On Bergman's Diamond Lemma for Ring Theory

Takao Inoué*

AI Nutrition Project National Institutes of Biomedical Innovation, Health and Nutrition, Japan

Abstract
This expository and review paper deals with the Diamond Lemma for ring theory, which is proved in the first section of G. M. Bergman, The Diamond Lemma for Ring Theory, Advances in Mathematics, 29 (1978), pp. 178–218. No originality of the present note is claimed on the part of the author, except for some suggestions and figures. Throughout this paper, I shall mostly use Bergman’s expressions in his paper. In Remarks and Notes, the reader will find some useful information on this topic.

Keywords: Diamond Lemma, Ring Theory, Rewriting, Algebras, Associative Algebras, 2020 Mathematics Subject Classification: 16S15, 16-02.

1. Introduction
This is an expository and review paper which deals with the Diamond Lemma for ring theory, which is proved in the first section of G. M. Bergman, The Diamond Lemma for Ring Theory, Advances in Mathematics, 29 (1978) [1,2]. No originality of the present note is claimed on the part of the author, except for some suggestions and figures. Throughout this paper, I shall mostly use Bergman’s expressions in his paper. In Remarks and Notes, the reader will find some useful information on this topic.

Suppose that \( R \) is an associative algebra with 1 over the commutative ring \( k \), and that we have a presentation of \( R \) by a family \( X \) of generators and a family \( S \) of relations. Suppose that each relation \( \sigma \in S \) has been written in the form \( W_{\sigma} = f_{\sigma} \), where \( W_{\sigma} \) is a monomial (a product of elements of \( X \)) and \( f_{\sigma} \) is a \( k \)-linear combination of monomials, and that we want to use these relations as instructions for reducing expressions \( r \) for elements of \( R \). That is, if any of the monomials occurring in the expression \( r \) contains one of the \( W_{\sigma} \) as a subword, we substitute \( f_{\sigma} \) for that subword, and we iterate this procedure as long as possible. In general, this process is not always well defined: at each step we must choose which reduction to apply to which subword of which monomial. Etcetera. So we are naturally led to the following questions:

(1) Under what conditions will such a procedure bring every expression to a unique irreducible form?
(2) Suppose that we have a set of suitable conditions satisfying (1). Does this yield then a canonical form for elements of \( R \)?

The Diamond Lemma is a general result of this sort due to Newman which was obtained in a graph-theoretic context [3]. Let \( G \) be an oriented graph. Here the vertices of \( G \) may be thought as expressions for the elements of some algebraic object (in our case, an associative algebra with 1 over the commutative ring \( k \)) and the edges as reduction steps (in our case, reductions using such a rule as \( W_{\sigma} = f_{\sigma} \)) going from one such expression to another one. Newman’s result is the following [1]. suppose that
(i) The oriented graph satisfies the descending chain condition. That is, all positively oriented path in \( G \) terminate; and
(ii) Whenever two edges, \( e \) and \( e' \), proceed from one vertex \( a \) of \( G \), there exists positively oriented paths \( p, p' \) in \( G \) leading from the end points \( b, b' \) of these edges to a common vertex \( c \). (This condition is called the diamond condition.)
Then every connected component $C$ of $G$ has a unique minimal vertex $m_C$.
This means that every maximal positively oriented path beginning at a point of $C$ will terminate at $m_C$; in other words (in our context) that the given reduction procedures yield unique canonical forms for elements of the original algebraic object. The main theorem to be proved in the third section, namely the Diamond Lemma for Ring Theory, is an analogue of the above observations for the case of associative rings, with reduction procedures of the form mentioned earlier. (For our argument in the sequel, we do not follow Newman’s graph-theoretic formulation).

In the following section 2, we introduce a lot of definitions and prove some lemmas and propositions used for the proof of the Diamond Lemma. In the last fourth section, we give some suggestions on literatures and so on. We have Notes and Appendix at the end of this paper.

2. Preliminaries
Let $k$ be a commutative associative ring with 1, $X$ a set, $<X>$ the free semigroup with 1 on $X$, and $k<X>$ the free associative $k$-algebra on $X$, which is the semigroup algebra of $<X>$ over $k$.

Let $S$ be a set of pairs of the form $\sigma = (W, f)$

where $W, f \in <X>$, $f \in k <X>$. For any $\sigma \in S$ and $A, B \in <X>$, let $r_{\sigma AB}$ denote the $k$-module endomorphism of $k <X>$ that fixes all elements of $<X>$ other than $A_{\sigma AB}$ and that sends this basis element to $A_{\sigma AB}$. We call the given set $S$ a reduction system, and the maps $r_{\sigma AB} : k <X> \rightarrow k <X>$ reductions.

