ELSEVIER

Topology and its Applications 129 (2003) 111–127

TOPOLOGY AND ITS APPLICATIONS

www.elsevier.com/locate/topol

Heegaard diagrams and handlebody groups

Feng Luo

Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

Received 25 February 2001; received in revised form 16 May 2002

Abstract

Heegaard diagrams on the boundary of a handlebody are studied from the dynamics systems point of view. A relationship between the strongly irreducible condition of Casson–Gordan and the Masur's domain of discontinuity for the action of the handlebody group is established.

MSC: 57M99; 57N70

Keywords: Handlebody; Heegaard diagram; Handlebody group

1. Introduction

By a *Heegaard diagram on a handlebody* we mean the isotopy class of a maximal collection of disjoint, pairwise non-isotopic essential simple loops (a pants-decomposition) on the boundary surface. We say a Heegaard diagram is *strongly irreducible* if each component of the diagram intersects all meridian discs. The notion of strongly irreducibility is motivated by the work of Casson and Gordon [2] on the strongly irreducible Heegaard splittings in which any two meridian discs from different handlebodies intersect. The aim of the paper is to study the space of Heegaard diagrams from the dynamics system point of view. The dynamics system consists of the action of the handlebody group on Thurston's space of measured laminations [6,11,13,15,19]. H. Masur made a deep study of the dynamics system and found the maximal open subset on which the handlebody group acts properly discontinuously. Our result is the following.

Theorem 1. (a) *A Heegaard diagram is in Masur's domain of discontinuity if and only if it is strongly irreducible in the above sense.*

E-mail address: fluo@math.rutgers.edu (F. Luo).

0166-8641/02/$ – see front matter
doi:10.1016/S0166-8641(02)00142-6

(b) *If α is a strongly irreducible Heegaard diagram and d is a hyperbolic metric on the boundary of the handlebody of a genus at least two, there is a computable constant $K > 0$ (depending only on α, d) so that $I(\alpha, \partial D) \geqslant K l_d(\partial D)$ for all meridian discs D in the handlebody where $l_d(x)$ is the length of the geodesic representative isotopic to x and $I(x, y)$ is the geometric intersection number.*

It is well known that for any hyperbolic metric on a closed surface and for any number n, there is an algorithm to list the isotopy classes of loops of length at most n. As a consequence, one sees that for any number n there is an algorithm to find the set of isotopy classes of meridian discs whose intersection number with a strongly irreducible Heegaard diagram α is at most n. Thus we obtain the following.

Corollary 2 (Johannson [5]). *Given two Heegaard diagrams so that one of them is strongly irreducible, there is an algorithm to decide if these two diagrams are related by a handlebody homeomorphism.*

The part (a) of Theorem 1 follows easily from a theorem of Starr [18] which characterizes irreducible curves systems on the boundary of a handlebody. (See [20, p. 689], for a short proof of Starr's theorem.) We were not aware of Starr's theorem when we worked on Theorem 1 and produced a proof Starr's theorem using the results obtained in [8]. This proof may be of some interests as it uses defining equations for the geometric intersection number functions.

The proof of part (b) of Theorem 1 uses a simple fact on counting the intersection points of curve systems on surfaces (Lemma 2.1). Namely, a string of a straight arcs and a string of b straight arcs in a convex planar region intersect at most ab points unless some arcs overlap. This counting lemma is used repeatedly to obtain the estimate on the constant K.

The paper is organized as follows. In Section 2, we introduce some background material. In particular, the intersection number with a pair of surface filling curve systems is emphasized. This intersection number is the combinatorial analogue of the length of the geodesics. We prove Lemma 2.1 which is the counterpart of an inequality of Thurston [3, p. 58]. In Section 3, we give a new proof of Starr's theorem. Theorem 1 and Corollary 2 are proven in Section 4. Some questions about Heegaard splittings are raised in Section 5. Also in Section 5, we discuss the relationship between the strongly irreducible Heegaard splittings and the Heegaard diagrams. In Appendix A we give a second proof of Starr's theorem.

2. Preliminaries on the measured lamination space

Let us fix a set of notations.

$\Sigma_{g,r}$ is the compact orientable surface of genus g with r boundary components;

$\mathcal{S} = \mathcal{S}(\Sigma_{g,r})$ is the set of isotopy classes of essential simple loops on $\Sigma_{g,r}$;

$\mathcal{C}S = \mathcal{C}S(\Sigma_{g,r})$ is the set of isotopy classes of curve systems on $\Sigma_{g,r}$ where a *curve system* is a finite disjoint union of essential non-boundary parallel simple loops and essential proper arcs on the surface;

H_g is the handlebody of genus $g \geqslant 2$;

Unless stated otherwise, we take the surface Σ to be the boundary of the handlebody H_g in the rest of the paper;

$\mathcal{F}N = \mathcal{F}N(\Sigma)$ is the set of isotopy classes of pants-decompositions on the surface;

$\mathcal{C}S_t = \mathcal{C}S_t(\Sigma)$ is the subset of $\mathcal{C}S = \mathcal{C}S(\Sigma)$ consisting of curve systems so that each component of the system is null homotopic in H_g;

$\mathcal{S}_t = \mathcal{C}S_t(\Sigma) \cap \mathcal{S}(\Sigma)$ is the set of isotopy classes of the boundary of meridian discs;

$\mathcal{F}N_t = \mathcal{C}S_t \cap \mathcal{F}N$ is the set of pants-decompositions of the handlebody by the meridian discs;

$\mathrm{Mod}(\Sigma)$ is the mapping class group $\mathrm{Homeo}^+(\Sigma)/\mathrm{Iso}$ of the surface;

$\Gamma = \Gamma_g$ is the handlebody group consisting of isotopy classes of homeomorphisms of the surfaces which extend to homeomorphisms of the handlebody;

$\mathcal{M}L = \mathcal{M}L(\Sigma)$ is Thurston's space of measured laminations on the surface.

The isotopy class of a 1-dimensional submanifold a is denoted by $[a]$. The geometric intersection number between two isotopy classes α, β is denoted by $I(\alpha, \beta) = \min\{|a \cap b|\colon\ a \in \alpha,\ b \in \beta\}$. We also use $I(a, b)$, $I([a], b)$ and $I(a, [b])$ to denote $I([a], [b])$. The intersection number function on the measured lamination spaces will also be denoted by I. A component of an isotopy class $[a] \in \mathcal{C}S$ is the isotopy class of a component of a. A regular neighborhood of a 1-dimensional submanifold a is denoted by $N(a)$. For details on the space of measured laminations, see [1,3,16,19] and the references cited therein.

Definition (*Masur* [11]). The limit set L for the action of the handlebody group Γ on the space of measured laminations is defined to be the closure of the set $\mathbf{Q}_{>0} \times \mathcal{S}_t$ in the space of measured laminations $\mathcal{M}L$. Let $\Omega = \{\alpha \in \mathcal{M}L(\Sigma)\colon\ I(\alpha, \beta) > 0 \text{ for all } \beta \in L - 0\}$.

Theorem (Masur [11]). *The set Ω is the maximal open subset on which the handlebody group Γ acts properly discontinuously.*

An equivalent definition of elements in the Masur domain Ω is as follows.

Definition. A set $A = \{\alpha_1, \ldots, \alpha_n\} \subset \mathcal{C}S(\Sigma)$ is called *surface filling* if $\sum_{i=1}^n I(x, \alpha_i) > 0$ for all $x \in \mathcal{C}S(\Sigma)$. In this case, define the *norm* induced by A on $\mathcal{C}S(\Sigma)$ to be $|x|_A = \sum_{i=1}^n I(x, \alpha_i)$.

Lemma 2.1. *If $A = \{\alpha_1, \ldots, \alpha_k\} \subset \mathcal{C}S(\Sigma)$ is surface filling, then for any $\alpha, \beta \in \mathcal{C}S(\Sigma)$, we have $I(\alpha, \beta) \leqslant |\alpha|_A |\beta|_A$.*

Proof. Choose representatives $a_i \in \alpha_i$, $a \in \alpha$ and $b \in \beta$ as curve systems so that their pairwise intersection numbers are minimal within the isotopy classes and there are no triple intersection points. Since A is surface filling, each component of $\Sigma - \bigcup_{i=1}^n a_i$ is contractible. Say $R_1, \ldots, R_m$ are these components. Let x_i (respectively y_i) be the number of connected components of a (respectively b) in R_i. Since there are no bi-gons in $a \cup b$ inside R_i, the number of intersection points $a \cap b$ inside R_i is at most $x_i y_i$. Thus, $I(a, b) \leqslant \sum_{i=1}^m x_i y_i \leqslant (\sum_{i=1}^m x_i)(\sum_{i=1}^m y_i) = |\alpha|_A |\beta|_A$. □

Remark 2.1. The above lemma still holds if one replaces the surface filling set A by a surface filling graph G, i.e., the components of $\Sigma - G$ are contractible and define the norm of α to be $\min\{|a \cap G|\colon\ a \in \alpha$ and a does not pass through the vertices of $G\}$.

