Gröbner Bases Theory and The Diamond Lemma

by

Wenfeng Ge

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Commutative Gröbner bases theory is well known and widely used. In this thesis, we will discuss thoroughly its generalization to noncommutative polynomial ring $k<X>$ which is also an associative free algebra. We introduce some results on monomial orders due to John Lawrence and the author. We show that a noncommutative monomial order is a well order while a one-sided noncommutative monomial order may not be. Then we discuss the generalization of polynomial reductions, S-polynomials and the characterizations of noncommutative Gröbner bases. Some results due to Mora are also discussed, such as the generalized Buchberger's algorithm and the solvability of ideal membership problem for homogeneous ideals. At last, we introduce Newman's diamond lemma and Bergman's diamond lemma and show their relations with Gröbner bases theory.

Acknowledgments

I would like to give sincere thanks to my supervisor Professor John Lawrence who gave me a lot of advice and great ideas during my work on this thesis. Also I would like to thank my thesis readers Professor Ross Willard and Professor Peter Hoffman for taking time to read my thesis and giving me helpful comments and suggestions. Thanks also go to Professor Che Tat Ng, Shonn Martin and Lis D'Alessio for all the kindness and help they gave me.

I would also like to thank Kai Cheong Chan, Denglin Zhou and his wife Ying Li, Shengli Wu for their help and friendship during my three years at Waterloo.

Finally I would like to thank my family: my grandfather, my parents, my wife and my little sister. Without their love and support, I would never have achieved my goals.

Dedication

This thesis is dedicated to my wife Yali Hu and our coming baby.

Contents

Chapter 1

Introduction

Gröbner bases and Buchberger's algorithm were introduced by B.Buchberger in 1965[2]. Today they are well-known and widely applied to many problems in mathematics, computer science and engineering. For a basic example in commutative algebra, ideal membership problem for commutative polynomial rings, or equivalently saying, word problem for commutative algebra presentations can be solved by Gröbner bases theory(see section 2.5). Since Buchberger's Gröbner bases theory mostly concerns commutative algebra, we call it *commutative Gröbner bases theory*.

In 1978, G.M.Bergman introduced his diamond lemma for ring theory [3], which is an analogue and strengthening of Newman's diamond lemma [5]. As T.Mora has pointed out in [9], Bergman's diamond lemma essentially contains a generalization of commutative Gröbner bases theory to general noncommutative polynomial rings which are also associative free algebras. In [9](1986) and [10](1994), T.Mora made the generalization precise. In this thesis, we call this generalization[1] *noncommutative Gröbner bases theory*.

[1]There are different generalizations of Gröbner bases theory to noncommutative areas. See [1] and "introduction" in [4].

The aim of this thesis is to discuss thoroughly the above noncommutative Gröbner bases theory and show explicitly the relations among commutative Gröbner bases theory, noncommutative Gröbner bases theory, Bergman's diamond lemma and Newman's diamond lemma.

The thesis is organized as follows.

In chapter 2, we give a brief introduction to commutative Gröbner bases theory as the background. Most important definitions, results and algorithms of the theory are included but some proofs are omitted. Interested readers are referred to [12] [8] [7] for more information on this theory.

In chapter 3, we generalize the definitions, results and algorithms given in chapter 2 to general noncommutative polynomial rings. Most results are based on Mora's work [9] [10], but we give more complete proofs of the results and explain more details of the generalization such as non-commutative polynomial reductions and noncommutative S-polynomials. In particular, we believe the results on monomial orders are new and they are due to my supervisor Prof. John Lawrence. We show a result about monomial partial order and then we prove that a noncommutative monomial order is a well order. We also give an example which shows that a one-sided noncommutative monomial order may not be a well order.

In chapter 4, we introduce Newman's diamond lemma firstly and then Bergman's diamond lemma. After that we show the relation between Gröbner bases theory and diamond lemma(s). We give a brief comment on the relation between Gröbner bases theory and Newman's diamond lemma and then deduce most characterizations of noncommutative Gröbner bases from Bergman's diamond lemma.

We need point out that, the emphasis of this thesis is on theoretical aspect not on computational aspect, although the latter is also very important, especially in practice. All the algorithms in this thesis are only explanatory, not written in formal programming languages. Topics on how to improve the efficiency of related algorithms are not covered. Readers are assumed to have a basic knowledge of rings(especially polynomial rings), vector spaces, modules and algebras.

Chapter 2

Commutative Gröbner Bases Theory

2.1 Notations and Basic Definitions

We let $\mathbb{N}$ denote the set of natural numbers with $0 \in \mathbb{N}$. Let k be a field, $k[x_1, x_2, \ldots, x_n]$ denote the commutative polynomial ring in n variables over k, $n \in \mathbb{N} - \{0\}$. For $k[x_1, x_2, \ldots, x_n]$, the following facts are known:

$$(1)\ \forall f \in k[x_1, x_2, \ldots, x_n] - \{0\}, f = \sum_{i=1}^{t} c_i {x_1}^{\beta_{i_1}} {x_2}^{\beta_{i_2}} \ldots {x_j}^{\beta_{i_j}} \ldots {x_n}^{\beta_{i_n}},$$

where $t \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, \beta_{i_j} \in \mathbb{N}, 1 \leqslant j \leqslant n, 1 \leqslant i \leqslant t$. Conventionally, $c_i {x_1}^{\beta_{i_1}} {x_2}^{\beta_{i_2}} \ldots {x_j}^{\beta_{i_j}} \ldots {x_n}^{\beta_{i_n}}$ is called a *term*, c_i is called the *coefficient* of the term, ${x_1}^{\beta_{i_1}} {x_2}^{\beta_{i_2}} \ldots {x_j}^{\beta_{i_j}} \ldots {x_n}^{\beta_{i_n}}$ is called a *monomial*, $\sum_{j=1}^{n} \beta_{i_j}$ is called the *degree* of the monomial. For any monomial m, the degree of m is denoted by $deg(m)$. The set of all monomials in n variables is denoted by M_n or simply M when n is not necessary to be indicated. Note that $\forall j$ with $1 \leqslant j \leqslant n$, we let ${x_j}^0 = 1 \in M$. Given $m_1, m_2 \in M$, if $\exists m_3 \in M$ such that $m_2 = m_3 m_1$, we say m_2 is a *multiple* of m_1, or m_1 *divides* m_2, denoted by $m_1 \mid m_2$.

(2) $\forall f \in k[x_1, x_2, \ldots, x_n] - \{0\}$, after all possible coalescence and cancel-

lation of terms, f has the *uniqnue form* as follows,

$$f = \sum_{i=1}^{t} c_i m_i, \tag{2.1.1}$$

where $t \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, m_i \in M$ and $m_i \neq m_j \; \forall 1 \leqslant i \neq j \leqslant t$. Here, the uniqueness is up to a permutation on the terms in the form.

Next, before introducing the definition of monomial order, let's look at some prerequisite definitions.

Definitions 2.1.1. (i) Let S be a nonempty set, $S \times S$ denote the set of all ordered pairs (a, b) of elements a,b in S. A subset R of S is called a (binary) *relation* on S. Usually, when $(a, b) \in R$, we write aRb.

(ii) Relation $\preceq$ on S is called a *partial order* if it satisfies the following properties:

- $reflexivity : a \preceq a, \; \forall a \in S$;
- $transitivity : a \preceq b \; and \; b \preceq c \; \Rightarrow \; a \preceq c, \quad \forall a, b, c \in S$;
- $antisymmetry : a \preceq b \; and \; b \preceq a \; \Rightarrow \; a = b, \; \forall a, b \in S.$

The *strict* part of $\preceq$, denoted by $\prec$, is defined by $a \prec b \; \Leftrightarrow \; a \preceq b \; and \; a \neq b$. The *inverse* of $\preceq$, denoted by $\succeq$, is defined by $a \succeq b \; \Leftrightarrow \; b \preceq a$.

(iii) A partial order on S is said to be a *total order* ,usually denoted by $\leq$, if it satisfies: $\forall a, b \in S, \; a \leq b \; or \; b \leq a$. A total order $\leq$ on S is said to be a *well order* if it satisfies descending chain condition(DCC), $i.e.$,there is no infinite strictly descending chain $a_1 > a_2 > \ldots$ in S with respect to(w.r.t.) $\leq$.

(iv) Let $\preceq$ be a partial order on S, $T \subseteq S$, if for some $t \in T$, $t \preceq a \; \forall a \in T$, we say t is a least element of T. Then a *well order* on S is also defined by a partial order $\preceq$ with each nonempty subset of S having a least element w.r.t. $\preceq$.

Remarks 2.1.2. (i) For the proof of equivalence of two definitions of well order, see [8]. (ii) Usually, S is said to be partially ordered(totally ordered, well ordered) by $\preceq$, if $\preceq$ is a partial order(total order,well order) defined on S. The ordered set S w.r.t. $\preceq$ is denoted by $(S, \preceq)$.

Now let's return to $k[x_1, x_2, \ldots, x_n]$.

Definition 2.1.3. $\leq$ is said to be a *monomial order* on the set of all monomials M, if it satisfies the following conditions:

(i) M is totally ordered by $\leq$;

(ii) $1 \leq m$, $\forall m \in M$;

(iii) $m_1 \leq m_2 \ \Rightarrow \ mm_1 \leq mm_2$, $\forall m, m_1, m_2 \in M$.

Let $\leq$ be a monomial order on M, suppose $m_1, m_2, m_3 \in M$, and $m_2 = m_3 m_1$, by the above condition(ii) $1 \leq m_3$, then by the condition(iii) $m_1 \leq m_3 m_1 = m_2$. Hence we can say $m_1 | m_2 \Rightarrow m_1 \leq m_2$. This shows the monomial order relation is an extension of the division relation.

Examples 2.1.4. Let $<_N$ be the natural order on $\mathbb{N}$. The following are three frequently used monomial orders. (Verifications of conditions (i)(ii)(iii) in the above definition are omitted.)

(1)The lexicographical order(abbreviated as *lex*) on M with $x_1 > x_2 > \ldots > x_n$. In the lex, ${x_1}^{\alpha_1} {x_2}^{\alpha_2} \ldots {x_n}^{\alpha_n} < {x_1}^{\beta_1} {x_2}^{\beta_2} \ldots {x_n}^{\beta_n} \Leftrightarrow \alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_l = \beta_l, \alpha_{l+1} <_N \beta_{l+1}$, for some l.

(2)The degree lexicographical order(abbreviated as *deglex*) on M with $x_1 > x_2 > \ldots > x_n$. In the deglex, for all m_1, m_2 in M, $m_1 < m_2 \Leftrightarrow$ either $deg(m_1) <_N deg(m_2)$ or $deg(m_1) = deg(m_2)$ *and* $m_1 <_{lex} m_2$, where

$<_{lex}$ is the lexicographical order with $x_1 > x_2 > \ldots > \ x_n$.

(3)The degree reverse lexicographical order(abbreviated as *degrevlex*) on M with $x_1 > x_2 > \ldots > \ x_n$. In the degrevlex, let $m_1 = {x_1}^{\alpha_1}{x_2}^{\alpha_2}\ldots{x_n}^{\alpha_n}$, $m_2 = {x_1}^{\beta_1}{x_2}^{\beta_2}\ldots{x_n}^{\beta_n}$, then $m_1 < m_2 \Leftrightarrow$ either $deg(m_1) <_N deg(m_2)$ or $deg(m_1) = deg(m_2)$ *and* $\alpha_n = \beta_n, \alpha_{n-1} = \beta_{n-1}, \ldots, \alpha_l = \beta_l, \alpha_{l-1} >_N \beta_{l-1}$, for some l.

Given any nonzero polynomial f, f can be written in the unique form (2.1.1). Now let M be totally ordered by some monomial order $\leq$, clearly there is a permutation on all terms in (2.1.1) such that $f = \sum_{i=1}^{t} c_i m_i$, and $m_1 > m_2 > \ldots > m_t$. In this case, we call $c_1 m_1$ the *leading term* of f, denoted by $lt(f)$; m_1 the *leading monomial* of f, denoted by $lm(f)$; c_1 the *leading coefficient* of f, denoted by $lc(f)$.

Definition 2.1.5. Given any subset G of $k[x_1, x_2, \ldots, x_n]$, we define the *leading monomial ideal* of G w.r.t. some monomial order $\leq$ to be

$$\begin{aligned} lm(G) := & < lm(g) |\ g \in G > \\ = & \{\sum_{i=1}^{t} f_i lm(g_i) |\ t \in \mathbb{N} - \{0\}, f_i \in k[x_1, x_2, \ldots, x_n], g_i \in G\} \end{aligned}$$

Definition 2.1.6. Given a monomial order on M, let G be a finite subset of $k[x_1, x_2, \ldots, x_n]$, if $lm(G) = lm(< G >)$, we say G is a *Gröbner basis* (of the ideal $< G >$). If a finite set $G \subseteq$ ideal $I \subseteq k[x_1, x_2, \ldots, x_n]$ and $lm(G) = lm(I)$, G is called a *Gröbner basis of I*.

Remarks 2.1.7. (i) We do not need $I =< G >$ in the above definition. Instead, we will prove $lm(G) = lm(I) \Rightarrow I =< G >$ in theorem 2.2.11. (ii) Gröbner basis has different characterizations (see section 2.4) and every characterization can work as the definition.

Next, let's discuss some fundamental results.

2.2 Noetherian Rings and Dickson's Lemma

We start from general commutative rings. Let R always denote a commutative ring in this section.

Definition 2.2.1. Ring R is said to be *Noetherian* if it satisfies the ascending chain condition(ACC) on ideals $i.e.$,there is no infinite properly ascending chain of ideals $I_1 \subsetneq I_2 \subsetneq \ldots$ in R.

Definition 2.2.2. An ideal I of ring R is said to be *finitely generated*, if $\exists a_1, a_2, \ldots, a_s \in I$, such that $I =< a_1, a_2, \ldots, a_s >= \{\sum_{i=1}^{s} r_i a_i \mid r_i \in R\}$.

Lemma 2.2.3. Ring R is Noetherian iff every ideal of R is finitely generated.

Proof:"$\Rightarrow$" Zero ideal is trivially finitely generated. Suppose a nonzero ideal I is not finitely generated. Select $a_1 \in I$, then $< a_1 > \subsetneq I$. Next select $a_2 \in I- < a_1 >$, we see $< a_1 > \subsetneq < a_1, a_2 > \subsetneq I$. Then we can select $a_3 \in I- < a_1, a_2 >, \ldots$ Obviously, since I is not finitely generated, the process can be continued without termination. Hence we would have an infinite ascending chain of ideals $< a_1 > \subsetneq < a_1, a_2 > \subsetneq \ldots$ in R. But R is Noetherian, a contradiction. So every ideal of R is finitely generated.

"$\Leftarrow$" Suppose R is not Noetherian, then there is an infinite chain $I_1 \subsetneq I_2 \subsetneq \ldots$ in R. Let $I = \bigcup_{i=1}^{\infty} I_i$, it's easy to see I is also an ideal of R. Then $\exists a_1, a_2, \ldots, a_s \in I$, such that $I =< a_1, a_2, \ldots, a_s >$. Notice that for some sufficient large l, $a_1, a_2, \ldots, a_s \in I_l$, then $I \subseteq I_l \subsetneq I_{l+1} \subseteq I$. It's a contradiction. Therefore R is Noetherian. $\Box$

The next theorem is well-known and a complete proof for it is not short, thus the proof is omitted here. Readers can refer to [11] or [12] for the proof.

Theorem 2.2.4.(Hilbert Basis Theorem) If R is Noetherian, so is $R[x]$.

Notice that $k[x_1, x_2, \ldots, x_n]= k[x_1, x_2, \ldots, x_{n-1}][x_n]$ $\forall n \in \mathbb{N}-\{0\}$ and the field k is trivially noetherian, by applying Hilbert Basis Theorem recursively, we can see the polynomial ring $k[x_1, x_2, \ldots, x_n]$ is Noetherian $\forall n \in \mathbb{N} - \{0\}$. By combining this result with lemma 2.2.3, we have the following theorem.

Theorem 2.2.5. $\forall n \in \mathbb{N} - \{0\}$, the polynomial ring $k[x_1, x_2, \ldots, x_n]$ is Noetherian and every ideal of $k[x_1, x_2, \ldots, x_n]$ is finitely generated.

Corollary 2.2.6. Let I be a nonzero ideal of $k[x_1, x_2, \ldots, x_n]$, suppose I is generated by a nonempty set S, then $\exists$ finite $S' \subseteq S$ such that $I \ = <\ S >$ $= <\ S' >$.

Proof: By theorem 2.2.5, $\exists a_1, a_2, \ldots, a_t \in I$, such that $I =< a_1, a_2, \ldots, a_t >$. Since $I =< S >$, $\exists$ finite $S_i \subseteq S$ such that $a_i \in< S_i >, \forall i$. Then let $S' = \bigcup_{i=1}^{t} S_i$, clearly finite $S' \subseteq S$ and $I =< a_1, a_2, \ldots, a_t > \subseteq < S' > \subseteq < S >= I$. We're done. $\square$

If we let the set S in corollary 2.2.6 contain only monomials, we have a result called Dickson's Lemma.

Theorem 2.2.7.(Dickson's Lemma) Let S be a nonempty set of monomials in $k[x_1, x_2, \ldots, x_n]$, then $\exists$ finite $S' \subseteq S$ such that $< S >=< S' >$, or equivalently saying, $\exists$ finite $S' \subseteq S$, $\forall m \in S, \exists m' \in S'$ such that m is a multiple of m'.

Proof: The first statement $< S >=< S' >$ is obvious by the corollary 2.2.6. Notice the fact that

$$m \in< S' > \quad \Rightarrow \quad m = \sum_{i=1}^{t} c_i a_i m_i$$

where a_i, m_i are monomials, $m_i \in S'$, $c_i \in k$ and $c_i = 0$ except for those i with $a_i m_i = m$, then we can see $m \in< S' > \Leftrightarrow m$ is a multiple of some member m' of S'. Therefore the second statement is equivalent to $< S >=< S' >$. $\square$

Next we prove some results based on Dickson's Lemma.

Theorem 2.2.8.(Existence of Gröbner Bases for Ideals) For any nonzero ideal I of $k[x_1, x_2, \ldots, x_n]$, given any monomial order $\leq$ on M, there exists some finite $G \subseteq I$ such that $lm(G) = lm(I)$

Proof: By definition, $lm(I) =< S >$ where $S = \{lm(f) \mid f \in I\}$. By Dickson's Lemma, $\exists$ finite $S' = \{lm(f_1), lm(f_2), \ldots, lm(f_l)\} \subseteq S$ such that $< S >=< S' >$. Let $G = \{f_1, f_2, \ldots, f_l\}$, clearly $G \subseteq I$ and $lm(G) = lm(I)$. $\square$

Theorem 2.2.9. Every monomial order $\leq$ on M is a well order.

Proof: Let S be a nonempty subset of M, we will show S has a least element w.r.t. $\leq$, then by the definition 2.1.1(iv), $\leq$ is a well order.

By Dickson's Lemma, $\exists$ finite $S' \subseteq S$, $\forall m \in S, \exists m' \in S'$ such that m is a multiple of m'. Because S' is finite, $\leq$ is known to be a total order, then S' has a least element m_0. Now $\forall m \in S, \exists m' \in S'$,such that $m = m'h$, where $h \in M$. Clearly $h \geq 1$ by condition(ii) in the definition 2.1.3. By condition (iii), $m = m'h \geq m' \geq m_0$. Hence m_0 is a least element of S. We're done. $\square$

Remark 2.2.10. In chapter 3 we will give another proof of the above theorem that does not use the Dickson's Lemma or Hilbert Basis Theorem.

Theorem 2.2.11. Let $\leq$ be a monomial order on the set of all monomials M and let I be an ideal of $k[x_1, x_2, \ldots, x_n]$. If a finite set G satisfies $G \subseteq I$ and $lm(G) = lm(I)$ (so G is a Gröbner basis of the ideal I), then $I =< G >$.

