On Dehn's Algorithm

Roger C. Lyndon*

1. Introduction

The Word Problem for groups was formulated by Dehn in 1912, who gave a solution for the fundamental groups of two dimensional manifolds. Dehn's method was geometrical. (See Dehn [5, 6], also Reidemeister [17].) Related results have since been obtained by other methods, in particular by Tartakovskii [20—24], Schiek [18], Britton [2, 3], and Greendlinger [9—12]. Here we return to Dehn's geometrical method, extending it to obtain most of these more recent results, together with some results that are new.

Dehn's method is based on the Cayley *color diagram*, or *Gruppenbild*, M, associated with a presentation of a group G. This is a certain abstract two dimensional complex (or, what comes to the same in the present context, the one skeleton of the complex). Let $G = F/N$, with F free on a basis of generators $x_1, x_2, \ldots$, and N the normal closure in F of a set R of relators $r_1, r_2, \ldots$ The group diagram M is defined as follows. It has vertices $P(g)$ in one to one correspondence with the elements g of G. For each g in G and each $y = x_i^{\pm 1}$, it has an oriented edge E leading from $P(g)$ to $P(g\bar{y})$, where $\bar{y} = yN$; to E is attached the *label* $\phi(E) = y$, and to the oppositely oriented edge E^{-1} the label $\phi(E^{-1}) = y^{-1}$. The two cells D of M are as follows: for each closed path W in M, such that the labels on the successive edges of W are the successive letters of the reduced word for some r_i in R, we introduce a two cell D with boundary W.

The words w in the $x_i^{\pm 1}$ correspond one to one to the paths W in M with base point $P(1) = 0$, and w represents an element of N iff W is a closed path. If w is moreover reduced, then W is spanned by a possibly singular union of two cells D, together with spines attaching them to the base point 0. In the cases considered by Dehn, the group diagram M can be embedded in the plane. Then, if w is the reduced word for a non trivial element of N, the path W will contain a simple closed subpath W_0 at some point P_0, and W_0 will be the boundary $U^{\cdot}$, at P_0, of a simple finite union U of two cells D in the plane, each corresponding to some relator r_i or its inverse r_i^{-1}.

Now restrictions on the amount that two reduced words $r_i^{\pm 1}$ can have in common provide conditions on the amount of common boundary between two adjoining discs D in U. At this point Dehn used the fact that (apart from trivial cases) M can be represented as a regular tesselation of the hyperbolic plane with Poincaré metric, and inferred on metric grounds that the boundary W_0 of U must contain more than half the boundary $D^{\cdot}$ of some disc D. (Reide-

MEISTER [17] later showed that this conclusion can be obtained purely com-
binatorially, without appeal to a metric.) It follows immediately that the reduced
word w must contain more than half of some cyclic permutation of the reduced
word for some $r_i^{\pm 1}$.

The Word Problem asks whether arbitrary w in F belongs in fact to N.
If $w = 1$, then w is in N. Otherwise, given a non trivial reduced word w in F,
it can be decided by inspection whether w contains more than half of some
cyclically reduced conjugate s of an $r_i^{\pm 1}$. If not, then, by the above, w is not in N.
If so, say, $w = abc$ with $s = bd$, and d shorter than b, then w is in N iff the shorter
word $w' = ad^{-1}c$ is in N. A finite number of such reductions then leads to a
decision whether w is in N.

We depart from DEHN's method mainly in not requiring that the entire
group diagram M be embeddable in the plane. For quite general $G = F/N$, and
arbitrary w in N, there exists a finite complex M' in the plane, with boundary
described by a closed path on which the product of the labels is w, and with
interior a union of discs corresponding to relators, in exactly the same way as
in DEHN's case, where all of M can simultaneously be embedded in the plane.
Conditions on the amount of agreement between relators now lead as before
to conditions on common boundaries between discs, and to a solution of the
Word Problem. Indeed, by a modified combinatorial argument, we are able
to solve the Word Problem in additional cases where non trivial w in N need
not contain as much as half of any cyclically reduced conjugate of a relator.

In addition to the Word Problem, the present methods yield a proof that,
in the cases treated, the commutator quotient group $N/[N, N]$ of N, viewed as
a natural $G = F/N$ module, if not free, is at least a direct sum of cyclic sub-
modules, whence the cohomology of G, beginning with dimension 3, has period 2.

The present methods, with minor modification, apply equally to the case
that F is an arbitrary free product, rather than a free group. This case, already
considered by SCHIEK [18] and BRITTON [3], has bearing on the problem of
adjunction of solutions to equations over a group.

The main argument falls naturally into three stages. In Section 2 we discuss
finite complexes (graphs, or maps) in the plane. In Section 3, we associate plane
diagrams with elements of a free group. In Section 4, this connection is used to
translate properties of plane graphs into properties of groups: this contains
our main results, Theorems I, II, and III. A further Section 5 deals with free
products, with examples illustrating applications to the Adjunction Problem.

In addition to references cited explicitly, I want to acknowledge suggestions
from B. J. BIRCH, D. E. COHEN, H. SCHIEK, and H. ZIESCHANG.

2. Maps

A *map M* in the plane Π is determined by a finite set of points, called *vertices*,
and of simple arcs, called *edges*. Both ends of an edge are vertices (not necessarily
distinct), and an edge is not incident elsewhere with any vertex or any other
edge. A *region D* of a map M is a bounded component of the complement $\Pi - M$

of M in the plane Π; we consider only maps in which every region is simply connected.

The *degree* $d(P)$ of a vertex P, in M, is the number of edges incident at P, any edge with both ends at P being counted twice. The *degree* $d(D)$ of a region D is the number of edges on the boundary $D^{\cdot}$ of D, any edge such that D lies on both sides of it being counted twice.

The *boundary* $M^{\cdot}$ of M is the topological boundary of the unbounded component of the complement $\Pi - M$. A *boundary vertex* or *edge* is one contained in $M^{\cdot}$, and a *boundary region* D is one whose boundary $D^{\cdot}$ contains a boundary edge. Those not on the boundary are *interior vertices, edges*, and *regions*.

Our main results about maps can be viewed as combinatorial versions of two elementary metric properties of sufficiently regular domains in the Euclidean plane. The first is that the integral of curvature along the boundary $M^{\cdot}$ of M is 2π, and the second that the area of M does not exceed a constant $(1/4\pi)$ times the square of the length of the boundary $M^{\cdot}$. To introduce the ideas, we suppose that M is a union of regions in a regular tesselation T of the Euclidean plane. There are, of course, three types of regular tesselation, by hexagons, by squares, and by triangles, and our later results fall into three corresponding cases; but, for the precent heuristic remarks, we confine attention to a regular tesselation T by hexagons. We suppose, for the moment, that the vertices, edges, and regions of the map M belong to the regular tesselation T, and that $M^{\cdot}$ is a simple closed polygon, containing more than one hexagon from T.

To calculate curvature along $M^{\cdot}$, we pass to the closed polygon C (in the tesselation T^{*} dual to T) whose successive vertices are the centers D^{*} of the successive regions D lying on the boundary $M^{\cdot}$ of M. It is easy to see that, if $i(D)$ is the number of edges on $D^{\cdot}$ that are interior to M, then the change of direction in C at the point D^{*} is $[4 - i(D)]\dfrac{2\pi}{6}$. It follows that

$$\Sigma'[4 - i(D)]\frac{2\pi}{6} = 2\pi,$$

where the notation indicates summation over all boundary regions D of M, and hence that
$$\Sigma'[4 - i(D)] = 6.$$

In the map M, contained in T, we evidently have $d(D) = 6$ for each of the hexagonal regions D of M, and $d(P) = 3$ for every interior vertex P of M. We now pass to more general maps M, requiring only that M contain more than one region, that $d(D) \geqq 6$ for every region D of M, and that $d(P) \geqq 3$ for every interior vertex P of M. We shall see that, in this case, the equation above may be replaced by an inequality,
$$\Sigma'[4 - i(D)] \geqq 6.$$

An alternative formulation of this result refers to a map M^{*} *dual* to M. The combinatorial relation of M^{*} to M is determined uniquely by the condition that M^{*} has exactly one vertex D^{*} within each region D of M, and one edge E^{*} corresponding to each interior edge E of M, crossing E to join the vertices D_1^{*} and D_2^{*} of M^{*} contained in the regions D_1 and D_2 of M abutting along E. Note

that M^{**} is not isomorphic to M, but can be identified with the submap of M obtained by deleting all vertices on M' and all edges incident with such a vertex. Replacing M by M^*, the above hypotheses require that M contain more than one vertex, that $d(D) \geqq 3$ for all D, and that $d(P) \geqq 6$ for all interior vertices P. The conclusion then takes the form that

$$\Sigma' [4 - d(P)] \geqq 6 .$$

One immediate consequence of this is that, under the hypotheses, M' contains at least two convex points P, that is, with $d(P) \leqq 3$. A second is that there exists no map M on the sphere, in the obvious sense, with all $d(D) \geqq 6$ and all $d(P) \geqq 3$.