We say that a reduction $r_{\sigma AB}$ acts trivially on an element $a \in k <X>$ if the coefficient of $r_{\sigma AB}$ in $a$ is zero. An element $a \in k <X>$ is said to be irreducible if every reduction act trivially on $a$.

**Proposition 2.1** The irreducible elements of $k <X>$ form a $k$-submodule of $k <X>$, denoted by $k <X>_{irr}$.
Proof. Let $a, b$ be any irreducible elements of $k <X>$ and $\lambda$ any element of $k$. Let $r$ be a reduction, say $r = r_{\sigma AB}$. The coefficient of $r_{\sigma AB}$ in $a$ and $b$ is zero, respectively. Thus so is that of $r_{\sigma AB}$ in $a - b$ and $\lambda a$. Trivially 0 is irreducible. This completes the proof.
A finite sequence of reduction \( r_{i_1} \ldots r_{i_l} (r_i = r_{i \text{ end}}) \) is said to be final on \( a \in k < X > \) if \( r_{i_1} \ldots r_{i_l} (a) \in k < X >_{i_l} \). An element \( a \) of \( k < X > \) is called reduction-finite if for every infinite sequence \( r_1 r_2 \ldots, \) of reductions, \( r_i \) acts trivially on \( r_{i+1} \ldots r_{i_l} (a) \) for all sufficiently large \( i \). If \( a \) is reduction-finite, then any maximal sequence of reductions \( r_i \) acts nontrivially on \( r_{i_1} \ldots r_{i_l} (a) \) is finite, and hence a final sequence.

**Proposition 2.2** The reduction-finite elements of \( k < X > \) form a \( k \)-submodule of \( k < X > \).

Proof: Suppose that \( a \) and \( b \) are reduction-finite elements and \( \lambda \) an element of \( k \). Then there are natural number \( i \) and \( j \) such that for every infinite sequence \( r_1 r_2 \ldots, \) of reductions, \( r_i \) and \( r_j \) act trivially on \( r_{i_1} \ldots r_{i_l} (a) \) and \( r_{j_1} \ldots r_{j_l} (b) \), respectively. Take \( l = \max(i, j) \). For every infinite sequence \( r_1 r_2 \ldots, \) of reductions, \( r_i \) and \( r_j \) act trivially on \( r_{i_1} \ldots r_{i_l} (a - b) \) and \( r_{j_1} \ldots r_{j_l} (\lambda a) \), respectively. Thus, \( a - b \) and \( \lambda a \) are reduction-finite. 0 is clearly reduction-finite. This completes the proof.

We call an element \( a \in k < X > \) reduction-unique if

1. it is reduction-finite; and
2. its images under all final sequences are the same. (This common value is denoted by \( r_i (a) \)).

**Lemma 2.3** (i) The set of reduction-finite elements of \( k < X > \) form a \( k \)-submodule of \( k < X > \), and \( rs \) is a \( k \)-linear map of this submodule into \( k < X >_{irr} \).

(ii) Suppose \( a, b, c \in k < X > \) are such that for all monomials \( A, B, C \) occurring with nonzero coefficient in \( a, b, c \), respectively, the product \( ABC \) is reduction-unique. (In particular this implies that \( abc \) is reduction-unique.)

(iii) Suppose \( a, b, c \in k < X > \) are such that for all monomials \( A, B, C \) occurring with nonzero coefficient in \( a, b, c \), respectively, the product \( ABC \) is reduction-unique. (In particular this implies that \( abc \) is reduction-unique.)

**Proof**

(i) Suppose that \( a, b \in k < X > \) are reduction-unique, and \( a \in k \). By Proposition 2.2, \( aa + b \) is reduction-finite. Let \( r \) be any composition of (finite) reductions final on \( aa + b \). Since \( a \) is reduction-unique, we can find a composition of (finite) reduction \( r' \) such that \( r'r(a) = r(a) \), and similarly there is a composition of reductions \( r'' \) such that \( r''r'(b) = r(b) \). Because \( r(aa + b) \in k < X >_{irr} \), we have

\[
    r(\alpha a + b) = r''r'(\alpha a + b) \\
    = \alpha r''r'(a) + r''r'(b) \\
    = \alpha r''s(a) + r_s(b) \\
    = \alpha r_s(a) + r_s(b)
\]

That is, images of \( \alpha a + b \) under all such final sequences of reductions are the same, i.e. \( \alpha r_s(a) + r_s(b) \). Thus, \( \alpha a + b \) is reduction-unique and so is \( b - a \) with \( \alpha = -1 \). 0 is clearly reduction-unique. Therefore, the set of reduction-unique elements of \( k < X > \) forms a \( k \)-submodule of \( k < X > \). Since \( r_s(aa + b) = r(aa + b) \), \( r_s(aa + b) = \alpha r_s(a) + r_s(b) \). For any reduction-unique elements of \( k < X > \), \( r_s(s) \in k < X >_{irr} \) is clear. Thus, \( r_s \) is a \( k \)-linear map of the module into \( k < X >_{irr} \).

