SOLVABILITY OF THE CONJUGACY PROBLEM FOR HNN EXTENSIONS OF FINITELY GENERATED FREE ABELIAN GROUPS

By

FARROKH ABEDI

Bachelor of Science in Mathematics
Pars College
Tehran, Iran
1973

Master of Arts
Eastern New Mexico University
Portales, New Mexico
1975

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1983

SOLVABILITY OF THE CONJUGACY PROBLEM FOR HNN EXTENSIONS OF FINITELY GENERATED FREE ABELIAN GROUPS

Thesis Approved:

Benny Evans
Thesis Adviser

Paul Duvall

Wayne B. Powell

Mark Samuel

Norman N Durham
Dean of the Graduate College

PREFACE

This study is concerned with the solvability of the conjugacy problem for HNN extensions of finitely generated free abelian groups. In order to prove the theorem, normal forms and a solution of the word problem for such groups are required. In Chapter II we show that the word problem and generalized word problem for a finitely generated abelian group are solvable. Then the normal form for HNN extension of finitely generated abelian groups immediately follows, and by Corollary 2.1 the word problem is solvable for these groups.

In Chapter III the conjugacy problem for long words is shown to be solvable. To clarify the theorem an example is given at the end of Chapter III.

To complete the algorithm we show that solving the conjugacy problem for short words of the elements of these groups is equivalent to a certain decision problem for polynomials with complex rational coefficients.

The author wishes to express his appreciation to his advisor Professor Benny Evans for his invaluable help and limitless patience in the preparation of this paper. Appreciation is also expressed to Professor Jim Boler for his encouragement and guidance.

An expression of gratitude is also due to the members of my committee and the faculty of the Mathematics Department of Oklahoma State University as each has played a part in making this possible.

A special note of thanks is due to Miss Barbara Newport for an excellent job of typing the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Groups are very often described as quotient groups of free groups: $G = \frac{F}{N}$, where G is generated by X, F is the free group on X and N is the kernel of the epimorphism $F \longrightarrow G$. Now if $w = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n} \in F$ is a normal generator of N, then under the epimorphism $F \longrightarrow G$, $w \longmapsto x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n} = 1 \in G$. The equation $x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n} = 1$ in G is called a relation on the generators x_i. Let R be a set of relations determined by normal generators $\{w_i\}$ for N. We say that the pair <X,R> is a presentation for G, and by a mild abuse of language, we may write G = <X,R>. It is customary to speak of X as a set of defining generators for G and of the equations $r = 1$, for $r \in R$, as a set of defining relations. A presentation <X,R> is finitely generated if X is finite, and is finitely related if R is finite. One says then also that G = <X,R> is finitely generated or finitely related. A presentation <X,R> is finite if both X and R are finite; in this case G = <X,R> is finitely presented. Perhaps the simplest example occurs in the case that G is finite. The multiplication table provides a finite presentation. For generators we take all the elements g_i of G and for defining relations all the equations $g_i g_j = g_k$ that are valid in G. In general we can get a presentation of a group G by taking a distinct generating symbol for each element in the group and using all relations on these generators that are valid in G as the set of defining relators. Thus, every group has a presentation. A presentation can then be thought of as a possibly abbreviated generalization of a multiplication table.

Many important infinite groups have a finite presentation. On the other hand, B. H. Neumann [10, p. 188] has shown that there are uncountably many non-isomorphic groups generated by two elements, from this it follows that there are many finitely generated groups that admit no finite presentation. A group G is determined up to isomorphism by a presentation. However, in general, even a finite presentation may not provide much knowledge about G. For example, the problem of deciding whether a word in G defines the identity element is the first of the following three fundamental decision problems formulated by Max Dehn in 1911 [5]. These problems are important for presentation theory as well as for its applications.

Given a presentation G = <X,R>, we say that G has solvable word problem if for an arbitrary element w of the free group generated by X, we can decide whether or not w defines the identity element of G. Clearly, this is equivalent to deciding whether an arbitrary element w of the free group on X lies in N, the normal closure in F of the set R. The second of Dehn's problems is the conjugacy problem, to decide if arbitrary w_1 and w_2 in G represent conjugate elements of G. The third problem is the isomorphism problem, to decide whether two given finite presentations define isomorphic groups. Dehn (1912) solved all three of these problems for fundamental groups of 2-manifolds.

The word problem has been solved for many classes of presentations. However, there exists finitely presented groups with unsolvable word problems (Boone and Higman [2]). Magnus [10] showed the word problem is solvable for groups with a single defining relation. It is not known whether every presentation with two defining relations has solvable word problem. Nor is it known whether every presentation with a single defining relation has solvable conjugacy problem. It is clear that a solution of the conjugacy problem contains a solution of the word problem. It has been shown that there exist finitely presented groups with solvable word problem but unsolvable conjugacy problem [7].

The isomorphism problem is the most difficult of the three problems of Dehn, and so it is no surprise that this problem is also in general not solvable. There does not even exist an algorithm to decide whether a finitely presented group is trivial [15].

Another decision problem that arises naturally in studying Dehn's three basic problems is the generalized word problem with respect to a finitely generated subgroup H of G. One says the generalized word problem for H in G is solvable if there exists an algorithm which, when given any word w in G, determines whether or not $w \in H$. We shall simply say that H _is solvable in_ G if the generalized word problem for H in G is solvable. Algebraic constructions provide a link between the word problem and generalized word problem. For suppose H is a finitely presented group which has a finitely generated subgroup A with unsolvable generalized word problem (note that H may have solvable word problem). Let H_1 be another copy of H with corresponding subgroup A_1. Now form the free product with amalgamation $G = \langle H * H_1 / u = u_1 \text{ for all } u \in A \rangle$. If w is an arbitrary word of H and w_1 is the corresponding word in H_1, then $ww_1^{-1} = 1$ in G if and only if $w \in A$. Thus if the generalized word problem for A in H is unsolvable, this shows that the word problem for the finitely presented group G is unsolvable. Even though H may have had solvable word problem. On the other hand, it is clear that if the generalized word problem for a group G is solvable, then G has solvable word problem.

Toh [16] has shown that polycyclic groups are subgroup separable. That is if H is a subgroup of the polycyclic group G with $w \notin H$, then there exists an epimorphism $\phi: G \longrightarrow F$ where F is a finite group such that $\phi(w) \notin \phi(H)$. Thus by an algorithm similar to that used by Dyson [5] to solve the word problem for finitely generated residually finite groups, Toh has shown the generalized word problem is solvable for polycyclic groups. However a construction of Mikhlailova [10], [13] shows that the direct product of two free groups may have unsolvable generalized word problem. (In view of the preceding remarks, one can see that there is in fact a fairly elementary group with unsolvable word problem, namely, a certain generalized free product of two copies of $F_1 \times F_2$ where F_1 and F_2 are free.)

In 1949 G. Higman, B. H. Neumann and H. Neumann [10], [12] introduced a construction which is basic to combinatorial group theory. This construction is the Higman-Neumann-Neumann extension, which we shall shorten to HNN extension. Let G be a group, and let A and B be subgroups of G with $\phi: A \longrightarrow B$ an isomorphism. The HNN extension of G relative to A, B and ϕ is the group

$$G^* = <G, t / t^{-1}\ a\ t = \phi(a),\ a \in A>.$$

The group G is called the base of G^*, t is called the stable letter, and A and B are called the associated subgroups.

The conjugacy problem for the HNN extension of finitely generated free groups relative to cyclic subgroups A and B has been shown to be solvable by Larsen [9]. Let

$$G = <a_0, a_1, \ldots, a_k \mid [a_i, a_j] = 1 \quad \forall\ i,j, \quad a_0^{s_0}\ a_1^{s_1} \ldots a_k^{s_k} = 1>$$

where $k \geq 1$, $s_0 \neq 0$, $s_k \neq 0$. Thus G is a free abelian group with one additional relation. That the conjugacy problem for an HNN extension of G (arising from G by adding a new generator t and defining relations $t^{-1}a_i t = a_{i+1}$, $i = 0,1,2,\ldots,k-1$) is solvable has been proved by Britton [4]. A group G is said to be conjugate separable if, whenever x and y are elements of G which are not conjugate, there is a finite homomorphic image of G in which the images of x and y are not conjugate. Formanek [6] showed that polycyclic-by-finite groups are conjugate separable. A group G is said to be abelian-by-cyclic if G has an abelian normal subgroup A with $\frac{G}{A} = T$ cyclic. Boler [1] has shown that finitely generated torsion-free abelian-by-cyclic groups have solvable conjugacy problem. However, there are many examples [11] of groups with solvable conjugacy problem (e.g., free groups) and HNN extensions of these groups that do not have solvable conjugacy problem.

The main result of this paper is that the conjugacy problem for each HNN extension of a finitely generated free abelian group is solvable. (This is a generalization of Britton's result mentioned above.) Let $G^* = \langle G,t \,|\, t^{-1}\ A\ t = \phi(A)\rangle$ be an HNN extension of a finitely generated abelian group G. The proof divides naturally into two cases.

First, for two "long words" u and v in G^* (that is for words that contain the stable letter t), an algorithm is produced which effectively determines whether or not u and v are conjugate in G^*. This part of the proof makes use of Collins' Lemma which provides normal forms for conjugate elements in HNN extensions. The bulk of the work at this stage involves a study of normal forms for integral matrices.

In the second case, for two short words (t-free) u and v, it is the case that u and v are conjugate in G^* if and only if $\phi^n(u) = v$ for some

n. Interestingly, the solution to the conjugacy problem for short words is more complex than one might expect. In order to decide whether or not u and v are conjugate, one takes values n = 1,2,... and checks to see if $\phi^n(u) = v$ for any n. In order to complete the algorithm, some computable upper bound for n must be determined. The search for such bounds is carried on primarily in the ring of polynomials with complex rational coefficients. Techniques from complex analysis and numerical analysis (The Lehmer-Schur algorithm is essential.) are required to complete the work.

CHAPTER II

COLLIN'S LEMMA AND HNN EXTENSIONS

Let $G^* = \langle G, t / t^{-1} a\, t = \phi(a),\ a \in A\rangle$ be an HNN extension. A word in G^* is any sequence $g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ such that $g_i \in G$, $\varepsilon_i = \pm 1$. For the rest of this paper, the letter g, with or without subscripts, will denote an element of G. That is g contains no occurrences of $t^{\pm 1}$. The letter ε with or without subscripts, will denote 1 or -1.

Definition 2.1: *A word* $g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ $(n \geq 0)$ *is said to be reduced if there is no consecutive subword* $t^{-1} g_i t$ *with* $g_i \in A$ *or* $t g_j t^{-1}$ *with* $g_j \in B$.

We shall consider another definition which will allow us to formulate a normal form theorem for HNN extensions. In their original paper, Higman, Neumann and Neumann proved that G is embedded in G^* by the map $g \mapsto g$. The rest of the Normal Form Theorem for HNN extensions was proved by J. L. Britton and is usually referred as Britton's Lemma [3].

Lemma 2.1 (Britton's Lemma): *If the word* $w = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ *is reduced and* $n \geq 1$, *then* $w \neq 1$ *in* G^*.

We need further refinement to actually get normal forms. Choose a set of representatives of the right cosets of A in G, and a set of representatives of the right cosets of B in G. We shall assume that 1 is the representative of both A and B. If $g \in G$, $\overline{g}$ will denote the representative of the coset Ag, and $\hat{g}$ will denote the representative of the coset Bg.

Definition 2.2: *A word* $g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ ($n \geq 0$) *is said to be in normal form if*:

i) g_o *is an arbitrary element of* G,

ii) *if* $\varepsilon_i = -1$, *then* g_i *is a representative of a coset of* A *in* G,

iii) *if* $\varepsilon_i = 1$, *then* g_i *is a representative of a coset of* B *in* G, *and*

iv) *there is no consecutive subword* $t^{\varepsilon} 1 t^{-\varepsilon}$.

The following discussion and example will explain our definition of a normal form. The defining relations

$$t^{-1}at = \phi(a), \quad a \in A \tag{1}$$

of the HNN extension, can be written as

$$t^{-1}a = \phi(a)t^{-1}. \tag{2}$$

By conjugating both sides of (1) by t, the relations (1) can also be written as

$$tbt^{-1} = \phi^{-1}(b), \quad b \in B \tag{3}$$

which are equivalent to

$$tb = \phi^{-1}(b)t \tag{4}$$

These relations allow us to move an element $a \in A$ to the left of at^{-1} by changing a to $\phi(a)$. Similarly, we can move $b \in B$ to the left of at, changing b to $\phi^{-1}(b)$. By working from right to left, we can show that every element of G^* can be written in normal form.

Example. Let $F = <x,y>$ and let $F^* = <x,y,t/t^{-1}xt = y^2>$ be an HNN extension of F. As representatives of cosets of $A = <x>$, choose all freely reduced words on x and y which do not begin with x. As representatives of cosets of $B = <y^2>$, choose all freely reduced words on x and y which do not begin with a power of y except possibly y^1.

Let

$$w = xyt^{-1}x^3yty^5xyt^{-1}x^3y^3.$$

Note that w is reduced. Now we calculate the normal form of w by working from right to left. Since the representative of Ax^3y^3 is y^3 and $t^{-1}x^3 = y^6t^{-1}$, we have

$$w = xyt^{-1}x^3yty^5xy^7t^{-1}y^3.$$

Since the representative of By^5xy^7 is yxy^7, and $ty^4 = x^2t$, we have

$$w = xyt^{-1}x^3yx^2tyxy^7t^{-1}y^3.$$

Since $t^{-1}x^3 = y^6t^{-1}$, we have

$$w = xy^7t^{-1}yx^2tyxy^7t^{-1}y^3$$

is a normal form.

The Normal Form Theorem has two equivalent statements (I) and (II) below. Note that (I) is the combination of the theorem of Higman, Neumann, and Neumann and Britton's Lemma.

Theorem 2.1: (The Normal Form Theorem for HNN Extensions). *Let* $G^* = <G,t/t^{-1}at = \phi(a), \ a \in A>$ *be an HNN extension*. *Then*

(I) *The group* G *is embedded in* G^* *by the map* g ⟶ g. *If* $w = g_0t^{\varepsilon_1}\dots t^{\varepsilon_n}g_n = 1$ *in* G^* *where* $n \geq 1$, *then* w *is not reduced*.

(II) *Every element* w *of* G^* *has a unique representation as a word in normal form*.