2.2. A finite subset $A \subset \mathcal{C}S$ is surface filling if and only if $\{\gamma \in \mathrm{Mod}(\Sigma)\colon\ \gamma(A) = A\}$ is a finite group.

Corollary 2.2. *Any two norms arising in this way are Lipschitz related.*

Indeed, say $|\alpha| = |\alpha|_A$ and $||\alpha|| = |\alpha|_B$ for two surface filling sets A and B. Then

$$|\alpha| = \sum_{i=1}^{n} I(\alpha, \alpha_i) \leqslant \sum_{i=1}^{n} \|\alpha\|\|\alpha_i\| = \left(\sum_{i=1}^{n} \|\alpha_i\|\right)\|\alpha\|.$$

Remark 2.3. Fix a hyperbolic metric on the surface and let $l(x)$ be the length of the geodesic in $x \in \mathcal{C}S$. Then for any norm $|x|$ on $\mathcal{C}S(\Sigma)$, there is a constant K_1 so that $\frac{1}{K_1}|x| \leqslant l(x) \leqslant K_1|x|$ for all $x \in \mathcal{C}S(\Sigma)$. Thus, the lemma above is a combinatorial analogous to Thurston's inequality that $I(x, y) \leqslant K_2 l(x) l(y)$ for all $x, y \in \mathcal{C}S(\Sigma)$ [3, Lemma 2, p. 58].

2.4. Fix a norm $|x|$ on $\mathcal{C}S(\Sigma)$. For each $r \in \mathbf{Z}$, let $N(r)$ be the number of elements in $\mathcal{C}S(\Sigma)$ of norm r. It can be shown easily that $N(r)$ has polynomial growth in r. Thus the function $\sum_{r=1}^{\infty} N(r)t^r$ is convergent for $|t| < 1$. Is the function rational?

2.5. Using Lemma 2.1, one can give a proof of Thurston's result that the projective measured lamination space $P\mathcal{M}L(\Sigma)$ is compact. Indeed, given a sequence $\{x_n\}$ in $\mathcal{C}S(\Sigma) - 0$, then for any $\beta \in \mathcal{S}$, the sequence $I(x_n/|x_n|, \beta)$ is bounded by $|\beta|$ by Lemma 2.1. By the standard Cantor diagonal process, we find a subsequence, still denoted by x_n so that $I(x_n/|x_n|, \beta)$ converges to a function $f(\beta)$ for all $\beta \in \mathcal{S}$. To show that the function f is not identically zero, consider the sum of the values of f on the elements α_i in the set defining the norm. The sum is 1 by definition.

Fix a norm $|x|$ on $\mathcal{C}S(\Sigma)$. Then an element $x \in \mathcal{M}L(\Sigma)$ satisfies $I(x, y) > 0$ for all $y \in L - 0$ if and only if $I(x, y) \geqslant K|y|$ for all $y \in L$ by the compactness of $PL = \{t/|t|\colon\ t \in L - 0\}$ in the projective measured lamination space $PML(\Sigma)$. Thus an element $x \in \mathcal{M}L(\Sigma)$ is in the Masur domain Ω if and only if the restriction of the intersection number function $I(x, .)$ on the limit set L is Lipschitz equivalent to a norm. We may rephrase the part (a) of Theorem 1 as follows.

Theorem 2.3. *Suppose $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{F}N(\Sigma)$ so that $I(\alpha_i, \beta) > 0$ for all i and all $\beta \in \mathcal{S}_t$. Then there is a constant $K > 0$ so that $I(\alpha, \beta) \geqslant K|\beta|$ for all $\beta \in \mathcal{S}_t$.*

The following lemma was known to many mathematicians [12].

Lemma 2.4. *If $\alpha \in \mathcal{S}(\Sigma)$ so that $I(\alpha, \alpha') = 0$ for some meridian disc $\alpha' \in \mathcal{S}_t$, then α is in the limit set L.*

Proof. We need to consider two cases: either α' is non-separating or α' is separating.

Case 1. If α' is non-separating, choose $\beta \in \mathcal{S}$ so that $I(\alpha', \beta) = 1$ and $I(\alpha, \beta) \neq 0$. Then $\beta_n = D_\alpha^n(\beta)$, where D_α is the positive Dehn twist on α, converges projectively to α in $\mathcal{M}L(\Sigma)$ and $I(\beta_n, \alpha') = 1$. Let $\alpha_n = \partial N(\alpha' \cup \beta_n)$ be the isotopy class of a regular neighborhood of a 1-holed torus which contains both α' and β_n. Since $\alpha' \in \mathcal{S}_t$ and $I(\beta_n, \alpha') = 1$, α_n is in $\mathcal{S}_t$. Furthermore α_n converges projectively to α. It follows that α is in the limit set L.

Case 2. If α' is non-separating, then the meridian disc bounded by α' cuts the handlebody into two handlebodies. Choose α'' to be a non-separating meridian disc in one of the handlebody which does not contain α. Then α is in L by case 1 applied to $\{\alpha, \alpha''\}$. □

The following lemma shows the main advantage of using pants-decompositions as Heegaard diagrams. The proofs are evident except part (c).

Lemma 2.5. (a) *If h is a homeomorphism leaving a Heegaard diagram invariant, then $h^{(3g-3)!}$ is a composition of Dehn twists on the components of the Heegaard diagram.*

(b) *If $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3}$ is a Heegaard diagram and $\beta \in \mathcal{M}L(\Sigma)$ so that $I(\alpha, \beta) = 0$, then $\beta = k_1\alpha_1 \sqcup \cdots \sqcup k_{3g-3}\alpha_{3g-3}$ where $k_i \in \mathbf{R}_{\geqslant 0}$.*

(c) *Given any integer n, there are only finitely many pairs of Heegaard diagrams $(\alpha, \beta) \in \mathcal{F}N(\Sigma) \times \mathcal{F}N(\Sigma)$ up to homeomorphisms of the surface so that $I(\alpha, \beta) \leqslant n$. Furthermore, these finitely many pairs can be listed algorithmically.*

To show (c), we first note that there are only finitely many Heegaard diagrams up to homeomorphisms of the surface. Thus, it suffices to count the set $\{\beta \in \mathcal{F}N \mid I(\alpha, \beta) \leqslant n\}$ modulo Dehn twists on α for a fixed $\alpha \in \mathcal{F}N$. Consider the Dehn–Thurston coordinate of β with respect to the pants-decomposition based on α (see [16] and [10] for more details on Dehn–Thurston coordinate). Then each β has the coordinate of the form $(x_1, t_1, \ldots, x_{3g-3}, t_{3g-3})$ where $x_i \leqslant n$ is the intersection number coordinate and t_i is the twisting coordinate. If $|t_i| > n$, then we may use the Dehn twist on the ith component of α to change β so that the new twisting number is within the interval $[0, n]$. Thus the result follows.

Remark 2.6. A stronger form of Lemma 2.5(c) holds. Namely, for any n the set $\{(\alpha, \beta) \in \mathcal{S} \times \mathcal{S}\colon\ I(\alpha, \beta) \leqslant n\}/\mathrm{Mod}(\Sigma)$ is finite.

We end this section by giving a proof of Masur's theorem in terms of norms. The basic ideas are due to Masur. We begin with a lemma characterizing compact sets in the Masur domain Ω in terms of norms. Fix a norm $|x| = |x|_A$ where $A = \{\alpha, \beta\} \subset \mathcal{C}S_t$ is surface filling.