(ii) Suppose that the assumption of (ii) holds. And say

\[
a = \sum_i \alpha_i A_i, \quad b = \sum_j \beta_j B_j, \quad c = \sum_l \gamma_l C_l.
\]

So \( abc = \sum_{i,j,l} \alpha_i \beta_j \gamma_l A_i B_j C_l \) and for any triple \( (i, j, l) \), \( A_i B_j C_l \) is reduction-unique. So \( abc \) is reduction-unique and \( r_s(abc) = \sum_{i,j,l} \alpha_i \beta_j \gamma_l r_s(A_i B_j C_l) \). Let \( r \) be any finite composition of reductions. It is sufficient to consider the case where \( a, b, c \) are monomials and \( r \) is a single reduction \( r_{D \sigma E} \). In this case, \( Ar_{D \sigma E} (B)C = r_{D \sigma E} BC \), which is the image of \( ABC \) under a reduction, hence is reduction-unique if \( ABC \) is so, with the reduced form. □

By an overlap ambiguity of \( S \) we mean a 5-tuple \( (\sigma, \tau, A, B, C) \) with \( \sigma, \tau \in S \) and \( A, B, C \in k < X > \) \(-\{1\} \), such that \( W_\sigma = AB, W_\tau = BC \). We say that the overlap ambiguity \( (\sigma, \tau, A, B, C) \) is resolvable if there exist compositions of reductions, \( r, r' \), such that \( r(f_C) = r'(A f_\tau) \); in other words, \( f_C \) and \( A f_\tau \) can be reduced to a common expression (This corresponds to the diamond condition seen in the introduction).
Let $\sigma \neq \tau \in S$ and $A, B, C \in < X >$ is called and inclusion ambiguity if $W_\sigma = B, W_\tau = ABC$. The inclusion ambiguity is called resolvable if there exists compositions of reductions, $r$ and $r'$, such that $r(Af_\sigma B) = r'(f_\tau)$.

By a semigroup partial ordering on $< X >$, we mean a partial order $<$ such that $B < B' \iff ABC < AB'C$ for any $A, B, B', C \in k < X >$, and it is called compatible with $S$ if for all $\sigma \in S$, $f_\sigma$ is a (finite) linear combination of monomials $< W_\sigma$.

If $\leq$ is a semigroup partial ordering on $< X >$ compatible with the reduction system $S$, and $A$ is any element of $< X >$, let $I_A$ denote the submodule of $k < X >$ spanned by all elements $B(W_\sigma - f_\sigma)C$ such that $BW_\sigma C < A$. We say that an overlap (inclusion) ambiguity $(\sigma, \tau, A, B, C)$ is resolvable relative to $\leq$ if $f_\sigma C - Af_\tau \in I_{ABC}$ ($A(f_\sigma C - f_\tau) \in I_{ABC}$).

The following lemma is trivial. But it will be useful in what follows.

**Lemma 2.4** Let $a \in k < X >$. Suppose that $a$ contains a monomial of the form $AW_\sigma B$ with a coefficient $\lambda (\neq 0) \in k$. Then we have

$$r_{A\sigma B}(a) = a - \lambda A(W_\sigma - f_\sigma)B.$$

**Proof.** The lemma immediately follows from the following observation:

$$r_{A\sigma B}(\lambda AW_\sigma B) = \lambda Af_\sigma B = \lambda AW_\sigma B - \lambda A(W_\sigma - f_\sigma)B.$$

Let $I$ denote the two-sided ideal of $k < X >$ generated by the elements $W_\sigma - f_\sigma (\sigma \in S)$. As a $k$-module, $I$ is spanned by the products $A(W_\sigma - f_\sigma)B$.

**Proposition 2.5** Let $a \in k < X >$. Suppose that $a$ is reduction-unique. Then, if $r_s(a) = 0$, then $a$ is an element of $I$.

**Proof.** If $a = 0$, the $0 \in I$ is trivial. So assume $a \neq 0$. Suppose that $r$ is the composition of a sequence final on $a$ and say

$$r = r_{A_n \sigma_n B_n} \cdots r_{A_2 \sigma_2 B_2} r_{A_1 \sigma_1 B_1}.$$
Then, By Lemma 2.4, $r(a)$ is of the form $a - \sum \lambda_i A_i(W_{\sigma_i} - f_{\sigma_i})B_i$ with $\lambda_i(\neq 0) \in k$ for all $i$. Because $a$ is reduction-unique, $r(a) = r_s(a)$, which implies $a = \sum \lambda_i A_i(W_{\sigma_i} - f_{\sigma_i})B_i$. Thus, $a \in I$. \hfill \Box

3. The Diamond Lemma

The following theorem is called the Diamond Lemma for Ring Theory.

