COMMUNICATIONS IN ALGEBRA, 29(11), 4831–4851 (2001)

# NON-COMMUTATIVE GRÖBNER BASES UNDER COMPOSITION

**Patrik Nordbeck**

Lund University, Centre for Mathematical Sciences,
P.O. Box 118, SE-221 00 Lund, Sweden
E-mail: nordbeck@maths.lth.se

## ABSTRACT

Polynomial composition is the operation of replacing the variables in a polynomial with other polynomials. In this paper we give sufficient and necessary conditions on a set $\Theta$ of non-commutative polynomials to assure that the set $G \circ \Theta$ of composed polynomials is a Gröbner basis in the free associative algebra whenever $G$ is. The subject was initiated by Hong, treating the commutative analogue in (1998, J. Symb. Comput. *25*, 643–663).

*Key Words:* Non-commutative Gröbner bases; Composition of polynomials

## INTRODUCTION

In the recent paper [6], Hoon Hong addresses the problem of the behavior of (commutative) Gröbner bases under composition of polynomials. More precisely, let $\Theta$ be a set of polynomials, as many as the variables in our polynomial ring. The question then is under which conditions on these polynomials it is true that for an arbitrary Gröbner basis $G$ (under some term ordering), also the composed set $G \circ \Theta$ is a Gröbner basis (under the same ordering). The main result in [6] is that this happens if and only if the composition is "compatible" with the ordering and the non-divisibility (see Section 2).

Since many of the basic properties of Gröbner bases transfer to the non-commutative polynomial ring, it seems natural to ask under which conditions non-commutative Gröbner bases are preserved after composition by a set of (non-commutative) polynomials. The main contribution of this paper is to show that this is the case if and only if the composition is compatible with the ordering, and the set of leading words of $\Theta$ is combinatorially free.

Not surprisingly, the non-commutative case gets considerably more complicated, but we can still use some of Hong's ideas; *e.g.* Lemma 3 (and its proof) below is identical with the corresponding in [6].

Finally we mention that the subject, for the commutative case, has been studied further by Hong in [5], and by Gutierrez and San Miguel in [4]. The first of these papers is devoted to the case where the composed Gröbner bases may be under a possibly different ordering, the second concerns *reduced* Gröbner bases under composition.

## 1 BASIC DEFINITIONS AND NOTATION

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite alphabet, and let $K\langle X\rangle$ denote the free associative algebra over the arbitrary field $K$. We will assume that $n \geq 2$, for $K\langle x_1\rangle$ is equal to $K[x_1]$ (the commutative polynomial ring), and this case is covered by [6]. Denote by $W$ the set of all words in $X$, including the empty word **1** (*i.e.* $W$ is the free monoid generated by $X$).

We will always in what follows assume that $W$ is given an *admissible ordering*, *i.e.* a well-ordering preserving multiplication: $f < g$ implies $hfk < hgk$ for all $f, g, h, k \in W$, and the smallest word is the unity **1**.

In the examples below we will use the following admissible ordering called *deglex* (degree lexicographical): If $|u|$ denotes the length of $u \in W$, then we let $u > v$ if either $|u| > |v|$, or $|u| = |v|$ but $u$ is larger than $v$ lexicographically with $x_n > \cdots > x_1$.

When we have chosen an admissible ordering we can, if terms with identical words are collected together using the operations over $K$, with every non-zero element $f \in K\langle X\rangle$ associate its *leading word* $\hat{f} \in W$, *i.e.* the word in $f$ that is larger (relative the ordering) than every other word occurring in $f$. We also define, for a subset $F \subset K\langle X\rangle$, $\widehat{F} = \{\hat{f} \,|\, f \in F\}$. The *leading coefficient* of $f \in K\langle X\rangle$, *i.e.* the coefficient of $\hat{f}$, will be denoted $\mathrm{lc}_f$.

If $u, v \in W$, and $u$ is a (not necessarily proper) subword of $v$, then we write $u \mid v$. In this case we have, since our ordering is preserved by multiplication, $u \leq v$. We will use this frequently in the form

$$u > v \quad \Longrightarrow \quad u \nmid v \tag{1}$$

($\nmid$ meaning that $u$ is not a subword of $v$).

### 1.1 Gröbner Bases

We here gather the theory concerning non-commutative Gröbner bases that we will need. For a more complete exposition we refer to [8].

**Definition 1.** *A subset $G$ of an ideal $I$ (always two-sided) in $K\langle X\rangle$ is called a* Gröbner basis *for $I$ if $0 \notin G$, and for every $f \in I$, $f \neq 0$, there is $g \in G$ such that $\hat{g} \mid \hat{f}$.*

We can show that if $G$ is a Gröbner basis for $I$, then $G$ generates $I$. We may (and will) therefore simply say that $G$ is a Gröbner basis, meaning that $G$ is a Gröbner basis for the ideal generated by $G$.

We borrow the following terminology from [1] and [3].

**Definition 2.** *We will call a subset $\{w_1, w_2, \ldots\} \subset W$* combinatorially free *if*

1. *whenever $w_i = u_i v_i$ and $w_j = u_j v_j$* for $u_i, v_i, u_j, v_j \in W \setminus \{\mathbf{1}\}$ *we have $v_i \neq u_j$, and*
2. *no $w_i$ is a subword of $w_j$ for $i \neq j$.*

*A pair $w_i, w_j$ violating one of the conditions above is said to form an* overlap.

**Remark 1.** If two words $w_i, w_j \in W$ form an overlap, then we have

1. $w_i u = v w_j$ (alt. $u w_i = w_j v$), $|u| < |w_j|, |v| < |w_i|$, if $w_i, w_j$ violate the first condition above, or
2. $w_i = u w_j v$ (alt. $u w_i v = w_j$) if the second condition is violated.

Here $u, v \in W$, not equal to $\mathbf{1}$ in case 1, possibly equal to $\mathbf{1}$ in case 2. If we are only interested in whether a given set $W' \subset W$ is combinatorially free, then we can of course interchange the meaning of $w_i, w_j \in W'$. We will thus in the sequel consider only those overlaps above without brackets.

We will also call the explicit representations $w_i u = v w_j$ and $w_i = u w_j v$ overlaps. Note that a pair of words then can form several overlaps; *e.g.* $w_1 = x_1^2$ and $w_2 = x_1^2 x_2$ form the overlaps $w_1 x_1 x_2 = x_1 w_2$ and $w_1 x_2 = w_2$. Moreover, a single word, *e.g.* $w = x_1^3$, can form (several) overlaps with itself.

**Definition 3.** *If the leading words of $f_1, f_2 \in K\langle X\rangle$ form an overlap, then we call $(f_1, f_2)$ a critical pair. We then define the overlap relations of $(f_1, f_2)$ as $f_1 u - c v f_2$ for each overlap $\hat{f}_1 u = v \hat{f}_2$, and $f_1 - c u f_2 v$ for each overlap $\hat{f}_1 = u \hat{f}_2 v$. Here $c \in K$ is equal to $\mathrm{lc}_{f_1}$ divided by $\mathrm{lc}_{f_2}$, i.e. such that the leading words in the relations cancel.*

**Example 1.** Consider $F = \{f_1 = x_2^2 - x_1^2, f_2 = x_2 x_1^2 - x_1^2 x_2\}$ using deglex $(x_2 > x_1)$. We have $\widehat{F} = \{\hat{f}_1 = x_2^2, \hat{f}_2 = x_2 x_1^2\}$, and the overlaps of $\widehat{F}$ are $\hat{f}_1 x_2 = x_2^3 = x_2 \hat{f}_1$ and $\hat{f}_1 x_1^2 = x_2^2 x_1^2 = x_2 \hat{f}_2$. The overlap relations become

$$f_1 x_2 - x_2 f_1 = (x_2^2 - x_1^2) x_2 - x_2 (x_2^2 - x_1^2) = x_2 x_1^2 - x_1^2 x_2$$

and

$$f_1 x_1^2 - x_2 f_2 = (x_2^2 - x_1^2) x_1^2 - x_2 (x_2 x_1^2 - x_1^2 x_2) = x_2 x_1^2 x_2 - x_1^4.$$

We will also use the characterization of Gröbner bases presented in the following theorem. The proof is essentially the same as the proof of *e.g.* Theorem 3.3 (in particular Lemma 2.4) in [8].

