leads to the important consequence of Corollary 17; this consequence went unremarked in the previous work, but will show itself useful in § 4.

Besides a proof of the main theorem, this section develops several results that are interesting or useful in other contexts. Lemma 11, for example, took us completely by surprise. Lemma 16 generalizes a relationship between the gcd of two polynomials and their $S$-polynomial. Theorem 18 is similar to a well-known theorem regarding Buchberger’s lcm criterion; it will prove useful in § 4, whereas the main theorem does not.

We turn to the proof. We regularly make implicit use of Proposition 9 below. The proof is easy and well known, so we do not repeat it here.

**Proposition 9.** For all $f, g \in \mathcal{R}$, each of the following holds.

(A) If $f + g \neq 0$, then $\text{lt}_<(f + g) \leq \max_<(\text{lt}_<(f), \text{lt}_<(g))$.

(B) $\text{lt}_<(f \cdot g) = \text{lt}_<(f) \cdot \text{lt}_<(g)$.

(C) If $f/g$ is a polynomial, then $\text{lt}_<(f/g) = \text{lt}_<(f)/\text{lt}_<(g)$.

At this point we introduce the concept of an $S$-representation, which is essential to the proof.

**Definition 10.** Let $p \in \mathcal{R}$, $t$ a term of $\mathcal{R}$, and $G \in \mathcal{R}^m$. We say that $h \in \mathcal{R}^m$ is a $t$-representation of $p$ with respect to $G$ if $p = h_1g_1 + \ldots + h_mg_m$, and for all $i$ such that $1 \leq i \leq m$, we have $h_i = 0$ or $\text{lt}_<(h_ig_i) \leq t$.  

https://doi.org/10.1112/S1461157008000193 Published online by Cambridge University Press
Furthermore, let \( g_i, g_j \in G \). If \( t \prec \text{lcm}(\text{lt}_{\prec}(g_i), \text{lt}_{\prec}(g_j)) \) and \( \mathbf{h} \) is a \( t \)-representation of \( S_{\prec}(g_i, g_j) \) with respect to \( G \), then we say that \( S_{\prec}(g_i, g_j) \) has an \( S \)-representation with respect to \( G \), and that \( \mathbf{h} \) is an \( S \)-representation of \( S_{\prec}(g_i, g_j) \) with respect to \( G \). We may omit ‘with respect to \( G \)’ if it is clear from the context.

The notion of \( S \)-representation is related, but not equivalent, to the notion of reduction to zero. We discuss this relationship near the end of the section, where it becomes important for the main theorem. For the time being, we content ourselves with exploring how the extended criterion can link a chain of \( S \)-representations.

To do that, we will need Lemma 11, which identifies a useful and interesting structure in a certain chain of \( S \)-representations.

**Lemma 11.** Let \( G \in \mathbb{R}^m \). Then (A) \( \implies \) (B), where the following hold.

(A) \( S_{\prec}(g_1, g_2), S_{\prec}(g_2, g_3), \ldots, S_{\prec}(g_{m-1}, g_m) \) all have \( S \)-representations with respect to \( G \).

(B) There exist \( P, Q \in \mathbb{R} \) such that
\[
\text{lt}_{\prec}(P) = \sigma_{g_1, g_2} \sigma_{g_2, g_3} \cdots \sigma_{g_{m-1}, g_m},
\]
and
\[
\text{lt}_{\prec}(Q) = \sigma_{g_2, g_1} \sigma_{g_3, g_2} \cdots \sigma_{g_m, g_{m-1}}.
\]

The proof of Lemma 11 requires some non-trivial linear algebra, so we defer it to page 121. Lemmas 12 and 14 provide the necessary results. Lemma 12 describes a relationship between the elimination of variables in a linear system and the coefficients of those variables.

**Lemma 12.** Let \( n \in \mathbb{N}^+ \). Consider the system of \( n-1 \) linear equations in the \( n \) variables \( y_1, \ldots, y_n \):

\[
S_1 = \left\{ \sum_{j=1}^{n} a_{i,j} y_j \right\}_{i=1}^{n-1}.
\]

For \( k = 1, \ldots, n-2 \), define the matrix
\[
A_k = \begin{pmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,k} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,k} \\
\vdots & \vdots & \ddots & \vdots \\
a_{k,1} & a_{k,2} & \cdots & a_{k,k}
\end{pmatrix}.
\]

If each \( A_k \) has non-zero determinant, then, for each \( k = 2, \ldots, n-1 \), the system

\[
S_k = \left\{ \sum_{j=i}^{n} b_{i,j}^{(k)} y_j = 0 \right\}_{i=k}^{n-1}
\]

with
\[
b_{i,j}^{(k)} = \begin{vmatrix}
a_{1,1} & a_{1,j} \\
a_{2,1} & a_{2,j} \\
\vdots & \vdots \\
a_{k-1,1} & a_{k-1,j} \\
a_{i,1} & a_{i,2} & \cdots & a_{i,k-1}
\end{vmatrix}
\]

is consistent.

To prove Lemma 12, we use the following special case of Jacobi’s theorem on determinants, whose proof we do not reproduce here [14, 19].
THEOREM 13. Let $A$ be an $n \times n$ matrix, $M$ a $2 \times 2$ minor of $A$, $M'$ the corresponding $2 \times 2$ minor of the adjugate of $A$, and $M^*$ the $(n-2) \times (n-2)$ minor of $A$ that is complementary to $M$. Then

$$\det M' = \det A \cdot \det M^*.$$ 

We will use Theorem 13 by putting $M$ as the corners of the matrix, making $M^*$ the interior.

