PERGAMON

Topology 39 (2000) 1239–1251

TOPOLOGY

www.elsevier.com/locate/top

A Whitehead algorithm for surface groups

Gilbert Levitt[a], Karen Vogtmann[b,1]

[a]Laboratoire Émile Picard, UMR CNRS 5580, Université Paul Sabatier, 31062 Toulouse Cedex 4, France
[b]Department of Mathematics, White Hall, Cornell University Ithaca, NY 14853, USA

Abstract

For G the fundamental group of a closed surface, we produce an algorithm which decides whether there is an element of the automorphism group of G which takes one specified finite set of elements to another. The algorithm finds such an automorphism if it exists. The methods are geometric and also apply to surfaces with boundary. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Whitehead algorithm; Surface group; Automorphism; Diffeomorphism

1. Introduction

Consider the following basic problem about a group and its automorphisms:

Given finite sets $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ *of elements of a group G, is there an automorphism of G taking* a_i *to* b_i *for all i? If so, find one.*

In 1936, J.H.C. Whitehead found an elegant algorithm which solves this problem for a finitely generated free group [11] (see also [8]). In 1984, Collins and Zieschang extended Whitehead's methods to free products of finitely many freely indecomposable groups, assuming that the above problem, which we will call the *Whitehead problem*, can be solved in each factor [1,2]. In this paper, we give an algorithm to solve the Whitehead problem for the fundamental group of a closed surface Σ.

The Whitehead problem is easy to solve if $\chi(\Sigma) \geqslant 0$, so we assume that Σ is hyperbolic. The problem has a nice geometric translation, as follows. A diffeomorphism f of Σ induces an isomorphism $\pi_1(\Sigma, p) \to \pi_1(\Sigma, f(p))$, where p is a basepoint for Σ. If $f(p) = p$, this gives an automorphism of $\pi_1(\Sigma, p)$; otherwise, we have only an outer automorphism. The maps

E-mail addresses: levitt@picard.ups.tlse.fr (G. Levitt), vogtmann@math.cornell.edu (K. Vogtmann)

[1] Partially supported by NSF grant DMS-9307313.

PII: S0040-9383(99)00027-0

$\pi_0(\mathit{Diff}(\Sigma, p)) \to \mathit{Aut}(\pi_1(\Sigma, p))$ and $\pi_0(\mathit{Diff}(\Sigma)) \to \mathit{Out}(\pi_1(\Sigma))$ are isomorphisms for any closed surface Σ other than the sphere (see, e.g. [12]). If we represent the elements a_i and b_i in $\pi_1(\Sigma,p)$ by loops α_i and β_i in Σ based at p, the Whitehead problem then translates into the following problem about closed curves on a surface:

Is there a diffeomorphism of Σ fixing p which takes each α_i to a loop homotopic to β_i? If so, find one.

Our solution to this geometric problem relies in part on work of Hass and Scott on straightening curves on surfaces [7]. It also applies to surfaces with boundary.

We expect that the methods of this paper can be used to solve the Whitehead problem for torsion-free hyperbolic groups for which we are given the decomposition into freely indecomposable factors and the JSJ canonical splittings of these factors.

2. The Out problem

For a free group F, Whitehead began by solving the problem for finite sets of conjugacy classes of elements of F. Thus, instead of looking for an automorphism which sends a_i to b_i for all i, we just look for one which sends a_i to a conjugate of b_i. This simplifies the problem considerably, and we begin for surface groups in the same way. Represent the elements a_i and b_i in $G = \pi_1(\Sigma, p)$ by loops α_i and β_i based at p in Σ. Using the isomorphism $\pi_0(\mathit{Diff}(\Sigma)) \to \mathit{Out}\,(\pi_1(\Sigma))$, the Whitehead problem becomes

Is there a diffeomorphism of Σ (not necessarily fixing p) which takes each α_i to a loop freely homotopic to β_i? If so, find one.

Again, our solution to this geometric problem also applies to surfaces with boundary. However, this already follows from Whitehead's work on free groups, since the mapping class group of a bounded surface has finite index in the subgroup of outer automorphisms of the (free) fundamental group of the surface which send conjugacy classes represented by the boundary curves to other such classes (see, e.g. [12]).

2.1. Curve straightening

We briefly recall the facts we will need about the "disk flow" of Hass and Scott [7]. A closed curve α is said to be in *general position* if it is immersed in Σ and intersects itself transversely with no triple points. Starting with a curve in general position, Hass and Scott find a "straighter" curve homotopic to α. During the course of the homotopy, only three types of combinatorial changes take place:

1. A monogon bounding a disk in $\Sigma - \alpha$ is eliminated.
2. A bigon bounding a disk in $\Sigma - \alpha$ is eliminated.
3. One edge of a triangle in $\Sigma - \alpha$ is pushed across the opposite vertex. This move eliminates the triangle, while creating a new triangle on the opposite side of the vertex (see Fig. 1). The effect is independent of which edge is chosen to push.

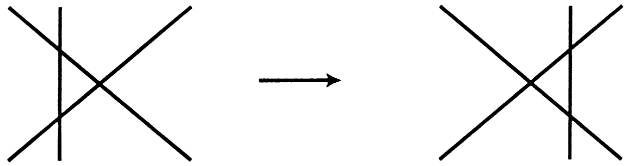

Fig. 1. Triangle move.

Here by a monogon, bigon or triangle we mean one whose closure is imbedded in Σ. The first two types of changes decrease the number of self-intersection points of the curve. The third type of change is called a *triangle move*, and leaves the self-intersecion number unchanged. Moves 1 and 2 are never done "backwards", so that the self-intersection number of the curve never increases during the homotopy.

