CONNECTED SUMS OF UNSTABILIZED HEEGAARD SPLITTINGS ARE UNSTABILIZED

DAVID BACHMAN

ABSTRACT. Let M_1 and M_2 be closed, orientable 3-manifolds. Let H_i denote an unstabilized Heegaard surface in M_i. We prove that if $H_1 \# H_2$ comes from stabilizing a lower genus splitting of $M_1 \# M_2$ then either H_1 or H_2 comes from stabilizing a lower genus splitting. If H_i and G_i are non-isotopic Heegaard surfaces in M_i, and H_i is unstabilized, then we show $H_1 \# H_2$ is not isotopic to $G_1 \# G_2$ in $M_1 \# M_2$. The former result answers a question of C. Gordon (Problem 3.91 from [Kir97]).

Keywords: Heegaard Splitting, Incompressible Surface

1. INTRODUCTION

Suppose M_1 and M_2 are closed, orientable 3-manifolds and H_i is a Heegaard surface in M_i. Then one can form the connected sum $H = H_1 \# H_2$ in the 3-manifold $M = M_1 \# M_2$ to obtain a new Heegaard surface. If H_1, say, came from stabilizing some lower genus Heegaard surface in M_1 then it immediately follows that H comes from stabilizing a lower genus Heegaard surface in M. In 1997 C. McA. Gordon conjectured that the converse must also be true (see Problem 3.91 from [Kir97]):

Conjecture 1.1 (Gordon's Conjecture). *Suppose H_i is a Heegaard surface in M_i, for $i = 1, 2$. If $H_1 \# H_2$ is a stabilized Heegaard surface in $M_1 \# M_2$ then either H_1 or H_2 is stabilized.*

This paper contains a proof of Gordon's conjecture (Theorem 10.1). In a previous version of this paper a proof of the conjecture was announced with the additional assumption that M_1 and M_2 are irreducible. At the same time Ruifeng Qui announced a complete proof of Gordon's conjecture [Qiu]. The assumption of irreducibility is not used in the present version of this paper. In addition, we prove the following:

Theorem 10.2. *Let M_1 and M_2 be two closed, orientable 3-manifolds. Suppose H_i and G_i are non-isotopic Heegaard splittings of M_i and H_i is unstabilized. Then $H_1 \# H_2$ is not isotopic to $G_1 \# G_2$ in $M_1 \# M_2$.*

The surfaces H_1 and H_2 of Gordon's conjecture together form a structure called a *generalized Heegaard splitting (GHS)*. Loosely speaking, a GHS H is a pair of sets of surfaces, $\text{Thick}(H)$ and $\text{Thin}(H)$, such that each element of $\text{Thick}(H)$ is a Heegaard surface for some component of $M - \text{Thin}(H)$ (see Section 4 for a more precise definition).

We can transform one GHS into another by a process called *weak reduction*. This can be done whenever there are disjoint compressing disks on opposite sides of some thick surface (see Section 5). So, for example, if some thick surface of a GHS comes from stabilization or connected sum then there is a weak reduction for the entire GHS. Definition 5.5 gives a simple complexity for GHSs under which weak reduction represents a decrease.

Once we can relate GHSs by weak reduction we can start examining an entire *Sequence of GHSs (SOG)*. This is defined to be a sequence $\{H^i\}$ such that for each i either the GHS H^{i+1} or the GHS H^i can be obtained from the other by weak reduction. Now, given fixed GHSs H and H' and an SOG which connects them, one can ask if there is some sense in which there is a "more efficient" SOG which connects them. In Section 8 we define several ways to find such a new SOG. Any SOG obtained by one of these operations is said to have been obtained by a *reduction*. If a simpler SOG cannot be found then the given one is said to be *irreducible*.

Section 8 concludes with a crucial result about irreducible SOGs. This is given by Lemma 8.9, which states that the thick surfaces of the maximal GHSs of an irreducible SOG satisfy one of two combinatorial conditions. These conditions are called *strong irreducibility* and *criticality*. Strongly irreducible Heegaard splittings were introduced by Casson and Gordon in [CG87]. Criticality was introduced by the author in [Bac02], although the definition given here in Section 3 is considerably simpler. By Lemma 4.7 any GHS whose thick surfaces are strongly irreducible or critical must be of an irreducible 3-manifold. This lemma relies on a deep result about the intersections of strongly irreducible and critical surfaces with incompressible surfaces (Lemma 3.5).

Our proof of Gordon's conjecture (Theorem 10.1) begins with the construction of a SOG $\{H^i\}_{i=1}^n$ as follows. Each element of the set $\text{Thick}(H^1)$ is a Heegaard surface of an irreducible 3-manifold. For some $k \geq 1$ the set $\text{Thick}(H^k)$ has two elements, namely H_1 and H_2,

where H_i is the connected sum of a subset of $\text{Thick}(H^1)$ for $i = 1, 2$. The set $\text{Thick}(H^{k+1})$ consists of a single element, $H = H_1 \# H_2$. We assume that H destabilizes to a Heegaard surface G, and define H^{k+2} to be the GHS with $\text{Thick}(H^{k+2}) = \{G\}$. Finally, H^n is a GHS where again each element of $\text{Thick}(H^n)$ is a Heegaard surface in an irreducible 3-manifold, and G is the connected sum of these surfaces. It follows that the sum of the genera of the thick surfaces of H^n is strictly less than the sum of the genera of the thick surfaces of H^1.

The SOG $\{H^i\}$ thus defined has a single maximal GHS, H^{k+1}. Since this is not a GHS of an irreducible 3-manifold, there must be a more efficient SOG from H^1 to H^n. In other words, there is some reduction for $\{H^i\}$. Reducing as much as possible yields a SOG in which every element is a GHS of an irreducible 3-manifold. Finally, it follows from the assumption that H_1 and H_2 are unstabilized that there are no destabilizations in this final SOG. This gives our contradiction, as it implies that the sum of the genera of the thick sufaces of H^1 is equal to the sum of the genera of the thick surfaces of H^n.

The author would like to thank Saul Schleimer for helpful conversations regarding the proof of Claim 3.10, and the referee for several extremely helpful suggestions, including many of the examples given in the paper.

2. Definitions.

2.1. Essential loops, disks and spheres. A 2-sphere in a 3-manifold which does not bound a 3-ball is called *essential*. If a manifold does not contain an essential 2-sphere then it is referred to as *irreducible*.

A loop on a surface is called *essential* if it does not bound a disk in the surface. An arc which is properly embedded in a surface F is *essential* if it does not cobound, with a subarc of ∂F, a subdisk of F.

Suppose F is a surface embedded in a 3-manifold M, D is a disk in M, and $D \cap F = \partial D$. There is an embedding $h : D \times I \to M$ such that $h(D \times \{\frac{1}{2}\}) = D$ and $h(D \times I) \cap F = h(\partial D \times I)$. To *surger* F *along* D is to remove $h(\partial D \times I)$ from F and replace it with $h(D \times \partial I)$. We denote the result of such a surgery as F/D. As a shorthand we will denote $(F/D)/E$ as F/DE when D and E are disjoint disks with boundary on F.

If ∂D is an essential loop on F then D is referred to as a *compressing disk for* F, and surgery along D is referred to as *compression*. A surface F is said to be *incompressible* if there are no compressing disks for F. A properly embedded disk in a 3-manifold M is *essential* if it is a compressing disk for ∂M.

Now suppose F is a properly embedded surface in a 3-manifold M with boundary, and D is a disk such that $\partial D = \alpha \cup \beta$, $F \cap D = \alpha$ is an arc on F, and $D \cap \partial M = \beta$. Then there is an embedding $h : D \times I \to M$ such that $h(D \times \{\frac{1}{2}\}) = D$ and $h(D \times I) \cap F = h(\alpha \times I)$ and $h(D \times I) \cap \partial M = h(\beta \times I)$. To *surger F along D* is to remove $h(\alpha \times I)$ from F and replace it with $h(D \times \partial I)$. If, furthermore, α is an essential arc on F then D is referred to as a *∂-compressing disk for F*, and surgery along D is referred to as *∂-compression*. A surface F is said to be *∂-incompressible* if there are no ∂-compressing disks for F.

If M_1 and M_2 are n-manifolds then the connected sum, denoted $M_1 \# M_2$, is constructed as follows. First, obtain M_i^* by removing the interior of an n-ball from the interior of M_i. Each M_i^* will thus have a new $(n-1)$-sphere boundary component, S_i. The manifold $M_1 \# M_2$ is then obtained by identifying S_1 with S_2. The image of S_i in $M_1 \# M_2$ is referred to as the *summing sphere*. Note that the summing sphere is essential if and only if neither M_i is an n-sphere.

2.2. Heegaard splittings.

Definition 2.1. A *compression body* is a 3-manifold which can be obtained by starting with some closed, orientable, connected surface, H, forming the product $H \times I$, attaching some number of 2-handles to $H \times \{1\}$, and capping off all resulting 2-sphere boundary components with 3-balls. The boundary component $H \times \{0\}$ is referred to as ∂_+. The rest of the boundary is referred to as ∂_-.

Definition 2.2. A *Heegaard splitting* of a 3-manifold M is an expression of M as a union $\mathcal{V} \cup_H \mathcal{W}$, where $\mathcal{V}$ and $\mathcal{W}$ are compression bodies that intersect in an oriented surface $H = \partial_+ \mathcal{V} = \partial_+ \mathcal{W}$. If $\mathcal{V} \cup_H \mathcal{W}$ is a Heegaard splitting of M then we say H is a *Heegaard surface*.

Note. The assumption that H is oriented in the above definition will play an important role. For example, $L(p,q)$, where $q \neq \pm 1 \mod p$, contains a pair of non-isotopic Heegaard tori [BO83], but as unoriented surfaces they are isotopic. In contrast, S^3 contains a unique Heegaard torus [Wal68], namely the boundary of a regular neighborhood of an unknotted loop. Hence, we may talk about the *standard* genus one Heegaard surface in S^3.

Definition 2.3. Suppose M_1 and M_2 are 3-manifolds and H_i is a Heegaard surface in M_i. Recall that the first step in defining the connected sum $M_1 \# M_2$ is to remove the interior of a ball B_i from M_i, resulting in a new 2-sphere boundary component S_i of the punctured 3-manifold M_i^*. If B_i is chosen to meet H_i in a disk then $H_i^* = H_i \cap M_i^*$ will divide

S_i into disks D_i and D_i'. Now the identification of S_1 with S_2 can be done in two ways; D_1 is glued to D_2 or to D_2'. However, only one such identification will make the orientation of H_1^* agree with that of H_2^*. When this identification is used the surface $H_1^* \cup H_2^*$ is referred to as the connected sum $H_1 \# H_2$ of H_1 and H_2.

Example 2.4. Let H_1 and H_2 be Heegaard tori in $L(p,q)$. Let $\overline{H_1}$ denote H_1 with the opposite orientation. Then $H_1 \# H_2$ and $\overline{H_1} \# H_2$ are non-isotopic genus two Heegaard surfaces in $L(p,q) \# L(p,q)$, even as unoriented surfaces [Eng70] (see also [Bir74]).

Definition 2.5. A *stabilization* of a Heegaard surface H is a new Heegaard surface which is the connected sum of H with the standard genus one Heegaard surface in S^3.

3. STRONG IRREDUCIBILITY AND CRITICALITY

The main technical tools of this paper are *strongly irreducible* [CG87] and *critical* [Bac02] surfaces. Both strong irreducibilty and criticality are combinatorial conditions satisfied by the compressing disks for a Heegaard surface.

Definition 3.1. Let $\mathcal{V} \cup_H \mathcal{W}$ be a Heegaard splitting of a 3-manifold M. Then we say the pair (V, W) is a *reducing pair* for H if V and W are disjoint compressing disks on opposite sides of H.

Definition 3.2. A Heegaard surface is *strongly irreducible* if it is compressible to both sides but has no reducing pairs.

Definition 3.3. Let H be a Heegaard surface in some 3-manifold which is compressible to both sides. The surface H is *critical* if the set of all compressing disks for H can be partitioned into subsets C_0 and C_1 such that

  (1) For each $i = 0, 1$ there is at least one reducing pair (V_i, W_i), where $V_i, W_i \in C_i$.
  (2) If $V \in C_0$ and $W \in C_1$ then (V, W) is not a reducing pair.

Definition 3.3 is significantly simpler, and slightly weaker, than the one given in [Bac02]. In other words, anything that was considered critical in [Bac02] is considered critical here as well. Hence, a result such as Theorem 7.1 of [Bac02] still holds. This result says that in a non-Haken 3-manifold the minimal genus common stabilization of any pair of non-isotopic, unstabilized Heegaard splittings is critical. The basic idea of the proof is as follows: suppose H_0 and H_1 are non-isotopic Heegaard splittings in a 3-manifold M which are isotopic to a surface

K after one stabilization. Another way to say this is that there are reducing pairs (V_0, W_0) and (V_1, W_1) representing destabilizations of K that lead to H_0 and H_1. We then show that either we can use the disks V_i and W_i to create a partition of the compressing disks for K that satisfies the conditions of Definition 3.3, or there is an incompressible surface in M.

Example 3.4. Let M be a Seifert fibered space which fibers over the sphere with three exceptional fibers. There are three vertical splittings H, H', and H'' of M and these are generally not isotopic (see [MS98] for the relevant definitions). Let K be the genus three splitting which is the common stabilization of these three. Since M is non-Haken, it follows from Theorem 7.1 of [Bac02] that K is critical.

Example 3.4 shows that the partition of disks into just two sets in Definition 3.3 is a bit misleading. For a critical Heegaard surface in a non-Haken 3-manifold one can make a partition with a set for each distinct destabilization. This point is made more explicit in the definition of criticality given in [Bac02]. Somehow the useful results (such as Lemma 3.5 below) about critical surfaces only require a partition with at least two sets. This is one reason for the more streamlined definition given here.

Lemma 3.5. *Suppose H is an incompressible, strongly irreducible, or critical Heegaard surface in a 3-manifold M and S is an essential sphere or disk in M. Then there is an essential surface S', obtained from S by surgery, such that $S' \cap H = \emptyset$.*

Proof. We break the proof into three cases, depending on whether H is incompressible, strongly irreducible, or critical.

Case 1. *H is incompressible.* The incompressible case follows from a standard innermost disk argument. We leave the proof to the reader.

Case 2. *H is strongly irreducible.* The strongly irreducible case also follows from standard arguments. For completeness we present a proof here.

