SOME PROPERTIES OF FREE GROUPS

BY

HERBERT FEDERER AND BJARNI JÓNSSON

1. **Introduction.** In this paper the methods introduced into the theory of free groups by Jacob Nielsen(1) are applied and extended. It is shown (3.9) that if Nielsen reductions are applied *at random* to a finite sequence of elements of a free group, then the process will terminate after finitely many steps and yield a free generating set for the subgroup generated by the given elements. It is proved (4.2) that if G is freely generated by the set X and if a subgroup H of G is well ordered in any manner consistent with X-length, then H is freely generated by the set of all those elements which are not expressible in terms of preceding elements(2). The homomorphisms of free groups into free groups are completely described by the theorem (6.4) that if f is a homomorphism of a free group G onto a free group H, then G is expressible as a free product of two subgroups such that f maps one factor isomorphically onto H and f maps the other factor onto the identity element of H(3). The answer to the question whether a given finite subset of a free group generates a free factor of this group is shown (5.1) to be finitely computable.

Since Nielsen's fundamental paper is written in Danish, the present paper includes an exposition (§3) of that part of his work which it uses—with certain modifications.

The paper is so arranged that the reader may pass from §3 to any of the last three sections, which are substantially independent of each other.

2. **Elementary facts.** We say that the group G *is freely generated by* X if

Presented to the Society, April 30, 1949; received by the editors April 1, 1949.

(1) J. Nielsen, *Om Regning med ikke-kommutative Faktorer og dens Anvendelse i Gruppeteorien*, Matematisk Tidsskrift (1921) pp. 77–94. Related papers by the same author are *Über die Isomorphismen unendlicher Gruppen ohne Relation* and *Die Isomorphismengruppe der freien Gruppen* in vols. 79 and 91 of Math. Ann.

(2) This implies the theorem of O. Schreier, *Die Untergruppen der freien Gruppen*, Abh. Math. Sem. Hamburgischen Univ. vol. 5 (1926) pp. 161–183, which states that every subgroup of a free group is a free group.

(3) The authors wish to thank Professor R. Baer for informing them that, in the special case in which G is finitely generated, this can be proved by a method used by F. Levi, *Über die Untergruppen der freien Gruppen* (2. Mitteilung), Math. Zeit. vol. 37 (1933) §3, pp. 95–97. See also J. H. C. Whitehead, *On equivalent sets of elements of a free group*, Ann. of Math. vol. 37 (1936) pp. 782–800, in particular §4, I. Gruschko, *Über die Basen eines freien Produktes von Gruppen*, Matematicheskii Sbornik vol. 50 (1940) pp. 169–182 (Russian with German summary), and A. G. Kurosch, *Teoriya Grupp*. OGIZ, Moscow-Leningrad, 1944 (Russian), in particular p. 314. We believe that the two theorems stated on p. 303 by Kurosch, and substantially credited to Gruschko, can probably be extended to the case of infinitely many arbitrary free factors by a convergence process very much like the one which we develop in §6.

and only if X generates G and for every function f on X into any group K there is a homomorphism F of G into K such that $F(x)=f(x)$ for $x\in X$.

A *free group* is a group which is freely generated by some set.

It follows at once that if G and G' are freely generated by X and X' and if card X = card X'(4), then every one-to-one correspondence between X and X' can be extended to an isomorphism between G and G'. Hence *the isomorphism type of a free group is determined by the cardinal number of any one of its free generating sets.*

It follows that G is freely generated by X if and only if $X\cap X^{-1}=0$ and every element of G is uniquely expressible in the form

$$\prod_{i=1}^{n} y_i,$$

where n is a nonnegative integer (in case $n=0$, the product is the identity element of G), $y_1, \cdots, y_n\in X\cup X^{-1}$ and $y_i\neq y_{i+1}^{-1}$ for $i=1, \cdots, n-1$.

Clearly the unique representation property implies the homomorphism property. In order to prove the converse it is sufficient, in view of our preceding remark, to construct a single group which is freely generated by a set with prescribed cardinal number and such that the unique representation property holds. Such a group is readily constructed as a homomorphic image of a semigroup of finite sequences.

It further follows that G is freely generated by X if and only if G is generated by X, $X\cap X^{-1}=0$, and there do not exist elements $y_1, \cdots, y_n \in X\cup X^{-1}$ such that

$$\prod_{i=1}^{n} y_i = e, \quad n > 0, \quad y_i \neq y_{i+1}^{-1} \text{ for } i = 1, \cdots, n-1.$$

We observe that $\{e\}$ is freely generated by 0.

If G is freely generated by X and C is the commutator subgroup of G, then G/C is a free abelian group generated by the image of X under the natural homomorphism. If S is the subgroup of all squares in G/C, then the cardinal number of $(G/C)/S$ equals $2^{\text{card } X}$ or card X according as X is finite or infinite. Hence *the cardinal number of a free generating set of a free group is determined by the group.*

In case G is freely generated by X we associate with each $a\in G$ the nonnegative integer $L_X(a)$, the *length of* a *with respect to* X, by the conditions

$$L_X(a) = n, \qquad a = \prod_{i=1}^{n} y_i,$$

$$y_1, \cdots, y_n \in X \cup X^{-1}, \qquad y_i \neq y_{i+1}^{-1} \text{ for } i = 1, \cdots, n-1.$$

(4) We agree that card X = the cardinal number of X.

In particular $L_X(e)=0$.

If $a_1, \cdots, a_m \in G$, then

$$L_X\left(\prod_{i=1}^{m} a_i\right) \leqq \sum_{i=1}^{m} L_X(a_i)$$

with equality holding if and only if

$$L_X(a_i a_{i+1}) = L_X(a_i) + L_X(a_{i+1}) \text{ for } i = 1, \cdots, m-1.$$

If $a, b \in G$, then there exist unique $u, v, w \in G$ such that

$$\begin{aligned} a &= uv^{-1}, & L_X(a) &= L_X(u) + L_X(v), \\ b &= vw^{-1}, & L_X(b) &= L_X(v) + L_X(w), \\ ab &= uw^{-1}, & L_X(ab) &= L_X(u) + L_X(w). \end{aligned}$$

We say that u is an *initial X segment of a* if and only if there is a v such that

$$a = uv^{-1}, \qquad L_X(a) = L_X(u) + L_X(v).$$

We observe that in this case v is an initial X segment of a^{-1}.

Extending the notion of length, we *define*

$$L_X(\alpha) = \sum_{i=1}^{p} L_X(\alpha_i)$$

for every p-termed sequence α of elements of G.

We say that the group G is the *free product* of its subgroups $H_1, \cdots, H_s$, and write

$$G = H_1 * H_2 * \cdots * H_s,$$

if and only if $H_1 \cup H_2 \cup \cdots \cup H_s$ generates G and for any homomorphisms $f_1, \cdots, f_s$ of $H_1, \cdots, H_s$ into any group K there is a homomorphism F of G into K such that $F(x) = f_j(x)$ for $x \in H_j$, $j = 1, \cdots, s$.

It follows that $G = H_1 * H_2 * \cdots * H_s$ if and only if every element of G is uniquely expressible in the form

$$\prod_{i=1}^{n} x_i,$$

with $e \neq x_i \in H_{p(i)}$ for $i = 1, \cdots, n$ and $p(i) \neq p(i+1)$ for $i = 1, \cdots, n-1$.

For each subset A of a group G we let $[A]$ be the subgroup of G generated by A.

Among the sets which generate a group there are some whose cardinal number is minimal. This cardinal number is the *rank* of the group. We observe that rank $\{e\} = 0$.

If G is freely generated by $A \cup B$ and $A \cap B = 0$, then $[A]$, $[B]$ are freely generated by A, B and

$$G = [A] * [B].$$

If H, K are freely generated by A, B and if $G = H * K$, then G is freely generated by $A \cup B$.

The class of *elementary transformations of order n* is the least class C with the following properties:

(1) Every element of C is a function which maps the class of all n-termed sequences of group elements into itself.

(2) C is closed to superposition.

(3) If $1 \leqq i \leqq n$ and g is the function such that

$$g(\alpha) = \langle \alpha_1, \cdots, \alpha_{i-1}, \alpha_i^{-1}, \alpha_{i+1}, \cdots, \alpha_n \rangle$$

for every n-termed sequence α of group elements, then $g \in C$.

(4) If $1 \leqq i \neq j \leqq n$ and g is the function such that

$$g(\alpha) = \langle \alpha_1, \cdots, \alpha_{i-1}, \alpha_j \alpha_i, \alpha_{i+1}, \cdots, \alpha_n \rangle$$

for every n-termed sequence α of group elements, then $g \in C$.

It is easy to check that if the group H is freely generated by the range of a univalent n-termed sequence α and if g is an elementary transformation of order n, then $g(\alpha)$ is a univalent sequence and H is freely generated by the range of $g(\alpha)$.