**Theorem 1.** *A set $G$ is a Gröbner basis if and only if each overlap relation of every critical pair $(g, g')$ of $G$ either is equal to zero, or can be written*

$$g u - c v g' = \sum_{k=1}^{t} c_k w_{k_L} g_k w_{k_R}, \quad w_{k_L} \hat{g}_k w_{k_R} < \hat{g} u = v \hat{g'} \quad \forall k \tag{2}$$

*alternatively*

$$g - c u g' v = \sum_{k=1}^{t} c_k w_{k_L} g_k w_{k_R}, \quad w_{k_L} \hat{g}_k w_{k_R} < \hat{g} = u \hat{g'} v \quad \forall k, \tag{3}$$

*where $c_k \in K$, $g_k \in G$ and $w_{k_L}, w_{k_R} \in W$ for all $k$. Furthermore, we can then assume that $w_{1_L} \hat{g}_1 w_{1_R} > \cdots > w_{t_L} \hat{g}_t w_{t_R}$.*

We will need the following Gröbner bases later on.

**Lemma 1.** *The following subsets of $K\langle X\rangle$ are all Gröbner bases (in every admissible ordering):*

*i.* $\{x_i^k\},\ k \geq 1$
*ii.* $\{x_i + \mathbf{1}\}$

*iii.* $\{(x_i x_j)^k\},\ k \geq 1$
*iv.* $\{x_i x_j + \mathbf{1}\},\ i \neq j$
*v.* $\{x_i, x_j\}$
*vi.* $\{x_i + \mathbf{1}, x_j\},\ i \neq j$
*vii.* $\{(x_i x_j x_i)^k\},\ k \geq 1$
*viii.* $\{f_1 = x_i x_j x_i + x_j,\ f_2 = x_i x_j x_j - x_j x_j x_i\}$
*ix.* $\{u - cv, v\},\ c \in K,\ u, v \in W,\ u > v$

*Proof.* Every set consisting only of words is a Gröbner basis, for every overlap relation is then equal to zero. This proves i, iii, v and vii. In ii, iv and vi, we have no overlaps, and thus no overlap relations. For ix, we note that the ideal generated by $F_1 = \{u - cv, v\}$ is the same as the one generated by $F_2 = \{u, v\}$, and $\widehat{F}_1 = \widehat{F}_2$ $(u > v)$. Since $\{u, v\}$ is a Gröbner basis, the same must then be true for ix. In viii, either $x_i x_j x_j > x_j x_j x_i$ or vice versa. Assuming the first (the other being symmetrical), we can write the two possible overlap relations as

$$f_1 x_j x_i - x_i x_j f_1 = -f_2, \quad f_1 x_j x_j - x_i x_j f_2 = f_2 x_j x_i + x_j x_j f_1.$$

Using $x_i x_j x_j > x_j x_j x_i$, we see that these are representations as in Theorem 1, and our lemma is proved.

### 1.2 Composition of Polynomials

We now define the process of composition of polynomials.

**Definition 4.** *Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a subset of $K\langle X\rangle$ $(= K\langle x_1, \ldots, x_n\rangle)$, and let $f \in K\langle X\rangle$. We define the* composition *of $f$ by $\Theta$, written $f \circ \Theta$, as the polynomial obtained from $f$ by replacing each occurrence of the $x_i$ with $\theta_i$. We also define, for $F \subset K\langle X\rangle$, $F \circ \Theta = \{f \circ \Theta \mid f \in F\}$.*

We clearly have, for $f, g \in K\langle X\rangle$,

$$(fg) \circ \Theta = f \circ \Theta\, g \circ \Theta, \tag{4}$$
$$(f + g) \circ \Theta = f \circ \Theta + g \circ \Theta. \tag{5}$$

Since our ordering is preserved by multiplication, we also have, for every word $w \in W$,

$$\widehat{w \circ \Theta} = w \circ \widehat{\Theta}. \tag{6}$$

**Definition 5.** *We say that composition by* $\Theta$ commutes with Gröbner bases computation *if for every Gröbner basis $G$, also $G \circ \Theta$ is a Gröbner basis (under the same ordering as $G$).*

As mentioned before, our main task in this paper will be to decide under which conditions composition by $\Theta$ commutes with Gröbner bases computation.

**Remark 2.** In Hong's paper, the counterpart of Definition 5 requires that if $G$ is a Gröbner basis for the ideal generated by a set of polynomials $F$, then $G \circ \Theta$ is a Gröbner basis for the ideal generated by $F \circ \Theta$. That this implies the statement in Definition 5 is clear (take $F = G$). The two formulations are in fact equivalent since it is easy to prove that $\langle G \rangle = \langle F \rangle$ implies $\langle G \circ \Theta \rangle = \langle F \circ \Theta \rangle$. Here we have used the notation $\langle F \rangle$ for the ideal generated by $F$.

**Definition 6.** *We say that composition by* $\Theta$ *is* compatible with our given ordering *if for all words* $u, v \in W$, *we have*

$$u > v \implies u \circ \widehat{\Theta} > v \circ \widehat{\Theta}. \tag{7}$$

Now, let $f \in K\langle X \rangle$ be written as a linear combination of words in decreasing order: $f = c_1 w_1 + \cdots + c_s w_s$, $w_1 > \cdots > w_s$. If composition by $\Theta$ is compatible with our ordering, then we have $w_1 \circ \widehat{\Theta} > \cdots > w_s \circ \widehat{\Theta}$, so using (5) and (6) we get

$$\widehat{f \circ \Theta} = \hat{f} \circ \widehat{\Theta}. \tag{8}$$

## 2 COMPARISON WITH THE COMMUTATIVE CASE

The main theorem in Hong's paper [6] is the following, where of course all statements are in a commutative meaning.

**Theorem 2** (Theorem 3.1. in [6]). *Composition by* $\Theta$ *commutes with Gröbner bases computation if and only if both of the following conditions hold:*

1. *Composition by* $\Theta$ *is compatible with the ordering and*
2. *for all monomials (commutative words)* $m_1, m_2 \in K[X]$, $m_1 \nmid m_2$ *implies* $m_1 \circ \widehat{\Theta} \nmid m_2 \circ \widehat{\Theta}$.

In [6], it is constantly used that condition [2] above is equivalent to that $\widehat{\Theta}$ is a *permuted powering*, *i.e.* $\widehat{\Theta} = \{x_{\pi(1)}^{\lambda_1}, \ldots, x_{\pi(n)}^{\lambda_n}\}$ for some permutation $\pi \in S^n$ and some $\lambda_1, \ldots, \lambda_n > 0$. This is not the case in our non-commutative setting; it is not hard to see that *e.g.* $\widehat{\Theta} = \{x_1^2 x_2, x_1 x_2^2\} \subset K\langle x_1, x_2 \rangle$ fulfills the non-commutative version of condition 2 above.

However, if $\widehat{\Theta}$ is a permuted powering, then it is easy to see that condition 2 (in a non-commutative sense) is true for $\Theta$. The following

example then shows that conditions 1 and 2 above are not sufficient in the non-commutative case.

**Example 2.** Let $G = \{x_2 - x_1\} \subset K\langle x_1, x_2\rangle$ and $\Theta = \{\theta_1 = x_1^2, \theta_2 = x_2^2\}$ $(= \widehat{\Theta})$. It is obvious that $G$ is a Gröbner basis, and that composition by $\Theta$ is compatible with the deglex ordering $(x_2 > x_1)$. But $G \circ \Theta = \{x_2^2 - x_1^2\}$ is not a Gröbner basis since for example

$$f = (x_2^2 - x_1^2)x_2 - x_2(x_2^2 - x_1^2) = x_2x_1^2 - x_1^2x_2 \in G \circ \Theta\,,$$

and $x_2^2 \nmid \hat{f} = x_2x_1^2$. The (reduced) Gröbner basis for $G \circ \Theta$ is the completed set $\{x_2^2 - x_1^2, x_2x_1^2 - x_1^2x_2\}$ (compare Example 1).

The crucial condition we *can* generalize from the commutative case is whether the least common multiples coincide in the sense that

$$\mathrm{lcm}(m_1, m_2) \circ \widehat{\Theta} = \mathrm{lcm}(m_1 \circ \widehat{\Theta}, m_2 \circ \widehat{\Theta}) \tag{9}$$

for all monomials $m_1, m_2 \in K[X]$. We can not define the least common multiple in the non-commutative case, but this role is in some sense played by the overlaps in Definition 2. The counterpart of (9) will be handled in Lemma 2 below.

That $\widehat{\Theta}$ is a permuted powering implies, in our non-commutative setting, that there are no overlaps formed by two different words from $\widehat{\Theta}$. But this is not sufficient; we saw in Example 2 that a word can form an overlap with itself. To handle such "self-overlaps", we need to replace condition 2 with the statement that $\widehat{\Theta}$ must be combinatorially free.

The fact that a word can form an overlap with itself will cause most difficulties in Section 3.2 below. Compared to the commutative case, we will need a lot more Gröbner bases as counterexamples to show the necessity in our main theorem.

## 3 MAIN THEOREM

We now state our main result.

