THE UNIVERSITY OF MANITOBA

SOLUTION OF CERTAIN EQUATIONS IN FREE GROUPS

by

CHARLES C. EDMUNDS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

WINNIPEG, MANITOBA

October, 1973

SOLUTION OF CERTAIN EQUATIONS IN FREE GROUPS

by

CHARLES C. EDMUNDS

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

A C K N O W L E D G E M E N T S

This thesis represents part of the research done while the author was a graduate student at the University of Manitoba. I wish to take this opportunity to express my deepest thanks to the University and the National Research Council for their generous financial support. Above all, I wish to express my deepest appreciation to my supervisor, Professor N.D. Gupta, for his steadfast encouragement and guidance throughout my graduate studies. Thanks are also due to the staff and graduate students of the University of Manitoba Mathematics Department for providing a comfortable and stimulating environment in which to work. I would also like to thank my wife, Jean, for her patience and encouragement. Finally, I wish to express my appreciation to Miss Debbie Gessner for the excellent work she did typing this manuscript.

C O N T E N T S

P R E F A C E

Our main reference for this thesis is Combinatorial Group Theory by Magnus, Karrass, and Solitar; otherwise our exposition is self-contained. Definitions, Lemmas, and Theorems are numbered first with respect to chapter and then consecutively; thus in Chapter 1 the fourth and fifth numbered items are Definition 1.4 and Lemma 1.5. The third part of Lemma 1.5 will be referred to as Lemma 1.5(c). The symbol "■" indicates the end of a proof. In the proof of a Lemma or Theorem with several parts, the end of proof symbol will appear at the conclusion of the proof of the last part. An index of numbered items and an index of special symbols are included at the end of the manuscript. A bibliography is also included; our convention will be to refer to the n^{th} item in the Bibliography as [n].

ABSTRACT

If F is a free group and (w, u) is a pair of (reduced) words in F, the _endomorphism problem for the pair (w, u)_ is the problem of effectively deciding whether or not there is an endomorphism of F which sends w to u. Theorem 2.2 gives a necessary and sufficient condition that the endomorphism problem be solvable for w and arbitrary u. This condition is based on a set, $C(w)$, of words in F. To obtain our applications we prove (Theorem 5.1) that if w_1 and w_2 are quadratic words in F in non-overlapping variables for which $C(w_1)$ and $C(w_2)$ are finite sets, then $C(w_1 w_2)$ is also a finite set. Using this we show (Theorems 6.2 and 6.3) that the endomorphism problem is solvable for (w, u) where u is arbitrary for w a product of distinct squares or a product of commutators with non-overlapping entries.

CHAPTER 0

INTRODUCTION

Tarski originally raised the question of the decidability of the elementary theory of free groups. It seems that this theory could be decidable since the solutions of the classical decision problems for groups — the word problem, the conjugacy problem, the isomorphism problem — are nearly trivial for free groups. However, many investigations ([1] - [21]) made in the past fifteen to twenty years show Tarski's question to be very difficult. For the most part, these investigations involve special decision problems for free groups. For example, there is the <u>endomorphism problem</u> for a pair (w, u) of elements in a free group F :

<u>Problem I</u>. Can it be effectively decided (by a finite procedure) whether or not there is an endomorphism of F which sends w to u ?

An alternate form of this is the <u>substitution problem</u> [16] for the pair (w, u) :

<u>Problem II</u>. If F is freely generated by $\{x_1, x_2, \cdots, x_n\}$ and $w = w(x_1, x_2, \cdots, x_n)$, can it be effectively decided whether or not there are elements $w_1, w_2, \cdots, w_n \in F$ such that $w(w_1, w_2, \cdots, w_n) = u$?

In response to Problem II, raised by Lyndon in a note in [11], Schupp [16] solved the problem positively for every pair (w, u) in a free group of rank 2. This covered the only other results known at that time viz. (w, u) where w is a power or a commutator and u is arbitrary. The commutator result is due to Wicks [21] . In Chapter 2 we will give a necessary and sufficient condition (Theorem 2.2) that Problem II be solvable for any given w and arbitrary u's. Using this criterion, we will solve Problem II for a new class of words which is not covered by any previous result (Theorem 6.1) . As consequences of this, we will show that Problem II is solvable for (w, u) if w is a product of distinct squares or a product of commutators with distinct entries (Theorem 6.2 and 6.3). Our result for one commutator covers Wicks' result [21], and our solution to Problem II for w a product of n (≥ 2) commutators answers a question of Wicks (written communication).

In 1959, Baumslag, Boone, and Neumann [3] showed that there are groups for which it is impossible to decide whether or not a given word w is a commutator. Wicks [21] reports that in 1960 Boone asked if it could be effectively decided whether or not a given element of a free group is a commutator (i.e. whether Problem II is solvable for the pair $([x_1, x_2], u)$.) Wicks gives an elegant solution to this problem: A word w in a free group is a commutator if and only if some cyclic permutation of w has one of the forms $xyx^{-1}y^{-1}$ or $xyzx^{-1}y^{-1}z^{-1}$. Our Chapter 2 is inspired by Wicks' idea. Given a word, w, we will generate a set of words $C(w)$, the closure of w. This set will be used to test whether or not a word u can be derived from w by a substitution (i.e. whether or not there is an endomorphism of the free group which sends w to u.)

The process of finding $w_1, w_2, \cdots, w_n$ in F such that $w(w_1, w_2, \cdots, w_n) = u$, can be viewed as solving the equation

$$w(x_1, x_2, \cdots, x_n) = u$$

where the x_i's are thought of as variables and u is a constant. This is the approach taken by Lyndon in his study of equations in free groups ([8] - [14]). Lyndon was prompted to undertake this study by Tarski's question about the decidability of the elementary theory of free groups and by a specialized (unpublished) conjecture of Vaught, viz. if a, b, and c are elements of a free group F such that $a^2b^2 = c^2$, then $ab = ba$. In [8], Lyndon showed that under these conditions a, b, and c all lie within the same cyclic subgroup of F. Many generalizations of this result have been made [1], [2], [4], [14], [17], [18], [19], [20].

Along these lines we will reprove (Theorem 6.4) a new result of Lyndon and Newman [13]; namely, that if F is a free group freely generated by x_1 and x_2, then there are no words a and $b \in F$ such that $x_1x_2x_1^{-1}x_2^{-1} = a^2b^2$.

In Chapter 1 we describe our notation and introduce some definitions. The most important new idea here is that of a c-free (cancellation free) substitution (Definition 1.1(e)) for a freely reduced word w. This is a monoid endomorphism (from the free monoid X, freely generated by $\{x_1, x_1^{-1}, x_2, x_2^{-1}, \cdots\}$, to itself) which preserves formal inverses, sends no x-symbol occurring in w to the empty word, and sends w to a word which is freely reduced as written.

In Chapter 2 we define the set $C(w)$ and show that $C(w)$ is the minimal complete set of images of w in the sense that each element of $C(w)$ comes from w by a substitution, each element of X which comes from w by a substitution comes from some element of $C(w)$ by a c-free substitution, and any other set with these properties contains a copy of $C(w)$ (Theorems 2.13 and 2.14). From this we show (Theorem 2.2) that Problem III (which is equivalent to Problem II) is solvable for w and arbitrary u's if and only if membership in $C(w)$ is effectively decidable.

Having reduced Problem II for (w, u) to a study of $C(w)$, we develop the machinery to prove that $C(w)$ is finite if w is a quadratic word such that $w = \prod_{i=1}^{n} w_i$, where the w_i's are words in mutually disjoint symbols and each $C(w_i)$ is finite. In Chapter 3 we introduce certain diagrams in Euclidean 2-space called *-graphs (star-graphs). We also introduce a procedure for refining these *-graphs which in Chapter 4 is shown to yield a complete set of c-free solutions (Definition 4.6) to certain equations in X. The use we make of "graph" theory could probably be eliminated in favour of purely algebraic techniques, but the geometric nature of the c-free solution of equations makes the use of *-graphs convenient and appropriate. Approaching the problem in this way enables us to use our geometric as well as our algebraic intuition to study the solutions of equations in free groups. Perhaps further investigation will show that our main application (Theorem 6.1) could be proved replacing Chapter 2 by a more involved discussion of *-graphs in Chapters 3 and 4; however, we will not attempt to do this here.

Chapter 4 is a technical chapter in which the link is established between the *-graphs of Chapter 3 and the c-free solution of certain equations in X (see Corollary 4.12).

In Chapter 5 we prove Theorem 5.1: if w_1 and w_2 are freely reduced words in X such that

(i) w_1 and w_2 are quadratic (i.e. each word contains x_i, with exponent $+1$ or -1, at most twice,)

(ii) w_1 and w_2 are words in non-overlapping variables, and

(iii) both $C(w_1)$ and $C(w_2)$ are finite, then $C(w_1 w_2)$ is finite.

This implies that Problem II is solvable for $(w_1 w_2, u)$. We list some of the consequences of our work in Chapter 6.

CHAPTER 1

DEFINITIONS AND NOTATION

We begin by discussing our notational conventions regarding countably generated absolutely free groups. Any details which we omit can be found in Magnus, Karrass, and Solitar [15].

Let X denote the set of (reduced and unreduced) words in the symbols $\{x_1^{+1}, x_1^{-1}, x_2^{+1}, x_2^{-1}, \cdots\}$. For notational convenience we let the symbol "1" denote the empty word and abbreviate the symbols x_i^{+1} $(i \geq 1)$ as x_i.

It will be convenient to write a non-empty word w as $x_{i(1)}^{\epsilon(1)} x_{i(2)}^{\epsilon(2)} \cdots x_{i(n)}^{\epsilon(n)}$ where each $i(j)$ is a positive integer and $\epsilon(j)$ is $+1$ or -1. The length of w will be denoted by $\ell(w)$, and the <u>support</u> of w, which is the set

$$\{x_s : x_s \text{ or } x_s^{-1} \text{ occurs in } w\},$$

will be denoted by $\mathrm{Supp}(w)$.

If v and w name the same word, $x_{i(1)}^{\epsilon(1)} x_{i(2)}^{\epsilon(2)} \cdots x_{i(n)}^{\epsilon(n)}$, we say that v and w are <u>identically</u> (or <u>schematically</u>) equal and write $v \equiv w \equiv x_{i(1)}^{\epsilon(1)} x_{i(2)}^{\epsilon(2)} \cdots x_{i(n)}^{\epsilon(n)}$. Note that the words $x_1 x_1^{-1}$ and 1 are not identically equal.

If $v \equiv x_{i(1)}^{\epsilon(1)} x_{i(2)}^{\epsilon(2)} \cdots x_{i(m)}^{\epsilon(m)}$ and $w \equiv x_{j(1)}^{\eta(1)} x_{j(2)}^{\eta(2)} \cdots x_{j(n)}^{\eta(n)}$, we define the <u>product</u> of v and w as follows:

$$v \cdot w = x_{i(1)}^{\varepsilon(1)} x_{i(2)}^{\varepsilon(2)} \cdots x_{i(m)}^{\varepsilon(m)} x_{j(1)}^{\eta(1)} x_{j(2)}^{\eta(2)} \cdots x_{j(n)}^{\eta(n)},$$

(with $1 \cdot w = w \cdot 1 = w$). Clearly $(X, \cdot)$ is a semigroup with identity element 1; we call $(X, \cdot)$ <u>the free monoid in x-symbols</u>. If $w = x_{i(1)}^{\varepsilon(1)} x_{i(2)}^{\varepsilon(2)} \cdots x_{i(n)}^{\varepsilon(n)}$, we define the (<u>formal</u>) <u>inverse</u> of w as follows:

$$w^{-1} = x_{i(n)}^{-\varepsilon(n)} x_{i(n-1)}^{-\varepsilon(n-1)} \cdots x_{i(1)}^{-\varepsilon(1)},$$

(with $1^{-1} = 1$).

If $w = a \cdot u \cdot v \cdot b$ where a and b are (possibly empty) segments of w, u is a segment of w which ends with the symbol x_i^{ε}, and v is a segment of w which begins with the symbol x_j^{η}, we call the segment $x_i^{\varepsilon} x_j^{\eta}$ the <u>junction</u> of the segments u and v.

The words $x_i x_i^{-1}$ and $x_i^{-1} x_i$ $(i \geq 1)$ are called <u>trivial relators</u>. If w is a word in X, and no segment of w is a trivial relator, w is said to be <u>freely reduced</u>; we let $\overline{X}$ denote the set of all freely reduced words in X. Given any word w in X, repeated deletion of trivial relators finally yields a unique freely reduced word $\overline{w}$. We let $d(w)$ denote the number of deletions of trivial relators required to reduce w to $\overline{w}$; note that $d(w) = \frac{1}{2}(\ell(w) - \ell(\overline{w}))$. A word that results from w by the deletion of some (but not necessarily all) trivial relators is called <u>a partially reduced form</u> of w; we also call w <u>a partially reduced form</u> of itself.

Two words $v, w \in X$ are freely equal, denoted $v \approx w$, if $\overline{v} = \overline{w}$. It can be shown that "$\approx$" is an equivalence relation on X, and that the set $X/\approx$ with the multiplication and inverse operations induced by "$\cdot$" and "-1" forms the absolutely free group freely generated by $\{x_1, x_2, \cdots\}$. Furthermore, the freely reduced words in X comprise a set of unique equivalence class representatives for $X/\approx$. Henceforth we will denote the free group in x-symbols by $(\overline{X}, \circ)$ where $v \circ w = \overline{v \cdot w}$.

We now introduce some definitions of a more specialized nature.

Definition 1.1. (a) Given the monoid $(X, \cdot)$, a substitution from X into X is a homomorphism $\sigma : (X, \cdot) \to (X, \cdot)$ such that $1\sigma = 1$ and $x_s^{-1}\sigma = (x_s\sigma)^{-1}$ for each x_s. Note that if σ and τ are substitutions from X into X, $\sigma = \tau$ if and only if $x_s\sigma = x_s\tau$ for each x_s.

(b) If σ is a substitution from X into X such that $x_s\sigma \in \overline{X}$ for each x_s, we call σ a reduced substitution.

(c) The substitution which sends each word to the empty word is called the trivial substitution.

(d) The substitution which sends each word to itself is called the identity substitution.

(e) If $w \in \overline{X}$ and σ is a substitution from X into X, we say σ is c-free (cancellation free) for w if

$w\sigma \in \overline{X}$ and $x_s\sigma \neq 1$ for each $x_s \in \mathrm{Supp}(w)$. We denote by $S_X(w)$ the set of all c-free substitutions for w from X into X . Note that if σ is c-free for w , then $\ell(w) \leq \ell(w\sigma)$.

Definition 1.2. Given $w \in X$ and $x_s \in \mathrm{Supp}(w)$, w is said to be linear (quadratic) in x_s if the subscript s occurs exactly exactly once (twice) in the word w . The word w is said to be quadratic if for each $x_s \in \mathrm{Supp}(w)$ the subscript s occurs at most twice in w .

Definition 1.3. (a) If $v, w \in \overline{X}$, we call the expression $v \overset{V}{=} w$ a verbal equation; the elements of $\mathrm{Supp}(v)$ and $\mathrm{Supp}(w)$ are called verbal variables.

(b) A solution in X to the verbal equation $v \overset{V}{=} w$ is a pair (σ, τ) , with σ and τ substitutions from X into X such that $\overline{v\sigma} = \overline{w\tau}$.

(c) A c-free solution in X to the verbal equation $v \overset{V}{=} w$ is a solution, (σ, τ) , such that σ and τ are c-free for v and w respectively. We denote by $S_X(v, w)$ the set of all c-free solutions in X to the verbal equation $v \overset{V}{=} w$.

(d) A verbal equation $v \overset{V}{=} w$ is called quadratic if both v and w are quadratic words.

Definition 1.4. (a) The elementary level substitutions, λ_i and $\lambda_{i,j}$ $(i < j)$, from X into X are defined as follows:

$$x_s\lambda_i = \begin{cases} x_s & \text{if } s \neq i \\ x_i^{-1} & \text{if } s = i \end{cases} \quad \text{and} \quad x_s\lambda_{i,j} = \begin{cases} x_s & \text{if } s \neq i, j \\ x_j & \text{if } s = i \\ x_i & \text{if } s = j\,. \end{cases}$$

(b) A substitution λ is called level if it is a product of elementary level substitutions. See [15] p. 163 for a similar use of the term "level".

(c) Ordering the x-symbols $x_1 < x_1^{-1} < x_2 < x_2^{-1} < \cdots$, the words in X can be ordered lexicographically (see [15] p. 26). It is easy to see that $(X, \leq)$ is a well-ordered set, where $v \leq w$ means either v precedes w in the lexicographic ordering or $v = w$. A word $w \in X$ is said to be in (special) normal form if, for each level substitution λ from X into X, $w \leq w\lambda$. Henceforth, we will use the term "normal form" in place of "special normal form".

Lemma 1.5. (a) If λ is a level substitution from X into X and $w \in \overline{X}$, then $\lambda \in S_X(w)$.

(b) If λ is a level substitution from X into X and $w \in X$, then $\ell(w) = \ell(w\lambda)$.

(c) For each level substitution λ from X into X, there exists a level substitution λ^{-1} from X into X such that $\lambda\lambda^{-1} = \lambda^{-1}\lambda$ is the identity substitution from X into X.

(d) If $w \in X$, there exists a level substitution ν from X into X such that $w\nu$ is in normal form, and this normal form is unique.

(e) If $w \in \overline{X}$, $\mathrm{card}(\mathrm{Supp}(w)) = c$, and w is in normal form, then $\mathrm{Supp}(w) = \{x_1, x_2, \cdots, x_c\}$, where $\mathrm{card}(S)$ denotes the cardinality of the set S.

Proof. (a) Since λ is level, it is a product of elementary level substitutions. Each of these is obviously c-free for any freely reduced word, and since a product of substitutions which are c-free for any reduced word is c-free for any reduced word, λ is c-free for any reduced word.

(b) This result is obvious for any elementary level substitution and follows for any level substitution, λ, by induction on the number of elementary level substitutions of which λ is a product.

(c) Clearly $\lambda_i^{-1} = \lambda_i$ and $\lambda_{i,j}^{-1} = \lambda_{i,j}$. If $\lambda = \mu_1\mu_2 \cdots \mu_n$, where each μ_i is an elementary level substitution, then $\lambda^{-1} = \mu_n^{-1}\mu_{n-1}^{-1} \cdots \mu_1^{-1}$.

(d) The lexicographic ordering of the set X is a well ordering; thus the set $\{w\lambda : \lambda$ is a level substitution from X into $X\}$ has a least element, say $w\nu$. By Definition 1.4 (c), $w\nu$ is in normal form. Suppose that ν_1 and ν_2 were level substitutions from X into X such that both $w\nu_1$ and $w\nu_2$ were in normal form with $w\nu_1 \neq w\nu_2$. It would follow that

$w\nu_1 \le w\nu_2$ and $w\nu_2 \le w\nu_1$ by the definition of normal form applied to $w\nu_1$ and $w\nu_2$ respectively. But since "$\le$" is antisymmetric, $w\nu_1 = w\nu_2$, which is a contradiction. Thus the normal form is unique.

(e) Suppose, by way of contradiction, that $x_i \notin \mathrm{Supp}(w)$ for some $i (1 \le i \le c)$; thus for some j $(j > i)$ we have $x_j \in \mathrm{Supp}(w)$. It is clear that the word $w\lambda_{i,j}$ precedes w in the lexicographic ordering, since these two words are the same up to the leftmost occurrance of x_j^ε in w and at this point $w\lambda_{i,j}$ contains x_i^ε and w contains x_j^ε . Since $\lambda_{i,j}$ is a level substitution and $w\lambda_{i,j}$ precedes w in the lexicographic ordering, we have a contradiction to the hypothesis that w was in normal form. ∎

Definition 1.6. If $W \subseteq X$, the normalization of W , denoted $N(W)$, is the set $\{w\nu : w \in W$ and ν is a level substitution which brings w to its unique normal form$\}$.

Definition 1.7. For $v, w \in \overline{X}$, we say v is less than or equal to w , denoted $v \lesssim w$, if $v\sigma = w$ for some $\sigma \in S_X(v)$; with equality, denoted $v \simeq w$, if and only if some such σ is level.

Lemma 1.8. (a) The relation "$\simeq$" is an equivalence relation on $\overline{X}$.

(b) The relation "$\lesssim$" is a partial order relation on $\overline{X}/\simeq$.

(c) If both v and w are in normal form and $v \simeq w$, then $v = w$.

(d) If $w_1 \gtrsim w_2 \gtrsim \cdots \gtrsim w_k \gtrsim \cdots$ is a descending chain of words in $\overline{X}$, then for some $n \geq 1$,
$w_n \simeq w_{n+1} \simeq \cdots \simeq w_{n+i} \simeq \cdots$.

Proof. (a) The reflexive property follows from the fact that the identity substitution is level, the symmetric property follows by Lemma 1.5 (c), and the transitive property follows since the product of two level substitutions is clearly level (see Definition 1.4 (b)).

(b) The reflexive property follows since the identity substitution is c-free for any reduced word. To verify the antisymmetric property, we consider the words $v, w \in \overline{X}$ with $v \lesssim w$ and $w \lesssim v$, and show that $v \simeq w$.

Since $v \lesssim w$ and $w \lesssim v$, there exist $\sigma \in S_X(v)$ and $\tau \in S_X(w)$ such that $v\sigma = w$ and $w\tau = v$. Being c-free, σ and τ do not decrease length, thus $\ell(v) \leq \ell(v\sigma) = \ell(w) \leq \ell(w\tau) = \ell(v)$. Let us write $v = x_{i(1)}^{\epsilon(1)} x_{i(2)}^{\epsilon(2)} \cdots x_{i(n)}^{\epsilon(n)}$ and $w = x_{j(1)}^{\eta(1)} x_{j(2)}^{\eta(2)} \cdots x_{j(n)}^{\eta(n)}$ where $n = \ell(v) = \ell(w)$. Since $v\sigma = w$ we can also write $w = w_1 \cdot w_2$ where $w_1 = x_{i(1)}^{\epsilon(1)}\sigma$ and $w_2 = (x_{i(2)}^{\epsilon(2)} x_{i(3)}^{\epsilon(3)} \cdots x_{i(n)}^{\epsilon(n)})\sigma$. If $\ell(w_1) = 0$, then $w_1 = 1$ which contradicts the fact that σ is c-free (and hence sends no x_s to the empty word). Thus $\ell(w_1) \geq 1$ and so $\ell(w_2) \leq n-1$.

Since σ is c-free for v, it is c-free for any segment of v; in particular σ is c-free for $x_{i(2)}^{\varepsilon(2)} x_{i(3)}^{\varepsilon(3)} \cdots x_{i(n)}^{\varepsilon(n)}$. Therefore $n - 1 = \ell(x_{i(2)}^{\varepsilon(2)} x_{i(3)}^{\varepsilon(3)} \cdots x_{i(n)}^{\varepsilon(n)}) \leq$ $\ell((x_{i(2)}^{\varepsilon(2)} x_{i(3)}^{\varepsilon(3)} \cdots x_{i(n)}^{\varepsilon(n)})\sigma) = \ell(w_2)$, and so $\ell(w_2) = n - 1$ and $\ell(w_1) = 1$. As a result, we have that $x_{i(1)}^{\varepsilon(1)}\sigma = x_{j(1)}^{\eta(1)}$, $(x_{i(2)}^{\varepsilon(2)} x_{i(3)}^{\varepsilon(3)} \cdots x_{i(n)}^{\varepsilon(n)})\sigma = x_{j(2)}^{\eta(2)} x_{j(3)}^{\eta(3)} \cdots x_{j(n)}^{\eta(n)}$, and, applying the same argument using τ instead of σ, that $x_{j(1)}^{\eta(1)}\tau = x_{i(1)}^{\varepsilon(1)}$ and $(x_{j(2)}^{\eta(2)} x_{j(3)}^{\eta(3)} \cdots x_{j(n)}^{\eta(n)})\tau = x_{i(2)}^{\varepsilon(2)} x_{i(3)}^{\varepsilon(3)} \cdots x_{i(n)}^{\varepsilon(n)}$. By a straightforward induction on n, we arrive at the fact that $x_{i(k)}^{\varepsilon(k)}\sigma = x_{j(k)}^{\eta(k)}$ and $x_{j(k)}^{\eta(k)}\tau = x_{i(k)}^{\varepsilon(k)}$ for each $k = 1, 2, \cdots, n$.

Define the mapping $\pi : \{s : x_s \in \mathrm{Supp}(v)\} \to \{k : x_k \in \mathrm{Supp}(w)\}$ by $s\pi = k$ where $x_s\sigma = x_k^{\eta}$ for $\eta \in \{-1, 1\}$. We claim that π is one-to-one. To see this, suppose that $s\pi = t\pi = k$, then $x_s\sigma = x_k^{\alpha}$ and $x_t\sigma = x_k^{\beta}$ for some $\alpha, \beta \in \{-1, 1\}$. Since $x_s, x_t \in \mathrm{Supp}(v)$, there exist positive integers p and q such that $x_s^{\varepsilon(p)} = x_{i(p)}^{\varepsilon(p)}$ and $x_t^{\varepsilon(q)} = x_{i(q)}^{\varepsilon(q)}$. Therefore,

$$\begin{aligned}
x_{j(p)}^{\eta(p)} &= x_{i(p)}^{\varepsilon(p)}\sigma = (x_{i(p)}\sigma)^{\varepsilon(p)} = (x_s\sigma)^{\varepsilon(p)} = \\
(x_k)^{\alpha\varepsilon(p)} &= (x_k^{\beta})^{\alpha\beta\varepsilon(p)} = (x_t\sigma)^{\alpha\beta\varepsilon(p)} = (x_t\sigma)^{\varepsilon(q)\alpha\beta\varepsilon(p)\varepsilon(q)} \\
&= (x_t^{\varepsilon(q)}\sigma)^{\alpha\beta\varepsilon(p)\varepsilon(q)} = (x_{i(q)}^{\varepsilon(q)}\sigma)^{\alpha\beta\varepsilon(p)\varepsilon(q)} \\
&= (x_{j(q)}^{\eta(q)})^{\alpha\beta\varepsilon(p)\varepsilon(q)} .
\end{aligned}$$

Thus,

$$x_s^{\varepsilon(p)} = x_{i(p)}^{\varepsilon(p)} = x_{j(p)}^{\eta(p)}\tau = (x_{j(q)}^{\eta(q)}\tau)^{\alpha\beta\varepsilon(p)\varepsilon(q)}$$
$$= (x_{i(q)}^{\varepsilon(q)})^{\alpha\beta\varepsilon(p)\varepsilon(q)} = (x_t^{\varepsilon(q)})^{\alpha\beta\varepsilon(p)\varepsilon(q)} .$$

Since $x_s = x_t^{\pm 1}$, we have $s = t$; it follows that π is one-to-one. Since π is one-to-one, some permutation of the set $\{s : x_s \in \mathrm{Supp}(v) \cup \mathrm{Supp}(w)\}$ sends $i(k)$ to $j(k)$ for each $k = 1, 2, \cdots, n$. Any permutation of a finite set can be written as a product of transpositions, thus there is a level substitution from X into X, λ (a product of elementary level substitutions of the form $\lambda_{i,j}$), which sends $x_{i(k)}$ to $x_{j(k)}$ for each $k = 1, 2, \cdots, n$. Since the substitutions $\lambda_{i,j}$ affect only subscripts and leave exponents unchanged, we have $v\lambda = x_{j(1)}^{\varepsilon(1)} x_{j(2)}^{\varepsilon(2)} \cdots x_{j(n)}^{\varepsilon(n)}$.

If $\varepsilon(k) = \eta(k)$ for each $k = 1, 2, \cdots, n$, then $v\lambda = w$; therefore $v \simeq w$ and we are done. Otherwise, let s be the least integer such that $\varepsilon(s) \neq \eta(s)$ and write $v\lambda\lambda_{j(s)} = x_{j(1)}^{\gamma(1)} x_{j(2)}^{\gamma(2)} \cdots x_{j(n)}^{\gamma(n)}$. Note that $\gamma(s) = \eta(s)$. We will prove by contradiction that if t is the least integer such that $\gamma(t) \neq \eta(t)$, then $t > s$. Since $\gamma(s) = \eta(s)$ it is clear that $t \neq s$. Suppose that $t < s$. If $j(t) \neq j(s)$, then $\lambda_{j(s)}$ sends $x_{j(t)}^{\varepsilon(t)}$ to $x_{j(t)}^{\varepsilon(t)}$ $(= x_{j(t)}^{\gamma(t)})$, therefore $\gamma(t) = \varepsilon(t)$. But since $t < s$ and s was chosen to be the least integer such that $\varepsilon(s) \neq \eta(s)$, we have $\varepsilon(t) = \eta(t)$.

