title

Transforming Curves on Surfaces*

Tamal K. Dey† Sumanta Guha‡

Abstract

We describe an optimal algorithm to decide if one closed curve on a triangulated 2-manifold can be continuously transformed to another, i.e., if they are homotopic. Suppose C_1 and C_2 are two closed curves on a surface M of genus g. Further, suppose T is a triangulation of M of size n such that C_1 and C_2 are represented as edge-vertex sequences of lengths k_1 and k_2 in T, respectively. Then, our algorithm decides if C_1 and C_2 are homotopic in $O(n+k_1+k_2)$ time and space, provided $g \neq 2$ if M orientable, and $g \neq 3, 4$ if M is non-orientable. This as well implies an optimal algorithm to decide if a closed curve on a surface can be continuously contracted to a point. Except for three low genus cases, our algorithm completes an investigation into the computational complexity of two classical problems for surfaces posed by the mathematician Max Dehn at the beginning of this century. The novelty of our approach is in the application of methods from modern combinatorial group theory.

Keywords: Combinatorial group theory, computation, curve, fundamental group, homotopy, surface, topology.

1 Introduction

Computational topology is an emerging new subdiscipline of computational geometry. There are often situations in topology when the existence of some structure $\mathcal{S}$ or the decidability of some problem $\mathcal{P}$ has been proved by mathematicians, but the complexity of $\mathcal{S}$ or the efficiency of algorithms to decide $\mathcal{P}$ remains to be thoroughly investigated. Computational topology deals with these algorithmic aspects of topology. Some work in the area related to this paper may be found in [5, 6, 8, 9, 17, 22]. Two recent survey articles on computational topology are [7, 21].

The topological objects that we consider in this paper are surfaces or, equivalently, 2-manifolds. By a surface or 2-manifold we shall always mean a compact, connected, and boundaryless 2-manifold. Everyday examples of such surfaces include spheres and tori (doughnuts). In fact, any finite object with volume that we care to examine in the three-dimensional world around us is bounded by an orientable surface, i.e., a surface with two distinct sides. A well-known exotic surface is the Klein bottle which is non-orientable and cannot be physically realized in three-dimensional space.

Vegter and Yap [22] first examined the computational problems associated with surfaces, in particular with the combinatorial representation of surfaces. A combinatorial representation is a representation as a discrete structure, a necessary preliminary in any discrete algorithm for a topological object. An example of a combinatorial representation of a surface M is a triangulation of M.

The particular problems for surfaces that we consider here date back to the beginning of this century when Max Dehn [2, 3, 4] formulated and mathematically solved two now-classical

*An extended abstract appeared in the Proceedings of the 32nd FOCS, Milwaukee, 1995.

†Department of Computer Science and Engineering, I.I.T., Kharagpur, India 721302. Email: dey@cse.iitkgp.ernet.in. Fax: (91-3222) 55239.

‡Department of Electrical Engineering and Computer Science, University of Wisconsin-Milwaukee, P.O. Box 784, Milwaukee, WI 53201, USA. Email: guha@cs.uwm.edu. Fax: (1-414) 229-6958.

problems, the contractability and transformability problems, as he termed them (see also Poincaré [16]).

The contractability problem is to decide if a closed curve, or cycle, C on a surface M can be continuously contracted to a point, i.e., if C is null-homotopic. Schipper [17] first investigated the complexity of an algorithm for the contractability problem that dynamically maintains a part of the universal covering space of M. Subsequently, Dey [6] and Dey and Schipper [8] gave improved implementations of this algorithm. All use complex data structures, and the best result heretofore is in the latter paper, using $O(n+k\log g)$ time, which is suboptimal, and $O(n+k)$ space to decide contractability, where C is of length k on a surface of genus g with a triangulation of size n.

The second and harder of Dehn's problems, the transformability problem, asks if two closed curves on a surface can be continuously transformed one to the other, i.e., if they are homotopic. For example, in Figure 1, C_1 is homotopic to C_2 but not C_3. In this paper we describe a time and space optimal algorithm for the transformability problem on a surface M of genus g, where $g \neq 2$ if M is orientable, and $g \neq 3, 4$ if M is non-orientable. Given a triangulation T of size n of M, and closed curves C_1 and C_2 on M presented as edge-vertex sequences in T of lengths k_1 and k_2, respectively, our algorithm decides if C_1 and C_2 are homotopic in $O(n + k_1 + k_2)$ time and space. This immediately implies an optimal algorithm to decide the contractability of C_1 by choosing C_2 to be a point.

Our approach however abandons universal covering spaces in favor of non-metric combinatorial methods. We find a canonical representation for closed curves on M as elements of the fundamental group $\pi(M)$, observe that deciding the transformability of two curves is equivalent to deciding if their canonical representatives in $\pi(M)$ are conjugate, and use methods from modern combinatorial group theory to efficiently solve this conjugacy problem. Specifically, we use results of Greendlinger [10] to formulate a Dehn-type algorithm. Our algorithm is consequently altogether different from previous ones [6, 8, 17]. Moreover, it is simpler to implement, using data structures no more complex than required to manipulate graphs, stacks and strings.

In Section 2 we discuss some preliminaries. Our algorithm is described in Section 3, and we conclude in Section 4.

2 Preliminaries

In this section we review notions from combinatorial group theory, topology, and surface theory that we later employ. As these are mostly standard, we shall be terse and provide references as needed.

2.1 Combinatorial Group Theory

Given a set of symbols X, let X^{-1} denote the set of symbols $\{a^{-1} : a \in X\}$. A *letter* is an element of $X \cup X^{-1}$. A *word* w on X is a finite sequence $a_1 \dots a_k, k \geq 0$, of letters. The *length* of w, denoted $|w|$, is k. An *elementary transformation* of a word w consists of inserting or deleting a subword of the form aa^{-1} or $a^{-1}a$.

The set of all words on X is denoted $W(X)$. The equivalence relation $\sim$ on $W(X)$ is defined by $w_1 \sim w_2$ if w_2 can be derived from w_1 by a finite sequence of elementary transformations. The *free group* $F(X)$ on X is the set $W(X)/\sim$ of equivalence classes modulo $\sim$, endowed with the binary operation induced by concatenation. The *unit* of $F(X)$ is the equivalence class of the empty word ϵ, and the *inverse* of the equivalence class of $a_1 \dots a_k$ is the equivalence class of $a_k^{-1} \dots a_1^{-1}$. We shall, henceforth, always identify a word $w \in W(X)$ with its equivalence class in $F(X)$, and denote the (group) inverse of w by w^{-1}.

A word $w = a_1 \dots a_k \in W(X)$ is *reduced* if it does not contain two successive letters that are inverses of each other; if, in addition, a_k is not the inverse of a_1, then w is *cyclically reduced*. Each element in $F(X)$ has a unique representation as a reduced word. If $w_i, 1 \le i \le q$, are words such that in forming the product $z = w_1 \dots w_q$ there is no *cancellation* – i.e., no $w_i, 1 \le i \le q-1$, ends in a letter s.t. w_{i+1} begins with the inverse of that letter – write $z \equiv w_1 \dots w_q$. A *conjugate* of a word w is a word of the form ywy^{-1}, where $y \in W(X)$ is arbitrary. A subset R of $F(X)$ is called *symmetrized* if all elements of R are cyclically reduced and, for each $r \in R$, all cyclic permutations, i.e., cyclically reduced conjugates, of both r and r^{-1} are in R as well.

A group G has a *finite presentation* $(X; R)$, where X is a finite set of symbols and $R \subset W(X)$ is also finite, if G is isomorphic to the quotient group $F(X)/N$, where N is the smallest normal subgroup of $F(X)$ containing R. N is termed the *normal closure* of R in $F(X)$. We say that G is generated by X with the *relations* in R, and write $G = (X; R)$ (it may be seen that if R is not already symmetrized it may be extended to be so without changing G).

For example, $(x; x^3)$ is the 3-element cyclic group, and $(x, y; xyx^{-1}y^{-1})$ is the product of two infinite cyclic groups, i.e, a free abelian group on two generators.

The *word problem* for a group $G = (X; R)$ asks for an algorithm to decide if an element $w \in W(X)$ represents the identity of G. In general, the word problem is unsolvable – P. S. Novikov [15] shows a finitely presented group with an unsolvable word problem. The *conjugacy problem*, generally unsolvable as well, asks to decide if two elements $w_1, w_2 \in W(X)$ represent conjugate elements of G, i.e., if there is a $c \in W(X)$ such that $w_1 = cw_2c^{-1}$ in G.

Assume now that $G = (X; R)$ is a finite presentation where R is symmetrized. If r_1 and r_2 are distinct elements of R such that $r_1 \equiv bc_1$ and $r_2 \equiv bc_2$, then b is called a *piece* of R. Consider the product $r_1^{-1}r_2$ to see that a piece is a subword of an element of R that can be non-trivially cancelled by multiplication with another element of R. R satisfies the *small cancellation* condition $C'(\lambda)$, for a real $\lambda > 0$, if $r \equiv bc$, where $r \in R$ and b is a piece of R, implies that $|b| < \lambda|r|$. A word $w \in F(X)$ is *R-reduced* if it is reduced and does not contain a subword w' such that there exists a relation $r \in R$ with $r \equiv w'w''$ and $|w'| > \frac{1}{2}|r|$.

The following consequence of Greendlinger's Lemma for Sixth-Groups [10] (see also [13]) is crucial to an efficient solution of the word problem for fundamental groups of surfaces:

Proposition 1 *If $G = (X; R)$ and R satisfies $C'(\frac{1}{6})$, then a non-empty reduced word $w \in W(X)$ that represents the identity element of G must contain a subword w' such that there exists a relation $r \in R$ with $r \equiv w'w''$ and $|w'| > \frac{1}{2}|r|$. In other words, a non-empty R-reduced word cannot represent the identity.* ♣

A word $w \in F(X)$ is *cyclically R-reduced* if it is cyclically reduced and all its cyclic permutations are R-reduced.

Another consequence of Greendlinger's results that we shall use to solve the conjugacy problem for fundamental groups of surfaces is:

Proposition 2 *If $G = (X; R)$ and R satisfies $C'(\frac{1}{8})$, then two non-empty cyclically R-reduced words $w_1, w_2 \in W(X)$ represent conjugate elements of G if and only if the equation $w_1^* = hw_2^*h^{-1}$ holds in G, where w_1^* and w_2^* are cyclically reduced conjugates of w_1 and w_2, respectively, and h is a subword of some relation $r \in R$.* ♣

Note that Greendlinger's original results imply much more than Propositions 1 and 2, but these are sufficient for our purposes.

For further discussions of group theory refer to Rotman [20], and for combinatorial group theory, in particular, to Lyndon and Schupp [13].

2.2 Homotopy and Fundamental Groups

In the following, by a *map* we shall always mean a continuous function from one topological space to another. Two maps $\phi : T_1 \to T_2$ and $\psi : T_1 \to T_2$ are *(freely) homotopic*, denoted $\phi \simeq \psi$, if there exists a map $H : [0,1] \times T_1 \to T_2$ such that $H(0,*) = \phi(*)$ and $H(1,*) = \psi(*)$. If $p \in T_1$, then the homotopy H *fixes* p if $H(*,p)$ is constant.

Two spaces T_1 and T_2 are of the same *homotopy type* if there exist maps $\phi : T_1 \to T_2$ and $\psi : T_2 \to T_1$, called *homotopy equivalences*, such the $\psi\phi \simeq 1_{T_1}$, the identity on T_1, and $\phi\psi \simeq 1_{T_2}$. If $T_2 \subset T_1$, then T_2 is a *deformation retract* of T_1 if there is a *retract* $r : T_1 \to T_2$ (i.e., a map with $r(t) = t$, for all $t \in T_2$) which is homotopic to 1_{T_1} by a homotopy that fixes points of T_2. In this case, the inclusion map from T_2 to T_1 is a homotopy equivalence.

Let $I = [0,1]$ be the unit interval. Then, a *path* in T from q_1 to q_2 is a map $g : I \to T$, such that $g(0) = q_1$ and $g(1) = q_2$. T is *path-connected* if, given arbitrary points $q_1, q_2 \in T$, there is a path in T from q_1 to q_2.

