Can. J. Math. Vol. 45 (2), 1993 pp. 225–254

ON ALMOST PRIMITIVE ELEMENTS OF FREE GROUPS WITH AN APPLICATION TO FUCHSIAN GROUPS

A. M. BRUNNER, R. G. BURNS AND SHEILA OATES-WILLIAMS

ABSTRACT An element of a free group F is called *almost primitive* in F, if it is primitive in every proper subgroup containing it, though not in F itself Several examples of almost primitive elements (APEs) are exhibited The main results concern the behaviour of proper powers w^{ℓ} of certain APEs w in a free group F (and, more generally, in free products of cycles) with respect to any subgroup H containing such a power "minimally" these assert, in essence, that either such powers of w behave in H as do powers of primitives of F, or, if not, then they "almost" do so and furthermore H must then have finite index in F precisely determined by the smallest positive powers of conjugates of w lying in H Finally, these results are applied to show that the groups of a certain class (potentially larger than that of finitely generated Fuchsian groups) have the property that all their subgroups of infinite index are free products of cyclic groups

1. **Introduction.** The concept of "primitivity" of an element w of a free group F is well-known: w is *primitive* in F, if it can be included in some free basis for F. It is natural then to define an element of a free group F to be an *almost primitive element* of F (briefly, APE) if it is primitive in every proper subgroup containing it, though not in the whole group F. Here are some examples of inequivalent (*i.e.* not transformable one into another by means of an automorphism) "irreducible" APEs.

1.1 EXAMPLES OF ALMOST PRIMITIVE ELEMENTS. (i) x^p, p prime, in the infinite cyclic group $F(x)$;

(ii) the commutator $[x,y](:= x^{-1}y^{-1}xy)$ in the free group $F(x,y)$ of rank 2;

(iii) $x[y,x]$, xyx^2y^3 in $F(x,y)$.

(Here (i) is easy. We indicate briefly a proof of (ii): Let $[x,y] \in H < F(x,y)$. If $x \notin H$, then there is a right Schreier transversal for H in F containing x^{-1}, and then the Schreier rewrite of $x^{-1}y^{-1}xy$ will yield a word of length ≤ 3 in the corresponding Schreier free generators of H. As squares or cubes of such generators are easily ruled out, it follows that $[x,y]$ is primitive in H. On the other hand if $x \in H$, then $y \notin H$, and the preceding argument applies to $[y,x] = [x,y]^{-1}$ with the roles of x and y interchanged. That the elements in (iii) are primitive in every proper subgroup containing them can be established along similar lines. That they are not primitive is an easy consequence of J. Nielsen's result that the natural epimorphism $\operatorname{Aut} F(x,y) \longrightarrow \operatorname{Aut} A_2$, where A_2 is the abelianization of $F(x,y)$, has kernel $\operatorname{Inn} F(x,y)$. That they are inequivalent follows by applying Whitehead's algorithm (see *e.g.* [4], p. 166).)

Received by the editors July 10, 1992
AMS subject classification Primary 20E05, secondary 20F38

The notion of "primitivity", and so also "almost primitivity", may be extended to arbitrary groups: We shall say that an element g of an arbitrary group G is *primitive* in G if g generates an infinite cyclic free factor of G, *i.e.* if g has infinite order, and $G = \langle g \rangle * G_1$ for some $G_1 < G$. An element of G is then *almost primitive* in G if, as before, it is primitive in every proper subgroup containing it, but not in the whole group G. An example germane to the application we have in view is that of a finite cyclic group C_n: any generator is (vacuously) an APE in C_n.

The following simple result (proved in §2) furnishes a method of obtaining new APEs from known "irreducible" ones.

PROPOSITION 1.2. *Let A, B be arbitrary groups containing* APE*s a, b respectively. Then the product ab is either primitive or an* APE *in the free product $A * B$.*

(It is highly likely that in fact ab is an APE in $A * B$ if and only if a, b are APEs in A, B respectively; this may be provable for instance by modifying part of the proof of the Grushko-Neumann theorem in [5]. As this stronger result will not be needed, we shall not pursue the matter further here.)

Of the resulting examples we single out two particular types for attention.

1.3 EXAMPLES OF REDUCIBLE APEs. (i) (*cf.* [7, Theorem 2.2]). In the free group $F = F(x_i, y_i, z_j; 1 \leq i \leq m, 1 \leq j \leq n)$ of rank $2m + n$ the element

$$[x_1, y_1] \cdots [x_m, y_m] z_1^{p_1} \cdots z_n^{p_n}, \tag{1}$$

(where the p_i are not necessarily distinct primes) is an almost primitive element of F.

(ii) Let $c_1, \ldots, c_k$ be generators of finite cyclic groups $C_{n_1}, \ldots, C_{n_k}$. In the group

$$K = C_{n_1} * \cdots * C_{n_k} * F(x_i, y_i; 1 \leq i \leq m), \tag{2}$$

the free product of the k finite cycles and $2m$ infinite cycles, the element

$$c_1 \cdots c_k [x_1, y_1] \cdots [x_m, y_m] \tag{3}$$

is almost primitive.

The example (ii) is of interest (see also below) since setting the element (3) equal to 1 in K yields the general presentation of the finitely generated Fuchsian groups (other than those that are free products of cyclic groups).

Our main results concern what might be termed the "primitive" behaviour of powers of APEs in subgroups containing such powers minimally. If x is primitive in a group G, and x^n is the smallest positive power of x in a subgroup H, then it is not difficult to show that x^n is primitive in H. In fact much more is true: there is a full set of representatives g_i, $i \in I$, for the double cosets $Hg\langle x \rangle$, with the property that the smallest positive powers $g_i x^{\ell_i} g_i^{-1}$ in H (for those i for which $g_i \langle x \rangle g_i^{-1} \cap H \neq \{1\}$) form a "coprimitive" subset of H, *i.e.* freely generate a free factor of H. (This can be established easily using the Kurosh subgroup theorem.) The situation for an almost primitive element is more complex, although, at least in some cases, still tractable. For the APE $[x, y]$ in $F(x, y)$ we have the following result.

THEOREM 1.4. *Let H be any subgroup of $F(x,y)$, and let $\{a_i \mid i \in I\}$ be a full set of representatives (including 1) of the double cosets $Hg\langle[x,y]\rangle$, $g \in F$. Let $J \subseteq I$ consist of those indices j for which*

$$a_j\langle[x,y]\rangle a_j^{-1} \cap H \neq \{1\},$$

and for each such j denote by m_j the least positive integer such that $a_j[x,y]^{m_j}a_j^{-1} \in H$. Then the elements $a_j[x,y]^{m_j}a_j^{-1}$ are all distinct, and the a_i, $i \in I$, may be so chosen that either:

(i) the $a_j[x,y]^{m_j}a_j^{-1}$, $j \in J$, are coprimitive in H; or

(ii) they form a finite, non-empty, minimally non-coprimitive set in H (in the sense that every proper subset, but not the whole set, is coprimitive—one might use the term "almost coprimitive"), and furthermore

$$|F:H| = \sum_{j\in J} m_j \quad (<\infty),$$

(so that in fact $J = I$).

Motivated by this result, we make the following definition.

DEFINITION 1.5. Let w be an APE in an arbitrary group G. For any subgroup $H \leq G$, let $\{g_i \mid i \in I\}$ be a full set of representatives, including 1, of the double cosets $Hg\langle w\rangle$, let $\{g_j \mid j \in J \subseteq I\}$ be the subset of those representatives for which

$$g_j\langle w\rangle g_j^{-1} \cap H \neq \{1\},$$

and for each $j \in J$ denote by ℓ_j the least positive integer such that $g_jw_j^{\ell}g_j^{-1} \in H$. We shall say that the APE w is *tame* in G if the representatives g_i, $i \in I$, can always (*i.e.* for every $H \leq G$) be chosen so that the $g_jw^{\ell_j}g_j^{-1}$, $j \in J$, are all distinct, and either:

(i) are coprimitive in H, or

(ii) form a non-empty, finite, minimally non-coprimitive subset of H, which furthermore has finite index in G, given by

$$|G:H| = \sum_{j\in J} \ell_j \quad (<\infty),$$

(whence, in particular, $J = I$).

Thus by Theorem 1.4, the commutator $[x,y]$ is a tame APE in $F(x,y)$, and our object is to identify as many more tame APEs as possible. It is easy to verify that a generator of a finite cyclic group is (vacuously) tame. However x^p, p prime, is not tame in $F(x)$ and, as we shall see below, neither is x^py^q (p,q primes) in $F(x,y)$. However the "free product" of tame APEs is again tame.

THEOREM 1.6. *If a and b are tame* APE*s in groups A and B respectively, then their product ab is tame in $A * B$.*

COROLLARY 1.7. *In the group (2), namely*

$$K = C_{n_1} * \cdots * C_{n_k} * F(x_i, y_i; 1 \leq i \leq m),$$

the element $c_1 \cdots c_k[x_1, y_1] \cdots [x_m, y_m]$ *is a tame* APE.

COROLLARY 1.8 (*cf.* ROSENBERGER [6]). *If w is a tame* APE *in a group G (e.g. as in the preceding corollary) and $H \leq G$ contains w^λ, $\lambda > 0$, but no smaller positive power of w, then either*

(i) w^λ is primitive in H, or

(ii) w^λ is an APE *of H, and H has index λ in G (so that $\{1, w, \ldots, w^{\lambda-1}\}$ is a complete left or right coset representative system for H in G).*

(This is immediate from the definition of "tame APE", except for the assertion in (ii) that w^λ is an APE in H. To see this, suppose $H_1 < H$ is such that $w^\lambda \in H_1$. Then H_1 does not satisfy (ii) since the index is wrong, so that (i) must hold for this subgroup H_1, *i.e.* w^λ must be primitive in H_1.)

We note that the example on p. 172 of [7] shows that the alternative (ii) can arise.

COROLLARY 1.9 (HOARE, KARRASS, SOLITAR [1]). *In a finitely generated Fuchsian group, the subgroups of infinite index are free products of cyclic groups.*

PROOF. If the Fuchsian group is a free product of cycles, then by the Kurosh subgroup theorem so is every subgroup. Hence we may suppose that the Fuchsian group is isomorphic to

(4)
$$\hat{K} := K/\langle w\rangle^K$$
$$\cong \langle c_1, \ldots, c_k; x_1, y_1, \ldots, x_m, y_m \mid c_1^{n_1} = \cdots = c_k^{n_k} = c_1 \cdots c_k[x_1, y_1] \cdots [x_m, y_m] = 1\rangle,$$

where K is as in (2) and w as in (3). Write N for the normal closure in K of the word w. Let $\hat{H}$ be any subgroup of infinite index in $K = K/N$ (the group (4)); its complete inverse image, H say, under the natural epimorphism $K \longrightarrow K/N$, is then a subgroup of infinite index in K, containing N. Since $H \geq N$, we have $gwg^{-1} \in H$ for all $g \in K$, so that by Corollary 1.7 there is a full set $\{g_\iota \mid i \in I\}$ of representative of double cosets $Hg\langle w\rangle$ in K such that the elements

$$h_\iota := g_\iota w g_\iota^{-1}, \quad i \in I,$$

are coprimitive in H. Let $H_1 \leq H$ be such that

$$H = \langle h_\iota \mid i \in I\rangle * H_1.$$

Now each $g \in G$ has the form $g = hg_\iota w^q$ for some $h \in H$, $i \in I$, $q \in \mathbb{Z}$ (all depending on g), whence

$$gwg^{-1} = hg_\iota w g_\iota^{-1} h^{-1} = hh_\iota h^{-1}.$$

From this it is clear that N is the normal closure in H of the set $\{h_\iota \mid i \in I\}$, so that $H/N \cong H_1$, which has the requisite structure (by the Kurosh subgroup theorem—see below). ∎

REMARKS. 1. This argument applies (potentially) more generally: If any tame APE in a free product of cycles is set equal to 1, then in the resulting group all subgroups of infinite index will be free products of cycles.

2. We may conclude from this proof that $x^p y^q$ is not a tame APE, since the group $\langle x, y \mid x^p y^q = 1\rangle$ has infinite-index subgroups that are free abelian of rank 2.

The proofs of our main results (Theorems 1.4 and 1.6), though lengthy, are elementary (given some familiarity with the Schreier and Kurosh subgroup theorems), and so provide a possible alternative approach to the exploration of structure of a natural class of groups potentially wider than that of the finitely generated Fuchsian groups.

The following questions naturally suggest themselves.

1. Are there tame APEs other than those of the form (3) (and their automorphic images)? Are, for instance, $[x, y]z^p$, $x^p y^q z^r$ (p, q, r prime) tame APEs in $F(x, y, z)$?

2. Is it possible to classify all irreducible APEs, at least of $F(x, y)$?

2. **Preliminaries to the proof of Theorem 1.6: The Kurosh subgroup theorem.** Let A, B be groups. It is well-known that each element g of their free product $A * B$ has a unique *normal* (or *reduced*) form as an "alternating" product $g = d_1 \cdots d_n$, where $n \geq 0$ and the d_i belong to $A \setminus \{1\}$ or $B \setminus \{1\}$ with adjacent d_i from different factors. For $n \geq 1$ we call the $d_1 d_2 \cdots d_i$, $i \geq 0$, *initial segments* of g, and d_1 the *beginning* of g; *terminal segments*, and the *ending* of g are defined analogously.

