

Connected sums of unstabilized Heegaard splittings are unstabilized

DAVID BACHMAN

Let M_1 and M_2 be closed, orientable 3–manifolds. Let H_i denote a 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 one of H_1 or H_2 comes from stabilizing a lower genus splitting. This answers a question of C Gordon [9, Problem 3.91]. We also show that every unstabilized Heegaard splitting has a unique expression as the connected sum of Heegaard splittings of prime 3–manifolds.

57N10; 57M27

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 Gordon conjectured that the converse must also be true (see [9, Problem 3.91]).

Conjecture 1.1 (Gordon’s Conjecture) *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). This result has also been announced by R Qiu [12] (see also Scharlemann’s exposition of Qiu’s proof [14]). In addition, we prove that the prime decomposition of unstabilized Heegaard splittings is unique.

Theorem (Theorem 10.2) *Let M be a closed, orientable 3–manifold. Let H be an unstabilized Heegaard splitting of M . Then H has a unique expression as the connected sum of Heegaard splittings of prime 3–manifolds.*

Both Theorem 10.1 and Theorem 10.2 have algebraic analogues established by R Weidmann [19]. In Weidmann's work, the analogue of an unstabilized Heegaard splitting is a generating set (g_1, \dots, g_k) for a group that is not Nielsen equivalent to a generating set of the form $(g'_1, \dots, g'_{k-1}, 1)$. The analogue of connected sum of Heegaard splittings is free product of groups.

The surfaces H_1 and H_2 of Gordon's conjecture together form a structure called a *generalized Heegaard splitting (GHS)*. Loosely speaking, H is a GHS if it 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 $M \setminus \text{Thin}(H)$ (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 5.13 gives a simple complexity for GHSs under which weak reduction represents a decrease.

A coarser measure of complexity of a GHS is its *irreducible genus*. Each GHS of a connected 3-manifold M defines a Heegaard splitting in a unique way. The irreducible genus is the genus of this Heegaard splitting, minus the number of $S^2 \times S^1$ summands in a prime decomposition of M . The key fact that we use here is given by Corollary 7.9, which says that weak reduction of an unstabilized GHS preserves irreducible genus.

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 of GHSs $\{H^i\}$ such that for each i either H^{i+1} or H^i can be obtained from the other by weak reduction. Now, given fixed GHSs H and H' and a SOG which connects them, one can ask if 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*. Most reductions involve a process of "peak reduction." That is, we examine maximal GHSs of a SOG and determine when we can swap the order of weak reductions in the sequence, trading a maximal GHS for possibly several smaller ones. An analysis of when one can switch the order of weak reductions in a SOG is done in Section 6.

Section 8 contains 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 [4]. Criticality was introduced by the author in [1], 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 essential spheres and disks (Theorem 3.5).

Our proof of Gordon’s Conjecture (Theorem 10.1) begins with the construction of a SOG $\{Y^i\}_{i=1}^n$ as follows (see Figure 18). The Heegaard surfaces H_1 and H_2 of the conjecture can be considered to be the thick levels of a single GHS Y^1 of the manifold $M_1 \cup M_2$. Similarly, $H = H_1 \# H_2$ is the unique thick level of a GHS Y^2 of $M = M_1 \# M_2$. We can obtain Y^1 from Y^2 by a weak reduction. By way of contradiction, we can also obtain a GHS Y^3 from Y^2 by destabilizing H . The three GHSs Y^1, Y^2 and Y^3 can then be used to build a SOG \mathbf{Y} . Note that the irreducible genus of Y^1 is strictly larger than the irreducible genus of Y^3 .

By construction, the GHS Y^2 is the single maximal GHS in \mathbf{Y} . As this is a GHS of a reducible manifold, by Lemma 4.7 and Lemma 8.9 there must be a way to reduce the entire SOG \mathbf{Y} . Reducing as much as possible yields a SOG \mathbf{X} in which every element is a GHS of an irreducible 3–manifold.

Since the irreducible genus of the GHSs of \mathbf{Y} was monotonically decreasing, Lemma 8.12 implies that there is a value of k such that the irreducible genus of X^l is constant for $l \leq k$ and drops from X^k to X^{k+1} . As X^1 is unstabilized, it follows from Corollary 7.5 that X^k is unstabilized as well. (Here is where we use the fact that each element of \mathbf{X} is a GHS of an irreducible manifold.) But this implies that the irreducible genus of X^k can not drop from X^k to X^{k+1} , a contradiction.

The author would like to thank Saul Schleimer for helpful conversations regarding the proof of Claim 3.10 and the referee for countless extremely helpful suggestions, including a restructuring of the proof of Lemma 7.2.

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 an orientable 3-manifold M , D is a disk in M and $D \cap F = \partial D$. There is an embedding $h: D \times I \rightarrow M$ 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 . If D and E are disjoint disks that both have boundary on F , then we let F/DE denote the surface obtained from F by simultaneous surgery along D and E .

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 $h: D \times I \rightarrow M$ such that $h(D \times \{\frac{1}{2}\}) = D$, $h(D \times I) \cap F = h(\alpha \times I)$ 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 oriented n -manifolds (with $n \leq 3$) 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 , which will inherit some induced orientation from the orientation of M_i . The manifold $M_1 \# M_2$ is then obtained by identifying S_1 with S_2 via an orientation reversing map. 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 that are not contained in $H \times \{0\}$ 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 a transversally 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 transversally oriented in the above definition will play an important role. For example, $L(p, q)$, where $q \not\equiv \pm 1 \pmod{p}$, contains a pair of non-isotopic Heegaard tori (Bonahon and Otal [3]), but as unoriented surfaces they are isotopic. In contrast, S^3 contains a unique Heegaard torus (Waldhausen [18]), 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 makes the transverse orientation of H_1^* compatible 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 (Engmann [6] see also Birman [2]).

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* [4] and *critical* [1] surfaces. Both strong irreducibility 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 *weak 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 weak reducing pairs.

The restriction to surfaces that are compressible to both sides rules out various possibilities, such as trivial Heegaard splittings of compression bodies and product manifolds.

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 the following hold.

- (1) For each $i = 0, 1$ there is at least one weak 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 weak reducing pair.

Definition 3.3 is significantly simpler and slightly weaker, than the one given in [1]. In other words, anything that was considered critical in [1] is considered critical here as well. Hence, a result such as [1, Theorem 7.1] 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 weak 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 a two-sphere with three exceptional fibers. There are three vertical splittings H, H' and H'' of M and these are generally not isotopic (see Moriah and Schultens [11] 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 [1, Theorem 7.1] 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 Heegaard splitting obtained by destabilization. This point is made more explicit in the definition of criticality given in [1]. Somehow the useful results (such as Theorem 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.

Theorem 3.5 *If a 3–manifold M admits a strongly irreducible or critical Heegaard splitting then M is irreducible and ∂M is incompressible.*

Proof Let S be an essential sphere or disk in M and H be a Heegaard surface. We break the proof into two cases, depending on whether H is strongly irreducible or critical.

Case 1 *H is strongly irreducible.* The strongly irreducible case is a direct consequence of the “Haken Lemma” [8]. For completeness (and perhaps of independent interest) we present a proof here.

Let H be a strongly irreducible Heegaard surface, separating M into compression bodies \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 supported in this 3-ball 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 (ie a map $\gamma: S \times [-1, 1] \rightarrow M$ where $S_t = \gamma(S, t)$) such that

- $S_{-1} \cap V = \emptyset$,
- $S_0 = S$ and
- $S_1 \cap W = \emptyset$.

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} .*

The proof is essentially the same as that of Gabai [7, Lemma 4.4].

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 $S_{t_+} \cap H$. 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.

Claim 3.9 *Every loop of $S_t \cap H$ is inessential on H .*

Proof 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 . Our goal is to show the loop α bounds a compressing disk for H , which contradicts the fact that t lies in an unlabeled interval.

As S_t is a sphere or disk the loop α bounds a subdisk A of S_t so that all curves of $\text{int}(A) \cap H$ 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 . Note that $\gamma \neq \alpha$, as α is essential on H and γ is not. The loop 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 $S_t \cap H$ consists of loops that are inessential on both surfaces. Let γ now be a loop of $H \cap S_t$ that is innermost on H , in the sense that γ bounds a subdisk $C \subset H$ whose interior is disjoint from S_t . Then we may use the disk C to surger S_t to obtain S'_t and S''_t . At least one of these surfaces will also be essential and will meet H fewer times. Continuing in this way we obtain an essential sphere or disk that is disjoint from H and thus lies in a compression body. If the surface is an essential sphere, we contradict the fact that compression bodies are irreducible. If the surface is a disk, we contradict the fact that the negative boundary of a compression body is incompressible.

Case 2 H is critical. In this case Theorem 3.5 essentially follows from [1, Theorem 5.1]. 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 weak 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 is homotopy equivalent to a circle 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 Rubinstein and Scharlemann [13], although the details have little in common.

3.1 Constructing the map, $\Phi: S \times D^2 \rightarrow M$

We begin by defining a two-parameter family of surfaces in M isotopic to S . For any map $\Phi: S \times D^2 \rightarrow M$ 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 $\Phi: S \times D^2 \rightarrow M$ 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 $V^i \cap V^{i+1} = \emptyset$ 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 $m - 1$.

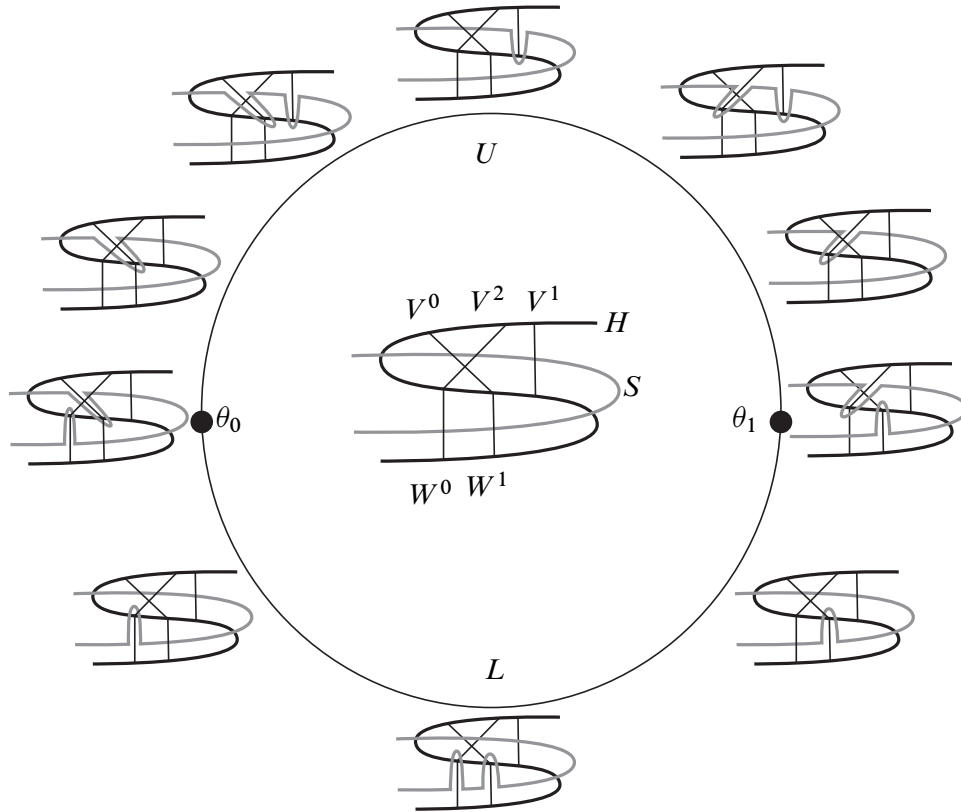


Figure 1

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 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 $\gamma^i: \overline{A^i} \times I \rightarrow \overline{A^i}$ be an isotopy which pushes S off of V^i , inside A^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 .