Let $S^1 = \{(x,y) \in \Re^2 : x^2 + y^2 = 1\}$ be the circle of unit radius. Then, a *closed curve* (or *cycle*) in T is a map $f : S^1 \to T$. Choose a *base-point* $p \in S^1$, and consider it fixed for the rest of this discussion. Then, for a point $q \in T$, the set of (base-point preserving) cycles through q consists of maps $f : S^1 \to T$, such that $f(p) = q$.

The product of two cycles f and g through q is the cycle $f \circ g$ that is defined as the concatenation of f followed by g. The inverse f^{-1} of a cycle f through q is obtained by reversing the orientation of the map $f : S^1 \to T$. Two cycles through q are *equivalent* if there is a homotopy between them that fixes p. The set of equivalence classes of cycles through q forms a group under the induced product and inverse operations. This group is the *fundamental group* of T at q, and denoted $\pi(T,q)$. We shall, henceforth, not distinguish between a cycle and its equivalence class in $\pi(T,q)$, or, often, its image on T. The identity element of $\pi(T,q)$ is the trivial cycle, which is a constant function. If T is path-connected, then its fundamental groups at all points are isomorphic, so we can refer simply to $\pi(T)$. If, in addition, $\pi(T)$ is trivial, then T is *simply connected*.

A map $\phi : T_1 \to T_2$ induces a homomorphism $\phi_* : \pi(T_1) \to \pi(T_2)$ by $(\phi_*(f))(x) = \phi(f(x))$. If ϕ is a homotopy equivalence then ϕ_* is an isomorphism.

The following proposition describes how to translate certain topological properties of cycles into discrete properties of corresponding elements in the fundamental group (see Singer and Thorpe [18]):

Proposition 3 *Let $f : S^1 \to T$ be an arbitrary cycle. Then f is (freely) homotopic to some trivial cycle if and only if $f = 1$ in $\pi(T, f(p))$, and this is independent of the choice of the base-point $p \in S^1$. In this case, f is said to be null-homotopic or contractable.*

Let $f_1, f_2 : S^1 \to T$ be two arbitrary cycles. Then $f_1 \simeq f_2$, i.e. f_1 and f_2 are freely homotopic, if and only if f_1 and $g \circ f_2 \circ g^{-1}$ represent conjugate elements in $\pi(T, f_1(p))$, where g is any path in T from $f_1(p)$ to $f_2(p)$. Again, this is independent of both the choice of the base-point $p \in S^1$ and the choice of the path g from $f_1(p)$ to $f_2(p)$. ♣

The following proposition describing a method to compute the fundamental group of a union of two spaces is a consequence of the theorem of Seifert and Van Kampen (see Massey [14]):

Proposition 4 *If $T = U \cup V$, where U, V, and $U \cap V$ are path-connected open subspaces of T, and $x \in U \cap V$, and V is simply connected, then $\pi(T,x)$ is the quotient group of $\pi(U,x)$ modulo the normal closure of the image of $\pi(U \cap V, x)$ in $\pi(U,x)$ by inclusion.* ♣

For more complete discussions of algebraic topology refer to Massey [14], Singer and Thorpe [18], and Stillwell [19].

2.3 Closed Surfaces

A *surface*, or *2-manifold*, is a topological space where each point has a neighborhood homeomorphic to an open disk in $\Re^2$. With this definition we restrict ourselves to *boundaryless*, or *closed*, surfaces. Further, we shall consider only surfaces that are compact and connected.

A *triangulation* T of a surface M consists of a decomposition of M into triangles (or, more precisely, subspaces homeomorphic to closed triangles on the plane) such that

(a) any two distinct triangles intersect in either one common vertex, or along one common edge, or are disjoint,

(b) each edge belongs to exactly two triangles, and,

(c) the triangles intersecting at each vertex can be ordered circularly so that two triangles in this ordering intersect along an edge if and only if they are adjacent in the ordering.

All surfaces are known to be *triangulable*, i.e., possess a triangulation.

A surface M is *orientable* if, given a triangulation T of M, an orientation can be chosen of the boundary of each triangle in T such that any two triangles always induce opposite orientations on a common edge, in case they share one. Otherwise, M is *non-orientable*. For example, the sphere and torus are orientable, while the projective plane and Klein bottle are non-orientable.

The *connected sum* of two surfaces M_1 and M_2 is the surface, denoted $M_1 \# M_2$, obtained by cutting a disk-like hole in each of M_1 and M_2 and attaching them along the boundaries of these holes. A famous classification theorem (see [14]), due to Brahana, and Dehn and Heegaard, states that an orientable surface is either a sphere or a connected sum of finitely many tori, while a non-orientable surface is a connected sum of finitely many projective planes. A sphere has *genus* 0, while any other surface is said to have genus g (≥ 1) if it is the connected sum of either g tori or g projective planes.

A classic representation for surfaces is by a *polygonal schema*. A polygonal schema consists of a polygon P with an even number, say $2m$, of edges. The edges are labeled by symbols from the set $\{x_1, x_1^{-1}, \ldots, x_m, x_m^{-1}\}$ such that each unsigned symbol, i.e., ignoring inverse signs, occurs exactly twice. Two edges with the same unsigned label are *partners*. Edges labeled by a symbol with an inverse sign are oriented in a direction opposite, along the boundary $bd(P)$ of P, to that of edges labeled by a symbol without an inverse sign. A polygonal schema P represents the surface M that is obtained from P by attaching each partnered pair of edges of P so that orientations match. A polygonal schema for a double torus is shown in Figure 1. Since the surface M is identified up to homeomorphism by the sequence of labels around the boundary of P in, say, clockwise order, we identify the the polygonal schema itself by this sequence. For example, the polygonal schema of Figure 1 may be defined by the sequence $x_1 y_1 x_1^{-1} y_1^{-1} x_2 y_2 x_2^{-1} y_2^{-1}$, or any of its cyclic permutations.

An orientable surface of genus $g > 0$ can be represented *canonically* by the polygonal schema $x_1 y_1 x_1^{-1} y_1^{-1} \ldots x_g y_g x_g^{-1} y_g^{-1}$ of size $4g$; in fact, it cannot be represented by a shorter polygonal schema. Similarly, a non-orientable surface can be represented canonically and minimally by the polygonal schema $x_1 x_1 x_2 x_2 \ldots x_g x_g$ of size $2g$. The exceptional canonical schema representing the sphere is xx^{-1}.

The following proposition describes the fundamental groups of surfaces (see Massey [14] and Stillwell [19]):

Proposition 5 *The fundamental group of the sphere is trivial.*

For an orientable surface M of genus $g \geq 1$, $\pi(M)$ has a finite presentation

$$\pi(M) = (x_1, y_1, \ldots, x_g, y_g \ ; \ x_1 y_1 x_1^{-1} y_1^{-1} \ldots x_g y_g x_g^{-1} y_g^{-1}).$$

For a non-orientable surface M of genus $g \geq 1$, $\pi(M)$ has a finite presentation

$$\pi(M) = (x_1, \ldots, x_g \; ; \; x_1^2 \ldots x_g^2).$$

Surfaces are identified up to homeomorphism by their fundamental groups, and surfaces of the same homotopy type are homeomorphic. ♣

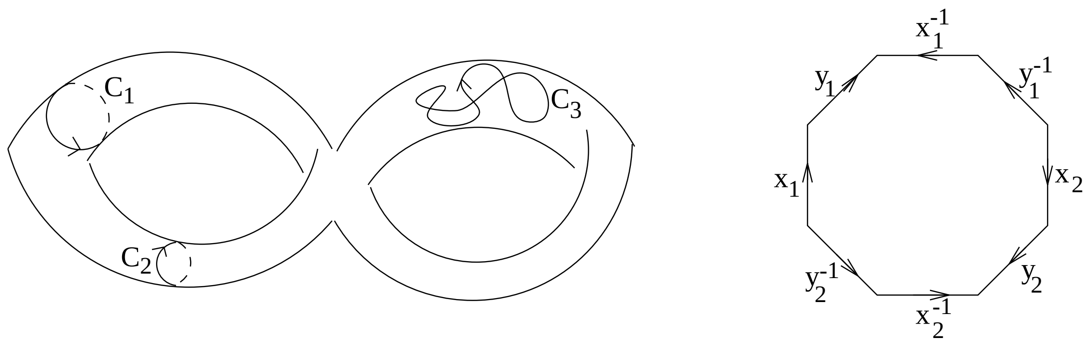

Figure 1: A double torus, an orientable surface of genus 2, and a polygonal schema in *canonical* form for the double torus. Cycles C_1 and C_2 are homotopic to each other but not to cycle C_3.

3 The Algorithm

The input to the algorithm includes a triangulation T, of size n, of a surface M, together with cycles C_1 and C_2 on M presented as edge-vertex sequences in T of length k_1 and k_2, respectively. Assume that T is represented by a data structure that allows access to the edges of a triangle, as well as the triangles incident on an edge, in $O(1)$ time. The quad-edge data structure of Guibas and Stolfi [11] can be used for this purpose. Further, for a reason that will be apparent later, assume M to be a manifold of genus g, where $g \geq 3$ if M is orientable and $g \geq 5$ if M is non-orientable.

Say v_1 and v_2 are two vertices on C_1 and C_2, respectively. Find *any* path D from v_1 to v_2, by, say, an $O(n)$-time breadth-first search of the 1-skeleton of T, so that D is an edge-vertex sequence of length $O(n)$. By Proposition 3, $C_1 \simeq C_2$ *if and only if* C_1 and $D \circ C_2 \circ D^{-1}$ represent conjugate elements in the fundamental group $\pi(M, v_1)$ $(= \pi(M)$, as M is path-connected). Therefore, to avoid clumsy notation later, we shall fudge a little now and assume that we are, in fact, given C_1 and $D \circ C_2 \circ D^{-1}$ as input, and denote the latter as C_2. Assume as well that the edge-vertex sequence representing C_1 is $v_{1,1}e_{1,1}v_{1,2} \ldots e_{1,k_1}v_{1,k_1+1}$ and that representing C_2 is $v_{2,1}e_{2,1}v_{2,2} \ldots e_{2,k_2}v_{2,k_2+1}$, where $v_1 = v_{1,1} = v_{2,1} = v_{1,k_1+1} = v_{2,k_2+1}$. Neither the extra time to find D nor the extra $O(n)$ part in k_2 due to the "hidden" D and D^{-1} affect our future claims on time and space, as these claims are of the form $O(n + k_1 + k_2)$.

The algorithm consists of two phases: the first phase converts the geometric problem of deciding transformability to the algebraic one of deciding if two elements in a group are conjugate, while the second phase solves this conjugacy problem. It is in the second phase that we apply our new combinatorial approach.

3.1 Phase 1 - Geometric

This phase consists of two subphases similar to procedures in Dey [6]. However, in order to make this discussion self-contained, we include brief descriptions.

3.1.1 Subphase 1a - Finding a Polygonal Schema

In this subphase we find a *polygonal schema* (see [14, 19, 22]) P representing M such that P has a triangulation T' containing the same number of triangles as T. We also find representations of C_1 and C_2 on P.

The procedure is to construct a sequence of polygons $P_1, \ldots, P_n$ incrementally on the plane such that finally $P = P_n$. Initially, set $P_1 = \sigma'_1$, a triangle in the plane that corresponds to an arbitrarily chosen triangle $\sigma_1 \in T$. The correspondence between σ_1 and σ'_1 specifies an identification between their vertices as well.

Inductively, assume that $P_i = \sigma'_1 \cup \ldots \cup \sigma'_i$ after the ith step, where each triangle $\sigma'_r, 1 \leq r \leq i$, corresponds to a distinct triangle $\sigma_r \in T$. At the $i+1$ th step choose a triangle $\sigma_{i+1} \in T$ such that

(a) no triangle corresponding to σ_{i+1} has been included in P_i, and,

(b) a triangle σ'_j corresponding to some triangle σ_j adjacent to σ_{i+1} in T has been included in P_i.

These two conditions imply that there is an edge $e = \sigma_j \cap \sigma_{i+1}$ such that its corresponding edge e' in P_i appears on $bd(P_i)$. Attach a triangle σ'_{i+1}, corresponding to σ_{i+1}, to $bd(P_i)$, so that $\sigma'_{i+1} \cap P_i = e'$, and so that the identification of vertices of e' specified by σ'_j matches that specified by σ'_{i+1}. This gives P_{i+1}.