Therefore $\gamma(t) = \eta(t)$, which contradicts the choice of t . Thus we can assume that $j(t) = j(s)$. Since π is one-to-one, we also have $i(t) = i(s)$, so $x_{i(t)}\sigma = x_{i(s)}\sigma$. Thus $x_{j(t)}^{\eta(t)\varepsilon(t)} = x_{i(t)}\sigma = x_{i(s)}\sigma = x_{j(s)}^{\eta(s)\varepsilon(s)}$, and since $j(t) = j(s)$, $\eta(t)\varepsilon(t) = \eta(s)\varepsilon(s)$. But since $\eta(s)\varepsilon(s) = -1$, we have $\eta(t)\varepsilon(t) = -1$ and, as a result, $\varepsilon(t) \neq \eta(t)$. This is a contradiction, thus $t > s$.

Now by induction on n there exists a sequence of level transformations $\lambda_{j(k_1)}, \lambda_{j(k_2)}, \cdots, \lambda_{j(k_m)}$ such that, letting $\mu = \lambda_{j(k_1)}\lambda_{j(k_2)} \cdots \lambda_{j(k_m)}$, we have $v\lambda\mu =$ $x_{j(1)}^{\eta(1)} x_{j(2)}^{\eta(2)} \cdots x_{j(n)}^{\eta(n)} = w$. Thus we have the level substitution $\lambda\mu \in S_X(v)$ such that $v\lambda\mu = w$, and hence $v \simeq w$.

In order to verify the transitive property it suffices to show that if u, v, and w are in $\overline{X}$ with $u \lesssim v$ and $v \lesssim w$, then $u \lesssim w$. If $u \lesssim v$ and $v \lesssim w$, there exist $\sigma \in S_X(u)$ and $\tau \in S_X(v)$ such that $u\sigma = v$ and $v\tau = w$. Clearly $u(\sigma\tau) = (u\sigma)\tau = v\tau = w \in \overline{X}$. Thus, in order to show that $\sigma\tau$ is c-free for u, it suffices to show that for each $x_s \in \mathrm{Supp}(u)$, $x_s\sigma\tau \neq 1$. Since $\sigma \in S_X(u)$, $x_s\sigma \neq 1$; thus there is an $x_t \in \mathrm{Supp}(x_s\sigma)$. Since $\tau \in S_X(v)$, $x_t\tau \neq 1$; thus there is an $x_r \in \mathrm{Supp}(x_t\tau) \subseteq \mathrm{Supp}(x_s\sigma\tau)$. Therefore $x_s\sigma\tau \neq 1$.

(c) If $v(x) \simeq w(x)$, there exist level substitutions λ and λ^{-1} from X into X such that $v\lambda = w$ and $v = w\lambda^{-1}$. Since v is in normal form we know that, in the lexicographic ordering, $v \leq v\lambda$ $(= w)$, and since w is in normal form, we also know that $w \leq w\lambda^{-1}$ $(= v)$. Therefore, we have $v = w$.

(d) Let us suppose that there is an infinite properly descending chain, $w_1 \gtrsim\!\!\!\!/\ w_2 \gtrsim\!\!\!\!/ \cdots \gtrsim\!\!\!\!/\ w_k \gtrsim\!\!\!\!/ \cdots$, of words in $\overline{X}$.

Since $w_k \gtrsim\!\!\!\!/\ w_{k+1}$ implies that $\ell(w_k) \geq \ell(w_{k+1})$ and since $\ell(w_1)$ is finite, there is an integer $i \geq 1$ such that $\ell(w_i) = \ell(w_{i+1}) = \cdots = \ell(w_{i+j}) = \cdots$.

Letting ν_k be a level substitution which brings w_k into normal form and letting $v_k = w_k\nu_k$ for each $k \geq 1$, we note that,

$$v_k \simeq w_k\nu_k \gtrsim\!\!\!\!/\ w_{k+1}\nu_{k+1} \simeq v_{k+1} .$$

Therefore we obtain the chain $v_1 \gtrsim\!\!\!\!/\ v_2 \gtrsim\!\!\!\!/ \cdots \gtrsim\!\!\!\!/\ v_k \gtrsim\!\!\!\!/ \cdots$ where each v_k is a word in $\overline{X}$ which is in normal form and all v_k $(k \geq i)$ have the same length, say n . It is clear that $\mathrm{card}(\mathrm{Supp}(v_k)) \leq n$ for each $k \geq i$; furthermore, since each v_k is in normal form, it follows by Lemma 1.5 (e), that $\mathrm{Supp}(v_k) = \{x_1, x_2, \cdots, x_c\}$ where $c = \mathrm{card}(\mathrm{Supp}(v_k))$. Thus the set of words $\{v_i, v_{i+1}, \ldots\}$ is a subset of the finite set of all words of length n whose support is a subset of $\{x_1, x_2, \cdots, x_n\}$; therefore there exist integers k and j $(i \leq k < j)$ such that

$v_k = v_j$. But if $k < j$, $v_k \underset{\not\simeq}{>} v_j$ and, in particular, $v_k \not\simeq v_j$. This contradicts the fact that $v_k = v_j$ (and hence that $v_k \simeq v_j$) . Therefore, no such chain exists. ∎

CHAPTER 2

THE SUBSTITUTION PROBLEM

Given the free monoid $(X, \cdot)$ and two words $w, u \in \overline{X}$, we pose the following problem.

Problem III. Can it be decided by a finite procedure whether or not there is a substitution σ from X into X such that $\overline{w\sigma} = u$?

This will be called the substitution problem for the pair (w, u). It is easy to see that Problem III is an alternate form of Problems I and II of the Introduction.

Since the condition $\overline{w\sigma} = u$ involves a free reduction, Problem III is really a problem about free groups rather than monoids. In later sections our method of solving Problem III for certain words will centre around the associativity of $(X, \cdot)$, i.e., the fact that $(X, \cdot)$ is a monoid. This will result from the reduction of Problem III, which involves an arbitrary substitution, to problems involving only c-free substitutions. Thus the free group problem will truly become a problem about free monoids.

The following problem is much simpler than Problem III.

Problem IV. Can it be decided by a finite procedure whether or not there is a c-free substitution $\sigma \in S_X(w)$ such that $w\sigma = u$?

This will be called the c-free substitution problem for the pair (w, u), and can be solved as follows.

<u>Solution of Problem IV</u>. If there is a $\sigma \in S_X(w)$ such that $w\sigma = u$, then for each $x_s \in \mathrm{Supp}(w)$, $\ell(x_s\sigma) \leq \ell(u)$ and $\mathrm{Supp}(x_s\sigma) \subseteq \mathrm{Supp}(u)$. Clearly, if there is such a σ, then there is one which sends x_t to itself for each $x_t \notin \mathrm{Supp}(w)$. Since there are only finitely many such substitutions, τ, we can list all the words, $w\tau$. If some $w\tau = u$, the problem is solved by setting $\sigma = \tau$; otherwise, there is no c-free $\sigma \in S_X(w)$ such that $w\sigma = u$. In either case, the problem has been solved by a finite procedure. ∎

As an obvious corollary to the solution of Problem IV, we note that, given a finite set of words $\{w_1, w_2, \cdots, w_n\}$ and another word u, it can be decided by a finite procedure whether or not there is a w_i and a $\sigma \in S_X(w_i)$ such that $w_i\sigma = u$.

In this chapter we will give a procedure for generating a certain set $C(w)$ (see Definition 2.5(b)), of images of a word $w \in \overline{X}$. This set will be shown (Theorem 2.13) to be "complete" in the following sense.

<u>Definition 2.1</u>. <u>A complete set of images</u> of a word $w \in \overline{X}$ is a set $W \subseteq \overline{X}$ such that:

(i) for each $v \in W$ there is a substitution μ from X into X such that $\overline{w\mu} = v$, and (ii) for any substitution σ from X into X for which $\overline{w\sigma} \neq 1$, there is a word $v \in W$ such that $v \lesssim w\sigma$ (see Definition 1.7).

The set $C(w)$ will also be shown (Theorem 2.14) to be "minimal" in the sense that $C(w) = \bigcap_{I \in \mathcal{J}} N(I)$, where $\mathcal{J}$ is the class of all complete sets of images of w and $N(I)$ is the normalization of I (see Definition 1.6).

The central role played by the set $C(w)$ throughout this thesis is illustrated by the following important result.

Theorem 2.2. Given $w \in \overline{X}$ and assuming that $C(w)$ is a complete set of images of w (Theorem 2.13) and that $C(w) = \bigcap_{I \in \mathcal{J}} N(I)$ (Theorem 2.14), it follows that Problem III is solvable for w and arbitrary u's if and only if membership in $C(w)$ is effectively decidable. (i.e. if and only if, given $v \in \overline{X}$, it can be decided by a finite procedure whether or not v is an element of $C(w)$.)

Proof of Theorem 2.2. Given $v \in \overline{X}$, let us suppose that we can effectively decide whether or not there is a substitution, μ, from X into X such that $\overline{w\mu} = v$. If no such μ exists, then by Theorem 2.13 and Definition 2.1(i), $v \notin C(w)$. If v is an image of w, then by the corollary to the solution of Problem IV, given the finite set, $\{w_1, w_2, \cdots, w_n\}$, of all normalized words in $\overline{X}$ of length not exceeding $\ell(v)$, we can effectively list those w_i's for which $w_i \not\lesssim v$ (i.e. $w_i \lesssim v$

but $w_i \not\simeq v$.) Furthermore, by our assumption that Problem III is solvable for w and arbitrary u's , we can effectively list those w_i's from the list above which are images of w . Let V denote the set of all such words. If v is not in normal form, $v \notin C(w)$; thus w.ℓ.o.g. we may assume that v is in normal form.

We claim that $v \in C(w)$ if and only if $V = \emptyset$; since V can be effectively listed, this implies that it can be effectively decided whether or not $v \in C(w)$. First let us suppose that $v \in C(w)$ and $V \neq \emptyset$. Thus there is a word $u \in V$ which is an image of w such that $u \precsim v$. Since $C(w)$ is a complete set of images of w there is a $u' \in C(w)$ such that $u' \lesssim u$; therefore $S = C(w) \setminus \{v\}$ is also a complete set of images of w . But then $S = N(S)$ is a proper subset of $C(w)$ in contradiction to Theorem 2.14. Thus we have $V = \emptyset$. Now let us suppose that $V = \emptyset$ and $v \notin C(w)$. Since $C(w)$ is a complete set of images of w , there is a word $u \in C(w)$ such that $u \lesssim v$. If $u \simeq v$, it follows that $u = v$ by Lemma 1.7(c), but this contradicts the assumption that $v \notin C(w)$. Thus u is a normalized image of w such that $u \precsim v$; therefore $u \in V$ in contradiction to our assumption that $V = \emptyset$.

Suppose that membership in $C(w)$ is effectively decidable and that $u \in \overline{X}$. By Theorem 2.13 and Definition 2.1(ii), in order to determine whether or not there is a substitution σ such that $\overline{w\sigma} = u$, it is enough to check all $\mu \in S_X(v)$, where $v \in C(w)$, to see if $v\mu = u$. If $v \in C(w)$, then v is in normal form, and if $v\mu = u$, then $\ell(v) \leq \ell(u)$. Since there are only finitely many normalized words of length not exceeding $\ell(u)$, there are only finitely many "candidates" for v . Since membership in $C(w)$ is effectively decidable, the set $\{v_1, v_2, \cdots, v_n\}$ of "candidates" for v which lie in $C(w)$ can be found effectively. The solution of Problem III for the pair (w, u) then reduces to the solution of Problem IV for the set $\{v_1, v_2, \cdots, v_n\}$ and the word u . Since this is solvable, Problem III is solvable as well. ∎

Remark. If $C(w)$ is a finite set, Lemma 2.6(d) will imply that the elements of $C(w)$ can be listed by a finite procedure. As a corollary to Theorem 2.2, assuming Lemma 2.6(d), it follows that if $C(w)$ is finite, Problem III is solvable for w and arbitrary u's .

THE SET $C(w)$.

We now proceed with the definition of the set $C(w)$ and the proofs of Theorems 2.13 and 2.14. First we make some preliminary definitions.

Definition 2.3. (a) Let τ_i be the substitution from X into X defined by

$$x_s\tau_i = \begin{cases} x_s & \text{if } s \neq i \\ 1 & \text{if } s = i . \end{cases}$$

By convention, we let τ_0 be the identity substitution.

(b) For each segment $x_i^{\epsilon} x_j^{\eta}$ of w, let $\rho_{(w;\, i,\epsilon;\, j,\eta)}$ be the substitution from X into X defined by

$$x_s \rho_{(w;\, i,\epsilon;\, j,\eta)} = \begin{cases} x_s & \text{if } s \neq i, j \\ x_m^{-1} x_s x_m & \text{if } s = i = j \\ x_m^{\frac{\epsilon-1}{2}} x_i x_m^{\frac{\epsilon+1}{2}} & \text{if } s = i \neq j \\ x_m^{-\frac{\eta+1}{2}} x_j x_m^{\frac{1-\eta}{2}} & \text{if } s = j \neq i \end{cases}$$

where $m = \max\{s\colon\ x_s \in \mathrm{Supp}(w)\} + 1$ and $x_m^0 = 1$. By convention, we let $\rho_{(w;\, 0,0;\, 0,0)}$ be the identity substitution.

<u>Example</u>. Given $w = x_2^{-1} x_3^{-1} x_2 x_3$, we have

$$w\tau_2 = (x_2\tau_2)^{-1} \cdot (x_3\tau_2)^{-1} \cdot (x_2\tau_2) \cdot (x_3\tau_2) = x_3^{-1} x_3 ,$$

and letting $\rho = \rho_{(w;\, 3,-1;\, 2,1)}$, we have

$$\begin{aligned} w\rho &= (x_2\rho)^{-1} \cdot (x_3\rho)^{-1} \cdot (x_2\rho) \cdot (x_3\rho) \\ &= (x_4^{-1}x_2)^{-1} \cdot (x_4^{-1}x_3)^{-1} \cdot (x_4^{-1}x_2) \cdot (x_4^{-1}x_3) \\ &= x_2^{-1} x_4 x_3^{-1} x_4 x_4^{-1} x_2 x_4^{-1} x_3 . \end{aligned}$$

Remark. The substitution τ_i has the effect of deleting each occurrence of the symbols x_i and x_i^{-1} from w. The substitution $\rho_{(w;\, i,\, \epsilon;\, j,\, \eta)}$ has the effect of replacing each occurrence of the symbols x_i, x_i^{-1}, x_j, and x_j^{-1} by words involving x_i and x_m or x_j and x_m and, in particular, replacing the segment $x_i^{\epsilon} x_j^{\eta}$ by $x_i^{\epsilon} x_m x_m^{-1} x_j^{\eta}$.

Notation. If $\alpha = \{j(1), j(2), \cdots, j(t)\}$ is a set of distinct positive integers, we denote the composition $\tau_{j(1)}\tau_{j(2)} \cdots \tau_{j(t)}$ by τ_α, and we let $\tau_\emptyset = \tau_0$.

Definition 2.4. For $W \subseteq \overline{X}$ we define the sets of reduced words $T(W)$ (trivialization), $R(W)$ (replacement), and $M(W)$ (minimization) as follows:

(a) $T(W) = \bigcup_{w \in W} \bigcup_{\alpha} \{\overline{w\tau_\alpha} : \alpha \subseteq \{s : x_s \in \mathrm{Supp}(w)\}\}$.

(b) $R(W) = \bigcup_{w \in W} \bigcup_{(w;\, i,\, \epsilon;\, j,\, \eta)} \{\overline{w\rho}_{(w;\, i,\, \epsilon;\, j,\, \eta)} : x_i^{\epsilon} x_j^{\eta}$ is a segment of w or $i = \epsilon = j = \eta = 0\}$.

(c) $M(W) = \{w \in W :\ w$ is minimal in W with respect to "$\lesssim$"$\}$ (see Definition 1.7).

Examples. If $W = \{x_1 x_2 x_1^{-1} x_3 x_1\}$,

$$T(W) = \{x_1x_2x_1^{-1}x_3x_1,\ x_2x_3,\ x_3x_1,\ x_1x_2,\ x_3,\ x_2,\ x_1,\ 1\}$$

and
$$R(W) = \{x_1x_2x_1^{-1}x_3x_1,\ x_1x_2x_4^{-1}x_1^{-1}x_3x_1x_4,\ x_1x_4x_2x_1^{-1}x_3x_1x_4,$$
$$x_4^{-1}x_1x_2x_1^{-1}x_3x_4^{-1}x_1,\ x_4^{-1}x_1x_2x_1^{-1}x_4x_3x_1\}.$$

If $W = \{x_1^{-1}x_2x_1\}$, $N(W) = \{x_1x_2x_1^{-1}\}$, and if $W = \{x_2x_1\}$, $N(W) = \{x_1x_2\}$.

If $W = \{x_1x_2x_3,\ x_1x_2x_2,\ x_2^{-1}x_2^{-1},\ x_1x_3^{-1}x_1x_2,\ x_3x_3\}$, $M(W) = \{x_2^{-1}x_2^{-1},\ x_3x_3,\ x_1x_2x_3\}$, $N(W) = \{x_1x_1,\ x_1x_2x_2,\ x_1x_2x_3,\ x_1x_2x_1x_3\}$, and $NM(W) = \{x_1x_1,\ x_1x_2x_3\}$.

<u>Remarks</u>. (1) It should be noted that, although we write substitutions on the right (<u>e.g.</u> $w\tau$) , operations on sets of words will always be written on the left (<u>e.g.</u> $T(W)$).

(2) Given any $w \in W$ there exists a $v \in M(W)$ such that $v \lesssim w$. This is clear if $w \in M(W)$. If $w \notin M(W)$, there is a word $w_1 \in W$ such that $w \gtrsim\!\!\!\!/\ w_1$ (ie. such that $w \gtrsim w_1$ but $w \not\sim w_1$) . If $w_1 \in M(W)$, we let $v = w_1$, and we are done; otherwise, we continue the process. Either we produce an infinite descending chain $w \gtrsim\!\!\!\!/\ w_1 \gtrsim\!\!\!\!/\ w_2 \gtrsim\!\!\!\!/\ \cdots \gtrsim\!\!\!\!/\ w_k \gtrsim\!\!\!\!/\ \cdots$ or at some stage we obtain a chain $w \gtrsim\!\!\!\!/\ w_1 \gtrsim\!\!\!\!/\ w_2 \gtrsim\!\!\!\!/\ \cdots \gtrsim\!\!\!\!/\ w_k$ where $w_k \in M(W)$. The first case cannot occur according to Lemma 1.8(d), therefore the latter case must occur for some integer k . Thus we let $v = w_k$.

(3) If W is a finite subset of $\overline{X}$, the sets $T(W)$, $R(W)$, $N(W)$, and $M(W)$ are finite and can be constructed by finite procedures. This is obvious for $T(W)$, $R(W)$ and $N(W)$. The complement of $M(W)$ can be constructed from W by first performing all possible substitutions which are c-free for some word in W and which give words, u , for which

$\mathrm{Supp}(u) \subseteq \mathrm{Supp}(W)$ $(= \bigcup_{w \in W} \mathrm{Supp}(w))$ and $\ell(u) \leq \max\{\ell(v): v \in W\}$. Those elements of W which arise through non-level substitutions of the above type comprise the complement of $M(W)$ in W , thus we can construct $M(W)$.

The "minimal" complete set, $C(w)$, of images of w will be generated by normalizing w and then successively performing the operations R, T, N, and M, repeatedly. This will generate a sequence of sets $C^0(w) = N(w)$, $C^1(w) = M(N(T(R(C^0(w)))))$, $C^2(w) = M(N(T(R(C^1(w)))))$, $\cdots$; $C(w)$ will consist of those words, v , for which there is an integer $n(v)$ such that $v \in \bigcap_{i \geq n(v)} C^i(w)$.

Definition 2.5. If $w \in \overline{X}$ we define the sets of reduced words $C^i(w)$ $(i \geq 0)$ and $C(w)$ (closure) as follows:

(a) $C^0(w) = N\{w\}$ and $C^{i+1}(w) = MNTR(C^i(w))$ for $i \geq 0$.

(b) $C(w) = \bigcup_{j=1}^{\infty} \bigcap_{i=j}^{\infty} C^i(w)$.

Each set $C^i(w)$ $(i \geq 0)$ is finite and can be constructed by the finite process (see Remark 3 following Definition 2.4) of first normalizing w , then replacing, trivializing, normalizing, and minimizing, in that order, i times.

Examples. (1) First we calculate $C(x_2^{-1})$:

$C^0(x_2^{-1}) = N\{x_2^{-1}\} = \{x_1\}$,

$RC^0(x_2^{-1}) = R\{x_1\} = \{x_1\}$,

$TRC^0(x_2^{-1}) = T\{x_1\} = \{x_1, 1\}$,

$NTRC^0(x_2^{-1}) = N\{x_1, 1\} = \{x_1, 1\}$,

$MNTRC^0(x_2^{-1}) = M\{x_1, 1\} = \{x_1, 1\}$; therefore

$C^1(x_2^{-1}) = \{x_1, 1\}$.

It is easy to see that $C^i(x_2^{-1}) = \{x_1, 1\}$ for $i \geq 1$, thus $C(x_2^{-1}) = \{x_1, 1\}$. Since $C(x_2^{-1})$ is a finite set, we apply the corollary to Theorem 2.2 to obtain the fact that Problem III is solvable for (x_2^{-1}, u), where u is arbitrary. Of course, this is obvious since we could choose the substitution, σ , which sends x_2 to u^{-1} and x_s to x_s , for $s \neq 2$. The next Example will give a less obvious result.

(2) We calculate $C(x_1^n)$ (where $x_1^n = \overbrace{x_1 x_1 \cdots x_1}^{n\text{-times}}$) :

$C^0(x_1^n) = N\{x_1^n\} = \{x_1^n\}$,

$RC^0(x_1^n) = R\{x_1^n\} = \{x_1^n, x_2^{-1}x_1^n x_2\}$,

$TRC^0(x_1^n) = \{x_1^n, x_2^{-1}x_1^n x_2, 1\}$,

$NTRC^0(x_1^n) = \{x_1^n, x_1 x_2^n x_1^{-1}, 1\}$, and

$C^1(x_1^n) = MNTRC^0(x_1^n) = \{x_1^n, x_1 x_2^n x_1^{-1}, 1\}$.

$RC^1(x_1^n) = R\{x_1^n, x_1 x_2^n x_1^{-1}, 1\}$

$= \{x_1^n, x_2^{-1}x_1^n x_2, x_1 x_2^n x_1^{-1}, x_1(x_2 x_3^{-1})^n x_1^{-1}, (x_1 x_3^{-1})x_2^n(x_3 x_1^{-1}), 1\}$,

$TRC^1(x_1^n) = \{x_1^n, x_2^{-1}x_1^n x_2, x_1 x_2^n x_1^{-1}, x_2^n, x_1(x_2 x_3^{-1})^n x_1^{-1}, (x_2 x_3^{-1})^n$,

$x_1(x_3^{-1})^n x_1^{-1}, (x_3^{-1})^n, (x_1 x_3^{-1})x_2^n(x_3 x_1^{-1}), x_3^{-1}x_2^n x_3, 1\}$,

$NTRC^1(x_1^n) = \{x_1^n,\ x_1x_2^nx_1^{-1},\ x_1(x_2x_3)^nx_1^{-1},\ (x_1x_2)^n,\ (x_1x_2)x_3^n(x_2^{-1}x_1^{-1}), 1\}$,

and $C^2(x_1^n) = MNTRC^1(x_1^n) = \{x_1^n,\ x_1x_2^nx_1^{-1},\ 1\}$.

It is easy to see that $C^i(x_1^n) = \{x_1^n,\ x_1x_2^nx_1^{-1},\ 1\}$ for each $i \geq 1$, thus $C(x_1^n) = \{x_1^n,\ x_1x_2^nx_1^{-1},\ 1\}$. By the corollary to Theorem 2.2, we obtain the fact that Problem III is solvable for (x_1^n, u) , where $n \geq 1$ and u is arbitrary. Thus it can be decided whether or not a given word, u , is an $n^{\underline{th}}$ power. Although this is well known, it illustrates the way we will apply our results in Chapter 6.

(3) If $w \in \overline{X}$ and there is an $x_s \in Supp(w)$ such that w is linear in x_s , then $C(w) = \{x_1, 1\}$. To see this, suppose that ν brings w into normal form. Clearly $w\nu$ is linear in x_t , where $\{x_t\} = Supp(x_s\nu)$; thus, letting $\alpha = Supp(w\nu) - \{x_t\}$, we have $w\nu\rho_{(w\nu;\,0,\,0;\,0,\,0)}\tau_\alpha = x_t$. Since x_1 is the normal form of x_t , it follows that $x_1 \in NRT(C^0(w))$; and since $x_1 \lesssim u$ for any non empty word $u \in \overline{X}$, $C^1(w) = MNTR(C^0(w)) = \{x_1, 1\}$. It is then clear that $C^i(w) = \{x_1, 1\}$ for $i \geq 1$; therefore $C(w) = \{x_1, 1\}$. Thus Problem III is trivial for w's which are linear in some x_s .

Lemma 2.6. (a) If the word u is produced at any stage in the construction of $C^i(w)$ $(i \geq 0)$, then there exists a word $v \in C(w)$ such that $v \lesssim u$.

(b) If w is a quadratic word in $\overline{X}$, then each word in $C(w)$ is also quadratic.

(c) The set $C(w)$ is finite if and only if there exists an integer $m \geq 0$ such that $C^m(w) = C^{m+1}(w)$; and if this occurs, then $C(w) = C^{m+k}(w)$ for each $k \geq 0$.

(d) If $C(w)$ is known to be a finite set, $C(w)$ can be constructed by a finite procedure.

Proof. (a) Since every word produced in the construction of $C^{i+1}(w)$ appears first in either $C^i(w)$, $R(C^i(w))$, $TR(C^i(w))$, or $NTR(C^i(w))$ for some $i \geq 0$ and since $C^i(w) \subseteq R(C^i(w)) \subseteq TR(C^i(w))$, it suffices to verify this Lemma for words in $TR(C^i(w))$ or $NTR(C^i(w))$ for each $i \geq 0$. Given a word $u \in TR(C^i(w))$ there is a word $v \in NTR(C^i(w))$ such that $v \simeq u$ (choose v to be the normal form of u .) Thus it suffices to verify this Lemma for each word $v_0 \in NTR(C^i(w))$ $(i \geq 0)$.

By the definition of M, there is a word $v_1 \in MNTR(C^i(w)) = C^{i+1}(w)$ such that $v_1 \lesssim v_0$. Since $v_1 \in C^{i+1}(w) \subseteq TR(C^{i+1}(w))$, there is a word $v_1' \in NTR(C^{i+1}(w))$ such that $v_1' \simeq v_1$. By the definition of M, there is a word $v_2 \in MNTR(C^{i+1}(w)) = C^{i+2}(w)$ such that $v_2 \lesssim v_1'$ $(\lesssim v_1)$. Using this argument repeatedly we get a chain of words $v_0 \gtrsim v_1 \gtrsim v_2 \gtrsim \cdots$ with $v_j \in C^{i+j}(w)$ for each $j \geq 1$. From Lemma 1.8(d), we deduce that there is an integer n such that $v_n \simeq v_{n+1} \cdots \simeq v_{n+k} \simeq \cdots$. By

Lemma 1.8(c) we also deduce that $v_n = v_{n+1} = \cdots = v_{n+k} = \cdots$, since each v_j is in normal form. But $v_n = v_{n+k} \in C^{i+n+k}(w)$ for each $k \geq 0$, and so $v_n \in \bigcap_{j \geq i+n} C^j(w)$. This implies that $v_n \in C(w)$ and, by the transitivity of $\lesssim$, we have $v_n \lesssim v_0$.

(b) The words in $C(w)$ are produced by repeated applications of the substitutions τ_i , $\rho_{(v;\, i,\, \epsilon;\, j,\, \eta)}$, and λ (where λ is elementary level) together with free reduction. It is obvious that quadratic words go to quadratic words under trivialization (τ_i), normalization (λ) , and the deletion of trivial relators. The substitution $\rho_{(v;\, i,\, \epsilon;\, j,\, \eta)}$ introduces at most four occurrences of symbols with the subscript m into the quadratic word, v , but, upon free reduction, at least two of these form a trivial relator (which is deleted). Therefore the result of replacement into a quadratic word followed by free reduction is also a quadratic word.