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 2¹. For the i th pair of arcs chosen, centered on a point of U , define $f^i: D^2 \rightarrow [0, 1]$ to be the continuous function depicted in Figure 3. Let $g^j: D^2 \rightarrow [0, 1]$ be the function similarly defined for the j th pair of arcs centered on a point of L .

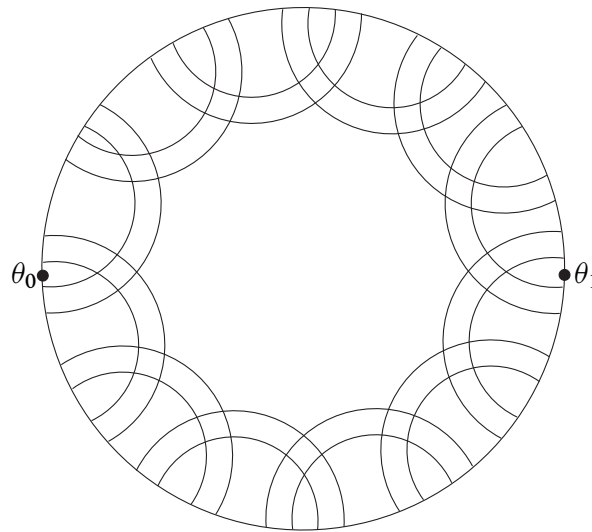


Figure 2: “Linked” pairs of circular arcs in D^2 .

Finally, we define $\Phi: S \times D^2 \rightarrow M$. 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 $\Phi_p(x) = x$.

¹Figure suggested by Saul Schleimer.

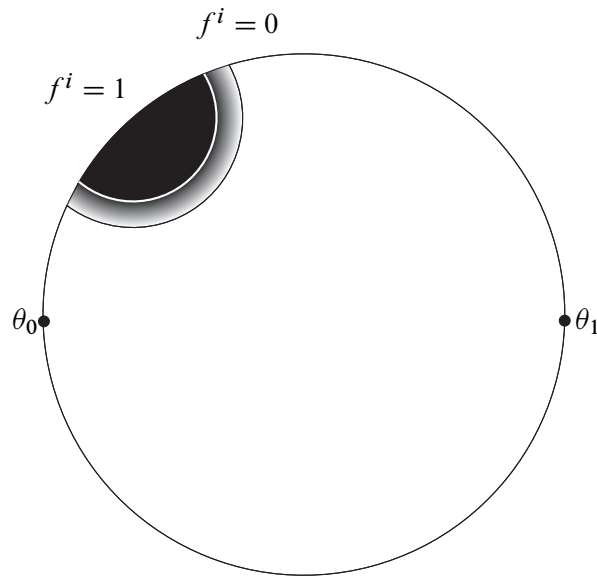


Figure 3: In the black region f^i takes on the value 1. In the white region $f^i = 0$. The shading between these two regions is meant to indicate that between the arcs the function f^i continuously varies from 0 to 1.

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 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 in $\{V^i\}$ 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 $\Sigma = \{x \in D^2 | S_x \text{ is not transverse to } H\}$. (This set is called the *graphic* in [13].) If Φ is in general position with respect to H , then Cerf theory (see [5]) 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 $V \in C_1$ or
- \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 Theorem 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 an argument identical to that which appears in the last paragraph of Case 1. The conclusion of Theorem 3.5 thus follows.

The argument as to why every loop of $H \cap S_x$ is inessential on both surfaces is identical to that of Claim 3.9. 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. □

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 weak reducing pair. \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 $p: I \rightarrow D^2$ be an embedded path connecting x_0 to x_1 meeting Σ once, transversally. 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 (ie 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

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

as in 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 $W_t \in C_1$. \square

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

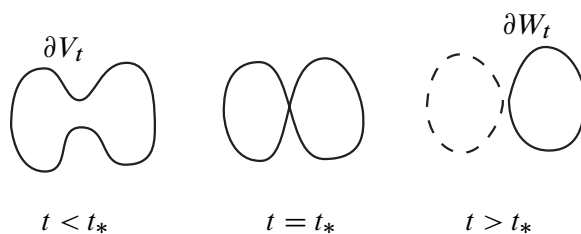


Figure 4

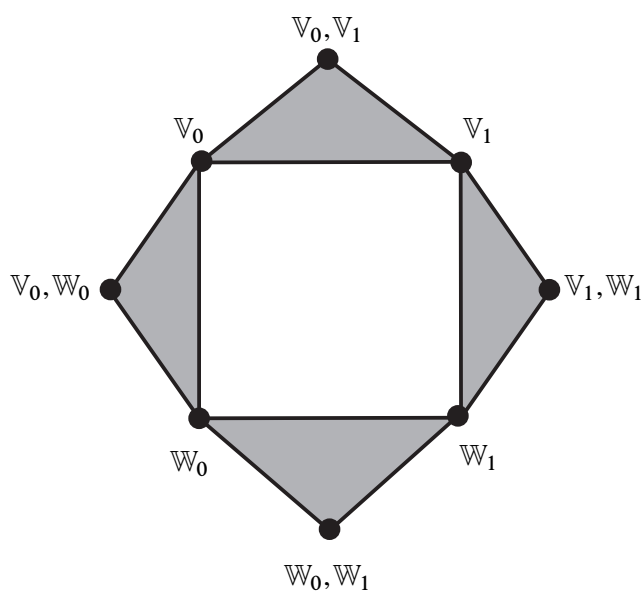


Figure 5: The 2-complex, Π .

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.

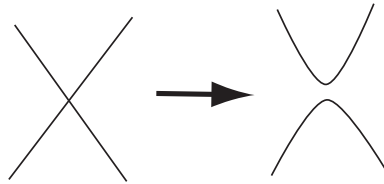


Figure 6: Resolving a vertex of $S_{x_*} \cap H$.

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 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. \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 weak reducing pair. The proof is complete by symmetry. \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. \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 a homotopy equivalence, a contradiction. We conclude that there must have been an unlabeled region, as this was the only assumption we made in constructing the map. Theorem 3.5 now follows from Claim 3.11. \square

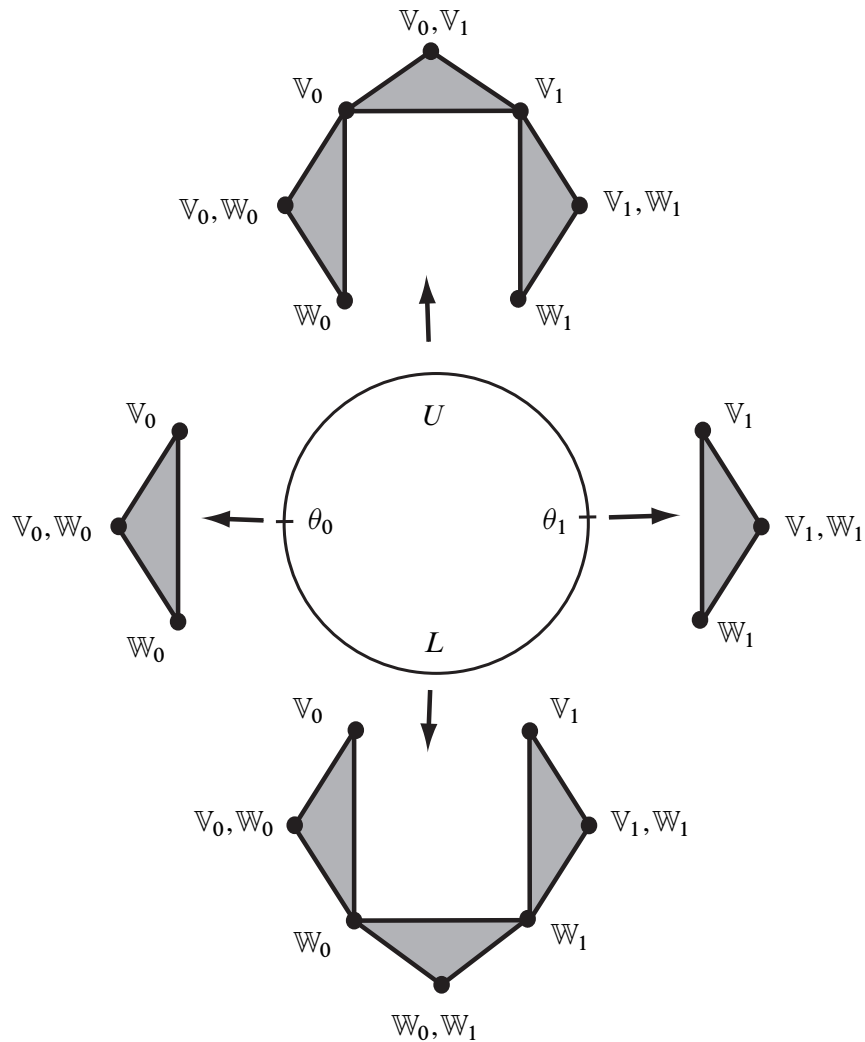


Figure 7: The map from ∂D^2 to Π is non-trivial on homology.

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, transversally oriented, connected surfaces, $\text{Thick}(H)$ and $\text{Thin}(H)$ (called the *thick levels* and *thin levels*, respectively), which satisfies 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 transversally oriented, 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_- . 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 transverse orientations are chosen for H_+^1 and H_+^2 that 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.

Example 4.4 The original formulation of a GHS was given by Scharlemann and Thompson in [15]. 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.