A closed curve α in general position in Σ is called a *minimal representative* if it has the minimal possible number of self-intersections among curves in general position freely homotopic to α. Hass and Scott prove the following results, for Σ a closed orientable surface:

Theorem 2.1 (Hass and Scott [7, Theorem 2.2]). *A closed curve α in Σ is homotopic to a minimal representative by a homotopy through moves 1, 2 and 3 above.*

Theorem 2.2 (Hass and Scott [7, Theorem 2.1]). *Let α and β be homotopic minimal representatives. Then there is a curve α' obtained from α by triangle moves and an isotopy of Σ sending α' to β.*

The above definitions readily extend to finite n-tuples of closed curves $(\alpha_1, \dots, \alpha_n)$. Hass and Scott remark that Theorems 2.1 and 2.2 apply to finite n-tuples of curves, provided the curves in each n-tuple are distinct and none of them is a proper power [7, top of p. 34]. The proofs given in [7] of Theorems 2.1 and 2.2 work for such an n-tuple of curves even if Σ is not orientable.

2.2. Proper powers

In order to use the results of Hass and Scott to solve the Whitehead problem, we must show that we can assume that none of the a_i or b_i are proper powers.

The following lemma is well known. We sketch a proof for completeness.

Lemma 2.3. *Let G be a torsion free hyperbolic group. Given $g \in G$ nontrivial, there exists a unique $x \in G$, computable effectively, such that x is not a proper power and $g = x^t$ for some $t \geqslant 1$.*

Proof. Existence and uniqueness follow from the fact that the centralizer of g is virtually cyclic, hence cyclic since G is torsion free. To find x and t explicitly, choose a finite generating system and recall the following facts about hyperbolic groups.

Denote by $|h|$ the length of an element $h \in G$, by $\|h\|$ the minimal length of a conjugate of h, and by $\|h\|_s$ the limit (which is also the infimum) of $|h^m|/m$ (this is the *stable norm of h*). Suppose G is

δ-hyperbolic for the chosen generating system. The constant δ may be computed explicitly from a presentation, and $\|h\|_s \geqslant \|h\| - 16\delta$ for every $h \in G$ [3, chapter 10, Proposition 6.4]. Furthermore, $\|h\|_s$ is bounded below by some explicitly computable $\varepsilon > 0$ for h of infinite order ([6, 8.5.T], see also [4, Proposition 3.1]).

If $g = x^t$, then $|g|/t \geqslant \|x\|_s > \varepsilon$, hence $t \leqslant |g|/\varepsilon$. We also have $|g|/t \geqslant \|x\|_s \geqslant \|x\| - 16\delta$, hence a bound for $\|x\|$. Now we only need to check whether h^m is conjugate to g, for a finite number of elements h and exponents m. $\square$

Corollary 2.4. *In the Whitehead problems, we may assume that no a_i or b_i is a proper power.*

Proof. Write $a_i = x_i^{t_i}$ and $b_i = y_i^{u_i}$ as in Lemma 2.3. If $t_i \neq u_i$ for some i, the problem has no solution. Otherwise, an automorphism φ takes a_i to (a conjugate of) b_i if and only if it takes x_i to (a conjugate of) y_i. $\square$

2.3. n-tuples of curves

Fix integers $n > 0$ and $m \geqslant 0$. We consider n-tuples $\gamma = (\gamma_1, \dots \gamma_n)$ of oriented closed curves in Σ in general position, intersecting in m points. We let $T_{m,n}$ be the set whose elements are such n-tuples of curves, up to diffeomorphism of Σ.

Lemma 2.5. *The set $T_{m,n}$ is finite and may be constructed explicitly.*

Proof. Given a set γ, we first consider $\Gamma = \gamma_1 \cup \dots \cup \gamma_n$ as an abstract topological space. There are finitely many possibilities, since the number of components and the number of vertices (all 4-valent) are bounded.

Next we consider a regular neighborhood P of Γ in Σ, up to a diffeomorphism preserving the orientation and the ordering of the curves $\gamma_1, \dots, \gamma_n$. It may be constructed by gluing finitely many pieces: one disc for each vertex of Γ, one strip $e \times [0, 1]$ for each edge e, and either an annulus or a Möbius band for each circle component of Γ. We may list explicity all the possibilities (of course some of the surfaces we obtain do not embed into Σ, e.g. if P is not orientable whereas Σ is).

To each P thus constructed, we attach bounded surfaces so as to obtain a closed surface. This is described by the topological type of each component of $\Sigma - P$ (orientability and Euler characteristic), and the way each boundary component is attached to the boundary of P. Only finitely many possibilities (if any) may lead to a surface diffeomorphic to Σ, since the Euler characteristic of each component of $\Sigma - P$ is bounded by $-\chi(\Sigma)$. $\square$

We make $T_{m,n}$ into a finite, possibly disconnected, graph by placing an edge between two vertices if they correspond to n-tuples of curves differing by a triangle move.

2.4. Solution of the Whitehead problem for conjugacy classes

Proposition 2.6. *Let $G = \pi_1(\Sigma, p)$, and let $a_1, \dots, a_n$ and $b_1, \dots, b_n$ be elements of G. There is an algorithm which decides whether there is an automorphism φ of G taking a_i to a conjugate of b_i for all $i = 1, \dots, n$, and finds such a φ if it exists.*

Proof. We may assume that no a_i or b_i is a proper power by Corollary 2.4, that no a_i is conjugate to $a_j^{\pm 1}$ for $i \neq j$, and similarly for the b_i's (recall that the conjugacy problem is solvable in G). Represent the a_i and b_i by n-tuples of closed curves $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$ in Σ. The assumptions just made allow us to apply Hass–Scott, and after straightening we may assume α and β are minimal representatives.

