Geometry & Topology Monographs 12 (2007) 157–190 **157**

On the Heegaard splittings of amalgamated 3–manifolds

TAO LI

We give a combinatorial proof of a theorem first proved by Souto which says the following. Let M_1 and M_2 be simple 3–manifolds with connected boundary of genus $g > 0$. If M_1 and M_2 are glued via a complicated map, then every minimal Heegaard splitting of the resulting closed 3–manifold is an amalgamation. This proof also provides an algorithm to find a bound on the complexity of the gluing map.

57N10; 57M50

1 Introduction

The study of Heegaard splitting has been dramatically changed since Casson and Gordon introduced the notion of strongly irreducible Heegaard splitting [4]. Casson and Gordon proved in [4] that if a Heegaard splitting is irreducible but weakly reducible, then one can perform some compressions on both sides of the Heegaard surface and obtain an incompressible surface.

Conversely, let F be a connected separating incompressible surface in a closed 3–manifold M' and M_1 and M_2 the two manifolds obtained by cutting open M' along F. Then one can construct a weakly reducible Heegaard splitting by amalgamating two splittings of M_1 and M_2 along F, see Scharlemann [13] for more detailed discussion.

In [10] Lackenby showed that if M_1 and M_2 are simple and the gluing map is a high power of a pseudo-Ansov homeomorphism of F (F is connected), then the minimal genus Heegaard splitting of M' is obtained from splittings of M_1 and M_2 by amalgamation. This implies that the genus of M' is $g(M_1)+g(M_2)-g(F)$.

As pointed out in [10], it is generally believed that the same is true if the gluing map is of high distance in the curve complex, see Theorem 1.1. Note that a high power of a pseudo-Ansov map has high distance in the curve complex. Souto [18] proved this first using the same principles as in [10] by analyzing the geometry near the incompressible surface. In this paper, we give a combinatorial proof of this result and this proof also provides an algorithm to find the bound on the distance for the gluing map.

Published: 3 December 2007 DOI: 10.2140/gtm.2007.12.157

Theorem 1.1 *Let M_1 and M_2 be orientable simple 3–manifolds with connected boundary and suppose $\partial M_1 \cong \partial M_2 \cong F$. Then there is a finite set of curves $\mathcal{C}_i \subset \partial M_i$ and a number N such that, if a homeomorphism $\phi\colon \partial M_1 \to \partial M_2$ satisfies $d_{\mathcal{C}(F)}(\phi(\mathcal{C}_1), \mathcal{C}_2) > N$, where $d_{\mathcal{C}(F)}$ is the distance in the curve complex $\mathcal{C}(F)$ of F, then*

(1) *every minimal genus Heegaard splitting of $M_1 \cup_\phi M_2$ is an amalgamation,*

(2) *the Heegaard genus satisfies $g(M_1 \cup_\phi M_2) = g(M_1) + g(M_2) - g(F)$.*

Moreover, there is an algorithm to find $\mathcal{C}_i$ and N.

In this paper, we will study 0–efficient triangulations for 3–manifolds with connected boundary. A 0–efficient triangulation for a 3–manifold with connected boundary is a triangulation with only one vertex (on the boundary), the only normal disk is vertex linking, and there is no normal S^2. By an in-depth analysis of normal annuli in such triangulations, we prove the following theorem which can be viewed as a generalization of Hatcher's theorem [6] and a theorem of Jaco and Sedgwick [9] to manifolds with higher genus boundary.

Theorem 1.2 *Let M be a simple 3–manifold with connected boundary and a 0–efficient triangulation. Let S_k be the set of normal and almost normal surfaces satisfying the following two conditions*

(1) *the boundary of each surface in S_k consists of essential curves in ∂M*

(2) *the Euler characteristic of each surface in S_k is at least $-k$.*

Let C_k be the set of boundary curves of surfaces in S_k. Then C_k has bounded diameter in the curve complex of ∂M. Moreover, there is an algorithm to find the diameter.

Since every incompressible and ∂–incompressible surface in M is isotopic to a normal surface in any triangulation, an immediate corollary of Theorem 1.2 is that the set of boundary curves of essential surfaces with bounded Euler characteristic has bounded diameter in the curve complex of ∂M.

It seems that a version of Theorem 1.1 is true without the assumption that M_i is atoroidal or ∂M_i is incompressible.

Conjecture 1.3 *Let M_i $(i = 1, 2)$ be an irreducible 3–manifold with connected boundary $\partial M_1 \cong \partial M_2 \cong F$. Let $\mathcal{D}_i$ be the set of essential curves in F that bound disks in M_i. Then there is an essential curve $\mathcal{C}_i$ $(i = 1, 2)$ in ∂M_i such that if the distance between $\mathcal{D}_2 \cup \mathcal{C}_2$ and $\phi(\mathcal{D}_1 \cup \mathcal{C}_1)$ in the curve complex $\mathcal{C}(F)$ is sufficiently large, then any minimal-genus Heegaard splitting of $M_1 \cup_\phi M_2$ can be constructed from an amalgamation.*

This conjecture can be viewed as a generalization of Theorem 1.1 and a theorem of Scharlemann and Tomova [16]. Note that in the case that both M_1 and M_2 are handlebodies, $\mathcal{C}_1$ and $\mathcal{C}_2$ can be chosen to be empty and the theorem of Scharlemann and Tomova [16] can be formulated as: if the distance between $\mathcal{D}_2$ and $\phi(\mathcal{D}_1)$ (ie, the Hempel distance) is large, then the genus of any other Heegaard splitting must be large unless it is a stabilized copy of F.

The proof in this paper is different from the original proof presented in the Haifa workshop in 2005, though both proofs use Jaco and Rubinstein's theory on 0–efficient triangulation [8]. This proof is a byproduct of an effort of finding an algorithmic proof of the generalized Waldhausen conjecture [12] and it gives a much clearer algorithm than the original proof.

I would like to thank Saul Schleimer and Dave Bachman for useful conversation about their work [3] with Eric Sedgwick. I also thank the referee for many helpful comments and corrections. The research was partially supported by an NSF grant.

Throughout this paper, we will denote the interior of X by $\mathrm{int}(X)$, the closure (under path metric) of X by $\overline{X}$, and the number of components of X by $|X|$.

2 Strongly irreducible Heegaard surfaces

Let M_1 and M_2 be orientable simple 3–manifolds with connected boundary and suppose $\partial M_1 \cong \partial M_2 \cong F$. Let $\phi\colon \partial M_1 \to \partial M_2$ be a homeomorphism and $M' = M_1 \cup_\phi M_2$ the closed manifold by gluing M_1 and M_2 via ϕ. Thus there is an embedded surface F in M' such that $\overline{M' - F}$ is the disjoint union of M_1 and M_2. Since each M_i is irreducible and ∂M_i is incompressible in M_i, M' is irreducible and F is incompressible in M'. We may regard M_i as a submanifold of M'.

From any Heegaard splittings of M_1 and M_2, one can naturally construct a Heegaard splitting of M', called amalgamation. This operation was defined by Schultens [17]. We give a brief description below, see [13; 17] for details. Any Heegaard surface S_i of M_i ($i = 1, 2$) decomposes M_i into a handlebody and a compression body. Each compression body can be obtained by attaching 1–handles to $F \times I$, a product neighborhood of F. One can extend the 1–handles of the compression body of M_1 vertically through the product region $F \times I$ and attach these extended 1–handles to the handlebody in the splitting of M_2. This operation produces a handlebody of genus $g(S_1)+g(S_2)-g(F)$. It is easy to check that its complement is also a handlebody and we get a Heegaard splitting of M'. This Heegaard splitting is called an *amalgamation* of S_1 and S_2. Clearly, the resulting Heegaard splitting from amalgamation is weakly reducible, see [4; 13] for definitions and basic properties of weakly reducible and strongly irreducible Heegaard splittings.

Given a weakly reducible but irreducible Heegaard surface S, Casson and Gordon [4] showed that one can compress S on both sides along a maximal collection of disjoint compressing disks and obtain an incompressible surface. Scharlemann and Thompson generalized this construction and gave a construction of *untelescoping* of a weakly reducible Heegaard splitting, see [13; 15] for details. The following lemma follows trivially from the untelescoping construction. We say two surfaces intersect nontrivially if they cannot be made disjoint by an isotopy.

Lemma 2.1 *Let S be an irreducible Heegaard surface of $M_1 \cup_\phi M_2$. Then either*

(1) *S is an amalgamation of two splittings of M_1 and M_2, or*

(2) *there is a submanifold M_F of $M_1 \cup_\phi M_2$ (M_F may be $M_1 \cup_\phi M_2$) such that $F \subset M_F$ and ∂M_F, if non-empty, is incompressible, and there is a strongly irreducible Heegaard surface S' of M_F such that the genus of S' is at most $g(S)$ and S' nontrivially intersects F, or*

(3) *there is an incompressible surface S' with genus less than $g(S)$ such that S' nontrivially intersects F.* $\square$

3 Intersection with F

Suppose S is a minimal genus Heegaard surface of $M' = M_1 \cup_\phi M_2$. So the genus $g(S)$ is at most $g(M_1) + g(M_2) - g(F)$. To prove Theorem 1.1, ie, S is an amalgamation, we need to rule out case (2) and case (3) in Lemma 2.1.

We first consider the case (3) in Lemma 2.1. The following lemma is easy to prove.

Lemma 3.1 *Suppose there is an incompressible surface S' that nontrivially intersects F. Then there is an incompressible and ∂–incompressible surface S_i in M_i such that $d_{\mathcal{C}(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S')$.*

Proof Since both S' and F are incompressible, we may assume that $S' \cap F$ consists of essential curves. Let $S_i' = S' \cap M_i$ ($i = 1, 2$). Hence S_i' is incompressible in M_i and $d_{\mathcal{C}(F)}(\phi(\partial S_1'), \partial S_2') = 0$.

If S_1' is ∂–compressible, then we perform a ∂–compression on S_1' and get a new incompressible surface S_1''. Clearly $d_{\mathcal{C}(F)}(\partial S_1', \partial S_1'') \leq 1$. Note that S_1' and S_1'' are not

∂–parallel in M_1, because otherwise S' can be isotoped to be disjoint from F, contradicting our hypothesis. Thus, after fewer than $-\chi(S_1')$ ∂–compressions, we obtain an incompressible and ∂–incompressible surface S_1 in M_1. So $d_{\mathcal{C}(F)}(\partial S_1', \partial S_1) < -\chi(S_1')$. Similarly, we can find an incompressible and ∂–incompressible surface S_2 in M_2 with $d_{\mathcal{C}(F)}(\partial S_2', \partial S_2) < -\chi(S_2')$. Therefore, $d_{\mathcal{C}(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S_1') - \chi(S_2') = -\chi(S')$. □

Next we consider the case (2) in Lemma 2.1. Bachman, Schleimer and Sedgwick [3] proved a version of Lemma 3.1 for strongly irreducible Heegaard surfaces, see Lemma 3.3 below and [3, Lemma 3.3].

We first give some definitions using the terminology in [3].

Definition 3.2 A properly embedded surface is *essential* if it is incompressible and ∂–incompressible. A properly embedded, separating surface is *strongly irreducible* if there are compressing disks for it on both sides, and each compressing disk on one side meets each compressing disk on the other side. It is *∂–strongly irreducible* if

(1) every compressing and ∂–compressing disk on one side meets every compressing and ∂–compressing disk on the other side, and

(2) there is at least one compressing or ∂–compressing disk on each side.

Lemma 3.3 (Bachman–Schleimer–Sedgwick [3]) *Let M_F be a compact, irreducible, orientable 3–manifold with ∂M_F incompressible, if non-empty. Suppose $M_F = V \cup_S W$, where S is a strongly irreducible Heegaard surface. Suppose further that M_F contains an incompressible, orientable, closed, non-boundary parallel surface F. Then either*

- *S may be isotoped to be transverse to F, with every component of $S - N(F)$ incompressible in the respective submanifold of $M_F - N(F)$, where $N(F)$ is a small neighborhood of F in M_F,*
- *S may be isotoped to be transverse to F, with every component of $S - N(F)$ incompressible in the respective submanifold of $M_F - N(F)$ except for exactly one strongly irreducible component, or*
- *S may be isotoped to be almost transverse to F (ie, S is transverse to F except for one saddle point), with every component of $S - N(F)$ incompressible in the respective submanifold of $M_F - N(F)$.*

Corollary 3.4 *Let M_F be a 3–manifold with incompressible boundary and let F be a separating incompressible and non-boundary parallel surface in M_F. Let M_1' and M_2'*

be the 3–manifolds obtained by cutting M_F open along F and $\phi\colon \partial M'_1 \to \partial M'_2$ the gluing map so that $M_F = M'_1 \cup_\phi M'_2$. Suppose S is a strongly irreducible Heegaard surface of M_F. Then there are surfaces S_i in M'_i such that $d_{\mathcal{C}(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S)$ and each S_i is either essential or strongly irreducible and ∂–strongly irreducible in M'_i.