We observe that to every permutation of the integers $1, \cdots, n$ there corresponds an elementary transformation of order n.

3. **Nielsen reductions.** Jacob Nielsen([5]) discovered a process which allows one to compute from each finite subset A of a free group, in finitely many steps, a set B such that $[A]$ is freely generated by B. This section describes Nielsen's method; in particular the theorems 3.10, 3.11, 3.13 are due to him, and 3.9 generalizes one of his results by allowing greater freedom in the computational process without impairing its finitary character.

Suppose G is freely generated by X. A subset A of G is said to have the *Nielsen property with respect to X* if and only if

(1) $A \cap A^{-1} = 0$,

(2) $a, b \in A \cup A^{-1}$, $L_X(ab) < L_X(a)$ implies that $b = a^{-1}$,

(3) $a, b, c \in A \cup A^{-1}$, $L_X(abc) \leqq L_X(a) - L_X(b) + L_X(c)$ *implies that either* $b = a^{-1}$ or $c = b^{-1}$.

3.1. LEMMA. *If G is freely generated by X, $A \subset G$, A has the Nielsen property with respect to X, and if $a_1, \cdots, a_s \in A \cup A^{-1}$, $a_i \neq a_{i+1}^{-1}$ for $i = 1, \cdots, s-1$, and $b = \prod_{i=1}^{s} a_i$, then there exist $y_1, \cdots, y_s, z_0, \cdots, z_s \in G$ such that*

([5]) See footnote 1.

$$z_0 = z_s = e, \qquad a_i = z_{i-1}y_i z_i^{-1}, \qquad y_i \neq e,$$
$$L_X(a_i) = L_X(z_{i-1}) + L_X(y_i) + L_X(z_i),$$
$$L_X(z_{i-1}) \leqq L_X(a_i)/2, \qquad L_X(z_i) \leqq L_X(a_i)/2,$$
$$b = \prod_{i=1}^{s} y_i, \qquad L_X(b) = \sum_{i=1}^{s} L_X(y_i),$$
$$L_X(b) \geqq L_X(a_1)/2 + s - 2 + L_X(a_s)/2.$$

Proof. The statement is trivial in case $s=1$. Using induction, suppose the statement holds for s and let

$$a_1, \cdots, a_{s+1} \in G, \ a_i \neq a_{i+1}^{-1} \quad \text{for} \quad i = 1, \cdots, s, \qquad \beta = \prod_{i=1}^{s+1} a_i.$$

Determine b, $y_1, \cdots, y_s$, $z_0, \cdots, z_s$ in terms of $a_1, \cdots, a_s$.
Choose u, v, $w \in G$ so that

$$a_s = uv^{-1}, \qquad L_X(a_s) = L_X(u) + L_X(v),$$
$$a_{s+1} = vw^{-1}, \qquad L_X(a_{s+1}) = L_X(v) + L_X(w),$$
$$a_s a_{s+1} = uw^{-1}, \qquad L_X(a_s a_{s+1}) = L_X(u) + L_X(w).$$

Subtracting the last of the three equations on the right from the sum of the first two and using (2) in the definition of the Nielsen property, we find that

$$2L_X(v) = L_X(a_s) + L_X(a_{s+1}) - L_X(a_s a_{s+1})$$
$$\leqq \inf\,(L_X(a_s), L_X(a_{s+1})),$$
$$L_X(u) \geqq L_X(a_s)/2, \qquad L_X(w) \geqq L_X(a_{s+1})/2.$$

Hence u and z_{s-1} are initial X segments of a_s with $L_X(z_{s-1}) \leqq L_X(u)$, and we choose $r \in G$ so that

$$u = z_{s-1}r, \qquad L_X(u) = L_X(z_{s-1}) + L_X(r).$$

We shall prove that $r \neq e$. In fact, if $r = e$, then

$$a_{s-1}a_s a_{s+1} = z_{s-2}y_{s-1}w^{-1},$$
$$L_X(a_{s-1}a_s a_{s+1}) \leqq L_X(z_{s-2}) + L_X(y_{s-1}) + L_X(w)$$
$$= L_X(a_{s-1}) - L_X(z_{s-1}) + L_X(a_{s+1}) - L_X(v)$$
$$= L_X(a_{s-1}) - L_X(a_s) + L_X(a_{s+1}),$$

and we would get a contradiction from (3) in the definition of the Nielsen property.

Defining $\eta_1, \cdots, \eta_{s+1}$, $\zeta_0, \cdots, \zeta_{s+1}$ by the formulae

$$\eta_i = y_i \quad \text{for} \quad i = 1, \cdots, s-1; \qquad \eta_s = r, \quad \eta_{s+1} = w^{-1},$$
$$\zeta_i = z_i \quad \text{for} \quad i = 0, \cdots, s-1; \qquad \zeta_s = v, \quad \zeta_{s+1} = e,$$

we readily check that the statement of the lemma holds with s, b, y, z replaced by $s+1, \beta, \eta, \zeta$.

3.2. COROLLARY. *The hypothesis of Lemma* 3.1 *implies that*

$$L_X\Big(\prod_{i=j}^{k} a_i\Big) \leqq L_X(b) \text{ for } 1 \leqq j \leqq k \leqq s.$$

Proof. For $k<s$ we have

$$L_X(z_k) \leqq L_X(a_{k+1}) - L_X(z_k) = L_X(y_{k+1}) + L_X(z_{k+1}),$$
$$L_X\Big(\prod_{i=j}^{k} a_i\Big) = L_X(z_{j-1}) + \sum_{i=j}^{k} L_X(y_i) + L_X(z_k)$$
$$\leqq L_X(z_{j-1}) + \sum_{i=j}^{k+1} L_X(y_i) + L_X(z_{k+1}) = L_X\Big(\prod_{i=j}^{k+1} a_i\Big).$$

It follows that

$$L_X\Big(\prod_{i=j}^{k} a_i\Big) \leqq L_X\Big(\prod_{i=j}^{s} a_i\Big).$$

A similar argument yields the inequality

$$L_X\Big(\prod_{i=j}^{s} a_i\Big) \leqq L_X\Big(\prod_{i=1}^{s} a_i\Big) = L_X(b).$$

3.3. COROLLARY. *The hypothesis of Lemma* 3.1 *implies that*

$$L_X(a_i) \leqq L_X(b) \text{ for } i = 1, \cdots, s.$$

3.4. COROLLARY. *If A has the Nielsen property with respect to X, then $[A]$ is freely generated by A.*

Furthermore

$$L_A(b) \leqq L_X(b) \text{ for } b \in [A].$$

Proof. Since the hypothesis of Lemma 3.1 cannot hold with $s>0$ and $b=e$, $[A]$ is freely generated by A.

3.5. REMARK. It is possible to show that, even if A satisfies only (1) and (3) of the definition of the Nielsen property, $[A]$ is freely generated by A. However, we shall never use this fact, proof of which is somewhat more involved.

3.6. DEFINITION. *Suppose G is freely generated by X and α is a finite sequence of elements of G.*

For $a \in G$, let $M_X(a, \alpha)$ be the number of ordered pairs (i, v) for which $v = \pm 1$ and a is an initial X segment of α_i^v.

Also let $N_X(a, \alpha)$ be the number of ordered pairs (i, v) for which $v = \pm 1$, a is an initial X segment of α_i^v, and $L_X(a) \leqq L_X(\alpha_i)/2$.

For each positive integer k, let

$$P_X^k(\alpha) = \coprod_{a \in S} (N_X(a, \alpha) + 1),$$

where $S = \{a \mid a \in G \text{ and } L_X(a) = k\}$.

We observe that *if a is an initial X segment of both c and c^{-1} and if $c \neq e$, then $L_X(a) < L_X(c)/2$.*

Otherwise we could find $y \in G$ such that

$$c = ay^{-1}, \qquad L_X(y) = L_X(c) - L_X(a) \leqq L_X(a).$$

Then a and y would be initial X segments of c^{-1} and we could choose $z \in G$ so that

$$a = yz, \qquad L_X(a) = L_X(y) + L(z).$$

Consequently

$$c = yzy^{-1}, \qquad L_X(c) = L_X(y) + L_X(z) + L_X(y).$$

Thus $a = yz$ and yz^{-1} would be initial X segments of c^{-1} with equal X length. Hence $yz = yz^{-1}$, $z = e$, $c = e$.