Proof: Given any $f \in I - \{0\}$, do the following process: Let $f_1 = f,\ f_1 \in I,\ lm(f_1) \in lm(I) = lm(G) \Rightarrow \exists g_1 \in G \subseteq I,\ q_1 \in M,\ c_1 \in k - \{0\}$, such that $lt(f_1) = lt(c_1 g_1 q_1)$. Let $f_2 = f_1 - c_1 g_1 q_1$, clearly $f_2 \in I$, when $f_2 \neq 0,\ lm(f_2) \in lm(I) = lm(G) \Rightarrow \exists g_2 \in G \subseteq I,\ q_2 \in M,\ c_2 \in k - \{0\}$, such that $lt(f_2) = lt(c_2 g_2 q_2)$. Let $f_3 = f_2 - c_2 g_2 q_2$, clearly $f_3 \in I$, and when

$f_3 \neq 0$, we can move to $f_4 \ldots$

Notice that $lm(f_1) > lm(f_2) > lm(f_3) > \ldots$, the process must terminate, since the monomial order $\leq$ is a well order. Moreover, the last f_l must be 0, otherwise, we would be able to continue the process to f_{l+1}. Hence we have

$$\begin{aligned} 0 &= f_l = f_{l-1} - c_{l-1}g_{l-1}q_{l-1} = f_{l-2} - c_{l-2}g_{l-2}q_{l-2} - c_{l-1}g_{l-1}q_{l-1} = \ldots \\ &= f_1 - \sum_{i=1}^{l-1} c_i g_i q_i. \end{aligned}$$

So $f = f_1 = \sum_{i=1}^{l-1} c_i g_i q_i \in < G >$. This implies $I \subseteq < G >$. It's obvious that $< G > \subseteq I$. Therefore $I =< G >$. □

At last we point out the following fact about the Dickson's Lemma and the theorem 2.2.5.

Claim 2.2.12. Theorem 2.2.5 $\Leftrightarrow$ Dickson's Lemma.

Proof: "$\Rightarrow$" We have shown that Theorem 2.2.5$\Rightarrow$Corollary 2.2.6$\Rightarrow$Theorem 2.2.7(Dickson's Lemma).

"$\Leftarrow$" We can define some monomial order on M. Notice that Dickson's Lemma$\Rightarrow$Theorem 2.2.8(Existence of Gröbner Bases for Ideals). Also Dickson's Lemma$\Rightarrow$Theorem 2.2.9$\Rightarrow$Theorem 2.2.11. Therefore, every ideal of $k[x_1, x_2, \ldots, x_n]$ is finitely generated (by its Gröbner basis). The theorem 2.2.5 is proved. □

2.3 Polynomial Reduction

We assume a monomial order $\leq$ has been defined on M in the following discussion.

Definition 2.3.1. For any $f, g \in k[x_1, x_2, \ldots, x_n]$, if $lm(g)$ divides some nonzero term cm in f, let $h = f - \frac{cm}{lt(g)} g$, then it's easy to see the term cm

in f is replaced by a linear combination of monomials $< m$. We call this manipulation a *polynomial reduction*, denoted by $f \xrightarrow{g} h$, and say f *reduces to h modulo g.*

Definition 2.3.2. If there is a finite sequence of polynomial reductions $f \xrightarrow{g_1} h_1 \xrightarrow{g_2} h_2 \ldots \xrightarrow{g_t} h_t$, where $g_i \in$ finite set $G \subseteq k[x_1, x_2, \ldots, x_n]$ and g_i not necessarily pairwise distinct for $1 \leqslant i \leqslant t$, we say f *reduces to h_t modulo G*, denoted by $f \xrightarrow{G} h_t$.

Definition 2.3.3. Polynomial r is called *reduced* or *in the reduced form* w.r.t. some finite set $G \subseteq k[x_1, x_2, \ldots, x_n]$, if $r = 0$ or no monomial occurring in the unique form (2.1.1) of r is divisible by $lm(g)$ $\forall g \in G$. If $f \xrightarrow{G} r$ and r is reduced w.r.t. G, we say r is *a reduced form of f* w.r.t. G. When the reduced form of f w.r.t. G is unique, we denote it by $R(f, G)$. In particular, when f is reduced w.r.t. G, $R(f, G) = f$.

Remarks 2.3.4. (i) In our definition of polynomial reductions(Definition 2.3.1), g is assumed in the unique form (2.1.1) but f is not. In general we do NOT require the polynomial which is to be reduced is given in the unique form. But in the case we need consider the leading monomial of f (such as the case in the following algorithm 2.3.6), clearly f will be assumed in the unique form.

(ii) The equivalence "monomial $m \in lm(G) \Leftrightarrow \exists g \in G,\ lm(g)|\ m$" is often used in our discussion. For example, when we say $r \neq 0$ is reduced w.r.t. G, it is equivalent to say no monomial in the unique form of r is in $lm(G)$.

(iii) Given a finite set $G \subseteq k[x_1, x_2, \ldots, x_n]$, we let $M(G)$ denote the set of all monomials in $lm(G)$,*i.e.*,$M(G) = M \bigcap lm(G)$, let $k_R(G)$ denote the set of all reduced polynomials w.r.t. G, then it's easy to verify that $k_R(G) =$

$span_k\{M - M(G)\}$ as a k-vector space.

By the above definitions, the following proposition is obvious.

Proposition 2.3.5. Given polynomials f and r, if r is a reduced form of f w.r.t. some G, then either $f = r$ or $\exists s \in \mathbb{N} - \{0\}, c_u \in k - \{0\}, m_u \in M, g_u \in G$ and not necessarily pairwise distinct $\forall\, u,\ 1 \leqslant u \leqslant s$, such that

$$f = \sum_{u=1}^{s} c_u g_u m_u + r. \tag{2.3.1}$$

Algorithm 2.3.6.(Reduction Algorithm) Given a nonzero polynomial f and a finite $G = \{g_1, g_2, \ldots, g_j, \ldots, g_l\} \subseteq k[x_1, x_2, \ldots, x_n]$, the following algorithm provides one way to compute a reduced form of f w.r.t. G.

$i := 1, r := 0, f_i := f$

$(*)$while $f_i \neq 0$ *do*

if $\exists lm(g_j)|\ lm(f_i)$, *choose the least* j *such that* $lm(g_j)|\ lm(f_i)$ *and do*

$f_{i+1} := f_i - \frac{lt(f_i)}{lt(g_j)} g_j$

$i := i + 1$ *and goto* $(*)$

else $r := r + lt(f_i)$

$f_{i+1} := f_i - lt(f_i)$

$i := i + 1$ *and goto* $(*)$ $\square$

Remarks 2.3.7. (i) Notice that in the above algorithm $lm(f_i) > lm(f_{i+1}) \forall i$, since the monomial order $\leq$ is a well order, the algorithm must terminate. Moreover, when it terminates at some $i = t$, f_t must be 0 and every monomial occurring in the final r is not divisible by any $lm(g), g \in G$. Therefore the final r is a reduced form of f w.r.t. G.

(ii) It's not hard to see the algorithm actually produces the following representation for f:

$$f = \sum_{u=1}^{s} c_u g'_u m_u + r, \tag{2.3.2}$$

where $s \in \mathbb{N}$(when $s = 0$, $f = r$), $c_u \in k - \{0\}$, $m_u \in M$, $g'_u \in G$ and not necessarily pairwise distinct $\forall\ u$, $1 \leqslant u \leqslant s$, r is the reduced form of f w.r.t. G, **and** $\mathbf{lm(f) = max\{lm(g'_1)m_1, lm(g'_2)m_2, \ldots, lm(g'_s)m_s, lm(r)\}}$. (Compare with 2.3.1)

In particular, when $r = 0$ in the above representation, (2.3.2) is said to be a *standard representation* of f w.r.t. G.

(iii) In the algorithm, we always choose the least j such that $lm(g_j) |\ lm(f_i)$. This implies that when we change the index order of elements in G, the final r produced by the algorithm may change too, hence the reduced form of a polynomial w.r.t. a general G may not be unique.

Example 2.3.8. Let $f = x_1^2 x_2^3$, $G = \{x_1^2, x_1x_2 - x_2^2\}$, $\leq$ be the *lex* with $x_1 > x_2$. Apply the algorithm 2.3.6, we have

$$f \xrightarrow{x_1^2} x_1^2 x_2^3 - x_1^2 x_2^3 = 0, \tag{2.3.3}$$

when $G = \{g_1 = x_1^2, g_2 = x_1x_2 - x_2^2\}$; or

$$f \xrightarrow{x_1x_2 - x_2^2} x_1^2x_2^3 - (x_1x_2 - x_2^2)x_1x_2^2 = x_1x_2^4 \xrightarrow{x_1x_2 - x_2^2} x_1x_2^4 - (x_1x_2 - x_2^2)x_2^3 = x_2^5, \tag{2.3.4}$$

when $G = \{g_1 = x_1x_2 - x_2^2, g_2 = x_1^2\}$. That is to say, 0 and x_2^5 are two different reduced forms of f w.r.t. G.

Moreover, from (2.3.3) we see $f \in< G >$, then from (2.3.4) we see $x_2^5 = f - (x_1x_2 - x_2^2)x_1x_2^2 - (x_1x_2 - x_2^2)x_2^3 \in< G >$, but clearly x_2^5 is reduced w.r.t. G. Then x_2^5 is in $lm(< G >)$ but not in $lm(G)$. This implies that G is not a Gröbner basis.

2.4 Characterizations of Gröbner Bases

We need a new definition before introducing the characterizations of Gröbner bases. Given $m_1, m_2 \in M$, it is known there exists the least common multiple of m_1 and m_2, denoted by $lcm(m_1, m_2)$, such that $m_1|\ lcm(m_1, m_2)$, $m_2|\ lcm(m_1, m_2)$ and $\forall m \in M$ with $m_1|\ m$ and $m_2|\ m$, $lcm(m_1, m_2)|\ m$.

Definition 2.4.1. Let $f, g \in k[x_1, x_2, \ldots, x_n] - \{0\}$, $L = lcm(lm(f), lm(g))$, then $S(f, g) = \frac{L}{lt(f)} f - \frac{L}{lt(g)} g$ is called the *S-polynomial* of f and g. Clearly, $S(g, f) = -S(f, g)$.

Theorem 2.4.2.(Characterizations of Gröbner Bases) Given a finite $G = \{g_1, g_2, \ldots, g_j, \ldots, g_l\} \subseteq k[x_1, x_2, \ldots, x_n]$ and $g_j \neq 0\ \forall j = 1, 2, \ldots, l$, let $I =< G >$ be the ideal generated by G, let $\leq$ be a monomial order on M. The following conditions are equivalent:

(a) $lm(G) = lm(I)$;

(b) $\forall f \in I - \{0\}$, $\exists j \in \{1, 2, \ldots, l\}$, such that $lm(g_j)|\ lm(f)$;

(c) $f \in I \Leftrightarrow R(f, G) = 0$;

(d) $f \in I \Leftrightarrow f$ has a standard representation w.r.t. G;

(e) $\forall f \in k[x_1, x_2, \ldots, x_n]$, the reduced form of f w.r.t. G is unique;

(f) As k-vector spaces, $k[x_1, x_2, \ldots, x_n] = k_R(G) \bigoplus I$;

(g) $\forall g'_1, g'_2 \in G$, $R(S(g'_1, g'_2), G) = 0$;

(h) $\forall g'_1, g'_2 \in G$, $S(g'_1, g'_2)$ has a standard representation w.r.t. G.

Proof: The proof can be found in [12] [8]. Or you may refer to chapter 3 for the proof of the characterizations of noncommutative Gröbner bases. The basic idea of that proof also works here. □

Remarks 2.4.3. (i) In these characterizations, **(a)** and **(b)** are essentially the same. When G is a Gröbner basis of the ideal $<G>$(or I), **(c)** and **(d)**

show the property of elements in the ideal $< G >$ and **(c)** is used to solve the word problem (see problem 2.5.4). By **(e)** and **(f)**, when G is a Gröbner basis, every polynomial f in $k[x_1, x_2, \ldots, x_n]$ has a unique representative $R(f, G)$ in the quotient ring $k[x_1, x_2, \ldots, x_n]/ <G>$, and as k-vector spaces, $k[x_1, x_2, \ldots, x_n]/ < G >$ is isomorphic to $k_R(G)$ which is spanned by the monomials not in $lm(G)$. Characterizations **(g)** and **(h)** are the foundations of Buchberger's Algorithm.

(ii) In the theorem, the condition "$g_j \neq 0 \ \forall j = 1, 2, \ldots, l$" has no effect on the characterizations but deletes trivial element in our Gröbner basis. In fact, we can add more conditions to G such that the Gröbner basis of an ideal I is *unique* w.r.t. the idea I and the monomial order $\leq$. This Gröbner basis is called a *reduced Gröbner basis.* In this thesis, we won't discuss this subject. See [12] [8] [7] for more information.

2.5 Buchberger's Algorithm and One Application

Given a finite $G = \{g_1, g_2, \ldots, g_j, \ldots, g_l\} \subseteq k[x_1, x_2, \ldots, x_n]$ and $g_j \neq 0 \ \forall j = 1, 2, \ldots, l$, let $I = <G>$ be the ideal generated by G, let $\leq$ be a monomial order on M. The example 2.3.8 shows G is not necessary to be a Gröbner basis of the ideal I. However, the theorem 2.2.8 and theorem 2.2.11 tell us there does exist some finite $G' \subseteq I$ such that $I = <G'>$ and $lm(G') = lm(I)$. In this section we will introduce Buchberger's algorithm which can decide whether the given $G(= G_1)$ is a Gröbner basis of $< G >$ and find out a Gröbner basis $G'(= G_i)$ of $<G>$ if G is not.

Algorithm 2.5.1.(Buchberger's Algorithm)

$i := 1, G_1 := \{g_1, g_2, \ldots, g_l\}, H := \{(g_{j_1}, g_{j_2}) | g_{j_1}, g_{j_2} \in G_1, 1 \leqslant j_1 < j_2 \leqslant l\}$

(*)while $H \neq \emptyset$ *do*

choose $(g'_{t1}, g'_{t2}) \in H$, *then let* $H := H - \{(g'_{t1}, g'_{t2})\}$

do algorithm 2.3.6 to compute a reduced form r *of* $S(g'_{t1}, g'_{t2})$ *w.r.t.* G_i

if $r \neq 0$ *then*

$H := H \bigcup \{(g, r) | \ g \in G_i\}$

$G_{i+1} := G_i \bigcup \{r\}$

$i := i + 1$

goto (*) $\quad \square$

Claim 2.5.2. The Buchberger's algorithm terminates at some i, $i \geqslant 1$, and the final G_i is a Gröbner basis of the ideal $< G_1 >$.

Proof: Suppose the algorithm doesn't terminate, then for each i, we must have some $r \neq 0$ and $lm(r) \in lm(G_{i+1})$ but not in $lm(G_i)$. This implies we would have an infinite properly ascending chain in $k[x_1, x_2, \ldots, x_n]$: $lm(G_1) \subsetneqq lm(G_2) \subsetneqq \ldots$. But $k[x_1, x_2, \ldots, x_n]$ is Noetherian, this is impossible. Hence the algorithm terminates at some i.

Clearly when the algorithm terminates, $H = \emptyset$. We have two cases:

Case 1: the initial $H = \emptyset$, the algorithm ends at $i = 1$ without doing anything. $H = \emptyset$ implies $l = 1$, *i.e.*, G_1 contains a single element g. Clearly $S(g, g) = 0$, thus by the characterization (g) or (h), $G_1 = \{g\}$ is a Gröbner basis of $< g >$.

Case 2: the initial $H \neq \emptyset$ and the algorithm ends at $i \geqslant 1$. Obviously, $G_1 \subseteq G_i \subseteq < G_1 >$, then $I =< G_1 >=< G_i >$. From the algorithm, we can see the reduced form of $S(g'_1, g'_2)$ that is produced by the algorithm 2.3.6 must be 0, $\forall g'_1, g'_2 \in G_i$. Then every $S(g'_1, g'_2)$ must have a standard representation w.r.t. G_i. By the characterization (h), G_i is a Gröbner basis of $< G_i >=< G_1 >$. $\square$

Remark 2.5.3. The proof actually gives us more information. (i) When the algorithm terminates at $i = 1$, G_1 is a Gröbner basis. Otherwise, the algorithm terminates at $i > 1$ and $lm(G_1) \subsetneq lm(G_2) \subsetneq \ldots \subsetneq lm(G_i) = lm(I)$. (ii) A generator of a principle ideal of $k[x_1, x_2, \ldots, x_n]$ is a Gröbner basis of that ideal.

At last, we introduce a basic application of commutative Gröbner bases theory.

Problem 2.5.4.(Word Problem for Commutative Algebra Presentations) For a k-algebra presentation $R = k[x_1, x_2, \ldots, x_n]/ < g_1, g_2, \ldots, g_l >$, is there an algorithm which, given $f \in k[x_1, x_2, \ldots, x_n]$, decides whether $f = \bar{0}$ in R ? (Clearly, $f = \bar{0}$ in $R \Leftrightarrow f \in < g_1, g_2, \ldots, g_l >$, so it is actually an ideal membership problem for $k[x_1, x_2, \ldots, x_n]$).

Solution: $f = 0$ is trivial. If $f \neq 0$, let $I =< g_1, g_2, \ldots, g_l >$ and define a monomial order $\leq$ on M. Then, apply the Buchberger's algorithm to compute out a Gröbner basis of I, denoted by G. Given G, apply the reduction algorithm 2.3.6 to compute a reduced form of f w.r.t. G. By the characterization (e), the reduced form is unique, then by the characterization (c), $R(f, G) = 0 \Leftrightarrow f \in I$. $\square$

Chapter 3

Generalization to Noncommutative Polynomial Rings

In this chapter, we generalize the commutative Gröbner bases theory to general noncommutative polynomial rings which are also associative free algebras. We will mainly discuss the generalization for two-sided ideals(called ideals simply). The generalization for one-sided ideals can be discussed in a similar way thus is omitted mostly.

3.1 Notations and Basic Definitions

Let $<X_n>$ denote the *free* monoid generated by set $X_n = \{x_1, x_2, \ldots, x_n\}$, let ${x_j}^0 = 1 \in <X_n> \ \forall\, 1 \leqslant j \leqslant n$, then

$$<X_n>= \{x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \ldots x_{i_j}^{\beta_j} \ldots x_{i_l}^{\beta_l} \mid \beta_j \in \mathbb{N}, x_{i_j} \in X_n, 1 \leqslant j \leqslant l, l \in \mathbb{N} - \{0\}\}.$$

When n is not important, we simply write $<X>$ instead of $< X_n >$. A typical element $x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \ldots x_{i_j}^{\beta_j} \ldots x_{i_l}^{\beta_l}$ in $<X>$ is called a *monomial*, $\sum_{j=1}^{l} \beta_j$ is called the *degree* of the monomial. For any monomial m, the degree of m

is denoted by $deg(m)$. Each $m \in <X> - \{1\}$ can be written as $u_1 u_2 \dots u_s$, where $u_i \in \{x_1, x_2, \dots, x_n\}, 1 \leqslant i \leqslant s, s \in \mathbb{N} - \{0\}$ and $deg(m) = s$. In addition, for $m_1, m_2 \in <X>$, if $\exists l, r \in <X>$ such that $m_2 = l m_1 r$, we say m_2 is a *multiple* of m_1, or m_1 *divides* m_2, denoted by $m_1 | \, m_2$.

Let k be a field, we use $k<X_n>$ or simply $k<X>$ to denote the associative *free* k-algebra generated by set X_n. When $n = 1$, the algebra is also a commutative polynomial ring in one variable. In the following discussion, we always assume $n \geqslant 2$, then $k<X>$ is known to be a *noncommutative* polynomial ring. For $k<X>$, we have the following facts:

(1) $\forall f \in k<X> - \{0\}, f = \sum_{i=1}^{s} c_i m_i$, where $s \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, m_i \in <X>, 1 \leqslant i \leqslant s$. We call $c_i m_i$ a *term*, c_i the *coefficient* of the term.

(2) $\forall f \in k<X> - \{0\}$, after all possible coalescence and cancellation of terms, f has the *unique form*,

$$f = \sum_{i=1}^{t} c_i m_i, \tag{3.1.1}$$

where $t \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, m_i \in <X>$ and $m_i \neq m_j \ \forall 1 \leqslant i \neq j \leqslant t$. The uniqueness is up to a permutation on the terms.