For a precise formulation of the Kurosh subgroup theorem for $A * B$, the concept of a "Kurosh system", analogous to that of a Schreier transversal for a subgroup of a free group, is useful.

DEFINITION 2.1. Let $H \leq A * B$. A *Kurosh system* for H in $A * B$, is a pair (T_A, T_B) of right transversals for H in $A * B$, both containing 1, with the following properties:

(i) each of the sets T_A, T_B is closed under taking initial segments;

(ii) if $t \in T_A \cup T_B$ ends in an element of A, then $t \in T_A$ (and, similarly, if the ending of t is in B, then $t \in T_B$);

(iii) for each fixed $s \in T_A$ such that either $s = 1$ or s ends in an element of $B \setminus \{1\}$, the set of elements $\alpha \in A$ such that $s\alpha \in T_A$, forms a right transversal for $s^{-1}Hs \cap A$ in A (and analogously with A replaced by B throughout).

We shall also need the notation S_A for the subset of T_A containing 1 and those elements of T_A whose ending is from $B \setminus \{1\}$ (and S_B analogously).

THEOREM 2.2 (THE KUROSH SUBGROUP THEOREM). *Let $H \leq A * B$. There exists a Kurosh system for H in $A * B$, and for any such system the following assertions are valid:*

(i) the set

$$\Phi := \{t[\varphi(t)]^{-1} \mid t \in T_A \setminus T_B\}, \tag{5}$$

where $\varphi\colon T_A \longrightarrow T_B$ is the bijection defined by $Ht = H\varphi(t)$, is a free basis for the subgroup F, say, it generates;

(ii) the subgroup H is generated by the subgroup F together with the subgroups $s_A A s_A^{-1} \cap H$, $s_A \in S_A$, and $s_B B s_B^{-1} \cap H$, $s_B \in S_B$, as the free product of these subgroups:

$$H = F * \Big[\prod_{s_A \in S_A}^{*} (s_A A s_A^{-1} \cap H)\Big] * \prod_{s_B \in S_B}^{*} (s_B B s_B^{-1} \cap H). \tag{6}$$

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

This theorem may be deduced from a "rigidity" or "non-cancellation" property of the generators of H given by (ii). As we shall be exploiting this property, we shall now describe it in some detail. We shall call the non-trivial elements of the groups $s_A A s_A^{-1} \cap H$, $s_A \in S_A$, of the $s_B B s_B^{-1} \cap H$, $s_B \in S_B$, and the elements of Φ *Kurosh generators* of H (relative to a given Kurosh system (T_A, T_B)). We define the "significant", *i.e.* non-cancelling, syllables of a Kurosh generator as follows: Taking any $t = t_1 a_1 \in T_A \setminus T_B$, and writing $\varphi(t) = \tau_1 b_1 \in T_B \setminus T_A$ (so that a_1, b_1 are the endings of t, $\varphi(t)$ respectively), we have

$$t[\varphi(t)]^{-1} = t_1 a_1 b_1^{-1} \tau_1^{-1}$$

as the reduced form of a typical element of Φ; the *significant symbols* of this element are then defined to be the symbols a_1 and b_1^{-1}. For a Kurosh generator of the form $s_A \alpha s_A^{-1}$, $\alpha \in A \setminus \{1\}$, on the other hand, we define the *significant symbol* to be α, the central one, and make the analogous definition for Kurosh generators of the form $s_B \beta s_B^{-1}$, $\beta \in B \setminus \{1\}$. The well-known non-cancellation property of the significant symbols is as follows.

LEMMA 2.3. *(i)* $\Phi \cap \Phi^{-1} = \emptyset$.

(ii) Let $h_1 \cdots h_n$ be a product of $n \geq 1$ Kurosh generators of $H \leq A * B$ or their inverses, where adjacent h_i are not mutually inverse elements of $\Phi \cup \Phi^{-1}$, nor in the same subgroup $s_A A s_A^{-1} \cap H$ or $s_B B s_B^{-1} \cap H$. (In other words $h_1 \cdots h_n$ is, potentially or formally, the normal form of a non-trivial element of the free product on the right-hand side of (6).) Then in reducing $h_1 \cdots h_n$ to its normal form as an element of the free product $A * B$, the significant symbol(s) of each h_i remain uncancelled, although they may be consolidated (i.e. merge without cancelling, with symbols from the same free factor A or B).

For our proof of Theorem 1.6 to follow, we shall be needing a rather special Kurosh system (T_A, T_B) for H in $A * B$, defined in terms of the given elements $a \in A$, $b \in B$, by modifying the usual construction. In that construction, one first chooses the identity element 1 as representative of each of the double cosets HA, HB, and then, assuming inductively that A- and B-representatives have been chosen (*i.e.* to go into T_A and T_B) for the right cosets of H containing elements of length $< k$ (where $k \geq 1$), one chooses double-coset representatives of the (H, A)-double cosets of length k by choosing from each such double coset HgA an element $\hat{t}$ of smallest length k, say $\hat{t} = t_1 b_1$ in reduced form, where $b_1 \in B \setminus \{1\}$ and then replacing t_1 by its A-representative t_2 say, already chosen (by the inductive hypothesis), obtaining thereby $t = t_2 b_1$ as the new member of T_A (and T_B). The A-representatives of the cosets $Hg_1 \subseteq HgA = HtA$ are then obtained, according to the usual procedure, by arbitrarily choosing a right transversal $\{\alpha\}$, containing 1, for $t^{-1}Ht \cap A$ in A, and including in T_A all multiples of t on the right by the members of this transversal, *i.e.* all $t\alpha$ (*cf.* the defining condition 2.1(iii)). The procedure for constructing T_B is analogous.

Relative to prescribed elements $a \in A$, $b \in B$ (ultimately to be APEs in the groups A, B respectively), the inductive step in the foregoing construction may be modified as follows: Rather than choose the transversal $\{\alpha\}$ arbitrarily, we first choose representatives a_ι, $i \in I_t^A$ (some index set) of the $(t^{-1}Ht \cap A, \langle a \rangle)$-double cosets in A, choosing 1 to represent $(t^{-1}Ht \cap A)\langle a \rangle$, and then augment these to the desired full transversal for $t^{-1}Ht \cap A$ in A given by adjoining the elements

$$
(7) \qquad \begin{gathered} a_\iota a^{j_\iota},\ 0 \leq j_\iota < m_\iota(t) \text{ if } a_\iota^{-1}t^{-1}Hta_\iota \cap \langle a \rangle = \langle a^{m_\iota(t)} \rangle,\ m_\iota(t) > 0, \\ a_\iota a^m,\ m \in \mathbb{Z} \text{ if } a_\iota^{-1}t^{-1}Hta_\iota \cap \langle a \rangle = \{1\}. \end{gathered}
$$

(The B-endings to be attached to $s \in S_B$ (s assumed already constructed) are chosen analogously using $\langle b \rangle$ in place of $\langle a \rangle$.)

We end this section with the

PROOF OF PROPOSITION 1.2. Let $H < A * B$ be such that $ab \in H$. Suppose first that $a \in H$ (whence also $b \in H$). By the Kurosh subgroup Theorem (2.2) H can be decomposed as a free product of the form

$$
H = (A \cap H) * (B \cap H) * H_1.
$$

If $A \cap H < A$, then a is primitive in $A \cap H$, and therefore, in view of the above free decomposition of H, is also primitive in H. If on the other hand $A \cap H = A$, then we must have $B \cap H < B$, and we infer in the same way that b is primitive in H. In either case it follows that ab is primitive in H.

If now $a \notin H$, then also $b \notin H$, and we may choose a Kurosh system (T_A, T_B) for H in $A * B$ with $a \in T_A$, $b^{-1} \in T_B$ (by the usual procedure for constructing a Kurosh system described above). This done, we shall have $a(b^{-1})^{-1}$ as an element of the set Φ (see (5)), so that ab is primitive in H. ∎

3. **Proof of Theorem 1.6: The free product of tame** APEs **is a tame** APE. As in that theorem, let a, b be tame APEs in the groups A, B respectively, and let $H \leq A * B$. We construct a "minimal" Kurosh system for H as above (see in particular (7)), but now being more particular in our choice, at the inductive step, of the double-coset representatives a_ι to be attached as A-endings to an element t already constructed in S_A (and likewise in our choice of the B-endings b_ι to be attached to the $s \in S_B$). The choice of these a_ι, $i \in I_t^A$, is made as follows: By hypothesis, Definition 1.5 applies to the APE a of A, in particular with respect to the subgroup $t^{-1}Ht \cap A$. Choose a_ι, $i \in I_t^A$, to form a full set of representatives of double cosets $(t^{-1}Ht \cap A)\alpha\langle a \rangle$, $\alpha \in A$, for which one of the conditions (i) or (ii) of Definition 1.5 is satisfied. Thus if the subset $\{a_j \mid j \in J_t^A \subseteq I_t^A\}$ consists of those representatives for which $ta_j\langle a \rangle a_j^{-1}t^{-1} \cap H \neq \{1\}$, and for each $j \in J_t^A$, $m_j(t)$ denotes the least positive integer for which $ta_ja^{m_j(t)}a_j^{-1}t^{-1} \in H$, then, as well as being distinct,

(i) the elements

$$
(8) \qquad ta_ja^{m_j(t)}a_j^{-1}t^{-1},\ j \in J_t^A, \text{ are coprimitive in } H \cap tAt^{-1}, \text{ or}
$$

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

(ii) these elements form a finite minimally non-coprimitive set in $H \cap tAt^{-1}$, and

(9) $$|tAt^{-1} : H \cap tAt^{-1}| = \sum_{j \in J_t^A} m_j(t) \quad (< \infty).$$

Similarly, in the inductive step in the construction of T_B (carried out in step with the construction of T_A), we choose $(s^{-1}Hs \cap B, \langle b \rangle)$-double coset representatives b_ι, $i \in I_s^B$ (s assumed already constructed in S_B) to be attached as B-endings to s, such that Definition 1.5 is satisfied for these b_ι with respect to the subgroup $s^{-1}Hs \cap B \leq B$ and the APE b in B; *i.e.* if $\{b_j \mid j \in J_s^B \subseteq I_s^B\}$ is the subset of those representatives among the b_ι, $i \in I_s^B$, for which $sb_j\langle b \rangle b_j^{-1}s^{-1} \cap H \neq \{1\}$ and $n_j(s)$ denotes the least positive integer such that $sb_j b^{n_j(s)} b_j^{-1} s^{-1} \in H$, then one of the analogues of (i), (ii) above occurs.

Having thus carefully chosen the (minimal) *special* Kurosh system (T_A, T_B) for H, let $\{g_\iota \mid i \in I\}$ be a full set of $(H, \langle ab \rangle)$-double coset representatives of least length subject to being in T_A. Let $\{g_j \mid j \in J \subseteq I\}$ be the subset of these representatives for which $g_j\langle ab \rangle g_j^{-1} \cap H \neq \{1\}$, and for each $j \in J$ denote by l_j the least positive integer satisfying $g_j(ab)^{l_j} g_j^{-1} \in H$. We shall show that the defining conditions for a tame APE hold for $ab \in A * B$ with respect to $H \leq A * B$ and this choice of g_ι, $i \in I$, that is that the $g_j(ab)^{l_j} g_j^{-1}$ are distinct, and:

(10) (i) they are coprimitive in H, or

(11) (ii) they form a finite, minimally non-coprimitive set in H, and
$$|F : H| = \sum_{j \in J} l_j \quad (< \infty).$$

We shall henceforth *assume that (i) does not occur* and deduce from this that (ii) must occur.

We separate out portions of the proof as lemmas (the first of which uses only the assumption that the appropriate conjugates are distinct).

LEMMA 3.1. *Let $H \leq A * B$ and let (T_A, T_B) be a special Kurosh system chosen as above relative to the* APE*s $a \in A$, $b \in B$. The following assertions are valid.*

(i) In the reduced rewritten expressions

(12) $$g_j(ab)^{l_j} g_j^{-1} = h_{j1} \cdots h_{jr_j}, \quad j \in J,$$

for the $g_j(ab)^{l_j} g_j^{-1}$ (see above) in terms of the Kurosh generators determined by the special Kurosh system (T_B, T_B), any Kurosh generator h_{jk} of "conjugate form", e.g. from some subgroup $tAt^{-1} \cap H$, $t \in S_A$, has the form $ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$ for some representative $a_\iota \in A$ (chosen as above) of a $(t^{-1}Ht \cap A, \langle a \rangle)$-double coset. Moreover each such Kurosh generator occurs at most once in the totality of the right-hand-side expressions in (12). (The analogous assertion is valid with B, b in place of A, a.)

(ii) Any Kurosh generator from Φ (see (5)) appearing in a rewritten expression in (12), occurs at most twice in the totality of such expressions, and if twice then once to each of the exponents ± 1.

PROOF. Applying to $g_j(ab)^{l_j}g_j^{-1}$ the left-to-right "Kurosh rewriting process" for systematically expressing an element of H in terms of our Kurosh generators (see [4, p. 230]), one obtains generators of the following forms:

$$(13)\qquad \begin{aligned} &\text{of conjugate form (or trivial)} \begin{cases} \varphi_A[g_j(ab)^{k_j}]a\{\varphi_A[g_j(ab)^{k_j}a]\}^{-1}; \\ \varphi_B[g_j(ab)^{k_j}a]b\{\varphi_B[g_j(ab)^{k_j+1}]\}^{-1}; \end{cases} \\ &\text{from } \Phi\cup\Phi^{-1} \text{ (if nontrivial)} \begin{cases} \varphi_A[g_j(ab)^{k_j}a]\{\varphi_B[g_j(ab)^{k_j}a]\}^{-1}; \\ \varphi_B[g_j(ab)^{k_j+1}]\{\varphi_A[g_j(ab)^{k_j+1}]\}^{-1}; \end{cases} \end{aligned}$$

where $0 \le k_j < l_j$, and $\varphi_A: A*B \longrightarrow T_A$, $\varphi_B: A*B \longrightarrow T_B$, are the right-coset representative functions for H in $A * B$.