3.7. DEFINITION. *Let G be freely generated by X.*

Suppose α and β are n-termed sequences of elements of G. We say that

β is an X reduction of α

if and only if there are integers i, u and elements a, b of G such that

(1) $u = \pm 1$, $\alpha_i^u = ab^{-1}$, $L_X(\alpha_i) = L_X(a) + L_X(b) > 0$,

either (2) $L_X(a) < L_X(b)$ *and* $M_X(b, \alpha) \geqq 2$,

or (3) $L_X(a) = L_X(b)$, $N_X(a, \alpha) \geqq 2$ *and* $N_X(b, \alpha) \geqq 2$,

(4) $\beta_j = \alpha_i^u \alpha_j$ *in case* $j \neq 1$ *and b is an initial X segment of α_j but not of α_j^{-1},*

(5) $\beta_j = \alpha_j \alpha_i^{-u}$ *in case* $j \neq i$ *and b is an initial X segment of α_j^{-1} but not of α_j,*

(6) $\beta_j = \alpha_i^u \alpha_j \alpha_i^{-u}$ *in case* $j \neq i$ *and b is an initial X segment of both α_j and α_j^{-1},*

(7) $\beta_j = \alpha_j$ *for all j not covered by* (4), (5), (6).

3.8. COROLLARY. *Under the conditions of Definition 3.7 we have:*

(8) *There is an elementary transformation g such that $g(\alpha) = \beta$.*

(9) $L_X(\beta_j) \leqq L_X(\alpha_j)$ *for* $j = 1, \cdots, n$.

(10) *If* $L_X(a) < L_X(b)$, *then* $L_X(\beta) < L_X(\alpha)$.

(11) *If* $L_X(a) = L_X(b) = m$ *and* $L_X(\beta) = L_X(\alpha)$, *then* $P_X^m(\beta) < P_X^m(\alpha)$ *and*

$P_X^k(\beta) \leqq P_X^k(\alpha)$ *for* $k > m$.

Proof. The statements (8), (9), (10) are obvious. Suppose now that the hypotheses of (11) hold. Then

$$\begin{aligned} L_X(\beta_j) &= L_X(\alpha_j) \text{ for } j = 1, \cdots, n, \\ N_X(b, \beta) &= 1, \\ N_X(a, \beta) &= N_X(b, \alpha) + N_X(a, \alpha) - 1, \\ N_X(c, \beta) &= N_X(c, \alpha) \quad \text{whenever} \quad L_X(c) = m,\ c \neq a,\ c \neq b, \end{aligned}$$

and since

$$2 \cdot (p + q) < (p + 1)(q + 1) \text{ for } p > 1,\ q > 1,$$

we have

$$\begin{gathered} (N_X(b, \beta) + 1)(N_X(a, \beta) + 1) < (N_X(b, \alpha) + 1)(N_X(a, \alpha) + 1), \\ P_X^m(\beta) < P_X^m(\alpha). \end{gathered}$$

Next suppose $k > m$. If $L_X(c) = k$ and neither a nor b is an initial X segment of c, then

$$N_X(c, \beta) = N_X(c, \alpha).$$

All other c for which $L_X(c) = k$ are of the form ad or bd with $L_X(d) = k - m$. In case $L_X(ad) = k$ but $L_X(bd) \neq k$, then

$$N_X(ad, \beta) = N_X(ad, \alpha).$$

In case $L_X(bd) = k$, but $L_X(ad) \neq k$, then

$$N_X(bd, \beta) = 0 = N_X(bd, \alpha).$$

In case $L_X(ad) = k$ and $L_X(bd) = k$, then

$$\begin{aligned} N_X(bd, \beta) &= 0, \\ N_X(ad, \beta) &= N_X(ad, \alpha) + N_X(bd, \alpha), \end{aligned}$$

and since

$$p + q + 1 \leqq (p + 1)(q + 1) \text{ whenever } p \geqq 0,\ q \geqq 0,$$

we have

$$(N_X(ad, \beta) + 1)(N_X(bd, \beta) + 1) \leqq (N_X(ad, \alpha) + 1)(N_X(bd, \alpha) + 1).$$

Pairing ad and bd whenever the last of these cases occurs, we obtain the inequality

$$P_X^k(\beta) \leqq P_X^k(\alpha).$$

3.9 THEOREM. *If G is freely generated by X and $\alpha^{(1)}, \cdots, \alpha^{(p)}$ are such n-termed sequences of elements of G that $\alpha^{(i+1)}$ is an X reduction of $\alpha^{(i)}$ for $i=1, \cdots, p-1$, and if $q=L_X(\alpha^{(1)})$, then*

$$p \leqq (q+1)(2n)^{2nq}.$$

Proof. Let V be the set of all those $q+1$ termed sequences v of integers such that

$$0 \leqq v_1 \leqq q, \qquad 1 \leqq v_i \leqq (2n)^{2n} \text{ for } i = 2, \cdots, q+1.$$

We order V lexicographically so that $v<w$ if and only if there is an integer j such that $v_i=w_i$ for $1 \leqq i<j$ and $v_j<w_j$. Let

$$v^{(i)} = \langle L_X(\alpha^{(i)}), P_X^q(\alpha^{(i)}), P_X^{q-1}(\alpha^{(i)}), \cdots, P_X^1(\alpha^{(i)}) \rangle$$

for $i=1, \cdots, p$ and note that

$$v^{(i)} \in V \quad \text{for} \quad i = 1, \cdots, p, \qquad v^{(i+1)} < v^{(i)} \quad \text{for} \quad i = 1, \cdots, p-1,$$

$$p \leqq \text{card } V = (q+1)(2n)^{2nq}.$$

3.10. THEOREM. *If G is freely generated by X, α is a finite sequence of elements of G, and if there exists no X reduction of α, then* (range $\alpha - \{e\}$) *has the Nielsen property with respect to X and*

$$\alpha_i = \alpha_j \text{ implies } i = j \text{ or } \alpha_i = e.$$

3.11. THEOREM. *Suppose G is freely generated by X and A is a finite subset of G. Then $[A]$ is a free group.*

In fact, $[A]$ is freely generated by a set B which has the Nielsen property with respect to X. Furthermore

$$\text{card } B \leqq \text{card } A,$$

with equality holding if and only if $[A]$ is freely generated by A.

Proof. Let $n=$card A and let $\alpha^{(1)}$ be a univalent n-termed sequence whose range is A.

Let $\alpha^{(2)}, \cdots, \alpha^{(p)}$ be n-termed sequences such that $\alpha^{(i+1)}$ is an X reduction of $\alpha^{(i)}$ for $i=1, \cdots, p-1$ and such that there exists no X reduction of $\alpha^{(p)}$. Then the set

$$B = \text{range } \alpha^{(p)} - \{e\}$$

has the Nielsen property with respect to X. Therefore the group

$$[A] = [\text{range } \alpha^{(1)}] = [\text{range } \alpha^{(p)}] = [B]$$

is freely generated by B, and

$$\text{card } B \leqq n = \text{card } A.$$

If $[A]$ is freely generated by A, then card $B=$card A.

If card $B=$card A, then card $B=n$; hence $\alpha^{(p)}$ is univalent, range $\alpha^{(p)}=B$, and $[A]$ is freely generated by range $\alpha^{(p)}$. Since there exists an elementary transformation g of order n for which $g(\alpha^{(p)})=\alpha^{(1)}$, we conclude that $[A]$ is freely generated by range $\alpha^{(1)}=A$.

3.12. REMARK. *Suppose A and B are as above. Then*

$$\sum_{b\in B} L_X(b) \leqq \sum_{a\in A} L_X(a),$$

and if equality holds, then $[A]$ is freely generated by $A-\{e\}$.

This inequality holds because the left member equals $L_X(\alpha^{(p)})$, the right member equals $L_X(\alpha^{(1)})$, and

$$L_X(\alpha_j^{(p)}) \leqq L_X(\alpha_j^{(1)}) \text{ for } j = 1, \cdots, n.$$

If $[A]$ is not freely generated by $A-\{e\}$, then

$$\begin{aligned}\text{card } \{j \mid \alpha_j^{(1)} \neq e\} &= \text{card } (A - \{e\}) > \text{card } B \\ &= \text{card } \{j \mid \alpha_j^{(p)} \neq e\}.\end{aligned}$$

We therefore have $\alpha_j^{(1)}\neq e$ and $\alpha_j^{(p)}=e$ for some j, and strict inequality holds.

3.13. THEOREM(6). *If G is freely generated by A, then* rank $G=$card A.

If G is a free group, $G=[A]$, rank G *is finite and* rank $G=$card A, *then G is freely generated by A.*

Proof. In case rank G is infinite, the conclusion follows from the fact that card $G=$card A for every set A such that $G=[A]$.

From now on suppose rank G is finite, and suppose G is freely generated by X. Clearly rank $G\leqq$card X.

First we choose a finite set A_0 for which $G=[A_0]$ and rank $G=$card A_0. We use Theorem 3.11 to obtain a set B_0 such that card $B_0\leqq$card A_0 and G is freely generated by B_0. Hence card $X=$card $B_0\leqq$rank G, and we conclude that

$$\text{card } X = \text{rank } G.$$

Now, if G is freely generated by A, then card $A=$card $X=$rank G.