To relate area and length of boundary, we again suppose for the moment that M is part of the regular tesselation T. Then the area of M is proportionate to the number $a(M)$ of regions D of M, and the length of M' to the number of edges on M', that is, to $b(M) = \Sigma' [6 - i(D)]$. In this case it is obvious that $a(M)$ is bounded by a constant multiple of $b(M)^2$. It will be seen that for more general M, such that $d(D) \geqq 6$ for all D and that $d(P) \geqq 3$ for all interior P, we have indeed

$$a(M) \leqq \frac{1}{36} b(M)^2 .$$

We turn now to exact statements and proofs of these results. We begin with a rather general formula; the present method of proof was suggested by BIRCH. If M is any map in the plane, we write Q for the number of connected components of M, and Q_0 for the number that consist of a single isolated point. We define a generalization of the expression found above for curvature, depending on a real parameter t, by the formula

$$K(t) = \Sigma' [t - d(P)] ,$$

where summation is over all P on the boundary M' of M. As a measure of the excess of degree over t, on all interior vertices, we take

$$V(t) = \Sigma^{\circ} [d(P) - t] ,$$

sum over all interior vertices, P on $M - M'$. As a measure of excess degree of regions we take

$$R(t) = \Sigma [d(D) - t] ,$$

sum over all regions D of M. In these terms we state our first basic formula.

Lemma 2.1. *Let M be any map, and p and q real numbers such that $1/p + 1/q = 1/2$. Then*

$$K\left(\frac{p}{q} + 2\right) \geqq pQ + V(p) + \frac{p}{q} R(q) - \frac{p}{q} Q_0 .$$

We begin the proof with EULER's formula, in the form

(1) $$Q = v - e + d ,$$

where v, e, and d are the numbers of vertices, edges, and regions. It is immediate that

(2) $2e = \Sigma d(P)$, sum over all P in M, and

(3) $2e = \Sigma d(D) + d_\infty$, sum over all D in M, with d_∞ the number of edges $d(D_\infty)$ on the unbounded component of $\Pi - M$.

To eliminate e from these three equations we combine (2) multiplied by a real number m, and (3) multiplied by a real number n, with (1) multiplied by $2(m+n)$. Putting aside the cases that $m=0$, $n=0$, or $m+n=0$, we may normalize, taking $m=1$ and $n \neq 0$, -1. We obtain then

$$2(n+1)Q = 2(n+1)v - \Sigma d(P) + 2(n+1)d - n\Sigma d(D) - nd_\infty$$

(4)
$$= \Sigma[2(n+1) - d(P)] + n\Sigma\left[\frac{2(n+1)}{n} - d(D)\right] - nd_\infty.$$

Now the real solutions of $1/p + 1/q = 1/2$ are given parametrically by $p = 2(n+1)$ and $q = 2(n+1)/n$, for $n \neq 0$, -1. With these substitutions we have

$$pQ = \Sigma[p - d(P)] + \frac{p}{q} \Sigma[q - d(D)] - nd_\infty$$

$$= -V(p) + \Sigma[2(n+1) - d(P)] - \frac{p}{q} R(q) - nd_\infty.$$

To remove the term in d_∞ we note that the number v_∞ of vertices on M satisfies the relation

$$v_\infty \leqq d_\infty + Q_0,$$

since, in traversing M in the positive sense, each non isolated point of M is encountered as the beginning of at least one distinct edge on M. With this we have

$$pQ = -V(p) + \Sigma[n+2 - d(P)] - \frac{p}{q} R(q) + nv_\infty - nd_\infty \leqq$$

$$\leqq -V(p) + K\left(\frac{p}{q} + 2\right) - \frac{p}{q} R(q) + \frac{p}{q} Q_0.$$

This completes the proof.

The essence of this formula is contained in the case that M is a simple closed curve. Here we have $Q = 1$, $Q_0 = 0$, $v_\infty = d_\infty$, and obtain an exact equality

$$K\left(\frac{p}{q} + 2\right) = p + V(p) + \frac{p}{q} R(q),$$

which may be viewed as a combinatorial generalization of the Gauss-Bonnet formula for tesselations of the hyperbolic plane. Suppose further that $d(D) > 6$ for all D and $d(P) \geqq 3$ for all interior P. Taking $p = 3$ and $q = 6$ we have

$$K\left(\frac{p}{q} + 2\right) = \Sigma\left[\frac{5}{2} - d(P)\right] \leqq \Sigma\left[\frac{5}{2} - 2\right] = \frac{1}{2} b(M),$$

where $b(M)$ is the number of edges on M. We have $V(p) = \Sigma^0[d(P) - 3] \geqq 0$, and

$$\frac{p}{q} R(q) = \frac{1}{2} \Sigma[d(D) - 6] \geqq \frac{1}{2} \Sigma[7 - 6] = \frac{1}{2} a(M),$$

where $a(M)$ is the number of regions of M. From this we have $a(M) \leqq b(M) - 6$, a linear relation between area and length of boundary, which can also be inferred from DEHN's algorithm for this case.

We comment also that the proof of Lemma 2.1 indicates how the condition $1/p + 1/q = 1/2$ enters into our considerations, independent of any metric, Euclidean or otherwise. In the sequel, we shall always take p and q to be positive integers, hence (p, q) one of the pairs $(6, 3)$, $(4, 4)$, and $(3, 6)$, corresponding to the three Euclidean tesselations.

Corollary 2.2. *Let $(p, q) = (6, 3), (4, 4), or (3, 6)$. Let M be a map containing more than one vertex, and such that $d(D) \geqq q$ for every region D and that $d(P) \geqq p$ for every interior vertex P. Then, according to the three cases,*

$$\Sigma\,[4 - d(P)] \geqq 6, \qquad \Sigma\,[3 - d(P)] \geqq 4, \qquad or \qquad \Sigma\left[\frac{5}{2} - d(P)\right] \geqq 3.$$

Corollary 2.3. *Under the same hypotheses, the boundary $M^{\cdot}$ of M contains at least two points P such that, according to the case, $d(P) \leqq 3$, $d(P) \leqq 2$, or $d(P) \leqq 2$.*

Corollary 2.4. *Let $(p, q) = (6, 3), (4, 4), or (3, 6)$. Let M be a map containing more than one region, and such that $d(D) \geqq p$ for each interior region D and that $d(P) \geqq q$ for each interior vertex P. Then there are at least two regions D on the boundary $M^{\cdot}$ of M such that, according to the case, $D^{\cdot}$ contains no more than 3, 2, or 2 interior edges.*

Corollary 2.2 is an immediate consequence of Lemma 2.1, and Corollary 2.3 is an immediate consequence of Corollary 2.2. To prove Corollary 2.4, let M satisfy the given hypotheses, and let M^* be a map dual to M. Since M contains more that one region D, M^* contains at least two corresponding vertices D^*. Each region of M^* contains some interior vertex P of M, and has the same degree, $d(P) \geqq q$. Each interior vertex of M^* is the D^* contained in some interior region D of M, and has the same degree, $d(D) \geqq p$. Thus M^* satisfies the hypotheses, and so the conclusion, of Corollary 2.3. Finally, a vertex D^* of M^* with $d(D^*) \leqq 3$, 2, or 2, is contained in a region D of M with no more than 3, 2, or 2 interior edges.

We note that the case $(p, q) = (4, 4)$ of Corollary 2.4 is closely related to a theorem of REIDEMEISTER [17], page 201.

A map on the two dimensional sphere is defined exactly as a map on the plane, except that now every component of the complement is counted as a region.

Corollary 2.5. *Let $(p, q) = (6, 3), (4, 4), or (3, 6)$. Then there does not exist any map M on the sphere such that $d(D) \geqq p$ for all regions D and that $d(P) \geqq q$ for all vertices P.*

To see this, suppose such a map M on the sphere given. Deleting a point p_∞ of the sphere, not on M, we obtain a plane, and deleting from M the region D_∞ containing p_∞ we obtain a map M_0 on the plane. We may suppose D_∞ chosen to have no more than one edge in common with any other region D. By Corollary 2.4, M_0 contains a region D with no more than 3, 2, or 2 interior edges in M_0; then D, as a region of M, has degree $d(D) \leq 4$, 3, or 3. In the first two cases this contradicts the hypothesis on M. The third case follows from the first by duality: if M satisfied the hypotheses with $(p, q) = (3, 6)$, than a dual map M^* would satisfy the hypothesis with $(p, q) = (6, 3)$.

Although Corollary 2.5 must surely be long known, the nearest related result we can find is in BERGE [1], page 208.