Lemma 2.6. *If K is a compact subset in Ω, then there is a constant $c > 0$ so that for all $x \in K$ and $t \in L$, we have $\frac{1}{c}|t| \leqslant I(x, t) \leqslant c|t|$.*

The proof uses the standard compactness argument. For instance, if the left-hand-side inequality fails, then there are $x_n \in K$ and $t_n \in L - 0$ so that $\frac{1}{n}|t_n| \geqslant I(x_n, t_n)$ for all integer n. Then by choosing a subsequence (still denoted by the same index), we may assume that x_n converges to $x \in K$ and $t_n/|t_n|$ converges to $t \in L - 0$. By the continuity of the intersection number function $I(.,.)$, we obtain that $I(x,t) = 0$. But this contradicts the assumption that $x \in \Omega$ and $A \subset \mathcal{CS}_t$.

Now to prove Masur's theorem, take a compact set K in Ω. We shall prove that there are only finitely many elements $\gamma \in \Gamma$ so that $\gamma(K) \cap K \neq \emptyset$. By Lemma 2.6, there is a constant $c > 0$ so that $\frac{1}{c}|t| \leqslant I(x,t) \leqslant c|t|$ for all $x \in K$ and $t \in L$. Suppose $\gamma(K) \cap K \neq \emptyset$. Then there is $x \in K$ so that $\gamma(x) \in K$. Thus we have, $\frac{1}{c}|t| \leqslant I(x,t) \leqslant c|t|$, and $\frac{1}{c}|t| \leqslant I(\gamma(x), t) \leqslant c|t|$ for all $t \in L$. By the choice of the surface filling set $A = \{\alpha, \beta\} \subset L$, we have,

$$\begin{aligned} \frac{1}{c}|\gamma(A)| &= \frac{1}{c}\big(|\gamma(\alpha)| + |\gamma(\beta)|\big) \\ &\leqslant I\big(\gamma(x), \gamma(\alpha)\big) + I\big(\gamma(x), \gamma(\beta)\big) \\ &= I(x,\alpha) + I(x,\beta) \leqslant c(|\alpha| + |\beta|). \end{aligned}$$

This shows that the norm of $\gamma(A)$ is bounded by a constant independent of γ. There are only finitely many elements in $\mathcal{CS}$ of norm at most a given number and also the set $\{\gamma \in \Gamma\colon \gamma(A) = A\}$ is finite due to the surface filling property of A. Therefore, we see that there are only finitely many $\gamma \in \Gamma$ with $\gamma(K) \cap K \neq \emptyset$.

3. A proof of Starr's theorem

We begin by introducing some notations. Let

$$\Delta = \big\{(x_1, x_2, x_3) \in \mathbf{R}^3_{\geqslant 0}\colon\ x_i + x_j \geqslant x_k,\ i \neq j \neq k \neq i\big\};$$

$$\Delta^+ = \big\{(x_1, x_2, x_3) \in \mathbf{R}^3_{>0}\colon\ x_i + x_j > x_k,\ i \neq j \neq k \neq i\big\};$$

$$W = \big\{(x_1, x_2, x_3) \in \mathbf{R}^3_{\geqslant 0}\colon \text{ there is an index } i \text{ so that } x_i \geqslant x_j + x_k\big\};$$

$$W^+ = \big\{(x_1, x_2, x_3) \in \mathbf{R}^3_{\geqslant 0}\colon \text{ there is an index } i \text{ so that } x_i > x_j + x_k\big\}.$$

Evidently we have $\Delta \cap W^+ = \Delta^+ \cap W = \emptyset$ and $\Delta \cup W^+ = \Delta^+ \cup W = \mathbf{R}^3_{\geqslant 0}$. We say that three elements $\alpha, \beta, \gamma \in \mathcal{S}$ *bound* a 3-holed sphere, denoted by $(\alpha, \beta, \gamma) \in \mathcal{P}$, if there are representatives a, b, c in α, β, γ respectively so that a, b, c bound a 3-holed sphere in the surface. Note that two of the elements $\{\alpha, \beta, \gamma\}$ may be the same. See Fig. 1.

Definition. (a) An element $\alpha \in \mathcal{CS}$ is called irreducible with respect to the handlebody if $I(\alpha, \beta) > 0$ for all $\beta \in \mathcal{S}_t$. Let $\mathcal{CS}^+(\Sigma)$ be the set of all isotopy classes of irreducible curve systems.

(b) Given a Heegaard diagram $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{FN}(\Sigma)$, we associate to α the following sets: $\Delta(\alpha) = \{\beta \in \mathcal{ML}(\Sigma)\colon$ if $(\alpha_i, \alpha_j, \alpha_k) \in \mathcal{P}$, then $(I(\alpha_i, \beta), I(\alpha_j, \beta),$

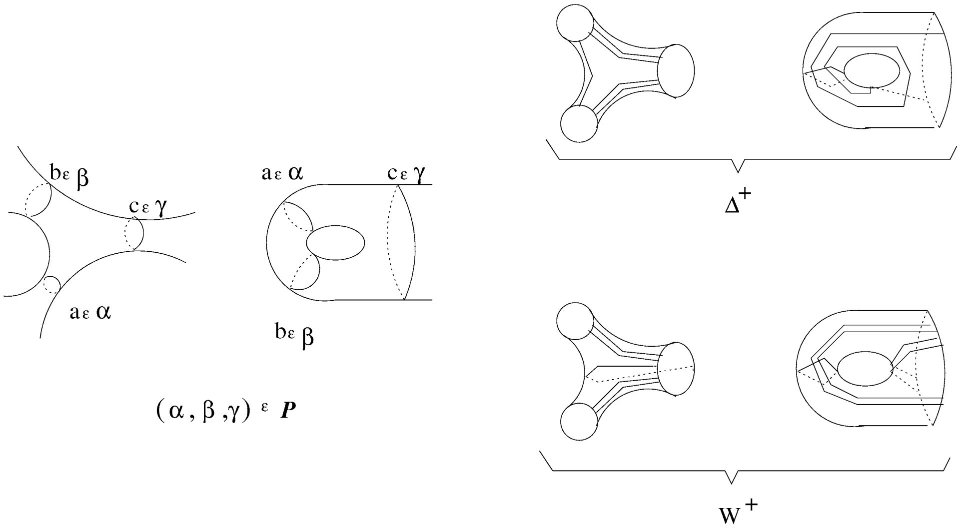


Fig. 1.

$I(\alpha_k,\beta))\in\Delta\}$; define $\Delta^+(\alpha)$ similarly. Define $W(\alpha)=\{\beta\in\mathcal{ML}(\Sigma)$: there is $(\alpha_i,\alpha_j,\alpha_k)\in\mathcal{P}$ so that $(I(\alpha_i,\beta),I(\alpha_j,\beta),I(\alpha_k,\beta))\in W\}$; define $W^+(\alpha)$ similarly.

The goal of this section is to show the following theorem.

Theorem 3.1 (Starr [18]). $\mathcal{CS}^+(\Sigma)=\bigcup_{\alpha\in\mathcal{FN}_t}(\Delta^+(\alpha)\cap\mathcal{CS}(\Sigma))$.

A similar result that $\mathcal{CS}(\Sigma)=\bigcup_{\alpha\in\mathcal{FN}_t}(\Delta(\alpha)\cap\mathcal{CS}(\Sigma))$ was proved in [9].

One may interpret the theorem as follows. Given two curve systems a,b with $|a\cap b|=I(a,b)$, we say that a (respectively $[a]$) contains a *wave* with respect to b (respectively $[b]$) if there exists an arc x in a and a component b' of b so that (1) $\partial x\subset b'$ and (2) $x\cap b=x\cap b'=\partial x$ and x approaches its end points from the same side of b'. For a 3-holed sphere P with $\partial P=a_1\sqcup a_2\sqcup a_3$ and a curve system b on P, then b contains a wave with respect to ∂P means that $(I(a_1,b),I(a_2,b),I(a_3,b))\in W^+$. The curve system b has components joining each of the three pairs of boundary components $\{a_i,a_j\}$ if and only if $(I(a_1,b),I(a_2,b),I(a_3,b))\in\Delta^+$. Thus Starr's theorem states that for each irreducible curve system β, there is a pants-decomposition α of the handlebody so that in each of the 3-holed sphere determined α there are arcs in β which join any pair of the boundary components.

Proof of Theorem 3.1 that $\bigcup_{\alpha\in\mathcal{FN}_t}(\Delta^+(\alpha)\cap\mathcal{CS})\subset\mathcal{CS}^+$. This follows from the parts (b) and (c) of the following lemma.