**Theorem 3.1** Let $S$ be a reduction system for a free associative algebra $k < X >$ (a subset of $< X > \times k < X >$), and $\leq$ a semigroup partial ordering on $< X >$, compatible with $S$, and satisfies the descending chain condition. Then the following conditions are equivalent:

(a) All ambiguities of $S$ are resolvable.
(b) All elements of $k < X >$ are reduction-unique under $S$.
(c) A set of representatives in $k < X >$ for the elements of the algebra $R = k < X > / I$ determined by the generators and the relations $W_{\sigma} = f_{\sigma} (\sigma \in S)$ is given by the $k$-submodule $k < X >_{irr}$ spanned by the $S$-irreducible monomials of $< X >$.

When these conditions hold, $R$ may be identified with the $k$-module $k < X >_{irr}$, made a $k$-algebra by the multiplication $a \cdot b = r_S(ab)$.

**Proof.** First we see from our general hypothesis, that every element of $< X >$ is reduction-finite. We can prove this formally by induction with respect to the partial ordering with the descending chain condition $\leq$. But here we prove it informally to make the situation clearer. For illustrative purposes, suppose that $a \in k < X >$ has a monomial of the form $AW_{\sigma}B$ ($\sigma \in S$). By a reduction $r_{A\sigma B}$, $r_{A\sigma B}(a)$ has a monomial of the form $Af_{\sigma}B$. Since $\leq$ is compatible with $S$, $f_{\sigma}$ is of the form $\sum \lambda_i W_{\tau_i}$ with $W_{\tau_i} < W_{\sigma}$ for any $i$. So every monomial of $Af_{\sigma}B$ is of the form $AW_{\tau}B$ for some $i$. If the monomial $AW_{\tau}B$ contains a subword $W_{\tau}$ ($\tau \in S$), say $AW_{\tau}B = AW_{\tau}B'$, by a reduction $r_{A\tau' B'}$, $r_{A\tau' B'}(a)$ has a monomial of the form $A'f_{\tau}B'$. By compatibility of $\leq$, $f_{\tau}$ is again a (finite) linear combination of monomials $< W_{\tau} >$, say $f_{\tau} = \sum \mu_j W_{\tau_j}$ with $W_{\tau_j} < W_{\tau}$ for any $j$. So $r_{A\tau' B'}(a)$ has a monomial of the form $A'W_{\tau_j}B'$ for all $j$. If we iterate this process, we will get a sequence of monomials, for example $W_{\sigma} > W_{\sigma'} > W_{\tau} > \cdots$. All of such sequences must be finite because of the descending chain condition. It is also clear that the number of all the possible sequences is finite. (See the following figure.) Therefore, $a$ is reduction-finite. Since every element of $< X >$ is reduction-finite, hence so is every element of $k < X >$.

Next we prove $(b) \Leftrightarrow (c)$. We note first that $(c)$ simply says

$$k < X > \simeq k < X >_{irr} \oplus I.$$

Assuming $(b)$, we show that the following sequence,

$$(*) \quad 0 \rightarrow I \xrightarrow{i} k < X > \xrightarrow{r_S} k < X >_{irr} \rightarrow 0,$$

is a short exact sequence $k$-module homomorphisms, where $i$ is inclusion map. If so, it is immediate to conclude $k < X > \simeq k < X >_{irr} \oplus I$. That is, it follows from $r_S i = id_{k < X >_{irr}}$ (see Appendix).

Case 1: inclusion map $i$ is injective. So $0 \rightarrow I \rightarrow k < X >$ is exact.
By Proposition 2.5, above, (b) is nothing but the exactness of \( I \) that we may have then, \( i \leq r \) for any \( r \) is, there are compositions of reductions \( \lambda \) are obviously final on \( \mathcal{W} \) and with similar remarks, (2) For any \( x \in k \langle X \rangle \) that holds, because we see that

\[
(aa) \cdot b = r_S((aa) \cdot b) = r_S(a(ab)) = (aa)b = ab, \quad (\text{aa})b = a(ab) \text{ holds since } k < X > \text{ is a } k\text{-algebra.}
\]

By Lemma 2.3.(i). Thus it is a \( k \)-module.

(2) For any \( a, b \in k \langle X \rangle \) and any \( a \in k \),

\[
\alpha(a \cdot b) = (\alpha a) \cdot b = a \cdot (ab),
\]

holds, because we see that

\[
(aa) \cdot b = r_S((aa) \cdot b) = r_S(a(ab)) = (aa)b = ab. \quad (\text{aa})b = a(ab) \text{ holds since } k < X > \text{ is a } k\text{-algebra.}
\]

and with similar remarks, \( a \cdot (ab) = (\alpha a) \cdot b \).