We turn now to our remaining lemma on maps, relating area to length of boundary.

Lemma 2.6. *Let* $(p, q) = (6, 3), (4, 4), or (3, 6)$. *Let M be a map in the plane such that $d(D) \geqq q$ for each region D and that $d(P) \geqq p$ for each interior vertex P. Define*

$$a(M) = \text{number of vertices of } M,$$
$$b(M) = \Sigma^{\cdot}[p - d(P)], \text{ all } P \text{ on } M^{\cdot}.$$

Then

$$a(M) \leqq \frac{1}{K} b(M)^2$$

where, according to the case, K is 36, 16, or 6.

The proof is by induction on the number of vertices of M. In the initial case that M contains only a single point, and hence, by the condition $d(D) \geqq q$, no edges, the relation $Ka \leqq b^2$ evidently holds provided only that $K \leqq p^2$. But, it will be seen, the inductive step requires that $K \leqq 2p(p - 2)$. Now the given number 36, 16, or 6 is the smaller of p^2 and $2p(p - 2)$, according as p is 6, 4, or 3.

As a preliminary to the inductive step, suppose that M consists of a non empty part M_0, together with one further vertex P, either isolated or attached to M_0 by a single edge. Then $a = a_0 + 1$, with $a_0 \geqq 1$, while either $b = b_0 + p$ or $b = b_0 + p - 2$. Now, assuming inductively that $Ka_0 \leqq b_0^2$, and using the fact that $K \leqq 2p(p - 2)$ and that b_0 is an integer, it is easy to deduce that $Ka = Ka_0 + K \leqq (b_0 + p - 2)^2 \leqq b^2$. With this, we may assume henceforth that M contains more than one vertex, and that every vertex on $M^{\cdot}$ is incident with at least two edges on $M^{\cdot}$.

Now Corollary 2.2 applies to give $\Sigma^{\cdot}\left[\dfrac{p+2}{2} - d(P)\right] > p$. If we write $a^{\cdot}$ for the number of vertices on $M^{\cdot}$ and $d^{\cdot} = \Sigma^{\cdot} d(P)$, this inequality takes the form

$$(1) \qquad\qquad \frac{p+2}{2} a^{\cdot} - d^{\cdot} \geqq p.$$

Since $b = \Sigma^{\cdot}[p - d(P)] = pa^{\cdot} - d^{\cdot}$, form (1) we have

$$(2) \qquad\qquad b \geqq \frac{p-2}{2} a^{\cdot} + p.$$

Let $M_1 = M^{**}$, the map obtained from M by deleting all vertices on $M^{\cdot}$, together with the set F of all edges with one or both ends on $M^{\cdot}$. Let f_1 be the number of such edges with one end on $M^{\cdot}$, and f_2 of those with both ends on $M^{\cdot}$. Evidently

$$(2) \qquad\qquad d^{\cdot} = f_1 + 2f_2.$$

Since every vertex on $M^{\cdot}$ is incident with at least two edges on $M^{\cdot}$, we have $2a^{\cdot} \leqq 2f_2$, that is,

$$(4) \qquad\qquad a^{\cdot} \leqq f_2.$$

We now estimate $b_1 = b(M_1) = \Sigma_1' [p - d_1(P)]$, where the sum is over all P on M_1' and where $d_1(P)$ is the degree of P in M_1. First, if P is in M_1, then P is not on M', whence $d(P) \geqq p$. Second, from the definition of M_1, if P is in M_1, then $d_1(P) = d(P) - f(P)$, where $f(P)$ is the number of edges of the set F incident at P. Evidently $\Sigma_1' f(P) = f_1$. Now

$$b_1 = \Sigma_1' [p - d_1(P)] \leqq \Sigma_1' [d(P) - d_1(P)] = \Sigma_1' f(P) = f_1,$$

whence

$$b - b_1 = pa' - d' - f_1$$

and using successively (3), (4), and (1), we have

$$b - b_1 = pa' - 2d' + 2f_2 \geqq (p + 2)a' - 2d' \geqq 2p,$$

that is,

(5) $$b_1 \leqq b - 2p.$$

Assume inductively that $Ka_1 \leqq b_1^2$, where $K \leqq 2p(p - 2)$. Since $a = a' + a_1$, we have $Ka = Ka' + Ka_1 \leqq 2p(p - 2)a' + b_1^2$, and estimating a' and b_1 by (2) and (5) gives $Ka \leqq 4p(b - p) + (b - 2p)^2 = b^2$.

3. Diagrams

A *diagram*, over a group F, consists of a finite map M equipped with a function ϕ assigning to each oriented edge E as *label* a non trivial element $\phi(E)$ of F, and assigning the inverse element to the oppositely oriented edge E^{-1}. We shall associate a diagram M with each finite sequence $p_1, \ldots, p_n$ of non trivial elements from F, a free group with given basis. Note that each p_i has a unique representation $p_i = u_i r_i u_i^{-1}$ in reduced form, where r_i is cyclically reduced.

A diagram M will be called a *diagram for the sequence* $p_1, \ldots, p_n$ if it satisfies the following two conditions.

(1) The boundary M' is connected and contains a vertex 0 such that, if $E_1, \ldots, E_k$ are the successive oriented edges traversed in running around M' from 0 back to 0 in the positive sense, then the product of the labels $\phi(E_1), \ldots, \phi(E_k)$ is reduced without cancellation, and is the reduced word for $w = p_1 \ldots p_n$.

(2) If D is any region of M, then its boundary D' contains a point P such that, if $E_1, \ldots, E_k$ are the successive oriented edges traversed in running around D' from P back to P in the positive sense, then the product of the labels $\phi(E_1), \ldots, \phi(E_k)$ is reduced without cancellation, and is the reduced word for one of $r_1, \ldots, r_n$.

Lemma 3.1. *There exists a diagram for every sequence.*

We shall construct, for each sequence, a canonical diagram.

If $n = 0$, and the sequence is empty, we take M to consist of a single point 0.

Let $n = 1$, and $w = p_1 = u_1 r_1 u_1^{-1}$. We draw a loop, consisting of a single edge E with both ends at the same vertex P, and label E, in the positive sense, with r_1. If $u_1 = 1$, we take $0 = P$ and are done. If $u_1 \neq 1$, we connect an external point 0 to P with an edge E', with label $\phi(E') = u_1$.

Let $n > 1$. As in the case $n = 1$, we construct diagrams $M_1, \ldots, M_n$ for the single factors $p_1, \ldots, p_n$. We may take these with common base point 0, each external to the rest, and arranged in order about the point 0. If M is the resulting diagram, then (2) is already satisfied. Moreover, M is connected and contains 0, and the product of the labels on M from 0 back to 0 (beginning with the first edge of M_1) is equal to $w = p_1 \ldots p_n$. However, this product need not be reduced without cancellation. We shall show how to modify a diagram M, satisfying the conditions just stated, until ultimately (1) is satisfied.

For simplicity, if an edge E bears a label with reduced form $v = y_1, \ldots, y_k$, each y_i a generator or the inverse of a generator, we divide E into successive edges $E_1, \ldots, E_k$ with labels $y_1, \ldots, y_k$. Thus we may suppose each label is a generator or the inverse of a generator. Let α be an arc, from 0 to 0, describing M, and such that the product of the labels on the successive edges of α is w. If (1) is not already satisfied, then α will have the form $\alpha = \beta E_1 E_2 \gamma$, where E_1 and E_2 are consecutive edges bearing inverse labels. Let E_1 run from P_1 to P_0 and E_2 from P_0 to P_2. Suppose first that P_1 is distinct from P_0 and P_2. Then, deforming that part of M lying in a small neighborhood of E_1, and keeping P_0 fixed, we can pivot $E_1 = P_1 P_0$ about P_0 in the exterior of M until it comes into coincidence with $E_2 = P_2 P_0$. Then P_1 and P_2 coincide, as do the oriented edges E_1 and E_2^{-1}. In resulting M' the arc α will be replaced by $\alpha' = \beta \gamma$, containing fewer edges than α. If P_2 is distinct from P_0 and P_1, we proceed similarly. The case remains that $P_1 = P_2$. In this case the part $E_1 E^2$ of α is a loop, attached to the rest of α only at the point P_0. We form M' by deleting this loop, together with all of M interior to the loop; again α is replaced by $\alpha' = \beta \gamma$. In all cases, M' satisfies the same hypotheses as M, whence, by iteration, we must ultimately arrive at some M'' satisfying (1) as well as (2).