(c) Suppose $C(w) = \{w_1, w_2, \cdots, w_n\}$ is a finite set. By the definition of $C(w)$, for each i $(1 \leq i \leq n)$ there is an integer $m(i)$ such that $w_i \in \bigcap_{j \geq m(i)} C^j(w)$. Let $m = \max_{1 \leq i \leq n} m(i)$; therefore $w_i \in C^m(w)$ for each i and, as a result,

$$C(w) \subseteq C^m(w) = MNTR(C^{m-1}(w)) \subseteq NTR(C^{m-1}(w)) .$$

Recall that $C^{m}(w)$ is the set of elements in $NTR(C^{m-1}(w))$ minimal with respect to " $\lesssim$ " . Suppose there were a word $u \in C^{m}(w) - C(w)$; then by Lemma 2.6(a), there is a $v \in C(w)$ such that $v \lesssim u$. But $v \in C(w)$ implies that $v \in NTR(C^{m-1}(w))$, thus u dominates an element of $NTR(C^{m-1}(w))$ (viz. v) . By the minimality of u we have $u \simeq v$ and further by Lemma 1.8(c) since both words are in normal form, we have $u = v$. This contradicts the assumption that $u \notin C(w)$. Thus $C(w) = C^{m}(w)$. In fact, by the same proof, if i is any number for which $C(w) \subseteq NTR(C^{i}(w))$, then $C(w) = C^{i+1}(w)$. Thus $C(w) = C^{m}(w)$ implies that $C(w) = C^{m+1}(w)$; therefore by induction,

$$C(w) = C^{m}(w) = C^{m+1}(w) = \cdots .$$

Now suppose there is an integer m for which $C^{m}(w) = C^{m+1}(w)$; then

$$C^{m+2}(w) = MNTR(C^{m+1}(w)) = MNTR(C^{m}(w)) = C^{m+1}(w) ,$$

and similarly, $C^{m}(w) = C^{m+1}(w) = C^{m+2}(w) = \cdots$. Given $u \in C^{m}(w)$, we have $u \in \bigcap_{j\geq m} C^{j}(w)$; thus $u \in C(w)$. As a result we know that $C^{m}(w) \subseteq C(w)$. Suppose $v \in C(w)$; by the definition of C , there is an integer n such that $v \in \bigcap_{i\geq n} C^{i}(w)$. Choose $n' = \max\{m, n\}$, then since $n' \geq m$, $C^{n'}(w) = C^{m}(w)$ and since $n' \geq n$, $v \in C^{n'}(w)$; therefore $v \in C^{m}(w)$. Hence we have the inclusion $C(w) \subseteq C^{m}(w)$, and thus $C(w) = C^{m}(w)$.

When applied to a finite set, each of R, T, N and M yields a finite set. Thus beginning with the finite set $C^0(w)$ and applying the combined operation, MNTR, m times we produce the finite set $C^m(w)$. In the above case $C^m(w) = C(w)$, thus $C(w)$ is finite.

(d) If $C(w)$ is known to be finite, Lemma 2.6(c) shows that the procedure of successively constructing the sets $C^0(w), C^1(w), C^2(w), \cdots$ must terminate in the sense that for some m, $C^m(w) = C^{m+1}(w)$. When this occurs, we stop the construction, knowing by Lemma 2.6(c), again, that $C(w) = C^m(w)$. ∎

PRELIMINARY LEMMAS.

It remains to be shown that $C(w)$ is a complete set of images for w (Theorem 2.13) and that $C(w) = \bigcap_{I \in \mathcal{J}} N(I)$ where $\mathcal{J}$ is the class of all complete sets of images for w (Theorem 2.14). The following lemmas lead to the proofs of these theorems.

Recall that $d(w)$ is the number of deletions of trivial relators required in the free reduction of w to $\overline{w}$.

Lemma 2.7. Suppose $w \in X$ and σ is a substitution from X into X, then the word $\overline{w}\sigma$ is a partially reduced form of $w\sigma$, and if w is not freely reduced, $d(\overline{w}\sigma) < d(w\sigma)$.

Proof. We induct on $d = d(w)$. If $d = 0$, then $w = \overline{w}$ and thus $\overline{w}\sigma = w\sigma$. Since $w\sigma$ is a partially reduced form of itself, this case is complete.

If $d(w) = d + 1 > 0$, then w contains a trivial relator. Let us write $w = v_1 x_i^{\epsilon} x_i^{-\epsilon} v_2$ where $\epsilon \in \{-1, +1\}$ and v_1 is a freely reduced initial segment of w not ending in the symbol $x_i^{-\epsilon}$ (i.e. $x_i^{\epsilon} x_i^{-\epsilon}$ is the leftmost trivial relator in w .)

Since $w\sigma = (v_1\sigma) \cdot (x_i\sigma)^{\epsilon} \cdot (x_i\sigma)^{-\epsilon} \cdot (v_2\sigma)$, the word $(v_1\sigma) \cdot (v_2\sigma)$ is clearly a partially reduced form of $w\sigma$, and $(v_1\sigma) \cdot (v_2\sigma) = (v_1v_2)\sigma$. Since $d(v_1v_2) = d$, we can apply the induction hypothesis to the word v_1v_2 and the substitution σ . Thus $\overline{v_1v_2}\sigma$ is a partially reduced form of $(v_1v_2)\sigma$ which in turn is a partially reduced form of $w\sigma$. But since v_1v_2 is a partially reduced form of w , $\overline{v_1v_2} = \overline{w}$; therefore $\overline{w}\sigma$ is a partially reduced form of $w\sigma$. ∎

<u>Lemma 2.8.</u> If $w \in \overline{X}$ and μ is a substitution from X into X , then there exists a reduced substitution σ from X into X (see Definition 1.1(b)) such that $\overline{w\sigma} = \overline{w\mu}$ and $d(w\sigma) \leq d(w\mu)$.

<u>Proof</u>. Let σ be defined by $x_s\sigma = \overline{x_s\mu}$ for each x_s ; clearly σ is a reduced substitution from X into X . It is also clear that $w\sigma$ can be obtained from $w\mu$ by the deletion of those trivial relators occurring within the segments $x_s\mu$. Thus $w\sigma$ is a partially reduced form of $w\mu$. It follows that $\overline{w\sigma} = \overline{w\mu}$ and $d(w\sigma) \leq d(w\mu)$. ∎

Before we proceed, we introduce the following Definition.

Definition 2.9. If $w \in \overline{X}$ and σ is a substitution from X into X, the set of singularities of σ with respect to w, denoted $\text{sing}(\sigma, w)$, is the set

$$\{x_s \in \text{Supp}(w): \overline{x_s\sigma} = 1\} .$$

Lemma 2.10. If $w \in \overline{X}$, $v \in C^n(w)$ for some integer n, and σ is a reduced substitution from X into X such that $\overline{v\sigma} \neq 1$, then there exists a word $v_0 \in C^{n+n_0}(w)$. (for some integer n_0) and a reduced substitution σ_0 from X into X such that $\overline{v_0\sigma_0} = \overline{v\sigma}$, $d(v_0\sigma_0) \leq d(v\sigma)$, and $\text{sing}(\sigma_0, v_0) = \emptyset$.

Proof. If $\text{sing}(\sigma, v) = \emptyset$, choose $n_0 = 0$, $v_0 = v$, and $\sigma_0 = \sigma$; otherwise, let $\emptyset \neq \alpha_1 = \text{sing}(\sigma, v)$ and let $w_1 = \overline{v\tau_{\alpha_1}}$. Note that $\ell(w_1) < \ell(v)$. Since $v \in C^n(w) \subseteq R(C^n(w))$, it follows that $w_1 \in TR(C^n(w))$. It is clear that $\overline{w_1\sigma} = \overline{v\sigma}$, $d(w_1\sigma) \leq d(v\sigma)$, and $\text{sing}(\sigma, w_1) = \emptyset$.

Let ν_1 be a level substitution from X into X which brings w_1 to normal form. If $w_1\nu_1 \in MNTR(C^n(w))$ $(= C^{n+1}(w))$, we let $n_0 = 1$, $v_0 = w_1\nu_1$, and $\sigma_0 = \nu_1^{-1}\sigma$ and the proof is complete. Otherwise we have $w_1\nu_1 \in NTR(C^n(w)) \setminus MNTR(C^n(w))$; thus there exists a word $v_1 \in MNTR(C^n(w))$ $(= C^{n+1}(w))$ such that $v_1 \lesssim w_1 v_1 (\simeq w_1)$ and hence $\ell(v_1) \leq \ell(w_1) < \ell(v)$. Let $\gamma_1 \in S_X(v_1)$ so that $v_1\gamma_1 = w_1$. Therefore $v_1\gamma_1\sigma = w_1\sigma$ and, by Lemma 2.8, there exists a reduced substitution σ_1 from X into X such that $\overline{v_1\sigma_1} = \overline{v_1\gamma_1\sigma}$ and $d(v_1\sigma_1) \leq d(v_1\gamma_1\sigma)$. But since $\overline{v_1\gamma_1\sigma} = \overline{w_1\sigma} = \overline{v\sigma}$, we have $v_1 \in C^{n+1}(w)$ and σ_1 a reduced substitution from X into X such that $\overline{v_1\sigma_1} = \overline{v\sigma}$ and $d(v_1\sigma_1) \leq d(v\sigma)$. If $\mathrm{sing}(\sigma_1, v_1) = \phi$, we choose $n_0 = 1$, $v_0 = v_1$, and $\sigma_0 = \sigma_1$ and the proof is complete. Otherwise we let $\alpha_2 = \mathrm{sing}(\sigma_1, v_1)$ and repeat the above process, thereby obtaining a word $v_2 \in C^{n+2}(w)$ with $\ell(v_2) < \ell(v_1)$ and a reduced substitution σ_2 from X into X such that $\overline{v_2\sigma_2} = \overline{v_1\sigma_1}$ and $d(v_2\sigma_2) \leq d(v_1\sigma_1)$. If $\mathrm{sing}(\sigma_2, v_2) = \phi$, the proof is complete; if not, we continue this process. Since each v_{k+1} that is produced has length strictly less than its predecessor, v_k , the process must terminate after a finite number of steps, say k . Thus we have $v_k \in C^{n+k}(w)$, and a reduced substitution σ_k from X into X such that $\overline{v_k\sigma_k} = \overline{v\sigma}$, $d(v_k\sigma_k) \leq d(v\sigma)$, and $\mathrm{sing}(\sigma_k, v_k) = \phi$. ∎

Note that a reduced substitution is, in general, c-free only for the words, x_s^ϵ $(s \geq 1)$, and a substitution which is c-free for a word w need not be a reduced substitution for those $x_t \notin \mathrm{Supp}(w)$. Even if a reduced substitution sends a word v to a reduced word, the substitution may not be c-free for v , since it may send some $x_s \in \mathrm{Supp}(v)$ to 1 . However, we do have the following.

<u>Corollary 2.11</u>. If $w \in \overline{X}$, $v \in C^n(w)$ for some integer n , and σ is a reduced substitution from X into X with $1 \neq v\sigma \in \overline{X}$, then there exists a word $v_0 \in C^{n+n_0}(w)$, for some integer $n_0 \geq 0$, and a substitution $\sigma_0 \in S_X(v_0)$ such that $v_0\sigma_0 = v\sigma$.

<u>Proof</u>. By Lemma 2.10, we have a $v_0 \in C^{n+n_0}(w)$ and a reduced substitution σ_0 from X into Y such that $\overline{v_0\sigma_0} = \overline{v\sigma}$, $d(v_0\sigma_0) \leq d(v\sigma)$, and $\mathrm{sing}(\sigma_0, v_0) = \emptyset$. Since $d(v\sigma) = 0$, it is clear that $d(v_0\sigma_0) = 0$ and hence that $v_0\sigma_0$ is freely reduced as written. Thus the equation $\overline{v_0\sigma_0} = \overline{v\sigma}$ is the same as $v_0\sigma_0 = v\sigma$. Since no cancellation occurs in $v_0\sigma_0$ and no symbol, $x_s \in \mathrm{Supp}(v_0)$, is sent to 1 by σ_0 , σ_0 is c-free for v_0 .

<u>THE KEY LEMMA</u>.

The following lemma is central in our proof of Theorem 2.13. Starting with a word, w , and a substitution, σ , we will apply Lemma 2.12 repeatedly to obtain a sequence of words, $v_1, v_2, \ldots, v_m$ ($v_i \in C^i(w)$), and substitutions, $\sigma_1, \sigma_2, \cdots, \sigma_m$, such that $d(v_m\sigma_m) = 0$. This will yield a substitution which is c-free for v_m ($\in C^m(w)$) and which sends v_m to $\overline{w\sigma}$, thus verifying the second half of Definition 2.1 for $C(w)$.

Lemma 2.12. If $w \in \overline{X}$, $v_0 \in C^{n_0}(w)$ for some integer $n_0 \geq 0$, and σ_0 is a reduced substitution from X into X such that $\overline{v_0\sigma_0} \neq 1$, then either,

(a) $d(v_0\sigma_0) = 0$

or (b) there exists a word $v_1 \in C^{n_0+1}(w)$ and a reduced substitution σ_1 from X into X such that $\overline{v_1\sigma_1} = \overline{v_0\sigma_0}$ and $d(v_1\sigma_1) < d(v_0\sigma_0)$.

Proof. We will suppose that $d(v_0\sigma_0) \neq 0$ and verify condition (b). Since σ_0 is a reduced substitution and $v_0\sigma_0$ is not freely reduced there must be a segment of v_0 , $x_i^\epsilon x_j^\eta$, such that a trivial relator (w.ℓ.o.g. $x_1x_1^{-1}$) is the junction of the segments $x_i^\epsilon\sigma_0$ and $x_j^\eta\sigma_0$ in $v_0\sigma_0$. We note that the segments $x_i^\epsilon x_j^\eta$ and $x_j^{-\eta} x_i^{-\epsilon}$ may occur more than once in v_0, and so all the junctions of $x_i^\epsilon\sigma_0$ and $x_j^\eta\sigma_0$ (or $x_j^{-\eta}\sigma_0$ and $x_i^{-\epsilon}\sigma_0$) are the trivial relators $x_1x_1^{-1}$ (or $x_1^{-1}x_1$) . Thus there exist possibly empty words u_ℓ and u_r in $\overline{X}$ such that $x_i^\epsilon\sigma_0 = u_\ell x_1$ and $x_j^\eta\sigma_0 = x_1^{-1}u_r$.

In the event that $i = j$, we clearly have $\epsilon = \eta$, for otherwise v_0 is not freely reduced. Also there is a non-empty word $u \in \overline{X}$ such that $x_i^\epsilon\sigma_0 = x_j^\eta\sigma_0 = x_1^{-1} ux_1$. This is true because $i = j$ and $\epsilon = \eta$ imply that $x_i^\epsilon\sigma_0 = x_j^\eta\sigma_0$ and hence $u_\ell x_1 = x_1^{-1}u_r$. Thus, the first symbol of u_ℓ is x_1^{-1} and the last symbol of u_r is x_1. Therefore we can write $u_\ell = x_1^{-1}u_\ell'$ and $u_r = u_r'x_1$. But then $u_\ell x_1 = x_1^{-1}u_r$ implies that $x_1^{-1}u_\ell' x_1 = x_1^{-1}u_r' x_1$ and so $u_\ell' = u_r'$. If $u_\ell' = u_r' = 1$, then $x_i^\epsilon \sigma_0 = u_\ell x_1 = x_1^{-1}u_\ell' x_1 = x_1^{-1}x_1$ and hence σ_0 is not reduced, contrary to our hypothesis; thus $u_\ell' = u_r' \neq 1$.

We will treat as separate cases the possibilities that $i = j$ or $i \neq j$ and in the latter case that either or both of u_ℓ and u_r may be empty. In each case we give a word v_1 and a substitution σ_1 which fulfill the requirements of (b) in this Lemma.

<u>Case 1</u>. Let $i = j$ and abbreviate $\rho_{(v_0;\, i,\, \epsilon;\, i,\, \epsilon)}$ by ρ. Recall that $x_j^\eta = x_i^\epsilon$, and so $x_j^\eta \sigma_0 = x_i^\epsilon \sigma_0 = x_1^{-1}ux_1$ with u a freely reduced non-empty word. Since $v_0 \in C^{n_0}(w)$, it is clear that $\overline{v_0\rho} \in R(C^{n_0}(w)) \subseteq TR(C^{n_0}(w))$. Thus there exists a word $v_1 \in MNTR(C^{n_0}(w))$ $(= C^{n_0+1}(w))$ such that $v_1 \lesssim \overline{v_0\rho}$. This implies that there is a substitution $\mu_0 \in S_X(v_1)$ with $v_1\mu_0 = \overline{v_0\rho}$. By Definition 2.3, we have $\mathrm{Supp}(v_0\rho) \subseteq \mathrm{Supp}(v_0) \cup \{x_m\}$, thus we can define a reduced substitution μ_1 from X into X by

$$x_s\mu_1 = \begin{cases} x_s\sigma_0 & \text{if } s \neq i, m \\ u^\epsilon & \text{if } s = i \\ x_1 & \text{if } s = m . \end{cases}$$

We note that $v_0\rho\mu_1 = v_0\sigma_0$, since $x_s\rho\mu_1 = x_s\sigma_0$ for each $x_s \in \mathrm{Supp}(v_0)$. This is true because if $s \neq i$, $x_s\rho = x_s$ and thus $x_s\rho\mu_1 = x_s\mu_1 = x_s\sigma_0$; and if $s = i$, $x_i\rho = x_m^{-1}x_ix_m$ and thus $x_i\rho\mu_1 = (x_m^{-1}x_ix_m)\mu_1 = (x_m\mu_1)^{-1} \cdot (x_i\mu_1) \cdot (x_m\mu_1) = x_1^{-1}u^{\epsilon}x_1$. We began with $x_i^{\epsilon}\sigma_0 = x_1^{-1}ux_1$, therefore $(x_i\sigma_0)^{\epsilon} = x_1^{-1}ux_1$ and $x_i\sigma_0 = (x_1^{-1}ux_1)^{\epsilon} = x_1^{-1}u^{\epsilon}x_1$. Thus we have $x_i\sigma_0 = x_1^{-1}u^{\epsilon}x_1 = x_i\rho\mu_1$.

By Lemma 2.7, $\overline{v_0\rho}\,\mu_1$ is a partially reduced form of $v_0\rho\mu_1$. But $\overline{v_0\rho}\,\mu_1 = v_1\mu_0\mu_1$, and so $v_1\mu_0\mu_1$ is a partially reduced form of $v_0\rho\mu_1$ $(= v_0\sigma_0)$. Therefore $\overline{v_1\mu_0\mu_1} = \overline{v_0\sigma_0}$, and since $v_0\rho$ was not freely reduced as written, Lemma 2.7 implies that $d(v_1\mu_0\mu_1) < d(v_0\sigma_0)$.

Letting $\mu = \mu_0\mu_1$ we apply Lemma 2.8 to v_1 and μ to get a reduced substitution σ_1 from X into X such that $\overline{v_1\sigma_1} = \overline{v_1\mu}$ and $d(v_1\sigma_1) \leq d(v_1\mu)$. Now $v_1 \in C^{n_0+1}(w)$, σ_1 is a reduced substitution for v_1 into X, $\overline{v_1\sigma_1} = \overline{v_1\mu} = \overline{v_1\mu_0\mu_1} = \overline{v_0\sigma_0}$, and $d(v_1\sigma_1) \leq d(v_1\mu_0\mu_1) < d(v_0\sigma_0)$; therefore, v_1 and σ_1 satisfy the requirements of (b).

Case 2(a). Let $i \neq j$, let u_{ℓ} and u_r be non-empty, and abbreviate the substitution $\rho_{(v_0;\, i,\, \epsilon;\, j,\, \eta)}$ by ρ. The proof here is exactly the same as that of Case 1 except that the substitution μ_1 is defined by

$$x_s\mu_1 = \begin{cases} x_s\sigma_0 & \text{if } s \neq i, j, m \\ u_\ell^\epsilon & \text{if } s = i \\ u_r^\eta & \text{if } s = j \\ x_1 & \text{if } s = m . \end{cases}$$

It remains to verify that $v_0\rho\mu_1 = v_0\sigma_0$ in this case. This will follow from the fact that $x_s\rho\mu_1 = x_s\sigma_0$ for each x_s. We have three possibilities; $s = i$, $s = j$, and $s \neq i, j$. If $s \neq i, j$, then $x_s\rho = x_s$ and so $x_s\rho\mu_1 = x_s\mu_1 = x_s\sigma_0$. If $s = i$ and $\epsilon = 1$, then $x_i\rho = x_i x_m$ and

$$x_i\rho\mu_1 = (x_i x_m)\mu_1 = (x_i\mu_1)\cdot(x_m\mu_1) = u_\ell^\epsilon x_1 = u_\ell x_1 .$$

If $s = i$ and $\epsilon = -1$, then $x_i\rho = x_m^{-1} x_i$ and

$$x_i\rho\mu_1 = (x_m^{-1} x_i)\mu_1 = (x_m\mu_1)^{-1}\cdot(x_i\mu_1) = x_1^{-1} u_\ell^\epsilon = x_1^{-1} u_\ell^{-1} .$$

Recalling that $x_i^\epsilon\sigma_0 = u_\ell x_1$, we have $x_i\sigma_0 = (u_\ell x_1)^\epsilon$ which is identically equal to $u_\ell x_1$ or $x_1^{-1} u_\ell^{-1}$ as ϵ is $+1$ or -1. Thus if $s = i$, $x_s\rho\mu_1 = x_s\sigma_0$. If $s = j$ and $\eta = 1$, then $x_j\rho = x_m^{-1}x_j$ and $x_j\rho\mu_1 = (x_m^{-1}x_j)\mu_1 = (x_m\mu_1)^{-1}\cdot(x_j\mu_1) = x_1^{-1}u_r^\eta = x_1^{-1}u_r$. If $s = j$ and $\eta = -1$, then $x_j\rho = x_jx_m$ and

$$x_j\rho\mu_1 = (x_jx_m)\mu_1 = (x_j\mu_1)\cdot(x_m\mu_1) = u_r^\eta x_1 = u_r^{-1}x_1 .$$

Recalling that $x_j^{\eta}\sigma_0 = x_1^{-1}u_r$, we have $x_j\sigma_0 = (x_1^{-1}u_r)^{\eta}$ which is identically equal to $x_1^{-1}u_r$ or $u_r^{-1}x_1$ as η is $+1$ or -1 . Thus if $s = j$, $x_s\rho\mu_1 = x_s\sigma_0$.

Case 2(b). Let $i \neq j$, let $u_\ell = 1$ and $u_r \neq 1$ and abbreviate the substitution $\rho_{(v_0;\, i,\epsilon;\, j,\eta)}\tau_i$ by ρ . Since $v_0 \in C^{n_0}(w)$, it is clear that $\overline{v_0\rho} \in TR(C^{n_0}(w))$. Thus the same proof used in Case 2(a) will go through with μ_1 defined by

$$x_s\mu_1 = \begin{cases} x_s\sigma_0 & \text{if } s \neq j,\ m \\ u_r^{\eta} & \text{if } s = j \\ x_1 & \text{if } s = m. \end{cases}$$

Case 2(c). Let $i \neq j$, let $u_\ell \neq 1$ and $u_r = 1$, and abbreviate the substitution $\rho_{(v_0;\, i,\epsilon;\, j,\eta)}\tau_j$ by ρ . As in Case 2(b) it is clear that $\overline{v_0\rho} \in TR(C^{n_0}(w))$. Thus the same proof used in Case 2(a) will go through with μ_1 defined by

$$x_s\mu_1 = \begin{cases} x_s\sigma_0 & \text{if } s \neq i,\ m \\ u_\ell^{\epsilon} & \text{if } s = i \\ x_1 & \text{if } s = m. \end{cases}$$

Case 2(d). Let $i \neq j$, let $u_\ell = u_r = 1$, and abbreviate the substitution $\rho_{(v_0;\, i, \epsilon;\, j, \eta)} \tau_i \tau_j$ by ρ. Since $v_0 \in C^{n_0}(w)$, it is clear that $\overline{v_0\rho} \in TR(C^{n_0}(w))$. Again we use the same proof as in Case 2(a) but with μ_1 defined by

$$x_s \mu_1 = \begin{cases} x_s \sigma_0 & \text{if } s \neq m \\ \\ x_1 & \text{if } s = m . \end{cases} \qquad \blacksquare$$

THE COMPLETENESS OF $C(w)$.

We are now in a position to prove the following Theorem.

Theorem 2.13. For each $w \in \overline{X}$, $C(w)$ is a complete set of images for w.

Proof. We must verify that the set $C(w)$ satisfies Definition 2.1(i) and (ii).

(i) Let $u \in C(w)$, then there is a least integer k such that $u \in C^k(w)$. We will prove by induction on k that for each $u \in C^k(w)$ there is a substitution μ from X into X such that $\overline{w\mu} = u$.

If $k = 0$, then $u \in C^0(w) = \{w\nu_0\}$ where ν_0 brings w into normal form. Since $w \in \overline{X}$, we know that $w\nu_0$ is freely reduced; therefore, $u = \overline{w\nu_0}$ and we are done. Now suppose

$u \in C^{k+1}(w) = \mathrm{MNTR}(C^k(w))$, then there is a word $v \in C^k(w)$ and there exist substitutions ρ, τ_α, and ν (any of which may be the identity substitution) such that $u = \overline{\overline{\overline{v\rho}\;\tau_\alpha}\;\nu}$. Here ρ is a replacement defined for v , τ_α is a trivialization (see Definitions 2.3 and 2.4), and ν is the level substitution which brings $\overline{\overline{v\rho}\;\tau_\alpha}$ to normal form. By the induction hypothesis applied to v , there exists a substitution μ_1 from X into X such that $\overline{w\mu_1} = v$. By Lemma 2.7, $\overline{w\mu_1}\,\rho$ is a partially reduced form of $w\mu_1\rho$; therefore, $\overline{v\rho} = \overline{\overline{w\mu_1}\,\rho} = \overline{w\mu_1\rho}$. Again by Lemma 2.7 $\overline{w\mu_1\rho}\,\tau_\alpha$ is a partially reduced form of $w\mu_1\rho\tau_\alpha$ thus, $\overline{\overline{v\rho}\;\tau_\alpha} = \overline{\overline{w\mu_1\rho}\,\tau_\alpha} = \overline{w\mu_1\rho\tau_\alpha}$. Applying Lemma 2.7 once more we see that $\overline{w\mu_1\rho\tau_\alpha}\,\nu$ is a partially reduced form of $w\mu_1\rho\tau_\alpha\nu$; therefore $u = \overline{\overline{\overline{v\rho}\;\tau_\alpha}\;\nu} = \overline{\overline{w\mu_1\rho\tau_\alpha}\;\nu} = \overline{w\mu_1\rho\tau_\alpha\;\nu}$. Thus $\mu = \mu_1\rho\;\tau_\alpha\nu$ is a substitution from X into X such that $\overline{w\mu} = u$.

(ii) Let μ be a substitution from X into X such that $\overline{w\mu} \neq 1$. We may assume w.ℓ.o.g. that w is in normal form since the following proof could be given beginning with $w\nu$ (where ν brings w to normal form) and $\nu^{-1}\mu$ instead of w and μ . By Lemma 2.8, there exists a reduced substitution σ from X into X such that $\overline{w\sigma} = \overline{w\mu}$ $(\neq 1)$

and $d(w\sigma) \leq d(w\mu)$. Since we have assumed that w is in normal form we have $w \in C^0(w)$. We induct on $d(w\sigma)$ to obtain the final result.

If $d(w\sigma) = 0$, by Corollary 2.11, there exists a word $v \in C^n(w)$ for some integer $n \geq 0$ and a substitution $\tau \in S_X(v)$ such that $v\tau = w\sigma$. But $w\sigma = \overline{w\sigma} = \overline{w\mu}$, and so $v\tau = \overline{w\mu}$ and the proof is complete.

If $d(w\sigma) > 0$, we apply Lemma 2.12 to w and σ . Since $d(w\sigma) \neq 0$, there exists a word $v_1 \in C^1(w)$ and a reduced substitution σ_1 from X into X such that $\overline{v_1\sigma_1} = \overline{w\sigma}$ and $d(v_1\sigma_1) < d(w\sigma)$. If $d(v_1\sigma_1) > 0$ we can apply Lemma 2.12 to v_1 and σ_1 . By repeated applications of Lemma 2.12, we arrive at a word $v_m \in C^m(w)$ and a reduced substitution σ_m from X into X such that $\overline{v_m\sigma_m} = \overline{w\sigma}$ and $d(v_m\sigma_m) = 0$. Thus, by Corollary 2.11, there is a word $v \in C^n(w)$ for some integer $n \geq m$ and a substitution $\tau \in S_X(v)$ such that $v\tau = \overline{v_m\sigma_m} = \overline{w\sigma} = \overline{w\mu}$.

Since $v \in C^n(w)$, from Lemma 2.6(a) we see that there exists a word $u \in C(w)$ such that $u \lesssim v$. Thus there is a substitution $\gamma \in S_X(u)$ such that $u\gamma = v$. The substitution $\gamma\tau$ is the composition of two c-free substitutions and hence is c-free. Thus we have $u \in C(w)$ and $\gamma\tau \in S_X(u)$ such that $u\gamma\tau = v\tau = \overline{w\mu}$, as required. ∎

CHARACTERIZATION OF $C(w)$ FROM ABOVE.