Proof of Lemma 12. We proceed by induction on $k$. For the inductive base $k = 2$, eliminate $y_1$ from the equations $i = 2, \ldots, n-1$ in $S_1$ by subtracting the product of the first equation and $a_{i,1}$ from the product of the second equation and $a_{1,1}$. It is routine to verify that for $i = 2, \ldots, n-1$ and $j = 2, \ldots, n$, we have

$$b^{(k)}_{i,j} = \begin{vmatrix} a_{1,1} & a_{1,j} \\ a_{i,1} & a_{i,j} \end{vmatrix}.$$ 

Now assume that the assertion is true for all $\ell$, where $1 \leq \ell < k$. In the system $S_{k-1}$, use the equation $k-1$ to eliminate the variable $y_{k-1}$ from the equations $k, \ldots, n-1$. We obtain a new system of equations

$$S_k = \left\{ \sum_{j=i}^{n} \beta_{i,j} y_j = 0 \right\}_{i=k}^{n-1},$$ 

where, for each $i, j, k$, we have

$$\beta_{i,j} = \begin{vmatrix} a_{1,k-1} & \cdots & a_{1,j} \\ \vdots & \ddots & \vdots \\ a_{i,k-1} & \cdots & a_{i,j} \end{vmatrix} A_{k-2}^{-1} \begin{vmatrix} a_{1,k-1} & \cdots & a_{1,j} \\ \vdots & \ddots & \vdots \\ a_{i,k-1} & \cdots & a_{i,j} \end{vmatrix}.$$ 

Perform the following row and column swaps:

- in $b^{(k-1)}_{k-1,k-1}$, move the bottom row to the top, and the right-most row to the left-most;
- in $b^{(k-1)}_{k-1,j}$, do nothing;
- in $b^{(k-1)}_{i,k-1}$, move the right-most row to the left-most; and
- in $b^{(k-1)}_{i,j}$, move the bottom row to the top.

Denote the resulting matrices by $B_1$, $B_2$, $B_3$, and $B_4$; the negations introduced by the row and column swaps cancel, so that $\beta_{i,j} = B_1 B_2 - B_3 B_4$.

Let

$$C = \begin{vmatrix} a_{k-1,k-1} & \cdots & a_{k-1,k-2} & a_{k-1,j} \\ a_{1,k-1} & \vdots & \vdots & a_{1,j} \\ a_{k-2,k-1} & \cdots & a_{k-2,j} \\ a_{i,k-1} & a_{i,k-2} & a_{i,j} \end{vmatrix}.$$
Theorem 13 with
\[ M = \begin{pmatrix} a_{k-1,k-1} & a_{k-1,j} \\ a_{i,k-1} & a_{i,j} \end{pmatrix} \] and \( M^* = A_{k-2} \)
implies that
\[ \beta_{i,j} = |C| \cdot |A_{k-2}|. \]

Move the top row of \( C \) to the next-to-last row, and the left-most row of \( C \) to the next-to-last column; the negations introduced by the row and column swaps cancel, so that
\[ \beta_{i,j} = \left| \begin{array}{ccc} a_{1,j} & \cdots & a_{k-1,j} \\ A_{k-1} & \cdots & a_{i,k-1} \\ a_{i,1} & \cdots & a_{i,j} \end{array} \right| |A_{k-2}|. \]
From the assumption that \( A_{k-2} \) is non-zero, we can divide each equation of \( S_k \) by \( A_{k-2} \), obtaining the desired linear system.

From this point on, the presence of several \( S \)-representations requires a notation that will allow us to distinguish them.

**Notation.** Let \( G \in \mathbb{R}^m \). Let \( i, j \in \{1, \ldots, m-1\} \) be distinct. We write
\[
h^{(i,j)} = (h_1^{(i,j)}, h_2^{(i,j)}, \ldots, h_m^{(i,j)})
\]
for an \( S \)-representation of \( S_<(g_i, g_j) \) with respect to \( G \). In addition, when \( i < j \), we write
\[
Z_{i,j} = -\text{lrc}_<(g_j)\sigma_{g_i,g_j} + h_i^{(i,j)},
\]
\[
Z_{j,i} = \text{lrc}_<(g_i)\sigma_{g_j,g_i} + h_j^{(i,j)}.
\]
Note that \( \text{lt}_<(Z_{i,j}) = \sigma_{g_i,g_j} \) and \( \text{lt}_<(Z_{j,i}) = \sigma_{g_j,g_i} \).

In the proof of Lemma 11, we will simplify a linear system of the form shown in Lemma 12. To perform this simplification, we must ascertain that the matrices \( A_k \) in that context have non-zero determinant.

**Lemma 14.** Let \( G \in \mathbb{R}^m \). Then \( (A) \implies [(B) and (C)] \), where:

(A) \( S_<(g_1, g_2), S_<(g_2, g_3), \ldots, S_<(g_{m-1}, g_m) \) all have \( S \)-representations with respect to \( G \);

(B) for each \( k = 2, \ldots, m-1 \), the \( k \times k \) matrix
\[
P_k = \begin{pmatrix}
Z_{1,2} & Z_{2,1} & h_3^{(1,2)} & \cdots & h_k^{(1,2)} \\
h_1^{(1,2)} & Z_{2,3} & Z_3 & \cdots & h_k^{(2,3)} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
h_2^{(3,4)} & \ddots & \ddots & \ddots & \vdots \\
h_1^{(k,k+1)} & \cdots & h_{k-1}^{(k,k+1)} & Z_{k,k+1}
\end{pmatrix}
\]
has non-zero determinant; indeed, \( \text{lt}_<(\det P_k) = \sigma_{1,2}\sigma_{2,3}\cdots\sigma_{k,k+1} \).
(C) for each \( k = 2, \ldots, m - 1 \), the \( k \times k \) matrix

\[
Q_k = \begin{pmatrix}
Z_{2,1} & h_{3}^{(1,2)} & h_{4}^{(1,2)} & \cdots & h_{k+1}^{(1,2)} \\
Z_{2,3} & Z_{3,2} & h_{4}^{(2,3)} & \cdots & h_{k+1}^{(2,3)} \\
h_{2}^{(3,4)} & \cdots & \cdots & \cdots & h_{k+1}^{(3,4)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
h_{2}^{(k,k+1)} & \cdots & h_{k-2}^{(k+1)} & Z_{k,k+1} & Z_{k+1,k+1}
\end{pmatrix}
\]

has non-zero determinant; indeed, \( \lt_{\prec} (\det Q_k) = \sigma_{2,1} \sigma_{3,2} \cdots \sigma_{k+1,k} \).

The proof of Lemma 14 is tricky, so we present a simple but non-trivial example to illustrate the strategy.

Example 15. Suppose \( m > 3 \) and the system \( G \in \mathbb{R}^m \) satisfies (A) of Lemma 14. We show that (C) is satisfied for \( k = 3 \). A determinant is a sum of elementary products; since

\[
Q_3 = \begin{pmatrix}
Z_{2,1} & h_{3}^{(1,2)} & h_{4}^{(1,2)} \\
Z_{2,3} & Z_{3,2} & h_{4}^{(2,3)} \\
h_{2}^{(3,4)} & Z_{3,4} & Z_{4,3}
\end{pmatrix}
\]

and the leading term of \( Z_{2,1} Z_{3,2} Z_{4,3} \) is \( \tau = \sigma_{2,1} \sigma_{3,2} \sigma_{4,3} \), the leading term of at least one elementary product of \( \det Q_3 \) has the desired form.

