COSET REPRESENTATIONS IN FREE GROUPS[1]

BY

MARSHALL HALL, JR.

1. **Introduction.** Schreier [5][2] proved that every subgroup U of a free group F is free by constructing generators of U from the generators of F and the coset representatives of U in F. In a recent paper *On Schreier systems in free groups* [1] written jointly by T. Radó and the author, subgroups of free groups were completely characterized by left coset representatives and a certain function. This paper will be referred to as SS.

In the present paper the study of a subgroup U of a free group F is further advanced, primarily in terms of the representation $U = U[\{G\}, \phi(H)]$ established in SS. §2 shows the relation of the decision problem to this representation. In §3 a canonical representation, called the alphabetical representation, is defined. This is based on an alphabetical ordering of the elements of F. A more complicated ordering, the semi-alphabetical ordering, is defined in §4, and this provides a link between the coset representations of U and a canonical system of free generators for U. This generalizes the Nielsen construction [4]. As an application of techniques available it is shown in §5 that in a free group F there is a subgroup of finite index which contains given elements $\alpha_1, \alpha_2, \cdots, \alpha_m$ but none of given elements $\beta_1, \cdots, \beta_n$ under the obviously necessary condition that no β is directly expressible by the α's. The special case in which $n=1$, $m=0$ was first stated by von Neumann and Wigner [6] and later proved by Iwasawa [2]. In §6 properties are given which distinguish the representation of a normal subgroup U from a non-normal subgroup.

Except in §5 the assumption is made that the free group F is finitely generated. If this assumption is dropped similar results may be obtained but at the expense of more complicated statement of the theorems and straightforward but more complicated proofs. Enumerations must be replaced by well orderings, and inductions must be replaced by transfinite inductions. If this more general treatment had been used throughout it would still have been necessary to give the more precise formulations for finitely generated F.

2. **The standard representation and the decision problem.** Let us take a subgroup U of a free group F and the decomposition of F into left cosets of U

$$F = U\cdot 1 + Ug_2 + \cdots + Ug_i + \cdots . \tag{2.1}$$

Presented to the Society, April 16, 1948; received by the editors July 16, 1948, and, in revised form, January 7, 1949.

[1] This paper is part of a project for the Office of Naval Research under the administration of the Ohio State University Research Foundation.

[2] Numbers in brackets refer to the bibliography at the end of the paper.

Using G as a generic term for the representatives $1, g_2, \cdots, g_i, \cdots$, Schreier [5] showed that the G's may be chosen so that if in reduced form $a_1a_2 \cdots a_{t-1}a_t$ is a G, each a_i being $s_j^{\epsilon_i}$, $\epsilon_i = \pm 1$, s_j a generator of F, then $a_1a_2 \cdots a_{t-1}$ is also a G. A set $\{G\}$ with this property is called a Schreier system. Let S be a generic term for a generator of F. In the paper SS it was shown that the Schreier system $\{G\}$ and a function $\phi(H)$ defined for arguments $H = GS^\epsilon$, $\epsilon = \pm 1$, determines U if $\phi(H)$ satisfies

$$(2.1.1) \qquad \phi(H) \text{ is a } G,$$

$$(2.1.2) \qquad \phi(H) = G \quad \text{if} \quad H = G,$$

$$(2.1.3) \qquad \phi[\phi(GS^\epsilon)S^{-\epsilon}] = G.$$

Thus $\{G\}$ and $\phi(H)$ give a means of describing U which we shall call a standard representation $U = U[\{G\}, \phi(H)]$.

In §5 of SS a function $\Phi(f)$ was defined for an arbitrary $f \in F$ so that if $Uf = UG$, then $\Phi(f) = G$. Here if

$$(2.2) \qquad f = a_1a_2 \cdots a_{t-1}a_t,$$

put

$$f_0 = 1, f_1 = a_1, f_2 = a_1a_2, \cdots, f_i = a_1a_2 \cdots a_i, \cdots, f_t = f.$$

Then

$$(2.3) \qquad \begin{aligned} \Phi(f_0) &= 1, \\ \Phi(f_1) &= \phi(1 \cdot a_1) = G_1, \\ &\cdots\cdots\cdots\cdots, \\ \Phi(f_i) &= \phi(G_{i-1}a_i) = G_i, \\ &\cdots\cdots\cdots\cdots, \\ \Phi(f) &= \phi(G_{t-1}a_t) = G_t = G. \end{aligned}$$

Thus (2.3) gives a constructive method of finding $\Phi(f)$ for an arbitrary f, given a standard representation of U. In particular $f \in U$ if and only if $\Phi(f) = 1$. Hence a standard representation of U yields a solution of the decision problem for U, that is, gives a finite constructive method for deciding whether or not a given element f belongs to U.

3. **The alphabetical representation.** Consider a free group F_r with r generators, which we may number $s_1, s_2, \cdots, s_r$. We wish to establish a simple ordering of all the elements of F_r. We begin by putting

$$(3.1) \qquad s_1 < s_1^{-1} < s_2 < \cdots < s_i < s_i^{-1} < s_{i+1} < \cdots < s_r < s_r^{-1}.$$

Any other simple ordering of the s's and their inverses could be used. Then we order an arbitrary $f \in F_r$ written in reduced form

$$f = a_1a_2 \cdots a_t$$

respectively on (1) $l(f)=t$, (2) a_1, (3) $a_2, \cdots, (t+1)a_t$. It is readily verified that this is a simple ordering, which we shall call the alphabetical ordering. In fact it is a well ordering since an arbitrary set of elements $\{f\}$ contains a first in this ordering. The first of a set $\{f\}$ we shall call the earliest of this set. A useful property of this ordering is the following: If $x \leqq y$, then $xz \leqq yz$ if there is no cancellation between y and z, and similarly $zx \leqq zy$ if there is no cancellation between z and y.