After the nth step we have $P_n = P$ with a triangulation T', consisting of triangles $\sigma'_r, 1 \leq r \leq n$, such that

(a) there is a one-to-one correspondence between the triangles of T and T', together with a vertex-to-vertex identification specified for each pair of corresponding triangles,

(b) each edge e' on $bd(P)$ has a partner edge e'' on $bd(P)$ such that they both correspond to a single edge e in T; further, assuming some arbitrary orientation on the edges of T, we have an orientation on the edges of $bd(P)$ induced by the vertex-to-vertex identification specified in the correspondence between triangles of T and T', and,

(c) we can obtain M by attaching partnered edges of $bd(P)$, taking care to match orientations, as given in (b), when attaching edges. More precisely, there is a homeomorphism ϕ from M to the quotient space of P modulo the identification of partnered edges.

Thus, appropriately labeling the edges of $bd(P)$, so that partnered edges have the same unsigned symbol and signs represent orientations, P is indeed a polygonal schema for M such that P and M have equal sized triangulation. Say, $bd(P)$ has edges labeled by symbols from the set $\{x_1, x_1^{-1}, \ldots, x_m, x_m^{-1}\}$, such that each unsigned symbol occurs exactly twice – either as a pair x_i, x_i^{-1} or as a pair x_i, x_i.

Next, considering first the cycle C_1, we see that its homeomorphic image by ϕ is an edge-vertex sequence C'_1 that consists of a possibly "disconnected" circular sequence $C'_{1,1}, \ldots, C'_{1,h_1}$, $h_1 \leq k_1$, of *arcs* such that

(a) each arc $C'_{1,j}$ consists of a connected edge-vertex sequence $v_{i_j} e_{i_j} v_{i_j+1} \ldots v_{i'_j}$, where only the first vertex v_{i_j} and last vertex $v_{i'_j}$ of the sequence lie on $bd(P)$, and,

(b) For $1 \leq j \leq h_1$, the last vertex $v_{i'_j}$ of $C_{1,j}$ and the first vertex $v_{i_{j+1}}$ of $C_{1,j+1}$ on $bd(P)$ are identified by the partnering of oriented edges of $bd(P)$ (of course, "$h_1 + 1 = 1$").

Similar remarks apply to C_2 so that its homeomorphic image by ϕ is an edge-vertex sequence C'_2 that consists of the circular sequence of arcs $C'_{2,1}, \ldots, C'_{2,h_2}$, $h_2 \leq k_2$.

See Figure 3, forgetting for purpose of convenient illustration the restriction that genus $g \geq 3$ if M is orientable.

It is easily verified that this subphase completes in time $O(n + k_1 + k_2)$.

3.1.2 Subphase 1b - Reducing the Polygonal Schema

The size of the polygonal schema P, i.e., the number of edges on $bd(P)(= 2m)$, found in Subphase 1a may be $\Theta(n)$. In this subphase we find a polygonal schema Q for M which is of minimal size. Such a polygonal schema Q is called a *reduced* polygonal schema for M (see [6, 22]), and, in fact, $bd(Q)$ has $4g$ or $2g$ edges, according as M is orientable or not.

Denote by G the 1-complex, i.e., graph, formed by taking $bd(P)$ and identifying partnered edges so that orientations match along identified edges. Let Y be a spanning tree of G. See Figure 2 for a simple example.

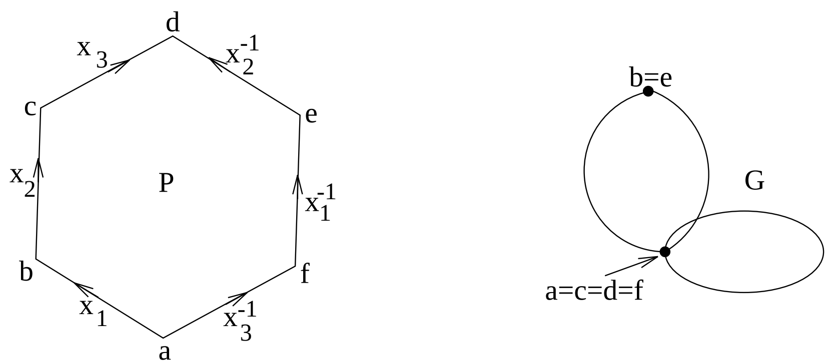

Figure 2: A polygonal schema P for a torus, and the 1-complex G corresponding to P: a possible spanning tree Y of G has exactly one edge ab.

Let $B = \{b_1, \ldots, b_l\}$ be the set of edges of G *not in* Y. Call the edges of G *in* Y *excess*. Form the polygonal schema Q as follows: proceed through the sequence of symbols that define P, i.e., the labels of $bd(P)$, deleting those that correspond to an excess edge of G. Recall that each edge of G was formed by identifying two partnered edges, so that each deletion of an excess edge will result in the deletion of a partnered pair of symbols. Call such deleted symbols *excess symbols*. Declare Q to be the polygonal schema defined by the sequence of symbols that remain after deleting excess symbols from P. Clearly, the length of this sequence is $2l$, as 2 symbols remain for each edge in B. Let us write this sequence as $y_1 \ldots y_{2l}$, where each unsigned symbol y_i is one of $\{b_1, \ldots, b_l\}$, and the sign is assigned according to orientation. See Figure 3 for a less trivial example (G is not shown).

Now, the projection map from G to the quotient space G/Y may be verified to be a homotopy equivalence – G/Y may be thought of as G with spanning tree Y contracted to a point. Considering G as a subspace of M via the homeomorphism ϕ (of Subphase 1a), this homotopy equivalence extends to the projection from M to M/Y. However, M/Y is homeomorphic to the manifold M' represented by the polygonal schema Q. Thus, we have a projection $\psi : M \to M'$ which is a homotopy equivalence. It follows that, as surfaces, M and M' are homeomorphic, and Q may be considered a polygonal schema for M. We show next that Q is, in fact, a reduced polygonal schema for M, so that $l = 2g$ if M is orientable, and $l = g$ if M is not orientable.

Consider M' as the appropriate quotient space of polygonal schema Q. Now, the subspace $D = Q - bd(Q)$ is an open disk in M'. Choose a point $p \in D$, and let $U = M' - \{p\}$. Since

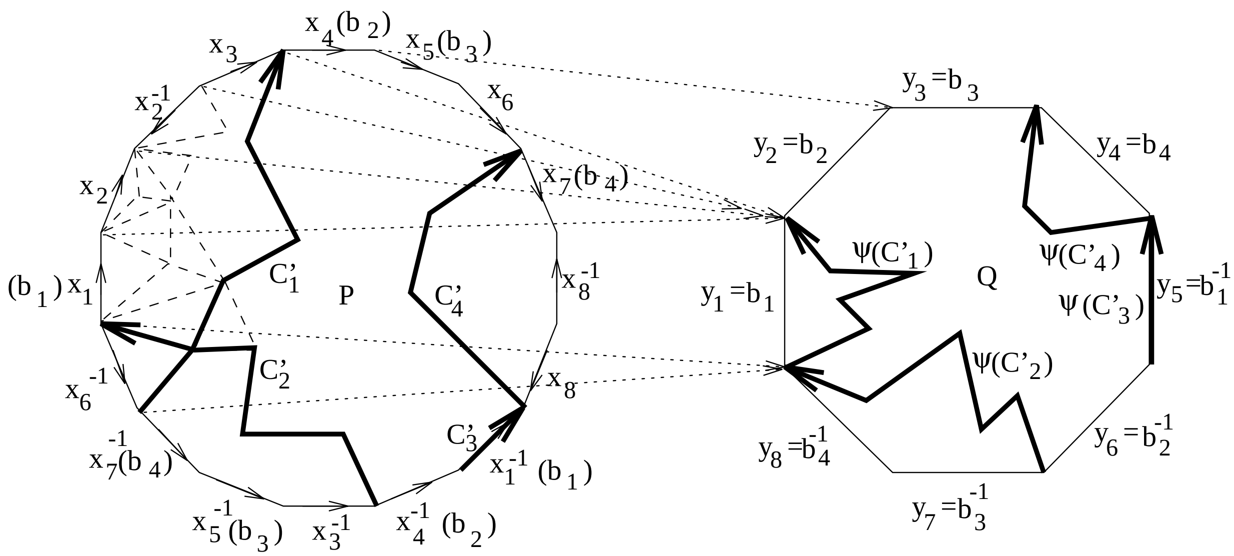

Figure 3: A polygonal schema P for a double torus and its reduction Q, *not* in canonical form: (i) The dashed lines show part of the triangulation T' of P, (ii) The bold lines in P and Q show cycles C' and $\psi(C')$, respectively, (iii) A symbol b_i in parentheses indicates that the corresponding edge of P is associated with the edge b_i of B, (iv) The dotted lines indicate some vertex mappings by ψ, (v) $\psi(C') = (1,1)(7,8)(6,4)(5,3)$.

$M' = U \cup D$, and D is simply connected, Proposition 4 implies that $\pi(M')$ is $\pi(U)$ modulo the normal closure of the image of $\pi(U \cap D)$ in $\pi(U)$ by the inclusion map. It may be seen that, after identification of partnered edges, $bd(Q)$ is a subspace of M', call it G', where all vertices of $bd(Q)$ are identified to a single point, say $q \in G'$. Furthermore, G' itself is the union of l cycles, through q, each cycle arising from the identification of a pair of edges of $bd(Q)$ that are labeled with the same unsigned symbol in $B = \{b_1, \ldots, b_l\}$. Thus, $\pi(G')$ is the free group on generators $\{b_1, \ldots, b_l\}$ (see [14]), where b_i represents the cycle derived from the identification of the pair of edges of $bd(Q)$ with unsigned label b_i. Since G' is a deformation retract of U, $\pi(U)$ is also the free group on generators $\{b_1, \ldots, b_l\}$. Observe now that $U \cap D$ has the homotopy type of a circle, so $\pi(U \cap D)$ is the infinite cyclic group with generator a cycle c in $U \cap D$. The image of c in $\pi(U)$ by the inclusion map is the product $y_1 \ldots y_{2l}$. To see this, "push" the cycle c all the way to $bd(Q)$.

Thus, $\pi(M') = (b_1, \ldots, b_l; y_1 \ldots y_{2l})$, and from this finite presentation, and the fact the M' has the same genus g and orientability as M, it is deduced that $l = 2g$ if M is orientable, and $l = g$ if M is not orientable, so Q is indeed a reduced polygonal schema for M.

Since ψ is a homotopy equivalence, deciding the conjugacy of C_1 and C_2 in $\pi(M)$ is equivalent to deciding the conjugacy of $\psi(C_1)$ and $\psi(C_2)$ in $\pi(M')$. Let us now compute the element $z_\alpha \in \pi(M')$ that the cycle $\psi(C_\alpha)$ represents, $\alpha = 1, 2$.

First, we make some observations about the projection ψ. Consider ψ as a map from P to Q, by identifying points of P and Q with the corresponding points in the respective quotient spaces M and M'. Then, ψ projects $bd(P)$ onto $bd(Q)$, and the interior $int(P)$ of P onto $int(Q)$. It is only $\psi|bd(P)$ that concerns us. Specifically, the behavior of ψ on $bd(P)$ is as follows: ψ projects each edge on $bd(P)$, labeled with a non-excess symbol, onto the corresponding edge on $bd(Q)$; each sequence of edges on $bd(P)$ labeled with excess symbols, that lies between two edges e and e' labeled with non-excess symbols, is projected to the common endpoint between the edges corresponding to e and e' on $bd(Q)$.