We claim that the leading term of every other elementary product of \( \det Q_3 \) is smaller than \( \tau \). We proceed by way of contradiction. Assume that some other term in the elementary product has a leading term greater than or equal to \( \tau \). Consider the leading terms of the other five polynomials, denoting \( \text{lcm}(\lt_{\prec} (g_i), \lt_{\prec} (g_j)) \) by \( L_{i,j} \) and \( \lt_{\prec} (g_i) \) by \( t_i \).

Case 1: Suppose \( \tau \preceq \lt_{\prec} (h_{3}^{(1,2)} \cdot h_{2}^{(2,3)} \cdot h_{2}^{(3,4)}) \). Multiply both sides of the inequality by \( t_2 t_3 t_4 \) to obtain

\[
L_{1,2} L_{2,3} L_{3,4} \preceq [t_3 \cdot \lt_{\prec} (h_{3}^{(1,2)})][t_4 \cdot \lt_{\prec} (h_{2}^{(2,3)})][t_2 \cdot \lt_{\prec} (h_{2}^{(3,4)})],
\]

which contradicts the definition of an \( S \)-representation.

Case 2: Suppose \( \tau \preceq \lt_{\prec} (h_{4}^{(1,2)} \cdot Z_{2,3} \cdot Z_{3,4}) \). Multiply both sides of the inequality by \( t_2 t_3 t_4 \) to obtain

\[
L_{1,2} L_{2,3} L_{3,4} \preceq [t_4 \cdot \lt_{\prec} (h_{4}^{(1,2)})] \cdot L_{2,3} \cdot L_{3,4},
\]

and divide both sides by the common \( \text{lcm} \)'s to obtain

\[
L_{1,2} \preceq t_4 \cdot \lt_{\prec} (h_{4}^{(1,2)}),
\]

which contradicts the definition of an \( S \)-representation.

Case 3: Suppose \( \tau \preceq \lt_{\prec} (h_{2}^{(3,4)} \cdot Z_{3,2} \cdot h_{4}^{(1,2)}) \). Multiply both sides of the inequality by \( t_2 t_3 t_4 \) to obtain

\[
L_{1,2} L_{2,3} L_{3,4} \preceq [t_2 \cdot \lt_{\prec} (h_{2}^{(3,4)})] \cdot L_{2,3} \cdot [t_4 \cdot \lt_{\prec} (h_{4}^{(1,2)})],
\]

and divide both sides by the common \( \text{lcm} \)'s to obtain

\[
L_{1,2} L_{3,4} \preceq [t_2 \cdot \lt_{\prec} (h_{2}^{(3,4)})][t_4 \cdot \lt_{\prec} (h_{4}^{(1,2)})],
\]

which contradicts the definition of an \( S \)-representation.
Case 4: Suppose \( \tau \leq \text{lt}_\prec(Z_{(3,4)} \cdot h_4^{(2,3)} \cdot Z_{2,1}) \). Multiply both sides of the inequality by \( t_2t_3t_4 \) to obtain
\[
L_{1,2}L_{2,3}L_{3,4} \leq L_{3,4} \cdot [t_4 \cdot \text{lt}_\prec(h_4^{(2,3)})] \cdot L_{1,2},
\]
and divide both sides by the common lcm’s to obtain
\[
L_{2,3} \leq t_4 \cdot \text{lt}_\prec(h_4^{(2,3)}),
\]
which contradicts the definition of an \( S \)-representation.

Case 5: Suppose \( \tau \leq \text{lt}_\prec(Z_{4,3} \cdot Z_{2,3} \cdot h_3^{(1,2)}) \). Multiply both sides of the inequality by \( t_2t_3t_4 \) to obtain
\[
L_{1,2}L_{2,3}L_{3,4} \leq L_{3,4} \cdot L_{2,3} \cdot [t_3 \cdot \text{lt}_\prec(h_3^{(1,2)})],
\]
and divide both sides by the common lcm’s to obtain
\[
L_{1,2} \leq t_3 \cdot \text{lt}_\prec(h_3^{(1,2)}),
\]
which contradicts the definition of an \( S \)-representation.

The proof of Lemma 14 follows this strategy. It is clear from the main diagonal of each \( Q_k \) that the leading term \( t \) of one elementary product of the determinant of \( Q_k \) has the desired form; assume by way of contradiction that the leading term of another elementary product is greater than or equal to \( t \); simplify the equivalent inequality by clearing the denominators and dividing by the lcm’s; the resulting inequality will contradict the definition of an \( S \)-representation.

**Proof of Lemma 14.** We prove that \((A) \implies (C)\). The proof that \((A) \implies (B)\) is similar.

It is clear that \( \det Q_k \) is a polynomial, each of whose terms is an elementary product of the matrix. We can write any elementary product as \( T = \prod_{i=1}^{k} B_i \) such that:

– each \( B_i \) is an element of row \( i \); and

– if \( i \neq j \), then \( B_i \) and \( B_j \) are elements of different columns.

As noted above, the main diagonal of \( Q_k \) produces an elementary product whose leading term has the desired form; we claim that every other elementary product has a smaller leading term.

We proceed by way of contradiction. Assume that some elementary product \( T \) besides the main diagonal satisfies
\[
\prod_{i=1}^{k} \sigma_{i+1,i} \leq \text{lt}_\prec(T), \tag{3.1}
\]
Partition the set of factors of \( T \) into three sets:

– \( \mathcal{D} \), containing those factors which are on the main diagonal, which have the form \( Z_{i,i+1} \) for some \( i = 1, \ldots, k \);

– \( \mathcal{L} \), containing those factors which are immediately below the main diagonal, which have the form \( Z_{i,i+1} \) for some \( i = 1, \ldots, k \); and

– \( \mathcal{O} \), containing the other factors, which have the form \( h_i^{(j,j+1)} \) for appropriate \( i, j \).

Since \( T \) is not the product of the main diagonal, the uniqueness of row and column representatives among the factors of \( T \) implies that \( \mathcal{O} \) is guaranteed to be non-empty.