THEOREM 3.1. *Given a free group F_r, and a subgroup U. If the representative G of a left coset UG is the earliest element of the coset, then the set $\{G\}$ is a Schreier system. If U is a normal subgroup, then the set $\{G\}$ has the stronger property: If $a_1a_2 \cdots a_{t-1}a_t$ is a G, then both $a_1a_2 \cdots a_{t-1}$ and $a_2 \cdots a_{t-1}a_t$ are G's.*

Proof. Let $g_i=a_1 \cdots a_t$ be the earliest element in its coset Ug_i. Then if $x=a_1 \cdots a_{t-1}$, let $\Phi(x)=g_j$. Since g_j is the earliest element in the coset Ug_j, $g_j \leqq x$ in the alphabetical ordering. Hence $g_ja_t \leqq xa_t=g_i$, noting that xa_t is in reduced form. But $g_ja_t \in Ug_i$, whence $g_i \leqq g_ja_t$. Hence $g_ja_t=xa_t=g_i$, whence $x=g_j$ and $x=a_1 \cdots a_{t-1}$ is a G. Now if U is a normal subgroup, let $y=a_2 \cdots a_t$, and $\Phi(y)=g_k$. Hence $g_k \leqq y$ and, as $g_i=a_1y$ is in reduced form, $a_1g_k \leqq a_1y=g_i$. But, since U is normal,

$$Ua_1g_k = a_1Ug_k = a_1Uy = Ua_1y = Ug_i,$$

whence $g_i=\Phi(a_1g_k)$ and $g_i \leqq a_1g_k$. Hence $g_i=a_1g_k=a_1y$ and $y=a_2 \cdots a_t=g_k$ is a G.

Thus, once having chosen an order as in (3.1) for the generators of F_r and their inverses, the alphabetical ordering of F_r determines a unique Schreier system for a subgroup U, and hence a unique representation $U=U[\{G\}, \phi(H)]$ where each G is the earliest element in its coset UG. Such a representation will be called the alphabetical representation for U.

THEOREM 3.2. *A standard representation $U=U[\{G\}, \phi(H)]$ is the alphabetical representation if and only if for every $H=GS^\epsilon$, $\phi(H) \leqq H$ in the alphabetical ordering.*

Proof. We require $\Phi(f) \leqq f$ for every $f \in F$. Hence $\Phi(H)=\phi(H) \leqq H$ is clearly a necessary condition for the representation to be the alphabetical representation. For sufficiency, suppose $\phi(H) \leqq H$ for all $H=GS^\epsilon$. Let us take an arbitrary $f=a_1 \cdots a_t$ in reduced form and define $f_0=1$, $f_1=a_1$, $f_2=a_1a_2, \cdots, f_t=a_t \cdots a_t=f$. Here

$$\Phi(f_0) = 1 = f_0,$$
$$\Phi(f_1) = \phi(1 \cdot a_1) = g_1 \leqq 1 \cdot a_1 = f_1,$$

$$\Phi(f_2) = \phi(g_1a_2) = g_2 \leqq g_1a_2 \leqq a_1a_2 = f_2,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$\Phi(f_{i+1}) = \phi(g_ia_{i+1}) = g_{i+1} \leqq g_ia_{i+1} \leqq f_ia_{i+1} = f_{i+1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$\Phi(f_t) = \phi(g_{t-1}a_t) = g_t \leqq g_{t-1}a_t \leqq f_{t-1}a_t = f_t.$$

As $f=f_t$, $\Phi(f)\leqq f$. Note that since f is in reduced form we may conclude from $g_i\leqq f_i$ that $g_ia_{i+1}\leqq f_ia_{i+1}$.