Suppose that, in obtaining the reduced form for the product $w = p_1 \ldots p_n$, there is cancellation between *corresponding parts* of some p_i and p_j, for $1 \leq i < j \leq n$, in the following sense. First, r_i and r_j have reduced forms $r_i = axb$ and $r_j = cx^{-1}d$, where the parts x and x^{-1} correspond in the sense that cyclic conjugates xba and $xc^{-1}d^{-1}$ of r_i and r_j^{-1} are equal, hence $ba = c^{-1}d^{-1}$. Second, the corresponding parts cancel, that is, $xbu_i^{-1}qu_jcx^{-1} = 1$ where $q = p_{i+1} \ldots p_{j-1}$. Then we have $p_i q p_j = u_i a d u_j^{-1}$, and since $ad = b^{-1}c^{-1}$ and $q = u_i b^{-1} c^{-1} u_j^{-1}$, we have $p_i q p_j = q$. It follows that $w = p_1 \ldots p_{i-1} p_{i+1} \ldots p_{j-1} \times \times p_{j+1} \ldots p_n$.

We shall call a sequence $p_1, \ldots, p_n$ *reduced* if, for no i and j, where $1 \leq i < < j \leq n$, do we have $p_i \ldots p_j = p_{i+1} \ldots p_{j-1}$. Then if w is the product of a sequence, it is the product of some reduced subsequence.

We note that reduced sequences have been considered by Greendlinger [9], page 68, and by Gladkii [8].

A diagram M will be called *reduced* if it contains no interior edge E, separating regions D_1 and D_2, such that the product of the labels on the boundary D of the combined region $D = D_1 \cup D_2 \cup E$ reduces to 1. Suppose the canonical diagram M, constructed above, for a sequence $p_1, \ldots, p_n$ is not reduced. For E, D_1 and D_2 as above we have D_1' and D_2' described by arcs $E \delta_1$ and

$E^{-1}\delta_2$, where the product of the labels on the arc $\delta_1\delta_2$ describing $D^{\cdot}$ is 1. Now E must have entered M through identification of edges E_1 from a certain M_i and E_2^{-1} from a certain M_j, where we may suppose $i < j$. Thus M_i contains D_1 with $D_1^{\cdot}$ described by $E_1\delta_1$ where $r_i = axb$, for $\phi(E_1) = x$, and ba the product of the labels on δ_1. Similarly, $r_j \equiv cx^{-1}d$, with dc the product of the labels on δ_2. The product of the labels $\delta_1\delta_2$ is then $badc = 1$. Moreover, in order that E_1 and E_2 should become successive edges, the product of the intervening edges on $M_1 \cup \cdots \cup M_n$ must have reduced to 1, that is, $bu_i^{-1}p_{i+1}\ldots p_{j-1}u_jc = 1$. But this implies that the sequence $p_1, \ldots, p_n$ is not reduced. We have proved the following.

Lemma 3.2. *Every reduced sequence possesses a reduced diagram.*

In fact, if $p_1, \ldots, p_n$ is any sequence, it contains a reduced subsequence with the same product w, and any reduced diagram for the subsequence will be a reduced diagram for the original sequence. One can, alternatively, obtained a reduced diagram for any sequence by successively shrinking to a slit any combined region D as above with the product of the labels on $D^{\cdot}$ equal to 1. This can be done as well in the case that D is the union of more than two regions, together with their separating edges.

We next examine the possibility of the loss of some of the n regions of $M_1 \cup \cdots \cup M_n$ in the course of constructing the canonical diagram M for a reduced sequence $p_1, \ldots, p_n$. This can happen only in case, in the construction of M, we encounter a loop E_1E_2 on the boundary. Let L_0 be the part of the diagram consisting of this loop and everything interior to it. Completing the plane to a sphere by adjoining a point p_∞, we can flatten the new region D_∞ containing p_∞ to a slit, bringing the two edges E_1 and E_2^{-1} into coincidence. We then have a diagram L on the sphere which satisfies (2). Now L will be a reduced diagram unless identification of E_1 and E_2^{-1} violates the condition for a reduced diagram. But then exactly the same argument as before would show that the labels x on E_1 and x^{-1} on E_2 were corresponding parts of certain r_i and r_j, which could be cancelled in reducing the product $p_1 \ldots p_n$. This would contradict the assumption that $p_1, \ldots, p_n$ is reduced. This proves the following.

Lemma 3.3. *If $p_1, \ldots, p_n$ is a reduced sequence, then there exists a reduced diagram M for this sequence, together with some number $q \geqq 0$ of reduced diagrams $L_1, \ldots, L_q$ on spheres, such that these diagrams contain in all exactly n regions.*

A subdiagram M_0 of a reduced diagram M for $p_1, \ldots, p_n$ will be called *critical* if either (i) M_0 contains a single region D, and a closed arc bounding D is part of a closed arc at 0 bounding M, or else if (ii) M_0 contains more than one region, and with the possible exception of one region D_0, for every D in M_0, such that $D^{\cdot} \cap M_0^{\cdot}$ contains more than a single point, $D^{\cdot} \cap M_0^{\cdot}$ is a simple arc and a part of a closed arc at 0 bounding M.

Lemma 3.4. *If $p_1, \ldots, p_n$ is a reduced sequence, and $p_1 \ldots p_n \neq 1$, then the sequence possesses a critical subdiagram.*

Let M be a reduced diagram. From (1) it follows, since $p_1 \ldots p_n \neq 1$, that $M^{\cdot}$ consists of more than the single point 0, and that $M^{\cdot}$ contains no spines, that is, successive edges E and E^{-1}, except possibly at 0. It follows that, if α is an arc

with base point 0 that describes M', then α must contain a simple loop β at some base point P. Let M_0 be the subdiagram of M consisting of β together with all of M contained within β; then β describes M'_0, and M_0 has at most the point P in common with the remainder of M. If M_0 contains a single region D, then β describes D', and, since β is part of α, M_0 is critical. Suppose M_0 contains more than one region. Choose a region D_0 of M_0 such that P lies on D'_0. If D is a region of M_0 other than D_0, and if $D' \cap M'$ is connected and consists of more than one point, then $D' \cap M'$ is an arc δ, not containing P in its interior, and hence is a simple subarc of β and so of α. Thus again M_0 is critical provided every $D' \cap M'_0$ is connected.

Suppose D_1 is a region of M_0 such that $D'_1 \cap M'_0$ is not connected. Then removing D'_1 separates M_0 into more than one component, and we may choose a component K not containing P. Let $M_1 = K \cup D'_1$, and suppose first D_1 and then K chosen to make the number of regions in M_1 as small as possible. Then $D' \cap M'_1$ is connected for all D in M_1. Moreover, M_1 contains some region other than D_1, and if D is such a region, and $D' \cap M'_1$ contains more than one point, it is an arc not containing P in its interior, and hence a simple subarc of α. Therefore M_1 is critical. This establishes Lemma 3.4.

It is perhaps worth commenting on the relation between the diagram M associated with a sequence $p_1, \ldots, p_n$ and the full group diagram $\overline{M}$ for the presentation determined by the relators $r_1, \ldots, r_n$. An incidence preserving map ψ from M into $\overline{M}$ is determined by mapping the vertex 0 of M into the vertex $P(1)$ of $\overline{M}$, and each edge E of M with starting point P and label $y = x_i^{\pm 1}$ into the unique edge ψE in $\overline{M}$ with starting point ψP and label y. Moreover, this map ψ can be extended to map the regions of M onto two cells of $\overline{M}$. The map ψ need not be one to one: if p_1, p_2 has the form r_1, r_1, then M consists of two loops bounding disjoint regions at 0, which doubly cover a loop at $P(1)$ in $\overline{M}$ bounding a single two cell. The possibility of embedding M in the plane, and the consequent multiplicity of ψ, results from the fact that, in constructing M, we have not taken into account all coincidences between parts of the $p_i^{\pm 1}$, but only those that play a role in cancellations leading to the reduced form of the product $w = p_1 \ldots p_n$.

4. Free groups

We now suppose that F is a free group, R a set of non trivial elements of F, and N the normal closure of R in F. If w is in N, then w can be written as a product $w = p_1 \ldots p_n$, for some $n \geqq 0$, of conjugates p_i of elements $r^{\pm 1}$ for r in R; indeed, each $p_i = u_i r_i u_i^{-1}$ reduced, where each r_i is a cyclically reduced conjugate of some r in R. Moreover, we can suppose the sequence $p_1, \ldots, p_n$ reduced. If M is a reduced diagram for this sequence, we shall impose conditions on R which imply conditions on the degrees of interior regions and vertices of M, and hence, by our lemmas on maps, imply conditions on M'.

First, it is convenient to suppose R *symmetrized* in the sense that every element r of R is cyclically reduced and that, if r is in R, then every cyclically reduced conjugate of r or r^{-1} is in R. Second, we can without loss delete from

M every vertex of degree 2, except possibly 0, with the result that $d(P) \geqq 3$ for each interior vertex P.

