PERGAMON

Topology 39 (2000) 283–298

TOPOLOGY

www.elsevier.com/locate/top

Automorphisms of the complex of curves

Feng Luo

Department of Mathematics, Rutgers University, New Brunswick, Busch Campus, 110 Frelinghuysen Road, Piscataway 08854-8019, NJ 08903, USA

Received 25 August 1997; received in revised form 25 March 1998; accepted 14 December 1998

1. Introduction

Given a compact orientable surface $\Sigma = \Sigma_{g,n}$ of genus g with n boundary components (possibly $n = 0$), let $\mathscr{S}(\Sigma)$ be the set of isotopy classes of essential unoriented non-boundary parallel simple loops in Σ. Two classes in $\mathscr{S}(\Sigma)$ are called *disjoint* if they are distinct and have disjoint representatives. In [7], Harvey introduced the complex of curves $\mathscr{C}(\Sigma)$ for Σ as follows. The vertices of $\mathscr{C}(\Sigma)$ are elements in $\mathscr{S}(\Sigma)$ and the simplexes of $\mathscr{C}(\Sigma)$ are $\langle \alpha_1, \ldots, \alpha_k \rangle$ where α_i is disjoint from α_j for $i \neq j$. This complex encodes the asymptotic geometry of the Teichmüller space in analogy with Tits buildings for symmetric spaces. The mapping class group acts on the curve complex preserving the simplicial structure. A natural question one would like to ask is whether every automorphism of the curve complex is induced by a homeomorphism of the surface.

In 1989 (see [11]), Ivanov sketched a proof of the result that if the genus of the surface g is at least 2, then any automorphism of the curve complex $\mathscr{C}(\Sigma_{g,n})$ is induced by a homeomorphism of the surface.

The aim of the paper is to settle the automorphism problem for the rest of the surfaces. Our proof does not distinguish the case genus $g \geqslant 2$ from the case $g \leqslant 1$. We have the following.

Theorem. (a) *If the dimension $3g + n - 4$ of the curve complex $\mathscr{C}(\Sigma_{g,n})$ at least one and $(g, n) \neq (1, 2)$, then any automorphism of $\mathscr{C}(\Sigma_{g,n})$ is induced by a self-homeomorphism of the surface.*

(b) *Any automorphism of $\mathscr{C}(\Sigma_{1,2})$ preserving the set of vertices represented by separating loops is induced by a self-homeomorphism of the surface.*

(c) *There is an automorphism of $\mathscr{C}(\Sigma_{1,2})$ which is not induced by any homeomorphisms.*

E-mail address: fluo@math.rutgers.edu.

PII: S0040-9383(99)00008-7

We remark that Korkmaz [12] has independently proved part (a) of the theorem for genus $g \leqslant 1$ using different methods.

We use the induction on the dimension of the curve complex $\mathscr{C}(\Sigma)$ to prove the theorem. The strategy behind the proof fits extremely well with Grothendieck's philosophy (see [4]) that in the hierarchy of surfaces of negative Euler numbers under inclusion, the "generators" are the 1-holed torus and 4-holed sphere and the "relators" are the 2-holed torus and 5-holed sphere. Indeed, the most difficult and crucial cases in the proof are the 2-holed torus and 5-holed sphere whose curve complexes have dimension one. The proof for these two specific surfaces Σ depends on an extremely simple fact that given two distinct elements in $\mathscr{S}(\Sigma)$, there is at most one element in $\mathscr{S}(\Sigma)$ which is disjoint from both of them (Lemma 4.3). Our proof makes extensive use of the work of several other authors [1, 2, 6, 11, 13, 21]. In particular the work of Harer on the homotopy type of the curve complex is essential to our approach.

The theorem is an analogy to a result of Tits that all automorphisms of Tits buildings are induced by the automorphisms of the corresponding algebraic groups.

Let $\mathrm{Mod}(\Sigma)$ be the mapping class group $\mathrm{Home}(\Sigma)/\mathrm{Iso}$ of the surface of negative Euler number. There is a natural homomorphism $\pi: \mathrm{Mod}(\Sigma) \to \mathrm{Aut}(\mathscr{C}(\Sigma))$ sending the isotopy class of a homeomorphism to the induced map on the curve complex. The theorem shows that π is an epimorphism except $\Sigma = \Sigma_{1,2}$. The image $\pi(\mathrm{Mod}(\Sigma_{1,2}))$ is a subgroup of index 5 in $\mathrm{Aut}(\mathscr{C}(\Sigma_{0,5}))$. By the work of Birman [1] and Viro [21], the kernel of π is known to be trivial unless $(g, n) = (1, 1), (0, 4), (1, 2), (2, 0)$. If $(g, n) = (1, 1), (1, 2)$ and $(2, 0)$, then the kernel is $\mathbf{Z}_2$ generated by a hyperelliptic involution. If $(g, n) = (0, 4)$, then $\ker(\pi) \cong \mathbf{Z}_2 + \mathbf{Z}_2$ and is generated by two hyperelliptic involutions (see Fig. 6). Thus the theorem gives a new characterization of the mapping class group.

The complex of curves also arises in the study of 3-manifolds and mapping class groups. This complex was considered by Harer [5, 6] from homological point of view (with applications to the homology of the mapping class group). In particular, Harer determined the homotopy type of the curve complex [6, Theorem 3.5]. Ivanov [9, 10] used the curve complex to determine the structure of the mapping class group. Masur and Minsky [17] showed that the curve complex is δ-hyperbolic in Gromov's sense. And Hempel [8] used the curve complex for studying 3-manifolds. See also [19].

The paper is organized as follows. In Section 2, we establish basic properties of $\mathscr{C}(\Sigma)$. In particular, using a result of Harer on the homotopy type of $\mathscr{C}(\Sigma)$, it is shown that the curve complexes are pairwise non-isomorphic unless $\mathscr{C}(\Sigma_{1,1}) \cong \mathscr{C}(\Sigma_{0,4})$, $\mathscr{C}(\Sigma_{1,2}) \cong \mathscr{C}(\Sigma_{0,5})$ and $\mathscr{C}(\Sigma_{0,6}) \cong \mathscr{C}(\Sigma_{2,0})$. This is an analogy to Patterson's theorem for Teichmüller spaces. Part (c) of the theorem follows easily from $\mathscr{C}(\Sigma_{1,2}) \cong \mathscr{C}(\Sigma_{0,5})$. In Section 3, we introduce a multiplicative structure on $\mathscr{S}(\Sigma)$. In particular, we define a $(\mathbf{Q}P^1, PSL(2, \mathbf{Z}))$ modular structure on $\mathscr{S}(\Sigma)$ (Definition 3.4). The $PSL(2, \mathbf{Z})$ modular structure is fundamental to our approach to the automorphism problem. In Section 4, we show that any automorphism of $\mathscr{C}(\Sigma)$ takes two curves intersecting at one point (resp. two points of different signs) to two curves intersecting at one point (resp. two points of different signs). Finally, in Section 5, we prove the main theorem by showing that any automorphism of $\mathscr{C}(\Sigma)$ preserving the multiplicative structure is induced by a homeomorphism of the surface. To achieve this, we make extensive use of the modular structure (Lemma 3.1).

2. Preliminaries

2.1. Notations and conventions

We work in the piecewise linear category. All surfaces are oriented, connected and have negative Euler number. The isotopy class of a one-dimensional submanifold s is denoted by $[s]$. The group of homeomorphisms (resp. orientation preserving homeomorphisms) of Σ is denoted by $\mathrm{Home}(\Sigma)$ (resp. $\mathrm{Home}^+(\Sigma)$). The group $\mathrm{Home}(\Sigma)$ acts on $\mathscr{S}(\Sigma)$ as follows: $h([a]) = [h(a)]$ where $h \in \mathrm{Home}(\Sigma)$ and $[a] \in \mathscr{S}(\Sigma)$. Given $\alpha, \beta \in \mathscr{S}(\Sigma)$, the *geometric intersection number* $I(\alpha, \beta)$ between the two classes is defined to be $\min\{|a \cap b| | a \in \alpha, b \in \beta\}$. If F is a function defined on $\mathscr{S}(\Sigma)$, we shall use $F(a)$ to denote $F([a])$ where $[a] \in \mathscr{S}(\Sigma)$. In particular, if $a \in \alpha$, $b \in \beta$, then $I(a, b) = I(a, \beta) = I(\alpha, b) = I(\alpha, \beta)$. We shall use $\alpha \cap \beta = \emptyset$ to denote two disjoint elements α and β, i.e., $I(\alpha, \beta) = 0$ and $\alpha \neq \beta$. If two elements α and β satisfies $I(\alpha, \beta) \neq 0$, we say that they *intersect* and denote them by $\alpha \cap \beta \neq \emptyset$. We use $\alpha \perp \beta$ to denote the relation $I(\alpha, \beta) = 1$. And we use $\alpha \perp_0 \beta$ to denote two elements α and β so that $I(\alpha, \beta) = 2$ and their algebraic intersection number is zero.

A subsurface Σ' in Σ is called *incompressible* if the inclusion map i: $\Sigma' \to \Sigma$ induces a monomorphism in the fundamental groups. It is well-known that Σ' is incompressible if and only if each component of $\partial\Sigma'$ is essential in Σ. Assume that Σ' is incompressible in Σ. Then the map i_*: $\mathscr{S}(\Sigma') \to \mathscr{S}(\Sigma)$ sending $[a]$ to $[i(a)]$ is injective so that $\alpha \cap \beta = \emptyset$, $\alpha \perp \beta$, or $\alpha \perp_0 \beta$ in $\mathscr{S}(\Sigma')$ if and only if their images under i_* satisfy the same relations. Due to this property, we shall identify $\mathscr{S}(\Sigma')$ with the subset $i_*(\mathscr{S}(\Sigma'))$. An element $\alpha \in \mathscr{S}(\Sigma)$ is said to be in Σ' if $\alpha \in i_*(\mathscr{S}(\Sigma'))$. We say an element α in $\mathscr{S}(\Sigma)$ *decomposes* Σ into two subsurfaces Σ' and Σ'' if $\Sigma = \Sigma' \cup \Sigma''$ and $\Sigma' \cap \Sigma'' \in \alpha$. If a class $\alpha \in \mathscr{S}(\Sigma)$ decomposes the surface into a $\Sigma_{0,3}$ and Σ', we say α is a *boundary class*. A class $\alpha \in \mathscr{S}(\Sigma)$ is called *separating* if it has a representative which is a separating loop.