Denote \( \text{lcm}(\text{lt}_\prec(g_i), \text{lt}_\prec(g_j)) \) by \( L_{i,j} \) and \( \text{lt}_\prec(g_i) \) by \( t_i \). Multiply both sides of (3.1) by \( \prod_{\ell=2}^{k+1} t_\ell \).

This results in the equation
\[
\prod_{i=1}^{k} t_{i+1} \cdot \sigma_{i+1,i} \leq \prod_{\ell=2}^{k+1} t_\ell \cdot \prod_{z_{i+1,i} \in \mathcal{D}} \sigma_{i+1,i} \cdot \prod_{z_{i,i+1} \in \mathcal{L}} \sigma_{i,i+1} \cdot \prod_{h_i^{(j,j+1)} \in \mathcal{O}} h_i^{(j,j+1)}.
\]
Simplify the left-hand side to obtain
\[
\prod_{i=1}^{k} L_{i,i+1} \leq \prod_{\ell=2}^{k+1} t_{\ell} \cdot \prod_{i=1}^{\sigma_{i+1,i}} \sigma_{i+1,i} \cdot \prod_{Z_{i,i+1} \in \mathcal{L}} \sigma_{i,i+1} \cdot \prod_{h_{i}^{(j,j+1)} \in \mathcal{O}} h_{i}^{(j,j+1)}.
\]  
(3.2)

Rearrange the right-hand side of (3.2) by pairing each \( t_{\ell} \) with the corresponding factor taken from column \( \ell - 1 \). The uniqueness of column representatives among the factors of an elementary product of a matrix guarantees a one-to-one pairing. If \( t_{\ell} \) is paired with an element of:
- \( \mathcal{D} \), it is paired with \( Z_{\ell,\ell-1} \), and the product simplifies to \( L_{\ell-1,\ell} \);
- \( \mathcal{L} \), it is paired with \( Z_{\ell,\ell+1} \), and the product simplifies to \( L_{\ell,\ell+1} \);
- \( \mathcal{O} \), it is paired with \( h_{i}^{(j,j+1)} \) for appropriate \( j \).

In addition, the uniqueness of row representatives among the factors of an elementary product implies that for each \( i \), at most one pairing simplifies to \( L_{i,i+1} \). Thus, if we simplify the right-hand side of (3.2), we have
\[
\prod_{i=1}^{k} L_{i,i+1} \leq \prod_{h_{i}^{(j,j+1)} \in \mathcal{O}} L_{i,i+1} \cdot \prod_{h_{i}^{(j,j+1)} \in \mathcal{O}} t_{i} h_{i}^{(j,j+1)}.
\]

Divide both sides by \( \prod_{h_{i} \notin \mathcal{O}} L_{i,i+1} \) and we have
\[
\prod_{h_{i}^{(j,j+1)} \in \mathcal{O}} L_{i,i+1} \leq \prod_{h_{i}^{(j,j+1)} \in \mathcal{O}} t_{i} h_{i}^{(j,j+1)}.
\]

Recall that \( \mathcal{O} \) was guaranteed to be non-empty, so these products are greater than 1. This contradicts the definition of an \( S \)-representation.

We have shown that the leading term of the elementary product of \( \text{det} Q_{k} \) formed on the main diagonal is \( \prod_{i=1}^{k} \sigma_{i+1,i} \), while the leading terms of the remaining elementary products are strictly smaller. The sum of the elementary products thus derives its leading term from the main diagonal, whose leading term is the form described by (B).

Finally, we turn to the proof of Lemma 11.

Proof of Lemma 11. Assume (A). We must show (B).

For each \( i = 1, \ldots, m - 1 \), fix \( h^{(i,i+1)} \), an \( S \)-representation of \( S_{<}(g_{i}, g_{i+1}) \). We have the system of \( m - 1 \) equations
\[
Z_{1,2} g_{1} + Z_{2,1} g_{2} + h_{3}^{(1,2)} g_{3} + \ldots + h_{m}^{(1,2)} g_{m} = 0,
\]
\[
h_{1}^{(2,3)} g_{1} + Z_{2,3} g_{2} + Z_{3,2} g_{3} + \ldots + h_{m}^{(2,3)} g_{m} = 0,
\]
\[
\vdots
\]
\[
h_{1}^{(m-1,m)} g_{1} + \ldots + h_{m-1,m}^{(m-1,m)} g_{m-1} + Z_{m-1,m} g_{m-1} + Z_{m,m} g_{m} = 0.
\]

We will study this system in the context of Lemmas 12 and 14.

To apply Lemma 12, we put \( y_{k} = g_{k+1} \) for \( k = 1, \ldots, m - 2 \), \( y_{m-1} = g_{1} \), and \( y_{m} = g_{m} \). For each \( k = 1, \ldots, m - 2 \),
\[
A_{k} = \begin{pmatrix}
Z_{2,1} & h_{3}^{(1,2)} & h_{4}^{(1,2)} & \cdots & h_{m}^{(1,2)} \\
Z_{2,3} & Z_{3,2} & h_{4}^{(2,3)} & \cdots & h_{k+1}^{(2,3)} \\
h_{2}^{(3,4)} & Z_{3,4} & Z_{4,3} & h_{k+1}^{(3,4)} & \cdots & h_{k+1}^{(k+1,k)} \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
h_{2}^{(k,k+1)} & \cdots & h_{k-3}^{(k,k+1)} & Z_{k,k+1} & Z_{k+1,k}
\end{pmatrix}
\]
It is evident that $A_1$ is non-singular; by Lemma 14, $A_k$ is non-singular for $k > 1$. By Lemma 12, the system

$$S_{m-1} = \left\{ \sum_{j=1}^{m} b_{i,j}^{(m-1)} y_j = 0 \right\}_{i=m-1}^{m-1}$$

with

$$b_{m-1,m-1}^{(m-1)} = \begin{vmatrix} Z_{1,2} & A_{m-2} & h_1^{(2,3)} \\ & h_{1}^{(m-2,m-1)} & h_1^{(m-2,m-1)} \\ & h_2^{(m-1,m)} & Z_{m-1,m} & h_1^{(m-1,m)} \end{vmatrix}$$

and

$$b_{m-1,m}^{(m-1)} = \begin{vmatrix} A_{m-2} & h_1^{(1,2)} & h_1^{(2,3)} & \vdots \\ & h_2^{(m-1,m)} & \ldots & Z_{m-1,m} & h_1^{(m-1,m)} \\ & h_2^{(m-1,m)} & \ldots & Z_{m-1,m} & Z_{m,m-1} \end{vmatrix}$$

is consistent. In other words,

$$-b_{m-1,m-1}^{(m-1)} \cdot g_1 = b_{m-1,m}^{(m-1)} \cdot g_m.$$ 

Let

$$P = \det(b_{m-1,m-1}^{(m-1)}) \quad \text{and} \quad Q = \det(b_{m-1,m}^{(m-1)}).$$

We claim that $\text{lt}_{\prec}(P)$ and $\text{lt}_{\prec}(Q)$ have the form specified. For $\text{lt}_{\prec}(Q)$, it is clear that Lemma 14(C) applies. For $\text{lt}_{\prec}(P)$, use column swaps to shift the right-most column to the left-most, while shifting the other columns one position to the right; at this point Lemma 14(B) applies.