Given a standard representation of $U=U[\{G\}, \phi(H)]$ which is not the alphabetical representation, it is not difficult to construct the alphabetical representation. Since this is not the alphabetical representation, there will be some $g_is_j^\epsilon$ such that $\phi(g_is_j^\epsilon)=g_k>g_is_j^\epsilon$. The procedure is to replace the Schreier system $\{G\}$ by another Schreier system $\{G'\}$ in which the new representative of the coset Ug_k is $g_k'=g_is_j^\epsilon$. We define $g_u'=g_u$ for all g_u which do *not* have g_k as a beginning section. We define $g_k'=g_is_j^\epsilon=g_i's_j^\epsilon$. Note that $g_i's_j^\epsilon=g_is_j^\epsilon$ is reduced as it stands, since if s_j^ϵ cancels with g_i, then by (2.1.2) $\phi(g_i\,s_j^\epsilon)=g_is_j^\epsilon$, contrary to our assumption that $g_k>g_is_j^\epsilon$. If g_v has g_k as a beginning section, then $g_v=g_kw$ is the reduced form of g_v. Here w does not begin with $s_j^{-\epsilon}$, for then $g_ks_j^{-\epsilon}$ is a G, and by (2.1.3) $\phi(g_ks_j^{-\epsilon})=\phi[\phi(g_is_j^\epsilon)s_j^{-\epsilon}]=g_i$, whence by (2.1.2) $g_ks_j^{-\epsilon}=g_i$ and $g_k=g_is_j^\epsilon$ again contrary to the assumption $g_k>g_is_j^\epsilon$. Hence if we put $g_v'=g_is_j^\epsilon w$ whenever $g_v=g_kw$, we are sure that g_v' is reduced as written, since we have verified that s_j^ϵ does not cancel with either g_i or w. Also since $Ug_k=Ug_is_j^\epsilon$, it follows that $Ug_kw=Ug_is_j^\epsilon w$ or $Ug_v=Ug_v'$. The beginning sections of $g_u'=g_u$ will be of the form $g_r'=g_r$. The beginning sections of $g_v'=g_is_j^\epsilon w$ will be either of the same form or beginning sections of $g_i=g_i'$. Hence $\{G'\}$ is a Schreier system. The new ϕ function is easily determined by the rule $\phi'(G'S^\epsilon)=[\phi(GS^\epsilon)]'$. This replacement of $\{G\}$ by $\{G'\}$ has the effect of replacing every G which is altered by an earlier G', since from $g_is_j^\epsilon<g_k$ we have $g_is_j^\epsilon w<g_kw$. Since any sequence $g_k>g_k'>g_k''>\cdots$ must be finite, after a finite number of alterations the representative of a given coset remains unaltered, no matter how often this process is repeated. Moreover, if an $f=a_1a_2\cdots a_t$ is the earliest element in its coset, each of $f_0=1$, $f_1=a_1$, $f_2=a_1a_2$, $\cdots$, $f_t=f$ is the earliest element in its coset and so this process will in t applications replace $\Phi(f_1)$ $\Phi(f_2)$ $\cdots$ $\Phi(f)$ by f_1, f_2, $\cdots$, $f_t=f$. Hence repeated applications of the process will yield the alphabetical representation. In general the process is infinite. But if $F=F_r$ has a finite number of generators, finding the earliest representative for each coset which contains an element of length less than some given N is a finite process.

4. **Semi-alphabetical ordering.** J. Nielsen [4] has devised a process for finding the free generators of a finitely generated subgroup U of a free group F_r. The properties of the alphabetical representation suggest a refinement of the Neilsen process. Two advantages of the process given here are that we

find a canonical set of generators for U, and that it is not necessary that U be finitely generated.

If U is given its alphabetical representation $U=U[\{G\}, \phi(H)]$, then the free generators of U are of the form $u=g_i s_\alpha^\epsilon g_j^{-1}\neq 1$ where $\phi(g_i s_\alpha^\epsilon)=g_j$, and $\phi(g_j s_\alpha^{-\epsilon})=g_i$. From the property $\phi(H)\leqq H$ in the alphabetical representation we have

$$(4.1)\qquad g_j < g_i s_\alpha^{\epsilon}, \qquad g_i < g_j s_\alpha^{-\epsilon}.$$

Since $u\neq 1$ we may exclude $\phi(H)=H$ and have the proper inequalities of (4.1). From (4.1) we have directly: (A) If u is of odd length, $l(u)=2k+1$, $l(g_i)=l(g_j)=k$. (B) If u is of even length, $l(u)=2k$, then either: (1) $l(g_i s_\alpha^\epsilon)=l(g_j)=k$ and $g_j<g_i s_\alpha^\epsilon$ or (2) $l(g_i)=l(g_j s_\alpha^{-\epsilon})=k$ and $g_i<g_j s_\alpha^{-\epsilon}$. In case $l(u)$ is even and the first alternative holds, we may replace u by $u^{-1}=g_j s_\alpha^{-\epsilon} g_i^{-1}$ and the second alternative holds. In other words, we choose whichever of u or u^{-1} has the significant factor s_α^ϵ or $s_\alpha^{-\epsilon}$ in position $k+1$.

Suppose now that we are given a subgroup U of F_r, *not* in terms of any standard representation but as generated by elements $\alpha_1, \alpha_2, \cdots, \alpha_r, \cdots$ of F_r not assumed to be free generators of U. We define a semi-alphabetical ordering for all elements α of F_r. If in its reduced form α is of odd length, $l(\alpha)=2k+1$, write α in the form $\alpha=\beta_1 s_i^\epsilon \beta_2^{-1}$, where $l(\beta_1)=l(\beta_2)=k$. We order all α's on (1) $l(\alpha)=2k+1$, (2) the alphabetical order of β_1, (3) the alphabetical order of β_2, (4) the alphabetical order of s_i^ϵ. If in its reduced form the length of α is even, $l(\alpha)=2k$, write α in the form $\alpha=\beta_1\beta_2^{-1}$ where $l(\beta_1)=l(\beta_2)$. We order all α's on (1) $l(\alpha)=2k$, (2) the alphabetical order of β_1, (3) the alphabetical order of β_2. We combine these two orderings to define the semi-alphabetical ordering of F_r. There is no difficulty about the ordering between elements of even and odd lengths, since the lengths are different. It is not difficult to show that the semi-alphabetical ordering is a simple ordering of the elements of F_r, and in fact an enumeration since r is finite. We shall write the semi-alphabetical ordering in the form $x\ll y$ and also write $x\leqq\leqq y$ to mean that x precedes or is equal to y. The semi-alphabetical ordering may be used to characterize the free generators of a subgroup U given by its alphabetical representation.