On the other hand, if $G=[A]$ and rank $G=$card A, then A is finite. If G were not freely generated by A, then Theorem 3.11 would yield a set B such

(6) An immediate consequence of this result, due to J. Nielsen, loc. cit. footnote 1, is the theorem, conjectured by H. Hopf, that every homomorphism of a finitely generated free group onto itself is an isomorphism. In fact, if G is freely generated by the finite set X and f is a homomorphism of G onto G, then $G=[f(X)]$ and rank $G\leqq$card $f(X)\leqq$card $X=$rank G, hence G is freely generated by $f(X)$. Consequently f is an isomorphism. W. Magnus, *Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring*, Math. Ann. vol. 111 (1935) pp. 259–280, has published a proof of this theorem.

that $G=[B]$ and card $A>$card $B\geqq$rank G.

4. **Level sets.** The main result of this section is Theorem 4.2, and the reader may better understand our motivation by glancing at 4.2 and 4.3 before reading 4.1.

Suppose G is freely generated by X. A subset A of G will be called a *level set with respect to* X if and only if $[A]$ is freely generated by A and for each $b\in[A]$ we have

$$b \in [\{a \mid a \in A \quad \text{and} \quad L_X(a) \leqq L_X(b)\}].$$

It follows from Corollary 3.3 and the first statement in Corollary 3.4 that *every set which has the Nielsen property with respect to* X *is a level set with respect to* X.

In case $[A]=G$, we see that A is a level set with respect to X if and only if there is a subset Y of X such that $A=(X-Y)\cup Y^{-1}$.

For each subset A *of* G and each nonnegative integer n we *define*

$$Q_X(A, n)$$

as the set of all n-termed sequences α of elements of G such that

(1) $[\text{range } \alpha]=[A]$,

(2) $\alpha_i\notin[\{\alpha_j \mid j<i\}]$ for $1\leqq i\leqq n$,

(3) If $1\leqq i\leqq n$, $b\in[A]$, and $L_X(b)<L_X(\alpha_i)$, then $b\in[\{\alpha_j \mid j<i\}]$.

We observe that if A is a level set with n elements, then there is a univalent sequence $\alpha\in Q_X(A, n)$ for which range $\alpha=A$.

In case $[A]$ *is freely generated by* A and α, β are elements of $Q_X(A, n)$, we say that β *is an admissible reduction of* α *with respect to* A *and* X if and only if there are integers i, j, u, v such that

$$1 \leqq i \neq j \leqq n, \qquad u = \pm 1, \qquad v = \pm 1, \qquad \alpha_i \in A \cup A^{-1},$$

$$\beta_j = \alpha_i^u \alpha_j^v, \qquad L_X(\beta_j) = L_X(\alpha_j), \qquad L_A(\beta_j) < L_A(\alpha_j),$$

$$\beta_k = \alpha_k \text{ for } k \neq j.$$

4.1. LEMMA. *Suppose* $A\subset G$ *and* $\alpha\in Q_X(A, n)$. *We then have*:

(1) $[A]$ *is freely generated by* range α.

(2) *If* A *has the Nielsen property with respect to* X, *then there exist* $\alpha^{(1)}, \cdots, \alpha^{(p)}\in Q_X(A, n)$ *such that* $\alpha^{(1)}=\alpha$, $\alpha^{(i+1)}$ *is an admissible reduction of* $\alpha^{(i)}$ *with respect to* A *and* X *for* $i=1, \cdots, p-1$, *and*

$$\text{range } \alpha^{(p)} \subset A \cup A^{-1}.$$

Proof. These statements are trivial in case $n=0$. We shall use induction with respect to n.

First we show that, for each integer n, (2) implies (1).

In fact, if $A\subset G$ and $\alpha\in Q_X(A, n)$, then $[A]$ is finitely generated and there is a set $B\subset G$ such that $[B]=[A]$ and B has the Nielsen property

with respect to X. Then $\alpha \in Q_X(B, n)$ and there exist successive admissible reductions $\alpha = \alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(p)} \in Q_X(B, n)$ with respect to B and X such that range $\alpha^{(p)} \subset B \cup B^{-1}$. Since $\beta_i^u \neq \beta_j$ for $\beta \in Q_X(B, n)$, $1 \leqq i < j \leqq n$, $u = \pm 1$, we conclude that $[B]$ is freely generated by range $\alpha^{(p)}$. Since α can be obtained from $\alpha^{(p)}$ by an elementary transformation, it follows that $[B]$ is freely generated by range α.

Now assume n is a positive integer and that the lemma holds with n replaced by any smaller integer.

In order to prove (2), suppose A has the Nielsen property with respect to X and $\alpha \in Q_X(A, n)$.

If $\beta, \gamma \in Q_X(A, n)$ and γ is an admissible reduction of β with respect to A and X, then $L_A(\gamma) < L_A(\beta)$. Hence every sequence of successive admissible reductions with respect to A and X is finite, and we may just as well *assume that there is no admissible reduction of α with respect to A and X, and prove that, in this case,* range $\alpha \subset A \cup A^{-1}$.

Since $\alpha \in Q_X(A, n)$, we have

$$L_X(\alpha_i) \leqq L_X(\alpha_{i+1}) \text{ for } i = 1, \cdots, n-1.$$

Let $t = L_X(\alpha_n)$, let m be the least integer such that $L_X(\alpha_{m+1}) = t$ and let

$$\beta = \langle \alpha_1, \cdots, \alpha_m \rangle, \qquad B = \{a \mid a \in A, L_X(a) < t\}.$$

Since $\alpha \in Q_X(A, n)$ and A has the Nielsen property with respect to X, which implies that A is a level set with respect to X, it follows that

$$[\text{range } \beta] = [[A] \cap \{b \mid L_X(b) < t\}] = [B],$$

$\beta \in Q_X(B, m)$, B has the Nielsen property with respect to X.

We observe that there exists no admissible reduction of β with respect to B and X because every such reduction would induce an admissible reduction of α with respect to A and X. Since $m < n$, we may apply (2) to B, β, m and conclude that

$$\text{range } \beta \subset B \cup B^{-1},$$

$$\text{range } \beta \cup (\text{range } \beta)^{-1} = B \cup B^{-1}.$$

Next we shall prove the following statement:

If $L_X(\alpha_j) = t$, then $L_A(\alpha_j) \leqq 2$ and $\alpha_j \in [A - B]$.

Suppose $L_A(\alpha_j) = s$ and

$$\alpha_j = \prod_{k=1}^{s} a_k, \qquad a_k \in A \cup A^{-1} \text{ for } k = 1, \cdots, s.$$

In case $a_1 \in B \cup B^{-1}$, we have $a_1 = \alpha_i^{-u}$ with $i \leqq m < j$, $u = \pm 1$, $L_X(\alpha_i^u \alpha_j) \leqq L_X(\alpha_j)$ by Lemma 3.2, and $L_X(\alpha_i^u \alpha_j) \geqq t$ because $\alpha_j \notin [B]$. Then the sequence

$$\langle \alpha_1, \cdots, \alpha_{j-1}, \alpha_i^u \alpha_j, \alpha_{j+1}, \cdots, \alpha_n \rangle$$

would be an element of $Q_X(A, n)$ and an admissible reduction of α with respect to A and X.

Hence $L_X(a_1) = t$. Similarly $L_X(a_s) = t$.

By Lemma 3.1 we have

$$t = L_X(\alpha_j) \geqq L_X(a_1)/2 + s - 2 + L_X(a_s)/2 = s - 2 + t,$$

hence $s \leqq 2$. This proves our statement.

Let $C = \{\alpha_j | L_A(\alpha_j) = 2\}$.

Then card $C < n$, because otherwise $L_A(b) = 2$ for $b \in$ range α, hence $L_A(b)$ would be even for $b \in [\text{range } \alpha] = [A]$.

Let γ be the subsequence of α such that range $\gamma = C$.

For each $b \in C$ we successively infer that $b \notin B \cup B^{-1}$, $b \notin$ range β, $L_X(b) = t$, $b \in [A - B]$. It follows that $[C] \subset [A - B]$ and that $\gamma \in Q_X(C, \text{card } C)$. We apply (1) to C, γ, card C and conclude that

$$[C] \textit{ is freely generated by } C.$$

Let D be the set of all those elements of $A \cup A^{-1}$ which are initial A segments of elements of $C \cup C^{-1}$. Then

$$D \cap D^{-1} = 0,$$

because otherwise there would exist a, b, c such that

$$a, b, c \in (A - B) \cup (A - B)^{-1} \quad \text{and} \quad ab, a^{-1}c \in C \cup C^{-1},$$

and we could apply Lemma 3.1, with $s = 3$, $a_1 = c^{-1}$, $a_2 = a$, $a_3 = b$ to infer that $c^{-1}a = y_1 y_2 z_2^{-1}$, $ab = z_1 y_2 y_3$,

$$L_X(y_1) + L_X(y_2) + L_X(z_2) = L_X(c^{-1}a) = t,$$
$$L_X(z_1) + L_X(y_2) + L_X(y_3) = L_X(ab) = t.$$

Adding the last two equations we would obtain

$$L_X(c^{-1}ab) + L_X(a) = 2t = L_X(c^{-1}) + L_X(b),$$

contrary to the fact that A has the Nielsen property with respect to X.