Proof Note that if $S \cap M'_i$ consists of ∂–parallel surfaces in M'_i ($i = 1, 2$), then we can perform an isotopy on S so that $S \cap F = \varnothing$ after the isotopy, contradicting our hypotheses. So at least one component of $S \cap M'_i$ is not ∂–parallel. Moreover, if a component of $S \cap M'_i$ is a disk, since F is incompressible, the disk must be ∂–parallel in M'_i and we can perform an isotopy on S and remove a trivial-curve component of $S \cap F$. Thus, after some isotopies on S, we may assume that no component of $S \cap M'_i$ is a disk. This implies that every curve of $S \cap F$ is essential in S and hence every component of $S - F$ or $S \cap M'_i$ is an essential subsurface of S.

By Lemma 3.3, we can find a component of $S \cap M'_i$, denoted by S'_i ($i = 1, 2$), such that (1) each S'_i is an essential subsurface of S and not ∂–parallel in M'_i, (2) each S'_i is either incompressible or strongly irreducible in M'_i, and (3) $d_{\mathcal{C}(F)}(\phi(\partial S'_1), \partial S'_2) \le 1$. If the third case in Lemma 3.3 occurs, then both S'_1 and S'_2 are incompressible. In the other 2 cases, it follows from Lemma 3.3 that every curve in $S \cap F$ must be essential in F. To see this, suppose $\gamma \subset S \cap F$ is an innermost trivial curve in F, then the disk bounded by γ in F is a compressing disk for S. This means that both $S \cap M'_1$ and $S \cap M'_2$ have compressible components, a contradiction to Lemma 3.3. Therefore, $\partial S'_1$ and $\partial S'_2$ are essential in F in any case.

Suppose S'_i is incompressible but ∂–compressible in M'_i. As in the proof of Lemma 3.1, after fewer than $-\chi(S'_i)$ ∂–compressions, we obtain an essential surface S_i with $d_{\mathcal{C}(F)}(\partial S'_i, \partial S_i) < -\chi(S'_i)$.

Suppose S'_i is strongly irreducible but not ∂–strongly irreducible in M'_i. We say a ∂–compressing disk D of S'_i is *disk-busting* if every compressing disk on the other side of S'_i intersects ∂D.

We first consider the case that S'_i contains a ∂–compressing disk D that is not disk-busting. So there is a compressing disk D' on the other side of S'_i with $D \cap D' = \varnothing$. Now we perform a ∂–compression along D and get a new surface, which we denote by S''_i. Since $D' \cap D = \varnothing$, after the isotopy, D' remains a compressing disk of S''_i. Note that since S'_i is strongly irreducible, by definition, there is a compressing disk of S'_i on the same side as D, in fact, a simple cutting-and-pasting argument can show that there is a compressing disk on the same side as D and disjoint from D. This means that S''_i is still strongly irreducible.

After a finite number of such ∂–compressions, we may assume every ∂–compressing disk of S_i'' is disk-busting. If S_i'' is not ∂–strongly irreducible, then there must be a pair of disjoint ∂–compressing disks D and D' on different sides of S_i''. Since $D \cap D' = \varnothing$, we can perform ∂–compressions along D and D' simultaneously. Since both D and D' are disk-busting, the resulting surface after ∂–compressions along D and D' is incompressible.

Therefore, after fewer than $-\chi(S_i')$ ∂–compressions, we obtain a surface S_i in M_i' $(i = 1, 2)$ such that each S_i is either essential or strongly irreducible and ∂–strongly irreducible in M_i' and $d_{\mathcal{C}(F)}(\partial S_i', \partial S_i) < -\chi(S_i')$. Similar to the proof of Lemma 3.1, we have $d_{\mathcal{C}(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S)$. □

The next corollary follows trivially from Lemma 2.1, Lemma 3.1 and Corollary 3.4.

Corollary 3.5 *Let S be an irreducible Heegaard surface of $M_1 \cup_\phi M_2$. Suppose S is not a amalgamation of two splittings of M_1 and M_2. Then there is a properly embedded surface with boundary S_i in M_i such that $d_{\mathcal{C}(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S)$ and each S_i is either essential or strongly irreducible and ∂–strongly irreducible in M_i.* □

Remark 3.6 Let S_1 and S_2 be components of $S \cap M_i$ as in Corollary 3.5. It follows from the construction above and Definition 3.2 that the boundary of S_1 and S_2 consists of essential curves. We fix a 0–efficient triangulation (described below) for each M_i. If S_i is essential, then S_i is isotopic to a normal surface. If S_i is ∂–strongly irreducible, then by a theorem of Bachman [2], S_i is isotopic to a normal or an almost normal surface with boundary. The referee pointed out a controversy in a theorem in [2]. In our proof, we will use Bachman's theorem, but give a workaround in the appendix avoiding the controversial part of Bachman's argument. If S_i is ∂–strongly irreducible, the general case follows from the appendix is that, after isotopy or possible ∂–compressions, (1) S_i is normal or almost normal and ∂S_i consists of normal curves in ∂M_i, and (2) at most one component of ∂S_i is a trivial curve and at least one component of ∂S_i is an essential curve. Note that a trivial normal curve in a one-vertex triangulation of ∂M_i is vertex-linking, see Proposition 4.3 below. For simplicity, we will assume that S_i is almost normal and ∂S_i consists of essential normal curves and the proof for the general case is basically the same.

4 The 0–efficient triangulation

Let S be a minimal genus Heegaard surface. By Corollary 3.5, if a Heegaard surface S is not obtained from amalgamation, then there is a surface S_i properly embedded in M_i such that S_i is either essential or ∂–strongly irreducible in M_i and $d_{C(F)}(\phi(\partial S_1), \partial S_2) < -\chi(S) \leq 2(g(M_1) + g(M_2) - g(F)) - 2$. Fixing a 0–efficient triangulation (described below) of M_i, as in Remark 3.6, we may assume S_i is a normal or an almost normal surface with respect to the triangulation and ∂S_i consists of essential normal curves in ∂M_i. Our goal is to prove that the boundary curves of such (almost) normal surfaces have bounded diameter in the curve complex of $F = \partial M_i$, see Theorem 1.2.

The 0–efficient triangulation, introduced by Jaco and Rubinstein [8], is a very convienient tool, see for example [12]. In this paper we are mainly interested in 0–efficient triangulation for manifolds with connected and incompressible boundary. We first give an overview of the definition and special properties of such a triangulation.

Since ∂M_i is connected and incompressible in M_i, by [8], M_i admits a special triangulation with the following properties:

(1) the triangulation has only one vertex which lies in ∂M_i

(2) the only normal disk is the vertex-linking one,

(3) there is no normal S^2 in M_i

We call such a triangulation a *0–efficient triangulation* for M_i. It is also shown in [8] that there is an algorithm to find such a triangulation.

Similar to 0–efficient triangulations for closed 3–manifolds, such triangulations have some remarkable properties. The following lemma was proved by Jaco and Rubinstein and the proof is basically the same as the closed case, also see [12, Lemma 5.1]. The proof of the Lemma 4.1 uses a technique in [8] called *barrier*. A barrier is basically a 2–complex barrier for the normalization operations. We refer the reader to [12, section 5] for a brief explanation and [8, section 3.1] for more details. The proof of Lemma 4.1 is similar in spirit to that of [12, Lemma 5.1].

Lemma 4.1 *Let M_i be a simple 3–manifold with connected boundary and $\mathcal{T}$ a 0–efficient triangulation. Then every properly embedded normal annulus with respect to $\mathcal{T}$ is ∂–parallel and incompressible.*

Remark 4.2 M_i does not contain any normal Möbius band, since the boundary of a neighborhood of a normal Möbius band is a normal annulus, which contradicts the above lemma and the assumption that M_i is simple.

Proof Let A be a properly embedded normal annulus in M_i. Since M_i is simple, every incompressible annulus is ∂–parallel. So it suffices to prove that A is incompressible.

Suppose A is compressible, then ∂A must be trivial curves in ∂M_i since ∂M_i is incompressible in M_i. Note that the induced triangulation of ∂M_i has only one vertex. The only trivial normal curve in a one-vertex triangulation of ∂M_i is vertex-linking (see part (a) of Proposition 4.3). Hence ∂A is a pair of parallel vertex-linking curves. Let γ_1 and γ_2 be the two components of ∂A and D_j ($j = 1, 2$) the disk bounded by γ_j in ∂M_i. As γ_1 and γ_2 are parallel, we may suppose $D_1 \subset D_2$.

Note that the disk $A \cup D_1$ may not be normal, but A is a barrier for the normalization operations that make $A \cup D_1$ normal. So we can normalize $A \cup D_1$ to a normal disk Δ. Since the triangulation is 0–efficient, Δ is a vertex-linking disk. Since A is a barrier for the normalizing operations, A must lie in the 3–ball bounded by Δ and a disk of ∂M_i. However, there is no normal annulus in a small neighborhood of the vertex, a contradiction. □

Notation To simplify notation, in the remaining of the paper, we use M to denote either M_1 or M_2 and $F = \partial M_i$. Unless specified, we use S to denote the surface S_i in Corollary 3.5. We fix a 0–efficient triangulation of M and assume S is a normal or an almost normal surface in M with respect to the 0–efficient triangulation and ∂S consists of essential normal curves in ∂M.

Now we consider all the properly embedded normal and almost normal surfaces in M whose boundary consists of essential curves in ∂M. Similar to [5; 12], there is a finite collection of branched surfaces each obtained by gluing normal disks and at most one almost normal piece, such that each of these normal or almost normal surfaces is fully carried by a branched surface in this collection. Moreover, similar to [1; 12], since there is no normal S^2 and the only normal disk in this triangulation is vertex-linking, after taking sub-branched surfaces if necessary, we may assume no branched surface in this collection carries a normal disk or normal S^2.

Let B be a branched surface in this collection that fully carries S. So ∂B is a train track in ∂M. We call a train track a *normal train track* if every curve carried by the train track is normal with respect to the induced triangulation of ∂M. By the construction, ∂B is a normal train track.

Note that in the general case from the appendix, ∂S may contain a single trivial curve, though at least one component of ∂S is essential. In this case, we may split B so that a component of ∂B is an isolated trivial circle and each other component of ∂B fully carries an essential curve (as required by part (c) of Proposition 4.3). For simplicity, as mentioned in Remark 3.6, we assume ∂S is essential and ∂B fully carries ∂S.

Proposition 4.3

(a) *A normal simple closed curve in ∂M is trivial if and only if it is vertex-linking.*

(b) *At most one component of $\partial M - \partial B$ is a monogon.*

(c) *The train track ∂B does not carry any trivial curve.*

Proof Part (a) follows from the fact that the induced triangulation of ∂M has only one vertex. To see this, let γ be a normal trivial curve and D the disk bounded by γ in ∂M. Let e be any edge (or 1–simplex) in the induced triangulation of ∂M. If a component α of $e \cap D$ is an arc in $\mathrm{int}(e)$, then α is properly embedded in D and cuts D into two subdisks D_1 and D_2. As there is only one vertex, at least one subdisk, say D_1, does not contain the vertex. Hence the intersection of D_1 and the 1–skeleton of the triangulation consists of arcs properly embedded in D_1. These arcs cut D_1 into subdisks and an outermost subdisk is a bigon with one edge in ∂D and the other edge in the 1–skeleton. This means that $\gamma = \partial D$ is not a normal curve, a contradiction. Therefore, every component of $e \cap D$ is an arc with one endpoint the vertex of the triangulation and the other endpoint in ∂D. This implies that $\gamma = \partial D$ is vertex-linking.

The proof of part (b) is similar. Since every curve carried by ∂B is a normal curve, the argument above implies that each monogon component of $\partial M - \partial B$ must contain the vertex of the triangulation. Part (b) follows from that assumption that there is only one vertex in the triangulation.

Part (c) follows from the assumption that B fully carries S and ∂S consists of essential curves. Let $N(\partial B)$ be a fibered neighborhood of the train track ∂B in ∂M. We may assume ∂S lies in $N(\partial B)$ and is transverse to the interval fibers of $N(\partial B)$. Since ∂B fully carries ∂S, after some isotopy and taking multiple copies of ∂S if necessary, we may assume that the horizontal boundary of $N(\partial B)$ lies in ∂S. Since each component of ∂S is essential, this means that no horizontal boundary component of $N(\partial B)$ is a trivial circle. In other words, no component of $\overline{\partial M - \partial B}$ (or $\partial M - \mathrm{int}(N(\partial B))$) is a disk with smooth boundary. If ∂B carries a trivial circle γ, then a trivial index argument implies that the disk bounded by γ contains either a disk component of $\overline{\partial M - \partial B}$ with smooth boundary or at least two monogons. The first case is impossible by the argument above and the second case is ruled out by part (b). So ∂B does not carry any trivial curve. □

Each surface carried by B is corresponding to a nonnegative integer solution to the system of branch equations, see [1; 5; 12] for more detailed discussion. To simplify notation, we will not distinguish between a surface carried by B and its corresponding integer solution to the system of branch equations.