THEOREM 4.1. *Let the generators of U given by the alphabetical representation $U=U[\{G\}, \phi(H)]$ be the set $\{u\}$. Replacing, if necessary, each u by u^{-1} so that $u\ll u^{-1}$ and numbering the u's so that $1\ll u_1\ll u_2\ll u_3\ll \cdots$, then we have $u_j\ll u_i^\epsilon u_j^\eta\neq 1$, $u_j\ll u_j^\epsilon u_i^\eta\neq 1$,* for $\epsilon=\pm 1$, $\eta=\pm 1$. *Conversely if $\alpha_1, \alpha_2, \cdots$ are elements of F_r satisfying*

$$(4.2.1)\qquad 1\ll\alpha_1\ll\alpha_2\ll\alpha_3\ll\cdots,$$

$$(4.2.2)\qquad \alpha_i\ll\alpha_i^{-1},$$

$$(4.2.3)\qquad \alpha_j\ll\alpha_i^{\epsilon}\alpha_j^{\eta}\neq 1, \qquad \alpha_j\ll\alpha_j^{\epsilon}\alpha_i^{\eta}\neq 1, \textit{ all } i, j,$$

then the α's *are the free generators of a subgroup* U *of* F_r *given by the alphabetical representation of* U.

Proof. The direct part of the theorem is relatively easy. Let $u = g_i s_\alpha^\epsilon g_j^{-1} \neq 1$ where $\phi(g_i s_\alpha^\epsilon) = g_j$ and $\phi(g_j s_\alpha^{-\epsilon}) = g_i$. Then as remarked above the condition $u \ll u^{-1}$ means $g_i < g_j$ for u's of odd length, $2k+1$, and for u's of even length, $2k$, it means $l(g_i) = k$, $l(g_j) = k-1$ and $g_i < g_j s_\alpha^{-\epsilon}$. Here the significant factor s_α^ϵ is in the middle of a u of odd length and in position $k+1$ of a u of even length, $2k$. Now consider the semi-alphabetical ordering of a u_j in comparison with that of a product $u_i^\epsilon u_j^\eta$ or $u_j^\epsilon u_i^\eta$. Since the cancellation in these products cannot include the significant factor of either term by SS 3.11, the number of cancellations from each term is at most half the length of the shorter of u_i or u_j. Hence u_j will be shorter than either product unless $l(u_i) \leqq l(u_j)$ and exactly half of u_i has been cancelled with part of u_j. If $u_i = \beta_1 \beta_2^{-1}$, $l(\beta_1) = l(\beta_2)$, $\beta_1 < \beta_2$, the significant factor of u_i is the first term in β_2^{-1}, and so the β_1 terms have been cancelled. If this happens in $u_i^\epsilon u_j^\eta$, then $\epsilon = -1$, $u_j^\eta = \beta_1 z$ with $l(z) \geqq l(\beta_1)$ since $l(u_j) \geqq l(u_i) = 2\ l(\beta_1)$. Here $u_i^{-1} u_j^\eta = \beta_2 \beta_1^{-1} \beta_1 z = \beta_2 z$. But as $\beta_1 < \beta_2$ and $l(\beta_1) = l(\beta_2) \leqq l(z)$ it follows that $u_j^\eta \ll u_i^\epsilon u_j^\eta$, and as $u_j \ll u_j^{-1}$ a fortiori that $u_j \ll u_i^\epsilon u_j^\eta$. If the cancellation of half of u_i occurs in $u_j^\epsilon u_i^\eta$ it follows that $\eta = +1$, $u_i = \beta_1 \beta_2^{-1}$, $u_j^\epsilon = z\beta_1^{-1}$, $u_j^\epsilon u_i = z\beta_2^{-1}$ where $l(\beta_1) = l(\beta_2) = \leqq l(z)$. Since $\beta_1 < \beta_2$ it follows that $u_j^\epsilon = z\beta_1^{-1} \ll z\beta_2^{-1} = u_j^\epsilon u_i$. Note that we have proved slightly more than was required having shown that $u_j^\eta \ll u_i^\epsilon u_j^\eta$ and $u_j^\epsilon \ll u_j^\epsilon u_i^\eta$.

The converse is somewhat more difficult though the ideas are similar. Let $\alpha_i = \beta_1 s_l^\epsilon \beta_2^{-1}$ where if $l(\alpha_i) = 2k+1$, $l(\beta_1) = l(\beta_2) = k$ and if $l(\alpha_i) = 2k$, $l(\beta_1) = k$, $l(\beta_2) = k-1$, and $\beta_1 < \beta_2 s_l^{-\epsilon}$ since $\alpha_i \ll \alpha_i^{-1}$. Now define the s_l^ϵ as the significant factor of α_i, $s_l^{-\epsilon}$ as the significant factor of $\alpha_i^{-1} = \beta_2 s_l^{-\epsilon} \beta_1^{-1}$. We wish to show that from (4.2) it follows that in a product $\alpha_i^\epsilon \alpha_j^\eta$ neither significant factor is cancelled. Arguing on length alone (4.2.3) will be violated if either α_i or α_j is of odd length and its significant factor is cancelled in the product. Similarly (4.2.3) is violated by length alone if more than half of either α_i^ϵ or α_j^η is cancelled in the product. Suppose then the significant factor of α_i^ϵ is cancelled. Then we need only consider the case in which α_i is of even length $2k$, $\epsilon = +1$, $\alpha_i = \beta_1 s_l^\rho \beta_2^{-1}$, $l(\beta_1) = k = l(\beta_2 s_l^{-\rho})$, $\beta_1 < \beta_2 s_l^{-\rho}$ and $l(\alpha_j^\eta) \geqq 2k$, $\alpha_j^\eta = \beta_2 s_l^{-\rho} z$, $l(z) \geqq k$. Here $\alpha_i \alpha_j^\eta = \beta_1 z \ll \alpha_j^\eta = \beta_2 s_l^{-\rho} z$ since $\beta_1 < \beta_2 s_l^{-\rho}$. This is in conflict with (4.2.3) if $\eta = +1$. If $\eta = -1$ then $\alpha_j \alpha_i^{-1} = z^{-1} \beta_1^{-1} \ll \alpha_j = z^{-1} s_l^\rho \beta_2^{-1}$. A similar conflict is reached if the significant factor of α_j^η is cancelled.