Definition 3.1.1. $\leq$ is said to be a (noncommutative) *monomial order* on $<X>$, if it satisfies the following conditions:

(i) $<X>$ is totally ordered by $\leq$;

(ii) $1 \leq m$, $\forall m \in <X>$;

(iii) $m_1 \leq m_2 \ \Rightarrow \ l m_1 r \leq l m_2 r$, $\forall \, l, r, m_1, m_2 \in <X>$.

Let $\leq$ be a monomial order on $<X>$, suppose $m_1, m_2 \in <X>$, and $m_2 = l m_1 r$ for some $l, r \in <X>$. By the above condition(ii) $1 \leq l$ and

$1 \leq r$, then by the condition(iii) $m_1 \leq lm_1 \leq lm_1 r = m_2$. Hence we can say $m_1 | m_2 \Rightarrow m_1 \leq m_2$, *i.e.*, the generalized monomial order is still an extension of the division relation.

Examples 3.1.2. We let $m_1 = u_1 u_2 \ldots u_s$ and $m_2 = v_1 v_2 \ldots v_t$ denote two monomials in $<X>$, where $u_i, v_j \in \{x_1, x_2, \ldots, x_n\}$, $1 \leqslant i \leqslant s, 1 \leqslant j \leqslant t$ and $s, t \in \mathbb{N}$. In particular, we let $s = 0 (t = 0)$ imply $m_1 = 1 (m_2 = 1)$.

(1)In the *lex* on $<X>$ with $x_1 > x_2 > \ldots > x_n$,

$$m_1 = u_1 u_2 \ldots u_s < m_2 = v_1 v_2 \ldots v_t$$

$$\Leftrightarrow \begin{cases} u_1 = v_1, u_2 = v_2, \ldots, u_l = v_l, u_{1+1} < v_{l+1},\ 0 \leqslant l < s; & \text{or,} \\ 0 < s < t \quad \textit{and} \quad u_1 = v_1, u_2 = v_2, \ldots, u_s = v_s; & \text{or,} \\ s = 0 < t. & \end{cases}$$

Now let $m_1 = x_2 x_2 x_1$, $m_2 = x_2 x_1$, then $m_1 = x_2 x_2 x_1 < x_2 x_1 = m_2$ since $x_2 < x_1$. But this contradicts with conditions (ii) and (iii) in definition 3.1.1 which require that $m_2 = 1 \cdot x_2 x_1 < x_2 \cdot x_2 x_1 = m_1$. Hence the lexicographical order is NOT a monomial order on $<X>$.

(2)In the *deglex* on $<X>$ with $x_1 > x_2 > \ldots > x_n$,

$$m_1 = u_1 u_2 \ldots u_s < m_2 = v_1 v_2 \ldots v_t$$

$$\Leftrightarrow \begin{cases} deg(m_1) = s < t = deg(m_2); & \text{or,} \\ deg(m_1) = s = t = deg(m_2) \quad \textit{and} \quad m_1 <_{lex} m_2; & \end{cases}$$

where the $<_{lex}$ is the lexicographical order on $<X>$ with $x_1 > x_2 > \ldots > x_n$.

(3)In the *degrevlex* on $<X>$ with $x_1 > x_2 > \ldots > x_n$,

$$m_1 = u_1 u_2 \ldots u_s < m_2 = v_1 v_2 \ldots v_t$$

$$\Leftrightarrow \begin{cases} deg(m_1) = s < t = deg(m_2); & \text{or,} \\ deg(m_1) = s = t = deg(m_2) \quad \textit{and} \quad u_s > v_s; & \text{or,} \\ deg(m_1) = s = t = deg(m_2) \quad \textit{and} & \\ u_s = v_s, \ldots, u_{l+1} = v_{l+1}, u_l > v_l, 1 \leqslant l < s. & \end{cases}$$

It's easy to verify the *deglex* and the *degrevlex* are still monomial orders on $<X>$.

Given any nonzero noncommutative polynomial f, f can be written in the unique form (3.1.1). Again we can arrange the terms by a monomial order, such that $f = \sum_{i=1}^{t} c_i m_i$, and $m_1 > m_2 > \ldots > m_t$. In this case, we call $c_1 m_1$ the *leading term* of f, denoted by $lt(f)$; m_1 the *leading monomial* of f, denoted by $lm(f)$; c_1 the *leading coefficient* of f, denoted by $lc(f)$.

Definition 3.1.3. Given any $G \subseteq k{<}X{>}$, we define the *leading monomial ideal* of G w.r.t. some monomial order $\leq$ to be

$$\begin{aligned} lm(G) : &= \ < lm(g) |\ g \in G > \\ &= \ \{\sum_{i=1}^{t} f_i lm(g_i) h_i |\ t \in \mathbb{N} - \{0\}, f_i, h_i \in k{<}X{>}, g_i \in G\} \end{aligned}$$

Definition 3.1.4. Given a monomial order on $<X>$, let G be a subset of $k{<}X{>}$, if $lm(G) = lm(< G >)$, we say G is a *Gröbner basis* (of the ideal $< G >$). If a set $G \subseteq$ ideal $I \subseteq k{<}X{>}$ and $lm(G) = lm(I)$, G is called a *Gröbner basis of I*.

Remark 3.1.5. Unlike commutative Gröbner bases, noncommutative Gröbner bases are allowed to be infinite. There are two reasons for this.

Firstly, finite Gröbner bases do not exist for some ideals in $k{<}X{>}$. In the next section we will show that some ideals of $k{<}X{>}$ cannot be finitely generated. Clearly those ideals have no finite Gröbner bases(see theorem 3.3.8). Moreover, in section 3.4 we will explain why there even exist finitely generated ideals which have no finite Gröbner bases.

Secondly, there do exist infinite sets of noncommutative polynomials which can play the same role as finite Gröbner bases in our theory(see claim 3.4.15 and theorem 3.7.6).

3.2 Infinitely Generated Ideals of $k{<}X{>}$

Recall that we based most fundamental results of commutative Gröbner bases theory on the fact that $k[x_1, x_2, \ldots, x_n]$ is Noetherian. For noncommutative polynomial ring $k{<}X{>}$, things are not as nicely behaved.

Definition 3.2.1. A noncommutative ring R is said to be *left(right) Noetherian* if R satisfies the ACC on left(right) ideals. R is said to be *Noetherian* if R is both left and right Noetherian.

Lemma 3.2.2. Noncommutative ring R satisfies the ACC on ideals(left ideals, right ideals) iff every ideal(left ideal, right ideal, respectively) of R is finitely generated.

Proof: For ideals, the proof of lemma 2.2.3 still works here. For left and right ideals, only appropriate modification on terminologies are needed. □

Next we give an example of infinite properly ascending chain of ideals in $k{<}X{>}$. Since ideals are both left and right ideals, the example shows that $k{<}X{>}$ satisfies ACC neither on left nor on right ideals.

Example 3.2.3. For convenience, we use x, y to denote two noncommutative variables in $k{<}X{>}$. $\forall i \in \mathbb{N}$, define I_i to be the (two-sided) ideal generated by the set $\{xy^jx | 0 \leqslant j \leqslant i\}$, *i.e.*, $I_i = < x^2, xyx, \ldots, xy^ix >$. Obviously, $xy^{i+1}x \in I_{i+1}$ but not in I_i, and $I_i \subseteq I_{i+1}$, for all $i \in \mathbb{N}$. Hence we have an infinite ascending chain in $k{<}X{>}$: $I_1 \subsetneq I_2 \subsetneq \cdots I_i \subsetneq I_{i+1} \cdots$.

By the definition 3.2.1 and the above example, we have the claim:

Claim 3.2.4. $k{<}X{>}$ is neither left nor right Noetherian, thus not Noetherian.

By the lemma 3.2.2 and the above example, we have the claim:

Claim 3.2.5. In $k{<}X{>}$ there exists an ideal(left ideal, right ideal) which cannot be finitely generated. (We call such an ideal an *infinitely generated ideal.*)

Actually, in the proof of lemma 2.2.3, we suggest a way to construct explicitly an infinitely generated ideal.

Example 3.2.6. Let $I = \bigcup_{i=0}^{\infty} I_i$, where I_i is the ideal defined in the example 3.2.3. It's easy to see I is also an ideal of $k{<}X{>}$. Suppose $\exists a_1, a_2, \ldots, a_s \in I$, such that $I =< a_1, a_2, \ldots, a_s >$, then for some sufficient large l, all $a_1, a_2, \ldots, a_s \in I_l$, then $I \subseteq I_l \subsetneq I_{l+1} \subseteq I$. This is impossible. Therefore, the ideal I of $k{<}X{>}$ cannot be finitely generated.

Clearly we can find many other infinitely generated ideals of $k{<}X{>}$ in the same way.

Notice that the above ideal I is actually generated by a set of monomials $S = \{xy^ix | i \in \mathbb{N}\}$, but I cannot be generated by any finite subset of S. So example 3.2.6 is also a counter example in $k{<}X{>}$ for Dickson's Lemma.

Claim 3.2.7. Dickson's Lemma doesn't hold in $k{<}X{>}$.

3.3 Discussion on Monomial Orders

In this section we will show the generalized monomial order on ${<}X{>}$ is still a well order, although $k{<}X{>}$ is not Noetherian and the Dickson's Lemma doesn't hold in $k{<}X{>}$.

First, let's see a result about monomial partial order on ${<}X{>}$, which is given by Prof. John Lawrence.

Definition 3.3.1. $\preceq$ is said to be a *monomial partial order* on ${<}X{>}$, if it satisfies the following conditions:

(i) $<X>$ is partially ordered by $\preceq$;

(ii) $1 \preceq m,\ \forall m \in <X>$;

(iii) $m_1 \preceq m_2 \Rightarrow lm_1r \preceq lm_2r,\ \forall\, l, r, m_1, m_2 \in <X>$.

Theorem 3.3.2. Let $\preceq$ be a monomial partial order on the free monoid $<X_n> = <x_1, x_2, \ldots, x_n>$. When $n = 2$, $(<X_n>, \preceq)$ satisfies DCC.

Before proving the above theorem, we introduce the following lemma.

Lemma 3.3.3. Let $\{a_n\}_{n=0}^{\infty}$ be a sequence in the well-ordered set $(A, \leq)$. Then there exists a subsequence $\{a_{n(j)}\}_{j=0}^{\infty}$ of $\{a_n\}_{n=0}^{\infty}$ such that $a_{n(j)} \leq a_{n(j+1)}\ \forall j \in \mathbb{N}$.

Proof: Since $\leq$ is a well order, we can find a least element $a_{n(0)}$ in the sequence $\{a_n\}_{n=0}^{\infty}$, then we define $a_{n(j)}$ recursively to be a least element in $\{a_n\}_{n=n(j-1)+1}^{\infty}$, for all $j > 0$. Clearly $\{a_{n(j)}\}_{j=0}^{\infty}$ is non-descending. $\square$

Proof of Theorem 3.3.2: (Due to John Lawrence) To simplify notations, we let $< X_2 >= \Delta =< u, v >$. Suppose we have an infinite properly descending chain of monomials in $(\Delta, \preceq)$: $w_0 \succ w_1 \succ \cdots$. Multiply each monomial of the chain by u on the left. By the property of the monomial partial order, the chain $uw_0 \succ uw_1 \succ \cdots$ is still infinite properly descending. Hence, without loss of generality(WLOG), assume w_i starts with the same variable u, for all $i \in \mathbb{N}$.

For any monomial $w \in \Delta$, define $\phi(w)$ to be the number of times "uv" occurs in w, *e.g.*, $\phi(uvuv) = 2$, $\phi(uuvv) = 1$ and $\phi(uuu) = 0$. Now we have two cases for $\{\phi(w_i)\}_{i=0}^{\infty}$.

Case 1: $\{\phi(w_i)\}_{i=0}^{\infty}$ is unbounded. Let $w_0 = y_1y_2 \cdots y_l$, $y_j \in \{u, v\}(y_0 = u)$, $1 \leqslant j \leqslant l$. Then $\exists t > 0$ such that $\phi(w_t) > l$. Clearly, w_t can be written as $w_t = u \cdots uvm_1uvm_2 \cdots uvm_l$, where $m_j \in \Delta$, $1 \leqslant j \leqslant l$. Then $y_j \prec uvm_j$

for all $1 \leqslant j \leqslant l$. Hence $w_0 \prec w_t$, a contradiction.

Case 2: $\{\phi(w_i)\}_{i=0}^{\infty}$ is bounded.

Claim: there exists a subchain of $w_0 \succ w_1 \succ \cdots$, denoted by $w'_0 \succ w'_1 \succ \cdots$, and some $s \geqslant 1$, such that, $\forall i \in \mathbb{N}$, $w'_i = u^{\alpha_{i1}}v^{\alpha_{i2}}u^{\alpha_{i3}}v^{\alpha_{i4}}\cdots u^{\alpha_{i(2s-1)}}v^{\alpha_{i(2s)}}$, $\alpha_{ij} \in \mathbb{N}$, $1 \leqslant j \leqslant 2s$, and $\forall j$, the sequence $\{\alpha_{ij}\}_{i=0}^{\infty}$ is non-descending.

Proof of the claim: Let's start with the original chain $w_0 \succ w_1 \succ \cdots$. Since $\{\phi(w_i)\}_{i=0}^{\infty}$ is bounded, clearly there is some bound $s \geqslant 1$, such that $\forall i \in \mathbb{N}$, $w_i = u^{\alpha_{i1}}v^{\alpha_{i2}}u^{\alpha_{i3}}v^{\alpha_{i4}}\cdots u^{\alpha_{i(2s-1)}}v^{\alpha_{i(2s)}}$, where $\alpha_{ij} \in \mathbb{N}$, $1 \leqslant j \leqslant 2s$. Suppose for some $j = J$, the sequence $\{\alpha_{iJ}\}_{i=0}^{\infty}$ is not non-descending. Since it is a sequence in $(\mathbb{N}, <_N)$ and the natural order $<_N$ is a well order on $\mathbb{N}$, by the lemma 3.3.3, there exists a non-descending subsequence $\{\alpha_{i(p)J}\}_{p=0}^{\infty}$ of $\{\alpha_{iJ}\}_{i=0}^{\infty}$. Hence, there exists a subchain of $w_0 \succ w_1 \succ \cdots$, denoted by $w_{i(0)} \succ w_{i(1)} \succ \cdots$, such that $\forall p \in \mathbb{N}$, $w_{i(p)} = u^{\alpha_{i(p)1}}v^{\alpha_{i(p)2}}u^{\alpha_{i(p)3}}v^{\alpha_{i(p)4}}\cdots u^{\alpha_{i(p)(2s-1)}}v^{\alpha_{i(p)(2s)}}$, $\alpha_{i(p)j} \in \mathbb{N}$, $1 \leqslant j \leqslant 2s$ and $\alpha_{i(p)J} \leqslant \alpha_{i(p+1)J}$. Clearly, after repeating the above and applying lemma 3.3.3 recursively at most $2s$ times, we can get the subchain required in the claim. Thus the claim is proved.

In the claim, $\alpha_{ij} \leqslant \alpha_{(i+1)j}$ for all i, j. By the property of the monomial partial order, this implies that, in the subchain, for all $i \in \mathbb{N}$,

$$\begin{array}{rcl} w'_i & = & u^{\alpha_{i1}}v^{\alpha_{i2}}u^{\alpha_{i3}}v^{\alpha_{i4}}\cdots u^{\alpha_{i(2s-1)}}v^{\alpha_{i(2s)}} \\ \preceq \quad w'_{i+1} & = & u^{\alpha_{(i+1)1}}v^{\alpha_{(i+1)2}}u^{\alpha_{(i+1)3}}v^{\alpha_{(i+1)4}}\cdots u^{\alpha_{(i+1)(2s-1)}}v^{\alpha_{(i+1)(2s)}}. \end{array}$$

It's impossible since $w'_0 \succ w'_1 \succ \cdots$ is properly descending.

To sum up, it's impossible to have an infinite properly descending chain of monomials in $(\Delta, \preceq)$, so $(<X_2>, \preceq) = (\Delta, \preceq)$ satisfies DCC. $\square$

Theorem 3.3.4. Every monomial order $\leq$ on $<X_n>$ is a well order. (As we have pointed out earlier, we always assume $n \geqslant 2$.)

Proof: We prove the statement by induction on n.

Assume $n = 2$. Let $\leq$ be a monomial order on $<X_2>$. Obviously, it is also a monomial partial order on $<X_2>$, then by theorem 3.3.2, $(<X_2>, \leq)$ satisfies DCC. Since the monomial order $\leq$ is a total order, it is a well order.

Suppose the statement is true for $n = l$ $(n \geqslant 2, n \in \mathbb{N})$, let's look at $n = l + 1$. Let $\leq$ be a monomial order on $<X_{l+1}> = <x_1, x_2, \ldots, x_{l+1}>$. Since $\leq$ is a total order, WLOG, we assume $x_{l+1} \geq x_i, \forall i = 1, 2, \ldots, l + 1$. Let $<X_l>$ denote the free monoid generated by $\{x_1, x_2, \ldots, x_l\}$. Obviously $<X_l> \subset <X_{l+1}>$ and $\leq$ is also a monomial order on $<X_l>$. By our induction hypothesis, $\leq$ is a well order on $<X_l>$.

For any $w \in <X_{l+1}>$, define $\phi(w)$ to be the number of "x_{l+1}" occurring in w. For example, $\phi(x_{l+1}^2) = 2$, $\phi(x_1 x_{l+1} x_2) = 1$ and $\phi(x_1 x_2) = 0$. Suppose we have an infinite properly descending chain of monomials in $(<X_{l+1}>, \leq)$: $w_0 > w_1 > \cdots$. Then for $\{\phi(w_i)\}_{i=0}^{\infty}$, we still have two cases.

Case 1: $\{\phi(w_i)\}_{i=0}^{\infty}$ is unbounded. Let $w_0 = u_1 u_2 \cdots u_s$, where $u_j \in \{x_1, x_2, \cdots, x_{l+1}\}$, $1 \leqslant j \leqslant s$. Then $\exists t > 0$ such that $\phi(w_t) > s$. Clearly, w_t can be written as $w_t = m_1 x_{l+1} m_2 x_{l+1} \cdots m_s x_{l+1} m_{s+1}$, where $m_j \in <X_l>$, $1 \leqslant j \leqslant s+1$. Then $u_j \leq m_j x_{l+1}$ for all $j = 1, 2, \ldots, s$ and $1 \leq m_{s+1}$. Hence $w_0 \leq w_t$, a contradiction.

Case 2: $\{\phi(w_i)\}_{i=0}^{\infty}$ is bounded. Then there is some bound $b \geqslant 1$, such that $\forall i \in \mathbb{N}$, $w_i = m_{i1} x_{l+1}^{\beta_{i1}} m_{i2} x_{l+1}^{\beta_{i2}} \cdots m_{ib} x_{l+1}^{\beta_{ib}} m_{i(b+1)}$, where $m_{ij} \in <X_l>$ for all $j = 1, 2, \ldots, b + 1$, and $\beta_{ij} \in \{0, 1\}$ for all $j = 1, 2, \ldots, b$. Notice that for all j, $\{m_{ij}\}_{i=0}^{\infty}$ is a sequence in the well-ordered set $(<X_l>, \leq)$, $\{x_{l+1}^{\beta_{ij}}\}_{i=0}^{\infty}$ is a sequence in the well-ordered set $(\{1, x_{l+1}\}, \leq)$, we can apply lemma 3.3.3 recursively with finite times, like we did in the proof of theorem

3.3.2, and find a subchain of $w_0 > w_1 > \cdots$, denoted by $w'_0 > w'_1 > \cdots$, such that $\forall i \in \mathbb{N}$, $w'_i = m'_{i1}x_{l+1}^{\beta'_{i1}}m'_{i2}x_{l+1}^{\beta'_{i2}}\cdots m'_{ib}x_{l+1}^{\beta'_{ib}}m'_{i(b+1)}$ and for all j, the sequence $\{m'_{ij}\}_{i=0}^{\infty}$ and $\{x_{l+1}^{\beta'_{ij}}\}_{i=0}^{\infty}$ are non-descending. But by the property of the monomial order, this implies that, for all $i \in \mathbb{N}$,

$$\begin{array}{rcl} & w'_i & = & m'_{i1}x_{l+1}^{\beta'_{i1}}m'_{i2}x_{l+1}^{\beta'_{i2}}\cdots m'_{ib}x_{l+1}^{\beta'_{ib}}m'_{i(b+1)} \\ \leq & w'_{i+1} & = & m'_{(i+1)1}x_{l+1}^{\beta'_{(i+1)1}}m'_{(i+1)2}x_{l+1}^{\beta'_{(i+1)2}}\cdots m'_{(i+1)b}x_{l+1}^{\beta'_{(i+1)b}}m'_{(i+1)(b+1)}. \end{array}$$

It's impossible since $w'_0 > w'_1 > \cdots$ is properly descending.