By the normal surface theory, there is a finite set of fundamental solutions of the system of branch equations such that any surface carried by B is a linear combination of the fundamental solutions with nonnegative integer coefficients. We denote the fundamental solutions by $F_1,\dots,F_s,C_1,\dots,C_t,A_1,\dots A_n$, where each A_j is a normal annulus carried by B, each C_j is a closed surface carried by B and the F_j's are the other fundamental solutions. So the surface S can be written as $S=\sum s_jF_j+\sum t_jC_j+\sum n_jA_j$ where each s_j, t_j or n_j is a nonnegative integer.

Proposition 4.4 $\sum s_j\le 2-\chi(S)$.

Proof Since S is a normal or an almost normal surface, we may assume that at most one fundamental solution contains an almost normal piece and its coefficient in the linear combination above is either 0 or 1.

Note that M does not contain any normal projective plane, since the boundary of a twisted I–bundle over a normal P^2 is a normal S^2 and M does not contain any normal S^2. Moreover B does not carry any normal disk by our assumption. These imply that B does not carry any normal surface with positive Euler characteristic.

We first consider the case that S is a normal surface. First, we have $\chi(S)=\sum s_j\chi(F_j)+\sum t_j\chi(C_j)+\sum n_j\chi(A_j)$. Since S is normal, each fundamental solution with positive coefficient in the linear combination above is a normal surface. Since B does not carry a normal surface with positive Euler characteristic, we have

$$\chi(S)=\sum s_j\chi(F_j)+\sum t_j\chi(C_j)\le\sum s_j\chi(F_j)\le-\sum s_j.$$

So in the case that S is a normal surface, we have $\sum s_j\le-\chi(S)$. If S is almost normal, we may suppose some C_k (or F_k) is almost normal and the coefficient of C_k (or F_k) is 1. Note that since $\chi(C_k)$ (or $\chi(F_k)$) is at most 2, we have $\chi(S)\le 2-\sum s_j$ and $\sum s_j\le 2-\chi(S)$. □

Since there are only finitely many such branched surfaces B, to prove Theorem 1.2, it suffices to show that the set of boundary curves of surfaces carried by B with bounded Euler characteristic has bounded diameter in the curve complex of F. Since each C_j is a closed surface, $\partial S=\sum s_j\partial F_j+\sum n_j\partial A_j$. As $\sum s_j$ is bounded by Proposition 4.4, there are only finitely many possibilities for curves $\sum s_j\partial F_j$. Thus the key part of the proof is to study normal annuli carried by B.

5 Normal annuli

We use the same notation. Let B be a branched surface in M that fully carries S as above and $A_1,\dots,A_n$ the fundamental solutions that correspond to normal annuli carried by B. Since B does not carry any normal surface of positive Euler characteristic, each component of the normal sum $\sum n_i A_i$ must have Euler characteristic 0 and hence is either a normal torus or a normal annulus carried by B. Note that there is no normal Klein bottle in the 0–efficient triangulation, see Lemma 5.1 and Corollary 5.2 in [12].

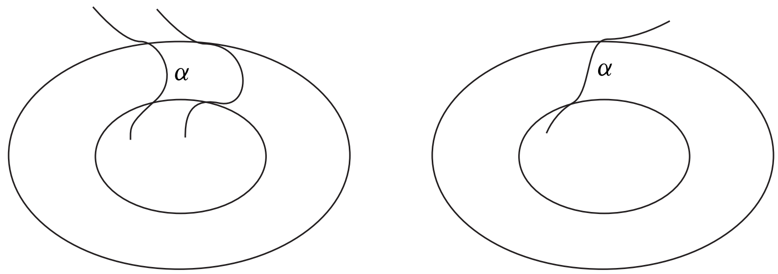
Let $N(B)$ be a fibered neighborhood of B and $\pi\colon N(B) \to B$ the map collapsing each I–fiber to a point, see [5; 12] for more details. We may view $A_1, \dots A_n$ as embedded annuli in $N(B)$. Then $\pi(\sum n_i A_i)$ is a sub-branched surface of B fully carrying $\sum n_i A_i$. Since each A_j is ∂–parallel, there is an annulus $\Gamma_j \subset \partial M$ such that $\partial\Gamma_j = \partial A_j$ and A_j is isotopic to Γ_j relative to ∂A_j. Throughout this paper, we will use T_j to denote the solid torus bounded by $A_j \cup \Gamma_j$.

Next we study the intersection of two normal annuli carried by B. Let A_1 and A_2 be two annuli carried by B and suppose $A_1 \cap A_2 \neq \varnothing$. As above, let Γ_1 and Γ_2 be the annuli in ∂M bounded by ∂A_1 and ∂A_2 respectively.

If $A_1 \cap A_2$ contains a closed curve γ, then since every normal annulus is incompressible by Lemma 4.1, γ is either trivial in both A_1 and A_2 or essential in both A_1 and A_2. Let Γ be the union of closed curves in $A_1 \cap A_2$ that are trivial in both A_1 and A_2. Let P_i be the component of $A_i - \Gamma$ that contains ∂A_i ($i = 1, 2$). Clearly $P_1 \cap P_2 = (A_1 \cap A_2) - \Gamma$. Now we perform standard cutting and pasting along Γ and denote by A_i' the resulting component that contains P_i ($i = 1, 2$). If $A_1' = A_2'$, then $\chi(A_1') < 0$, which means that the cutting and pasting above also produces an embedded normal surface with positive Euler characteristic, a contradiction to the assumptions of the branched surface B. Thus $A_1' \neq A_2'$, each A_i' is an embedded normal annulus, and $A_1' \cap A_2'$ does not contain any trivial closed curves. Note that the cutting and pasting above may produce a normal torus. Therefore, after some cutting and pasting above, we may assume the intersection of two normal annuli does not contain trivial curves.

Definition 5.1 Suppose $\partial A_1 \cap \partial A_2 \neq \varnothing$. This means that $\Gamma_1 \cap \partial A_2 \neq \varnothing$. We consider an arc α of $\Gamma_1 \cap \partial A_2$ with endpoints in different components of $\partial\Gamma_1$. Since ∂A_1 and ∂A_2 are carried by B, $\partial A_1 \cup \partial A_2$ naturally forms a train track. We say α is of type I in Γ_1 if $\partial\Gamma_1 \cup \alpha$ form a train track of a Reeb annulus, as shown in Figure 5.1(a). Otherwise, the train track $\partial A_1 \cup \alpha$ is as shown in Figure 5.1(b) and we say α is of type II in Γ_1. We say A_1 is of type I relative to A_2 if a component of $\Gamma_1 \cap \partial A_2$ is of type I, otherwise we say A_1 is of type II relative to A_2.

Note that if there are two type I arcs of $\Gamma_1 \cap \partial A_2$ with opposite switching directions along $\partial\Gamma_1$, then the train track $\pi(\partial A_1 \cup \partial A_2)$ carries a trivial circle. By part (c) of Proposition 4.3, ∂B does not carry any trivial circle. So all the type I arcs of $\Gamma_1 \cap \partial A_2$ must have coherent switching directions as shown in Figure 5.1(a), ie, the train track formed by ∂A_1 and these type I arcs carries a Reeb lamination of an annulus.



type I (a) type II (b)

Figure 5.1

Proposition 5.2 *Let Γ be an annulus in ∂M and suppose $\partial\Gamma$ consists of normal curves with respect to a one-vertex triangulation of ∂M. Let α be a properly embedded essential arc in Γ. Suppose $\partial\Gamma \cup \alpha$ forms a Reeb train track as shown in Figure 5.1(a) which carries a normal Reeb lamination. Then Γ contains the vertex of the triangulation.*

Proof We may deform $\partial\Gamma \cup \alpha$ into a train track τ. Note that our hypothesis says that the Reeb lamination carried by τ is normal with respect to the one-vertex triangulation of ∂M.

Suppose that Γ does not contain the vertex. Let e be an edge intersecting Γ. Let β be a component of $e \cap \Gamma$. Since Γ does not contain the vertex, there are only two possibilities: (1) $\partial\beta$ lies in the same circle of $\partial\Gamma$ and (2) the endpoints of β lie in different circles of $\partial\Gamma$. If $\partial\beta$ lies in the same component of $\partial\Gamma$, then similar to the proof of Proposition 4.3, this component of $\partial\Gamma$ must have trivial intersection with the edge e and hence cannot be a normal curve. Similarly, if the endpoints of β lie in different circles of $\partial\Gamma$, then every non-compact leaf of the Reeb lamination has trivial intersection with the edge e and hence the Reeb lamination carried by τ cannot be normal. Thus Γ must contain the vertex. □

An isotopy is called a *normal isotopy* if it is invariant on each simplex of the triangulation. Next we will perform some normal isotopies on ∂A_i. If γ is normally isotopic to ∂A_i, then A_i is normally isotopic to a normal annulus A_i' with $\partial A_i' = \gamma$. Moreover, for any surface X carried by B, we may assume $A_i + X$ is normally isotopic to $A_i' + X$. Next we will perform some normal isotopies and these normal isotopies do not change the surface under normal sum.

Definition 5.3 Let X be a point of $\partial A_1 \cap \partial A_2$. A small neighborhood of X is cut into 4 corners by $\partial A_1 \cup \partial A_2$. A corner is called a cusp if it becomes a cusp after deforming $\partial A_1 \cup \partial A_2$ into a train track. We call a disk D in ∂M a *bigon* if (1) ∂D consists of two arcs, one from ∂A_1 and the other from ∂A_2, and (2) the two corners of D at $\partial A_1 \cap \partial A_2$ are both cusps. We say D is an innermost bigon if $\mathrm{int}(D) \cap (\partial A_1 \cup \partial A_2) = \varnothing$. A bigon is said to be *trivial* if it does not contain the vertex of the triangulation.

Eliminate a trivial bigon Let D be an innermost trivial bigon and $b_1 \subset \partial A_1$ and $b_2 \subset \partial A_2$ the two edges of $\partial D = b_1 \cup b_2$. Since D does not contain the vertex and both ∂A_1 and ∂A_2 are normal curves, the intersection of D and the 1–skeleton of the triangulation consists of arcs with one endpoint in b_1 and the other endpoint in b_2. This means that b_1 and b_2 are normally isotopic. Hence we can perform a normal isotopy on A_i near ∂A_i, changing ∂A_1 to $(\partial A_1 - b_1) \cup b_2$ and ∂A_2 to $(\partial A_2 - b_2) \cup b_1$. After the normal isotopy and a small perturbation, $\partial A_1 \cap \partial A_2$ has fewer intersection points. We may successively eliminate all the trivial bigons using such normal isotopies.

For a given finite set of annuli carried by B, after some normal isotopies as above, we may assume that for any pair A_i and A_j, $\partial A_i \cup \partial A_j$ does not form any trivial bigon.

Definition 5.4 Let α be an arc component of $A_1 \cap A_2$ that is trivial (ie ∂–parallel) in both A_1 and A_2. Then α together with a subarc β_i of ∂A_i $(i = 1, 2)$ bounds a subdisk D_i of A_i. If $D_1 \cap D_2 = \alpha$ then $\beta_1 \cup \beta_2$ bounds a disk Δ in ∂M and $D_1 \cup D_2 \cup \Delta$ is a 2–sphere bounding a 3–ball. We call such a 3–ball a *football region*. Note that since the endpoints of β_i are also the endpoints of α and since A_1 and A_2 are carried by the same branched surface B, after deforming $\beta_1 \cup \beta_2$ into train track, $\beta_1 \cup \beta_2$ cannot form a monogon. Since the train track ∂B does not carry any trivial circle, Δ must be a bigon. Moreover, since we have assumed that there is no trivial bigon, the bigon Δ must contain the vertex of the triangulation. A football region is said to be *innermost* if it does not contains any other football region. A football region bounded by $D_1 \cup D_2 \cup \Delta$ said to be *trivial* if $D_1 \cap A_2 = D_2 \cap A_1 = \alpha$. Clearly a trivial football region must be innermost.

Eliminate a trivial football region Suppose the football region bounded by $D_1 \cup D_2 \cup \Delta$ as above is trivial. Let $\alpha = \partial D_1 \cap \partial D_2$. Since $D_1 \cap A_2 = D_2 \cap A_1 = \alpha$, we can perform a canonical cutting and pasting along α and obtain annuli $(A_1 - D_1) \cup D_2$ and $(A_2 - D_2) \cup D_1$. Clearly $(A_1 - D_1) \cup D_2$ and $(A_2 - D_2) \cup D_1$ are embedded annuli carried by B and are isotopic to A_1 and A_2 respectively. After a slight perturbation, the resulting annuli have fewer intersection curves. Thus, after a finite number of such operations, we may assume there is no trivial football region.

Definition 5.5 We say $A_1 \cup A_2$ is bigon-efficient if $A_1 \cap A_2$ contains no trivial closed curve, $\partial A_1 \cup \partial A_2$ does not form any trivial bigon in ∂M, and $A_1 \cup A_2$ does not form any trivial football region.