Lemma 4.1. *The elements* α *satisfying* 4.2 *are free generators of a subgroup* U *of* F_r.

Proof. Consider any product of α's, $A_1 A_2 \cdots A_t$, where $A_i A_{i+1} \neq 1$, $i = 1, \cdots, t-1$, and each A_i is some α^ϵ. From what has just been shown neither significant factor is included in the cancellation between consecutive terms $A_i A_{i+1}$. Thus all cancellation in $A_1 A_2 \cdots A_t$ leaves the significant factors intact and so $A_1A_2 \cdots A_t \neq 1$, whence the α's are free generators of a subgroup U of F_r.

To prove that the α's are the generators of U given by the alphabetical representation of U we need a stronger lemma.

LEMMA 4.2. *If $f = A_1A_2 \cdots A_t$, $t \geqq 2$, where $A_iA_{i+1} \neq 1$ and each A_i is some α^ϵ of* (4.2), *then $A_1A_2 \cdots A_{t-1} \ll f$ and $A_2 \cdots A_t \ll f$. Consequently $A_i \ll f$, $i = 1, \cdots, t$.*

The special case $t = 2$ has already been treated in showing that the significant factor of $\alpha_i^\epsilon\alpha_j^\eta \neq 1$ is not cancelled and that $\alpha_i^\epsilon \ll \alpha_i^\epsilon\alpha_j^\eta$, $\alpha_j^\eta \ll \alpha_i^\epsilon\alpha_j^\eta$. Let us proceed by induction on t. f will be longer than $A_1A_2 \cdots A_{t-1}$ unless exactly half of A_t has been cancelled in the product. In this case $A_t = \alpha_i = \beta_1\beta_2^{-1}$ where $l(\beta_1) = l(\beta_2) = k$ and $\beta_1 < \beta_2$. Here $A_1A_2 \cdots A_{t-1} = z\beta_1^{-1}$ where β_1^{-1} is part of A_{t-1} and at most half of A_{t-1} is cancelled by A_t. Hence $l(\beta_1^{-1}) \leqq l(A_{t-1})/2$ $\leqq l(A_1 \cdots A_{t-1})/2$ and so $l(z) \geqq l(\beta_1^{-1})$. Here $f = A_1A_2 \cdots A_t = z\beta_1^{-1}\beta_1\beta_2^{-1}$ $= z\beta_2^{-1}$ whence $A_1 \cdots A_{t-1} = z\beta_1^{-1} \ll z\beta_2^{-1} = f$ since $\beta_1 < \beta_2$. A similar argument shows $A_2 \cdots A_t \ll f$. Applying the lemma repeatedly we find $A_i \ll A_{i-1}A_i$ $\ll \cdots \ll A_1 \cdots A_i \ll A_1 \cdots A_{i+1} \ll \cdots \ll f$, for $i = 1, \cdots, t$.

Now let $1 \ll u_1 \ll u_2 \ll \cdots$ be the generators of U given by the alphabetical representation of U where $u_i \ll u_i^{-1}$. Both the u's and the α's satisfy (4.3) and so Lemma 4.2 is applicable. If we express u_1 as a product of α's, and some α_i occurs in the product, then $\alpha_i \ll u_1$ if the product has more than two terms and $\alpha_i^\epsilon = u_1$ if there is but one term. Here $\alpha_i \ll \alpha_i^{-1}u_1 \ll u_1^{-1}$, whence $\epsilon = +1$ and $\alpha_i = u_1$. In either event $\alpha_1 \leqq \leqq \alpha_i \leqq \leqq u_1$. By reversing the roles of u's and α's we find $u_1 \leqq \leqq \alpha_1$, whence $\alpha_1 = u_1$. Now we proceed by induction. Having shown that $\alpha_1 = u_1, \alpha_2 = u_2, \cdots, \alpha_{i-1} = u_{i-1}$ let us consider α_i and u_i. Expressing α_i in terms of u's we must have some $u_j, j \geqq 1$, in the form of α_i, for since the α's are free generators α_i cannot be a product of u_j's, $j < i$. Hence by Lemma 4.2, $u_i \leqq \leqq u_j \leqq \leqq \alpha_i$. Similarly, $\alpha_i \leqq \leqq u_i$ and so $\alpha_i = u_i$. Hence the α's have been identified with the u's and the proof of the theorem is complete.

In still another way the generators given by the alphabetical representation are unique among all possible sets of free generators. They are the earliest possible system of generators in a very strong sense.

THEOREM 4.2. *Let $\alpha_1, \alpha_2, \alpha_3, \cdots$ be the generators of U given by its alphabetical representation indexed so that relations* (4.2) *hold. Then if $\beta_1, \beta_2, \cdots$ are any set of free generators of U indexed so that $\beta_1 \ll \beta_2 \ll \beta_3 \ll \cdots$, then α_1 $\leqq \leqq \beta_1, \cdots, \alpha_i \leqq \leqq \beta_i, \cdots$.*

Proof. The commutator subgroup U' of U can be characterized in terms of any set of free generators of U. An element $f \in U$ will belong to U' if and only if in the expression for f each generator occurs with exponents whose algebraic sum is zero. See SS, §7.6. Hence in expressing $\beta_1, \cdots, \beta_n$ in terms of α's at least n different α's will occur, since otherwise we could find an element which should belong to U' according to its expression in α's but not according to its expression in β's. If α_r is the latest α in the product for β_j, then, by Lemma 4.2, $\alpha_r \leqq \leqq \beta_j \leqq \leqq \beta_n$. But some $r \geqq n$, whence $\alpha_n \leqq \leqq \alpha_r \leqq \leqq \beta_n$. This holds for all finite n, proving the theorem.