Hence when $n = l+1$, the statement still holds.

By induction on n, the statement holds for all $n \geqslant 2$. □

Corollary 3.3.5.(Theorem 2.2.9.) Every monomial order $\leq$ on M(the set of commutative monomials) is a well order.

Proof: It can be proved in the same way as above. (Here we do not need the Hilbert Basis Theorem or Dickson's Lemma.) □

It is known that every partial order can be refined to a total order. We may ask the following question: can every monomial partial order $\preceq$ be refined to a monomial (total) order $\leq$? If the answer is yes, then, by the theorem 3.3.4, $\leq$ satisfies DCC on $<X_n>$ for all $n \geqslant 2$, so would $\preceq$. Hence the statement in theorem 3.3.2 could be proved true for all $n \geqslant 2$ in this way. However, the following example gives a negative answer to our question.

Example 3.3.6.(Due to Bergman [3]) Consider $<X> = < u, v, x, y >$, let $\preceq$ be a monomial partial order which is generated by the basic relations $yu \prec xu$ and $xv \prec yv$. Then $\preceq$ can be refined to a total order but can NOT be refined to a monomial order. Because either $x < y$ or $y < x$ will bring about contradiction with $yu < xu$ or $xv < yv$ respectively.

Remark 3.3.7. Recently Prof. John Lawrence has proved a generalized

Dickson's Lemma for finitely generated free monoids. We point out that, as a corollary of that new result, the statement in theorem 3.3.2 does hold for all n. Then theorem 3.3.4 and corollary 3.3.5 (theorem 2.2.9) are immediate results from the complete version of theorem 3.3.2.

Next let's apply theorem 3.3.4 to prove the noncommutative version of theorem 2.2.11.

Theorem 3.3.8. Let $\leq$ be a monomial order on $<X>$ and let I be an ideal of $k<X>$. If a set G satisfies $G \subseteq I$ and $lm(G) = lm(I)$ (so G is a Gröbner basis of the ideal I), then $I =<G>$.

Proof: The proof is similar to the one for theorem 2.2.11.

Given any $f \in I - \{0\}$, do the following process: Let $f_1 = f,\ f_1 \in I,\ lm(f_1) \in lm(I) = lm(G) \Rightarrow \exists g_1 \in G \subseteq I,\ l_1, r_1 \in <X>,\ c_1 \in k - \{0\}$, such that $lt(f_1) = lt(c_1 l_1 g_1 r_1)$. Let $f_2 = f_1 - c_1 l_1 g_1 r_1$, clearly $f_2 \in I$, when $f_2 \neq 0,\ lm(f_2) \in lm(I) = lm(G) \Rightarrow \exists g_2 \in G \subseteq I,\ l_2, r_2 \in <X>,\ c_2 \in k - \{0\}$, such that $lt(f_2) = lt(c_2 l_2 g_2 r_2)$. Let $f_3 = f_2 - c_2 l_2 g_2 r_2$, clearly $f_3 \in I$, and when $f_3 \neq 0$, we can move to $f_4 \ldots$

Notice that $lm(f_1) > lm(f_2) > lm(f_3) > \ldots$, since we have proved the monomial order $\leq$ is a well order, the process must terminate at some $f_l = 0$. Hence we have

$$\begin{aligned} 0 &= f_l = f_{l-1} - c_{l-1} l_{l-1} g_{l-1} r_{l-1} = f_{l-2} - c_{l-2} l_{l-2} g_{l-2} r_{l-2} - c_{l-1} l_{l-1} g_{l-1} r_{l-1} \\ &= \ldots = f_1 - \sum_{i=1}^{l-1} c_i l_i g_i r_i. \end{aligned}$$

So $f = f_1 = \sum_{i=1}^{l-1} c_i l_i g_i r_i \in<G>$. This implies $I \subseteq<G>$. It's obvious that $<G>\subseteq I$. Therefore $I =<G>$. $\square$

We now show that a one-sided noncommutative monomial order on $<X>$ may not be a well order.

Definition 3.3.9. $\leq$ is said to be a *right monomial order* on $<X>$, if it satisfies the following conditions:

(i) $<X>$ is totally ordered by $\leq$;

(ii) $1 \leq m,\ \forall m \in <X>$;

(iii) $m_1 \leq m_2 \;\Rightarrow\; m_1 r \leq m_2 r,\ \forall\ m_1, m_2, r \in <X>$.(But $m_1 \leq m_2$ doesn't imply $lm_1 \leq lm_2$, for some $m_1, m_2, l \in <X>$.)

Example 3.3.10.(Due to John Lawrence) Let's consider the free monoid $\Delta =< x, y >$. For any $m \in \Delta$, define $deg_x(m) =$ the number of "x" occurring in m, define $\phi(m) =$ the number of "y" which is to the right of an "x" in m. For example, let $m = yxyxyy$, then $deg_x(m) = 2$ and $\phi(m) = 3$.

Now for any two monomials $m_1, m_2 \in \Delta$, we define a total order $\leq$ on Δ by $m_1 < m_2$

$$\Leftrightarrow \begin{cases} deg_x(m_1) < deg_x(m_2); & \text{or,} \\ deg_x(m_1) = deg_x(m_2) \quad and \quad \phi(m_1) > \phi(m_2); & \text{or,} \\ deg_x(m_1) = deg_x(m_2), \phi(m_1) = \phi(m_2) \quad and \quad deg(m_1) < deg(m_2); & \text{or,} \\ deg_x(m_1) = deg_x(m_2), \phi(m_1) = \phi(m_2), deg(m_1) = deg(m_2) \quad and & \\ m_1 <_{lex} m_2, & \end{cases}$$

where $<_{lex}$ is the lexicographical order on Δ with $y < x$.

It's easy to verify that the above order $\leq$ satisfies the conditions of the definition 3.3.9, *i.e.*, $\leq$ is a right monomial order on Δ. But we have an infinite properly descending chain in $(\Delta, \leq)$: $x > xy > xy^2 > \cdots$. (Notice that $deg_x(xy^i) = deg_x(xy^{i+1}) = 1$ but $\phi(xy^i) = i < \phi(xy^{i+1}) = i+1$, therefore $xy^i > xy^{i+1}$, $\forall i \in \mathbb{N}$).

Hence the right monomial order $\leq$ is not a well order.

3.4 Generalization of Polynomial Reduction

Let $\leq$ be a monomial order on $<X>$, we will discuss noncommutative polynomial reductions in this section.

Definition 3.4.1. For any $f, g \in k<X>$, if $lm(g)$ divides some nonzero term cm in f, then $\exists l, r \in <X>$ such that $cm = c \cdot l \cdot lm(g) \cdot r$. Let $h = f - \frac{c}{lc(g)} lgr$, then it's easy to see the term cm in f is replaced by a linear combination of monomials $< m$. We call this manipulation a *polynomial reduction*, denoted by $f \xrightarrow{g} h$, and say f *reduces to* h *modulo* g.

Definition 3.4.2. If there is a finite sequence of polynomial reductions $f \xrightarrow{g_1} h_1 \xrightarrow{g_2} h_2 \ldots \xrightarrow{g_t} h_t$, where $g_i \in G \subseteq k<X>$ and g_i not necessarily pairwise distinct for $1 \leqslant i \leqslant t$, we say f *reduces to* h_t *modulo* G, denoted by $f \xrightarrow{G} h_t$.

Definition 3.4.3. Polynomial d is called *reduced* or *in the reduced form* w.r.t. some $G \subseteq k<X>$, if $d = 0$ or no monomial occurring in the unique form (3.1.1) of d is divisible by $lm(g)$ $\forall g \in G$. If $f \xrightarrow{G} d$ and d is reduced w.r.t. G, we say d is a *reduced form of* f w.r.t. G. When the reduced form of f w.r.t. G is unique(in general it is not, see examples 3.4.16), we denote it by $R(f, G)$. In particular, when f is reduced w.r.t. G, $R(f, G) = f$.

Remarks 3.4.4. (i) We do NOT require f in the definitions to be in the unique form except in the case we need consider its leading term or leading monomial, like in the following reduction process.

(ii) The equivalence "monomial $m \in lm(G) \Leftrightarrow \exists g \in G,\ lm(g) | \ m$" is still true. Hence, when we say $d \neq 0$ is reduced w.r.t G, it is equivalent to say no monomial in the unique form of d is in $lm(G)$.

(iii) Given $G \subseteq k<X>$, let $M(G)$ denote the set of all monomials in $lm(G)$, *i.e.*,$M(G) = <X> \bigcap lm(G)$ and let $k_R(G)$ denote the set of all reduced polynomials w.r.t. G. Then it's easy to verify that, as a k-vector space, $k_R(G) = span_k\{<X> - M(G)\}$, *i.e.*,$k_R(G)$ is spanned by monomials not in $lm(G)$.

By the above definitions, the following proposition is obvious. It will be useful in proofs.

Proposition 3.4.5. Given polynomials f and d, if d is a reduced form of f w.r.t. some G, then either $f = d$ or $\exists s \in \mathbb{N} - \{0\}, c_u \in k - \{0\}, l_u, r_u \in <X>, g_u \in G$ and not necessarily pairwise distinct $\forall\, u,\ 1 \leqslant u \leqslant s$, such that

$$f = \sum_{u=1}^{s} c_u l_u g_u r_u + d. \tag{3.4.1}$$

Next we introduce a reduction process which shows that, for any nonzero polynomial f and a set $G \subseteq k<X>$, the reduced form of f w.r.t. G always exists.

Reduction Process 3.4.6.

$i := 1, d := 0, f_i := f$

$(*)$while $f_i \neq 0$ *do*

if $\exists g_i \in G, l_i, r_i \in <X>$ *such that* $lm(f_i) = l_i \cdot lm(g_i) \cdot r_i$, *do*

$f_{i+1} := f_i - \frac{lc(f_i)}{lc(g_i)} l_i g_i r_i$

$i := i + 1$ *and goto* $(*)$

else $d := d + lt(f_i)$

$f_{i+1} := f_i - lt(f_i)$

$i := i + 1$ *and goto* $(*)$ $\square$

Remarks 3.4.7. (i) In the above process $lm(f_i) > lm(f_{i+1}) \forall i$, since the monomial order $\leq$ is a well order, the process must terminate at some $i = t$ and $f_t = 0$. It's easy to see every monomial occurring in the final d is not divisible by any $lm(g), g \in G$. Therefore the final d is a reduced form of f w.r.t. G.

(ii) The process actually shows that f has the following representation:

$$f = \sum_{u=1}^{s} c_u l_u g_u r_u + d, \tag{3.4.2}$$

where $s \in \mathbb{N}$(when $s = 0,\ f = d$), $c_u \in k - \{0\},\ l_u, r_u \in <X>,\ g_u \in G$ and not necessarily pairwise distinct $\forall\ u,\ 1 \leqslant u \leqslant s$, d is the reduced form of f w.r.t. G, **and** $\mathbf{lm(f) = max\{l_1 lm(g_1) r_1, l_2 lm(g_2) r_2, \ldots, l_s lm(g_s) r_s, lm(d)\}}$. (Compare with 3.4.1)

In particular, when $d = 0$ in the above representation, (3.4.2) is said to be a *standard representation* of f w.r.t. G.

(iii) Obviously, the process is a generalization of reduction algorithm 2.3.6. But we don't call the above process an "algorithm". When we apply the process to reduce f modulo some infinite G, we may not be able to decide whether "*$\exists g_i \in G, l_i, r_i \in <X>$ such that $lm(f_i) = l_i \cdot lm(g_i) \cdot r_i$*" or not, although one of two cases must be true theoretically. In other words, the process only shows the reduced form of f exists in theory, but in practice we may not be able to compute it for some infinite G.

Next let's explain why there does exist such an infinite G in $k<X>$.

Problem 3.4.8.(Word Problem for Noncommutative Algebra Presentations) For a k-algebra presentation $R = k<X>/ < g_1, g_2, \ldots, g_l >$, is there an algorithm which, given $f \in k<X>$, decides whether $f = \bar{0}$ in R ? (Clearly, $f = \bar{0}$ in $R \Leftrightarrow f \in < g_1, g_2, \ldots, g_l >$, so it is also an ideal membership problem for $k<X>$).

Claim 3.4.9. The word problem for noncommutative algebra presentations, or the ideal membership problem for $k<X>$, is unsolvable in general.

Proof: See [1].

Theorem 3.4.10. Let $G \subseteq$ ideal $I \subseteq k{<}X{>}$, $lm(G) = lm(I)$,*i.e.*,G is a Gröbner basis of the ideal I, then $f \in I \Leftrightarrow R(f, G) = 0$.

Proof: See the characterizations of noncommutative Gröbner bases.

By the above results, there exists a finitely generated ideal $I \subseteq k{<}X{>}$ for which the ideal membership problem is unsolvable. I can be regarded as an infinite set and by our definition I is a Gröbner basis of itself. Now suppose the reduction process 3.4.6 could compute a reduced form of any given f w.r.t. I, then by theorem 3.4.10, we could decide whether $f \in I$ or not. Hence the ideal membership problem would be solved. This is a contradiction. Therefore we have the following claim:

Claim 3.4.11. (i) There exists an infinite $G \subseteq k{<}X{>}$ such that the reduction process 3.4.6 cannot be implemented. (ii) There exists a finitely generated ideal I of $k{<}X{>}$ for which we cannot find a finite Gröbner basis.

Moreover, with the above results, the following claim is also obvious.

Definition 3.4.12. For a set $G \subseteq k{<}X{>}$, if given any $f \in k{<}X{>}$, one can compute a reduced form of f w.r.t. G, then we say G is *computable*.

Claim 3.4.13. For any ideal I of $k{<}X{>}$, if we can find a computable Gröbner basis of I, then we can solve the ideal membership problem for that ideal. On the other hand, there does exist an ideal I in $k{<}X{>}$ whose ideal membership problem is unsolvable. For such an ideal I, there is no algorithm which can find a computable Gröbner basis of I.

Now let's focus on the computable sets in $k{<}X{>}$. Clearly, finite sets are always computable. Given a finite set $G = \{g_1, g_2, \ldots, g_j, \ldots, g_l\}$, the reduction process 3.4.6 can be refined to the following algorithm.

Algorithm 3.4.14.(Reduction Algorithm)

$i := 1, d := 0, f_i := f$

(*)while $f_i \neq 0$ *do*

if $\exists lm(g_j) | \ lm(f_i)$, *choose the least* j *from* 1 *to* l, $l_i, r_i \in <X>$ *such that* $lm(f_i) = l_i lm(g_j) r_i$, *and do*

$f_{i+1} := f_i - \frac{lc(f_i)}{lc(g_j)} l_i g_j r_i$

$i := i + 1$ *and goto (*)*

else $d := d + lt(f_i)$

$f_{i+1} := f_i - lt(f_i)$

$i := i + 1$ *and goto (*)* □

For infinite sets, we have the following claim.

Claim 3.4.15. Given an infinite $G \subseteq k<X>$, if for any $D \in \mathbb{N}$, the subset $G(D) = \{g \in G | \ deg(lm(g)) \leqslant D\}$ is finite and every element of $G(D)$ can be calculated explicitly, then G is computable.

Proof: In the reduction process 3.4.6, to decide whether "$\exists g_i \in G, l_i, r_i \in <X>$ such that $lm(f_i) = l_i lm(g_i) r_i$" or not, we only need compare $lm(f_i)$ to every $lm(g)$ with $deg(g) \leqslant deg(f_i)$ and $g \in G$, *i.e.*, we only need know every element in $G(D)$, where $D = deg(f_i)$. Clearly G in our claim is satisfactory. Then given any f in $k<X>$, we can apply the reduction process 3.4.6 to compute a reduced form of f w.r.t. G, *i.e.*, G is computable. □

At last, we give some examples which show that the reduced form of f is not unique w.r.t. general computable G.

Examples 3.4.16. (1) Let $f = x_1^2 x_2 x_1 - x_1 x_2^2 x_1$, $G = \{g_1, g_2\} \subseteq k<X>$, where $g_1 = x_1^2 - x_1 x_2$,$g_2 = x_1 x_2 x_1 - x_2 x_1 x_2$. Let $\leq$ be the *deglex* with $x_1 > x_2$, then $lm(f) = x_1^2 x_2 x_1$, $lm(g_1) = x_1^2$, $lm(g_2) = x_1 x_2 x_1$. Applying the algorithm 3.4.14 to reduce f modulo $\{g_1, g_2\}$ and $\{g_2, g_1\}$, we have

$$f \xrightarrow{g_1} f - g_1 x_2 x_1 = 0 \tag{3.4.3}$$

and

$$f \xrightarrow{g_2} f - x_1g_2 = x_1x_2x_1x_2 - x_1x_2^2x_1 \xrightarrow{g_2} (f - x_1g_2) - g_2x_2 = -x_1x_2^2x_1 + x_2x_1x_2^2. \tag{3.4.4}$$

That is to say, we find two different reduced forms of f w.r.t. G, 0 and $-x_1x_2^2x_1 + x_2x_1x_2^2$.

Moreover, from (3.4.3) we see $f \in < G >$, then from (3.4.4) we see $d = -x_1x_2^2x_1 + x_2x_1x_2^2 = f - x_1g_2 - g_2x_2 \in < G >$, then $lm(d) \in lm(< G >)$ but clearly not in $lm(G)$. Hence $lm(G) \subsetneq lm(< G >)$, so G is not a Gröbner basis.

(2) Let $f = x_1^3$, $G = \{g = x_1^2 - x_2\}$. If they are considered in commutative Gröbner bases theory, G will be a Gröbner basis since it contains a single polynomial(see remark 2.5.3(ii)). Then by a characterization of commutative Gröbner bases, the reduced form of f w.r.t. G will be unique.

But now, let's consider them in noncommutative polynomial rings. Let $\leq$ be the *deglex* with $x_1 > x_2$, then $lm(f) = x_1^3$, $lm(g) = x_1^2$. We apply the algorithm 3.4.14 to reduce f modulo G. Clearly we have only one choice of g_j such that $lm(g_j)|\, lm(f_i)$, but we do have different choices of l_i, r_i such that $lm(f_i) = l_i lm(g_j) r_i$. Thus we have the following results.

$$f \xrightarrow{g} f - x_1g = x_1^3 - x_1(x_1^2 - x_2) = x_1x_2, \tag{3.4.5}$$

and

$$f \xrightarrow{g} f - gx_1 = x_1^3 - (x_1^2 - x_2)x_1 = x_2x_1. \tag{3.4.6}$$

That is to say, we find two different reduced forms of f w.r.t. G, x_1x_2 and x_2x_1.

Moreover, from (3.4.5) and (3.4.6), we see that

$$x_1x_2 - x_2x_1 = gx_1 - x_1g \in < G > \tag{3.4.7}$$

Therefore, $lm(x_1x_2 - x_2x_1) = x_1x_2 \in lm(<G>)$ but x_1x_2 is not in $lm(G)$, hence $lm(G) \subsetneq lm(<G>)$, so G is not a Gröbner basis.

3.5 Noncommutative S-polynomials

This section is devoted to the generalization of S-polynomials(see definition 2.4.1). Two problems arise. Firstly, given $m_1, m_2 \in <X>$, it's often impossible to find the least common multiple of m_1, m_2 like we did for commutative monomials. For example, both xyx and yxy are common multiples of xy and yx, but we cannot find $lcm(xy, yx)$, such that $lcm(xy, yx)|\ xyx$ and $lcm(xy, yx)|\ yxy$. Secondly, for noncommutative m_1, m_2, even if we have found $L = lcm(m_1, m_2)$, the expression $\frac{L}{m_1}$ (or $\frac{L}{m_2}$) is ambiguous since we may have different choices of l, r such that $L = lm_1r$ (or $L = lm_2r$). In the example 3.4.16(2), we have seen the ambiguity of $\frac{x_1^3}{x_1^2}$ caused by two possibilities $x_1^3 = x_1 \cdot x_1^2$ or $x_1^2 \cdot x_1$.