Lemma 3.2. *Suppose $\alpha\in\mathcal{FN}$. Then*

(a) *$\Delta(\alpha)\cap W^+(\alpha)=\Delta^+(\alpha)\cap W(\alpha)=\emptyset$ and $\Delta(\alpha)\cup W^+(\alpha)=\Delta^+(\alpha)\cup W(\alpha)=\mathcal{CS}(\Sigma)$. Furthermore, $\Delta(\alpha)$ and $W(\alpha)$ are closed subsets in $\mathcal{ML}(\Sigma)$.*

(b) *If* $\beta \in \Delta^+(\alpha)$ *and* $\gamma \in W^+(\alpha)$, *then* $I(\beta,\gamma) > 0$.

(c) *If* $\alpha \in \mathcal{F}N_t$ *and* $\beta \in \mathcal{C}S_t(\Sigma)$, *then* $\beta \in W(\alpha)$. *Furthermore, if* $I(\alpha,\beta) > 0$, *then* $\beta \in W^+(\alpha)$. *In particular, the limit set* L *is in* $W(\alpha)$ *for each* $\alpha \in \mathcal{F}N_t$.

Proof. Part (a) follows from the definition. Part (b) follows from the fact that if b, c are curve systems on a 3-holed sphere P with $\partial P = a_1 \sqcup a_2 \sqcup a_3$ so that c contains a wave with respect to ∂P and $(I(a_1,b), I(a_2,b), I(a_3,b)) \in \Delta^+$, then $I(b,c) > 0$. Part (c) follows from the outmost disc argument applied to the meridian discs bounded by α and β. To see the last statement, we have $\mathcal{S}_t \subset \bigcap_{\alpha\in\mathcal{F}N_t} W(\alpha)$. Thus $\mathbf{Q}_{>0} \times \mathcal{S}_t \subset \bigcap_{\alpha\in\mathcal{F}N_t} W(\alpha)$. But $\bigcap_{\alpha\in\mathcal{F}N_t} W(\alpha)$ is closed. Thus the limit set $L \subset \bigcap_{\alpha\in\mathcal{F}N_t} W(\alpha)$.

Proof of Theorem 3.1 that $\mathcal{C}S^+ \subset \bigcup_{\alpha\in\mathcal{F}N_t} \Delta^+(\alpha)$. Take an element $\beta \in \mathcal{C}S^+$ and take $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{F}N_t$. Define the complexity of α with respect to β to be $C(\alpha) = \sum_{(i,j,k)\in\mathcal{P}} 2^{I(\alpha_i,\beta)+I(\alpha_j,\beta)+I(\alpha_k,\beta)}$ where $(i,j,k) \in \mathcal{P}$ means $(\alpha_i,\alpha_j,\alpha_k) \in \mathcal{P}$. We use induction on the complexity to prove the theorem.

Suppose for some $(i,j,k) \in \mathcal{P}$, $(I(\alpha_i,\beta), I(\alpha_j,\beta), I(\alpha_k,\beta)$ is not in Δ^+, say $I(\alpha_i,\beta)$ is at least $I(\alpha_j,\beta) + I(\alpha_k,\beta)$. Let α_i be the component of α so that $I(\alpha_i,\beta)$ is the largest among all triples $(i,j,k) \in \mathcal{P}$ so that $I(\alpha_i,\beta) \geqslant I(\alpha_j,\beta) + I(\alpha_k,\beta)$. To simplify notations, let us assume that $(i,j,k) = (5,1,2)$. Since $\beta \in \mathcal{C}S^+$, $I(\alpha_r,\beta) > 0$ for all r. Thus $I(\alpha_5,\beta) > \max(I(\alpha_1,\beta), I(\alpha_2,\beta))$. In particular, $\alpha_5 \neq \alpha_1, \alpha_2$. Choose a representative $a = a_1 \sqcup \cdots \sqcup a_{3g-3} \in \alpha$ and $b \in \beta$ so that $|a \cap b| = I(a,b)$. Let P_i be the 3-holed sphere components of $\Sigma - \operatorname{int}(N(a))$. Then since $\alpha_5 \neq \alpha_1, \alpha_2$, $N(a_5)$ is adjacent to two distinct 3-holed spheres, say P_1 and P_2. Let $\Sigma_{0,4} = P_1 \cup N(a_5) \cup P_2$ be the 4-holed sphere and $\mathcal{S}' = \mathcal{S}'(\Sigma_{0,4})$ be the set of isotopy classes of essential non-boundary parallel simple loops on $\Sigma_{0,4}$. We claim that there exits an element $\alpha_5' \in \mathcal{S}'$ so that for the new Heegaard diagram $\alpha' = \alpha_1 \sqcup \cdots \sqcup \alpha_4 \sqcup \alpha_5' \sqcup \alpha_6 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{F}N_t$, $C(\alpha') < C(\alpha)$.

To see this, let the boundary components of $\Sigma_{0,4}$ correspond to $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Since $I(\alpha_5,\beta)$ is maximal, we have either $(I(\alpha_5,\beta), I(\alpha_3,\beta), I(\alpha_4,\beta)) \in \Delta^+$ or $I(\alpha_5,\beta) \geqslant I(\alpha_3,\beta) + I(\alpha_4,\beta)$. Thus the claim follows from the lemma below by taking $\alpha_5 = \gamma_1$ and $\alpha_5' = \gamma$.

Lemma 3.3. *Suppose* $\partial\Sigma_{0,4} = \alpha_1 \sqcup \alpha_2 \sqcup \alpha_3 \sqcup \alpha_4$, $\gamma_1 \in \mathcal{S}'(\Sigma_{0,4})$ *so that* $(\gamma_1,\alpha_1,\alpha_2) \in \mathcal{P}$, *and* $f(x) = I(x,\beta) : \mathcal{S}(\Sigma_{0,4}) \to \mathbf{Z}_{\geqslant 0}$ *is the geometric intersection function associated to* $\beta \in \mathcal{C}S(\Sigma_{0,4})$ *so that* $f(\alpha_i) > 0$. *If either* (a) $f(\gamma_1) \geqslant \max(f(\alpha_1) + f(\alpha_2), f(\alpha_3) + f(\alpha_4))$ *or* (b) $f(\gamma_1) \geqslant f(\alpha_1) + f(\alpha_2)$ *and* $(f(\gamma_1), f(\alpha_3), f(\alpha_4)) \in \Delta^+$, *then there exists* $\gamma \in \mathcal{S}'(\Sigma_{0,4})$ *so that*

$$\max_{(\gamma,\alpha_r,\alpha_s)\in\mathcal{P}} \bigl(f(\gamma) + f(\alpha_r) + f(\alpha_s)\bigr) < \max_{(\gamma_1,\alpha_r,\alpha_s)\in\mathcal{P}} \bigl(f(\gamma_1) + f(\alpha_r) + f(\alpha_s)\bigr).$$

Proof. The proof is based on a theorem proved in [8] which characterizes geometric intersection number functions. We shall recall the relevant result.

Three elements $\gamma_1, \gamma_2, \gamma_3 \in \mathcal{S}' = \mathcal{S}'(\Sigma_{0,4})$ are said to form an *ideal triangle* if $I(\gamma_i,\gamma_j) = 2$ for $i \neq j$. Given two elements $\gamma, \gamma' \in \mathcal{S}'$ with $I(\gamma,\gamma') = 2$, there are exactly two distinct ideal triangles of the form $(\gamma, \gamma', \gamma'')$. The following theorem was proved in [8]. As a convention, we use $(i, r, s) \in \mathcal{P}$ to denote $(\gamma_i, \alpha_r, \alpha_s) \in \mathcal{P}$, and unless indicated otherwise, the index i runs from 1 to 3, and indices r, s run from 1 to 4.