**Theorem 3.** *Composition by $\Theta$ commutes with Gröbner bases computation if and only if both of the following conditions hold:*

1. *Composition by $\Theta$ is compatible with the ordering and*
2. *$\widehat{\Theta}$ is combinatorially free.*

We will prove this theorem in the following two sections. We start with the easiest part; to show that the two conditions above are sufficient for the commutation. In the second section we show that these conditions are also necessary.

### 3.1 Proof of Sufficiency

Condition 1 is used to obtain (8): $\widehat{f \circ \Theta} = \widehat{f \circ \widehat{\Theta}}$ for all $f \in K\langle X \rangle$. The second condition will ensure that every overlap relation of $G \circ \Theta$ corresponds to an overlap relation of $G$.

The key to the sufficiency in Theorem 3 is the following lemma.

**Lemma 2.** *Assume that composition by $\Theta$ is compatible with the ordering, and that $\widehat{\Theta}$ is combinatorially free. If $(g \circ \Theta, g' \circ \Theta)$ is a critical pair of $G \circ \Theta$, then $(g, g')$ is a critical pair of $G$. Moreover, each overlap relation of $(g \circ \Theta, g' \circ \Theta)$ is of the form*

$$g \circ \Theta\, u \circ \widehat{\Theta} - c_1 v \circ \widehat{\Theta} g' \circ \Theta \quad (\textit{alt. } g \circ \Theta - c_2 u \circ \widehat{\Theta} g' \circ \Theta\, v \circ \widehat{\Theta}),$$

*where $gu - cvg'$ $(g - cug'v)$ is an overlap relation of $(g, g')$, $c_1 = c\, lc_{v \circ \Theta} / lc_{u \circ \Theta}$ and $c_2 = c\, lc_{u \circ \Theta}\, lc_{v \circ \Theta}$.*

*Proof.* Assume that $\hat{g} = x_{i_1} \cdots x_{i_s}$ and $\hat{g}' = x_{j_1} \cdots x_{j_t}$. By (8) we have *e.g.* $\widehat{g \circ \Theta} = \hat{g} \circ \widehat{\Theta}$, so the leading words of $g \circ \Theta$ and $g' \circ \Theta$ are products of words from $\widehat{\Theta}$; in our case $\widehat{g \circ \Theta} = \hat{\theta}_{i_1} \cdots \hat{\theta}_{i_s}$ and $\widehat{g' \circ \Theta} = \hat{\theta}_{j_1} \cdots \hat{\theta}_{j_t}$. If $(g \circ \Theta, g' \circ \Theta)$ is a critical pair, then $\widehat{g \circ \Theta}$ and $\widehat{g' \circ \Theta}$ form at least one overlap, *i.e.* the two products of words from $\widehat{\Theta}$ "intersect" (recall Definition 2). Since $\widehat{\Theta}$ is combinatorially free, there can not be any overlaps among the $\hat{\theta}_k$. This means that if an overlap is of the form $\widehat{g \circ \Theta}\, u' = v' \widehat{g' \circ \Theta}$, then we must have the situation

$$\begin{aligned} \widehat{g \overset{\circ}{\Theta}}\, u' &= \overbrace{\quad}^{\hat{\theta}_{i_1}} \cdots \overbrace{\quad}^{\hat{\theta}_{i_{k-1}}} \overbrace{\quad}^{\hat{\theta}_{i_k}} \cdots \overbrace{\quad}^{\hat{\theta}_{i_s}} \overbrace{\qquad}^{u'} \\ v' \widehat{g' \overset{\circ}{\Theta}} &= \underbrace{\qquad}_{v'} \underbrace{\quad}_{\hat{\theta}_{j_1}} \cdots \underbrace{\quad}_{\hat{\theta}_{j_l}} \underbrace{\quad}_{\hat{\theta}_{j_{l+1}}} \cdots \underbrace{\quad}_{\hat{\theta}_{j_t}}, \end{aligned}$$

and $\theta_{i_k} = \theta_{j_1}, \ldots, \theta_{i_s} = \theta_{j_l}$. The latter implies $x_{i_k} = x_{j_1}, \ldots, x_{i_s} = x_{j_l}$, so $(g, g')$ is a critical pair of $G$. If we let $u = x_{i_1} \cdots x_{i_{k-1}}$ and $v = x_{j_l} \cdots x_{j_t}$, then we see that $u' = u \circ \widehat{\Theta}, v' = v \circ \widehat{\Theta}$ and $\hat{g}u = v\hat{g}'$. The statement about the overlap relations is thus clear, except maybe for the constant $c_1$. Since $\mathrm{lc}_{gu} = \mathrm{lc}_{cvg'}$ we also have $\mathrm{lc}_{g \circ \Theta u \circ \Theta} = \mathrm{lc}_{cv \circ \Theta g' \circ \Theta}$. (The compositions of $gu$ and $vg'$ involves, since $\hat{g}u = v\hat{g}'$, exactly the same $\theta_i$ at the leading words.) It now follows that $\mathrm{lc}_{u \circ \Theta} g \circ \Theta\, u \circ \widehat{\Theta}$ and $c\, \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta} g' \circ \Theta$ have equal leading coefficients ($u \circ \Theta$ and $v \circ \widehat{\Theta}$ are words). Dividing by $\mathrm{lc}_{u \circ \Theta}$ we get the constant $c_1$.

The overlaps $\widehat{g \circ \Theta} = u' \widehat{g' \circ \Theta}\, v'$ are treated by the same principle.

**Proposition 1.** *If composition by $\Theta$ is compatible with the ordering and $\widehat{\Theta}$ is combinatorially free, then composition by $\Theta$ commutes with Gröbner bases computation.*

*Proof.* For an arbitrary Gröbner basis $G$, we need to show that also $G \circ \Theta$ is a Gröbner basis. We will use Theorem 1, so let $(g \circ \Theta, g' \circ \Theta)$ be a critical pair of $G \circ \Theta$.

Assume first that this critical pair has overlaps of the first type in Remark 1. By Lemma 2, each overlap relation is of the form $g \circ \Theta\, u \circ \widehat{\Theta} - c_1 v \circ \widehat{\Theta} g' \circ \Theta$, where $gu - cvg'$ is an overlap relation of the critical pair $(g, g')$ of $G$, and $c_1$ is as above. Since $G$ is a Gröbner basis, we get from (2) in Theorem 1, composing by $\Theta$ and using (4), (5), (6) and (7),

$$\begin{aligned} & g \circ \Theta\, u \circ \Theta - cv \circ \Theta\, g' \circ \Theta \\ &\quad = \sum_k c_k w_{k_L} \circ \Theta\, g_k \circ \Theta\, w_{k_R} \circ \Theta, w_{k_L} \circ \widehat{\Theta} \widehat{g_k \circ \Theta}\, w_{k_R} \circ \widehat{\Theta} < \widehat{g \circ \Theta}\, u \circ \widehat{\Theta} \\ &\quad = v \circ \widehat{\Theta} \widehat{g' \circ \Theta} \quad \forall k. \end{aligned}$$

Rewriting, we see that this can be written

$$\begin{aligned} & \mathrm{lc}_{u \circ \Theta} g \circ \Theta\, u \circ \widehat{\Theta} - c\, \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta} g' \circ \Theta \\ &\quad = \sum c_k w_{k_L} \circ \Theta\, g_k \circ \Theta\, w_{k_R} \circ \Theta - g \circ \Theta\, (u \circ \Theta - \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta}) \\ &\qquad + c(v \circ \Theta - \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta}) g' \circ \Theta. \end{aligned}$$

It is clear that both $g \circ \Theta\, (u \circ \Theta - \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta})$ and $(v \circ \Theta - \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta})$ $g' \circ \Theta$ have leading words smaller than $\widehat{g \circ \Theta}\, u \circ \widehat{\Theta} = v \circ \widehat{\Theta} \widehat{g' \circ \Theta}$. (The leading words cancel in the parentheses.) By expanding $(u \circ \Theta - \mathrm{lc}_{u \circ \Theta}$ $u \circ \widehat{\Theta}), (v \circ \Theta - \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta})$ and each $w_{k_L} \circ \Theta, w_{k_R} \circ \Theta$ to words, and dividing by $\mathrm{lc}_{u \circ \Theta}$, we then have a representation of our overlap relation $g \circ \Theta\, u \circ \widehat{\Theta} - c_1 v \circ \widehat{\Theta} g' \circ \Theta$ as in Theorem 1.