As above, we can perform some canonical cutting and pasting along $A_1 \cap A_2$ and get a pair of new annuli A'_1 and A'_2 such that $A'_1 \cup A'_2$ is bigon-efficient. By our construction, A'_1 and A'_2 are also carried by B and $A_1 + A_2 = A'_1 + A'_2$.

Next we will assume that $A_1 \cup A_2$ is bigon-efficient and consider the intersection pattern of $A_1 \cap A_2$.

Lemma 5.6 *Let β_0 be an arc in $A_1 \cap A_2$ and suppose β_0 is ∂–parallel in A_1. Let Δ_0 be the subdisk of A_1 bounded by β_0 and a subarc of ∂A_1. Let $\beta_1, \dots, \beta_k$ be the components of* $\operatorname{int}(\Delta_0) \cap A_2$. *Suppose each β_i ($i \geq 1$) is outermost in A_1. Then at least one β_i ($i \geq 1$) is ∂–parallel in A_2.*

Proof Suppose each β_i ($i \geq 1$) is an essential arc in A_2. Let δ_i be the subdisk of Δ_0 bounded by β_i ($i \geq 1$) and a subarc of ∂A_1. Since β_i ($i \geq 1$) is essential in A_2 and outermost in A_1, each δ_i is a ∂–compressing disk for A_2. This implies that $\partial \delta_i \cap \partial M$ is a type I arc in Γ_2. By Proposition 5.2, Γ_2 contains the vertex of the triangulation.

Since A_2 is ∂–parallel in M, $A_2 \cup \Gamma_2$ bounds a solid torus T_2. Let M' be the closure of $M - T_2$. So $M' \cong M$ and we may view A_2 as an annulus in $\partial M'$.

We use D to denote the closure of $\Delta_0 - \cup_{i=1}^{k} \delta_i$. Thus we may view D as a disk properly embedded in M'. Since $\partial M'$ is incompressible in M', ∂D bounds a disk D' in $\partial M'$. We view A_2 as a subannulus of $\partial M'$. So $D' \cap A_2 \neq \varnothing$.

Note that ∂A_2 cuts D' into disks and at least two such disks are outermost in D' (an outermost disk is a disk whose boundary consists of a subarc of $\partial D'$ and a subarc of ∂A_2). Let Δ be such an outermost disk. If $\Delta \subset A_2 \subset \partial M'$, then since each β_i ($i \geq 1$) is essential in A_2, β_0 must be an arc in $\partial \Delta$. Since there are at least two outermost disks, we may choose Δ to be outside A_2. In other words, $\Delta \subset \partial M' - \operatorname{int}(A_2) = \partial M - \operatorname{int}(\Gamma_2)$. Since Γ_2 contains the vertex of the triangulation, this means that Δ does not contain the vertex. If we deform $\partial A_1 \cup \partial A_2$ into a train track, then Δ becomes either a bigon or a monogon or a smooth disk. As in the proof of Proposition 4.3, a monogon or a smooth disk must contain the vertex. Since Δ does not contain the vertex, Δ must be a trivial bigon, which contradicts our assumption that $A_1 \cup A_2$ is bigon-efficient. □

Lemma 5.7 *Let A_1 and A_2 be as above and suppose $A_1 \cup A_2$ is bigon-efficient. Then A_1 and A_2 do not form any football region.*

Proof Suppose there is a football region X bounded by $D_1 \cup D_2 \cup \Delta$, where $D_i \subset A_i$ is a disk bounded by a component α of $A_1 \cap A_2$ and a subarc of ∂A_i and $\Delta \subset \partial M$. We use β_i ($\beta_i \subset \partial A_i$) to denote $\partial D_i - \mathrm{int}(\alpha)$ ($i = 1, 2$). Note that Δ must contain the vertex of the triangulation, because otherwise Δ is a trivial bigon contradicting that $A_1 \cup A_2$ is bigon-efficient. Without loss of generality, we may assume X does not contain any other football region.

If $D_1 \cap A_2 = D_2 \cap A_1 = \alpha$, then the 3–ball bounded by $D_1 \cup D_2 \cup \Delta$ is a trivial football region, contradicting the assumption that $A_1 \cup A_2$ is bigon-efficient. So we may assume $\mathrm{int}(D_1) \cap A_2 \neq \varnothing$.

Since $\mathrm{int}(D_1) \cap A_2 \neq \varnothing$, we can always find a component β_0 of $D_1 \cap A_2$ such that β_0 is not outermost in A_1 but every component of $\mathrm{int}(D_1) \cap A_2$ inside the disk bounded by β_0 and a subarc of ∂A_1 is outermost in A_1. By Lemma 5.6, there is at least one arc $\alpha' \subset \mathrm{int}(D_1) \cap A_2$ that is outermost in D_1 and ∂–parallel in A_2. Since α' is outermost, α' and a subarc of β_1, say β_1', bound a subdisk d_1 of D_1 and $d_1 \cap A_2 = \alpha'$. Since α' is ∂–parallel in A_2, α' and a subarc of ∂A_2, say β_2', bound a subdisk d_2 of A_2. Moreover, since $d_1 \cap A_2 = \alpha'$, $\beta_1' \cup \beta_2'$ bounds an embedded bigon Δ' in ∂M and $d_1 \cup d_2 \cup \Delta'$ bounds a football region, which we denote by X'.

If $\mathrm{int}(d_2) \cap D_1 = \varnothing$, then either $X' \subset X$ or $\mathrm{int}(X) \cap \mathrm{int}(X') = \varnothing$. Since the football region X is assumed to be innermost, X' does not lie in X. Moreover, since $\partial A_1 \cup \partial A_2$ does not form any trivial bigon, both football regions X and X' must contain the vertex of the triangulation. This means that $\mathrm{int}(X) \cap \mathrm{int}(X') \neq \varnothing$. Thus $\mathrm{int}(d_2) \cap D_1 \neq \varnothing$.

Let $\alpha'' \subset d_2 \cap D_1$ be an outermost intersection arc in d_2. We use e_2 to denote the subdisk of d_2 bounded by α'' and β_2' ($e_2 \cap D_1 = \alpha''$). As $\alpha'' \subset D_1$, the arc α'' and a subarc of β_1 bound a subdisk of D_1 which we denoted by e_1. As before, e_1, e_2 and a bigon in ∂M bound another football region, which we denote by X''. Since $e_1 \subset D_1$ and $e_2 \cap D_1 = \alpha''$, if e_2 lies in the football region X, then $X'' \subset X$ contradicting the assumption the X is innermost. Similarly, if e_2 is outside X, then since $e_2 \cap D_1 = \alpha''$, X'' must be outside X and $X'' \cap \mathrm{int}(X) = \varnothing$. As before, this is also impossible because by our assumptions every football region must contain the vertex of the triangulation, which implies $X'' \cap \mathrm{int}(X) \neq \varnothing$. □

Corollary 5.8 *Let α be an arc component of $A_1 \cap A_2$ and suppose α is ∂–parallel in A_1. Then the following are true.*

(1) *α must be outermost in A_1.*

(2) *α must be an essential arc in A_2.*

Proof We first prove that if α is outermost in A_1 then α must be an essential arc in A_2. Suppose otherwise that α is ∂–parallel in A_2. Since α is outermost in A_1, the two subdisks of A_1 and A_2 cut off by α form an embedded disk and bound a football region, which contradicts Lemma 5.7.

Since α is ∂–parallel in A_1, α and a subarc of ∂A_1 bound a subdisk D of A_1. Suppose α is not outermost. Then we can choose α so that every component of $\mathrm{int}(D) \cap A_2$ is outermost in A_1. Let $\alpha_1, \dots, \alpha_k$ be the components of $\mathrm{int}(D) \cap A_2$. Since each α_i is outermost, by the argument above, every α_i is an essential arc in A_2. This is an immediate contradiction to Lemma 5.6.

Part (2) follows from part (1) and the argument above. □

Lemma 5.9 *Suppose $A_1 \cup A_2$ is bigon-efficient. If $A_1 \cap A_2$ contains an arc that is ∂–parallel in A_1 then*

(1) *every arc of $A_1 \cap A_2$ is ∂–parallel and outermost in A_1 but essential in A_2,*

(2) *$A_1 \cap T_2$ consists of ∂–compressing disks of A_2,*

(3) *every arc of $\partial A_1 \cap \Gamma_2$ is of type I in Γ_2 and every arc of $\partial A_2 \cap \Gamma_1$ is of type II in Γ_1, see Figure 5.1*

Proof We first claim that $A_1 \cap A_2$ contains no closed curve. Suppose otherwise $A_1 \cap A_2$ contains a closed curve. Since A_1 and A_2 are incompressible by Lemma 4.1, every closed curve in $A_1 \cap A_2$ is either essential in both A_1 and A_2 or trivial in both A_1 and A_2. Since $A_1 \cup A_2$ is bigon-efficient, a closed curve in $A_1 \cap A_2$ is essential in both annuli. This implies that every arc component of $A_1 \cap A_2$ is ∂–parallel in both A_1 and A_2, a contradiction to Corollary 5.8.

Suppose $A_1 \cap A_2$ contains an arc which is essential in A_1 and let $\gamma_1, \dots, \gamma_k$ be all the components of $A_1 \cap A_2$ that are essential in A_1. Then $\gamma_1, \dots, \gamma_k$ cut A_1 into a collection of rectangles $R_1, \dots, R_k$ and we can suppose R_i is the rectangle between γ_i and γ_{i+1} (setting $\gamma_{k+1} = \gamma_1$). In other words, γ_i and γ_{i+1} are two opposite edges of R_i and the other two edges of R_i are subarcs of ∂A_1.

Since $A_1 \cap A_2$ contains an arc trivial in A_1, at least one R_i contains other arcs of $A_1 \cap A_2$. Let $\alpha_1, \dots, \alpha_m$ be the components of $\mathrm{int}(R_i) \cap A_2$. By our construction of R_i, each α_j is ∂–parallel in A_1. By Corollary 5.8, each α_j is ∂–parallel and outermost in A_1. Hence each α_j and a subarc of ∂A_1 bound a disk Δ_j in R_i and these Δ_j's are pairwise disjoint. Moreover, each Δ_j is a ∂–compressing disk of A_2, in particular $\Delta_j \subset T_2$. This implies that ∂A_2 and the arcs $\partial \Delta_j \cap \Gamma_2$ naturally deform into a Reeb train track. By Proposition 5.2, Γ_2 contains the vertex of the triangulation.

Let P and M' be the closures of $R_i - \bigcup_{j=1}^{m} \Delta_j$ and $M - T_2$ respectively. So P is a disk properly embedded in M'. Let P' be the disk bounded by ∂P in $\partial M'$. We may consider A_2 as an annulus in $\partial M'$ and $P' \cap A_2 \neq \varnothing$. Similar to the proof of Lemma 5.6, ∂A_2 cuts P' into a collection of disks and there are at least two outermost such disks. If an outermost disk Δ lies in $\partial M' - \mathrm{int}(A_2) = \partial M - \mathrm{int}(\Gamma_2)$, as in the proof of Lemma 5.6, Δ must contain the vertex, which contradicts the previous conclusion that the vertex lies in Γ_2. This means that every outermost disk in $P' - \partial A_2$ lies in A_2. Since each α_j ($j = 1, \ldots, m$) is essential in A_2, this implies that there are exactly two outermost disks and both γ_i and γ_{i+1} must be ∂–parallel arcs in A_2.

Let β_i and β_{i+1} be subarcs of ∂A_2 such that $\partial\gamma_i = \partial\beta_i$, $\partial\gamma_{i+1} = \partial\beta_{i+1}$, and $\gamma_i \cup \beta_i$ and $\gamma_{i+1} \cup \beta_{i+1}$ bound subdisks δ_i and δ_{i+1} of A_2 respectively. By Corollary 5.8, β_i and β_{i+1} must both be outermost in A_2 and δ_i and δ_{i+1} are disjoint ∂–compressing disks for A_1. This implies that β_i and β_{i+1} are of type I in Γ_1.

Note that $R_i \cup \delta_i \cup \delta_{i+1}$ is a disk properly embedded in M. Moreover, $\partial A_1 \cup \beta_i \cup \beta_{i+1}$ naturally deforms into a Reeb train track and $\partial(R_i \cup \delta_i \cup \delta_{i+1})$ deforms into a bigon in the Reeb train track. Let Q' be the disk bounded by $\partial(R_i \cup \delta_i \cup \delta_{i+1})$ in ∂M, see the shaded region in Figure 5.2(a) for a picture of Q'. Clearly $Q' \subset \Gamma_1$. As above, we say a disk in $Q' - \mathrm{int}(\Gamma_2)$ is outermost if its boundary consists of an arc from ∂A_1 and an arc from ∂A_2. As in the proof of Lemma 5.6, any outermost disk must contain the vertex of the triangulation. Since Γ_2 contains the vertex, $Q' - \mathrm{int}(\Gamma_2)$ contains no outermost disk. This implies that $Q' \cap \Gamma_2$ consists of rectangles which naturally deform into bigons in the Reeb annulus Γ_2, see Figure 5.2(a) for a picture. As shown in Figure 5.2(a), at least one component of $Q' - \mathrm{int}(\Gamma_2)$ is a monogon (after deforming into a train track). Since a monogon contains the vertex of the triangulation, this implies that the vertex of the triangulation lies outside Γ_2, a contradiction. This proves part (1).