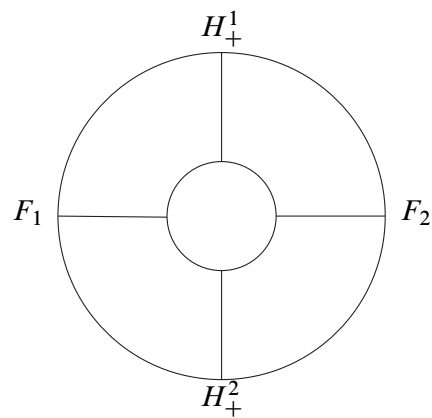


Figure 8: Transverse orientations on H_+^1 and H_+^2 are given by normal vectors. If the orientations on F_1 and F_2 are chosen to be consistent with condition (2) of Definition 4.1 (as pictured) then they will be inconsistent with condition (3).

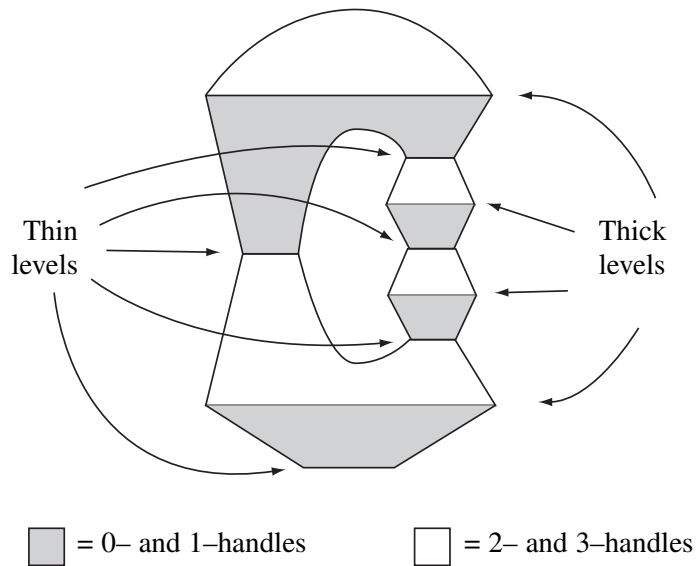


Figure 9: The GHS associated with a handle structure.

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 done in [15].

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. If any loop of $D \cap \text{Thin}(H)$ is inessential on $\text{Thin}(H)$ then let β denote such a loop that is innermost on $\text{Thin}(H)$. Then β bounds a subdisk C of $\text{Thin}(H)$ that we can use to surger D , lowering $|D \cap \text{Thin}(H)|$.

We conclude that every loop of $D \cap \text{Thin}(H)$ is essential on $\text{Thin}(H)$. Let α now denote such a loop that is innermost on D . 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 $M(H_+)$. This now contradicts Theorem 3.5, as A is a compressing disk for $\partial M(H_+)$. \square

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

Proof Let S be an essential sphere in M . 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 (possibly disconnected) incompressible surface in M . Let γ be a loop of $\text{Thin}(H) \cap S$ that is innermost on $\text{Thin}(H)$, in the sense that γ bounds a subdisk $C \subset \text{Thin}(H)$ whose interior is disjoint from S . Then we may use the disk C to surger S to obtain S' and S'' . At least one of these surfaces will also be essential and will meet $\text{Thin}(H)$ fewer times. Continuing in this way we obtain an essential sphere that is disjoint from $\text{Thin}(H)$ and thus lies in $M(H_+)$, for some $H_+ \in \text{Thick}(H)$. This now contradicts Theorem 3.5. \square

Definition 4.8 *A stabilization of a GHS G is a new GHS, obtained from G by forming the connected sum of some element $G_+ \in \text{Thick}(G)$ with the standard genus one Heegaard surface in S^3 .*

Note that in the case where G has a unique thick level, this definition is consistent with Definition 2.5.

5 Reducing GHSs

We now define *weak reduction*: a way to take a GHS of a 3-manifold M which is not strongly irreducible and obtain a “simpler” GHS. The new GHS will be of a manifold which is either homeomorphic to M , or is obtained from M by cutting along a 2-sphere and capping off the resulting boundary components with a pair of 3-balls. The procedure is broken into two stages, called *pre-weak reduction* and *cleaning*. Half-way between the two we have a structure called a *pseudo-GHS*, defined as follows.

Definition 5.1 Let M be a compact, orientable 3-manifold with no 3-sphere components. Let $G = \{T(G), t(G)\}$ be a pair of sets of transversally oriented, connected surfaces in M such that the elements of $T(G) \cup t(G)$ are pairwise disjoint. Then we say G is a *pseudo-GHS* if the following hold.

- (1) Each component M' of $M - t(G)$ meets exactly one element G_+ of $T(G)$. We denote the closure of M' as $M(G_+)$.
- (2) Each element $G_+ \in T(G)$ separates $M(G_+)$ into punctured (possibly trivial) compression bodies \mathcal{W} and \mathcal{W}' , where $\partial_+ \mathcal{W} = \partial_+ \mathcal{W}' = G_+$.
- (3) There is a partial ordering of the elements of $t(G)$ that satisfies similar properties to the partial ordering of the thin levels of a GHS given in Definition 4.1.

We now define the first stage of weak reduction.

Definition 5.2 Let M be a compact, orientable 3-manifold with no 3-sphere components. Let H be a pseudo-GHS. Let D and E be disjoint disks in the complement of $t(H)$ which meet $T(H)$ in their boundaries. If H_+ is the union of the surfaces in $T(H)$ that meet $\partial D \cup \partial E$ then D and E are assumed to be on opposite sides of H_+ . (Hence, if H_+ is connected and ∂D and ∂E are essential on this surface then (D, E) is a weak reducing pair for H_+ .) Define

$$\begin{aligned} T(G) &= (T(H) - \{H_+\}) \cup \{H_+/D, H_+/E\}, \text{ and} \\ t(G) &= t(H) \cup \{H_+/DE\}. \end{aligned}$$

We say the pseudo-GHS $G = \{T(G), t(G)\}$ is obtained from H by *pre-weak reduction* along (D, E) .

Pre-weak reduction is illustrated in Figure 10.

Definition 5.3 Let M be a compact, orientable 3-manifold with no 3-sphere components. Let $H = \{T(H), t(H)\}$ be a pseudo-GHS in M . We say the pseudo-GHS,

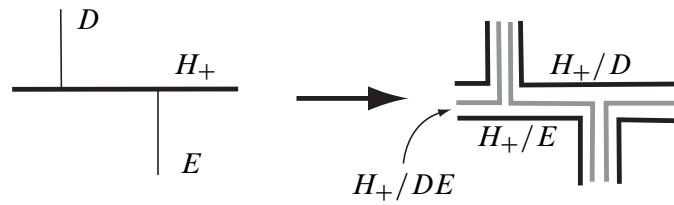


Figure 10: Pre-weak reduction.

along with the 3-manifold M' , given by the following algorithm is obtained from H by *cleaning*. The algorithm starts by initially setting $M' = M$. We then follow these steps.

- (1) For each sphere $S \in t(H)$, cut M' along S , cap off the resulting sphere boundary components with 3-balls and remove S from $t(H)$.
- (2) If M' now contains 3-sphere components then delete them, along with all elements of $T(H)$ and $t(H)$ that they contain. (So if $M \cong S^2 \times S^1$ then cleaning may delete the entire manifold. See Example 5.11.)
- (3) If there are elements $H_+ \in T(H)$ and $H_- \in t(H)$ that cobound a product region P of M' such that $P \cap T(H) = H_+$ and $P \cap t(H) = H_-$ then remove H_+ from $T(H)$ and H_- from $t(H)$.

Lemma 5.4 *Cleaning a pseudo-GHS produces a GHS.*

Proof Suppose a pseudo-GHS G has been obtained from H by cleaning. As the surfaces of $t(G)$ are appropriately partially ordered, we must only check that each element G_+ of $T(G)$ bounds a non-trivial compression-body to both sides in $M(G_+)$. As G is a pseudo-GHS, G_+ bounds punctured, possibly trivial compression bodies on both sides in $M(G_+)$. But the existence of punctures implies that there are sphere elements of $t(G)$. Hence, G could not have been obtained from cleaning. If there are no punctures, but some compression body is trivial, then a surface of $t(G)$ and a surface of $T(G)$ cobound a product region that should have been removed during cleaning. \square

It will be a useful fact that the order one performs the steps of cleaning often does not change the final result. We prove this now.

Lemma 5.5 *Let M be a compact, orientable 3-manifold with no 3-sphere components. Let $H = \{T(H), t(H)\}$ be a pseudo-GHS in M . Suppose that for some*

$H_+ \in T(H)$ there is a surface $H_- \subset \partial M(H_+)$ such that H_+ and H_- cobound a product region. Removing H_+ and H_- from $T(H)$ and $t(H)$, then cleaning the resulting pseudo-GHS, yields the same pseudo-GHS as just cleaning H .

Proof If H_+ is not a sphere, then it will be unaffected if we cut M along a sphere component S of $t(H)$ and then cap off the resulting boundaries by 3-balls (Step 1). If $H_+ \cup H_-$ then ends up in a 3-sphere component of the resulting manifold, then it does not matter if we had first removed them from $T(H)$ and $t(H)$, or if we remove the whole S^3 component first.

The only remaining possibility, then, is that H_+ and H_- are spheres. Since H is a pseudo-GHS, H_+ divides $M(H_+)$ into two punctured 3-balls. Hence, all of the components of $\partial M(H_+)$ are spheres. If we first perform Step 1 of Definition 5.3 then $M(H_+)$ will get cut out of M and all of its boundary components will get capped off by 3-balls. The result will be S^3 and so will be removed in Step 2. This yields the same result as if we first removed H_+ and H_- from $T(H)$ and $t(H)$, cut M along all of the boundary components of $M(H_+)$ that were not H_- and capped off the resulting boundaries by 3-balls. See Figure 11. \square

Definition 5.6 Let M be a compact, orientable 3-manifold with no 3-sphere components. Let H be a pseudo-GHS of M . Let D and E be disjoint disks in the complement of $t(H)$, with boundaries on elements of $T(H)$. We say the GHS G of the 3-manifold M' given by pre-weak reduction along (D, E) , followed by cleaning, is obtained from H by *weak reduction* along (D, E) .

One thing that may seem curious to the experts about the way we have defined weak reduction is that there is no assumption that D or E are compressing disks, or that their boundaries are on the same surface. This technical point will be useful in the next section. For now, however, we prove that nothing has been lost by stating things in such generality.

Definition 5.7 Let D and E be disks as given in Definition 5.6. If ∂D and ∂E are essential loops on the same element of $T(H)$ then we say (D, E) is a *weak reducing pair* for the pseudo-GHS H .