Define a *piece* to be an element x of F such that R contains two distinct elements with reduced forms $r = xa$ and $s = xb$. Suppose that E is an interior edge of M, separating regions D_1 and D_2. If x is the label on E, then $D_1^{\cdot}$ and $D_2^{\cdot}$ will have labels, starting at the starting point of E, which are elements of R of reduced forms $r_1 = xa$ and $r_2 = bx^{-1}$. Since the label on $D^{\cdot}$, for $D = D_1 \cup D_2 \cup E$, is ab, and M is reduced, $ab \neq 1$, and R contains distinct elements with reduced forms xa and xb^{-1}. It follows that every label on an interior edge is a piece.

If D is an interior region of M, and $d(D) = k$, then $D^{\cdot}$ is described by an arc $E_1, \ldots, E_k$ where each E_i is an interior edge, and $D^{\cdot}$ has a label r in R with reduced form $r = x_1 \ldots x_k$, where each x_i is a piece. Consequently the condition that $d(D) \geqq p$ for every interior region D will follow from the following condition on R:

$C(p)$:   *No element of R is a product of fewer than p pieces.*

Case i. Assume that R satisfies the condition $C(6)$. Then $d(D) \geqq 6$ for each interior region D. Moreover, $d(P) \geqq 3$ for each interior vertex. Suppose $w \neq 1$, whence M contains a critical subdiagram M_0 satisfying the same conditions on interior degrees. If M_0 consists of a single region D, then some simple closed arc δ bounding D is a subarc of the arc α describing $M^{\cdot}$ at 0, whence the label r in R on δ is a part of the reduced form of the label w on α. If M_0 contains more than one region, then, by Corollary 2.3, it contains at least two boundary regions D with at most 3 interior edges each. Taking one of these D distinct from the exceptional region D_0, we have that $D^{\cdot} \cap M_0^{\cdot}$ is a simple arc δ, part of α. Now $D^{\cdot}$ is described by an arc $\delta E_1 \ldots E_k$, where $k \leqq 3$ and the E_i are interior edges. Then R contains an element r with reduced form $r = ax_1 \ldots x_k$, $k \leqq 3$, where a is the label on δ and $x_1, \ldots, x_k$ are pieces. Moreover, a is a part of the reduced word w. In this sense, w contains some r from R with (at most) 3 pieces missing. (We can delete 'at most', since 1 is always a piece.)

Case ii. Assume $C(4)$, whence $d(D) \geqq 4$ of each interior region D. If we can ensure that $d(P) \geqq 4$ for each interior vertex P, then we can conclude, exactly as in Case i, that each non trivial element w in N contains some element r from R with two pieces missing. It remains to formulate a condition on R excluding interior vertices of degree 3. Suppose that P were such a vertex, with $E_1, D_1, E_2, D_2, E_3, D_3$ in order the edges leading into P and the included regions. The boundary of D_i, starting at P, then has label with reduced form $r_i = x_i^{-1} a_i x_{i+1}$ in R, where i is taken modulo 3 and x_i is the label on E_i. Thus there is cancellation in each of the products $r_1 r_2$, $r_2 r_3$, and $r_3 r_1$. To exclude such vertices we impose on R the following condition:

T_3:   *If r_1, r_2, and r_3 are in R, then in at least one of the products $r_1 r_2$, $r_2 r_3$, $r_3 r_1$ there is no cancellation.*

Case iii. This case is of very limited interest, since the hypotheses needed for our argument seem both artificial and very restrictive; we know of no

natural application, although we shall give later a natural, if special, application of the analógous case for free products. In fact, the condition $C(3)$ is very mild; it can be given the simpler form that, if r_1, r_2, and r_3 are elements of R, with $r_1 r_2 \neq 1$ and $r_2 r_3 \neq 1$, then r_2 does not cancel entirely in the product $r_1 r_2 r_3$. The condition that all interior vertices have degree $d(P) \geq 6$ is, on the other hand, very strong. We can ensure this by assuming T_3 together with analogous T_4 and T_5. In fact, since M is reduced, we can modify each of these conditions T_k to T'_k by inserting the additional hypothesis that no $r_i r_{i+1} = 1$. It is easy to see that T'_3 is in fact equivalent to T_3, and that, in the presence of T_3, T'_5 is equivalent to T_5. However T'_4 is significantly weaker than T_4. Even so, if each r in R has length at·least 4, T'_4 cannot be realized in any case that does not already fall under Case ii. The best we can do to improve this is to replace T_4 by the following:

T''_4:   If $r_1, \ldots, r_5$ are elements of R, with each $r_i r_{i+1} \neq 1$, then there exists a in F such that each $r'_i = a r a^{-1}$ is in R, and in at least one of the products $r'_i r'_{i+1}$ there is no cancellation.

For, it can be shown by an argument similar to that we shall use in the next section, if M contains an interior vertex P of degree 4, assuming T''_4 we can modify M to eliminate this vertex. In all, under hypotheses $C(3)$, T_3, T''_4, and T_5, we can conclude as before that each non trivial element of N contains some r in R with two pieces missing.

We have proved the following theorem.

Theorem I. *Let F be a free group, R a symmetrized subset of F, N the normal closure of R in F, and w a non trivial element of N. Under the additional hypothesis*

(i) $C(6)$,

it follows that w contains some r from R with three pieces missing : w and r have reduced forms $w = bac$ and $r = a x_1 x_2 x_3$, where x_1, x_2, and x_3 are pieces.

Under either of the hypotheses

(ii) $C(4)$ and T_3,

or

(iii) $C(3)$, T_3, T''_4,   and T_5,

it follows that w contains some r from R with two pieces missing.

With each positive real number λ we associate the following condition:

$C'(\lambda)$:   *If r in R has reduced form $r = ab$ where a is a piece, then a has length $|a| < \lambda |r|$.*

If $\lambda \leq \dfrac{1}{p-1}$, then r cannot be the product of as few as $p-1$ pieces, and $C(p)$ holds. Thus Theorem I has the following corollary.

Corollary 4.1. *Let F, R, N, and w be as before. If we have*

(i) $C'(\lambda)$ *for some* $\lambda \leq \dfrac{1}{5}$,

then w contains a part a of some r in R, with $|a| > (1 - 3\lambda)|r|$. If we have either

(ii) $C'(\lambda)$,   *for some* $\lambda \leq \dfrac{1}{3}$,   *and* T_3,

or

(iii) $C'(\lambda)$,   *for some* $\lambda \leqq \dfrac{1}{2}$,   *and* T_3, T_4'',   *and* T_5,

then w contains a part a of some r in R, with $|a| > (1 - 2\lambda)|r|$.

Most of these results are known. Both SCHIEK [18] and GREENDLINGER [10] have stated results in terms of pieces missing. Case i of the corollary, for $\lambda = 1/6$, was stated explicitly by GREENDLINGER [9]; earlier TARTAKOVSKII [22] had obtained a closely related result, and BRITTON [3] had obtained the analogous result with the free group F replaced by a free product. It is mentioned in a review by BOONE [2] that BRITTON had conjectured Case i with $\lambda = 1/5$, and BRITTON now has a proof by his methods for this case. Case ii, for $\lambda = 1/4$, was stated explicitly by GREENDLINGER [11]; earlier a related result had been given by SCHIEK [18], who introduced the condition T_3, and BRITTON [3] appears to allude to a similar result.

DEHN's argument gives an immediate solution of the Word Problem for R finite whenever every non trivial element of N contains more than half of some element from R, in particular, under Case i of the corollary provided $\lambda \leqq 1/6$, or Case ii provided $\lambda \leqq 1/4$. Under a more general hypothesis, a different argument is needed.

Theorem II. *Let F be a free group, R a finite symmetrized subset of F, and N the normal closure of R in F. Assume one of the three hypotheses:*

(i) $C(6)$,

(ii) $C(4)$ *and* T_3,

(iii) $C(3)$, T_3, T_4'',   *and* T_5.

Then there exists an algorithm for deciding whether arbitrary w in F belongs to N.

We show first that if w is in N, hence $w = p_1 \ldots p_n$ for some $n \geqq 0$ and $p_i = u_i r_i u_i^{-1}$ reduced, r_i in R, then some such sequence $p_1, \ldots, p_n$ has a diagram M with no more than $p^2 |w|^2$ vertices, where $p = 6, 4, 3$ according to the case. Indeed, let $p_1, \ldots, p_n$ be a reduced such that $p_1 \ldots p_n = w$, and let M be the associated reduced diagram. If a closed arc α at 0 describing M has k edges, then w is the product without cancellation of the k non trivial labels on these edges, whence $k \leqq |w|$. Each vertex on M occurs at least once as the starting vertex of some edge in α, whence the number a of vertices on M is no greater than k, that is, $a \leqq k$. The length of M, in the sense of Lemma 2.6, is given by $b = \Sigma [p - d(P)]$, whence it follows that $b = \Sigma [p - d(P)] \leqq \Sigma p = pa \leqq pk \leqq p|w|$. Now Lemma 2.6 gives $a \leqq b^2 \leqq p^2 |w|^2$.