Part (2) is an immediate corollary of part (1).

Part (1) also implies that every arc of $\Gamma_2 \cap \partial A_1$ is of type I in Γ_2 and $\partial A_2 \cup (\Gamma_2 \cap \partial A_1)$ forms a standard Reeb train track. Now we consider $\partial A_2 \cap \Gamma_1$.

As above, since Γ_2 contains the vertex, $\Gamma_1 - \Gamma_2$ has no outermost disk (an outermost disk is a component with a boundary edge in ∂A_1 and a boundary edge in ∂A_2). This implies that every arc in $\partial A_2 \cap \Gamma_1$ is an essential arc in Γ_1. Since $\partial A_2 \cup (\Gamma_2 \cap \partial A_1)$ form a standard Reeb train track, as shown in Figure 5.2(b), every arc of $\partial A_2 \cap \Gamma_1$ must be of type II in Γ_1. □

Lemma 5.10 *Suppose $A_1 \cap A_2$ is bigon-efficient and $A_1 \cap A_2 \neq \varnothing$. Then no arc component of $A_1 \cap A_2$ is essential in both A_1 and A_2.*

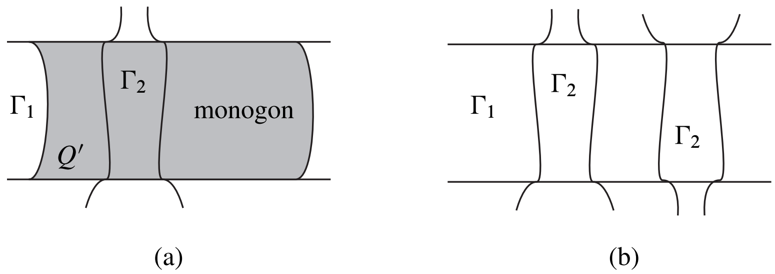


Figure 5.2

Proof Suppose there is an arc component of $A_1 \cap A_2$ that is essential in both A_1 and A_2. As in the proof of Lemma 5.9, $A_1 \cap A_2$ contains no closed curve. If there is a component of $A_1 \cap A_2$ that is trivial in A_1 then by Lemma 5.9 every component of $A_1 \cap A_2$ is trivial in A_1. Thus every arc of $A_1 \cap A_2$ must be essential in both A_1 and A_2.

So $A_1 \cap A_2$ cuts both A_1 and A_2 into a collections of rectangles. Let R be a component of $A_1 \cap T_2$. Two opposite boundary edges of the rectangle R are essential arcs in A_2 and the other two edges of ∂R, denoted by γ_1 and γ_2, are properly embedded in Γ_2. Since R is a disk properly embedded in the solid torus T_2, both γ_1 and γ_2 must be ∂–parallel in Γ_2. Moreover, since each arc in $A_1 \cap A_2$ is essential in both A_1 and A_2, $\partial\gamma_1$ and $\partial\gamma_2$ lie in different components of $\partial\Gamma_2$. Thus γ_1 and γ_2 and two subarcs of ∂A_2 (from different components of ∂A_2) bound two disjoint disks d_1 and d_2 in ∂M respectively. After naturally deforming $\partial A_1 \cup \partial A_2$ into a train track, d_1 and d_2 become bigons or monogons. Since $\partial A_1 \cap \partial A_2$ is bigon-efficient, every bigon contains the vertex of the triangulation. Since every monogon also must contain the vertex, this contradicts that d_1 and d_2 are disjoint and there is only one vertex in the triangulation. □

Corollary 5.11 *Suppose $A_1 \cap A_2$ is bigon-efficient. Suppose A_1 is of type I relative to A_2. Then every arc of $\Gamma_1 \cap \partial A_2$ is of type I in Γ_1 and every arc of $\Gamma_2 \cap \partial A_1$ is of type II in Γ_2.*

Proof By Lemma 5.10, no arc component of $A_1 \cap A_2$ is essential in both A_1 and A_2. Now the corollary follows from Lemma 5.9. □

Next we study the intersection patterns of 3 normal annuli carried by B.

Lemma 5.12 *Let A_1, A_2 and A_3 be pairwise bigon-efficient normal annuli carried by B. Suppose A_1 is of type I relative to A_2 and $\partial A_1 \cap \partial A_3 \neq \varnothing$. Then,*

(1) A_1 *must be of type* I *relative to* A_3 *and*

(2) $\partial A_2 \cap \partial A_3 = \varnothing$.

Proof Since $\partial A_1 \cap \partial A_3 \neq \varnothing$, by Lemmas 5.10 and 5.9, either A_1 is of type I relative to A_3 or A_3 is of type I relative to A_1. Suppose part (1) is not true and A_3 is of type I relative to A_1. So by Lemma 5.9 and Proposition 5.2, Γ_3 contains the vertex of the triangulation. Moreover, since A_1 is of type I relative to A_2, both Γ_1 and Γ_3 contain the vertex.

Let R be the component of $\Gamma_1 \cap \Gamma_3$ that contains the vertex of the triangulation. By Lemma 5.9, $\partial A_1 \cap \Gamma_3$ consists of type I arcs in Γ_3, so R is a quadrilateral that naturally deforms into a bigon. Two opposite edges of ∂R, denoted by r_1 and r_2, are components of $\Gamma_1 \cap \partial A_3$. By Lemma 5.9, r_1 and r_2 are type II arcs in Γ_1, see Figure 5.3(a) for a picture of R. Let r_3 and r_4 be the other two edges of R. Hence, $r_3 \cup r_4$ are two components of $\partial A_1 \cap \Gamma_3$ and r_3 and r_4 are of type I in Γ_3.

Since A_1 is of type I relative to A_2, every component of $\partial A_2 \cap \Gamma_1$ is of type I in Γ_1 and $\partial A_1 \cup (\partial A_2 \cap \Gamma_1)$ forms a standard Reeb train track.

Case 1 $(\partial A_2 \cap \Gamma_1) \bigcap (r_1 \cup r_2) = \varnothing$

If a component of $\partial A_2 \cap \Gamma_1$ lies outside R, as shown in Figure 5.3(a), it creates a monogon region outside R. Since any monogon region contains the vertex, this contradicts that R contains the vertex. Thus $\partial A_2 \cap \Gamma_1 \subset R$.

Next we view R as a quadrilateral in Γ_3. Hence r_3 and r_4 are type I arcs in Γ_3. Each component of $\partial A_2 \cap R$ is an arc with one endpoint in r_3 and the other endpoint in r_4. Moreover, as shown in Figure 5.3(b), after deforming into a train track, $\partial A_2 \cap R$ cuts R into a monogon region X, a 3–prong triangle Y, and a collection of bigons.

Now we consider the disk $R' = \overline{\Gamma_3 - R}$. We first consider the possibility that there is an arc α of $\partial A_2 \cap R'$ with both endpoints in $r_3 \cup r_4$. Note that $\partial A_2 \cap R$ is as shown in Figure 5.3(b), so this configuration fixes the switching direction of $\partial\alpha$ in the train track. There are two cases to consider: (1) both endpoints of α lie in r_3 (or r_4) and (2) one endpoint of α lies in r_3 and the other lies in r_4. As shown in Figure 5.3(c) and (d), in either case, α produces a monogon in R', which means the vertex of the triangulation lies in R' and contradicts the assumption that R contains the vertex. Thus, every component of $\partial A_2 \cap R'$ has one endpoint in $r_3 \cup r_4$ and the other endpoint in $\partial\Gamma_3 \cap \partial R'$.

After deforming into a train track, R' becomes a bigon. Since R' does not contain the vertex, ∂A_2 cuts R' into a collection of disks, each of which becomes a bigon

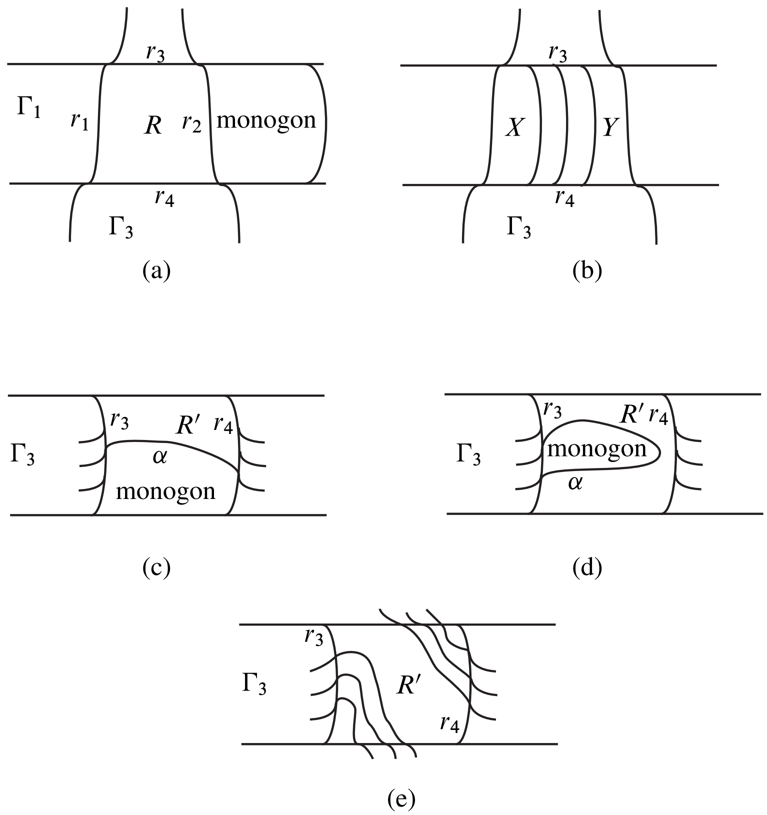


Figure 5.3

after deformed into a train track. Because of the switching direction of the train track at $\partial A_2 \cap (r_3 \cup r_4)$, as shown in Figure 5.3(e), the arcs with an endpoint in r_4 must have the same configuration. Otherwise, these arcs would create a monogon in R'. Furthermore, since every arc of $\partial A_2 \cap \Gamma_3$ is essential in Γ_3 by Lemma 5.9, the arcs with an endpoint in r_3 must also have the same configuration as shown in Figure 5.3(e). In other words, Figure 5.3(e) is the only possible configuration for $\partial A_2 \cap R'$.

As shown in Figure 6.1(a) and (b), given a component α of ∂A_3 and any arc β intersecting α, there are essentially two different switching directions at $\alpha \cap \beta$ along α. By examining the switching directions of the train track at $\partial A_2 \cap \partial A_3$ in $\partial R'$ along ∂A_3 as shown in Figure 5.3(e), we can see that each component of $\partial A_2 \cap \Gamma_3$ must be of type II in Γ_3. Moreover, as shown in Figure 5.3(e), the argument above implies that the switching directions (of the train track) at the intersection points of ∂A_2 with any component of ∂A_3 are all the same. However, by part (3) of Lemma 5.9, the conclusion that $\partial A_2 \cap \Gamma_3$ contains a type II arc in Γ_3 implies that $\partial A_3 \cap \Gamma_2$ consists of type I arcs in Γ_2. This means that there are two arcs of $\partial A_2 \cap \Gamma_3$, similar to the r_1 and r_2 in Figure 5.3(a), whose endpoints on a component of $\partial \Gamma_3$ have opposite switching direction. This contradicts the previous conclusion (as depicted in Figure 5.3(e)) that all the switching directions at such points are the same.

Case 2 $(\partial A_2 \cap \Gamma_1) \bigcap (r_1 \cup r_2) \neq \varnothing$

We will perform some normal isotopies so that $(\partial A_2 \cap \Gamma_1) \bigcap (r_1 \cup r_2) = \varnothing$ after the isotopies.

Let $\alpha_i \subset \partial A_i$ $(i = 1, 2, 3)$ be 3 arcs intersecting each other and forming a triangle Δ as shown in Figure 5.4. Suppose Δ naturally deforms into a bigon and Δ does not contain the vertex of the triangulation. Then, as shown in Figure 5.4, the isotopy on α_3, fixing α_1 and α_2, is a normal isotopy. Next we will fix $\partial A_1 \cup \partial A_3$ and perform some isotopies as in Figure 5.4 so that $(\partial A_2 \cap \Gamma_1) \bigcap (r_1 \cup r_2) = \varnothing$ after the isotopies. Each isotopy pushes an intersection point of $\partial A_2 \cap \partial A_3$ out of Γ_1.