Let H be a strongly irreducible Heegaard surface, separating M into $\mathcal{V}$ and $\mathcal{W}$. By the definition of strong irreducibility there are compressing disks $V \subset \mathcal{V}$ and $W \subset \mathcal{W}$ for H. A neighborhood of V is homeomorphic to a 3-ball, and hence we can define an isotopy which pushes S off V. Similarly, we can define an isotopy of S which pushes S off of W. Putting such isotopies together gives us an isotopy S_t (*i.e.* a map $\gamma : S \times [-1, 1] \to M$ where $S_t = \gamma(S, t)$) such that

- $S_{-1} \cap V = \emptyset$

- $S_0 = S$
- $S_1 \cap W = \emptyset$

Let $t_0 = -1$, $\{t_i\}_{i=1}^{n-1}$ be the set of points in $[-1, 1]$ where S_t is not transverse to H, and $t_n = 1$. These points break $[-1, 1]$ up into subintervals, which we now label. If there is a $t \in (t_i, t_{i+1})$ such that $H \cap S_t$ contains the boundary of a compressing disk for H in $\mathcal{V}$ then this interval gets the label $\mathbb{V}$. Similarly, if there is a $t \in (t_i, t_{i+1})$ such that $H \cap S_t$ contains the boundary of a compressing disk for H in $\mathcal{W}$ then this interval gets the label $\mathbb{W}$.

We now present several claims which produce the desired result.

Claim 3.6. *The interval (t_0, t_1) is either labeled $\mathbb{V}$ or has no label. Similarly, the interval (t_{n-1}, t_n) is labeled $\mathbb{W}$ or has no label.*

Proof. For t near -1 the surface S_t is disjoint from V. Suppose the interval containing t is labeled $\mathbb{W}$. Then there is a loop $\alpha \subset H \cap S_t$ which bounds a compressing disk W' for H in $\mathcal{W}$. But then $\partial W' \cap \partial V = \emptyset$, contradicting the strong irreducibility of H. A symmetric argument completes the proof. $\square$

Claim 3.7. *No interval has both labels.*

Proof. If, for some t, there are loops of $S_t \cap H$ that bound disks in $\mathcal{V}$ and $\mathcal{W}$ then we immediately contradict the strong irreducibility of H. $\square$

Claim 3.8. *An interval labeled $\mathbb{V}$ cannot be adjacent to an interval labeled $\mathbb{W}$.*

Proof. Suppose t_i is the intersection point of adjacent intervals with different labels. Let $t_- = t_i - \epsilon$ and $t_+ = t_i + \epsilon$. As H is orientable the loops of $S_{t_-} \cap H$ can be made disjoint (on H) from the loops of $S_{t_+} \cap H$ (see, for example, Lemma 4.4 of [Gab87]). Hence, if $S_{t_-} \cap H$ contains the boundary of a compressing disk for H in $\mathcal{V}$ it can be made disjoint from the boundary of a compressing disk in $\mathcal{W}$ that is contained in $S_{t_+} \cap H$. This again contradicts the strong irreducibility of H. $\square$

Following these claims we conclude there is an unlabeled interval. Henceforth we assume t is in such an interval. We now claim that every loop of $S_t \cap H$ is inessential on H. Suppose this is not the case. Let α be a loop of $S_t \cap H$ which is innermost on S_t among all such loops that are also essential on H.

Claim 3.9. *The loop α bounds a compressing disk for H.*

Proof. As S is a sphere or disk the loop α bounds a subdisk A of S_t so that all curves of $int(A) \cap H$ are inessential on H. If the interior

of A misses H then A is a compressing disk for H. If not then there is some loop β where the interior of A meets H. By assumption β is inessential on H. Hence, β bounds a subdisk B of H. Let γ denote a loop of $A \cap B$ which is innermost on B. Then γ bounds a subdisk C of H whose interior is disjoint from A. We now use C to surger A, removing one loop of $A \cap H$. Continuing in this way we may remove all loops of $A \cap H$, besides ∂A. The resulting disk is a compressing disk for H. $\square$

By the previous claim the loop α bounds a compressing disk D for H. This contradicts the fact that t lies in an unlabeled interval. We conclude $S_t \cap H$ consists of loops that are inessential on both surfaces. A standard innermost disk argument now completes the proof of the strongly irreducible case of Lemma 3.5.

Case 3. H *is critical.* In this case Lemma 3.5 essentially follows from Theorem 5.1 of [Bac02]. As we are using a slightly weaker definition of the term "critical" here, we reproduce the proof. Henceforth we will assume that H is a critical Heegaard surface which separates M into compression bodies $\mathcal{V}$ and $\mathcal{W}$. Let C_0 and C_1 be the sets in Definition 3.3. By definition there are compressing disks $V_i \subset \mathcal{V}$ and $W_i \subset \mathcal{W}$ where (V_i, W_i) is a reducing pair and $V_i, W_i \in C_i$.

The proof is in several stages. First, we construct a map Φ from $S \times D^2$ into M. We then use Φ to break up D^2 into regions and label them in such a way so that if any region remains unlabelled then the conclusion of the lemma follows. Finally, we construct a map from D^2 to a labelled 2-complex Π which has non-trivial first homology, and show that if there is no unlabelled region then the induced map on homology is nontrivial, a contradiction. This general strategy is similar to that used in [RS96], although the details have little in common.

3.1. **Constructing the map, $\Phi : S \times D^2 \to M$.**

We begin by defining a two parameter family of surfaces in M isotopic to S. For any map $\Phi : S \times D^2 \to M$ we let S_x denote the image of $\Phi(S, x)$.

Claim 3.10. *Let θ_0 and θ_1 be distinct points on ∂D^2. Let U and L be arcs of ∂D^2 such that $\partial D^2 = U \cup L$, where $U \cap L = \theta_0 \cup \theta_1$. There is a continuous map $\Phi : S \times D^2 \to M$ such that*

- *for all $x \in D^2$ the surface S_x is embedded,*
- *for $i = 0, 1$ the surface S_{θ_i} is disjoint from both V_i and W_i,*
- *for each $\theta \in U$ the surface S_θ is disjoint from at least one compressing disk for H in $\mathcal{V}$, and*

- *for each $\theta \in L$ the surface S_θ is disjoint from at least one compressing disk for H in $\mathcal{W}$.*

Proof. We start by inductively defining a sequence of compressing disks for H, $\{V^i\}_{i=0}^n$, such that $V^i \cap V^{i+1} = 0$, for all i between 0 and $n-1$, $V^0 = V_0$, and $V^n = V_1$.

(1) Define $V^0 = V_0$.
(2) Let V^i denote the last disk defined, and suppose there is a simple closed curve in $V_1 \cap V^i$. Let v denote an innermost subdisk of V_1 bounded by a loop of $V_1 \cap V^i$. Now surger V^i along v. The result is a disk and a sphere. Throw away the sphere and denote the disk as V^{i+1}. Note that $|V_1 \cap V^{i+1}| < |V_1 \cap V^i|$.
(3) Let V^i denote the last disk defined, and suppose there are only arcs in $V_1 \cap V^i$. Let v denote an outermost subdisk of V_1 cut off by an arc of $V_1 \cap V^i$. Now surger V^i along v. The result is two disks, at least one of which is a compressing disk for H. Call such a disk V^{i+1}. Again, note that $|V_1 \cap V^{i+1}| < |V_1 \cap V^i|$.
(4) If neither of the previous two cases apply then we have arrived at a disk V^{n-1} such that $V^{n-1} \cap V_1 = \emptyset$. Now let $V^n = V_1$ and we are done.

We can apply a symmetric construction to produce a sequence of compressing disks, $\{W^j\}_{j=0}^m$, such that $W^0 = W_0$, $W^m = W_1$, and $W^j \cap W^{j+1} = \emptyset$, for all j between 0 and $m-1$.

The map Φ can now be described as follows. See Figure 1. For x at the center of D^2 the surface S_x is identical to S. Near θ_0 the surface S_x is disjoint from both V^0 and W^0. As x progresses along U toward θ_1 the surface S_x ceases to be disjoint from W^0, but becomes disjoint from V^1. Progressing further the surface S_x ceases to be disjoint from V^0, but becomes disjoint from V^2. This continues until x gets to θ_1, when S_x is disjoint from both V^n and W^m. This is illustrated for $n = 2$ and $m = 1$ in Figure 1.

To rigorously define Φ requires a considerable amount of further work (and, unfortunately, notation). For each i between 0 and n let A^i denote a neighborhood of the disk V^i. For each j between 0 and m let B^j denote a neighborhood of the disk W^j. Because $V^i \cap V^{i+1} = W^j \cap W^{j+1} = V^0 \cap W^0 = V^n \cap W^m = \emptyset$ we may assume $A^i \cap A^{i+1} = B^j \cap B^{j+1} = A^0 \cap B^0 = A^n \cap B^m = \emptyset$.

For each i between 0 and n, let $\gamma^i : \overline{A^i} \times I \to \overline{A^i}$ be an isotopy which pushes S off of V^i. In other words, $\gamma^i(x,0) = x$ for all $x \in \overline{A^i}$, $\gamma^i(S,1) \cap V^i = \emptyset$, and $\gamma^i(x,t) = x$ for all $x \in \partial \overline{A^i}$. Similarly, for each j between 0 and m, let δ^j be an isotopy which pushes S off of W^j, inside B^j.
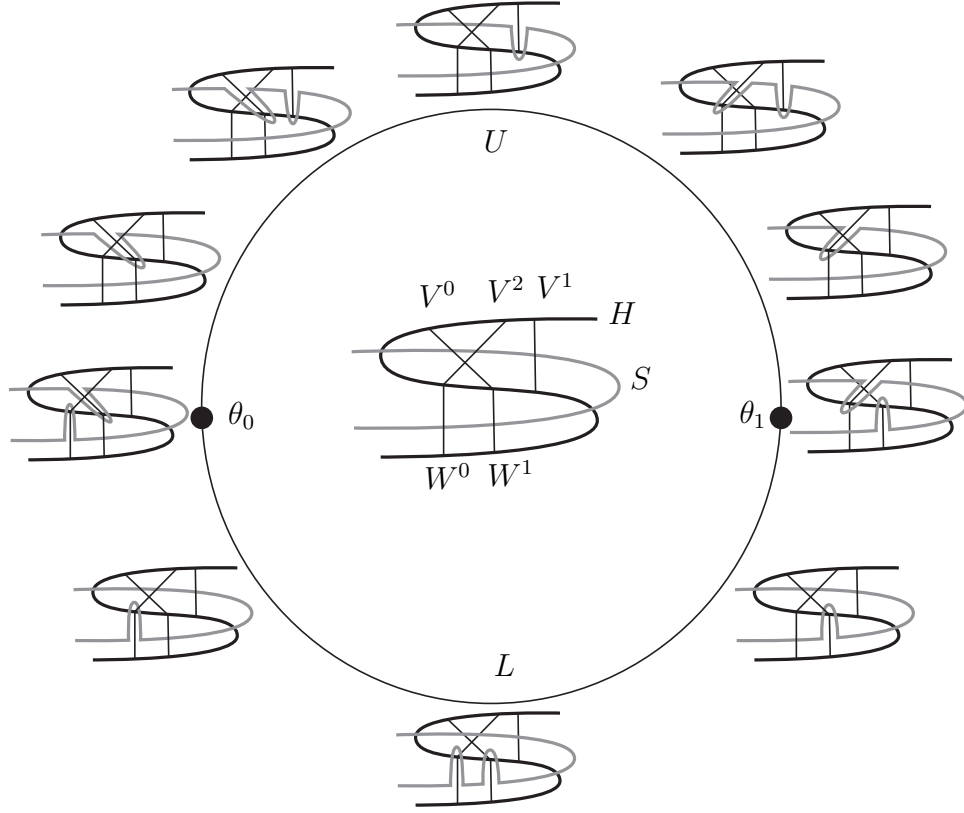
$$\text{FIGURE 1}$$

Choose n pairs of circular arcs centered on points of U, and m pairs of arcs centered on points of L, "linked" as in Figure 2[1]. For the ith pair of arcs chosen, centered on a point of U, define $f^i : D^2 \to [0,1]$ to be the continuous function depicted in Figure 3. Let $g^j : D^2 \to [0,1]$ be the function similarly defined for the jth pair of arcs centered on a point of L.

Finally, we define $\Phi : S \times D^2 \to M$. Suppose $x \in S$ and $p \in D^2$. If f^i is non-zero at p and $x \in A^i$ then we define $\Phi_p(x) = \gamma^i(x, f^i(p))$, where $\Phi_p(x) = \Phi(x,p)$. Similarly, if g^j is non-zero at p and $x \in B^j$ then we define $\Phi_p(x) = \delta^j(x, g^j(p))$. If x is a point of S not in such an A^i or B^j then we define $\Phi_p(x) = x$.

The proof is now complete by making the following observations:

- If f^i and f^k are non-zero at p (where $k > i$) then $k = i+1$ and $g^j(p) = 0$ for all j. Since A^i and A^{i+1} are disjoint the function
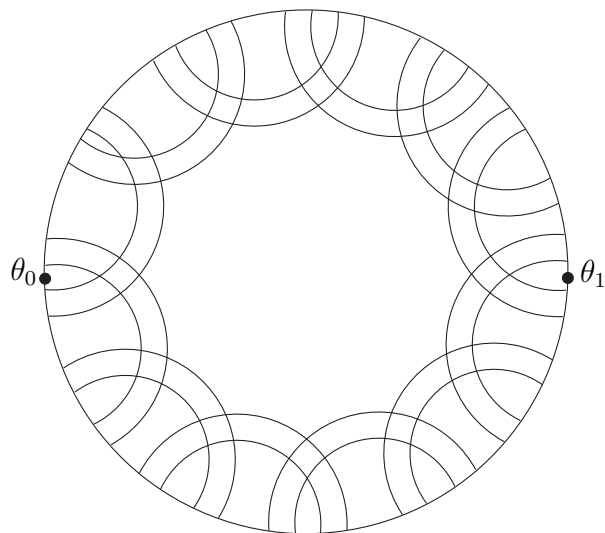
---

[1]Figure suggested by Saul Schleimer.
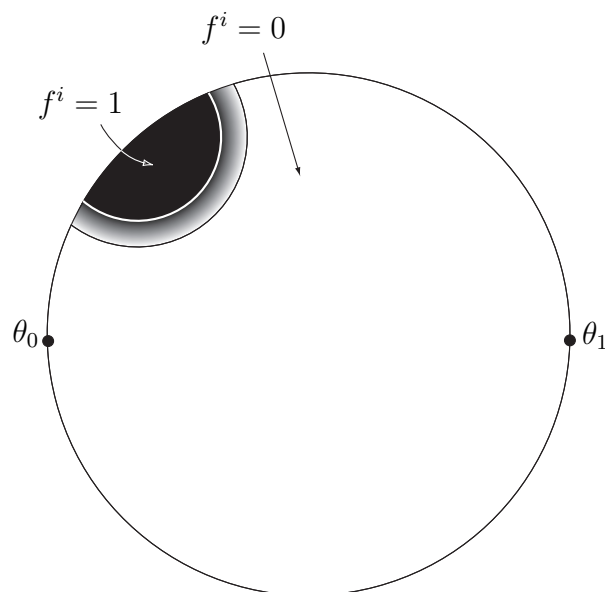
FIGURE 2. "Linked" pairs of circular arcs in D^2.