Proof: See R. Lyndon and P. Schupp [10].

If w is a word of G^* we can write

$$w = g_0 t^{\varepsilon_1} \dots t^{\varepsilon_n} g_n$$

where the above sequence is not necessarily reduced. Consider operations, called t-reductions, of the form

(i) replace a subword of the form $t^{-1}gt$, where $g \in A$, by $\phi(g)$, or

(ii) replace a subword of the form tgt^{-1}, where $g \in B$, by $\phi^{-1}(g)$.

A finite number of t-reductions leads from w to a word

$$w' = g_0' t^{\gamma_1} \dots t^{\gamma_k} g_k'$$

where the indicated sequence is reduced. If $k > 0$, Britton's Lemma says that w', and thus w, is not equal to 1 in G^*. If $k = 0$, then the theorem of Higman, Neumann, and Neumann says that $w' = 1$ in G^* only if $w' = 1$ in G. The process of performing t-reductions is effective if we can tell what words of G represent elements of A or B, and if we can effectively calculate the functions ϕ and ϕ^{-1}. (This later condition will always be satisfied if A and B are finitely generated.) Thus we have:

Corollary 2.1. Let $G^* = \langle G, t / t^{-1}At = B, \phi \rangle$ be an HNN extension. If G has solvable word problem and the generalized word problems for A and B in G are solvable, and ϕ and ϕ^{-1} are effectively calculable, then G^* has solvable word problem.

Lemma 2.2. Let $u = g_0 t^{\varepsilon_1} \dots t^{\varepsilon_n} g_n$ and $v = h_0 t^{\gamma_1} \dots t^{\gamma_m} h_m$ be reduced words, and suppose that $u = v$ in G^*. Then $m = n$ and $\varepsilon_i = \gamma_i$, $i = 1,2, \dots, n$.

Proof: Since u = v, we have

$$1 = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n h_m^{-1} t^{-\gamma_m} \ldots t^{-\gamma_1} h_0^{-1}$$

Since u and v are reduced, the only way the indicated sequence can fail to be reduced is that $\varepsilon_n = \gamma_n$ and $g_n h_m^{-1}$ is in the appropriate subgroup A or B. Making successive t-reductions we see that each $\varepsilon_i = \gamma_i$ and n = m.

Definition 2.3. *Let* z *be an element of* G^*, *and let* w *be any reduced word of* G^* *which represents* z. *If* $w = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$, *the length of* z, *written* $|z|$, *is the number* n *of occurrences* of $t^{\pm 1}$ *in* w. In view of the Lemma 2, $|z|$ is well defined. Under this definition, all elements g of the base G of G^* have length zero.

Definition 2.4. *An element* $w = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ *is cyclically reduced if all cyclic permutations of the word* $g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_n$ *are reduced*. Clearly every element of G^* is conjugate to a cyclically reduced element.

Theorem 2.2. (The Torsion Theorem for HNN Extensions). *Let* $G^* = \langle G, t / t^{-1} A t = B, \phi \rangle$ *be an HNN extension*. *Then every element of finite order in* G^* *is a conjugate of an element of finite order in the base group* G. *Thus* G^* *has elements of finite order* n *only if* G *has elements of order* n.

Proof: If u is an element of G^*, let $v = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n}$ be a cyclically reduced conjugate of u. If $n \geq 1$, then

$$v^m = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} \ldots g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n}$$

is reduced, and so by Britton's Lemma, $v^m \neq 1$.

The conjugacy theorem for HNN extensions is due to D. J. Collins [10], and is usually called Collins' Lemma. Due to the importance of the Collins' Lemma, which plays an important role in solving our problem, we exhibit the complete proof.

Theorem 2.3. (The Conjugacy Theorem for HNN Extensions). *Let* $G^* = \langle G, t/t^{-1}At = B, \phi\rangle$ *be an HNN extension*. *Let* $u = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n}$, $n \geq 1$, *and* v *be conjugate cyclically reduced elements of* G. *Then* $|u| = |v|$, *and* u *can be obtained from* v *by taking a suitable cyclic permutation* v^* *of* v, *which ends in* t^{ε_n}, *and then conjugating by an element* z, *where* $z \in A$ *if* $\varepsilon_n = -1$, *and* $z \in B$ *if* $\varepsilon_n = 1$.

Proof: We will prove by induction on $|c|$ that if v^* is any cyclic permutation of v which ends in a t-symbol and $cv^*c^{-1} = u$, then the conclusion of the theorem holds. If $|c| = 0$ we have

$$g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} = ch_0 t^{\delta_1} \ldots t^{\delta_m} c^{-1} \qquad \text{or}$$

$$1 = g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} c t^{-\delta_m} \ldots t^{-\delta_1} h_0^{-1} c^{-1} .$$

Since the only possible t-reduction is $t^{\varepsilon_n} c t^{-\delta_m}$, we must have $c \in A$ if $\varepsilon_n = -1$, and $c \in B$ if $\varepsilon_n = 1$. By considering successive t-reductions, we have exactly as in the proof of Lemma 2, that $n = m$ and, indeed, that $\delta_i = \varepsilon_i$, $i = 1,2,\ldots,n$.

Now suppose that c has reduced form

$$c = c_0 t^{\gamma_1} \ldots t^{\gamma_{k-1}} c_{k-1} t^{\gamma_k} c_k$$

where $k \geq 1$. We have

$$u = c_0 t^{\gamma_1} \ldots t^{\gamma_{k-1}} c_{k-1} t^{\gamma_k} c_k h_0 t^{\delta_1} h_1 \ldots h_{m-1} t^{\delta_m} c_k^{-1} t^{-\delta_k} c_{k-1}^{-1} \ldots t^{-\delta_1} c_0^{-1} . \quad (1)$$

Since u is cyclically reduced some t-reduction must be applicable to the right hand side of the above equation. The only possiblities are $t^{\gamma_k}c_kh_0t^{\delta_1}$ and $t^{\delta_m}c_k^{-1}t^{-\delta_k}$. For definiteness, assume that $\gamma_k = -1$ and that $c_kh_0 \in A$. Then $\delta_1 = 1$ and

$$t^{-1}c_kh_0t = \phi(c_kh_0) = b \in B. \tag{2}$$

Using the above equation, and the fact that $\gamma_k = -1$, we have

$$c_k^{-1}t^{-\gamma_k} = h_0tb^{-1}. \tag{3}$$

Replacing equations (2) and (3) in equation (1) we have

$$u = c_0t^{\gamma_1}\ldots t^{\gamma_{k-1}}c_{k-1}b(h_1t^{\delta_2}\ldots h_{m-1}t^{\delta_m}h_0t)b^{-1}c_{k-1}^{-1}t^{-\gamma_{k-1}}\ldots t^{-\gamma_1}c^{-1}.$$

Since the term in the middle is a cyclic permutation of v, the result follows by the induction hypothesis.

Finally, when $u = zv^*z^{-1}$ where $z \in A$ or B, Lemma 2 shows that the sequence of $t^{\pm 1}$ in v^* is exactly the same as in u.

Remark. If $|u| = 0$ and v is conjugate cyclically reduced element of G, then $|u| = |v| = 0$ and we have

$$u = g_0t^{\varepsilon_1}\ldots t^{\varepsilon_n}g_nvg_n^{-1}t^{-\varepsilon_n}\ldots t^{-\varepsilon_1}g_0^{-1}$$

for some $z = g_0t^{\varepsilon_1}\ldots t^{\varepsilon_n}g_n$ in G^*. Furthermore there exists a sequence $a_n, a_{n-1}, \ldots, a_1, a_0$ where a_i's are in the appropriate subgroups A or B such that $t^{\varepsilon_n}g_nvg_n^{-1}t^{-\varepsilon_n} = a_n$ and $t^{\varepsilon_i}g_ia_ig_i^{-1}t^{-\varepsilon_i} = a_{i-1}$ for $i = 1,2,\ldots,$ n-1, and $a_0 = g_0^{-1}ug_0$.

Decision Problems for Finitely Generated Abelian Groups

The main purpose of this section is to show that the generalized word problem for finitely generated abelian groups is solvable. Before proving results about the finitely generated abelian groups, let us recall some essentials in linear algebra. Part of the subject consists of the study of homomorphisms of finitely generated free modules, and the relationship between such homomorphisms and matrices. The investigation of the connection between two matrices that represent the same homomorphism relative to different bases leads to the concepts of equivalence and similarity of matrices.

Theorem 2.4. *Let* R *be a commutative ring with identity. Let* E *and* F *be free left* R-*modules with finite bases of* n *and* m *elements respectively. Let* M *be the left* R-*module of all* n×m *matrices over* R. *Then there is an isomorphism of left* R-*modules*

$$\mathrm{Hom}_R(E,F) \cong M.$$

Proof: Let $\{u_1,\ldots,u_n\}$ be a basis of E, $\{v_1,\ldots,v_m\}$ a basis of F and $f \in \mathrm{Hom}_R(F,E)$. There are elements r_{ij} of R such that

$$f(u_i) = \sum_{j=1}^{m} r_{ij}v_j, \qquad i = 1,2,\ldots,n.$$

Define a map β: $\mathrm{Hom}_R(E,F) \longrightarrow M$ by $f \longmapsto A$, where A is the n×m matrix (r_{ij}). The r_{ij} are uniquely determined since $\{v_1,\ldots,v_m\}$ is a basis of F, hence β is well-defined. Obviously β is an additive homomorphism. If $\beta(f) = 0$, then $f(u_i) = 0$ for every basis element u_i, whence $f = 0$. Thus β is a monomorphism. Given a matrix $(r_{ij}) \in M$, define f: $E \longrightarrow F$

by $f(u_i) = r_{i1}v_1 + r_{i2}v_2 + \dots + r_{im}v_m$ $(i = 1,2,\dots,n)$. Since E is free, this uniquely determines f as an element of $\mathrm{Hom}_R(E,F)$. By construction $\beta(f) = (r_{ij})$. Therefore β is surjective and hence an isomorphism. It is easy to verify that $\beta(rf) = r\beta(f)$, and hence β is an R-module isomorphism.

Remark. If $u = a_1u_1 + \dots + a_nu_n \in E$ $(a_i \in R)$, then

$$f(u) = f(\sum_{i=1}^{n} a_iu_i) = \sum_{i=1}^{n} a_if(u_i) = \sum_{i=1}^{n} a_i(\sum_{j=1}^{m} r_{ij}v_j)$$

$$= \sum_{j=1}^{m} (\sum_{i=1}^{n} a_ir_{ij})v_j = \sum_{j=1}^{m} b_jv_j,$$

where $b_j = \sum_{i=1}^{n} a_ir_{ij}$. Thus if X is the 1×n matrix $(a_1,a_2,\dots,a_n)$ and Y is the 1×n matrix $(b_1,b_2,\dots,b_m)$, then Y is precisely the matrix product XA, where A is the matrix (r_{ij}).

Definition 2.5: *Two* n×m *matrices* A *and* B *are said to be equivalent if and only if there exist invertible matrices* P *and* Q *such that* A = PBQ.

Definition 2.6: *Let* A *be a matrix over a commutative ring* R *with identity. Each of the following is called an elementary row (column) operation on* A.

i) permuting rows (columns);

ii) multiply a row (column) of A by a unit $c \in R$;

iii) for $r \in R$ and $i \neq j$, add r time row (column) j to row (column) i.

An n×n *elementary matrix (transformation)* is a matrix that is obtained by performing exactly one elementary row (or column) operation on the identity matrix I_n.

Theorem 2.5. *If* B *is the matrix obtained from an* n×m *matrix* A *by performing a finite sequence of elementary row and column operations, then* B *is equivalent to* A.

Proof: Since each row (column) operation used to obtain B from A is given by left (right) multiplication by an appropriate elementary matrix, we have $B = (E_p \ldots E_1)A(F_1 \ldots F_q) = PAQ$ with each E_i, F_j an elementary matrix and $P = E_p \ldots E_1, Q = F_1 \ldots F_q$. Since P and Q are products of invertable matrices, P and Q are invertable.

Theorem 2.6. *If* A *is an* n×m *matrix of rank* $r > 0$ *over a Euclidean domain* R, *then* A *is equivalent to a matrix of the form* $\begin{pmatrix} L_r & 0 \\ 0 & 0 \end{pmatrix}$, *where* L_r *is an* r×r *diagonal matrix with nonzero diagonal entries* $d_1, d_2, \ldots, d_r$ *such that* $d_1 | d_2 | \ldots | d_r$.

Sketch of Proof: For a constructive solution to this problem with more detail refer to [8]. The following is an outline of the process:

Permuting rows, columns and changing signs if necessary, we arrive at a matrix (a_{ij}) such that $0 < a_{11} \leq |a_{ij}|$ for $a_{ij} \neq 0$. Suppose

$$a_{21} = k_2 a_{11} + r_{21} \qquad 0 \leq r_{21} < a_{11}.$$

Then if we add $(-k_2)R^{(1)}$ to $R^{(2)}$ (where $R^{(i)}$ denotes the ith row) our new $a_{21} = r_{21}$. If $r_{21} \neq 0$, it is smaller than a_{11} and we interchange the new $R^{(2)}$ and $R^{(1)}$. We repeat the above procedure until we obtain an $a_{21} = 0$. Applying the procedure to the third row, we obtain an $a_{31} = 0$. Continuing in this way we have a new matrix (a_{ij}) in which $a_{21} = a_{31} = \ldots = a_{n1} = 0$. Next apply the procedure to the first row using column operations provided that each time the first column is replaced by another column, we apply row operations so that as before the new first column has $a_{21} = a_{31} = \dots = a_{n1} = 0$. Continuing in this way, we obtain a new matrix (a_{ij}) in which $a_{i1} = a_{1j} = 0$, $i,j \neq 1$. If now there is some a_{ij} such that

$$a_{ij} = k_{ij}a_{11} + r_{ij} \qquad 0 < r_{ij} < a_{11}$$

add the jth column to the first column, making the new $a_{i1} = a_{ij}$ and a_{11} is unchanged. Proceeding as before, we obtain a new $a_{11} = r_{ij}$. Note that each of the above operations which changes a_{11} decreases it. Hence, after finitely many replacements, we arrive at a matrix (a_{ij}) in which

$$a_{i1} = a_{1j} = 0, \qquad i,j \neq 1,$$

and a_{11} divides a_{ij}. Starting over again with the submatrix obtained by ignoring the first row and column, we arrive at a matrix (a_{ij}) in which

$$a_{i1} = a_{1j} = 0, \qquad i,j \neq 1 \qquad a_{i2} = a_{2j} = 0, \qquad i,j \neq 2,$$

and a_{11} divides a_{22} divides a_{ij}, $i,j > 1$. Continuing in this way, we finally arrive at a matrix in which a_{ii} divides a_{jj}, $i < j \leq n$ and $a_{ij} = 0$, $i \neq j$.