Lemma 5.8 Let M be a compact, orientable 3-manifold with no 3-sphere components. Let H be a pseudo-GHS of M . Suppose G has been obtained from H by weak reduction along (D, E) . If (D, E) is not a weak reducing pair then G can be obtained from H by cleaning.

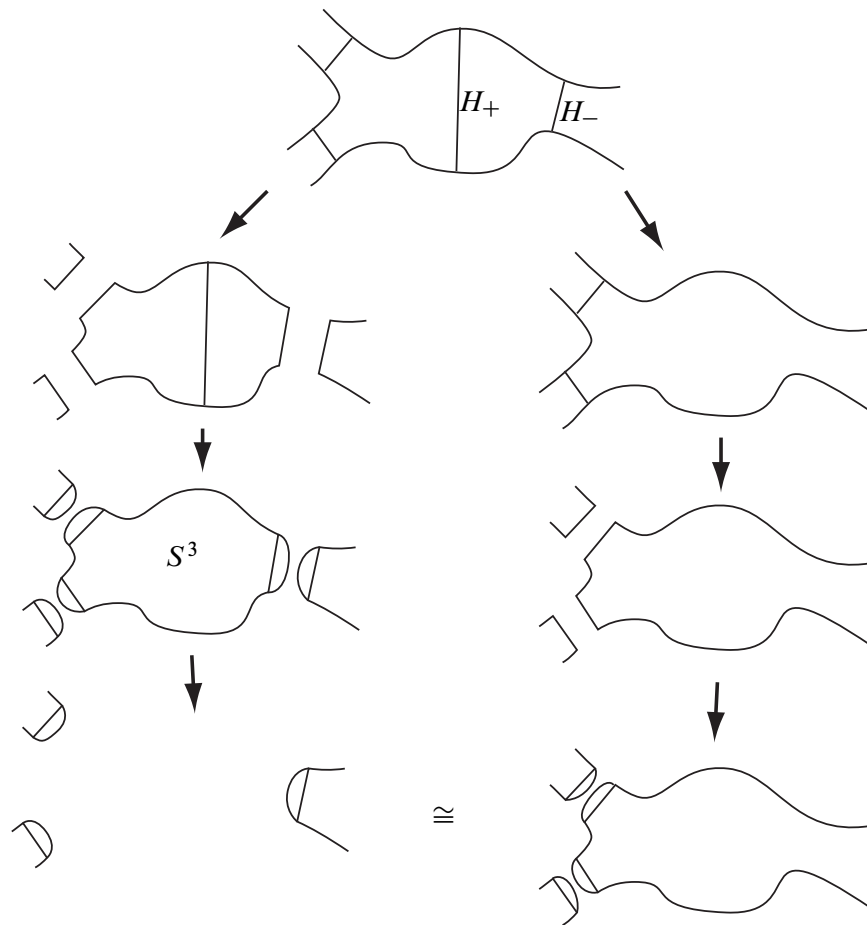


Figure 11: On the left we cut along the components of $\partial M(H_+)$, cap off with 3-balls and remove S^3 . On the right we remove H_+ and H_- , and then cut and cap off. The results are homeomorphic.

It follows that if H is a GHS and (D, E) is not a weak reducing pair then weak reducing along (D, E) has no effect.

Proof There are two cases, with very similar proofs. Suppose first that $H_D \in T(H)$, the surface that contains ∂D , is not the same as $H_E \in T(H)$, the surface that contains ∂E . Let $H_+ = H_D \cup H_E$. Then $H_+/D = (H_D/D) \cup H_E$ and $H_+/E = H_D \cup$

(H_E/E) . Hence, pre-weak reduction produces the sets

$$\begin{aligned} T(G) &= (T(H) - \{H_+\}) \cup \{H_+/D, H_+/E\} \\ &= (T(H) - \{H_D, H_E\}) \cup \{H_D/D, H_E, H_D, H_E/E\} \\ t(G) &= t(H) \cup \{H_+/DE\} \\ &= t(H) \cup \{H_D/D, H_E/E\}. \end{aligned}$$

By Lemma 5.5, we may assume the first step of cleaning is the removal of repetitions between $T(H)$ and $t(H)$. This leaves

$$\begin{aligned} T(G) &= (T(H) - \{H_D, H_E\}) \cup \{H_E, H_D\} \\ t(G) &= t(H). \end{aligned}$$

This is the original pseudo-GHS H . Hence, completion of cleaning by removing any remaining spheres and parallelisms between $t(G)$ and $T(G)$ will produce the same thing as cleaning H .

The second case is when $H_D = H_E = H_+$, but ∂D (say) is not essential on H_+ . The loop ∂D thus bounds a disk $D' \subset H_+$. The surface H_+/D has two components. One of these is the sphere $D \cup D'$ and the other is a surface H'_+ .

Pre-weak reduction produces the sets

$$\begin{aligned} T(G) &= (T(H) - \{H_+\}) \cup \{H_+/D, H_+/E\} \\ &= (T(H) - \{H_+\}) \cup \{D \cup D', H'_+, H_+/E\} \\ t(G) &= t(H) \cup \{H_+/DE\} \\ &= t(H) \cup \{D \cup D', H'_+/E\}. \end{aligned}$$

Let M' denote the manifold obtained from M by cutting along the sphere $D \cup D'$ and capping off the resulting boundary components with 3-balls. There are three cases for the sphere $D \cup D'$. Either it bounds a ball B in M which is disjoint from H , it bounds a ball B in M which meets H , or it is essential in M . In the first case the surfaces H'_+ and H'_+/E are isotopic to H_+ and H_+/E in M' . In the latter two cases there is a collection of spheres $S \subset t(H)$. Cutting along S and capping off the resulting boundary components yields M' as well and again the surfaces H'_+ and H'_+/E are isotopic to H_+ and H_+/E in M' . Hence, in all cases after cleaning H we may assume H'_+ and H'_+/E are the same as H_+ and H_+/E in M' .

This gives us

$$\begin{aligned} T(G) &= (T(H) - \{H_+\}) \cup \{H_+, H_+/E\} \\ t(G) &= t(H)^* \cup \{H_+/E\} \end{aligned}$$

where $t(H)^*$ is either the original set $t(H)$, or the one obtained from it by removing the collection of spheres S .

By Lemma 5.5 we may assume repetitions between $T(G)$ and $t(G)$ are removed next. (That is, before we cut along any more spheres.) This leaves

$$\begin{aligned} T(G) &= (T(H) - \{H_+\}) \cup \{H_+\} \\ t(G) &= t(H)^*. \end{aligned}$$

This is either the original pseudo-GHS H , or the one obtained from H by cutting M along S and capping off the resulting boundary components. \square

Example 5.9 Suppose $\text{Thick}(H)$ contains a single element H_+ and $\text{Thin}(H) = \emptyset$, so that H_+ is a Heegaard surface for M and M is not S^3 . Suppose (D, E) is a weak reducing pair 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 performs a destabilization.

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 appear at thin levels during weak reduction.

Example 5.10 Let M_1 and M_2 be 3-manifolds with genus one Heegaard splittings, at least one of which is not 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 the union of whose boundaries do not separate H_+ . Hence the pair (D, E) is a weak reducing pair. During pre-weak reduction we add H_+/DE to $\text{Thin}(H)$. But this surface is a sphere, which will get removed when cleaning.

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.

Example 5.10 is interesting because it says that there are weak reductions (D, E) that perform destabilizations, where ∂D is not parallel to ∂E .

In Example 5.9 and Example 5.10 the surface H_+/DE had a separating sphere component. In the next two examples we explore what happens when H_+/DE contains a non-separating sphere.

Example 5.11 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.

Our next example is, in some sense, a combination of the situations described in Example 5.9 and Example 5.11.

Example 5.12 Suppose, as in Example 5.9, that $\text{Thick}(H)$ contains a single element H_+ and $\text{Thin}(H) = \emptyset$, so that H_+ is a Heegaard surface for M . Suppose again that (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 non-separating in M then either $M \cong M' \# (S^2 \times S^1)$ or $M \cong S^2 \times S^1$. In the first case the weak reduction produces a Heegaard splitting of M' whose genus is one less than the genus of H_+ . In the second case weak reduction 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 [15].

Definition 5.13 If H is a GHS then let $c(H)$ denote the set of genera of the elements of $\text{Thick}(H)$, where repeated integers are included and the set is put in non-increasing order. We compare two such sets lexicographically. If G and H are two GHSs then we say $G < H$ if $c(G) < c(H)$.

Lemma 5.14 Suppose G and H are GHSs and G is obtained from H by a weak reduction. Then $G < H$.

Proof In the definition of weak reduction we remove H_+ from $\text{Thick}(H)$ 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 prove an important lemma that tells us when we can do this. Before proceeding to this we will need the following result.

Lemma 6.1 Suppose $H = \{T(H), t(H)\}$ is a pseudo-GHS of M . Then cleaning H , followed by a weak reduction along (D, E) , yields the same result as pre-weak reducing along (D, E) , followed by cleaning.

Proof The assertion of the lemma only has content when H is not clean. That is, there is a sphere in $t(H)$ or a surface in $T(H)$ that is parallel to a surface in $t(H)$, with no other elements of $t(H) \cup T(H)$ between them. We address each of these cases independently.

- (1) There is a sphere S in $t(H)$. If D or E meets S then (D, E) is not a weak reducing pair. Thus, by Lemma 5.8 pre-weak reduction, followed by cleaning, yields the same result as just cleaning. On the other hand, if neither D nor E meet S then pre-weak reduction along (D, E) will not effect S . During cleaning M will be cut along S and the resulting boundary components will be capped off by 3–balls. This can happen before or after pre-weak reduction with the same result.
- (2) There is a surface $H_+ \in T(H)$ which is parallel to a surface $H_- \in t(H)$, with no other elements of $t(H) \cup T(H)$ between them. Then H_+ is incompressible in the submanifold of M cobounded by H_+ and H_- . It follows that if (D, E) is a weak reducing pair then neither disk meets H_+ and thus the removal of $H_+ \cup H_-$ can happen before or after pre-weak reduction along (D, E) . On the other hand, if (D, E) is not a weak reducing pair then Lemma 5.8 implies the pre-weak reduction, followed by cleaning, yields the same result as just cleaning. \square

Lemma 6.2 Suppose (D, E) and (D', E) are weak reductions for a GHS H , where $D \cap D' = \emptyset$. Let G be the GHS obtained from H by the weak reduction (D, E) and G' the GHS obtained from H by the weak reduction (D', E) . Then one of the following holds:

- (1) there is a GHS F that can be obtained from G by the weak reduction (D', E) and from G' by the weak reduction (D, E) ,
- (2) G can be obtained from G' by the weak reduction (D, E) ,
- (3) G' can be obtained from G by the weak reduction (D', E) or
- (4) G is isotopic to G' .