Now suppose arbitrary w in F given. If w is in N, it must be the product of a sequence with a reduced diagram with no more than $p^2 |w|^2$ vertices. There are, within isomorphism, only finitely many unlabeled maps with no more than this many vertices, whence it will suffice, given w in F, to decide whether a given such map can be labeled to yield a diagram of the required sort. Now a label on an interior edge is a piece, and, since R is finite, there are only finitely many pieces. A label on a boundary edge must be part of the reduced word w,

and there are only finitely many such parts. Thus, given a map, there are only finitely many labelings of its edges to consider, and it suffices to decide whether a given labeled map in fact constitutes a diagram for a sequence $p_1, \ldots, p_n$ of conjugates of elements r_i in R, with $p_1 \ldots p_n = w$. But this last equation can clearly be settled by inspection.

We state now the last of our three main results.

Theorem III. *Let F be a free group, R a symmetrized subset of F, and N the normal closure of R in F. The operation of F on N by conjugation induces on the quotient group $N_0 = N/[N,N]$ of N by its derived group the structure of an F/N module. Assume one of the following three hypotheses:*

(i) $C(6)$,

(ii) $C(4)$ *and* T_3,

(iii) $C(3)$, T_3, T_4'',   *and* T_5.

Then the F/N module N_0 is the direct sum of cyclic submodules.

For each r in R, the coset $r[N, N]$ in N_0 generates a cyclic submodule, containing all $p[N, N]$ for p conjugate to r; clearly conjugate elements r of R generate the same submodule. We shall show that N_0 is the direct sum of these cyclic submodules. This amounts to showing that if $w = p_1 \ldots p_n$ is in $[N, N]$, where each p_i is a conjugate of some element from R, then the factors p_i fall into pairs p_i and p_j, with $p_i = q p_j^{-1} q^{-1}$ for some q in N. Since a commutator $p_i^{-1} p_j^{-1} p_i p_j$ of conjugates of elements of R obviously has this pairing property, it suffices, by an induction on n, to show that if $p_1 \ldots p_n = 1$, and $n > 0$, then there is one pair of factors p_i and p_j of the required sort. We shall show more, that $p_1 \ldots p_n = 1$ implies the sequence $p_1, \ldots, p_n$ is not reduced.

Suppose, $w = p_1 \ldots p_n$, $n > 0$, each p_i conjugate to some element of R, and that the sequence $p_1, \ldots, p_n$ is reduced. By Lemma 3.3, there exists a reduced diagram M for this sequence, together with some number $q \geqq 0$ of reduced diagrams $L_1, \ldots, L_q$ on the sphere, containing in all n regions. By Corollary 2.5, there cannot exist any map L on the sphere satisfying those conditions on degrees that follow from the hypotheses of Theorem II. Therefore $q = 0$ and M contains n regions. But then M, and so also M', consists of more than a single point. It follows that $w \neq 1$.

In the cases where one can conclude that every non trivial element of N contains more than half of some element from R, COHEN and LYNDON [4] derived the conclusion of Theorem III from a stronger result, that N has a free basis consisting of conjugates of elements from R. It is not clear whether this stronger result holds in the more general situation considered in Theorem III.

5. Free products

We now suppose that F is no longer a free group, but the free product of a given family of subgroups F_i, with distinct F_i having trivial intersection. We shall again take R to be a subset of F and N the normal closure of R in F. The construction of diagrams and the proofs of Theorems I, II, and III then go

over with very little change. We shall confine ourselves to the proof of an analog of Theorem I, and to giving two simple applications.

Every element w of F now has a unique *normal form*, $w = w_1 \ldots w_n$, $n \geqq 0$, where each of the *letters* w_i is a non trivial element of some component F_j, and adjacent w_i, w_{i+1} do not belong to the same F_j. The length of such w is $|w| = n$. Given two elements with normal forms $u = u_1 \ldots u_m$ and $v = v_1 \ldots v_n$, in obtaining the normal form for the product $w = uv$ there may be cancellation of a certain number of letters from the end of u against their inverses, at the beginning of v; it may also happen that the last remaining letter u_i of u is in the same F_k as the first remaining letter v_j of v, in which case they are consolidated to give a single letter $u_i v_j = w_i$ in the normal form for w. We shall speak of w as having *reduced form* $w = uv$ if there is no cancellation in forming the product; we do not exclude the possibility of consolidation. It should be emphasized that for a to be a part of w with reduced form $w = bac$ does not require that the letters of a should be a subsequence of those of w.

We shall call an element r *cyclically reduced* if it has normal form $r = r_1 \ldots r_n$ where $r_n r_1 \neq 1$; if r_1 and r_n belong to different F_j, we call r *strictly cyclically reduced*. A set R will be called *symmetrized* if every r in R is cyclically reduced and if, together with each r in R, every cyclically reduced conjugate of r on r^{-1} also belongs to R. The definition of a *piece* is modified as follows: x is a piece if R contains elements with reduced forms axb and cxd where $xba \neq xdc$.

We turn now to the construction of a diagram M associated with a sequence $p_1, \ldots, p_n$. In all diagrams considered, the vertices will fall into two classes, primary and secondary, and the two ends of every edge will fall into different classes. The label on every edge will belong to some F_i, with the labels on two edges meeting successively at a primary vertex belonging to different F_i, and those on edges meeting at a secondary vertex belonging to the same F_j. If $E_1, \ldots, E_k$ are in order the edges running into a secondary vertex P, then we shall want to specify the values of the labels $\phi(E_i^{-1} E_{i+1}) = \phi(E_i)^{-1} \phi(E_{i+1})$, but we shall not be concerned with the particular choice of the $\phi(E_i) = x_i$. For the x_i we shall then have a system of equations $x_i^{-1} x_{i+1} = \phi(E_i^{-1} E_{i+1}) = a_i$, in F_j, for i modulo k, and of inequations $x_i \neq 1$. We shall have always the compatibility condition $a_1 \ldots a_k = 1$; with this one of the equations $x_i^{-1} x_{i+1} \equiv a_i$ becomes redundant, and the rest can be solved to express the remaining x_i in terms of x_1. We are then left with k inequations $x_1 \neq b_1, \ldots, b_k$. Assuming, as we may in this context, that each F_j is infinite, a solution can always be found.

With a sequence of length $n = 0$ we associate as before a diagram M consisting of the single (primary) point O. Let $n = 1$; then p_1 has normal form $p_1 = a_1 \ldots a_k b_1 \ldots b_m a_k^{-1} \ldots a_1^{-1}$ where $k \geqq 0$, $m \geqq 1$, and $b_m b_1 \neq 1$. If b_1 and b_m are in different F_j we take M to be a loop at P joined to O by a path. The path OP consists of $2k$ edges $E_1, E_1', \ldots, E_k, E_k'$, with each $\phi(E_i E_i') = a_i$. The loop at P consists of $2m$ edges $B_1, B_1', \ldots, B_m, B_m'$, with each $\phi(B_i B_i') = b_i$. If b_1 and b_2 lie in the same F_j, we take OP to end with an additional edge E, and the loop at P to consist of edges $B, B_2, B_2', \ldots, B_{m-1}, B_{m-1}', C$, with $\phi(B_i B_i') = b_i$ as before for $2 \leqq i \leqq m - 1$. The three edges E, B, and C at the secondary vertex P are to be

labeled in such a way as to satisfy the compatible conditions $\phi(EB) = a_1$, $\phi(CE^{-1}) = a_m$ and $\phi(CB) = a_m a_1$.

The argument by induction for $n > 1$ follows the same plan as before. Suppose at some stage we have successive edges E, E', G, and G', where E and G are separated by a primary vertex P, and yet $\phi(EE') = a$ and $\phi(GG') = b$ lie in the same F_j. If $ab = 1$, we either delete a loop bounded by these four edges, or else, as before, we bring E' into coincidence with G^{-1} (their secondary vertices cannot coincide) and E with G'^{-1}. If $ab \neq 1$, we bring E' into coincidence with G^{-1}, leaving on the boundary consecutive edges to be labeled $\phi(EG') = ab$. In either case it is necessary, but possible, to readjust the labels on the four edges involved.

With this, we obtain analogs of everything in Section 3.

To begin with an analog, Theorem IV, of Theorem I, we again assume a condition $C(p)$ to ensure that every interior region D of a reduced diagram M has degree $d(D) \geq p$. As before, we can also suppose that every interior vertex P has degree $d(P) \geq 3$. For the case that $p = 6$, no more is required.