Count the number of self-intersections of each set of curves. If these numbers are different, there is no automorphism and we are done. Otherwise, let m be the self-intersection number. We claim that the automorphism φ exists if and only if the vertices of $T_{m,n}$ determined by α and β are connected in $T_{m,n}$.

If the two vertices are connected in $T_{m,n}$, then we can construct a diffeomorphism taking α to a set of curves obtained from β by triangle moves (and therefore homotopic to β). Conversely, suppose f is a diffeomorphism of Σ with $f(\alpha_i)$ homotopic to β_i for all i. Then Theorem 2.2 says that the corresponding vertices of $T_{m,n}$ are connected. $\square$

Corollary 2.7. (The Whitehead problem for single elements). *Let $G = \pi_1(\Sigma, p)$, and let $a, b \in G$. There is an algorithm which decides whether there is an automorphism of G taking a to b, and finds such an automorphism if it exists.*

Proof. By Proposition 2.6, we can decide whether there is an automorphism φ taking a to a conjugate of b, and find φ if it exists. We can then solve for x in the equation $\varphi(a) = xbx^{-1}$. The composition of φ with conjugation by x sends a to b. $\square$

3. The Aut problem

We now return to the Whitehead problem for words, as opposed to conjugacy classes of words. Using Corollary 2.7 we reduce to the following special case:

> *Given elements $c, a_1, \dots, a_n, b_1, \dots b_n$ in $G = \pi_1(\Sigma, p)$, is there an automorphism of G which fixes c and sends a_i to b_i for all $i = 1, \dots, n$? If so, find one.*

We may assume that c is not a proper power, by Lemma 2.3. We represent the elements c, a_i and b_i by loops θ, α_i and β_i based at p in Σ.

We may also assume that θ has minimal self-intersection number among all loops freely homotopic to θ, i.e. θ is a minimal representative. To see this, notice that each of the Hass–Scott moves which must be done to find a minimal representative for θ may be performed without moving the point p, as long as p is not a double point of θ. Alternatively, find a minimal representative of θ using the Hass–Scott algorithm; the result is conjugate to θ in $\pi_1(\Sigma)$, by a (computable) element δ. Now choose a diffeomorphism of Σ isotopic to the identity which fixes p and induces conjugation by δ on $\pi_1(\Sigma, p)$. The image of the minimal representative is then homotopic to the original θ, and is still a minimal representative.

Let N be a subsurface of Σ consisting of a regular neighborhood of θ together with all disk components of the complement of this neighborhood. Then N is an invariant of the homotopy class

c of θ in the following sense: if $\theta' \simeq \theta$ is another minimal representative of c, then there is an isotopy of Σ sending N diffeomorphically onto the subsurface N' corresponding to θ'. This follows since, by Theorem 2.2, θ' can be obtained from θ by triangle moves, which do not change the isotopy class of N.

3.1. Reduction to a subgroup of the stabilizer of c

Let *Stab(c)* denote the stabilizer of c in $Aut\,(\pi_1(\Sigma, p)) \cong \pi_0(Diff\,(\Sigma, p))$. Then *Stab(c)* contains the subgroup $\pi_0(Diff\,(\Sigma, N))$ of (isotopy classes of) diffeomorphisms of Σ which fix N pointwise; it also contains the inner automorphism $i_c(x) = cxc^{-1}$, which can be represented by a diffeomorphism of Σ fixing p and $\Sigma - N$. Note that i_c commutes with each element g of $\pi_0(Diff\,(\Sigma, N))$, since i_c and g can be represented by diffeomorphisms with disjoint support. We define A_c to be the subgroup of $Stab(c)$ generated by $\pi_0(Diff\,(\Sigma, N))$ and i_c.

Lemma 3.1. *A_c has finite index in Stab(c).*

Proof. The result is easy if N is an annulus or a Möbius band, so we assume $\chi(N) < 0$.

If f is a diffeomorphism of Σ with $f(\theta) \simeq \theta$, then up to isotopy, f restricts to a diffeomorphism f_N of N. Thus we have a map ρ from $Stab\,(c)$ to $\pi_0(Diff\,(N))$, which we view as a subgroup of $Out\,(\pi_1(N, p))$.

We claim that the image of $Stab\,(c)$ under this map is finite. For any fixed positive integer m, the set S_m of (homotopy classes of) simple closed curves in N with intersection number at most m with θ is finite, since each component of $N - \theta$ is either a disk or an annulus. For $f \in Stab(c)$, the associated diffeomorphism f_N of N preserves the intersection number with θ, so acts on S_m. We can write down a finite set Δ of simple closed curves in N with the property that an element of $\pi_0(Diff\,(N))$ is determined completely by the images of the curves in Δ. Now choose m larger than the geometric intersection number of θ with any curve in Δ, so that the action of f_N on S_m completely determines the class of f_N in $\pi_0(Diff\,(N))$. Since there are only a finite number of permutations of the finite set S_m, the image of $Stab(c)$ in $\pi_0(Diff\,(N))$ is finite.