Theorem 3.4. *For surface $\Sigma_{0,4}$ with $\partial\Sigma_{0,4} = \alpha_1 \sqcup \alpha_2 \sqcup \alpha_3 \sqcup \alpha_4$, a function $f(\delta) : \mathcal{S}(\Sigma_{0,4}) \to$ $\mathbf{Z}$ is a function of the form $I(\beta, \delta)$ for some fixed $\beta \in \mathcal{CS}(\Sigma_{0,4})$ if and only if the following three conditions hold*:

(1) *If $(\gamma_1, \gamma_2, \gamma_3)$ forms an ideal triangle, then*

$$\sum_{i=1}^{4} f(\gamma_i) = \max_{(i,r,s)\in\mathcal{P}}\left(2f(\gamma_i), 2f(\alpha_r), \sum_{r=1}^{4} f(\alpha_r), f(\gamma_i) + f(\alpha_r) + f(\alpha_s)\right); \tag{1}$$

(2) *If $(\gamma_1, \gamma_2, \gamma_3)$ and $(\gamma_1, \gamma_2, \gamma_3')$ both form ideal triangles with $\gamma_3 \neq \gamma_3'$, then*

$$f(\gamma_3) + f(\gamma_3') = \max_{i=1,2;\,(i,r,s)\in\mathcal{P}}\left(2f(\gamma_i), 2f(\alpha_r), \sum_{r=1}^{4} f(\alpha_r), f(\gamma_i) + f(\alpha_r) + f(\alpha_s)\right); \tag{2}$$

(3) $$f(\gamma_i) + f(\alpha_r) + f(\alpha_s) \in 2\mathbf{Z} \quad \textit{when } (i, r, s) \in \mathcal{P}. \tag{3}$$

Applying the theorem to our situation, we have $f(\delta) = I(\beta, \delta)$ takes positive values on α_i and γ_1. Let $(\gamma_1, \gamma_2, \gamma_3)$ be the ideal triangle so that $f(\gamma_2, \gamma_3)$ is the smallest among all ideal triangles of the form $(\gamma_1, \gamma', \gamma'')$. For simplicity, let $x_i = f(\gamma_i)$, $a_r = f(\alpha_r)$ and $x_i' = \max_{(i,r,s)\in\mathcal{P}}(x_i + a_r + a_s)$. Then Eq. (1) in Theorem 3.4 becomes

$$\sum_{i=1}^{3} x_i = \max_{i,r}\left(2x_i, x_i', 2a_r, \sum_{r=1}^{4} a_r\right). \tag{4}$$

Due to the minimality of $x_2 + x_3$, we claim that Eq. (5) holds.

$$\sum_{i=1}^{3} x_i = \max_{i,r}\left(2x_1, x_i', 2a_r, \sum_{r=1}^{4} a_r\right). \tag{5}$$

Indeed, if otherwise, by Eq. (4), $\sum_{i=1}^{3} x_i = \max(2x_2, 2x_3)$, say, $\sum_{i=1}^{3} x_i = 2x_3$ and $2x_3 > \max_{i,r}(2x_1, x_i', 2a_r, \sum_{r=1}^{4} a_r)$. Since each $x_i > 0$, we obtain $x_3 = x_1 + x_2 > x_2$. Consider a new ideal triangle $(\gamma_1, \gamma_2, \gamma_3')$ where $\gamma_3' \neq \gamma_3$ and let $y_3 = f(\gamma_3')$. Then Eq. (2) in Theorem 3.4 shows $x_3 + y_3 = \max(2x_1, 2x_2, x_1', x_2', 2a_r, \sum_{r=1}^{4} a_r)$. By the assumption $\max(2x_1, 2x_2, x_1', x_2', 2a_r, \sum_{r=1}^{4} a_r) < 2x_3$. Thus $y_3 < x_3$ which contradicts the minimality of $(\gamma_1, \gamma_2, \gamma_3)$.

To finish the proof of the lemma, we claim that for γ to be one of γ_2 or γ_3, the conclusion of the lemma holds. If otherwise, we would have

$$\min(x_2', x_3') \geqslant x_1'. \tag{6}$$

We shall derive a contradiction that $\min_r(a_r) = 0$ from (6) by considering two separate cases: case 1. $x_1 \geqslant \max(a_1 + a_2, a_3 + a_4)$ and case 2. $x_1 \geqslant a_1 + a_2$ and $(x_1, a_3, a_4) \in \Delta^+$.

Case 1. $x_1 \geqslant \max(a_1 + a_2, a_3, +a_4)$. Then $2x_1 \geqslant \max_r(x_1', 2a_r, \sum_{r=1}^4 a_r)$. Thus Eq. (5) becomes $\sum_{i=1}^3 x_i = \max(2x_1, x_2', x_3')$.

Subcase 1.1. $\sum_{i=1}^3 x_i = 2x_1$, i.e., $x_2 + x_3 = x_1$. By Eq. (6), we may write

$$x_2 + a_r + a_s \geqslant x_1' \quad (2, r, s) \in \mathcal{P}. \tag{7}$$

$$x_3 + a_r + a_{s'} \geqslant x_1' \quad (3, r, s') \in \mathcal{P}. \tag{8}$$

We have $(1, s, s') \in \mathcal{P}$. The sum of (7) and (8) gives

$$x_2 + x_3 + 2a_r + a_s + a_{s'} \geqslant 2x_1'. \tag{9}$$

But $x_2 + x_3 = x_1$ and $2x_1' \geqslant 2x_1 + \sum_{r=1}^4 a_r$. Thus (9) implies that $2a_r + a_s + a_{s'} \geqslant x_1 + \sum_{t=1}^4 a_t$. This shows $a_r \geqslant x_1 + a_{r'}$ where $(1, r, r') \in \mathcal{P}$. Due to $x_1 \geqslant \max(a_1 + a_2, a_3 + a_4)$, we obtain that $\min_t(a_t) = 0$.

Subcase 1.2. $\sum_{i=1}^3 x_i = \max(x_2', x_3')$, say $\sum_{i=1}^3 x_i = x_2 + a_r + a_s$ with $(2, r, s) \in \mathcal{P}$. Then we have

$$x_1 + x_3 = a_r + a_s. \tag{10}$$

By (6), we have

$$x_3 + a_{r'} + a_s \geqslant x_1' \quad (3, r', s) \in \mathcal{P}. \tag{11}$$

Adding x_1 to both sides of (11) and using (10), we obtain

$$a_r + a_{r'} + 2a_s \geqslant x_1 + x_1' \quad (1, r, r') \in \mathcal{P}. \tag{12}$$

But $x_1 + x_1' \geqslant x_1 + x_1 + a_r + a_{r'}$. Thus by (12), we obtain $2a_s \geqslant 2x_1$. Due to $x_1 \geqslant \max(a_1 + a_2, a_3 + a_4)$, this implies $\min_t(a_t) = 0$.

Case 2. $x_1 \geqslant a_1 + a_2$ and $(x_1, a_3, a_4) \in \Delta^+$. Then $x_1 + a_3 + a_4 \geqslant \max(2x_1, x_1', 2a_r, \sum_{r=1}^4 a_r)$. Thus Eq. (5) becomes $\sum_{i=1}^3 x_i = \max(x_1 + a_3 + a_4, x_2', x_3')$.

Subcase 2.1. $\sum_{i=1}^3 x_i = \max(x_2', x_3')$, say $\sum_{i=1}^3 x_i = x_2 + a_r + a_s$ where $(2, r, s) \in \mathcal{P}$. Then

$$x_1 + x_3 = a_r + a_s. \tag{13}$$

By Eq. (6), we may assume that

$$x_3 + a_r + a_{s'} \geqslant x_1 + a_3 + a_4. \tag{14}$$

Adding x_1 to both sides of (14) and using (13), we obtain

$$2a_r + a_s + a_{s'} \geqslant 2x_1 + a_3 + a_4. \tag{15}$$

Note that $\{s, s'\} = \{1, 2\}$ or $\{3, 4\}$. If $\{s, s'\} = \{3, 4\}$, then Eq. (15) becomes $2a_r \geqslant 2x_1$ where $r \in \{1, 2\}$. Due to $x_1 \geqslant a_1 + a_2$, we have $\min(a_1, a_2) = 0$. If $\{s, s'\} = \{1, 2\}$, then $x_1 \geqslant a_s + a_{s'}$ and $r \in \{3, 4\}$. Thus (15) implies $2a_r + x_1 \geqslant 2x_1 + a_3 + a_4$, i.e., $2a_r \geqslant x_1 + a_3 + a_4$. This contradicts $(x_1, a_3, a_4) \in \Delta^+$.