Given a submanifold s, we use $N(s)$ to denote a small regular neighborhood of s. We use $\mathrm{int}(X)$ to denote the interior of a surface X. The symbol $\cong$ is used to denote the homeomorphisms between surfaces, the isomorphisms between simplicial complexes, and isotopy.

Simple loops on surfaces will be denoted by small letters $a, b, \ldots, x, y, z$ and isotopy classes will be denoted by Greek letters α, β, γ etc.

2.2. Basic properties of the curve complex

The homotopy type of the curve complex $\mathscr{C}(\Sigma)$ was determined by Harer [6, Theorem 3.5].

Theorem (Harer [6]). *The curve complex $\mathscr{C}(\Sigma_{g,n})$ is homotopic to a wedge of spheres of dimension r where* (i) $r = 2g + n - 3$, *if* $g > 0$ *and* $n > 0$, (ii) $r = 2g - 2$, *if* $n = 0$ *and* (iii) $r = n - 4$, *if* $g = 0$.

A simplex in $\mathscr{C}(\Sigma_{g,n})$ of maximal dimension $3g + n - 4$ is called a *Fenchel–Nielsen system* (or a pants-decomposition). The following lemma is an easy consequence of Harer's theorem and Birman and Viro's work on the hyperelliptic involutions. The lemma is an analogous to a result of Patterson [18] that the Teichmüller spaces are pairwise nonisomorphic except $T_{1,1} \cong T_{0,4}$, $T_{1,2} \cong T_{0,5}$, and $T_{2,0} \cong T_{0,6}$. Note that since $\mathscr{C}(\Sigma_{1,1})$ and $\mathscr{C}(\Sigma_{0,4})$ are zero-dimensional, by an isomorphism between them we mean a bijection ϕ from $\mathscr{C}(\Sigma_{1,1})$ to $\mathscr{C}(\Sigma_{0,4})$ respecting the relations $\perp$ and $\perp_0$, i.e., $\alpha \perp \beta$ if and only if $\phi(\alpha) \perp_0 \phi(\beta)$.

Lemma 2.1. (a) $\mathscr{C}(\Sigma_{2,0}) \cong \mathscr{C}(\Sigma_{0,6})$, $\mathscr{C}(\Sigma_{1,2}) \cong \mathscr{C}(\Sigma_{0,5})$ *and* $(\mathscr{C}(\Sigma_{1,1}), \perp) \cong (\mathscr{C}(\Sigma_{0,4}), \perp_0)$.

(b) *If* $(g, n) \neq (g', n')$ *and* $\mathscr{C}(\Sigma_{g,n})$ *is not one of the six complexes above, then the curve complexes* $\mathscr{C}(\Sigma_{g,n})$ *and* $\mathscr{C}(\Sigma_{g',n'})$ *are not isomorphic.*

Proof. To show (a), let us first consider $\mathscr{C}(\Sigma_{0,6}) \cong \mathscr{C}(\Sigma_{2,0})$. We construct a bijection from $\mathscr{S}(\Sigma_{0,6})$ to $\mathscr{S}(\Sigma_{2,0})$ preserving the disjointness as follows. Let r: $\Sigma_{2,0} \to \Sigma_{2,0}$ be a hyperelliptic involution. Then $r(\alpha) = \alpha$ for all $\alpha \in \mathscr{S}(\Sigma_{2,0})$ by a result of Birman [1] and Viro [21]. Indeed, r commutes with all Dehn twists up to isotopy. Let P: $\Sigma_{2,0} \to \Sigma_{2,0}/r \cong S^2$ be the quotient map which is a branched covering branched over a six-point set B. Define P^*: $\mathscr{S}(S^2 - \mathrm{int}(N(B))) \to \mathscr{S}(\Sigma_{2,0})$ by sending $[a]$ to $[b]$ where b is a component of $P^{-1}(a)$. Then P^* is a bijection preserving disjointness. Now, $S^2 - \mathrm{int}(N(B)) \cong \Sigma_{0,6}$. Thus $\mathscr{C}(\Sigma_{0,6}) \cong \mathscr{C}(\Sigma_{2,0})$.

For $\Sigma_{1,2}$, take a non-separating r-invariant simple loop s in $\Sigma_{2,0}$ and let $\Sigma_{1,2} = \Sigma_{2,0} - \mathrm{int}(N(s))$. Then $P(\Sigma_{1,2})$ is a disc with 4-cone points B_4 of order 2. Let $\Sigma_{0,5}$ be $P(\Sigma_{1,2}) - \mathrm{int}(N(B_4))$. Then $P^*|_{S(\Sigma_{0,5})}$ is a bijection from $\mathscr{S}(\Sigma_{0,5})$ onto $\mathscr{S}(\Sigma_{1,2})$ preserving disjointness. Finally, identify $\Sigma_{1,1}$ with an r-invariant subsurface of $\Sigma_{2,0}$. Then $P(\Sigma_{1,1})$ is a disc with three cone points of order 2. The same argument shows that the restriction of P^* gives a bijection between $\mathscr{S}(\Sigma_{0,4})$ and $\mathscr{S}(\Sigma_{1,1})$ which respects the relations $\perp_0$ and $\perp$.

To see (b), take $(g, n) \neq (g', n')$. Using Harer's theorem and counting the dimension of the curve complex, we conclude that $\mathscr{C}(\Sigma_{g,n})$ and $\mathscr{C}(\Sigma_{g',n'})$ are not isomorphic except possibly the following cases: (i) $(g', n') = (0, n')$ with $n' \geqslant 7$ and (g, n) with $g \geqslant 1$, and (ii) $(g, n) = (g, 3)$ and $(g', n') = (g + 1, 0)$.

In case (i), suppose otherwise that ϕ: $\mathscr{S}(\Sigma_{g,n}) \to \mathscr{S}(\Sigma_{0,n'})$ is a bijection preserving disjointness. Since $g \geqslant 1$, take a non-separating class $\alpha \in \mathscr{S}(\Sigma_{g,n})$. Then $\phi(\alpha)$ must be a boundary class, i.e., it decomposes $\Sigma_{0,n'}$ into an $\Sigma_{0,3}$ and Σ'. To see this, for any two classes β and γ disjoint from α, there exists a sequence $\alpha_1 = \beta, \alpha_2, \ldots, \alpha_k = \gamma$ so that $\alpha_i \cap \alpha = \emptyset$ and $\alpha_i \cap \alpha_{i+1} \neq \emptyset$. However, if $\phi(\alpha)$ is not a boundary class, there exist two classes β' and γ' disjoint from $\phi(\alpha)$ which cannot be joint by such a sequence. Since $g \geqslant 1$, there exists a maximal dimension simplex $\langle \alpha_1, \ldots, \alpha_k \rangle$ in $\mathscr{C}(\Sigma_{g,n})$ so that each α_i is non-separating. Its image under ϕ is a maximal dimension simplex in $\mathscr{C}(\Sigma_{0,n'})$ so that each vertex is a boundary class. This is impossible unless $n' = 4$, 5, or 6.

In case (ii), suppose otherwise that ϕ: $\mathscr{C}(\Sigma_{g,3}) \to \mathscr{C}(\Sigma_{g+1,0})$ is an isomorphism where $g \geqslant 1$. Take a non-separating class $\alpha \in \mathscr{S}(\Sigma_{g,3})$ and consider its image under ϕ. Since there are no boundary classes in $\mathscr{S}(\Sigma_{g+1,0})$, the image $\phi(\alpha)$ must be non-separating by the same argument as before. By considering the classes disjoint from α, we obtain the following isomorphism $\mathscr{C}(\Sigma_{g-1,5}) \cong \mathscr{C}(\Sigma_{g,2})$. By the result just proved above, this shows $g = 1$, i.e., we have $\mathscr{C}(\Sigma_{1,3}) \cong \mathscr{C}(\Sigma_{2,0})$. But by part (a), we have $\mathscr{C}(\Sigma_{2,0}) \cong \mathscr{C}(\Sigma_{0,6})$. Thus we obtain $\mathscr{C}(\Sigma_{1,3}) \cong \mathscr{C}(\Sigma_{0,6})$. This contradicts the conclusion of case (i). □

Remark. The maximal dimension of those simplexes $\langle \alpha_1, \ldots, \alpha_k \rangle$ in $\mathscr{C}(\Sigma_{g,n})$ so that each α_i is separating is $2g + n - 4$.