The kernel of the map $\rho : Stab(c) \to Out(\pi_1(N, p))$ is generated by $\pi_0(Diff\,(\Sigma, N))$ together with diffeomorphisms which induce inner automorphisms on $\pi_1(N, p)$. But the only inner automorphisms of $\pi_1(N, p)$ which fix c are conjugation by powers of c. (If $aca^{-1} = c$, then a commutes with c. Since c is not a proper power, a is a power of c.) Therefore A_c is equal to the kernel of ρ, so has finite index in $Stab(c)$. □

One can determine explicitly the image of ρ, and thus write down explicit representatives $\varphi_1, \ldots, \varphi_s$ for cosets of A_c in $Stab(c)$, as follows. Using the solution to the Out problem, we can list those permutations of S_m that are induced by some diffeomorphism of N (which is then unique up to isotopy). For each such diffeomorphism, we check whether it fixes θ and extends to a diffeomorphism of the whole of Σ.

If $\psi \in Stab(c)$ sends a_i to b_i for all i, then there exist $j \in \{1, \ldots, s\}$ and $\varphi \in A_c$ such that $\varphi(a_i) = \varphi_j^{-1}(b_i)$ for all i. Our question in $Stab(c)$ is therefore equivalent to a finite set of questions of the form:

Is there an element φ of A_c with $\varphi(a_i) = b_i$ for all $i = 1, \ldots, n$? If so, find one.

We now proceed to solve this question.

3.2. Geometric translation

On each component C_k of the boundary of N, fix an orientation and choose a basepoint q_k. We denote by γ_k the oriented simple loop based at q_k with image C_k.

Let α be any loop based at p in Σ, and homotope α so that it is immersed and intersects ∂N transversely with minimal possible intersection number. The intersection points of α with ∂N divide α into arcs: $\alpha = \eta_0\sigma_1\eta_1\sigma_2 \ldots \sigma_r\eta_r$ with $\eta_j \subset N$ and $\sigma_j \subset \overline{\Sigma - N}$. Associated to this decomposition are indexing functions o, e, where σ_j begins in the component $C_{o(j)}$ of ∂N and ends in $C_{e(j)}$. We may assume each σ_j begins at $q_{o(j)}$ and ends at $q_{e(j)}$ (see Fig. 2).

We say that $\eta_0\sigma_1\eta_1 \ldots \sigma_r\eta_r$ is a *normal form* for α with indexing functions o and e. Note that σ_j is not homotopic to a power of $\gamma_{o(j)}$ if $o(j) = e(j)$ since otherwise we could push σ_j across $C_{o(j)}$ and decrease the number of intersections of α with ∂N. Similarly η_j is not a power of $\gamma_{e(j)}$ if $o(j+1) = e(j)$.

Lemma 3.2. *Let α and β be two loops at p, with normal forms $\eta_0\sigma_1\eta_1 \ldots \sigma_r\eta_r$ and $\mu_0\tau_1\mu_1 \ldots \tau_s\mu_s$ and indexing functions o,e and o′, e′, respectively. Then α is homotopic to β if and only if $r = s$, $o = o'$, $e = e'$ and there are integers x_j and y_j for $1 \leqslant j \leqslant r$ with*

$$\begin{aligned}
&\eta_0 \simeq \mu_0\gamma_{o(1)}^{-x_1}, \\
&\sigma_j \simeq \gamma_{o(j)}^{x_j}\tau_j\gamma_{e(j)}^{-y_j} \qquad \textit{for } 1 \leqslant j \leqslant r, \\
&\eta_j \simeq \gamma_{e(j)}^{y_j}\mu_j\gamma_{o(j+1)}^{-x_{j+1}} \quad \textit{for } 1 \leqslant j \leqslant r-1, \\
&\eta_r \simeq \gamma_{e(r)}^{y_r}\mu_r.
\end{aligned}$$

Proof. In each component Σ_l of $\overline{\Sigma - N}$, fix a basepoint p_l and a path from q_k to p_l for each component C_k of the boundary of Σ_l. Then $\pi_1(\Sigma, p)$ is the fundamental group of a graph of groups, with vertex stabilizers $\pi_1(N, p)$ and $\pi_1(\Sigma_l, p_l)$, and cyclic edge stabilizers generated by the γ_k. The expressions $\eta_0\sigma_1\eta_1 \ldots \sigma_r\eta_r$ and $\mu_0\tau_1\mu_1 \ldots \tau_s\mu_s$ are reduced for the graph-of-groups representation of $\pi_1(\Sigma, p)$, and so are uniquely determined up to the equivalence relation in the statement of the Lemma (see [10, Exercise I.5.2]). $\square$

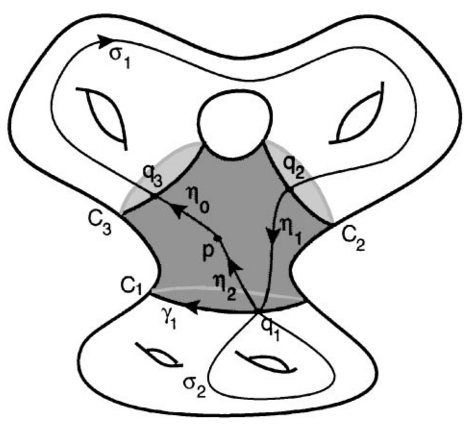

Fig. 2. Decomposition of α.