Subcase 2.2. $\sum_{i=1}^3 x_i = x_1 + a_3 + a_4$. Then $x_2 + x_3 = a_3 + a_4$. By (6), we have

$$x_2 + a_r + a_s \geqslant x_1 + a_3 + a_4 \quad (2, r, s) \in \mathcal{P}. \tag{16}$$

$$x_3 + a_r + a_{s'} \geqslant x_1 + a_3 + a_4 \quad (3, r, s') \in \mathcal{P}. \tag{17}$$

Adding (16) to (17) gives $x_2 + x_3 + 2a_r + a_s + a_{s'} \geqslant 2x_1 + 2a_3 + 2a_4$. Using $x_2 + x_3 = a_3 + a_4$, we obtain

$$2a_r + a_s + a_{s'} \geqslant 2x_1 + a_3 + a_4 \quad (1, s, s') \in \mathcal{P}. \tag{18}$$

Inequality (18) is the same as (15). By the same argument as above, we obtain a contradiction again.

Remarks 3.1. One can give a new proof of the main theorem in [9] that $\mathcal{CS}(\Sigma) = \bigcup_{\alpha \in \mathcal{FN}_t} \Delta(\alpha) \cap \mathcal{CS}$ using the same argument as above. Indeed, the goal in this case is to eliminate the waves of β with respect to $\alpha \in \mathcal{FN}_t$. Suppose there are waves. Then as in the proof above, we choose α_i so that $I(\alpha_i, \beta) > I(\alpha_j, \beta) + I(\alpha_k, \beta)$ where $(i, j, k) \in \mathcal{P}$ and $I(\alpha_i, \beta)$ is the largest. Assume again that $(i, j, k) = (5, 1, 2)$. Then $(I(\alpha_5, \beta), I(\alpha_3, \beta), I(\alpha_4, \beta)) \in \Delta$ as in the proof above. To construct the move on α, we prove a lemma similar to Lemma 3.3 where the conditions (a) and (b) are replaced by (a') $f(\gamma_1) > f(\alpha_1) + f(\alpha_2)$ and $f(\gamma_1) \geqslant f(\alpha_3) + f(\alpha_4)$ or (b') $f(\gamma_1) > f(\alpha_1) + f(\alpha_2)$ and $(f(\gamma_1), f(\alpha_3), f(\alpha_4)) \in \Delta$. The proof of the lemma is the same as above.

3.2. A different proof of Starr's theorem using $\mathcal{CS}(\Sigma) = \bigcup_{\alpha \in \mathcal{FN}_t} \Delta(\alpha) \cap \mathcal{CS}$ and diagram chasing is given in Appendix A.

3.3. One may quantify the part (b) of Lemma 3.2 as follows. For $\varepsilon > 0$, let $\Delta_\varepsilon = \{(x_1, x_2, x_3) \in \mathbf{R}^3_{>0}\colon x_i + x_j \geqslant (1 + \varepsilon)x_k, i \neq j \neq k \neq i\}$. Evidently, $\Delta^+ = \bigcup_{\varepsilon > 0} \Delta_\varepsilon$. For $\alpha \in \mathcal{FN}$, we define the set $\Delta_\varepsilon(\alpha)$ in the same way as in the definition of $\Delta(\alpha)$. Then we have the following stronger version of part (b) of Lemma 3.2. Namely, if $\beta \in \Delta_\varepsilon$ and $\gamma \in W^+(\alpha)$, then $I(\beta, \gamma) \geqslant \frac{1}{3g-3}(\frac{\varepsilon}{2})^{2g-1} W(\gamma, \alpha) I(\alpha, \beta)$ where $W(\gamma, \alpha)$ is the number of waves of γ with respect to α and is explicitly given by $\sum_{(\alpha_i, \alpha_j, \alpha_k) \in \mathcal{P}} \frac{1}{2}(I(\alpha_i, \gamma) - I(\alpha_j, \gamma) - I(\alpha_k, \beta))_+$ with $x_+ = \frac{1}{2}(|x| + x)$.

4. Proofs of Theorem 1 and Corollary 2

Proof of part (a) of Theorem 1. To see the necessity, take $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{FN}_t \cap \Omega$. We claim that each α_i intersects all elements in $\mathcal{S}_t$. Indeed, if otherwise, say $I(\alpha_i, \beta) = 0$ for some $\beta \in \mathcal{S}_t$, then by Lemma 2.4, α_i is in the limit set L. But we also have $I(\alpha, \alpha_i) = 0$. This implies that α is not in Ω which contradicts the assumption.

To see the sufficiency part of part (a), take $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{FN}$ which is strongly irreducible, i.e., $I(\alpha_i, \beta) > 0$ for all $\beta \in \mathcal{S}_t$ and all i. If α is not in Ω, then there is $\beta \in L - 0$ so that $I(\alpha, \beta) = 0$. But $\alpha \in \mathcal{FN}$, thus $\beta = k_1\alpha_1 \sqcup \cdots \sqcup k_{3g-3}\alpha_{3g-3}$ for some numbers $k_i \in \mathbf{R}_{\geqslant 0}$ and one of them, say $k_1 > 0$. By Starr's theorem, there exists $\gamma \in \mathcal{FN}_t$ so that $\alpha_1 \in \Delta^+(\gamma)$. Thus $k_1\alpha_1 \in \Delta^+(\gamma)$. This implies $\beta \in \Delta^+(\gamma)$. But $\beta \in L \subset W(\gamma)$ which is the complement of $\Delta^+(\gamma)$. This is a contradiction.

Proof of part (b) of Theorem 1. We now show that $I(\alpha, x) \geqslant K|x|$ for all $[x] \in \mathcal{S}_t$ for some computable constant $K > 0$ and some fixed norm $|.|$.

Take $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{F}N$ which is strongly irreducible. For each index i, by Starr's theorem, we find (algorithmically) $\beta_i \in \mathcal{F}N_t$ so that $\alpha_i \in \Delta^+(\beta_i)$. Since $I(\alpha_i, \alpha_j) = 0$, we also have $\alpha \in \Delta^+(\beta_i)$. Evidently $\{\alpha, \beta_i\}$ is a surface filling system. Take the norm to be the one associated to $\{\alpha, \beta_1\}$, i.e., $|x| = I(\alpha, x) + I(\beta_1, x)$. We claim the following holds.

Claim. *There is a computable constant $K > 0$ depending on α_i and β_i so that for $[x] \in \bigcap_{i=1}^{3g-3} W^+(\beta_i)$,*

$$I(x, \alpha) \geqslant K\, I(x, \beta_1).$$

As a consequence, for all $[x] \in \bigcap_{i=1}^{3g-3} W^+(\beta_i)$ we have,

$$I(x, \alpha) \geqslant \frac{K}{K+1}|x|.$$

But only the other hand, $(\mathbf{Q}_{>0} \times \mathcal{S}_t) \cap \bigcap_{i=1}^{3g-3} W^+(\beta_i)$ is dense in the limit set L. Thus the above inequality still holds for all $[x] \in L$.

To prove the claim, take $a_i \in \alpha_i$, $b_j \in \beta_j$ and x so that their pairwise intersection numbers are minimal within the isotopy classes, and there are no triple intersection points. For each index i, due to $[x] \in W^+(\beta_i)$, the curve system x contains a wave with respect to b_i. But $[a_i] \in \Delta^+(\beta_i)$, thus $x \cap a_i \neq \emptyset$. Let x_i be an arc in x with end points on a so that $\mathrm{int}(x_i) \cap a_i \neq \emptyset$ and $|x_i \cap a| \leqslant 3$. Let $x' = \bigcup_{i=1}^{3g-3} x_i$. We note that $a \cup x'$ is a surface filling 1-dimensional cell complex (i.e., each component of $\Sigma - (a \cup x')$ is contractible). To see this, it suffices to show that for each 3-holed sphere component P of $\Sigma - \mathrm{int}(N(a))$ the components of $P - x'$ are contractible. This is equivalent to show that each component of ∂P intersects x'. But the last statement follows from the construction of x'.