For the first problem, we will investigate all cases of common multiples of m_1, m_2. Instead to look for a *least* common multiple, we try to find out the *minimal* common multiple in each case. To avoid the ambiguity in the second problem, for an ordered pair of monomials $(m_1, m_2) \in <X>^2$, we let $T(m_1, m_2)$ denote the set of 4-tuples $(l_1, r_1, l_2, r_2) \in <X>^4$ satisfying $l_1m_1r_1 = l_2m_2r_2$.

The Cases of Common Multiples of m_1 and m_2

Given an ordered pair of monomials $(m_1, m_2) \in <X>^2$, if $(l_1, r_1, l_2, r_2) \in T(m_1, m_2)$, then according to the relative locations of m_1 and m_2 in the common multiple W, we have the following three cases.

Case 1: $\exists w \in <X>$ between m_1 and m_2.

Case 1-1: $W = l_1 m_1 w m_2 r_2$, the minimal common multiple is $m_1 w m_2$.

Case 1-2: $W = l_2 m_2 w m_1 r_1$, the minimal common multiple is $m_2 w m_1$.

Since w can be any monomial, we have an infinite set of minimal common multiples in Case 1. Fortunately, in later discussion, we will see that we don't need the generalization of S-polynomials for this case.

Case 2: There is no w between m_1 and m_2. The monomials m_1, m_2 overlap but no one contains the other.

Case 2-1: $\exists (1, R_1, L_2, 1) \in T(m_1, m_2)$, $R_1 \neq 1$, $L_2 \neq 1$ such that $m_1 R_1 = L_2 m_2 = m_0$. Then $W = l_1 m_0 r_2$, the minimal common multiple is m_0.

Case 2-2: $\exists (L_1, 1, 1, R_2) \in T(m_1, m_2)$, $L_1 \neq 1$, $R_2 \neq 1$ such that $L_1 m_1 = m_2 R_2 = m_0$. Then $W = l_2 m_0 r_1$, the minimal common multiple is m_0.

Since m_1, m_2 overlap, $deg(m_0) < deg(m_1) + deg(m_2)$, hence we have finite minimal common multiples of m_1 and m_2 in case 2.

Case 3: There is no w between m_1 and m_2, and one contains the other.

Case 3-1: m_1 contains m_2. Then the minimal common multiple is $m_0 = m_1$, $W = l_1 m_0 r_1$ and $\exists (1, 1, L_2, R_2) \in T(m_1, m_2)$ such that $m_0 = m_1 = L_2 m_2 R_2$.

Case 3-2: m_2 contains m_1. Then the minimal common multiple is $m_0 = m_2$, $W = l_2 m_0 r_2$ and $\exists (L_1, R_1, 1, 1) \in T(m_1, m_2)$ such that $m_0 = m_2 = L_1 m_1 R_1$.

Clearly, we have finite minimal common multiples in case 3.

With the above discussion, we can define noncommutative S-polynomials.

Definition 3.5.1. Given an ordered pair of monomials $(m_1, m_2) \in <X>^2$, *the set of matches of* (m_1, m_2), denoted by $MS(m_1, m_2)$, is the finite set of all ordered 4-tuples $(L_1, R_1, L_2, R_2) \in T(m_1, m_2)$, such that, either

(i) $(L_1, R_1, L_2, R_2) = (1, R_1, L_2, 1)$ with $R_1 \neq 1$, $L_2 \neq 1$ and $\exists w \neq 1$ such that $wR_1 = m_2$ and $L_2 w = m_1$;

or (ii) $(L_1, R_1, L_2, R_2) = (L_1, 1, 1, R_2)$ with $L_1 \neq 1$, $R_2 \neq 1$ and $\exists w \neq 1$ such that $wR_2 = m_1$ and $L_1 w = m_2$;

or (iii) $(L_1, R_1, L_2, R_2) = (1, 1, L_2, R_2)$ with $m_1 = L_2 m_2 R_2$;

or (iv) $(L_1, R_1, L_2, R_2) = (L_1, R_1, 1, 1)$ with $m_2 = L_1 m_1 R_1$.

Definition 3.5.2. Given $f, g \in k{<}X{>}-\{0\}$, if $MS(lm(f), lm(g)) \neq \emptyset$, then

$$S(f,g)[L_1, R_1, L_2, R_2] := \frac{1}{lc(f)} L_1 f R_1 - \frac{1}{lc(g)} L_2 g R_2$$

is called an *S-polynomial* of f and g w.r.t. (L_1, R_1, L_2, R_2), where $(L_1, R_1, L_2, R_2) \in MS(lm(f), lm(g))$.

Remarks 3.5.3. (i)The four cases in the definition of $MS(m_1, m_2)$ clearly corresponds to minimal common multiples in cases 2-1, 2-2, 3-1 and 3-2. By the discussion there, $MS(m_1, m_2)$ is finite. Note that $MS(x, y)$ may be empty.

(ii)By symmetry,

$$(L_1, R_1, L_2, R_2) \in MS(lm(f), lm(g)) \Leftrightarrow (L_2, R_2, L_1, R_1) \in MS(lm(g), lm(f)),$$

thus $S(f,g)[L_1, R_1, L_2, R_2] = -S(g,f)[L_2, R_2, L_1, R_1]$.

We conclude the discussion in this section with the following observation.

Theorem 3.5.4. For any polynomial

$$\frac{1}{lc(f)} l_1 f r_1 - \frac{1}{lc(g)} l_2 g r_2,$$

where $(l_1, r_1, l_2, r_2) \in T(lm(f), lm(g))$, $f, g \in k{<}X{>}-\{0\}$ and $l_1 lm(f) r_1 = l_2 lm(g) r_2 = W$, we have the following three cases.

Case 1: $\exists w \in \langle X \rangle$ such that $W = l_1 \cdot lm(f) \cdot w \cdot lm(g) \cdot r_2$.

Case 2: $\exists w \in \langle X \rangle$ such that $W = l_2 \cdot lm(g) \cdot w \cdot lm(f) \cdot r_1$.

Case 3: $\exists l, r \in \langle X \rangle$, $\exists (L_1, R_1, L_2, R_2) \in MS(lm(f), lm(g))$, such that $W = l(L_1 lm(f) R_1) r = l(L_2 lm(g) R_2) r$, and

$$\frac{1}{lc(f)} l_1 f r_1 - \frac{1}{lc(g)} l_2 g r_2 = l \cdot S(f,g)[L_1, R_1, L_2, R_2] \cdot r.$$

Proof: By our previous discussion, the result is obvious. $\square$

3.6 Characterizations of Noncommutative Gröbner Bases

In this section we will prove several characterizations of noncommutative Gröbner bases.

Theorem 3.6.1.(Characterizations of Noncommutative Gröbner Bases)Given $G \subseteq k\langle X \rangle$, assume 0 is not in G, let $I = \langle G \rangle$ be the ideal generated by G, let $\leq$ be a monomial order on $\langle X \rangle$. The following conditions are equivalent:

(a) $lm(G) = lm(I)$;

(b) $\forall f \in I - \{0\}$, $\exists g \in G$ such that $lm(g) | \, lm(f)$;

(c) $f \in I \Leftrightarrow R(f, G) = 0$;

(d) $f \in I \Leftrightarrow f$ has a standard representation w.r.t. G;

(e) $\forall f \in k\langle X \rangle$, the reduced form of f w.r.t. G is unique;

(f) As k-vector spaces, $k\langle X \rangle = k_R(G) \bigoplus I$;

(g) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$, $R(S(g_1, g_2)[L_1, R_1, L_2, R_2], G) = 0$;

(h) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$, $S(g_1, g_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G.

Proof: We show the cycle **(a)⇒(b)⇒(c)⇒(f)⇒(e)⇒(g)⇒(h)⇒(d)⇒(a)** as follows.

(a)⇒(b): If $f \in I - \{0\}$, then $lm(f) \in lm(I) = lm(G)$.

(b)⇒(c): By prop.3.4.5, $R(f,G) = 0 \Rightarrow f = 0$ or $f = \sum_{u=1}^{s} c_u l_u g_u r_u$, where $s \in \mathbb{N} - \{0\}, c_u \in k - \{0\}, l_u, r_u \in <X>, g_u \in G$. Clearly, $f \in I$.

Conversely, when $f = 0$, $R(f,G) = f = 0$. For $f \in I - \{0\}$, if $d \neq 0$ is a reduced form of f w.r.t. G, again by prop.3.4.5, $d = f \in I - \{0\}$ or $d = f - \sum_{u=1}^{s} c_u l_u g_u r_u \in I - \{0\}$, then by **(b)**, $\exists g \in G$ such that $lm(g) |\, lm(d)$. But d is reduced w.r.t. G, a contradiction. So the reduced form of f w.r.t. G must be 0 and thus must be unique, *i.e.*, $R(f,G) = 0$.

(c)⇒(f): $\forall f \in k<X>$, the process 3.4.6 shows that $f = d + \sum_{u=1}^{s} c_u l_u g_u r_u$(see (3.4.2) in remark 3.4.7(ii)), where d is a reduced form of f w.r.t. G. Hence, $k<X> = k_R(G) + I$. We only need show $k_R(G) \bigcap I = \{0\}$. $\forall d \in k_R(G) \bigcap I$, $d \in k_R(G) \Rightarrow d = R(d,G)$; also, $d \in I \Rightarrow R(d,G) = 0$ by **(c)**. Therefore $d = 0$, *i.e.*, $k_R(G) \bigcap I = \{0\}$.

(f)⇒(e): $\forall f \in k<X>$, if f is reduced w.r.t. G, then $R(f,G) = f$. When f is not reduced, let d_1 and d_2 be two reduced forms of f w.r.t. G. By prop.3.4.5, $f = \sum_{u=1}^{s} c_u l_u g_u r_u + d_1 = \sum_{v=1}^{t} c'_v l'_v g'_v r'_v + d_2$ (see 3.4.1). Then $d_1 - d_2 = \sum_{v=1}^{t} c'_v l'_v g'_v r'_v - \sum_{u=1}^{s} c_u l_u g_u r_u \in k_R(G) \bigcap I$. By **(f)**, $d_1 - d_2 = 0$. Hence, the reduced form of f w.r.t. G is unique.

(e)⇒(g): $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$, by **(e)**, the reduced form of $S(g_1,g_2)[L_1,R_1,L_2,R_2]$ w.r.t. G is unique. We only need show it is zero.

Let $W = L_1 lm(g_1) R_1 = L_2 lm(g_2) R_2$, $h_1 = W - \frac{1}{lc(g_1)} L_1 g_1 R_1$, $h_2 = W - \frac{1}{lc(g_2)} L_2 g_2 R_2$. By reduction process 3.4.6, there exist two finite sequence of reductions as follows:(Although we possibly cannot compute them out, they do exist!)

$$W \xrightarrow{g_1} h_1 = W - \frac{1}{lc(g_1)} L_1 g_1 R_1 \xrightarrow{g_{11}} \dots \xrightarrow{g_{1a}} d_1 = R(W, G)$$

and

$$W \xrightarrow{g_2} h_2 = W - \frac{1}{lc(g_2)} L_2 g_2 R_2 \xrightarrow{g_{21}} \dots \xrightarrow{g_{2b}} d_2 = R(W, G).$$

Then for $S(g_1, g_2)[L_1, R_1, L_2, R_2]$, there is a finite sequence of reductions:

$$S(g_1, g_2)[L_1, R_1, L_2, R_2] = h_2 - h_1 \xrightarrow{g_{11}} \dots \xrightarrow{g_{1a}} h_2 - d_1 \xrightarrow{g_{21}} \dots \xrightarrow{g_{2b}} d_2 - d_1. \tag{3.6.1}$$

Since $R(W, G)$ is unique, $d_2 - d_1 = 0$, hence,

$$R(S(g_1, g_2)[L_1, R_1, L_2, R_2], G) = 0.$$

(g)⇒(h): By **(g)**, $R(S(g_1, g_2)[L_1, R_1, L_2, R_2], G) = 0$, then by remark 3.4.7(ii), each corresponding $S(g_1, g_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G.

(h)⇒(d): Obviously, "f has a standard representation w.r.t. G"⇒ "$f \in I$"(see 3.4.2 in remark 3.4.7(ii)). We only show "$f \in I$"⇒ "f has a standard representation w.r.t. G".

Now 0 has a trivial standard representation w.r.t. G. $\forall f \in I - \{0\}$, I is generated by G, thus f can be written as $f = \sum_{i=1}^{t} c_i l_i g_i r_i$, where $t \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, l_i, r_i \in <X>, g_i \in G$. Then

$$\begin{aligned} \Gamma &= \Big\{ \max_{1 \leqslant i \leqslant t} \{l_i \cdot lm(g_i) \cdot r_i\} \mid f = \sum_{i=1}^{t} c_i l_i g_i r_i, t \in \mathbb{N} - \{0\}, \\ &\qquad c_i \in k - \{0\}, l_i, r_i \in <X>, g_i \in G \Big\} \\ &\neq \emptyset. \end{aligned}$$

Since the monomial order $\leq$ is a well order, Γ has a least element m. This implies, there is a representation of f such that,

$$f = \sum_{i=1}^{t} c_i l_i g_i r_i, \tag{3.6.2}$$

$t \in \mathbb{N} - \{0\}$, $c_i \in k - \{0\}$, $l_i, r_i \in <X>$, $g_i \in G$, and

$$\max_{1 \leqslant i \leqslant t} \{l_i \cdot lm(g_i) \cdot r_i\} = m.$$

Claim: the above (3.6.2) is a standard representation of f w.r.t. G.

Proof of the claim: Obviously, $lm(f) \leq m$. By the definition of standard representation, we need show $lm(f) = m$. WLOG, assume

$$l_i \cdot lm(g_i) \cdot r_i \geq l_{i+1} \cdot lm(g_{i+1}) \cdot r_{i+1}, \forall i = 1, 2, \ldots, t-1, \text{ in (3.6.2)}.$$

Let $J = \max\{i |\ l_i \cdot lm(g_i) \cdot r_i = m, 1 \leqslant i \leqslant t\}$, then (3.6.2) looks like

$$f = c_1 l_1 g_1 r_1 + \cdots + c_J l_J g_J r_J + \sum_{i=J+1}^{t} c_i l_i g_i r_i, \tag{3.6.3}$$

where

$$m = l_i \cdot lm(g_i) \cdot r_i, \forall i = 1, 2, \ldots, J,$$

and

$$m > l_i \cdot lm(g_i) \cdot r_i \geq l_{i+1} \cdot lm(g_{i+1}) \cdot r_{i+1}, \forall i = J+1, \ldots, t-1.$$

Let $a_i = lc(g_i)$, $\forall i = 1, 2, \ldots, J$. If $J = 1$ or $c_1 a_1 + \cdots + c_J a_J \neq 0$, since m cannot be canceled on the right side of (3.6.3)(*i.e.*,(3.6.2)), $lm(f)$ has to be m. Then (3.6.3)(*i.e.*,(3.6.2)) is a standard representation of f.

Assume now that $J \geqslant 2$ and $c_1a_1 + \cdots + c_Ja_J = 0$. Let's show this case is impossible. Notice that (3.6.3) can be rewritten as follows.

$$\begin{aligned} f &= c_1a_1(\frac{1}{a_1}l_1g_1r_1 - \frac{1}{a_2}l_2g_2r_2) + (c_1a_1 + c_2a_2)(\frac{1}{a_2}l_2g_2r_2 - \frac{1}{a_3}l_3g_3r_3) \\ &\quad + \cdots + (c_1a_1 + \cdots + c_{J-1}a_{J-1})(\frac{1}{a_{J-1}}l_{J-1}g_{J-1}r_{J-1} - \frac{1}{a_J}l_Jg_Jr_J) \\ &\quad +(c_1a_1 + \cdots + c_Ja_J)\frac{1}{a_J}l_Jg_Jr_J + \sum_{i=J+1}^{t} c_il_ig_ir_i. \end{aligned}$$

Since $c_1a_1 + \cdots + c_Ja_J = 0$, we have

$$\begin{aligned} f &= c_1a_1(\frac{1}{a_1}l_1g_1r_1 - \frac{1}{a_2}l_2g_2r_2) + (c_1a_1 + c_2a_2)(\frac{1}{a_2}l_2g_2r_2 - \frac{1}{a_3}l_3g_3r_3) \\ &\quad + \cdots + (c_1a_1 + \cdots + c_{J-1}a_{J-1})(\frac{1}{a_{J-1}}l_{J-1}g_{J-1}r_{J-1} - \frac{1}{a_J}l_Jg_Jr_J) \\ &\quad + \sum_{i=J+1}^{t} c_il_ig_ir_i. \end{aligned} \tag{3.6.4}$$

Now consider

$$\frac{1}{a_i}l_ig_ir_i - \frac{1}{a_{i+1}}l_{i+1}g_{i+1}r_{i+1}, \forall i = 1, 2, \ldots, J-1.$$

Since

$$m = l_i \cdot lm(g_i) \cdot r_i = l_{i+1} \cdot lm(g_{i+1}) \cdot r_{i+1}, \forall i = 1, 2, \ldots, J-1,$$

by theorem 3.5.4, we have three cases.

Case 1: $\exists w \in <X>$ such that $m = l_i \cdot lm(g_i) \cdot w \cdot lm(g_{i+1}) \cdot r_{i+1}$. Then

$$r_i = w \cdot lm(g_{i+1}) \cdot r_{i+1}, \tag{3.6.5}$$

$$l_{i+1} = l_i \cdot lm(g_i) \cdot w. \tag{3.6.6}$$

Notice that the unique forms(see (3.1.1)) of g_i and g_{i+1} look like:

$$g_i = a_i lm(g_i) + \sum_{p=2}^{t_1} c_{ip} m_{ip} \ , \tag{3.6.7}$$

$$g_{i+1} = a_{i+1} lm(g_{i+1}) + \sum_{q=2}^{t_2} c_{(i+1)q} m_{(i+1)q} \ , \tag{3.6.8}$$

where $t_1, t_2 \in \mathbb{N} - \{0\}$, $c_{ip}, c_{(i+1)q} \in k - \{0\}$, $m_{ip}, m_{(i+1)q} \in <X>$ and

$$lm(g_i) > m_{ip} > m_{p+1}, \ \forall p = 2, \ldots, t_1 - 1,$$

$$lm(g_{i+1}) > m_{(i+1)q} > m_{(i+1)(q+1)}, \ \forall q = 2, \ldots, t_2 - 1.$$

With (3.6.5)—(3.6.8), we have

$$\begin{aligned}
& \frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1} \\
= \ & \frac{1}{a_i a_{i+1}} \Bigg\{ l_i g_i w a_{i+1} lm(g_{i+1}) r_{i+1} - l_i a_i lm(g_i) w g_{i+1} r_{i+1} \\
& + l_i g_i w g_{i+1} r_{i+1} - l_i g_i w g_{i+1} r_{i+1} \Bigg\} \\
= \ & \frac{1}{a_i a_{i+1}} \Bigg\{ l_i [g_i - a_i lm(g_i)] w g_{i+1} r_{i+1} - l_i g_i w [g_{i+1} - a_{i+1} lm(g_{i+1})] r_{i+1} \Bigg\} \\
= \ & \frac{1}{a_i a_{i+1}} \Bigg\{ l_i [\sum_{p=2}^{t_1} c_{ip} m_{ip}] w g_{i+1} r_{i+1} - l_i g_i w [\sum_{q=2}^{t_2} c_{(i+1)q} m_{(i+1)q}] r_{i+1} \Bigg\}
\end{aligned}$$

where

$$l_i \cdot m_{ip} \cdot w \cdot lm(g_{i+1}) \cdot r_{i+1} < l_i \cdot lm(g_i) \cdot w \cdot lm(g_{i+1}) \cdot r_{i+1} = m,$$

$$l_i \cdot lm(g_i) \cdot w \cdot m_{(i+1)q} \cdot r_{i+1} < l_i \cdot lm(g_i) \cdot w \cdot lm(g_{i+1}) \cdot r_{i+1} = m,$$

for all $p = 2, \ldots, t_1$, and $q = 2, \ldots, t_2$. In other words, we have rewritten

$$\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1} = \sum_{j=1}^{s_i} c_j l_j g_j r_j \tag{3.6.9}$$

where $s_i \in \mathbb{N} - \{0\}$, $c_j \in k - \{0\}$, $l_j, r_j \in <X>$, $g_j \in G$, and

$$\max_{1 \leqslant j \leqslant s_i} \{l_j \cdot lm(g_j) \cdot r_j\} < m.$$

Case 2: $\exists w \in <X>$ such that $m = l_{i+1} \cdot lm(g_{i+1}) \cdot w \cdot lm(g_i) \cdot r_i$. Clearly this case is symmetric to the case 1. With the same method, we also can rewrite

$$\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1} = \sum_{j=1}^{s_i} c_j l_j g_j r_j \tag{3.6.10}$$

where $s_i \in \mathbb{N} - \{0\}$, $c_j \in k - \{0\}$, $l_j, r_j \in <X>$, $g_j \in G$, and

$$\max_{1 \leqslant j \leqslant s_i} \{l_j \cdot lm(g_j) \cdot r_j\} < m.$$

Case 3: $\exists l, r \in <X>$ and $\exists (L_1, R_1, L_2, R_2) \in MS(lm(g_i), lm(g_{i+1}))$, such that

$$\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1} = l \cdot S(g_i, g_{i+1})[L_1, R_1, L_2, R_2] \cdot r.$$

The condition **(h)** says that $S(g_i, g_{i+1})[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G, as does $l \cdot S(g_i, g_{i+1})[L_1, R_1, L_2, R_2] \cdot r$. Hence we can rewrite

$$\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1} = \sum_{j=1}^{s_i} c_j l_j g_j r_j \tag{3.6.11}$$

where $s_i \in \mathbb{N} - \{0\}$, $c_j \in k - \{0\}$, $l_j, r_j \in <X>$, $g_j \in G$, and

$$\max_{1 \leqslant j \leqslant s_i} \{l_j \cdot lm(g_j) \cdot r_j\} = lm(\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1}) < m.$$

With above results, let's revisit (3.6.4). For all $i = 1, 2, \ldots, J-1$, we can rewrite

$$\frac{1}{a_i} l_i g_i r_i - \frac{1}{a_{i+1}} l_{i+1} g_{i+1} r_{i+1}$$

to the form (3.6.9) or (3.6.10) or (3.6.11). For all $i = J+1, \ldots, t$, it is known

$$l_i \cdot lm(g_i) \cdot r_i < m.$$

Hence (3.6.4) can be rewritten as

$$f = \sum_{j=1}^{t'} c'_j l'_j g'_j r'_j \tag{3.6.12}$$

where $t' \in \mathbb{N} - \{0\}$, $c'_j \in k - \{0\}$, $l'_j, r'_j \in <X>$, $g'_j \in G$, and

$$\max_{1 \leqslant j \leqslant t'} \{l'_j \cdot lm(g'_j) \cdot r'_j\} = m' < m.$$

Obviously, $m' \in \Gamma$, but m is the least element of Γ, it's impossible that $m' < m$.