In *Case ii*, assuming $C(4)$, we again need a condition to exclude interior vertices of degree 4. To exclude such primary vertices, we require the following variant of T_3:

T_3^1:   *If r_1, r_2, and r_3 are strictly cyclically reduced elements of R, then, in at least one of the products $r_1 r_2$, $r_2 r_3$, $r_3 r_1$, there is non cancellation.*

At a secondary vertex P of degree 3, with successive edges E_1, E_2, and E_3 leading into P, let x_1, x_2, x_3 be the labels on these edges. Then the three regions D_i meeting at P will bear on their boundaries labels r_i in R, strictly cyclically reduced, where r_i contains a letter $a_i = x_{i+1} x_i^{-1}$. Thus $a_3 a_2 a_1 = 1$. To exclude this, we require the following condition:

T_3^2:   *If a_1, a_2, and a_3 are letters occurring in strictly cyclically reduced elements of R, then $a_1 a_2 a_3 \neq 1$.*

This suffices for the case $p = 4$.

The remaining *Case iii* is again awkward. We again assume $C(3)$, and also T_3^1, T_4^1, and T_5^1, obtained from T_3, T_4'', and T_5 by restricting consideration to strictly cyclically reduced r_i in R. These imply the condition $d(D) \geq 3$ on interior regions D, and the condition $d(P) \geq 6$ on interior primary vertices. For interior secondary vertices, T_3^2 together with analogous T_4^2 and T_5^2 would suffice, but are needlessly strong. If we restrict these conditions to the case where all $a_i a_{i+1} \neq 1$, the content of T_3^2 and T_5^2 taken together is not altered, but we obtain T_4^2 in weaker form:

T_4^2:   *If a_1, a_2, a_3, a_4 are letters occuring in strictly cyclically reduced elements of R, and $a_1 a_2$, $a_2 a_3$, $a_3 a_4 \neq 1$, then $a_1 a_2 a_3 a_4 \neq 1$.*

This condition does not in fact exclude that the diagram M should contain an interior secondary vertex P of degree $d(P) = 4$, but does ensure that M can be modified in such a way as to eliminate all such vertices. Suppose then that P is such a vertex, with four edges $E_i = P_i P$ running into P, in order, and with

regions D_i lying between E_i and E_{i+1}, where i is taken modulo 4. Since P is secondary, the labels $x_i = \phi(E_i)$ all lie in the same F_j, and the $a_i = x_{i+1}x_i^{-1}$ satisfy $a_4 a_3 a_2 a_1 = 1$. By T_4^2 and symmetry we can suppose $a_2 a_1 = 1$, hence $x_1 = x_3$. In the diagram $K = D_1 \cup \cdots \cup D_4$, we think of the pair of edges $E_1 = P_1 P$ and $E_3^{-1} = PP_3$ as obtained by flattening a rhombus $P_1 P^2 P_3 P^4$ onto a diagonal in such a way that $P^2 = P^4 = P$. We obtain K' by flattening this rhombus onto its other diagonal, so that P_1 and P_3 coincide at a point Q. Since K' has the same boundary as K, we can form M' from M by replacing the part K by K'. Now M' does not contain the vertex P, but contains three vertices that did not appear in M. Two of these, P^2 and P^4, are of degree 2, and can be deleted, while the third, Q, is a primary vertex. Since M' contains fewer interior vertices of degree 4 than M, by iteration of this process we can eliminate all such vertices.

With this, we have proved an analog of Theorem I.

Theorem IV. *Let F be a free product, R a symmetrized subset of F, and w, a non trivial element in the normal closure N of R in F. Assuming*

(i) $C(6)$

we conclude that $w = bac$ reduced, where $r = ax_1 x_2 x_3$ reduced, for some r in R and pieces x_1, x_2, x_3.

Assuming either

(ii) $C(4)$, T_3^1 *and* T_3^2,

or

(iii) $C(3)$ *and* T_i^1, T_i^2, $i = 3,\ 4,\ 5$,

we conclude that $w = bac$ reduced, where $r = ax_1 x_2$ reduced, for some r in R and pieces x_1 and x_2.

If we replace the hypothesis $C(p)$ by $C'(\lambda)$, for $\lambda < \dfrac{1}{p-1}$, then the exact analog of Corollary 4.1 holds, with the conclusion that $|a| > (1 - 3\lambda)|r|$ or $|a| > (1 - 2\lambda)|r|$ according to the case. This is not quite immediate since, for example, we may have $|x_1 x_2 x_3| < |x_1| + |x_2| + |x_3|$ if there is consolidation in forming the product. However, it is easy to see that corresponding considerations strengthen the hypothesis $C'(\lambda)$ in a way that more than compensates for this.

It should be emphasized that, in Theorem IV, the part a of w need not consist precisely of a consecutive sequence of factors from the normal form of w, or of r. If we require this, and also restrict a piece to be a consecutive part in this sense of the normal forms of two distinct strictly cyclically reduced elements of R, then, for the reasons indicated, we can assure only that $|a| > (1 - 3\lambda)|r| - 4$ or $|a| > (1 - 2\lambda)|r| - 3$, according to the case. A procedure that is in some sense more natural is to replace $|a|$, as part of $w = bac$, by the number of full factors from w in a, and adding $1/2$ for a factor on either end of a that is consolidated in forming w; if this is done consistently, the number $\|a\|$ exceeds $(1 - 3\lambda)|r|$ or $(1 - 2\lambda)|r|$, r strictly cyclically reduced. This is similar to the procedure of BRITTON [3].

We conclude with two examples. The case with $C(6)$, in particular with $\lambda \leq 1/6$, either for F a free group or F a free product, has already been adequately illustrated by Dehn and his successors. We shall illustrate the two remaining cases, for F a free product. It must be confessed that our example for the third case is the only reasonable application of this case that we have been able to discover.

Example 5.1. Let $F = A * B$, free product of subgroups A and B with $A \cap B = 1$. Let $r = a_1 b_1 a_2 b_2$ with a_1, a_2 in A and b_1, b_2 in B, and let N be the normal closure of r in F. Assume that the set of $a_i^{\pm 1}$ and $b_i^{\pm 1}$ consists of eight distinct elements, and that no one is the square of another. Then $N \cap A = N \cap B = 1$.

We take R to consist of all cyclically reduced conjugates of r and r^{-1}. If s is any strictly cyclically reduced element of R, it follows from the assumption that the $a_i^{\pm 1}$ and $b_i^{\pm 1}$ are distinct, that a piece of s cannot contain all of any letter from s, and hence consists at most of parts of two adjoining letters. Thus s cannot be a product of three pieces, and $C(4)$ holds. Condition T_3^1 is immediate on the same grounds: there cannot be cancellation in the product st of two strictly cyclically reduced elements of R unless $st = 1$. For T_3^2 we require that no product of three factors $a_i^{\pm 1}$ or $b_i^{\pm 1}$ is trivial, and this has been incorporated in our hypothesis. It follows now that if w is any non trivial element of N, then $w = bac$ reduced, where some $s = ax_1 x_2$, reduced, s in R and x_1, x_2 pieces. Evidently the part a of s cannot lie in A or in B, whence the same holds for w.

Example 5.2. Let $F = A * B$, free product of subgroups A and B with $A \cap B = 1$. Let $r = a_1 b_1 a_2 b_2 a_3 b_3$ with the a_i in A and the b_i in B, and let N be the normal closure of r in F. Assume that the b_i are strictly 5-free in the following sense: if $b_{i_1}^{e_1} \ldots b_{i_m}^{e_m} = 1$, where $1 \leq m \leq 5$, each $i_h = 1, 2,$ or 3, and each $e_h = \pm 1$, then there exists h, $1 \leq h < m$, such that $i_h = i_{h+1}$ and $e_h = -e_{h+1}$. Assume that the a_i are weakly 5-free in the following sense: if $a_{i_1}^{e_1} \ldots a_{i_m}^{e_m} = 1$; $1 \leq m \leq 5$; $i_h = 1, 2,$ or 3; $e_h = \pm 1$; then $a_{i_h}^{e_h} a_{i_{h+1}}^{e_{h+1}} = 1$ for some h, $1 \leq h < m$. Then $N \cap A = N \cap B = 1$.