Theorem 2.14. If $w \in \overline{X}$, then $C(w) = \bigcap_{I \in \mathcal{J}} N(I)$ where $\mathcal{J}$ is the class of all complete sets of images for w.

Proof. Since each element of $C(w)$ is in normal form, we have $NC(w) = C(w)$. By Theorem 2.13, $C(w) \in \mathcal{J}$; thus it follows that $\bigcap_{I \in \mathcal{J}} N(I) \subseteq NC(w) = C(w)$. Therefore it suffices to show that the reverse inclusion holds.

Let $I \in \mathcal{J}$ and consider the set $N(I)$; this is defined to be the set of words in I put into normal form. It is easy to see that since $I \in \mathcal{J}$, we also have $N(I) \in \mathcal{J}$. Thus both $C(w)$ and $N(I)$ satisfy Definition 2.1(i) and (ii).

Given any word $u \in C(w)$ we know that there exists a substitution σ from X into X such that $\overline{w\sigma} = u$, by 2.1(i) for $C(w)$. By condition 2.1(ii) for $N(I)$, there is a word $v \in N(I)$ and a substitution $\tau \in S_X(v)$ such that $v\tau = u$, i.e. $v \lesssim u$. By condition 2.1(i) applied to $N(I)$, there is a substitution σ' from X into X such that $\overline{w\sigma'} = v$. Thus, by 2.1(ii) applied to $C(w)$, there is a word $u' \in C(w)$ and a substitution $\tau' \in S_X(u')$ such that $u'\tau' = v$, i.e. $u' \lesssim v$. Thus we have $u' \lesssim v \lesssim u$, and by the minimality of the elements in $C(w)$ it follows that $u' \simeq v \simeq u$. Now by Lemma 1.8(c), we have $v = u$; thus $C(w) \subseteq N(I)$ for each $I \in \mathcal{J}$ and therefore $C(w) \subseteq \bigcap_{I \in \mathcal{J}} N(I)$ as required. ∎

CHAPTER 3

*-GRAPHS

In this chapter we discuss certain diagrams in Euclidean 2-space which we refer to as *-graphs (see Definition 3.2). A procedure, Γ^1 , for "simplifying" these *-graphs, is given and several facts are proved about this procedure. The results of this section will be applied, in Chapter 4, to the c-free solution of certain quadratic verbal equations.

Definition 3.1. Let π be a Euclidean plane. A diagram, $\mathfrak{D}$, in π is a pair (V, E) where V is a finite set of points in π and E is a set of (not necessarily straight) line segments in π whose endpoints form a subset of V . Furthermore, two distinct points $v, v' \in V$ determine at most one line segment, $(v, v') \in E$, joining them. If such a line segment exists, we say that v and v' are adjacent. The sets V and E will be referred to as, respectively, as the vertices and edges of $\mathfrak{D}$. If $\mathfrak{D}' = (V', E')$ is a diagram in π with $V' \subseteq V$ and $E' \subseteq E$, we call $\mathfrak{D}'$ a subdiagram of $\mathfrak{D}$.

Definition 3.2. Let ℓ_a and ℓ_b be two horizontal parallel lines in the plane π with ℓ_a above and ℓ_b below, and let $\{a_1, a_2, \cdots, a_m\}$ and $\{b_1, b_2, \cdots, b_n\}$ be non-empty finite subsets of points on ℓ_a and ℓ_b respectively, linearly ordered from left to right (i.e. $a_i < a_j$ means a_i is to the left of a_j on ℓ_a). A *-graph (star-graph) $\mathcal{G} = (V, E)$ is a diagram in π such that:

(1) $V = \{a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_n\}$,

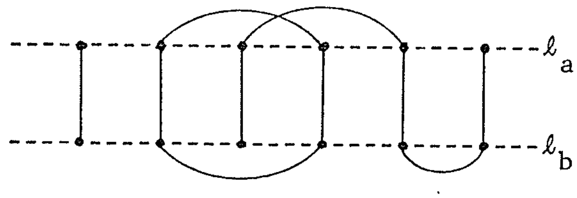
and (2) the set E satisfies (i) and (ii) below.

(i) A vertex on ℓ_a is joined by an edge to at most one other vertex on ℓ_a and at least one other vertex on ℓ_b. Similarly, a vertex on ℓ_b is joined by an edge to at most one other vertex on ℓ_b and at least one other vertex on ℓ_a.

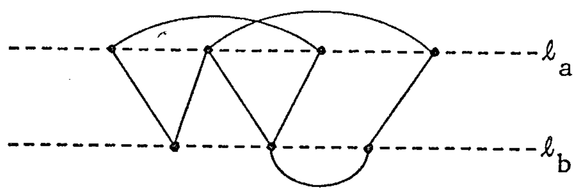
(ii) The edges joining vertices on opposite lines are straight line segments which intersect only at the vertices of $\mathcal{G}$.

Examples. In the following diagrams: $\mathcal{G}_1$ and $\mathcal{G}_2$ are *-graphs, $\mathcal{G}_3$ and $\mathcal{G}_4$ are not *-graphs, since they fail to satisfy property 2(i), and $\mathcal{G}_5$ is not a *-graph since it fails to satisfy property 2(ii).

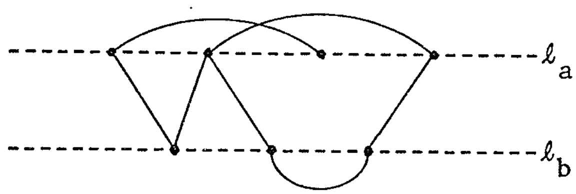
$\mathcal{G}_1$:



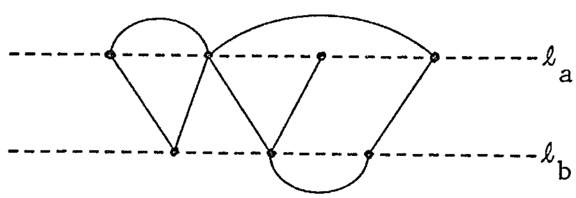
$\mathcal{G}_2$:



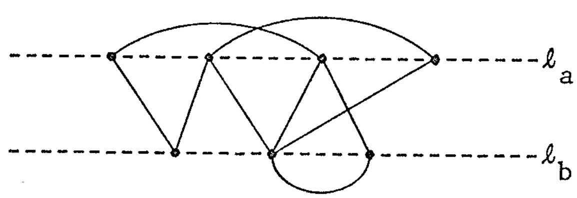
$\mathcal{G}_3$:



$\mathcal{G}_4$:



$\mathcal{G}_5$:



Definition 3.3. If v is a vertex of a *-graph, $\mathcal{G} = (V, E)$, we define the opposite-line degree and the same-line degree of v as follows:

$$\deg_o(v, \mathcal{G}) = \text{card}\{w \text{ on the line opposite to } v \text{ for which } (v, w) \in E\},$$

and

$$\deg_s(v, \mathcal{G}) = \text{card}\{w \text{ on the same line as } v \text{ for which } (v, w) \in E\}.$$

Remarks. When the meaning is clear from context, we will abbreviate $\deg_o(v, \mathcal{G})$ and $\deg_s(v, \mathcal{G})$ by $\deg_o(v)$ and $\deg_s(v)$ respectively. We note that condition 2(i) of Definition 3.2, then takes the form:

$$2(i)': \text{ For each } v \in V, \ \deg_o(v) \geq 1 \text{ and } \deg_s(v) \leq 1 .$$

Definition 3.4. If $\mathcal{G}$ is a *-graph, the deviation of $\mathcal{G}$, denoted $\text{dev}(\mathcal{G})$, is $\sum_{v \in V} (\deg_o(v) - 1)$. If $\text{dev}(\mathcal{G}) = 0$, we say that $\mathcal{G}$ is a simple *-graph. (The *-graph, $\mathcal{G}_1$, in the previous example is simple.)

It is straightforward to verify the following lemma.

Lemma 3.5. (a) $\mathcal{G} = (V, E)$ is simple if and only if $\deg_o(v) = 1$ for each $v \in V$.

(b) If $a_1 < a_2 < \cdots < a_m$ and $b_1 < b_2 < \cdots < b_n$ are respectively the vertices on ℓ_a and ℓ_b of a simple *-graph, $\mathcal{G}$, then $m = n$ and the set of edges joining ℓ_a and ℓ_b in $\mathcal{G}$ is precisely $\{(a_i, b_i): i = 1, 2, \cdots, m\}$.

(c) If v is a vertex of a simple *-graph, $\mathcal{G}$, then v lies on a unique, maximal path in $\mathcal{G}$; this path is simple and may be closed. That is, there exists a maximal set of pairwise adjacent edges of $\mathcal{G}$, $\{(v_1, v_2), (v_2, v_3), \cdots, (v_{q-1}, v_q)\}$, such that $v = v_k$, for some k , and $v_i = v_j$ implies that $i = j$ or (in the event that this path is closed) that $i = 1$ and $j = q$ or $i = q$ and $j = 1$. By "maximal" we mean that any other path containing v is a subpath of this maximal one.

Definition 3.6. If $\mathcal{G} = (V, E)$ is a diagram with vertices V and edges E ,

(a) Removal of a vertex $v \in V$ from $\mathcal{G}$ results in the diagram $\mathcal{G} - v = (V', E')$ where $V' = V \setminus \{v\}$ and $E' = E \setminus \{(v, w) : w \in V\}$.

(b) Removal of an edge $(v, v') \in E$ from $\mathcal{G}$ results in the diagram $\mathcal{G} - (v, v') = (V', E')$ where $V' = V$ and $E' = E \setminus \{(v, v')\}$.

(c) Addition of a vertex v to $\mathcal{G}$ results in the diagram $\mathcal{G} + v = (V', E')$ where $V' = V \cup \{v\}$ and $E' = E$.

(d) Addition of an edge (v, v') to $\mathcal{G}$, providing $v, v' \in V$, results in the diagram $\mathcal{G} + (v, v') = (V', E')$ where $V' = V$ and $E' = E \cup \{(v, v')\}$. (Addition of a set, E_1 , of edges to $\mathcal{G}$ results in $\mathcal{G} + E_1 = (V, E')$ where $E' = E \cup E_1$) .

(e) The sum of the diagrams $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$ is the diagram $\mathcal{G} + \mathcal{H} = (V \cup V', E \cup E')$.

THE *-GRAPH SIMPLIFICATION PROCEDURE.

In the following discussion we will be dealing with *-graphs whose vertices are labelled (not necessarily uniquely) by the symbols $\{x_\alpha^\epsilon : \epsilon \in \{-1, +1\}$ and $\alpha = (i_1, i_2, \cdots, i_t)$ is a finite sequence of positive integers$\}$. At this stage, our requirement for the labelling is that if v and v' are adjacent vertices of $\mathcal{G}$ on the same line (i.e. both on ℓ_a or ℓ_b), then they must be labelled by symbols, x_α^ϵ and x_α^η (indexed by the same sequence). Later we will apply the results of this section to find the c-free solutions of certain quadratic verbal equations in free groups. We will see that, upon "translating" an equation into the context of *-graphs, "simplifying" these *-graphs according to the procedure defined below, and then "translating" back into a verbal setting, we get a procedure for solving the equation in question. Below we define this simplification procedure and prove some useful facts about its effect on a *-graph, labelled as above.

Definition 3.7. Let $\mathcal{G} = (V, E)$ be a *-graph, labelled as above, and let v be a vertex of $\mathcal{G}$. The *-graph simplification procedure, Γ_v^1, is a mapping which sends the labelled *-graph $\mathcal{G}$ to a finite set of labelled *-graphs which we denote by $\Gamma_v^1(\mathcal{G})$ and define below.

Without loss of generality we assume that v lies on ℓ_a, since the procedure is defined symmetrically if v is on ℓ_b (<u>i.e.</u> we interchange x and y, a and b, and α and β in the following definition, to obtain the definition of $\Gamma_v^1(\mathcal{G})$, when v is on ℓ_b). Suppose v is labelled by the symbol x_α^ϵ, where α is a finite sequence of positive integers and $\epsilon \in \{-1, +1\}$. There are two cases to consider.

<u>Case I</u>. $(\deg_s(v) = 0)$. Let $\deg_o(v) = m(\geq 1)$ and let $w_1 < w_2 < \cdots < w_m$ be the m vertices of $\mathcal{G}$ on ℓ_b which are adjacent to v. $\Gamma_v^1(\mathcal{G})$ consists of the set containing the single *-graph defined as follows:

If v is not the leftmost vertex on ℓ_a in $\mathcal{G}$, then there exists a vertex, x, of $\mathcal{G}$ in ℓ_a such that $x < v$, and $y<v$ implies that $y \leq x$. We say that x is <u>immediately to the left</u> of v. We define <u>immediately to the right</u> symmetrically.

To the diagram $\mathcal{G} - v$, we add m vertices, $v_1 < v_2 < \cdots < v_m$, such that $x < v_1$ and $v_m < y$, where x and y are the vertices of $\mathcal{G}$ which are, respectively, immediately to the left and right of v in $\mathcal{G}$. (If v is leftmost or rightmost on ℓ_a, we omit the meaningless condition). According to whether $\epsilon = 1$ or -1 we label $v_1, v_2, \cdots, v_m$ by the symbols $x_{(\alpha, 1)}, x_{(\alpha, 2)}, \cdots, x_{(\alpha, m)}$ or $x^{-1}_{(\alpha, m)}, x^{-1}_{(\alpha, m-1)}, \cdots, x^{-1}_{(\alpha, 1)}$. To this diagram, we add the edges $\{(v_i, w_i) : i = 1, \cdots, m\}$, and call the resulting labelled $*$-graph $\mathcal{G}_1$. We then denote the set $\{\mathcal{G}_1\}$ by $\Gamma_v^1(\mathcal{G})$.

Case II. $(\deg_s(v) = 1)$. Let (v, v') be the edge with v' on ℓ_a, let $\deg_o(v) = m$, $\deg_o(v') = n$, and let $w_1 < w_2 < \cdots < w_m$ and $w_1' < w_2' < \cdots < w_n'$ be the vertices of $\mathcal{G}$ on ℓ_b which are adjacent to v and v' respectively.

To the diagram $(\mathcal{G} - v) - v'$ we add the $2m$ vertices $v_1 < v_2 < \cdots < v_m$ and $v_1' < v_2' < \cdots < v_m'$ such that $x < v_1$, $v_m < y$, $x' < v_1'$ and $v_m' < y'$ where x and y and x' and y' are, respectively, immediately to the left and right of v and v' in $\mathcal{G}$. (If either v or v' is leftmost or rightmost on ℓ_a, we omit the meaningless conditions.) Suppose that v and v' are labelled by x_α^ϵ and x_α^η in $\mathcal{G}$, where $\epsilon, \eta \in \{-1, 1\}$. According to whether $\epsilon = 1$ or -1 we label $v_1, v_2, \cdots, v_m$ by the symbols $x_{(\alpha, 1)}, x_{(\alpha, 2)}, \cdots, x_{(\alpha, m)}$ or $x_{(\alpha, m)}^{-1}, x_{(\alpha, m-1)}^{-1}, \cdots, x_{(\alpha, 1)}^{-1}$, and according to whether $\eta = 1$ or -1 we label $v_1', v_2', \cdots, v_m'$ by the symbols $x_{(\alpha, 1)}, x_{(\alpha, 2)}, \cdots, x_{(\alpha, m)}$ or $x_{(\alpha, m)}^{-1}, x_{(\alpha, m-1)}^{-1}, \cdots, x_{(\alpha, 1)}^{-1}$.

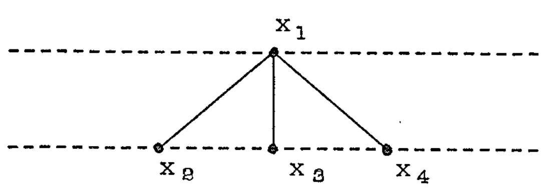
To this diagram we add the edges $\{(v_i, v_i') : i = 1, 2, \cdots, m\}$ or $\{(v_i, v_{m-i+1}') : i = 1, 2, \cdots, m\}$

respectively as $\epsilon = \eta$ or $\epsilon \neq \eta$. Also we add the (straight line) edges $\{(v_i, w_i) : i = 1, 2, \cdots, m\}$ and $\{(v_1', w_1'), (v_m', w_n')\}$ and call the resulting diagram $\mathcal{D}$.

The set, $\{(v_i', w_j') : i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n\}$, is the set of all possible (straight line) edges between the v_i' and w_j' vertices and is clearly finite, thus it has only finitely many subsets. Let $E_1, E_2, \cdots, E_k$ be those subsets such that each diagram $\mathcal{D} + E_i$, is a *-graph. Let $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_k$ denote these *-graphs and let $\Gamma_v^1(\mathcal{G})$ denote the set $\{\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_k\}$.

<u>Examples</u>. Given the following *-graph,

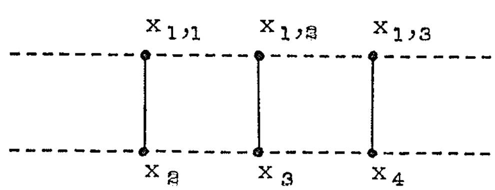
$\mathcal{G}_1$:



,

applying the simplification procedure (Case I) to the upper vertex, we get

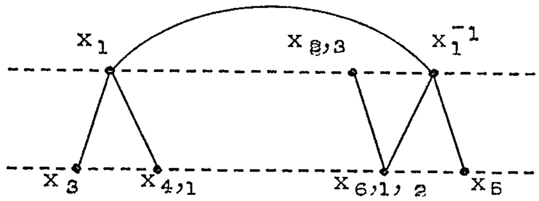
$\mathcal{G}_{(1,1)}$:



.

And, given the *-graph

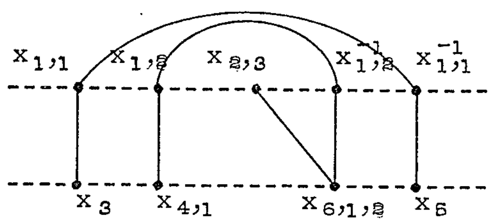
$\mathcal{G}_2$:



,

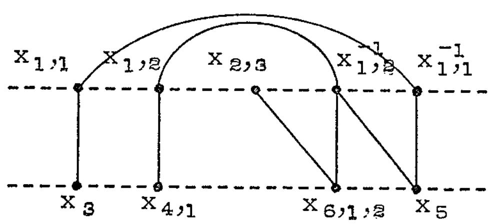
applying the simplification procedure to the upper left vertex, we get,

$\mathcal{G}_{(2,\,1)}$:



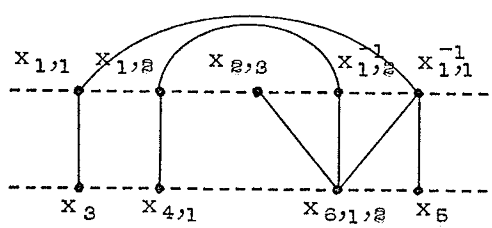
,

$\mathcal{G}_{(2,\,2)}$:



, and

$\mathcal{G}_{(2,\,3)}$:



.

Notation. If $\mathcal{G}$ is a labelled *-graph with vertices V , we will use the notation, $\Gamma^1(\mathcal{G})$, to denote the finite set $\bigcup_{v \in V} \Gamma_v^1(\mathcal{G})$, and if $\{\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_n\}$ is a finite set of labelled *-graphs, we define $\Gamma^1(\{\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_n\})$ to be the finite set $\bigcup_{i=1}^{n} \Gamma^1(\mathcal{G}_i)$. Also we let $\Gamma^n(\mathcal{G}) = \Gamma^1(\Gamma^{n-1}(\mathcal{G}))$ for $n > 1$; note that, for each n , $\Gamma^n(\mathcal{G})$ is a finite set. Finally, we let $\Gamma(\mathcal{G})$ denote the set $\bigcup_{n=1}^{\infty} \Gamma^n(\mathcal{G})$.

The study of what types of *-graphs these sets contain will prove interesting and useful. The labelling of the *-graphs will assume importance only for purposes of application; thus we will ignore the labelling for the rest of this chapter, remembering that since some labelling exists, we can apply the simplification procedure. The following Lemma justifies the use of the word "simplification" to describe the procedure Γ_v^1 .

Lemma 3.8. Let v be a vertex of a *-graph, $\mathcal{G}$.

(a) If $\deg_s(v) = 0$, then

$$\mathrm{dev}(\mathcal{H}) = \mathrm{dev}(\mathcal{G}) - (\deg_o(v) - 1)$$

for each $\mathcal{H} \in \Gamma_v^1(\mathcal{G})$ (in this case there is only one such $\mathcal{H}$.)

(b) If $\deg_s(v) = 1$, and v' is adjacent to v and on the same line as v , then

$$\operatorname{dev}(\mathcal{G}) - f(\deg_o(v), \deg_o(v')) \le \operatorname{dev}(\mathcal{H}) \le \operatorname{dev}(\mathcal{G})$$

for each $\mathcal{H} \in \Gamma_v^1(\mathcal{G})$, where $f(x, y) = (x - 1) + (y - 1) - |(x - 1) - (y - 1)|$. Furthermore, these bounds are best possible.

Note that if $\mathcal{H} \in \Gamma_v^1(\mathcal{G})$, then $\operatorname{dev}(\mathcal{H}) \le \operatorname{dev}(\mathcal{G})$; thus the procedure Γ_v^1 does tend to "simplify" $\mathcal{G}$ in the sense that Γ_v^1 produces *-graphs which are of deviation no larger than that of $\mathcal{G}$.

Proof. (a) Since $\deg_s(v) = 0$, the single *-graph, $\mathcal{H}$, in $\Gamma_v^1(\mathcal{G})$ results from $\mathcal{G}$ by replacing the vertex v by the $m(= \deg_o(v))$ vertices $v_1, v_2, \cdots, v_m$ and the edges $\{(v, w_i) : i = 1, 2, \cdots, m\}$ by the edges $\{(v_i, w_i) : i = 1, 2, \cdots, m\}$. Thus $\deg_o(w, \mathcal{G}) = \deg_o(w, \mathcal{H})$ if w is not v (in $\mathcal{G}$) or some v_i (in $\mathcal{H}$) for $i = 1, 2, \cdots, m$. Therefore, since $\deg_o(v_i, \mathcal{H}) = 1$, it follows that

$$\mathrm{dev}(\mathcal{H}) = \sum_{w \neq v_i} (\deg_o(w, \mathcal{H}) - 1) + \sum_{i=1}^{m} (\deg_o(v_i, \mathcal{H}) - 1)$$

$$= \sum_{w \neq v_i} (\deg_o(w, \mathcal{H}) - 1) = \sum_{w \neq v} (\deg_o(w, \mathcal{G}) - 1)$$

$$= \mathrm{dev}(\mathcal{G}) - (\deg_o(v, \mathcal{G}) - 1) .$$

(b) Let $\mathcal{H}$ be a *-graph in $\Gamma_v^1(\mathcal{G})$ produced by replacing v and v' by $v_1, v_2, \cdots, v_m$ and $v_1', v_2', \cdots, v_m'$ and choosing (according to Definition 3.7, Case II), E' a subset of $\{(v_i', w_j') : i = 1, \cdots, m$ and $j = 1, \cdots, n\}$, where $w_1', w_2', \cdots, w_n'$ are the $n(= \deg_o(v'))$ vertices on ℓ_b which are adjacent to v' in $\mathcal{G}$. Let $\mathcal{K}$ be the subdiagram of $\mathcal{H}$ consisting of the vertices $v_1', v_2', \cdots, v_m'$ and $w_1', w_2', \cdots, w_n'$ with edges $E' \cup \{(v_1', w_1'), (v_m', w_n')\}$.

First we establish the upper bound. It is easy to see that if $m = 1$ or $n = 1$, then $\mathrm{dev}(\mathcal{H}) = \mathrm{dev}(\mathcal{G})$; thus w.ℓ.o.g. we assume that $m, n > 1$. Note that $\deg_o(w, \mathcal{G}) = \deg_o(w, \mathcal{H})$ if w is not v, v', or a w_j' $(j = 1, 2, \cdots, n)$ (in $\mathcal{G}$) or if w is not a v_i, v_i' $(i = 1, 2, \cdots, m)$, or a w_j' $(j = 1, 2, \cdots, n)$ (in $\mathcal{H}$) . Thus we have

(1) $$\sum_{w \neq v,\, v',\, w_j'} (\deg_o(w, \mathcal{G}) - 1) = \sum_{w \neq v_i,\, v_i',\, w_j'} (\deg_o(w, \mathcal{H}) - 1) \ .$$

The following equations also hold:

(2) $\deg_o(v_i, \mathcal{H}) = 1 \qquad (1 \le i \le m)$,

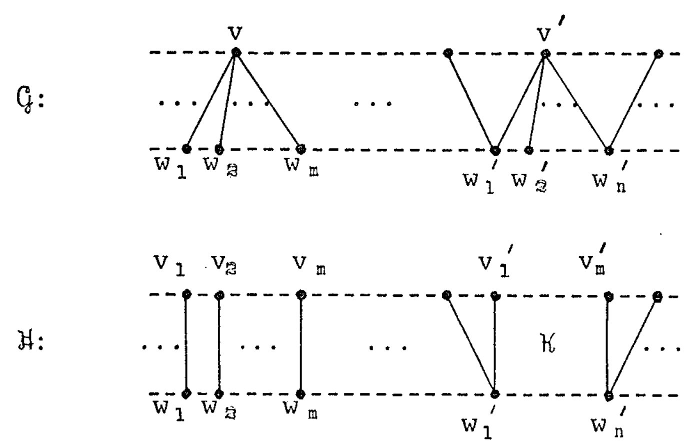
(3) $\deg_o(v_i', \mathcal{H}) = \deg_o(v_i', \mathcal{K}) \qquad (1 \le i \le m)$,

(4) $\deg_o(w_j', \mathcal{H}) = \deg_o(w_j', \mathcal{K}) \qquad (1 < i < n)$,

(5) $\deg_o(w_1', \mathcal{H}) = (\deg_o(w_1', \mathcal{G}) - 1) + \deg_o(w_1', \mathcal{K})$,

(6) $\deg_o(w_n', \mathcal{H}) = (\deg_o(w_n', \mathcal{G}) - 1) + \deg_o(w_n', \mathcal{K})$.

These equalities are easily verified by counting the appropriate edges added to $(\mathcal{G} - \{v\}) - \{v'\}$ during the simplification process (see Case II Definition 3.7). The following diagrams serve to illustrate the situation. (Note that we have drawn only some of the edges between ℓ_a and ℓ_b and none of the edges connecting vertices on the same line.)



Now,

$$\operatorname{dev}(\mathcal{G}) = \sum_{w \neq v, v', w'_j} (\deg_o(w, \mathcal{G}) - 1) + (\deg_o(v, \mathcal{G}) - 1) +$$

$$(\deg_o(v', \mathcal{G}) - 1) + (\deg_o(w'_1, \mathcal{G}) - 1) + (\deg_o(w'_n, \mathcal{G}) - 1) ,$$

since $\deg_o(w'_j, \mathcal{G}) - 1 = 0$ for $1 < j < n$ . .

Thus, by (1), (5), and (6), we have

$$\operatorname{dev}(\mathcal{G}) = \sum_{w \neq v_i, v'_i, w'_j} (\deg_o(w, \mathcal{H}) - 1) + (m - 1) + (n - 1) +$$

$$\deg_o(w'_1, \mathcal{H}) - \deg_o(w'_1, \mathcal{K}) + \deg_o(w'_n, \mathcal{H}) - \deg_o(w'_n, \mathcal{K}) .$$

Now,

$$\operatorname{dev}(\mathcal{H}) = \sum_{w \neq v_i, v'_i, w'_j} (\deg_o(w, \mathcal{H}) - 1) +$$

$$\sum_{i=1}^{m} (\deg_o(v_i, \mathcal{H}) - 1) + \sum_{i=1}^{m} (\deg_o(v'_i, \mathcal{H}) - 1) +$$

$$\sum_{j=1}^{n} (\deg_o(w'_j, \mathcal{H}) - 1) .$$

And, by (2), (3), and (4), we have

$$\operatorname{dev}(\mathcal{H}) = \sum_{w \neq v_i, v'_i, w'_j} (\deg_o(w, \mathcal{H}) - 1) +$$

$$\sum_{i=1}^{m} (\deg_o(v'_i, \mathcal{K}) - 1) + (\deg_o(w'_1, \mathcal{H}) - 1) +$$

$$\sum_{j=2}^{n-1} (\deg_o(w'_j, \mathcal{K}) - 1) + (\deg_o(w'_n, \mathcal{H}) - 1) .$$

Thus,

$$\mathrm{dev}(\mathcal{G}) - \mathrm{dev}(\mathcal{H}) = (m-1) + (n-1) - \deg_o(w_1', \mathcal{K}) + 1 -$$

$$\deg_o(w_n', \mathcal{K}) + 1 - \sum_{i=1}^{m} (\deg_o(v_i', \mathcal{K}) - 1) -$$

$$\sum_{j=2}^{n-1} (\deg_o(w_j', \mathcal{K}) - 1)$$

$$= m + n - \left(\sum_{i=1}^{m} (\deg_o(v_i', \mathcal{K}) - 1) + \sum_{j=1}^{n} (\deg_o(w_j', \mathcal{K}) - 1) \right) - 2$$

$$= (m + n - 2) - \mathrm{dev}(\mathcal{K}) .$$

Therefore it suffices to prove that $\mathrm{dev}(\mathcal{K}) \le m + n - 2$.