Proof of part (c) of the main theorem. Since $\mathscr{C}(\Sigma_{1,2})$ is isomorphic to $\mathscr{C}(\Sigma_{0,5})$ and $\mathrm{Home}(\Sigma_{0,5})$ acts transitively on $\mathscr{S}(\Sigma_{0,5})$, the automorphism group of $\mathscr{C}(\Sigma_{1,2})$ acts transitively on $\mathscr{S}(\Sigma_{1,2})$. In particular, there is an automorphism of $\mathscr{C}(\Sigma_{1,2})$ which sends a separating class to a non-separating class. □

Lemma 2.2. *Suppose $3g + n \geqslant 5$ and $(g, n) \neq (1, 2)$. If ϕ: $\mathscr{S}(\Sigma_{g,n}) \to \mathscr{S}(\Sigma_{g,n})$ is a bijection preserving disjointness, then ϕ preserves the separating classes.*

Proof. Suppose otherwise that α is non-separating and $\phi(\alpha)$ is separating. Then by the same argument as in the proof of Lemma 2.1, $\phi(\alpha)$ is a boundary class. By considering the isotopy classes disjoint from α and $\phi(\alpha)$, respectively, we obtain the following isomorphisms $\mathscr{C}(\Sigma_{g-1,n+2}) \cong \mathscr{C}(\Sigma_{g,n-1})$. By Lemma 2.1, we conclude that $(g, n) = (1, 2)$ or $(1, 3)$. By the assumption, thus $(g, n) = (1, 3)$. Extend α to a Fenchel–Nielsen system $\{\alpha, \beta, \gamma\}$ so that both β and γ are non-separating. Thus $\{\phi(\alpha), \phi(\beta), \phi(\gamma)\}$ is again a Fenchel–Nielsen system in $\Sigma_{1,3}$ where $\phi(\alpha)$ is a boundary class. Say, $\phi(\alpha)$ bounds a subsurface $\Sigma_{1,2}$. Since any Fenchel–Nielsen system on $\Sigma_{1,2}$ contains a non-separating element, thus one of the element $\phi(\beta)$ or $\phi(\gamma)$ is non-separating in the subsurface $\Sigma_{1,2}$. Say $\phi(\beta)$ is non-separating. Find a class δ in $\Sigma_{1,2}$ which is disjoint from $\phi(\beta)$ so that δ bounds a $\Sigma_{1,1}$ in $\Sigma_{1,2}$. Thus δ decomposes $\Sigma_{1,3}$ into a union of $\Sigma_{1,1}$ and $\Sigma_{0,4}$. By Lemma 2.1, $\phi^{-1}(\delta)$ bounds a $\Sigma_{1,1}$ in $\Sigma_{1,3}$. Thus we have a Fenchel–Nielsen system $\{\alpha, \beta, \phi^{-1}(\delta)\}$ so that α and β are non-separating and $\phi^{-1}(\delta)$ bounds $\Sigma_{1,1}$. This is absurd. Thus the lemma follows. $\square$

The following generalizes an observation of Ivanov which he proved for genus at least 2.

Lemma 2.3. *If ϕ: $\mathscr{S}(\Sigma) \to \mathscr{S}(\Sigma)$ is a bijection preserving disjointness so that in the case $\Sigma = \Sigma_{1,2}$, ϕ preserves the separating classes, then for any $\alpha \in \mathscr{S}(\Sigma)$ there exists $h \in$ Home(Σ) so that $h(\alpha) = \phi(\alpha)$.*

Proof. If $\Sigma = \Sigma_{g,n}$ satisfies $3g + n \leqslant 5$, then the lemma is evident. Assume now that $3g + n \geqslant 6$. By Lemma 2.2, we may assume further that α is separating and is not a boundary class. Take $a \in \alpha$ and $b \in \phi(\alpha)$. By Lemma 2.2, we have $\Sigma - \mathrm{int}(N(a)) = \Sigma_{g_1,n_1} \cup \Sigma_{g_2,n_2}$ and $\Sigma - \mathrm{int}(N(b)) = \Sigma_{g'_1,n'_1} \cup \Sigma_{g'_2,n'_2}$ so that the curve complexes $\mathscr{C}(\Sigma_{g_i,n_i})$ and $\mathscr{C}(\Sigma_{g'_i,n'_i})$ are both non-empty. The goal is to show that the two decompositions are homeomorphic. To this end, we note first that $g_1 + g_2 = g = g'_1 + g'_2$ and $n_1 + n_2 = n = n'_1 + n'_2$. Second, the bijection ϕ sends the pair $\{\mathscr{S}(\Sigma_{g_1,n_1}), \mathscr{S}(\Sigma_{g_2,n_2})\}$ to $\{\mathscr{S}(\Sigma_{g'_1,n'_1}), \mathscr{S}(\Sigma_{g'_2,n'_2})\}$ by the same argument as in the proof of Lemma 2.1. Assume without loss of generality that $\phi(\mathscr{C}(\Sigma_{g_i,n_i})) = \mathscr{C}(\Sigma_{g'_i,n'_i})$. It remains to show that $(g_i, n_i) = (g_i', n_i')$ for $i = 1, 2$ in order to finish the proof.

By Lemma 2.1, we obtain $(g_i, n_i) = (g_i', n_i')$ except the following three decompositions which need to be checked specifically. Namely (i) $(g_1, n_1) = (g'_2, n'_2) = (1, 1)$ and $(g_2, n_2) = (g'_1, n'_1) = (0, 4)$; (ii) $(g_1, n_1) = (g'_2, n'_2) = (0, 5)$ and $(g_2, n_2) = (g'_1, n'_1) = (1, 2)$; and (iii) $(g_1, n_1) = (1, 1)$, $(g'_1, n'_1) = (0, 4)$, $(g_2, n_2) = (0, 5)$ and $(g'_2, n'_2) = (1, 2)$. None of these three cases occurs due to Lemma 2.2. Indeed, if there were $\alpha \in \mathscr{S}(\Sigma)$ decomposing the surface Σ into a genus 0 and a genus 1 subsurfaces and ϕ interchanges the two subsurfaces, then ϕ would send a non-separating class to a separating class. $\square$.

Given an incompressible subsurface Σ' in a surface Σ, then each non-separating class in $\mathscr{S}(\Sigma')$ is again non-separating in $\mathscr{S}(\Sigma)$. But separating classes in $\mathscr{S}(\Sigma')$ may become nonseparating in $\mathscr{S}(\Sigma)$. However, if $\Sigma_{1,2}$ is an incompressible subsurface in a surface Σ, then each separating class in $\Sigma_{1,2}$ remains separating in Σ. Thus given an incompressible subsurface $\Sigma_{1,2}$, the inclusion map from $\Sigma_{1,2}$ to Σ preserves both separating and nonseparating classes. Combining with Lemma 2.3, we have the following.

Corollary 2.4. *Suppose the dimension of $\mathscr{C}(\Sigma)$ is at least 2 and $\Sigma_{1,2}$ is an incompressible subsurface in Σ. If ϕ: $\mathscr{S}(\Sigma) \to \mathscr{S}(\Sigma)$ is a bijection preserving the disjointness so that $\phi(\mathscr{S}(\Sigma_{1,2})) = \mathscr{S}(\Sigma_{1,2})$, then $\phi|_{S(\Sigma_{1,2})}$ preserves the separating classes.*

Corollary 2.5. *Suppose $\alpha_1 \in \mathscr{S}(\Sigma)$ decomposes Σ into a union of Σ' and Σ'' so that $\Sigma' \cong \Sigma_{1,1}$ or $\Sigma_{0,4}$ and suppose $\alpha_2 \in \mathscr{S}(\Sigma')$. Then given any bijection ϕ of $\mathscr{S}(\Sigma)$ preserving the disjointness and the separating classes, there is a homeomorphism h of the surface Σ so that $h(\alpha_i) = \phi(\alpha_i)$ for $i = 1, 2$.*

Proof. First, we claim that there is a homeomorphism h_1 of the surface Σ so that $h_1(\alpha_1) = \phi(\alpha_1)$ and $h_1(\mathscr{S}(\Sigma')) = \phi(\mathscr{S}(\Sigma'))$. Indeed, by Lemma 2.3, we find $h_2 \in$ Home(Σ) so that $h_2(\alpha_1) = \phi(\alpha_1)$. By the proof of Lemma 2.3, $h_2(\mathscr{S}(\Sigma')) = \phi(\mathscr{S}(\Sigma'))$ unless $\Sigma'' \cong \Sigma_{1,1}$ or $\Sigma_{0,4}$. If $h_2(\mathscr{S}(\Sigma')) = \mathscr{S}(\Sigma'')$, then $\Sigma' \cong \Sigma''$ because ϕ preserves separating classes. Now let h_3 be an involution of Σ interchanging Σ' and Σ'' and fixing α_1. Then $h_1 = h_3 \circ h_2$ is a required homeomorphism.

Second, let a_1 be the component of $\partial\Sigma'$ corresponding to α_1. The group of homeomorphisms of Σ' leaving a_1 pointwise fixed acts transitively on $\mathscr{S}(\Sigma')$. Thus, we find a homeomorphism h_4 of Σ which is the identity map on Σ'' so that $h_4(\alpha_2) = \phi(\alpha_2)$. The required homeomorphism $h = h_4 \circ h_1$. $\square$

3. A ($\mathbf{Q}P^1$, $PSL(2, \mathbf{Z})$) Structure on $\mathscr{S}(\Sigma)$

3.1.

Recall that surfaces are oriented. Suppose x and y are two open arcs intersecting transversely at a point p in a surface Σ. Then the *resolution of the intersection point p from x to y* is defined as follows. Fix an orientation on x. Use the orientations on the surface Σ and x to determine an orientation on y. Finally resolve the intersection according to the orientations (see Fig. 1(a)). This

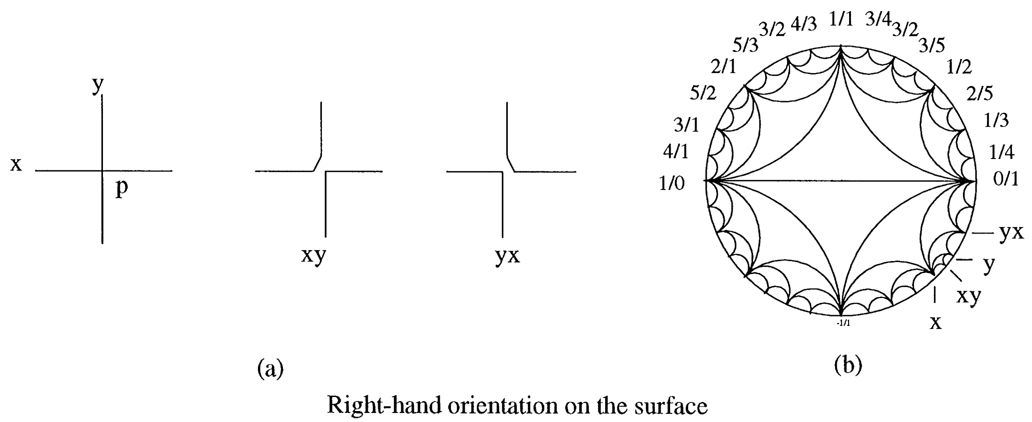

Fig. 1. Right-hand orientation on the surface.

resolution is independent of the choice of orientations on x. Suppose now that α and β are two elements in $\mathscr{S}(\Sigma)$ with $\alpha \perp \beta$ or $\alpha \perp_0 \beta$, we define the multiplication $\alpha\beta$ as follows. Take $a \in \alpha$ and $b \in \beta$ so that $|a \cap b| = I(\alpha, \beta)$. Then $\alpha\beta$ is the isotopy class $[ab]$ where ab is the simple loop obtained by resolving all intersection points in $a \cap b$ from a to b. See Fig. 1(a). One checks easily that if $\alpha \perp \beta$ (resp. $\alpha \perp_0 \beta$) then $\alpha\beta \in \mathscr{S}(\Sigma)$ and $\alpha\beta \perp \alpha, \beta$ (resp. $\alpha\beta \perp_0 \alpha, \beta$).