Proof Let $\{T(G), t(G)\}$ be the pseudo-GHS obtained from H by pre-weak reducing along (D, E) . By Lemma 6.1 if we complete the weak reduction by cleaning and follow with weak reduction along (D', E) , then we obtain the same thing as if we

pre-weak reduce $\{T(G), t(G)\}$ along (D', E) and then clean. Now consider the effect of pre-weak reducing along (D, E) , followed by immediately pre-weak reducing along (D', E) . Pre-weak reducing along (D, E) yields

$$\begin{aligned} T(G) &= (\text{Thick}(H) - \{H_+\}) \cup \{H_+/D, H_+/E\}, \text{ and} \\ t(G) &= \text{Thin}(H) \cup \{H_+/DE\}. \end{aligned}$$

We now follow this with pre-weak reduction along (D', E) . Notice that both D' and E are disks with their boundaries on H_+/D . Hence, pre-weak reduction yields:

$$\begin{aligned} T(F) &= (\text{Thick}(H) - \{H_+\}) \cup \{H_+/DD', H_+/DE, H_+/E\}, \text{ and} \\ t(F) &= \text{Thin}(H) \cup \{H_+/DE, H_+/DD'E\}. \end{aligned}$$

By Lemma 5.5 we may start cleaning by removing repetitions between $T(F)$ and $t(F)$. This yields

$$\begin{aligned} T(F) &= (\text{Thick}(H) - \{H_+\}) \cup \{H_+/DD', H_+/E\}, \text{ and} \\ t(F) &= \text{Thin}(H) \cup \{H_+/DD'E\}. \end{aligned}$$

But notice that this result is symmetric in D and D' , so we would obtain the same result if we pre-weak reduce along (D', E) first and then follow with pre-weak reduction along (D, E) .

If the pair (D', E) is not a weak reducing pair for G , then by Lemma 5.8 weak reduction of $\{\text{Thick}(G), \text{Thin}(G)\}$ along the pair (D', E) produces sets isotopic to $\{\text{Thick}(G), \text{Thin}(G)\}$. In this case F is isotopic to G' and hence G is obtained from G' by the weak reduction (D, E) . By symmetry, it may also be possible that G' is obtained from G by weak reduction along (D', E) . Finally, it may be that (D', E) is not a weak reducing pair for G and (D, E) is not a weak reducing pair for G' . It would thus follow from Lemma 5.8 that G is isotopic to G' . \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 . Note that this is a generalization of a standard procedure of the same name, first defined by Schultens [16]. 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 transversally 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 $\mathcal{V}(H_+)$ and $\mathcal{W}(H_+)$, where $\mathcal{V}(H_+)$ is below H_+ and $\mathcal{W}(H_+)$ is above. 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{V} and \mathcal{W} . Define $\partial_-M(H_+)$ to be $\partial_-\mathcal{V}(H_+)$ and $\partial_+M(H_+)$ to be $\partial_-\mathcal{W}(H_+)$. That is, $\partial_-M(H_+)$ and $\partial_+M(H_+)$ are the boundary components of $M(H_+)$ that are below and above H_+ , respectively.

We now inductively build Σ . The intersection of Σ with some $M(H_+)$ is depicted in Figure 12. First, we define a sequence of manifolds $\{M_i\}$ where

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

The submanifold M_0 is defined to be the disjoint union of all manifolds of the form $M(H_+)$, such that $\partial_-M(H_+) = \emptyset$. 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 manifolds $M(H_+)$ such that $\partial_-M(H_+) \subset \partial M_{i-1}$. 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 $\mathcal{V} \subset M_0$ 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 $\mathcal{W} \subset M_0$. 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 $\mathcal{V} \subset M'_i$ choose a set of arcs $\Gamma \subset \mathcal{V}$ such that $\partial\Gamma \subset \Sigma_{i-1} \cap \partial M_{i-1}$ and such that the complement of Γ in \mathcal{V} 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 .

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 We must show that the complement of Σ in M is a union of handlebodies. For each subset X of M , we denote $X - \text{Nbhd}(\Sigma)$ as X_Σ .

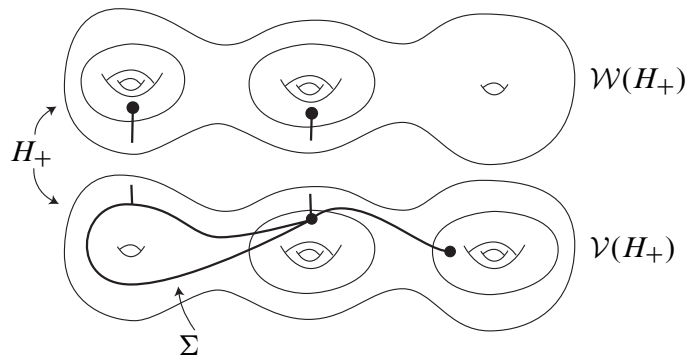


Figure 12: The intersection of Σ with $\mathcal{V}(H_+)$ and $\mathcal{W}(H_+)$.

Suppose that $N = M(H_+)$, where $H_+ \in \text{Thick}(H)$ and $\partial_+ N = \emptyset$. Then $\mathcal{W}(H_+)$ is a handlebody and is disjoint from Σ . Also, $\mathcal{V}(H_+)_{\Sigma}$ is homeomorphic to a product. Thus N_{Σ} is homeomorphic to $\mathcal{W}(H_+)$.

Now suppose that we have accumulated a submanifold $N \subset M$ such that

- (1) N is a union of submanifolds of the form $M(H_+)$, for thick levels $H_+ \in \text{Thick}(H)$,
- (2) $\partial_+ N = \emptyset$ and
- (3) N_{Σ} is a disjoint union of handlebodies.

Suppose that H_+ is a thick level which is not contained in N , such that $\partial_+ M(H_+) \subset \partial_- N$. Let $N' = N \cup M(H_+)$. We must check that N'_{Σ} is a disjoint union of handlebodies. If $\Sigma \cap \mathcal{W}(H_+) = \emptyset$ then this is easy; in fact N' is the disjoint union of N and $M(H_+)$. If not, then $\mathcal{W}(H_+)$ is of the form $\{\text{punctured surface}\} \times I$, together with one-handles. So $N_{\Sigma} \cup \mathcal{W}(H_+)_{\Sigma}$ is a disjoint collection of handlebodies. Again, $\mathcal{V}(H_+)_{\Sigma}$ is homeomorphic to a product and is correctly glued to $N_{\Sigma} \cup \mathcal{W}(H_+)_{\Sigma}$. \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 was uniquely defined up to isotopy at each stage. Hence, the amalgamation of a GHS is well defined, up to isotopy.

Lemma 7.4 Suppose H is a GHS of M , G is obtained from H by weak reduction, and H is not a stabilization of G . Then each component of $\mathcal{A}(H)$ is either

- (1) isotopic to a component of $\mathcal{A}(G)$,
- (2) the connected sum of two components of $\mathcal{A}(G)$,
- (3) the connected sum of a component of $\mathcal{A}(G)$ and the genus one Heegaard splitting of $S^2 \times S^1$ or
- (4) a Heegaard splitting of $S^2 \times S^1$ and $\mathcal{A}(G)$ is obtained from $\mathcal{A}(H)$ by deleting this splitting.

Proof It suffices to assume M is connected. 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 (pre-weak reduction) we replace H_+ in $\text{Thick}(H)$ with the components of H_+/D and H_+/E . We also add the components of H_+/DE to $\text{Thin}(H)$.

Suppose first that H_+/DE contains a sphere S and that S is inessential. Let B be the ball bounded by S . Then B contains a component H'_+ of either H_+/D or H_+/E . During cleaning M will get cut along S and the ball B will get capped off to form S^3 . The surface H'_+ then becomes a Heegaard surface in S^3 and is subsequently deleted. It follows that H is a stabilization of G .

Now suppose H_+/DE contains a sphere S that is essential in M . During cleaning we cut along this sphere and cap off the resulting boundary components with 3-balls. Thus, if S is separating then $M = M_1 \# M_2$ and G is a GHS of $M_1 \cup M_2$. If S is non-separating then either $M = M' \# (S^2 \times S^1)$ or $M \cong S^2 \times S^1$. In the first case G is a GHS of M' and H is obtained from G by taking the connected sum with the genus one splitting of $S^2 \times S^1$. In the second case G is empty.

If H_+/DE does not contain a sphere component then the second stage of weak reduction (cleaning) involves only the removal of 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 $M(H_+)$.

The manifold $M(H_+)$ can be built as follows. Begin with $\mathcal{V}(H_+/E)$. Attach a 1-handle to the positive boundary whose co-core is E . This is now the manifold $\mathcal{V}(H_+)$. To continue, attach a 2-handle whose core is D and then attach the manifold $\mathcal{W}(H_+/D)$. The 2-handle and $\mathcal{W}(H_+/D)$ is precisely $\mathcal{W}(H_+)$. See Figure 13. Since D and E are disjoint we may build $M(H_+)$ in an alternate way: Begin with $\mathcal{V}(H_+/E)$, attach the 2-handle, then attach the 1-handle and finally attach $\mathcal{W}(H_+/D)$.

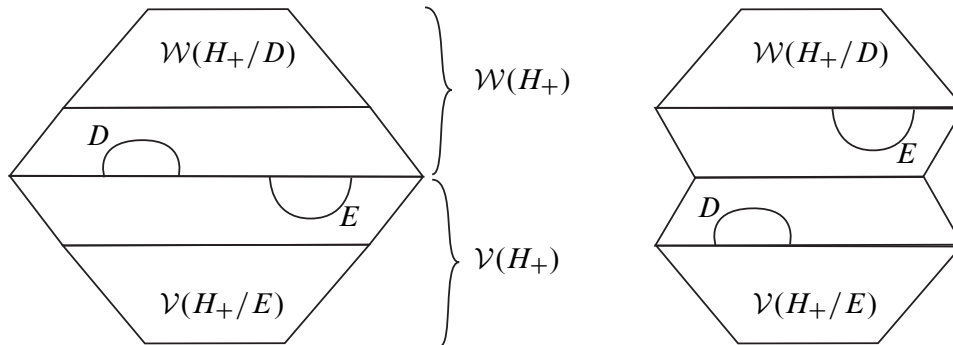


Figure 13: Because D and E are disjoint the manifold $M(H_+)$ can be built in two different ways.