An element of A_c sends a normal form decomposition

$$\eta_0\sigma_1\eta_1 \dots \sigma_r\eta_r$$

to

$$(\theta^w\eta_0)g(\sigma_1)\eta_1 \dots g(\sigma_r)(\eta_r\theta^{-w})$$

for some $w \in \boldsymbol{Z}$ and diffeomorphism g of Σ fixing N. This expression is still in normal form, so by Lemma 3.2, we may rephrase the Whitehead problem as:

Question 3.3. *For $1 \leqslant i \leqslant n$, let $\eta_{i0}\sigma_{i1}\eta_{i1} \dots \sigma_{ir_i}\eta_{ir_i}$ and $\mu_{i0}\tau_{i1}\mu_{i1} \dots \tau_{ir_i}\mu_{ir_i}$ be normal forms for α_i and β_i, respectively, with indexing functions o_i and e_i (the same for α_i and for β_i). Do there exist integers w, x_{ij} and y_{ij} and a diffeomorphism g of Σ fixing N with*

$$\begin{aligned} &\theta^w\eta_{i0} \simeq \mu_{i0}\gamma_{o_i(1)}^{-x_{i1}}, \\ &g(\sigma_{ij}) \simeq \gamma_{o_i(j)}^{x_{ij}}\tau_{ij}\gamma_{e_i(j)}^{-y_{ij}} \quad \text{for } 1 \leqslant j \leqslant r_i, \\ &\eta_{ij} \simeq \gamma_{e_i(j)}^{y_{ij}}\mu_{ij}\gamma_{oi(j+1)}^{-x_{i,j+1}} \quad \text{for } 1 \leqslant j \leqslant r_i - 1, \\ &\eta_{ir_i}\theta^{-w} \simeq \gamma_{e_i(r_i)}^{y_{ir_i}}\mu_{ir_i}. \end{aligned} \tag{$*$}$$

Note that an equation of the form $\lambda_* \simeq \gamma_{o(*)}^{x_*}\tau_*\gamma_{e(*)}^{-y_*}$ means that λ_* and τ_* are homotopic as paths in $\overline{\Sigma - N}$, with endpoints required to remain on the boundary during the homotopy.

3.3. Finding diffeomorphisms of the complementary components

We begin by trying to find the diffeomorphism g of Σ fixing N. This is equivalent to finding a diffeomorphism of each component M of $\overline{\Sigma - N}$ which fixes the boundary of M pointwise. In this subsection we fix such a component M; i.e. M is a connected, bounded surface which is not a disk.

Lemma 3.4. *Let $\{\sigma_1, \dots, \sigma_r\}$ and $\{\tau_1, \dots, \tau_r\}$ be two finite sets of oriented arcs in M such that for each j both σ_j and τ_j connect $q_{o(j)}$ to $q_{e(j)}$. We can decide whether there exists a diffeomorphism $h: M \to M$ equal to the identity on the boundary with $h(\sigma_j) \simeq \gamma_{o(j)}^{x_j}\tau_j\gamma_{e(j)}^{-y_j}$ for some integers x_j and y_j, and find one if it exists.*

Proof. Let M_1 and M_2 be two copies of M and form the double MM by gluing M_1 to M_2 by the identity on the boundary. Let $\sigma_j\bar{\sigma}_j$ denote the oriented closed curve in MM obtained by gluing the copy of σ_j in M_1 to the copy of $\bar{\sigma}_j$ in M_2, where $\bar{\sigma}_j$ denotes σ_j with the opposite orientation; similarly, let $\sigma_j\bar{\tau}_j$ denote the result of gluing the copy of σ_j in M_1 to the copy of $\bar{\tau}_j$ in M_2. By the outer version of the Whitehead algorithm, we can decide whether there is a diffeomorphism f of MM taking $\sigma_j\bar{\sigma}_j$ to $\sigma_j\bar{\tau}_j$ for all j and the oriented boundary curves γ_k to γ_k for all k. If f exists, it must restrict to a diffeomorphism of M_2 sending σ_j to $\gamma_{o(j)}^{x_j}\tau_j\gamma_{e(j)}^{-y_j}$ for some integers x_j and y_j as desired (we may have to compose f with the canonical involution of MM if f sends M_2 to M_1). $\square$

The diffeomorphism h in Lemma 3.4, if it exists, is not unique. In particular, we may compose it with Dehn twists around the components of ∂M. This has the effect of increasing each x_j and y_j by an integer depending only on the component of ∂M containing the corresponding endpoint of σ_j. Our next goal will be to show that, as far as the numbers x_j and y_j are concerned, this is the only freedom we have.

First we show the following:

Lemma 3.5. *Let M be a surface with boundary, with $\chi(M) < 0$. Let C be an oriented component of the boundary, $q \in C$ a basepoint and γ the corresponding boundary loop. Let f be a diffeomorphism of M which fixes C pointwise.*

(i) *If α is a loop at q with $f(\alpha) \simeq \gamma^u \alpha \gamma^{-v}$, then $u = v$.*
(ii) *If α, β are loops not homotopic to powers of γ, with $f(\alpha) \simeq \gamma^u \alpha \gamma^{-u}$ and $f(\beta) \simeq \gamma^t \beta \gamma^{-t}$, then $u = t$.*

Proof. (i) After possibly composing f with a power of the Dehn twist D about C we may assume $u = 0$.

If α is simple, then $f(\alpha) \simeq \alpha\gamma^{-v}$ is also simple, as is $f^{-1}(\alpha) \simeq \alpha\gamma^{v}$. But, one of $\alpha\gamma^{v}$ or $\alpha\gamma^{-v}$ has minimal self-intersection number exactly v, since the obvious transverse representative has no *embedded* monogons, bigons or triangles with which to do Hass-Scott moves (see Fig. 3). This proves $v = 0$ when α is simple.