Next, consider $\psi(C_1)$: regarding ψ as a map from P to Q, $\psi(C_1)$ is the same as $\psi(C_1')$. And, from Subphase 1a we have the arc sequence $C_{1,1}', \ldots, C_{1,h_1}'$ of C_1'. Now, the image $\psi(C_{1,j}') = \psi(v_{i_j} e_{i_j} v_{i_j+1} \ldots v_{i_j'})$ starts at the vertex $\psi(v_{i_j})$ and ends at the vertex $\psi(v_{i_j'})$, both on $bd(Q)$, and no interior vertex of $\psi(C_j')$ lies on $bd(Q)$. Let the sequence of symbols labeling edges of $bd(Q)$ *clockwise* between $\psi(v_{i_j})$ and $\psi(v_{i_j'})$ be $y_{r_{1,j}}, \ldots, y_{s_{1,j}}$ Note that $r_{1,j}$ may be greater than $s_{1,j}$ – the list $y_1 y_2 \ldots y_{2l}$ is *clockwise circular* around $bd(Q)$. Then, the path $\psi(C_{1,j}')$ is homotopic, with end-points fixed, to the cycle on M' that is represented in $\pi(M')$ by the product $y_{r_{1,j}} \ldots y_{s_{1,j}}$. Denote this product by the term $(r_{1,j}, s_{1,j})$. Completing the traversal of the arc sequence $C_{1,1}', \ldots, C_{1,h_1}'$ of C_1', we have a representation of the cycle $\psi(C_1') = \psi(C_1)$ as a product $z_1 = (r_{1,1}, s_{1,1}) \ldots (r_{1,h_1}, s_{1,h_1})$, where $h_1 \leq k_1$, in $\pi(M') = (b_1, \ldots, b_l; y_1 \ldots y_{2l})$. See Figure 3.

Similarly, we can find a representation of the cycle $\psi(C_2') = \psi(C_2)$ as a product
$z_2 = (r_{2,1}, s_{2,1}) \ldots (r_{2,h_2}, s_{2,h_2})$, where $h_2 \leq k_2$.

It may be verified that this subphase completes in time $O(n + k_1 + k_2)$ as well.

Remark. Subphase 1b could as well have been designed around the more "surgical" algorithm to find a reduced polygonal schema given by Vegter and Yap [22].

To solve the original homotopy problem we now have to determine if z_1 and z_2 represent conjugate elements in $\pi(M')$. We solve this algebraic problem in Phase 2.

3.2 Phase 2 - Algebraic

We have from Phase 1 words, each expressed as a product of terms,

$z_1 = (r_{1,1}, s_{1,1}) \ldots (r_{1,h_1}, s_{1,h_1})$ and $z_2 = (r_{2,1}, s_{2,1}) \ldots (r_{2,h_2}, s_{2,h_2})$ in $F(b_1 \ldots, b_l)$,
and must decide if they represent conjugate elements in the group $\pi(M') = (b_1, \ldots, b_l; y_1 \ldots y_{2l})$.

3.2.1 Subphase 2a - Rectification, and More Theory

First, we describe certain notations.

If $w_1, w_2 \in F(b_1, \ldots, b_l)$ are identical strings denote this as $w_1 = w_2$; if w_1 and w_2 represent equal elements of $\pi(M')$ denote this as $w_1 \approx w_2$. Of course, $w_1 = w_2$ implies that $w_1 \approx w_2$.

Denote by $\overline{(u,v)}$ the product given by the subword $y_u^{-1} \ldots y_v^{-1}$ of the circular list of symbols $y_{2l}^{-1} y_{2l-1}^{-1} \ldots y_1^{-1}$ (with y_{2l}^{-1} following y_1^{-1}: imagine traversing Q *circularly counter-clockwise*). If either (r,s) or $\overline{(u,v)}$ consists of a single symbol, always write it as that symbol (e.g., (r,r) is written as y_r, $\overline{(u,u)}$ as y_u^{-1}, where, of course, both y_r and y_u^{-1} are symbols of the form b_i or b_i^{-1}, $1 \leq i \leq l$). Let $|(r,s)|$ and $|\overline{(u,v)}|$ denote the length of the product that each represents. A consequence of the representation $\pi(M') = (b_1, \ldots, b_l; y_1 \ldots y_{2l})$ is:

Lemma 1 *The following hold:*

1. $(r,s) \approx \overline{(r-1, s+1)}$.
2. $(r,s)^{-1} \approx (s+1, r-1) \approx \overline{(s,r)}$.
3. $\overline{(r,s)} \approx (r+1, s-1)$.
4. $\overline{(r,s)}^{-1} \approx \overline{(s-1, r+1)} \approx (s,r)$
5. $|\overline{(r-1, s+1)}| = |(s+1, r-1)| = 2l - |(r,s)|$. ♣.

Subphase 2a consists simply of preprocessing z_α, $\alpha = 1, 2$, as follows:

For each term $(r_{\alpha,j}, s_{\alpha,j})$ of z_α,

(a) if $1 \leq |(r_{\alpha,j}, s_{\alpha,j})| \leq l$, leave it unchanged (if $|(r_{\alpha,j}, s_{\alpha,j})| = 1$ it is written as a single symbol, of course),

(b) if $l < |(r_{\alpha,j}, s_{\alpha,j})| < 2l-1$, replace it by $\overline{(r_{\alpha,j} - 1, s_{\alpha,j} + 1)}$,

(c) if $|(r_{\alpha,j}, s_{\alpha,j})| = 2l-1$, replace it by the single symbol $y^{-1}_{r_{\alpha,j}-1}$, and,

(d) if $|(r_{\alpha,j}, s_{\alpha,j})| = 2l$, delete it.

The preprocessing in Subphase 2a gives, for $\alpha = 1, 2$, a product $z'_\alpha = c_{\alpha,1} \ldots c_{\alpha,h'_\alpha}$, s.t. $h'_\alpha \leq h_\alpha$ and $z'_\alpha \approx z_\alpha$ by Lemma 1, of terms $c_{\alpha,i}$ either of the form (r, s), s.t. $2 \leq |(r, s)| \leq l$, *or* $\overline{(r, s)}$, s.t. $2 \leq |\overline{(r, s)}| \leq l-1$, *or* y_r. Call such terms *rectified terms*, and such a product a *rectified product*. Denote by $height(w)$ the number of rectified terms in a rectified product w.

E.g., rectification of the product $(1,1)(7,8)(6,4)(5,3) = \psi(C')$ in Figure 3 gives $b_1(7,8)b_1b_4^{-1}$.

Subphase 2a completes in time $O(k_1 + k_2)$.

More Theory

Let $y = y_1 \ldots y_{2l}$ and let R denote the set of $4l$ relations, each of length $2l$, consisting of y, y^{-1}, and all their cyclic permutations. It may be seen that $\pi(M') = (b_1, \ldots, b_l; R)$, where R is now symmetrized (see Section 2.1). For, clearly $(b_1, \ldots, b_l; y) = (b_1, \ldots, b_l; R)$. Further, y and, therefore, y^{-1} and all their cyclic permutations are cyclically reduced as, otherwise, it contradicts that Q is a reduced polygonal schema.

Another crucial consequence of Q being a reduced polygonal schema is:

Lemma 2 *Considering y to be circular sequence of symbols of length $2l$, a given pair of adjacent symbols y_iy_{i+1} at position i in y cannot occur at any other position in y, and the pair $y^{-1}_{i+1}y^{-1}_i$ cannot occur at all in y (i.e., y_iy_{i+1} cannot occur at all in y^{-1}).*

Proof. The reason is geometric. If this were not true, then $bd(Q)$ would have four edges in either one of the two forms depicted in Figure 4. In either case the edge pairs y_iy_{i+1} could both be replaced by one edge, contradicting that Q is a reduced polygonal schema. ♣

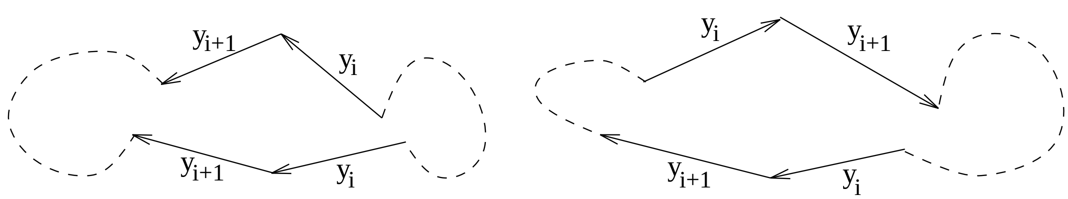

Figure 4: Illustration for Lemma 2.

This allows us to employ Greendlinger's powerful results:

Lemma 3 *The set of relations R satisfies the small cancellation condition $C'(\frac{1}{4g-1})$ if M' is orientable and $C'(\frac{1}{2g-1})$ if M' is non-orientable. Therefore, with our restriction that $g \geq 3$ if M is orientable and $g \geq 5$ if M is non-orientable, R always satisfies $C'(\frac{1}{8})$, and so also $C'(\frac{1}{6})$.*

Proof. Suppose $r_1, r_2 \in R$ are distinct relations such that there is a word b with $r_i \equiv bc_i$, $i = 1, 2$, so that b is a piece (see Section 2.1) of R. Since r_1 and r_2 are both cyclic permutations of either y or y^{-1}, it follows from Lemma 2 that $|b| \leq 1$. If M' is orientable, when $2l = 4g$ so $|r| = 4g$ for each $r \in R$, we have, therefore, $|b| < \frac{1}{4g-1}|r|$. This implies that R satisfies the small cancellation condition $C'(\frac{1}{4g-1})$. If M' is non-orientable, when $2l = 2g$, R satisfies $C'(\frac{1}{2g-1})$. ♣

To motivate the next definition consider a non-empty rectified product w such that $w \approx 1$. Assume that w is reduced. Then, by Proposition 1, w contains a subword w' such that there exists a relation $r \in R$ with $r \equiv w'w''$ and $|w'| > \frac{1}{2}|r|$. In which case, w can be "shortened" to a word $\hat{w} \approx w$ of smaller length, by replacing w' in w using Lemma 1 with a shorter subword $\hat{w}'$ such that $\hat{w}' \approx w'$.

Let us say that an ordered pair dc of rectified terms *react* if the rectified product $w = dc$ can be replaced by a rectified product $\hat{w}(d,c) \approx w$ such that $|\hat{w}(d,c)| \leq w$ and $height(\hat{w}(d,c)) \leq height(w)$, and such that *at least one* of the two inequalities holds strictly. If d and c are clear from the context we write $\hat{w}(d,c)$ simply as $\hat{w}$.

We determine in the following table when an ordered pair of rectified terms react, and give a C-like procedure that uniquely determines $\hat{w}$ in each case. There are 9 mutually exclusive possibilities for the forms of d and c corresponding to d being one of the forms (u,v), $\overline{(u,v)}$, or y_u, together with c being one of the forms (r,s), $\overline{(r,s)}$, or y_r. These are labeled 1-9. For each, the cases when dc react are labeled with letters.