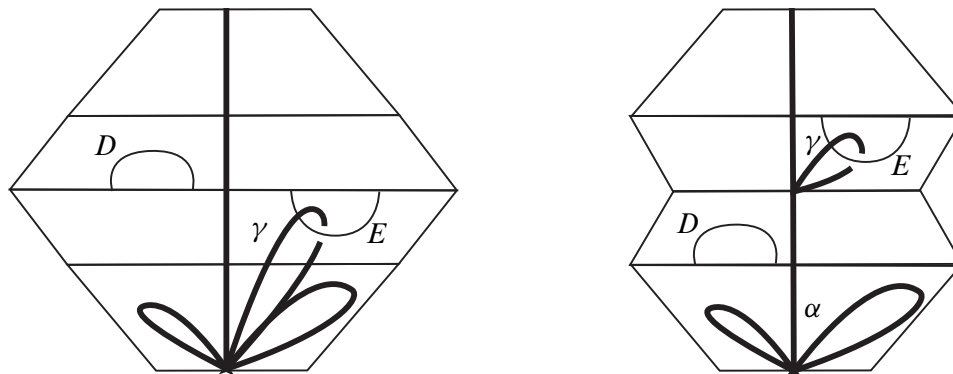


Figure 14: The graph Σ can be obtained from Σ' by contracting α .

Now note that we can build Σ by starting with a set of arcs in $\mathcal{V}(H_+/E)$ 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 $\mathcal{V}(H_+/E)$, attaching vertical arcs α , then attaching γ , and finishing with more vertical arcs. See Figure 14. By contracting the vertical arcs α we achieve the desired isotopy of $N(\Sigma')$. \square

Corollary 7.5 Suppose H is a GHS of an irreducible 3–manifold M , G is obtained from H by a weak reduction and H is not a stabilization of G . Then $\mathcal{A}(H)$ is isotopic to $\mathcal{A}(G)$.

Definition 7.6 The *genus* of a GHS H is the sum of the genera of the components of $\mathcal{A}(H)$.

Corollary 7.7 Suppose H is a GHS of M that is not a stabilization of another GHS. Let G be a GHS obtained from H by a weak reduction and suppose G is a GHS of M' . Then either $M = M'$ and $\text{genus}(H) = \text{genus}(G)$, or $M = M' \# (S^2 \times S^1)$ and $\text{genus}(H) = \text{genus}(G) + 1$.

Proof By Lemma 7.4 the only way the genus can drop by more than one is if $\mathcal{A}(H)$ has a component that is a genus g Heegaard splitting of $S^2 \times S^1$ and $g \geq 2$. But any such splitting is stabilized. \square

The previous corollary motivates the following definition.

Definition 7.8 Suppose H is a GHS of a 3-manifold M . The *irreducible genus* of H , denoted $\text{genus}^-(H)$, is equal to $\text{genus}(H)$ minus the number of $S^2 \times S^1$ summands in the prime decomposition of M .

With this definition in hand we can restate the previous corollary in a much simpler way.

Corollary 7.9 Suppose H is a GHS of M that is not a stabilization of another GHS. Let G be a GHS obtained from H by a weak reduction. Then $\text{genus}^-(H) = \text{genus}^-(G)$.

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 j th GHS of the SOG \mathbf{H} .

Definition 8.2 If \mathbf{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 \mathbf{H} .

It follows from Lemma 5.14 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 weak 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 weak reducing pair for H ,
- (4) for $1 \leq j \leq n$, D_{j-1} is disjoint from, or equal to, 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 weak 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 15(a) we have depicted a path in the curve complex from ∂V_0 to ∂W_1 . In Figure 15(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 Masur and Minsky [10]. 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 weak reducing pairs (D, E) and (D, E') for H then the distance between (D, E) and (D, E') is finite.*

Proof We 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 weak reducing pair and $E_i \cap E_{i+1} = \emptyset$. The sequence

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

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

The sequence $\{E_i\}$ is constructed inductively. The construction is essentially the same as that given in the proof of Claim 3.10.

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

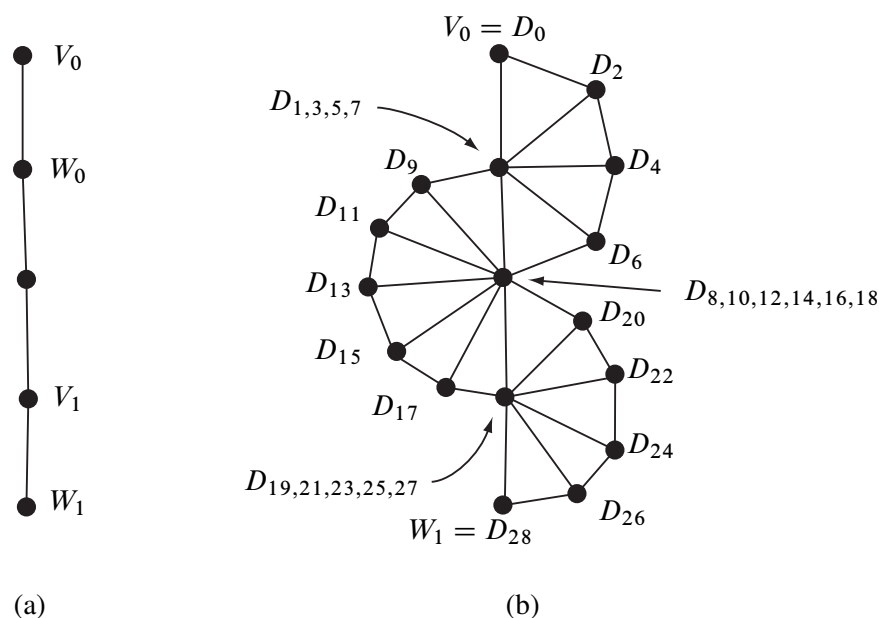


Figure 15: (a) The shortest path through alternately compressible loops from ∂V_0 to ∂W_1 , in the curve complex of H . (b) The distance from (V_0, W_0) to (V_1, W_1) is 27.

- (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 weak 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 a Heegaard surface in a 3-manifold. If there are weak 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 is not in C_1 then there is an E which forms a weak reducing pair with V_1 such that the distance between (V_0, W_0) and (V_1, E) is finite. But then, by concatenating sequences, 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 weak 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 again from Lemma 8.4 that the distance between (V_0, W_0) and (D, E) is finite and hence, $E \in C_0$. \square

We now define a complexity on maximal GHSs of SOGs. The definition is illustrated in Figure 16.

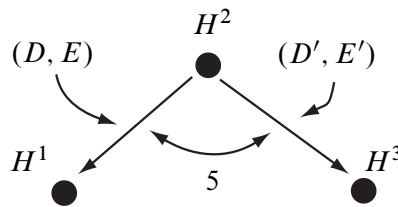


Figure 16: If the distance between (D, E) and (D', E') is 5 then the angle at H^2 is 5.

Definition 8.6 Suppose H^k is a maximal GHS of a SOG \mathbf{H} . Suppose further that H^{k-1} is obtained from H^k by the reduction given by the weak reducing pair (D, E) for the surface $H_+^k \in \text{Thick}(H^k)$ and that H^{k+1} is obtained from H^k by the reduction given by the weak reducing pair (D', E') for the surface $H_*^k \in \text{Thick}(H^k)$. If $H_+^k = H_*^k$ then define the *angle* $\angle(H^k)$ to be the distance between (D, E) and (D', E') . (In particular, if $(D, E) = (D', E')$ then the angle is 0.) If $H_+^k \neq H_*^k$ then we define $\angle(H^k)$ 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{H} be a SOG. Then the complexity of \mathbf{H} is given by

$$\left\{ \left(H^k, \angle(H^k) \right) \mid H^k \text{ is a maximal GHS} \right\}.$$

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. For comparison purposes we assume $\infty > n$ for all $n \in \mathbb{N}$.

Example 8.7 Consider the SOG pictured in Figure 17. If $H > K$ then the complexity of this SOG would be

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

If $K > H$ then the complexity would be

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

Finally, if H and K are not comparable (ie $c(H) = c(K)$) then the complexity would be

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

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 H^k is maximal in \mathbf{H} . We assume H^k is a GHS of a manifold with no S^3 components. There is then some thick level H^k_+ such that H^{k-1} is obtained from H^k by a reduction given by the weak reducing pair (D, E) for the surface H^k_+ . There is also a thick level H^k_* such that H^{k+1} is obtained from H^k by a reduction given by the weak reducing pair (D', E') for the surface H^k_* .

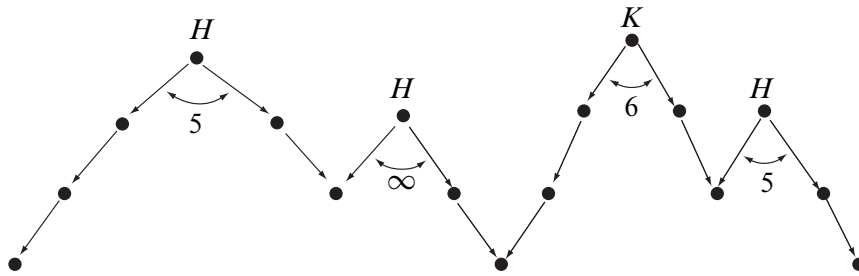
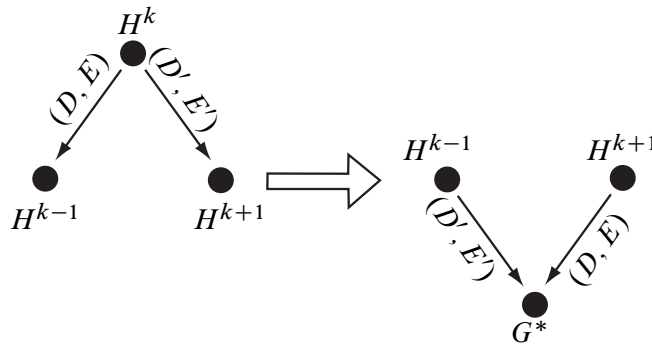


Figure 17: A SOG.

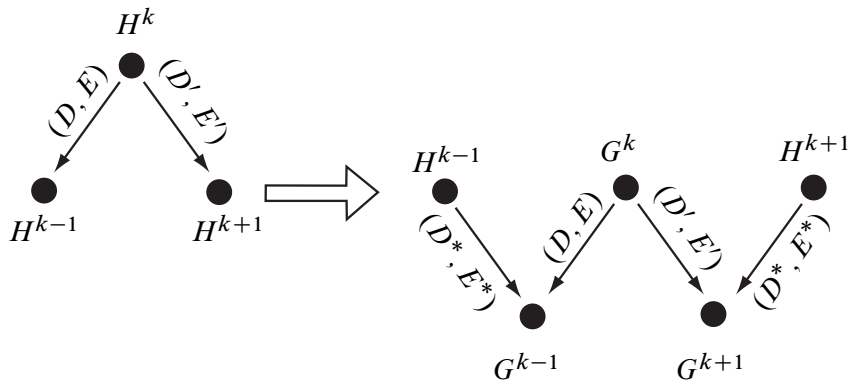
8.1 Reductions of type I

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

- (1) If $H_+^k \neq H_*^k$ then we may replace H^k with G^* in \mathbf{H} , where G^* is the GHS obtained from H^{k-1} by the weak reduction (D', E') . Since G^* can also be obtained from H^{k+1} by the weak reduction (D, E) our substitution has defined a smaller SOG \mathbf{H}' .