As an illustration of the above process, let

$A = \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix}$ be a 2×3 matrix over Z. We then have successively,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \overset{c^{(3)} \longrightarrow c^{(1)}}{\begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix}}$$

$$R^{(2)} \longrightarrow R^{(1)}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -4 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$

$$4R^{(1)}+R^{(2)} \longrightarrow R^{(2)}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ -4 & 0 & 2 \end{pmatrix}$$

$$-2C^{(1)}+C^{(2)} \longrightarrow C^{(2)}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 8 & 2 \end{pmatrix}$$

$$C^{(3)} \longrightarrow C^{(2)}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 2 \end{pmatrix}$$

$$-4C^{(2)}\ C^{(2)} \longrightarrow C^{(3)}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 & -4 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -4 \\ 0 & 0 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

Lemma 2.3. *Let* G *be a finitely generated free abelian group with a basis* $\{x_1,\ldots,x_n\}$ *and* H *a subgroup of* G *with a basis* $\{y_1,\ldots,y_m\}$. *Then there exist a basis* $\{u_1,\ldots,u_n\}$ *of* G *and integers* $d_1, d_2, \ldots d_n$ *with* $d_i \geq 0$ *and* $d_1 | d_2 | \ldots | d_n$ *such that* H *is generated by* $d_1u_1, d_2u_2, \ldots, d_nu_n$.

Proof: Define a map $f\colon G \longrightarrow H$, by

$$f(x_i) = y_i = \sum_{j=1}^{m} r_{ij}x_j, \qquad i = 1,2,\ldots,n.$$

The matrix of f with respect to the basis $\{x_1,x_2,\ldots,x_n\}$ is $A = (r_{ij})$. The rows of A represent the generators of the subgroup H relative to the basis $\{x_1,x_2,\ldots,x_n\}$. If $H \neq 0$, then by Theorem 2.6 A may be changed via a finite sequence of elementary row and column operations, to a diagonal matrix

$$D = \begin{pmatrix} d_1 & 0 & & & \\ & d_2 & 0 & & \\ 0 & & \ddots & & \\ & 0 & & \ddots & \\ & & & & d_n \end{pmatrix}$$

such that $d_i \geq 0$ for all i and $d_1 | d_2 \ldots | d_n$. Furthermore D is equivalent to A, that is $PAQ = D$ for some invertible matrices P and Q. Let $g\colon G \longrightarrow G$ be the isomorphism with matrix P^{-1} relative to $X = \{x_1,x_2,\ldots,x_n\}$ and $h\colon H \longrightarrow H$ the isomorphism with matrix Q relative to $Y = \{y_1,\ldots y_n\}$. Then $g(X) = \{g(x_1),\ldots,g(x_n)\}$ is also a basis of G and P^{-1} is the matrix of 1_G relative to the bases $g(X) = \{g(x_1),\ldots,g(x_n)\}$ and $X = \{x_1,\ldots,x_n\}$. Similarly Q is the matrix of 1_H relative to the bases $h(Y)$ and Y, whence $P = (P^{-1})^{-1}$ is the matrix of 1_G relative to X and $g(X)$. Schematically we have

$$\begin{tikzcd}
G \arrow[r, "1_G"] & G \arrow[r, "f"] & H \arrow[r, "1_H"] & H \\
\cup\mid & \cup\mid & \cup\mid & \cup\mid \\
g(X) \arrow[r, phantom, "P"'] & X \arrow[r, phantom, "A"'] & Y \arrow[r, phantom, "Q"'] & h(Y)
\end{tikzcd}$$

Thus the matrix of $f = 1_H \circ f \circ 1_G$ relative to the bases g(X) and h(Y) is PAQ = D, whence $f(g(x_i)) = d_i(g(x_i)) = h(y_i)$ for i = 1,2,...,n. Let $u_i = g(x_i)$, i = 1,2,...,n, then G is generated by $\{u_1,\dots,u_n\}$ and H is generated by $\{d_1u_1,d_2u_2,\dots,d_nu_n\}$ such that $d_i \geq 0$ and $d_1|d_2\dots|d_n$.

Lemma 2.4. *Every finitely generated abelian group of rank* n *is the direct sum of* r *infinite cyclic groups and* n-r *finite cyclic groups of orders* $d_1,d_2,\dots,d_{n-r}$ *where* $d_1|d_2|\dots|d_{n-r}$.

Proof: Let $G = <x_1,\dots,x_n/y_1 = 0,\dots,y_m = 0>$ be a presentation of a finitely generated abelian group. Let F be the free abelian group on the set $\{x_1,\dots,x_n\}$ and K be the subgroup of F generated by $y_1,y_2,\dots,y_m$. Note that $G \cong \frac{F}{K}$. By Lemma 2.2, F is generated by elements $u_1,u_2,\dots,u_n$ such that K is generated by $d_1u_1,\dots,d_ru_r$ where $d_i > 0$ and d_i/d_{i+1}. Consequently

$$G \cong \frac{F}{K} \approx (Uu_1 \oplus \dots \oplus Uu_n)/(Ud_1u_1 \oplus \dots \oplus Ud_ru_r \oplus 0 \oplus \dots \oplus 0)$$

$$\cong U/d_1U \oplus \dots \oplus U/d_rU \oplus U/0 \oplus \dots \oplus U/0$$

$$= U_{d_1} \oplus \dots \oplus U_{d_r} \oplus U \oplus \dots \oplus U$$

where the rank of $(U \oplus \dots \oplus U)$ is n-r and $d_1|d_2\dots|d_r$.

Lemma 2.5. *A finitely generated abelian group* G *has solvable word problem*.

Proof: Let G be a finitely generated abelian group of rank n. By Lemma 2.3, G is generated by elements $x_1,\dots,x_n$ with relators $d_1x_1,\dots,$ d_nx_n where $d_i \geq 0$ and d_i divides d_{i+1}. An algorithm for the word problem of G is the following. Given $g \in G$, then g can be uniquely written in the form $g = \sum_{i=1}^{n} r_i x_i$ for some integers $r_1, \ldots, r_n$. Now $g \neq 0$ in G if and only if $d_j \nmid r_j$ for some j and $r_j \neq 0$.

Corollary 2.2. The generalized word problem for a finitely generated free abelian group G is solvable.

Proof: Let H be a subgroup of G. Then the generalized word problem for G relative to H is solvable if and only if the word problem is solvable for the finitely generated group $K \cong \frac{G}{H}$.

Definition 2.7. A group G is residually finite, if for every non-trivial element $g \neq 0$ of G, there is a homomorphism ϕ from G into a finite group K such that $\phi(g) \neq 0$. The choice of K and ϕ depends, of course, on the element g.

Definition 2.8. A group G is said to be subgroup separable if, H is a subgroup of G and $w \notin H$, then there exists an epimorphism $\phi\colon G \longrightarrow K$ where K is a finite group such that $\phi(w) \notin \phi(H)$.

Lemma 2.6. A finitely generated abelian group G is residually finite.

Proof: Let us assume that G is a finitely generated abelian group with a presentation $G = \langle x_1, \ldots, x_n / d_1 x_1, \ldots, d_n x_n \rangle$ where $d_i \geq 0$ and $d_i | d_{i+1}$. Let $g \neq 0$ be a non-trivial element of G, then $g = \sum_{i=1}^{n} r_i x_i$, where $d_j \nmid r_j$ for some j and $r_j \neq 0$. If x_j is a free generator, map G onto the cyclic group on y of order r_j+1 by sending $x_j \longmapsto y$ and $x_i \longmapsto 1$, $j \neq i$. If $d_j \neq 0$, map G onto the cyclic group on y of order d_j by sending $x_j \longmapsto y$ and $x_i \longmapsto 1$, $j \neq i$. This clearly gives the required epimorphism.

Lemma 2.7. *A finitely generated abelian group* G *is subgroup separable*.

Proof: Let H be a subgroup of G and $g \notin H$. Let $p: G \longrightarrow \frac{G}{H}$ be the projection map, then $p(g) \neq 0$ in the finitely generated abelian group $\frac{G}{H}$. By Lemma 4 there exists an epimorphism f from $\frac{G}{H}$ into a finite group K such that $f(p(g)) \neq 0$ in K. Let $\phi = f \circ p$, then ϕ is an epimorphism from G into K such that $\phi(g) \notin \phi(H)$.

Lemma 2.8. *The generalized word problem is solvable for a finitely generated abelian group* G.

Proof: Let $G = <x_1,x_2,\ldots,x_n/[x_i,x_j] = 1 \; \forall \; i,jr_1 = 1,\ldots,r_m = 1>$, and let H be a subgroup of G. Since H is finitely generated abelian group, the set of elements in H can be effectively enumerated. We show that the set of elements not in H is also effectively enumerable. A homomorphism from G into a finite abelian group K is completely determined by its effect on the generators $x_1,\ldots,x_n$. Thus a candidate for a homomorphism of G into K is an n-tuple $(k_1,\ldots,k_n)$ of elements of K. The map $x_i \longmapsto k_i$ is a homomorphism if and only if each r_i goes to 0. Since we can solve the word problem in K, we can check this condition. Thus we can effectively enumerate the set $\phi_1,\phi_2,\ldots$ of all homomorphisms of G into finite groups. We can thus enumerate all images $\phi_i(g)$ where g is an element on the generators of G. If some $\phi_i(g) \notin \phi_i(H)$, then $g \notin H$, and we put g on the list of elements not in H. Since G is subgroup separable, if g is any element of G not in H, there exists some ϕ_i with $\phi_i(g) \notin \phi_i(H)$. Thus we list all elements in G not in H.

CHAPTER III

CONJUGACY PROBLEM FOR LONG WORDS

Our intention in this chapter is to study the conjugacy problem for long words in HNN extension of finitely generated free abelian groups.

We use the notation

$$G^* = \langle G, t / t^{-1}At = B, \phi \rangle$$

for an HNN extension of a finitely generated free abelian group G relative to subgroups A and B of G with ϕ: A $\longrightarrow$ B an isomorphism.

Before proving results about these HNN constructions, let us recall some terminology about elements of such groups. Let $w \in G^*$ be a given word. If neither t nor t^{-1} occurs in w, then w is called t-free (short word). If w is not t-free (long word), then $w = w_0 t^{\varepsilon_1} w_1 \ldots t^{\varepsilon_n} w_n$ where each w_i is t-free. If $w \in G^*$ contains no subword $t^{\varepsilon} w_i t^{-\varepsilon}$ with $w_i \in A$ for $\varepsilon = 1$ or $w_i \in B$ for $\varepsilon = -1$, then w is called t-reduced. If all cyclic permutations of w are t-reduced, then w is called cyclically t-reduced. Now if $w = w_0 t^{\varepsilon_1} \ldots t^{\varepsilon_n} w_n$ $(n \geq 0)$ with w_i a coset representative of A if $\varepsilon_i = -1$ and w_j a coset representative of B if $\varepsilon_j = 1$, then w is said to be in normal form.

Our study of conjugacy begins with considering two trivial cases. Let $G^* = \langle G, t / t^{-1}At = B, \phi \rangle$ be an HNN extension of a finitely generated free abelian group. By Corollary 2.1, Lemma 2.5 and Corollary 2.2 the word problem for G^* is solvable. Let u' and v' be any two elements of

G^*. We can effectively replace u' and v' by conjugates u and v which are cyclically reduced. If both u and v are 1, they are certainly conjugate in G, while if one of u or v is 1 and the other is not, they are not conjugate in G^*. The conjugacy problem is thus reduced to considering pairs of non-trivial cyclically reduced words.

Let $u = u_0 t^{\varepsilon_1} \ldots u_{n-1} t^{\varepsilon_n}$ $(n \geq 1)$, and v be conjugate cyclically reduced elements of G^*. Then by Collins' Lemma (Theorem 2.3) we know that $|u| = |v|$ and $u = z^{-1} v^* z$ for some $z \in A$ or $z \in B$ and cyclic permutation v^* of v. Thus if we can decide for any cyclic permutation v^* of v whether or not any z satisfies

$$u = z^{-1} v^* z,$$

then we can solve the conjugacy problem for all cyclically reduced elements of G^* that are not t-free.

Before proceeding with more complex words we need two lemmas.

Lemma 3.1. _Let_ G _be the free abelian group generated by_ $x_1, x_2, \ldots, x_n$ _and_ H _the subgroup generated by_ $d_1x_1, d_2x_2, \ldots, d_nx_n$ _where_ $d_i \geq 0$ _and_ $d_i | d_{i+1}$. _There is an algorithm that will construct for each_ g _in_ G _elements_ h _of_ H _and_ c _of_ $\{\sum_{i=1}^{n} r_i x_i \,|\, 0 \leq r_i \leq d_i\}$ _such that_ $g = h + c$.

Proof: Let $g \in G$. Then g can be written uniquely in the form $g = \sum_{i=1}^{n} u_i x_i$ for some integers $u_i \geq 0$, $i = 1,2,\ldots,n$. By division algorithm, $u_i = q_i d_i + r_i$, $0 \leq r_i < d_i$, $i = 1,2,\ldots,n$. Then

$$g = \sum_{i=1}^{n} u_i x_i = \sum_{i=1}^{n} (q_i d_i + r_i) x_i = \sum_{i=1}^{n} q_i (d_i x_i) + \sum_{i=1}^{n} r_i x_i = h + c,$$

where $h = \sum_{i=1}^{n} q_i (d_i x_i) \in H$ and $\sum_{i=1}^{n} r_i x_i = c$.

Lemma 3.2. *Let* G *be a finitely generated abelian group*, A *and* B *subgroups of* G *and* ϕ: $A \longrightarrow B$ *an isomorphism*. *Given* g *in* G, *there is an algorithm that will determine all* $a \in A$ *and* $z \in G$ *such that* $g + \phi(z) = a$.