1. if $d = (u,v)$ && $c = (r,s)$ {
   (a) : if $(v + 1 = r)$
   {
   if $(|(u,s)| \leq l)$ $\hat{w} = (u,s)$;
   if $(l < |(u,s)| < 2l)$ $\hat{w} = \overline{(u-1,s+1)}$;
   if $(|(u,s)| = 2l)$ $\hat{w} = \epsilon$;
   }
   (b) : else if $(y_{v+1} = y_r$ && $|(u,v)| = l)$ $\hat{w} = \overline{(u-1,v+2)}(r+1,s)$;
   (c) : else if $(y_v = y_{r-1}$ && $|(r,s)| = l)$ $\hat{w} = (u,v-1)\overline{(r-2,s+1)}$;
   (d) : else if $(y_v = y_r^{-1})$ $\hat{w} = (u,v-1)(r+1,s)$; }

2. if $d = \overline{(u,v)}$ && $c = (r,s)$ {
   (a) : if $(v = r)$
   {
   if (u appears before s going clockwise from r) $\hat{w} = (u+1,s)$;
   if (u appears after s going clockwise from r) $\hat{w} = \overline{(u,s+1)}$;
   if $(u = s)$ then $\hat{w} = \epsilon$;
   }
   (b) : else if $(y_v^{-1} = y_{r-1}$ && $|(r,s)| = l)$ $\hat{w} = \overline{(u,v+1)}\,\overline{(r-2,s+1)}$;
   (c) : else if $(y_v = y_r)$ $\hat{w} = \overline{(u,v+1)}(r+1,s)$; }

3. if $d = y_u$ && $c = (r,s)$ {
   (a) : if $(y_u = y_{r-1}$ && $|(r,s)| < l)$ $\hat{w} = (r-1,s)$;
   (b) : else if $(y_u = y_{r-1}$ && $|(r,s)| = l)$ $\hat{w} = \overline{(r-2,s+1)}$;
   (c) : else if $(\ y_u = y_r^{-1})$ $\hat{w} = (r+1,s)$; }

4. if $d = (u,v)$ && $c = \overline{(r,s)}$ {
   (a) : if $(v = r)$
   {
   if (u appears before s going counter-clockwise from r) $\hat{w} = \overline{(u-1,s)}$;
   if (u appears after s going counter-clockwise from r) $\hat{w} = (u,s-1)$;
   if $(u = s)$ $\hat{w} = \epsilon$;
   }

(b) : else if $(y_{v+1} = y_r^{-1}$ && $|(u,v)| = l)$ $\hat{w} = \overline{(u-1, v+2)}\,\overline{(r-1, s)}$;

(c) : else if $(y_v = y_r)$ $\hat{w} = (u, v-1)\overline{(r-1, s)}$; }

5. if $d = \overline{(u,v)}$ && $c = \overline{(r,s)}$ {

(a) : if $(v - 1 = r)$

{

if $(|\overline{(u,s)}| \leq l-1)$ $\hat{w} = \overline{(u,s)}$;

if $(l - 1 < |\overline{(u,s)}| < 2l)$ $\hat{w} = (u+1, s-1)$;

if $(|\overline{(u,s)}| = 2l)$ $\hat{w} = \epsilon$;

}

(b) : else if $(y_v = y_r^{-1})$ $\hat{w} = \overline{(u, v+1)}\,\overline{(r-1, s)}$; }

6. if $d = y_u$ && $c = \overline{(r,s)}$ {

(a) : if $(y_u = y_{r+1}^{-1}$ && $|\overline{(r,s)}| < l-1)$ $\hat{w} = \overline{(r+1, s)}$;

(b) : else if $(y_u = y_{r+1}^{-1}$ && $|\overline{(r,s)}| = l-1)$ $\hat{w} = (r+2, s-1)$;

(c) : else if $(y_u = y_r)$ $\hat{w} = \overline{(r-1, s)}$; }

7. if $d = (u,v)$ && $c = y_r$ {

(a) : if $(y_{v+1} = y_r$ && $|(u,v)| < l)$ $\hat{w} = (u, v+1)$;

(b) : else if $(y_{v+1} = y_r$ && $|(u,v)| = l)$ $\hat{w} = \overline{(u-1, v+2)}$;

(c) : else if $(y_v = y_r^{-1})$ $\hat{w} = (u, v-1)$; }

8. if $d = \overline{(u,v)}$ && $c = y_r$ {

(a) : if $(y_{v-1}^{-1} = y_r$ && $|\overline{(u,v)}| < l-1)$ $\hat{w} = \overline{(u, v-1)}$;

(b) : else if $(y_{v-1}^{-1} = y_r$ && $|\overline{(u,v)}| = l-1)$ $\hat{w} = (u+1, v-2)$;

(c) : else if $(y_v = y_r)$ $\hat{w} = \overline{(u, v+1)}$; }

9. if $d = y_u$ && $c = y_r$ {

(a) : if (the sequence $y_u y_r$ occurs in y or y^{-1})

{

if $(y_u y_r = (u', r'))$ $\hat{w} = (u', r')$;

if $(y_u y_r = \overline{(u', r')})$ $\hat{w} = \overline{(u', r')}$;

}

(b) : else if $(y_u = y_r^{-1})$ $\hat{w} = \epsilon$; }

Table 1: Procedure to determine when dc react and corresponding $\hat{w}(d, c)$.

Remark 1. The conditional clauses in each of the 9 cases are, of course, checked sequentially as they are not exclusive. E.g., suppose

$$(1,6),\ (7,9),\ y_8\ \in\ (b_1, b_2, \ldots, b_6\ ;\ y_1 y_2 \ldots y_{12} = b_1^2 b_2^2 \ldots b_6^2).$$

Then, $w = (1,6)(7,9)$ satisfies the *if*-clause of both 1(a) and 1(b), but 1(a) is used to give $\hat{w} = \overline{(12, 10)}$; if $w = (1,6)y_8$, 1(b) gives $\hat{w} = \overline{(12, 8)}$.

Remark 2. An exact symmetry does not exist between (u,v) and $\overline{(u,v)}$, or (r,s) and $\overline{(r,s)}$, because of the restriction that $|\overline{(u,v)}|$, $|\overline{(r,s)}| < l$. E.g., the clauses corresponding to 1(b) and 1(c) are missing in case 5.

A rectified product $w = c_1 \dots c_h$ is called *stable* if no ordered pair $c_i c_{i+1}$, $1 \leq i \leq h-1$, of successive terms of w react.

Remark 3. It is easy to check in Table 1 that when dc react not only is $\hat{w} \approx dc$ and $\hat{w}$ of at least smaller length or height (or both) than dc, but $\hat{w}$ is stable.

If the ordered pair dc of rectified terms react, they are said to *0-react*, *1-react*, or *2-react* according as, from Table 1, $\hat{w}$ contains 0, 1, or 2 terms, respectively. If dc 0-react, of course, $d = c^{-1}$. The following technical lemma will prove of importance:

Lemma 4 *Suppose the rectified product ed is stable and c is a rectified term. Then, if dc react, either one of the following 3 occurs:*

1. *dc 0-react.*
2. *dc 1-react, in which case $\hat{w}(dc) = d_1$, say, where ed_1 either do not react or 2-react.*
3. *dc 2-react, in which case $\hat{w}(dc) = d_1 d_2$, say, where ed_1 either do not react or 2-react. Further, if ed_1 2-react then $\hat{w}(ed_1) = d_3 d_4$ where $d_4 d_2$ do not react.*

Proof. The proof consists of a somewhat tedious analysis of several cases using Table 1. There are 27 possibilities for the forms of e, d, and c corresponding to each being one of the forms (p,q), $\overline{(p,q)}$, or y_p. Analysis is quite easy in the cases when one of e, d or c is of the form y_u. Of the remaining 8 cases, we discuss one below. The others are similar and we can often avoid repeating arguments by noting symmetries between cases.

Consider the case when $e = (p,q)$, $d = (u,v)$, and $c = (r,s)$.

Let us check the four possibilities when dc react, corresponding to the subcases of case 1 of Table 1.

1(a). If $v+1 = r$:

if $(|(u,s)| \leq l)$ $d_1 = \hat{w}(d,c) = (u,s)$: ed_1 do not 0- or 1-react as that requires (case 1(a)) $q+1 = u$ contradicting that ed is stable.

if $(l < |(u,s)| < 2l)$ $d_1 = \hat{w}(d,c) = \overline{(u-1,s+1)}$: ed_1 do not 0- or 1-react as that requires (case 4(a)) $q = u-1$ contradicting that ed is stable.

if $(|(u,s)| = 2l)$ $\hat{w} = \epsilon$: dc 0-react.

1(b). If $(y_{v+1} = y_r$ && $|(u,v)| = l)$ $d_1 d_2 = \hat{w}(d,c) = \overline{(u-1,v+2)}(r+1,s)$: ed_1 do not 0- or 1-react as that requires (case 4(a)) $q = u-1$ contradicting that ed is stable. There are two possibilities if ed_1 2-react:

(i) if $(y_{q+1} = y_{u-1}^{-1}$ && $|(p,q)| = l)$ $d_3 d_4 = \hat{w}(ed_1) = \overline{(p-1,q+2)}\,\overline{(u-2,v+2)}$: $d_4 d_2$ do not react as that requires either (case 2(a)) $v+2 = r+1$, i.e. $r = v+1$, contradicting that dc 2-react, *or* (case 2(c)) $y_{v+2} = y_{r+1}$ which implies that the pair of symbols $y_r y_{r+1}$ occurs at two distinct positions, r and $v+1$, in y, contradicting Lemma 2. Note that the case corresponding to 2(b) cannot arise as $|(r+1,s)| < l$.

(ii) if $(y_q = y_{u-1})$ $d_3 d_4 = \hat{w}(ed_1) = (p,q-1)\overline{(u-2,v+2)}$: $d_4 d_2$ are as in the previous case and do not react.

1(c). If $(y_v = y_{r-1}$ && $|(r,s)| = l)$ $d_1 d_2 = \hat{w}(d,c) = (u,v-1)\overline{(r-2,s+1)}$: ed_1 do not react as that, in each of cases 1(a)-(d), contradicts that ed is stable.

1(d). If $(y_v = y_r^{-1})$ $d_1 d_2 = \hat{w}(d,c) = (u,v-1)(r+1,s)$: ed_1 do not react again as that, in each of cases 1(a)-(d), contradicts that ed is stable.

This completes the analysis of the case $e = (p,q)$, $d = (u,v)$, and $c = (r,s)$. ♣

Imagine a stable rectified product S to be in the form of a stack where the leftmost term is at the bottom. Let c be a rectified term. We define a function *apply*(S,c) which returns a stable rectified product as follows.

```
function apply(S, c)
if (S = ε) push c into S;
if (S ≠ ε)
    {
    remove the top term d of S;
    if (dc do not react) push dc into S;
    if (dc react)
        {
         if (dc 0-react) do nothing;
         if (dc 1-react && ŵ(dc) = d1) apply(S, d1);
         if (dc 2-react && ŵ(dc) = d1d2) {apply(S, d1); push d2 into S;}
        }
    }
return S;
```

The following establishes the correctness of the definition of *apply*:

Lemma 5 *For any stable rectified product S and rectified term c the function $apply(S, c)$ does terminate and return a stable rectified product.*

Proof. Assume $S \neq \epsilon$, otherwise the claim follows trivially.

To verify that $apply(S, c)$ terminates it suffices to observe that if $apply(S, c)$ does issue a recursive call to another (one) instance of *apply*, then the value of S in the called instance is 1 smaller in height than that in the calling instance.

We prove that $apply(S, c)$ returns a stable product by an induction on the height of S. We shall, in fact, prove a stronger statement:

$apply(S, c)$ returns a stable product $\tilde{S}$ *and*, *further*, if the last term of S, when $S \neq \epsilon$, is d, and dc 1- or 2-react, then the last term of $\tilde{S}$ is the last term of $\hat{w}(d, c)$.

Starting the induction is trivial.

Assume inductively that the statement above is true if $height(S) \leq h - 1$.

Suppose $S = c_1 \dots c_{h-2} c_{h-1} c_h$ is of height h.

Now, $c_h c$ may either not react, or 0-, 1-, or 2-react.

Firstly, suppose $c_h c$ 2-react and $\hat{w}(c_h c) = d_1 d_2$. By the inductive hypothesis $apply(c_1 \dots c_{h-2} c_{h-1}, d_1)$ returns a stable product. Further, by Lemma 4, either $c_{h-1} d_1$ do not react or 2-react. We consider these two cases below.

(i) If $c_{h-1} d_1$ do not react then $apply(S, c)$ returns $c_1 \dots c_{h-2} c_{h-1} d_1 d_2$ which is clearly stable and whose last term d_2 is the last term of $\hat{w}(c_h c) = d_1 d_2$.

(ii) If $c_{h-1} d_1$ 2-react then, by Lemma 4 again, $\hat{w}(c_{h-1} d_1) = d_3 d_4$ where $d_4 d_2$ do not react and, moreover, by the inductive hypothesis, the last term of the product returned by $apply(c_1 \dots c_{h-2} c_{h-1}, d_1)$ is d_4. It follows that $apply(S, c)$ does indeed return a stable product of the form $S' d_4 d_2$ whose last term is again d_2.

The other cases when $c_h c$ do not react, 0-react, or 1-react are more easily resolved. ♣

We can say more about *apply*:

Lemma 6 *For any S and c the function $apply(S, c)$ returns a stable rectified product $\tilde{S}$ such that $\tilde{S} \approx Sc$.*

Proof. This may be proved as well by a simple induction on the height of S. We leave details to the reader. ♣

A *canonical* form for a rectified product w is a stable rectified product $\tilde{w}$ such $\tilde{w} \approx w$. For a rectified product w define a function $canonical(w)$ as follows.

function $canonical(w)$
if $(w = c_1 \ldots c_h)$
{
$S = \epsilon$;
for $(i = 1;\ i \leq h;\ i{+}{+})$ $S = apply(S, c_i)$;
}
return S;

Lemma 7 *For any rectified product w the function $canonical(w)$ returns a canonical form for w. It follows that $canonical(canonical(w)) = canonical(w)$.*

Proof. This follows from from Lemma 6 and a simple induction on the height of w. ♣

Lemma 8 *A stable rectified product w is R-reduced.*

Proof. Clearly w is reduced or it would not be stable.