We now apply Theorem 4.1 to obtain a refinement of the Nielsen process. Suppose we are given elements $\beta_1, \beta_2, \cdots$ of F_r which generate some subgroup U of F_r. We do not assume that the β's are free generators of U or that they are finite in number. We may replace the set $\beta_1, \beta_2, \cdots$ by another set $\beta_1', \beta_2', \cdots$ which generate U applying any one of the four following elementary changes:

Type A. If some $\beta_i = 1$, put $\beta_j' = \beta_j$ for $j < i$, $\beta_j' = \beta_{j+1}$ for $j \geqq i$. This suppresses a $\beta_i = 1$.

Type B. For a fixed pair of indices i and j, put $\beta_i' = \beta_j$, $\beta_j' = \beta_i$ and for $k \neq i, j$, put $\beta_k' = \beta_k$.

Type C. For a fixed i, put $\beta_i' = \beta_i^{-1}$ and for $j \neq i$, put $\beta_j' = \beta_j$.

Type D. For a fixed pair i and j, $i \neq j$, put $\beta_j' = \beta_i^\epsilon \beta_j^\eta$ or $\beta_j^\epsilon \beta_i^\eta$, ϵ, $\eta = \pm 1$, $\beta_i' = \beta_i$, and $\beta_k' = \beta_k$ all $k \neq i, j$.

Theorem 4.3. *By a well defined succession of elementary changes of types* A, B, C, D *we may replace arbitrary generators* $\beta_1, \beta_2, \cdots$ *of a subgroup* U *of* F_r *by the free generators* $\alpha_1, \alpha_2, \cdots$ *of* U *which satisfy relations* 4.3.

Proof. We shall operate in turn on the sets $\{\beta_1\}$, $\{\beta_1, \beta_2\}, \cdots$ $\{\beta_1, \cdots, \beta_n\}, \cdots$. More precisely we begin with the set $V_1 = \{\beta_1\}$, and proceed in a definite way. Given a set V_k we transform it into a set W_k, and then take a set V_{k+1} which will be W_k with β_{k+1} adjoined. In a set V (1) we apply the change of type A to the first β_i to which it applies; (2) we apply type B if $\beta_j \ll \beta_i$, using the smallest i to which this is applicable and for the given i the smallest j; (3) we apply type C to the smallest i for which $\beta_i^{-1} \ll \beta_i$; (4) we apply type D if $\beta_j' \ll \beta_j$, using the smallest possible j and the smallest i consistent with this j. Each V_k is a finite set and so type A changes can be applied only a finite number of times. For the rest of the changes the first β altered is replaced by an earlier β'. Hence these processes must come to an end for a set V_k. When no further alterations are possible we call the set W_k. Now W_k will satisfy the relations (4.2). We may verify trivially that $\beta_i \ll \beta_i^\epsilon \beta_i^\epsilon$ always and so any failure of relations (4.2) to hold will lead to one of the elementary changes. Now what relation do the elements obtained in the W_k's bear to the generators $\alpha_1, \alpha_2, \cdots$ of U given by the alphabetical representation which satisfy (4.2)? By application of Lemma 4.2 we see that α_{i+1} is the earliest element of U not a product of $\alpha_1, \cdots, \alpha_i$. Hence if α_1 is expressible as a product of $\beta_1, \cdots, \beta_k$, then, in W_k, $\beta_1' = \alpha_1$. For β_1' is the earliest element ($\neq 1$ naturally) in the group generated by $\beta_1, \cdots, \beta_k$ and α_1 is the earliest element in U. Similarly as k increases $\beta_1' = \alpha_1$, $\beta_2' = \alpha_2$, and so on. In this way we may construct the α's from the β's. If the number of β's is finite, then the process is finite. If the number of β's is infinite, then the values of β_1', β_2', $\cdots$ become stationary for increasing k when and only when $\beta_1' = \alpha_1$, $\beta_2 = \alpha_2$, and so on.

THEOREM 4.4. *Given a subgroup U of F_r generated by elements $\alpha_1, \alpha_2, \alpha_3, \cdots$ satisfying the relations* (4.2). *Then we may solve the decision problem for U, that is, we may decide whether or not a given γ belongs to U.*

Proof. If γ does belong to U, then $\gamma = A_1A_2 \cdots A_t$ where each $A_i =$ some α_j^ϵ and $A_iA_{i+1} \neq 1$, $i = 1, \cdots, t-1$. Since the α's may be identified with the generators given by the alphabetical representation of U, even though this representation is not assumed to be known, the cancellation in the product $A_1A_2 \cdots A_t$ does not include any significant factor. Hence the reduced form of γ begins with $\beta_1 s_i^\epsilon$ if $A_1 = \alpha = \beta_1 s_i^\epsilon \beta_2^{-1}$ and with $\beta_2 s_i^{-\epsilon}$ if $A_1 = \alpha^{-1}$. Hence γ does not belong to U unless its reduced form begins with either $\beta_1 s_i^\epsilon$ or $\beta_2 s_i^{-\epsilon}$ belonging to some $\alpha = \beta_1 s_i^\epsilon \beta_2^{-1}$. If γ is of reduced form $\beta_1 s_i^\epsilon z$, put $\gamma_1 = \alpha^{-1}\gamma = \beta_2^{-1} z$, and if γ is of reduced form $\beta_2^{-1} s_i^{-\epsilon} z$, put $\gamma_1 = \alpha\gamma = \beta_1 z$. In both cases γ belong to U if and only if γ_1 belongs to U. But $\gamma_1 \ll \gamma$ by the same arguments as used in Lemma 4.2, and so in a finite number of steps the question may be settled.