FIGURE 3. In the black region f^i takes on the value 1. In the white region $f^i = 0$. The shading between these two regions is meant to indicate that between the arcs the function f^i continuously varies from 0 to 1.

Φ_p is well defined. (Ambient isotopies with disjoint supports commute.)

- A similar statement holds if g^j and g^l are non-zero at p.
- If f^i and g^j are non-zero at p then either $i = j = 0$ or $i = n$ and $j = m$. Since $A^0 \cap B^0 = A^n \cap B^m = \emptyset$ the function is again well defined.
- For p near the center of D^2 the function Φ_p is the identity on S.
- At each $p \in \partial D^2$ at least one of the functions $\{f^i\}$ or $\{g^j\}$ is 1. Assume this is true of $f^i(p)$. If $x \in S \cap A^i$ then $\Phi_p(x) = \gamma^i(x, 1)$. Note that this is disjoint from the disk V^i. Hence, for every point $p \in \partial D^2$ the surface S_p is disjoint from at least one of the disks $\{V^i\}$ or $\{W^j\}$.
- At θ_0 both f^0 and g^0 are 1. Hence, the surface S_{θ_0} is disjoint from both $V^0 = V_0$ and $W^0 = W_0$. Similarly, at θ_1 both f^n and g^m are 1. Hence, the surface S_{θ_1} is disjoint from both $V^n = V_1$ and $W^m = W_1$.

$\square$

We now perturb Φ slightly so that it is smooth and in general position with respect to H, and again denote the new function as Φ. Consider the set $\Sigma = \{x \in D^2 | S_x \text{ is not transverse to } H\}$. If Φ is in general position with respect to H, then Cerf theory (see [Cer68]) tells us that Σ is homeomorphic to a graph, and the maximum valence of each vertex of this graph is 4. We will use these facts later.

3.2. Labelling D^2.

A *region* of D^2 is a component of $D^2 - \Sigma$. Let x be any point in some region. We will label this region from the set $\{\mathbb{V}_0, \mathbb{W}_0, \mathbb{V}_1, \mathbb{W}_1\}$ as follows. The region containing x will have the label

- $\mathbb{V}_0$ if there is a loop of $H \cap S_x$ which bounds a compressing disk $V \subset \mathcal{V}$ for H such that $V \in C_0$.
- $\mathbb{W}_0$ if there is a loop of $H \cap S_x$ which bounds a compressing disk $W \subset \mathcal{W}$ for H such that $W \in C_0$.
- $\mathbb{V}_1$ if there is a loop of $H \cap S_x$ which bounds a compressing disk $V \subset \mathcal{V}$ for H such that $V \in C_1$.
- $\mathbb{W}_1$ if there is a loop of $H \cap S_x$ which bounds a compressing disk $W \subset \mathcal{W}$ for H such that $W \in C_1$.

Claim 3.11. *If some region is unlabeled then the conclusion of Lemma 3.5 follows.*

Proof. If any region remains unlabeled then there is no loop of $H \cap S_x$ which bounds a compressing disk for H. We claim that in such a

situation every loop of $H \cap S_x$ is inessential on both surfaces, and hence we can remove all intersections by a sequence of surgeries, using a standard innermost disk argument. The conclusion of Lemma 3.5 thus follows.

Suppose $H \cap S_x$ contains a loop γ which is essential on H. As S is a sphere or disk the loop γ bounds a subdisk D of S. Let γ' denote a loop of $H \cap S_x$ which is innermost on D among all loops that are essential on H (possibly $\gamma' = \gamma$). The loop γ' thus bounds a subdisk D' of S whose interior meets H in a collection of loops that are inessential on both surfaces. Hence we may do a sequence of surgeries on D' to obtain a compressing disk for H. It would follow that the region containing x has a label. $\qquad\square$

Claim 3.12. *No region can have both of the labels $\mathbb{V}_i$ and $\mathbb{W}_{1-i}$.*

Proof. Let x be a point in a region with the labels $\mathbb{V}_0$ and $\mathbb{W}_1$. Let $V \subset \mathcal{V}$ and $W \subset \mathcal{W}$ be disks whose existence is implied by these labels. Hence, $V \in C_0$ and $W \in C_1$. But both ∂V and ∂W are contained in $H \cap S_x$. Thus they are either disjoint or equal (in which case they can be made disjoint). This contradicts the definition of C_0 and C_1, as then (V, W) would be a reducing pair. $\qquad\square$

Claim 3.13. *If a region has the label $\mathbb{V}_i$ then no adjacent region can have the label $\mathbb{W}_{1-i}$.*

The proof is similar to the argument of Gabai used in Claim 3.8.

Proof. Suppose the region $\mathcal{R}_0$ has the label $\mathbb{V}_0$ and $\mathcal{R}_1$ is an adjacent region with the label $\mathbb{W}_1$. Let x_i be some point in $\mathcal{R}_i$. Let $p : I \to D^2$ be an embedded path connecting x_0 to x_1, which does not wander into any region other than $\mathcal{R}_0$ or $\mathcal{R}_1$. The fact that $\mathcal{R}_0$ has the label $\mathbb{V}_0$ implies that for each t for which $p(t)$ is in $\mathcal{R}_0$ there is a compressing disk $V_t \subset \mathcal{V}$ for H such that $V_t \in C_0$ and $\partial V_t \subset H \cap S_{p(t)}$. Similarly, for each t for which $p(t)$ is in $\mathcal{R}_1$ there is a compressing disk $W_t \subset \mathcal{W}$ for H such that $W_t \in C_1$ and $\partial W_t \subset H \cap S_{p(t)}$.

As t increases from 0 to 1, we see a moment, t_*, when $S_{p(t_*)}$ does not meet H transversely (*i.e.* t_* corresponds to the place where the path p crosses an edge of Σ). At t_* we simultaneously see the disappearance of ∂V_t, and the appearance of ∂W_t. (Otherwise, $\mathcal{R}_1$ would have both the labels $\mathbb{V}_0$ and $\mathbb{W}_1$, which we ruled out in Claim 3.12.) We conclude that as t approaches t_* from below we see ∂V_t become tangent to itself, or to another loop of $H \cap S_{p(t)}$. Similarly, as t approaches t_* from above we see ∂W_t become tangent to itself, or to another loop of $H \cap S_{p(t)}$. As only one such tangency occurs for each t on an edge of Σ we see

that

$$\lim_{t \to t_*^-} \partial V_t \cap \lim_{t \to t_*^+} \partial W_t \neq \emptyset$$

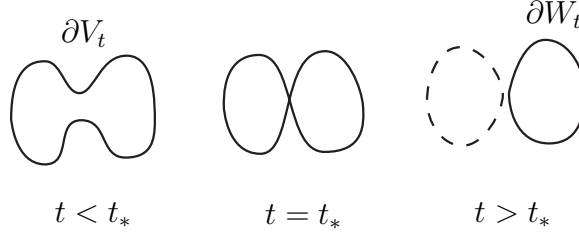as in Figure 4.

$$t < t_* \qquad t = t_* \qquad t > t_*$$

FIGURE 4

Since V_t and W_t are on opposite sides of H, we see from Figure 4 that $\partial V_{t_*-\epsilon}$ can be made disjoint from $\partial W_{t_*+\epsilon}$ (since H is orientable). This contradicts the fact that $V_t \in C_0$ and $W_t \in C_1$. □

We now assume, to obtain a contradiction, that there are no unlabeled regions.

3.3. The 2-complex Π and a map from D^2 to Π.

Let Π be the labelled 2-complex depicted in Figure 5. Let Σ' be the dual graph of Σ. Map each vertex of Σ' to the point of Π with the same label(s) as the region of D^2 in which it sits. Claim 3.12 assures that this map is well defined on the vertices of Σ'.

Similarly, map each edge of Σ' to the 1-simplex of Π whose endpoints are labelled the same. Claim 3.13 guarantees that this, too, is well defined.

Claim 3.14. *The map to Π extends to all of D^2.*

Proof. Note that the maximum valence of a vertex of Σ is four. Hence, the boundary of each region in the complement of Σ' gets mapped to a 1-cycle with at most four vertices in Π. Inspection of Figure 5 shows that there is only one such cycle which is not null homologous. Hence, we must rule out the possibility that there are four regions around a common vertex x_* of Σ, each with only one label, where all such labels are distinct. As in the proof of Claim 3.13 this implies that each edge of Σ incident to x_* corresponds to a saddle tangency. Hence, $S_{x_*} \cap H$ is a graph with exactly two valence four vertices and simple closed curves.

For x in the interior of a region which meets x_* the set $S_x \cap H$ is obtained from $S_{x_*} \cap H$ by some resolution of its vertices (see Figure 6). There are exactly four possible ways to resolve two vertices. As
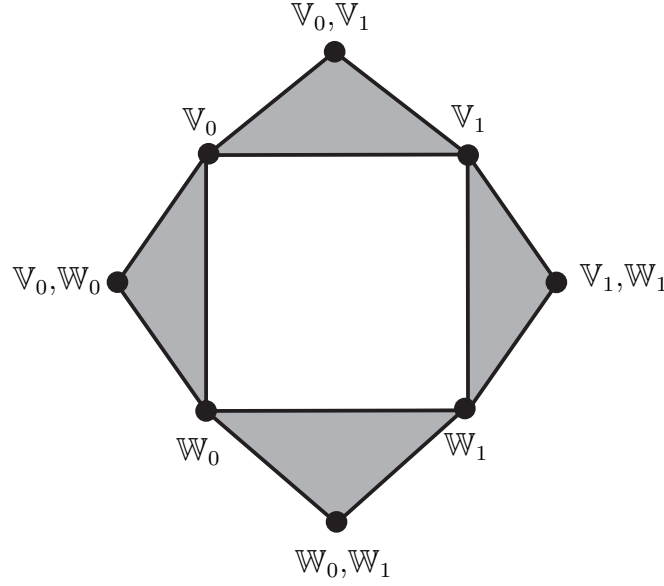
FIGURE 5. The 2-complex, Π.
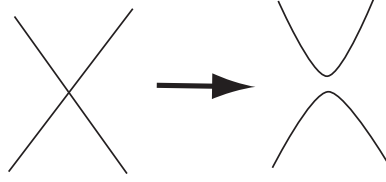
FIGURE 6. Resolving a vertex of $S_{x_*} \cap H$.

there are four regions with different labels around x_* we must see all four resolutions. However, the orientability of H guarantees that *some* resolution will consist of loops that can be made disjoint from all components of all other resolutions. There are now four symmetric cases. Suppose, for example, that such a resolution contains the boundary of a disk $V \subset \mathcal{V}$ such that $V \in C_0$. We know some other resolution contains that boundary of a disk $W \subset \mathcal{W}$ such that $W \in C_1$. But this contradicts the fact that ∂V can be made disjoint from ∂W. The proof is now complete by symmetry. $\qquad \square$

3.4. **Finding an unlabelled region.**

To obtain a contradiction to our assumption that all regions are labeled it suffices to prove that the map from D^2 to Π, when restricted

to ∂D^2, induces a non-trivial map on homology. To this end we must examine the possibilities for the labels of the regions adjacent to ∂D^2.

Claim 3.15. *If S_x is disjoint from a compressing disk $V \subset \mathcal{V}$ for H such that $V \in C_i$ then the region containing x does not have the label $\mathbb{W}_{1-i}$. Similarly, if S_x is disjoint from a compressing disk $W \subset \mathcal{W}$ for H such that $W \in C_i$ then the region containing x does not have the label $\mathbb{V}_{1-i}$.*

Proof. Assume S_x is disjoint from a disk $V \subset \mathcal{V}$, where $V \in C_i$. If the region containing x has the label $\mathbb{W}_{1-i}$ then there is a loop γ of $S_x \cap H$ which bounds a compressing disk $W \subset \mathcal{W}$ for H such that $W \in C_{1-i}$. Then $\partial W \cap \partial V = \emptyset$. This contradicts the fact that (V, W) is not a reducing pair. The proof is complete by symmetry. $\qquad\square$

Recall the arcs U and L of ∂D^2 from Claim 3.10. For any point x near U the surface S_x is disjoint from some compressing disk for H in $\mathcal{V}$. Similarly, for any x near L the surface S_x is disjoint from a compressing disk for H in $\mathcal{W}$.

Claim 3.16. *Suppose $\mathcal{R}_0$ and $\mathcal{R}_1$ are regions such that $\mathcal{R}_0 \cap U$ is adjacent to $\mathcal{R}_1 \cap U$. If $\mathcal{R}_0$ has the label $\mathbb{W}_0$ then $\mathcal{R}_1$ cannot have the label $\mathbb{W}_1$. Similarly, if $\mathcal{R}_0 \cap L$ is adjacent to $\mathcal{R}_1 \cap L$ and $\mathcal{R}_0$ has the label $\mathbb{V}_0$ then $\mathcal{R}_1$ cannot have the label $\mathbb{V}_1$.*

Proof. Suppose $\mathcal{R}_0 \cap U$ is adjacent to $\mathcal{R}_1 \cap U$, the label of $\mathcal{R}_0$ is $\mathbb{W}_0$, and the label of $\mathcal{R}_1$ is $\mathbb{W}_1$. Let p be the point $\mathcal{R}_0 \cap \mathcal{R}_1 \cap U$. By Claim 3.10 the surface S_p is disjoint from a compressing disk $V \subset \mathcal{V}$ for H. Hence, for all points x near p the surface S_x will be disjoint from V. But every neighborhood of p contains points in both $\mathcal{R}_0$ and $\mathcal{R}_1$. For x near p and in $\mathcal{R}_0$ Claim 3.15 implies $V \in C_0$. But if $x \in \mathcal{R}_1$ then Claim 3.15 implies $V \in C_1$. As C_0 and C_1 partition the compressing disks for H the disk V cannot be in both. $\qquad\square$