Therefore the case "$J \geqslant 2$ and $c_1 a_1 + \cdots + c_J a_J = 0$" is impossible for (3.6.3)(*i.e.*,(3.6.2)). Then (3.6.3)(*i.e.*,(3.6.2)) is a standard representation of f w.r.t. G. The claim is proved, thus "**(h)⇒(d)**" is proved.

(d)⇒(a): $\forall f \in I$, by **(d)**, f has a standard representation w.r.t. G, then $lm(f) = l_i \cdot lm(g_i) \cdot r_i$ for some $g_i \in G$, $l_i, r_i \in <X>$. Then $lm(f) \in lm(G)$, which implies $lm(I) \subseteq lm(G)$. It's obvious that $lm(G) \subseteq lm(I)$. Thus $lm(I) = lm(G)$. $\square$

Remark 3.6.2. As we can see, the above theorem is almost the same as its counterpart in chapter 2. In chapter 4, we will explain why.

3.7 Generalization of Buchberger's Algorithm

As we have pointed out, in noncommutative polynomial ring $k<X>$, there are ideals which cannot be finitely generated. For such an ideal I, we do not know if there is an algorithm which can find an infinite computable Gröbner basis G of I when G does exist. For a finitely generated ideal of $k{<}X{>}$, we have the following semi-decision algorithm, which is a generalization of Buchberger's algorithm and is one of the best results known so far.

Algorithm 3.7.1.(Generalized Buchberger's Algorithm Due to Mora)

$i := 1$, $H_1 := F$, $G_1 := F$, (F is a given finite subset of $k{<}X{>}$)

($*$)while $H_i \neq \emptyset$ do

$H_{i+1} := \emptyset$

$B_i := \{(f, g, L_1, R_1, L_2, R_2) | f \in G_i, g \in H_i, (L_1, R_1, L_2, R_2) \in MS(lm(f), lm(g))\}$

($\star$)while $B_i \neq \emptyset$ do

choose $(f_1, f_2, L_1, R_1, L_2, R_2) \in B_i$

$B_i := B_i - \{(f_1, f_2, L_1, R_1, L_2, R_2)\}$

$f := \frac{1}{lc(f_1)} L_1 f_1 R_1 - \frac{1}{lc(f_2)} L_2 f_2 R_2$

do reduction algorithm 3.4.14 to compute a reduced form d of f w.r.t. $G_i \bigcup H_{i+1}$

if $d \neq 0$ then $H_{i+1} := H_{i+1} \bigcup \{d\}$

goto ($\star$)

$G_{i+1} = G_i \bigcup H_{i+1}$

$i := i + 1$ and goto ($*$) $\square$

In the above algorithm, we assume F is nonempty and a monomial order $\leq$ has been defined on ${<}X{>}$. About the above algorithm, we now make the following claims.

Claim 3.7.2. $\forall i \in \mathbb{N} - \{0\}$, $\forall (f_1, f_2) \in G_i^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(f_1), lm(f_2))$, $S(f_1, f_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G_{i+1}. (If the algorithm terminates at i, let $G_j = G_i, \forall j > i$.)

Proof: We prove the statement by induction on i.

Consider $i = 1$. Notice that $G_1 = H_1$, $\forall (f_1, f_2) \in G_1 \times G_1 = G_1 \times H_1$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(f_1), lm(f_2))$, the algorithm reduces $f = S(f_1, f_2)[L_1, R_1, L_2, R_2]$ to a reduced form d w.r.t. $G_1 \bigcup H_2$. If $d = 0$, clearly, f has a standard representation w.r.t. $G_2 \supseteq G_1 \bigcup H_2$. If $d \neq 0$, then $d \in G_2$, again, f has a standard representation w.r.t. G_2.

Suppose the statement holds for i, let's prove that it holds for $i + 1$.

If the algorithm terminates at $i + 1$, $H_{i+1} = \emptyset$, and $G_{i+2} = G_{i+1} = G_i$. Then by our induction hypothesis, $\forall (f_1, f_2) \in G_{i+1}^2 = G_i^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(f_1), lm(f_2))$, $S(f_1, f_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t.$G_{i+2} = G_{i+1}$. The statement is true.

If the algorithm doesn't terminate at $i + 1$, $G_{i+1} = G_i \bigcup H_{i+1}$, we then have two cases.

Case 1: $(f_1, f_2) \in G_{i+1} \times H_{i+1}$. Like the case $i = 1$, the algorithm will reduce $f = S(f_1, f_2)[L_1, R_1, L_2, R_2]$ to d. Whether $d = 0$ or not, the algorithm ensures f has a standard representation w.r.t.G_{i+2}.

Case 2: $(f_1, f_2) \in G_{i+1} \times G_i$. We have two sub-cases:

Case 2-1: $(f_1, f_2) \in G_i \times G_i$. By hypothesis inductions, every S-polynomial of (f_1, f_2) has a standard representation w.r.t.$G_{i+2} \supseteq G_{i+1}$.

Case 2-2: $(f_1, f_2) \in H_{i+1} \times G_i$. Then (f_2, f_1) is in the case 1, every S-polynomial of (f_2, f_1) has a standard representation w.r.t.G_{i+2}. By remark 3.5.3(ii), so does every S-polynomial of (f_1, f_2).

Hence, the statement holds for $i + 1$. By induction, the statement holds for all i. $\square$

Claim 3.7.3. (i)The algorithm terminates at some $i + 1$, $i \in \mathbb{N} - \{0\}$, if and only if, G_i is a finite Gröbner basis of the ideal $I = <F>$.

(ii)If the algorithm never terminates, then $\bigcup_{i=1}^{\infty} G_i$ is an infinite Gröbner basis of the ideal $I =< F >$.

Proof: Notice that $F = G_1 \subseteq \ldots \subseteq G_i \subseteq \ldots \subseteq I =< F >$. Then

$$< F > \subseteq \ldots \subseteq < G_i > \subseteq \ldots \subseteq < F > .$$

Hence, $< G_i >= I =< F >$, $\forall i \in \mathbb{N} - \{0\}$.

(i) If the algorithm terminates at some $i+1$, then $H_{i+1} = \emptyset$, and $G_{i+1} = G_i \bigcup H_{i+1} = G_i$. By claim 3.7.2, $\forall (f_1, f_2) \in G_i^2$, every S-polynomial of (f_1, f_2) has a standard representation w.r.t.$G_{i+1} = G_i$. Since $< G_i >= I$, by the characterization (h), G_i is a Gröbner basis of I. Obviously, G_i is finite.

Conversely, if G_i is a finite Gröbner basis of the ideal I, then H_{i+1} will be empty since all $d = R(f, G_i \bigcup H_{i+1}) = 0$. Thus the algorithm terminates at $i+1$.

(ii) If the algorithm never terminates, let $G = \bigcup_{i=1}^{\infty} G_i$, then $\forall (f_1, f_2) \in G^2$, there is sufficient large J such that $(f_1, f_2) \in G_J^2$, then every S-polynomial of (f_1, f_2) has a standard representation w.r.t.$G_{J+1} \subseteq G$. Obviously $< G >= I$, then by the characterization (h), G is a Gröbner basis of the ideal I. Since the algorithm never terminates, $G_i \subsetneq G_{i+1}$, G must be infinite. $\square$

Claim 3.7.4. If $I =< F >$ has a finite Gröbner basis w.r.t. $\leq$, then the algorithm must terminate.

Proof: Let $G' = \{g_1, g_2, \ldots, g_l\}$ be a finite Gröbner basis of I w.r.t. $\leq$. Suppose the algorithm never terminates. By claim 3.7.3(ii), $G = \bigcup_{i=1}^{\infty} G_i$ is an infinite Gröbner basis of I. Then $\forall j = 1, 2, \ldots, l$, $\exists J(j) \in \mathbb{N} - \{0\}$ such that $lm(g_j) \in lm(G_{J(j)})$. Let $J = \max_{1 \leqslant j \leqslant l}\{J(j)\}$, then $lm(I) = lm(G') \subseteq lm(G_J) \subseteq lm(I)$, thus $lm(G_J) = lm(I)$, *i.e.*, G_J is a finite Gröbner basis of I. By claim 3.7.3(i), the algorithm terminates, a contradiction! Hence the algorithm does terminate. $\square$

Given a finitely generated ideal, if algorithm 3.7.1 terminates and produces a finite Gröbner basis of the ideal, then we can solve the ideal membership problem for the ideal by characterizations of noncommutative Gröbner bases, like we did for problem 2.5.4.

In the last part of this chapter, we show that when the ideal is finitely generated by homogenous polynomials, the ideal membership problem is still solvable even if the algorithm 3.7.1 never terminates.

Definition 3.7.5. A polynomial $f \in k{<}X{>}$ is said to be *homogeneous* if in the unique form (3.1.1) of f,

$$f = \sum_{i=1}^{t} c_i m_i,\ deg(m_i) = deg(m_j)\ \forall 1 \leqslant i \neq j \leqslant t,$$

$i.e.,f$ is a linear combination of monomials of the same degree. If an ideal is generated by homogeneous polynomials, it is said a *homogeneous ideal*.

Theorem 3.7.6. Given a homogeneous ideal $I = {<} f_1, f_2, \ldots, f_l {>} \subseteq k{<}X{>}$, given a monomial order $\leq$ on ${<}X{>}$, algorithm 3.7.1 always produces a computable Gröbner basis of the ideal I. (Thus the ideal membership problem for the ideal is solvable.)

Proof: If the algorithm terminates, then by claim 3.7.3(i), it produces a finite Gröbner basis G of the ideal I. Clearly G is computable.

Next we assume the algorithm never terminates. By claim 3.7.3(ii), $G = \bigcup_{i=1}^{\infty} G_i$ is an infinite Gröbner basis of the ideal I. We will show that G satisfies the conditions in claim 3.4.15, *i.e.,*"for any $D \in \mathbb{N}$, the subset $G(D) = \{g \in G |\ deg(lm(g)) \leqslant D\}$ is finite and every element of $G(D)$ can be calculated explicitly", thus G will be computable by the claim.

For convenience, if g is a homogeneous polynomial, we let $deg(g)$ denote the degree of the leading monomial of g, *i.e.,* $deg(g) = deg(lm(g))$. By computing, it's easy to see the following claim is true.

Claim 1: If f_1, f_2 are homogeneous, then $f = S(f_1, f_2)[L_1, R_1, L_2, R_2]$ is homogeneous $\forall (L_1, R_1, L_2, R_2) \in MS(lm(f_1), lm(f_2))$, and

$$deg(f) = deg(L_1 f_1 R_1) = deg(L_2 f_2 R_2) \geqslant \max\{deg(f_1), deg(f_2)\}.$$

Claim 2: If G' is a finite set of homogeneous polynomials and f is homogeneous, then d, which is the reduced form of f w.r.t. G' produced by reduction algorithm 3.4.14, is also homogeneous and $deg(d) = deg(f)$ if $d \neq 0$.

Proof of the claim 2: Notice that in reduction algorithm 3.4.14, there are totaly three types of computation, either "$f_{i+1} := f_i - \frac{lc(f_i)}{lc(g_j)} l_i g_j r_i$", or "$d := d + lt(f_i)$", or "$f_{i+1} := f_i - lt(f_i)$" preserves $deg(f)$ and homogeneity. Hence the claim is true.

Now let's look at G_i, H_i in algorithm 3.7.1.

Claim 3: $\forall i \in \mathbb{N} - \{0\}$, $G_{i+1} = G_i \dot{\bigcup} H_{i+1}$ (*i.e.*, G_{i+1} is a *disjoint* union of G_i and H_{i+1}). Moreover, $\forall d_1 \in H_{i+1}$ and $\forall d_2 \in H_i$, $lm(d_1) \neq lm(d_2)$.

Proof of the claim 3: By the algorithm, G_{i+1} is a union of G_i and H_{i+1}. $\forall d_1 \in H_{i+1}$, since d_1 is a reduced form of f w.r.t. $G_i \bigcup H_{i+1}$ and $d_1 \neq 0$, clearly, d_1 is not in G_i. So the union is disjoint. Also, notice that $H_i \subseteq G_i$ and obviously there is no $g \in G_i$ such that $lm(g) | lm(d_1)$, so there is no $d_2 \in H_i$ with $lm(d_1) = lm(d_2)$. Thus the claim 3 is proved.

Applying claims 1 and 2 recursively, we can see all elements in G_i and H_i are homogeneous. Then all elements in the infinite Gröbner basis G are homogeneous. Moreover, by claim 3, we can see

$$G = H_1(G_1) \dot{\bigcup} H_2 \dot{\bigcup} \cdots \dot{\bigcup} H_i \dot{\bigcup} H_{i+1} \cdots,$$

i.e., G is a pairwise disjoint union of H_i, $i \in \mathbb{N} - \{0\}$. Define $D_i := \min\{deg(d) |\ d \in H_i\}$.

Claim 4: $D_i \leqslant D_{i+1}$, $\forall i \in \mathbb{N} - \{0\}$.

Proof of the claim 4: $\forall d \in H_{i+1}$, since d is the reduced form of f w.r.t. $G_i \bigcup H_{i+1}$ and $d \neq 0$, where $f = S(f_1, f_2)[L_1, R_1, L_2, R_2]$ with $f_1 \in G_i$, $f_2 \in H_i$, by claims 1 and 2,

$$deg(d) = deg(f) \geqslant \max\{deg(f_1), deg(f_2)\} \geqslant deg(f_2) \geqslant D_i.$$

Hence, $D_i \leqslant D_{i+1}$.

Now let's prove G is computable. Since the algorithm 3.7.1 never terminates, $D_i \geqslant D_1 \geqslant 1$, $\forall i \in \mathbb{N} - \{0\}$. Thus $G(0)$ is empty. Given $D \in \mathbb{N} - \{0\}$, suppose $\forall i \in \mathbb{N} - \{0\}$, $D_i \leqslant D$, then there is at least one d_i in H_i such that $deg(d_i) \leqslant D$. By claim 3, $lm(d_i) \neq lm(d_j)$ for all $i \neq j$, then we would have infinite different monomials with degree $\leqslant D$. This is impossible. Hence, there exists some $J \in \mathbb{N} - \{0\}$ such that $D_J > D$. By claim 4, $D_i > D$ for all $i \geqslant J$. Hence, $G(D) = \{g \in G |\ deg(lm(g)) \leqslant D\} \subseteq \dot{\bigcup}_{i=1}^{J-1} H_i = G_{J-1}$. Clearly, $G(D)$ is finite and every element of $G(D)$ can be calculated explicitly by algorithm 3.7.1. $\square$

Remark 3.7.7. In practice, since we need check each D_i, we need modify algorithm 3.7.1 to make it pause after computing out each G_i. That is easy to realize and not the topic of this thesis.

Chapter 4

Diamond Lemma(s)

4.1 Newman's Diamond Lemma

Newman's diamond lemma was firstly introduced by M.H.A.Newman in [5]. Readers are referred to [6] for an introduction to the lemma in the terminology of graph theory. In this section, we will introduce the lemma in the terminology of reduction theory. Our introduction is based on [8]. We also point out that [8] actually has shown the relations between Newman's diamond lemma and commutative Gröbner bases theory.

Definition 4.1.1. We define a general *reduction* on a nonempty set S to be a strictly antisymmetric relation on S, *i.e.*, a *reduction* on S is a subset R of $S \times S$ such that $(a, b) \in R \Rightarrow (b, a)$ not in R, $\forall (a, b) \in R$.

Notations 4.1.2. For a reduction relation R on nonempty set S, we will write:

(i)$a \to b \Leftrightarrow (a, b) \in R$.

(ii)

$$a \xrightarrow{n} b \; (n \in \mathbb{N}) \Leftrightarrow \begin{cases} a = b & \text{when} \quad n = 0, \\ \text{or } \exists a_0, a_1, \ldots, a_n \in S, \text{ such that} & \\ a = a_0 \to a_1 \to \cdots \to a_n = b & \text{when} \quad n \geqslant 1. \end{cases}$$

(iii)$a \xrightarrow{*} b \Leftrightarrow \exists n \in \mathbb{N}, a \xrightarrow{n} b$.

(iv)$a \leftrightarrow b \Leftrightarrow a \to b \quad \text{or} \quad b \to a$.

(v)

$$a \overset{n}{\leftrightarrow} b \ (n \in \mathbb{N}) \Leftrightarrow \begin{cases} a = b & \text{when} \quad n = 0, \\ \text{or } \exists a_0, a_1, \ldots, a_n \in S, \text{ such that} & \\ a = a_0 \leftrightarrow a_1 \leftrightarrow \cdots \leftrightarrow a_n = b & \text{when} \quad n \geqslant 1. \end{cases}$$

(vi)$a \overset{*}{\leftrightarrow} b \Leftrightarrow \exists n \in \mathbb{N}, a \overset{n}{\leftrightarrow} b$.

(vii)$b \leftarrow a \to c \Leftrightarrow a \to c$ and $a \to b$. (This notation rule also applies to $\xrightarrow{n}$ and $\xrightarrow{*}$.)

(viii)$a \downarrow b \Leftrightarrow \exists c \in S$ such that $a \xrightarrow{*} c \xleftarrow{*} b$.

The following claim is obvious.

Claim 4.1.3. "$\overset{*}{\leftrightarrow}$" is an equivalence relation on S.

Definition 4.1.4. Let $\to$ be a reduction defined on a nonempty set S.

(i) If there is no infinite reduction chain w.r.t. $\to$ in S, *i.e.*, every reduction chain in S is finite, then we say $\to$ satisfies DCC.