We prove that $\mathrm{dev}(\mathcal{K}) \le m + n - 2$ by induction on n . If $n = 1$, the result is clear. If $1 < k < n$, our inductive hypothesis is; given a $*$-graph with k vertices on one line and m vertices on the other, the deviation of such a graph is bounded by $m + k - 2$. We let $\mathcal{K}$ be a $*$-graph, with m vertices on one line and n on the other, which has maximal deviation. Let us call the vertices of $\mathcal{K}$, $v_1', v_2', \cdots, v_m'$ and $w_1', w_2', \cdots, w_n'$.

We claim that there is an i $(1 \le i \le m)$ for which $\deg_o(v_i') \ge 2$. Suppose the claim is not true, then $\deg_o(v_i') = 1$ for each i . Since $n > 1$, we have w_1' and w_2' , distinct vertices, and since $\mathcal{K}$ is a $*$-graph, there exist vertices v_i'

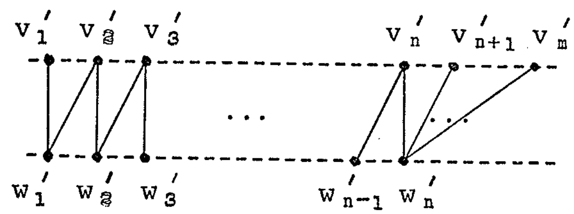
and v'_{i+k} such that (v'_i, w'_1) and (v'_{i+k}, w'_2) are edges of $\mathcal{K}$. If $0 \le q \le k$, then since $\mathcal{K}$ is a *-graph, each vertex v'_{i+q}, is adjacent to w'_1 or w'_2. Thus for some q between 0 and k, we have that (v'_{i+q}, w'_1) and (v'_{i+q+1}, w'_2) are edges of $\mathcal{K}$; but since $\deg_o(v'_i) = 1$ for each i, (v'_{i+q}, w'_2) and (v'_{i+q+1}, w'_1) are not edges of $\mathcal{K}$. The graph $\mathcal{K} + (v'_{i+q}, w'_2)$ is clearly still a *-graph, but $\operatorname{dev}(\mathcal{K} + (v'_{i+q}, w'_2)) = \operatorname{dev}(\mathcal{K}) + 2 > \operatorname{dev}(\mathcal{K})$. This contradicts the maximality of $\operatorname{dev}(\mathcal{K})$; therefore the claim is true.

Let w'_j and w'_{j+1} be adjacent to v'_i. Let $\mathcal{P}$ be the subdiagram of $\mathcal{K}$ with vertices $v'_1, v'_2, \cdots, v'_i$ and $w'_1, w'_2, \cdots, w'_j$ and with all edges of $\mathcal{K}$ which join these vertices; and let $\mathcal{Q}$ be the subdiagram of $\mathcal{K}$ with vertices $v'_i, v'_{i+1}, \cdots, v'_m$ and $w'_{j+1}, w'_{j+2}, \cdots, w'_m$ and with all edges of $\mathcal{K}$ which join these vertices. The deviations of $\mathcal{P}$ and $\mathcal{Q}$ must be maximal since $\mathcal{K}$ is chosen to have maximal deviation. Also $\operatorname{dev}(\mathcal{K}) = \operatorname{dev}(\mathcal{P}) + \operatorname{dev}(\mathcal{Q}) + 1$ where the extra "1" is added since the vertex v'_i in $\mathcal{K}$ will have degree one more than we count in $\mathcal{P}$ and $\mathcal{Q}$ together. By induction, $\operatorname{dev}(\mathcal{P}) = i + j - 2$ and $\operatorname{dev}(\mathcal{Q}) = m - (i - 1) + (n - j) - 2$, so

$$\operatorname{dev}(\mathcal{K}) = (i + j - 2) + (m - (i - 1) + (n - j) - 2) + 1 = m + n - 2 .$$

In order to show that this upper bound is best possible, it suffices to exhibit a *-graph, $\mathcal{K}$, with m points on one line and n points on the other such that $dev(\mathcal{K}) = m + n - 2$. It is easily verified that the following *-graph has this property.

$\mathcal{K}$:



(assuming w.ℓ.o.g. that $m \geq n$)

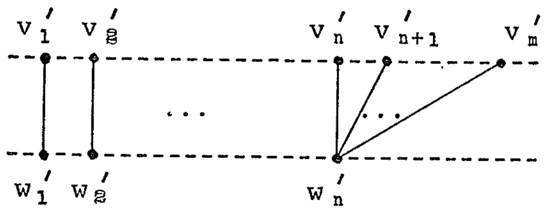
Lastly, we establish the lower bound. Proceeding as in the case of the upper bound, we note that it suffices to show that a *-graph $\mathcal{K}$ with m vertices on one line and n on the other must have $dev(\mathcal{K}) \geq (m + n - 2) - f(m, n) = |m - n|$. We assume w.ℓ.o.g. that $m \geq n$, thus $|m - n| = m - n$. Since $\mathcal{K}$ is a *-graph, each of the v_i' vertices is adjacent to at least one of the w_j' vertices, thus $\sum_{j=1}^{n} deg_o(w_j') \geq m$. Therefore,

$$dev(\mathcal{K}) = \sum_{i=1}^{m} (deg_o(v_i') - 1) + \sum_{j=1}^{n} (deg_o(w_j') - 1) \geq$$

$$\sum_{j=1}^{n} (deg_o(w_j') - 1) = (\sum_{j=1}^{n} deg_o(w_j')) - n \geq m - n = |m - n| .$$

This bound is best possible, as illustrated by the fact that the following $*$-graph has deviation $m - n$ (where $m \geq n$ w.ℓ.o.g.).

$\mathcal{H}$:



∎

<u>Definition 3.9</u>. (a) If v is a vertex of a $*$-graph $\mathcal{G}$ for which $\deg_s(v) = 0$, we will call v an <u>end</u> of $\mathcal{G}$.

(b) We denote the number of ends of $\mathcal{G}$ by ends($\mathcal{G}$).

<u>Lemma 3.10</u>. If v is a vertex of a $*$-graph, $\mathcal{G}$, and $\mathcal{H} \in \Gamma_v^1(\mathcal{G})$, then $\mathrm{dev}(\mathcal{H}) + \mathrm{ends}(\mathcal{H}) \leq \mathrm{dev}(\mathcal{G}) + \mathrm{ends}(\mathcal{G})$, with equality if and only if either v is an end or $\mathrm{dev}(\mathcal{H}) = \mathrm{dev}(\mathcal{G})$.

<u>Proof</u>. If v is an end of $\mathcal{G}$, the number of "new" ends in $\mathcal{H}$ is $\deg_o(v, \mathcal{G}) - 1$. By Lemma 3.8(a), $\mathrm{dev}(\mathcal{H}) = \mathrm{dev}(\mathcal{G}) - (\deg_o(v, \mathcal{G}) - 1)$, and, by Case I of Definition 3.7, $\mathrm{ends}(\mathcal{H}) = \mathrm{ends}(\mathcal{G}) + (\deg_o(v, \mathcal{G}) - 1)$. Therefore,

$$\mathrm{dev}(\mathcal{H}) + \mathrm{ends}(\mathcal{H}) = \mathrm{dev}(\mathcal{G}) - (\deg_o(v, \mathcal{G}) - 1) + \mathrm{ends}(\mathcal{G}) + (\deg_o(v, \mathcal{G}) - 1) = \mathrm{dev}(\mathcal{G}) + \mathrm{ends}(\mathcal{G}) .$$

If v is not an end of $\mathcal{G}$, there is some vertex v' on the same line as v and adjacent to v. Since no "new" ends are added in going from $\mathcal{G}$ to $\mathcal{H}$ by Γ_v^1, we have

$\text{ends}(\mathcal{H}) = \text{ends}(\mathcal{G})$. By Lemma 3.8(b), $\text{dev}(\mathcal{H}) \le \text{dev}(\mathcal{G})$. Thus $\text{dev}(\mathcal{H}) + \text{ends}(\mathcal{H}) \le \text{dev}(\mathcal{G}) + \text{ends}(\mathcal{G})$ with equality if and only if $\text{dev}(\mathcal{H}) = \text{dev}(\mathcal{G})$. ∎

We are now in a position to make the following important observation:

<u>Theorem 3.11</u>. If $\mathcal{G}$ is a $*$-graph and $\mathcal{H}$ is a simple $*$-graph in $\Gamma(\mathcal{G})$, then $\text{ends}(\mathcal{H}) \le \text{dev}(\mathcal{G}) + \text{ends}(\mathcal{G}) \le 2(\text{card}(V) - 1)$, where V is the set of vertices of $\mathcal{G}$.

<u>Proof</u>. Since $\mathcal{H} \in \Gamma(\mathcal{G})$ it follows that $\mathcal{H} \in \Gamma^n(\mathcal{G})$ for some n ; so, by induction on n , using the previous Lemma we have $\text{dev}(\mathcal{H}) + \text{ends}(\mathcal{H}) \le \text{dev}(\mathcal{G}) + \text{ends}(\mathcal{G})$. But since $\mathcal{H}$ is simple, $\text{dev}(\mathcal{H}) = 0$; thus $\text{ends}(\mathcal{H}) \le \text{dev}(\mathcal{G}) + \text{ends}(\mathcal{G})$. If $\mathcal{G}$ is a $*$-graph with m vertices on one line and n vertices on the other, we could have at most $m + n$ ends, and we have already shown, during the proof of Lemma 3.8(b), that $\text{dev}(\mathcal{G}) \le m + n - 2$. Therefore $\text{dev}(\mathcal{G}) + \text{ends}(\mathcal{G}) \le (m + n - 2) + (m + n) = 2(m + n) - 2 = 2(\text{card}(V) - 1)$. ∎

The above Theorem says that when a simple $*$-graph, $\mathcal{H}$, is produced from a $*$-graph, $\mathcal{G}$, by repeated applications of Γ^1 , the number of ends of $\mathcal{H}$ is bounded by a constant associated with the original $*$-graph. When a quadratic verbal equation is "translated" into a set of $*$-graphs and the simplification procedure is applied repeatedly, the simple *-graphs resulting from this process can then be "translated" into solutions of the original equation. Theorem 3.11 will allow us to bound these solutions in some sense (see Corollary 4.12).

CHAPTER 4

THE c-FREE SOLUTIONS TO CERTAIN EQUATIONS

In this chapter we will be interested in the c-free solutions to any quadratic verbal equation, $w_1 \overset{V}{=} w_2$, for which $\mathrm{Supp}(w_1) \cap \mathrm{Supp}(w_2) = \emptyset$. We begin by defining a set, $\mathcal{G}(w_1, w_2)$ (Definition 4.1), of labelled *-graphs associated with the equation, $w_1 \overset{V}{=} w_2$. Applying the simplification procedure repeatedly, we obtain the set $\Gamma(\mathcal{G}(w_1, w_2))$. Next we give a procedure (Definition 4.2) for relabelling the simple (labelled) *-graphs in $\Gamma(\mathcal{G}(w_1, w_2))$. This procedure may produce an "incomplete" relabelling for some simple *-graphs, but for those simple *-graphs, $\mathcal{G} \in \Gamma(\mathcal{G}(w_1, w_2))$, for which the relabelling is complete, we can define a pair, (μ_1, μ_2) , of substitutions said to be "derived" (Definition 4.3) from $\mathcal{G}$. This pair may or may not be a c-free solution to $w_1 \overset{V}{=} w_2$. We define the set $K_X(w_1, w_2)$ (Definition 4.5) to be the set of all those pairs, (μ_1, μ_2) , which are also c-free solutions to the equation, $w_1 \overset{V}{=} w_2$. Our main result (Theorem 4.7) is that $K_X(w_1, w_2)$ is a complete (Definition 4.6) set of c-free solutions in X to $w_1 \overset{V}{=} w_2$.

THE *-GRAPHS OF AN EQUATION.

Throughout this chapter we fix the non-empty reduced quadratic words

$$w_1 = x_{i(1)}^{\varepsilon(1)} x_{i(2)}^{\varepsilon(2)} \cdots x_{i(m)}^{\varepsilon(m)} \quad \text{and} \quad w_2 = x_{j(1)}^{\eta(1)} x_{j(2)}^{\eta(2)} \cdots x_{j(n)}^{\eta(n)} .$$

Definition 4.1. The set of labelled *-graphs of the equation $w_1 \stackrel{\vee}{=} w_2$, denoted $\mathcal{G}(w_1, w_2)$, is the finite set of labelled *-graphs, $\{\mathcal{H}_1, \mathcal{H}_2, \cdots, \mathcal{H}_k\}$ $(k \geq 1)$, defined as follows:

With π, ℓ_a, and ℓ_b as in Chapter 3, we let $\mathcal{H}_0 = (V, \emptyset)$ be a diagram in π with vertices, respectively, $a_1 < a_2 < \cdots < a_m$ and $b_1 < b_2 < \cdots < b_n$. We label a_r by $x_{i(r)}^{\varepsilon(r)}$ $(r = 1, 2, \cdots, m)$ and b_s by $x_{j(s)}^{\eta(s)}$ $(s = 1, 2, \cdots, n)$ and let $\mathcal{H}$ be the diagram in π produced by adding edges to $\mathcal{H}_0$ according to the following rule:

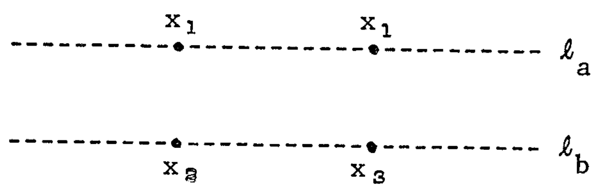
For each $x_{i(r)} = x_{i(s)}$ $(r \neq s)$, add the edge (a_r, a_s), and for each $x_{j(r)} = x_{j(s)}$ $(r \neq s)$, add the edge (b_r, b_s).

Thus, in $\mathcal{H}$, vertices on the same line are adjacent if and only if their labels have the same subscript. Note that, since w_1 and w_2 are quadratic words, $\deg_s(v) \leq 1$ for each $v \in V$ (see Definition 3.3).

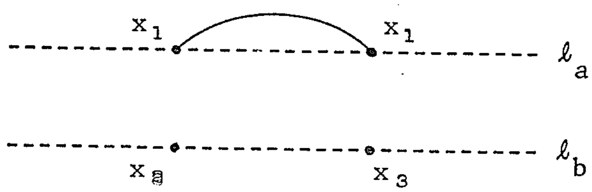
Letting $E = \{(a_i, b_j) : i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, n\}$, we choose those subsets, $E_1, E_2, \cdots, E_k$, of E for which $\mathcal{H} + E_i$ $(i = 1, 2, \cdots, k)$ is a $*$-graph, in accordance with $2(i)'$ and $2(ii)$ of Definition 3.2. Letting $\mathcal{H}_i = \mathcal{H} + E_i$ for each $i = 1, 2, \cdots, k$, we denote the set of labelled $*$-graphs, $\{\mathcal{H}_1, \mathcal{H}_2, \cdots, \mathcal{H}_k\}$, by $\mathcal{G}(w_1, w_2)$.

<u>Example</u>. Given the equation $x_1x_1 \overset{\vee}{=} x_2x_3$, the diagrams mentioned above are as follows:

$\mathcal{H}_0$:



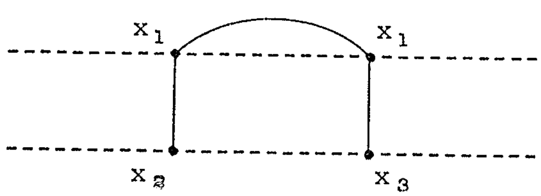
$\mathcal{H}$:



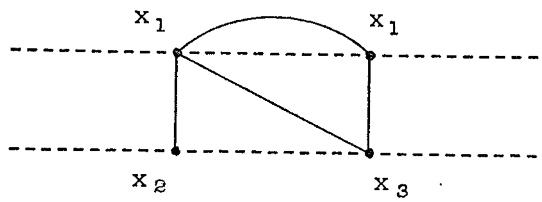
.

$\mathcal{G}(x_1x_1, x_2x_3) = \{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3\}$:

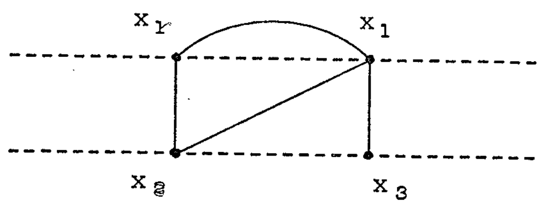
$\mathcal{H}_1$:



$\mathcal{H}_2$:



$\mathcal{H}_3$:



THE RELABELLING PROCEDURE.

Definition 4.2. Below we define a relabelling procedure for any simple labelled $*$-graph $\mathcal{G} \in \Gamma(\mathcal{G}(w_1, w_2))$. Since $\mathcal{G}$ is simple, Lemma 3.5(b) implies that $\mathcal{G}$ has the same number of vertices on ℓ_a and ℓ_b , thus we let the vertices of $\mathcal{G}$ be

$$a_1 < a_2 < \cdots < a_p \quad \text{and} \quad b_1 < b_2 < \cdots < b_p ,$$

respectively labelled by

$$x_{\alpha(1)}^{\gamma(1)} , x_{\alpha(2)}^{\gamma(2)} , \ldots , x_{\alpha(p)}^{\gamma(p)} \quad \text{and} \quad x_{\beta(1)}^{\delta(1)} , x_{\beta(2)}^{\delta(2)} , \ldots x_{\beta(p)}^{\delta(p)} ,$$

where the $\alpha(i)$'s and $\beta(i)$'s are finite sequences of positive integers and $\gamma(i), \delta(i) \in \{-1, +1\}$. We define a mapping θ from the edges of $\mathcal{G}$ into $\{-1, +1\}$ as follows:

$\theta(a_r, a_s) = \gamma(r) \cdot \gamma(s)$, $\theta(b_r, b_s) = \delta(r) \cdot \delta(s)$, and

$\theta(a_r, b_r) = 1$. (Note that (a_r, b_s) $(r \neq s)$ is not an edge of $\mathcal{G}$, since $\mathcal{G}$ is simple.) By Lemma 3.5(c), $\mathcal{G}$ can be viewed as a sum (see Definition 3.6(e)) of disjoint maximal paths, $\mathcal{G} = \mathcal{P}_1 + \mathcal{P}_2 + \cdots + \mathcal{P}_k$. The relabelling procedure will be defined for $\mathcal{G}$ in terms of the relabelling of each path $\mathcal{P}_i$.

We relabel $\mathcal{P}_i$ $(1 \leq i \leq k)$ as follows:

Let a_r be the left most vertex on ℓ_a in the path $\mathcal{P}_i$; relabel a_r by x_i . Now suppose that $v(\neq a_r)$ is a vertex in $\mathcal{P}_i$; if $\mathcal{P}_i$ is not closed it has a unique subpath,

$$\{(v_{1,1}, v_{1,2}), (v_{1,2}, v_{1,3}), \cdots, (v_{1,s-1}, v_{1,s})\} ,$$

where $v_{1,1} = a_r$ and $v_{1,s} = v$. If $\mathcal{P}_i$ is closed it has two subpaths joining a_r and v , in this case denote the second path by

$$\{(v_{2,1}, v_{2,2}), (v_{2,2}, v_{2,3}), \cdots, (v_{2,t-1}, v_{2,t})\} ,$$

where $v_{2,1} = a_r$ and $v_{2,t} = v$. We let

$$e_1 = \theta(v_{1,1}, v_{1,2}) \cdot \theta(v_{1,2}, v_{1,3}) \cdot \cdots \cdot \theta(v_{1,s-1}, v_{1,s})$$

and

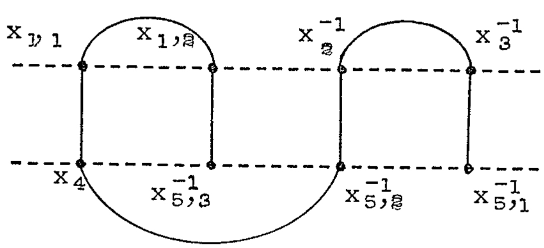
$$e_2 = \theta(v_{2,1}, v_{2,2}) \cdot \theta(v_{2,2}, v_{2,3}) \cdot \cdots \cdot \theta(v_{2,t-1}, v_{2,t}) ,$$

(e_2 is defined only if $\mathcal{P}_i$ is closed.) If $\mathcal{P}_i$ is closed and $e_1 \neq e_2$, the path $\mathcal{P}_i$ cannot be relabelled. If $e_1 = e_2$ or if $\mathcal{P}_i$ is not closed, we relabel v by $x_i^{e_1}$. We continue relabelling the vertices of $\mathcal{P}_i$ until $\mathcal{P}_i$ is completely relabelled or until we come to a vertex, v , for which $e_1 \neq e_2$, in which case we stop.

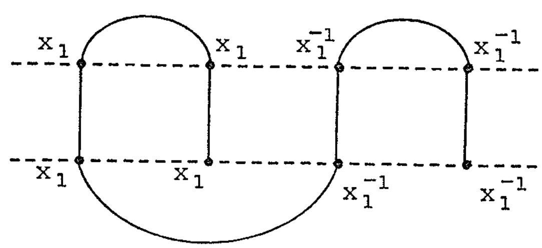
The relabelling of $\mathcal{G}$ is accomplished by relabelling each of the paths $\mathcal{P}_i$ $(1 \leq i \leq k)$. If all of these paths can be relabelled, we say the relabelling procedure _completely_ relabels $\mathcal{G}$; otherwise we say the relabelling procedure is _incomplete_ for $\mathcal{G}$.

Examples. The paths $\mathcal{P}_1$ and $\mathcal{P}_2$ are relabelled as indicated, $\mathcal{P}_3$ cannot be relabelled.

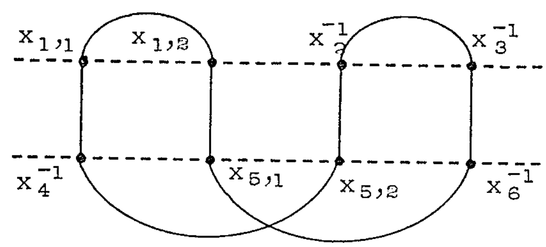
$\mathcal{P}_1$:



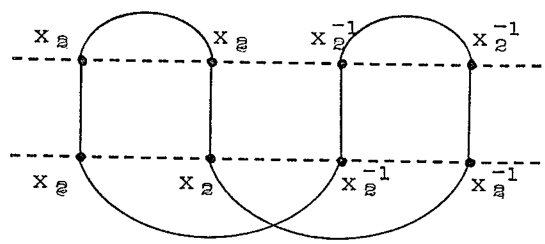
$\mathcal{P}_1$: (relabelled)



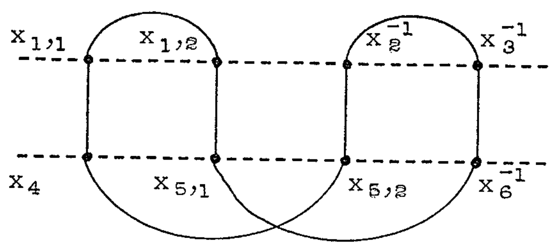
$\mathcal{P}_2$:



$\mathcal{P}_2$: (relabelled)



$\mathcal{P}_3$:



Remarks. (1) If the relabelling procedure is incomplete for a simple $*$-graph, $\mathcal{G}$, then there is a closed maximal path in $\mathcal{G}$,

$$\mathcal{P}_i = \{(v_1, v_2), (v_2, v_3), \cdots, (v_{q-1}, v_q)\} \quad (v_1 = v_q) ,$$

such that

$$\theta(v_1, v_2) \cdot \theta(v_2, v_3) \cdot \cdots \cdot \theta(v_{q-1}, v_q) = -1 .$$

This is clear, since for some vertex v on $\mathcal{P}_i$, $e_1 \neq e_2$; thus $e_1 e_2$ $(= \theta(v_1, v_2) \cdot \theta(v_2, v_3) \cdot \cdots \cdot \theta(v_{q-1}, v_q)) = -1$.

(2) If $\mathcal{G}$ is completely relabelled by the relabelling procedure, then a_t and b_t $(1 \le t \le p)$ are relabelled by the same x-symbol.

(3) The number of distinct x-symbols (not counting inverses) used in the relabelling of $\mathcal{G}$ is precisely k, the number of disjoint maximal paths in $\mathcal{G}$.

<u>THE SET $K_X(w_1, w_2)$</u>.

Let $\mathcal{G}$ be a simple labelled *-graph in $\Gamma(\mathcal{G}(w_1, w_2))$ which is completely relabelled by the relabelling procedure. Suppose the vertices of $\mathcal{G}$,

$$a_1 < a_2 < \cdots < a_p \quad \text{and} \quad b_1 < b_2 < \cdots < b_p ,$$

were originally labelled

$$x_{\alpha(1)}^{\gamma(1)}, x_{\alpha(2)}^{\gamma(2)}, \ldots, x_{\alpha(p)}^{\gamma(p)} \quad \text{and} \quad x_{\beta(1)}^{\delta(1)}, x_{\beta(2)}^{\delta(2)}, \ldots, x_{\beta(p)}^{\delta(p)}$$

and are relabelled

$$x_{\nu(1)}^{e(1)}, x_{\nu(2)}^{e(2)}, \ldots, x_{\nu(p)}^{e(p)} \quad \text{and} \quad x_{\nu(1)}^{e(1)}, x_{\nu(2)}^{e(2)}, \ldots, x_{\nu(p)}^{e(p)},$$

where the $\nu(i)$'s are positive integers and $e(i) \in \{-1, +1\}$. (Note that the new labels for a_r and b_r $(1 \le r \le p)$ are the same by Remark 2 following Definition 4.2.)

Since $\mathcal{G} \in \Gamma(\mathcal{G}(w_1, w_2))$, there is a labelled $*$-graph $\mathcal{G}_0 \in \mathcal{G}(w_1, w_2)$ (see Definition 4.1), with vertices

$$a_1' < a_2' < \cdots < a_m' \quad \text{and} \quad b_1' < b_2' < \cdots < b_n'$$

respectively labelled

$$x_{i(1)}^{\epsilon(1)}, x_{i(2)}^{\epsilon(2)}, \ldots, x_{i(m)}^{\epsilon(m)} \quad \text{and} \quad x_{j(1)}^{\eta(1)}, x_{j(2)}^{\eta(2)}, \ldots, x_{j(n)}^{\eta(n)},$$

and there is a sequence of $*$-graphs, $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_\ell$, such that $\mathcal{G}_{i+1} \in \Gamma^1_{v_i}(\mathcal{G}_i)$ $(0 \le i \le \ell - 1$, v_i a vertex of $\mathcal{G}_i)$ and $\mathcal{G}_\ell = \mathcal{G}$.

We define a map, $\mu_{0,1}$, from the labels of $\mathcal{G}_0$ to sequences of labels of $\mathcal{G}_1$ $(\in \Gamma^1_{v_0}(\mathcal{G}_0))$ as follows:

We will assume w.ℓ.o.g. that $v_0 = a_r'$ (if v_0 is on ℓ_b , the mapping $\mu_{0,1}$ is defined similarly).

1. If $\deg_s(a_r') = 0$ we let

(i) $x_{j(s)}^{\eta(s)} \mu_{0,1} = x_{j(s)}^{\eta(s)}$ (the sequence of length 1) for $1 \le s \le n$

(ii) $x_{i(s)}^{\epsilon(s)} \mu_{0,1} = x_{i(s)}^{\epsilon(s)}$ for $s \ne r$ $(1 \le s \le m)$

(iii) $$x_{i(r)}^{\epsilon(r)} \mu_{0,1} = \begin{cases} x_{(i(r),1)}, x_{(i(r),2)}, \ldots, x_{(i(r),d)} & \text{if } \epsilon(r) = 1 \\[2ex] x^{-1}_{(i(r),d)}, x^{-1}_{(i(r),d-1)}, \ldots, x^{-1}_{(i(r),1)} & \text{if } \epsilon(r) = -1 \end{cases},$$

where $d = \deg_0(a_r', \mathcal{G}_0)$.