Next we shall prove the following statement:

If $C \neq 0$, *then* card $C \leqq$ card $D - 1$.

Consider a graph whose vertices correspond to the elements of D and whose edges correspond to the elements of C in such a way that the end points of the edge corresponding to c are the vertices corresponding to the initial A segments of c and c^{-1}. Since every simple closed polygon corresponds to a sequence of elements of $C \cup C^{-1}$ whose product is e and since $[C]$ is freely generated by C, our graph is a tree. Since the difference between the number of vertices and the number of edges of a tree equals the number of its components, we conclude that

$$\text{card } D - \text{card } C \geqq 1.$$

Next we define

$$E = A \cap (D \cup D^{-1}),$$
$$F = A \cap (\text{range } \alpha \cup (\text{range } \alpha)^{-1}),$$

and we shall prove the following statement:

If $C \neq 0$, *then* $n \leqq \text{card } A + \text{card } (E \cap F) - 1$.

This inequality holds because

$$n = \text{card } F + \text{card } C \leqq \text{card } F + \text{card } D - 1$$
$$= \text{card } F + \text{card } E - 1 = \text{card } (E \cup F) + \text{card } (E \cap F) - 1.$$

The remainder of our argument will hinge on the following proposition:

If $1 \leqq i \leqq n$, $\delta = \langle \alpha_1, \cdots, \alpha_i \rangle$, *and if* [range δ] *is freely generated by* range δ, *then* $\alpha_i \notin D \cup D^{-1}$.

In fact, if α_i^{-u} were an initial A segment of α_j^v, with $\alpha_j \in C$, then the sequence

$$\langle \alpha_1, \cdots, \alpha_{j-1}, \overset{u}{\alpha_i} \overset{v}{\alpha_j}, \alpha_{j+1}, \cdots, \alpha_n \rangle$$

would be an element of $Q_X(A, n)$ and an admissible reduction of α with respect to A and X.

Now if $1 \leqq i < n$ and $\delta = \langle \alpha_1, \cdots, \alpha_i \rangle$, then $\delta \in Q_X(\text{range } \delta, i)$. Applying (1) to range δ, δ, i, we conclude that [range δ] is freely generated by range δ. Hence

$$\alpha_i \notin D \cup D^{-1} \text{ for } i = 1, \cdots, n-1,$$

and $(E \cap F)$ has no element except possibly α_n or α_n^{-1}. It follows that card $(E \cap F) \leqq 1$.

If $C \neq 0$, then $n \leqq \text{card } A$, card range $\alpha \leqq \text{rank } [\text{range } \alpha]$ and we infer from Theorem 3.13 that [range α] is freely generated by range α. Taking $i = n$, $\delta = \alpha$, we conclude that $\alpha_n \notin D \cup D^{-1}$, hence $D = 0$, $C = 0$.

Consequently $C = 0$, range $\alpha \subset A \cup A^{-1}$.

4.2. THEOREM. *Suppose* G *is freely generated by* X *and* H *is a subgroup of* G. *If* H *is so well-ordered by the relation* $<$ *that*

$$a \in H, b \in H, a < b \text{ implies } L_X(a) \leqq L_X(b)$$

and if A *is the set of all* $a \in H$, *for which*

$$a \notin [\{b \mid b \in [H] \text{ and } b < a\}],$$

then H *is freely generated by* A; *in fact* A *is a level set with respect to* X.

Proof. Clearly $H = [A]$.

Suppose B is any finite subset of A.

Let C be the set of all $a \in [B]$ for which

$$a \notin [\{b \mid b \in [B] \text{ and } b < a\}].$$

Then $B \subset C$ and no element of C comes after the last element of B. Choosing such a finite subset Y of X that $B \subset [Y]$, we have

$$C \subset \{a \mid a \in [Y] \text{ and } L_Y(a) \leqq \sup_{b \in B} L_Y(b)\}$$

and conclude that C is finite. Let γ be the finite sequence such that range $\gamma = C$ and $\gamma_i < \gamma_j$ for $i < j$. Then $\gamma \in Q_X(B, \text{card } C)$ and we infer from Lemma 4.1 that $[B]$ is freely generated by C. Since $B \subset C$, we have $B = C$.

Hence $[B]$ is freely generated by B.

4.3. REMARK. An immediate consequence of the preceding theorem is the statement that *every subgroup of a free group is a free group*, which was first proved by Schreier in a quite different way(7). We note that, though Schreier's construction is not similar to ours, his free generating set is a level set with respect to X, like ours; in fact, Schreier's set has the Nielsen property with respect to X.

F. Levi has constructed a free generating set of a subgroup of a free group by well-ordering the subgroup, as we do, but, unlike ours, Levi's well-ordering is a special one(8). Levi's generating set has the Nielsen property with respect to X.

In view of the arbitrary nature of the well-ordering which we use, it is clear that every level set can be obtained, from some well-ordering of the subgroup which it generates, in the manner described in Theorem 4.2.

4.4. THEOREM. *If G is a free group, $G = A * B$, H is a subgroup of G and if $A \subset H$, then there is a subgroup C of H for which $H = A * C$.*

Proof. Choose E and F so that A and B are freely generated by E and F respectively, let $X = E \cup F$, and well-order H by the relation $<$ in such a way that

$$a \in H, b \in H, a < b \quad \text{implies} \quad L_X(a) \leqq L_X(b),$$
$$x < x^{-1} \text{ for } x \in E.$$

Let Y be the set of all $a \in H$ for which

$$a \notin [\{b \mid b \in H \text{ and } b < a\}].$$

It follows from Theorem 4.2 that H is freely generated by Y. Furthermore $E \subset Y$.

We let $C = [Y - E]$ and conclude that $H = [Y] = [E] * [Y - E] = A * C$.

(7) See footnote 2.

(8) F. Levi, *Über die Untergruppen der freien Gruppen*, Math. Zeit. vol. 32 (1930) pp. 315–318.

4.5. THEOREM(9). *Suppose G is freely generated by X, H is a finitely generated subgroup of G, and*

$$F_k = [\{a \mid a \in H \text{ and } L_X(a) \leqq k\}] \text{ for } k = 0, 1, 2, \cdots .$$

For every set A such that $[A]=H$ we have

$$\sum_{a\in A} L_X(a) \geqq \sum_{k=1}^{\infty} k\cdot(\operatorname{rank} F_k - \operatorname{rank} F_{k-1})$$

with equality holding if and only if $A-\{e\}$ is a level set with respect to X.

We observe that $F_0=\{e\}$ and $F_k=H$ for large k.

Proof. We shall always assume that A is finite and that $e\notin A$.

First suppose H is freely generated by A.

For $k=0, 1, 2, \cdots$, let

$$A_k = \{a \mid a \in A \text{ and } L_X(a) \leqq k\},$$

note that $H=[A_k]*[A-A_k]$, $[A_k]\subset F_k$, hence, by Theorem 4.4, there is a group C_k such that $F_k=[A_k]*C_k$, and we have

$$\operatorname{rank} F_k \geqq \operatorname{rank} [A_k] = \operatorname{card} A_k$$

with equality holding if and only if $F_k=[A_k]$.

Taking n so large that $A_n=A$ and $F_n=H$, we obtain

$$\begin{aligned}\sum_{a\in A} L_X(a) &= \sum_{k=1}^{n} \sum_{a\in A_k-A_{k-1}} L_X(a)\\ &= \sum_{k=1}^{n} k\cdot(\operatorname{card} A_k - \operatorname{card} A_{k-1}) = n \operatorname{card} A - \sum_{k=1}^{n-1} \operatorname{card} A_k\\ &\geqq n \operatorname{rank} H - \sum_{k=1}^{n-1} \operatorname{rank} F_k = \sum_{k=1}^{n} k\cdot(\operatorname{rank} F_k - \operatorname{rank} F_{k-1})\end{aligned}$$

with equality holding if and only if

$$F_k = [A_k] \text{ for } k = 1, \cdots, n-1,$$

that is, if and only if A is a level set with respect to X.

Next suppose A generates H, but not freely.

From 3.12 we obtain a set B which generates H freely and for which

$$\sum_{a\in A} L_X(a) > \sum_{b\in B} L_X(b) \geqq \sum_{k=1}^{\infty} k\cdot(\operatorname{rank} F_k - \operatorname{rank} F_{k-1}).$$

(9) The referee has pointed out a connection between this theorem and pp. 406–407 of *On Schreier systems in free groups* by M. Hall and T. Radó, Trans. Amer. Math. Soc. vol. 64 (1948).