Gröbner basis theory generalizes many algorithms for univariate polynomials to systems of multivariate polynomials; one oft-cited example is how Buchberger’s algorithm to compute a Gröbner basis can be viewed as a generalization of the Euclidean algorithm to compute the gcd. We likewise expect relationships to exist between the $S$-polynomials and the gcd’s of polynomials.

Moreover, the construction of $S$-polynomials relies on the computation of

$$\sigma_{g_i,g_j} = \frac{\text{lcm}(\text{lt}_{\prec}(g_i), \text{lt}_{\prec}(g_j))}{\text{lt}_{\prec}(g_i)},$$

which can be rewritten as

$$\sigma_{g_i,g_j} = \frac{\text{lt}_{\prec}(g_j)}{\gcd(\text{lt}_{\prec}(g_i), \text{lt}_{\prec}(g_j))}.$$ 

Based on this, one might expect the existence of criteria on $S$-polynomials that relate the gcd of two polynomials with the gcd of their leading terms.

One such criterion exists for two polynomials: if $G = \{g_1, g_2\}$ is a Gröbner basis, then the $S$-polynomial of $g_1$ and $g_2$ reduces to zero, and in addition $g_1 = f_1 p$ and $g_2 = f_2 p$, where $p = \gcd(g_1, g_2)$ and the leading terms of $f_1$ and $f_2$ are relatively prime [1]. In this case, we infer a surprising fact. Observe that

$$\text{lt}_{\prec}(\gcd(g_1, g_2)) = \text{lt}_{\prec}(\gcd(f_1 p, f_2 p)) = \text{lt}_{\prec}(\gcd(f_1, f_2) \cdot p).$$
Since $p$ is the gcd of $g_1$ and $g_2$, we know that $f_1$ and $f_2$ must be relatively prime, so
\[
\text{lt}((\gcd(g_1, g_2)) = \text{lt}(1) \cdot \text{lt}(p)
\]
\[
= \gcd(\text{lt}(f_1), \text{lt}(f_2)) \cdot \text{lt}(p)
\]
\[
= \gcd(\text{lt}(f_1) \text{lt}(p), \text{lt}(f_2) \text{lt}(p))
\]
\[
= \gcd(\text{lt}(g_1), \text{lt}(g_2)).
\]

Lemma 16 generalizes this observation in a way that does not require a Gröbner basis, but does require the extended criterion.

**Lemma 16.** Let $G \in \mathcal{R}^m$, and suppose that the leading terms of $G$ satisfy the extended criterion. Then $(A) \implies (B)$, where:

(A) each of $S_\prec(g_1, g_2)$, $S_\prec(g_2, g_3), \ldots, S_\prec(g_{m-1}, g_m)$ has an $S$-representation with respect to $G$;

(B) $\gcd(\text{lt}(g_1), \text{lt}(g_m)) \neq \gcd(\text{lt}(g_1), \text{lt}(g_m))$.

**Proof.** Assume (A). We must show (B). For the sake of convenience, denote $\text{lt}(g_i)$ by $t_i$.

By Lemma 11, we have
\[
g_1 P = g_m Q,
\]
where
\[
\text{lt}(P) = \sigma_{g_1:9_2} \sigma_{g_2:9_3} \cdots \sigma_{g_{m-1}:9_m} \quad \text{and} \quad \text{lt}(Q) = \sigma_{g_2:9_1} \sigma_{g_3:9_2} \cdots \sigma_{g_m:9_{m-1}}.
\]

Let $p = \gcd(g_1, g_m)$ and put $f_1 = g_1/p$ and $f_m = g_m/p$. Then
\[
f_1 P = f_m Q.
\]

(3.3)

Since $f_1, f_m$ are relatively prime, $f_1 \mid Q$. Thus, $\text{lt}(f_1)$ divides $\text{lt}(Q)$.

Observe that for any $i = 1, \ldots, m - 1$, we have
\[
\sigma_{g_{i+1}:9_i} = \frac{\text{lcm}(t_i, t_{i+1})}{t_{i+1}} = \frac{t_i}{\gcd(t_i, t_{i+1})}.
\]

Thus,
\[
\text{lt}(f_1) \mid \frac{t_1 t_2 \cdots t_{m-1}}{\gcd(t_1, t_2) \gcd(t_2, t_3) \cdots \gcd(t_{m-1}, t_m)}.
\]

Denote $\gcd(t_i, t_j)$ by $d_{i,j}$. For all variables $x$, we have
\[
\deg_x \text{lt}(f_1) \leq \deg_x \frac{t_1 \cdots t_{m-1}}{d_{1,2}d_{2,3} \cdots d_{m-1,m}}.
\]

Recall that $f_1 = g_1/p$. For all variables $x$, we have
\[
\deg_x t_1 - \deg_x \text{lt}(p) \leq \sum_{1 \leq i < m} \deg_x t_i - \sum_{1 \leq i < m} \deg_x d_{i,i+1}.
\]

\[
\sum_{1 \leq i < m} \deg_x d_{i,i+1} \leq \deg_x \text{lt}(p) + \sum_{1 \leq i < m} \deg_x t_i.
\]

(3.4)

We claim that for all variables $x$, $\deg_x d_{1,m} \leq \deg_x \text{lt}(p)$. Let $x$ be arbitrary, but fixed. If $\deg_x t_1 = 0$ or $\deg_x t_m = 0$, the claim is trivially true. So, assume $\deg_x t_1 \neq 0$ and $\deg_x t_m \neq 0$. We consider two cases.