In view of our special choice of Kurosh system, each element $\varphi_A[g_j(ab)^{k_j}]$ must have the form $ta_\iota a^l$ for some $t \in S_A$, specially chosen $(t^{-1}Ht \cap A, \langle a\rangle)$-double coset representative a_ι in A, and integer l satisfying $0 \le l < m_\iota(t)$ if $m_\iota(t)$ is defined, otherwise arbitrary (*i.e.* if $ta_\iota\langle a\rangle a_\iota^{-1}t^{-1} \cap H = \{1\}$), and the same is true of $\varphi_A[g_j(ab)^{k_j}a]$ (with the same t and a_ι for the same k_j). It follows that for an expression of the first type in (13) to represent a non-trivial generator we must have $\varphi_A[g_j(ab)^{k_j}] = ta_\iota a^{m_\iota(t)-1}$ (for some t, a_ι as before), and $\varphi_A[g_j(ab)^{k_j}a] = ta_\iota$ (and then that generator will have the form $ta_\iota a^{m_\iota(t)}a_\iota^{-1}t^{-1}$). (Analogously, a non-trivial generator of the second type in (13) must have the corresponding form $sb_\iota b^{n_\iota(s)}b_\iota^{-1}s^{-1}$.) Thus a generator of the first type can arise twice in the course of applying the Kurosh rewriting process to the $g_j(ab)^{l_j}g_j^{-1}$, if and only if there exist $j, l \in J$ such that

$$(14)\qquad \varphi_A[g_j(ab)^{k_j}] = \varphi_A[g_l(ab)^{\hat{k}_l}] \quad (= ta_\iota a^{m_\iota(t)-1}),$$

for some appropriate k_j, $\hat{k}_l$. However then $Hg_j\langle ab\rangle = Hg_l\langle ab\rangle$, so that we must in fact have $j = l$; but then by (14) again, with $j = l$, we have

$$g_j(ab)^{|k_j-\hat{k}_j|}g_j^{-1} \in H,$$

and $k_j \ne \hat{k}_j$ would entail $0 < |k_j - \hat{k}_j| < l_j$, contradicting the minimality of l_j. Hence $k_j = \hat{k}_j$, and we see that in fact no such generator can arise twice during the rewriting of the totality of the $g_j(ab)^{l_j}g_j^{-1}$. (Clearly the same argument establishes that a generator of the form $sb_\iota b^{n_\iota(s)}b_\iota^{-1}s^{-1}$ can likewise arise at most once.)

It is on the face of it, however, conceivable that two non-trivial distinct generators

$$(15)\qquad ta_\iota a^{m_\iota(t)}a_\iota^{-1}t^{-1},\ ta_k a^{m_k(t)}a_k^{-1}t^{-1} \quad (\text{same } t)$$

might emerge as adjacent in the course of the rewriting process, in which case the rewritten expression would not be reduced in the Kurosh generators. We now show that this cannot happen. Suppose that (as above)

$$\varphi_A[g_j(ab)^{k_j}] = ta_\iota a^{m_\iota(t)-1},\ \varphi_A[g_j(ab)^{k_j}a] = ta_\iota,$$

and that the next non-trivial generator arising (to the right of this one) is of this form for the same t, so that for some $l > 0$, $0 < k_j + l < l_j$,

$$\varphi_A[g_j(ab)^{k_j+l}] = ta_k a^{m_k(t)-1}, \ \varphi_A[g_j(ab)^{k_j+l}a] = ta_k. \tag{16}$$

The triviality of the intervening expressions (13) would then mean that

$$\varphi_B[g_j(ab)^{k_j+l}] = \varphi_A[g_j(ab)^{k_j+l}] = \varphi_A[g_j(ab)^{k_j}a]b(ab)^{l-1} = ta_i b(ab)^{l-1},$$

which is incompatible with (16). (The same argument shows that, analogously, non-trivial generators from $sBs^{-1} \cap H$ cannot arise next to one another in the rewriting process.)

If a non-trivial generator of the third type in (13) (*i.e.* an element of Φ) were to arise more than once in rewriting the $g_j(ab)^{l_j}g_j^{-1}$, then, the "first half" of such a generator being uniquely determined, we should have

$$\varphi_A[g_j(ab)^{k_j}a] = \varphi_A[g_l(ab)^{\hat{k}_l}a],$$

for some $j, l \in J$ and appropriate k_j, $\hat{k}_l$. However this yields $g_j(ab)^{k_j-\hat{k}_l}g_l^{-1} \in H$, so that $j = l$, and then it follows as before that $k_j = \hat{k}_l$. A similar argument shows that a nontrivial generator of the fourth type in (13), an element of Φ^{-1}, likewise occurs at most once as an h-symbol in the totality of reduced rewritten expressions in (12). ∎

COROLLARY 3.2. *With $a \in A$, $b \in B$, $H \leq A * B$, (T_A, T_B) a special Kurosh system for $H, J \subseteq I$, etc., as above, suppose that the possibility (10) does not occur. There then exists a finite (non-empty) minimally non-coprimitive subset M of the $g_j(ab)^{\ell_j}g_j^{-1}$, and any such subset M, consisting of elements*

$$g_1(ab)^{\ell_1}g_1^{-1}, \ldots, g_k(ab)^{\ell_k}g_k^{-1}, \tag{17}$$

say, with reduced rewritten expressions in terms of the Kurosh generators of H determined by (T_A, T_B) given by (cf. (12))

$$g_\mu(ab)^{\ell_\mu}g_\mu^{-1} = h_{\mu 1} \cdots h_{\mu r_\mu}, \quad \mu = 1, \ldots, k, \tag{18}$$

must have the following properties.

(i) If a Kurosh generator $h_{\mu\nu}$ in (18) is of conjugate form, say (invoking Lemma 3.1)

$$h_{\mu\nu} = ta_i a^{m_i(t)} a_i^{-1} t^{-1},$$

where $t \in S_A$ and a_i, $i \in I_t^A$, is a specially chosen $(t^{-1}Ht \cap A, \langle a \rangle)$-double coset representative in A, etc., then every "cognate" generator

$$ta_j a^{m_j(t)} a_j^{-1} t^{-1}, \quad j \in J_t^A,$$

must also occur as an h-symbol in (18) (and then just once by Lemma 3.1), and furthermore (9) must hold for this $t \in S_A$. (The analogous assertion is valid for generators of the form $sb_j b^{n_j(s)} b_j^{-1} s^{-1}$.)

(ii) Each Kurosh generator $h_{\mu\nu} \in \Phi$ appearing in (18) occurs exactly twice, once to each of the exponents ± 1.

PROOF. We first establish (i) and (ii) for such a finite non-empty set M, and then give the (considerably longer) proof that such a set M exists.

Thus let M be as in the statement of the corollary. To see that (i) holds, suppose the contrary, and let $h_{\mu_1\nu_1}, \ldots, h_{\mu_r\nu_r}$ be all the Kurosh generators from $tAt^{-1} \cap H$ ($t \in S_A$ fixed) occuring in the right-hand sides of (18). Assuming, as we are, that the conclusion of (i) fails, then either because (8) holds or because $h_{\mu_1\nu_1}, \ldots, h_{\mu_r\nu_r}$ do not exhaust all of the $ta_j a^{m_j(t)} a_j^{-1} t^{-1}$, $j \in J_t^A$, the elements $h_{\mu_1\nu_1}, \ldots, h_{\mu_r\nu_r}$ must generate freely a free factor H_1 say, of H, which furthermore by the Kurosh subgroup theorem has a "free complement" H_2 (*i.e.* such that $H = H_1 * H_2$) containing all other $h_{\mu\nu}$ occurring in (18). It follows that if $g_1(ab)^{\ell_1} g_1^{-1}$ say, is the (unique) element on the left-hand side of (18) involving $h_{\mu_1\nu_1}$, then $\langle g_1(ab)^{\ell_1} g_1^{-1} \rangle$ has a free complement in H containing all $h_{\mu\nu} \neq h_{\mu_1\nu_1}$, and therefore $g_2(ab)^{\ell_2} g_2^{-1}, \ldots, g_k(ab)^{\ell_k} g_k^{-1}$. Hence the subset $M \setminus \{g_1(ab)^{\ell_1} g_1^{-1}\}$ must be non-coprimitive in H, contradicting the minimality of M.

For (ii), observe that if there were no $h_{\gamma\delta} = h_{\mu\nu}^{-1} \in \Phi^{-1}$, then, supposing without loss of generality that $g_1(ab)^{\ell_1} g_1^{-1}$ is the (unique) element in (18) involving $h_{\mu\nu}$, we could, as in the above proof of (i), decompose H as a free product $\langle g_1(ab)^{\ell_1} g_1^{-1} \rangle * H_3$ where H_3 contains all $h_{\gamma\delta} \neq h_{\mu\nu}$ in (18), thereby obtaining once again a contradiction of the minimality property of M.

It follows in much the same way that if for any $t \in S_A$ the set

$$X_t := \{ta_j a^{m_j(t)} a_j^{-1} t^{-1} \mid j \in J_t^A\} \tag{19}$$

is coprimitive, then the set of those $g_j(ab)^{\ell_j} g_j^{-1}$ involving generators from X_t, freely generates a subgroup H_1 of H freely complemented in H by the subgroup H_2 generated by all other $X_{\hat{t}}$ together with Φ, which clearly contains all $g_j(ab)^{\ell_j} g_j^{-1}$ not involving generators from X_t. Hence in view of our assumption that the totality of the $g_j(ab)^{\ell_j} g_j^{-1}$ are not coprimitive, these latter $g_j(ab)^{\ell_j} g_j^{-1}$ cannot be coprimitive, and we may restrict attention to them. Thus we may suppose without loss of generality that every X_t is a finite minimally non-coprimitive subset of $tAt^{-1} \cap H$. The existence of a (non-empty) finite, non-coprimitive subset of the $g_j(ab)^{\ell_j} g_j^{-1}$ (and therefore a minimal such subset), under the assumption of the corollary, is then immediate from the following

LEMMA 3.3. *Let $\{H_\lambda \mid \lambda \in \Lambda\}$ be a family of groups and for each λ let X_λ be a finite almost coprimitive set of generators of H_λ. Let Σ be a subset of the free product*

$$H := \prod_{\lambda \in \Lambda}^{*} H_\lambda,$$

with the property that there exist expressions for the elements of Σ as (semigroup) products of the generators from $\bigcup_{\lambda \in \Lambda} X_\lambda$, with each such generator appearing at most once in the totality of these expressions. Then either Σ is coprimitive in H, or Σ contains a (non-empty) finite non-coprimitive subset.

(In the context of Corollary 3.2, the role of the X_λ is played by the X_ι (assumed finite and almost coprimitive) together with the sets $\{h, h^{-1}\}$, $h \in \Phi$, also finite and almost coprimitive.)

PROOF OF LEMMA 3.3. (As an aid to understanding the argument the following two examples may be useful: Let x_n, $n \in \mathbb{Z}$, be free generators of a free group F; then the set $\{x_n^{-1}x_{n+1}^2 \mid n \in \mathbb{Z}\}$, though locally coprimitive in F is not coprimitive (and the assumptions of the lemma, with $X_n := \{x_n^{-1}, x_n^2\}$, do not hold for it), while the set $\{x_n x_{n+1}^{-1} \mid n \in \mathbb{Z}\}$ is coprimitive in F.)

Writing Σ_λ for the subset of Σ consisting of those elements having at least one syllable from X_λ, we define a graph Γ by taking as vertices the Σ_λ, and joining two distinct vertices Σ_{λ_1}, Σ_{λ_2}, by an edge precisely if $\Sigma_{\lambda_1} \cap \Sigma_{\lambda_2} \neq \emptyset$. (Note that since the Σ_λ are finite, each vertex has finite valency in Γ.) We shall show that the union of the vertices of any infinite connected component of Γ is a coprimitive subset of H, in fact of the free product of those H_λ such that Σ_λ belongs to the component. It will then follow that if the whole set Σ is not coprimitive in H, that there must be a finite connected component of Γ the union of whose vertices is a (finite) non-coprimitive subset of H.