4.6([10]). THEOREM. *If α and β are univalent n-termed sequences whose ranges generate the same group freely, then there exists an elementary transformation g of order n such that $g(\alpha)=\beta$.*

Proof. Let $B=$range β. From Theorem 3.8 we obtain sequences $\alpha=\gamma^{(1)}$, $\gamma^{(2)}, \cdots, \gamma^{(p)}$ such that $\gamma^{(i+1)}$ is a B reduction of $\gamma^{(i)}$ for $i=1, \cdots, p-1$, and such that there exists no B reduction of $\gamma^{(p)}$.

Clearly $\gamma^{(p)}$ can be obtained from α by an elementary transformation.

Since the range of each of these sequences generates $[B]$ freely, we have $e \notin$range $\gamma^{(p)}$, and we conclude from Theorem 3.10 that range $\gamma^{(p)}$ has the Nielsen property with respect to B; hence it is a level set with respect to B, and there is a subset Y of B such that

$$\text{range } \gamma^{(p)} = (B - Y) \cup Y^{-1}.$$

Consequently β can be obtained from $\gamma^{(p)}$ by an elementary transformation which permutes the terms of $\gamma^{(p)}$ and replaces some of them by their inverses.

4.7. THEOREM([11]). *Suppose G is freely generated by X. If α and β are univalent n-termed sequences whose ranges generate the same subgroup of G and are level sets with respect to X, then there exist univalent-n-termed sequences $\alpha=\gamma^{(1)}, \gamma^{(2)}, \cdots, \gamma^{(p)}=\beta$ such that*

$$L_X(\alpha) = L_X(\gamma^{(i)}) = L_X(\beta) \text{ for } i = 1, \cdots, p,$$

and such that, for $i=1, \cdots, p-1$, the sequence $\gamma^{(i+1)}$ is obtained from the sequence $\gamma^{(i)}$ by an elementary transformation which either is a permutation or is of one of the two special types described in (3) *and* (4) *of the definition of the term "elementary transformation" at the end of* §2.

Proof. Choose A so that $[A]=[\text{range } \alpha]$ and so that A has the Nielsen property with respect to X. Then $\alpha, \beta \in Q_X(A, n)$ and by Lemma 4.1(2) there exist sequences $\alpha=\alpha^{(1)}, \cdots, \alpha^{(q)}$ and $\beta=\beta^{(1)}, \cdots, \beta^{(r)}$ such that

$\alpha^{(i+1)}$ is an admissible reduction of $\alpha^{(i)}$ with respect to A and X, for $i=1, \cdots, q-1$,

$\beta^{(i+1)}$ is an admissible reduction of $\beta^{(i)}$ with respect to A and X, for $i=1, \cdots, r-1$,

$$\text{range } \alpha^{(q)} \subset A \cup A^{-1}, \qquad \text{range } \beta^{(r)} \subset A \cup A^{-1}.$$

Since $\alpha^{(q)}$ and $\beta^{(r)}$ are univalent sequences whose ranges generate $[A]$ freely, $\beta^{(r)}$ can be obtained from $\alpha^{(q)}$ by replacing certain terms of $\alpha^{(q)}$ by their inverses and then applying a permutation.

([10]) See footnote 1.

([11]) This theorem appears to be substantially the same as Theorem 2 of J. H. C. Whitehead, loc. cit. footnote 3.

Since $L_X(\alpha^{(i)}) = L_X(\alpha)$ for $i = 1, \cdots, q$ and $L_X(\beta^{(i)}) = L_X(\beta)$ for $i = 1, \cdots, r$, completion of the proof is immediate.

5. **Free factors.** In this section we show how to compute the answer to the question whether a given finite subset A of a group G, which is freely generated by X, generates a *free factor* of G; that is, whether there exists a subgroup H of G such that $G = [A] * H$.

We may and shall henceforth assume that $[A]$ is freely generated by A, because otherwise we could arrive at this situation by a finite number of X reductions, without changing $[A]$.

Clearly $[A]$ is then a free factor of G if and only if A is a subset of some set which generates G freely.

We may just as well assume that X is finite, for if it were not, we could effectively choose a finite subset Y of X such that $A \subset [Y]$, and by Theorem 4.4 the question whether $[A]$ is a free factor of G is equivalent to the question whether $[A]$ is a free factor of $[Y]$.

In view of Theorem 5.1, which follows, we need, in case X is finite, only consider all those subsets of G which contain A and have no element which is longer than the longest element of A; there are only finitely many such sets and for each of them we can decide by a finite number of X reductions whether or not it generates G freely.

Thus we have a method to decide the question.

We have not considered the problem whether Theorem 5.1 remains true with the word "finite" deleted throughout, because Theorem 5.1 interests us mainly for its computational applications.

5.1. THEOREM. *If G is freely generated by the finite set X, A is a finite subset of G, and if there exists such a subgroup H of G that*

$$G = [A] * H,$$

then there exists a finite subset B of G for which

$$\sup_{b \in B} L_X(b) \leqq \sup_{a \in A} L_X(a), \qquad G = [A] * [B].$$

Proof. Suppose the theorem is false. Accordingly choose A and H so that the conclusion does not hold and so that rank H is as small as possible.

In view of 3.11 and 3.12 we may assume that A has the Nielsen property. Let

$$s = \sup_{a \in A} L_X(a).$$

The hypothesis of the theorem implies the existence of some finite sets B such that G is freely generated by $A \cup B$ and $A \cap B = 0$. *If B is any such set,* we have

$$L_X(b) > s \text{ for } b \in B,$$

because otherwise we could define

$$A' = A \cup \{b \mid b \in B \text{ and } L_X(b) \leqq s\}, \qquad H' = [B - A'],$$

infer that

$$G = [A'] * H', \qquad \text{rank } H' = \text{card } (B - A') < \text{card } B = \text{rank } H,$$

and apply the minimal property of rank H to obtain a finite subset B' of G for which

$$\sup_{b \in B'} L_X(b) \leqq \sup_{a \in A'} L_X(a) \leqq s,$$

$$G = [A'] * [B'] = [A] * [B \cap A'] * [B']$$
$$= [A] * [(B \cap A') \cup B'],$$

$$\sup_{b \in (B \cap A') \cup B'} L_X(b) \leqq s,$$

contrary to the choice of A and H.

We now fix such a set B, define

$$m = \text{card } A, \qquad n = \text{card } (A \cup B),$$

and let α be a univalent n-termed sequence for which

$$A = \{\alpha_i \mid 1 \leqq i \leqq m\}, \qquad B = \{\alpha_i \mid m < i \leqq n\}.$$

We may just as well assume that there exists no X reduction β of α such that

$$[\{\beta_i \mid 1 \leqq i \leqq m\}] = [A].$$

Since rank $H > 0$, $B \neq 0$, $m < n$,

$$L_X(\alpha_i) > s \geqq 1 \text{ for } m < i \leqq n,$$

there exists, in view of Theorem 3.10 and the third and fourth paragraphs of §4, an X reduction β of α. Reverting to the notation of Definition 3.7, we observe that $i > m$ and that we can choose $j < m$ so that either (4) or (5) or (6) holds.

Now (6) is excluded because $L_X(\alpha_j) \leqq s < L_X(\alpha_i)$.

Hence (4) or (5) holds, we have $L_X(\beta_j) \leqq L_X(\alpha_j) \leqq s$, define

$$C = (B - \{\alpha_i\}) \cup \{\beta_j\},$$

and conclude that G is freely generated by $A \cup C$ with $A \cap C = 0$. It follows from the first part of this proof (with C replacing B) that

$$L_X(b) > s \text{ for } b \in C,$$

which is impossible because $\beta_j \in C$ and $L_X(\beta_j) \leqq s$.

5.2. EXAMPLE. We shall illustrate our method by applying it to the case in

which

$$X = \{x, y\}, \qquad a = x^2y^3, \qquad A = \{a\}.$$

The answer will turn out to be negative, that is, $[A]$ is not a free factor of G.

If A is a free factor of G, then there is an element b of G such that

$$L_X(b) \leqq L_X(a) = 5$$

and G is freely generated by $\{a, b\}$.

We shall, as is usually possible, shorten our computations by considering G/C, where C is the commutator group of G, and the natural homomorphism f of G onto G/C.

Then

$$f(a) = 2f(x) + 3f(y),$$
$$f(b) = mf(x) + nf(y),$$

where m and n are integers and

$$|m| + |n| \leqq 5.$$

Since $\{f(x), f(y)\}$ and $\{f(a), f(b)\}$ are two bases of the free abelian group G/C, we have

$$\begin{vmatrix} 2 & 3 \\ m & n \end{vmatrix} = \pm 1$$

and (m, n) is one of the four pairs

$$(1, 1), \quad (1, 2), \quad (-1, -1), \quad (-1, -2).$$

Consequently $L_X(b) \geqq 2$.