Note, by Definition 3.7, Case I, $\mu_{0,1}$ sends a label of $\mathcal{G}_0$ to the label, or sequence of labels, which it corresponds to in $\mathcal{G}_1$.

2. If $\deg_s(a_r') = 1$, say (a_r', a_t') is an edge of $\mathcal{G}_0$, we let $\mu_{0,1}$ be defined as in part 1, except that when $s = t$ we let

$$\text{(iv)} \qquad x_{i(t)}^{\varepsilon(t)} \mu_{0,1} = \begin{cases} x_{(i(t),1)}, x_{(i(t),2)}, \cdots, x_{(i(t),d)} & \text{if } \varepsilon(t) = 1 \\ x^{-1}_{(i(t),d)}, x^{-1}_{(i(t),d-1)}, \cdots, x^{-1}_{(i(t),1)} & \text{if } \varepsilon(t) = -1 . \end{cases}$$

In this case, also, $\mu_{0,1}$ sends a label in $\mathcal{G}_0$ to the corresponding label, or sequence of labels, in $\mathcal{G}_1$ (see Definition 3.7, Case II.)

Similarly, we can define $\mu_{i,i+1}$ from the labels of $\mathcal{G}_i$ to the labels of $\mathcal{G}_{i+1}$ $(\in \Gamma^1_{v_i}(\mathcal{G}_i))$ (<u>i.e.</u> we send each label of $\mathcal{G}_i$ to the corresponding label, or sequence of labels in $\mathcal{G}_{i+1}$.) Thus the composition $\mu_{0,1}\mu_{1,2} \cdots \mu_{\ell-1,\ell}$ sends each label of $\mathcal{G}_0$ to the sequence of labels in $\mathcal{G}_\ell$ which ultimately replace it.

<u>Definition 4.3</u>. If $\mathcal{G}$ (as above) is a simple $*$-graph in $\Gamma(\mathcal{G}(w_1, w_2))$ which is completely relabelled by the relabelling procedure, we define <u>the pair</u> (μ_1, μ_2), <u>derived from</u> $\mathcal{G}$ as follows:

μ_1: If $x_s \in \mathrm{Supp}(w_1)$ then there is a least r $(1 \le r \le m)$ such that $x_s = x_{i(r)}$. We let

$$x_s\mu_1 = (x_{\nu(i)}^{e(i)}\, x_{\nu(i+1)}^{e(i+1)} \cdots x_{\nu(i+j)}^{e(i+j)})^{\epsilon(r)} ,$$

where $x_{\nu(i)}^{e(i)}, x_{\nu(i+1)}^{e(i+1)}, \ldots, x_{\nu(i+j)}^{e(i+j)}$ $(1 \le i < i + j \le p)$ is the sequence of new labels which replaces the sequence

$$x_{i(r)}^{\epsilon(r)}\, \mu_{0,1}\mu_{1,2} \cdots \mu_{\ell-1,\ell} \quad (= x_{\alpha(i)}^{\gamma(i)}, x_{\alpha(i+1)}^{\gamma(i+1)}, \ldots, x_{\alpha(i+j)}^{\gamma(i+j)})$$

when $\mathcal{G}_\ell$ is relabelled.

If $x_s \notin \mathrm{Supp}(w_1)$, let $x_s\mu_1 = x_s$.

μ_2: If $x_s \in \mathrm{Supp}(w_2)$, $x_s\mu_2$ is defined similarly; if $x_s \notin \mathrm{Supp}(w_2)$ we let $x_s\mu_2 = x_s$.

Lemma 4.4. With $\mathcal{G}$ and (μ_1, μ_2) as above, we have

$$w_1\mu_1 = x_{\nu(1)}^{e(1)}\, x_{\nu(2)}^{e(2)} \cdots x_{\nu(p)}^{e(p)} = w_2\mu_2 .$$

Proof. To see the lefthand equality, we first note that if w_1 is quadratic in x_s, say $s = i(r) = i(t)$ $(1 \le r < t \le m)$, then (a_r', a_t') is an edge of $\mathcal{G}_0$ and for some k $(> i + j)$ the vertices $a_k, a_{k+1}, \ldots, a_{k+j}$ are precisely the vertices of $\mathcal{G}_\ell$ which ultimately replace a_t'. Furthermore, the labels

$$x_{\alpha(k)}^{\gamma(k)}, x_{\alpha(k+1)}^{\gamma(k+1)}, \ldots, x_{\alpha(k+j)}^{\gamma(k+j)}$$

equal either

$$x_{\alpha(i)}^{\gamma(i)},\ x_{\alpha(i+1)}^{\gamma(i+1)},\ \ldots,\ x_{\alpha(i+j)}^{\gamma(i+j)} \quad \text{or}$$

$$x_{\alpha(i+j)}^{-\gamma(i+j)},\ x_{\alpha(i+j-1)}^{-\gamma(i+j-1)},\ \ldots,\ x_{\alpha(i)}^{-\gamma(i)}$$

as $\epsilon(r)\epsilon(t) = +1$ or -1. All this follows from Definition 3.7 by a straightforward induction on ℓ (the number of times the simplification procedure was applied to obtain $\mathcal{G}$.) Since either

$$\{(a_{i+q}, a_{k+q}) : 0 \leq q \leq j\} \quad \text{or} \quad \{(a_{i+q}, a_{k+j-q}) : 0 \leq q \leq j\}$$

are edges of $\mathcal{G}$, when $\mathcal{G}$ is relabelled we have

$$(x_{\nu(i)}^{e(i)} x_{\nu(i+1)}^{e(i+1)} \cdots x_{\nu(i+j)}^{e(i+j)})^{\epsilon(r)} = (x_{\nu(k)}^{e(k)} x_{\nu(k+1)}^{e(k+1)} \cdots x_{\nu(k+j)}^{e(k+j)})^{\epsilon(t)}.$$

Thus $x_{i(t)}^{\epsilon(t)}\mu_1 = x_{\nu(k)}^{e(k)} x_{\nu(k+1)}^{e(k+1)} \cdots x_{\nu(k+j)}^{e(k+j)}$. Since μ_1 sends each $x_{i(r)}^{\epsilon(r)}$ in w_1 to the product of the new labels of the vertices in $\mathcal{G}$ which ultimately replace the vertex a_r' in $\mathcal{G}_0$ (even if w_1 is quadratic in $x_{i(r)}$), the lefthand equality follows.

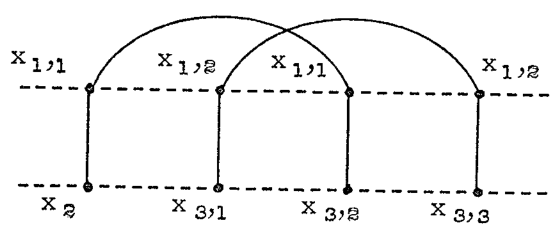
The righthand equality is verified similarly. ∎

<u>Definition 4.5</u>. Letting $S = \{(\mu_1, \mu_2) : (\mu_1, \mu_2)$ is derived from a simple $*$-graph in $\Gamma(\mathcal{G}(w_1, w_2))$ which can be completely relabelled$\}$, we set $K_X(w_1, w_2) = S \cap S_X(w_1, w_2)$ (see Definition 1.3(c)).

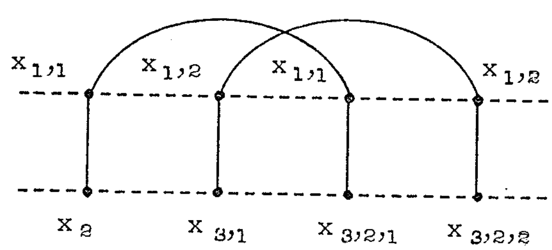
<u>Example</u>. Suppose we wish to find $K_X(x_1x_1, x_2x_3)$. Referring to the example following Definition 4.1, we have $\mathcal{G}(x_1x_1, x_2x_3) = \{\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3\}$, where $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_3$ are given in that example. $\mathcal{H}_1$ is simple, thus when the simplification procedure is applied to $\mathcal{H}_1$, only $\mathcal{H}_1$ results.

When the simplification procedure is applied to $\mathcal{H}_2$, we get the two simple $*$-graphs, $\mathcal{H}_{2,1}$ and $\mathcal{H}_{2,2}$, below:

$\mathcal{H}_{2,1}$:

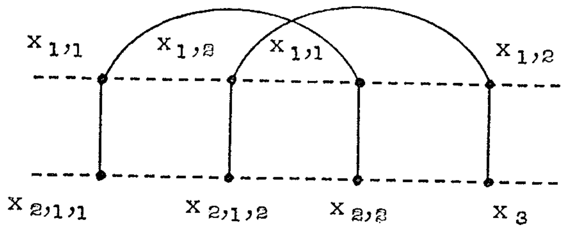


$\mathcal{H}_{2,2}$:

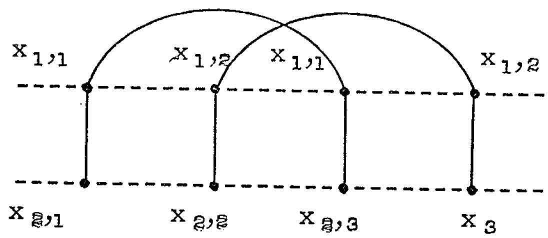


And when the simplification procedure is applied to $\mathcal{H}_3$, we get the two simple $*$-graphs, $\mathcal{H}_{3,1}$ and $\mathcal{H}_{3,2}$, below:

$\mathcal{H}_{3,1}$:



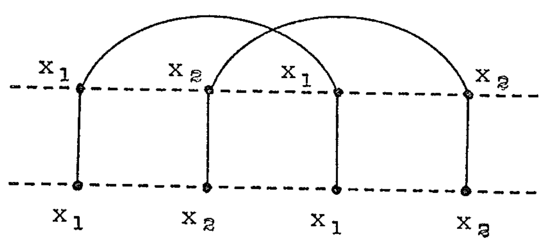
$\mathcal{H}_{3,2}$:



.

Applying the Relabelling Procedure to $\mathcal{H}_{2,1}$, $\mathcal{H}_{2,2}$, $\mathcal{H}_{3,1}$, and $\mathcal{H}_{3,2}$, we get the single simple $*$-graph, $\mathcal{H}$, below (see the Example following Definition 4.2):

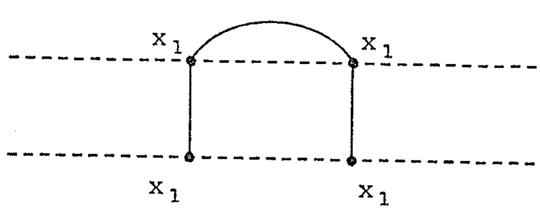
$\mathcal{H}$:



.

And Relabelling $\mathcal{H}_1$, we get $\mathcal{K}$ below:

$\mathcal{K}$:



.

The pair, (μ_1, μ_2) , derived from $\mathcal{H}_1$ is defined (Definition 4.3) as follows:

$$x_s\mu_1 = \begin{cases} x_s & \text{if } s \neq 1 \\ x_1 & \text{if } s = 1 \end{cases}$$

$$x_s\mu_2 = \begin{cases} x_s & \text{if } s \neq 2, 3 \\ x_1 & \text{if } s = 2 \\ x_1 & \text{if } s = 3 \,. \end{cases}$$

The pair (μ_3, μ_4) derived from $\mathcal{H}_{2,1}$ is defined by:

$$x_s\mu_3 = \begin{cases} x_s & \text{if } s \neq 1 \\ x_1x_2 & \text{if } s = 1 \end{cases}$$

$$x_s\mu_4 = \begin{cases} x_s & \text{if } s \neq 2, 3 \\ x_1 & \text{if } s = 2 \\ x_2x_1x_2 & \text{if } s = 3 \,. \end{cases}$$

The same pair, (μ_3, μ_4) , is derived from $\mathcal{H}_{2,2}$.

The pair, (μ_5, μ_6) , derived from $\mathcal{H}_{3,1}$, is defined by:

$$x_s\mu_5 = \begin{cases} x_s & \text{if} \quad s \neq 1 \\ \\ x_1x_2 & \text{if} \quad s = 1 \end{cases}$$

$$x_s\mu_6 = \begin{cases} x_s & \text{if} \quad s \neq 2,\ 3 \\ x_1x_2x_1 & \text{if} \quad x = 2 \\ x_2 & \text{if} \quad s = 3 \end{cases}$$

The same pair is derived from $\mathcal{H}_{3,2}$.

Thus by Definition 4.5,

$$K_X(x_1x_1,\ x_2x_3) = S_X(x_1x_1,\ x_2x_3) \cap \{(\mu_1, \mu_2),\ (\mu_3, \mu_4),\ (\mu_5, \mu_6)\} .$$

Since $(x_1x_1)\mu_1 = x_1x_1 = (x_2x_3)\mu_2$ and μ_1 and μ_2 are c-free for x_1x_1 and x_2x_3 respectively, $(\mu_1, \mu_2) \in S_X(x_1x_1, x_2x_3)$. Similarly, (μ_3, μ_4), $(\mu_5, \mu_6) \in S_X(x_1x_1, x_2x_3)$. Therefore,

$K_X(x_1x_1,\ x_2x_3) = \{(\mu_1, \mu_2),\ (\mu_3, \mu_4),\ (\mu_5, \mu_6)\}$.

Note that in general $K_X(w_1, w_2)$ is not a finite set.

THE MAIN THEOREM.

Before we can state our main result, we need the following definition.

Definition 4.6. If $w_1, w_2 \in \overline{X}$, a subset, K, of $S_X(w_1, w_2)$ (see Definition 1.3(c)) is a complete set of c-free solutions in X to the verbal equation $w_1 \overset{\vee}{=} w_2$, if for each $(\sigma_1, \sigma_2) \in S_X(w_1, w_2)$, there exists a pair $(\mu_1, \mu_2) \in K$ and a c-free substitution $\delta \in S_X(w_1 \mu_1)$ such that:

$$x_s \mu_1 \delta = x_s \sigma_1 \qquad \text{for each } x_s \in \mathrm{Supp}(w_1)\ ,$$

and $$x_t \mu_2 \delta = x_t \sigma_2 \qquad \text{for each } x_t \in \mathrm{Supp}(w_2)\ .$$

(Note that since (μ_1, μ_2) is a c-free solution to $w_1 \overset{\vee}{=} w_2$, we have $w_1 \mu_1 = w_2 \mu_2$; therefore, δ is c-free for $w_2 \mu_2$ as well as $w_1 \mu_1$.)

Our main result in this chapter is the following Theorem.

Theorem 4.7. The set $K_X(w_1, w_2)$ (see Definition 4.5) is a complete set of c-free solutions in X to the verbal equation $w_1 \overset{\vee}{=} w_2$.

The proof of this theorem is rather involved; therefore we will begin with an example of how the proof works.

ABOUT THE PROOF OF THEOREM 4.7.

Though trivial, the equation $x_1x_1 \veeeq x_2x_3$ provides a good example with which to illustrate the idea of the proof of Theorem 4.7. Starting with a specific solution $(\sigma_1, \sigma_2) \in S_X(x_1x_1, x_2x_3)$, we will choose a pair, (μ_i, μ_{i+1}) $(i = 1, 3, \text{ or } 5)$, from $K_X(x_1x_1, x_2x_3)$ and define a substitution, δ, with the required properties (see Definition 4.6).

Suppose (σ_1, σ_2) is the pair defined by:

$$x_s\sigma_1 = \begin{cases} x_s & \text{if } s \neq 1 \\ x_1x_2x_3 & \text{if } s = 1 \end{cases}$$

$$x_s\sigma_2 = \begin{cases} x_s & \text{if } s \neq 2, 3 \\ x_1x_2 & \text{if } s = 2 \\ x_3x_1x_2x_3 & \text{if } s = 3 \, . \end{cases}$$

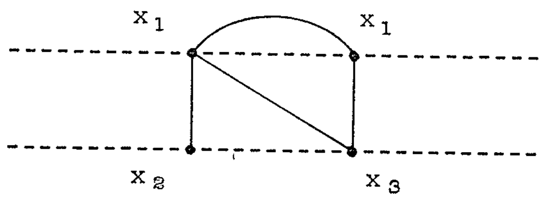
It is easy to check that $(\sigma_1, \sigma_2) \in S_X(x_1x_1, x_2x_3)$. Consider the word, $(x_1x_1)\sigma_1$ $(= (x_2x_3)\sigma_2) = x_1x_2x_3x_1x_2x_3$; in our proof we will call this word w. The two substitutions σ_1 and σ_2 partition w in two different ways (call these partitions $\mathfrak{u}_0$ and $\mathfrak{v}_0$.)

$\mathcal{U}_0 : \ (x_1x_2x_3) \cdot (x_1x_2x_3)$

$\mathcal{V}_0 : \ (x_1x_2) \cdot (x_3x_1x_2x_3)$.

In $\mathcal{G}(x_1x_1, x_2x_3)$, the $*$-graph,

$\mathcal{H}_2$:



,

corresponds to the pair of partitions $(\mathcal{U}_0, \mathcal{V}_0)$ in the sense that the vertices on ℓ_a and ℓ_b in $\mathcal{H}_2$ are in one-to-one correspondence with the segments in $\mathcal{U}_0$ and $\mathcal{V}_0$ respectively and vertices on opposite lines are adjacent if and only if the corresponding segments overlap. In this case we have an edge between the upper left vertex and lower right vertex of $\mathcal{H}_2$ since the lefthand segment, $x_1x_2x_3$, of $\mathcal{U}_0$ overlaps the righthand segment, $x_3x_1x_2x_3$, of $\mathcal{V}_0$.

We will simultaneously refine the partitions $(\mathcal{U}_0, \mathcal{V}_0)$ and simplify $\mathcal{H}_2$ so that there is still this correspondence between them.

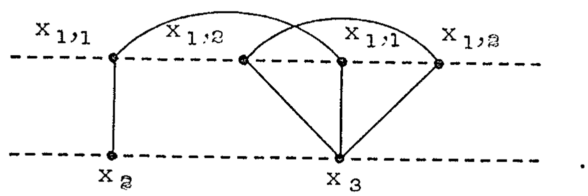
In this case our first refinement of $(\mathcal{U}_0, \mathcal{V}_0)$ results in $(\mathcal{U}_1, \mathcal{V}_1)$

$\mathcal{U}_1$: $(x_1x_2) \cdot (x_3) \cdot (x_1x_2) \cdot (x_3)$

$\mathcal{V}_1$: $(x_1x_2) \cdot (x_3x_1x_2x_3)$.

This is obtained by partitioning the segment, $x_1x_2x_3$, of $\mathcal{U}_0$ into $(x_1x_2) \cdot (x_3)$. The corresponding $*$-graph is $\mathcal{H}_2'$ below.

$\mathcal{H}_2'$:



Note that $\mathcal{H}_2' \in \Gamma^1(\mathcal{H}_2)$ and is produced by applying the simplification procedure to the upper left vertex of $\mathcal{H}_2$.

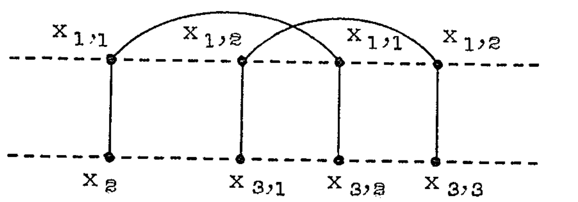
The next refinement partitions $x_3x_1x_2x_3$ into $(x_3) \cdot (x_1x_2) \cdot (x_3)$, thus yielding

$\mathcal{U}_2$: $(x_1x_2) \cdot (x_3) \cdot (x_1x_2) \cdot (x_3)$

$\mathcal{V}_2$: $(x_1x_2) \cdot (x_3) \cdot (x_1x_2) \cdot (x_3)$

and the corresponding $*$-graph, $\mathcal{H}_2''$ (which is simple.)

$\mathcal{H}_2''$:



At this point we stop refining the partitions since $\mathcal{U}_2 = \mathcal{V}_2$ and $\mathcal{H}_2''$ is simple. Note that if we had started with another solution, (σ_1', σ_2') , that had the same pattern of overlapping as (σ_1, σ_2) , we would also arrive at $\mathcal{H}_2''$ as the simple $*$-graph produced by our procedure. Now, we have already derived the pair (μ_3, μ_4) from $\mathcal{H}_2''$ ($= \mathcal{H}_{2,1}$ in the example following Definition 4.5.) This will turn out to be the correct pair to choose. The substitution, δ , is defined by sending the new labels of $\mathcal{H}_2''$ to their corresponding segments in $(\mathfrak{U}_2, \mathfrak{V}_2)$, i.e.

$$x_s\delta = \begin{cases} x_s & \text{if} \quad s \neq 1, 2 \\ x_1x_2 & \text{if} \quad s = 1 \\ x_3 & \text{if} \quad s = 2 \,. \end{cases}$$

It is easy to check that (μ_3, μ_4) and δ are the correct choices given this particular solution (σ_1, σ_2) .

Before we prove Theorem 4.7, we will make the nature of the correspondence between a pair of partitions, $(\mathfrak{U}, \mathfrak{V})$, of a word and a labelled $*$-graph, $\mathcal{G}$, more precise (Definition 4.9). Also, we will define the procedure for simultaneously refining $(\mathfrak{U}, \mathfrak{V})$ and simplifying $\mathcal{G}$ (Definition 4.11).

VALID CORRESPONDENCE.

Definition 4.8. Given a word $w \in \overline{X}$ and two partitions, $\mathfrak{U}$ and $\mathfrak{V}$, of w into non-empty segments,

$$\mathfrak{U}: \; u_1u_2 \cdots u_p \quad (= w)$$
$$\mathfrak{V}: \; v_1v_2 \cdots v_q \quad (= w) \;,$$

we say the segments u_i $(1 \le i \le p)$ and v_j $(1 \le j \le q)$ overlap in $(\mathfrak{U}, \mathfrak{V})$ if either

$$\ell(v_1 v_2 \cdots v_{j-1}) \le \ell(u_1 u_2 \cdots u_{i-1}) < \ell(v_1 v_2 \cdots v_{j-1} v_j) ,$$

or $$\ell(u_1 u_2 \cdots u_{i-1}) \le \ell(v_1 v_2 \cdots v_{j-1}) < \ell(u_1 u_2 \cdots u_{i-1} u_i) .$$

Definition 4.9. Let $\mathcal{G} = (V, E)$ be a labelled $*$-graph, with vertices

$$a_1 < a_2 < \cdots < a_p \quad \text{and} \quad b_1 < b_2 < \cdots < b_q \qquad (p, q > 0),$$

respectively labelled by

$$x_{\alpha(1)}^{\gamma(1)}, x_{\alpha(2)}^{\gamma(2)}, \ldots, x_{\alpha(p)}^{\gamma(p)} \quad \text{and} \quad x_{\beta(1)}^{\delta(1)}, x_{\beta(2)}^{\delta(2)}, \ldots, x_{\beta(q)}^{\delta(q)} ,$$

where the $\alpha(i)$'s and $\beta(i)$'s are finite sequences of positive integers and $\gamma(i), \delta(i) \in \{-1, 1\}$. Further, let $\mathfrak{U}$ and $\mathfrak{V}$ be two partitions of a word $w \in \overline{X}$ into p and q non-empty segments respectively. Letting φ be the map which sends a_i and b_j $(1 \le i \le p , 1 \le j \le q)$, respectively, to the $i^{\underline{th}}$ segment of $\mathfrak{U}$ and the $j^{\underline{th}}$ segment of $\mathfrak{V}$, we say that $\mathcal{G}$ and the pair $(\mathfrak{U}, \mathfrak{V})$ are in valid correspondence if the following properties hold:

(i) $(a_i, b_j) \in E$ if and only if $\varphi(a_i)$ overlaps $\varphi(b_j)$,

(ii) $(a_i, a_j) \in E$ implies that $\varphi(a_i) = (\varphi(a_j))^{\gamma(i)\cdot\gamma(j)}$,

and (iii) $(b_i, b_j) \in E$ implies that $\varphi(b_i) = (\varphi(b_j))^{\delta(i)\cdot\delta(j)}$.

Example. Referring to our example illustrating the proof of Theorem 4.7, the $*$-graphs $\mathcal{H}_2$, $\mathcal{H}_2'$, and $\mathcal{H}_2''$ are easily seen to be in valid correspondence with the pairs $(\mathfrak{U}_0, \mathfrak{V}_0)$, $(\mathfrak{U}_1, \mathfrak{V}_1)$, and $(\mathfrak{U}_2, \mathfrak{V}_2)$ respectively.

Lemma 4.10. (a) If $\mathcal{G}$ and $(\mathfrak{U}, \mathfrak{V})$ are in valid correspondence, then $\mathfrak{U} = \mathfrak{V}$ if and only if $\mathcal{G}$ is simple.

(b) If $\mathcal{G}$ is a simple $*$-graph in $\Gamma(\mathcal{G}(w_1, w_2))$ which is completely relabelled by the relabelling procedure and $\mathcal{G}$ is in valid correspondence with a pair $(\mathfrak{U}, \mathfrak{V})$, then if v and v' are vertices of $\mathcal{G}$, relabelled x_r^{ϵ} and x_r^{η}, we have $\varphi(v) = (\varphi(v'))^{\epsilon\eta}$.

Proof. The proof of part (a) follows immediately from Definition 4.9. Part (b) follows easily by induction on the length of the shortest path in $\mathcal{G}$ joining v and v' (they lie on the same maximal path since they are relabelled by x-symbols with the same subscript.)

S-R-S PROCEDURE.

Definition 4.11. Suppose $\mathcal{G}$ and $(\mathfrak{U}, \mathfrak{V})$ are as in Definition 4.9 and that there is a valid correspondence between $\mathcal{G}$ and $(\mathfrak{U}, \mathfrak{V})$; below we define a procedure for finding a new pair of partitions, $(\mathfrak{U}_1, \mathfrak{V}_1)$, and a new $*$-graph, $\mathcal{G}_1$. This will be accomplished by simultaneously refining one of the partitions, $\mathfrak{U}$ or $\mathfrak{V}$, and applying the simplification procedure to $\mathcal{G}$, hence the name S-R-S (simultaneous refinement and simplification) procedure.

To begin with, if $\mathfrak{U} = \mathfrak{V}$ we let $\mathfrak{U}_1 = \mathfrak{U}$, $\mathfrak{V}_1 = \mathfrak{V}$, and $\mathcal{G}_1 = \mathcal{G}$ and stop the procedure; otherwise there is a leftmost pair of vertices, a_r and b_r, such that $\ell(\varphi(a_r)) \neq \ell(\varphi(b_r))$. We assume, w.$\ell$.o.g., that $\ell(\varphi(a_r)) > \ell(\varphi(b_r))$ since the other possibility will be handled symmetrically.

Since $\ell(\varphi(a_i)) = \ell(\varphi(b_i))$ $(1 \leq i < r)$ and $\ell(\varphi(a_r)) > \ell(\varphi(b_r))$, there exist segments $\varphi(b_r)$, $\varphi(b_{r+1})$, $\ldots$, $\varphi(b_{r+m-1})$ which overlap $\varphi(a_r)$. (Note that $m = \deg_0(a_r, \mathcal{G})$ since $\mathcal{G}$ and $(\mathfrak{U}, \mathfrak{V})$ are in valid correspondence.) Some terminal segment of $\varphi(b_{r+m-1})$ may not overlap $\varphi(a_r)$, thus we partition $\varphi(a_r)$ into the segments $\varphi(b_r) \cdot \varphi(b_{r+1}) \circ \cdots \cdot \varphi(b_{r+m-2}) \cdot y$, where $\varphi(b_{r+m-1}) = y \cdot z$ with z the longest terminal segment of $\varphi(b_{r+m-1})$ which does not overlap $\varphi(a_r)$. Note that $y \neq 1$ but z may be empty.

If $\deg_s(a_r) = 0$, we let $\mathfrak{U}_1$ be the refinement of $\mathfrak{U}$ with $\varphi(a_r)$ partitioned into m segments as above. Further, we let $\mathfrak{V}_1 = \mathfrak{V}$ and we let $\mathcal{G}_1$ be the single $*$-graph in $\Gamma^1_{a_r}(\mathcal{G})$.

If $\deg_s(a_r) = 1$, there is an edge (a_r, a_t) in $\mathcal{G}$. By property (ii) of the valid correspondence, this implies that $\varphi(a_t) = (\varphi(a_r))^{\gamma(r)\cdot\gamma(t)} =$

$$\varphi(b_r)\cdot\varphi(b_{r+1})\cdot\cdots\cdot\varphi(b_{r+m-2})\cdot y \quad \text{or}$$

$$y^{-1}\cdot(\varphi(b_{r+m-2}))^{-1}\cdot\cdots\cdot(\varphi(b_r))^{-1}$$

as $\gamma(r)\cdot\gamma(t) = 1$ or -1 . Thus, in this case, we let $\mathcal{U}_1$ be the refinement of $\mathcal{U}$ with $\varphi(a_r)$ and $\varphi(a_t)$ each partitioned into m segments as we have described. As in the previous case, we let $\mathcal{V}_1 = \mathcal{V}$ and choose $\mathcal{G}_1$ from $\Gamma^1_{a_r}(\mathcal{G})$. It is easy to see that there is exactly one $*$-graph in $\Gamma^1_{a_r}(\mathcal{G})$ which is in valid correspondence with $(\mathcal{U}_1, \mathcal{V}_1)$; this is the $*$-graph produced by choosing the edges between the vertices that replace a_t and those on ℓ_b in accordance with property (i) of the valid correspondence.