If our critical pair has overlaps of the second type in Remark 1, then each overlap relation is of the form $g \circ \Theta - c_2 u \circ \widehat{\Theta} g' \circ \Theta\, v \circ \widehat{\Theta}$ with $c_2$ as above, corresponding to the overlap relation $g - cug'v$ (again by Lemma 2). In the same way as above we get, now using (3) in Theorem 1,

$$
\begin{aligned}
& g \circ \Theta - cu \circ \Theta\, g' \circ \Theta\, v \circ \Theta \\
& \quad = \sum_{k=1}^{t} c_k w_{k_L} \circ \Theta\, g_k \circ \Theta\, w_{k_R} \circ \Theta\,, w_{k_L} \circ \widehat{\Theta} \widehat{g_k \circ \Theta}\, w_{k_R} \circ \widehat{\Theta} < \widehat{g \circ \Theta} \\
& \quad = u \circ \widehat{\Theta} \widehat{g' \circ \Theta}\, v \circ \widehat{\Theta} \quad \forall k.
\end{aligned}
$$

We then use the rewriting

$$
\begin{aligned}
& u \circ \Theta\, g' \circ \Theta\, v \circ \Theta \\
& \quad = \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta} g' \circ \Theta\, v \circ \Theta + (u \circ \Theta - \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta}) g' \circ \Theta\, v \circ \Theta \\
& \quad = \mathrm{lc}_{u \circ \Theta} \mathrm{lc}_{v \circ \Theta} u \circ \widehat{\Theta} g' \circ \Theta\, v \circ \widehat{\Theta} + (u \circ \Theta - \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta}) g' \circ \Theta\, v \circ \Theta \\
& \qquad + \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta} g' \circ \Theta\, (v \circ \Theta - \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta})
\end{aligned}
$$

to obtain

$$
\begin{aligned}
& g \circ \Theta - c\, \mathrm{lc}_{u \circ \Theta} \mathrm{lc}_{v \circ \Theta} u \circ \widehat{\Theta} g' \circ \Theta\, v \circ \widehat{\Theta} \\
& \quad = \sum c_k w_{k_L} \circ \Theta\, g_k \circ \Theta\, w_{k_R} \circ \Theta - c(u \circ \Theta - \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta}) g' \circ \Theta\, v \circ \Theta \\
& \qquad - c\, \mathrm{lc}_{u \circ \Theta} u \circ \widehat{\Theta} g' \circ \Theta\, (v \circ \Theta - \mathrm{lc}_{v \circ \Theta} v \circ \widehat{\Theta}).
\end{aligned}
$$

In the same way as before, we see that this is a representation as (3) in Theorem 1, and the proposition is proved.

### 3.2 Proof of Necessity

The strategy we have to use in this section is clear: Assuming one of the two conditions in Theorem 3 not true, we need to find a suitable Gröbner basis, contradicting the commutation of Gröbner bases computation.

To prove the necessity of condition 1 in Theorem 3, we can use the same counterexample as in the commutative case.

**Lemma 3.** *If composition by $\Theta$ commutes with Gröbner bases computation, then composition by $\Theta$ is compatible with the ordering.*

*Proof.* If $u, v \in W$ are two words with $u > v$, then we have to show that $u \circ \widehat{\Theta} > v \circ \widehat{\Theta}$. Let $c \in K$ be such that $\mathrm{lc}_{u \circ \Theta} = c\, \mathrm{lc}_{v \circ \Theta}$. By Lemma 1-ix, $G = \{u - cv, v\}$ is a Gröbner basis, so $G \circ \Theta = \{u \circ \Theta - cv \circ \Theta\,, v \circ \Theta\}$ must also be a Gröbner basis.

Assume first that $u \circ \widehat{\Theta} < v \circ \widehat{\Theta}$. We then have $\widehat{G \circ \Theta} = \{v \circ \widehat{\Theta}\}$, and $v \circ \widehat{\Theta} \nmid u \circ \widehat{\Theta}$ (by (1)). Since

$$u \circ \Theta = (u \circ \Theta - cv \circ \Theta) + cv \circ \Theta \in \langle G \circ \Theta \rangle,$$

$G \circ \Theta$ is not a Gröbner basis, so composition by $\Theta$ does not commute with Gröbner bases computation.

The case remaining to exclude is when $u \circ \widehat{\Theta} = v \circ \widehat{\Theta}$. If $u \circ \Theta = cv \circ \Theta$, then $0 = u \circ \Theta - cv \circ \Theta \in G \circ \Theta$, so $G \circ \Theta$ is not a Gröbner basis. Otherwise we have $f = u \circ \Theta - cv \circ \Theta \neq 0$ and $\hat{f} < u \circ \widehat{\Theta}$ (by our choice of $c$). Since $f \in \langle G \circ \Theta \rangle$ and $\widehat{G \circ \Theta} = \{u \circ \widehat{\Theta}\}$, $G \circ \Theta$ is not a Gröbner basis (again using (1)).

Also in the following proof we use a technique inspired by Hong. A similar argument is sufficient to show the necessity of the counterpart of condition 2 in the commutative case.

**Lemma 4.** *If composition by $\Theta$ commutes with Gröbner bases computation, then $\widehat{\Theta}$ does not contain any overlaps of the second type in Remark 1.*

*Proof.* We begin by noting that $\hat{\theta}_i \neq \mathbf{1}$ for all $\theta_i \in \Theta$. This follows from Lemma 3, since $x_i > \mathbf{1}$ implies $\hat{\theta}_i = x_i \circ \widehat{\Theta} > \mathbf{1} \circ \widehat{\Theta} = \mathbf{1}$.

Now assume that we have an overlap $\hat{\theta}_i = u\hat{\theta}_j v$ for some $\theta_i, \theta_j \in \Theta$ and $u, v \in W$. We know that both $\{\theta_i, \theta_j\}$ and $\{\theta_i + \mathbf{1}, \theta_j\}$ must be Gröbner bases (Lemma 1-v, vi). Applying (3) of Theorem 1 on $\{\theta_i + \mathbf{1}, \theta_j\}$, we can write

$$(\theta_i + \mathbf{1}) - cu\theta_j v = \sum_k c_k w_{k_L} \theta_j w_{k_R},$$

where the sum in the right hand side might be empty (*i.e.* equal to zero). We have no terms involving $\theta_i + \mathbf{1}$ in this sum since this clearly would contradict the inequality in (3). Rewriting, we get

$$\mathbf{1} = \sum c_k w_{k_L} \theta_j w_{k_R} - \theta_i + cu\theta_j v \in \langle \theta_i, \theta_j \rangle.$$

Since also $\{\theta_i, \theta_j\}$ is a Gröbner basis we then have $\hat{\theta}_i \mid \mathbf{1}$ or $\hat{\theta}_j \mid \mathbf{1}$, which is absurd. We conclude that $\{\theta_i, \theta_j\}$ and $\{\theta_i + \mathbf{1}, \theta_j\}$ can not both be Gröbner bases, a contradiction.

To handle the overlaps of the first type in Remark 1, we need some combinatorics. From [2] we borrow the following definition and proposition.

**Definition 7.** *For every word $u \in W$, we define the* root *of $u$ as the (unique) shortest word of which $u$ is a power.*

So for example the root of $u = x_1x_2x_1x_2x_1x_2$ is $x_1x_2$, but for $u = x_1x_2^2x_1x_2$ it is $u$ itself.

**Proposition 2.** *For two non-empty words $u, v \in W$, $uv = vu$ if and only if $u$ and $v$ have the same root.*

Since the following situation will appear several times, we give it an own definition.

**Definition 8.** *We say that $u \in W$ is a* non-trivial subword *of $u^2$ if $u$ can be placed in $u^2$ in a position different from the beginning and the ending of $u^2$, i.e. $u$ is not merely a left and right subword of $u^2$.*

**Lemma 5.** *Let $u \in W$ be a non-empty word (i.e. $\neq \mathbf{1}$). If $u$ is a non-trivial subword of $u^2$, then $u = w^k$, $k \geq 2$, $w$ the root of $u$. (I.e. $u$ is not the root of itself.)*

*Proof.* We have $u^2 = v_1uv_2$ for some $v_1, v_2 \in W, v_1, v_2 \neq \mathbf{1}$. It is easy to see that $u = v_1w_1 = w_2v_2$, $w_1, w_2 \in W, w_1, w_2 \neq \mathbf{1}$, where $u = w_1w_2$ is a decomposition of the $u$ in $v_1uv_2$. Since $uw_2 = v_1u$, we have $|v_1| = |w_2|$. But since also $u = v_1w_1 = w_2v_2$, we must then have $v_1 = w_2$. We conclude that $u$ commutes with $v_1 = w_2$. Since $v_1$ clearly is shorter than $u$, it follows from Proposition 2 that $u$ can not be the root of itself.