Thus let $\mathcal{C}$ be an infinite connected component of Γ. Choose a subtree $\mathcal{T}$ of $\mathcal{C}$ inductively as follows: Choose a "root", or level 1, vertex Σ_1 say, arbitrarily. (We shall re-index the $\Sigma_\lambda \in \operatorname{Vert} \mathcal{T}$ with the natural numbers as we progressively choose them.) The level 2 vertices are chosen next in the following manner: let Σ_2 ($= \Sigma_2^{(2)}$—we shall occasionally use superscripts to indicate the level of a vertex in $\mathcal{T}$) be any vertex of $\mathcal{C}$ adjacent to Σ_1 and not contained in Σ_1 (*i.e.* $\Sigma_2 \not\subseteq \Sigma_1$), and Σ_3 ($= \Sigma_3^{(2)}$) any other vertex of $\mathcal{C}$ adjacent to Σ_1 and not contained in $\Sigma_1 \cup \Sigma_2$, and so on, until all vertices adjacent to Σ_1 have been used up. We shall then have as our level 2 vertices in $\mathcal{T}$, all joined to Σ_1,

$$\Sigma_2^{(2)}, \ldots, \Sigma_k^{(2)}$$

where $\Sigma_\iota^{(2)} \not\subseteq \Sigma_1 \cup \Sigma_2 \cup \cdots \cup \Sigma_{\iota-1}$, for each $i = 2, \ldots, k$. To obtain the level 3 vertices, first join to $\Sigma_2^{(2)}$ a vertex $\Sigma_{k+1}^{(3)}$ adjacent to it in $\mathcal{C}$ and not contained in $\bigcup_{\iota=1}^{k} \Sigma_\iota$, and continue as for the adjunction of the level 2 vertices to Σ_1. Repeat this procedure for $\Sigma_3^{(2)}, \ldots, \Sigma_k^{(2)}$, in order, to obtain all of the level 3 vertices, say

$$\Sigma_{k+1}^{(3)}, \ldots, \Sigma_{k+\ell}^{(3)},$$

where again $\Sigma_r^{(3)} \not\subseteq \bigcup_{\iota=1}^{r-1} \Sigma_\iota$ for each $r \leq k + \ell$. The tree $\mathcal{T}$ is constructed by continuing in this way (inductively) *ad infinitum*, by joining to each level i vertex (in order) certain of the adjacent vertices (in some order) to obtain the level $(i+1)$ vertices (observing the requirement that each vertex adjoined should not be contained in the union of its predecessors). Denote by H_ι and X_ι ($\subseteq H_\iota$) the free factor of H and specified subset corresponding to each vertex Σ_ι of $\mathcal{T}$ (*i.e.* re-index also the relevant H_ι and X_ι to accord with the re-indexing of the $\Sigma_\iota \in \operatorname{Vert} \mathcal{T}$). Note that the union of the vertices of $\mathcal{T}$ is the union of those of $\mathcal{C}$.

Consider now the subtree $\bar{\mathcal{T}}$ of $\mathcal{T}$ maximal with respect to the property of having no extremal vertices (with the exception of the root vertex Σ_1 if this is extremal); thus $\bar{\mathcal{T}}$

is obtained from $\mathcal{T}$ by shearing off all finite branches compatibly with leaving Σ_1 as a vertex of $\bar{\mathcal{T}}$. From each vertex $\Sigma_u^{(j)}$ of the "trunk" $\bar{\mathcal{T}}$, choose an element

$$\sigma_u = \sigma_u^{(j)} \in \Sigma_u^{(j)} \cap \Sigma_v^{(j+1)},$$

for some vertex $\Sigma_v^{(j+1)}$ adjacent to $\Sigma_u^{(j)}$ in $\bar{\mathcal{T}}$ and at the next higher level in $\bar{\mathcal{T}}$, such that

$$\sigma_u \notin \bigcup_{\iota<u} \Sigma_\iota.$$

(This is possible by virtue of the fact that $\Sigma_u \not\subseteq \bigcup_{\iota<u} \Sigma_\iota$.) Corresponding to each such σ_u, let x_u, $\hat{x}_v$ be elements of X_u, X_v respectively, occurring as syllables of σ_u. (Note that, by definition of Σ_u as consisting of all elements of Σ with one or more syllables from X_u, the element $\sigma_u \in \Sigma_u$ must have at least one of its syllables from X_u, and, similarly, since also $\sigma_u \in \Sigma_v^{(j+1)}$, a syllable from X_v. Note also that $\hat{x}_1$ is not defined, and that there may be other v for which $\hat{x}_v$ is not defined.

For each subscript $i = 1, 2, \ldots$, define a subset $\bar{X}_\iota$ of X_ι as follows:

(i) if Σ_ι is a vertex of $\bar{\mathcal{T}}$, take $\bar{X}_\iota$ to consist of all elements of $X_\iota \setminus \{\hat{x}_\iota, x_\iota\}$ occurring as syllables of vertices Σ_j with $j \leq i$;

(ii) if Σ_ι is a vertex of $\mathcal{T}$ but not of $\bar{\mathcal{T}}$, set $\bar{X}_\iota := X_\iota \cap \Sigma_\iota$.

Note that in either case $\bar{X}_\iota$ is a proper subset of X_ι (in case (i) since $x_\iota \notin \bar{X}_\iota$, and in (ii) since $\mathcal{T}$ is connected). Hence $\bar{X}_\iota$ is a coprimitive subset of H_ι. Write

$$\bar{X} := \bigcup_{\iota=1}^{\infty} \bar{X}_\iota.$$

For each i let $\hat{X}_\iota$ be a subset of H_ι satisfying

$$H_\iota = \langle X_\iota \setminus \{x_\iota\}\rangle * \langle \hat{X}_\iota \rangle,$$

in the case that $\Sigma_\iota \in \operatorname{Vert} \bar{\mathcal{T}}$, and $H_\iota = \langle \bar{X}_\iota \rangle * \langle \hat{X}_\iota \rangle$ otherwise.

Write U for the set of indices u such that $\Sigma_u \in \operatorname{Vert} \bar{\mathcal{T}}$. The set

$$X := \bigcup_{u \in U} (X_u \setminus \{x_u\}) \cup \bigcup_{\iota \notin U} \bar{X}_\iota \tag{20}$$

is a coprimitive subset of H, in view of the assumption that each X_ι is a finite minimally non-coprimitive subset of H_ι. Setting

$$\hat{X} := \bigcup_{\iota=1}^{\infty} \hat{X}_\iota,$$

we have in fact

$$\prod_{\iota=1}^{\infty}{}^{*} H_\iota = \langle X \rangle * \langle \hat{X} \rangle.$$

We shall now show how the set $X \cup \hat{X}$ can be transformed to a set containing

$$\bigcup_{\iota=1}^{\infty} \Sigma_\iota \cup \bar{X} \cup \hat{X},$$

by means of a Nielsen transformation (*i.e.* induced by a free automorphism of the free group on $X \cup \hat{X}$) leaving $\bar{X} \cup \hat{X}$ fixed elementwise. (It will then follow that $\bigcup_{i=1}^{\infty} \Sigma_i$ is coprimitive in $\prod_{i=1}^{\infty *} H_i$.) The Nielsen transformation in question is defined in two stages, as a product of two simpler such transformations. First consider the specially chosen elements σ_u from the $\Sigma_u \in \operatorname{Vert} \bar{\mathcal{T}}$. Recall that each such σ_u was chosen from $\Sigma_u \cap \Sigma_v$ where Σ_v is some vertex of $\bar{\mathcal{T}}$ one level higher than Σ_u, and that we singled out syllables $x_u, \hat{x}_v$ of σ_u, coming from X_u, X_v respectively. Hence σ_u has the form

$$\sigma_u = w_u \hat{x}_v z_u, \tag{21}$$

where $\hat{x}_v \in X_v \setminus \{x_v\}$ (since $\sigma_u \neq \sigma_v$), and the syllables of w_u and z_u other than x_u all lie in various $\bar{X}_i$ with $i \geq u$, since σ_u was chosen outside $\bigcup_{i<u} \Sigma_i$. As the first of our Nielsen transformations of $X \cup \hat{X}$, we pre- or post-multiply each $\hat{x}_v$ by the appropriate elements of these $\bar{X}_i$ and by

$$x_u \in H_u = \langle X_u \setminus \{x_u\} \rangle * \langle \hat{X}_u \rangle.$$

However since of course $\hat{x}_u \in H_u$ (if $\hat{x}_u$ is defined), this replacement of each $\hat{x}_v$ by the chosen σ_u involving it, must be carried out inductively, in the natural order of the indices v.

Thus to begin with, we obtain

$$\sigma_1 = w_1 \hat{x}_2 z_1,$$

(in the notation of (21)), by pre- or post-multiplying $\hat{x}_2$ by elements of various $\bar{X}_i$ with $i \geq 1$, and by

$$x_1 \in H_1 = \langle X_1 \setminus \{x_1\} \rangle * \langle \hat{X}_1 \rangle;$$

since $\hat{x}_1$ is not defined, this presents no difficulty. The inverse transformation simply pre- or post-multiplies σ_1 by the inverses of the appropriate elements of

$$(X_1 \setminus \{x_1\}) \cup \hat{X}_1$$

(all of which are left fixed), to yield back $\hat{x}_2$. Now suppose $v > 2$, and, inductively, that we have defined a Nielsen transformation φ of $X \cup \hat{X}$ replacing all $\hat{x}_{v_1}$ with $v_1 < v$ by the corresponding σ_{u_1}, and fixing all other elements of $X \cup \hat{X}$. Then in extending this Nielsen transformation to $\hat{x}_v$ as described above (see (21) *et seq.*), the only $\hat{x}_{v_1}$ involved (if any) are those with $v_1 < v$, and these are generated by $(X \cup \hat{X})\varphi \setminus \{\hat{x}_v\}$. It follows that the replacement of each $\hat{x}_v \in X \cup \hat{X}$ by the corresponding σ_u is a Nielsen transformation θ_1 say.

The second stage, involving the construction of a Nielsen transformation of $(X \cup \hat{X})\theta_1$, is, though simpler, also carried out inductively. Each element σ of $\Sigma_1 \setminus \{\sigma_1\}$ has the form

$$\sigma = \bar{w}_1 \bar{x}_1 \bar{z}_1,$$

where $\bar{x}_1 \in X_1 \setminus \{x_1\}$, and the syllables of $\bar{w}_1$ and $\bar{z}_1$ all come from $\bar{X}$. Thus to obtain each such σ from a corresponding $\bar{x}_1 \in X_1 \setminus \{x_1\}$, one simply pre- or post-multiplies $\bar{x}_1$

by various elements of $\bar{X}$. Suppose inductively that u (> 1) is such that $\Sigma_u \in \operatorname{Vert} \bar{\mathcal{T}}$, and that for all $u_1 < u$ (with $\Sigma_{u_1} \in \operatorname{Vert} \bar{\mathcal{T}}$) we have defined a Nielsen transformation of $(X \cup \hat{X})\theta_1$ replacing certain $\bar{x}_{u_1} \in X_{u_1} \setminus \{x_{u_1}\}$ (via pre- or post-multiplication by elements of $\bar{X}$) by the elements of

$$\bigcup_{u_1 < u} (\Sigma_{u_1} \setminus \{\sigma_{u_1}\}).$$

Each element σ in $\Sigma_u \setminus \{\sigma_u\}$ but outside $\bigcup_{u_1<u} \Sigma_{u_1}$, has the form

$$\sigma = \bar{w}_u \bar{x}_u \bar{z}_u,$$

where $\bar{x}_u \in X_u \setminus \{x_u\}$ has not yet been replaced, and the syllables of $\bar{w}_u$ and $\bar{z}_u$ all come from $\bar{X}$. Hence proceeding as before we may extend our current Nielsen transformation to include $\Sigma_u \setminus \{\sigma_u\}$ in its range. This completes the definition of the second Nielsen transformation θ_2. The product $\theta_1\theta_2$ is then a Nielsen transformation of $X \cup \hat{X}$ fixing $\bar{X} \cup \hat{X}$, whose image contains all $\Sigma_u \in \operatorname{Vert} \bar{\mathcal{T}}$.

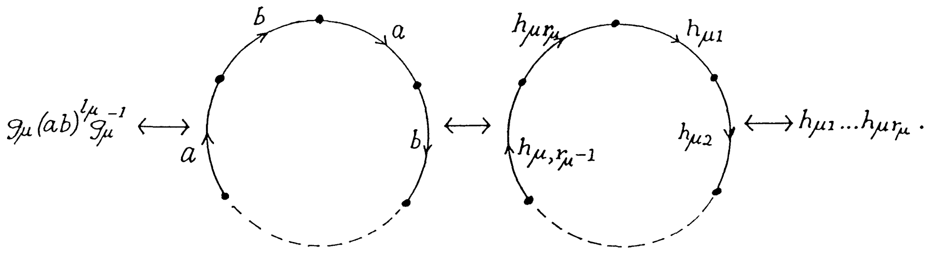
There remains the question of the $\Sigma_\iota \notin \operatorname{Vert} \bar{\mathcal{T}}$. Now it is not difficult to see (from the definition of $\mathcal{T}$) that if Σ_ι is a terminal vertex of the tree $\mathcal{T}$ (different from Σ_1), then the elements of $\Sigma_\iota \setminus \bar{X}_\iota$ all lie in earlier $\Sigma_\jmath$ in $\operatorname{Vert} \mathcal{T}$. Working backwards from each such Σ_ι towards the "trunk" $\bar{\mathcal{T}}$, one then sees that each $\Sigma_\jmath$ connected to Σ_ι in the complement $\mathcal{T} \setminus \bar{\mathcal{T}}$, has all of its elements outside $\bar{X}_\jmath$ contained in vertices of $\bar{\mathcal{T}}$, so that in fact

$$\bar{X} \cup \Big(\bigcup_{u \in U} \Sigma_u \Big) \supseteq \bigcup_{\iota=1}^{\infty} \Sigma_\iota.$$

Since the union of the vertices of $\mathcal{T}$ is the union of the vertices of $\mathcal{C}$, it follows that the latter set is coprimitive in the free product of those H_λ such that Σ_λ is a vertex of $\mathcal{C}$, as claimed. ∎