We next deal with the proof of (b) \( \iff (a) \). Suppose (b). We consider only the case of overlap ambiguities, because those of inclusion ones are similarly taken case of. Let \( (\sigma, \tau, A, B, C) \) be any overlap ambiguity. \( f_\sigma C, Af_\tau \) are reduction-unique by (b). So we may take compositions of reductions \( r \) and \( r' \) which are final on \( f_\sigma C \) and \( A f_\tau \), respectively. By (b), \( ABC \) is reduction-unique. Moreover, \( rr_{1\sigma} C \) and \( r' r_{A1} \) are obviously final on \( ABC \). So we see that
$r(f_{r}C) = rr_{1}aC(ABC) = r_s(ABC) = r^r r_{A^r 1}(ABC) = r(Af_{r})$.  

This means that the ambiguity is resolvable.

In this paragraph, we prove (a) $\Rightarrow$ (a'). We assume (a). First we consider the case of overlap ambiguities.

Let $(\sigma, \tau, A, B, C)$ be any overlap ambiguity of $S$. By (a), it is resolvable. That is, there are compositions of reductions $r$ and $r'$ such that $r(f_{r}C) = r^r(Af_{r})$, say

$$r = rD_{n\sigma n}E_{n} \cdots rD_{2\sigma 2}E_{1} \quad \text{and} \quad r' = rD_{m^r\tau m^r}E_{m^r} \cdots rD_{1\tau 1}E_{1}.$$  

By Lemma 2.4, we obtain,

$$r(f_{r}C) = f_{r}C = \sum_{i=1}^{n} \lambda_{i} D_{i}(W_{\sigma_{i}} - f_{\sigma_{i}})E_{i}$$

and

$$r'(Af_{r}) = Af_{r} - \sum_{i=1}^{m} \mu_{i} D'_{i}(W_{\tau_{i}} - f_{\tau_{i}})E'_{i}$$

with $\lambda_{i}(\neq 0), \mu_{j}(\neq 0) \in k$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. Under the diamond condition in our sense, we may have then,

$$f_{r}C - Af_{r} = \sum_{i=1}^{n} \lambda_{i} D_{i}(W_{\sigma_{i}} - f_{\sigma_{i}})E_{i} - \sum_{i=1}^{m} \mu_{i} D'_{i}(W_{\tau_{i}} - f_{\tau_{i}})E'_{i}.$$  

Further, it is not so difficult to (I) $D_{i}W_{\sigma_{i}}E_{i} < W_{\sigma}C = ABC$ and (II) $D'_{j}W_{\tau_{j}}E'_{j} < W_{\tau} = ABC$ for any $1 \leq i \leq n$ and any $1 \leq j \leq m$. For the verification of (I), we show only the case of $i = 1$. The rest of the proof is taken care of by induction. Since $\leq$ is compatible with $S$, if $f_{r}$ is of the form $\sum \alpha_{i} Z_{i}$, then $Z_{1} < W_{\sigma}$ holds for any $i$. Further, $Z_{1} < W_{\sigma}$ leads to $Z_{i}C < W_{\sigma}C = ABC$ for all $i$. So we must have $D_{1}W_{\sigma_{1}}E_{1} = Z_{1}C$ for some $i$. Hence, $D_{1}W_{\sigma_{1}}E_{1} < ABC$. We can verify (II) similarly. So we omit the verification. For inclusion ambiguity, we can also show resolvability relative to $\leq$ in a completely similar way. So this is left to the reader.

In this paragraph, we take care of the last implication to be shown, i.e. (a') $\Leftarrow$ (b). It suffices to prove all monomials $D < X$ reduction-unique, since the reduction-unique elements of $k < X$ form a submodule (Lemma 2.3(i)). That is, if every monomial of $k < X$ is reduction-unique, then $k < X \Rightarrow_{r_{s}} k < X$. We assume inductively that all monomials $< D$ are reduction-unique. Thus the domain of $r_{s}$ includes the submodule spanned by all these monomials, so the kernel of $r_{s}$ contains $I_{D}$. That is, if $a \in ker(r_{s})$, then by Proposition 2.5, $a$ is of the form $\sum \lambda_{i} A_{i}(W_{\sigma_{i}} - f_{\sigma_{i}})B_{i}$, with $A_{i}W_{\sigma_{i}}B_{i} < D$ for any $i$, which means $a \in I_{D}$. We must now show that given any two reductions $r_{L\sigma M'}$ and $r_{L'r_{M}}$ each acting nontrivially on $D$ (and hence each sending $D$ to a linear combination of monomials $< D$), we will have

$$r_{s}(r_{L\sigma M'}(D)) = r_{s}(r_{L'r_{M}}(D)).$$

We have to check three cases for that, according to the relative locations of the subwords $W_{\sigma}$ and $W_{\tau}$ in the monomial $D$. We may assume without loss of generality that $\text{length}(L) \leq \text{length}(L')$, in other words, that the indicated copy of $W_{\sigma}$ in $D$ begins no later than that of $W_{\tau}$.