Proof: Assume rank of $G = n$. By Lemma 2.3 G is generated by elements $x_1, x_2, \ldots, x_n$ such that A is generated by $\delta_1 x_1, \ldots, \delta_n x_n$ where $\delta_i \geq 0$ and δ_i divides δ_{i+1}.

Let $M = (d_{ij})$ be the matrix of ϕ with respect to this basis. Notice that the matrix M is effectively computable from the presentation of G^*. Let $g = b_1 x_1 + \ldots + b_n x_n$ be a given element of G. Suppose that

$$g + \phi(z) = a,$$

for some z and a in A, and suppose that

$$z = c_1(\delta_1 x_1) + c_2(\delta_2 x_2) + \ldots + c_n(\delta_n x_n) \text{ and}$$

$$a = a_1(\delta_1 x_1) + a_2(\delta_2 x_2) + \ldots + a_n(\delta_n x_n)$$

for some integers a_i and c_i, $i = 1,2,\ldots,n$. We will find an algorithm to enumerate c_is and a_is.

Consider the equation

$$b_1 x_1 + \ldots + b_n x_n + \phi(c_1(\delta_1 x_1) + \ldots + c_n(\delta_n x_n)) = a_1(\delta_1 x_1) + \ldots + a_n(\delta_n x_n). \qquad (1)$$

Since ϕ is a homomorphism, we have

$$\phi(c_1(\delta_1 x_1) + \ldots + c_n(\delta_n x_n)) = c_1\phi(\delta_1 x_1) + c_2\phi(\delta_2 x_2) + \ldots + c_n\phi(\delta_n x_n)$$

$$= c_1(\alpha_{11}x_1 + \ldots + \alpha_{1n}x_n) + \ldots + c_n(\alpha_{n1}x_1 + \ldots + \alpha_{nn}x_n)$$

$$= (c_1, c_2, \ldots, c_n)\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} .$$

Substituting the above equation in Equation (1), we have

$$(b_1, \ldots, b_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + (c_1, \ldots, c_n)\begin{bmatrix} \alpha_{11} & \alpha_{1n} \\ \vdots & \vdots \\ \alpha_{n1} & \alpha_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (a_1\delta_1, \ldots, a_n\delta_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$(b_1, \ldots, b_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + (c_1, \ldots, c_n, a_1, \ldots, a_n)\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \\ \alpha_{n1} & \cdots & \alpha_{nn} \\ -\delta_1 & & \\ & -\delta_2 & 0 \\ 0 & \cdots & -\delta_n \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = 0$$

where the matrix $N = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \\ \alpha_{n1} & \cdots & \alpha_{nn} \\ -\delta_1 & \ddots & 0 \\ 0 & & -\delta_n \end{bmatrix}$ is $2n \times n$.

Since $x_1, \ldots, x_n$ freely generate G, we have

$$(c_1,\ldots,c_n,a_1,\ldots,a_n)\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \\ \alpha_{n1} & \cdots & \alpha_{nn} \\ -\delta_1 & \ddots & 0 \\ 0 & & -\delta_n \end{bmatrix} = (-b_1,\ldots,-b_n). \qquad (2)$$

By Prop. 2.6 the matrix N may be changed via a finite sequence of elementary row and column operations to a diagonal matrix

$$D = \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & d_2 & \\ 0 & \ddots & d_n \\ 0 & 0 & 0 \end{bmatrix}$$

such that $d_i \geq 0$, $d_i | d_{i+1}$ and $PNQ = D$ where $P = (\gamma_{ij})$ and $Q = (\beta_{ij})$ are invertable $2n\times 2n$ and $n\times n$ matrices, respectively. If U is the $1\times 2n$ matrix $(c_1,c_2,\ldots,c_n,a_1,\ldots,a_n)$ and V is the $1\times n$ matrix $(-b_1,\ldots,-b_n)$, then V is precisely the product $U\cdot N$. Using the above observation, and the fact that $N = P^{-1}DQ^{-1}$ equation (2) becomes:

$$UP^{-1}DQ^{-1} = V. \qquad (3)$$

Let $UP^{-1} = W$, where here W is a $1\times 2n$ matrix $(f_1,\ldots,f_n,f_{n+1},\ldots,f_{2n})$. Thus, equation (3) becomes

$$WD = VQ \qquad (4)$$

Equivalently,

$$(f_1,\ldots,f_n,f_{n+1},\ldots,f_{2n})\begin{bmatrix} d_1 & \cdots & 0 \\ & \ddots & \\ 0 & & d_n \\ 0 & \cdots & 0 \end{bmatrix} = (-b_1,\ldots,-b_n)\begin{bmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{bmatrix} \quad (5)$$

From Equation (5) we obtain the following system of n linear equations.

$$f_i d_i = \sum_{j=1}^{n} -b_j \beta_{ji}, \quad i = 1,2,\ldots,n \text{ and } f_{n+1},f_{n+2},\ldots,f_{2n}$$

are completely arbitrary. Equation (1) is valid if and only if the above system of n linear equations is consistant. Solving for f_i, $i = 1,2,\ldots,n$, we obtain $f_1 = z_1$, $f_2 = z_2,\ldots,f_n = z_n$ with at most n arbitrary values. Of course $f_{n+1} = z_{n+1},\ldots,f_{2n} = z_{2n}$ are all arbitrary. Substituting these solutions in the equation $UP^{-1} = W$ and then solving for U, we have

$$(c_1,c_2,\ldots,c_n,a_1,\ldots,a_n) = (z_1,\ldots,z_n,z_{n+1},\ldots,z_{2n})\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1,2n} \\ \vdots & \cdots & \\ \alpha_{2n,1} & \cdots & \alpha_{2n,2n} \end{bmatrix} \quad (6)$$

Solving the above equation leads us to the following solutions.

$$c_i = \sum_{j=1}^{2n} z_j \alpha_{ji} \qquad i = 1,2,\ldots,n \text{ and}$$

$$a_k = \sum_{j=1}^{2n} z_j \alpha_{j,n+k} \qquad k = 1,2,\ldots,n.$$

Theorem 3.1. *Let* $u = u_0 t^{\varepsilon_1} \ldots u_{m-1} t^{\varepsilon_m}$ $(m \geq 1)$ *and* v *be cyclically reduced elements of the HNN group*

$$G^* = \langle G, t / t^{-1}At = B, \phi\rangle$$

with G *a finitely generated free abelian group of rank* n. *There is an algorithm that will decide whether* u *and* v *are conjugate in* G^*.

Proof: Suppose that $|u| = |v|$ and there is a cyclic permutation v^* of v which ends in t^{ε_n} and has the same number of t-symbols as in u. Now if either of these two conditions fail, we conclude that u and v are not conjugate. We shall proceed by normalizing u and v^*. The process of working from right to left yields normal forms

$$u = g_0 t^{\varepsilon_1} \ldots g_{m-1} t^{\varepsilon_m} \quad \text{and} \quad v^* = h_0 t^{\varepsilon_1} \ldots h_{m-1} t^{\varepsilon_m}.$$

These can be done using Lemma 3.1. It remains to consider whether

$$u = z^{-1} v^* z \tag{1}$$

for any z in the appropriate subgroup A or B. For definiteness, assume that $\varepsilon_m = -1$ and that $z \in A$. Thus, Equation (1) becomes

$$g_0 t^{\varepsilon_1} \ldots g_{m-1} t^{\varepsilon_m} = z^{-1} h_0 t^{\varepsilon_1} \ldots t^{\varepsilon_{m-1}} h_{m-1} \phi(z) t^{\varepsilon_m} \tag{2}$$

Step 1. $\varepsilon_{m-1} = 1$. We have $t^{\varepsilon_{m-1}} h_{m-1} \phi(z) = z t^{\varepsilon_{m-1}} h_{m-1}$ and hence Equation (2) becomes

$$g_0 t^{\varepsilon_1} \ldots t^{\varepsilon_{m-1}} g_{m-1} t^{\varepsilon_m} = z^{-1} h_0 t^{\varepsilon_1} \ldots t^{\varepsilon_{m-2}} h_{m-2} z t^{\varepsilon_{m-1}} h_{m-1} t^{\varepsilon_m} \tag{3}$$

To proceed we must have $h_{m-1} = g_{m-1}$. If this condition fails we stop and conclude that $u \neq z^{-1} v^* z$ for any $z \in A$, and this particular cyclic permutation v^* of v. If $\varepsilon_{m-2} = -1$, by moving z to the left of $t^{\varepsilon_{m-2}}$ we obtain

$$g_0 t^{\varepsilon_1} \ldots g_{m-2} t^{\varepsilon_{m-1}} g_{m-1} t^{\varepsilon_m} = z^{-1} h_0 t^{\varepsilon_1} \ldots h_{m-3} \phi(z) t^{\varepsilon_{m-2}} h_{m-2} t^{\varepsilon_{m-1}} g_{m-1} t^{\varepsilon_m}. \quad (4)$$

We must have $g_{m-2} = h_{m-2}$, otherwise we write $u \neq z^{-1} v^* z$ and stop. Note that Equation (4) is of the type of Equation (2). If $\varepsilon_{m-2} = 1$, proceed to step 2.

Step 2. $\varepsilon_{m-1} = -1$. Let $h_{m-1}\phi(z) = \overline{h_{m-1}\phi(z)} \cdot a_{m-1}$ where $\overline{h_{m-1}\phi(z)}$ is a representative of a coset of A in G and a_{m-1} is an element of A. Since the left hand side of Equation (2) is in normal form, we must have

$$h_{m-1}\phi(z) = g_{m-1} \cdot a_{m-1} \qquad \text{or}$$

$$g_{m-1}^{-1} h_{m-1} \phi(z) = a_{m-1} \in A \quad (5)$$

According to Lemma 3.2 elements z and a_{m-1} can be found, and also if we think of them as n-tupls, then some of the coordinates will be arbitrary. Replacing $h_{m-1}\phi(z)$ by $g_{m-1} a_{m-1}$ in Equation (2) and moving a_{m-1} to the left of $t^{\varepsilon_{m-2}}$, we obtain

$$g_0 t^{\varepsilon_1} \ldots g_{m-2} t^{\varepsilon_{m-1}} g_{m-1} t^{\varepsilon_m} = z^{-1} h_0 t^{\varepsilon_1} \ldots t^{\varepsilon_{m-2}} h_{m-2} \theta(a_{m-1}) t^{\varepsilon_{m-1}} g_{m-1} t^{\varepsilon_m}. \quad (6)$$

Note that again the above equation is of the type of Equation (2). Now if $\varepsilon_{m-2} = -1$, then proceed to step 1 and if $\varepsilon_{m-2} = -1$, then proceed to step 2 and apply Lemma 1 to find elements a_{m-2} in A such that $h_{m-2}\phi(a_{m-1}) = g_{m-2} \cdot a_{m-2}$.

Since the length of v^* is finite, a finite number of steps of this process produces the following equation:

$$g_0 t^{\varepsilon_1} g_1 \ldots g_{n-1} t^{\varepsilon_m} = z^{-1} h_0 \Phi(z) t^{\varepsilon_1} g_1 \ldots g_{n-1} t^{\varepsilon_m} \quad (7)$$

where $\Phi(z)$ is an element of G, which is obtained by moving z from its position in Equation (1) to its position in Equation (7). Since ε_is, h_is and g_is are known, the operation Φ is known precisely. By reversing steps to obtain Φ, then Φ^{-1} is known. Suppose $\Phi(z) = \sum_{i=1}^{n} r_i x_i$ where some of integers r_i are possibly arbitrary, then

$$z = \Phi^{-1}\left(\sum_{i=1}^{n} r_i x_i\right) = \sum_{i=1}^{n} p_i x_i$$

for some p_i. From Equation (7) we have

$$g_0 = z^{-1} h_0 \Phi(z). \tag{8}$$

Using additive notation, we have

$$z = h_0 - g_0 + \Phi(z),$$

$$\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} (q_i + r_i) x_i$$

where $h_0 - g_0 = \sum_{i=1}^{n} q_i x_i$ for some integers $q_1, \ldots, q_n$. To decide whether $u = z^{-1} v^* z$ for this particular cyclic permutation v^* of v, we must be able to find a solution to the system of equations:

$$\begin{aligned} p_1 &= q_1 + r_1 \\ p_2 &= q_2 + r_2 \\ &\vdots \\ p_n &= q_n + r_n. \end{aligned}$$

If the above system of linear equations is inconsistent, we will check whether u is conjugate to any other cyclic permutation of v by an element z in appropriate subgroups A or B. Since v has only finite number of cyclic permutations with the same t-symbols as in u, we can effectively decide whether u and v are conjugate in G^*.

An example may help to clarify the above. Let G be a finitely generated free abelian group, generated by $\{x_1,x_2,x_3,x_4\}$, and let $A = \langle x_1^2, x_2^3, x_3\rangle$ and $B = \langle x_2^4, x_3^3, x_4\rangle$ be subgroups of G with isomorphism $\phi\colon A \longrightarrow B$ defined by $\phi(x_1^2) = x_2^4$, $\phi(x_2^3) = x_3^3$ and $\phi(x_3) = x_4$. Let $u = x_1^{-3}x_4^9t^{-1}x_1t^{-1}x_2t^{-1}$ and $v = x_1t^{-1}x_1t^{-1}x_2^2t^{-1}$ be elements of

$$G^* = \langle G, t/t^{-1}At = B, \phi\rangle,$$

the HNN extension of G. We will determine whether u and v are conjugate in G^*. Note that u and v both are in normal form.