If $\deg_x t_1 \leq \deg_x t_m$, then $\deg_x d_{1,m} = \deg_x t_1$. Recall that $t_1, \ldots, t_m$ satisfy EC. Therefore, $\deg_x t_1 \leq \deg_x t_2 \leq \cdots \leq \deg_x t_m$. Thus, $\deg_x d_{i,i+1} = \deg_x t_i$ for all $i$ such that $1 \leq i \leq m - 1$. Apply this to (3.4) to obtain
\[
\deg_x d_{1,m} = \deg_x t_1 \leq \deg_x \text{lt}(p).
\]
If \( \deg_x t_1 \geq \deg_x t_m \), a similar argument gives \( \deg_x d_{1,m} \leq \deg_x \text{lt}_\prec(p) \).

Since \( x \) is arbitrary, \( d_{1,m} \) divides \( \text{lt}_\prec(p) \) or, equivalently, \( \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) \) divides \( \text{lt}_\prec(\gcd(g_1, g_m)) \). That \( \text{lt}_\prec(\gcd(g_1, g_m)) \) divides \( \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) \) is trivial. Hence, \( \text{lt}_\prec(\gcd(g_1, g_m)) = \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) \).

The following result will be useful both for the proof of the main theorem and for § 4.

**Corollary 17.** Let \( G \in \mathbb{R}^m \), and suppose that the leading terms of \( G \) satisfy the extended criterion. Then \( (A) \Rightarrow (B) \), where:

(A) \( S_\prec(g_1, g_2), S_\prec(g_2, g_3), \ldots, S_\prec(g_{m-1}, g_m) \) all have \( S \)-representations with respect to \( G \);

(B) if \( p = \gcd(g_1, g_m) \), then \( \text{lt}_\prec(g_1/p) \) and \( \text{lt}_\prec(g_m/p) \) are relatively prime.

**Proof.** Assume \( (A) \). Let \( p = \gcd(g_1, g_m) \), and denote \( g_1/p \) and \( g_m/p \) by \( f_1 \) and \( f_m \), respectively. From Lemma 16, we know that

\[
\gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) = \text{lt}_\prec(p).
\]

Thus, for any variable \( x \),

\[
\deg_x \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) = \deg_x \text{lt}_\prec(g_1) - \deg_x \text{lt}_\prec(f_1) = \deg_x \text{lt}_\prec(g_m) - \deg_x \text{lt}_\prec(f_m).
\]

Let \( x \) be arbitrary, but fixed. If \( \deg_x \text{lt}_\prec(g_1) \leq \deg_x \text{lt}_\prec(g_m) \), then

\[
\deg_x \text{lt}_\prec(g_1) = \deg_x \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) = \deg_x \text{lt}_\prec(g_1) - \deg_x \text{lt}_\prec(f_1),
\]

so \( \deg_x \text{lt}_\prec(f_1) = 0 \). Similar reasoning shows that if \( \deg_x \text{lt}_\prec(g_1) \geq \deg_x \text{lt}_\prec(g_m) \), then \( \deg_x \text{lt}_\prec(f_m) = 0 \). It follows that \( \text{lt}_\prec(g_1/p) \) and \( \text{lt}_\prec(g_m/p) \) are relatively prime.

**Theorem 18.** Let \( G \in \mathbb{R}^m \), and suppose that the leading terms of \( G \) satisfy the extended criterion. Then \( (A) \Rightarrow (B) \), where:

(A) \( S_\prec(g_1, g_2), S_\prec(g_2, g_3), \ldots, S_\prec(g_{m-1}, g_m) \) all have \( S \)-representations with respect to \( G \);

(B) \( S_\prec(g_1, g_m) \) has an \( S \)-representation with respect to \( G \).

**Proof.** Assume \( (A) \). We want to show \( (B) \). For the sake of convenience, denote \( \text{lt}_\prec(g_1) \) by \( t_1 \).

Recall that

\[
S_\prec(g_1, g_m) = \text{lcm}(g_m) \cdot \frac{\text{lcm}(t_1, t_m)}{t_1} \cdot g_1 - \text{lcm}(g_1) \cdot \frac{\text{lcm}(t_1, t_m)}{t_m} \cdot g_m.
\]

Let \( p = \gcd(g_1, g_m) \), where \( \text{lcm}(p) = 1 \). Put \( f_1 = g_1/p \) and \( f_m = g_m/p \). From Lemma 16, we know that \( \gcd(\text{lt}_\prec(g_1), \text{lt}_\prec(g_m)) = \text{lt}_\prec(\gcd(g_1, g_m)) \). This and the facts that \( \text{lcm}(f_1) = \text{lcm}(g_1) \) and \( \text{lcm}(f_m) = \text{lcm}(g_m) \) give

\[
\text{lcm}(g_1) \cdot \frac{\text{lcm}(t_1, t_m)}{t_m} = \text{lcm}(g_1) \cdot \frac{t_1 t_m}{t_m \gcd(t_1, t_m)} = \text{lcm}(f_1) \cdot \text{lt}_\prec(f_1)
\]

and

\[
\text{lcm}(g_m) \cdot \frac{\text{lcm}(t_1, t_m)}{t_1} = \text{lcm}(g_m) \cdot \frac{t_1 t_m}{t_1 \gcd(t_1, t_m)} = \text{lcm}(f_m) \cdot \text{lt}_\prec(f_m).
\]
This allows us to rewrite (3.5) as

\[ S_<(g_1, g_m) = lc_<(f_m) \cdot lt_<(f_m) \cdot g_1 - lc_<(f_1) \cdot lt_<(f_1) \cdot g_m = p \cdot S_<(f_1, f_m). \]

By Corollary 17, the leading terms of \( f_1 \) and \( f_m \) are relatively prime; by Buchberger’s gcd criterion, \( S_<(f_1, f_m) \) has an \( S \)-representation \( h \). It follows that \( h_p = (h_1 p, \ldots, h_m p) \) is an \( S \)-representation of \( S_<(g_1, g_m) \).

Theorem 18 provides us with sufficient information to conclude that the main theorem is true. This may not be clear, because we have discussed only \( S \)-representations, and not reduction to zero. To show how the two come together, we need to recall two additional results. The first is the characterization of Gröbner bases due to Lazard [17].

**Theorem 19 (Lazard’s characterization).** Let \( G \in \mathcal{R}^m \). The following are equivalent:

(A) \( G \) is a Gröbner basis with respect to \( \prec \);

(B) for every \( i, j \) such that \( 1 \leq i < j \leq m \), \( S_<(g_i, g_j) \) has an \( S \)-representation with respect to \( G \).