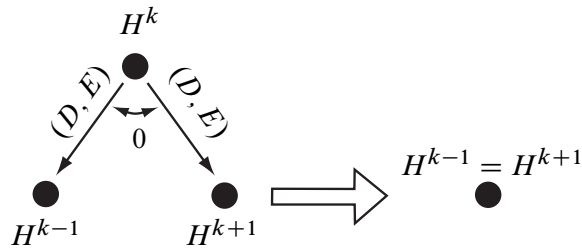
- (2) Next, assume $H_+^k = H_*^k$, but there is a thick level $H_0^k \neq H_+^k$ in $\text{Thick}(H^k)$ which is not strongly irreducible. Let (D^*, E^*) be a weak reducing pair for H_0^k . Let G^{k-1} , G^k and G^{k+1} denote the GHSs obtained from H^{k-1} , H^k and H^{k+1} by the weak reduction corresponding to (D^*, E^*) . Now replace H^k in \mathbf{H} with the subsequence $\{G^{k-1}, G^k, G^{k+1}\}$ to define a new, smaller SOG \mathbf{H}' .



8.2 Reductions of type II

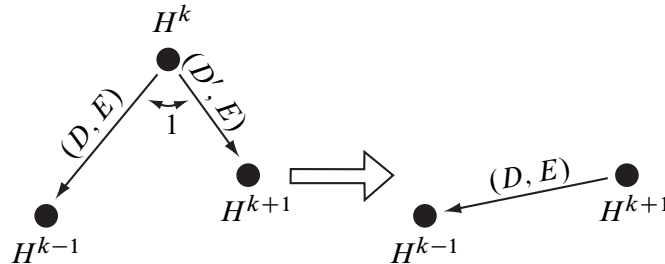
In all Type II reductions $H_+^k = H_*^k$ and we focus on $\angle(H^k)$.

- (1) $\angle(H^k) = 0$. Then (D, E) is the same as (D', E') , so removal of the subsequence $\{H^k, H^{k+1}\}$ from \mathbf{H} defines a new, smaller SOG \mathbf{H}' .

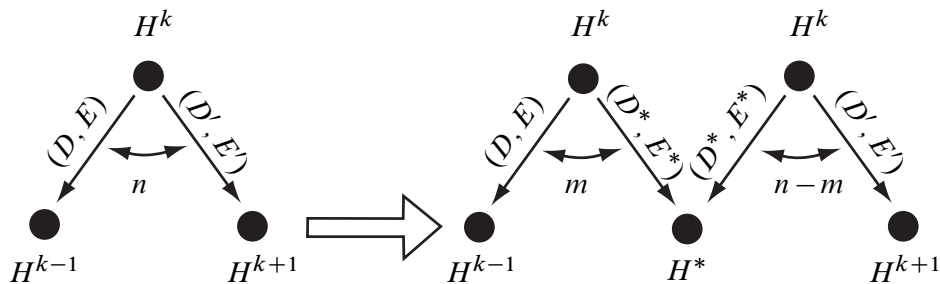


- (2) $\angle(H^k) = 1$. In this case either $D = D'$ and $E \cap E' = \emptyset$ or $E = E'$ and $D \cap D' = \emptyset$. Assume the latter. Then we may apply Lemma 6.2 to conclude one of the following holds.
 - (a) There is a GHS G^* that can be obtained from both H^{k-1} and H^{k+1} by weak reductions. Replacing H^k with G^* in \mathbf{H} thus produces a smaller SOG. The accompanying figure is the same as that for a Type I(1) reduction, with $E = E'$.

- (b) One of H^{k-1} or H^{k+1} can be obtained from the other by weak reduction. We may thus remove H^k from \mathbf{H} to obtain a smaller SOG.



- (c) H^{k-1} is isotopic to H^{k+1} . Hence we may remove the subsequence $\{H^k, H^{k+1}\}$ from \mathbf{H} to obtain a smaller SOG. The accompanying figure is the same as that for a Type II(1) reduction, with $E = E'$.
- (3) $\angle(H^k) = n$ 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 weak reducing pair (D_m, D_{m+1}) cannot be equal to either (D, E) or (D', E') . Let H^* denote the GHS obtained from H^k by the weak reduction corresponding to $(D_m, D_{m+1}) = (D^*, E^*)$. Now, let \mathbf{H}' denote the SOG obtained from \mathbf{H} by inserting the subsequence $\{H^*, H^k\}$ just after H^k . Note that the maximal GHS H^k appears one more time in \mathbf{H}' than in \mathbf{H} . However, the angle at the old occurrence of H^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 ∞ . \square

Definition 8.10 Let \mathbf{H} be a SOG. If $\text{genus}^-(H^i)$ is non-constant and the first change in irreducible genus is an increase, then we say \mathbf{H} is *initially increasing*. Similarly, if $\text{genus}^-(H^i)$ is non-constant and the first change in irreducible genus is a decrease, then we say \mathbf{H} is *initially decreasing*.

Hence, if \mathbf{H} is neither initially increasing nor initially decreasing, then $\text{genus}^-(H^i)$ is the same for all i .

Definition 8.11 Suppose a GHS G is obtained from a GHS H by weak reduction. Then we say H was obtained from G by *partial amalgamation*.

So, if H is a GHS of a connected 3-manifold M , then the GHS whose only thick level is $\mathcal{A}(H)$ is obtained from H by a sequence of partial amalgamations.

The advantage of this terminology is that it allows us to always talk about an element of a SOG in terms of how it was obtained from its predecessor. So, if \mathbf{H} is a SOG, then \mathbf{H} is defined by an initial GHS, together with a sequence of weak reductions and partial amalgamations. This will be a useful point of view for the proof of the next lemma.

Lemma 8.12 *Suppose \mathbf{G} is a SOG that is obtained from a SOG \mathbf{H} by a reduction. If \mathbf{G} is initially increasing then \mathbf{H} was initially increasing.*

Proof Assume \mathbf{H} is not initially increasing. We show \mathbf{G} is also not initially increasing. To prove the lemma it suffices to establish the following facts.

- (1) If a reduction of \mathbf{H} inserts a partial amalgamation that makes the irreducible genus go up, then it also inserts a weak reduction before it that makes the irreducible genus go down.
- (2) No reduction of \mathbf{H} will ever remove the first weak reduction that lowers the irreducible genus, if it comes before every partial amalgamation that raises irreducible genus.

The first assertion above is almost immediate: inspection of the different types of reductions shows that the only time a partial amalgamation (D^*, E^*) is introduced

into a SOG is when the weak reduction (D^*, E^*) is also introduced sometime before it. If the partial amalgamation (D^*, E^*) makes the irreducible genus go up, then the weak reduction defined by (D^*, E^*) makes the irreducible genus go down. This can happen only in Type I(2) and Type II(3) reductions.

The second assertion is not quite as immediate. In Type II(1) and Type II(2c) reductions a weak reduction that makes the irreducible genus go down may be removed, but only if it had been preceded by its inverse. Hence, these cases can not effect the first time a weak reduction makes the irreducible genus go down. In Type I(1) and Type II(2a) reductions the order of the partial amalgamation (D, E) and the weak reduction (D', E') may be reversed. Hence, if (D', E') is the first weak reduction in \mathbf{H} that reduces irreducible genus, then it will also correspond to the first weak reduction in \mathbf{G} that reduces irreducible genus.

There is one final case that needs to be addressed to establish the second assertion above. In a Type II(2b) reduction the decrease (D', E) is removed from \mathbf{H} . We claim, however, that if it reduced irreducible genus then the partial amalgamation (D, E) made the irreducible genus increase. Hence, (D', E) was not the *first* weak reduction that reduced irreducible genus. (Since we are assuming \mathbf{H} is not initially increasing, the first weak reduction that reduces irreducible genus has no partial amalgamations before it that increase irreducible genus.)

As the weak reduction (D', E) is assumed to reduce irreducible genus, we have

$$\text{genus}^-(H^{k+1}) < \text{genus}^-(H^k).$$

By hypothesis, H^{k-1} is obtained from H^{k+1} by the weak reduction (D, E) ; it follows that

$$\text{genus}^-(H^{k-1}) \leq \text{genus}^-(H^{k+1}).$$

Putting these inequalities together, we find

$$\text{genus}^-(H^{k-1}) < \text{genus}^-(H^k).$$

Hence, the partial amalgamation (D, E) did make the irreducible genus go up. \square

9 Examples

We now present examples suggested at a conference at Oberwolfach in 2005, and related to the author by the referee. These examples illustrate some of the key concepts of this paper.

9.1 SOGs of connected sums of lens spaces

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 $\mathbf{Y} = \{(Y^i, M^i)\}$ as follows:

- $M^1 = L_1 \cup L_2$, $\text{Thick}(Y^1) = \{T_1, T_2\}$, $\text{Thin}(Y^1) = \emptyset$,
- $M^2 = L_1\#L_2$, $\text{Thick}(Y^2) = \{T_1\#T_2\}$, $\text{Thin}(Y^2) = \emptyset$,
- $M^3 = L_1\#L_2\#S^3$, $\text{Thick}(Y^3) = \{T_1\#T_2\#T\}$, $\text{Thin}(Y^3) = \emptyset$,
- $M^4 = L_1\#L_2$, $\text{Thick}(Y^4) = \{\overline{T}_1\#T_2\}$, $\text{Thin}(Y^4) = \emptyset$,
- $M^5 = L_1 \cup L_2$, $\text{Thick}(Y^5) = \{\overline{T}_1, T_2\}$, $\text{Thin}(Y^5) = \emptyset$.

The GHS Y^3 is maximal in \mathbf{Y} . 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{Y} . 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 Y^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 weak reducing pair which we use to go from Y^3 to Y^2 and let (A', B') denote the weak reducing pair which we use to go from Y^3 to Y^4 . It follows that D is disjoint from both B and B' . The sequence $\{A, B, D, B', A'\}$ thus satisfies the conditions of Definition 8.3. We conclude $\angle(Y^3) = 3$. (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(3) to \mathbf{Y} , followed by several Type II(2) reductions. After such a sequence of reductions we are left with the following SOG $\mathbf{X} = \{(X^i, N^i)\}$:

- $N^1 = L_1 \cup L_2$, $\text{Thick}(X^1) = \{T_1, T_2\}$, $\text{Thin}(X^1) = \emptyset$,
- $N^2 = L_1\#S^3 \cup L_2$, $\text{Thick}(X^2) = \{T_1\#T, T_2\}$, $\text{Thin}(X^2) = \emptyset$,
- $N^3 = L_1 \cup L_2$, $\text{Thick}(X^3) = \{\overline{T}_1, T_2\}$, $\text{Thin}(X^3) = \emptyset$.