We now examine the properties of Φ listed in Claim 3.10 and the complex Π. If x is in a region containing the point θ_0 then S_x is disjoint from V_0 and W_0. It follows from Claim 3.15 that this region can not have either of the labels $\mathbb{V}_1$ or $\mathbb{W}_1$, and hence must get mapped to the left triangle of Π. Similarly, a region containing the point θ_1 gets mapped to the right triangle of Π. If x is in a region containing a point of U then S_x is disjoint from some compressing disk for H in $\mathcal{V}$. It follows from Claim 3.15 that such a region cannot get both of the labels $\mathbb{W}_0$ and $\mathbb{W}_1$, and hence does not get mapped to the bottom point of Π. It follows from Claim 3.16 that no adjacent pair of regions next to U gets mapped to the horizontal edge of the bottom triangle

of Π. We conclude that the arc U gets mapped to the union of the left, right, and top triangles of Π. A symmetric argument shows that L gets mapped to the union of the left, right, and bottom triangles of Π. We summarize these observations in Figure 7. It follows that the map from D^2 to Π, when restricted to ∂D^2, is non-trivial on homology, a contradiction. We conclude that there must have been an unlabeled region, as this was the only assumption we made in constructing the map. Lemma 3.5 now follows from Claim 3.11. $\qquad\square$
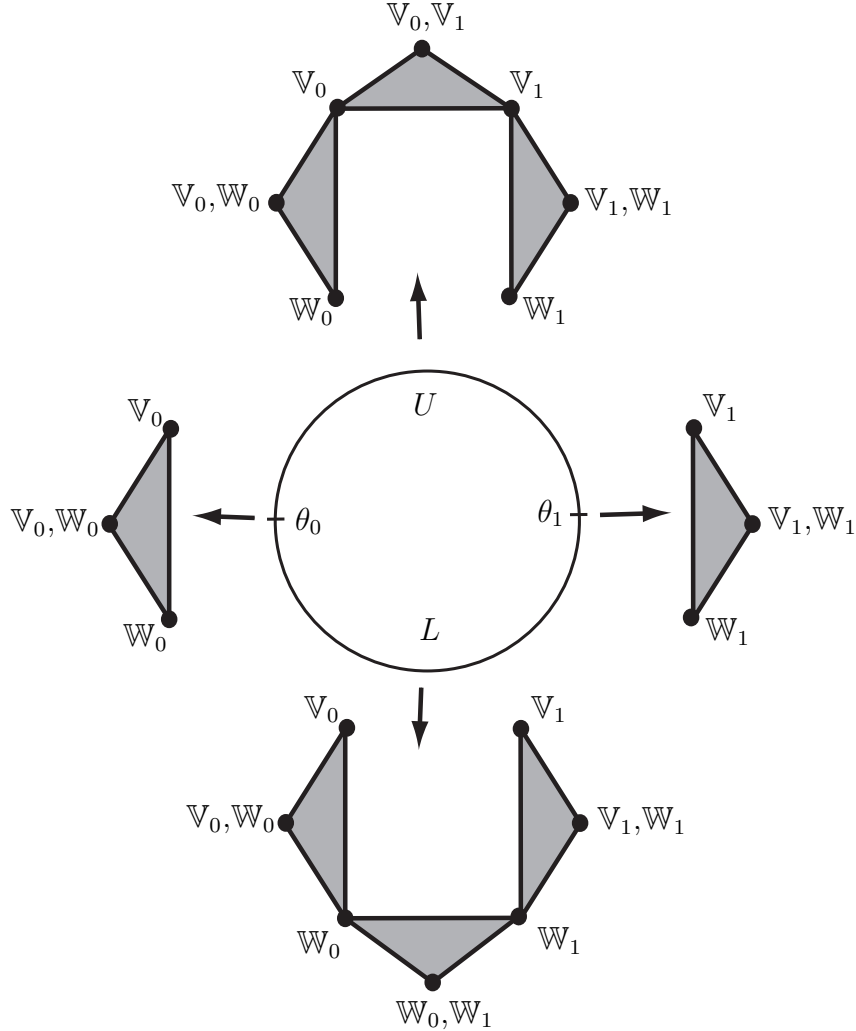

FIGURE 7. The map from ∂D^2 to Π is non-trivial on homology.

4. Generalized Heegaard Splittings

Theorem 10.1 is essentially proved by a complex sequence of handle slides. However, we find that a handle structure is more cumbersome to deal with than a *generalized Heegaard splitting*, defined presently. The relationship between handle structures and generalized Heegaard splittings is made explicit in Example 4.4.

Definition 4.1. A *generalized Heegaard splitting (GHS)* H of a 3-manifold M is a pair of sets of pairwise disjoint, connected surfaces, $\text{Thick}(H)$ and $\text{Thin}(H)$ (called the *thick levels* and *thin levels*, respectively), which satisfy the following conditions.

(1) Each component M' of $M - \text{Thin}(H)$ meets a unique element H_+ of $\text{Thick}(H)$, and H_+ is a Heegaard surface in M'. Henceforth we will denote the closure of the component of $M - \text{Thin}(H)$ that contains an element $H_+ \in \text{Thick}(H)$ as $M(H_+)$.

(2) As each Heegaard surface $H_+ \subset M(H_+)$ is oriented, say, by an outward pointing normal vector, we can consistently talk about the points of $M(H_+)$ that are "above" H_+ or "below" H_+. Suppose $H_- \in \text{Thin}(H)$. Let $M(H_+)$ and $M(H'_+)$ be the submanifolds on each side of H_- (a priori it is possible that $M(H_+) = M(H'_+)$). Then H_- is below H_+ if and only if it is above H'_+.

(3) There is a partial ordering on the elements of $\text{Thin}(H)$ which satisfies the following: Suppose H_+ is an element of $\text{Thick}(H)$, H_- is a component of $\partial M(H_+)$ above H_+, and H'_- is a component of $\partial M(H_+)$ below H_+. Then $H_- > H'_-$.

Example 4.2. Suppose H_+ is a Heegaard surface in a closed 3-manifold M. Then a GHS H of M is given by $\text{Thick}(H) = \{H_+\}$ and $\text{Thin}(H) = \emptyset$.

Example 4.3. Suppose M is a 3-manifold, H_+ is a Heegaard surface in M, and ∂M has two homeomorphic components separated by H_+. Let M^1 and M^2 denote two homeomorphic copies of M, and H^i_+ the image of H_+ in M^i. Let N denote the connected 3-manifold obtained from M^1 and M^2 by identifying the boundary components in pairs by some homeomorphisms. The image of the boundary components in N are F_1 and F_2. We may attempt to define a GHS H by setting $\text{Thick}(H) = \{H^1_+, H^2_+\}$ and $\text{Thin}(H) = \{F_1, F_2\}$. But this H is NOT a GHS of N, since there is no way to consistently satisfy conditions (2) and (3) of the definition. If orientations are chosen on H^1_+ and H^2_+ to satisfy condition (2), then the relation specified by condition (3) will imply both $F_1 > F_2$ and $F_2 > F_1$. See Figure 8.
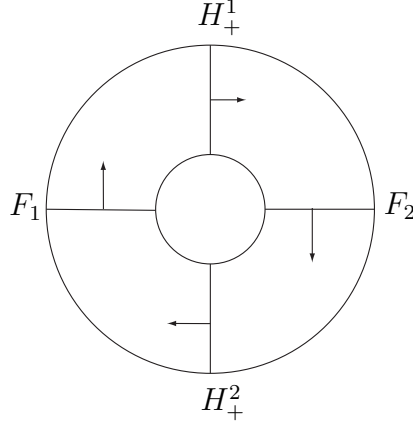
FIGURE 8. The orientations on H_+^1 and H_+^2 are given by normal vectors. If the orientations on F_1 and F_2 are chosen to be consistent with condition (2) of Definition 4.1 (as pictured) then they will be inconsistent with condition (3).

Example 4.4. The original formulation of a GHS was given by Scharlemann and Thompson in [ST94]. They defined these structures as being "dual" to handle structures. To be explicit, assume M is built by starting with 0-handles and attaching 1-handles, 2-handles, and 3-handles in any order. The set of thick and thin levels of the GHS associated with this handle structure appears at the various interfaces of the 1- and 2-handles, as indicated in Figure 9. Note that some of the horizontal lines in the figure may represent a disconnected surface. Each component of such a surface will be either a thin or thick level.
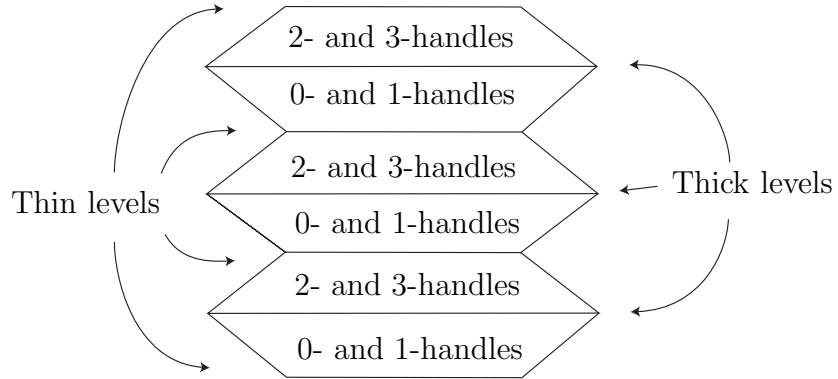

FIGURE 9. The GHS associated with a handle structure.

Definition 4.5. Suppose H is a GHS of a 3-manifold M with no 3-sphere components. Then H is *strongly irreducible* if each element $H_+ \in \text{Thick}(H)$ is strongly irreducible in $M(H_+)$. We say H is *critical* if there is a unique element $H_* \in \text{Thick}(H)$ which is critical in $M(H_*)$, and every other element $H_+ \in \text{Thick}(H)$ is strongly irreducible in $M(H_+)$.

Lemma 4.6. *Suppose H is a strongly irreducible or critical GHS of M. Then each element of $\text{Thin}(H)$ is incompressible in M.*

The strongly irreducible case is motivated by [ST94].

Proof. Let H be a strongly irreducible or critical GHS. Choose a compressing disk D for some thin level whose interior meets the union of all thin levels a minimal number of times. Let α denote an intersection loop of D with the set of thin levels which is innermost on D (possibly $\alpha = \partial D$). Let H_- denote the thin level that contains α. Suppose first that α is inessential on H_-, bounding a subdisk A of this surface. Let B be a subdisk of A bounded by a loop of $D \cap A$ which is innermost on A. Then we may use A to surger D, producing a new compressing disk for some thin level which meets the union of all thin levels fewer times.

Now suppose α is essential on H_-. The loop α bounds a subdisk A' of D which lies in $M(H_+)$, for some $H_+ \in \text{Thick}(H)$. By assumption H_+ is either a strongly irreducible or a critical Heegaard surface in $M(H_+)$.

By Lemma 3.5 there is a disk in $M(H_+)$, with the same boundary as A', which misses H_+. This disk is thus a compressing disk for the negative boundary of a compression body, a contradiction. $\square$

Lemma 4.7. *If M admits a strongly irreducible or critical GHS then M is irreducible.*

Proof. Let H denote a strongly irreducible or critical GHS. By Lemma 4.6 the union of the set of thin levels of H forms a (disconnected) incompressible surface in M. By Lemma 3.5 there is thus an essential sphere S' which is disjoint from $\text{Thin}(H)$. Hence $S' \subset M(H_+)$ for some $H_+ \in \text{Thick}(H)$. Again by Lemma 3.5 there must be an essential sphere $S'' \subset M(H_+)$ that misses H_+. But then S'' is an essential sphere in a compression body, which can not exist. $\square$

5. REDUCING GHSS

We now define a way to take a GHS G of a 3-manifold M which is not strongly irreducible and obtain a "simpler" GHS H. The new GHS

will be of a manifold M' which is obtained from M by cutting along a sphere and capping off the resulting boundary components with a pair of 3-balls. This procedure will be called *weak reduction*, and can be performed whenever there is a thick level $G_+ \in \text{Thick}(G)$ and a reducing pair (D, E) for G_+ in $M(G_+)$. We assume M has no 3-sphere components. We define the new GHS by the following algorithm:

(1) Initially set
$$\text{Thick}(H) = \text{Thick}(G) - \{G_+\} \cup \{G_+/D, G_+/E\},$$
$$\text{Thin}(H) = \text{Thin}(G) \cup \{G_+/DE\}, \text{ and } M' = M.$$
See Figure 10.

(2) If there is now a sphere $S \in \text{Thin}(H)$ then cut M' along S, cap off the resulting sphere boundary components with 3-balls, and remove S from $\text{Thin}(H)$.

(3) If M' now contains a 3-sphere component then delete it, along with all elements of $\text{Thick}(H)$ and $\text{Thin}(H)$ that lie in this component.

(4) If there are elements $H_+ \in \text{Thick}(H)$ and $H_- \in \text{Thin}(H)$ that cobound a product region P of M' such that $P \cap \text{Thick}(H) = H_+$ and $P \cap \text{Thin}(H) = H_-$ then remove H_+ from $\text{Thick}(H)$ and H_- from $\text{Thin}(H)$.

The first step of the above algorithm is illustrated in Figure 10.
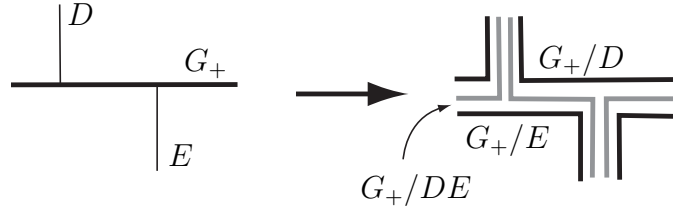

FIGURE 10. The first step in defining a weak reduction.

Example 5.1. Suppose $\text{Thick}(H)$ contains a single element H_+ and $\text{Thin}(H) = \emptyset$, so that H_+ is a Heegaard surface for M. Suppose (D, E) is a weak reduction for H and ∂D and ∂E cobound an annulus A of H_+. If the sphere $D \cup A \cup E$ is separating in M then the inverse of the weak reduction is a connected sum, in the sense of Definition 2.3.

Suppose H_+ is obtained from some lower genus Heegaard surface by a stabilization. Recall that a stabilization is a connected sum with the standard genus one Heegaard surface in S^3. Hence there will be a weak reduction for H which undoes this.

In the previous example we saw that a sphere may appear at a thin level in the process of weak reduction when ∂D and ∂E were parallel. The next example shows that there is another way for spheres to crop up at thin levels during weak reduction.

Example 5.2. Let M_1 and M_2 be 3-manifolds with genus one Heegaard splittings (possibly one or both are S^3). Let $M = M_1 \# M_2$. Then, as in Example 2.4, there is a genus two Heegaard splitting H_+ of M which is the connected sum of genus one splittings of M_1 and M_2. There are compressing disks $D \subset M_1$ and $E \subset M_2$ on opposite sides of H_+ in M. Hence the pair (D, E) is a reducing pair. In forming the weak reduction one of the first steps is to initially add H_+/DE to Thin(H). But this surface is a sphere, which will get removed in subsequent steps.

If $M_2 \cong S^3$ then H_+ was a stabilization of the splitting of M_1 that we started with. The weak reduction given by (D, E) has undone this stabilization.

The preceding pair of examples motivates us to make the following definition.

Definition 5.3. The weak reduction of a GHS given by the reducing pair (D, E) for the thick level H_+ is called a *destabilization* if H_+/DE contains a sphere which bounds a ball.