3.2.

Let $\hat{\mathbf{Q}} = \mathbf{Q} \cup \{\infty\}$. Two rational numbers p/q and p'/q' satisfying $pq' - p'q = \pm 1$ are denoted by $p/q \perp p'/q'$. The relation $(\hat{\mathbf{Q}}, \perp)$ is called the *modular configuration.* A standard way of presenting the configuration is to consider $\hat{\mathbf{Q}}$ as a subset of the boundary of the upper half plane $\mathbf{H}$ and to draw a hyperbolic geodesic ending at p/q and p'/q' if $p/q \perp p'/q'$. Fig. 1(b) is the configuration after a Möbius transformation. It was known to Max Dehn [2] that both $(\mathscr{S}(\Sigma_{1,1}), \perp)$ and $(\mathscr{S}(\Sigma_{0,4}), \perp_0)$ are isomorphic to the modular configuration, i.e., there exists a bijection π between $\mathscr{S}(\Sigma_{1,1})$ (resp. $\mathscr{S}(\Sigma_{0,4})$) and $\hat{\mathbf{Q}}$ so that $\alpha \perp \beta$ (resp. $\alpha \perp_0 \beta$) if and only if $\phi(\alpha) \perp \phi(\beta)$. Furthermore, if $\pi(\alpha) = p/q$ and $\pi(\beta) = p'/q'$ so that $p/q \perp p'/q'$, then $\pi(\alpha\beta) = (p + \lambda q)/(p' + \lambda q')$ where $\lambda = pq' - p'q$. Note that $(\pi(\alpha), \pi(\alpha\beta), \pi(\beta))$ determines the right-hand orientation on the circle.

The following lemma is an easy consequence of the modular configuration.

Lemma 3.1. (a) *If* ϕ: $\mathscr{S}(\Sigma_{1,1}) \to \mathscr{S}(\Sigma_{1,1})$ (*resp.* $\mathscr{S}(\Sigma_{0,4}) \to \mathscr{S}(\Sigma_{0,4})$) *is a bijection preserving the relation* $\perp$ (*resp.* $\perp_0$), *then* ϕ *is induced by a homeomorphism of the surface.*

(b) *Two elements* $\alpha_1, \alpha_2 \in \hat{\mathbf{Q}}$ *satisfy* $\alpha_1 \perp \alpha_2$ *if and only if there are two distinct elements* γ_1, γ_2 *so that* $\gamma_i \perp \alpha_j$ *and* γ_1 *and* γ_2 *are not related by* $\perp$. *Furthermore,* $\{\gamma_1, \gamma_2\} = \{\alpha_1\alpha_2, \alpha_2\alpha_1\}$.

3.3.

We begin by introducing some notations. If $\alpha \perp \beta$ or $\alpha \perp_0 \beta$, we denote it by $\alpha \top \beta$. Given a subset $\mathscr{X} \subset \mathscr{S}(\Sigma)$, let $\mathscr{X}_\infty = \bigcup_{n=0}^{\infty} \mathscr{X}_n$ where $\mathscr{X}_0 = \mathscr{X}$, and $\mathscr{X}_{n+1} = \mathscr{X}_n \cup \{\alpha | \alpha = \beta\gamma$, where $\beta \top \gamma$, and $\beta, \gamma, \gamma\beta$ are in $\mathscr{X}_n\}$. If $\mathscr{X}_\infty = \mathscr{S}(\Sigma)$, we say that $\mathscr{X}$ *generates* $\mathscr{S}(\Sigma)$. For instance, the three-element set $\{\alpha, \beta, \alpha\beta\}$ generates the sets $\mathscr{S}(\Sigma_{1,1})$ and $\mathscr{S}(\Sigma_{0,4})$. The following lemma is motivated by the proof of Lemma 2 in [13]. See also [15].

Lemma 3.2. *Suppose* $\{\alpha_1, \ldots, \alpha_k\}$ *are pairwise disjoint elements in* $\mathscr{S}(\Sigma)$ *and* $\alpha \in \mathscr{S}(\Sigma)$ *so that* $I(\alpha, \alpha_1) \geqslant 2$ *and* α *and* α_1 *are not related by* $\perp_0$. *Then* $\alpha = \beta_1\beta_2$ *where* $\beta_1 \top \beta_2$ *so that* $I(\beta_i, \alpha_1) < I(\alpha, \alpha_1)$, $I(\beta_2\beta_1, \alpha_1) < I(\alpha, \alpha_1)$, $I(\beta_i, \alpha_j) \leqslant I(\alpha, \alpha_j)$, *and* $I(\beta_2\beta_1, \alpha_j) \leqslant I(\alpha, \alpha_j)$ *for* $i = 1, 2$ *and* $j \geqslant 2$. *In particular,* $\mathscr{S}(\Sigma)$ *is generated by the set* $\mathscr{G} = \{\alpha \in \mathscr{S}(\Sigma) |$ *for each i, either* $\alpha \top \alpha_i$ *or* $\alpha \cap \alpha_i = \emptyset\}$.

Proof. Take $a \in \alpha$ and $a_i \in \alpha_i$ so that $|a \cap a_i| = I(\alpha, \alpha_i)$ and $a_i \cap a_j = \emptyset$. Now consider the following two cases.

Case 1. There exist two points $p, q \in a \cap a_1$ which are adjacent in a_1 so that they have the same intersection sign. See Fig. 2.

Assuming that the surface has the right-hand orientation, we take β_1 and β_2 as indicated. Then $\beta_1 \perp \beta_2$, and $\alpha = \beta_1\beta_2$. We verify the required conditions for β_1 and β_2 as in Fig. 2. If the surface has the left-hand orientation, we interchange β_1 and β_2.

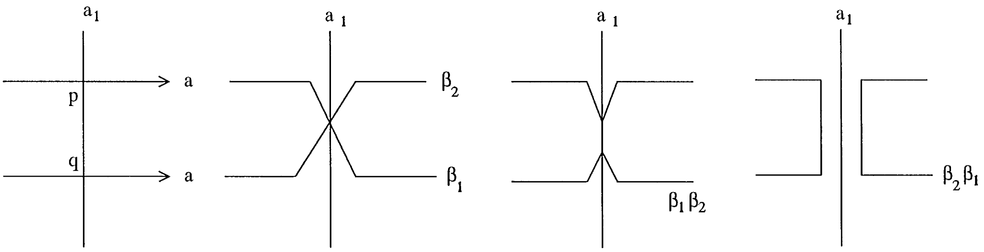

Fig. 2.

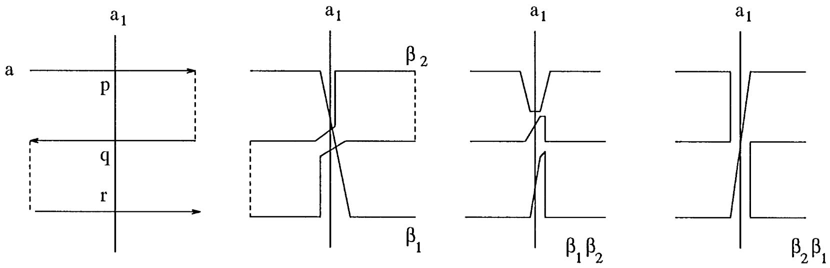

Fig. 3.

Case 2. If case 1 does not occur, then there are three points p, q and r in $a \cap a_1$ which are adjacent in a_1 so that their intersection signs alternate. See Fig. 3. Fix an orientation on a. Assume without loss of generality that the arc in a from p to q does not contain the point r. If the surface has the right-hand orientation, we choose β_1 and β_2 as in Fig. 3. Since $|a \cap a_1| = I(\alpha, \alpha_1)$, β_1 and β_2 are both in $\mathscr{S}(\Sigma)$ and $\beta_1 \perp_0 \beta_2$. We have $\alpha = \beta_1\beta_2$. The required conditions for β_i's are verified as in Fig. 3. If the surface has the left-hand orientation, we interchange β_1 and β_2. $\square$

Remark. A stronger version of the lemma still holds. See [15, Lemma 7].

Corollary 3.3. *Under the same assumption as in Lemma* 3.2, *suppose ϕ and ψ are two bijections of $\mathscr{S}(\Sigma)$ satisfying the following conditions*:

(1) *Both ϕ and ψ preserve the disjointness and relations $\perp$ and $\perp_0$.*
(2) *If $\alpha \top \beta$, then $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$ and $\{\psi(\alpha\beta), \psi(\beta\alpha)\} = \{\psi(\alpha)\psi(\beta), \psi(\beta)\psi(\alpha).\}$*
(3) $\phi|_{\mathscr{G}} = \phi|_{\mathscr{G}}$.
Then $\phi = \psi$.

3.4.

We have mentioned in several places the notion of modular structure on a discrete set. Here is a formal definition after Thurston's geometric structures on manifolds.