It turns out that Buchberger’s characterization implies Lazard’s, thanks to the following lemma [3].

**Lemma 20.** Let \( G \in \mathcal{R}^m \) and let \( i, j \) satisfy \( 1 \leq i < j \leq m \). Then (A) \( \implies \) (B), where:

(A) \( S_<(g_i, g_j) \) reduces to zero with respect to \( G \);

(B) \( S_<(g_i, g_j) \) has an \( S \)-representation with respect to \( G \).

However, the converse of Lemma 20 is known to be false, so the fact that Lazard’s characterization implies Buchberger’s is not obvious. It depends on the fact that in Lazard’s characterization, every pair \( (i, j) \) has an \( S \)-representation for \( S_<(g_i, g_j) \), whereas Lemma 20 deals only with one \( S \)-representation.

We can now show how Theorem 18 proves the main theorem.

**Proof of the main theorem.** That (A) implies (B) is trivial, so we assume (B) and show (A). To prove (A), we will employ Lazard’s characterization.

From (B), every pair \( (i, j) \) satisfies one of (B0)—(B3). Let \( i, j \) be such that \( 1 \leq i < j \leq m \). Clearly, \( S_<(g_i, g_j) \) has an \( S \)-representation:

- if \( (i, j) \) satisfies (B0), then, by Lemma 20;
- if \( (i, j) \) satisfies (B1) or (B2), then, by well-known results [1, 3, 10];
- if \( (i, j) \) satisfies (B3), then, by Theorem 18.

By Lazard’s characterization (Theorem 19), \( G \) is a Gröbner basis with respect to \( \prec \).

4. ‘Pham-like’ systems

In this section, we describe a class of polynomial systems for which the extended criterion provides a dramatic reduction in the number of \( S \)-polynomial computations required for verification (Corollary 23).

A well-studied system of polynomials is the Pham system [9, Chapter 6, p. 147].

**Definition 21 (Pham system).** Let \( P \in \mathbb{F}[x_1, x_2, \ldots, x_n]^n \). We say that \( P \) is a Pham system if \( \text{lt}_<(p_i) \) and \( \text{lt}_<(p_j) \) are relatively prime whenever \( i \neq j \).
Thanks to Theorem 3, one can verify that any Pham system is a Gröbner basis without checking any $S$-polynomials at all. Now we obfuscate matters somewhat through multiplication.

**Definition 22 (Pham-like systems).** Suppose that $G = (g_1, \ldots, g_m)$ has leading terms $(c_1d, \ldots, c_md)$, where, for all $i = 1, \ldots, m$,
- $c_i$ and $d$ are relatively prime; and
- for all $j \neq i$, $c_i$ and $c_j$ are relatively prime.
We call such a $G$ a Pham-like system.

Consider the following question.

*Is a Pham-like system a Gröbner basis?*

The temptation may arise to answer in the affirmative, because the cofactors of the leading terms’ gcd are relatively prime, which through some manipulation might allow Buchberger’s gcd criterion to apply. It does not. Numerous systems are not Gröbner bases even though this property is true; for example,

$$g_1 = xy + y, \quad g_2 = xz.$$ 

So, deciding whether $G$ is a Gröbner basis requires us to check whether the $S$-polynomials reduce to zero. We would like to avoid checking all of them if possible.

To that end, we turn first to Buchberger’s criteria, but:
- none of the leading terms $c_id, c_jd$ are relatively prime; and
- for any pair $c_id$ and $c_jd$, no $c_kd$ divides their lcm.

If we were to rely only on Buchberger’s criteria, we would have to reduce all $m(m - 1)/2$ $S$-polynomials to zero to see that a Pham-like system is a Gröbner basis.

However, the extended criterion allows us to decide whether a Pham-like system is a Gröbner basis by checking at most $m - 1$ $S$-polynomials, even though Buchberger’s criteria provide no benefit.

**Corollary 23.** Let $G \in \mathbb{R}^m$ be a Pham-like system. The following are equivalent:
(A) $G$ is a Gröbner basis with respect to $\prec$;
(B) the $S$-polynomials $S_<(g_1, g_2), S_<(g_2, g_3), \ldots, S_<(g_{m-1}, g_m)$ reduce to zero with respect to $G$.

*Proof.* That (A) implies (B) is trivial, so we assume (B) and show (A). From (B), we know that $S_<(g_1, g_2), S_<(g_2, g_3), \ldots, S_<(g_{m-1}, g_m)$ reduce to zero with respect to $G$. It follows from Lemma 20 that they have $S$-representations with respect to $G$.

For the sake of convenience, denote $\text{lt}_<(g_i)$ by $t_i$. Write $t_i = c_it$, where $c_i$ and $d$ are as in Definition 22. Recall that $\gcd(c_i, t_j) = 1$ whenever $i \neq j$; inspection shows that the list of terms $(t_1, t_2, \ldots, t_m)$ satisfies the extended criterion. By Theorem 18, $S_<(g_1, g_m)$ has an $S$-representation with respect to $G$. Let $p_{1,m} = \gcd(g_1, g_m)$ and choose $f_1, f_m \in \mathbb{R}$ such that:
- $g_1 = f_1p_{1,m}$; and
- $g_m = f_mp_{1,m}$.

Recall Lemma 16 and the assumption that $c_1$ is relatively prime to $t_m$; then

$$d = \gcd(c_1d, c_md) = \gcd(\text{lt}_<(g_1), \text{lt}_<(g_m)) = \text{lt}_<(p_{1,m}).$$

Thus,

$$c_1d = t_1 = \text{lt}_<(g_1) = \text{lt}_<(f_1p_{1,m}) = \text{lt}_<(f_1) \text{lt}_<(p_{1,m}) = \text{lt}_<(f_1) d,$$

whence $c_1 = \text{lt}_<(f_1)$. Similarly, $c_m = \text{lt}_<(f_m)$.
Inspection shows that the list of terms \((t_1, t_m, t_{m-1}, \ldots, t_3, t_2)\) also satisfies the extended criterion. We now know that \(S_<(g_1, g_m)\) has an \(S\)-representation with respect to \(G\), so we can reason as before that there exist \(\varphi_1, \varphi_2, p_{1,2} \in R\) such that:

- \(g_1 = \varphi_1 p_1 2\);
- \(g_2 = \varphi_2 p_2 12\);
- \(p_{1,2} = \gcd(g_1, g_2)\); and
- the leading terms of \(\varphi_1\) and \(\varphi_2\) are relatively prime.