COROLLARY. *For the subgroup U of the theorem we may construct in a finite number of steps all left coset representatives of the alphabetical representation of U whose length does not exceed a given finite value N.*

For we may test in turn the finite number of elements of F_r whose length does not exceed N. The first element $g_2 \neq g_1 = 1$ which does not belong to U will be an earliest coset representative. Then an x not belonging to U will belong to Ug_2 if and only if xg_2^{-1} belongs to U. In general two elements x and y will belong to the same left coset of U if and only if xy^{-1} belongs to U. Thus applying the methods of the theorem we may subdivide any finite set of elements of F according to the left cosets of U and find from these the earliest element in each coset.

5. **A separation theorem.** Using results of Magnus [3] on the complex commutator series in a free group, Iwasawa [2] has shown that there exists a subgroup U of finite index in a free group F which does not contain a given $\beta \neq 1$. With the machinery available here we may prove a much stronger result.

THEOREM 5.1. *Given a free group F with an arbitrary number of generators, and a finite number of elements $\alpha_1, \alpha_2, \cdots, \alpha_m$ of F. Suppose we are also given a finite number of elements $\beta_1, \beta_2, \cdots, \beta_n$ such that no β belongs to the subgroup H generated by $\alpha_1, \alpha_2, \cdots, \alpha_m$. Then we may construct a subgroup $\overline{H}$ of finite index in F containing $\alpha_1, \alpha_2, \cdots, \alpha_m$ (and hence H) but no one of β_1 $\beta_2, \cdots, \beta_n$.*

Proof. Let r be the number of generators of F which appear in any of the words for $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n$. Then without loss of generality we may confine our attention to F_r. For if we may find an $\bar{H}$ in F_r as required, then we may construct a subgroup of the same properties in F by adjoining to $\bar{H}$ all of the generators of F not in F_r and also all conjugates of these generators. By Theorem 4.3 we may in F_r replace $\alpha_1, \cdots, \alpha_m$ by other elements generating H which satisfy relations (4.2). Since the number of α's is finite this is a finite process. Let us suppose this already done. By the corollary to Theorem 4.4 we may construct the coset representatives of any limited length for the alphabetical representation of U. Let us suppose this done for all cosets containing elements of lengths not exceeding the longest of $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n$. Let these be $g_1=1, g_2, \cdots, g_s$. From Theorem 4.1 every α is of the form $g_i s_l^\epsilon g_j^{-1}$. Now let us consider the $\phi(GS^\epsilon)$ table treated in §§3 and 4 of SS. Any finite part of this table for H may be constructed by application of Theorem 4.4. This means that for every $\alpha = g_i s_l^\epsilon g_j^{-1}$ we find g_j in the s_l^ϵ column and g_i row and g_i in the $s_l^{-\epsilon}$ column and g_j row. Now in the manner of §4 let us construct a subgroup $\bar{H}$ of F_r of index s with coset representatives $g_1=1, g_2, \cdots, g_s$. The G's as earliest coset representatives of H of lengths not exceeding N will form a Schreier system. In filling in the $\phi(GS^\epsilon)$ table for $\bar{H}$ there cannot be more compulsory entries than there were in the corresponding part of the H table. For each of $\alpha_1, \cdots, \alpha_m$, we have $\alpha = g_i s_l^\epsilon g_j^{-1}$. If in the $\bar{H}$ table we enter g_j for $\phi(g_i s_l^\epsilon)$ and g_i for $\phi(g_j s_l^{-\epsilon})$, then $\alpha_1, \cdots, \alpha_m$ will be among the generators of $\bar{H}$. Let us then complete the $\bar{H}$ table in any permissible way, and this will surely be possible since the Schreier system $\{G\}$ is finite. Hence $\bar{H}$ is a subgroup of finite index s in F_r and contains $\alpha_1, \cdots, \alpha_m$ and so also H. What about the $\beta_1, \cdots, \beta_n$? A β_i belongs to some coset of H whose representative does not exceed N in length. Hence for some g_j of $g_2, \cdots, g_s$, $\beta_i g_j^{-1} \in H$. As $H \leqq \bar{H}$ it follows that $\beta_i g_j^{-1} \in \bar{H}$. But if it were true that $\beta_i \in \bar{H}$ it would follow that $g_j \in \bar{H}$. This cannot happen since $g_j \neq 1$ is one of the coset representatives of $\bar{H}$ in F_r. Hence $\beta_i \in \bar{H}$ and the theorem is proved.

A comparison of this result with Iwasawa's is not out of order. Iwasawa proved the existence of a *normal* subgroup of finite index not containing a given $\beta \neq 1$. Since every subgroup of finite index contains a normal subgroup of finite index, the result given here trivially includes Iwasawa's. Moreover there is no real generality in the theorem above in taking more than one β, since we may exclude the β's one at a time and take the intersection of the resulting subroups. But it is not possible with the methods used here to find a *normal* subgroup of finite index containing the α's and excluding the β's. In fact it is conceivable that F may contain no proper normal subgroup of finite index containing the α's, even though the normal subgroup generated by the α's and their transforms should be of infinite index in F.