For the surface filling 1-dimensional complex $a \cup x'$, we introduce a norm $\|\gamma\| = \min\{|a \cap y| + |x' \cap y|\colon\ y \in \gamma\}$. By Lemma 2.1, we have

$$I(x, \beta_1) \leqslant \|x\| \|\beta_1\| \leqslant I(x, \alpha)\left(I(\beta_1, \alpha) + \sum_{i=1}^{3g-3} |x_i \cap b_1|\right). \tag{19}$$

It remains to show that for each index i, $|x_i \cap b_1|$ is bounded by a computable constant. To this end, we prove a stronger statement that for each index i and for all $[x] \in \bigcap_{j=1}^{3g-3} W^+(\beta_j)$, there are only finitely many constructible isotopy classes of arc x_i under isotopies leaving a invariant and fixing each point of intersection $a \cap (\bigcup_{i=1}^{3g-3} b_i)$. Indeed, take the closure P of a component of $\Sigma - a$. It suffices to prove there are only finitely many constructible isotopy classes of arcs $x_i' = x_i \cap P$ under isotopies leaving each component of ∂P invariant and fixing $\partial P \cap (\bigcup_{i=1}^{3g-3} b_i)$. By construction, $|x_i \cap \partial P| \leqslant 3$. Thus there are at most eighteen isotopy classes of arc x_i' in P under isotopies leaving ∂P invariant. The only possibility to have infinitely many isotopy classes of x_i' under isotopies of P fixing each point in $\partial P \cap (\bigcup_{i=1}^{3g-3} b_i)$ is that x_i' spirals toward its end points on ∂P, say x_i' spirals toward a_j. Then each arc t on a_j with end points in b_j is isotopic to an arc t' in x_i' so that the isotopy preserves the curve system $b_j \cap P$. Since $[a_j] \in \Delta^+(\beta_j)$, i.e., any two components of b_j which are the boundary of a pants is joint by an arc $t \subset a_j$ so that the interior of t is disjoint from b_j, this implies that $[x_i] \in \Delta^+(\beta_j)$ which contradicts the assumption. In terms of the Dehn–Thurston coordinate for x_i' with respect to α, one can constructively estimate the twisting coordinate of x_i at α_j (see [10,16] for details on the Dehn–Thurston coordinate). Thus, we obtain a computable upper bound on the term $|x_i \cap b_1|$ in (19). This ends the constructive proof.

Proof of Corollary 2. Suppose $\alpha, \alpha' \in \mathcal{F}N$ so that α is strongly irreducible. Take $\beta' \in \mathcal{F}N_t$ and let $C = I(\alpha', \beta')$. By the constructive proof above, we have $I(\alpha, x) \geqslant K|x|$ for all $x \in \mathcal{C}S_t$ for some computable constant K and a fixed norm $|.|$. Find all elements $\beta_1, \ldots, \beta_r \in \mathcal{F}N_t$ so that $|\beta_i| \leqslant C/K$ (this can be done algorithmically). Now given two pairs of Heegaard diagrams (a, b) and (a', b'), there is an algorithm to check if they are related by an element in the mapping class group of the surface (one may use Dehn–Thurston coordinate to do this). Check if (α, β_i) is related to (α', β') by an element in the mapping class group. If they are related by an element $\gamma \in \mathrm{Mod}(\Sigma)$, then $\gamma \in \Gamma$ since $\beta_i, \beta' \in \mathcal{F}N_t$. Thus α and α' are related by an element in Γ. If none of the pair (α, β_i) is related to (α', β') by an element in $\mathrm{Mod}(\Sigma)$, then α and α' are not related by any element in Γ. Indeed, if there were $\gamma \in \Gamma$ so that $\gamma(\alpha) = \alpha'$. Then $\gamma(\beta') \in \mathcal{F}N_t$. Furthermore, $|\gamma(\beta')| \leqslant \frac{1}{K} I(\alpha, \gamma(\beta')) = I(\alpha', \beta') = C/K$. This shows $\gamma(\beta')$ must be one of the elements β_i by construction.

Remark 4.1. If a pair $(\alpha, \beta) \in \mathcal{F}N \times \mathcal{F}N_t$ satisfies Casson–Gordon's rectangle condition (see [4,7]), then the inequality in (19) becomes $I(\alpha, x) \geqslant \frac{I(\beta,x)}{I(\alpha,\beta)}$ for all $[x] \in W^+(\beta)$.

Corollary 4.1. *If $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_{3g-3} \in \mathcal{F}N$ is strongly irreducible, then any non-trivial Dehn twist $h = D_{\alpha_1}^{k_1} \cdots D_{\alpha_{3g-3}}^{k_{3g-3}}$ with one of $k_i \neq 0$ is not in the handlebody group Γ.*

Indeed, if $h \in \Gamma$, there would be infinitely many distinct elements of the form $h^n(\beta)$ in $\mathcal{F}N_t$ whose intersection number with α is bounded. This contracts the fact that $\alpha \in \Omega$.

Remark 4.2. The related result to Corollary 4.1 is Corollary 2 in [9] which was mistakenly stated. The correct statement is that for $\alpha = \alpha_1 \sqcup \cdots \sqcup \alpha_k \in \mathcal{C}S$, then $\alpha \in \mathcal{C}S_t$ if and only if $D_{\alpha_1} \cdots D_{\alpha_k} \in \Gamma$ (i.e., the condition $a_i > 0$ is needed in the Corollary 2).

5. Some questions

We begin with some terminologies. Given two Heegaard diagrams α and β on the boundary Σ of a handlebody H, we say they determine the same handlebody structure if in the handlebody $\Sigma(\alpha)$ obtained by attaching 2-handles along α to Σ and then 3-handles, each component of β is null homotopic, i.e., $\Sigma(\alpha) = \Sigma(\beta)$. This is equivalent to the existence of a homeomorphism between $\Sigma(\alpha)$ and $\Sigma(\beta)$ which is the identity map on the boundary. For a Heegaard diagram α in a handlebody, let $\mathrm{sp}(\alpha)$ be the set of all Heegaard diagrams which determine the same handlebody structure as α. We call the pair $(\Sigma(\alpha), H)$ a *Heegaard splitting*. Casson and Gordon's strongly irreducible condition on Heegaard splitting $(\Sigma(\alpha), H)$ says that each meridian in $\Sigma(\alpha)$ intersects each meridian in H. Thus by Theorem 1, a Heegaard splitting $(\Sigma(\alpha), H)$ is strongly irreducible if and only if $\mathrm{sp}(\alpha)$ is a subset of the Masur domain Ω. Call a Heegaard splitting $(\Sigma(\alpha), H)$ *hyperbolic* if the closure of $\mathbf{Q}_{>0} \times \mathrm{sp}(\alpha)$ in the measured lamination space $\mathcal{ML}(\Sigma)$ is in Ω. Equivalently, $(\Sigma(\alpha), H)$ is hyperbolic if and only if there is a positive constant K so that $I(x, y) \geqslant K|x||y|$ for all meridian disc x in H and meridian disc y in $\Sigma(\alpha)$. For a hyperbolic Heegaard splitting $(\Sigma(\alpha), H)$ with a computable constant K, the homeomorphism problem for the manifold $M = H \cup_{\mathrm{id}} \Sigma(\alpha)$ is always solvable (using the work of Rubinstein [17] on the algorithmic construction of all strongly irreducible Heegaard splittings of a given genus and the same argument used in the proof of Corollary 2 in Section 4).

Lemma 5.1. *Suppose $(\Sigma(\alpha), H)$ is a hyperbolic Heegaard splitting. Let x, y be two meridians in different handlebodies in the Heegaard splitting. Then $([x], [y])$ forms a surface filling pair. In particular, this implies that the closed 3-manifold $M^3 = \sigma(\alpha) \cup_{\mathrm{id}} H$ is irreducible and atoroidal.*

Proof. Suppose otherwise that there is a simple loop c on the surface Σ which is disjoint from both $x \subset H$ and $y \subset \Sigma(\alpha)$. Then by Lemma 2.4, c is in the limit set L. Thus $I([y], [c]) = 0$ for some $[c] \in L$. This contradicts the assumption that $[y] \in \Omega$.

Evidently the 3-manifold M^3 is irreducible since the Heegaard splitting is strongly irreducible. To see that it is atoroidal, we use an argument by Hempel. Suppose otherwise that M^3 contains an incompressible torus T. Then due to the strongly irreducibility of the Heegaard splitting, we may find an incompressible torus T' so that T' in each of the handlebody consists of annuli which are incompressible in the handlebody. Let c be a component of the curve system $T' \cap \Sigma$ in the surface. Then c is disjoint from some merdians x and y from each handlebody. Thus we produce a meridian pair (x, y) which is not surface filling.