In the event that $\ell(\varphi(a_r)) < \ell(\varphi(b_r))$ we would refine $\mathcal{V}$, instead of $\mathcal{U}$, partitioning $\varphi(b_r)$ (and $\varphi(b_t)$, if (b_r, b_t) is an edge) into $\deg_0(b_r, \mathcal{G})$ segments. We would then choose $\mathcal{G}_1$ from $\Gamma^1_{b_r}(\mathcal{G})$ appropriately.

<u>Remark</u>. It is clear from the definition that the S-R-S procedure preserves valid correspondence, (<u>i.e.</u> if $\mathcal{G}$ and $(\mathcal{U}, \mathcal{V})$ are in valid correspondence, then so are $\mathcal{G}_1$ and $(\mathcal{U}_1, \mathcal{V}_1)$.)

PROOF OF THEOREM 4.7.

Since $K_X(w_1, w_2) \subseteq S_X(w_1, w_2)$, it suffices to show that if $(\sigma_1, \sigma_2) \in S_X(w_1, w_2)$, there is a pair $(\mu_1, \mu_2) \in K_X(w_1, w_2)$ and a substitution δ satisfying the conditions of Definition 4.6. We begin by finding the appropriate pair, (μ_1, μ_2).

Since (σ_1, σ_2) is a c-free solution to $w_1 \overset{\vee}{=} w_2$, the word $w = w_1\sigma_1 = w_2\sigma_2$ is freely reduced and the substitutions, σ_1 and σ_2, induce two (possibly different) partitions, $\mathcal{U}_0$ and $\mathcal{V}_0$, of w. It is easy to see, from Definition 4.1, that there is a unique labelled $*$-graph, $\mathcal{G}_0$, in $\mathcal{G}(w_1, w_2)$ such that $\mathcal{G}_0$ is in valid correspondence with $(\mathcal{U}_0, \mathcal{V}_0)$; $\mathcal{G}_0$ is the $*$-graph in $\mathcal{G}(w_1, w_2)$ such that (a_i, b_j) is an edge of $\mathcal{G}_0$ if and only if the $i^{\underline{th}}$ segment of $\mathcal{U}_0$ and the $j^{\underline{th}}$ segment of $\mathcal{V}_0$ overlap. Let φ_0 be the correspondence between $\mathcal{G}_0$ and $(\mathcal{U}_0, \mathcal{V}_0)$.

We apply the S-R-S procedure (Definition 4.11) repeatedly to obtain a sequence of pairs of partitions of w,

$$(\mathcal{U}_0, \mathcal{V}_0), (\mathcal{U}_1, \mathcal{V}_1), \ldots, (\mathcal{U}_\ell, \mathcal{V}_\ell),$$

and a sequence of $*$-graphs,

$$\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_\ell,$$

such that $\mathcal{G}_i$ is in valid correspondence with $(\mathcal{U}_i, \mathcal{V}_i)$, say under φ_i , for each i $(1 \le i \le \ell)$. The positive integer, ℓ , is chosen so that $\mathcal{U}_\ell = \mathcal{V}_\ell$, and hence $\mathcal{G}_\ell$ is simple (Lemma 4.10(a)); this is possible since, at each stage in the S-R-S procedure, either $\mathcal{U}_{i+1}$ is a proper refinement of $\mathcal{U}_i$ or $\mathcal{V}_{i+1}$ is a proper refinement of $\mathcal{V}_i$. The fact that $\ell(w)$ is finite then clearly implies that for some $\ell \ge 1$, $\mathcal{U}_\ell = \mathcal{V}_\ell$.

Since $\mathcal{G}_\ell \in \Gamma^\ell(\mathcal{G}_0)$ and $\mathcal{G}_0 \in \mathcal{G}(w_1, w_2)$, it follows that $\mathcal{G}_\ell \in \Gamma(\mathcal{G}(w_1, w_2))$. Letting the vertices of $\mathcal{G}_\ell$ be

$$a_1 < a_2 < \cdots < a_p \quad \text{and} \quad b_1 < b_2 < \cdots < b_p ,$$

labelled by

$$x_{\alpha(1)}^{\gamma(1)}, x_{\alpha(2)}^{\gamma(2)}, \ldots, x_{\alpha(p)}^{\gamma(p)} \quad \text{and} \quad x_{\beta(1)}^{\delta(1)}, x_{\beta(2)}^{\delta(2)}, \ldots, x_{\beta(p)}^{\delta(p)} ,$$

it is easy to see that the relabelling procedure (Definition 4.2) completely relabels $\mathcal{G}_\ell$. If the relabelling were incomplete, by Remark 1 following Definition 4.2, there is a closed path,

$$\mathcal{P}_i = \{(v_1, v_2), (v_2, v_3), \ldots, (v_{q-1}, v_q)\} \quad (v_1 = v_q) ,$$

in $\mathcal{G}_\ell$ such that

$$\theta(v_1, v_2) \cdot \theta(v_2, v_3) \cdot \cdots \cdot \theta(v_{q-1}, v_q) = -1 .$$

But since φ_ℓ is valid, using properties (i), (ii), and (iii) of Definition 4.9 and the fact that $\mathcal{G}_\ell$ is simple, we have:

$$\varphi_\ell(v_1) = (\varphi_\ell(v_2))^{\theta(v_1,\, v_2)} , \; \varphi_\ell(v_2) = (\varphi_\ell(v_3))^{\theta(v_2,\, v_3)} , \; \ldots ,$$

$$\varphi_\ell(v_{q-1}) = (\varphi_\ell(v_q))^{\theta(v_{q-1},\, v_q)} .$$

And since $v_1 = v_q$, it follows that $\varphi_\ell(v_1) = (\varphi_\ell(v_1))^e$ where

$$e = \theta(v_1, v_2) \cdot \theta(v_2, v_3) \cdot \cdots \cdot \theta(v_{q-1}, v_q) = -1 .$$

This is impossible unless $\varphi_\ell(v_1)$ is empty, but each segment in the partitions $\mathfrak{U}_\ell$ and $\mathfrak{V}_\ell$ is non-empty; thus we have a contradiction. Therefore, $\mathcal{G}_\ell$ is a simple $*$-graph in $\Gamma(\mathcal{G}(w_1, w_2))$ which is completely relabelled by the relabelling procedure.

Referring to Definition 4.3, there is a well defined pair, (μ_1, μ_2) , of substitutions derived from $\mathcal{G}_\ell$. First we will show that $(\mu_1, \mu_2) \in S_X(w_1, w_2)$ (and hence $(\mu_1, \mu_2) \in K_X(w_1, w_2))$, then we will define a substitution δ so that (μ_1, μ_2) and δ have the desired properties.

By Lemma 4.4, we have

$$w_1\mu_1 = x_{\nu(1)}^{e(1)} \; x_{\nu(2)}^{e(2)} \cdots x_{\nu(p)}^{e(p)} = w_2\mu_2 ,$$

thus, in order to show that $(\mu_1, \mu_2) \in S_X(w_1, w_2)$, it suffices to show that $\mu_1 \in S_X(w_1)$ and $\mu_2 \in S_X(w_2)$. We will prove only that $\mu_1 \in S_X(w_1)$, since the proof that $\mu_2 \in S_X(w_2)$ is similar. It is clear, from the definition of

μ_1 (Definition 4.3), that $x_s \mu_1 \neq 1$ for each x_s . Thus it suffices to show that $x_{\nu(1)}^{e(1)} x_{\nu(2)}^{e(2)} \cdots x_{\nu(p)}^{e(p)}$ is freely reduced as written. Suppose not; then there is a segment $x_{\nu(i)}^{e(i)} x_{\nu(i+1)}^{e(i+1)}$ such that $\nu(i) = \nu(i+1)$ and $e(i) = -e(i+1)$. By Lemma 4.10(b), we have $\varphi_\ell(a_i) = \varphi_\ell(a_{i+1})^{-1}$; thus w ($= w_1\sigma_1$) has two consecutive non-empty segments which cancel with each other. This contradicts the fact that $w \in \overline{X}$; therefore, $x_{\nu(1)}^{e(1)} x_{\nu(2)}^{e(2)} \cdots x_{\nu(p)}^{e(p)}$ is in $\overline{X}$ and μ_1 is c-free for w_1 . Thus we have $(\mu_1, \mu_2) \in S_X(w_1, w_2)$ (and hence $(\mu_1, \mu_2) \in K_X(w_1, w_2)$.)

We define the substitution, δ , as follows:

$$x_s\delta = \begin{cases} x_s & \text{if } x_s \notin \mathrm{Supp}(w_1 \mu_1) \\ \\ \varphi_\ell(a_r)^{e(r)} & \text{if } x_s \in \mathrm{Supp}(w_1 \mu_1) \end{cases} ,$$

where r is the least positive integer such that $\nu(r) = s$. (Note that by Remark (2) following Definition 4.2 and Lemma 4.10(b), $\varphi_\ell(a_r) = \varphi_\ell(b_r)$ $(1 \le r \le p)$.) First we claim that δ is c-free for $w_1 \mu_1$. This is clear since $x_s\delta \neq 1$ for each x_s

and $(w_1 \mu_1)\delta = (x_{\nu(1)}^{e(1)} \cdot x_{\nu(2)}^{e(2)} \cdot \cdots \cdot x_{\nu(p)}^{e(p)})\delta =$

$\varphi_\ell(a_1)\,\varphi_\ell(a_2) \cdots \varphi_\ell(a_p) = w_1\sigma_1 \in \overline{X}$. (Note that $\varphi_\ell(a_i)^{e(i)} = \varphi_\ell(a_j)^{e(j)}$ if $\nu(i) = \nu(j)$, by Lemma 4.10(b).) Thus $\delta \in S_X(w_1\mu_1)$.

Finally we show that, for $x_s \in \mathrm{Supp}(w_1)$ and $x_t \in \mathrm{Supp}(w_2)$,

$$x_s\mu_1\delta = x_s\sigma_1 \quad \text{and} \quad x_t\mu_2\delta = x_t\sigma_2 .$$

Again, both proofs are similar; thus we only prove the first equality. By the definition of μ_1 (Definition 4.3), if $x_s \in \mathrm{Supp}(w_1)$,

$$x_s\mu_1 = (x_{\nu(i)}^{e(i)}\, x_{\nu(i+1)}^{e(i+1)} \cdots x_{\nu(i+j)}^{e(i+j)})^{\varepsilon(r)} ,$$

where r is the least positive integer such that $x_{i(r)} = x_s$ (recall, $w_1 = x_{i(1)}^{\varepsilon(1)}\, x_{i(2)}^{\varepsilon(2)} \cdots x_{i(m)}^{\varepsilon(m)}$.) We then have,

$$(x_s\mu_1)\delta = (x_{\nu(i)}^{e(i)}\, x_{\nu(i+1)}^{e(i+1)} \cdots x_{\nu(i+j)}^{e(i+j)})^{\varepsilon(r)}\delta$$

$$= (\varphi_\ell(a_i)\,\varphi_\ell(a_{i+1}) \cdots \varphi_\ell(a_{i+j}))^{\varepsilon(r)} .$$

But $\varphi_\ell(a_i)\,\varphi_\ell(a_{i+1}) \cdots \varphi_\ell(a_{i+j}) = \varphi_0(a'_r)$, where a'_r is the $r^{\underline{th}}$ vertex on ℓ_a in $\mathcal{G}_0$. (This follows from

Definition 4.11.) By the definition of $\mathcal{G}_0$,

$\varphi_0(a'_r) = x_{i(r)}^{\varepsilon(r)} \sigma_1$; therefore

$$(x_s \mu_1)\delta = (\varphi_0(a'_r))^{\varepsilon(r)} = x_{i(r)}\sigma_1 = x_s\sigma_1 ,$$

as required. ∎

We will apply Theorem 4.7 in the next chapter by way of the following consequence.

<u>Corollary 4.12.</u> If $w_1 \stackrel{\vee}{=} w_2$ is a quadratic verbal equation with $\mathrm{Supp}(w_1) \cap \mathrm{Supp}(w_2) = \emptyset$, then there exists a complete set, K , of c-free solutions in X to $w_1 \stackrel{\vee}{=} w_2$ and there is an integer $n(w_1, w_2) \geq 0$ such that if w_i $(i = 1, 2)$ is linear in some $x_s \in \mathrm{Supp}(w_i)$, then $\ell(x_s \mu_i) \leq n(w_1, w_2)$ for each pair $(\mu_1, \mu_2) \in K$.

<u>Proof</u>. We let $K = K_X(w_1, w_2)$ and use Theorem 4.7 to show that K is complete. If w_1 is linear in x_s , let $(\mu_1, \mu_2) \in K$ and consider $x_s \mu_1$. By Definition 4.3, we have

$$x_s \mu_1 = (x_{\nu(i)}^{e(i)} x_{\nu(i+1)}^{e(i+1)} \cdots x_{\nu(i+j)}^{e(i+j)})^{\varepsilon(r)} ,$$

where r is the least integer such that $i(r) = s$ and

$$a_i < a_{i+1} < \cdots < a_{i+j}$$

are the consecutive vertices of $\mathcal{G}_\ell$ which ultimately replace a_r' in $\mathcal{G}_0$. Since w_1 is linear in x_s, $\deg_s(a_r', \mathcal{G}_0) = 0$ (i.e. a_r' is an end of $\mathcal{G}_0$). Since the simplification procedure (Definition 3.7) replaces ends by ends, $a_i, a_{i+1}, \cdots,$ and a_{i+j} are ends of $\mathcal{G}_\ell$. By Theorem 3.11, the number of ends of $\mathcal{G}_\ell$ is bounded by $2(\mathrm{card}(V_0)-1)$ where V_0 is the set of vertices of $\mathcal{G}_0$. By the definition of $\mathcal{G}_0$, $\mathrm{card}(V_0) = m + n$. Therefore,

$$\ell(x_s \mu_1) = j \le \mathrm{ends}(\mathcal{G}_\ell) \le 2(m + n - 1) .$$

A similar argument for w_2 (linear in some x_t) would show that it suffices to choose $n(w_1, w_2) \ge 2(m + n - 1)$. ∎

CHAPTER 5

THE PRODUCT OF TWO QUADRATIC WORDS

In this Chapter, we apply the results of Chapter 4, to obtain the following Theorem.

Theorem 5.1. If w_1 and w_2 are quadratic words in $\overline{X}$ such that $\mathrm{Supp}(w_1) \cap \mathrm{Supp}(w_2) = \phi$, $C(w_1)$ is finite, and $C(w_2)$ is finite, then $C(w_1 \cdot w_2)$ is finite.

We note that this theorem is trivial if either w_1 or w_2 is empty or linear in some x_s (see Example (3) following Definition 2.5); henceforth we will assume that w_1 and w_2 are non-empty and strictly quadratic, i.e. quadratic in each element of their supports.

We will draw several interesting conclusions from Theorem 5.1 in conjunction with the corollary to Theorem 2.2. We will list these consequences in Chapter 6.

To prove Theorem 5.1, we will exhibit a complete set, S , of images of $w_1 \cdot w_2$ (see Definition 2.1) which contains only words of length less than some fixed bound. By Lemma 1.5(e), it will follow that $N(S)$, the normalization of S , is finite. But once we have $N(S)$ finite for some complete set, S , of images of $w_1 \cdot w_2$, Theorem 2.14 will imply that $C(w_1 \cdot w_2)$ is finite. The following sequence of definitions leads to the definition of the set S .

Definition 5.2. If $w \in \overline{X}$ and $x_i \in \mathrm{Supp}(w)$, we define the substitutions $\beta_{(w;\, i,\, \gamma)}$ ($\gamma \in \{-1, +1\}$) from X into X by

$$x_s\beta_{(w;\, i,\, \gamma)} = \begin{cases} x_s & \text{if } s \neq i \\ (x_m x_{m+1})^{\gamma} & \text{if } s = i \end{cases}$$

where $m = \max\{s : x_s \in \mathrm{Supp}(w)\} + 1$.

If w is quadratic in x_i, we can write $w = u_1 x_i^{\epsilon} u_2 x_i^{\eta} u_3$ where $\epsilon, \eta \in \{-1, +1\}$, and u_1, u_2, and u_3 are (possibly empty) words in $\overline{X}$ whose support does not contain x_i.

If $\epsilon \neq \eta$, then

$$w\beta_{(w;\, i,\, \epsilon)} = u_1 x_m x_{m+1} u_2 x_{m+1}^{-1} x_m^{-1} u_3$$

and

$$w\beta_{(w;\, i,\, \eta)} = u_1 x_{m+1}^{-1} x_m^{-1} u_2 x_m x_{m+1} u_3 .$$

In this case, we call the words $u_1 x_m$ and $x_{m+1} u_2 x_{m+1}^{-1} x_m^{-1} u_3$, respectively, the <u>special initial</u> and <u>special terminal segments</u>

of $w\beta_{(w;\, i,\, \epsilon)}$; similarly, we call $u_1 x_{m+1}^{-1} x_m^{-1} u_2 x_m$ and $x_{m+1} u_3$ the special initial and special terminal segments of $w\beta_{(w;\, i,\, \eta)}$.

If $\epsilon = \eta$, then

$$w\beta_{(w;\, i,\, \epsilon)} = w\beta_{(w;\, i,\, \eta)} = u_1 x_m x_{m+1} u_2 x_m x_{m+1} u_3 \ .$$

In this case, we call the words $u_1 x_m$ and $u_1 x_m x_{m+1} u_2 x_m$ the special initial segments of $w\beta_{(w;\, i,\, \epsilon)}$ and we call the words $x_{m+1} u_2 x_m x_{m+1} u_3$ and $x_{m+1} u_3$ the special terminal segments of $w\beta_{(w;\, i,\, \epsilon)}$.

Definition 5.3. If w is a quadratic word in $\overline{X}$, we define the finite sets of reduced words Initial(w) and Terminal(w) as follows:

Initial(w) (Terminal(w)) consists of the initial (terminal) segments of w (including 1 and w) and, for all $x_i \in \mathrm{Supp}(w)$, the special initial (terminal) segments of $w\beta_{(w;\, i,\, \epsilon)}$ and $w\beta_{(w;\, i,\, \eta)}$ (as defined above) .

Example. If $w = x_2^{-1} x_3 x_2 x_3$, then

$$w\beta_{(w;\, 2,\, -1)} = x_4 x_5 x_3 x_5^{-1} x_4^{-1} x_3 \ ,$$

$$w\beta_{(w;\, 3,\, 1)} = x_2^{-1} x_4 x_5 x_2 x_4 x_5 \ , \text{ and}$$

$$w\beta_{(w;\,2,\,1)} = x_5^{-1}x_4^{-1}x_3x_4x_5x_3 \ .$$

Thus, for instance,

$$\text{Initial}(w) = \{x_2^{-1}x_3x_2x_3,\ x_2^{-1}x_3x_2,\ x_2^{-1}x_3,\ x_2^{-1},\ 1,\ x_4,\ x_2^{-1}x_4,$$

$$x_2^{-1}x_4x_5x_2x_4,\ x_5^{-1}x_4^{-1}x_3x_4\} \ .$$

According to the hypothesis of Theorem 5.1, w_2 is a quadratic word in $\overline{X}$ and $C(w_2)$ is finite. Henceforth, we will denote by $D(w_2)$ the set of words obtained by adding the constant m_0 to each subscript of each x-symbol occurring in $C(w_2)$, where

$$m_0 = \max\{s : \ x_s \in \text{Supp}(w) \ \text{ where } \ w \in C(w_1)\} + 2 \ .$$

This insures that, for each $w \in C(w_1)$ and $u \in D(w_2)$,

$$\text{Supp}(w) \cap \text{Supp}(u) = \phi$$

and

$$\text{Supp}(w\beta_{(w;\,i,\,\epsilon)}) \cap \text{Supp}(u\beta_{(u;\,j,\,\eta)}) = \phi \ .$$

Also, since $D(w_2)$ is produced by a level substitution, it follows clearly that $D(w_2)$ is a complete set of images of w_2 .

<u>Definition 5.4</u>. (a) The set of <u>cancellation</u> <u>equations</u> derived from the pair, (w_1, w_2) , is the set

$$\{w_T \stackrel{\vee}{=} u_I^{-1} : \ \ w \in C(w_1),\ u \in D(w_2),\ w_T \in \text{Terminal}(w),\ u_I \in \text{Initial}(u)\} \ .$$

(b) The set of <u>residual products</u> derived from the pair, (w_1, w_2) , is the set

$$\{w_I \cdot u_T : w \in C(w_1), u \in D(w_2), w_I \in \text{Initial}(w), u_T \in \text{Terminal}(u)\} .$$

We note that by Lemma 2.6(b) the sets $C(w_1)$ and $C(w_2)$ (and hence $D(w_2)$) contain only quadratic words. Since segments of quadratic words are quadratic and since $\beta_{(w;\, i,\, \varepsilon)}$ sends quadratic words to quadratic words, we also note that each cancellation equation derived from a pair of quadratic words is quadratic and that each residual product derived from a pair of quadratic words is a quadratic word.

<u>Proof of Theorem 5.1</u>. Beginning with the pair (w_1, w_2) , we derive the finite sets of (quadratic) cancellation equations,

$$\{w_T \veeeq u_I^{-1} : w \in C(w_1), u \in D(w_2), w_T \in \text{Terminal}(w), u_I \in \text{Initial}(u)\},$$

and (quadratic) residual products,

$$\{w_I \cdot u_T : w \in C(w_1), u \in D(w_2), w_I \in \text{Initial}(w), u_T \in \text{Terminal}(u)\} .$$

For each equation, $w_T \veeeq u_I^{-1}$, we use Corollary 4.12 to obtain the set $K_X(w_T, u_I^{-1})$ and the integer $n(w_T, u_I^{-1})$. We let m_1 be the maximum of all these integers.

If $(\mu_1, \mu_2) \in K_X(w_T, u_I^{-1})$ for some w_T and u_I^{-1} , we will assume w.ℓ.o.g. that

$$x_s \mu_1 = x_{s+m(w_T)} \quad \text{for} \quad x_s \notin \text{Supp}(w_T)$$

and $$x_s\mu_2 = x_{s+m(u_I^{-1})} \qquad \text{for } x_s \notin \mathrm{Supp}(u_I^{-1}) ,$$

where $m(v) = \max\{t : x_t \in \mathrm{Supp}(v)\}$ if v is a non-empty freely reduced word and $m(v) = 0$ if $v = 1$. If we redefine the pairs in $K_X(w_T, u_I^{-1})$ to be of this type, it is easy to see that the new set of pairs (which we will also refer to as $K_X(w_T, u_I^{-1})$, for convenience) has all the properties mentioned in Corollary 4.12. Therefore we lose no generality in making this assumption. Later we will assume that for each $(\mu_1, \mu_2) \in K_X(w_T, u_I^{-1})$, the associated substitution, δ (see Corollary 4.12), is also of a special type; again, no generality will be lost in this assumption.

We now define the set, S, as follows:

$$S = S' \cup S'' ,$$

where $S' = \{w \cdot u : w \in C(w_1) \text{ and } u \in D(w_2)\}$, and

$$\begin{aligned} S'' = \{(w_I\mu_1)\cdot(u_T\mu_2) \in \bar{X} :\ & w \in C(w_1),\ u \in D(w_2),\ w_I \in \mathrm{Initial}(w), \\ & u_T \in \mathrm{Terminal}(u), \text{ and } (\mu_1, \mu_2) \in K_X(w_T, u_I^{-1}) \text{ where} \\ & w_Iw_T = w \text{ or } w\beta_{(w;\, i,\, \epsilon)} \text{ for some } x_i^{\epsilon} \text{ in } w, \text{ and} \\ & u_Iu_T = u \quad \text{or} \quad u\beta_{(u;\, j,\, \eta)} \quad \text{for some } x_j^{\eta} \text{ in } u\} . \end{aligned}$$

We will show that S is a complete set of images of $w_1 \cdot w_2$ and that there is a bound on the length of the words in S. It will then be clear that $N(S)$ (see Definition 1.6) is a finite complete set of images of $w_1 \cdot w_2$ and thus, by Theorem 2.14, $C(w_1 \cdot w_2)$ is finite.

To show that S is a complete set of images of $w_1 \cdot w_2$, we must verify Definition 2.1(i) and (ii) for S.

VERIFICATION OF 2.1(i).

If $w \cdot u \in S'$, then $w \in C(w_1)$ and $u \in D(w_2)$ and, by the definition of $D(w_2)$, $\mathrm{Supp}(w) \cap \mathrm{Supp}(u) = \emptyset$. Since $C(w_1)$ and $D(w_2)$ are complete sets of images of w_1 and w_2 respectively, there exist substitutions σ and τ such that $\overline{w_1\sigma} = w$ and $\overline{w_2\tau} = u$. Defining the substitution γ by

$$x_s\gamma = \begin{cases} x_s\sigma & \text{if } x_s \in \mathrm{Supp}(w_1) \\ \\ x_s\tau & \text{otherwise}, \end{cases}$$

it is clear that γ is well defined and $(w_1 \cdot w_2)\gamma = (w_1\sigma) \cdot (w_2\tau)$. Therefore, $w \cdot u$ $(= (\overline{w_1\sigma}) \cdot (\overline{w_2\tau}))$ is a partially reduced form of $(w_1 \cdot w_2)\gamma$. But $w \cdot u$ is freely reduced, since $\mathrm{Supp}(w) \cap \mathrm{Supp}(u) = \emptyset$ implies that the junction of w and u cannot be a trivial relator. Thus, $\overline{(w_1 \cdot w_2)\gamma} = \overline{w \cdot u} = w \cdot u$, and hence $w \cdot u$ is an image of $w_1 \cdot w_2$.

Suppose $(w_I \mu_1) \cdot (u_T \mu_2) \in S''$. Since $w_I \cdot w_T$ equals either w or $w\beta_{(w;\, i,\, \epsilon)}$ for some $w \in C(w_1)$, it follows by Theorem 2.13 and Definition 2.1(i) that there is a substitution σ_1 such that $\overline{w_1 \sigma_1} = w_I \cdot w_T$. Similarly there is a σ_2 such that $\overline{w_2 \sigma_2} = u_I \cdot u_T$. Let the substitution σ be defined by

$$x_s\sigma = \begin{cases} x_s\sigma_1 \mu_1 & \text{if } x_s \in \mathrm{Supp}(w_1) \\ x_s\sigma_2 \mu_2 & \text{if } x_s \in \mathrm{Supp}(w_2) \\ x_s & \text{otherwise.} \end{cases}$$

The substitution σ is well defined since $\mathrm{Supp}(w_1) \cap \mathrm{Supp}(w_2) = \phi$.

Now

$$(\overline{w_1\sigma_1}\, \mu_1) \cdot (\overline{w_2\sigma_2}\, \mu_2) = ((w_I \cdot w_T)\mu_1) \cdot ((u_I \cdot u_T)\, \mu_2) =$$
$$(w_I \mu_1) \cdot (w_T \mu_1) \cdot (u_I \mu_2) \cdot (u_T \mu_2) ,$$

and since $(\mu_1, \mu_2) \in K_X\, (w_T, u_I^{-1})$, $w_T \mu_1 = (u_I \mu_2)^{-1}$; therefore $(w_I \mu_1) \cdot (u_T \mu_2)$ is a partially reduced form of $(\overline{w_1\sigma_1}\, \mu_1) \cdot (\overline{w_2\sigma_2}\, \mu_2)$ which is a partially reduced form of $(w_1\sigma) \cdot (w_2\sigma) = (w_1 \cdot w_2)\sigma$. By the definition of S'' , $(w_I \mu_1) \cdot (u_T \mu_2)$ is in $\overline{X}$ and thus is freely reduced as written. Therefore, $\overline{(w_1 \cdot w_2)\sigma} = (w_I \mu_1) \cdot (u_T \mu_2)$ as required.

VERIFICATION OF 2.1(ii).

This portion of the proof of Theorem 5.1 breaks up into four separate cases and, therefore, becomes somewhat long. The proof of the first case is given in full detail, but the proofs of the remaining cases are given in a more abbreviated form, since they are similar to the proof of the first case.