Definition. A *modular structure* on a discrete set X is a maximal collection of charts $\{(U_i, \phi_i) \mid i \in I\}$ where ϕ_i: $U_i \to \mathbf{Q}P^1$ is injective so that the following two conditions are satisfied:

(1) The union of the domains of the charts covers X, i.e., $\bigcup_{i \in I} U_i = X$.
(2) The transition functions $\phi_i \phi_j^{-1}$ are restrictions of elements in $PSL(2, \mathbf{Z})$.
A modular structure on X is *compact* if the following additional condition holds:
(3) The automorphism group of the structure $(X, \{(U_i, \phi_i)\})$ acts on X with finite orbits.

The last condition seems to be crucial. Examples of modular structure are $\mathscr{S}(\Sigma)$ and the set of all Fenchel–Nielsen systems (see [16]).

Lemma 3.4. *If Σ is an oriented surface with $\mathscr{S}(\Sigma) \neq \emptyset$, then $\mathscr{S}(\Sigma)$ has a natural modular structure invariant under the action of the orientation preserving mapping class group.*

In fact, as a consequence of the main theorem of the paper, one sees that the automorphism group of the modular structure on $\mathscr{S}(\Sigma)$ is the orientation preserving mapping class group for all surfaces.

Proof. If the dimension of $\mathscr{C}(\Sigma)$ is zero, then the surfaces are $\Sigma_{1,1}$, $\Sigma_{0,4}$ or $\Sigma_{1,0}$. The result follows by the proof of Lemma 2.1. Fix a standard oriented 1-holed torus $\Sigma_{1,1}$ and an identification between $\mathscr{S}(\Sigma_{1,1})$ and $\mathbf{Q}P^1$. If the dimension of the complex $\mathscr{C}(\Sigma)$ is at least one, then any element in $\mathscr{S}(\Sigma)$ lies in an incompressible subsurface Σ' homeomorphic to either $\Sigma_{1,1}$ or $\Sigma_{0,4}$. Assume the subsurface has the induced orientation. Then the charts are $(\mathscr{S}(\Sigma'), \phi)$ where ϕ: $\mathscr{S}(\Sigma') \to \mathscr{S}(\Sigma_{1,1})$ is a bijection produced in the proof of Lemma 2.1 so that ϕ respects the orientations. Extends these charts to be the maximal collection. One checks easily that all conditions are satisfied. □

4. A basic property of the automorphisms of $\mathscr{S}(\Sigma)$

The aim of this section is to prove the following proposition.

Proposition. *Suppose $3g + n \geqslant 5$ and $\phi : \mathscr{S}(\Sigma) \to \mathscr{S}(\Sigma)$ is a bijection preserving disjointness and the separating classes. Then ϕ preserves the relations $\perp$ and $\perp_0$ in $\mathscr{S}(\Sigma)$. Furthermore, if $\alpha \top \beta$, then $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$.*

Proof. We use induction on $|\Sigma_{g,n}| = 3g + n$. The main step is in the case where $|\Sigma| = 1$, i.e., $\Sigma = \Sigma_{0,5}$ (case 1) and $\Sigma_{1,2}$ (case 2).

Case 1. The surface is $\Sigma_{0,5}$. We first show that ϕ preserves the relation $\perp_0$. To this end, take two isotopy classes α_1 and α_2 so that $\alpha_1 \perp_0 \alpha_2$. To show that $\phi(\alpha_1) \perp_0 \phi(\alpha_2)$, we extend $\{\alpha_1, \alpha_2\}$ to a "pentagon" $\{\alpha_1, \ldots, \alpha_5\}$ where $\alpha_i \perp_0 \alpha_{i+1}$ and $\alpha_i \cap \alpha_{i+2} = \emptyset$ (indices i are counted mod 5) as shown in Fig. 4. Here we have used a simple fact that any two pairs of isotopy classes (α, β) with $\alpha \perp_0 \beta$ in $\mathscr{S}(\Sigma_{0,5})$ are related by a homeomorphism of the surface. Indeed, if we take $a \in \alpha$ and $b \in \beta$ with $|a \cap b| = 2$, then the regular neighborhood $N(a \cup b)$ is an incompressible subsurface $\Sigma_{0,4}$. These incompressible subsurfaces are unique up to homeomorphisms of the surface. Thus we may draw (α_1, α_2) as in Fig. 4. Then $\phi(\alpha_i)$'s satisfy the conditions that $\phi(\alpha_i) \cap \phi(\alpha_{i+1}) \neq \emptyset$ and $\phi(\alpha_i) \cap \phi(\alpha_{i+2}) = \emptyset$. Now $\phi(\alpha_1) \perp_0 \phi(\alpha_2)$ follows from the lemma below.

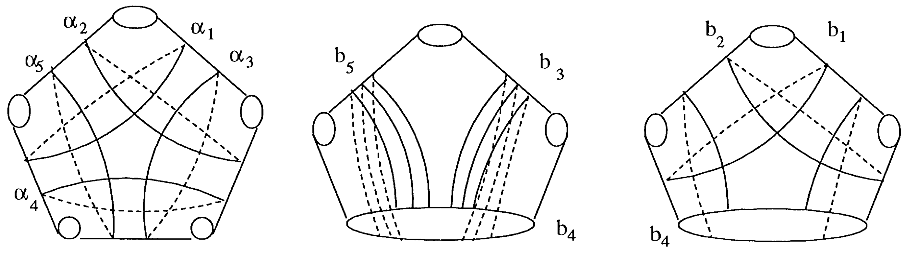

Fig. 4.

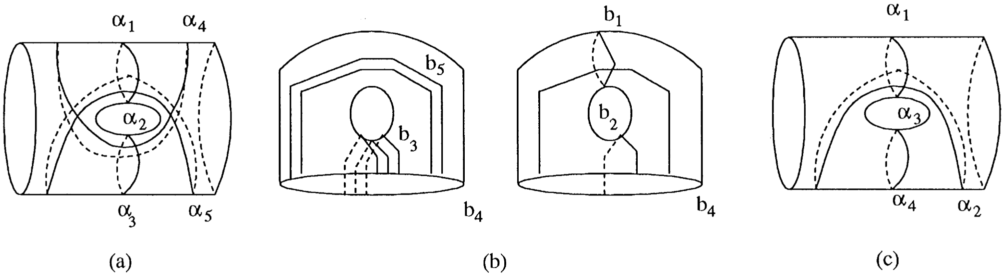

Fig. 5.

Lemma 4.2. *Suppose $\beta_1, \dots, \beta_5$ are five pairwise distinct elements in $\mathscr{S}(\Sigma_{0,5})$ so that $\beta_i \cap \beta_{i+1} \neq \emptyset$ and $\beta_i \cap \beta_{i+2} = \emptyset$ for all indices i (mod 5). Then $\beta_i \perp_0 \beta_{i+1}$ for all i.*

Proof. We shall prove $\beta_1 \perp_0 \beta_2$ only. Take $b_i \in \beta_i$ so that $|b_i \cap b_j| = I(b_i, b_j)$. Consider the subsurface $\Sigma_{0,4}$ bounded by b_4. The subsurface $\Sigma_{0,4}$ contains b_1 and b_2 by the assumption. Since $b_1 \cap b_3 = \emptyset$, we conclude that $b_3 \cap \Sigma_{0,4}$ consists of parallel copies of an arc in $\Sigma_{0,4}$. Furthermore, b_1 is determined up to isotopy by b_3 and b_4. Indeed, b_1 is isotopic to a boundary component of $N(b_3 \cup b_4)$. Another way to see it is to use the following lemma.

Lemma 4.3. *Given two distinct classes in $\mathscr{S}(\Sigma_{0,5})$ (resp. in $\mathscr{S}(\Sigma_{1,2})$), there is at most one class in $\mathscr{S}(\Sigma_{0,5})$ (resp. in $\mathscr{S}(\Sigma_{1,2})$) which is disjoint from both classes.*

To see the proof, we note that the only incompressible subsurfaces of negative Euler number in the surface are $\Sigma_{0,3}$, $\Sigma_{0,4}$ and $\Sigma_{1,1}$. Lemma 4.3 follows by considering the smallest subsurface containing the given classes.

Back to the proof of Lemma 4.2, we have the same conclusion that $b_5 \cap \Sigma_{0,4}$ consists of parallel copies of an arc in $\Sigma_{0,4}$ and b_5 is determined uniquely up to isotopy by b_2 and b_4. Since $b_3 \cap b_5 = \emptyset$, $b_1 \cap b_2$ consists of two points. This shows that $\beta_1 \perp_0 \beta_2$. $\square$.

Case 2. The surface is $\Sigma_{1,2}$. Take $\alpha_1 \top \alpha_2$. We shall discuss three subcases: (2.1) $\alpha_1 \perp \alpha_2$, (2.2) $\alpha_1 \perp_0 \alpha_2$ so that one of α_i is separating, (2.3) $\alpha_1 \perp_0 \alpha_2$ so that both elements α_i are non-separating.

Subcase 2.1. If $\alpha_1 \perp \alpha_2$, we extend $\{\alpha_1, \alpha_2\}$ to a "pentagon" set $\{\alpha_1, \ldots, \alpha_5\}$ as in Fig. 5(a) where $\alpha_i \cap \alpha_{i+2} = \emptyset$, $\alpha_1 \perp_0 \alpha_5$, $\alpha_2 \perp \alpha_3$, $\alpha_3 \perp_0 \alpha_4$, and $\alpha_4 \cap \alpha_5 \neq \emptyset$. Now $\phi(\alpha_1) \perp \phi(\alpha_2)$ follows by the same argument as in case 1 (applied to $\Sigma_{1,1}$ instead of $\Sigma_{0,4}$). See Fig. 5(b).