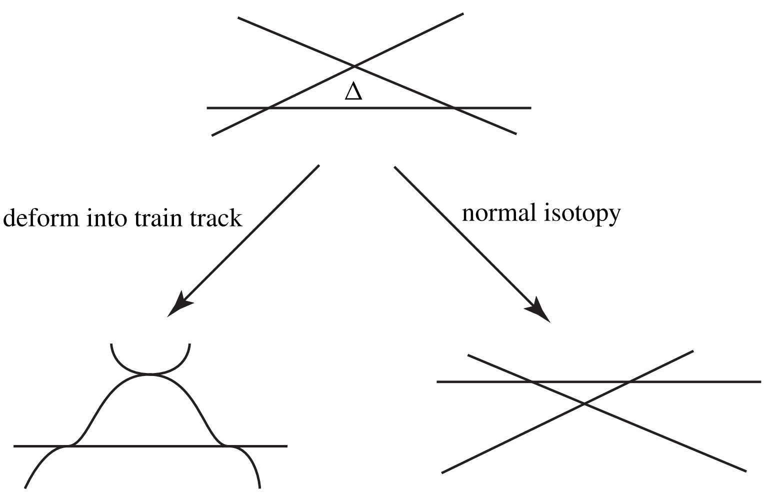


Figure 5.4

Let α be a component of $\partial A_2 \cap \Gamma_1$ and suppose $\alpha \cap (r_1 \cup r_2) \neq \varnothing$. Let α_1 and α_2 be the closure of the components of $\alpha - (r_1 \cup r_2)$ that contain $\partial \alpha$. So α_i $(i = 1, 2)$ has one endpoint in ∂A_1 and the other endpoint in $r_1 \cup r_2$. Thus α_1 and α_2 are the edges of two triangles Δ_1 and Δ_2 respectively formed by ∂A_1, ∂A_2 and ∂A_3. Since the two endpoints of α lie in different components of $\partial \Gamma_1$, Δ_1 and Δ_2 are not nested. Without loss of generality, we may assume each Δ_i is innermost, ie, $\mathrm{int}(\Delta_i)\cap(\partial A_1\cup\partial A_2\cup\partial A_3)=\varnothing$ for both $i=1,2$. After deforming $\partial A_1\cup\partial A_2\cup\partial A_3$ into a train track, each Δ_i becomes either a bigon or a monogon. Since each monogon contains the vertex of the triangulation, at least one Δ_i is a bigon that does not contain the vertex. Hence a normal isotopy on ∂A_2, as shown in Figure 5.4 pushes an intersection point of $\partial A_2\cap(r_1\cup r_2)$ out of Γ_1. So after finitely many such normal isotopies, $(\partial A_2\cap\Gamma_1)\bigcap(r_1\cup r_2)=\varnothing$ and we can apply Case 1 to obtain a contradiction.

Therefore, A_1 is of type I relative to both A_2 and A_3 and part (1) of the lemma holds. If $\partial A_2\cap\partial A_3\neq\varnothing$, then by Lemma 5.9 and Lemma 5.10, either A_2 is of type I relative to A_3, or A_3 of type I relative to A_2. Both possibilities contradict part (1), since A_1 is of type I relative to both A_2 and A_3. □

6 Boundary curves

Suppose $A_1\cap A_2$ is bigon-efficient. If $A_1\cap A_2$ contains a closed curve, by Lemma 5.9, all the components of $A_1\cap A_2$ must be closed essential curves. After performing canonical cutting and pasting along these curves, we get a pair of disjoint annuli A_1', A_2' and a possible collection of tori T. Clearly, $A_1+A_2=A_1'+A_2'+T$. In particular, $weight(A_1'+A_2')\leq weight(A_1+A_2)$.

Let $A_1,\dots,A_n$ be a fixed set of normal annuli carried by B. We consider m_i parallel copies of A_i $(i=1,\dots,n)$. Then we can perform the isotopy and cutting and pasting above on each pair of the $\sum_{i=1}^n m_i$ annuli, so that each pair of resulting set of annuli are bigon-efficient and have no closed intersection curve. So there is a set of normal annuli $\mathcal{A}$ such that for any set of nonnegative integers m_i, there is a collection of annuli $A_1',\dots,A_k'$ in $\mathcal{A}$ such that

(1) $\sum_{i=1}^n m_iA_i=T+\sum_{i=1}^k m_i'A_i'$, where T is a collection of normal tori.

(2) $A_1',\dots,A_k'$ are pairwise bigon-efficient

(3) $A_i'\cap A_j'$ contains no closed curve for any $i\neq j$.

We claim that one can choose $\mathcal{A}$ to be a finite set of annuli. Let $\partial\mathcal{A}$ be the set of boundary curves of all possible normal annuli resulting from the normal isotopies and canonical cutting and pasting as above (among all possible m_i's). Since the operations that make $A_k\cap A_j$ bigon-efficient, when restricted to ∂M, are simply cutting and pasting on bigons in ∂M, $\partial\mathcal{A}$ is a finite set of normal curves. Now suppose there is an infinite set of normal annuli, denoted by $\mathcal{D}$, in $\mathcal{A}$ with the same pair of boundary curves. Then by the normal surface theory, there must be two annuli in $\mathcal{D}$, say A_i' and

A'_j such that $A'_j = T' + A'_i$ where T' is a collection of normal tori. This means that A'_j is redundant as we can use $T' + A'_i$ instead. Therefore, we may choose $\mathcal{A}$ to be a finite set and there is an algorithm to find all the annuli in $\mathcal{A}$ using normal surface theory.

We are mainly interested in the boundary curves. In the conclusion (1) above, clearly $\sum_{i=1}^n m_i \partial A_i = \sum_{i=1}^k m'_i \partial A'_i$ in ∂M.

Since each A'_i is also a normal annulus, A'_i is ∂–parallel in ∂M. We use Γ'_i to denote the subannulus of ∂M isotopic to A'_i and with $\partial\Gamma'_i = \partial A'_i$.

Let $\mathcal{S}$ be the union of a fixed set of pairwise disjoint compact surfaces carried by B. Since $\partial\mathcal{S}$ is carried by the train track ∂B and ∂B does not carry any trivial circle, every component of $\partial\mathcal{S}$ is an essential normal curve. Next we consider $\partial\mathcal{S} + \sum m_i \partial A'_i$. Our goal is to prove the following lemma.

Lemma 6.1 *Let $\mathcal{S}$ and A'_i be as above. Then the diameter of the set $\{\partial\mathcal{S} + \sum m_i \partial A'_i\}$ (for all nonnegative integers m_i) in the curve complex $\mathcal{C}(F)$ is bounded.*

Suppose A'_1 and A'_2 are of type I relative to A'_i and A'_j respectively. If $\partial A'_1 \cap \partial A'_2 \neq \varnothing$, then by Lemma 5.9 and Lemma 5.10, one of A'_1 and A'_2 is of type II relative the other, contradicting Lemma 5.12. Thus $\partial A'_1 \cap \partial A'_2 = \varnothing$. Since both Γ'_1 and Γ'_2 contain the vertex of the triangulation by Proposition 5.2 and since $A'_1 \cap A'_2$ contains no closed curve, Γ'_1 and Γ'_2 must be nested and $\partial A'_1$ must be normally isotopic to $\partial A'_2$. Thus we have $m_1 \partial A'_1 + m_2 \partial A'_2 = (m_1 + m_2)\partial A'_1$.

We say A'_i is of type I if A'_i is of type I relative to one of $A'_1, \dots, A'_k$. The argument above implies that the boundary of all the type I annuli are normally parallel. Moreover, by Lemma 5.9 and Lemma 5.12, those annuli among $A'_1, \dots, A'_k$ that are not of type I are pairwise disjoint.

Next we will only focus on the boundary curves of $A'_1, \dots, A'_k$. If no A'_i is of type I, then Lemma 5.9 and Lemma 5.10 imply that these A'_i's are mutually disjoint. Suppose A'_1 a type I annulus. Since the boundary of other type I annuli are normally parallel to $\partial A'_1$, without loss of generality, we may assume A'_1 is the only type I annulus in $A'_1, \dots, A'_k$. By Lemma 5.9 and Lemma 5.12, this implies $A'_2 \dots, A'_k$ are pairwise disjoint. Let γ_i be a component of ∂A_i and k_i the number of intersection points of γ_i with $\partial\mathcal{S}$.

Lemma 6.2 *The distance between γ_j ($j \neq 1$) and $\partial\mathcal{S} + \sum_{i=2}^k m_i \partial A'_i$ is at most $2 + 2\log_2 k_j$.*

Proof By our earlier assumptions, $A'_2, \dots, A'_k$ are mutually disjoint. So $\sum_{i=2}^{k} m_i \partial A'_i$ is a union of disjoint curves and we may regard γ_j $(j \neq 1)$ as a component of $\sum_{i=2}^{k} m_i \partial A'_i$. Since the number of intersection points of γ_j with $\partial \mathcal{S}$ is k_j, the intersection number of γ_j and $\partial \mathcal{S} + \sum_{i=2}^{k} m_i \partial A'_i$ is at most k_j. Now it is clear that Lemma 6.2 follows from [7, Lemma 2.1], which says that the distance between any two curves with intersection number k is at most $2 + 2\log_2 k$. □

Note that Lemma 6.2 implies Lemma 6.1 in the case that no A'_i is of type I.

Lemma 6.3 *If there is some $\partial A'_j$ $(j \neq 1)$ disjoint from $\partial A'_1$, then the distance between γ_j and $\partial \mathcal{S} + \sum_{i=1}^{k} m_i \partial A'_i$ is at most $2 + 2\log_2 k_j$.*

Proof The proof is identical to that of Lemma 6.2. Since $A'_2, \dots, A'_k$ are mutually disjoint, γ_j can be viewed as a component of $\sum_{i=1}^{k} m_i \partial A'_i$. So the intersection number of γ_j and $\partial \mathcal{S} + \sum_{i=1}^{k} m_i \partial A'_i$ is at most k_j and Lemma 6.3 follows from Lemma 2.1 of [7]. □

So to prove Lemma 6.1, we may assume $\partial A'_i \cap \partial A'_1 \neq \varnothing$ for each $i \neq 1$. As A'_1 is of type I, every component of $\partial A'_i \cap \Gamma'_1$ is a type I arc in Γ'_1.

Let α_1 and α_2 be the two components of $\partial A'_1$. We fix a direction for the circle α_1 and assign the same direction to α_2. Let β be an arc carried by ∂B and intersecting α_i ($i = 1$ or 2) in one point. We say β and the point $\beta \cap \alpha_i$ are of positive (resp. negative) type if $\alpha_i \cup \beta$ deforms into a train track as in Figure 6.1(a) (resp. Figure 6.1(b)). Note that a curve carried by the train track Figure 6.1(a) or (b) is a spiral around α_i. We call a spiral carried by the train track in Figure 6.1(a) (resp. Figure 6.1(b)) a positive (resp. negative) spiral.

Let S be any compact surface carried by B and suppose $A'_1 \cap S$ contains an arc component γ. Then there are two cases (1) γ is ∂–parallel in A'_1 and (2) γ is an essential arc in A'_1. Since both S and A'_1 are carried by the same branched surface, as in [6], in either case, one endpoint of γ is of positive type and the other endpoint is of negative type. Let P_i $(i = 1, 2)$ be the number of points in $\partial S \cap \alpha_i$ of positive type and N_i the number of points in $\partial S \cap \alpha_i$ of negative type. The argument above implies that $P_1 + P_2 = N_1 + N_2$.