Since u and v have the same t-symbols, we first set

$$x_1^{-3}x_4^9t^{-1}x_1t^{-1}x_2t^{-1} = z^{-1}x_1t^{-1}x_1t^{-1}x_2^2t^{-1}z \tag{1}$$

for some z in A. By moving z across the t^{-1} we obtain

$$x_1^{-3}x_4^9t^{-1}x_1t^{-1}x_2t^{-1} = z^{-1}x_1t^{-1}x_1t^{-1}x_2^2\phi(z)t^{-1}. \tag{2}$$

Now let $x_2 \cdot a_1 = x_2^2\phi(z)$ for some $a_1 \in A$. If we use additive notation, we have

$$a_1 = x_2 + \phi(z), \tag{3}$$

Assume

$$z = c_1(2x_1) + c_2(3x_2) + c_3(x_3), \quad \text{and}$$

$$a_1 = r_1(2x_1) + r_2(3x_2) + r_3(x_3).$$

Substituting the above equations in Equation (3) we have

$$c_1\phi(2x_1) + c_2\phi(3x_2) + c_3\phi(x_3) - 2r_1x_1 - 3r_2x_2 - r_3x_3 = -x_2,$$

$$(c_1,c_2,c_3)\begin{bmatrix}0 & 4 & 0 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} + (r_1,r_2,r_3)\begin{bmatrix}-2 & 0 & 0 & 0\\ 0 & -3 & 0 & 0\\ 0 & 0 & -1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} =$$

$$(0,-1,0,0)\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}$$

$$(c_1,c_2,c_3,r_1,r_2,r_3)\begin{bmatrix}0 & 4 & 0 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 1\\ -2 & 0 & 0 & 0\\ 0 & -3 & 0 & 0\\ 0 & 0 & -1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} = (0,-1,0,0)\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}.$$

In the more compact form,

$$UM = V \tag{4}$$

where $U = (c_1,c_2,c_3,r_1,r_2,r_3)$, M is the above 6×4 matrix and V = (0,-1, 0,0). Performing suitable elementary row and column operations yields:

$$PMQ = D \tag{5}$$

where

$$P = \begin{bmatrix}0 & 0 & 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 3 & 0 & 0 & 0 & 4 & 0\\ 0 & 1 & 0 & 0 & 0 & 3\end{bmatrix}$$

$$Q = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and}$$

$$D = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now Equation (4) becomes

$$UP^{-1}DQ^{-1} = V \qquad (6)$$

Let $UP^{-1} = W$, where W is a 1×6 matrix $(w_1,w_2,w_3,w_4,w_5,w_6)$. Then Equation (6) becomes

$$WD = VQ \qquad (7)$$

Equivalently

$$(w_1,w_2,w_3,w_4,w_5,w_6) \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = (0,-1,0,0) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad (8)$$

From Equation (8) we obtain the following solutions

$w_1 = 0$

$w_2 = -1$

$w_3 = 0$

$w_4 = 0$

$w_5 = p$

$w_6 = q$

where p and q are arbitrary. Substituting these solutions in the equation $UP^{-1} = W$ and then solving for U, we have

$$(c_1,c_2,c_3,r_1,r_2,r_3) = (0,-1,0,0,p,q)\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 & 0 & 3 \end{bmatrix} \qquad (9)$$

Solving the above equation leads us to the following solutions:

$c_1 = -1 + 3p$

$c_2 = q$

$c_3 = 0$

$r_1 = 0$

$r_2 = -1 + 4p$

$r_3 = 3q$

Thus

$$z = (-1+3p)(2x_1) + q(3x_2) \quad \text{and}$$

$$a_1 = (-1 + 4p)(3x_2) + 3qx_3 \qquad (10)$$

with p and q arbitrary integers. Now we proceed normalizing right hand side of Equation (2). By replacing $\phi(z) = x_2^{-1}a_1$ in Equation (2) and moving a_1 to the left of t^{-1}, Equation (2) becomes:

$$x_1^{-3}x_4^{9}t^{-1}x_1t^{-1}x_2t^{-1} = z^{-1}x_1t^{-1}x_1\phi(a_1)t^{-1}x_2t^{-1} \tag{11}$$

Let $x_1 \cdot a_2 = x_1\phi(a_1)$ for some $a_2 \in A$. Using additive notation and applying Equation (10), we have

$$\begin{aligned} a_2 = \phi(a_1) &= \phi((-1+4p)(3x_2) + 3qx_3) \\ &= \phi((-1+4p)(3x_2)) + \phi(3qx_3) \\ &= (-1+4p)(3x_3) + 3qx_4 \in A \end{aligned}$$

so $q = 0$ and $a_2 = (4p-1)(3x_3)$. Replacing a_2 by $\phi(a_1)$ in Equation (11) and moving a_2 across t^{-1}, we have

$$x_1^{-3}x_4^{9}t^{-1}x_1t^{-1}x_2t^{-1} = z^{-1}x_1\phi(a_2)t^{-1}x_1t^{-1}x_2t^{-1} \tag{12}$$

Let $\Phi(z) = \phi(a_2) = \phi((4p-1)(3x_3)) = (4p-1)(3x_4)$, then

$$\Phi(z) = \phi(\phi(x_2+\phi(z)) = 3(4p-1)x_4,$$

$$z = \phi^{-1}(-x_2+\phi^{-1}(\phi^{-1}(3(4p-1)x_4)),$$

$$\begin{aligned} z &= \phi^{-1}(-x_2+\phi^{-1}(3(4p-1)x_3) \\ &= \phi^{-1}(-x_2+3(4p-1)x_2) \\ &= \phi^{-1}(4(3p-1)x_2) \\ &= 2(3p-1)x_1. \end{aligned}$$

From Equation (12) we have

$$x_1^{-3}x_4^{9} = z^{-1}x_1\phi(a_2)$$

or additively

$$-3x_1 + 9x_4 = -2(3p-1)x_1 + x_1 + 3(4p-1)x_4.$$

Equating the appropriate terms, we obtain the following system of two linear equations:

$$3x_1 = 3(2p-1)x_1$$

$$9x_4 = 3(4p-1)x_4.$$

The system is consistent and has only one solution $p = 1$. Therefore, $z = x_1^4$ and $u = x_1^{-4}vx_1^4$.

CHAPTER IV

LEHMER-SCHUR ALGORITHM

In this chapter we will discuss the results necessary to prove the conjugacy problem for short words (t-free) in an HNN extension of a finitely generated free abelian group. Our aim is to find an algorithm to determine whether a given complex polynomial $\lambda(z)$ with rational coefficients divides the polynomial $z^N P(z) - Q(z)$ for any N, where P(z) and Q(z) are given complex polynomials over the rationals. To search for this N, the Lehmer-Schur algorithm is essential. It will be used to approximate the roots of complex polynomials with real coefficients.

Theorem 4.1. (Lehmer-Schur). *There is an algorithm that will construct for any polynomial* P(z) *with rational coefficients and for any* $\delta > 0$, *a rational* $\varepsilon > 0$ *with* $\varepsilon < \delta$ *and a collection* B_i *of balls of radius* ε *in the complex plane with the following properties*:

i) *Each zero of* P(z) *lies in some* B_i.

ii) *Each* B_i *contains at least one zero of* P(z).

Proof: The proof is given in [15].

We shall refer to the balls B_i given by Theorem 4.1 as the ε-*Schur circles* for P(z).

Lemma 4.1. *Let* P(z) *be a polynomial with coefficients from* Q[i].

There is an algorithm that will determine if all zeros of P(z) *are real.*

Proof: First of all we assume this is a known algorithm for polynomial of degree smaller than P. With no loss of generality we may assume that the leading coefficient of P is 1. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the zeros of P(z) then $P(z) = \prod_{i=1}^{n} (z-\alpha_i)$. Clearly then, if P(z) has any complex coefficient then it has at least one complex zero. Proceed assuming all coefficients are rational. For $\varepsilon = 1, \frac{1}{2}, \frac{1}{4}, \ldots,$ construct ε-Schur circles for P(z). If P(z) has a complex zero eventually an ε-Schur circle will be constructed that misses the real axis. At the same time this search for nonreal zeros is being carried out, we simultaneously test to try and locate real zeros of P(z). Let Q(z) = G.C.D(P(z),Q(z)) and $S(z) = \frac{P(z)}{Q(z)}$. Then S has no multiple zeros (all zeros are of multiplicity 1). If α is a real number and $S(\alpha) = 0$, then S crosses the x-axis at $x = \alpha$. Suppose that the deg S = k, then S has all real zeros if and only if S has k sign changes on $-\infty < x < \infty$. For each N look at sign changes in

$$X_N = \{P(\frac{N-2^N}{2^N}),\ P(\frac{(N+1)-2^{N+1}}{2^{N+1}}), \ldots\} \quad .$$

If S has all real zeros eventually we see k sign changes in X_N. If this occurs than P has all real zeros if and only if Q does. But deg Q < deg P, and so it can be determined if all zeros of Q(Z) are real.

Lemma 4.2. *Let* P(z) *be a polynomial with coefficients in* Q[i]. *There is an algorithm that will determine if all zeros of* P(z) *lie on* $|z| = 1$.

Proof: Choose $\alpha \in Q[i]$ such that $|\alpha| = 1$, $\alpha \neq 1$, $\alpha \neq i$ and $P(\alpha) \neq 0$. We will find a mobius transformation which maps the lower half of the complex plane onto the unit circle in such a way that $z = 0$ is mapped into i and $z = 1$ is mapped into 1 while the point at infinity is mapped into α. From $\sigma(z) = \frac{az + b}{z + c}$ we have $i = \frac{b}{c}$, $1 = \frac{a + b}{1 + c}$ and $\alpha = a$. So that $a = \alpha$, $b = \frac{1 - \alpha}{i + 1}$ and $c = \frac{1 - \alpha}{i - 1}$. Hence the required transformation is

$$\sigma(z) = \frac{\alpha z + \frac{1-\alpha}{i+1}}{z + \frac{1-\alpha}{i-1}} .$$

Suppose that the degree of $P(z)$ is n, and let $S(z) = (z + \frac{1-\alpha}{i-1})^n P(\sigma(z))$. Then $S(z)$ is a polynomial with coefficients in $Q[i]$. We claim that all zeros of $S(z)$ are real if and only if all zeros of $P(z)$ are on $|z| = 1$.

Suppose that all zeros of $S(z)$ are real. Let β be a zero of P. Since $P(\sigma(\sigma^{-1}(\beta)) = P(\beta) = 0$, $\sigma^{-1}(\beta)$ is a zero of S. Thus $\sigma^{-1}(\beta)$ is a real number. Since σ maps the real line onto the $|z| = 1$, $\sigma(\sigma^{-1}(\beta)) = \beta$ is on the unit circle.

Now suppose that all zeros of P are on $|z| = 1$. Let γ be a zero of S, then $P(\sigma(\gamma)) = 0$. Hence $\sigma(\gamma)$ is a zero of P and $\sigma(\gamma)$ is on the unit circle. Therefore γ is a real number.

Lemma 4.3. *Let* $P(z)$ *be a polynomial over the rationals with no zero on* $|z| = 1$. *There is an algorithm that will construct a positive* ε *such that* $|P(z)| > \varepsilon$ *for all* z *on the unit circle*.

Proof: First of all use the Lehmer-Schur algorithm to put the zeros of P inside δ-balls that miss the unit circle. Let $P(z) = a\prod_i(z-\alpha_i)$ where $\alpha_i \in B_\delta(c_i)$. We have

$$|\alpha_i| > |c_i| - \delta,$$

$$|z-\alpha_i| > |z-c_i| - \delta \geq \left|\frac{c_i}{|c_i|} - c_i\right| - \delta$$

for all z with $|z| = 1$ (see Figure 1). Thus $|P(z)| = |a\prod_i(z-\alpha_i)| = |a|\prod_i|z-\alpha_i| > |a|\prod_i(|\frac{c_i}{|c_i|} - c_i| - \delta) = \varepsilon$ for all z on $|z| = 1$.

Corollary 4.1. *Let* $P(z)$ *be a polynomial with coefficients in* Q *and suppose that* $P(z) \neq 0$ *for all* z *on the circle* $|z-z_0| = R$. *Then a positive lower bound for* $|P(z)|$ *on* $|z-z_0| = R$ *can be effectively computed*.

Proof: Note that any point α on the circle $|z-z_0| = R$ is of the form $\alpha = z_0 + \beta R$ for some β with $|\beta| = 1$. Suppose that P has no zero on the circle $|z-z_0| = R$ and let $Q(z) = P(z_0+Rz)$. We claim that Q has no zero on the unit circle. Suppose that $Q(\beta) = 0 = P(z_0+R\beta)$ with $|\beta| = 1$, then $z_0 + R\beta$ is on the circle $|z-z_0| = R$. This contradicts $P(z) \neq 0$ on $|z-z_0| = R$. According to Lemma 4.3 it is possible to compute an $\varepsilon > 0$ such that $|P(z_0+Rz)| = |Q(z)| > \varepsilon$ for all z on $|z| = 1$. Therefore $|P(z)| > \varepsilon$ for all z on $|z-z_0| = R$.

Definition 4.1. A *circular domain* is a compact, connected subset of complex plane whose boundary components consist of circles.

Corollary 4.2. *Let* $P(z)$ *be a polynomial with rational coefficients. If* $P(z)$ *does not vanish on a circular domain* Ω, *then a positive lower bound for* $|P(z)|$ *on* Ω *can be effectively computed*.

Proof: Since P has no zero on Ω, $\frac{1}{P}$ is analytic on Ω. By the Maximum Modulus Theorem $\frac{1}{P}$ attains its maximum on the boundary of Ω, that is, P reaches its minimum on $\partial\Omega$. Corollary 4.1 can be used to compute lower bounds ε_i for $|P|$ on the boundary components of Ω. Clearly $\varepsilon = \min\{\varepsilon_i\}$ is the required lower bound for $|P|$ on Ω.