For the remainder of the proof of Theorem 1.6 we consider a subset M of the $g_\jmath(ab)^{l_\jmath} g_\jmath^{-1}$, $\jmath \in J$, made up of the elements (17) as in the statement of Corollary 3.2, with reduced rewritten expressions in terms of our special Kurosh generators for H as in (18), and having properties (i) and (ii) of that corollary (as consequences of denying the possibility (10)). We shall treat the words on either side of the equation (18) as "cyclic" words, *i.e.* we shall consider the final symbol of each such word as adjacent to, and preceding (in clockwise order—see (22) below) the initial symbol, carrying out any cancellations thereby made possible. Thus the cyclic words arising from the words $g_\mu(ab)^{l_\mu} g_\mu^{-1}$, $\mu = 1, \ldots, k$, in (18) will be the same as those arising from the words $(ab)^{l_\mu}$, $\mu = 1, \ldots, k$, but we preserve the correspondence between each such cyclic word in a and b and that arising similarly from its rewritten expression $h_{\mu 1} \cdots h_{\mu r_\mu}$ in the Kurosh generators. For each $\mu = 1, \ldots, k$, this may be indicated diagramatically as in (22):

REMARKS. 1. Note that in the cyclic word in the $h_{\mu\nu}$ depicted on the right side in (22) it has been assumed (for convenience only) that $h_{\mu 1} \neq h_{\mu r_\mu}^{-1}$, since in the contrary situation $h_{\mu 1}$ and $h_{\mu r_\mu}$ would have been cancelled, and any further cancellations effected until the cyclic word was reduced. However, whatever the additional cancellation incurred, the



(22)

reduced cyclic word arising from $h_{\mu 1}\cdots h_{\mu r_\mu}$ continues to have the property established in the course of proving Lemma 3.1, of having no two generators of the form (15) as adjacent h-symbols. For in the contrary case, application of the Kurosh rewriting process to $\left(g_\mu(ab)^{l_\mu}g_\mu^{-1}\right)^2 = g_\mu(ab)^{2l_\mu}g_\mu^{-1}$ would yield two such adjacent generators, whereas the argument beginning just prior to (15), which did not depend on the minimality of l_μ, applies to show that this is not possible.

2. Since each Kurosh generator of conjugate form (and so, by Lemma 3.1 (i), of the form $ta_ja^{m_j(t)}a_j^{-1}t^{-1}$ or $sb_jb^{n_j(s)}b_j^{-1}s^{-1}$) occurs at most once in the right-hand sides of the equations (18) (and if it does occur then so do all of its "t-cognates" by Corollary 3.2(i)), the above process of forming reduced cyclic words out of the $h_{\mu 1}\cdots h_{\mu r_\mu}$, $\mu = 1,\dots,k$, will not result in the disappearance of any of these generators (although conceivably some pairs of mutually inverse h-symbols of non-conjugate form may cancel).

In the rest of the proof we shall use the notation $h_{\mu,\nu-1}$, $h_{\mu,\nu+1}$ for the predecessor and successor respectively of an h-symbol $h_{\mu\nu}$ of a cyclic h-word in (22).

The following corollary of Lemma 3.1, like that lemma, does not require the full strength of our hypotheses.

COROLLARY 3.4. *With $a \in A$, $b \in B$, $H \le A*B$, (T_A, T_B), et cetera, as in Lemma 3.1, consolidation without cancellation of A-symbols, in reducing any cyclic h-word in (22) down to the corresponding cyclic word a and b, occurs in one of the following two ways:*

(i) at most two adjacent h-symbols, say $h_{\mu\nu}$ and $h_{\mu,\nu+1}$, are involved in the consolidation, and as elements of $A*B$ these have the reduced forms

$$h_{\mu\nu} = u(ta_ia^l)^{-1}, \quad h_{\mu,\nu+1} = (ta_ia^{l+1})v,$$

where $t \in S_A$, $i \in I_t^A$, and l satisfies $0 \le l < m_i(t) - 1$ if $m_i(t)$ is defined, and is otherwise arbitrary (see (7));

(ii) the central A-symbol of an h-symbol of conjugate form:

$$h_{\mu\nu} = ta_ia^{m_i(t)}a_i^{-1}t^{-1}, \tag{23}$$

is involved in the consolidation, and

$$h_{\mu,\nu-1} = u(ta_\iota a^{m_\iota(t)-1})^{-1}, \quad h_{\mu,\nu+1} = ta_\iota v,$$

(where u ends in an element of $B \setminus \{1\}$ *and v begins in an element of* $B \setminus \{1\}$*).*

(The analogues of (i) and (ii) for consolidation of B-symbols in the cyclic h-words in (22) are likewise valid.)

PROOF. (Recall from §2 the definition and non-cancellation property of the significant syllables of the Kurosh generators.) By Lemma 3.1, if a Kurosh generator of the form (23) occurs as an $h_{\mu\nu}$ in a cyclic h-word in (22), then it occurs just once. By the remark above, neither the preceding generator $h_{\mu,\nu-1}$ nor the succeeding generator $h_{\mu,\nu+1}$ can be from $tAt^{-1} \cap H$, whence (ii). Any consolidation of A-symbols not involving a generator of the form (23) is readily seen to be of the type indicated in (i). ∎

We are now in a position to prove the main lemma, from which Theorem 1.6 follows relatively easily. Relative to the given elements $a \in A$, $b \in B$, we choose, as above, a special Kurosh system (T_A, T_B) for $H \leq A*B$, and, assuming that the possibility (10) does not occur, we have a subset M as in (17) with properties 3.2(i), 3.2(ii), which, together with the rewritten expressions in the Kurosh generators (see (18)), are turned into cyclic words as indicated in (22). Write $P_A(\subseteq T_A)$ for the (finite) set of elements T_A that are initial segments of generators $h_{\mu\nu}$ occurring in the reduced cyclic h-words in (22), and define $P_B(\subseteq T_B)$ similarly.

LEMMA 3.5. *With notation and assumptions as above, the following statements are valid:*

(i) For each $t \in P_A \cap S_A$, $ta_\iota a^l$ *also belongs to* P_A *for every* $i \in I_t^A$ *and every admissible* l. *It follows that for every such t and for all* $i \in I_t^A$, *the positive integer* $m_\iota(t)$ *is defined (since* P_A *is of course finite), that is,* $I_t^A = J_t^A$. *For all such t, and all* $i \in I_t^A$, *the Kurosh generator* $ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$ *occurs (and then just once by Corollary 3.1(i)) as an h-symbol in the cyclic h-words in (22). (The analogous assertion is valid for* P_B*.)*

(ii) Corresponding to each element $ta_\iota a^l \in P_A$ *(as in (i)), with* $l \neq 0$, *there is an h-symbol* $h_{\mu\nu}$ *in some cyclic h-word in (22) with* $ta_\iota a^l$ *as initial segment, and an h-symbol* $h_{\gamma\delta}$ *with* $(ta_\iota a^l)^{-1}$ *as terminal segment, from neither of which these segments are wholly cancelled by the appropriate adjacent h-symbols (and analogously for* P_B*).*

PROOF. Using the representation (22) of the equations (18), we shall establish (i) and (ii) of the present lemma by reverse induction on the length of $t \in P_A \cup P_B$. (It may make for greater ease of understanding to visualize the elements of T_A as the vertices of a tree $\mathcal{T}_A$ where two vertices are joined by an edge if one is an initial segment of the other and their lengths (as elements of $A * B$) differ by 1; the tree $\mathcal{P}_A$ determined in the same way by P_A, is then a subtree of $\mathcal{T}_A$.)

As the first step of the induction, consider $t \in P_A \cap S_A$ maximal in P_A in the sense that it is not a proper initial segment of any other element of P_A (*i.e.* represents an extremal vertex, or vertex of valency 1, in the finite tree $\mathcal{P}_A$). Since t is either trivial or ends in an element of $B \setminus \{1\}$, we must have, by Definition 2.1, $t \in T_B$ also (so that $t \in T_A \cap T_B$). This, together with the maximality property of t, implies that any Kurosh generator $h_{\mu\nu}$ (in a cyclic h-word in (22)) having t as initial segment must have the form

$$h_{\mu\nu} = ta_1t^{-1}, \quad a_i \in A \setminus \{1\}. \tag{24}$$

Now the Kurosh generator $h_{\mu,\nu+1}$ immediately following $h_{\mu\nu}$ in its cyclic h-word, cannot have t as an initial segment, else for the same reasons as applied to $h_{\mu\nu}$ it would have to be of the form ta_2t^{-1}, $a_2 \in A \setminus \{1\}$, contradicting the assumption that the cyclic h-words in (22) are all reduced. Similarly, the preceding Kurosh generator $h_{\mu,\nu-1}$ cannot have t^{-1} as a terminal segment. Hence the syllable a_1 in (24) remains unconsolidated in reducing its cyclic h-word (the μ^{th}) down to the corresponding cyclic alternating word in a and b, whence we must have $a_1 = a$. Thus any $h_{\mu\nu}$ of the form $t\alpha t^{-1}$, $\alpha \in A \setminus \{1\}$, occurring in any cyclic h-word in (22) is actually tat^{-1}. It follows from Corollary 3.2(i) that the singleton $\{tat^{-1}\}$ is a minimally non-coprimitive set in $H \cap tAt^{-1}$, *i.e.* that tat^{-1} is not primitive in $H \cap tAt^{-1}$. Since tat^{-1} is almost primitive in tAt^{-1}, we must therefore have

$$tAt^{-1} \cap H = tAt^{-1},$$

whence it follows (via Definition 2.1(iii)) that t is actually maximal in T_A, *i.e.* an extremal vertex of the possibly larger tree $\mathcal{T}_A$, so that conditions (i) and (ii) of the lemma are satisfied vacuously for this t. (The analogous argument applies to maximal elements of P_B.)

Proceeding now to the inductive step, let $t \in P_A \cap S_A$ be an initial segment of some element $\hat{t} \in P_A$, whose length exceeds that of t by 1 (*i.e.* $|\hat{t}| = |t| + 1$, where lengths are taken with respect to the free product $A * B$), and assume inductively that the assertions (i), (ii) of the lemma are valid for all elements of $P_A \cap S_A$ and $P_B \cap S_B$ of length $> |t|$. Since $\hat{t}$ necessarily ends in an element of $A \setminus \{1\}$, we may write $\hat{t} = ta_1$ in reduced form, where $a_1 \in A \setminus \{1\}$.

Suppose first that $\hat{t}$ also belongs to T_B (and therefore to $P_B \cap S_B$). Then by the inductive hypothesis we have, for every $i \in I^B_{\hat{t}}$, that:

(i) $n_i(\hat{t})$ is defined, and $\hat{t}b_ib^{n_i(\hat{t})}b_i^{-1}\hat{t}^{-1}$ is an h-symbol in some cyclic h-word in (22); and

(ii) each of the elements

$$\hat{t}b_ib, \hat{t}b_ib^2, \ldots, \hat{t}b_ib^{n_i(\hat{t})-1}, \tag{25}$$

is an initial segment of an h-symbol in some cyclic h-word in (22), from which it is not all cancelled by the preceding h-symbol, and likewise its inverse is a terminal segment of some $h_{\mu\nu}$ not entirely cancelled by $h_{\mu,\nu+1}$.

These conditions are assumed to hold in particular for that i, say $i = l$, for which $b_l = 1$. If $n_l(\hat{t}) = 1$, then by (i) $\hat{t}b\hat{t}^{-1}$ occurs (just once by Lemma 3.1(i)) as an h-symbol, $h_{\mu\nu}$ say, in (22), and it is easy to see that its initial segment $\hat{t}$ cannot be wholly cancelled by $h_{\mu,\nu-1}$, nor all of its terminal segment by $h_{\mu,\nu+1}$. On the other hand if $n_l(\hat{t}) > 1$, then

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

by (ii) above (see (25)) with $i = l$ (so $b_\iota = 1$) there are h-symbols in (22) of reduced forms

$$(26) \qquad h_{\mu\nu} = \hat{t}bu, \; h_{\gamma\delta} = (\hat{t}b^{n_l(\hat{t})-1})^{-1}$$

(possibly one and the same h-symbol, or mutual inverses), such that the initial segment $\hat{t}b$ is not all cancelled from $h_{\mu\nu}$ (in (26)) by its predecessor $h_{\mu,\nu-1}$, nor the terminal segment $b^{1-n_l(\hat{t})}\hat{t}^{-1}$ of $h_{\gamma\delta}$ wholly cancelled by its successor $h_{\gamma,\delta+1}$. It follows (from the fact that the cyclic h-words in (22) reduce to cyclic words in the symbols a and b alone) that in fact not all of the initial segment $\hat{t}$ of $h_{\mu\nu}$ is cancelled by $h_{\mu,\nu-1}$, and that we must have

$$h_{\gamma,\delta+1} = \hat{t}b^{n_l(\hat{t})}\hat{t}^{-1},$$

with not all of the terminal segment $\hat{t}^{-1}$ of *this* generator $h_{\gamma,\delta+1}$ cancelled by $h_{\gamma,\delta+2}$.