Case 1: The subwords $W_{\sigma}$ and $W_{\tau}$ overlap in $D$, neither contains the other, figured as follows: under the condition $\text{length}(L) \leq \text{length}(L')$,

$$\begin{array}{c|c|c} L & W_{\sigma} & M' \\ \hline & & \\ L' & W_{\tau} & M \end{array}$$

Then $D = LABC\tau C$, where $(\sigma, \tau, A, B, C)$ is an overlap ambiguity of $S$, i.e. $W_{\sigma} = AB, W_{\tau} = BC, \sigma, \tau \in S, A, B, C \in < X > \setminus \{1\}$. Then,
\[ F := r_{L \sigma M}(D) - r_{L^\tau M}(D) \\
= Lf_\sigma CM - LAf_\tau M \\
= L(f_\sigma CM - Af_\tau M) \\
= L(f_\sigma C - Af_\tau M). \quad \cdots \cdots (1.1) \]

By (a’) every overlap ambiguity is resolvable relative to \( \geq \). So we have \( f_\sigma C - Af_\tau \in I_{ABC} \) by definition. That is,
\[ f_\sigma C - Af_\tau = \sum \lambda_i D_i(W_{\sigma_i} - f_{\sigma_i})E_i \]
with \( D_iW_{\sigma_i}E_i < ABC \) for any \( i \). Substitute this to (1.1). Then we get,
\[ F = \sum \lambda_i LD_i(W_{\sigma_i} - f_{\sigma_i})E_iM. \quad \cdots \cdots (1.2) \]

Since \( \geq \) is a semigroup ordering, the following inequality,
\[ LD_i(W_{\sigma_i})E_iM < LABCM. \quad \cdots \cdots (1.3) \]
holds by \( D_i(W_{\sigma_i})E_i < ABC \). From (1.2) and (1.3), it follows that \( F \in I_{LABCM} \).
Thus \( r_S(F) = 0 \), in other words,
\[ r_S(r_{L \sigma M'}(D) - r_{L^\tau M}(D)) = 0, \]
so \( r_S(r_{L \sigma M'}(D)) = r_S(r_{L^\tau M}(D)) \).

The next case is similarly dealt with as the case 1. But we shall work it out for the sake of the reader.

Case 2: One of the subwords \( W_\sigma \), \( W_\tau \) is contained in the other. By \( \text{length}(L) \leq \text{length}(L') \), we have the following case where \( W_\sigma \) contains \( W_\tau \), figured below.

\[
\begin{array}{c|c|c}
L & W_\sigma & M' \\
\hline
L' & W_\tau & M \\
\end{array}
\]

Then \( D = LABCM' \), \( CM' = M \) and \( L' = LA \), where \( (\sigma, \tau, A, B, C) \) is an inclusion ambiguity of \( S \), i.e. \( W_\tau = B, W_\sigma = ABC \) with \( \tau \neq \sigma \in S, A, B, C \in X \). Then,
\[ F := r_{L \sigma M'}(D) - r_{L^\tau M}(D) \\
= Lf_\sigma M' - LAf_\tau CM' \\
= L(f_\sigma - Af_\tau C)M'. \quad \cdots \cdots (2.1) \]

By (a’) we know that every inclusion ambiguity is resolvable relative to \( \geq \). So we get \( Af_\tau C - f_\sigma \in I_{ABC} \) by definition. That is,
\[ Af_\tau C - f_\sigma = \sum \lambda_i D_i(W_{\sigma_i} - f_{\sigma_i})E_i \]
with \( D_iW_{\sigma_i}E_i < ABC \) for any \( i \). Substituting this to (2.1), we obtain,
\[ F = \sum (-\lambda_i) LD_i(W_{\sigma_i} - f_{\sigma_i})E_iM'. \quad \cdots \cdots (2.2) \]

Because \( \geq \) is a semigroup ordering, the following inequality,
\[ LD_i(W_{\sigma_i})E_iM' < LABCM. \quad \cdots \cdots (2.3) \]
holds by \( D_i(W_{\sigma_i})E_i < ABC \). From (2.2) and (2.3), we get \( F \in I_{LABCM'} = I_D \), so
\[ r_S(F) = 0. \]
That is,
\[ r_S(r_{L \sigma M'}(D) - r_{L^\tau M}(D)) = 0, \]
thus
\[ r_S(r_{L \sigma M'}(D)) = r_S(r_{L^\tau M}(D)). \]

The following is our last case to check, with which we complete the whole proof of the Diamond Lemma.