Let $N(\alpha_i)$ $(i = 1, 2)$ be a small annular neighborhood of α_i in ∂M. We consider $\partial S + m\alpha_i$ restricted to $N(\alpha_i)$. As depicted in Figure 6.1(c), if $P_i \neq N_i$ and $m \geq \min\{N_i, P_i\}$, then $\partial S + m\alpha_i$ restricted to $N(\alpha_i)$ consists of $|P_i - N_i|$ spirals and $2\min\{N_i, P_i\}$ ∂–parallel arcs in $N(\alpha_i)$. As shown in Figure 6.1(d), if $N_i = P_i$ and

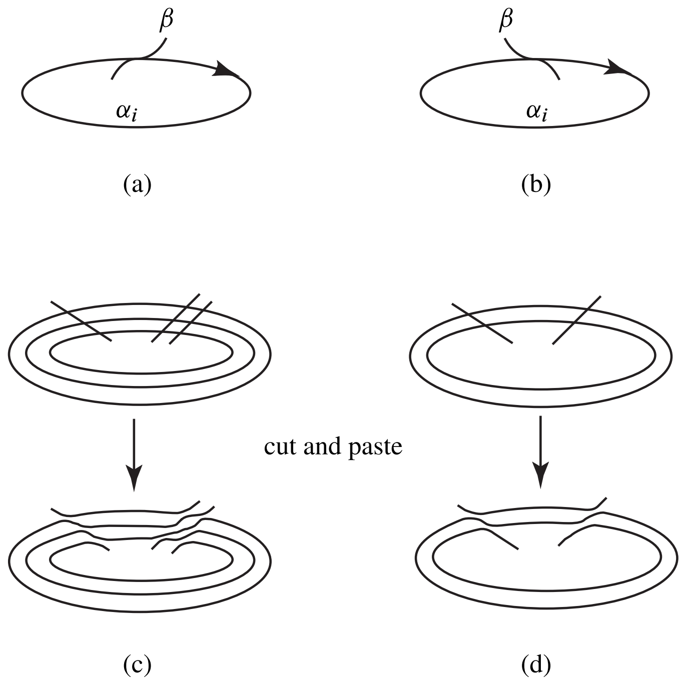
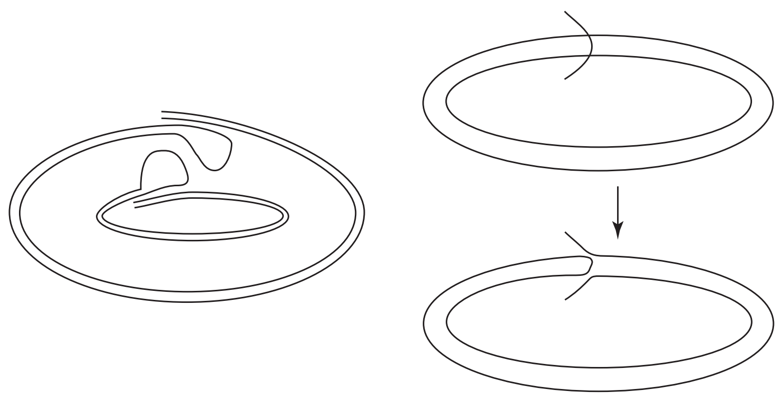


Figure 6.1

$m > \min\{N_i, P_i\}$, at least one component of $\partial S + m\alpha_i$ is parallel to α_i and hence we may view the distance $d_{\mathcal{C}(F)}(\partial S + m\alpha_i, \alpha_i) \leq 1$. Without loss of generality, we may assume $P_1 > N_1$. Since $P_1 + P_2 = N_1 + N_2$, $P_2 < N_2$. So if $m \geq \max\{N_1, P_2\}$, $\partial S + m\partial A_1'$ has $r = P_1 - N_1 = N_2 - P_2$ positive spirals in $N(\alpha_1)$ and r negative spirals in $N(\alpha_2)$.

Now we assume $m \geq \max\{N_1, P_2\}$ and consider $\partial S + m\partial A_1'$ restricted to $N(\Gamma_1')$, which is a small neighborhood of Γ_1' in ∂M. The positive and negative spirals in $N(\alpha_1)$ and $N(\alpha_2)$ are connected by some arcs of $\partial S \cap \Gamma_1'$. First suppose two positive spirals are connected by an arc in $\partial S \cap \Gamma_1'$. Then this arc and the two spirals in $N(\alpha_1)$ form a monogon whose "tail" spirals along α_1. Moreover, since the number of negative spirals equals the number of positive spirals, there must be an arc of $\partial S \cap \Gamma_1'$ connecting two negative spirals in $N(\alpha_2)$ and hence forming another monogon, as shown in Figure 6.2(a). Since each monogon must contain the vertex of the triangulation, this is a contradiction. Thus every positive spiral in $N(\alpha_1)$ is connected to a negative spiral in $N(\alpha_2)$ by an arc in $\partial S \cap \Gamma_1'$. The standard picture of these arcs are type

I arcs whose two ends spiraling around $\partial A_1'$. Therefore, as shown in Figure 6.2(b), $(\partial S + m\partial A_1') + \partial A_1'$ is isotopic to $\partial S + m\partial A_1'$.



(a) (b)

Figure 6.2

Now we assume the surface S in the argument above is the resulting surface of $\mathcal{S} + \sum_{i=2}^{n} m_i A_i'$ and let $\sigma = \sum_{i=2}^{n} m_i$. Clearly, there is a number K depending on $\mathcal{S} \cap A_1'$ and $A_i' \cap A_1'$, such that $K\sigma \geq \max\{P_1, N_1, P_2, N_2\}$. Thus by the discussion above, if $P_1 \neq N_1$ and $m_1 \geq K\sigma$, the set of curves $\{\partial S + m_1 \partial A_1'\}$ are all isotopic. Moreover, by Lemma 6.2, the set of curves $\{\partial \mathcal{S} + \sum_{i=2}^{n} m_i \partial A_i'\}$ for all m_i $(i = 2, \dots, n)$ has bounded diameter. As $S = \mathcal{S} + \sum_{i=2}^{n} m_i A_i'$, by the argument above, if $N_1 = P_1$ Lemma 6.1 holds, and if $P_1 \neq N_1$, Lemma 6.1 holds under the condition that $m_1 \geq K\sigma$.

Next we consider the case that $m_1 < K\sigma$. By our assumptions, A_1' is the only type I annulus and $\partial A_i' \cap \Gamma_1'$ consists of type I arcs in Γ_1'. So, as in Figure 6.2(b), $\sum_{i=1}^{n} m_i \partial A_i'$ is isotopic to $\sum_{i=2}^{n} m_i \partial A_i'$. Thus $\sum_{i=1}^{n} m_i \partial A_i'$ consists of 2σ closed curves.

Let ω be the maximal weight of $\partial A_i'$ among all i. So if $m_1 < K\sigma$, the total weight of $\sum_{i=1}^{n} m_i \partial A_i'$ is less than $K\sigma\omega + \sigma\omega = (K+1)\sigma\omega$. Since $\sum_{i=1}^{n} m_i \partial A_i'$ consists of 2σ closed curves, there is a component γ of $\sum_{i=1}^{n} m_i \partial A_i'$ with weight less than $(K+1)\omega/2$. Up to normal isotopy, there are only finitely many curves with weight under $(K+1)\omega/2$. So there is a number K' such that $|\partial \mathcal{S} \cap \gamma| \leq K'$. As in the proof of Lemma 6.2, by a theorem of Hempel [7], the distance between $\partial \mathcal{S} + \sum_{i=1}^{n} m_i \partial A_i'$ and γ is less than $2 + 2\log_2 K'$. As γ is isotopic to a component of $\partial A_i'$ for some i, in the case that $m_1 < K\sigma$, the distance between $\partial\mathcal{S} + \sum_{i=1}^{n} m_i \partial A_i'$ and some $\partial A_i'$ is bounded by a number that depends only on K, $\partial\mathcal{S}$ and the $\partial A_i'$'s.

Therefore, combining the two cases above, Lemma 6.1 holds. Moreover, it follows from the proof that the diameter of the set $\{\partial\mathcal{S} + \sum_{i=1}^{n} m_i \partial A_i'\}$ can be found algorithmically.

Now Theorem 1.2 follows from Lemma 6.1 and the discussions in Section 4.

Theorem 1.2 *Let M be a simple 3–manifold with connected boundary and a 0–efficient triangulation. Let S_k be the set of normal and almost normal surfaces satisfying the following two conditions*

(1) *the boundary of each surface in S_k consists of essential curves in ∂M*
(2) *the Euler characteristic of each surface in S_k is at least $-k$.*

Let C_k be the set of boundary curves of surfaces in S_k. Then C_k has bounded diameter in the curve complex of ∂M. Moreover, there is an algorithm to find the diameter.

Proof Let S be a normal or an almost normal surface with $-\chi(S) \le k$. So we have $S = \mathcal{S} + C + \sum m_i A_i$, where C is a closed surface and A_i is a normal annulus in the fundamental solution. Moreover, by Proposition 4.4, there are only finitely many possible surfaces for $\mathcal{S}$.

If we fix a $\mathcal{S}$, then Lemma 6.1 says that $\{\partial S = \partial\mathcal{S} + \sum m_i \partial A_i\}$ has bounded diameter. Since there are only finitely many choices for $\mathcal{S}$, C_k has bounded diameter. It follows from the proof that there is an algorithm to find this diameter. □

Proof of Theorem 1.1 Theorem 1.1 follows immediately from Theorem 1.2 and the discussions in Section 2 and Section 3. By Corollary 3.5, there is surface S_i ($i = 1, 2$) properly embedded in M_i, such that S_i is either essential or ∂–strongly irreducible in M_i and the distance $d_{\mathcal{C}(F)}(\phi(\partial S_1), S_2)$ is at most $2g-2$, where $g = g(M_1) + g(M_2) - g(F)$. By [5] and a theorem in [2] (see the Appendix below for a workaround for [2]), S_i is isotopic to a normal or an almost normal surface for any 0–efficient triangulation of M_i, see Remark 3.6. Now we choose a 0–efficient triangulation for M_i and Theorem 1.1 follows from Theorem 1.2. □

Appendix

The purpose of this appendix is to address an issue in the proof of [2, Corollary 8.9]. While Bachman insists the proof is correct, there is a concern on the thin-position argument for manifolds with boundary in the proof of [2, Lemma 8.5]. The following is a workaround suggested by the referee.

Note that an essential surface is isotopic to a normal surface with respect to any triangulation, so the issue here is on ∂–strongly irreducible surfaces. Suppose S_1 is a strongly irreducible and ∂–strongly irreducible surface properly embedded in M_1 as in Lemma 3.3 and Corollary 3.5. It follows from the sweepout argument in [3] that S_1 is compressible on both sides (see Definition 3.2), and ∂S_1 consists of essential curves in ∂M_1 (see the proof of Corollary 3.4). Next we show that S_1 does not admit nested ∂–compressions.

Suppose S_1 admits nested ∂–compressions, then we can find a disk D such that $\partial D = \alpha \cup \beta$, $\alpha \subset \partial M_1$, $\beta \subset S_1$, and $\mathrm{int}(D) \cap S_1 \neq \varnothing$ consists of non-nested arcs. Let $\beta_1, \dots, \beta_k$ be the arcs of $\mathrm{int}(D) \cap S_1$ and δ_i $(i = 1, \dots, k)$ the subdisk of D bounded by β_i and a subarc of α. By our assumption, $\delta_i \cap \delta_j = \varnothing$ if $i \neq j$. We may assume that each δ_i is a ∂–compressing disk on the same side of S_1. Moreover, we may choose D so that $k > 0$ and k is minimal among all such disks D. Let $Q = D - \bigcup_{i=1}^{k} \mathrm{int}(\delta_i)$. Since S_1 is compressible on both sides, there is a compressing disk D' on the opposite side of δ_i or equivalently on the same side as Q. We may assume $D' \cap Q$ contains no closed curve. Since S_1 is ∂–strongly irreducible, $D' \cap \beta_i \neq \varnothing$ for each i. Let γ be an arc of $D' \cap Q$ that is outermost in D' and Δ the subdisk of D' cut off by γ with $\Delta \cap Q = \gamma$. The arc γ cuts Q into two disks Q_1 and Q_2. Thus either (1) $Q_i \cup \Delta$ $(i = 1$ or $2)$ is a compressing disk disjoint from some δ_i, a contradiction to the ∂–strong irreducibility, or (2) the union of $Q_i \cup \Delta$ $(i = 1$ or $2)$ and some δ_j's form a new disk similar to D, which contradicts the assumption that k is minimal. Thus S_1 does not admit nested ∂–compressions.

We call the two sides of S_1 plus and minus sides. By the definition of strongly irreducible surfaces (Definition 3.2), S_1 is compressible on both sides. If we perform a maximal compression on the plus side of S_1 and discard the closed surface components, then we get a surface S_1^+. Since S_1 is ∂–strongly irreducible, S_1^+ is incompressible and ∂–incompressible on the minus side. This basically follows from [4], see part (1) of [14, Lemma 5.5] for a proof for surfaces with boundary. Note that the proof of part (1) of [14, Lemma 5.5] does not mention ∂–compressing disks because the surface in [14] is strongly irreducible but may not be ∂–strongly irreducible. However, with the assumption of ∂–strong irreducibility, the same proof of [14] shows S_1^+ is also ∂–incompressible on the minus side. Thus either S_1^+ consists of ∂–parallel surfaces, or after some ∂–compressions on the plus side, S_1^+ becomes an essential surface S_1' in M_1 with $d(\partial S_1^+, \partial S_1') \leq -\chi(S_1^+)$. As the argument for essential surfaces does not use Bachman's theorem [2], we may assume S_1^+ is ∂–parallel. Since S_1 has no nested

∂–compressions, the ∂–parallel components of S_1^+ are not nested. We can also apply the same argument on the minus side of S_1. Therefore we may assume that S_1 is a boundary-Heegaard surface as in [2].

Next we explain the controversial part of [2] which is pointed out by the referee. The proof of the main theorem in [2] is basically a thin-position argument in which the 1–skeleton of the triangulation is in thin position with respect a sweepout $\{\mathcal{S}_t\}$ of a boundary-Heegaard surface. A problem arises when a thick level surface admits a high disk D' and a low disk D with $D \cap D' = \varnothing$. If both D and D' lie in the interior of M_1 then a simple isotopy as in [2, Figure 4] can reduce the width. The controversial part in [2] is the case that D' lies in the interior of M_1 and D has a boundary arc in ∂M_1. Note that one can assume that D is a ∂–compressing disk, since otherwise there is a low disk totally in ∂M_1 disjoint from D' and the usual isotopy can reduce the width. However, if D is a ∂–compressing disk, the usual width-reduction operation as above would be a ∂–compression along D, which is not an isotopy on the level surface any more. Below is a workaround for this situation.