As before, we obtain \(d = \text{lt}_<(p_{1,2})\) and \(c_1 = \text{lt}_<(\varphi_1)\). Thus, \(\text{lt}_<(f_1) = \text{lt}_<(\varphi_1)\). We claim that in fact \(f_1 = \varphi_1\). By way of contradiction, assume that \(f_1 = \varphi_1\) and \(\varphi_1\) are not equal. From \(f_1 p_{1,m} = \varphi_1 p_{1,2}\), we conclude that \(f_1\) has a common factor with \(p_{1,2}\) or \(\varphi_1\) has a common factor with \(p_{1,m}\); but this contradicts the hypothesis that \(c_1\) is relatively prime to \(d\). Hence, \(f_1 = \varphi_1\) and \(p_{1,m} = p_{1,2}\). Write \(p = p_{1,m}, g_1 = f_1 p, g_2 = f_2 p,\) and \(g_m = f_m p\).

Proceeding in like fashion, we can factor every \(g_i\) as \(g_i = f_i p\) such that \(\text{lt}_<(f_i)\) and \(\text{lt}_<(f_j)\) are relatively prime whenever \(i \neq j\). By Theorem 3, \(F = (f_1, f_2, \ldots, f_m)\) is a Gröbner basis with respect to \(\prec\). Let \(i, j\) be arbitrary, but fixed. Assume \(1 \leq i < j \leq m\). By Lazard’s characterization, \(S_<(f_1, f_j)\) has an \(S\)-representation \(h^{(i,j)}\). This implies that \(S_<(g_i, g_j)\) has an \(S\)-representation \(p h^{(i,j)} = (p h_1^{(i,j)}, \ldots, p h_m^{(i,j)})\). Since \(i\) and \(j\) are arbitrary, by Lazard’s characterization \(G\) is a Gröbner basis with respect to \(\prec\).

Acknowledgements. The author would like to thank the anonymous referees whose comments greatly improved the quality of the paper.

Part of this work was conducted during the Special Semester on Gröbner Bases, 1 February–31 July 2006, organized by RICAM, Austrian Academy of Sciences, and RISC, Johannes Kepler University, Linz, Austria.

References

1. W. Adams and P. Loustaunau, An introduction to Gröbner bases, Graduate Studies in Mathematics 3 (American Mathematical Society, Providence, RI, 1994).
2. J. Backelin and R. Fröberg, ‘How we proved that there are exactly 924 cyclic-7 roots’, ISSAC ’91: Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation (ACM Press, New York, NY, 1991) 103–111, ISBN:0-89791-437-6.
3. T. Becker, V. Weispfenning and H. Kredel, Gröbner bases: a computational approach to commutative algebra (Springer, New York, 1993).
4. B. Buchberger, ‘Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal (an algorithm for finding the basis elements in the residue class ring modulo a zero dimensional polynomial ideal)’. PhD Thesis, Mathematical Institute, University of Innsbruck, Austria, 1965. English translation published in J. Symbolic Comput. 41 (2006) 475–511.
5. B. Buchberger, ‘An algorithmic criterion for the solvability of a system of algebraic equations’, Aequationes Math. 4 (1970) 374–383 Gröbner Bases and Applications, London Mathematical Society Lecture Note Series 251 (1998) (English translation).
6. B. Buchberger, ‘A criterion for detecting unnecessary reductions in the construction of Gröbner bases’, Proceedings of the EUROSAM 79 Symposium on Symbolic and Algebraic Manipulation, Marseille, 26–28 June 1979, Lecture Notes in Computer Science 72 (ed. E. W. Ng; Springer, Berlin, 1979) 3–21.
7. B. Buchberger, ‘Gröbner-bases: an algorithmic method in polynomial ideal theory’, Multidimensional systems theory — progress, directions and open problems in multidimensional systems, (ed. N. K. Bose; Reidel, Dordrecht, 1985) 184–232.
8. M. Caboara, M. Kreuzer and L. Robbiano, ‘Minimal sets of critical pairs’, Proceedings of the First Congress of Mathematical Software, (eds B. Cohen, X. Gao and N. Takayama; World Scientific, Singapore, 2002) 390–404.
9. A. M. Cohen, H. Cuypers and H. Sterk (eds) Some tapas of computer algebra, Algorithms and Computation in Mathematics 4 (Springer, Berlin, 1999).
10. D. Cox, J. Little and D. O’Shea, Ideals, varieties, and algorithms, 2nd edn (Springer, New York, 1997).
11. D. Cox, J. Little and D. O’Shea, Using algebraic geometry (Springer, New York, 1998).
12. R. Gebauer and H. Möller, ‘On an installation of Buchberger’s algorithm’, J. Symbolic Comput. 6 (1988) 275–286.
13. H. Hong and J. Perry, ‘Are Buchberger’s criteria necessary for the chain condition?’, J. Symbolic Comput. 42 (2007) 717–732.
14. C. G. J. Jacobi, ‘De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant; una cum variis theorematibus de transformatione et determinatione integralium multiplicium’, *J. Reine Angew. Math.* 12 (1833) 1–69.

15. C. Kollreider and B. Buchberger, ‘An improved algorithmic construction of Gröbner-bases for polynomial ideals’, *ACM SIGSAM Bull.* 12 (1978) 27–36.

16. M. Kreuzer and L. Robbiano, *Computational commutative algebra I* (Springer, Heidelberg, 2000).

17. D. Lazard, ‘Gröbner bases, Gaussian elimination, and resolution of systems of algebraic equations’, *EUROCAL ’83, European Computer Algebra Conference*, Lecture Notes in Computer Science 162 (ed. J. A. van Hulzen; Springer, Berlin, 1983) 146–156.

18. H. Möller, F. Mora and C. Traverso, ‘Gröbner bases computation using syzygies’, *Proceedings of the 1992 International Symposium on Symbolic and Algebraic Computation*, Association for Computing Machinery (ed. P. S. Wang; ACM Press, New York, NY, 1992) 320–328.

19. A. Rice and E. Torrence, ‘Shutting up like a telescope: Lewis Carroll’s curious condensation method for evaluating determinants’, *College Math. J.* 38 (2007) 85–95.

John Perry  
Department of Mathematics  
The University of Southern Mississippi  
118 College Drive, Box #5045  
Hattiesburg, MS 39406  
USA  
john.perry@usm.edu