**Lemma 6.** *Let $u, v_1, v_2 \in W$ be three words with $v_1, v_2 < u$. If $u^4$ is a subword of $u^3v_1uv_2u^3$, then $v_1u = uv$ or $uv_2 = v'u$ (or both) for some $v$, $v' \in W$.*

*Proof.* Let $u = w^k$, $w$ the root of $u$. Then our task is to place $w_1 = w^{4k}$ in $w_2 = w^{3k}v_1w^kv_2w^{3k}$. We know from the previous lemma that $w$ is not a non-trivial subword of $w^2$, for then $w$ can not be a root. This means that we must try to "fit the $w$ of $w_1$ exactly over the $w$ of $w_2$".

If we start to place the $w$ of $w_1$ over some of the first $w$ in $w_2$ (*i.e.* to the left of $v_1$), then we have at least $k$ copies of $w$ from $w_1$ left when we reach $v_1u$ in $w_2$. It then clearly follows that $v_1u = uv$ for some $v$. Using the same argument in the other direction, *i.e.* trying to place the end of $w_1$ over some of the $w$ to the right of $v_2$, then we obtain $uv_2 = v'u$.

But $w_1$ is sufficiently long for us to be able to conclude that we must have at least one of the situations above. (Since $v_1, v_2 < u = w^k$, it follows from (1) that $v_1, v_2$ can not contain more that $k-1$ copies of $w$.) To illustrate, we try to place $w_1$ in the "middle" of $w_2$. We must place some $w^k$ of $w_1$ exactly over the middle $u = w^k$ of $w_1$ (otherwise we have a non-trivial subword). Then we have $3k$ copies of $w$ left, so on one side of the middle $u$ we

have $\geq 3k/2$ copies (certainly $> 3k/2$ if $k$ is odd), let say to the left. Since $v_1 < u = w^k$, we then have $\geq k/2$ copies of $w$ to place to the left of $v_1$. Not allowed to have any non-trivial subword, we must then have the case further up in this proof.

We will need one more purely combinatorial result. The following proposition is the main result in [7]. (This paper is devoted to free groups, but the result is valid also in our free monoid $W$.)

**Proposition 3.** *Let $u, v, w \in W$ be three words satisfying the equality $u^k v^l = w^m$ with $k, l, m \geq 2$. Then u, v and w all have the same root.*

We now return to our Gröbner bases.

**Lemma 7.** *Let p and q be two polynomials with $p - q \neq 0$ and $\widehat{p - q} < \hat{\theta}_i$ for some $\theta_i \in \Theta$. If composition by $\Theta$ commutes with Gröbner bases computation, then $\theta_i p - q\theta_i \neq 0$.*

*Proof.* We will use that $\{\theta_i + \mathbf{1}\}$ must be a Gröbner basis (Lemma 1-ii). If $\theta_i p - q\theta_i = 0$, then

$$0 \neq p - q = (\theta_i + \mathbf{1})p - q(\theta_i + \mathbf{1}) \in \langle \theta_i + \mathbf{1} \rangle.$$

But since $\widehat{p - q} < \hat{\theta}_i$, we can not have $\widehat{\theta_i + \mathbf{1}} = \hat{\theta}_i \mid \widehat{p - q}$, a contradiction.

In the proof of the following lemma, we will use a technique that will be applicable also in the proofs of Lemma 10 and Lemma 11 below.

**Lemma 8.** *Assume that composition by $\Theta$ commutes with Gröbner bases computation. If some $\hat{\theta}_i$ ($\theta_i \in \Theta$) form an overlap with itself, then $\hat{\theta}_i = w^k$, $k \geq 2$, w the root of $\hat{\theta}_i$.*

*Proof.* Let the overlap be $\hat{\theta}_i u = v\hat{\theta}_i$. Using Proposition 2, we see that we are done if $u = v$ (since $|u| = |v| < |\hat{\theta}_i|$), so assume $u \neq v$. Assume further that we have chosen an overlap (if there are more than one possibility) with $|u| = |v|$ minimal. Since $\{\theta_i\}$ must be a Gröbner basis, we have by Theorem 1,

$$\theta_i u - v\theta_i = \sum_{k=1}^{t} c_k w_{k_L} \theta_i w_{k_R}, \quad w_{k_L} \hat{\theta}_i w_{k_R} < \hat{\theta}_i u = v\hat{\theta}_i \quad \forall k. \tag{10}$$

The inequality clearly implies $w_{k_L} < v\ (< \hat{\theta}_i)$ and $w_{k_R} < u\ (< \hat{\theta}_i)$ for all $k$.

The conditions in Lemma 7 are fulfilled with $p = u, q = v$, so we conclude that $\theta_i u - v\theta_i \neq 0$, *i.e.* the sum in the right hand side of (10) is non-empty. Recall that we can assume that this sum is written with its terms in decreasing order, so the leading word is $w_{1_L} \hat{\theta}_{j_1} w_{1_R}$.

If $w_{1_L}$ or $w_{1_R}$ is equal to **1**, let say $w_{1_L}$ (the other case being symmetrical), then we can write

$$\theta_i(u - c_1 w_{1_R}) - v\theta_i = \sum_{k=2}^{t} c_k w_{k_L} \theta_i w_{k_R}.$$

The new sum in the right hand side has leading word smaller than the original sum in (10). Continuing in the same way (first considering $w_{2_L}\hat{\theta}_i w_{2_R}$), we end up with

$$\theta_i(u - p) - (v - q)\theta_i = \sum_{k=m}^{t} c_k w_{k_L} \theta_i w_{k_R}, \tag{11}$$

where for the leading word $w_{m_L}\hat{\theta}_i w_{m_R}$ of the sum, $w_{m_L}, w_{m_R} \neq \mathbf{1}$. Here every word in the polynomial $p$ is smaller than $u$, and every word in $q$ is smaller than $v$, so we can again deduce from Lemma 7 that the sum in (11) is non-empty.

We now multiply equation (11) by $\theta_i^3$ from left and right to obtain

$$\theta_i^4(u - p)\theta_i^3 - \theta_i^3(v - q)\theta_i^4 = \sum_{k=m}^{t} c_k \theta_i^3 w_{k_L} \theta_i w_{k_R} \theta_i^3.$$

It follows that

$$\Sigma = \sum_{k=m}^{t} c_k \theta_i^3 w_{k_L} \theta_i w_{k_R} \theta_i^3$$

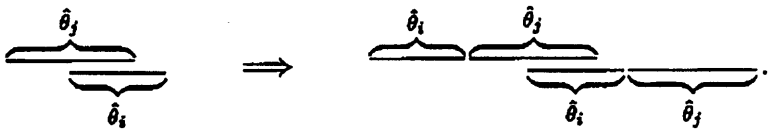
is in the ideal $\langle \theta_i^4 \rangle$. We then know, since $\{\theta_i^4\}$ must be a Gröbner basis, that $\hat{\theta}_i^4$ is a subword of $\hat{\Sigma} = \hat{\theta}_i^3 w_{m_L} \hat{\theta}_i w_{m_R} \hat{\theta}_i^3$. We can now use Lemma 6 (with $u = \hat{\theta}_i, v_1 = w_{m_L}, v_2 = w_{m_R}$) to conclude that $w_{m_L}\hat{\theta}_i = \hat{\theta}_i w$ or $\hat{\theta}_i w_{m_R} = w'\hat{\theta}_i$ for some $w, w' \in W$.

Assume first that $w_{m_L}\hat{\theta}_i = \hat{\theta}_i w$. Since both $w_{m_L}$ and $v$ then are left subwords of $\hat{\theta}_i$, and $w_{m_L} < v$, we have $|w_{m_L}| < |v|$. But since $w_{m_L} \neq \mathbf{1}$, $w_{m_L}\hat{\theta}_i = \hat{\theta}_i w$ is an overlap contradicting the minimality of $|u| = |v|$. The same argument applies if $\hat{\theta}_i w_{m_R} = w'\hat{\theta}_i$, now using that $w_{m_R} \neq \mathbf{1}$ is shorter than $u$. Thus the lemma is proved.

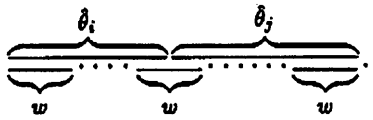
We now have enough tools to exclude more overlaps.