We claim that the general case follows from the simple case by passing to a finite cover. By a theorem of Scott [9], there is a finite cover $p : M_1 \to M$ and an incompressible subsurface $X \subseteq M_1$ with $p_*\pi_1(X)$ equal to the subgroup F of $\pi_1(M)$ generated by α and γ. We may assume that α is not a power of γ, so that F is free of rank two. Let $\tilde{\alpha}, \tilde{\gamma}$ be lifts of α, γ, respectively, so that $\pi_1(X) = \langle \tilde{\alpha}, \tilde{\gamma} \rangle$. Since γ lifts to a boundary curve $\tilde{\gamma} \subset \partial X$ which is primitive in $\pi_1(X)$, X is a pair of pants or a punctured Möbius band. Thus there is a *simple* curve $\tilde{\alpha}'$ in X with $\pi_1(X) = \langle \tilde{\alpha}', \tilde{\gamma} \rangle$. By a standard result about the free group of rank 2, there exist integers s, s' such that $\alpha' = \tilde{\gamma}^s \tilde{\alpha}^{\pm 1} \tilde{\gamma}^{s'}$. Replacing α by $\gamma^s \alpha^{\pm 1} \gamma^{s'}$ in the original problem, we may assume both α and γ lift to simple curves in M_1. The index of $p_*\pi_1(M_1)$ in $\pi_1(M)$ is finite. Since there are only finitely many subgroups of a given finite index in $\pi_1(M)$, we must have $f^l_*(p_*\pi_1(M_1)) = p_*\pi_1(M_1)$ for some $l > 0$, so that f^l lifts to a diffeomorphism of M_1. Replacing f by f^l in the original problem, we are reduced to the simple case.

(ii) By composing with a power of the Dehn twist D along C, we may assume $u = 0$.

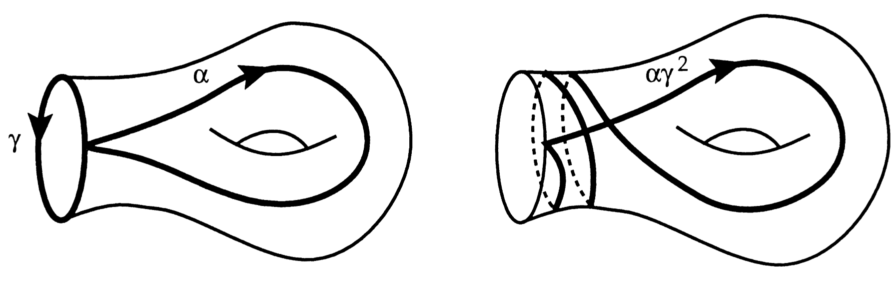

Fig. 3. Minimal representative for $\alpha\gamma^2$.

We first consider the case that α and β are both simple. Double M along its boundary and let A and B be simple curves in the doubled surface, obtained by doubling α, β and making them simple. Extend f to the doubled surface by the identity. Then f acts on B as D^t and on A as the identity (up to homotopy). If $t \neq 0$, the geometric intersection number between $f^m(A)$ and $f^m(B)$ goes to infinity as m goes to infinity by [5, Proposition 1 p. 68], a contradiction (note that the result of [5] applies even if M is not orientable, because C does not bound a Möbius band on the doubled surface).

Again, we reduce to the simple case by passing to a finite cover. Exactly as in part (i) we may lift to a cover in which α and γ are simple. In this cover, β may not lift to a closed loop; however some power of β is closed and we may replace β by this power without loss of generality.

We now apply the same trick to γ and β: we pass to a finite cover in which γ and $\gamma^s\beta^{\pm 1}\gamma^{s'}$ are both simple, for some s and s'. Replacing β by $\gamma^{s'}\beta^{\pm 1}\gamma^{s'}$ in the original problem, and α by some power of α, we are reduced to the situation that γ, α and β are all simple. □

Corollary 3.6. *Let $h: M \to M$ be a diffeomorphism equal to the identity on the boundary, with $h(\sigma_j) \simeq \gamma_{o(j)}^{x_j}\tau_j\gamma_{e(j)}^{-y_j}$ for some integers x_j and y_j (as in Lemma 3.4). Let h' be another diffeomorphism with the same properties (for possibly different integers x'_j, y'_j). Then there exists an integer z_k for each boundary component C_k such that $h'(\sigma_j) \simeq \gamma_{o(j)}^{z_{o(j)}}h(\sigma_j)\gamma_{e(j)}^{-z_{e(j)}}$ for all j.*

Proof. Consider the composition $f = h^{-1}h'$; this sends each σ_j to $\gamma_{o(j)}^{u_j}\sigma_j\gamma_{e(j)}^{-v_j}$ for some integers u_j and v_j. By Lemma 3.5, $u_j = v_j$ whenever $o(j) = e(j)$. We also need to show that if σ_j and σ_l both begin on $C = C_{o(j)} = C_{o(l)}$, then $u_j = u_l$. There are four cases to consider, depending on where the other endpoints are.

The result follows from Lemma 3.5 if σ_j and σ_l have both endpoints on C. If σ_j and σ_l have endpoints on the same component $C' \neq C$, collapse C' to a point and apply the first part of Lemma 3.5 to $\sigma_j\sigma_l^{-1}$ (we may assume M is not an annulus, so we don't get a disc after collapsing). If σ_j and σ_l have endpoints on distinct components C', C'' both different from C, perform Dehn twists around C' and C'' to get rid of v_j and v_l, glue C' to C'', and apply the first part of Lemma 3.5 to $\sigma_j\sigma_l^{-1}$. If $C' = C \neq C''$, glue a punctured torus on C'', extend f by the identity, and apply the second part of Lemma 3.5 to a loop of the form $\sigma_l\tau\sigma_l^{-1}$, where τ is a loop in the punctured torus which is not parallel to the boundary. □

In the situation of Lemma 3.4 we may now decide whether there exists h sending σ_j to $\gamma_{o(j)}^{x_j}\tau_j\gamma_{e(j)}^{-y_j}$, where some of the x_j, y_j are prescribed in advance.