A related notion on Heegaard diagrams was introduced by Hempel [4] as follows. Call a Heegaard splitting $\mathrm{sp}(\alpha)$ a *distance at least three* splitting if for each meridian x in H and meridian y in the handlebody $\Sigma(\alpha)$, the pair (x, y) is surface filling. Evidently, by the above lemma, if a Heegaard splitting is hyperbolic than it is of distance at least three. Using the work [14], Hempel showed that a 3-manifold with Heegaard splitting of distance at least three contains no incompressible tori and is not a Seifert fibered space. Thus according to Thurston's geometrization conjecture, the manifold M should be hyperbolic. Following this line one may ask if each hyperbolic 3-manifold supports a hyperbolic Heegaard splittings. Note that Hempel [4] has constructed many strongly irreducible Heegaard splittings of hyperbolic manifolds which are of distance at most two. A less ambitious question is the following.

Question. Is the fundamental group of a closed 3-manifold with a hyperbolic Heegaard splitting infinite?

Finally one may ask if Hempel's notion of distance at least three is the same as the hyperbolicity.

Acknowledgements

Part of the work was done while I was visiting Rice University. I would like to thank Richard Stong for inviting me to visit Rice. I thank John Hempel and Stong for discussions and I thank Ying-Qing Wu for informing me Starr's work. The work is supported in part by the NSF.

Appendix A. A diagram chasing proof of Starr's theorem

We give a different proof of the fact that $\mathcal{CS}^+ \subset \bigcup_{\alpha\in\mathcal{F}N_t} \Delta^+(\alpha)$ using the main theorem of [9] that $\mathcal{CS} \subset \bigcup_{\alpha\in\mathcal{F}N_t} \Delta(\alpha)$ and the diagram chasing argument.

Let $\beta \in \mathcal{CS}^+$. Take $\alpha \in \mathcal{F}N_t$ so that $\beta \in \Delta(\alpha)$. We now follow the same reduction as in the first two paragraphs in the proof of Theorem 3.1 to construct $\alpha_1, \ldots, \alpha_5$, P_1, P_2, and $\Sigma_{0,4}$.

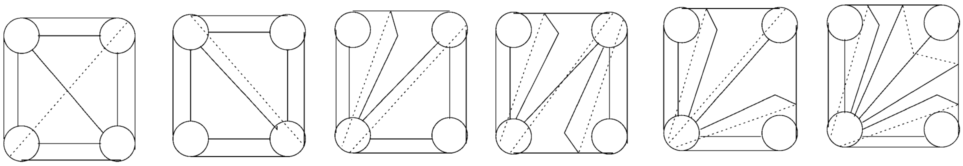
Recall that an *ideal triangulation* of a compact surface with boundary is a maximal collection of disjoint pairwise non-parallel essential arcs on the surface. It can be shown easily the following.

Lemma A1. *Any ideal triangulation of the* 4*-holed sphere is homeomorphic to one of the following six ideal triangulations.*

Now for the curve system $b' = b \cap \Sigma_{0,4}$, there is an ideal triangulation $T = t_1 \sqcup \cdots \sqcup t_6$ of $\Sigma_{0,4}$ so that b' is isotopic to $k_1t_1 \sqcup \cdots \sqcup k_6t_6$ for some $k_i \in \mathbf{Z}_{\geqslant 0}$. Let T' be the subset of T consisting of those $t_i's$ so that $k_i > 0$. Since β is irreducible, T' contains at least three components. By the diagram chasing argument, one shows the following lemma.

Lemma A2. *There is a homeomorphism h of the* 4*-holed sphere $\Sigma_{0,4}$ preserving two* 3*-holed spheres P_1 and P_2 so that $h(T')$ is one of the following seven curve systems.*

Now for each of the seven cases, choose α_5' as indicated. One see from Fig. 3 that $C(\alpha') < C(\alpha)$.



The set of all ideal triangulations of the 4-holed sphere

Fig. 2.

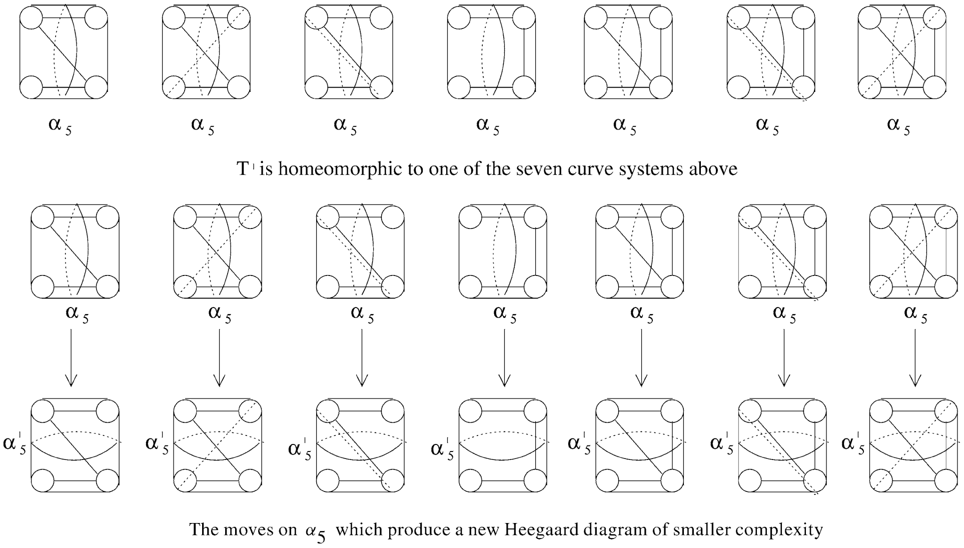


Fig. 3.

Remark. The complexity for the induction argument was suggested by the equations in Theorem 3.4.

References

[1] F. Bonahon, Bouts des varitis hyperboliques de dimension 3, Ann. of Math. (2) 124 (1) (1986) 71–158.
[2] A.J. Casson, C.McA. Gordon, Reducing Heegaard splittings, Topology Appl. 27 (3) (1987) 275–283.
[3] A. Fathi, F. Laudenbach, V. Poenaru, Travaux de Thurston sur les surfaces, in: Astérisque 66–67, Société Mathématique de France, 1979.
[4] J. Hempel, 3-manifolds as viewed from the curve complex, Preprint.
[5] K. Johannson, Topology and Combinatorics of 3-Manifolds, in: Lecture Notes in Math., Vol. 1599, Springer-Verlag, Berlin, 1995.
[6] S. Kerckhoff, The measure of the limit set of the handlebody group, Topology 29 (1) (1990) 27–40.
[7] T. Kobayashi, Casson–Gordon's rectangle condition of Heegaard diagrams and incompressible tori in 3-manifolds, Osaka J. Math. 25 (3) (1988) 553–573.
[8] F. Luo, Simple loops on surfaces and their intersection numbers, Math. Res. Lett. 5 (1998) 47–56.
[9] F. Luo, On Heegaard diagrams, Math. Res. Lett. 4 (1997) 365–373.
[10] F. Luo, R. Stong, Dehn–Thurston coordinates for curves on surfaces, Preprint, 2002.
[11] H. Masur, Measured foliations and handlebodies, Ergodic Theory Dynamical Systems 6 (1) (1986) 99–116.
[12] H. Masur, Private communication.
[13] J. McCarthy, A. Papadopoulos, Dynamics on Thurston's sphere of projective measured foliations, Comment. Math. Helv. 64 (1) (1989) 133–166.
[14] Y. Moriah, J. Schultens, Irreducible Heegaard splittings of Seifert fibered spaces are either vertical or horizontal, Topology 37 (5) (1998) 1089–1112.
[15] J.-P. Otal, Courants géodésiqes de surfaces, Thèse de Doctorat d'Etat, Université de Paris-Sud, Centre d'Orsay, 1989.
[16] R.C. Penner, J.L. Harer, Combinatorics of Train Tracks, in: Ann. of Math. Stud., Vol. 125, Princeton University Press, Princeton, NJ, 1992.

[17] J.H. Rubinstein, Polyhedral minimal surfaces, Heegaard splittings and decision problems for 3-dimensional manifolds, in: Geometric Topology, Athens, GA, 1993, in: AMS/IP Stud. Adv. Math., Vol. 2.1, American Mathematical Society, Providence, RI, 1997, pp. 1–20.
[18] E. Starr, Curves in handlebodies, Thesis, UC Berkeley, 1992.
[19] W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces, Bull. Amer. Math. Soc. (N.S.) 19 (2) (1988) 417–431.
[20] Y.-Q. Wu, Incompressible surfaces and Dehn surgery on 1-bridge knots in handlebodies, Math. Proc. Cambridge Philos. Soc. 120 (1996) 687–696.