Subcase 2.2. If $\alpha_1 \perp_0 \alpha_2$ so that α_2 is separating, then α_1 is non-separating. We extend it to a four-element set $\{\alpha_1, \ldots, \alpha_4\}$ as in Fig. 5(c) where $\alpha_3 \cap \alpha_2 = \alpha_1 \cap \alpha_4 = \alpha_4 \cap \alpha_2 = \emptyset$ and $\alpha_1 \perp \alpha_3 \perp \alpha_4$. By subcase 2.1, we conclude that $\phi(\alpha_1) \perp \phi(\alpha_3)$ and $\phi(\alpha_3) \perp \phi(\alpha_4)$. Furthermore, $\phi(\alpha_2) \cap \phi(\alpha_3) = \phi(\alpha_2) \cap \phi(\alpha_4) = \phi(\alpha_1) \cap \phi(\alpha_4) = \emptyset$. Now by Lemma 4.3, $\phi(\alpha_2)$ is determined by $\phi(\alpha_3)$ and $\phi(\alpha_4)$. Thus $\phi(\alpha_1) \perp_0 \phi(\alpha_2)$.

Subcase 2.3. If $\alpha_1 \perp_0 \alpha_2$ so that both α_i's are non-separating, then both $\alpha_1\alpha_2$ and $\alpha_2\alpha_1$ are separating. Since $\alpha_1\alpha_2 \perp_0 \alpha_i$ for $i = 1, 2$, by subcase 2.2, we obtain $\phi(\alpha_1\alpha_2) \perp_0 \phi(\alpha_i)$ for $i = 1, 2$. Similarly, $\phi(\alpha_2\alpha_1) \perp_0 \phi(\alpha_i)$ for $i = 1, 2$. Since α_1, α_2, $\alpha_1\alpha_2$ and $\alpha_2\alpha_1$ are in a subsurface homeomorphic to $\Sigma_{0,4}$, by Lemma 2.2, we conclude that classes $\phi(\alpha_1)$, $\phi(\alpha_2)$, $\phi(\alpha_1\alpha_2)$ and $\phi(\alpha_2\alpha_1)$ are in a subsurface homeomorphic to $\Sigma_{0,4}$ as well. Thus by Lemma 3.1(b) applied to the subsurface $\Sigma_{0,4}$, we have $\phi(\alpha_1) \perp_0 \phi(\alpha_2)$.

To show the last assertion in the proposition for $\Sigma_{1,2}$, take $\alpha_1 \top \alpha_2$. Then $\alpha_1\alpha_2$ is not $\top$-related to $\alpha_2\alpha_1$ and α_1, α_2, $\alpha_1\alpha_2$ and $\alpha_2\alpha_1$ are in a subsurface homeomorphic to $\Sigma_{1,1}$ or $\Sigma_{0,4}$. Since ϕ preserves disjointness and relations $\perp$ and $\perp_0$, $\phi(\alpha_1)$, $\phi(\alpha_2)$, $\phi(\alpha_1\alpha_2)$ and $\phi(\alpha_2\alpha_1)$ are in a subsurface homeomorphic to $\Sigma_{1,1}$ or $\Sigma_{0,4}$ and $\phi(\alpha_1\alpha_2)$ is not $\top$-related to $\phi(\alpha_2\alpha_1)$. Applying Lemma 3.1(b) to the subsurface, we conclude that $\{\phi(\alpha_1\alpha_2), \phi(\alpha_2\alpha_1)\} = \{\phi(\alpha_1)\phi(\alpha_2), \phi(\alpha_2)\phi(\alpha_1)\}$.

We now prove the proposition by induction on $|\Sigma_{g,n}| = 3g + n$. The result holds for $|\Sigma| = 5$ by the above two cases. If $|\Sigma| \geqslant 6$, take $\alpha \top \beta$ in $\mathscr{S}(\Sigma)$. Then α and β lie in an incompressible subsurface homeomorphic to $\Sigma_{1,1}$ or $\Sigma_{0,4}$. Choose a class γ disjoint from α and β so that either γ is non-separating or is a boundary class. Take $c \in \gamma$ and let Σ' be a component of $\Sigma - \mathrm{int}(N(c))$ which contains α and β. Then $|\Sigma'| \geqslant 5$ by the choice of γ. Since $|\Sigma| \geqslant 6$, by Lemma 2.3, there is a homeomorphism h of the surface sending γ to $\phi(\gamma)$. After composing ϕ by h^{-1}, we may assume that $\phi(\gamma) = \gamma$. It follows that $\phi(\mathscr{S}(\Sigma')) = \mathscr{S}(\Sigma')$ by the choice of γ. Thus by Lemma 2.3, $\phi|_{S(\Sigma')}$ preserves the separating classes if $\Sigma' \ncong \Sigma_{1,2}$. If $\Sigma' \cong \Sigma_{1,2}$, then by Corollary 2.4, $\phi|_{S(\Sigma')}$ again preserves the separating classes. Thus by the induction hypothesis applied to Σ', we conclude that if $\alpha \perp \beta$, then $\phi(\alpha) \perp \phi(\beta)$ and if $\alpha \perp_0 \beta$ then $\phi(\alpha) \perp_0 \phi(\beta)$. Furthermore, in both cases, we have $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$. $\square$

5. Proof of the main theorem

Recall that surfaces in this section have negative Euler number. By Proposition 4.1 and Lemma 2.2, it suffices to show the following in order to finishing proof of the main theorem.

Theorem. *Suppose ϕ: $\mathscr{S}(\Sigma) \to \mathscr{S}(\Sigma)$ is a bijection preserving disjointness, the separating classes, the relations $\perp$, $\perp_0$, and $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$. Then $\phi = h$ for some $h \in Home(\Sigma)$.*

Proof. We use induction on $|\Sigma|$. For $|\Sigma| = 4$, the result follows from Lemma 3.1. If $|\Sigma| \geqslant 5$, we decompose $\Sigma = X \cup Y$ where X and Y are compact incompressible subsurfaces so that the following conditions hold: (i) $X \cap Y \cong \Sigma_{0,3}$, (ii) if the genus $g = 0$, then $X \cong \Sigma_{0,4}$ and $Y \cong \Sigma_{0,n-1}$, (iii) if the genus $g \geqslant 1$, then $X \cong \Sigma_{1,1}$ and $Y \cong \Sigma_{g-1,n+2}$. See Fig. 6.

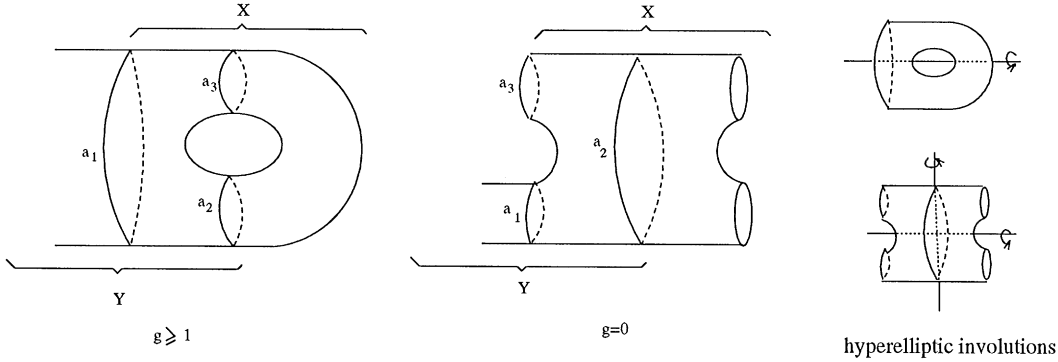

Fig. 6.

We write $\partial(X \cap Y) = a_1 \cup a_2 \cup a_3$ so that $a_1 \subset \partial X$, $a_2 \cup a_3 \subset \partial Y$, and if the genus $g = 0$, $a_3 \subset \partial\Sigma$. By Corollary 2.5, we find $h_1 \in \mathrm{Home}(\Sigma)$ so that $h_1(\phi([a_i])) = [a_i]$ for $i = 1, 2$. Thus, by replacing ϕ by $h_1\phi$, we may assume that $\phi([a_i]) = [a_i]$ for $i = 1, 2$. This implies that $\phi(\mathscr{S}(X)) = \mathscr{S}(X)$ and $\phi(\mathscr{S}(Y)) = \mathscr{S}(Y)$. Now by the construction, $|X|, |Y| < |\Sigma|$ and $Y \ncong \Sigma_{0,3}$. We claim that the restrictions of ϕ to $\mathscr{S}(X)$ and $\mathscr{S}(Y)$ satisfy the induction hypothesis. Evidently the restrictions preserve the disjointness, the relations $\perp$ and $\perp_0$ and $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$. By Lemma 2.3 and Corollary 2.4, the restriction of ϕ to $\mathscr{S}(Y)$ preserves the separating classes. Thus, by the induction hypothesis, we find $h_X \in \mathrm{Home}(X)$, $h_Y \in \mathrm{Home}(Y)$ so that $h_X = \phi|_{S(X)}$, and $h_Y = \phi|_{S(Y)}$.

We shall use the following results to finish the proof of the theorem. The proofs of these results are deferred to the end of this section.