3.4. A Whitehead algorithm for surface groups

We are now in a position to prove our main result.

Theorem 3.7. *Let $G = \pi_1(\Sigma, p)$, and let $\{a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ be elements of G. There is an algorithm which decides whether there is an automorphism φ of G taking a_i to b_i for each i, and finds such a φ if it exists.*

Proof. After solving the problem for conjugacy classes, we are reduced to the same problem in the stabilizer of an element $c \in G$. In the notation defined above, we have further reduced this problem to that of finding a diffeomorphism g of Σ fixing the subsurface N and integers x_{ij}, y_{ij} and w which satisfy equations $(*)$ of Question 3.3.

We first consider the case that c is represented by a simple two-sided curve, so that N is an annulus with two boundary components C_1 and C_2. We orient C_1 and C_2 in the same direction, and we denote by γ the corresponding homotopy class of loops (dropping all subscripts $e_i(j)$, $o_i(j)$ for simplicity). Conjugation by c can be realized by a diffeomorphism of Σ equal to the identity on N (opposite Dehn twists supported in neighborhoods of C_1 and C_2), so we may in fact assume that $w = 0$.

Choose an embedded arc v through p from C_1 and C_2. By pushing curves into $\Sigma - N$, we may assume that all of the η_{ij} and μ_{ij} in equations $(*)$ are contained in v. The equations

$$\eta_{i0} \simeq \mu_{i0}\gamma^{-x_{i1}},$$

$$\eta_{ij} \simeq \gamma^{y_{ij}}\mu_{ij}\gamma^{-x_{i,j+1}},$$

$$\eta_{ir_i} \simeq \gamma^{y_{ir_i}}\mu_{ir_i},$$

then imply that

$$x_{i1} = 0,$$

$$y_{ij} = x_{i,j+1},$$

$$y_{ir_i} = 0$$

for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant r_i - 1$.

By Lemma 3.4, we can decide whether there exists a diffeomorphism h of $\overline{\Sigma - N}$ fixing the boundary with $h(\sigma_{ij}) \simeq \gamma^{u_{ij}}\tau_{ij}\gamma^{-v_{ij}}$ for some integers u_{ij} and v_{ij}. If h does not exist, then g does not exist. If h exists, then by Corollary 3.6 any diffeomorphism g which satisfies our requirements has the same effect on the σ_{ij} as the product of h with Dehn twists z_1 times around C_1 and z_2 times around C_2, with $x_{ij} = u_{ij} + z_{o_i(j)}$ and $y_{ij} = v_{ij} + z_{e_i(j)}$ for all i and j. Thus g exists if and only if the following system of integral linear equations is solvable:

$$u_{i1} + z_{o_i(1)} = 0,$$

$$v_{ij} + z_{e_i(j)} = u_{i,j+1} + z_{o_i(j+1)},$$

$$v_{ir_i} + z_{e_i(r_i)} = 0$$

for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant r_i - 1$. This completes the proof when N is an annulus.

The proof when N is a Möbius band is fairly similar. We denote by $\gamma \simeq c^2$ the loop corresponding to ∂N. Since the Dehn twist around γ represents conjugation by c^2, we can assume $w = 0$ or 1. The precise value is determined by the equation $\theta^w\eta_{10} \simeq \mu_{10}\gamma_{o_1(1)}^{-x_{11}}$, considered modulo 2 in $\pi_1 N \simeq \boldsymbol{Z}$. By properties of normal forms, all loops η_{ij}, u_{ij} for $0 < j < r_i$ represent odd powers of c, and by pushing into $\Sigma - N$ we may assume $\eta_{ij} = \mu_{ij}$. For $j = 0$, we consider $\theta^w\eta_{i0}$ and μ_{i0}. If they are not equal modulo 2 in $\pi_1 N$, then g does not exist. If they are, then we may assume that they are actually equal (not only modulo 2). A similar argument applies for $j = r_i$. The rest of the proof is then the same as when N is an annulus.

If N is not an annulus or a Möbius band, we cannot ignore conjugation by c. The equation in $(*)$ involving η_{ij} and μ_{ij} can be interpreted as equations in the free group $\pi_1(N, p)$ by fixing paths from p to each q_k. We will need the following lemma about solving equations in free groups:

Lemma 3.8. *Let A, B, c and d be elements of a finitely generated free group F, with c and d nontrivial.*

(i) *If c and d are not contained in conjugate cyclic subgroups, the equation $B = c^t A d^u$ has at most one solution $(t, u) \in Z^2$. We can decide effectively whether there is a solution, and if so find it.*

(ii) *If A does not belong to the maximal cyclic subgroup containing c, the same conclusions hold for the equation $B = c^t A c^u$.*

Proof. (i) To show uniqueness, note that if $c^t A d^u = c^{t'} A d^{u'}$, then $(AdA^{-1})^{u-u'} = c^{t-t'}$. Since F is free, this implies that either AdA^{-1} and c are in the same cyclic subgroup or $u - u' = t - t' = 0$.

If t and u satisfy $B = c^t A d^u$ with $|t|$ (hence also $|u|$) large, then some central segment of the word $c^t A = B d^{-u}$ is a conjugate of both a power of c and power of d, contradicting our hypothesis. Consequently, we need only check whether the equation is true for finitely many values of t and u.