Definition 4.2. *Let* λ, P *and* Q *be complex polynomials with rational*

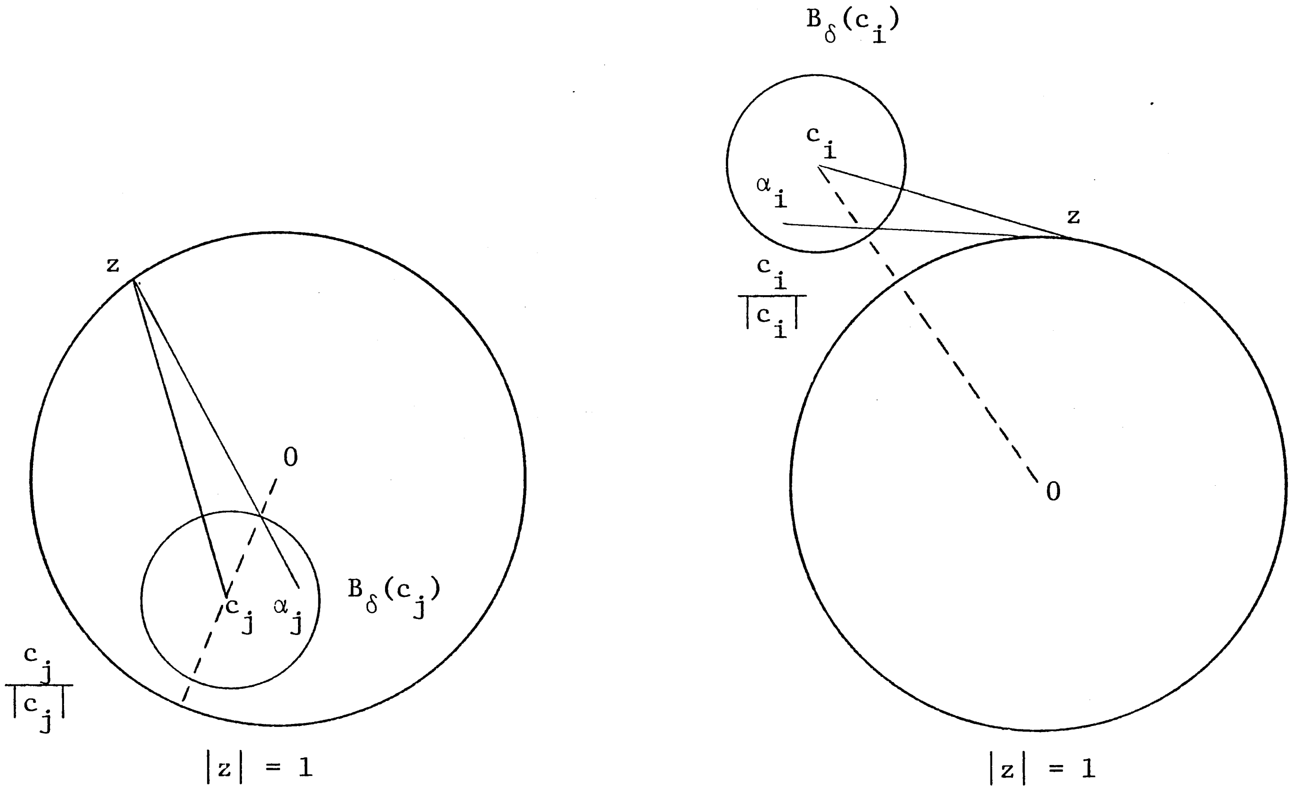

Figure 1. The Choices of α_i's

coefficients. Suppose that λ, P *and* Q *are relatively prime in pairs. We say* P *and* Q *are* λ-K *incompatible if* $\lambda(z)$ *does not divide* $z^N P(z) - Q(z)$ *for any* $N \geq K$.

Remark: Let $\lambda(t)$, $P(t)$ and $Q(t)$ be given polynomials with deg $P(t)$ and deg $Q(t)$ less than deg $\lambda(t)$. To decide whether $\lambda(t)$ divides $t^N P(t) - Q(t)$ for any N, we may assume that λ, P and Q are relatively prime in pairs. First if G C D $(\lambda,P,Q) = R$, with deg $R > 0$, then $\lambda(t)|t^N P(t)-Q(t)$ if and only if $\lambda_1|t^N P_1 - Q_1$ where $\lambda_1 R = \lambda$, $P_1 R = P$ and $Q_1 R = Q$ and G C D $(\lambda_1,P_1,Q_1) = 1$. Let us assume G C D $(\lambda,P,Q) = 1$. Now if G C D $(P,Q) = R$ with deg $R > 0$, then $\lambda|t^N P-Q$ if and only if $\lambda|R(t^N P_1-Q_1)$ where $P = RP_1$ and $Q = RQ_1$. But $(\lambda,R) = 1$, since G C D $(\lambda,P,Q) = 1$. Thus $\lambda|t^N P-Q$ if and only if $\lambda|\ t^N P_1-Q_1$ and $(P_1,Q_1) = 1$. So we may assume $(P,Q) = 1$.

Suppose G C D $(\lambda,P) = R$ with deg $R > 0$. If $\lambda|t^N P - Q$ then R must divide Q since R divides λ and P, contradicting G C D $(\lambda,P,Q) = 1$.

Finally suppose that G C D $(\lambda(t),Q(t)) = t^K\ R(t)$ with deg $R(t) > 0$ and $(R(t),t) = 1$. Then $\lambda(t)|t^N P(t) - Q(t)$ if and only if $\lambda_1(t)R(t)|\ t^{N-K}P(t) - Q_1(t)R(t)$ where $\lambda(t) = t^K\lambda_1(t)R(t)$ and $Q(t) = t^K Q_1(t)R(t)$. Thus $R(t)$ must divide $P(t)$. Contradicting $(\lambda(t),P(t),Q(t)) = 1$.

Theorem 4.1. *Suppose that* λ, P *and* Q *are rational polynomials that are relatively prime in pairs. Suppose* λ *has a zero not on* $|z| = 1$. *Then an integer* K *can be effectively computed so that* P *and* Q *are* λ-K *incompatible*.

Proof: Make ε-Schur circles for λ, P and Q for $\varepsilon = 1,\frac{1}{2},\frac{1}{4},\ldots.$ Since λ, P and Q are relatively prime in pairs we eventually find balls

$\{B_\varepsilon(C_i)\}_{i=1}^{n}$, $\{B_\varepsilon(D_j)\}_{j=1}^{\rho}$ and $\{B_\varepsilon(E_k)\}_{k=1}^{r}$ containing zeros of λ, P and Q respectively, in which their intersection is empty, and at least one $B_\varepsilon(C_i)$ misses $|z| = 1$. If necessary reduce ε such that the Schur circles of zeros of λ, P and Q off the unit circle do not intersect $|z| = 1$. Choose R large enough such that all the ε-Schur circles of λ, P and Q lie inside the annulus $0 \leq |z| \leq R$.

Case 1. Suppose $\lambda(z_\lambda) = 0$ and $|z_\lambda| > 1$. Let $\Omega = \{z|1 \leq |z| \leq R\} - \{z||z-D_j| < \varepsilon,\ j = 1,\ldots,\rho\}$. If P has a zero on the unit circle let

$$\delta = \frac{1}{2}\{[\min_{\substack{|C_i|>1 \\ |D_j|>1}}\{|C_i|,|D_j|\}] - \varepsilon - 1\} .$$

Then 2δ is the distance between the unit circle and the closest Schur circles of λ and P outside the unit circle. Thus the annulus $1 \leq |z| \leq 1+\delta$ does not intersect any Schur circle of λ and P outside $|z| = 1$. Now let $\Omega_1 = \Omega \setminus \{z|1 \leq |z| < 1+\delta\}$, then Ω_1 is a circular domain whose boundary components are the circles $|z| = R$, $|z| = 1+\delta$ and Schur circles of P outside $|z| = 1$. The shaded area below indicates the circular domain Ω_1 (see Figure 2). Since P has no zero on Ω_1, a positive lower bound m for $|P|$ on Ω_1 can be computed.

Suppose that $Q(z) = a_\rho z^\rho + a_{\rho-1}z^{\rho-1} + \ldots + a_0$. Then

$$|Q(z)| = |a_\rho z^\rho + \ldots + a_0| \leq |a_\rho z^\rho| + \ldots + |a_0|$$

$$= |a_\rho||z|^\rho + \ldots + |a_0|$$

$$\leq |a_\rho|R^\rho + \ldots + |a_0| = M$$

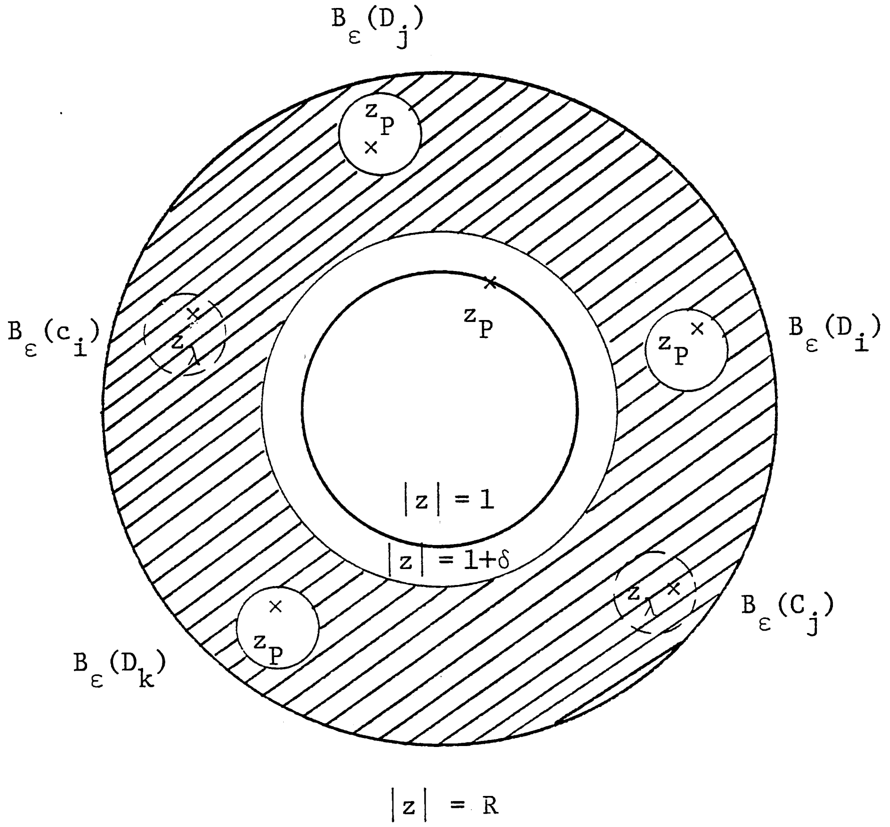

Figure 2. The Circular Domain Ω_1, Where λ Has a Zero Outside $|z| = 1$

We choose K such that $\frac{M}{(1+\delta)^K} < m$. Claim: P and Q are λ-K incompatable. Obviously λ has a zero on Ω_1. If $N \geq K$ then $Z^N P(z) - Q(z)$ does not vanish on Ω_1. For suppose $\alpha \in \Omega_1$ and $\alpha^N P(\alpha) - Q(\alpha) = 0$, then

$$|P(\alpha)| = \frac{|Q(\alpha)|}{|\alpha^N|} \leq \frac{M}{(1+\delta)^N} \leq \frac{M}{(1+\delta)^K} < m.$$

This is not possible since $|P(z)| > m$ on Ω.

Case 2. Suppose $\lambda(z_\lambda) = 0$ and $0 < |z_\beta| < 1$. Let $\Omega = \{z|0 \leq |z| \leq 1\} - \{z||z-E_k| < \varepsilon, k = 1,2,\ldots,r\}$, and $\delta = \frac{1}{2}\{1 - [\min_{\substack{|C_i|<1 \\ |E_k|<1}}\{|C_i|, |E_k|\} - \varepsilon]\}$.

Then 2δ is the distance between the unit circle and the closest Schur circles of λ and Q inside the unit circle. Thus the annulus $1-\delta \leq |z| \leq 1$ does not intersect any Schur circles of λ and Q inside $|z| = 1$. Now let $\Omega_1 = \Omega \setminus \{z|1-\delta < |z| \leq 1\}$, then Ω_1 is a circular domain containing all the zeros of λ inside $|z| = 1$, and missing the Schur circles of Q inside $|z| = 1$. The boundary components of Ω_1 are the circles $|z| = 1 - \delta$ and the Schur circles of Q inside $|z| = 1$. The shaded area below indicates the circular domain Ω_1 (see Figure 3). Since Q does not vanish on Ω_1, a positive lower bound m for $|Q|$ on Ω_1 can be computed. Let M be an upper bound for $|P|$ on Ω_1, and choose K such that $(1-\delta)^K M < m$. We claim that P and Q are λ-K incompatible. Since λ has a zero on Ω_1, it suffices to show that $Z^N P(z) - Q(z)$ does not vanish on Ω_1 for $N \geq K$. Suppose that $\alpha \in \Omega_1$ and $\alpha^N P(\alpha) = Q(\alpha)$, then

$$|Q(\alpha)| = |\alpha^N P(\alpha)| = |\alpha|^N \; |P(\alpha)| \leq (1-\delta)^N M \leq (1-\delta)^K M < m.$$

But $|Q| > m$ on Ω.

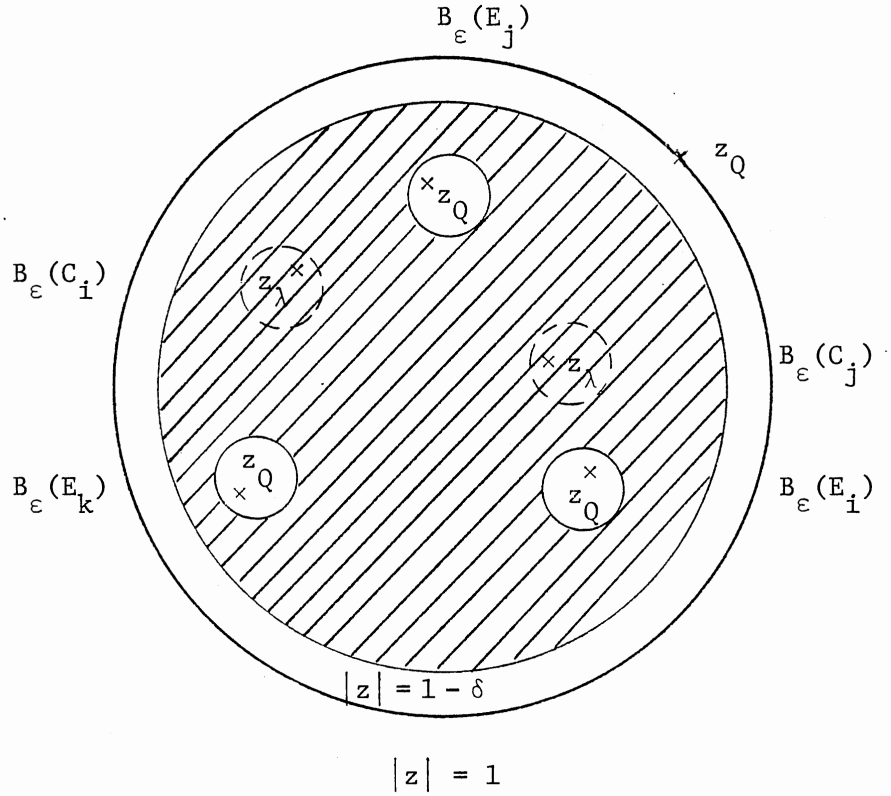

Figure 3. The Circular Domain Ω_1, Where λ Has a Zero Inside $|z| = 1$

Theorem 4.2. *Let* λ, P *and* Q *be complex polynomials with rational coefficients. Suppose that* λ, P *and* Q *are relatively prime in pairs. It is possible to construct a polynomial* $\overline{Q}(z)$ *such that if* $\lambda(z)$ *divides* $z^N P(z) - Q(z)$ *for any* N, *then* $\lambda(z)$ *divides* $z^N - \overline{Q}(z)$. ($\overline{Q}$ is independent of N).