Suppose that σ is a substitution from X into X and consider the (not necessarily reduced) word $(w_1 \cdot w_2)\sigma = (w_1\sigma) \cdot (w_2\sigma)$. Since $C(w_1)$ and $D(w_2)$ are complete sets of images for w_1 and w_2 respectively, there exist words $w \in C(w_1)$ and $u \in D(w_2)$ and substitutions $\gamma_1 \in S_X(w)$ and $\gamma_2 \in S_X(u)$ such that $w\gamma_1 = \overline{w_1\sigma}$ and $u\gamma_2 = \overline{w_2\sigma}$. Thus $(w\gamma_1) \cdot (u\gamma_2)$ is a reduced form of $(w_1 \cdot w_2)\sigma$ with the initial segment, $w\gamma_1$, and the terminal segment, $u\gamma_2$, both freely reduced.

If the word $(w\gamma_1) \cdot (u\gamma_2)$ is freely reduced as written, we choose $w_I = w$, $w_T = u_I = 1$, and $u_T = u$, and define the substitution $\gamma \in S_X(w \cdot u)$ by

$$x_s\gamma = \begin{cases} x_s\gamma_1 & \text{if } x_s \in \mathrm{Supp}(w) \\ x_s\gamma_2 & \text{otherwise.} \end{cases}$$

Clearly $x_s\gamma = x_s\gamma_2$ for $x_s \in \mathrm{Supp}(u)$, since $\mathrm{Supp}(w) \cap \mathrm{Supp}(u) = \emptyset$.

Thus we have $\overline{(w_1 \cdot w_2)\sigma} = (w\gamma_1) \cdot (u\gamma_2) = (w \cdot u)\gamma$ with $\gamma \in S_X(w \cdot u)$. Also since $w \in C(w_1)$ and $u \in D(w_2)$, we have $w \cdot u \in S' \subseteq S$; thus 2.1(ii) holds in this case.

Suppose $(w\gamma_1) \cdot (u\gamma_2)$ is not freely reduced as written; since $w\gamma_1$ and $u\gamma_2$ are freely reduced, the free reduction of $(w\gamma_1) \cdot (u\gamma_2)$ can be accomplished by cancelling a terminal segment of $(w\gamma_1)$ against an initial segment of $u\gamma_2$. Thus $w\gamma_1$ and $u\gamma_2$ can be partitioned into segments $w\gamma_1 = (w\gamma_1)_I \cdot (w\gamma_1)_T$ and $u\gamma_2 = (u\gamma_2)_I \cdot (u\gamma_2)_T$ so that $(w\gamma_1)_T = (u\gamma_2)_I^{-1}$ and $(w\gamma_1)_I \cdot (u\gamma_2)_T$ is freely reduced as written. Therefore $\overline{(w_1 \cdot w_2)\sigma} = \overline{(w\gamma_1) \cdot (u\gamma_2)} = \overline{(w\gamma_1)_I \cdot (w\gamma_1)_T \cdot (u\gamma_2)_I \cdot (u\gamma_2)_T} = (w\gamma_1)_I \cdot (u\gamma_2)_T$.

And thus it suffices to show that for some $(w_I\mu_1) \cdot (u_T\mu_2) \in S$, there is a substitution $\delta \in S_X((w_I\mu_1) \cdot (u_T\mu_2))$ such that $((w_I\mu_1) \cdot (u_T\mu_2))\delta = (w\gamma_1)_I \cdot (u\gamma_2)_T$.

We partition w and u as follows: $w = w_L \cdot x_i^\epsilon \cdot w_R$ and $u = u_L \cdot x_j^\eta \cdot u_R$ where

$$\ell(w_L\gamma_1) \leq \ell((w\gamma_1)_I) < \ell((w_L \cdot x_i^\epsilon)\gamma_1)$$

and

$$\ell((x_j^\eta \cdot u_R)\gamma_2) > \ell((u\gamma_2)_T) \geq \ell(u_R\gamma_2) .$$

We have four cases to consider.

Case 1. Suppose that $\ell(w_L\gamma_1) = \ell((w\gamma_1)_I)$ and $\ell((u\gamma_2)_T) = \ell(u_R\gamma_2)$. Then we let $w = w_I \cdot w_T$ and $u = u_I \cdot u_T$ where $w_I = w_L$, $w_T = x_i^\epsilon w_R$, $u_I = u_L x_j^\eta$, and $u_T = u_R$. Since $w \in C(w_1)$ and $u \in D(w_2)$, it follows that $w_T \stackrel{\vee}{=} u_I^{-1}$ is a cancellation equation derived from (w_1, w_2) and $w_I \cdot u_T$ is the residual product associated with it. Since $\ell(w_L\gamma_1) = \ell((w\gamma_1)_I)$ and both words are initial segments of $w\gamma_1$, it follows that $w_L\gamma_1 = (w\gamma_1)_I$ and similarly that $(u\gamma_2)_T = u_R\gamma_2$. Thus

$$w_I\gamma_1 = (w\gamma_1)_I \quad \text{and} \quad (u\gamma_2)_T = u_T\gamma_2 ,$$

therefore

$$w_T\gamma_1 = (w\gamma_1)_T \quad \text{and} \quad u_I\gamma_2 = (u\gamma_2)_I \ .$$

Since $w_T\gamma_1 = (w\gamma_1)_T = (u\gamma_2)_I^{-1} = u_I^{-1}\gamma_2$, the pair (γ_1, γ_2) is in $S_X(w_T, u_I^{-1})$. And since $K_X(w_T, u_I^{-1})$ is a complete set of c-free solutions to $w_T \stackrel{\vee}{=} u_I^{-1}$ in X , there is a pair $(\mu_1, \mu_2) \in K_X(w_T, u_I^{-1})$ and a substitution δ from X into X which is c-free for $w_T\mu_1$ $(= u_I^{-1}\mu_2)$ and for which $x_s\mu_1\delta = x_s\gamma_1$ for each $x_s \in \mathrm{Supp}(w_T)$ and $x_t\mu_2\delta = x_t\gamma_2$ for each $x_t \in \mathrm{Supp}(u_I)$.

Recall, we have assumed that $x_s\mu_1 = x_{s+m(w_T)}$ and $x_t\mu_2 = x_{t+m(u_I^{-1})}$ respectively, for $x_s \notin \mathrm{Supp}(w_T)$ and $x_t \notin \mathrm{Supp}(u_I)$. Similarly, w.ℓ.o.g. we may assume that $x_s\delta = x_{s-m(w_T)}\gamma_1$ for $x_s \in \mathrm{Supp}(w_I\mu_1) \setminus \mathrm{Supp}(w_T\mu_1)$ and $x_t\delta = x_{t-m(u_I^{-1})}\gamma_2$ for $x_t \in \mathrm{Supp}(u_T\mu_2) \setminus \mathrm{Supp}(u_I\mu_2)$. To see that no generality is lost when δ is chosen this way, first note that

$$(\mathrm{Supp}(w_I\mu_2) \setminus \mathrm{Supp}(w_T\mu_1)) \cap (\mathrm{Supp}(u_T\mu_2) \setminus \mathrm{Supp}(u_I\mu_2)) = \emptyset \ .$$

This is true, for suppose $x_s \in \mathrm{Supp}(w_I\mu_1) \setminus \mathrm{Supp}(w_T\mu_1)$, then $x_s \in \mathrm{Supp}(x_t\mu_1)$ for some $x_t \in \mathrm{Supp}(w_I) \setminus \mathrm{Supp}(w_T)$ and thus $x_s = x_t\mu_1$ where $t = s - m(w_T)$. Similarly, if $x_s \in \mathrm{Supp}(u_T\mu_2) \setminus \mathrm{Supp}(u_I\mu_2)$, we have $x_s \in \mathrm{Supp}(x_r\mu_2)$ for some $x_r \in \mathrm{Supp}(u_T) \setminus \mathrm{Supp}(u_I)$ and thus $x_s = x_r\mu_2$ where $r = s - m(u_I^{-1})$. Therefore, $t + m(w_T) = r + m(u_I^{-1})$. But, since $x_t \in \mathrm{Supp}(w)$ and $x_r \in \mathrm{Supp}(u)$ and $u \in D(w_2)$, the definition of $D(w_2)$ implies that $t < r$ and $m(w_T) < m(u_I^{-1})$. Therefore $t + m(w_T) \neq r + m(u_I^{-1})$, which is a contradiction.

Since

$$\mathrm{Supp}(w_I\mu_1) \setminus \mathrm{Supp}(w_T\mu_1) \quad \text{and} \quad \mathrm{Supp}(u_T\mu_2) \setminus \mathrm{Supp}(u_I\mu_2)$$

are disjoint sets, the values of $x_s\delta$ and $x_t\delta$ can be defined to be any non-empty reduced word we choose (viz. $x_{s-m(w_T)}\gamma_1$ and $x_{t-m(u_I^{-1})}\gamma_2$) on these sets and it is still true that $x_s\mu_1\delta = x_s\gamma_1$ on $\mathrm{Supp}(w_T)$, $x_t\mu_2\delta = x_t\gamma_2$ on $\mathrm{Supp}(u_I)$, and δ is c-free for $w_T\mu_1$ $(= u_I^{-1}\mu_2)$.

We claim that $w_I\mu_1\delta = w_I\gamma_1$ and $u_T\mu_2\delta = u_T\gamma_2$. To see this, let $x_s \in \mathrm{Supp}(w_I) \cap \mathrm{Supp}(w_T)$; then, since $x_s \in \mathrm{Supp}(w_T)$, we have $x_s\mu_1\delta = x_s\gamma_1$. If

$x_s \in \mathrm{Supp}(w_I) \setminus \mathrm{Supp}(w_T)$, then by our conventions concerning μ_1 and δ we have $x_s\mu_1 = x_{s+m(w_T)}$, and thus $x_s\mu_1\delta = x_{s+m(w_T)}\delta = x_{s+m(w_T)-m(w_T)}\gamma_1 = x_s\gamma_1$. Therefore, for each $x_s \in \mathrm{Supp}(w_I)$, $x_s\mu_1\delta = x_s\gamma_1$ and thus $w_I\mu_1\delta = w_I\gamma_1$. A similar argument shows that $u_T\mu_2\delta = u_T\gamma_2$.

We have shown that $((w_I\mu_1) \cdot (u_T\mu_2))\delta = (w_I\gamma_1) \circ (u_T\gamma_2) = (w\gamma_1)_I \cdot (u\gamma_2)_T$, which is freely reduced by assumption, thus, by Lemma 2.7, $(w_I\mu_1) \cdot (u_T\mu_2)$ is freely reduced as written and hence is in S''. Therefore, we have exhibited $(w_I\mu_1) \cdot (u_T\mu_2) \in S$ and $\delta \in S_X((w_I\mu_1) \cdot (u_T\mu_2))$ such that

$$((w_I\mu_1) \cdot (u_T\mu_2))\delta = (w\gamma_1)_I \cdot (u\gamma_2)_T = \overline{(w\gamma_1) \cdot (u\gamma_2)} = \overline{(w_1 \cdot w_2)\sigma} .$$

This completes the proof in Case 1.

Case 2. Suppose that $\ell(w_L\gamma_1) < \ell((w\gamma_1)_I)$ and $\ell((u\gamma_2)_T = \ell(u_R\gamma_2)$. Recall that we have written $w = w_L \circ x_i^{\epsilon} \cdot w_R$ and $u = u_L\, x_j^{\eta}\, u_R$, thus we can write

$$(w\gamma_1)_I = (w_L\gamma_1) \circ v_{1,1} \quad \text{and} \quad (w\gamma_1)_T = v_{1,2} \cdot (w_R\gamma_1) ,$$

where $x_i^{\epsilon}\gamma_1 = v_{1,1} \cdot v_{1,2}$ and neither $v_{1,1}$ nor $v_{1,2}$

is empty. Also we have

$w\beta_{(w;\, i,\, \epsilon)} = (w_L\beta_{(w;\, i,\, \epsilon)}) \cdot (x_m x_{m+1}) \cdot (w_R\beta_{(w;\, i,\, \epsilon)})$. Choose $u_I = u_L x_j^{\eta}$, $u_T = u_R$ and from Initial(w) and Terminal(w) choose $w_I = (w_L\beta_{(w;\, i,\, \epsilon)}) \cdot x_m$ and $w_T = x_{m+1} \cdot (w_R\beta_{(w;\, i,\, \epsilon)})$. Clearly $w_T \veeeq u_I^{-1}$ is a cancellation equation derived from (w_1, w_2) , and $w_I \cdot u_T$ is the residual pair associated with it. We can assume w.ℓ.o.g. that $x_s\gamma_1$ is any arbitrary non-empty freely reduced word in $\overline{X}$, provided $x_s \notin \mathrm{Supp}(w)$; thus w.ℓ.o.g. assume that $x_m\gamma_1 = v_{1,1}$ and $x_{m+1}\gamma_1 = v_{1,2}$.

We note that, under this assumption, $w\gamma_1 = w\beta_{(w;\, i,\, \epsilon)}\gamma_1$ and furthermore, $w_I\gamma_1 = (w\gamma_1)_I$, $(u\gamma_2)_T = u_T\gamma_2$, $w_T\gamma_1 = (w\gamma_1)_T$, and $u_I\gamma_2 = (u\gamma_2)_I$. The proof now proceeds as in Case 1.

<u>Case 3</u>. Now suppose $\ell(w_L\gamma_1) = \ell((w\gamma_1)_I)$ and $\ell((u\gamma_2)_T) < \ell(u_R\gamma_2)$. Here we choose $w_I = w_L$, $w_T = x_i^{\epsilon} w_R$, $u_I = (u_L\beta_{(u;\, j,\, \eta)}) \cdot x_n$, and $u_T = x_{n+1} \cdot (u_R\beta_{(u;\, j,\, \eta)})$. And, writing $x_j^{\eta}\gamma_2 = v_{2,1} \cdot v_{2,2}$, we assume w.ℓ.o.g. that $x_n\gamma_2 = v_{2,1}$ and $x_{n+1}\gamma_2 = v_{2,2}$, where $n = \max\{x_s : x_s \in \mathrm{Supp}(u)\} + 1$.

<u>Case 4</u>. Finally suppose $\ell(w_L\gamma_1) < \ell((w\gamma_1)_I)$ and $\ell((u\gamma_2)_T) > \ell(u_R\gamma_2)$.

Again the proof is essentially the same as in the previous cases, but here we choose $w_I = (w_L\beta_{(w;\, i,\, \epsilon)}) \cdot x_m$, $w_T = x_{m+1} \cdot (w_R\beta_{(w;\, i,\, \epsilon)})$, $u_I = (u_L\beta_{(u;\, j,\, \eta)}) \cdot x_n$, and $u_T = x_{n+1} \cdot (u_R\beta_{(u;\, j,\, \eta)})$. And w.$\ell$.o.g. we assume that $x_m\gamma_1 = v_{1,1}$, $x_{m+1}\gamma_1 = v_{1,2}$, $x_n\gamma_2 = v_{2,1}$, and $x_{n+1}\gamma_2 = v_{2,2}$. Note that since $u \in D(w_2)$ and the subscripts of u have been increased by $m_0 (= \max\{s : x_s \in \mathrm{Supp}(w)\,,\ w \in C(w_1)\} + 2)$, we are assured that $m+1$ is less than every subscript appearing in u or $u\beta_{(u;\, i,\, \eta)}$. Thus we do not lose generality in making the assumptions above.

This completes the verification of 2.1(ii) and, thus, shows that S is a complete set of images for $w_1 \cdot w_2$.

THE BOUND.

Finally, we will show that there is a bound on the lengths of the words in the set S.

By assumption, $C(w_1)$ and $D(w_2)$ are finite sets; therefore there are only finitely many cancellation equations, $w_T \underline{\underline{\vee}} u_I^{-1}$, and residual products $w_I \cdot u_T$, derived from the pair (w_1, w_2). Thus, it suffices to prove that for each residual product, $w_I \cdot u_T$, there is a bound on the length of the words in the set $\{(w_I\mu_1) \cdot (u_T\mu_2) : (\mu_1, \mu_2) \in K_X(w_T, u_I^{-1})\}$. If x_s occurs twice in w_T , then it cannot occur in w_I since w_Iw_T ($= w$ or $w\beta_{(w;\, i,\, \epsilon)}$) is quadratic. If x_s occurs once in w_T it can occur at most once in w_I , and thus $\ell(x_s\mu_1) \leq m_1$. (Recall, $m_1 = \max\{n(w_T, u_I^{-1})\}$) . If x_s is not present in w_T , we have assumed that μ_1 is defined so that $x_s\mu_1 = x_{s+m(w_T)}$, thus $\ell(x_s\mu_1) = 1 \leq m_1$. A similar argument shows that for each $x_s \in \mathrm{Supp}(u_T)$, $\ell(x_s\mu_2) \leq m_1$. Since there are only finitely many residual products, $w_I \cdot u_T$, there is an upper bound on the length of all w_I's and u_T's . And since m_1 bounds the length of each $x_s\mu_1$ and $x_t\mu_2$, where $x_s \in \mathrm{Supp}(w_I)$ and $x_t \in \mathrm{Supp}(u_T)$, it follows that there is a bound on the maximum length of the words in the set $\{(w_I\mu_1) \cdot (u_T\mu_2) : (\mu_1, \mu_2) \in K_X(w_T, u_I^{-1})\}$. This completes the proof of Theorem 5.1. ∎

CHAPTER 6

APPLICATIONS

In this chapter we apply the techniques developed in the preceding chapters. Our main application is the following theorem.

Theorem 6.1. If w is a freely reduced quadratic word in $(\overline{X}, \circ)$ such that $w = \prod_{i=1}^{n} w_i$ where

(i) $C(w_i)$ is finite for each i ,

and (ii) $\mathrm{Supp}(w_i) \cap \mathrm{Supp}(w_j) = \emptyset$ for $1 \leq i < j \leq n$,

then it can be effectively decided whether any given $u \in \overline{X}$ is an image of w under an endomorphism of $(\overline{X}, \circ)$.

This provides new information concerning the endomorphism problem (see Lyndon [11], p. 283). As consequences of this theorem we will obtain the following results.

Theorem 6.2. For each $n \geq 1$, there is an effective procedure for deciding whether or not a given word u , in a free group, is a product of n squares.

Theorem 6.3. For each $n \geq 1$, there is an effective procedure for deciding whether or not a given word u , in a free group, is a product of n commutators.

COROLLARY By the standard method used to put compact surfaces into normal form, due to Max Dehn, it is well known that every quadratic word can be effectively transformed by Nielsen transformations into the form in Theorems 6.2 or 6.3.

For the case $n = 1$, this latter result is a theorem of Wicks [21]; for $n \geq 2$, it answers positively a question of Wicks (written communication to N.D. Gupta).

Following the proof of these main results, we will apply our results to prove a new result of Lyndon and Newman [13] and to answer a question of our own.

PROOF OF THEOREMS 6.1, 6.2, and 6.3.

To prove Theorem 6.1, we note that since w is reduced and quadratic, each w_i is reduced and quadratic; thus, by a straightforward induction on n using Theorem 5.1, it follows that $C(w)$ is finite. But then by the corollary to Theorem 2.2, Problem III (and hence Problems I and II) is solvable for w and arbitrary u's.

Theorems 6.2 and 6.3 will follow easily from Theorem 6.1 by letting w_i $(1 \leq i \leq n)$ be respectively x_i^2 and $[x_{2i-1}, x_{2i}]$. It clearly suffices to show that $C(x_1^2)$ and $C([x_1, x_2])$ are finite sets $([x_1, x_2] = x_1x_2x_1^{-1}x_2^{-1})$. To see that $C(x_1^2)$ is finite, we refer the reader to Example 2 following Definition 2.5. It is not difficult to calculate $C([x_1, x_2])$:

$$C([x_1, x_2]) = \{1,\ x_1x_2x_1^{-1}x_2^{-1},\ x_1x_2x_3x_2^{-1}x_3^{-1}x_1^{-1},\ x_1x_2x_3x_1^{-1}x_2^{-1}x_3^{-1},$$
$$x_1x_2x_1^{-1}x_3x_2^{-1}x_3^{-1},\ x_1x_2x_3x_4x_2^{-1}x_3^{-1}x_4^{-1}x_1^{-1},$$
$$x_1x_2x_3x_2^{-1}x_4x_3^{-1}x_4^{-1}x_1^{-1},\ x_1x_2x_3x_1^{-1}x_4x_2^{-1}x_3^{-1}x_4^{-1},$$
$$x_1x_2x_3x_4x_2^{-1}x_5x_3^{-1}x_4^{-1}x_5^{-1}x_1^{-1}\}\ .$$

Since this set is finite, our proof is complete. ∎

In [13] Lyndon and Newman comment that if x and y are elements of any group, then

$$[x, y] = (xy)^2\ (y^{-1}x^{-1}y)^2\ (y^{-1})^2\ .$$

This leads to the natural question of whether there is a group in which some commutator cannot be written as a product of fewer than three squares. Their theorem, stated in our notation, is the following.

<u>Theorem 6.4</u>. If $(\overline{X}_2, {}^{\circ})$ is the free group of rank two, freely generated by x_1 and x_2, there are no words $a, b \in \overline{X}_2$ such that $[x_1, x_2] \approx a^2b^2$.

<u>Proof</u>. To prove this result using our techniques, we first suppose that there exist $a, b \in \overline{X}_2$ such that $[x_1, x_2] \approx a^2b^2$. Let σ be the substitution from X into X defined by

$$x_i\sigma = \begin{cases} a & \text{if } i = 1 \\ b & \text{if } i = 2 \\ x_i & \text{if } i \neq 1, 2\ . \end{cases}$$

It is clear that $\overline{(x_1^2x_2^2)\sigma} = [x_1, x_2]$; therefore, by Theorem 2.13, there is a word $u \in C(x_1^2x_2^2)$ such that $u \lesssim [x_1, x_2]$. Note that since $u \lesssim [x_1, x_2]$, it follows that $\ell(u) \leq 4$. The words in $C(x_1^2x_2^2)$ which are of length no more than four are:

$$1,\ x_1^2,\ x_1x_2x_1x_2^{-1},\ x_1x_2x_1^{-1}x_2,\ x_1x_2^2x_1^{-1},\ x_1^2x_2^2,\ x_1x_2^2x_1 .$$

It is easy to see (from the solution to Problem IV in Chapter 2) that, upon c-free substitution, none of these words yields $[x_1, x_2]$. Thus there is no $u \in C(x_1^2x_2^2)$ such that $u \lesssim [x_1, x_2]$; this is a contradiction. ∎

We conclude these applications by answering a question of our own. A word $w \in (\overline{X}, \circ)$ is said to be primitive if there is an automorphism of $(\overline{X}, \circ)$ which sends w to x_1. The following theorem completely characterizes those words which can be sent to x_1 under an endomorphism of $(\overline{X}, \circ)$.

Theorem 6.5. Given $w \in (\overline{X}, \circ)$, the following conditions are equivalent.

(i) There is an endomorphism of $(\overline{X}, \circ)$ which sends w to x_1.

(ii) $C(w) = \{1, x_1\}$.

(iii) $\gcd(w)$ $(= \gcd\{|\sigma_s(w)| : x_s \in \mathrm{Supp}(w)\}) = 1$, where $\sigma_s(w)$ is the exponent sum of w on x_s (see [15], p. 76).

<u>Proof</u>. It is clear that (i) is equivalent to (ii) since $C(w)$ is a complete set of images of w (see Definition 2.1). To see that (ii) is equivalent to (iii), first suppose that $\gcd(w) = 1$; therefore there are integers $\{n_s\}$ such that $\Sigma n_s \sigma_s(w) = 1$. Letting τ be the substitution from X into X defined by

$$x_s\tau = \begin{cases} x_1^{n_s} & \text{if } x_s \in \mathrm{Supp}(w) \\ x_s & \text{otherwise,} \end{cases}$$

it is easy to see that $\overline{w\tau} = x_1$; thus there is a word $u \in C(w)$ such that $u \lesssim x_1$. But this implies that $u = x_i^{\varepsilon}$ for some i. Since each word in $C(w)$ is in normal form we have $x_1 = u \ (\in C(w))$. Now the minimality of x_1 in $(\overline{X}, \lesssim)$ implies that $C(w) = \{1, x_1\}$. Therefore (iii) implies (ii).

To see that (ii) implies (iii), suppose that (iii) does not hold, <u>i.e.</u> that $\gcd(w) \neq 1$. We claim that for each $u \in N(w)$, $T(w)$, or $R(w)$ $\gcd(u) \neq 1$. First note that if u is a (partially) reduced form of w, $\sigma_s(w) = \sigma_s(u)$ for each s; thus $\gcd(u) = \gcd(w)$. If λ is an elementary level substitution, it is easy to see that

$$\{|\sigma_s(w\lambda)| : x_s \in \mathrm{Supp}(w\lambda)\} = \{|\sigma_s(w)| : x_s \in \mathrm{Supp}(w)\} ;$$

thus if ν is a level substitution such that $w\nu \in N(w)$, we have $\gcd(w\nu) = \gcd(w)$. If $u \in T(w)$, it is clear that

$$\{|\sigma_s(u)| : x_s \in \mathrm{Supp}(u)\} \subseteq \{|\sigma_s(w)| : x_s \in \mathrm{Supp}(w)\} ;$$

thus if $\gcd(w) \neq 1$, $\gcd(u) \neq 1$ also. Lastly, suppose that $u = w\rho_{(w;\, i,\, \epsilon;\, j,\, \eta)}$ and $\bar{u} \in R(w)$. Note that u is obtained by adding several occurrences of a "new" variable, x_m, into w (see Definition 2.3). It is easy to see that

$$\sigma_m(u) = \epsilon\,\sigma_i(w) + \eta\,\sigma_j(w)$$

and
$$\sigma_s(u) = \sigma_s(w) \qquad \text{for} \quad s \neq m\ .$$

Therefore, we have $\gcd(u) = \gcd(w)$. Since free reduction preserves gcd, it follows that $\gcd(\bar{u}) = \gcd(w) \quad (\neq 1)$.

Since $C(w)$ consists of words produced by repeated use of T, R, N (and M), it follows that if $\gcd(w) \neq 1$, then $\gcd(u) \neq 1$ for each $u \in C(w)$. But $\gcd(x_1) = 1$; therefore $x_1 \notin C(w)$ which contradicts (ii). ∎

INDEX OF NUMBERED ITEMS

INDEX OF SPECIAL SYMBOLS

B I B L I O G R A P H Y

[1] G. Baumslag, "On a problem of Lyndon", J. London Math. Soc. 35(1960), p. 30-32.

[2] __________, "Residual nilpotence and relations in free groups", J. Algebra 2(1965), p. 271-282.

[3] __________, W.W. Boone and B.H. Neumann, "Some unsolvable problems about elements and subgroups of groups", Math. Scand. 7(1959), p. 191-201.

[4] __________, A. Steinberg, "Residual nilpotence and relations in free groups", Bull. Amer. Math. Soc. 70(1964), p. 283-284.

[5] Ju. I. Hmelevskiĭ, "Systems of equations in a free group", Math. USSR Izvestija Vol. 5 (1971) No. 6, p. 1245-1276.

[6] S. Lipschutz, "Note on independent equation problems in groups", Archiv der Mathematik 22(1971), p. 113-116.

[7] F. Levin, "Solutions of equations over groups", Bull. Amer. Math. Soc. 68(1962), p. 603-604.

[8] R.C. Lyndon, "The equation $a^2b^2 = c^2$ in free groups", Michigan Math. J. 6(1959), p. 89-95.

[9] __________, "Equations in free groups", Trans. Amer. Math. Soc. 96(1960), p. 445-457.

[10] R.C. Lyndon, "Groups with parametric exponents", Trans Amer. Math. Soc. 96(1970), p. 518-533.

[11] __________, "Dependence in groups", Colloquium Mathematicum 14(1966), p. 275-283.

[12] __________, T. McDonough and M. Newman, "On products of powers in groups", Proc. Amer. Math. Soc. (to appear).

[13] __________, M. Newman, "Commutators as products of squares", Proc. Amer. Math. Soc. (to appear).

[14] __________, M.P. Schützenberger, "The equation $a^M = b^N c^P$ in a free group", Michigan Math. J. 9(1962), p. 289-298.

[15] W. Magnus, A. Karrass and D. Solitar, Combinatorial Group Theory, Pure and Applied Math. vol. 13, Interscience, New York, 1966.

[16] P.E. Schupp, "On the substitution problem for free groups", Proc. Amer. Math. Soc. 23(1969), p. 421-424.

[17] M.P. Schützenberger, "Sur l'equation $a^{2+n} = b^{2+m} c^{2+p}$ dans un group libre", C.R. Acad. Sci. Paris 248(1959), p. 2435-2436.

[18] E. Shenkman, "The equation $a^n b^n = c^n$ in a free group", Ann. of Math. (2)70 (1959), p. 562-564.

[19] J. Stallings, "On certain relations in free groups", Notices Amer. Math. Soc. 6(1959), p. 532.

[20] A. Steinberg, Ph.D. thesis, N.Y.U., 1962.

[21] M.J. Wicks, "Commutators in Free Products", J. London Math. Soc. 37(1962), p. 433-444.