There are two ways that the weak reduction given by (D, E) can be a destabilization. The first is when ∂D and ∂E are parallel on H_+, as in Example 5.1. The second is when ∂D and ∂E are non-parallel but H_+ has genus two, as in Example 5.2.

Example 5.4. Consider $S^2 \times S^1$. If l is a loop in S^2 then $l \times S^1$ is a Heegaard torus, H. The loop l bounds disks D and E on opposite sides of H, which can be isotoped to be disjoint. In the process of weak reduction the manifold will get cut along an essential 2-sphere, and the resulting boundary components will get capped off by 3-balls. The result is S^3, which will then get deleted. Hence, weak reducing along the pair (D, E) produces the empty GHS of the empty set.

We now define a partial ordering of GHSs. Any partial ordering will suffice, as long as a weak reduction produces a smaller GHS, and any monotonically decreasing sequence of GHSs must terminate. This is motivated by [ST94].

Definition 5.5. If H is a GHS then let $c(H)$ denote the set of genera of the elements of Thick(H), where repeated integers are included, and the set is put in non-increasing order. We compare two such sets

lexicographically. If $H^1 = \{H_i^1\}$ and $H^2 = \{H_j^2\}$ are two GHSs then we say $H^1 < H^2$ if $c(H^1) < c(H^2)$.

Lemma 5.6. *Suppose H^1 and H^2 are GHSs and H^1 is obtained from H^2 by a weak reduction. Then $H^1 < H^2$.*

Proof. In the definition of weak reduction we remove H_+ from $\mathrm{Thick}(H)$ and replace it with (possibly multiple) surfaces of smaller genus. Some of these surfaces may then appear at the boundary of a product submanifold or in an S^3 component, and so will be removed. In any case, $c(H)$ has gone down under the lexicographical ordering. $\square$

6. Swapping Weak Reductions

There will be many times when we will want to alter a sequence of weak reductions. In this section we list a few basic ways to do this.

Lemma 6.1. *Suppose (D, E) and (D', E) are different weak reductions for a GHS G. Let G_+ denote the thick level of G which contains ∂D, $\partial D'$, and ∂E, and suppose ∂D separates ∂E from $\partial D'$ on G_+. Then performing the weak reduction (D, E) on G yields the same GHS as performing (D', E) followed by (D, E).*

Proof. The proof is indicated in Figure 11. After performing the weak reduction (D, E) the components of the surface G_+/D are thick levels. (This is illustrated in the left side of the figure.) But $\partial D'$ and ∂E now lie on different components, so we cannot follow with the weak reduction (D', E).

Since ∂D separates $\partial D'$ from ∂E on G_+ it follows that ∂D and ∂E lie on the same component of G_+/D'. Hence we may follow the weak reduction (D', E) with the weak reduction (D, E). (This is indicated in the right side of the figure.) After removing parallel thick and thin levels (as indicated by the shaded regions in the figure) the two GHSs that we obtain are the same.

$\square$

Lemma 6.2. *Suppose (D, E) and (D', E) are different weak reductions for a GHS G. Let G_+ denote the thick level of G which contains ∂D, $\partial D'$, and ∂E, and suppose neither ∂D nor $\partial D'$ separates the other from ∂E on G_+. If, furthermore, neither ∂D nor $\partial D'$ are parallel to ∂E then performing the weak reduction (D, E) followed by (D', E) yields the same GHS as performing (D', E) followed by (D, E).*

Proof. The proof is indicated in Figure 12. After performing the weak reduction (D, E) the curves $\partial D'$ and ∂E lie on same component of G_+/D so we can follow with the weak reduction (D', E) (as indicated in
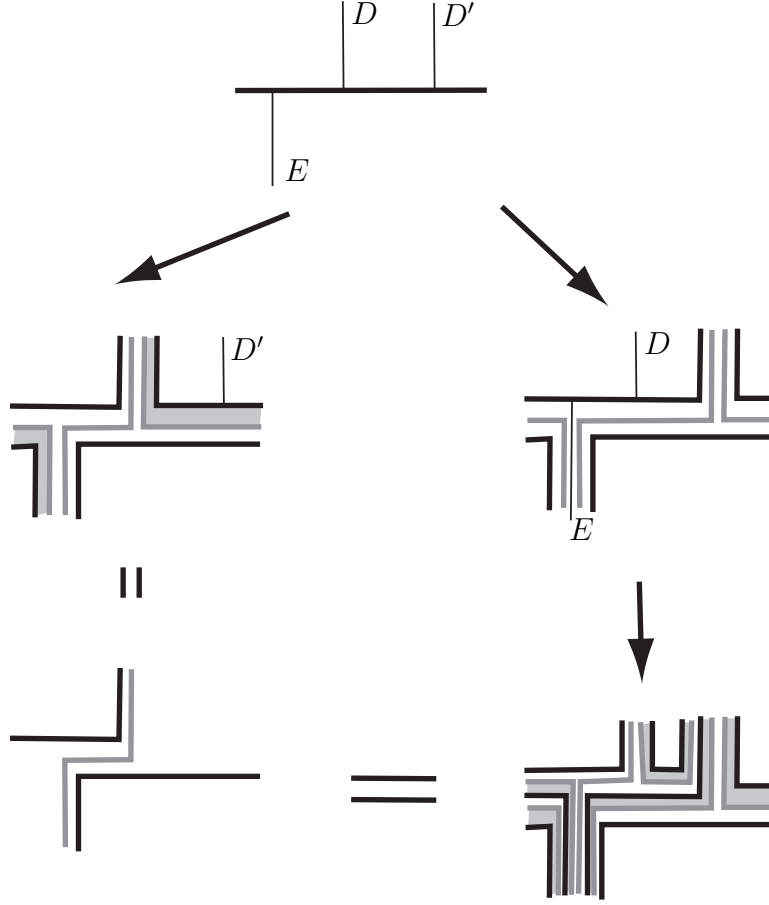
FIGURE 11. If ∂D separates $\partial D'$ from ∂E then perform-
ing the weak reduction (D, E) yields the same GHS as
performing (D', E) followed by (D, E).

the left side of the figure). Similarly, we may follow the weak reduction
$(D'E)$ the weak reduction with (D, E) (as indicated in the right side of
the figure). After removing parallel thick and thin levels the two GHSs
are the same. $\square$

Lemma 6.3. *Suppose (D, E) and (D', E) are different weak reductions
for a GHS G. Let G_+ denote the thick level of G which contains ∂D,
$\partial D'$, and ∂E. If $\partial D'$ is parallel to ∂E then performing the weak re-
duction (D, E) followed by (D', E) yields the same GHS as performing
(D', E).*

Proof. The proof is indicated in Figure 13. On the left side of this figure
we illustrate the following steps: (1) Performing the weak reduction
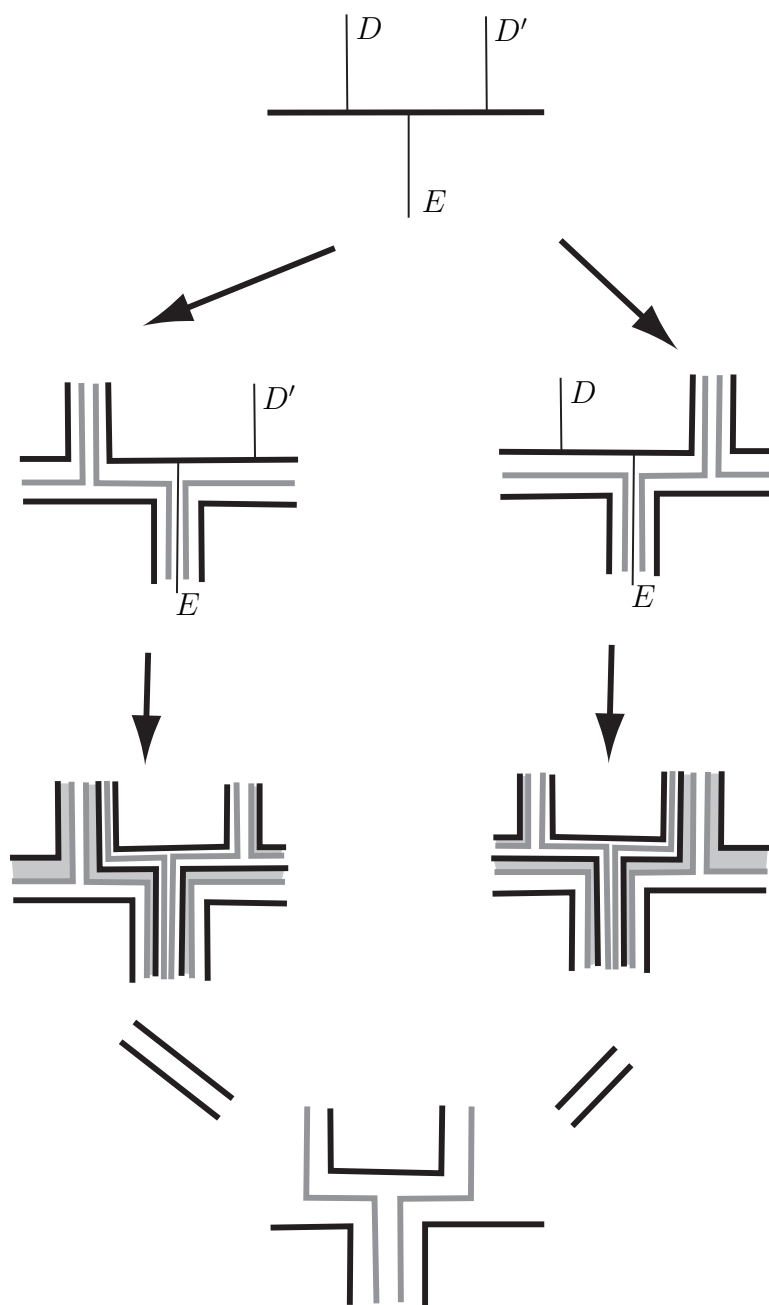
FIGURE 12. Performing (D, E) followed by (D', E) yields the same GHS as (D', E) followed by (D, E).

(D, E). After this the curves $\partial D'$ and ∂E lie on same component of G_+/D so we can follow with (2) the weak reduction (D', E). There are now parallel thick and thin levels that can be removed (3). There is also a sphere thin level, so to complete the formation of the GHS we must (4) cut along the sphere and cap off with 3-balls. When we do this some thin level becomes parallel to a thick level (5 and 6), and so they both get removed (e).

On the right side we (a) do (D', E) first. We can not follow with (D, E) because E is not a compressing disk for G_+/D'. However, there is a sphere thin level, so we must again (b) cut and cap off with 3-balls. Once again there is now a thin level which becomes parallel to a thick level (c and d), and so they both get removed (e).

$\square$

7. AMALGAMATIONS

Let H be a GHS of a 3-manifold M. In this section we use H to produce a graph Σ in M. Each component of M will contain exactly one component of Σ. Each component of Σ will be the spine of a Heegaard splitting of the component of M that contains it. We call the disjoint union of these Heegaard splittings the *amalgamation* of H. First, we must introduce some new notation.

Definition 7.1. Suppose H is a GHS of M and $H_+ \in \text{Thick}(H)$. Recall that H_+ is oriented, so that we may consistently talk about those points of $M(H_+)$ that are "above" H_+ and those points that are "below." The surface H_+ divides $M(H_+)$ into two compression bodies. Henceforth we will denote these compression bodies as $\mathcal{W}_\uparrow(H_+)$ and $\mathcal{W}_\downarrow(H_+)$, where $\mathcal{W}_\uparrow(H_+)$ is above H_+ and $\mathcal{W}_\downarrow(H_+)$ is below. When we wish to make reference to an arbitrary compression body which lies above or below some thick level we will use the notation $\mathcal{W}_\uparrow$ and $\mathcal{W}_\downarrow$. Define $\partial_\uparrow M(H_+)$ to be $\partial_- \mathcal{W}_\uparrow(H_+)$ and $\partial_\downarrow M(H_+)$ to be $\partial_- \mathcal{W}_\downarrow(H_+)$. That is, $\partial_\uparrow M(H_+)$ and $\partial_\downarrow M(H_+)$ are the boundary components of $M(H_+)$ that are above and below H_+, respectively.

We now inductively build Σ. The intersection of Σ with some $M(H_+)$ is depicted in Figure 14. First, we define a sequence of manifolds $\{M_i\}$ where

$$M_0 \subset M_1 \subset ... \subset M_n = M.$$

The submanifold M_0 is defined to be the disjoint union of all manifolds of the form $M(H_+)$, such that $\partial_\downarrow M(H_+) = \emptyset$. The fact that M is closed and the thin levels of H are partially ordered guarantees $M_0 \neq \emptyset$. Now, for each i we define M_i to be the union of M_{i-1} and all
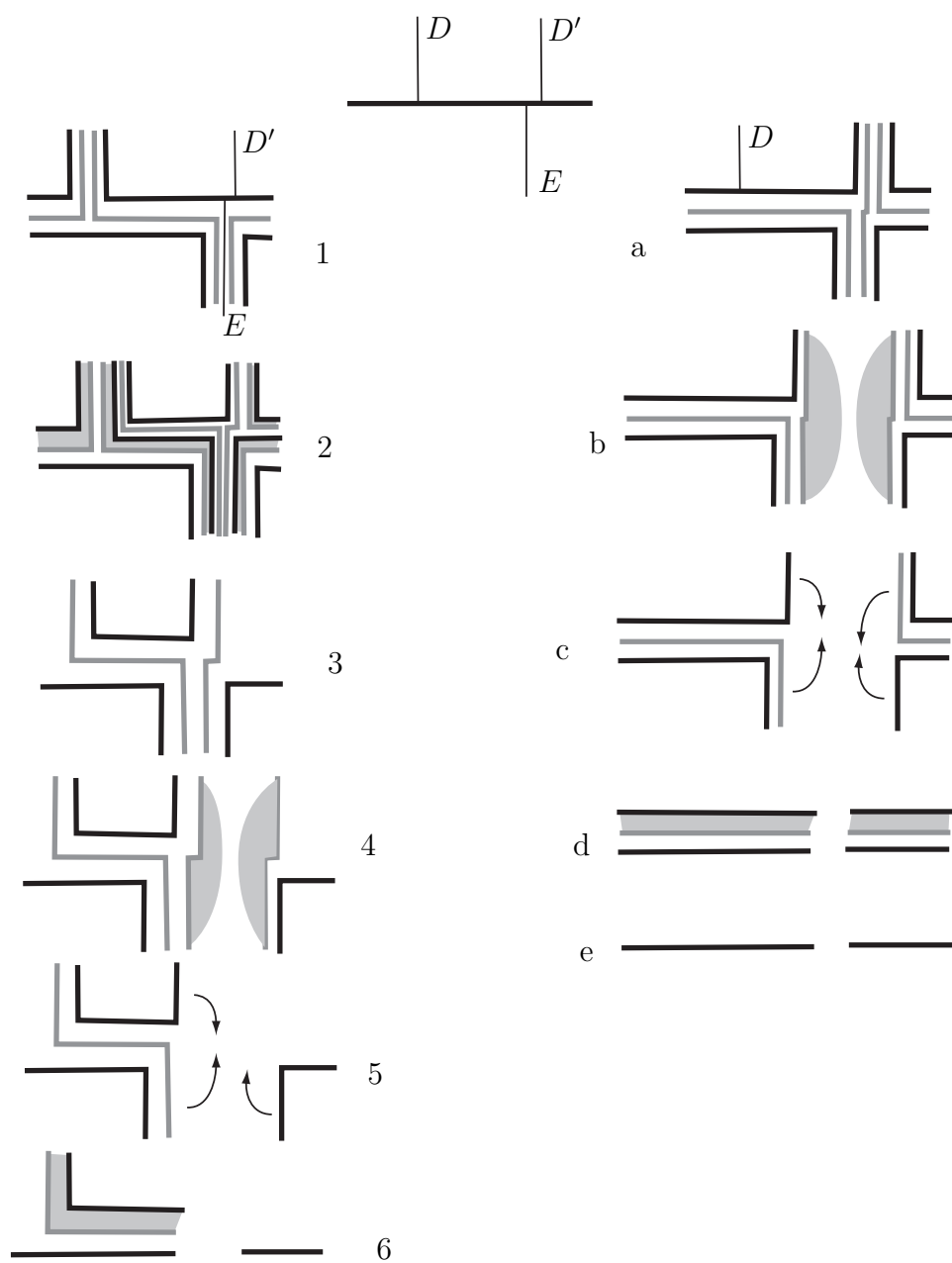
FIGURE 13. Performing (D, E) followed by (D', E) (left side) yields the same GHS as performing just (D', E) (right side).

manifolds $M(H_+)$ such that $\partial_\downarrow M(H_+) \subset \partial M_{i-1}$. Again, it follows from the partial ordering of thin levels that for some i the manifold $M_i = M$.