Proof: Since $(\lambda,P) = 1$, there exist polynomials R and S in $Q[z]$ such

$$\lambda R + PS = 1. \tag{1}$$

Suppose that $\lambda(z)$ divides $z^N P(z) - Q(z)$ for some N, then $\lambda(z)$ obviously divides

$$(z^N P(z) - Q(z))S(z) = z^N P(z)S(z) - Q(z)S(z). \tag{2}$$

From Equation (1) $PS = 1 - \lambda R$. By replacing $1 - \lambda R$ for PS in Equation (2) it follows that $\lambda(z)$ divides $z^N - z^N\lambda(z)R(z) - Q(z)S(z)$. Thus λ must divide $z^N - \overline{Q}(z)$ where $\overline{Q}(z) = Q(z)S(z)$.

Theorem 4.3. *Let* λ, P *and* Q *be as in Theorem* 4.2. *Suppose that all zeros of* λ *are on the unit circle and* λ *has a repeated zero. Then an integer* K *can be effectively computed so that* P *and* Q *are* λ-K *incompatible*.

Proof: Suppose that λ has a multiple zero on the unit circle and that $\lambda(z)$ divides $z^N P(z) - Q(z)$ for some N. By Theorem 4.2 $\lambda(z)$ divides $z^N - S(z)$ for some $S(z)$. Since λ has a multiple zero, $Nz^{N-1} - S'(z) = 0$ for some z with $|z| = 1$. Thus we have

$$1 = |z|^{N-1} = \frac{|S'(z)|}{N} .$$

Now choose K such that $K > \max_{|z|=1} |S'(z)|$. Then for $N \geq K$, $z^N - X(z)$ has no multiple zero on the $|z| = 1$. Thus $\lambda(z)$ does not divide $z^N P(z) - Q(z)$ for $N \geq K$.

CHAPTER V

DECOMPOSITION OF SINGLE LINEAR TRANSFORMATION

In this chapter we will investigate the structure of a finite dimensional vector space A over a field K relative to a linear transformation $A \longrightarrow A$. The linear transformation induces a decomposition of A as a direct sum of certain K[t]-modules.

Let K be a field and $\phi\colon A \longrightarrow A$ a linear transformation of an n-dimensional K-vector space A. Then ϕ induces a K[t]-module structure on A as follows. If $f = \Sigma k_i t^i$ is a polynomial in K[t] and $u \in A$, then $f(\phi) = \Sigma k_i \phi^i \in \mathrm{Hom}_K(A,A)$ (where $\phi^0 = 1_A$ as usual) and $fu = \Sigma k_i \phi^i(u) = f(\phi)(u)$.

<u>Theorem</u> 5.1. <u>Let</u> $\phi\colon A \longrightarrow A$ <u>be</u> <u>a</u> <u>linear</u> <u>transformation</u> <u>of</u> <u>an</u> n-<u>di</u><u>mensional</u> <u>vector</u> <u>space</u> A <u>over</u> <u>a</u> <u>field</u> K. <u>Then</u> $A = \frac{K[t]}{\langle\lambda_1\rangle} \oplus \frac{K[t]}{\langle\lambda_2\rangle} \oplus \cdots \oplus \frac{K[t]}{\langle\lambda_n\rangle}$ <u>where</u> $\lambda_1, \lambda_2, \ldots, \lambda_n \in K[t]$ <u>such</u> <u>that</u> <u>each</u> λ_i <u>is</u> <u>monic</u> <u>and</u> $\lambda_1 | \lambda_2 | \ldots | \lambda_n$.

Proof: Let $\{x_1, \ldots, x_n\}$ be a basis of A and $M = (a_{ij})$ be the matrix of ϕ relative to this basis. Let F be a free K[t]-module with the standard basis $\{\varepsilon_i\}_{i=1}^n$. Let $\pi\colon F \longrightarrow A$ be the unique K[t]-module homomorphism such that $\pi(\varepsilon_i) = x_i$ for $i = 1,2,\ldots,n$, and let $\psi\colon F \longrightarrow F$ be the unique K[t]-module homomorphism such that $\psi(u_i) = t\varepsilon_i - \sum_{j=1}^{n} a_{ij}\varepsilon_j$. Then the matrix of ψ relative to the basis $\{\varepsilon_i\}$ is $tI_n - M$.

We claim that the sequence of K[t]-module

$$F \xrightarrow{\psi} F \xrightarrow{\pi} A \longrightarrow 0$$

is exact. Clearly π is a K[t]-module epimorphism. Since M is the matrix of ϕ and the K[t]-module structure of A is induced by ϕ,

$$\pi(t\varepsilon_i) = t\pi(\varepsilon_i) = tx_i = \phi(x_i) = \sum_{j=1}^{n} a_{ij}x_j.$$

Consequently, for each i

$$\pi\psi(\varepsilon_i) = \pi(t\varepsilon_i - \sum_{j=1}^{n} a_{ij}\varepsilon_j) = \pi(t\varepsilon_i) - \sum_{j=1}^{n} a_{ij}\pi(\varepsilon_j)$$
$$= \sum_{j=1}^{n} a_{ij}x_j - \sum_{j=1}^{n} a_{ij}x_j = 0,$$

whence Im $\psi \subset$ ker π. To show that ker $\pi \subset$ Im ψ, it suffices to prove that every element w of F is of the form $w = \psi(v) + \sum_{j=1}^{n} k_j\varepsilon_j$ $(v \in F, k_j \in K)$. For in this case if $w \in \ker \pi$, then

$$0 = \pi(w) = \pi(\psi(v)) + \pi\Sigma k_j\varepsilon_i = 0 + \Sigma k_jx_j.$$

Since $\{x_j\}$ is a basis of A, $k_j = 0$ for all j, consequently, $w = \psi(v)$ and hence ker $\pi \subset$ Im ψ. Since every element of F is a sum of terms of the form $f\varepsilon_i$ with $f \in K[t]$, we need only show that for each i and r, there exist v_{ir} and $k_j \in K$ such that $t^r\varepsilon_i = \psi(v_{ir}) + \sum_{j=1}^{n} k_j\varepsilon_j$. For each i and r = 1, we have $t\varepsilon_i = t\varepsilon_i - \sum_{j=1}^{n} a_{ij}\varepsilon_j + \sum_{j=1}^{n} a_{ij}\varepsilon_j = \psi(\varepsilon_i) + \sum_{j=1}^{n} a_{ij}\varepsilon_j$. Proceeding inductively assume that for each j there exist $v_{j,r-1} \in F$ and $k_{j,\rho} \in K$ such that $t^{r-1}\varepsilon_j = \psi(v_{j,r-1}) + \sum_{\rho=1}^{n} k_{j,\rho}\varepsilon_\rho$. Then for each i

$$t^r\varepsilon_i = t^{r-1}(t\varepsilon_i) = t^{r-1}(\psi(\varepsilon_i) + \sum_{j=1}^{n} a_{ij}\varepsilon_j)$$
$$= \psi(t^{r-1}\varepsilon_i) + \sum_j a_{ij}t^{r-1}\varepsilon_j$$

$$= \psi(t^{r-1}\varepsilon_i) + \sum_j a_{ij}(\psi(v_{j,r-1}) + \sum_\rho k_{j\rho}\varepsilon_\rho)$$

$$= \psi(t^{r-1}\varepsilon_i) + \sum_j a_{ij}\psi(v_{j,r-1}) + \sum_j \sum_\rho a_{ij}k_{j\rho}\varepsilon_\rho$$

$$= \psi(t^{r-1}\varepsilon_i + \sum a_{ij}v_{j,r-1}) + \sum_\rho (\sum_j a_{ij}k_{j\rho})\varepsilon_\rho \ .$$

Thus $t^r\varepsilon_i = \psi(v_{i,r}) + \sum_{\rho=1}^{n} c_\rho \varepsilon_\rho$ with $v_{i,r} = t^{r-1}\varepsilon_i + \sum_j a_{ij}v_{j,r-1} \in F$ and $c_\rho = \sum_j a_{ij}k_{j\rho} \in K$ and the induction is complete. Therefore $F \xrightarrow{\psi} F \xrightarrow{\pi} A \longrightarrow 0$ is exact and hence $A \cong \frac{F}{\ker \pi} = \frac{F}{\operatorname{Im} \psi}$.

Since K[t] is a principal ideal domain, Theorem 2.6 shows that $tI_n - M$, the matrix of ψ relative to the basis $\{\varepsilon_i\}$, is equivalent to a diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ & 0 & \lambda_n \end{bmatrix}$ with $\lambda_1, \ldots, \lambda_n \in K[t]$ such that $\lambda_1 | \lambda_2 \ldots | \lambda_n$. We may assume each λ_i is monic (if necessary, perform suitable elementary row operations on D). Clearly the determinant $|tI_n - M|$ in K[t] is monic polynomial of degree n. In particular, det. $|tI_n - M| \neq 0$, whence det. $|D| \neq 0$. Consequently, all the diagonal entries of D are nonzero. Since D is equivalent to $tI_n - M$, D is the matrix of ψ relative to some pair of ordered bases $\{\gamma_1, \ldots, \gamma_n\}$ and $\{\delta_1, \delta_2, \ldots, \delta_n\}$ of F (exactly as in the proof of Lemma 2.3). This means that $\psi(\gamma_i) = \lambda_i\delta_i$ for each i and $\{\lambda_1\delta_1, \ldots, \lambda_n\delta_n\}$ spans Im ψ. Thus $\operatorname{Im} \psi = K[t]\lambda_1\delta_1 \oplus \ldots \oplus K[t]\lambda_n\delta_n$. Consequently

$$A \cong \frac{F}{\ker \pi} = \frac{F}{\operatorname{Im} \psi} = \frac{K[t]\delta_1 \oplus \ldots \oplus K[t]\delta_n}{K[t]\lambda_1\delta_1 \oplus \ldots \oplus K[t]\lambda_n\delta_n}$$

$$\cong \frac{K[t]\delta_1}{K[t]\lambda_1\delta_1} \oplus \ldots \oplus \frac{K[t]\delta_n}{K[t]\lambda_n\delta_n}$$

$$\cong \frac{K[t]}{<\lambda_1>} \oplus \frac{K[t]}{<\lambda_2>} \oplus \ldots \oplus \frac{K[t]}{<\lambda_n>} \ ,$$

where each λ_i is monic and $\lambda_1|\lambda_2\ldots|\lambda_n$. For some t ($0 \leq t \leq n$), $\lambda_1 = \lambda_2 = \ldots = \lambda_t = 1$ and $\lambda_{t+1},\ldots,\lambda_n$ are nonconstant. Thus for $i \leq t$, $\frac{K[t]}{<\lambda_i>} = \frac{K[t]}{(1)} = 0$ and for $i > t$, $\frac{K[t]}{<\lambda_i>} \cong A_i$ is a cyclic K[t]-module of order λ_i. Therefore A is the direct sum of nonzero torsian cyclic K[t]-submodules $A_{t+1},\ldots,A_n$ of orders $\lambda_{t+1},\ldots,\lambda_n$ respectively, such that $\lambda_{t+1}|\lambda_{t+2}|\ldots|\lambda_n$. Since the K[t]-module structure of A is induced by ϕ, $0 = \lambda_i \cdot A_i = \lambda_i(\phi)A_i$.

We remark that any one of the cyclic submodules in the direct sum of A such as $\frac{K[t]}{<\lambda_i(t)>}$ can be further decomposed into the direct sum of cyclic submodules according to the prime power factorization of the order $\lambda_i(t)$. If $\lambda_i(t) = (t-c_1)^{\alpha_1}\ldots(t-c_k)^{\alpha_k}$, then

$$\frac{K[t]}{<\lambda_i(t)>} \simeq \frac{K[t]}{<(t-c_i)^{\alpha_1}>} \oplus \ldots \oplus \frac{K[t]}{<(t-c_k)^{\alpha_k}>} .$$

<u>Definition</u> 5.2: The nth invariant factor $\lambda_n(t)$ of $tI_n - A$, is called a <u>minimal polynomial</u> of ϕ or M. Det $|tI_n - A|$ is called the <u>characteristic polynomial</u> of ϕ or M.

<u>Lemma</u> 5.1. <u>If in Theorem</u> 4.1 <u>each</u> λ_i <u>has distinct roots, then</u> M <u>is similar to a diagonal matrix</u>.

Proof: Assume that the minimal polynomial $\lambda_m(t)$ is a product of distinct linear factors $(t-c_1),(t-c_2),\ldots,(t-c_k)$. Then det $|D| = (t-c_1)^{d_1}\ldots(t-c_k)^{d_k}$ where $d_1 + d_2 + \ldots + d_k = n$. Thus we have

$$A \cong \frac{K[t]}{<(t-c_1)^{d_1}>} \oplus \ldots \oplus \frac{K[t]}{<(t-c_k)^{d_k}>} \cong E_1 \oplus \ldots \oplus E_k \text{ where dimension}$$

of E_i is d_i. Let N be a diagonal matrix which has for its diagonal entries the scalers c_i, each repeated d_i times. Then the matrix N has the block form

$$N = \begin{bmatrix} c_1I_1 & 0 & \cdots & 0 \\ 0 & c_2I_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & c_kI_k \end{bmatrix}$$

where I_j is the $d_j \times d_j$ identity matrix. Now for each i let Y_i be a basis for E_i such that matrix of $\phi|E_i$ relative to Y_i is c_iI_i. Since $E \tilde{=} E_1 \oplus E_2 \oplus \ldots \oplus E_k$, it follows that $Y = \bigcup_{i=1}^{k} Y_i$ is a basis of E. The basis $Y_1,\ldots,Y_k$ collectively string together to form the sequence of columns of a matrix P:

$$P = [P_1,P_2,\ldots] = (Y_1,\ldots,Y_k).$$

Since Y contains n linearly independent vectors, P is a n×n invertible matrix and $P^{-1}MP = N$ is diagonal.

<u>Definition</u> 5.3: <u>The scalers</u> $c_1,c_2,\ldots,c_k$ <u>are called the eigenvalues of the matrix</u> M.