Lemma 5.2. *We may modify h_X and h_Y by composing with hyperelliptic involutions which are in the center of the mapping class group so that after the modification $h_X(a_i) = h_Y(a_i)$, for $i = 1, 2, 3$.*

Proposition 5.3. *Both homeomorphisms h_X and h_Y are orientation preserving or both are orientation reversing.*

Lemma 5.4. *An orientation preserving homeomorphism of the 3-holed sphere leaving each boundary component invariant is isotopic to the identity map.*

By Lemmas 5.2, 5.4 and Proposition 5.3, we conclude that $h_X|_{X \cap Y}$: $X \cap Y \to \Sigma$ and $h_Y|_{X \cap Y}$: $X \cap Y \to \Sigma$ are isotopic. Thus there exists $h \in \mathrm{Home}(\Sigma)$ so that $h|_X \cong h_X$ and $h|_Y \cong h_Y$. We have $\phi|_{S(X) \cup S(Y)} = h|_{S(X) \cup S(Y)}$. The aim is to show that $\phi = h$. Since $\{\phi(\alpha\beta), \phi(\beta\alpha)\} = \{\phi(\alpha)\phi(\beta), \phi(\beta)\phi(\alpha)\}$ and $\{h(\alpha\beta), h(\beta\alpha)\} = \{h(\alpha)h(\beta), h(\beta)h(\alpha)\}$, by Corollary 3.3, it suffices to show that $h(\alpha) = \phi(\alpha)$ for all α so that $\alpha \perp_0 [a_1]$ and either $\alpha \top [a_i]$ or $\alpha \cap [a_i] = \emptyset$ for $i = 2, 3$. Since $[a_1]$ is separating, $I(\alpha, a_2) + I(\alpha, a_3)$ is even. Thus $(I(\alpha, a_2), I(\alpha, a_3))$ is one of the following four pairs (0, 2), (2, 0), (1, 1) or (2, 2). On the other hand, a_3 is either a boundary component or is isotopic to a_2 by the construction. Thus $(I(\alpha, a_2), I(\alpha, a_3)) = (0, 2)$ is impossible. We shall discuss the three cases $(I(\alpha, a_2), I(\alpha, a_3)) = (0, 2), (2, 0)$ and $(1, 1)$ separately.

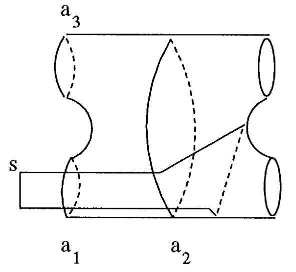

Fig. 7.

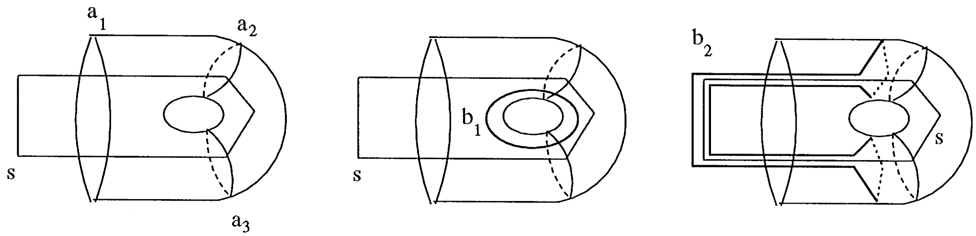

Fig. 8.

The strategy to show $h(\alpha) = \phi(\alpha)$ for these specific elements α is as follows. First we construct an incompressible subsurface $\Sigma' \cong \Sigma_{0,5}$ or $\Sigma_{1,2}$ which contains both X and α. Second, we shall construct two distinct elements $[b_1]$ and $[b_2]$ in $(\mathscr{S}(X) \cup \mathscr{S}(Y)) \cap \mathscr{S}(\Sigma')$ which are disjoint from α. By the assumption, $h([b_i]) = \phi([b_i])$ for $i = 1, 2$ and each elements of $\{\phi(\alpha), h(\alpha)\}$ is disjoint from $\{h([b_1]), h([b_2])\}$. Finally, we show that $\phi(\alpha)$ is in $\mathscr{S}(h(\Sigma'))$. By Lemma 4.3 applied to $h(\Sigma')$ and the pair $\{h([b_1]), h([b_2])\}$, we conclude that $h(\alpha) = \phi(\alpha)$.

Now take $s \in \alpha$ so that $|s \cap a_i| = I(\alpha, a_i)$.

Case 1. $(I(\alpha, a_2), I(\alpha, a_3)) = (2, 0)$ and $\alpha \perp [a_2]$. Then the surface $X \cong \Sigma_{0,4}$.

Let $\Sigma' = N(s) \cup X \cong \Sigma_{0,5}$. Then Σ' is incompressible. Take two essential non-boundary parallel simple loops b_1 and b_2 in Σ' so that (i) $b_1 \subset X$ and $b_2 \subset Y$, (ii) $[b_i] \cap \alpha = \emptyset$, for $i = 1, 2$ and (3) $[b_1] \neq [b_2]$ as shown in Fig. 7.

The isotopy classes of each boundary component of Σ' is either in $\partial\Sigma$ or is in $\mathscr{S}(X) \cup \mathscr{S}(Y)$. By the assumption, we have $h(\beta) = \phi(\beta)$ for each isotopy class β of the component of $\partial\Sigma'$ so that $\beta \in \mathscr{S}(\Sigma)$. Now $\phi(\alpha)$ is disjoint from the isotopy classes of the boundary components of $h(\Sigma')$ and $\phi(\alpha)$ intersets an isotopy class in $\mathscr{S}(h(\Sigma'))$. This shows that $\phi(\alpha)$ is in $\mathscr{S}(h(\Sigma'))$. Furthermore, $\phi(\alpha)$ and $h(\alpha)$ are disjoint from $h([b_i])$ ($= \phi([b_i])$) for $i = 1, 2$. Thus by Lemma 4.3 applied to $h(\Sigma')$, $\phi(\alpha) = h(\alpha)$.

Case 2. $(I(\alpha, a_2), I(\alpha, a_3)) = (1, 1)$. Then $X \cong \Sigma_{1,1}$. Let $\Sigma' = N(s) \cup X$. Then Σ' is incompressible and is homeomorphic to $\Sigma_{1,2}$. Choose two non-isotopic, non-boundary parallel curves b_1 and b_2 in Σ' as in Fig. 8. By the construction, $[b_i] \in \mathscr{S}(X) \cup \mathscr{S}(Y)$ and $[b_i] \cap \alpha = \emptyset$ for $i = 1, 2$. Furthermore, each component of $\partial\Sigma'$ is either in X, Y or in $\partial\Sigma$. Thus $\phi(\alpha) = h(\alpha)$ by the same argument as in case 1.

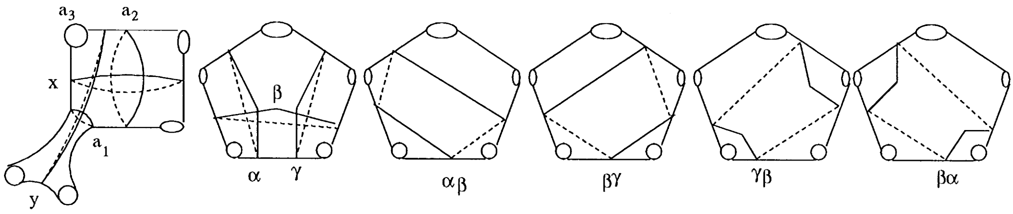

Fig. 9. Surfaces have the right hand orientation.

Case 3. $(I(\alpha, a_2), I(\alpha, a_3)) = (2, 2)$ and $\alpha \perp_0 [a_2]$. This case does not occur. Indeed, $\alpha \perp_0 [a_2]$ shows that the arc $s \cap X$ would intersect a_2 at two points of different signs. Since $X \cong \Sigma_{1,1}$, this shows that $I(s, a_1) = 0$. This contradicts the assumption that $\alpha \perp_0 [a_1]$. $\square$

We now prove Lemma 5.2 and Proposition 5.3. Lemma 5.4 is well known. See for instance [3], exposé 2.

Proof of Lemma 5.2. Since hyperelliptic involutions of X act trivially on $S(X)$, by composing h_X by an isotopy and hyperelliptic involutions, we may assume that $h_X(a_i) = a_i$ for $i = 1, 2, 3$. Since $h_Y(a_1) \cong a_1$, we may assume that $h_Y(a_1) = a_1$ after an isotopy.

If $Y \cong \Sigma_{0,4}$, then we may assume that $h_Y(a_i) = a_i$ for $i = 2, 3$ by composing h_Y by hyperelliptic involutions. Thus the lemma follows in this case.

If $Y \not\cong \Sigma_{0,4}$, then h_Y permutes $\{[a_2], [a_3]\}$. If $g \geqslant 1$, by composing h_X by hyperelliptic involutions if necessary, we obtain $h_X(a_i) = h_Y(a_i)$ for $i = 1, 2, 3$. If the genus $g = 0$, we shall prove that $h_Y(a_i) \cong a_i$ for $i = 2, 3$. Suppose otherwise that $h_Y(a_2) \cong a_3$. Choose a boundary class $\beta \in \mathscr{S}(Y)$ so that β, a_3 and a component b of $\partial Y \cap \partial \Sigma$ bound $\Sigma_{0,3}$ in Y. Thus β is also a boundary class in Σ. By Lemma 2.3, $h_Y(\beta)$ $(= \phi(\beta))$ is again a boundary class in Σ. But $h_Y(\beta)$ is also a boundary class in Y since $h_Y(\beta)$, $a_2 = h_Y(a_3)$ and $h_Y(b)$ bound a 3-holed sphere in Y. Since $[a_2] \in \mathscr{S}(\Sigma)$, this shows that $Y \cong \Sigma_{0,4}$ which contradicts the assumption. $\square$

Proof of Proposition 5.3. Suppose otherwise, we may assume that h_X is orientation reversing and h_Y is orientation preserving. Thus $\phi(\alpha)\phi(\beta) = \phi(\alpha\beta)$ for $\alpha \top \beta$ in $\mathscr{S}(Y)$ and $\phi(\beta)\phi(\alpha) = \phi(\alpha\beta)$ for $\alpha \top \beta$ for α, β in $\mathscr{S}(X)$.

If the genus $g = 0$, construct two curves x and y as in Fig. 9 so that $[x] \in \mathscr{S}(X)$, $[y] \in \mathscr{S}(Y)$, $[x] \perp_0 [y]$, $[x] \perp_0 [a_2]$, $[y] \perp_0 [a_1]$, and $|y \cap a_1| = 2$. Then $\phi(a_2 x) = \phi(x)\phi(a_2)$ and $\phi(a_1 y) = \phi(a_1)\phi(y)$. Furthermore, the subsurface $\Sigma' = N(y) \cup X \cong \Sigma_{0,5}$ is incompressible in Σ. Thus two classes $\alpha, \beta \in \mathscr{S}(\Sigma')$ are disjoint in $\mathscr{S}(\Sigma)$ if and only if they are disjoint in $\mathscr{S}(\Sigma')$. We now use Lemma 5.5 below to derive a contradiction.