6. **Representation of normal subgroups.** It has been shown in Theorem 3.1 that if U is a normal subgroup of a free group F, then U possesses coset

representatives $\{G\}$ which are a two-sided Schreier system, that is:

If $g=a_1a_2 \cdots a_t$ is a G, then

(1) $a_1 \cdots a_{t-1}$ is a G and

(2) $a_2 \cdots a_t$ is a G.

But there may not be any normal subgroup whose coset representatives are a given two-sided Schreier system. For example, if F is the free group on two generators a, b, then the set 1, a, b, ab, ba is a two-sided Schreier system, but there is no normal subgroup U of F with these as coset representatives. If there were then F/U would be a group of order 5 and hence abelian, and so $aba^{-1}b^{-1}\in U$ whence ab and ba should belong to the same coset.

It is not at all easy to distinguish normal subgroups of F from non-normal subgroups, nor do we possess as much information on the existence of normal subgroups such as we find in Theorem 5.1 for subgroups in general.

As a first step in the study of normal subgroups we find a criterion for a standard representation $U=U[\{G\}, \phi(H)]$ which assures us that U is a normal subgroup.

THEOREM 6.1. *Given a standard representation* $U=U[\{G\}, \phi(H)]$ *of a subgroup* U *of* F. *Then* U *is a normal subgroup of* F *if and only if* $\Phi(S^\epsilon H)=\Phi(S^\epsilon\phi(H))$ *for every* S^ϵ *and every* H.

Proof. *Necessity.* Suppose U is a normal subgroup of F. Here $H\phi(H)^{-1}\in U$ and as U is a normal subgroup, then also $S^\epsilon H\phi(H)^{-1}S^{-\epsilon}\in U$. Thus $S^\epsilon H$ and $S^\epsilon\phi(H)$ belong to the same coset of U and so $\Phi(S^\epsilon H)=\Phi(S^\epsilon\phi(H))$.

Sufficiency. Suppose $\Phi(S^\epsilon H)=\Phi(S^\epsilon\phi(H))$ for every S^ϵ and every H. This is equivalent to saying $S^\epsilon H\phi(H)^{-1}S^{-\epsilon}\in U$. But the elements $H\phi(H)^{-1}$ generate U. The condition assures us that the transform of any generator of U by an arbitrary S or S^{-1} is again an element of U. Hence U is transformed into itself by words of length one in F. By induction it follows that U is transformed into itself by every element of F and hence is a normal subgroup of F.

The standard representation of subgroups has been given in terms of left cosets. It is clear, however, that right cosets may be similarly used if the representatives $\{G\}$ are a reverse Schreier system, that is, if $a_1a_2 \cdots a_t$ is a G, then $a_2 \cdots a_t$ is a G. The analogue to $\phi(GS^\epsilon)$ is a function $\psi(S^\epsilon G)$ where

(6.1.1) $\psi(S^\epsilon G)$ is a G,

(6.1.2) If $S^\epsilon G$ is a G, then $\psi(S^\epsilon G)=S^\epsilon G$,

(6.1.3) $\psi[S^{-\epsilon}\psi(S^\epsilon G)]=G$.

It has been noted above that a normal subgroup U has a set of representatives $\{G\}$ which form a two-sided Schreier system. Hence such a set may be used for both a (left) standard representation and a right standard representation of U. If we write $\Phi(G^{-1})=G^I$, then we may easily show that $\psi(S^\epsilon G)=[\phi(G^IS^{-\epsilon})]^I$. This idea may be used to define another characterization of normal subgroups, which we shall give here without proof.

THEOREM 6.2. *Given a standard representation $U=U[\{G\},\phi(H)]$ of a subgroup U of F where $\{G\}$ is a two-sided Schreier system. Then U is a normal subgroup of F if and only if there is a mapping of G onto itself $G\rightarrow G^I$ with the following properties*:

(6.2.1) $1^I=1$.

(6.2.2) $(G^I)^I=G$.

(6.2.3) $\phi(S^{-\epsilon})^I=\phi(S^\epsilon)$.

(6.2.4) *If we put $\psi(S^\epsilon G)=[\phi(G^I S^{-\epsilon})]^I$, then $\phi[\psi(T^\eta G)S^\epsilon]=\psi[T^\eta\phi(GS^\epsilon)]$ for any G and any generators T, S* and η, $\epsilon=\pm 1$.

Both Theorem 6.1 and 6.2 are somewhat tedious in application, but the condition of normality is so strong that it seems unlikely that any criterion noticeably easier may be found.

BIBLIOGRAPHY

1. M. Hall, Jr., and T. Rado, *On Schreier systems in free groups*, Trans. Amer. Math. Soc. vol. 64 (1948) pp. 386–408.

2. K. Iwasawa, *Einige Sätze über freie Gruppen*, Proc. Imp. Acad. Tokyo vol. 19 (1943) pp. 272–274.

3. W. Magnus, *Über Beziehungen zwischen höheren Kommutatoren*, J. Reine Angew. Math. vol. 177 (1937) pp. 105–115.

4. J. Nielsen, *Om Regning med ikkekommutative faktorer og dens anvendelse i gruppenteorien*, Matematisk Tidsskrift B (1921) pp. 77–94.

5. O. Schreier, *Die Untergruppen der freien Gruppen*, Abh. Math. Sem. Hamburgischen Univ. vol. 5 (1926) pp. 161–183.

6. J. von Neumann and E. P. Wigner, *Minimally almost periodic groups*, Ann. of Math. vol. 41 (1940) pp. 746–750.

THE OHIO STATE UNIVERSITY,
COLUMBUS, OHIO.