Next, suppose, if possible, w contains a subword w' such that there exists a relation $r \in R$ with $r \equiv w'w''$ and $|w'| > \frac{1}{2}|r|$. Then w' must intersect at least two consecutive terms, say c_i and c_{i+1}, of w, as the length of a rectified term is at most $l = \frac{1}{2}|r|$. In which case, since w' is a subword of one of the circular lists $y_1 \ldots y_{2l}$ and $y_{2l}^{-1} \ldots y_1^{-1}$, it may be seen that $c_i c_{i+1}$ react, contradicting that w is stable. ♣

We defer discussion of the complexity of function *canonical*, but show now how it may be applied to solve the word problem for $\pi(M')$.

Proposition 6 $w \approx 1$ *if and only if $canonical(w) = \epsilon$, the empty word.*

Proof. Lemmas 7 and 8 imply that $canonical(w)$ is R-reduced. By Proposition 1, $canonical(w) \approx 1$ if and only if it is empty. The result follows as $canonical(w) \approx w$. ♣

Example. If

$$\pi(M') = (b_1, b_2, b_3, b_4; y_1 y_2 \ldots y_8 = b_1 b_2 b_3 b_4 b_1^{-1} b_2^{-1} b_3^{-1} b_4^{-1})$$

and

$$w = (2,3)(4,7)(8,2)(3,5)(6,7)(8,2)b_2^{-1},$$

then the following shows the transitions of the stack S in computing $canonical(w)$,

$$\epsilon \rightarrow (2,3) \rightarrow \overline{(1,8)} \rightarrow b_2 \rightarrow (2,5) \rightarrow \overline{(1,8)} \rightarrow b_2 \rightarrow \epsilon,$$

implying $w \approx 1$.

An important property of the function *canonical* is described in:

Lemma 9 *If w is a non-empty stable rectified product with last term d, and c is a rectified term, then $height(canonical(wc))$ is*

1. *$height(w) - 1$, when dc 0-react, i.e., when $d = c^{-1}$.*

2. *$height(w)$, when dc 1-react.*

3. *$height(w) + 1$, when dc do not react or 2-react.*

Proof. We shall prove the claim by an induction on the height of w, very similar to the induction in the proof of Lemma 5.

Starting the induction is trivial.

Assume inductively that the claim is true if $height(w) \leq h-1$.

Suppose $w = c_1 \dots c_{h-2}c_{h-1}c_h$ is of height h.

Consider first when $c_h c$ 2-react and $\hat{w}(c_h c) = d_1 d_2$. Then, by Lemma 4, either $c_{h-1}d_1$ do not react or 2-react. We consider these two cases below.

(i) If $c_{h-1}d_1$ do not react then $canonical(wc) = c_1 \dots c_{h-2}c_{h-1}d_1 d_2$ of height $h+1$ verifying the claim.

(ii) If $c_{h-1}d_1$ 2-react, by Lemma 4 again, $\hat{w}(c_{h-1}d_1) = d_3 d_4$ where $d_4 d_2$ do not react. Then

$$\begin{aligned} canonical(wc) &= canonical(c_1 \dots c_{h-2}c_{h-1}d_1 d_2) \\ &= canonical(canonical(c_1 \dots c_{h-2}c_{h-1}d_1)\ d_2) \\ &= canonical((w'd_4)d_2), \end{aligned}$$

using the stronger statement verified in the proof of Lemma 5 to deduce that $canonical(c_1 \dots c_{h-2}c_{h-1}d_1) = apply(c_1 \dots c_{h-2}c_{h-1}, d_1)$ must end with the term d_4 and, therefore, is of the form $w'd_4$. Further, by the inductive hypothesis, $w'd_4$ is of height h. Since $d_4 d_2$ do not react we conclude that $canonical(wc) = w'd_4 d_2$ is stable of height $h+1$, verifying the claim again.

The other cases when $c_h c$ do not react, 0-react, or 1-react are easily resolved. ♣

The following useful proposition shows that the word $canonical(w)$ does indeed determine w up to equality in $\pi(M')$.

Proposition 7 *$w_1 \approx w_2$ if and only if $canonical(w_1) = canonical(w_2)$.*

Proof. Since $canonical(w_\alpha) \approx w_\alpha$, $\alpha = 1, 2$, it follows that $canonical(w_1) = canonical(w_2)$ implies that $w_1 \approx w_2$.

Conversely, suppose that $w_1 \approx w_2$, and let

$$canonical(w_1) = c_{1,1} \dots c_{1,h_1} \text{ and } canonical(w_2) = c_{2,1} \dots c_{2,h_2}.$$

We shall prove the equality $canonical(w_1) = canonical(w_2)$ by induction on $\min(h_1, h_2)$.

To start the induction observe that if $\min(h_1, h_2) = 0$, say w.l.o.g. $h_1 = 0$, then $canonical(w_1) \approx 1$. Further, observe that, as $w_1 \approx w_2$ and $w_\alpha \approx canonical(w_\alpha)$, $\alpha = 1, 2$, we have $canonical(w_2) \approx canonical(w_1)$. Therefore, in this case, $canonical(w_2) \approx 1$ implying, by Lemma 7 and Proposition 6, that

$$canonical(w_2) = canonical(canonical(w_2)) = \epsilon.$$

Assume inductively that the equality is true if $\min(h_1, h_2) \leq N$, for some $N \geq 0$, and consider the case when $\min(h_1, h_2) = N+1$, where, w.l.o.g., we suppose that $h_1 = N+1$.

We have already observed

$$\begin{aligned} & canonical(w_1) \approx canonical(w_2) \\ \text{which implies} \quad & c_{1,1} \dots c_{1,h_1} \approx c_{2,1} \dots c_{2,h_2} \\ \text{which implies} \quad & c_{1,1} \dots c_{1,h_1-1} \approx c_{2,1} \dots c_{2,h_2}({c_{1,h_1}}^{-1}). \end{aligned}$$

Since the product on the LHS of the last equation is stable (being part of a stable product),

$$canonical(c_{1,1} \dots c_{1,h_1-1}) = c_{1,1} \dots c_{1,h_1-1}.$$

Therefore, by the inductive hypothesis, we must have

$$c_{1,1}\dots c_{1,h_1-1} = canonical(c_{2,1}\dots c_{2,h_2}({c_{1,h_1}}^{-1})).$$

For the preceding equation to hold we must further have, by Lemma 9, that

$$c_{2,h_2}{c_{1,h_1}}^{-1} \text{ 0-react implying that } c_{1,h_1} = c_{2,h_2} \text{ and } c_{1,1}\dots c_{1,h_1-1} = c_{2,1}\dots c_{2,h_2-1}.$$

It follows that

$$c_{1,1}\dots c_{1,h_1} = c_{2,1}\dots c_{2,h_2}, \text{ so, indeed,}$$
$$canonical(w_1) = canonical(w_2).$$

♣

Let us determine now the complexity of the function *canonical*.

Lemma 10 *$canonical(w)$ returns in time $O(height(w))$.*

Proof. This would be almost obvious except that a call to $apply(S, c_i)$ within $canonical(w)$, where $w = c_1 \dots c_h$, may, after one initial 1- or 2-reaction, trigger a "chain of 2-reactions" down the stack. See Figure 5.

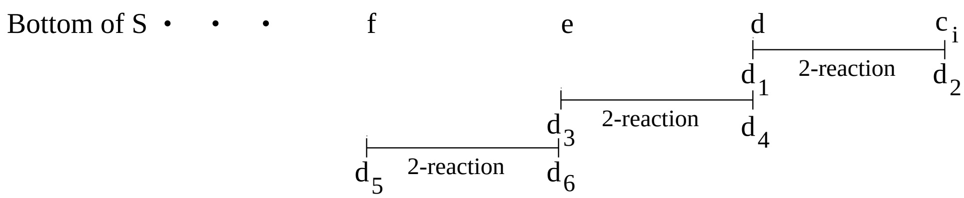

Figure 5: A chain of 2-reactions.

To investigate this possibility we first need an augmentation of Lemma 4. In cases 2 and 3 of Lemma 4, if ed_1 do indeed 2-react, then we say c *alters* e in edc. In particular, we need:

Lemma 11 (Augmentation of Lemma 4) *In cases 2 and 3 of Lemma 4, if c alters e, in particular if $\hat{w}(ed_1) = d_3d_4$, then* **no** *rectified term b can alter d_3 in the product d_3d_4b.*

Proof. This again is a technical exercise in analyzing various cases. We omit details. ♣

With Lemma 11 in hand now, determining the complexity of *canonical* is a task in accounting.

If, in a call to $apply(S, c_i)$ within $canonical(c_1 \dots c_h)$, c_i does not alter the second-last term of S, the cost of $apply(S, c_i)$ is $O(1)$. On the other hand, for those c_i that alter terms of S (which in turn may alter other terms, ...) we need to bound the number of alterations that may occur totally through a call to $canonical(w)$.

Consider a particular level k of the stack S, i.e., the position of the kth term from the bottom. Suppose at some step the term, say e, at level k is altered by c in the product, say edc inside S, so that $\hat{w}(ed_1) = d_3d_4$ (following notations of Lemma 11). Then, it follows from Lemma 11, that a subsequent alteration can occur of a term at level k *only after* a 0-reaction between terms at levels $k+1$ and $k+2$. For, if not, this would be *equivalent* to d_3 being altered by some term b in a product d_3d_4b, contradicting Lemma 11.

Now, by Lemma 9, a 0-reaction decrements $height(S)$ by 1, while $height(S)$ may increase by at most 1 after a call to $apply(S, c_i)$. As $w = c_1 \dots c_h$ we conclude that at most h 0-reactions can occur, and, therefore, at most h alterations can occur through a call to *canonical*(w). This implies that the complexity of *canonical*(w) is linear in *height*(w). ♣

We return to the problem of deciding if the rectified products $z'_\alpha = c_{\alpha,1} \dots c_{\alpha,h'_\alpha}$, $\alpha = 1, 2$, obtained from z_α from the preprocessing of Subphase 2a, represent conjugate elements of the group $\pi(M') = (b_1, \dots, b_l; y_1 \dots y_{2l})$.

3.2.2 Subphase 2b - Computing a Good Conjugate

The first step of this subphase is to compute $canonical(z'_\alpha) = c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, say, $\alpha = 1, 2$. Since $z'_\alpha \approx canonical(z'_\alpha)$, the problem of deciding the conjugacy of z'_α in $\pi(M')$ is equivalent to that of deciding the conjugacy of $canonical(z'_\alpha)$ in $\pi(M')$.

We intend to apply Proposition 2 to this problem. In doing so, the premise that $canonical(z'_\alpha) \neq \epsilon$, $\alpha = 1, 2$, may be assumed as, otherwise, the problem is trivial. However, the premise that both $canonical(z'_\alpha)$ are cyclically R-reduced may not hold. E.g., $canonical(z'_1) = (7, 15)(2, 3)(5, 6)$ may not be cyclically R-reduced if the cyclic conjugate $(2, 3)(5, 6)(7, 15)$ is not R-reduced (it is certainly not stable). We address this problem by proving:

Lemma 12 *For any stable rectified product w we can, in time $O(height(w))$, find a stable rectified product w' that is a cyclically R-reduced conjugate of w.*

Proof. Let $w = c_1 c_2 \dots c_h$ be a stable rectified product. Observe that if $c_2 \dots c_h c_1$ is stable, i.e., if $c_h c_1$ do no react, then w is, in fact, cyclically R-reduced. Call the transformation of $c_1 c_2 \dots c_h$ to $c_2 \dots c_h c_1$ a *rotation*. Consider the following function.

function *reduced_conjugate*(w)
while (*rotate*(w) is not stable) /* check if last term, first term of w react */
{
c = first term of w;
$v = w$ with first term deleted;
$w = apply(v, c)$; /* = $canonical(rotate(w))$ */
}
return w;

We claim that, for any stable rectified product w, *reduced_conjugate*(w) terminates in $O(height(w))$ steps to return a cyclically R-reduced conjugate w' of w. The verification is an analysis of various possible forms of $w = c_1 c_2 \dots c_h$. We shall consider only one in some detail to explain the underlying intuition.