We now define a sequence of graphs Σ_i in M. The final element of this sequence will be the desired graph Σ.

Each $\mathcal{W}_\downarrow \subset M_0$ is a handle-body. Choose a spine of each, and let Σ'_0 denote the union of these spines. The complement of Σ'_0 in M_0 is a (disconnected) compression body, homeomorphic to the union of the compression bodies $\mathcal{W}_\uparrow \subset M_0$. Now let Σ_0 be the union of Σ'_0 and one vertical arc for each component H_- of ∂M_0, connecting H_- to Σ'_0.

We now assume Σ_{i-1} has been constructed and we construct Σ_i. Let $M'_i = \overline{M_i - M_{i-1}}$. For each compression body $\mathcal{W}_\downarrow \subset M'_i$ choose a set of arcs $\Gamma \subset \mathcal{W}_\downarrow$ such that $\partial \Gamma \subset \Sigma_{i-1} \cap \partial M_{i-1}$, and such that the complement of Γ in $\mathcal{W}_\downarrow$ is a product. Let Σ'_i be the union of Σ_{i-1} with all such arcs Γ. Now let Σ_i be the union of Σ'_i and one vertical arc for each component H_- of ∂M_i, connecting H_- to Σ'_i.
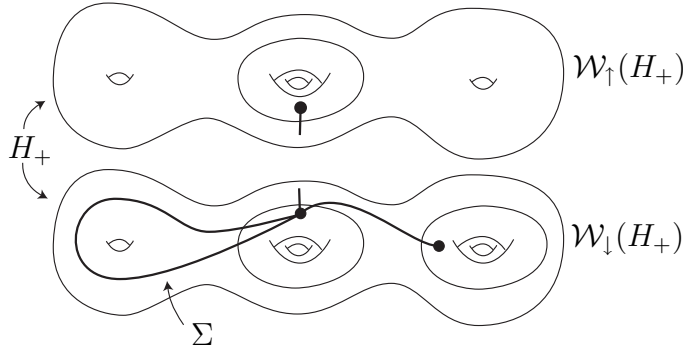

FIGURE 14. The intersection of Σ with $\mathcal{W}_\uparrow(H_+)$ and $\mathcal{W}_\downarrow(H_+)$.

Lemma 7.2. *If H is a GHS of M then each component of the graph Σ defined above is the spine of a Heegaard splitting of the component of M that contains it.*

Proof. Recall the sequence of manifolds $\{M_i\}$ above. We prove the lemma by showing that if the complement of $N(\Sigma)$ in M_{i-1} is a union of compression bodies then the complement of $N(\Sigma)$ in M_i is a union of compression bodies. For this it is convenient to set $M_{-1} = \emptyset$.

Suppose $M(H_+) \subset M'_i = \overline{M_i - M_{i-1}}$. The complement of $N(\Sigma)$ in $\mathcal{W}_\downarrow(H_+)$ is a product of the form $\{$punctured surface$\} \times I$, so when we glue this on to the complement of $N(\Sigma)$ in M_{i-1} we do not change the homeomorphism type of the manifold. The complement of $N(\Sigma)$

in $\mathcal{W}_\uparrow(H_+)$ is of the form {punctured surface} $\times I$, together with a collection of 1-handles. Gluing this to the complement of $N(\Sigma)$ in $M_{i-1} \cup \mathcal{W}_\downarrow(H_+)$ thus produces a bigger compression body. $\qquad\square$

Definition 7.3. Let H be a GHS and Σ the graph in M defined above. The union of the Heegaard splittings that each component of Σ is a spine of is called the *amalgamation* of H and will be denoted $\mathcal{A}(H)$.

Note that although the construction of the graph Σ involved some choices, its neighborhood $N(\Sigma)$ is uniquely defined.

Lemma 7.4. *Suppose M is irreducible, H is a GHS of M and G is obtained from H by a weak reduction which is not a destabilization. Then $\mathcal{A}(H)$ is isotopic to $\mathcal{A}(G)$.*

Proof. Suppose G is obtained from H by the weak reduction (D, E) for the thick level H_+, where D is above H_+ and E is below. Let Σ be the graph associated with H as defined above. In the first stage of weak reduction we replace H_+ in Thick(H) with the components of H_+/D and H_+/E. We also add the components of H_+/DE to Thin(H). Since (D, E) is not a destabilization and M is irreducible it follows that H_+/DE did not contain any sphere components. Hence, to complete the formation of the weak reduction it remains only to remove parallel thick and thin levels. Now observe that if we postpone this step and we first form the associated graph Σ' as above, then we end up with the same graph as we would have if we removed parallel thick and thin levels first.

We now claim that a neighborhood $N(\Sigma')$ is isotopic to $N(\Sigma)$. Outside $M(H_+)$ we may assume Σ and Σ' coincide, so we focus our attention inside $M(H_+)$.
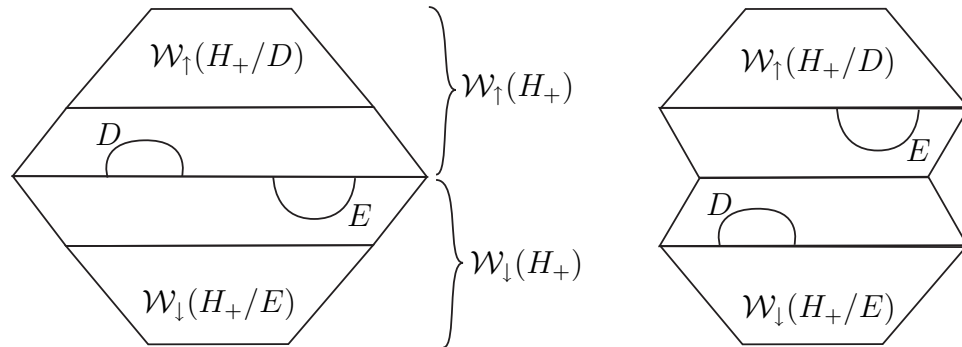

FIGURE 15. Because D and E are disjoint the manifold $M(H_+)$ can be built in two different ways.

The manifold $M(H_+)$ can be built as follows. Begin with $\mathcal{W}_\downarrow(H_+/E)$. Attach a 1-handle to the positive boundary whose co-core is E. This is now the manifold $\mathcal{W}_\downarrow(H_+)$. To continue, attach a 2-handle whose core is D, and then attach the manifold $\mathcal{W}_\uparrow(H_+/D)$. The 2-handle and $\mathcal{W}_\uparrow(H_+/D)$ is precisely $\mathcal{W}_\uparrow(H_+)$. See Figure 15. Since D and E are disjoint we may build $M(H_+)$ in an alternate way: Begin with $\mathcal{W}_\downarrow(H_+/E)$, attach the 2-handle, then attach the 1-handle, and finally attach $\mathcal{W}_\uparrow(H_+/D)$.
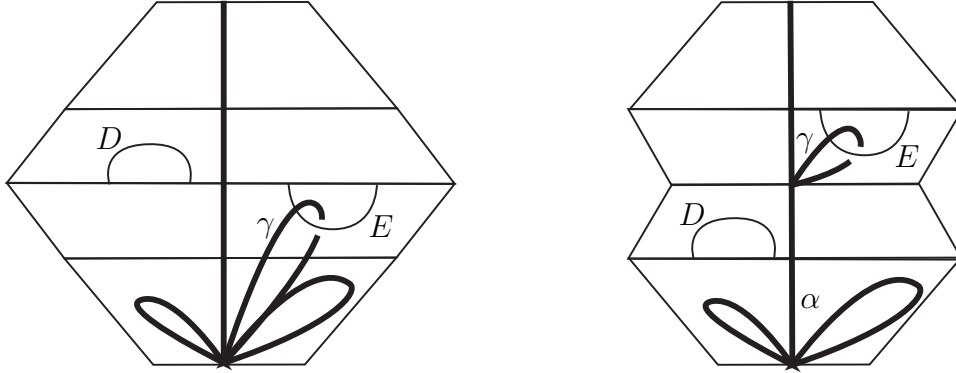

FIGURE 16. The graph Σ can be obtained from Σ' by contracting α.

Now note that we can build Σ by starting with a set of arcs in $\mathcal{W}_\downarrow(H_+/E)$ whose complement is a product, attaching a core γ of the 1-handle whose co-core is E, and then attaching some vertical arcs. Similarly, Σ' can be built by starting with the same set of arcs in $\mathcal{W}_\downarrow(H_+/E)$, attaching vertical arcs α, then attaching γ, and finishing with more vertical arcs. See Figure 16. By contracting the vertical arcs α we achieve the desired isotopy of $N(\Sigma')$.   □

8. SEQUENCES OF GENERALIZED HEEGAARD SPLITTINGS

Definition 8.1. A *Sequence Of GHSs* (SOG), $\{(H^i, M^i)\}$ is a finite sequence such that H^i is a GHS of M^i and either H^i or H^{i+1} is obtained from the other by a weak reduction.

Notation: We will always use superscripts to denote the GHSs of a SOG, and a boldface font to denote the entire SOG. Hence, H^j is the jth GHS of the SOG **H**.

Definition 8.2. If **H** is a SOG and k is such that H^{k-1} and H^{k+1} are obtained from H^k by a weak reduction then we say the GHS H^k is *maximal* in **H**.

It follows from Lemma 5.6 that maximal GHSs are larger than their immediate predecessor and immediate successor.

Definition 8.3. Suppose H is a Heegaard surface in a 3-manifold. Let (V_i, W_i) be a reducing pair for H for $i = 0, 1$. Then we define the *distance* between (V_0, W_0) and (V_1, W_1) to be the smallest n such that there is a sequence $\{D_j\}_{j=0}^{n+1}$ where

(1) $\{D_0, D_1\} = \{V_0, W_0\}$,
(2) $\{D_n, D_{n+1}\} = \{V_1, W_1\}$,
(3) for all j the pair (D_j, D_{j+1}) is a reducing pair for H,
(4) for $1 \le j \le n$, D_{j-1} is disjoint from D_{j+1}.

If there is no such sequence then we define the distance to be ∞.

The reader may wonder how this notion of distance between reducing pairs is related to the distance you would get by dropping the last condition. This is best visualized in the curve complex of H, where vertices correspond to isotopy classes of essential loops in H and edges correspond to disjoint pairs of loops. In Figure 17(a) we have depicted a path in the curve complex from ∂V_0 to ∂W_1. In Figure 17(b) we see how the boundaries of the disks D_i of Definition 8.3 are related. The picture is reminiscent of the *geodesic hierarchies* of [MM00]. The figure illustrates how the distance of Definition 8.3 is more closely related to the length of a chain of 2-simplices, rather than the length of a chain of 1-simplices.

Lemma 8.4. *Suppose H is an embedded surface in a 3-manifold. If there are reducing pairs $(V, W), (D, E)$ and (D, E') for H such that the distance between (V, W) and (D, E') is finite then the distance between (V, W) and (D, E) is finite.*

Proof. Suppose the distance between (V, W) and (D, E') is n. We now construct a sequence $\{E_i\}_{i=0}^m$ such that $E_0 = E'$, $E_m = E$, for all i the pair (D, E_i) is a reducing pair, and $E_i \cap E_{i+1} = \emptyset$. The sequence

$$\{D, E' = E_0, D, E_1, D, E_2, \ldots, D, E = E_m\}$$

then satisfies the conditions in Definition 8.3, establishing that the distance between (D, E') and (D, E) is at most $2m$. The distance between (V, W) and (D, E) is then at most $n + 2m$.

The sequence $\{E_i\}$ is defined inductively as follows:

(1) Define $E_0 = E'$.
(2) Let E_i denote the last disk defined, and suppose there is a simple closed curve in $E \cap E_i$. Let e denote an innermost subdisk of E bounded by a loop of $E \cap E_i$. Now surger E_i along e. The result
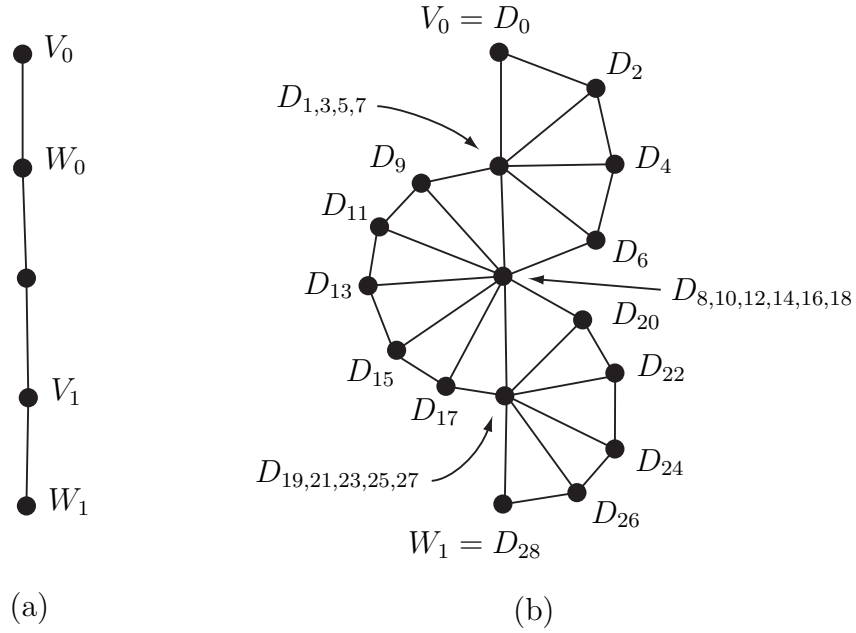
FIGURE 17. (a) The shortest path from ∂V_0 to ∂W_1 in the curve complex of H. (b) The distance from (V_0, W_0) to (V_1, W_1) is 27.

is a disk and a sphere. Throw away the sphere and denote the disk as E_{i+1}. Note that $|E \cap E_{i+1}| < |E \cap E_i|$.