<u>Lemma</u> 5.2. <u>Let</u> A <u>be a nonsingular integral matrix that is diagonalizable over the complex numbers. Suppose further that each eigenvalue of</u> A <u>has norm</u> 1. <u>Then for some</u> K, $A^K = I$.

Proof: We have $A = RDR^{-1}$ where

$$D = \begin{bmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \alpha_2 & \\ 0 & & \ddots \; \alpha_N \end{bmatrix}, \quad |\alpha_i| = 1.$$

Let $R = (x_{ij})$ and $R^{-1} = (y_{ij})$ and let M_1 and M_2 be upper bounds for the entries of $|R|$ and $|R^{-1}|$ respectively. That is $|x_{ij}| \leq M_1$ and $|y_{ij}| \leq M_2$ for all i and j. Now $A^K = RD^K R^{-1}$ and the entries of A^K are $\sum_{i=1}^{N} x_{ki}\alpha_i^K y_{i\rho}$, $1 \leq k \leq N$, $1 \leq \rho \leq N$. Then

$$\left| \sum_{i=1}^{N} x_{ki}\alpha_i^K y_{i\rho} \right| \leq \sum_{i=1}^{N} |x_{ki}||\alpha_i|^K |y_{i\rho}|$$

$$\leq \sum_{i=1}^{N} M_1 \cdot M_2 = N \cdot M_1 \cdot M_2 = N \cdot M.$$

Thus $|A^K| \leq N\,M$. Consider the matrices $A, A^2, \ldots, A^K, \ldots$. There are only finitely many matrices that can be made from integers of absolute value smaller than $N \cdot M$. Thus for some P, Q,

$$A^P = A^Q, \quad A^Q(A^{P-Q} - I) = 0, \quad A^{P-Q} = I.$$

Conjugacy Problem for Short Words

We now turn to a discussion of conjugacy problem for short words of HNN extension of a finitely generated abelian group.

Theorem 5.2. Let $G^* = \langle G, t/t^{-1}At = B\rangle$ be an HNN extension of a finitely generated free abelian group G. Then two words u and v in G^* with $|u| = |v| = 0$ are conjugate in G^* if and only if $\phi^N(u) = v$ for some N.

Proof: Suppose that u and v are conjugate in G^* and $|u| = |v| = 0$. Then

$$v = z_0 t^{\varepsilon_1} \ldots z_{n-1} t^{\varepsilon_n} u t^{-\varepsilon_n} z_{n-1}^{-1} \ldots t^{-\varepsilon_1} z_0^{-1} \qquad (1)$$

for some $w = z_0 t^{\varepsilon_1} \ldots z_{n-1} t^{\varepsilon_n}$ in G^*. Since v is t-free, some t-reduction must be applicable to the right hand side of the above equation. The only possible t-reduction is $t^{\varepsilon_n} u t^{-\varepsilon_n}$, thus we must have $u \in A$ if $\varepsilon_n = -1$, and $u \in B$ if $\varepsilon_n = 1$. Let us assume that $\varepsilon_n = -1$ and that $u \in A$. Then

Equation (1) becomes:

$$v = z_0 t^{\varepsilon_1} \dots z_{n-2} t^{\varepsilon_{n-1}} \phi(u) t^{-\varepsilon_{n-1}} z_{n-2}^{-1} \dots t^{-\varepsilon_1} z_0^{-1}. \tag{2}$$

If $\varepsilon_{n-1} = 1$, then the term $t^{\varepsilon_{n-1}} \phi(u) t^{-\varepsilon_{n-1}}$ can be replaced by u and Equation (2) becomes

$$v = z_0 t^{\varepsilon_1} \dots z_{n-3} t^{\varepsilon_{n-2}} u t^{-\varepsilon_{n-2}} z_{n-3}^{-1} \dots t^{-\varepsilon_1} z_0^{-1}. \tag{3}$$

If $\varepsilon_{n-1} = -1$, we must have $\phi(u) \in A$ and in this case Equation (2) reduces to:

$$v = z_0 t^{\varepsilon_1} \dots z_{n-3} t^{\varepsilon_{n-2}} \phi^2(u) t^{-\varepsilon_{n-2}} z_{n-3}^{-1} \dots t^{-\varepsilon_1} z^{-1}. \tag{4}$$

Note that Equation (3) is of the type of Equation (1). By considering successive t-reduction, we have $v = \phi^m(u)$ for some m.

Now suppose that $\phi^m(u) = v$ for some m, where u and v are t-free. Clearly $u, \phi(u), \dots, \phi^m(u)$ must all be in A. Let $w = z_0 t^{\varepsilon_1} \dots z_{m-1} t^{\varepsilon_m}$, where $\varepsilon_i = -1$ for $i = 1, 2, \dots, m$. Then

$$\begin{aligned} v = \phi^m(u) &= z_0 t^{\varepsilon_1} \phi^{m-1} t^{-\varepsilon_1} z_0^{-1} = z_0 t^{\varepsilon_1} z_1 t^{\varepsilon_2} \phi^{m-2}(u) t^{-\varepsilon_2} z_1^{-1} t^{-\varepsilon_1} z_0^{-1} \\ &= \dots z_0 t^{\varepsilon_1} \dots z_{m-2} t^{\varepsilon_{m-1}} \phi(u) t^{-\varepsilon_{m-1}} z_{m-2}^{-1} \dots t^{-\varepsilon_1} z_0^{1} \\ &= z_0 t^{\varepsilon_1} \dots z_m t^{\varepsilon_m} u t^{-\varepsilon_m} \dots t^{-\varepsilon_1} z_0^{-1} = wuw^{-1}. \end{aligned}$$

Lemma 5.3. *Let* G *be a finitely generated abelian group*, A *and* B *subgroups of* G *and* ϕ: A ⟶ B *a homomorphism*. *Let* $a \in A$. *There is an algorithm that will produce either an integer* K *such that* $\phi^K(a) \notin A$, *or a subgroup* C *of* A *such that* $\phi(C) \subseteq C$ *and* $a \in C$.

Proof: If $\{a, \phi(a), \dots, \phi^N(a)\} \subset A$, put $U_N = \langle a, \phi(a), \phi^2(a), \dots, \phi^N(a) \rangle$.

We have $U_i \subseteq U_{i+1}$. If $\phi^n(a) \in A$ for each n, then $U = \bigcup_{i=1}^{\infty} U_i$ is a finitely generated subgroup of A. Hence for some N_0, $U_K = U_{N_0}$ for all $K \geq N_0$. This happens if and only if $\phi^{N_0+1}(a) \in U_{N_0}$.

The algorithm for each N proceeds as follows. Check that $a, \phi(a), \ldots, \phi^{N+1}(a) \in A$. Then check if $\phi^{N+1}(a) \in U_N$. Eventually either $\phi^K(a) \notin A$, or $\phi^{K+1}(a) \in U_K$.

Theorem 5.3. Let u and v be t-free words of the HNN group

$$G^* = \langle G, t \mid t^{-1}At = B, \phi \rangle$$

with G a finitely generated free abelian group. There is an algorithm that will determine whether u and v are conjugate in G^*.

Proof: Let u and v be elements of G^* with $|u| = |v| = 0$. Then u and v are conjugate if and only if u and v are in appropriate subgroups A or B and $\phi^N(u) = v$ for some N. According to Lemma 5.3, there is an algorithm that will produce an integer K such that $\phi^K(u) \notin A$ or a subgroup C of A such that $\phi(C) \subseteq C$ and $u \in C$. If such K exists, it is sufficient to check the values $n = 1, 2, \ldots, K-1$ to see if $\phi^n(u) = v$. Thus let us assume there is a subgroup C of A containing u such that $\phi(C) \subseteq C$. The monomorphism $\phi|C: C \longrightarrow C$ induces an action of $Z[t]$ on C which gives C the structure of $Z[t]$-module. If we write C additively, then $t \cdot c = \phi(c)$ and $P(t) \cdot c = P(\phi(c))$ for $c \in C$ and $P(t) \in Z[t]$. Tensoring with the field of rational numbers, we obtain a $Q[t]$-module $C \otimes_Z Q$ via $t \cdot (c \otimes 1) = t \cdot c \otimes 1$. If C is generated as a Z-module by $x_1, \ldots, x_n$, then $C \otimes_Z Q$ is generated as a Q-module by $x_1 \otimes 1, \ldots, x_n \otimes 1$. There is an isomorphism $\alpha: C \otimes_Z Q \cong \bigoplus_{i=1}^{n} Q$ under the map $(c_1, \ldots, c_n) \otimes a \mapsto (ac_1, \ldots, ac_n)$. We have by Theorem

5.1

$$C \otimes_Z Q \cong \frac{Q[t]}{<\lambda_1>} \oplus \ldots \oplus \frac{Q[t]}{<\lambda_m>}$$

where $\lambda_i \in Q[t]$ and $\lambda_i | \lambda_{i+1}$. Since the map $c \mapsto c \otimes 1$ defines an embedding of C into $C \otimes_Z Q$, u and v can be identified by polynomials $\overline{P}(t) = (\overline{P}_1(t),\ldots,\overline{P}_n(t)$ and $\overline{Q}(t) = (\overline{Q}_1(t),\ldots,\overline{Q}_n(t))$ where $\overline{P}_i(t)$ and $\overline{Q}_i(t)$ are in $\frac{Q[t]}{<\lambda_i>}$ with degree P_i and degree Q_i less than degree λ_i for each i. Note that $t^N\overline{P}(t) = \overline{Q}(t)$ if and only if $t^N\overline{P}_i(t) = \overline{Q}_i(t)$ for each i. That is $t^N P_i(t) - Q_i(t) = 0$ in $\frac{Q[t]}{<\lambda_i>}$ for each i. Thus $t^N\overline{P}(t) = \overline{Q}(t)$ if and only if $t^N P_i(t) - Q_i(t)$ is a multiple of $\lambda_i(t)$ for each i. Now if $\phi^N(u) = v$ for some N, then $\lambda_i(t)$ divides $t^N P_i(t) - Q_i(t)$ for each i since $\phi^N(u) = t^N \cdot u \equiv t^N\overline{P}(t) = \overline{Q}(t) \equiv v$. To complete the algorithm we will find a lower bound K for N such that $\lambda_i(t)$ does not divide $t^N P_i(t) - Q_i(t)$ for some i and all $N \geq K$. Then it follows that $\phi^N(u) \neq v$ for $N \geq K$ By the Remark after the Definition 4.2, we may assume that λ_i, P_i and Q_i are relatively prime in pairs for each i.

Case 1. If $\lambda_i(t)$ has a zero off the unit circle for some i, then by Theorem 4.1 an integer K can be found such that $\lambda_i(t)$ does not divide $t^N P_i(t) - Q_i(t)$ for $N \geq K$.

It now remains that all the zeros of $\lambda_i(t)$ are on the unit circle for each i.

Case 2. $\lambda_i(t)$ has a multiple zero on the unit circle for some i. By applying Theorem 4.3 the result follows.

Case 3. Each λ_i has distinct zeros. According to Lemma 5.2 there exists an integer K such that $(\phi|C)^K(u) = u$. Thus if $(\phi|C)^n(u) \neq v$ for $n = 1,2,\ldots,K$, then $(\phi|C)^N(u) \neq v$ for any $N \geq K$.

BIBLIOGRAPHY

1. Boler, J., "Conjugacy in abelian-by-cyclic groups", American Mathematical Society, 55, 1 (1976).

2. Boone, W. W., "Certain simple unsolvable problems of group theory", Indiny. Math., 16 (1954), 231-237 & 492-497.

3. Britton, J. L., "The word problem", Ann. of Math., 77 (1963), 16-32.

4. Britton, J. L., "On the conjugacy problem and difference equation", J. London Math. Soc. (2) 17 (1978), No. 2, 240-245.

5. Dehn, M., "Uber unendliche diskontinuierliche Gruppen", Math. Annalen, 71 (1911), 114-116.

6. Formannek, E., "Conjugacy separability in polycyclic groups", Journal of Algebra, 42 (1979), 1-10.

7. Fridman, A. A., "On the relation between the word problem and the conjugacy problem in finitely defined groups", (Russian), Trudy Moskov, Math. Obsc., 9 (1960), 329-356.

8. Hungerford, T. W. Algebra. New York: Holt, 1974.

9. Larsen, L., "The conjugacy problem and cyclic HNN constructions", J. Austral. Math. Soc., 23, Series A (1977), 385-401.

10. Lyndon, R. C., P. E. Schupp. Combinatorial Group Theory. New York: Springer-Verlag, 1977.

11. Miller, C., III, "On group-theoretic decision problems and their classification", Annals of Mathematics Studies, 68 (1971).

12. Magnus, W., A. Karrass, D. Solitar. Combinatorial Group Theory. New York: Wiley, 1966.

13. Mikhlailova, K. A., "The occurrence problem for direct products of groups", (Russian), Dokl. Akad. Nauk SSSR, 119 (1958), 1103-1105.

14. Novikov, P. S., "On the algorithmic unsolvability of the word problem in groups", (Russian), Trudy Math. Inst. Steklov, Izdat Akad. Nauk. SSSR, No. 44, Moscow, 1955.

15. Ralston, A. A First Course in Numerical Analysis. McGraw-Hill, 1965.

16. Rabin, M. O., "Recursive unsolvability of group theoretic problems", Ann. of Math., 67 (1958), 172-194.

17. Toh, K. H., "Problems concerning residual finiteness in nilpotent groups", to be published.

VITA

2

Farrokh Abedi

Candidate for the Degree of

Doctor of Philosophy

Thesis: SOLVABILITY OF THE CONJUGACY PROBLEM FOR HNN EXTENSIONS OF FINITELY GENERATED FREE ABELIAN GROUPS

Major Field: Mathematics

Biographical:

Personal Data: Born in Tehran, Iran, August 20, 1949, the son of Mr. and Mrs. M. S. Abedi.

Education: Graduated from Arash High School, Tehran, Iran, in May, 1967; received Bachelor of Science in Mathematics degree from Pars College, Tehran, Iran, in August, 1973; received Master of Arts degree in Mathematics from Eastern New Mexico University in August, 1975; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in May, 1983.

Professional Experience: Grader in the Department of Mathematics, Eastern New Mexico University, 1974-1975; Graduate Teaching Associate in the Department of Mathematics, Oklahoma State University, 1975-1982.