(ii) Let S' be a nonempty subset of S, an element $a \in S'$ is called a *minimal element of* S' w.r.t. $\to$ if there is no $b \in S'$ such that $a \to b$. In particular, if a is a minimal element of S, we say a is a *normal form* or in *normal form* w.r.t. $\to$. If $a \xrightarrow{*} b$ and b is in normal form, we say b is *a normal form of* a.

Lemma 4.1.5. Let $\to$ be a reduction defined on nonempty set S. If $\to$ satisfies DCC, then every element $a \in S$ has at least one normal form in S.

Proof: For any $a \in S$, define $S'(a) = \{b \in S | a \xrightarrow{*} b\}$, then $a \in S'(a) \neq \emptyset$. Since $\to$ satisfies DCC, $S'(a)$ has a minimal element b_0 w.r.t $\to$. (Otherwise, we would have an infinite reduction chain in $S'(a)$). Suppose b_0 is not minimal in S, then we would have $b_0 \to b_1$. But then $a \xrightarrow{*} b_1$ implies $b_1 \in S'(a)$. Since

b_0 is minimal in $S'(a)$, this is impossible. Hence b_0 is minimal in S,*i.e.*,b_0 is a normal form of a. $\square$

The following theorem is introduced in [8] as "Newman's lemma", which is essentially a variation of Newman's diamond lemma.

Theorem 4.1.6. Let $\rightarrow$ be a reduction defined on nonempty set S. If $\rightarrow$ satisfies DCC, then the following conditions are equivalent:

(i)*Local confluence(diamond condition)*: $b \leftarrow a \rightarrow c \Rightarrow b \downarrow c, \forall a, b, c \in S$.

(ii)*Confluence*: $b \overset{*}{\leftarrow} a \overset{*}{\rightarrow} c \Rightarrow b \downarrow c, \forall a, b, c \in S$.

(iii)Every element in S has a *unique normal form*.

(iv)*Church-Rosser property*: $a \overset{*}{\leftrightarrow} b \Rightarrow a \downarrow b, \forall a, b \in S$.

Proof: see [8].

Theorem 4.1.7.(Newman's Diamond Lemma)Let $\rightarrow$ be a reduction defined on nonempty set S. If $\rightarrow$ satisfies two conditions (i)DCC and (ii)diamond condition, then every equivalence class of $\overset{*}{\leftrightarrow}$ contains a unique normal form.

Proof: Let $a \overset{*}{\leftrightarrow} b$. By lemma 4.1.5, a has a normal form a_0, b has a normal form b_0. By (iii) in theorem 4.1.6, a_0, b_0 are unique respectively of a and b. By transitivity of equivalence relation, $a_0 \overset{*}{\leftrightarrow} b_0$. Then by (iv) in theorem 4.1.6, $a_0 \downarrow b_0$, *i.e.*,$\exists c \in S$ such that $a_0 \overset{*}{\rightarrow} c \overset{*}{\leftarrow} b_0$. Since a_0, b_0 are normal forms, $a_0 = c = b_0$. Since a, b are arbitrary, every equivalence class of $\overset{*}{\leftrightarrow}$ contains a unique normal form. $\square$

4.2 Bergman's Diamond Lemma

In this section we still let $<X>$ denote the free monoid generated by set X, but X is allowed to be any nonempty set. Let $k<X>$ denote the free associative k-algebra on X, where k is allowed to be any commutative associative ring with 1. Without causing confusion, we will still use terminologies introduced in previous chapters, such as monomials, terms, polynomials, etc. In particular, we assume all polynomials are written in the unique form, *i.e.*,$\forall f \in k{<}X{>} - \{0\}$,

$$f = \sum_{i=1}^{t} c_i m_i,$$

where $t \in \mathbb{N} - \{0\}, c_i \in k - \{0\}, m_i \in {<}X{>}$ and $m_i \neq m_j \ \forall 1 \leqslant i \neq j \leqslant t$. 0 is the unique form of 0.

Notice that $k{<}X{>}$ is also a k-module, we have the following definitions.

Definitions 4.2.1.(i) For $k{<}X{>}$, we define a *reduction system* S to be a set of pairs $\sigma = (m_\sigma, f_\sigma)$ where $m_\sigma \in {<}X{>}$ and $f_\sigma \in k{<}X{>}$. For any $\sigma \in S$, any $l, r \in {<}X{>}$, we define a *reduction* $R_{l\sigma r}$ by a k-module endomorphism of $k{<}X{>}$ which, given any element f of $k{<}X{>}$, sends the monomial $lm_\sigma r$ in f to $lf_\sigma r$ but fixes all other monomials.

(ii) Let $f \in k{<}X{>}$. If the coefficient of $lm_\sigma r$ in f is 0, then $R_{l\sigma r}(f) = f$ and we say $R_{l\sigma r}$ is *trivial on* f. If every reduction under S is trivial on f, *i.e.*, $f = 0$ or no monomial in f is divisible by any m_σ, $\sigma \in S$, we say f is *reduced under* S or S*-reduced*. It's easy to see all S-reduced elements of $k{<}X{>}$ form a k-submodule, which is denoted by $k_R(S)$.

(iii) Let $f \in k{<}X{>}$, if there is a finite sequence of reductions $R_1, R_2, \ldots, R_l$ under S such that $R_l \ldots R_2 R_1(f) = d$ and d is S-reduced, then we say d is a *reduced form of* f under S.

(iv) Let $f \in k{<}X{>}$, if for every infinite sequence of reductions $R_1, R_2, \ldots$, $\exists i \in \mathbb{N}$ such that $\forall j > i$, R_{j+1} is trivial on $R_j \ldots R_2 R_1(f)$, then we say f is *reduction-finite*. It's easy to see, all reduction-finite elements form a k-submodule of $k{<}X{>}$ and if f is reduction-finite, f has a reduced form. We call f *reduction-unique* if it is reduction-finite and has a unique reduced form. The unique reduced form of f under S is denoted by $R_S(f)$.

Remark 4.2.2. A reduction system S actually defines a reduction relation $\to$ on $k{<}X{>}$ by

$$f \xrightarrow{l\sigma r} g \Leftrightarrow \sigma \in S,\ l, r \in {<}X{>}, R_{l\sigma r}(f) = g \text{ and } R_{l\sigma r} \text{ is not trivial on } f.$$

Hence we will make use of notations 4.1.2 to simplify the following discussion.

Definitions 4.2.3. (i) A 5-tuple (σ, τ, l, m, r) with $\sigma, \tau \in S$, $l, m, r \in {<}X{>} - \{1\}$, such that $m_\sigma = lm$, $m_\tau = mr$, is called *an overlap ambiguity of* S. If $f_\sigma r \downarrow l f_\tau$, we say the overlap ambiguity is *resolvable*.

(ii) A 5-tuple (σ, τ, l, m, r) with $\sigma, \tau \in S$, $\sigma \neq \tau$, $l, m, r \in {<}X{>}$, such that $m_\sigma = m$, $m_\tau = lmr$, is called *an inclusion ambiguity of* S. If $l f_\sigma r \downarrow f_\tau$, we say the overlap ambiguity is *resolvable*.

Definitions 4.2.4.(Weaker Monomial Partial Order) (i) In this chapter, by a *monomial partial order* we mean a partial order $\leq$ on ${<}X{>}$such that

$$m_1 \leq m_2 \Rightarrow l m_1 r \leq l m_2 r,\ \forall l, r, m_1, m_2 \in {<}X{>}.$$

(ii) We say $\leq$ satisfies *DCC* if there is no infinite properly descending chain in ${<}X{>}$ w.r.t. $\leq$.

(iii) Given a reduction system S, if for all $\sigma = (m_\sigma, f_\sigma) \in S$, f_σ is a linear combination of monomials $< m_\sigma$, then the monomial partial order $\leq$ is said to be *compatible with* S.

Definition 4.2.5. Let $\leq$ be a monomial partial order on ${<}X{>}$ and compatible with the reduction system S. For any $m \in {<}X{>}$, define I_m to be the submodule of $k{<}X{>}$ spanned by all $l(m_\sigma - f_\sigma)r$ such that $l m_\sigma r < m$. For an overlap(inclusion) ambiguity (σ, τ, l, m, r), if $f_\sigma r - l f_\tau \in I_{lmr}(l f_\sigma r - f_\tau \in I_{lmr})$, then we say the ambiguity is *resolvable relative to* $\leq$.

Lemma 4.2.6. Let $\leq$ be a monomial partial order on $<X>$ and compatible with the reduction system S. If $\leq$ satisfies DCC on $<X>$, then every element of $k<X>$ is reduction-finite.

Proof: Since reduction-finite elements form a k-submodule of $k<X>$, we only need show every monomial is reduction-finite. Assume that

$$N := \{m \in <X>| \ m \text{ is not reduction}-\text{finite}\} \neq \emptyset,$$

then there is a minimal monomial m_0 in N, since $\leq$ satisfies DCC. Then there is some $R_{l\sigma r}$ which is not trivial on m_0 such that $R_{l\sigma r}(m_0) = lf_\sigma r$. By the compatibility of $\leq$ with S and the minimality of m_0 in N, $lf_\sigma r$ is a linear combination of reduction-finite monomials $< m_0$. Then m_0 must be also reduction-finite, a contradiction. Hence $N = \emptyset$, *i.e.*, every monomial is reduction-finite. $\square$

Lemma 4.2.7.(i) $\forall f, g \in k<X>$, $c \in k$, if f, g are reduction-unique, so is $cf + g$. Hence reduction-unique elements also form a k-submodule of $k<X>$. Moreover, $R_S(cf+g) = cR_S(f)+R_S(g)$, thus we can regard R_S as a k-linear map from this submodule to the submodule $k_R(S)$ of S-reduce elements.

(ii) Let $f, g, h \in k<X>$, if for all monomials m_f, m_g, m_h occurring in f, g, h respectively, $m_f m_g m_h$ is reduction-unique, then for any finite composition of reductions, denoted by R for short, $fR(g)h$ is reduction-unique and $R_S(fR(g)h) = R_S(fgh)$.

Proof: A complete proof of (i) can be found in [3]. Here we give a complete proof of (ii).

Claim 1: $\forall m_a, m_b, m_c \in <X>$, if $m_a m_b m_c$ is reduction-unique, then for a single reduction R, $m_a R(m_b) m_c$ is reduction-unique too and $R_S(m_a R(m_b) m_c)$ $= R_S(m_a m_b m_c)$.

Proof of the claim 1: Let $R = R_{l\sigma r}$, notice that

$$m_a R_{l\sigma r}(m_b) m_c = R_{m_a l \sigma r m_c}(m_a m_b m_c),$$

thus $m_a R_{l\sigma r}(m_b) m_c$ must be reduction-finite. Moreover, suppose there is a finite composition of reductions R' such that $R'(m_a R_{l\sigma r}(m_b) m_c)$ is S-reduced, then

$$R'(m_a R_{l\sigma r}(m_b) m_c) = R' R_{m_a l \sigma r m_c}(m_a m_b m_c) = R_S(m_a m_b m_c).$$

Hence $R'(m_a R_{l\sigma r}(m_b) m_c)$ is unique and is $R_S(m_a m_b m_c)$.

The following two claims are immediate results from (i) and claim 1.

Claim 2: Let $f, g, h \in k{<}X{>}$, if for all monomials m_f, m_g, m_h occurring in f, g, h respectively, $m_f m_g m_h$ is reduction-unique, then for a single reduction R, for all monomials $m_f, m_{R(g)}, m_h$ occurring in $f, R(g), h$ respectively, $m_f m_{R(g)} m_h$ is reduction-unique.

Claim 3: Let $f, g, h \in k{<}X{>}$, if for all monomials m_f, m_g, m_h occurring in f, g, h respectively, $m_f m_g m_h$ is reduction-unique, then for a single reduction R, $fR(g)h$ is reduction-unique and $R_S(fR(g)h) = R_S(fgh)$

Now given a finite composition of reductions, by claim 2, we can apply claim 3 recursively, hence (ii) is proved. $\square$

Lemma 4.2.8. Let $\leq$ be a monomial partial order on ${<}X{>}$ and compatible with the reduction system S, then any resolvable ambiguity is resolvable relative to $\leq$.

Proof: The following fact is useful:

$$f \xrightarrow{l\sigma r} g \Rightarrow f - g = cl(m_\sigma - f_\sigma)r, \tag{4.2.1}$$

where $c \in k - \{0\}$, $\sigma \in S$ and $l, r \in {<}X{>}$.

Now for a resolvable overlap ambiguity (σ, τ, l, m, r), $f_\sigma r \downarrow l f_\tau$ implies

$$\begin{aligned} f_\sigma r = f_{10} \to f_{11} \to \ldots \to f_{1a} &= f_0, \\ l f_\tau = f_{20} \to f_{21} \to \ldots \to f_{2b} &= f_0. \end{aligned}$$

Hence

$$f_\sigma r - f_0 = \sum_{i=1}^{a} c_i l_i (m_{\sigma i} - f_{\sigma i}) r_i, \qquad \text{and}$$

$$l f_\tau - f_0 = \sum_{j=1}^{b} c_j l_j (m_{\sigma j} - f_{\sigma j}) r_j.$$

Since $f_\sigma r$ and $l f_\tau$ are linear combinations of monomials $< m_\sigma r = lmr$ or $< l m_\tau = lmr$, each $l_i m_{\sigma i} r_i < lmr$ and each $l_j m_{\sigma j} r_j < lmr$. Hence $f_\sigma r - l f_\tau \in I_{lmr}$, *i.e.*, the ambiguity is resolvable relative to $\leq$.

For a resolvable inclusion ambiguity, proof will be similar thus is omitted.

□

Example 4.2.9. The following example shows the converse of the above lemma is not true. Let

$$S := \{\sigma 1 = (x_2^4, x_1^2), \sigma 2 = (x_2^2, x_1), \sigma 3 = (x_1^2, x_1 x_2)\}$$

be a reduction system of $k{<}X{>}$, where ${<}X{>} = {<}x_1, x_2{>}$. $\leq$ is the *deglex* with $1 < x_2 < x_1$. Clearly $\leq$ is also a monomial partial order and compatible with S. Consider the overlap ambiguity $(\sigma 1, \sigma 1, x_2, x_2^3, x_2)$. Since we have

$$\begin{aligned} f_{\sigma 1} x_2 - x_2 f_{\sigma 1} &= x_1^2 x_2 - x_2 x_1^2 = x_2 (x_2^2 - x_1) x_1 - (x_2^2 - x_1) x_2 x_1 \\ &\quad + x_1 x_2 (x_2^2 - x_1) - x_1 (x_2^2 - x_1) x_2 \in I_{x_2^5}, \end{aligned}$$

the ambiguity is resolvable relative to $\leq$. However, the only possible nontrivial reduction sequences on $f_{\sigma 1}x_2$ and $x_2 f_{\sigma 1}$ are as follows:

$$
\begin{aligned}
f_{\sigma 1}x_2 &= x_1^2 x_2 \xrightarrow{\sigma 3 x_2} x_1 x_2 x_2 \xrightarrow{x_1 \sigma 2} x_1^2 \xrightarrow{\sigma 3} x_1 x_2, \\
x_2 f_{\sigma 1} &= x_2 x_1^2 \xrightarrow{x_2 \sigma 3} x_2 x_1 x_2.
\end{aligned}
$$

Therefore, we cannot have finite compositions of reductions such that $f_{\sigma 1}x_2 \downarrow x_2 f_{\sigma 1}$, *i.e.*, the ambiguity is not resolvable.

In fact the above example can be illustrated by the following figure. Notice that in order to make the ambiguity resolvable, we need $f_{\sigma 1}x_2 \downarrow x_2 f_{\sigma 1}$, *i.e.*, we need two finite sequences of reductions leading from $f_{\sigma 1}x_2$ and $x_2 f_{\sigma 1}$ to a common element of $k{<}X{>}$, but for the ambiguity resolvable relative to $\leq$, by the fact (4.2.1), we only need $f_{\sigma 1}x_2$ and $x_2 f_{\sigma 1}$ are connected by finite reductions staying "below" x_2^5. Clearly the latter is a more general condition.

$$
\begin{tikzcd}
 & & x_2^5 \arrow[dll] \arrow[drr] \arrow[ddll] \arrow[ddrr] & & \\
x_1 x_2^3 \arrow[d] \arrow[drr] & & & & x_2^3 x_1 \arrow[d] \arrow[dll] \\
f_{\sigma 1}x_2 = x_1^2 x_2 \arrow[d] & & x_1 x_2 x_1 & & x_2 x_1^2 = x_2 f_{\sigma 1} \arrow[d] \\
x_1 x_2^2 \arrow[d] & & & & x_2 x_1 x_2 \\
x_1^2 \arrow[d] & & & & \\
x_1 x_2 & & & &
\end{tikzcd}
$$

Theorem 4.2.10.(Bergman's Diamond Lemma) If S is a reduction system for $k{<}X{>}$, $\leq$ is a monomial partial order on ${<}X{>}$ and compatible with S,

$\leq$ satisfies DCC, then the following conditions are equivalent:

(a) All ambiguities of S are resolvable;

(a') All ambiguities of S are resolvable relative to $\leq$;

(b) All elements of $k{<}X{>}$ are reduction-unique under S;

(c) As k-modules, $k{<}X{>} = k_R(S) \bigoplus I$, where I is the two-sided ideal of $k{<}X{>}$ generated by $\{m_\sigma - f_\sigma | \sigma \in S\}$.

Sketch of Proof: We follow the proof given by Bergman. Firstly, by lemma 4.2.6, every element of $k{<}X{>}$ is reduction-finite, thus has a reduced form.

(b)$\Rightarrow$**(c)**: By lemma 4.2.7(i), R_S is a k-linear map from $k{<}X{>}$ onto $k_R(S)$. Hence, we only need show $ker(R_S) = I$, *i.e.*,

$$f \in I \Leftrightarrow R_S(f) = 0. \tag{4.2.2}$$

By lemma 4.2.7(i)(ii), $R_S(l(m_\sigma - f_\sigma)r) = R_S(lm_\sigma r) - R_S(lf_\sigma r) = 0$, thus "$\Rightarrow$" of (4.2.2) is proved. The other direction in (4.2.2) can be proved by the fact (4.2.1). Thus (b)$\Rightarrow$(c) is proved.

(c)$\Rightarrow$**(b)**: Let $R_S(f) = f_1$ or f_2, then by the fact (4.2.1), it's easy to see that $f_1 - f_2 \in I \bigcap k_R(S) = \{0\}$.

(b)$\Rightarrow$**(a)**: Given any overlap or inclusion ambiguity, by **(b)** we will have

$$f_\sigma r \xrightarrow{*} R_S(lmr) \xleftarrow{*} lf_\tau \quad \text{or} \quad lf_\sigma r \xrightarrow{*} R_S(lmr) \xleftarrow{*} f_\tau,$$

thus the ambiguity is resolvable.

(a)$\Rightarrow$**(a')**: Use lemma 4.2.8.

(a')$\Rightarrow$**(b)**: By lemma 4.2.7(i), it's sufficient to show all monomials are reduction-unique. Assume that

$$N := \{m \in {<}X{>} |\ m \text{ is not reduction}-\text{unique}\} \neq \emptyset,$$

then there is a minimal monomial m_0 in N, since $\leq$ satisfies DCC. If for any σ, τ in S, l_1, r_1, l_2, r_2 in $<X>$ such that $m_0 = l_1 m_\sigma r_1 = l_2 m_\tau r_2$ and $R_{l_1\sigma r_1}(m_0) \neq R_{l_2\tau r_2}(m_0)$, we still have

$$R_S(R_{l_1\sigma r_1}(m_0)) = R_S(R_{l_2\tau r_2}(m_0)), \tag{4.2.3}$$

then m_0 would be reduction-unique which leads to a contradiction.

To prove (4.2.3), we assume without loss of generality that $deg(l_1) \leqslant deg(l_2)$, then we have three cases for $m_0 = l_1 m_\sigma r_1 = l_2 m_\tau r_2$,

Case 1: $\exists w \in <X>$ such that $m_0 = l_1 m_\sigma w m_\tau r_2$. Then (4.2.3) can be proved by lemma 4.2.7(ii).

Case 2: $\exists$ an overlap ambiguity (σ, τ, l, m, r) such that $m_0 = l_1 lmr r_2$. Then by (a'), we can show

$$f = R_{l_1\sigma r_1}(m_0) - R_{l_2\tau r_2}(m_0) \in I_{m_0} \quad \text{and} \quad R_S(f) = 0,$$

hence (4.2.3) is proved.