(3) Let E_i denote the last disk defined, and suppose there are only arcs in $E \cap E_i$. Let e denote an outermost subdisk of E cut off by an arc of $E \cap E_i$. Now surger E_i along e. The result is two disks, at least one of which is a compressing disk for H. Call such a disk E_{i+1}. It follows from the fact that $e \cap D = \emptyset$ that $E_{i+1} \cap D = \emptyset$. Again, note that $|E \cap E_{i+1}| < |E \cap E_i|$.

(4) If neither of the previous two cases apply then we have arrived at a disk E_{m-1} such that (E_{m-1}, D) is a reducing pair and $E_{m-1} \cap E = \emptyset$. Now let $E_m = E$ and we are done.

$\square$

Lemma 8.5. *Suppose H is an embedded surface in a 3-manifold. If there are reducing pairs (V_0, W_0) and (V_1, W_1) for H such that the distance between (V_0, W_0) and (V_1, W_1) is ∞ then H is critical.*

Proof. Let C_0 be the set of compressing disks such that for each $D \in C_0$ there exists an E where the distance between (V_0, W_0) and (D, E) is finite. Let C_1 denote the set of compressing disks that are not in C_0.

We claim that the sets C_0 and C_1 satisfy the conditions of Definition 3.3.

Clearly, V_0 and W_0 are in C_0. We claim V_1 is in C_1. By symmetry it will follow that $W_1 \in C_1$, and hence Condition 1 of Definition 3.3 is satisfied. If $V_1 \notin C_1$ then there is an E which forms a reducing pair with V_1 such that the distance between (V_0, W_0) and (V_1, E) is finite. But then it would follow from Lemma 8.4 that the distance between (V_0, W_0) and (V_1, W_1) is finite, a contradiction.

Suppose now $D \in C_0$, and E is such that (D, E) is a reducing pair for H. To establish condition 2 we must show $E \in C_0$. By definition there is an E' such that the distance between (V_0, W_0) and (D, E') is finite. It thus follows from Lemma 8.4 that the distance between (V_0, W_0) and (D, E) is finite, and hence, $E \in C_0$. $\qquad\square$

We now define a complexity on maximal GHSs of SOGs. The definition is illustrated in Figure 18.
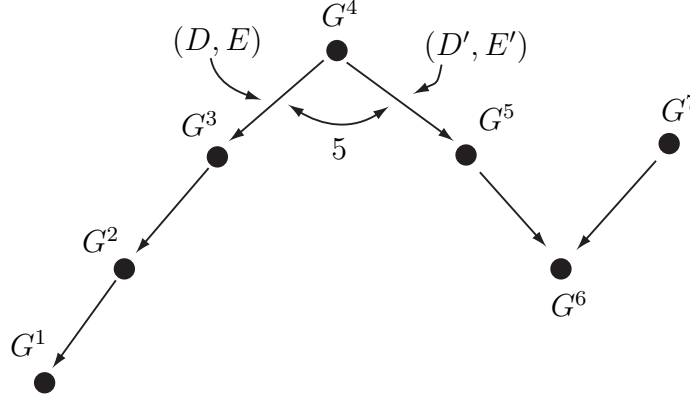

FIGURE 18. If the distance between (D, E) and (D', E') is 5 then the angle at G^4 is 5.

Definition 8.6. Suppose G^k is a maximal GHS of a SOG $\mathbf{G}$. Suppose further that G^{k-1} is obtained from G^k by the reduction given by the reducing pair (D, E) for the surface $G_+^k \in \mathrm{Thick}(G^k)$, and that G^{k+1} is obtained from G^k by the reduction given by the reducing pair (D', E') for the surface $G_*^k \in \mathrm{Thick}(G^k)$. If $G_+^k = G_*^k$ then define the *angle* $\angle(G^k)$ to be the distance between (D, E) and (D', E'). Otherwise we define $\angle(G^k)$ to be 1.

Recall that a weak reduction is a way to take a GHS and obtain a smaller one. Our goal here is to define several ways to take a SOG and obtain a smaller one. Any of these will be referred to as a *reduction*

of a SOG. To justify the statement that reduction produces something smaller, we must define a complexity for SOGs that induces a partial ordering. Furthermore, such a complexity should have the property that any decreasing sequence must terminate. This would immediately imply that any sequence of reductions must be finite. As our complexity is a lexicographically ordered multi-set of non-negative integers this latter property follows from a transfinite induction argument.

The actual complexity we define is a bit complicated. Fortunately, the only features that we will use are easy to list:

(1) Eliminating or replacing a maximal GHS with one or more smaller ones will represent a decrease in complexity.
(2) Replacing a maximal GHS with angle n with several identical maximal GHSs, each of which having angle less than n, will represent a decrease.

Any complexity one can define which behaves in this way will work for our purposes. We give one now for completeness. Let $\mathbf{G}$ be a SOG. Then the complexity of $\mathbf{G}$ is given by

$$\left\{\left(G^k, \angle(G^k)\right) | G^k \text{ is a maximal GHS}\right\}$$

Sets appearing in parentheses are ordered as written. Sets appearing in brackets are put in non-increasing (lexicographical) order and repetitions are included. When comparing the complexity of two SOGs one should make lexicographical comparisons at all levels.

Example 8.7. Consider the SOG pictured in Figure 19. If $H > K$ then the complexity of this SOG would be

$$\{(H, 7), (H, 5), (H, 5), (K, 6)\}.$$

If $K > H$ then the complexity would be

$$\{(K, 6), (H, 7), (H, 5), (H, 5)\}.$$

Finally, if H and K are not comparable (*i.e.* $c(H) = c(K)$) then the complexity would be

$$\{(H, 7), (K, 6), (H, 5), (H, 5)\}.$$

We now define the various complexity decreasing operations that one can perform on a SOG that will be referred to as *reductions*. As in Definition 8.6 assume G^k is maximal in $\mathbf{G}$, so that there is some thick level G^k_+ such that G^{k-1} is obtained from G^k by a reduction given by the reducing pair (D, E) for the surface G^k_+, and there is a thick level G^k_* such that G^{k+1} is obtained from G^k by a reduction given by the reducing pair (D', E') for the surface G^k_*.
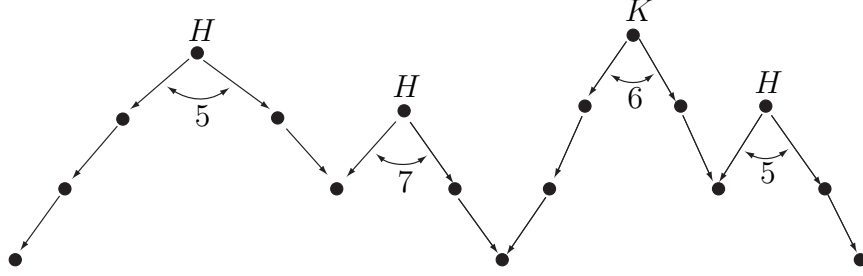
FIGURE 19. A SOG.

8.1. Reductions of Type I.

These are reductions which effect maximal GHSs without any consideration of their angles.

- If $G_+^k \neq G_*^k$ then we may replace G^k with H^* in $\mathbf{G}$, where H^* is the GHS obtained from G^{k-1} by the weak reduction (D', E'). Since H^* can also be obtained from G^{k+1} by the weak reduction (D, E) our substitution has defined a smaller SOG $\mathbf{G}'$.
- Next, assume $G_+^k = G_*^k$, but there is a thick level $G_0^k \neq G_+^k$ in Thick(G^k) which is not strongly irreducible. Let (D^*, E^*) be a reducing pair for G_0^k. Let H^{k-1}, H^k, and H^{k+1} denote the GHSs obtained from G^{k-1}, G^k, and G^{k+1} by the weak reduction corresponding to (D^*, E^*). Now replace G^k in $\mathbf{G}$ with the subsequence $\{H^{k-1}, H^k, H^{k+1}\}$ to define a new, smaller SOG $\mathbf{G}'$.

8.2. Reductions of Type II.

In all Type II reductions $G_+^k = G_*^k$ and we focus on $\angle(G^k)$.

- $\angle(G^k) = 0$. Then (D, E) is the same as (D', E'), so removal of the subsequence $\{G^k, G^{k+1}\}$ from $\mathbf{G}$ defines a new, smaller SOG $\mathbf{G}'$.
- $\angle(G^k) = 1$. In this case either $D = D'$ and $E \cap E' = \emptyset$ or $E = E'$ and $D \cap D' = \emptyset$. Assume the latter. There are now three subcases.

  First, assume neither ∂D nor $\partial D'$ is parallel to ∂E on G_+^k. If ∂D separates ∂E from $\partial D'$ then Lemma 6.1 implies that G^{k-1} can be obtained from G^{k+1} by the weak reduction (D, E). Hence, we may remove G^k from $\mathbf{G}$ to obtain a smaller SOG. A symmetric argument holds if $\partial D'$ separates ∂D from ∂E.

  If neither ∂D nor $\partial D'$ separates the other from ∂E, and neither is parallel to ∂E, then we may apply Lemma 6.2. This

implies that there is a GHS H that can be obtained from G^{k-1} by the weak reduction (D', E), and from G^{k+1} by the weak reduction (D, E). Hence, we may replace G^k in $\mathbf{G}$ with H to obtain a smaller SOG.

Finally, if $\partial D'$ is parallel to ∂E we may apply Lemma 6.3. This implies that G^{k+1} can be obtained from G^{k-1} by the weak reduction (D', E). Hence we may remove G^k from $\mathbf{G}$ to obtain a smaller SOG. A symmetric argument holds if ∂D is parallel to ∂E.

- $\angle(G^k) = n$ for some $n > 1$. Let $\{D_j\}_{j=0}^{n+1}$ be a sequence given by Definition 8.3. Choose some m between 1 and $n-1$. Then the reducing pair (D_m, D_{m+1}) cannot be equal to either (D, E) or (D', E'). Let G^* denote the GHS obtained from G^k by the weak reduction corresponding to (D_m, D_{m+1}). Now, let $\mathbf{G}'$ denote the SOG obtained from $\mathbf{G}$ by inserting the subsequence $\{G^*, G^k\}$ just after G^k. Note that the maximal GHS G^k appears one more time in $\mathbf{G}'$ than in $\mathbf{G}$. However, the angle at the old occurrence of G^k is now m, and the angle at the new occurrence is $n-m$. As both of these numbers are smaller than n we have produced a smaller SOG.

Definition 8.8. If the first and last GHS of a SOG are strongly irreducible and none of the above reductions can be performed then the SOG is said to be *irreducible*.

Lemma 8.9. *Every maximal GHS of an irreducible SOG is critical.*

Proof. The fact that every thick level but one is strongly irreducible follows immediately from the fact that one cannot perform any Type I reductions. The remaining thick level (the surface G_+^k in the definition of the reductions) must be critical by Lemma 8.5 since the lack of availability of Type II reductions implies that the distance between (D, E) and (D', E') is ∞. $\qquad\square$

Lemma 8.10. *Suppose $\mathbf{Y}$ is a SOG that is obtained from a SOG $\mathbf{X}$ by a reduction. If there is a destabilization in $\mathbf{X}$, and it comes before any stabilization, then there is a destabilization in $\mathbf{Y}$, and it comes before any stabilization in $\mathbf{Y}$.*

Proof. The fact that there is a destabilization in $\mathbf{Y}$ follows immediately from the fact that no reduction will ever remove a destabilization, just possibly exchange its order with some weak reduction. The remainder of the assertion can best be seen graphically. If X^i is a GHS of $\mathbf{X}$ then let $g(X^i)$ denote the genus of $\mathcal{A}(X^i)$. For the SOG $\mathbf{X}$ one can thus plot the point $(i, g(X^i))$ in the xy-plane. By Lemma 7.4 the

only increases and decreases in this graph are due to stabilizations and destabilizations.

Each type of reduction effects a local maximum of this graph (although in general it will not be a strict local maximum). In each case one of the following occurs.

(1) The reduction takes a local maximum of the graph and replaces it with points at the same height or lower.
(2) The reduction removes a local maximum and possibly its successor from the graph. The successor can only be removed if it is at the same height as the predecessor.
(3) The reduction inserts points just after a local maximum that are at the same height or lower.

In each case if the first decrease (i.e. destabilization) is effected then it is replaced with a decrease. If a new increase (i.e. stabilization) is introduced then a new decrease is also introduced before it. The result follows. $\square$

9. An Example

We now present an example suggested by the referee which is illustrative of some of the proof techniques used in the final section.

Let L_1 and L_2 denote two lens spaces. Let T_i denote a Heegaard torus in L_i. Let $\overline{T_1}$ denote the Heegaard torus obtained from T_1 by reversing orientation. Let T denote a Heegaard torus in S^3.

The genus two Heegaard surfaces $T_1 \# T_2$ and $\overline{T_1} \# T_2$ are equivalent after one stabilization in $M_1 \# M_2$. Hence, we may build a SOG $\mathbf{X} = \{(X^i, M^i)\}$ as follows:

- $M^1 = L_1 \cup L_2$, $\text{Thick}(X^1) = \{T_1, T_2\}$, $\text{Thin}(X^1) = \emptyset$.
- $M^2 = L_1 \# L_2$, $\text{Thick}(X^2) = \{T_1 \# T_2\}$, $\text{Thin}(X^2) = \emptyset$
- $M^3 = L_1 \# L_2 \# S^3$, $\text{Thick}(X^3) = \{T_1 \# T_2 \# T\}$, $\text{Thin}(X^2) = \emptyset$
- $M^4 = L_1 \# L_2$, $\text{Thick}(X^4) = \{\overline{T_1} \# T_2\}$, $\text{Thin}(X^2) = \emptyset$
- $M^5 = L_1 \cup L_2$, $\text{Thick}(X^5) = \{\overline{T_1}, T_2\}$, $\text{Thin}(X^1) = \emptyset$.