Take R as before. It follows from the distinctness of the $b_i^{\pm 1}$ that no piece can contain a letter $b_i^{\pm 1}$ except as its first or last letter, and hence cannot contain two letters $a_i^{\pm 1}$. Consequently no element of R is a product of two pieces, and $C(3)$ holds. (We note that $C(4)$ need not hold, for example if $a_1 = a_2 = a_3$.) Again the distinctness of the $b_i^{\pm 1}$ implies that one cannot have strictly cyclically reduced r_i in R such that for some i there is cancellation in both $r_{i-1} r_i$ and $r_i r_{i+1}$, with neither product trivial; thus T_3^1, T_4^1, and T_5^1 hold. The weak 5-freeness of the letters gives precisely T_3^2, T_4^2, and T_5^2. The conclusion of the theorem, that non trivial w in N contains some element from R with two pieces missing is here not strong enough. For let $a_1 = a_2 = a_3 = a$. Then $x_1 \equiv b_1 a b_2$ begins both $b_1 a b_2 a b_3 a$ and $b_1 a b_1 a b_2 a (b_3 b_1^{-1})$, $x_2 = a b_3$ begins both $a b_3 a b_1 a b_2$ and $a b_1 a b_2 a b_3$, and $r = a x_1 x_2$. We appeal to the diagram M, which we know to contain a region D with at most two interior edges, restoring to D' its three primary vertices and its three secondary vertices, dividing D' into what we shall call subedges. No interior edge can contain two successive subedges with labels in B, by virtue of the distinctness of the $b_i^{\pm 1}$. Therefore an edge can at most

contain two adjacent subedges with labels in A, together with a subedge with label in B at each end. Two such edges can exhaust at most eight of the twelve subedges of D', whence the part $D' \cap M'$ must contain at least four subedges, and hence bear a label that does not lie in A or in B. We conclude again that w is not in A or B.

Corollary 5.3. *Let B be any group, and b_1, b_2, and b_3 three elements of B that are 5-free in the above sense. Then B can be embedded in a group G containing an element x such that, in G, $xb_1xb_2x^{-1}b_3 = 1$.*

Let A be a cyclic group, with generator a, of order $n \neq 1$, 2, 3, 4, 5. Let $F = A * B$ and $r = ab_1ab_2a^{-1}b_3$, with N the normal closure of R in F. The hypotheses of Example 5.2 are satisfied, and we conclude that $N \cap B = 1$. Consequently $G = F/N$ contains B isomorphically, and the element $x = aN$ satisfied the required equation. We note that B is contained in no proper normal subgroup of G, for, if K is the normal closure of B in G, then G/K is generated by xK; with a defining relation implying that $xK = 1$.

We remark that the above corollary gives only a rather special example of a variety of similar results that can be obtained by similar methods. A stronger form of the result given is contained in a theorem of GERSTENHABER and ROTHAUS [7] under the assumption that B can be embedded in a compact connected Lie group. The analogous result for the equation $xb_1xb_2xb_3 = 1$ follows from a simpler argument of LEVIN [13]. If one considers more generally an equation $x^{n_1}b_1 \dots x^{n_k}b_k = 1$, the case $k = 1$ is trivial, while partial results for $k = 2$ and $k = 3$ can be derived from the two examples above. For $k = 4$ or $k = 5$, Case ii of Theorem IV can be applied under somewhat milder conditions on the n_i and b_i. For $k \geq 6$, we can use Case i of Theorem IV, and conclude that a solution exists provided all $n_i \neq 0$ and all $b_i^{\pm 1}$ are non trivial and distinct; in particular, one may well have $\Sigma n_i = 0$.

Bibliography

[1] BERGE, C.: Théorie des graphes et ses applications. Paris: Dunod 1958.
[2] BOONE, W. W.: Review of BRITTON [3]. Math. Reviews **23**, 440—441 (1962), A 2325.
[3] BRITTON, J. L.: Solution of the word problem for certain types of groups, I, II. Proc. Glasgow Math. Ass. **3**, 45—54, 68—90 (1956—58).
[4] COHEN, D. E., and R. C. LYNDON: Free bases for normal subgroups of free groups. Trans. Am. Math. Soc. **108**, 528—537 (1963).
[5] DEHN, M.: Über unendliche diskontinuierliche Gruppen. Math. Ann. **71**, 116—144 (1912).
[6] — Transformation der Kurven auf zweiseitigen Flächen. Math. Ann. **72**, 413—421 (1912).
[7] GERSTENHABER, M., and O. S. ROTHAUS: The solution of sets of equations in groups. Proc. Nat. Acad. Sci. U.S. **48**, 1531—1533 (1962).
[8] GLADKII, A. V.: On simple Dyck words. Sibirsk Mat. Zh. **2**, 36—45 (1961).
[9] GREENDLINGER, M.: Dehn's algorithm for the word problem. Comm. Pure Appl. Math. **13**, 67—83 (1960).
[10] — On Dehn's algorithms for the conjugacy and word problems with applications. Comm. Pure Appl. Math. **13**, 641—677 (1960).
[11] — Solutions of the word problem for a class of groups by means of Dehn's algorithm, and of the conjugacy problem by means of a generalization of Dehn's algorithm. Dokl. Akad. Nauk SSSR **154**, 507—509 (1964).

[12] — Solution of the conjugacy problem for a class of groups coinciding with their anti-centers, by means of the generalized Dehn algorithm. Dokl. Akad. Nauk SSSR **158**, 1254—1256 (1964).
[13] LEVIN, F.: Solutions of equations over groups. Bull. Am. Math. Soc. **68**, 603—604 (1962).
[14] LIPSCHUTZ, S.: Elements in *S*-groups with trivial centralizers. Comm. Pure Appl. Math. **13**, 679—683 (1960).
[15] LYNDON, R. C.: Cohomology theory of groups with a single defining relation. Ann. of Math. **52**, 650—665 (1950).
[16] — Dependence and independence in free groups. Crelles J. **210**, 148—174 (1962).
[17] REIDEMEISTER, K.: Einführung in die kombinatorische Topologie. Braunschweig: Vieweg 1932.
[18] SCHIEK, H.: Ähnlichkeitsanalyse von Gruppenrelationen. Acta Math. **96**, 157—251 (1956).
[19] — Das Adjunktionsproblem der Gruppentheorie. Math. Ann. **147**, 159—165 (1962).
[20] TARTAKOVSKII, V. A.: The sieve method in group theory. Mat. Sbornik (N.S.) **25**, (67), 3—50 (1949).
[21] — Application of the sieve method to the solution of the word problem for certain types of groups. Mat. Sbornik (N.S.) **25**, (67), 251—274 (1949).
[22] — Solution of the word problem for groups with a *k*-reduced basis for *k* > 6. Izvestiya Akad. Nauk SSSR, Ser. Mat. **13**, 483—494 (1949).
[23] — On primitive composition. Mat. Sbornik (N.S.) **30**, (72), 39—52 (1952).
[24] — Translations of [20, 21, 22]. Am. Math. Soc. Translations **60** (1952), reprint **1** (1962).
[25] ZIESCHANG, H.: Studien zur kombinatorischen Topologie von Flächen und ebenen diskontinuierlichen Gruppen (multigraphed). Frankfurt 1964 (priv. Veröffentlichung).

Added in proof. We would like to add to the bibliography several references that have come recently to our attention. First, independent work of C. WEINBAUM [34] appears to have much in common with the present paper; also, WEINBAUM drew our attention to a paper of E. VAN KAMPEN [33] containing essentially our concept of diagram. We are grateful to B. GRÜNBAUM for telling us of the papers [26, 27] of C. BLANC, and the paper [28] of F. FIALA, where geometric theorems for graphs that are similar to ours are used for a different purpose. We are grateful to M. GREENDLINGER for reference to the papers [29, 30] of A. V. GLADKII, where different proofs of some of our results may be found. We wish also to draw attention to the thesis [32] of P. E. SCHUPP, in which the present methods are extended to apply to the problems of conjugacy, of centralizers, and of torsion, and also to the thesis [31] of P. P. N. ORLIK, written under the guidance of F. RAYMOND, in which these ideas are applied to the theory of manifolds.

Additional bibliography

[26] BLANC, C.: Une interprétation élémentaire des théorèmes fondamentaux de M. Nevanlinna. Comm. Math. Helv. **12**, 153—163 (1940).
[27] — Les réseaux Riemanniens. Comm. Math. Helv. **13**, 54—67 (1941).
[28] FIALA, F.: Sur les polyèdres à faces triangulaires. Comm. Math. Helv. **19**, 83—90 (1946).
[29] GLADKII, A. V.: On groups with *k*-reducible bases. Dokl. Akad. Nauk SSSR **134**, 16—18 (1960).
[30] — On groups with *k*-reducible bases. Sibirsk Math. J. **2**, 366—383 (1961).
[31] ORLIK, P. P. N.: Necessary conditions for the homeomorphism of Seifert-manifolds. Thesis, University of Michigan, 1966.
[32] SCHUPP, P. E.: On Dehn's algorithm and the conjugacy problem. Thesis, University of Michigan, 1966. (To be submitted to Math. Ann.)
[33] VAN KAMPEN, E. R.: One some lemmas in the theory of groups. Ann. J. Math. **55**, 268—273 (1933).
[34] WEINBAUM, C. M.: Visualizing the word problem, with an application to sixth groups. Pacific J. Math. **16**, 557—578 (1966).

ROGER C. LYNDON
The University of Michigan, Department of Mathematics, Angell Hall
Ann Arbor, Michigan, USA