We infer that in either case ($n_l(\hat{t}) \geq 1$) there are h-symbols $h_{\kappa\lambda}$, $h_{\eta\tau}$ such that $\hat{t} = ta_1$ is an initial segment of $h_{\kappa\lambda}$ and $\hat{t}^{-1} = a_1^{-1}t^{-1}$ a terminal segment of $h_{\eta\tau}$, where neither segment is wholly cancelled by the appropriate adjacent h-symbol (namely $h_{\kappa,\lambda-1}, h_{\eta,\tau+1}$ respectively). Write, in reduced form,

$$h_{\kappa\lambda} = ta_1u, \; h_{\eta r} = va_1^{-1}t^{\ 1}.$$

By construction of our special Kurosh system (T_A, T_B) for H in $A * B$, the ending a_1 of $\hat{t} = ta_1$ must have the form $a_1 = a_\iota a^l$, $i \in I_t^A$, where a_ι is from the prescribed set of representatives of double cosets $(t^{-1}Ht \cap A)\alpha\langle a\rangle$ in A (*i.e.* such that (8) or (9) holds), and l is an integer satisfying $0 \leq l < m_\iota(t)$ if $m_\iota(t)$ is defined (*i.e.* if $ta_\iota\langle a\rangle a_\iota^{-1}t^{-1} \cap H \neq \{1\}$) and otherwise is arbitrary. By Corollary 3.4 consolidation (if any) of the symbol a_1 $(= a_\iota a^l)$ appearing explicitly in $h_{\kappa\lambda}$ (see above) must occur in the context of one of the following three situations:

(27)
$$\begin{cases} l = 1, \text{ and } a_\iota = l, \text{ that is } a_1 = a \text{ (no consolidation occurs);} \\ l = 0, h_{\kappa,\lambda-1} = ta_\iota a^{m_\iota(t)}a_\iota^{-1}t^{-1}, \text{ and } h_{\kappa,\lambda-2} = u_1a^{-(m_\iota(t)-1)}a_\iota^{-1}t^{-1} \text{ (in reduced form);} \\ l \neq 0, \text{ and if } l = 1 \text{ then } a_\iota \neq 1, \text{ and } h_{\kappa,\lambda-1} = u_2a^{-(l\ 1)}a_\iota^{-1}t^{-1} \text{ (in reduced form),} \end{cases}$$

(where u_1 and u_2 do not end in elements of $A \setminus \{1\}$).

Similarly, consolidation (if any) of the A-symbol a_1^{-1} $(= a^{-l}a_\iota^{-1})$ explicitly appearing in $h_{\eta\tau}$ (see above) must occur in the context of one of the following three situations:

(28)
$$\begin{cases} a_1^{-1} = a \text{ (}i.e.\ l = -1, a_\iota = 1\text{), in which case no consolidation occurs;} \\ l = m_\iota(t) - 1, h_{\eta,\tau+1} = ta_\iota a^{m_\iota(t)}a_\iota^{-1}t^{-1}, h_{\eta,\tau+2} = ta_\iota v_1 \text{ (in reduced form);} \\ l \neq m_\iota(t) - 1, \text{ and if } l = -1 \text{ then } a_\iota \neq 1, \text{ and } h_{\eta,\tau+1} = ta_\iota a^{l+1}v_2 \text{ (in reduced form).} \end{cases}$$

Combining the various possibilities in (27) and (28), we infer that one of the following four situations occurs:

$\hat{t} = ta$ and ($h_{\eta,\tau+1} = ta^2t^{-1}$ or $h_{\eta,\tau+1} = ta^2v_2$, in reduced form, where $ta^2 \in P_A$);
$\hat{t} = ta^{-1}$ and ($h_{\kappa,\lambda-1} = u_2a^2t^{-1}$ in reduced form, with $ta^{-2} \in P_A$);
$m_\iota(t)$ is not defined, and both $ta_\iota a^{l\pm1} \in P_A$;

$m_\iota(t)$ is defined and both $ta_\iota a^{(l\pm 1) \bmod m_\iota(t)} \in P_A$.
(Moreover, in the last case, if $l = 0$, then $h_{\kappa,\lambda-1} = ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$, and if $l = m_\iota(t) - 1$, then $h_{\eta,\tau+1} = ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$.) From these we draw immediately the following conclusion (from our original assumption that $\hat{t} = ta_\iota a^l \in P_A \cap P_B$);

$$(29) \qquad ta_\iota a^{l\pm 1} \in P_A,$$

(where, if $m_\iota(t)$ is defined, the integers $l \pm 1$ are reduced modulo $m_\iota(t)$ to residues from among $0, 1, \ldots, m_\iota(t) - 1$). Furthermore from (27) and (28) and the argument preceding them, we see that, provided $l \neq 1$, there is an h-symbol in some cyclic h-word in (22) with $(ta_\iota a^{l-1})^{-1}$ as terminal segment with the property that this terminal segment is not wholly cancelled by the succeeding h-symbol, and, similarly, provided $l \neq -1$, there is an h-symbol in (22) with $ta_\iota a^{l+1}$ as initial segment not wholly cancelled by the h-symbol preceding it. (Here, as in (29), the exponents $l \pm 1$ are to be reduced modulo $m_\iota(t)$ if this is defined.)

So far in the inductive step, we have been assuming that $\hat{t} = ta_\iota a^l$ belongs to P_B as well as P_A. Thus we need to consider the case that $\hat{t} \notin T_B$. In this case $\hat{t}$ must be extremal in $\mathcal{T}_A$ (see Definition 2.1), and only way it can occur as an initial segment of an h-symbol, $h_{\mu\nu}$ say, in some cyclic h-word in (22) is if

$$(30) \qquad \Phi \ni h_{\mu\nu} = (ta_\iota a^l)\beta^{-1}\tau^{-1}, \quad \tau\beta \in T_B \setminus T_A,\ \beta \in B \setminus \{1\},$$

in reduced form. By Corollary 3.2(ii) this Kurosh generator appears exactly twice in the totality of cyclic h-words in (22), once to each of the exponents ± 1; there is thus an h-symbol in (22) of the form

$$(31) \qquad h_{\gamma\delta} = (h_{\mu\nu})^{-1} = \tau\beta(ta_\iota a^l)^{-1}.$$

Now since the A-symbol $a_\iota a^l$ of $h_{\mu\nu}$ (see (30)) does not cancel, and consolidates, if at all, only from the left, we can argue in the present situation exactly as before to deduce that one of the possibilities (27) occurs. Similarly, since the terminal segment $a^{-l}a_\iota^{-1}t^{-1}$ of $h_{\gamma\delta}$ (see (31)) does not cancel, we can argue as before to infer that one of the situations (28) must occur. Hence we deduce that (29) and the statement following it (concerning non-cancellation of the wholes of an initial segment $ta_\iota a^{l+1}$ ($l \neq -1$) and a terminal segment $(ta_\iota a^{l-1})^{-1}$ ($l \neq 1$) from some h-symbols in (22)), hold also in the case that $\hat{t} \notin T_B$.

Thus in either case ($\hat{t} \in T_B$ or $\hat{t} \notin T_B$), starting with $ta_\iota a^l$ we have brought to light elements $ta_\iota a^{l\pm 1}$ in P_A (where the integers $l \pm 1$ are to be reduced modulo $m_\iota(t)$ if this is defined) with the property that, provided $l \neq -1$, $(ta_\iota a^{l+1})^{-1}$ is a "non-cancelling" initial segment of an h-symbol in some cyclic h-word in (22), and, provided $l \neq 1$, $(ta_\iota a^{l-1})^{-1}$ is a "non-cancelling" terminal segment of some such h-symbol.

Since the inductive hypothesis applies also to the elements $ta_\iota a^{l\pm 1}$, we may repeat the forgoing argument with each of these elements in place of $\hat{t} = ta_\iota a^l$ to infer the presence in P_A of $ta_\iota a^{l\pm 2}$, with the property (taking for instance $ta_\iota a^{l-1}$) that provided $l - 1 \neq -1$, there is an $h_{\mu\nu}$ with $ta_\iota a^l$ as initial segment not wholly cancelled by $h_{\mu,\nu-1}$, and provided $l-1 \neq 1$, there is an $h_{\gamma\delta}$ with $(ta_\iota a^{l-2})^{-1}$ as terminal segment not wholly cancelled by $h_{\gamma,\delta+1}$. (If instead we take $ta_\iota a^{l+1}$ and apply to it the argument used for $\hat{t} = ta_\iota a^l$, we deduce, apart from the existence in P_A of $ta_\iota a^{l-2}$, the existence of an $h_{\mu\nu}$ with $ta_\iota a^{l+2}$ as initial segment not wholly cancelled by $h_{\mu,\nu-1}$ (provided $l+1 \neq -1$), and an $h_{\gamma\delta}$ with $(ta_\iota a^l)^{-1}$ as terminal segment not wholly cancelled by $h_{\gamma,\delta+1}$ (provided $l+1 \neq 1$).)

We now reiterate the argument, this time for $ta_\iota a^{l\pm 2}$, and so on. If $m_\iota(t)$ were not defined then this iteration could be continued indefinitely, contradicting the finiteness of P_A; hence $m_\iota(t)$ is defined, and the possibilities in (27) and (28) (and 29) involving $l < 0$ do not actually arise since $0 \leq l < m_\iota(t)$. The above iteration continues until all $ta_\iota a^k$, $0 \leq k < m_\iota(t)$, are encountered in P_A. In particular when we reach ta_ι $(k=0)$ or $ta_\iota a^{m_\iota(t)-1}$, then the next iteration will disclose $ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$ as an h-symbol in some cyclic h-word in (22). Hence by Corollary 3.2(i), condition (9) holds for this t, (so that in particular $m_\iota(t)$ is defined for all $i \in I_t^A$), and for every $i \in I_t^A$, $ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$ occurs as an $h_{\mu\nu}$ in a cyclic h-word in (22). The inductive step is completed by repeating the whole of the above argument for each $t_1 \in P_A \cap S_A$ (and each $t_2 \in P_B \cap S_B$) of the same length as t. ∎

COMPLETION OF THE PROOF OF THEOREM 1.6. We shall deduce from Lemma 3.5, just established, that the index of H in $A * B$ is $\sum_{\mu=1}^k l_\mu$, from which it is immediate first that $J = I$ (since the index is finite), and secondly that the minimally non-coprimitive subset M (see (17)) in fact contains all $g_\iota(ab)^{l_\iota} g_\iota^{-1}$, $i \in J = I$.

It follows from Lemma 3.5 (and the definition of our Kurosh system) that $P_A = T_A$. Hence by that lemma each element $ta_\iota a^l \in T_A$ $(t \in S_A,\ i \in I_t^A,\ 0 < l < m_\iota(t))$ is an initial segment of an $h_{\mu\nu}$ occurring in some cyclic h-word in (22) from which it is not wholly cancelled by $h_{\mu,\nu-1}$, and analogously for $(ta_\iota a^l)^{-1}$ (as terminal segment of some h-symbol). Now by Corollary 3.4, consolidation of the A-endings of such initial and terminal segments occurs only in the following ways:

$$(32) \qquad h_{\mu,\nu-1} = u_1(ta_\iota a^{l-1})^{-1}, \quad h_{\mu\nu} = ta_\iota a^l u_2,$$

in reduced form (*i.e.* u_1 ends, and u_2 begins, in an element of $B \setminus \{1\}$), and then $h_{\mu,\nu-1}h_{\mu\nu} = u_1 a u_2$ in reduced form, without further consolidation of the symbol a appearing explicitly here; or

$$(33) \qquad h_{\gamma,\delta-1} = v_1(ta_\iota a^{m_\iota(t)-1}),\ h_{\gamma\delta} = ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1},\ h_{\gamma,\delta+1} = ta_\iota v_2,$$

in reduced form, so that $h_{\gamma,\delta-1}h_{\gamma\delta}h_{\gamma,\delta+1} = v_1 a v_2$, without any further consolidation of the symbol a by $h_{\gamma,\delta-2}$, $h_{\gamma,\delta+2}$, *et cetera*.