Let w be of the form $w = (r_1, s_1)(r_2, s_2) \dots (r_h, s_h)$. It is enough to consider only when the last two terms 2-react as we rotate. For, the cases when the last two terms do not react or 0-react are trivial, while the case when they 1-react reduces, after one step, by case 2 of Lemma 4, to that when they 2-react.

If $c_h c_1$ 2-react in *rotate*(w) according to cases 1(b) or 1(d) of Table 1, the last term of *canonical*(*rotate*(w)), in either case, is $(r_1 + 1, s_1)$. Then, another rotation, placing (r_2, s_2) after $(r_1 + 1, s_1)$, gives a stable product. Therefore, it may be seen that, as we rotate, we need only consider when the last 2 terms 2-react (if they react at all) according to case 1(c) at the first step, and then according to case 2(b) at each subsequent step. At most how many such 2-reactions can occur? Accordingly, in the following, after label (i), $1 \leq i \leq h - 1$, we indicate (i) conditions necessary for the last 2 terms to 2-react after i rotations, followed by (ii) the form of the product after the 2-reaction. The initial product is labeled (0).

(0) $(r_1, s_1)(r_2, s_2) \dots (r_h, s_h)$

$$
\begin{array}{lll}
(1) & \text{(i)} & y_{s_h} = y_{r_1-1} \;\&\&\; |(r_1,s_1)| = l\text{:} \\
 & \text{(ii)} & (r_2,s_2)\ldots(r_{h-1},s_{h-1})(r_h,s_h-1)\overline{(r_1-2,s_1+1)} \\
(2) & \text{(i)} & y^{-1}_{s_1+1} = y_{r_2-1} \;\&\&\; |(r_2,s_2)| = l\text{:} \\
 & \text{(ii)} & (r_3,s_3)\ldots(r_{h-1},s_{h-1})(r_h,s_h-1)\overline{(r_1-2,s_1+2)}\,\overline{(r_2-2,s_2+1)} \\
(3) & \text{(i)} & y^{-1}_{s_2+1} = y_{r_3-1} \;\&\&\; |(r_3,s_3)| = l\text{:} \\
 & \text{(ii)} & (r_4,s_4)\ldots(r_{h-1},s_{h-1})(r_h,s_h-1)\overline{(r_1-2,s_1+2)}\,\overline{(r_2-2,s_2+2)}\,\overline{(r_3-2,s_3+1)} \\
\ldots & \ldots & \\
(h-1) & \text{(i)} & y^{-1}_{s_{h-2}+1} = y_{r_{h-1}-1} \;\&\&\; |(r_{h-1},s_{h-1})| = l\text{:} \\
 & \text{(ii)} & (r_h,s_h-1)\overline{(r_1-2,s_1+2)}\,\overline{(r_2-2,s_2+2)}\ldots\overline{(r_{h-2}-2,s_{h-2}+2)}\,\overline{(r_{h-1}-2,s_{h-1}+1)}
\end{array}
$$

No further application can be made of case 2(b) after another rotation as $|(r_h, s_h - 1)| < l$. It follows that the number of 2-reactions is bounded by h. This verifies our claim for the function *reduced_conjugate* and the lemma follows. ♣

We must execute a second step in this subphase, where we actually need a stronger form of Lemma 12:

Lemma 13 *For any stable rectified product w we can, in time $O(height(w))$, find a conjugate $w' = c'_1 \ldots c'_{h'}$ of w such that*

(a) w' is a stable rectified product, and

(b) w' is a cyclically R-reduced, and

(c) for any i, $1 \leq i \leq h'$, and any rectified term c, c does not alter c'_{i-2} in $c'_{i-2}c'_{i-1}c$ and c does not alter c'_{i+1} in $cc'_ic'_{i+1}$ (for the second case, assume a definition of alter that is symmetric to the definition preceding Lemma 11). In other words, for any i, $1 \leq i \leq h'$, and any rectified terms b and d,

$$canonical(b\; c'_ic'_{i+1} \ldots c'_{h-1}c'_hc_1c_2 \ldots c'_{i-2}c'_{i-1}\; d) = B\; c'_{i+1} \ldots c'_{h-1}c'_hc_1c_2 \ldots c'_{i-2}\, D;$$

where B and D are empty or the product of 1 or 2 terms.

Proof. The proof is based on a function *good_conjugate* that utilizes *reduced_conjugate* as a subprocedure. Instead of formally defining *good_conjugate* and establishing its properties, which requires cumbersome notation, we give an informal description.

A *run of alterable terms* in $w = c_1c_2 \ldots c_h$ is a maximal subproduct $c_j \ldots c_k$ of w such that there exists a rectified term c that alters c_j in $c_j \ldots c_k\, c$ but not c_{j-1} in $c_{j-1}c_j \ldots c_k\, c$. Such runs can be determined in linear time by scanning down w. For example, a subproduct of the form $(r_j, s_j) \ldots (r_{k-1}, s_{k-1})(r_k, s_k)$ is a run if the conditions

$$y_{s_i+1} = y^{-1}_{r_{i+1}-1} \;\;\&\&\;\; |(r_i, s_i)| = l \qquad \text{(A)}$$

hold for i such that $j \leq i < h$, but not for $i = j - 1$.

Now, if $w = w_1w_2w_3$, where w_2 is a run such that the rectified term c alters terms of w_2, replace w with product $w_1w_2cc^{-1}w_3$. Repeat this procedure for each run to obtain $\tilde{w} \approx w$. Evaluate $\tilde{\tilde{w}} = canonical(\tilde{w})$. The idea is that, if $\tilde{w} = w_1w_2cc^{-1}w_3$, then a call to $apply(S, c)$ as we evaluate $canonical(\tilde{w})$ will "trigger" alterations through the run w_2, so that the corresponding subproduct of the altered product will no longer be part of a run. Finally, evaluate $w' = reduced\_conjugate(\tilde{\tilde{w}})$.

It can be verified using Lemmas 11 and 12 that *good_conjugate* indeed returns a desired w' in time $O(height(w))$. One needs to observe as well that the equations for "altering from the left" (use a definition symmetric to the one preceding Lemma 11) turn out to be *exactly identical* to the equations (A) above for "altering from the right," so that a subproduct that is no longer part of a run cannot be altered from the left either. ♣

The second step of this subphase is, therefore, to compute *good_conjugate*(*canonical*(z'_α)), $\alpha = 1, 2$. To save on notation, we denote *good_conjugate*(*canonical*(z'_α)) by $c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, as well.

By Lemmas 10 and 13, Subphase 2b completes in time $O(h_1 + h_2)$.

Returning to the application of Proposition 2, observe that deciding the conjugacy of z'_α is equivalent to deciding the conjugacy of *good_conjugate*(*canonical*(z'_α)) $= c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, $\alpha = 1, 2$. We may assume, by Lemma 13, that both latter products satisfy clauses (a), (b) and (c) of Lemma 13. Further, assume w.l.o.g. that $h''_1 \geq h''_2$.

From Lemma 3, R satisfies $C'(\frac{1}{8})$. We have, therefore, by Proposition 2, that $c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, $\alpha = 1, 2$, represent conjugate elements of $\pi(M')$ if and only if there exists c such that

$$(c^C_{2,1} \dots c^C_{2,h''_2})^* \approx c\ (c^C_{1,1} \dots c^C_{1,h''_1})^*\ c^{-1}, \tag{1}$$

where $(c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha})^*$ is a cyclic permutation of $c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, $\alpha = 1, 2$, and c is a subword of some relation $y' \in R$. By Lemma 3 we may assume that c is a rectified term.

Say, equation (1), in fact, holds with

$$(c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha})^* = c^C_{\alpha,i_\alpha} \dots c^C_{\alpha,h''_\alpha} c^C_{\alpha,1} \dots c^C_{\alpha,i_\alpha - 1},$$

where

$$1 \leq i_\alpha \leq h''_\alpha, \alpha = 1, 2.$$

Accordingly, rewrite equation (1) as

$$c^C_{2,i_2} \dots c^C_{2,h''_2} c^C_{2,1} \dots c^C_{2,i_2-1} \approx c\ c^C_{1,i_1} \dots c^C_{1,h''_1} c^C_{1,1} \dots c^C_{1,i_1-1}\ c^{-1}. \tag{2}$$

By Proposition 7, equation (2) holds if and only if

$$\mathit{canonical}(c^C_{2,i_2} \dots c^C_{2,h''_2} c^C_{2,1} \dots c^C_{2,i_2-1}) = \mathit{canonical}(c\ c^C_{1,i_1} \dots c^C_{1,h''_1} c^C_{1,1} \dots c^C_{1,i_1-1}\ c^{-1}). \tag{3}$$

Since $c^C_{\alpha,1} \dots c^C_{\alpha,h''_\alpha}$, $\alpha = 1, 2$, satisfy clauses (a), (b) and (c) of Lemma 13, it may be deduced that

$$\mathit{canonical}(c^C_{2,i_2} \dots c^C_{2,h''_2} c^C_{2,1} \dots c^C_{2,i_2-1}) = c^C_{2,i_2} \dots c^C_{2,h''_2} c^C_{2,1} \dots c^C_{2,i_2-1}, \tag{4}$$

and

$$\mathit{canonical}(c\ c^C_{1,i_1} \dots c^C_{1,h''_1} c^C_{1,1} \dots c^C_{1,i_1-1}\ c^{-1}) = B\ c^C_{1,i_1+1} \dots c^C_{1,h''_1} c^C_{1,1} \dots c^C_{1,i_1-2}\ D, \tag{5}$$

where, B and D are empty or the products of 1 or 2 terms.

With equations (4) and (5), rewrite (3) as

$$c^C_{2,i_2} \dots c^C_{2,h''_2} c^C_{2,1} \dots c^C_{2,i_2-1} = B\ c^C_{1,i_1+1} \dots c^C_{1,h''_1} c^C_{1,1} \dots c^C_{1,i_1-2}\ D. \tag{6}$$

Examining equation (6), we see that the string $c^C_{1,1} \dots c^C_{1,h''_1}$, with *some* two adjacent terms deleted, must occur as a substring of the *circular* string $c^C_{2,1} \dots c^C_{2,h''_2}$ (with the first term following the last). Whether this is true may be checked with a linear-time algorithm that uses methods of the Knuth-Morris-Pratt string matching algorithm [1, 12] as follows.

3.2.3 Subphase 2c - String Matching

The string matching problem that we have to solve in this subphase is:

Given a circular string $s^2 = s^2_1 s^2_2 ... s^2_{h''_2}$ ($s^2_{h''_2}$ is adjacent to s^2_1) and another string $s^1 = s^1_1 s^1_2 ... s^1_{h''_1}$ of rectified terms, determine all positions k such that s^1 occurs in s^2 (modulo circularity) with at most two terms s^2_k, s^2_{k+1} deleted.

First of all, we can determine in linear time all possible matchings without any term deletion using the KMP algorithm. We describe the method for matching with exactly two terms deleted from s^2; the case of one term deletion is similar.

To deal with the circularity of s^2, we linearize s^2 by concatenating two copies of s^2. To avoid cumbersome notation the new string is also denoted as s^2. It is not hard to see that all matchings of s^1 in the circular string s^2 are captured in the linearized version of s^2, and since the length is only doubled, the complexity of the algorithm is not affected by this linearization.

For any rectified term c we say $s^2_i s^2_{i+1}$ is *absorbed* if c reacts with s^2_i and c^-1 reacts with s^2_{i+1}. Two positions i, j in s^2 are called *equivalent* if $s^2 = s^2_1 s^2_2 ... s^2_i s^2_{i+1} ... s^2_j s^2_{j+1} ...$, where $s^2_i = s^2_j, s^2_{i+1} = s^2_{j+1}$. The motivation for defining equivalent positions is that for any rectified term c, $s^2_i s^2_{i+1}$ are absorbed due to reaction with it if and only if $s^2_j s^2_{j+1}$ are absorbed in a similar manner. Hence finding the match of s^1 with $s^2_i s^2_{i+1}$ deleted is equivalent to finding the match of s^1 with $s^2_j s^2_{j+1}$ deleted, and thus, detection of one is sufficient. This suggests that we need to find only one position from all equivalent positions where s^1 matches s^2.