Case 3: We consider the case where \( W_\sigma \) and \( W_\tau \) is disjoint. By the condition on the length of \( L \) and \( L' \), \( \text{length}(L) \leq \text{length}(L') \), the case is figured below.
So we may assume \( D = LW_\sigma NW_\tau M \), i.e.,

\[
\begin{array}{c|c|c}
L & W_\sigma & M' \\
\hline
L' & W_\tau & M \\
\end{array}
\]

is our present case with

\[
\begin{align*}
r_{L\sigma M}(D) &= LW_\sigma NW_\tau M, & (M' = NW_\tau M) \\
\text{and} \quad r_{L\tau M}(D) &= LW_\tau Nf_\tau M. & (L' = LW_\tau N)
\end{align*}
\]

By the general assumption, the ordering \( \leq \) is compatible with \( S \). So \( f_\sigma \) can be written as a linear combination of monomials \( < W_\sigma \), say \( f_\sigma = \sum \lambda_i Z_i \) with \( Z_i < W[i], \lambda_i(\neq 0) \in k \). The ordering is a semigroup ordering, so for any \( i \), we have \( LZ_iNW_\tau M < LW_\sigma NW_\tau M = D \) from \( Z_i < W_\sigma \). By induction hypothesis, \( LZ_iNW_\tau M \) is reduction-unique for all \( i \). Let \( a = 1, c = 1 \) and \( b = LF_\sigma NW_\tau M = \sum \lambda_i LZ_iNW_\tau M \). Then, for all monomials \( A, B, C \) occurring with non-zero coefficient in \( a, b, c \), i.e. 1, \( LZ_iNW_\tau M(\forall i) \), respectively, the product \( ABC \), namely \( LZ_iNW_\tau M(\forall i) \) is reduction-unique. Apply now Lemma 2.3,(ii) to such \( a, b, c \) with \( r = r_{L(\lambda_1N)\tau M} \). Then we have \( ar(b)c \) is reduction-unique and \( r_S(ar(b)c) = r_S(abc) \). In other words, \( LF_\sigma NW_\tau M \) is reduction-unique and

\[
r_S(LF_\sigma NW_\tau M) = r_S(LF_\sigma NW_\tau M). \quad \cdots (3.1)
\]

Similarly, we can obtain,

\[
r_S(LF_\sigma NW_\tau M) = r_S(LW_\sigma NW_\tau M). \quad \cdots (3.2)
\]

From (3.1) and (3.2), it follows that \( r_S(LF_\sigma NW_\tau M) = r_S(LW_\tau Nf_\tau M) \), which implies \( r_S(r_{L\sigma M}(D)) = r_S(r_{L\tau M}(D)) \). \( \square \)

The following corollary may be reserved for the reader to prove.

**Corollary 3.2** Let \( k < X > \) be a free associative algebra, and "\( \leq \)" a semigroup partial ordering of \( < X > \) with the descending chain condition.

If \( S \) is a reduction system on \( k < X > \) compatible with \( \leq \) and having no ambiguities, then the set of \( k \)-algebra relations \( W_\sigma = f_\sigma (\sigma \in S) \) is independent.

More generally, if \( S_1 \subseteq S_2 \) are reduction systems, such that \( S_2 \) is compatible with \( \leq \) and all its ambiguities are resolvable, and if \( S_2 \) contains some \( \sigma \) such that \( W_\sigma \) is irreducible with respect to \( S_1 \), then the inclusion of ideals associated with these systems, \( I_1 \subseteq I_2 \), is strict.

4. **Remarks**

First I would like to recommend the reader to read the original paper [2] of Bergman, because it is written with a broad perspective over a lot of algebraic structures, where the reader will find many interesting materials.

I have to mention at least that there is the correction and updates for the paper. Refer to Bergman [3].

The Diamond Lemma has another origin, although Newman [10] is already mentioned. For that, refer to Bokut et al [4] and Shirshov [11]. Also see Matveev [9]. Historically, Shirshov [11] gave the present lemma first for Lie algebras. Someone calls the Diamond Lemma Schirshov-Bergman’s diamond lemma.
To get more recent trends for the Diamond Lemma, I would like to cite, among others, Chenavier [5], Chenavier and Lucas [6], Elias [7] and Tsuchioka [12]. There the reader will find much more information on the lemma and see some practices as an application to representation theory, as an example.

Recently, there is a trend about Composition-Diamond Lemma. But I do not touch on it here. I shall have an opportunity in another occasion.

I think the Diamond Lemma and its techniques can be applied not only to mathematics but also to many scientific fields.