Since each Kurosh generator $ta_\iota a^{m_\iota(t)} a_\iota^{-1} t^{-1}$ does occur as an $h_{\mu\nu}$ (by Lemma 3.5), it follows from (32) and (33) that these situations (*i.e.* (32) and (33)) must occur in some cyclic h-word in (22) (for each relevant l in the case of (32)). Hence with each

$$ta_\iota a^l \in T_A \quad (t \in S_A, i \in I_t^A, 0 < l \leq m_\iota(t)),$$

we can associate, in one-to-one fashion, a symbol a in some cyclic word in a and b arising from one of the words $g_\mu(ab)^{l_\mu} g_\mu^{-1}$ in (17) Since the number of such symbols altogether in these words is $\sum_{\mu=1}^{k} l_\mu$, we deduce that

$$\sum_{\mu=1}^{k} l_\mu = |T_A|,$$

which is just the index of H in $A * B$ ∎

4 **Proof of Theorem 1.4: The** APE $[x, y]$ **is tame.** Our proof that $[x, y]$ is a tame APE in $F(x, y)$ is intricate, and parallels the proof occupying the preceding section (using Schreier free generators for the relevant subgroup $H \leq F(x, y)$ instead of Kurosh generators)

As with the proof of Theorem 1 6 in the preceding section, we subdivide the proof of Theorem 1 4 into a sequence of lemmas We first choose a set $\mathcal{H}$ of "minimal" Schreier free generators for H (relative to the basis $\{x, y\}$ of F), thus each $h \in \mathcal{H}$ can be written in reduced form as $tx\tau^{-1}$ or $ty\tau^{-1}$, where t, τ belong to a minimal right Schreier transversal T say, for H in $F = F(x, y)$ (Here by "minimal" we mean that each $t \in T$ has least length among all elements of its coset Ht) It is a well-known property of such free generators that their expression in one or the other of these forms is unique, and also that their "significant symbols" (i e the explicitly appearing x in $h = tx\tau^{-1}$, or y in $ty\tau^{-1}$), remain uncancelled in every appearance of $h^{\pm 1}$ in any reduced word in the generators $\mathcal{H}$, that is, remain uncancelled in reducing such a word down to a reduced word in x and y

We choose the $(H, \langle [x, y] \rangle)$-double coset representatives a_ι, $\iota \in I$, from T, ensuring that each a_ι has least length amongst all elements of its double coset $Ha_\iota\langle [x, y] \rangle$

LEMMA 4 1 (*cf* LEMMA 3 1, COROLLARY 3 2) *Let $H \leq F(x, y)$, $\mathcal{H}, \{a_\iota \mid \iota \in I\}$ $J \subseteq I$, et cetera, be as above The following assertions are valid*

(i) In the totality of rewritten expressions for the $a_j[x, y]^{m_j} a_j^{-1}$, $j \in J$, as reduced words in the $h \in \mathcal{H}$, each h that actually appears, does so at most twice, and if it does appear twice, then once with the exponent +1 *and once with exponent* -1

(ii) If condition (i) of Theorem 1 4 does not hold, then there is a finite subset M of the $a_j[x, y]^{m_j} a_j^{-1}$, $j \in J$, for simplicity say

$$M = \{a_1[x, y]^{m_1} a_1^{-1}, \quad , a_k[x, y]^{m_k} a_k^{-1}\},$$

which is minimally non-coprimitive in H, and consequently by (i) has the property that each generator $h \in \mathcal{H}$ appearing in (the totality of) the reduced rewritten expressions

$$a_\mu[x, y]^{m_\iota} a_\mu^{-1} = h_{\mu 1} \quad h_{\mu r}, \; h_{\mu\iota} \in \mathcal{H} \cup \mathcal{H}^{-1}, \quad \mu = 1, \quad k, \tag{34}$$

appears exactly twice (possibly in different right-hand sides in (34)), once to each of the exponents ± 1

PROOF (i) The free generators (or inverses thereof) $h_{\mu\iota} \in \mathcal{H} \cup \mathcal{H}^{-1}$, occurring in the rewritten expression for $a_j[x, y]^{m_j} a_j^{-1}$ in terms of those generators (obtained by means

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

of the "Schreier rewriting process"—see [4]), are all of one or other of the following forms:

$$(35)\qquad \begin{gathered}\varphi(a_j[x,y]^{k_j})x^{-1}\big[\varphi(a_j[x,y]^{k_j}x^{-1})\big]^{-1};\\ \varphi(a_j[x,y]^{k_j}x^{-1})y^{-1}\big[\varphi(a_j[x,y]^{k_j}x^{-1}y^{-1})\big]^{-1};\\ \varphi(a_j[x,y]^{k_j}x^{-1}y^{-1})x\big[\varphi(a_j[x,y]^{k_j}x^{-1}y^{-1}x)\big]^{-1}\\ \varphi(a_j[x,y]^{k_j}x^{-1}y^{-1}x)y\big[\varphi(a_j[x,y]^{k_j+1})\big]^{-1},\end{gathered}$$

where $\varphi\colon F \longrightarrow T$ is the coset representative function for H, and $0 \le k_j < m_j$. Now by the uniqueness of the form of Schreier free generators of a subgroup (noted above), two free generators $t_1z_1\tau_1^{-1}$, $t_2z_2\tau_2^{-1} \in \mathcal{H}$ are equal if and only if $t_1 = t_2$, $z_1 = z_2$ $(= x$ or $y)$, and $\tau_1 = \tau_2$. Hence by (35), given two (possibly equal) double coset representatives a_j, a_l, $(j, l \in J)$, the same free generator h with significant symbol x (for instance) can appear twice altogether, both times with exponent +1, in the reduced rewritten expressions for $a_j[x,y]^{m_j}a_j^{-1}$ and $a_l[x,y]^{m_l}a_l^{-1}$ if and only if

$$\varphi(a_j[x,y]^{k_j}x^{-1}y^{-1}) = \varphi(a_l[x,y]^{\hat{k}_l}x^{-1}y^{-1}),$$

for some $0 \le k_j < m_j$, $0 \le \hat{k}_l < m_l$. However then

$$a_j[x,y]^{k_j-\hat{k}_l}a_l^{-1} \in H,$$

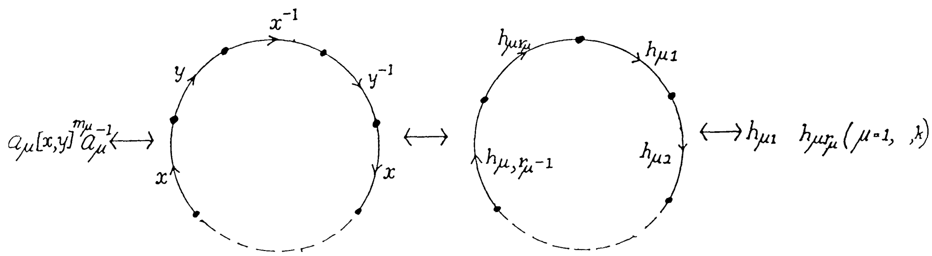
so that a_j, a_l lie in the same double coset $Hg\langle[x,y]\rangle$, whence $j = l$. However then if h actually occurs twice, we must have $k_j \ne \hat{k}_j$ $(= \hat{k}_l)$, and the fact that $0 < |k_j - \hat{k}_j| < m_j$ contradicts the minimality of m_j. One shows analogously that no $h^{-1} \in \mathcal{H}^{-1}$ can occur twice in the rewritten expressions for the $a_j[x,y]^{m_j}a_j^{-1}$, completing the proof of (i).

The proof of (ii) is analogous to that of Corollary 3.2, only somewhat simpler. We omit the details. ∎

For the remainder of this section we shall assume that the assertion (i) of Theorem 1.4 is invalid, and consider a fixed minimally non-coprimitive subset M, with elements denoted as in Lemma 4.1(ii) (and with reduced rewritten expressions as in (34)), which, by Lemma 4.1(ii), has the property that each $h \in \mathcal{H}$ appearing in those rewritten expressions, appears exactly twice, moreover once to each of the exponents ± 1. As in the preceding section we form, for convenience, reduced "cyclic" words from the words on either side of the equations (34), by regarding the terminal symbol of each word as adjacent to its initial symbol, depicting such words by labelling the edges of a corresponding cyclic graph, in clockwise order, as indicated in (36) below. Since consequent cancellations are to be carried out, each word $a_\mu[x,y]^{m_\mu}a_\mu^{-1}$ will yield the same cyclic word as $[x,y]^{m_\mu}$, and cancellations may conceivably occur also in forming the cyclic words in the $h_{\mu\nu}$. (In fact such cancellations do not occur, but we shall not need this fact.) However we preserve the correspondence between the cyclic word in x and y arising from each $a_\mu[x,y]^{m_\mu}a_\mu^{-1}$ in (34) and that in the $h \in \mathcal{H}$ arising from its rewritten expression

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

$h_{\mu 1} \quad h_{\mu r_\mu}$ in (34) (In the Figure (36) we have in fact assumed that $h_{\mu 1} \neq h_{\mu r}^{\;1}$, for notational convenience only) Thus each of the k reduced cyclic h-words in (36) yields the corresponding reduced cyclic word in x and y on replacing each $h_{\mu i}$ by its actual expression as a word in $F(x,y)$, and then reducing



(36)

We denote by S ($\subseteq T$) the (finite) set of elements of T that are initial segments of generators $h \in \mathcal{H}$ actually occurring in some reduced cyclic h-word in (36), *i e* of some $h_{\mu\nu}$ As in the previous section, it may be useful to visualize S as a tree $\mathcal{S}$ whose vertices are just the elements of S, and where $s_1, s_2 \in S$ are joined by an edge precisely if one is an initial segment of the other and their lengths differ by 1 We denote by $\mathcal{T}$ the tree determined similarly by T

We are now in a position to state and prove the main

LEMMA 4 2 (*cf* LEMMA 3 5) *Let $H \leq F(x,y)$, $\mathcal{H}$, M, S, et cetera, be as above The following assertions are valid*

(i) For each non-trivial $s \in S$, each of the three (four if $s = 1$) words in $F(x,y)$, of length $|s| + 1$ having s as initial segment, is again an initial segment of an $h_{\mu i}$ in some cyclic h-word in (36)

(ii) Each non-trivial element $s \in S \cup Sx^{\pm 1} \cup Sy^{\pm 1}$ is an initial segment of an $h_{\mu i}$ in some cyclic h-word in (36) from which it is not wholly cancelled by the preceding h-symbol $h_{\mu\, \nu - 1}$ (where if $\nu = 1$ we set $h_{\mu\, 0} = h_{\mu r}$) Similarly, the inverse of each such s is a terminal segment of an $h_{\gamma\delta}$ in some cyclic h-word in (36) from which it is not all cancelled by the succeeding h-symbol $h_{\gamma\, \delta+1}$ ($= h_{\gamma 0}$ if $\delta = r_\gamma$)

PROOF Much as in the proof of Lemma 3 5 we shall establish (i) and (ii) of the present lemma by means of reverse induction on the length of $s \in S$ As the first step of the induction, consider $s \in S$ maximal in S in the sense that it is not a proper initial segment of any other element of S (*i e* represents an end vertex, or vertex of valency 1, in the finite tree $\mathcal{S}$) By definition of S and by the property of M restated above, there must then exist $h_{\mu\nu}$, $h_{\gamma\delta}$ in the right-hand sides of (36) (possibly in different words) such that

$$h_{\mu\nu} = sx^{\varepsilon_{\mu\nu}}\sigma^{-1}, \quad \sigma \in S, \text{ (or } h_{\mu\nu} = sy^{\varepsilon_{\mu\nu}}\sigma^{-1}),$$

and

$$h_{\gamma\delta} = h_{\mu\nu}^{-1} = \sigma x^{-\varepsilon_{\mu\nu}} s^{-1}, \quad \text{(or } h_{\gamma\delta} = \sigma y^{-\varepsilon_{\mu\nu}} s^{-1}),$$

where $x^{\varepsilon_{\mu\nu}}$ (or $y^{\varepsilon_{\mu\nu}}$) is significant ($\varepsilon_{\mu\nu} = \pm 1$). We shall assume that $\varepsilon_{\mu\nu} = 1$, and also that the significant syllable of $h_{\mu\nu}$ is x rather then y, the other possibilities yielding to similar arguments; thus we assume that

$$h_{\mu\nu} = sx\sigma^{-1}, \ h_{\gamma\delta} = h_{\mu\nu}^{-1} = \sigma x^{-1}s^{-1}.$$

Now if the initial segment s of $h_{\mu\nu}$ does not cancel entirely into its predecessor $h_{\mu,\nu-1}$ in the μ^{th} cyclic h-word in (36), then, since x is always immediately preceded by y^{-1}, the word s must end in y^{-1}. (Note that, as in the statement of the lemma, the predecessor of $h_{\mu 1}$ in its cyclic word, is of course $h_{\mu r_\mu}$, assuming these do not cancel in their cyclic h-word.) On the other hand, if the terminal segment s^{-1} of $h_{\gamma\delta}$ does not cancel entirely into its successor $h_{\gamma,\delta+1}$ in the γ^{th} cyclic word on the right in (36), then for a similar reason, the word s^{-1} must begin in y^{-1}, *i.e.* s must end in y. (Again, note that, as in the statement of the lemma, the successor of the symbol $h_{\gamma r_\gamma}$ in the γ^{th} cyclic word on the right in (36) is, of course, just $h_{\gamma 1}$.) Hence at least one of the segments s of $h_{\mu\nu} = sx\sigma^{-1}$ and s^{-1} of $h_{\gamma\delta} = \sigma x^{-1}s^{-1}$, must cancel entirely into the appropriate adjacent h-symbol ($= h_{\mu,\nu-1}$ or $h_{\gamma,\delta+1}$ as the case may be) in the cyclic word in the $h \in \mathcal{H}$ in which it occurs. Suppose for instance that the segment s of $h_{\mu\nu} = sx\sigma^{-1}$ so cancels; by the maximality condition on s, and since y^{-1} always precedes x in the cyclic words on the left in (36), the element $h_{\mu,\nu-1}$ must then have the form

$$h_{\mu,\nu-1} = \sigma_1 y^{-1}s^{-1}, \quad \sigma_1 \in S.$$

By the basic property of the set M there then exists an h-symbol $h_{\kappa\lambda}$ in one of our cyclic words in the $h \in \mathcal{H}$, which is the inverse of $h_{\mu,\nu-1}$:

$$h_{\kappa\lambda} = sy\sigma_1^{-1}.$$

It is still open as to whether the terminal segment s^{-1} of $h_{\gamma\delta} = \sigma x^{-1}s^{-1}$ cancels completely into $h_{\gamma,\delta+1}$. If it does not, then, as already noted, s must end in y, whence we infer that the initial segment s of $h_{\kappa\lambda}$ cancels completely into $h_{\kappa,\lambda-1}$. Thus in this case we have

$$h_{\kappa,\lambda-1} = \sigma_2 xs^{-1}, \quad \sigma_2 \in S,$$

and we have found h-symbols in various of the cyclic h-words in (36), of the form

$$sx\sigma^{-1}, \ sy\sigma_1^{-1}, \ sx^{-1}\sigma_2^{-1}, \tag{37}$$

so that sx, sy, sx^{-1} are initial segments of various such h-symbols, as we wished to show. Furthermore since the syllables x, y, x^{-1} explicitly appearing in these h-symbols are all significant, none of the initial segments sx, sy, sx^{-1}, nor the corresponding terminal segments of the inverses of the h-symbols in (37), wherever they occur in the cyclic h-words in (36), cancels completely with adjacent h-symbols, so that condition (ii) of the lemma is also satisfied by this s.