Since $\{a, b\}$ cannot have the Nielsen property with respect to X, there are two possibilities:

In the first case

$$L_X(a^u b^v) < L_X(a) \text{ with } u, v = \pm 1,$$

hence either x^2 or y^{-2} must be an initial X segment of b or of b^{-1}, and b is either one of the elements

$$x^2yx^{-1}, \quad x^2y^2x^{-1}, \quad x^2yx^{-1}y, \quad xy^2, \quad y^{-1}xy^2, \quad y^{-1}xy^3,$$

or the inverse of one of these elements, a situation which we need not consider separately.

In the second case

$$L_X(a^u b a^v) = 2L_X(a) - L_X(b) \text{ with } u, v = \pm 1$$

and $L_X(b)$ is even, hence b is either the element

$$xy$$

or the inverse of this element, a situation which we need not consider separately.

We compute the X reductions:

$$\langle x^2y^3,\ x^2yx^{-1}\rangle \to \langle xy^2,\ x^2yx^{-1}\rangle,$$
$$\langle x^2y^3,\ x^2y^2x^{-1}\rangle \to \langle xy,\ x^2y^2x^{-1}\rangle,$$
$$\langle x^2y^3,\ x^2yx^{-1}y\rangle \to \langle y^{-1}xy^2,\ x^2yx^{-1}y\rangle,$$
$$\langle x^2y^3,\ xy^2\rangle \to \langle x^2yx^{-1},\ xy^2\rangle,$$
$$\langle x^2y^3,\ y^{-1}xy^3\rangle \to \langle xy,\ y^{-1}xy^3\rangle,$$
$$\langle x^2y^3,\ xy\rangle \to \langle x^2y^2x^{-1},\ xy\rangle.$$

There is no X reduction of any of the six sequences on the right, nor of the sequence $\langle x^2y^3,\ y^{-1}xy^2\rangle$, and the range of none of these sequences generates G because it is not a subset of $X\cup X^{-1}$.

Hence there exists no $b\in G$ such that G is freely generated by $\{a,\ b\}$.

5.3. REMARK. Suppose G is a free group. It is obvious that:

If H is a free factor of G, then H is a *retract* of G, that is, there is a homomorphism r of G onto H such that $r(x)=x$ for $x\in H$.

In order to show that the converse of this statement is false we revert to the preceding example, let $H=[A]$ and let r be the homomorphism of G onto H such that $r(x)=a^{-1}$ and $r(y)=a$. Then r is a retraction of G onto H, though H is not a free factor of G.

6. **Homomorphisms.** It is quite obvious that if G and H are free groups, $G=S*Z$ and rank $S=$rank H, then there exists a homomorphism f of G onto H such that f maps S isomorphically onto H and f maps Z onto the identity element of H. The main purpose of this section is to show, as a consequence of Theorem 6.4, that every homomorphism of a free group onto a free group is of this simple type.

Since the argument of this section, leading up to Theorem 6.4, is somewhat involved, it may help to know roughly how one could give a relatively simple proof of a special case of the result stated above:

In fact, assume that f maps the free group G homomorphically onto the free group H and that rank G is finite. We suppose that H is freely generated by Y. We take a finite univalent sequence γ whose range generates G freely, let α be the f image of γ, and pass by successive Y reductions from α to a sequence β such that each term of β is either the identity element of H or belongs to $Y\cup Y^{-1}$, each element of $Y\cup Y^{-1}$ occurs at most once as a term of β, and H is freely generated by these terms of β. There is an elementary transformation which carries α into β; applying the same elementary trans-

formation to γ, we obtain a univalent sequence δ whose range generates G freely, and such that the f image of δ is β. We then define Z as the group generated by all those terms of δ whose f image is the identity element of H, and S as the group generated by all those terms of δ whose f image is an element of $Y \cup Y^{-1}$. Then $G = S * Z$, f maps S isomorphically onto H, and f maps Z onto the identity element of H.

6.1. LEMMA. *If f is a homomorphism of G into H,*
H is freely generated by Y, $T \subset Y$,
γ is an n-termed sequence of elements of G, $\alpha = \langle f(\gamma_1), \cdots, f(\gamma_n)\rangle$,
$\langle \alpha_1, \cdots, \alpha_s\rangle$ is a univalent sequence of elements of Y,
and if there exists a Y reduction of α, then there exist δ, β, g such that
g is an elementary transformation of order n,
$g(\gamma) = \delta$, $g(\alpha) = \beta$, $\beta = \langle f(\delta_1), \cdots, f(\delta_n)\rangle$,
β is a Y reduction of α,
$\delta_k = \gamma_k$ for $k = 1, \cdots, s$,
and such that $\alpha_k \in [T]$ implies

$$\gamma_k \in [\{\delta_m \mid (\beta_m \in [T] \quad \text{and} \quad L_Y(\beta_m) \leqq L_Y(\alpha_k)) \quad \text{or} \quad (L_Y(\beta_m) < L_Y(\alpha_k))\}].$$

Proof. Our notation will be consistent with that of Definition 3.7 except that G and X will have to be replaced by H and Y.

In case there exists a Y reduction β of α with $\alpha_i \in [T]$, we choose β accordingly. In this case we may also assume that if $\alpha_i = \alpha_k^v$ with $k \leqq s$ and $v = \pm 1$, then $i = k$.

In the alternate case, in which the preceding procedure is inapplicable, we choose i and u so that if $L_Y(a) = L_Y(b)$, then $b \notin [T]$, and so that if $\alpha_i = \alpha_k^v$ with $k \leqq s$ and $v = \pm 1$, then $i = k$.

In both cases we let g be the elementary transformation of order n such that for each n-termed sequence ξ of group elements and for $j = 1, \cdots, n$ we have

$(g(\xi))_j = \xi_i^u \xi_j$ if the conditions of (4) hold,
$(g(\xi))_j = \xi_j \xi_i^{-u}$ if the conditions of (5) hold,
$(g(\xi))_j = \xi_i^u \xi_j \xi_i^{-u}$ if the conditions of (6) hold,
$(g(\xi))_j = \xi_j$ for all j not covered by (4), (5), (6).
We define $\delta = g(\gamma)$.

In order to check that δ, β, g have the required properties it is helpful to observe that in the alternate case the following statements are true:

If $\alpha_j \in [T]$ and $L_Y(a) = L_Y(b)$, then $\beta_j = \alpha_j$.

If $\alpha_j \in [T]$, $L_Y(a) < L_Y(b)$, and $\beta_j \neq \alpha_j$, then $L_Y(\beta_j) < L_Y(\alpha_j)$ and $L_Y(\alpha_i) < L_Y(\alpha_j)$.

The very last inequality holds because otherwise there would be an element c of H and an integer v such that $v = \pm 1$, $\alpha_j^v = cb^{-1}$, $L_Y(\alpha_j) = L_Y(c)$

$+L_Y(b)>0$, $M_Y(b, \alpha)\geqq 2$, $L_Y(c)\leqq L_Y(a)<L_Y(b)$ and the first case would be applicable.

6.2. COROLLARY. *If f is a homomorphism of G into H,*
H is freely generated by Y,
$C\subset W\subset G$, W is finite,
$[W]$ is freely generated by W,
$T\subset Y$, $T\subset[f(W)]$,
$f(C)\subset Y$, f is univalent on C,
then there exists a set D such that
$[W]$ is freely generated by D, $C\subset D$,
f is univalent on $\{b\,|\,b\in D \text{ and } f(b)\in T\}$,
and for each $a\in W$ for which $f(a)\in[T]$ we have

$$a\in[D\cap\{b\,|\,f(b)\in T \quad or \quad f(b)=e \quad or \quad L_Y(f(b))<L_Y(f(a))\}].$$

Proof. Let $n=\text{card } W$, $s=\text{card } C$, and let γ be such a univalent n-termed sequence that

$$W=\text{range }\gamma, \qquad C=\{\gamma_i\,|\,1\leqq i\leqq s\}.$$

We define

$$\alpha=\langle f(\gamma_1),\cdots,f(\gamma_n)\rangle$$

and apply Lemma 6.1 to obtain sequences

$$\gamma=\gamma^{(1)},\ \gamma^{(2)},\cdots,\gamma^{(p)} \text{ and } \alpha=\alpha^{(1)},\ \alpha^{(2)},\cdots,\alpha^{(p)}$$

such that $\gamma^{(i+1)}$ and $\alpha^{(i+1)}$ are related to $\gamma^{(i)}$ and $\alpha^{(i)}$ just as δ and β are related to γ and α in the lemma, and such that there exists no Y reduction of $\alpha^{(p)}$.

It is easy to see that, for $i=1,\cdots,p$,

$$[W] \text{ is freely generated by range } \gamma^{(i)},\ C\subset \text{range } \gamma^{(i)},$$

and that for each $a\in W$ for which $f(a)\in[T]$ there exists an integer k with $a=\gamma_k$, $f(a)=\alpha_k$, and we have

$$a\in[\{\gamma_m^{(i)}\,|\,(\alpha_m^{(i)}\in[T] \text{ and } L_Y(\alpha_m^{(i)})\leqq L_Y(f(a))) \quad or \quad (L_Y(\alpha_m^{(i)})<L_Y(f(a)))\}]$$

for $i=2,\cdots,p$.