Lemma 5.5. *Suppose* $\alpha \perp_0 \beta \perp_0 \gamma$, $\alpha \cap \gamma = \emptyset$ *in* $\mathscr{S}(\Sigma_{0,5})$. *Then* $\alpha\beta \cap \gamma\beta = \emptyset$, $\alpha\beta \cap \beta\gamma \neq \emptyset$, $\beta\alpha \cap \gamma\beta \neq \emptyset$, *and* $\beta\alpha \cap \beta\gamma = \emptyset$.

Proof. Take a triple (α, β, γ) as in Fig. 9. Then the lemma follows for the triple in Fig. 9 by the calculation in the figure. On the other hand, there is only one triple (α, β, γ) satisfying the conditions in the lemma up to self-homeomorphisms of the surface. Thus the lemma follows. To see the

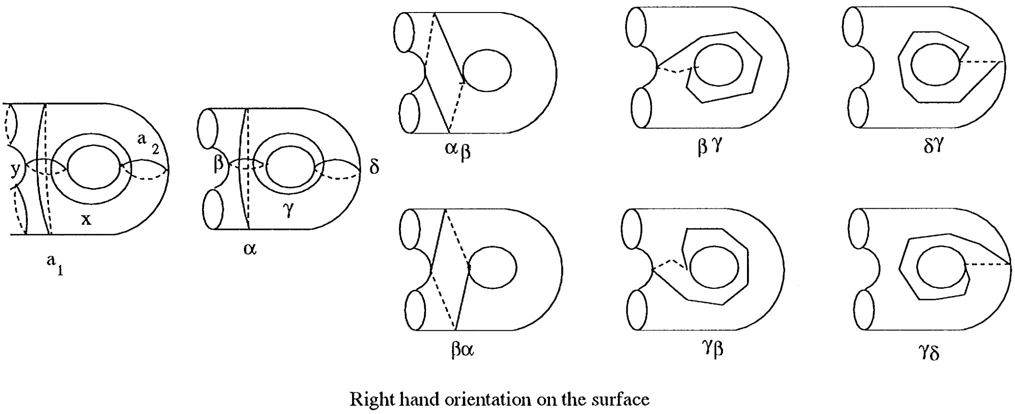

Fig. 10. Right-hand orientation on the surface.

uniqueness of the triple (α, β, γ) up to homeomorphisms, we take three representatives a, b, c in α, β, γ, respectively, so that they intersect minimally. Then the surface $\Sigma_{0,5}$ is homeomorphic to a regular neighborhood $N(a \cup b \cup c)$. Furthermore, the union $a \cup b \cup c$ is unique up to homeomorphisms. Thus the assertion follows. $\square$

Apply Lemma 5.5 to $(\alpha, \beta, \gamma) = ([a_1], [y], [x])$ and $(\phi(a_1), \phi(y), \phi(x))$. We obtain

$$[a_1][y] \cap [x][y] = \emptyset \qquad [a_1][y] \cap [y][x] \neq \emptyset \tag{1}$$

and

$$\phi(a_1)\phi(y) \cap \phi(x)\phi(y) = \emptyset \qquad \phi(a_1)\phi(y) \cap \phi(y)\phi(x) \neq \emptyset. \tag{2}$$

By applying ϕ to (1) and use $\phi(a_1 y) = \phi(a_1)\ \phi(y)$, we obtain

$$\phi(a_1)\phi(y) \cap \phi(xy) = \emptyset \qquad \phi(a_1)\ \phi(y)\ \cap \phi(yx) \neq \emptyset \tag{3}$$

Since $\{\phi(xy), \phi(yx)\} = \{\phi(x)\phi(y), \phi(y)\ \phi(x)\}$, by comparing Eqs. (2) and (3), we obtain

$$\phi(xy) = \phi(x)\phi(y). \tag{4}$$

If we apply Lemma 5.5 to $(\alpha, \beta, \gamma) = ([a_2], [x], [y])$ and $(\phi(a_2), \phi(x), \phi(y))$ and use $\phi(a_2 x) = \phi(a)\phi(a_2)$, we obtain $\phi(xy) = \phi(y)\phi(x)$. This contradicts (4).

If the genus $g = 1$, we construct two curves x, y as shown in Fig. 10 where $[x] \in \mathscr{S}(X)$, $[y] \in \mathscr{S}(Y)$ so that $[x] \perp [y]$, $[x] \perp [a_2]$, $[y] \perp_0 [a_1]$ and $|y \cap a_1| = 2$. The subsurface $\Sigma' = N(y) \cup X \cong \Sigma_{1,2}$ is incompressible in Σ. We use Lemma 5.6 below to obtain a contradiction.

Lemma 5.6. *If $\alpha \perp_0 \beta \perp \gamma \perp \delta$, $\alpha \cap \delta = \beta \cap \delta = \beta \cap \gamma = \emptyset$ in $\mathscr{S}(\Sigma_{1,2})$, then $\alpha\beta \perp \gamma\beta$, $\alpha\beta$ is not $\top$-related to $\beta\gamma$, $\beta\alpha$ is not $\top$-related to $\gamma\beta$, and $\beta\alpha \perp \beta\gamma$. Furthermore, $\delta\gamma \cap \beta\gamma = \gamma\delta \cap \gamma\beta = \emptyset$, and $\delta\gamma \cap \gamma\beta \neq \emptyset$, $\gamma\delta \cap \beta\gamma \neq \emptyset$.*

See Fig. 10 for a verification of the lemma for a specific choice of the quadruple $(\alpha, \beta, \gamma, \delta)$. But the quadruple satisfying the conditions in the lemma is unique up to self-homeomorphism of the surface. Indeed, by Lemma 4.3, δ is uniquely determined by α, β, and α is uniquely determined by δ, γ. The uniqueness of the triple (β, γ, δ) (resp. (α, β, γ)) follows by the same argument as in Lemma 5.5.

Now the proof is similar to the previous case. Namely, by the choice of x, y, we have $\phi(a_1 y) = \phi(a_1)\phi(y)$ and $\phi(a_2 x) = \phi(x)\phi(a_2)$. Applying Lemma 5.6 to the triples $([a_1], [x], [y])$ and $(\phi(a_1), \phi(y), \phi(x))$ (as (α, β, γ)), we obtain $\phi(xy) = \phi(x)\,\phi(y)$. If we apply the lemma to the different triples $([a_2], [x], [y])$ and $(\phi(a_2), \phi(x), \phi(y))$ (as (β, γ, δ)), we obtain $\phi(xy) = \phi(y)\phi(x)$. This is a contradiction. $\square$

Acknowledgements

I would like to thank X.S. Lin for many discussions, F. Bonahon for informing me Ivanov's work and N. Ivanov for sending his unpublished 1989 preprint. I thank the referee for many helpful suggestions. The work is supported in part by the NSF.

References

[1] J.S. Birman, Braids, links, and mapping class groups, Ann. Math. Stud., vol. 82, Princeton Univ. Press, Princeton, NJ, 1975.
[2] M. Dehn, in: J. Stillwell (Ed.), Papers on Group Theory and Topology, Springer, Berlin, 1987.
[3] A. Fathi, F. Laudenbach, V. Poenaru, Travaux de Thurston sur les surfaces, Asterisque 66–67, Societe Mathematique de France, 1979.
[4] A. Grothendieck, Esquisse d'un programme. In: Geometric Galois action, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, pp. 5–48.
[5] J. Harer, Stability of the homology of the mapping class groups of orientable surfaces, Ann. Math. 121 (2) (1985) 215–249.
[6] J. Harer, The virtual cohomological dimension of the mapping class group of an orientable surface, Invent. Math. 84 (1986) 157–176.
[7] W. Harvey, Boundary structure of the modular group, Riemann surfaces and related topics, in: Proceedings of the 1978 Stony Brook Conference, State Univ. New York, Stony Brook, N.Y., 1978, Ann. Math. Stud. 97, Princeton Univ. Press, Princeton, N.J., 1981, pp. 245–251.
[8] J. Hempel, 3-manifolds from curve complex point of view, preprint.
[9] N.V. Ivanov, Complexes of curves and Teichmüller modular groups, Uspekhi Mat. Nauk. 42 (3) (1987) 49–91.
[10] N.V. Ivanov, Complexes of curves and Teichmüller spaces, Math. Notes 49 (1991) 479–484.
[11] N.V. Ivanov, Automorphisms of complexes of curves and of Teichmüller spaces. Int. Math. Res. Notice 14 (1997) 651–666.
[12] M. Korkmaz, Complexes of curves and the mapping class groups, Topology Appl., to appear.
[13] R. Lickorish, A representation of oriented combinatorial 3-manifolds, Ann. Math. 72 (1962) 531–540.
[14] F. Luo, On non-separating simple closed curves in a compact surface, Topology 36 (2) (1997) 381–410.
[15] F. Luo, Geodesic length functions and Teichmüller spaces, J. Differential Geom. 48 (1998) 275–317.
[16] F. Luo, Grothendieck's reconstruction principle and 2-dimensional topology and geometry, preprint, 1998.
[17] H. Masur, Y. Minsky, Geometry of the complex of curves I: hyperbolicity, preprint 1996.
[18] D.B. Patterson, The Teichmüller spaces are distinct. Proc. Amer. Math. Soc. 35 (1972) 179–182; 38 (1973) 668–669.
[19] W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces. Bull. Amer. Math. Soc. 19 (2) (1988) 417–438.
[20] J. Tits, On buildings and their applications, in: R.D. James, (Ed.), International Congress of Mathematicians, Vancouver, B.C., 1974, pp. 221–228.
[21] O. Viro, Links, two-fold branched coverings and braids, Soviet Math. Sbornik 87 (2) (1972) 216–228.