**Lemma 9.** *If composition by* $\Theta$ *commutes with Gröbner bases computation, then no leading words of two different polynomials from* $\Theta$ *form an overlap of the first type in Remark 1.*

*Proof.* Assume that we have an overlap of this type, *i.e.* $\hat{\theta}_j u = v\hat{\theta}_i$ for some $\theta_i, \theta_j \in \Theta$ and $u, v \in W$. Then $\hat{\theta}_i\hat{\theta}_j$ forms an overlap with itself:



In the proof of the previous lemma, we used the Gröbner bases i and ii of Lemma 1 (the latter indirectly by Lemma 7). If we instead use iii and iv, then we see that the proof of the previous lemma goes through if we replace $\theta_i$ with $\theta_i\theta_j$. We conclude that $\hat{\theta}_i\hat{\theta}_j = w^k$, $k \geq 2$. If $\hat{\theta}_i = w^{k_1}$ and $\hat{\theta}_j = w^{k_2}$, then one of them is a subword of the other, contradicting Lemma 4. The same must be true if $\hat{\theta}_i$ or $\hat{\theta}_j$ is a subword of $w$. Thus $\hat{\theta}_i$ and $\hat{\theta}_j$ meet "in the middle" of a $w$ in the sense of the following figure:



We see that a (proper) right subword of $\hat{\theta}_i$ is a left subword of $w$, and then of course also a left subword of $\hat{\theta}_i$ itself. In the same way we see that a left subword of $\hat{\theta}_j$ is a right subword of $\hat{\theta}_j$. We conclude that $\hat{\theta}_i$ and $\hat{\theta}_j$ both must form overlaps with themselves. We can then use Lemma 8 again to get $\hat{\theta}_i = w_1^l$, $l \geq 2$, and $\hat{\theta}_j = w_2^m$, $m \geq 2$.

Summarizing, we get the equality $\hat{\theta}_i\hat{\theta}_j = w_1^l w_2^m = w^k$ with $k, l, m \geq 2$. But then $w_1$ and $w_2$ (and $w$) have the same root by Proposition 3, so $\hat{\theta}_i$ and $\hat{\theta}_j$ must also have the same root. It follows that one of $\hat{\theta}_i$ or $\hat{\theta}_j$ must be a subword of the other (or both if $\hat{\theta}_i = \hat{\theta}_j$), *i.e.* $\hat{\theta}_i$ and $\hat{\theta}_j$ form an overlap of the second type in Definition 2. We can consequently apply Lemma 4 to get a contradiction, and our lemma is proved.

We have now proved that two different words from $\widehat{\Theta}$ can not form an overlap (of either type) if composition by $\Theta$ commutes with Gröbner bases computation. The only thing we have left to show, to be able to conclude that $\widehat{\Theta}$ must be combinatorially free, is that no word from $\widehat{\Theta}$ forms an overlap with itself. We have proved in Lemma 8 that such a word must be of the form $w^l$ with $l \geq 2$. We will also need the following.

**Lemma 10.** *Assume that composition by $\Theta$ commutes with Gröbner bases computation. If some $\hat{\theta}_i$ ($\theta_i \in \Theta$) forms an overlap with itself, then the root of $\hat{\theta}_i$ can not form an overlap with itself.*

*Proof.* We know from Lemma 8 that $\hat{\theta}_i = u^l, l \geq 2$, where $u$ is the root of $\hat{\theta}_i$. Assume that $u$ forms an overlap with itself, *i.e.* $uw_1 = w_2u$ for some subwords $w_1, w_2$ of $u$. Assume further that we have chosen an overlap with

$|w_1| = |w_2|$ minimal. We must have $w_1 \neq w_2$. Otherwise $u$ commutes with $w_1 = w_2$, and by Proposition 2, $u$ can not be a root.

Multiplying the overlap above by $u^{l-1}$ from both left and right, we obtain $\hat{\theta}_i w_1 u^{l-1} = u^{l-1} w_2 \hat{\theta}_i$, which is an overlap of $\hat{\theta}_i$. We then have, since $\{\theta_i\}$ must be a Gröbner basis,

$$\theta_i w_1 u^{l-1} - u^{l-1} w_2 \theta_i = \sum_{k=1}^{t} c_k w_{k_L} \theta_i w_{k_R}, \quad w_{k_L} \hat{\theta}_i w_{k_R} < \hat{\theta}_i w_1 u^{l-1}$$

$$= u^{l-1} w_2 \hat{\theta}_i. \tag{12}$$

The inequality implies $w_{k_L} < u^{l-1} w_2$ and $w_{k_R} < w_1 u^{l-1}$ for all $k$.

We note that $w_1 u^{l-1} \neq u^{l-1} w_2$. Otherwise, since $u w_1 = w_2 u$, both $w_1$ and $w_2$ are left subwords of $u$, *i.e.* $w_1 = w_2$ (since $|w_1| = |w_2|$). We conclude, using Lemma 7, that the sum in (12) is non-empty ($w_1 u^{l-1}, u^{l-1} w_2 < \hat{\theta}_i$).

We will later need that some $w_{k_L}, w_{k_R}$ are not powers of $u$. So assume (the sum written in decreasing order) that *e.g.* $w_{1_L} = u^s$, the case $w_{1_R}$ being symmetrical. If $s = 0$ (*i.e.* $w_{1_L} = \mathbf{1}$), then we just move $c_1 \theta_i w_{1_R}$ to the left hand side of (12). Since $w_{1_L} < u^{l-1} w_2$, we can not have $s \geq l$, and for $1 \leq s \leq l-1$, $\theta_i u^s - u^s \theta_i$ is an overlap relation. This means (since $\{\theta_i\}$ is Gröbner basis) that we can replace $u^s \theta_i$ with $\theta_i u^s$ and a sum with terms $\tilde{c}_j v_{j_L} \theta_i v_{j_R}$ smaller than $u^s \hat{\theta}_i = w_{1_L} \hat{\theta}_i$:

$$w_{1_L} \theta_i w_{1_R} = u^s \theta_i w_{1_R} = \theta_i u^s w_{1_R} + \left( \sum \tilde{c}_j v_{j_L} \theta_i v_{j_R} \right) w_{1_R},$$
$$v_{j_L} \hat{\theta}_i v_{j_R} w_{1_R} < w_{1_L} \hat{\theta}_i w_{1_R}.$$

Moving $c_1 \theta_i u^s w_{1_R}$ to the left hand side of (12), we have in the right hand side of (12) a sum with terms smaller than $w_{1_L} \hat{\theta}_i w_{1_R}$. Continuing in the same way with this sum, we end up with

$$\theta_i (w_1 u^{l-1} - p) - (u^{l-1} w_2 - q) \theta_i = \sum_{k=m}^{t'} c'_k w'_{k_L} \theta_i w'_{k_R}, \tag{13}$$

where, assuming that $w'_{m_L} \hat{\theta}_i w'_{m_R}$ is the leading word of the sum, $w'_{m_L}$, $w'_{m_R} \neq u^s$. (It may happen, during the process just described, that we get terms in the sum with equal leading words. However, we will still end up with (13), but should maybe say that $w'_{m_L} \hat{\theta}_i w'_{m_R}$ is *one* representation of the leading word.) We must also have $w'_{m_L} < u^{l-1} w_2$ ($< \hat{\theta}_i$) and $w'_{m_R} < w_1 u^{l-1}$ ($< \hat{\theta}_i$). Since the leading word of $(w_1 u^{l-1} - p) - (u^{l-1} w_2 - q)$ is still $w_1 u^{l-1}$ or $u^{l-1} w_2$, the sum in the right hand side of (13) is non-empty by Lemma 7.

Multiplying by $\theta_i^3$ from left and right, we obtain

$$\theta_i^4(w_1u^{l-1}-p)\theta_i^3-\theta_i^3(u^{l-1}w_2-q)\theta_i^4=\sum_{k=m}^{t'}c'_k\theta_i^3w'_{k_L}\theta_iw'_{k_R}\theta_i^3.$$

We see that the sum is in $\langle\theta_i^4\rangle$, so $\hat{\theta}_i^4$ must be a subword of $\hat{\theta}_i^3w'_{m_L}\hat{\theta}_iw'_{m_R}\hat{\theta}_i^3$, and we can then use Lemma 6 to conclude that $w'_{m_L}\hat{\theta}_i=\hat{\theta}_iw$ or $\hat{\theta}_iw'_{m_R}=w'\hat{\theta}_i$.

Assume first that $w'_{m_L}\hat{\theta}_i=\hat{\theta}_iw$, *i.e.* $w'_{m_L}u^l=u^lw$. We know that $w'_{m_L}<u^{l-1}w_2$ and $w_2<u$. Since $w'_{m_L}\neq u^{l-1}$, we have $|w'_{m_L}|\neq|u^{l-1}|$. If $w'_{m_L}$ is shorter that $u^{l-1}$, then the first $u$ after $w'_{m_L}$ in $w'_{m_L}u^l$ must be a non-trivial subword of some $u^2$ in $u^lw$ ($w'_{m_L}\neq u^s$). But this is impossible, since $u$ then can not be a root by Lemma 5. On the other hand, if $|w_{m_L}|>|u^{l-1}|$, then we must have $w_{m_L}=u^{l-1}\tilde{w}$, $\tilde{w}\neq\mathbf{1}$, $|\tilde{w}|<|w_2|$ (since $\tilde{w},w_2$ both are left subwords of $u$ and $w'_{m_L}=u^{l-1}\tilde{w}<u^{l-1}w_2$). Cancelling the first $u^{l-1}$ in $w'_{m_L}u^l=u^lw$, we then get $\tilde{w}u^l=uw$. But here we have an overlap $\tilde{w}u=u\bar{w}$ contradicting the minimality of $|w_1|=|w_2|$.