Since range $\alpha^{(p)}-\{e\}$ has the Nielsen property with respect to Y and $T\subset[\text{range } \alpha^{(p)}]$, we see that

$\alpha_k^{(p)}=(\alpha_m^{(p)})^v$ with $v=\pm 1$ implies $k=m$ or $\alpha_k^{(p)}=e$,

$\alpha_m^{(p)}\in T\cup T^{-1}$ whenever $\alpha_m^{(p)}\in[T]$.

Let D be the set obtained from range $\gamma^{(p)}$ upon replacing $\gamma_m^{(p)}$ by its inverse in each case in which $\alpha_m^{(p)}\in T^{-1}$.

6.3. THEOREM. *If f is a homomorphism of the free group G onto the free group H and if*

H is freely generated by Y,
G is freely generated by $X' \cup U' \cup V$,
f is univalent on X', $f(X') \subset Y$,
$f(x) = e$ for $x \in U'$,

then there exist subsets X'' and U'' of G such that

$X' \subset X''$, $U' \subset U''$,
G is freely generated by $X'' \cup U''$,
f is univalent on X'', $f(X'') = Y$,
$f(x) = e$ for $x \in U''$.

Proof. Let $X = X' \cup U' \cup V$, and assume that X', U', V are disjoint.

Let F be the class of all four-termed sequences

$$A = \langle A_1, A_2, A_3, A_4 \rangle$$

with the following properties:

A_1, A_2, A_3, A_4 are disjoint subsets of G,
G is freely generated by $A_1 \cup A_2 \cup A_3 \cup A_4$,
$X' \subset A_1$, $f(A_1) \subset Y$, f is univalent on A_1,
$U' \subset A_2$, $f(A_2) \subset \{e\}$, A_3 is finite, $A_4 \subset V$,

$$[A_1 \cup A_2 \cup A_3] = [X - A_4].$$

We observe that $\langle X', U', 0, V \rangle \in F$.

We partially order F by the relation $<$ in such a way that, for $A, B \in F$, we have $A < B$ if and only if

$$A \neq B, \qquad A_1 \subset B_1, \qquad A_2 \subset B_2, \qquad A_4 \supset B_4,$$

and $a \in A_3$ implies

$$a \in [B_1 \cup B_2 \cup \{b \mid b \in B_3 \text{ and } L_Y(f(b)) < L_Y(f(a))\}].$$

We shall prove the following two statements:

(1) *Every simply ordered subclass of F has an upper bound in F.*

(2) *If A is a maximal element of F (that is, if $A \leqq B$ implies $A = B$), then*

$$f(A_1) = Y, \qquad A_3 = 0, \qquad A_4 = 0.$$

The theorem is an immediate consequence of these two propositions: Using a well known inductive principle[12] in conjunction with (1) and the fact that F is nonvacuous, we obtain a maximal element A of F. We then take $X'' = A_1$, $U'' = A_2$, and apply (2).

In order to prove (1), let S be a simply-ordered subclass of F. We may

(12) K. Kuratowski, *Une méthode d'élimination des nombres transfinis des raisonnement matématiques*, Fund. Math. vol. 3 (1922) pp. 76–108, in particular p. 89.

assume that S is nonvacuous and that no element of S is an upper bound of S. We let B be the four-termed sequence such that

$$B_1 = \bigcup_{A \in S} A_1, \qquad B_2 = \bigcup_{A \in S} A_2, \qquad B_3 = 0, \qquad B_4 = \bigcap_{A \in S} A_4.$$

The fact that B is an upper bound of S will be evident as soon as we have shown that

$$[X - B_4] \text{ is freely generated by } B_1 \cup B_2,$$

and that

$$A_3 \subset [B_1 \cup B_2] \text{ whenever } A \in S.$$

Suppose $A \in S$, $a \in A_3$ and $q = L_Y(f(a))$. Choose elements $A^{(0)}, A^{(1)}, \cdots, A^{(q)}$ of S such that

$$A = A^{(0)} < A^{(1)} < \cdots < A^{(q)},$$

and check by induction with respect to i that

$$a \in [A_1^{(i)} \cup A_2^{(i)} \cup \{b \mid b \in A_3^{(i)} \text{ and } L_Y(f(b)) \leqq q - i\}]$$

for $i = 0, 1, \cdots, q$. Consequently

$$a \in [A_1^{(q)} \cup A_2^{(q)}] \subset [B_1 \cup B_2].$$

It follows next that

$$X - B_4 = \bigcup_{A \in S} (X - A_4) \subset \bigcup_{A \in S} [A_1 \cup A_2 \cup A_3] \subset [B_1 \cup B_2].$$

On the other hand

$$B_1 \cup B_2 = \bigcup_{A \in S} (A_1 \cup A_2) \subset \bigcup_{A \in S} [X - A_4] \subset [X - B_4].$$

We conclude that

$$[X - B_4] = [B_1 \cup B_2].$$

Since every finite subset of $B_1 \cup B_2$ is contained in $A_1 \cup A_2$ for some $A \in S$, it follows that $[X - B_4]$ is freely generated by $B_1 \cup B_2$.

Last we prove (2). Let A be a maximal element of F. We observe that if $A_3 = 0$ and $A_4 = 0$, then $H = [f(A_1)]$ and $f(A_1) \subset Y$. Hence $f(A_1) = Y$.

In order to show that $A_3 = 0$ and $A_4 = 0$ we assume the contrary and choose a finite nonvacuous set R such that

$$A_3 \subset R \subset A_3 \cup A_4.$$

We successively choose finite sets T and W for which

$$T \subset Y, \qquad f(R) \subset [T], \qquad T \subset [f(W)],$$

$$R \cup \{a \mid a \in A_1 \text{ and } f(a) \in T\} \subset W \subset A_1 \cup A_3 \cup A_4,$$

let $C = A_1 \cap W$, and use Corollary 6.2 to obtain a set D with the properties listed there.

Let B be the four-termed sequence such that

$$B_1 = A_1 \cup \{b \mid b \in D \text{ and } f(b) \in T\},$$
$$B_2 = A_2 \cup \{b \mid b \in D \text{ and } f(b) = e\},$$
$$B_3 = D - (B_1 \cup B_2), \qquad B_4 = A_4 - W.$$

Clearly B_1, B_2, B_3 are disjoint,

$$\begin{aligned}[A_1 \cup A_2 \cup W] &= [(A_1 - C) \cup A_2 \cup W] = [(A_1 - C) \cup A_2] * [W] \\ &= [(A_1 - C) \cup A_2] * [D] = [(A_1 - C) \cup A_2 \cup D] \\ &= [B_1 \cup B_2 \cup B_3],\end{aligned}$$

and it is easy to check that $B \in F$. Furthermore

$$A_1 \subset B_1, \qquad A_2 \subset B_2, \qquad A_4 \supset B_4,$$

and $a \in A_3$ implies $a \in W$, $f(a) \in [T]$, hence

$$a \in [B_1 \cup B_2 \cup \{b \mid b \in B_3 \text{ and } L_T(f(b)) < L_T(f(a))\}].$$

In order to prove that $A \neq B$ we assume the contrary. Then $A_4 = B_4$, $W \cap A_4 = 0$, $R \cap A_4 = 0$, $R \subset A_3$, $A_3 \neq 0$. Picking $a \in A_3$, we have

$$a \in [B_1 \cup B_2 \cup (B_3 - \{a\})] = [A_1 \cup A_2 \cup (A_3 - \{a\})],$$

which is impossible because G is freely generated by $A_1 \cup A_2 \cup A_3 \cup A_4$.

Consequently $A \neq B$, $A < B$, contrary to the assumption that A is a maximal element of F.

6.4. THEOREM[13]. *If f is a homomorphism of the free group G onto the free group H and if*

$$G = G' * Z' * R, \qquad H = H' * K,$$

f maps G' isomorphically onto H',
f maps Z' onto the identity element of H,
then there exist subroups G'' and Z'' of G such that

$$G' \subset G'', \qquad Z' \subset Z'', \qquad G = G'' * Z'',$$

f maps G'' isomorphically onto H,
f maps Z'' onto the identity element of H.

Proof. Suppose

---

[13] See footnote 3.

$$G',\ Z',\ R,\ K$$

are freely generated by

$$X'\ U',\ V,\ W.$$

Then H' is freely generated by $f(X')$ and H is freely generated by the set

$$Y = f(X') \cup W.$$

Use Theorem 6.3 to obtain sets X'' and U'' with the properties listed there, and let

$$G'' = [X''], \qquad Z'' = [U''].$$

BROWN UNIVERSITY,
PROVIDENCE, R.I.