The GHS X^3 is maximal in $\mathbf{X}$. By Lemma 4.7 it cannot be critical, since M^3 is reducible. Hence, by Lemma 8.9 there must be a reduction for $\mathbf{X}$. Such a reduction is not difficult to find.

Let S denote the summing sphere in $L_1 \# L_2$. Now note that T_1 and $\overline{T_1}$ are equivalent after one stabilization in L_1. Hence, when we do the connected sum with T to form X^3 we may assume that it lies entirely on the L_1 side of S.

The surface $T_1 \# T_2 \# T$ cuts S into disks D and E. Let (A, B) denote the reducing pair which we use to go from X^3 to X^2, and let (A', B') denote the reducing pair pair which we use to go from X^3 to X^4. It

follows that D is disjoint from both B and B'. The fact that $B \cap B' = \emptyset$ implies the sequence $\{A, B, D, B', A'\}$ satisfies the conditions of Definition 8.3. We conclude $\angle(X^3) = 3$. (For it to be any less either $B = B'$, $A = A'$, $A \cap B' = \emptyset$, or $A' \cap B = \emptyset$. None of these are the case.) We may thus apply a reduction of Type II to $\mathbf{G}$. After a sequence of such reductions we are left with the following SOG $\mathbf{Y} = \{(Y^i, N^i)\}$:

- $N^1 = L_1 \cup L_2$, $\mathrm{Thick}(Y^1) = \{T_1, T_2\}$, $\mathrm{Thin}(Y^1) = \emptyset$.
- $N^2 = L_1 \# S^3 \cup L_2$, $\mathrm{Thick}(Y^2) = \{T_1 \# T, T_2\}$, $\mathrm{Thin}(Y^2) = \emptyset$.
- $N^3 = L_1 \cup L_2$, $\mathrm{Thick}(Y^3) = \{\overline{T_1}, T_2\}$, $\mathrm{Thin}(Y^3) = \emptyset$.

10. The Stability Theorem

We now proceed with our proof of Gordon's conjecture.

Theorem 10.1. *Let M_1 and M_2 be closed, orientable 3-manifolds. Suppose H_i is a Heegaard surface in M_i, for $i = 1, 2$. If $H_1 \# H_2$ is a stabilized Heegaard surface in $M_1 \# M_2$ then either H_1 or H_2 is stabilized.*

Proof. Let M_1 and M_2 be two closed, orientable 3-manifolds. Let $M = M_1 \# M_2$. Suppose H_i is an unstabilized Heegaard surface in M_i and let $H = H_1 \# H_2$. By way of contradiction we assume there is a Heegaard surface G in M such that H is a stabilization of G. Let X denote the GHS of $M_1 \cup M_2$ such that $\mathrm{Thick}(X) = \{H_1, H_2\}$ and $\mathrm{Thin}(X) = \emptyset$.

Let $\{M_i^j\}$ denote the irreducible manifolds in a prime decomposition of M_i. (If M_i is the connected sum of copies of $S^2 \times S^1$ then $\{M_i^j\} = \emptyset$.) By [Hak68] H_i is the connected sum of Heegaard splittings H_i^j of M_i^j and Heegaard splittings of copies of $S^2 \times S^1$. It follows that there is a GHS X^1 of $\bigcup M_i^j$ such that $\mathrm{Thick}(X^1) = \{H_i^j\}$, $\mathrm{Thin}(X^1) = \emptyset$, and X^1 is obtained from X by weak reduction. (Heegaard splittings of $S^2 \times S^1$ that crop up during weak reduction quickly disappear. See Example 5.4.) Similarly, the Heegaard surface G is a connected sum of Heegaard splittings G_i^j of M_i^j and Heegaard splittings of copies of $S^2 \times S^1$.

The first step is to build a SOG $\mathbf{X} = \{(X^i, M^i)\}_{i=1}^n$ as follows:

- $M^1 = \bigcup M_i^j$, $\mathrm{Thick}(X^1) = \{H_i^j\}$, $\mathrm{Thin}(X^1) = \emptyset$.
- $M^{i_1} = M_1 \cup M_2$, $\mathrm{Thick}(X^{i_1}) = \{H_1, H_2\}$, $\mathrm{Thin}(X^{i_1}) = \emptyset$.
- $M^{i_2} = M$, $\mathrm{Thick}(X^{i_2}) = \{H\}$, $\mathrm{Thin}(X^{i_2}) = \emptyset$.
- $M^{i_3} = M$, $\mathrm{Thick}(X^{i_3}) = \{G\}$, $\mathrm{Thin}(X^{i_3}) = \emptyset$.
- $M^n = \bigcup M_i^j$, $\mathrm{Thick}(X^n) = \{G_i^j\}$, $\mathrm{Thin}(X^n) = \emptyset$.

For $i < i_2$ in the above SOG the GHS X^i is obtained from X^{i+1} by weak reduction. For $i \geq i_2$ the GHS X^{i+1} is obtained from X^i by weak reduction. Hence X^{i_2} is maximal.

Apply a maximal sequence of weak reductions to X^1 and X^n to obtain strongly irreducible GHSs X^- and X^+. We now extend $\mathbf{X}$ in the natural way to a SOG $\overline{\mathbf{X}}$ such that:

- The first GHS of $\overline{\mathbf{X}}$ is X^- and the last GHS is X^+.
- The SOG $\overline{\mathbf{X}}$ contains $\mathbf{X}$ as a subsequence.

Now apply a maximal sequence of reductions to $\overline{\mathbf{X}}$ to obtain an irreducible SOG $\mathbf{Y} = \{Y^i\}_{i=1}^m$. By Lemma 8.9 the maximal GHSs of $\mathbf{Y}$ are critical. By Lemma 4.7 the manifolds that these are GHSs of are irreducible. It follows that every GHS in $\mathbf{Y}$ is a GHS of an irreducible manifold.

Since X^{i_1} is unstabilized (by assumption) it follows that X^1 is unstabilized. Since X^1 contains a single thick level in each M_i^j it immediately follows that $\mathcal{A}(X^1) = X^1$, and hence $\mathcal{A}(X^1)$ is unstabilized. The GHS X^- is obtained from X^1, a GHS of an irreducible manifold, by weak reductions. As X^1 is unstabilized these weak reductions can not be destabilizations. Hence, by Lemma 7.4 we conclude $\mathcal{A}(X^-) = \mathcal{A}(X^1)$, and so $\mathcal{A}(X^-)$ is unstabilized. Finally, since X^- is strongly irreducible, and $\mathbf{Y}$ is obtained from $\mathbf{X}$ by reductions, $Y^1 = X^-$. Hence $\mathcal{A}(Y^1)$ is unstabilized.

By construction there is a destabilization in $\mathbf{X}$, and this comes before any stabilization. By Lemma 8.10 there is thus an i such that Y^{i+1} is obtained from Y^i by destabilization. Also by Lemma 8.10 if Y^{j+1} is obtained from Y^j by stabilization then $j > i$. It now follows from Lemma 7.4 that $\mathcal{A}(Y^{k-1}) = \mathcal{A}(Y^k)$ for all $k \leq i$. In particular, $\mathcal{A}(Y^i) = \mathcal{A}(Y^1)$, and hence $\mathcal{A}(Y^i)$ is unstabilized. We have now reached a contradiction, since Y^{i+1} is obtained from Y^i by destabilization. $\qquad\square$

Theorem 10.2. *Let M_1 and M_2 be two closed, orientable 3-manifolds. Suppose H_i and G_i are non-isotopic Heegaard splittings of M_i and H_i is unstabilized. Then $H_1 \# H_2$ is not isotopic to $G_1 \# G_2$ in $M_1 \# M_2$.*

Proof. The proof is similar to that of Theorem 10.1. The main difference is that the Heegaard splittings G_1 and G_2 are now given in the hypotheses of the theorem. Also, since there will be no destabilizations in the SOG $\overline{\mathbf{X}}$ the final contradiction is slightly different.

Let $\{M_i^j\}$ denote the irreducible manifolds in a prime decomposition of M_i. By [Hak68] H_i is the connected sum of Heegaard splittings H_i^j of M_i^j and Heegaard splittings of copies of $S^2 \times S^1$. Also by [Hak68] the Heegaard surface G_i is a connected sum of Heegaard splittings G_i^j of M_i^j and Heegaard splittings of copies of $S^2 \times S^1$.

By way of contradiction, we now assume $H_1 \# H_2$ is isotopic to $G_1 \# G_2$ and build a SOG $\mathbf{X}$ as follows:

- $M^1 = \bigcup M_i^j$, Thick$(X^1) = \{H_i^j\}$, Thin$(X^1) = \emptyset$.
- $M^{i_1} = M_1 \cup M_2$, Thick$(X^{i_1}) = \{H_1, H_2\}$, Thin$(X^{i_1}) = \emptyset$.
- $M^{i_2} = M$, Thick$(X^{i_2}) = \{H_1 \# H_2 \simeq G_1 \# G_2\}$, Thin$(X^{i_2}) = \emptyset$.
- $M^{i_3} = M_1 \cup M_2$, Thick$(X^{i_3}) = \{G_1, G_2\}$, Thin$(X^{i_3}) = \emptyset$.
- $M^n = \bigcup M_i^j$, Thick$(X^n) = \{G_i^j\}$, Thin$(X^n) = \emptyset$.

For $i < i_2$ in the above SOG the GHS X^i is obtained from X^{i+1} by weak reduction. For $i \geq i_2$ the GHS X^{i+1} is obtained from X^i by weak reduction.

Apply a maximal sequence of weak reductions to X^1 and X^n to obtain strongly irreducible GHSs X^- and X^+. We now extend $\mathbf{X}$ in the natural way to a SOG $\overline{\mathbf{X}}$ such that:

- The SOG $\overline{\mathbf{X}}$ contains $\mathbf{X}$ as a subsequence.
- The first GHS of $\overline{\mathbf{X}}$ is X^- and the last GHS is X^+.

Now apply a maximal sequence of reductions to $\overline{\mathbf{X}}$ to obtain an irreducible SOG $\mathbf{Y} = \{Y^i\}_{i=1}^m$. Note that $Y^1 = X^-$ and $Y^m = X^+$. As in the proof of Theorem 10.1 it follows from Lemmas 8.9 and 4.7 that every GHS in $\mathbf{Y}$ is a GHS of an irreducible manifold.

Since X^{i_1} is unstabilized (by assumption) it follows that X^1 is unstabilized. Since X^1 contains a single thick level in each M_i^j it immediately follows that $\mathcal{A}(X^1) = X^1$. The GHS X^- is obtained from X^1, a GHS of an irreducible manifold, by weak reductions that are not destabilizations. By Lemma 7.4 we conclude $\mathcal{A}(X^-) = \mathcal{A}(X^1)$. Finally, since X^- is strongly irreducible, and $\mathbf{Y}$ is obtained from $\mathbf{X}$ by reductions, $Y^1 = X^-$. Hence $\mathcal{A}(Y^1) = \mathcal{A}(X^-) = \mathcal{A}(X_1) = X^1$.

By construction there are no stabilizations in $\mathbf{X}$. Hence, by Lemma 8.10 if there is a stabilization in $\mathbf{Y}$ it comes after a destabilization. Suppose now there is an i such that Y^{i+1} is obtained from Y^i by destabilization. It now follows from Lemma 7.4 that $\mathcal{A}(Y^{k-1}) = \mathcal{A}(Y^k)$ for all $k \leq i$. In particular, $\mathcal{A}(Y^i) = \mathcal{A}(Y^1)$, and hence $\mathcal{A}(Y^i)$ is unstabilized. We have now reached a contradiction, since Y^{i+1} is obtained from Y^i by destabilization. We conclude that there are no destabilizations, and hence no stabilizations, in $\mathbf{Y}$. But now it follows from Lemma 8.10 that there were no destabilizations in $\mathbf{X}$. We may thus conclude, as above, that $\mathcal{A}(Y^m) = \mathcal{A}(X^+) = \mathcal{A}(X^n) = X^n$.

It also follows from Lemma 7.4 that $\mathcal{A}(Y^i)$ is the same for all i, and hence $\mathcal{A}(Y^1) = \mathcal{A}(Y^m)$. Finally, this gives us $X^1 = X^n$, and the result follows. $\qquad\square$

REFERENCES

[Bac02] D. Bachman. Critical Heegaard surfaces. *Trans. Amer. Math. Soc.*, 354(10):4015–4042 (electronic), 2002.

[Bir74]   Joan S. Birman. *Braids, links, and mapping class groups*. Princeton University Press, Princeton, N.J., 1974. Annals of Mathematics Studies, No. 82.

[BO83]   F. Bonahon and J. P. Otal. Scindements de Heegaard des espaces lenticulaires. *Ann. scient. Éc. Norm. Sup.*, 16:451–466, 1983.

[Cer68]   J. Cerf. *Les diffeomorphismes de la sphere de dimension trois ($\Gamma(4) = 0$)*. Springer Verlag, 1968. Lecture Notes in Mathemaitcs #53.

[CG87]   A. J. Casson and C. McA. Gordon. Reducing Heegaard splittings. *Topology and its Applications*, 27:275–283, 1987.

[Eng70]   Renate Engmann. Nicht-homöomorphe Heegaard-Zerlegungen vom Geschlecht 2 der zusammenhängenden Summe zweier Linsenräume. *Abh. Math. Sem. Univ. Hamburg*, 35:33–38, 1970.

[Gab87]   D Gabai. Foliations and the topology of three-manifolds iii. *J. Diff. Geom.*, 26:479–536, 1987.

[Hak68]   W. Haken. *Some Results on Surfaces in 3-Manifolds*. M.A.A., Prentice Hall, 1968.

[Kir97]   Rob Kirby. Problems in low-dimensional topology. In *Geometric topology (Athens, GA, 1993)*, volume 2 of *AMS/IP Stud. Adv. Math.*, pages 35–473. Amer. Math. Soc., Providence, RI, 1997.

[MM00]   H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.

[MS98]   Y. Moriah and J. Schultens. Irreducible Heegaard splittings of Seifert fibered spaces are either vertical or horizontal. *Topology*, 37:1089–1112, 1998.

[Qiu]   Ruifeng Qiu. Stabilizations of Reducible Heegaard Splittings. Preprint. Available at `math.GT/0409497`.

[RS96]   H. Rubinstein and M. Scharleman. Comparing Heegaard Splittings of non-Haken 3-Manifolds. *Topology*, 35:1005–1026, 1996.

[ST94]   M. Scharlemann and A. Thompson. Thin position for 3-manifolds. *A.M.S. Contemporary Math.*, 164:231–238, 1994.

[Wal68]   F. Waldhausen. Heegaard Zerlegungen der 3-sphäre. *Topology*, 7:195–203, 1968.

Pitzer College, 1050 N. Mills Ave, Claremont, CA 91711
E-mail address: `bachman@pitzer.edu`