We first glue a product $F \times I$ ($F = \partial M_1$) to M_1 and obtain a manifold M_1' ($M_1' \cong M_1$). We can extend S_1 to a surface S_1' properly embedded in M_1' by adding vertical annuli in $F \times I$ along ∂S_1. We fix a 0–efficient triangulation of M_1 and suppose $F \times I$ is not triangulated.

We consider a special sweepout or foliation $\{\mathcal{S}_t\}$ as in [2] with the restriction that for each regular leaf $\mathcal{S}_t$, $\mathcal{S}_t \cap (F \times I)$ is obtained by pushing pairwise disjoint ∂–compressing disks of S_1 (on the same side) into $F \times I$, and $\mathcal{S}_t \cap M_1$ is obtained by ∂–compressions on one side. Note that if a ∂–compression on S_1 yields a ∂–parallel disk component, we also push the disk component into $F \times I$.

We now apply the thin-position argument on $\{\mathcal{S}_t \cap M_1\}$ and assume the 1–skeleton is in thin position. Suppose a thick level $\mathcal{S}_t$ admits a pair of disjoint high and low disks in the 2–skeleton. Let D be the low disk as explained above and suppose $\partial D = \alpha \cup \beta$ with $\alpha \subset \mathcal{S}_t$ and $\beta \subset \partial M_1$. We may assume D is a ∂–compressing disk for $\mathcal{S}_t \cap M_1$. Since S_1 is ∂–strongly irreducible, the high disk lies in the interior of M_1. Thus we can perform an isotopy on the triangulation as in [2, Figure 4] by pushing the high disk down and the low disk up, which leads to a contradiction to the thin-position assumption. Note that the isotopy of pushing the low disk D into $F \times I$ can be viewed as a ∂–compression on $\mathcal{S}_t$ in M_1. Moreover, by the assumptions on S_1, after a ∂–compression on one side, there is no ∂–compressing disk in M_1 on the other side, hence all the ∂–compressions are on the same side.

Therefore the arguments in [2; 19] imply that one can isotope S_1' into a surface Σ' so that (1) $\Sigma' \cap (F \times I)$ is obtained by pushing pairwise disjoint ∂–compressing disks of

S_1 (on the same side) into $F \times I$, and $\Sigma = \Sigma' \cap M_1$ is obtained by ∂–compressions on one side, note that we also push the possible trivial disk components into $F \times I$, and (2) Σ is normal or almost normal with respect to the triangulation of M_1. Furthermore, after one more ∂–compression in a tetrahedron, we may assume the special type of almost normal pieces in Figure 9 of [2] does not appear. This implies that $\partial\Sigma$ is normal in ∂M_1.

Now we study the property of Σ and $\Sigma' \cap (F \times I)$. First, since all the ∂–compressions occur on the same side of S_1, the trivial-circle components of $\partial\Sigma$ (if any) are not nested. Since $\partial\Sigma$ is normal and by part (a) of Proposition 4.3, this implies that $\partial\Sigma$ contains at most one trivial-curve component. Note that $\partial\Sigma \neq \varnothing$ since F is incompressible and S_1 is a subsurface of a Heegaard surface.

If $\partial\Sigma$ contains at least one essential curve, then as before, the distance $d(\partial\Sigma, \partial S_1) < -\chi(S_1)$, viewed in the curve complex $\mathcal{C}(F)$. Now we can prove the main theorem by applying the arguments in Sections 4, 5 and 6 on Σ, see Remark 3.6 and the remark before Proposition 4.3.

Therefore we may suppose $\partial\Sigma$ is a single trivial vertex-linking curve in ∂M_1. Let δ be the disk bounded by $\partial\Sigma$ in ∂M_1 and $P = \Sigma' \cap (F \times I)$. So $P \cup \Sigma = \Sigma'$ and by the construction of Σ, $\delta \cup P$ is ∂–parallel in $F \times I$, in other words, $P = \Sigma' - \mathrm{int}(\Sigma)$ can be constructed by adding a vertical tube to a (once-punctured) ∂–parallel surface in $F \times I$.

Let $F_- = \partial M_1 - \mathrm{int}(\delta)$. There is a natural projection from the arc-and-curve complex $\mathcal{AC}(F_-)$ to the curve complex $\mathcal{C}(\partial M_1) = \mathcal{C}(F)$ denoted by $\pi\colon \mathcal{AC}(F_-) \to \mathcal{C}(F)$ as follows. We view $F_- = F - \mathrm{int}(\delta)$. For any closed essential curve γ in F_-, γ is also an essential curve in F, we set $\pi([\gamma]) = [\gamma]$. For any essential arc α in F_-, let $\hat{\alpha}$ be the closed curve obtained by connecting $\partial\alpha$ by an arc properly embedded in the disk δ. We define $\pi([\alpha]) = [\hat{\alpha}]$. Note that if two arcs $\alpha \cap \beta = \varnothing$ in F_-, then $\hat{\alpha} \cap \hat{\beta}$ is either empty or a single point. This means that if $d(\alpha, \beta) = 1$ in $\mathcal{AC}(F_-)$ then $d(\pi(\alpha), \pi(\beta)) = d(\hat{\alpha}, \hat{\beta}) \leq 2$ in $\mathcal{C}(F)$.

Note that the disk δ is a compressing disk for $S_1' = \Sigma'$. We denote the two sides of S_1' using plus and minus and suppose δ is on the plus side. Since S_1' is compressible on both sides, there is another compressing disk D on the minus side and $\partial\delta \cap \partial D \neq \varnothing$ in S_1'. Since Σ is obtained by ∂–compressions on the plus side, Σ is ∂–incompressible on the minus side and $D \cap \partial M_1 \neq \varnothing$. Moreover, ∂M_1 cuts D into a collection of subdisks and all the bigon disks lie in $F \times I$ (since Σ is ∂–incompressible on the minus side in M_1). Let D_1 be such a bigon subdisk of D and suppose $\partial D_1 = \alpha \cup \beta$, where $\alpha \subset \partial M_1$ and $\beta \subset P$. Let D_0 be the subdisk of D adjacent to D_1 with $D_0 \cap D_1 = \alpha$ and $D_0 \subset M_1$. Since P can be obtained by adding a vertical tube to a punctured ∂–parallel surface in $F \times I$, $\hat{\alpha}$ and $\partial S_1'$ project to disjoint curves in F, ie, $d(\pi(\alpha), \partial S_1') \leq 1$.

Note that Σ cuts M_1 into two submanifolds and we denote the one on the minus side by N. Clearly D_0 is a compressing disk for N. For any compressing disk Δ of N, since Σ is incompressible on the minus side, $\partial\Delta \cap \partial\delta \neq \varnothing$, ie, $\partial\delta$ is disk-busting in N. For any compressing disk Δ of N, we suppose $|\partial\Delta \cap \partial\delta|$ is minimal among all disks in the isotopy class of Δ. We fix an arc component γ_Δ of $\partial\Delta \cap \partial M_1$ for each Δ. Let $\mathcal{D}$ be the disk complex of ∂N (ie, curves of ∂N bounding compressing disks in N). Define a projection $\pi_A\colon \mathcal{D} \to \mathcal{AC}(F_-)$ as $\pi_A([\partial\Delta]) = [\gamma_\Delta]$. The following theorem in [11] was also independently proved by Masur and Schleimer.

Theorem ([11]) *Let N be as above, $\mathcal{D}$ the disk complex, and F_- a compact essential subsurface of ∂N. Suppose ∂F_- is disk-busting in ∂N. Then either*

(1) *N is an I–bundle of which F_- is a horizontal boundary component, or*

(2) *the image $\pi_A(\mathcal{D})$ of the disk complex has diameter at most 10 in $\mathcal{AC}(F_-)$ and $\pi \circ \pi_A(\mathcal{D})$ has diameter at most 20 in $\mathcal{C}(F)$.*

Note that part (a) of the theorem cannot happen in our case because otherwise one could isotope F to be disjoint from the Heegaard surface. Thus for any compressing disk Δ of N, $d(\hat{\gamma}_\Delta, \hat{\alpha}) \leq 20$ in $\mathcal{C}(F)$, where α is the arc $D_1 \cap D_0$ above. Moreover, since $d(\hat{\alpha}, \partial S_1) \leq 1$, we have $d(\hat{\gamma}_\Delta, \partial S_1) \leq 21$ for any compressing disk Δ of N.

Let Γ be the set of almost normal surfaces in M_1 such that for each surface X in Γ, ∂X is a vertex linking circle in ∂M_1 and $\chi(X) \geq \chi(S_1)$. As in Section 4 and Section 5, there is a finite collection of branched surfaces such that each surface in Γ is fully carried by a branched surface in the collection, and for each branched surface B, ∂B a single trivial circle in ∂M_1. For any surface X in Γ, ∂X bounds a disk δ in ∂M_1. Let N be the closure of the component of $M_1 - X$ that contains $F_- = \partial M_1 - \delta$. If X is fully carried by B, then N can be constructed by connecting some components of $M_1 - \mathrm{int}(N(B))$ using I–bundles. Although there may be infinitely many surfaces in Γ, since there are only finitely many branched surfaces and $\chi(X)$ is bounded, there are only finitely many possible topological types for N and we can list them all. For each possible N, we randomly find a compressing disk Δ for N and fix an arc γ_Δ of $\partial\Delta \cap F_-$. So we can construct finitely many closed curves $\hat{\gamma}_\Delta$. By the discussion above, if the gluing map $\phi\colon \partial M_1 \to \partial M_2$ is so complex that $d(\hat{\gamma}_\Delta, \partial S_1) > 21$ for each possible N, then it is impossible to have a surface S_1 with all the requirements. This implies that the original Heegaard surface cannot be strongly irreducible and the theorem follows.

References

[1] **I Agol**, **T Li**, *An algorithm to detect laminar 3-manifolds*, Geom. Topol. 7 (2003) 287–309 MR1988287

[2] **D Bachman**, *Heegaard splittings with boundary and almost normal surfaces*, Topology Appl. 116 (2001) 153–184 MR1855961

[3] **D Bachman**, **S Schleimer**, **E Sedgwick**, *Sweepouts of amalgamated 3-manifolds*, Algebr. Geom. Topol. 6 (2006) 171–194 MR2199458

[4] **A J Casson**, **C M Gordon**, *Reducing Heegaard splittings*, Topology Appl. 27 (1987) 275–283 MR918537

[5] **W Floyd**, **U Oertel**, *Incompressible surfaces via branched surfaces*, Topology 23 (1984) 117–125 MR721458

[6] **A E Hatcher**, *On the boundary curves of incompressible surfaces*, Pacific J. Math. 99 (1982) 373–377 MR658066

[7] **J Hempel**, *3-manifolds as viewed from the curve complex*, Topology 40 (2001) 631–657 MR1838999

[8] **W Jaco**, **J H Rubinstein**, *0-efficient triangulations of 3-manifolds*, J. Differential Geom. 65 (2003) 61–168 MR2057531

[9] **W Jaco**, **E Sedgwick**, *Decision problems in the space of Dehn fillings*, Topology 42 (2003) 845–906 MR1958532

[10] **M Lackenby**, *The Heegaard genus of amalgamated 3-manifolds*, Geom. Dedicata 109 (2004) 139–145 MR2113191

[11] **T Li**, *Images of the disk complex*, preprint Available at `http://www2.bc.edu/~taoli/image.pdf`

[12] **T Li**, *Heegaard surfaces and measured laminations. I. The Waldhausen conjecture*, Invent. Math. 167 (2007) 135–177 MR2264807

[13] **M Scharlemann**, *Heegaard splittings of compact 3-manifolds*, from: "Handbook of geometric topology", North-Holland, Amsterdam (2002) 921–953 MR1886684

[14] **M Scharlemann**, *Proximity in the curve complex: boundary reduction and bicompressible surfaces*, Pacific J. Math. 228 (2006) 325–348 MR2274524

[15] **M Scharlemann**, **A Thompson**, *Thin position for 3-manifolds*, from: "Geometric topology (Haifa, 1992)", Contemp. Math. 164, Amer. Math. Soc., Providence, RI (1994) 231–238 MR1282766

[16] **M Scharlemann**, **M Tomova**, *Alternate Heegaard genus bounds distance*, Geom. Topol. 10 (2006) 593–617 MR2224466

[17] **J Schultens**, *The classification of Heegaard splittings for (compact orientable surface)* $\times S^1$, Proc. London Math. Soc. (3) 67 (1993) 425–448 MR1226608

[18] **J Souto**, *The Heegaard genus and distances in the curve complex*, preprint

[19] **M Stocking**, *Almost normal surfaces in 3-manifolds*, Trans. Amer. Math. Soc. 352 (2000) 171–207 MR1491877

Department of Mathematics, Boston College
Chestnut Hill, MA 02467, USA

taoli@bc.edu

http://www2.bc.edu/~taoli

Received: 19 November 2006 Revised: 31 August 2007