(ii) For the uniqueness statement, note that since A and c are not in the same maximal cyclic subgroup, they generate a free subgroup of F of rank 2, so are in fact a basis for this free subgroup. If $c^t A c^u = c^{t'} A c^{u'}$ then $t = t'$ and $u = u'$ by the normal form theorem for free groups.

Without loss of generality, we may assume that c is cyclically reduced, and that c is not a proper power. If $|t|$ (and therefore $|u|$) is large, the words Bc^{-u} and $c^t A$ have reduced form bc^{-s} ($|b| \leqslant |B| + |c|$) and $c^r a$ ($|a| \leqslant |A| + |c|$) respectively, with $|s|$ and $|r|$ large; if $B = c^t A c^u$, then either ca or $c^{-1}a$ is a terminal segment of c^{-s}. Since c is cyclically reduced and not a proper power, no cyclic permutation of c is equal to c or c^{-1}. Therefore a must be a power of c, i.e. A is a power of c, contradicting our hypothesis. □

Since N is not an annulus or a Möbius band, the maximal cyclic subgroups containing c and $[\gamma_k]$ are not conjugate in $\pi_1(N)$. Therefore the equations

$$c^w[\eta_{i0}] = [\mu_{i0}][\gamma_{o_i(1)}]^{-x_{i1}},$$

$$[\eta_{ir_i}]c^{-\omega} = [\gamma_{e_i(r_i)}]^{y_{ir_i}}[\mu_{ir_i}]$$

for $1 \leqslant i \leqslant n$, uniquely determine x_{i1}, y_{ir_i} and w, by part 1 of Lemma 3.8. Since N is not an annulus, the maximal cyclic subgroups containing $[\gamma_k]$ and $[\gamma_{k'}]$ are not conjugate for $k \neq k'$. Furthermore, neither $[\eta_{ij}]$ nor $[\mu_{ij}]$ is a power of $[\gamma_{e_i(j)}]$ if $e_i(j) = o_i(j+1)$, since the intersection number of the α_i and β_i with ∂N is minimal. Therefore the equations

$$[\eta_{ij}] = [\gamma_{e_i(j)}]^{y_{ij}}[\mu_{ij}][\gamma_{o_i(j+1)}]^{-x_{i,j+1}}$$

for $1 \leqslant i \leqslant n, 1 \leqslant j \leqslant r_i - 1$ uniquely determine the remaining values of y_{ij} and x_{ij} (if $e_i(j) \neq o_i(j+1)$, use part 1 of Lemma 3.8; otherwise part 2 applies).

Now apply Lemma 3.4 to determine whether there is a diffeomorphism h of Σ fixing N and sending each σ_{ij} to $\gamma_{o_i(j)}^{u_{ij}}\tau_{ij}\gamma_{e_i(j)}^{-v_{ij}}$ for some integers u_{ij} and v_{ij}. If such a diffeomorphism exists, we now check whether any choices of the twist factors z_k satisfy the equations $x_{ij} = u_{ij} + z_{o_i(j)}$ and $y_{ij} = v_{ij} + z_{e_i(j)}$ for all i and j. If so, composing h with the specified Dehn twists gives the required diffeomorphism g; the composition of the automorphism induced by g and conjugation by c^w solves the Whitehead problem in $Aut(\pi_1(N, p))$. If not, there is no solution. □

References

[1] D.J. Collins, H. Zieschang, Rescuing the Whitehead method for free products. I. Peak reduction, Mathematische Zeitschrift 185 (4) (1984) 487–504.

[2] D.J. Collins, H. Zieschang, Rescuing the Whitehead method for free products. II. The algorithm, Mathematische Zeitschrift 186 (3) (1984) 335–361.

[3] M. Coornaert, T. Delzant, A. Papadopoulos, Géométrie et théorie des groupes, Les groupes hyperboliques de Gromov, Lecture Notes in Mathematics, 1441, Springer, Berlin, 1990.

[4] T. Delzant, Sous-groupes distingués et quotients des groupes hyperboliques, Duke Mathematical Journal 83 (3) (1996) 661–682.

[5] A. Fathi, F. Laudenbach, V. Poenaru, Travaux de Thurston sur les surfaces, Séminaire Orsay, Astérisque, 66–67, Société Mathématique de France, Paris, 1979.

[6] M. Gromov, Hyperbolic groups, Essays in Group Theory, Mathematical Sciences Research Institute Publications, Vol. 8, Springer, New York, 1987, pp. 75–263.

[7] J. Hass, P. Scott, Shortening curves on surfaces, Topology 33 (1) (1994) 25–43.

[8] P.J. Higgins, R.C. Lyndon, Equivalence of elements under automorphisms of a free group, The Journal of the London Mathematical Society Second Series 8 (1974) 254–258.

[9] P. Scott, Subgroups of surface groups are almost geometric, The Journal of the London Mathematical Society Second Series 17 (3) (1978) 555–565.

[10] J.-P. Serre, Arbres, Amalgames, SL_2, Rédigé avec la collaboration de Hyman Bass, Astérisque, vol. 46, Société Mathématique de France, Paris, 1977.

[11] J.H.C. Whitehead, On equivalent sets of elements in free groups, Annals of Mathematics 37 (1936) 782–800.

[12] H. Zieschang, E. Vogt, H.-D. Coldewey, Surfaces and planar discontinuous groups, in: Lecture Notes in Mathematics, vol. 835, Springer, Berlin, 1980 Translated from the German by John Stillwell.