The same argument applies if $\hat{\theta}_iw'_{m_R}=w'\hat{\theta}_i$, now using that $w_{m_R}<w_1u^{l-1}$ and $w_1<u$. Thus the lemma is proved.

We can now show our last lemma, and thereby finishing the proof of Theorem 3. The proof will be similar to the proofs of Lemma 8 and Lemma 10. We can not apply Lemma 7 directly, but the reader will recognize the argument.

**Lemma 11.** *If composition by $\Theta$ commutes with Gröbner bases computation, then no $\hat{\theta}_i$ ($\theta_i\in\Theta$) forms an overlap with itself.*

*Proof.* Assume on the contrary that there is such a $\theta_i\in\Theta$. Since $n\geq 2$, there is at least one more polynomial $\theta_j\in\Theta$, and $\hat{\theta}_i,\hat{\theta}_j$ must not form an overlap.

We know from Lemma 8 that $\hat{\theta}_i=u^l$, $l\geq 2$, $u$ the root of $\hat{\theta}_i$. We then have an overlap $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iu\hat{\theta}_j\hat{\theta}_i=\hat{\theta}_i\hat{\theta}_ju\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$ (formed by $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$), and since $\{\theta_i\theta_j\theta_i\}$ is a Gröbner basis, an overlap relation that can be represented as

$$\theta_i\theta_j\theta_iu\hat{\theta}_j\hat{\theta}_i-\hat{\theta}_i\hat{\theta}_ju\theta_i\theta_j\theta_i=\sum_{k=1}^{t}c_kw_{k_L}\theta_i\theta_j\theta_iw_{k_R},\tag{14}$$

with an inequality that implies $w_{k_L}<\hat{\theta}_i\hat{\theta}_ju$ and $w_{k_R}<u\hat{\theta}_j\hat{\theta}_i$ for all $k$.

To show that the sum in the right hand side of (14) is non-empty, we will use the Gröbner basis

$$G=\{x_ix_jx_i+x_j,x_ix_jx_j-x_jx_jx_i\}$$

(Lemma 1-viii). Assuming the sum empty, we have

$$(\theta_i\theta_j\theta_i+\theta_j)u\hat{\theta}_j\hat{\theta}_i-\hat{\theta}_i\hat{\theta}_ju(\theta_i\theta_j\theta_i+\theta_j)=\theta_ju\hat{\theta}_j\hat{\theta}_i-\hat{\theta}_i\hat{\theta}_ju\theta_j\in\langle G\circ\Theta\rangle.$$

The leading words of the two terms in the right hand side can not be equal, for then $\hat{\theta}_i$ and $\hat{\theta}_j$ would form an overlap, which is impossible by previous lemmas. Thus some word of $\widehat{G\circ\Theta}$ must be a subword of the largest of $\hat{\theta}_ju\hat{\theta}_j\hat{\theta}_i, \hat{\theta}_i\hat{\theta}_ju\hat{\theta}_j$. But it is not hard to see that neither of $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i, \hat{\theta}_i\hat{\theta}_j\hat{\theta}_j, \hat{\theta}_j\hat{\theta}_j\hat{\theta}_i$ can be a subword of $\hat{\theta}_ju\hat{\theta}_j\hat{\theta}_i$ or $\hat{\theta}_i\hat{\theta}_ju\hat{\theta}_j$, without $\hat{\theta}_i, \hat{\theta}_j$ forming an overlap (recall that $\hat{\theta}_i=u^l$).

We will later need that $w_{k_L}\neq\hat{\theta}_i\hat{\theta}_j, w_{k_R}\neq\hat{\theta}_j\hat{\theta}_i$ and $w_{k_L}, w_{k_R}\neq\mathbf{1}$. Assume that the leading word of the sum in (14) is $w_{1_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw_{1_L}$. If $w_{1_L}$ or $w_{1_R}$ is equal to $\mathbf{1}$, then we can do as in previous proofs, *i.e.* just move $c_1\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw_{1_L}$ or $c_1w_{1_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$ to the left hand side of (14). So assume that $w_{1_L}=\hat{\theta}_i\hat{\theta}_j$, the case $w_{1_R}=\hat{\theta}_j\hat{\theta}_i$ being symmetrical. We then use the overlap relation $\theta_i\theta_j\theta_i\hat{\theta}_j\hat{\theta}_i-\hat{\theta}_i\hat{\theta}_j\theta_i\theta_j\theta_i$ to, as in the previous proof, replace $\hat{\theta}_i\hat{\theta}_j\theta_i\theta_j\theta_i$ with $\theta_i\theta_j\theta_i\hat{\theta}_j\hat{\theta}_i$ and a "smaller sum". Moving $c_1\theta_i\theta_j\theta_i\hat{\theta}_j\hat{\theta}_iw_{1_R}$ to the left hand side of (14), and continuing with the smaller terms in the new sum, we end up with

$$\theta_i\theta_j\theta_i(u\hat{\theta}_j\hat{\theta}_i-p)-(\hat{\theta}_i\hat{\theta}_ju-q)\theta_i\theta_j\theta_i=\sum_{k=m}^{t'}c'_kw'_{k_L}\theta_i\theta_j\theta_iw'_{k_R},$$

where, assuming that $w'_{m_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw'_{m_R}$ is the leading word (or one representation of the leading word), $w'_{m_L}\neq\hat{\theta}_i\hat{\theta}_j, w'_{m_R}\neq\hat{\theta}_j\hat{\theta}_i$ and $w'_{m_L}, w'_{m_R}\neq\mathbf{1}$. We also have $w'_{m_L}<\hat{\theta}_i\hat{\theta}_ju\ (<\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i), w'_{m_R}<u\hat{\theta}_j\hat{\theta}_i\ (<\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i)$. The sum can not be empty, for then the counterexample with $G$ above applies.

Multiplying by $(\theta_i\theta_j\theta_i)^3$ from left and right, we obtain

$$\begin{aligned}&(\theta_i\theta_j\theta_i)^4(u\hat{\theta}_j\hat{\theta}_i-p)(\theta_i\theta_j\theta_i)^3-(\theta_i\theta_j\theta_i)^3(\hat{\theta}_i\hat{\theta}_ju-q)(\theta_i\theta_j\theta_i)^4\\&\quad=\sum_{k=m}^{t'}c'_k(\theta_i\theta_j\theta_i)^3w'_{k_L}\theta_i\theta_j\theta_iw'_{k_R}(\theta_i\theta_j\theta_i)^3.\end{aligned}$$

We conclude, since $\{(\theta_i\theta_j\theta_i)^4\}$ is a Gröbner basis, that $(\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i)^4$ must be a subword of $(\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i)^3w'_{m_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw'_{m_R}(\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i)^3$. Applying Lemma 6 (with $u=\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$), we see that $w'_{m_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i=\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw$ or $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw'_{m_R}=w'\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$.

Assume first that $w'_{m_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i=\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw$. Since $w'_{m_L}\neq\hat{\theta}_i\hat{\theta}_j$, we have $|w'_{m_L}|\neq|\hat{\theta}_i\hat{\theta}_j|$. If $|w'_{m_L}|<|\hat{\theta}_i\hat{\theta}_j|$, then it is easy to see that $\hat{\theta}_i$ and $\hat{\theta}_j$ must form an overlap ($w'_{m_L}\neq\mathbf{1}$). If $|w'_{m_L}|>|\hat{\theta}_i\hat{\theta}_j|$, then $w'_{m_L}=\hat{\theta}_i\hat{\theta}_j\tilde{w}$ with $|\tilde{w}|<|u|$ ($w'_{m_L}<\hat{\theta}_i\hat{\theta}_ju$). It then follows that the first $u$ of the last $\hat{\theta}_i=u^l$ in $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_iw$ forms an overlap with the first $u$ of the $\hat{\theta}_i$ after $w'_{m_L}$ in $w'_{m_L}\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$. But this is impossible by the previous lemma.

The case when $\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i w_{m_R} = w'\hat{\theta}_i\hat{\theta}_j\hat{\theta}_i$ is symmetrical, now using that $w'_{m_R} \neq \hat{\theta}_j\hat{\theta}_i$ and $w'_{m_R} < u\hat{\theta}_j\hat{\theta}_i$. Thus this lemma, and also Theorem 3, is proved.

## 4 COMPATIBLE COMPOSITIONS

Given a finite set $\Theta \subset K\langle X\rangle$, it is of course not hard to decide whether $\widehat{\Theta}$ is combinatorially free. To check the other condition in our main theorem, *i.e.* if composition by $\Theta$ is compatible with the current ordering, is however not a trivial task. In Hong's paper, the following question is left as an open problem.