We have yet to consider the case that the terminal segment s^{-1} of $h_{\gamma\delta}$ *does* cancel completely into $h_{\gamma,\delta+1}$. In this case we must have (in reduced form)

$$h_{\gamma,\delta+1} = sy^{-1}\sigma_3^{-1}, \quad \sigma_3 \in S,$$

and we have found h-symbols appearing in cyclic h-words in (36) of the Schreier forms

$$sx\sigma^{-1},\ sy\sigma^{-1},\ sy^{-1}\sigma_3^{-1},$$

so that condition (i) of the lemma is satisfied. That condition (ii) is satisfied follows as before.

Turning now to the inductive step, suppose that s is an initial segment of some element $s_1 \in S$, where $|s_1| = |s| + 1$, and that both statements (i) and (ii) of the lemma are valid for all elements of S having s as a proper initial segment. We shall suppose that $s_1 = sy$ in reduced form, the argument in the other three cases being similar. Thus by virtue of the inductive hypothesis, the (reduced) words $s_1x = syx$, $s_1x^{-1} = syx^{-1}$, and $s_1y = sy^2$ are initial segments of h-symbols appearing in various of the cyclic h-words in (36), and in addition, corresponding to each of these three words $\hat{s}$ ($= s_1x, s_1x^{-1}$ or s_1y) there are factors $h_{\mu\nu}$, $h_{\gamma\delta}$ in some such (cyclic) words, such that $\hat{s}$ is an initial segment of $h_{\mu\nu}$, $\hat{s}^{-1}$ a terminal segment of $h_{\gamma\delta}$, neither segment cancelling into the appropriate adjacent h-symbol (*i.e.* into $h_{\mu,\nu-1}$ in the case of the initial segment $\hat{s}$ of $h_{\mu\nu}$, and into $h_{\gamma,\delta+1}$ in the case of the terminal segment $\hat{s}^{-1}$ of $h_{\gamma\delta}$).

To begin the argument for the inductive step, let $h_{\mu\nu}$ be an h-symbol in one of our (cyclic) words, of, for instance, the form

$$h_{\mu\nu} = s_1x^{-1}u = syx^{-1}u,$$

where not all of the initial segment s_1x^{-1} cancels into $h_{\mu,\nu-1}$. Since s_1 ends in y and x^{-1} is in fact always preceded by y in the cyclic words in x, y that we are considering (see the left-hand side of (36)), it follows that at most the initial segment s of $h_{\mu\nu}$ cancels into $h_{\mu,\nu-1}$. If s is not wholly cancelled by $h_{\mu,\nu-1}$, then (since y is always immediately preceded by x in our cyclic words) the word s must end in x.

Again by the inductive assumption there is a factor $h_{\gamma\delta}$ of one of our cyclic h-words, with the reduced form

$$h_{\gamma\delta} = v(s_1x)^{-1} = vx^{-1}s_1^{-1} = vx^{-1}y^{-1}s^{-1},$$

such that $x^{-1}s_1^{-1} = x^{-1}y^{-1}s^{-1}$ does not cancel entirely into $h_{\gamma,\delta+1}$. Since s_1^{-1} begins in y^{-1}, which symbol always immediately follows x^{-1} in the cyclic words in x and y in (36), not more than the segment s^{-1} can in fact cancel from the end of $h_{\gamma\delta}$. If s^{-1} is not wholly cancelled then since y^{-1} is always immediately followed by x in the aforesaid cyclic words, we conclude that in this case s^{-1} begins in x (*i.e.* s ends in x^{-1}). Since this contradicts the conclusion reached in the preceding paragraph, we infer that in fact either precisely the segment s is cancelled from $h_{\mu\nu} = syx^{-1}u$, or precisely s^{-1} from $h_{\gamma\delta} = vx^{-1}y^{-1}s^{-1}$, or both. In the former case we must have

$$h_{\mu,\nu-1} = u_1xs^{-1},$$

and in the latter,

$$h_{\gamma,\delta+1} = sxv_1,$$

in reduced form.

Summarizing the situation thus far, we have that:

(38) sy is an initial segment of $h_{\mu\nu}$, not all of which cancels into $h_{\mu,\nu-1}$;

(39) $y^{-1}s^{-1}$ is a terminal segment of $h_{\gamma\delta}$, not all of which cancels into $h_{\gamma,\delta+1}$;

and at least one of the following occurs:

$$(40)\qquad \begin{cases} xs^{-1} & \text{is reduced as written, and is a terminal segment of } h_{\mu,\nu-1} \\ & \text{not entirely cancelling into } h_{\mu\nu}; \\ sx & \text{is reduced as written, and is an initial segment of } h_{\gamma,\delta+1} \\ & \text{not entirely cancelling into } h_{\gamma\delta} \end{cases}$$

Suppose that the first of the possibilities (40) occurs. In view of the basic property of our set of cyclic words there is then a factor $h_{\kappa\lambda} = h_{\mu,\nu-1}^{-1} = sx^{-1}u_1^{-1}$ (in reduced form) occurring in one of those cyclic words, so that either $sx^{-1} \in S$ or the syllable x^{-1} appearing explicitly here is significant. In the former case, *i.e.* if $sx^{-1} \in S$, then the inductive hypothesis applies to $sx^{-1} = s_2$ say, and the above argument in terms of $sy = s_1 \in S$ adapts directly to yield the following analogues of (38), (39):

(41) sx^{-1} is an initial segment of some $h_{\eta\tau}$,
with the property that not all of sx^{-1} cancels into $h_{\eta,\tau-1}$;

(42) xs^{-1} is a terminal segment of some $h_{\alpha\beta}$,
with the property that not all of xs^{-1} cancels into $h_{\alpha,\beta+1}$.

In the other case, *i.e.* if x^{-1} is significant in $h_{\kappa\lambda} = sx^{-1}u_1^{-1}$, (and so also in $h_{\mu,\nu-1} = u_1xs^{-1}$), these two assertions are clear.

If the second of the assertions (40) also occurs then we obtain in a similar way the analogues of (38) and (39) for sx and $x^{-1}s^{-1}$. Hence if both of the possibilities (40) occur, then provided $s \neq 1$, we have established that s has the desired properties (and, incidentally, that s ends in y).

If in this situation (*i.e.* where both of the possibilities (40) occur) we have $s = 1$, then it remains to show that $sy^{-1} = y^{-1}$ and $(sy^{-1})^{-1} = y$ are respectively initial and terminal syllables of h-symbols in our cyclic h-words, that are not cancelled by the appropriate

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

adjacent factors Now by the analogue of (39) with y^{-1} replaced by x^{-1} (and with $s = 1$), which we are assuming to hold, we know that there is a factor $h_{\gamma\delta}$ of a cyclic h-word in (36) with $x^{-1}s^{-1} = x^{-1}$ as terminal syllable, not cancelled by $h_{\gamma,\delta+1}$ Since x^{-1} is always followed immediately by y^{-1} in our cyclic words in x and y, it follows that $h_{\gamma,\delta+1}$ must begin with y^{-1} Similarly by (41), which we are assuming to hold, we know that there is a factor $h_{\eta\tau}$ of a cyclic h-word in (36) with $sx^{-1} = x^{-1}$ as initial segment, not cancelled by $h_{\eta,\tau-1}$ Since x^{-1} is always preceded by y in the cyclic words in x and y in (36), it must therefore be the case that $h_{\eta,\tau-1}$ ends in y

We have yet to consider the case that just one of the possibilities (40) holds, say the first As noted earlier, the assumption that the second statement in (40) is not valid, implies that s must end in x^{-1}, say $s = \hat{s}x^{-1}$ in reduced form Since the first of the statements in (40) is valid, so also is (41) (and (42)), *i e* $sx^{-1} = \hat{s}x^{-2}$ is an initial segment of some factor $h_{\kappa\lambda}$, not all of which cancels into its predecessor $h_{\kappa,\lambda-1}$ Hence exactly s must cancel into $h_{\kappa,\lambda-1}$, since otherwise we should have a segment of one of our reduced cyclic words in x and y (namely the κ^{th}) of the form x^{-2} Thus $h_{\kappa,\lambda-1}$ has, for the usual sort of reason, the reduced form

$$h_{\kappa,\lambda-1} = wys^{-1},$$

where the terminal segment ys^{-1} does not wholly cancel into $h_{\kappa\lambda}$, *i e* the syllable y remains uncancelled By our basic assumption concerning the set M, or, equivalently, the set of cyclic h-words in (36), there is in some such cycle a syllable

$$h_{\alpha\beta} = h_{\kappa,\lambda-1}^{-1} = sy^{-1}w^{-1}$$

If the syllable y^{-1} is significant here, then we have found an h-symbol in some cyclic h-word, namely $h_{\alpha\beta}$, with non-cancelling initial segment sy^{-1} Otherwise $sy^{-1} \in S$, and the inductive hypothesis applies to $sy^{-1} = s_3$, to yield, via an argument similar to the earlier one for $s_1 = sy$, the existence of an h-symbol of some cyclic h-word, with non-cancelling initial segment sy^{-1}

The remaining case, namely that where just the second of the two possibilities in (40) occurs, being similar, the proof is concluded ∎

COMPLETION OF THE PROOF OF THEOREM 1 4 (Recall that we are assuming that condition 1 4(i) does not hold) Since $1 \in S$ and by Lemma 4 2(i) every vertex of $\mathcal{S}$ has the same valency in $\mathcal{S}$ as it does in $\mathcal{T}$, it follows that $\mathcal{S} = \mathcal{T}$, *i e* $S = T$ It is then also clear from Lemma 4 2(i) that every free generator h of H occurs in the rewritten expression of some $a_\mu[x,y]^m a_\mu^{-1} \in M$ (see (34)), moreover exactly twice (once to each of the exponents ± 1) In view of the facts that the significant symbols of the $h_{\mu i}$ in the reduced rewritten expression for the $a_\mu[x,y]^m a_\mu^{-1}$ (see (34)) do not cancel and that the

https://doi.org/10.4153/CJM-1993-011-9 Published online by Cambridge University Press

total of the lengths of all the corresponding cyclic words in x, y (see (36)) is $4\sum_{\mu=1}^{k} m_\mu$, it follows that

$$\operatorname{Rank} H \leq 2\sum_{\mu=1}^{k} m_\mu$$

However by assertion (ii) of Lemma 4 2 each non-trivial element s of S is an initial segment of some $h_{\mu\nu}$ in a cyclic h-word, from which it is not all cancelled by the predecessor $h_{\mu,\nu-1}$ and similarly for s^{-1} as terminal segment of some $h_{\gamma\delta}$ Hence there is a total of at least $2(|S|-1)$ non-significant syllables of various $h_{\alpha\beta}$ that remain uncancelled in reducing the cyclic h-words to cyclic words in x and y, and therefore there remain at most

$$4\sum_{\mu=1}^{k} m_\mu - 2(|S|-1)$$

syllables of those cyclic words in x and y, that are candidates for the status of significance Hence in fact

$$\operatorname{Rank} H \leq 2\sum m_\mu - (|S|-1) \tag{43}$$

Now by the Schreier rank formula,

$$\operatorname{Rank} H = |T| + 1 = |S| + 1,$$

and this and (43) together give

$$|S| + 1 \leq 2\sum m_\mu - |S| + 1,$$

or

$$|S| \leq \sum_{\mu=1}^{k} m_\mu$$

On the other hand since the elements $a_j[x,y]^{t_j}a_j^{-1}$, $j \in J$, $0 \leq t_j \leq m_j - 1$, lie in distinct cosets of H in $F(x,y)$, the index $|S|$ of H in $F(x,y)$ is at least $\sum_{j\in J} m_j$ Hence $|J| = k$, and H has (finite) index

$$|F : H| = \sum_{j\in J} m_j,$$

as claimed ∎

REFERENCES

1. A H M Hoare A Karrass and D Solitar *Subgroups of infinite index in Fuchsian groups* Math Z **125**(1972) 59–69
2. ______ *Subgroups of* NEC *groups* Comm Pure Appl Math **26**(1973) 731–744
3. R C Lyndon and P E Schupp *Combinatorial Group Theory* Springer Verlag Berlin Heidelberg New York 1977
4. W Magnus A Karass and D Solitar *Combinatorial Group Theory* Interscience New York 1966
5. B H Neumann *On the number of generators of a free product* J London Math Soc **18**(1943) 12–20

6. G Rosenberger, *A property of subgroups of free groups*, Bull Austral Math Soc **43**(1991)

7. ———, *Mınımal generatıng systems for plane dıscontınuous groups and an equatıon ın free groups*, Groups-Korea 1988, Lecture Notes ın Mathematıcs **1398**, Sprınger-Verlag, Berlın-Heıdelberg, 1989

Department of Mathematıcs
Unıversıty of Wısconsın Parksıde
Kenosha, Wısconsın
U S A 53141

Department of Mathematıcs
York Unıversıty
North York, Ontarıo
M3J 1P3

Department of Mathematıcs
The Unıversıty of Queensland
Queensland
Australıa 4072