Now, we describe the string matching algorithm – its correctness will be clear afterwards. First, for each $i = 1, \ldots, 2h''_2$, we determine the length k of all maximal prefixes of s^1 such that $s^2_{i-k-1} = s^1_1, \ldots, s^2_i = s^1_k$ and $s^2_{i+1} \neq s^1_{k+1}$. There can be more than one such maximal prefix. We attach all lengths of these prefixes in a list L_i to s^2_i. This can be done while running the KMP algorithm for matching the pattern s^1 in s^2. Since KMP runs in linear time, the total size of all L_i's together is linear, i.e., at most $O(h''_1 + h''_2)$. Similarly, with each s^2_i we attach another integer I_i where I_i is the length of the the maximum (not maximal) suffix of s^1 such that $s^2_i = s^1_{h''_1 - I_i - 1}, ..., s^2_{i+I_i-1} = s^1_{h''_1}$ and $s^2_{i-1} \neq s^1_{h''_1 - I_i - 2}$.

This can again be done by running KMP on the reversed strings of s^2 and s^1. We also compute a suffix function on s^1 which is exactly similar to prefix function computation for the KMP algorithm [1]. The only difference is that we run it on the reverse of s^1, i.e., start from the right end of s^1. This suffix function computes a number ℓ for each s^1_i such that ℓ is the length of the longest suffix that is also a prefix of $s^1_i s^1_{i+1} ... s^1_{h''_1}$. We build a forest F in which the node with integer value $h''_1 - i + 1$ is made a child of the node with integer value ℓ. A downward path in F contains the increasing order of lengths of the maximal suffixes that are also prefix of substrings of s^1. This forest is based on an array, where each location ℓ of the array has a list of indices (integer values) that are children of ℓ. We will describe later how we use this structure for our matching. First we need the following result.

Lemma 14 *Let s^2 match s^1 with two terms deleted at position i. Then there must exist an equivalent position j such that L_j or L_{j-1} is non-empty.*

Proof. Let $s^2_{i-\ell} = s^1_1, s^2_{i-\ell+1} = s^1_2, ..., s^2_{i-1} = s^1_\ell$ and $s^2_{i+2} = s^1_{\ell+1}, ..., s^2_{i+1+h''_1-\ell} = s^1_{h''_1}$. Then if $s^2_i \neq s^1_{\ell+1}$, we have $j = i$ (L_{i-1} is non-empty); otherwise we must have $s^2_i = s^1_{\ell+1}$. In this case, if $s^2_{i+1} \neq s^1_{\ell+2}$, we can take $j = i$ since L_i is non-empty. In the remaining case, where $s^2_i = s^1_{\ell+1}$ and $s^2_{i+1} = s^1_{\ell+2}$, let $k \geq 0$ be such that $s^2_i = s^1_{\ell+1}, s^2_{i+1} = s^1_{\ell+2}, ..., s^2_{i+2k-2} = s^1_{\ell+2k-1}, s^2_{i+2k-1} = s^1_{\ell+2k}$ and $s^2_{i+2k} \neq s^1_{\ell+2k+1}$ or $s^2_{i+2k+1} \neq s^1_{\ell+2k+2}$. By induction on k we can prove that s^1 matches s^2 at any position $i + 2m$, $0 \leq m \leq k$, where $i + 2m$ is equivalent to i. When $m = 0$, the assertion is trivially true. Let the assertion be true for $m = k-1$, i.e., s^1 matches s^2 at $i+2k-2$, where i and $i+2k-2$ are equivalent. Then, $s^2_{i+2k} = s^1_{\ell+2k-1}$ and $s^2_{i+2k+1} = s^1_{\ell+2k}$. But, $s^1_{\ell+2k-1} = s^2_{i+2k-2}$ and $s^1_{\ell+2k} = s^2_{i+2k-1}$, which implies $s^2_{i+2k} = s^2_{i+2k-2}$ and $s^2_{i+2k+1} = s^2_{i+2k-1}$. Hence, $i+2k-2$ and $i+2k$ are equivalent positions. Since i and $i+2k-2$ are equivalent by inductive hypothesis, we have i and $i+2k$ equivalent by definition. Also, it is straightforward to observe that s^1 matches s^2 at position $i+2k$. Hence $j = i+2k$ in this case. ♣

The above lemma ascertains that we can consider only those positions i where L_i or L_{i-1} is non-empty for possible matches. Below we describe the method assuming L_i is non-empty; the case of L_{i-1} not being empty is handled similarly.

First, we compute the L_i's , I_i's and F, all in linear time using methods of the KMP algorithm. We need to find all positions i such that I_{i+3} is a descendent of a number k, where $x + k = h''_1$ for some x in L_i. This means i is a possible position where s^2 and s^1 match with s^2_{i+1} and s^2_{i+2} deleted. To check this we produce the tuple (k, I_{i+3}) if there is a $c \in R$ such that c reacts with s^2_{i+1} to absorb it and s^2_{i+2} reacts with c^{-1} to get absorbed. It is takes constant time to determine if such c exists. Notice that there are at most $O(h''_1 + h''_2)$ such tuples created since the size of all L_i's together is at most $O(h''_1 + h''_2)$. The structure F is used to check if I_{i+3} is a descendent of k. We process all tuples together. The tuples are maintained in an array T such that $T[i]$ is a list of all ancestors of i that appeared with i in one of the tuples. We traverse F in *inorder* fashion and maintain a path array P of length h''_1 such that $P[i]$ is marked if and only if i is in the current path from the root to the node being visited. Whenever we visit the node with number i, we check first if $T[i]$ is non-empty, and if not, we check the indices in the list $T[i]$ that have been marked in the path array P. This determines which tuples are feasible. Updating P involves only deleting or adding marks in an array indexed by the visited node numbers. This entire process of traversing F and determining feasible tuples runs in time proportional to the size of F and the number of tuples to be checked, both of which are at most $O(h_1 + h_2)$.

This completes our description of Phase 2 where we solve the algebraic problem of deciding if z'_α, $\alpha = 1, 2$, represent conjugate elements of $\pi(M')$.

4 Conclusions and Remarks

The total cost of the algorithm described in the previous section is $O(n + k_1 + k_2)$ in both time and space.

In the context of the restriction on the genus g of M in our algorithm – $g \geq 3$ if M is orientable and $g \geq 5$ if M is non-orientable – we observe the following: when M is orientable with genus $g = 0$ (sphere) or 1 (torus), and when M is non-orientable with genus $g = 1$ (projective plane) or 2 (Klein bottle), the transformability problem can be solved in optimal time and space using Dehn's methods in [2].

The only exceptional cases that remain, therefore, are $g = 2$ (double torus) when M is orientable, and $g = 3, 4$ when M is non-orientable. Unfortunately, our primary algebraic tool of Proposition 2 is inapplicable in these cases. Nevertheless, the contractability problem for these cases can be solved in $O(n + (k_1 + k_2)\log g))$ time and $O(n + k_1 + k_2)$ space using the algorithm of [8], which is optimal as g is bounded.

We summarize in the following:

Theorem 1 *Given a triangulation T of size n of a genus g surface M,*
(i) it can be decided if a closed curve C presented as edge-vertex sequences of length k in T is contractable in optimal $O(n + k)$ time and space, and
(ii) it can be decided if two such closed curves C_1 and C_2 of lengths k_1 and k_2, respectively, are

homotopic in optimal $O(n+k_1+k_2)$ time and space except for the cases, where M is an orientable surface of genus 2, or a non-orientable surface of genus 3 or 4. ♣

We have thus used notions from modern combinatorial group theory to derive an optimal algorithm for the contractability problem for curves on any surface, as well as an optimal algorithm for the transformability problem for curves on any surface, except for three of low genus. This brings almost to a close the recent investigation [6, 8, 17] into the complexity of the problem originally posed by Dehn [2, 3].

Remark 1. It remains, of course, to settle the transformability problem for three surfaces of low genus. Further, our results suggest that it may be useful to examine other applications of combinatorial group theory to related problems in computational topology, obvious but difficult ones being the contractability and transformability problems for 3-manifolds. The corresponding problems for arbitrary manifolds of dimension ≥ 4 are known to be unsolvable [19].

Remark 2. Dehn's own combinatorial algorithm for the conjugacy problem [3] assumes as given a canonical form of the polygonal schema for M. Currently, the best known algorithm to obtain a canonical polygonal schema from an arbitrary triangulation takes $O(n \log n)$ time [22]. It is precisely to avoid this bottleneck that we deal with reduced polygonal schema, and require Greendlinger's results to formulate optimal algorithms that, nevertheless, may be said to be of "Dehn-type".

Acknowledgements. We thank Professor Tarun K. Mukherjee of Jadavpur University in Calcutta for several helpful discussions and pointing out a valuable reference. We also thank the anonymous referees for suggestions that helped improve the paper. The first author acknowledges the support of NSF grant CCR-93-21799 while he was at the Department of Computer Science of the Indiana-Purdue University in Indianapolis.

References

[1] T. H. Cormen, C. E. Leiserson, R. L. Rivest, *Introduction to algorithms*, MIT press, 1991.

[2] M. Dehn, Über unednliche diskontinuierliche gruppen, *Math. Ann.* **71** (1912), 116-144.

[3] M. Dehn, Transformation der Kurven auf zweiseitigen Flächen, *Math. Ann.* **72** (1912), 413-421.

[4] M. Dehn, *Papers on Group Theory and Topology* (translated and introduced by J. Stillwell), Springer-Verlag, 1987.

[5] C. Delfinado, H. Edelsbrunner, An incremental algorithm for betti numbers of simplicial complexes, *9th Ann. Symp. Comp. Geom.*, 1993, 232–239.

[6] T. K. Dey, A new technique to compute polygonal schema for 2-manifolds with application to null-homotopy detection, *Proc. 10th ACM Symp. Comp. Geom.*, 1994, 277-284.

[7] T. K. Dey, H. Edelsbrunner, S. Guha, Computational topology, *preprint*, 1997.

[8] T. K. Dey, H. Schipper, A new technique to compute polygonal schema for 2-manifolds with application to null-homotopy detection, *Disc. & Comp. Geom.* **14** (1995), 93–110.

[9] H. Edelsbrunner, Modeling with simplicial complexes: Topology, geometry, and algorithms, *Proc. 6th Canad. Conf. Comp. Geom.*, 1994, 36–44.

[10] M. D. Greendlinger, On Dehn's algorithm for the conjugacy and word problems with applications, *Comm. Pure Appl. Math.* **13** (1960), 641-677.

[11] L. Guibas, J. Stolfi, Primitives for the manipulation of general subdivisions and the computation of Voronoi diagrams. *ACM Trans. Graphics* **4** (1985), 74-123.

[12] D. E. Knuth, J. H. Morris, V. R. Pratt, Fast pattern matching on strings, *SIAM J. Comp.* **2** (1977), 323-350.

[13] R. C. Lyndon, P. E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, 1977.

[14] W. S. Massey, *Algebraic Topology: An Introduction*, Harcourt, Brace & World, 1967.

[15] P. S. Novikov, On the algorithmic unsolvability of the word problem in group theory, *Trudy Mat. Inst. Steklov* **44** (1955), 143.

[16] H. Poincaré, Cinquième complément à l'analysis situs, *Rend. circ. mat. Palermo* **18** (1904), 45-110.

[17] H. Schipper, Determining contractibility of curves, *Proc. 8th ACM Symp. Comp. Geom.*, 1992, 358-367.

[18] I. M. Singer, J. A. Thorpe, *Lecture Notes on Elementary Topology and Geometry*, Springer-Verlag, 1967.

[19] J. Stillwell, *Classical Topology and Combinatorial Group Theory*, Springer-Verlag, 1980.

[20] J. J. Rotman, *The Theory of Groups*, 2nd Ed., Allyn & Bacon, 1973.

[21] G. Vegter, Computational topology, *preprint*, 1995.

[22] G. Vegter, C. K. Yap, Computational complexity of combinatorial surfaces, *Proc. 6th ACM Symp. Comp. Geom.*, 1990, 102-111.