*Does there exist a decision procedure that will determine whether a given composition is compatible with a given term ordering?*

We will answer this question for two examples of non-commutative admissible orderings. We start with deglex.

**Proposition 4.** *Composition by $\Theta$ is compatible with the deglex ordering if and only if $|\hat{\theta}_i| = |\hat{\theta}_j|$ for all $\theta_i, \theta_j \in \Theta$, and $\hat{\theta}_i > \hat{\theta}_j$ if $x_i > x_j$.*

*Proof.* Assume first that composition by $\Theta$ is compatible with the deglex ordering, and for some $\hat{\theta}_i, \hat{\theta}_j$, their lengths are different. Let for example $x_i > x_j$ and $|\hat{\theta}_i| = k, |\hat{\theta}_j| = l$. We necessarily have $k > l$. Note that

$$k(l+1) = kl + k > l(k+1) = kl + l.$$

We have $x_j^{k+1} > x_i^{l+1}$. But the length of $x_j^{k+1} \circ \widehat{\Theta} = \hat{\theta}_j^{k+1}$ is $l(k+1)$, and then $x_j^{k+1} \circ \widehat{\Theta}$ is shorter than $x_i^{l+1} \circ \widehat{\Theta} = \hat{\theta}_i^{l+1}$, which has length $k(l+1)$. Thus $x_j^{k+1} \circ \widehat{\Theta} < x_i^{l+1} \circ \widehat{\Theta}$, a contradiction. That $x_i > x_j$ must imply $\hat{\theta}_i > \hat{\theta}_j$ is obvious.

Assume conversely that $|\hat{\theta}_i| = |\hat{\theta}_j|$ for all $\hat{\theta}_i, \hat{\theta}_j \in \Theta$, and $\hat{\theta}_i > \hat{\theta}_j$ if $x_i > x_j$. If $|u| > |v|$ $(u, v \in W)$, then clearly $|u \circ \widehat{\Theta}| > |v \circ \widehat{\Theta}|$. We thus have the case $|u| = |v|$ left, which implies $|u \circ \widehat{\Theta}| = |v \circ \widehat{\Theta}|$. Since $|\hat{\theta}_i| = |\hat{\theta}_j|$, $\hat{\theta}_i > \hat{\theta}_j$ is equivalent to that $\hat{\theta}_i$ is larger than $\hat{\theta}_j$ lexicographically. It is then easy to see that $u > v$ lexicographically implies $u \circ \widehat{\Theta} > v \circ \widehat{\Theta}$ lexicographically, and our proposition is proved.

In the commutative case, we have that a composition that allows commutation of Gröbner bases computation must contain all variables (see the discussion following Theorem 2). This is not necessary in our non-commutative setting. It follows from Proposition 5 that any subset $\Theta = \{\theta_1, \theta_2, \theta_3\}$ of $K\langle x_1, x_2, x_3\rangle$ with for example

$$\hat{\theta}_1 = x_1^3 x_2^3, \quad \hat{\theta}_2 = x_1^2 x_2 x_1 x_2^2, \quad \hat{\theta}_3 = x_1 x_2 x_1^2 x_2^2$$
$$(\Rightarrow \widehat{\Theta} \text{ combinatorially free})$$

fulfills the conditions in our main theorem if we use deglex ($x_3 > x_2 > x_1$).

We now consider the *elimination ordering* defined in [9]: For $u \in W$ and $1 \leq i \leq n$, let $\deg_{x_i} u$ denote the number of different occurrences of $x_i$ in $u$. We first use the commutative lexicographic ordering ($x_n > \cdots > x_1$), *i.e.* we let $u > v$ $(u, v \in W)$ if $\deg_{x_j} u > \deg_{x_j} v$ and $\deg_{x_i} u = \deg_{x_i} v$ for $i > j$. If $\deg_{x_i} u = \deg_{x_i} v$ for all $i$, then we use our non-commutative lexicographic ordering. This ordering is called an elimination ordering because it allows us to use elimination techniques.

**Proposition 5.** *Assume that $\Theta \subset K\langle x_1 \ldots x_n \rangle$ consists of n polynomials, and that $\widehat{\Theta}$ is combinatorially free. (This is clearly the case we are interested in.) Then composition by $\Theta$ is compatible with the elimination ordering above if and only if $\hat{\theta}_i = x_i$ for all i.*

*Proof.* Assume first that composition by $\Theta$ is compatible with the elimination ordering. Let $x_i > x_j$ (*i.e.* $i > j$), which implies $\hat{\theta}_i > \hat{\theta}_j$. Denote by $\max_x u$ the largest variable occurring in $u \in W$, *i.e.* the variable with highest index. Let $\max_x \hat{\theta}_i = x_{i'}$ and $\max_x \hat{\theta}_j = x_{j'}$; we clearly have $x_{i'} \geq x_{j'}$. If $x_{i'} = x_{j'}$, then let $\deg_{x_{i'}} \hat{\theta}_i = k$ and $\deg_{x_{i'}} \hat{\theta}_j = l$. If $m$ is an integer such that $ml > k$, then $\hat{\theta}_j^m > \hat{\theta}_i$. Since $x_i > x_j^m$ we get a contradiction, so $x_{i'} > x_{j'}$. It is then easy to see, since all $\hat{\theta}_j \neq \mathbf{1}$, that $\max_x \hat{\theta}_i = x_i$ for all $i$. (Since $\max_x \theta_i$ for the $n$ polynomials $\theta_i$ must form a decreasing sequence.)

Since $\max_x \hat{\theta}_1 = x_1$, we see that $\hat{\theta}_1$ contains only $x_1$, so we must have $\hat{\theta}_1 = x_1$ (otherwise $\hat{\theta}_1$ forms an overlap with itself). Moreover, $\hat{\theta}_2$ contains only $x_2$ and $x_1$, and since $\hat{\theta}_2$ can not form an overlap with $\hat{\theta}_1$, it follows that $\hat{\theta}_2$ must begin and end with $x_2$. We conclude that $\hat{\theta}_2 = x_2$. In the same way it follows that $\hat{\theta}_i = x_i$ for all $i$.

Conversely, it is clear that $\hat{\theta}_i = x_i$ for all $i$ implies compatibility with the ordering, and our proposition is proved.

We see that the elimination ordering is in some sense the worst case. Because for any ordering, $\hat{\theta}_i = x_i$ for all $i$ clearly implies that the two conditions in our main theorem are fulfilled. However, even for the elimination ordering our theorem is useful; we may *e.g.* after $\hat{\theta}_n$ in $\theta_n$ have any polynomial in $x_1, \ldots, x_{n-1}$.

Finally we mention that, for some Gröbner basis $G$, $G \circ \Theta$ can of course be a Gröbner basis even if $\Theta$ do not fulfill the two conditions in Theorem 3. By understanding the ideas in this paper we can sometimes, given $G$ and $\Theta$, decide whether $G \circ \Theta$ is a Gröbner basis without using our main theorem.

## ACKNOWLEDGMENT

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

## REFERENCES

1. Anick, D.J. Noncommutative Graded Algebras and Their Hilbert Series. J. Algebra **1982**, *78* (1), 120–140.
2. Cohn, P.M. *Free Rings and Their Relations*, 2nd Ed.; Academic Press Inc. [Harcourt Brace Jovanovich Publishers]: London, 1985.
3. Green, E.L. An Introduction to Noncommutative Gröbner Bases. In *Computational Algebra (Fairfax, VA, 1993)*; Dekker: New York, 1994; 167–190.
4. Gutierrez, J.; San Miguel, R.R. Reduced Gröbner Bases Under Composition. J. Symbolic Comput. **1998**, *26* (4), 433–444.
5. Hong, H. Groebner Basis Under Composition. II. In *Proc. ISSAC'96*; ACM Press, 1996.
6. Hong, H. Groebner Basis Under Composition. I. J. Symbolic Comput. **1998**, *25* (5), 643–663.
7. Lyndon, R.C.; Schützenberger, M.P. The Equation $a^M = b^N c^P$ in a Free Group. Michigan Math. J. **1962**, *9*, 289–298.
8. Mora, F. Groebner Bases for Noncommutative Polynomial rings. In *Algebraic Algorithms and Error Correcting Codes (Grenoble, 1985)*; vol. 229 of *Lecture Notes in Comput. Sci.* Springer: Berlin, 1986; 353–362.
9. Nordbeck, P. On Some Basic Applications of Gröbner Bases in Non-Commutative Polynomial Rings. In *Gröbner Bases and Applications*; vol. 251 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press: Cambridge, 1998; 463–472.

Received June 1999
Revised December 2000