Notes

1As a remark for logicians, Newman’s paper [10] is closely related to the theory of λ-calculi. This article also contains an interesting observation about a relation of weak Church-Rosser and Church-Rosser properties (see Barendregt [1, p. 58]).

2Newman’s original formulation and terminology for the Diamond Lemma is different from the one in the introduction of the present note (see Newman [10] for the details.)

3Given such k, X in the context, the semigroup algebra of < X > over k is called the free (associative) k-algebra on X.

4Inclusion ambiguities are, in a sense, always avoidable. Suppose that S is a reduction system for a free algebra k < X >. Let us construct a subset S′ ⊆ S by (1) deleting all σ ∈ S such that Wσ contains a proper subword of the form Wτ (τ ∈ S), and (2) whenever more than one element σ1, σ2, ..., σr ∈ S act on the same monomial (i.e. Wσ1 = Wσ2 = ...) dropping all but one of the σi from S. Then S′ ⊆ S will have the property that a ∈ k < X > is reducible under S. But from this it follows that if a ∈ k < X > is reduction-unique under S, then so is a under S′ and rS′(a) = rS(a). Hence if S is such that every element of k < X > is reduction-unique under it, then S′ has the same property and rS′ = rS. Therefore, S′, which has no inclusion ambiguities, defines the same ring and the same canonical form as S. (This remark is due to Bergman [2, p. 192].)

5As an application of the lemma, Bergman shows, for example an alternative proof of Poincaré-Birkoff-Witt Theorem in [2, p.186], which gives a basis of the universal enveloping algebra U(g) of g if we know a basis of a Lie algebra g. Also see Varadarajan [13].

Appendix

In this appendix, we show a well-known basic theorem for modules without a proof, which is used, in the proof of Theorem 3.1, namely the Diamond Lemma, in the third section.

**Theorem 4.1** Let R be a ring and 0 → A → B → C → 0 a short exact sequence of R-module homomorphism. Then the following statements are equivalent.

1. There is an R-module homomorphism h : C → B with gh = idC.
2. There is an R-module homomorphism k : B → A with kf = idA.
3. The given sequence is isomorphic (with identity maps on A and C) to the direct sum short exact sequence

   0 → A → A ⊕ C → C → 0,

in particular B ∼= A ⊕ C.

**Proof.** See for example Hungerford [8]. □
Acknowledgement
I would like to thank Professor Emeritus Dr. Michiel Hazewinkel for correcting a preliminary version of this note and his encouragements to me.

References
[1] H. P. Barendregt, The Lambda Calculus, its Syntax and Semantics, North-Holland, Amsterdam (1981).
[2] G. M. Bergman, The Diamond Lemma for Ring Theory, Advances in Mathematics, 29 (1978), pp. 178–218.
[3] G. M. Bergman, "CORRECTIONS AND UPDATES TO George M. Bergman, The diamond lemma for ring theory, Advances in Mathematics 29 (1978) 178-218". https://math.berkeley.edu/~gbergman/papers/updates/diamond.html
[4] L. A. Bokut, V. Latyshev, I. Shestakov, E. Zelmanov, M. Bremner and M. V. Kotchetov (eds.), Selected Works of A. I. Shirshov, Birkhäuser, Basel (2009).
[5] C. Chenavier, Reduction operators and completion of rewriting systems, Journal of Symbolic Computation, 84 (2018), pp. 57–83.
[6] C. Chenavier and M. Lucas, The Diamond Lemma for non-terminating rewriting systems using deterministic reduction strategies, IWC 2019 - 8th International Workshop on Confluence, Jun 2019, Dortmund, Germany. pp.1–5. ffhal-02385139f.
[7] B. Elias, A diamond lemma for Hecke-type algebra, arXiv:1907.10571v1 [math.RT], 2019.
[8] T. W. Hungerford, Algebra, Springer-Verlag, New-York (1974).
[9] S. Matveev, The Diamond Lemma and prime decompositions (Intelligence of Low dimensional Topology), RIMS Kˆokyˆuroku 2015, 1960: 69-72: KJ00009973824.
[10] M. H. A. Newman, On theories with a combinatorial definition of "equivalence", Annals of Mathematics, 43 (1942), pp. 223–243.
[11] A. I. Shirshov, Some algorithm problems for Lie algebras, (Russian), Sibirskii Matematicheskiii Zhurnal, 50 (1962), pp. 292–296. Also in [4], pp. 125–130.
[12] S. Tsuchioka, BMR freeness for icosahedral family, arXiv:1710.03868 [math.RT].
[13] V. S. Varadarajan, Supersymmetry for mathematicians: an introduction, Courant Lecture Notes in Mathematics, vol. 11, New York University Courant Institute of Mathematical Sciences, New York, (2004).