Case 3: The ambiguity is an inclusion ambiguity. The discussion is similar to the case 2. □

Remark 4.2.11. Comparing with Newman's diamond lemma, the strengthening of Bergman's diamond lemma lies in two aspects.

(i) We don't need verify DCC or diamond condition for all elements in $k<X>$. Instead, we only need DCC of a monomial partial order and the diamond condition on "minimal nontrivial ambiguously reducible monomials".

(ii) From the discussion in the example 4.2.9, we can see the condition (a') in Bergman's diamond lemma is a further improvement of the diamond condition.

4.3 Relations between Gröbner Bases and Diamond Lemma(s)

Firstly, we give a brief comment on the relation between Gröbner bases theory and Newman's diamond lemma.

T. Becker and V. Weispfenning have combined general reduction theory and Newman's diamond lemma in their introduction of commutative Gröbner bases theory [8]. The techniques used there are actually applicable to both commutative and noncommutative Gröbner bases theory.

1. We need to define polynomial reductions more carefully by requiring all polynomials given in the unique forms(see (2.1.1)(3.1.1)), then it's easy to see, w.r.t. a given set G and a given monomial order, the new definition of polynomial reductions does define a strictly antisymmetric (reduction) relation on $k[x_1, x_2, \ldots, x_n]$ or $k{<}X{>}$, denoted by $\xrightarrow{G}$.

2. To apply Newman's diamond lemma, we need some techniques to deduce the DCC of the reductions from the DCC of the monomial order. We can apply Bergman's technique which regards a reduction as an endomorphism and then proves every element is reduction-finite(see lemma 4.2.6). Or, we can extend the monomial order to a partial order or a quasi-order $\preceq$ on all polynomials(see [8]) such that $\preceq$ satisfies DCC and

$$f \xrightarrow{g} h \Rightarrow f \succ h, \quad \forall f, g, h \in k[x_1, x_2, \ldots, x_n] - \{0\} \text{ (or } k{<}X{>} - \{0\}),$$

then we have DCC on all reductions.

3. After the above preparations have been done, we can apply Newman's diamond lemma and get some new characterizations of Gröbner bases. The following result is for noncommutative Gröbner bases. For commutative Gröbner bases, see [8].

Theorem 4.3.1. Given $G \subseteq k{<}X{>}$, let $\leq$ be a monomial order on ${<}X{>}$, let $\xrightarrow{G}$ denote the polynomial reduction modulo G w.r.t. $\leq$. The following conditions are equivalent:

(1) $\xrightarrow{G}$ satisfies local confluence condition(diamond condition);

(2) $\xrightarrow{G}$ satisfies confluence condition;

(3) Every element in $k{<}X{>}$ has a unique reduced form w.r.t.G;

(4) $\xrightarrow{G}$ satisfies Church-Rosser property.

Proof: See theorem 4.1.6(a variation of Newman's diamond lemma). $\square$

Clearly, the above **(3)** is the characterization **(e)** in theorem 3.6.1. But notice that in the proof **(e)**$\Rightarrow$**(g)** in theorem 3.6.1, we did *selective* polynomial reductions when reducing $h_2 - h_1$. That is allowed there but not allowed by our new definition of polynomial reductions, since the new definition requires us to do coalescence and cancellation of terms to get *unique forms* of related polynomials before each reduction. However, even under the new definition, we will show that all the conditions in theorem 3.6.1 are still equivalent(see theorem 4.3.4 and remark 4.3.5(i)). Therefore, the above **(1)(2)(3)(4)** are indeed equivalent to characterizations **(a)**–**(h)** in theorem 3.6.1 and are new characterizations of noncommutative Gröbner bases.

From the above, we can see, among the characterizations **(a)**–**(h)** in theorem 3.6.1, only **(e)** is obviously contained in Newman's diamond lemma. Next let's turn to Bergman's diamond lemma. We will see that most important characterizations in theorem 3.6.1 are contained in Bergman's diamond lemma.

Let $k{<}X{>}$ denote the general noncommutative polynomial ring, where k is a field and ${<}X{>}$ is a free monoid generated by $X_n = \{x_1, x_2, \ldots, x_n\}$. To apply Bergman's diamond lemma, we assume all polynomials are given in the unique forms. Let $\leq$ be a monomial order on $<X>$. Let $G \subseteq k<X>$. Notice that: if G is a Gröbner basis in $k<X>$ w.r.t. $\leq$, then so is $G' = \{\frac{1}{lc(g)}g|\ g \in G\}$. Hence, without loss of generality, we assume all g in G is monic, *i.e.*, $lc(g) = 1, \forall g \in G$. Define a reduction system S w.r.t. $\leq$ and G,

$$S := \{\sigma = (\ m_\sigma = lm(g), f_\sigma = lm(g) - g\ )|\ g \in G\}.$$

Clearly $\leq$ is compatible with S and each polynomial reduction modulo G corresponds to a Bergman's "endomorphism" reduction under S. This implies that we may translate definitions and results in section 4.2 to our discussion here. In fact, most translations are obvious. For example, "S-reduced" is equivalent to "reduced w.r.t. G", $k_R(S) = k_R(G)$ and $R_S(f) = R(f, G)$. In particular, let's see the correspondence between ambiguities and S-polynomials.

Correspondence Between Ambiguities and S-polynomials

Given an S-polynomial of $(g_1, g_2) \in G^2$,

$$S(g_1, g_2)[L_1, R_1, L_2, R_2] = L_1 g_1 R_1 - L_2 g_2 R_2,$$

where $(L_1, R_1, L_2, R_2) \in MS(lm(f), lm(g))$.

Case 1: $(L_1, R_1, L_2, R_2) = (1, R_1, L_2, 1)$, $\exists w \neq 1$ such that $wR_1 = lm(g_2)$ and $L_2 w = lm(g_1)$. This corresponds to an overlap ambiguity $(\sigma, \tau, L_2, w, R_1)$ such that $m_\sigma = lm(g_1)$, $m_\tau = lm(g_2)$. Notice that

$$f_\sigma r - l f_\tau = -S(g_1, g_2)[L_1, R_1, L_2, R_2].$$

Case 2: $(L_1, R_1, L_2, R_2) = (L_1, 1, 1, R_2)$, $\exists w \neq 1$ such that $wR_2 = lm(g_1)$ and $L_1 w = lm(g_2)$. This corresponds to an overlap ambiguity $(\sigma, \tau, L_1, w, R_2)$ such that $m_\sigma = lm(g_2)$, $m_\tau = lm(g_1)$. Notice that

$$f_\sigma r - l f_\tau = S(g_1, g_2)[L_1, R_1, L_2, R_2].$$

Case 3: $(L_1, R_1, L_2, R_2) = (1, 1, L_2, R_2)$, $lm(g_1) = L_2 lm(g_2) R_2$. This corresponds to an inclusion ambiguity $(\sigma, \tau, L_2, lm(g_2), R_2)$ such that $m_\sigma = lm(g_2)$, $m_\tau = lm(g_1)$. Notice that

$$l f_\sigma r - f_\tau = S(g_1, g_2)[L_1, R_1, L_2, R_2].$$

Case 4: $(L_1, R_1, L_2, R_2) = (L_1, R_1, 1, 1)$, $lm(g_2) = L_1 lm(g_1) R_1$. This corresponds to an inclusion ambiguity $(\sigma, \tau, L_1, lm(g_1), R_1)$ such that $m_\sigma = lm(g_1)$, $m_\tau = lm(g_2)$. Notice that

$$l f_\sigma r - f_\tau = -S(g_1, g_2)[L_1, R_1, L_2, R_2].$$

Conversely, given an overlap ambiguity (σ, τ, l, m, r), by the definition of S, $\exists g_1, g_2 \in G$ such that $m_\sigma = lm(g_1) = lm$, $m_\tau = lm(g_2) = mr$. Then $(1, r, l, 1) \in MS(lm(g_1), lm(g_2))$ and notice that

$$f_\sigma r - l f_\tau = -S(g_1, g_2)[1, r, l, 1].$$

Given an inclusion ambiguity (σ, τ, l, m, r), $\exists g_1, g_2 \in G$ such that $m_\sigma = lm(g_1) = m$, $m_\tau = lm(g_2) = lmr$. Then $(l, r, 1, 1) \in MS(lm(g_1), lm(g_2))$ and notice that

$$l f_\sigma r - f_\tau = -S(g_1, g_2)[l, r, 1, 1].$$

Although the above correspondences are not required to be one-to-one, they are sufficient for us to deduce the following equivalence.

Claim 4.3.2.(i) All ambiguities of S are resolvable. $\Leftrightarrow$

$\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$L_1(lm(g_1) - g_1)R_1 \downarrow L_2(lm(g_2) - g_2)R_2.$$

(ii) All ambiguities of S are resolvable relative to $\leq$. $\Leftrightarrow$

$\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$S(g_1, g_2)[L_1, R_1, L_2, R_2] = \sum_{i=1}^{t} c_i l_i g_i r_i \qquad \text{and}$$

$$\max_{1 \leqslant i \leqslant t} \{l_i lm(g_i) r_i\} < L_1 lm(g_1) R_1 = L_2 lm(g_2) R_i,$$

where $t \in \mathbb{N} - \{0\}$, $c_i \in k - \{0\}$, $l_i, r_i \in <X>$, $g_i \in G$ and not necessarily pairwise distinct $\forall\, i$, $1 \leqslant i \leqslant t$.

With the above results, we can translate Bergman's diamond lemma as follows.

Theorem 4.3.3. Given $G \subseteq k<X>$, let $I =< G >$ be the ideal generated by G, let $\leq$ be a monomial order on $<X>$. The following conditions are equivalent:

(1) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$L_1(lm(g_1) - g_1)R_1 \downarrow L_2(lm(g_2) - g_2)R_2.$$

(2) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$S(g_1, g_2)[L_1, R_1, L_2, R_2] = \sum_{i=1}^{t} c_i l_i g_i r_i \qquad \text{and}$$

$$\max_{1 \leqslant i \leqslant t} \{l_i lm(g_i) r_i\} < L_1 lm(g_1) R_1 = L_2 lm(g_2) R_i,$$

where $t \in \mathbb{N} - \{0\}$, $c_i \in k - \{0\}$, $l_i, r_i \in <X>$, $g_i \in G$ and not necessarily pairwise distinct $\forall\, i$, $1 \leqslant i \leqslant t$.

(3) $\forall f \in k<X>$, the reduced form of f w.r.t. G is unique;

(4) As k-vector spaces, $k<X> = k_R(G) \bigoplus I$.

Proof: By the above discussion, the proof of Bergman's diamond lemma has actually shown that **(3)⇒(4)⇒(3)** and **(3)⇒(1)⇒(2)⇒(3)**. □

Moreover, in the proof of Bergman's diamond lemma, **(b)⇒(c)** contains the following condition(see (4.2.2)).

(5) $f \in I \Leftrightarrow R(f, G) = 0$.

By the proof there, we can see **(3)⇒(5)⇒(4)⇒(3)**.

Notice that all S-polynomials are in the ideal I, hence **(5)** implies the following condition.

(6) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$R(S(g_1, g_2)[L_1, R_1, L_2, R_2], G) = 0.$$

From **(6)**, it's easy to deduce a condition about standard representations.

(7) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$, $S(g_1, g_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G.

Notice that **(7)** actually is a strengthening of **(2)**, so **(7)** implies **(2)** obviously. Therefore we have a cycle **(3)⇒(5)⇒(6)⇒(7)⇒(2)⇒(3)**.

So far, most important characterizations of noncommutative Gröbner bases have been deduced from Bergman's diamond lemma.

At last, we conclude our discussion by listing all characterizations we have found for noncommutative Gröbner bases and summarize the proof based on diamond lemmas.

Theorem 4.3.4. Assume ($\star$) all polynomials of $k{<}X{>}$ are given in the unique forms (3.1.1). Given $G \subseteq k{<}X{>}$, let $I = <G>$ be the ideal generated by G, let $\leq$ be a monomial order on $<X>$, let $\xrightarrow{G}$ denote the polynomial reduction modulo G w.r.t.$\leq$. The following conditions are equivalent:

(1) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$L_1(lm(g_1) - g_1)R_1 \downarrow L_2(lm(g_2) - g_2)R_2.$$

(2) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$S(g_1, g_2)[L_1, R_1, L_2, R_2] = \sum_{i=1}^{t} c_i l_i g_i r_i \qquad \text{and}$$

$$\max_{1 \leqslant i \leqslant t} \{l_i lm(g_i) r_i\} < L_1 lm(g_1) R_1 = L_2 lm(g_2) R_i,$$

where $t \in \mathbb{N} - \{0\}$, $c_i \in k - \{0\}$, $l_i, r_i \in <X>$, $g_i \in G$ and not necessarily pairwise distinct $\forall\, i,\ 1 \leqslant i \leqslant t$.

(3) $\forall f \in k<X>$, the reduced form of f w.r.t. G is unique.

(4) As k-vector spaces, $k<X> = k_R(G) \bigoplus I$.

(5) $f \in I \Leftrightarrow R(f, G) = 0$.

(6) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$,

$$R(S(g_1, g_2)[L_1, R_1, L_2, R_2], G) = 0.$$

(7) $\forall (g_1, g_2) \in G^2$, $\forall (L_1, R_1, L_2, R_2) \in MS(lm(g_1), lm(g_2))$, $S(g_1, g_2)[L_1, R_1, L_2, R_2]$ has a standard representation w.r.t. G.

(8) $\xrightarrow{G}$ satisfies local confluence condition(diamond condition).

(9) $\xrightarrow{G}$ satisfies confluence condition.

(10) $\xrightarrow{G}$ satisfies Church-Rosser property.

(11) $lm(G) = lm(I)$.

(12) $\forall f \in I - \{0\}$, $\exists g \in G$ such that $lm(g) |\ lm(f)$.

(13) $f \in I \Leftrightarrow f$ has a standard representation w.r.t. G.

Proof: From Bergman's diamond lemma, we have deduced

$$(\mathbf{3}) \Rightarrow (\mathbf{5}) \Rightarrow (\mathbf{4}) \Rightarrow (\mathbf{3}),$$

$$(\mathbf{3}) \Rightarrow (\mathbf{1}) \Rightarrow (\mathbf{2}) \Rightarrow (\mathbf{3}) \qquad \text{and}$$

$$(\mathbf{3}) \Rightarrow (\mathbf{5}) \Rightarrow (\mathbf{6}) \Rightarrow (\mathbf{7}) \Rightarrow (\mathbf{2}) \Rightarrow (\mathbf{3}).$$

Hence **(1)**$\Leftrightarrow$**(2)**$\Leftrightarrow$**(3)**$\Leftrightarrow$**(4)**$\Leftrightarrow$**(5)**$\Leftrightarrow$**(6)**$\Leftrightarrow$**(7)**.

Newman's diamond lemma ensures **(3)**$\Leftrightarrow$**(8)**$\Leftrightarrow$**(9)**$\Leftrightarrow$**(10)**.

The proofs in theorem 3.6.1 for **(11)**$\Rightarrow$**(12)**$\Rightarrow$**(5)** and **(13)**$\Rightarrow$**(11)** are still effective under the assumption ($\star$). It's obvious that **(5)**$\Rightarrow$**(13)**. Hence, **(11)**$\Leftrightarrow$**(12)**$\Leftrightarrow$**(5)** $\Leftrightarrow$**(13)**$\Leftrightarrow$**(11)**.

To sum up, all the conditions are equivalent. $\square$

Remarks 4.3.5. (i) The above theorem contains all characterizations in theorem 3.6.1. This implies that theorem 3.6.1 is still true under the assumption ($\star$). In other word, the assumption ($\star$) has no effect on the characterizations of Gröbner bases.

(ii) Newman's diamond lemma and Bergman's diamond lemma actually form a common theoretical foundation of characterizations of both commutative and noncommutative Gröbner bases. This explains why theorem 2.4.2 and theorem 3.6.1 are almost the same.

Bibliography

[1] A.Kandri-Rody and V.Weispfenning. Non-commutative Gröbner bases in Algebras of Solvable Type. *J.Symbolic Computation*, 9:1–26, 1990.

[2] B.Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. Univ. of Innsbruck Math.Inst., PhD thesis, 1965.

[3] G.M.Bergman. The Diamond Lemma for Ring Theory. *Adv.Math*, 29:178–218, 1978.

[4] H.Li. *Noncommutative Gröbner bases and filtered-graded transfer*. Springer, Berlin,New York, 2002.

[5] M.H.A.Newman. On Theories with a Combinatorial Definition of "Equivalence". *Annals of Math.*, 43:223–243, 1942.

[6] P.M.Cohn. *Algebra Volume 3*. J.Wiley, Chichester,Toronto, second edition, 1991.

[7] R.Fröberg. *Introduction to Gröbner Bases*. John Wiley and Sons, Chichester,New York, 1997.

[8] T.Becker and V.Weispfenning. *Gröbner bases: a computational approach to commutative algebra*. Springer-Verlag, New York, 1991.

[9] T.Mora. Gröbner bases for non-commutative polynomial rings. *AAECC3,Springer LNCS*, 229:353–362, 1986.

[10] T.Mora. An introduction to commutative and noncommutative Gröbner bases. *Theoretical Computer Science*, 134:131–173, 1994.

[11] T.W.Hungerford. *Algebra*. Springer-Verlag, New York, 1980.

[12] W.W.Adams and P.Loustaunau. *An introduction to Gröbner bases.* AMS, Providence,R.I., 1994.

List of Notations

$\mathbb{N}$	set of natural numbers including 0, page 4
$k[x_1, x_2, \ldots, x_n]$	commutative polynomial ring, page 4
M_n or M	set of commutative monomials, page 4
$<X_n>$ or $<X>$	set of noncommutative monomials(free monoid), page 19, page 56
$k{<}X_n{>}$or $k{<}X{>}$	noncommutative polynomial ring(free algebra) page 20, page 56
$<X>^2$	set of ordered pairs of elements in $<X>$, page 37
$<X>^4$	set of ordered 4-tuples of elements in $<X>$, page 37
$<G>$	ideal generated by G, page 7, page 22
$S \times S$	set of ordered pairs of elements in set S, page 5
lex	lexicographical order, page 6, page 21
deglex	degree lexicographical order, page 6, page 21
degrevlex	degree reverse lexicographical order, page 7, page 21

$deg(m)$	degree of monomial m, page 4, page 20
$deg(g)$	degree of leading monomial of g, page 51
$lt(f)$	leading term of f, page 7, page 22
$lm(f)$	leading monomial of f, page 7, page 22
$lc(f)$	leading coefficient of f, page 7, page 22
$lm(G)$	leading monomial ideal of set G, page 7, page 22
$M(G)$	set of all monomials in $lm(G)$, page 12, page 31
$k_R(G)$	set of all reduced polynomials w.r.t. G, page 12, page 31
$f \xrightarrow{g} h$	f reduces to h modulo g, page 11, page 31
$f \xrightarrow{G} h_t$	f reduces to h_t modulo G, page 12, page 31
$R(f, G)$	unique reduced form of f w.r.t. G, page 12, page 31
$lcm(m_1, m_2)$	least common multiple, page 15
$S(f, g)$	S-polynomial of f and g, page 15
$T(m_1, m_2)$	set of 4-tuples $(l_1, r_1, l_2, r_2) \in <X>^4$ satisfying $l_1 m_1 r_1 = l_2 m_2 r_2$, page 37
$MS(m_1, m_2)$	set of matches of m_1 and m_2, page 38
$S(f, g)[L_1, R_1, L_2, R_2]$	noncommutative S-polynomial of f and g, page 39

$\dot{\bigcup}$	disjoint union, page 52
$a \downarrow b$, $a \overset{*}{\leftrightarrow} b$, etc.	page 55
$R_{l\sigma r}$	endomorphism reduction, page 57
$k_R(S)$	set of all S-reduced elements, page 57
$R_S(f)$	unique reduced form of f under S, page 58
$f \xrightarrow{l\sigma r} g$	$R_{l\sigma r}(f) = g$, page 58
(σ, τ, l, m, r)	overlap or inclusion ambiguity of S, page 58
I_m	page 58
$\xrightarrow{G}$	polynomial reduction modulo G, page 65, page 66
DCC	descending chain condition, page 5, page 55
ACC	ascending chain condition, page 8, page 23