9.2 SOGs of SFSs

Let M be a Seifert Fibered Space with four exceptional fibers and base a two-sphere. Let S be the base orbifold of M and $p: M \rightarrow S$ the projection map. Let e_i denote the exceptional fibers, for $i = 1, 2, 3, 4$. Thus, for each i , $p(e_i)$ is a point of S .

Let S^* denote S with a neighborhood of the points $p(e_i)$ removed, for $i = 1, 2, 3, 4$. Then there is a homeomorphism $h: S^* \times S^1 \rightarrow p^{-1}(S^*)$. The submanifold $p^{-1}(S^*)$ of M has four torus boundary components, T_i , where T_i bounds a solid torus in M that contains e_i . Choose some $x \in S^1$ and $y \in S^*$. Let α_i be an arc in $h(S^*, x)$ connecting T_i to the point $h(y, x)$.

For each i , we may now construct a genus three Heegaard splitting of M as follows. Remove a neighborhood of α_j from $p^{-1}(S^*)$, for each $j \neq i$. The resulting submanifold of M has two boundary components. One of these will be the torus T_i . The other will be a genus three “vertical” Heegaard surface H_i (see Moriah and Schultens [11]).

All four Heegaard splittings H_i have a common genus four stabilization (see Schultens [17]). We call this splitting simply H . The surface H is the boundary of the manifold obtained from $p^{-1}(S^*)$ by removing a neighborhood of α_i , for all i .

The Heegaard surface H separates M into \mathcal{V} and \mathcal{W} . Assume \mathcal{V} contains the exceptional fibers. A neighborhood of the arc α_i is a 1–handle in \mathcal{V} , for each i . Let V_i be a cocore of this 1–handle. To obtain H_i from H we destabilize along the pair (V_i, W_i) , where W_i is some compressing disk for H in \mathcal{W} that meets V_i in a point. Let W_{jk} be the vertical compressing disk for H in \mathcal{W} given by $p^{-1}(\alpha_j \cup \alpha_k) \cap \mathcal{W}$. Then for each $i \neq j \neq k$ the pair (V_i, W_{jk}) is a weak reducing pair for H .

We now define a SOG $\mathbf{Y} = \{Y^i\}_{i=1}^3$ of M , where $\text{Thin}(Y^i) = \emptyset$ for all i and

- $\text{Thick}(Y^1) = \{H_1\}$,
- $\text{Thick}(Y^2) = \{H\}$,
- $\text{Thick}(Y^3) = \{H_2\}$.

The GHS Y^2 is maximal in this SOG. The GHS Y^1 is obtained from Y^2 by the weak reduction (V_1, W_1) . The GHS Y^3 is obtained from Y^2 by (V_2, W_2) . Note also that the disks V_1 and V_2 are disjoint from the disk W_{34} . Because the sequence $\{W_1, V_1, W_{34}, V_2, W_2\}$ satisfies the conditions of Definition 8.3, the angle at Y^2 is 3. We may thus apply a Type II(3) reduction to \mathbf{Y} .

The resulting SOG now has two maximal GHSs, each a copy of Y^2 . Between them there is a new GHS, which is either obtained from Y^2 by the weak reduction (V_1, W_{34}) ,

or by the weak reduction (V_2, W_{34}) . In either case, this new GHS has a thin level that is an essential, vertical torus that separates $e_1 \cup e_2$ from $e_3 \cup e_4$. This illustrates the main result of [1]: if the minimal genus common stabilization of a pair of Heegaard splittings is not critical, then the manifold contains an essential surface.

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, oriented 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, oriented 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 $\{M_i^j\}$ 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 [8] H_i is the connected sum of Heegaard splittings H_i^j of M_i^j and Heegaard splittings of copies of $S^2 \times S^1$. 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 the SOG \mathbf{Y} defined by Figure 18. The GHSs Y^{i_1} through Y^{i_5} have no thin levels. The thick levels of Y^{i_1} and Y^{i_5} are precisely the sets $\{H_i^j\}$ and $\{G_i^j\}$. Both of these are GHSs of the disjoint union of the manifolds $\{M_i^j\}$. The GHSs Y^1 and Y^n are obtained from Y^{i_1} and Y^{i_5} (respectively) by maximal sequences of weak reductions and are hence strongly irreducible. Finally, by construction observe that Y^{i_3} is the unique maximal GHS of the SOG \mathbf{Y} .

Since the thick level of Y^{i_3} is the connected sum of the thick levels of Y^{i_2} , it follows that

$$\text{genus}^-(Y^{i_2}) = \text{genus}^-(Y^{i_3}).$$

As Y^{i_2} is unstabilized, it follows from Corollary 7.9 that

$$\text{genus}^-(Y^1) = \text{genus}^-(Y^{i_2}).$$

Putting these together gives

$$(1) \quad \text{genus}^-(Y^1) = \text{genus}^-(Y^{i_3}).$$

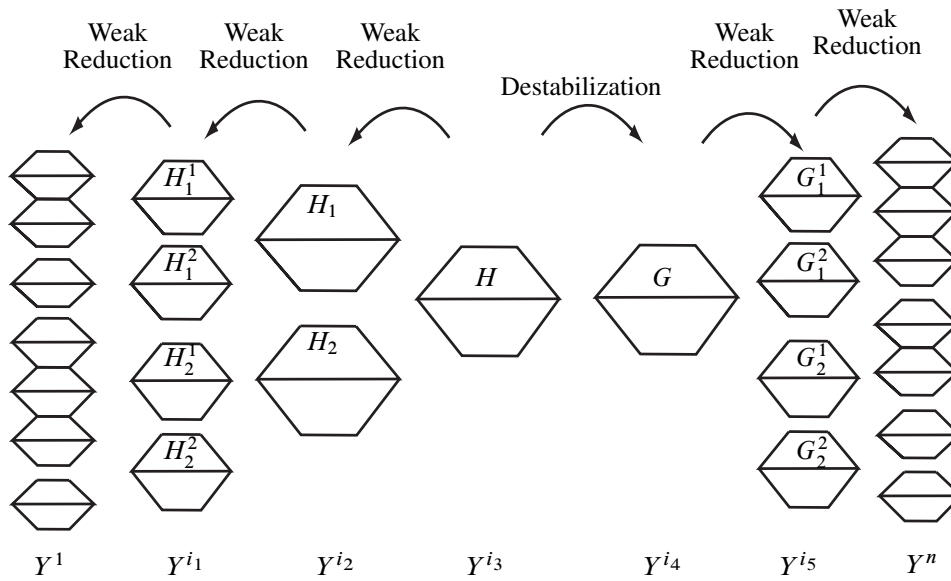


Figure 18

By hypothesis, the genus of the thick level of Y^{i_4} is less than the genus of the thick level of Y^{i_3} . Since the manifolds are the same, we conclude

$$\text{genus}^-(Y^{i_4}) < \text{genus}^-(Y^{i_3}).$$

It follows then, that the SOG \mathbf{Y} is initially decreasing.

As Y^n is obtained from Y^{i_4} by a sequence of weak reductions, we have

$$\text{genus}^-(Y^n) \leq \text{genus}^-(Y^{i_4}).$$

Hence,

$$(2) \quad \text{genus}^-(Y^n) < \text{genus}^-(Y^{i_3}).$$

Putting (1) and (2) together thus gives

$$\text{genus}^-(Y^n) < \text{genus}^-(Y^1).$$

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

Since \mathbf{X} is obtained from \mathbf{Y} by a sequence of reductions, it follows that $X^1 = Y^1$ and $X^m = Y^n$. Hence,

$$(3) \quad \text{genus}^-(X^m) < \text{genus}^-(X^1).$$

By construction, the SOG \mathbf{Y} is initially decreasing. It thus follows from Lemma 8.12 that \mathbf{X} is not initially increasing. Hence, if the irreducible genus changes in \mathbf{X} , it must go down first. By (3) the irreducible genus does change in \mathbf{X} , so there must be some number k such that

$$(4) \quad \text{genus}^-(X^{k+1}) < \text{genus}^-(X^k)$$

and for all $l < k$, $\text{genus}^-(X^l)$ is the same. As X^l is a GHS of an irreducible manifold for all l , it follows from Corollary 7.5 that $\mathcal{A}(X^l)$ is the same for all $l \leq k$. In particular, $\mathcal{A}(X^k) = \mathcal{A}(X^1)$.

As Y^{i_1} contains a single thick level in each component of the manifold it is a GHS of, it follows that $\text{Thick}(Y^{i_1}) = \mathcal{A}(Y^{i_1})$. As this manifold is irreducible, it follows from Corollary 7.5 that $\mathcal{A}(Y^1) = \mathcal{A}(Y^{i_1})$. Finally, as $X^1 = Y^1$, we have $\mathcal{A}(X^1) = \mathcal{A}(Y^1)$.

Putting all of this together, we have:

$$\text{Thick}(Y^{i_1}) = \mathcal{A}(Y^{i_1}) = \mathcal{A}(Y^1) = \mathcal{A}(X^1) = \mathcal{A}(X^k).$$

In particular, as Y^{i_1} is unstabilized, so is $\mathcal{A}(X^k)$. It follows that X^k is also unstabilized. But by Corollary 7.9, (4) implies that X^k was stabilized, so we have reached a contradiction. \square

As noted in the proof of the previous Theorem, it follows from [8] that every Heegaard splitting can be expressed as the connected sum of Heegaard splittings of prime 3-manifolds. The next theorem asserts that for unstabilized Heegaard splittings, this connected sum decomposition is unique (up to isotopy).

Theorem 10.2 *Let M be a closed, orientable 3-manifold. Let H be an unstabilized Heegaard splitting of M . Then H has a unique expression as the connected sum of Heegaard splittings of prime 3-manifolds.*

Proof Let $\{M_i\}$ denote the irreducible manifolds in a prime decomposition of M . By [8] H is the connected sum of Heegaard splittings H_i of M_i and Heegaard splittings of copies of $S^2 \times S^1$. Suppose H is also the connected sum of Heegaard splittings G_i of M_i and Heegaard splittings of copies of $S^2 \times S^1$. Our goal is to show that for each i , H_i is isotopic to G_i .

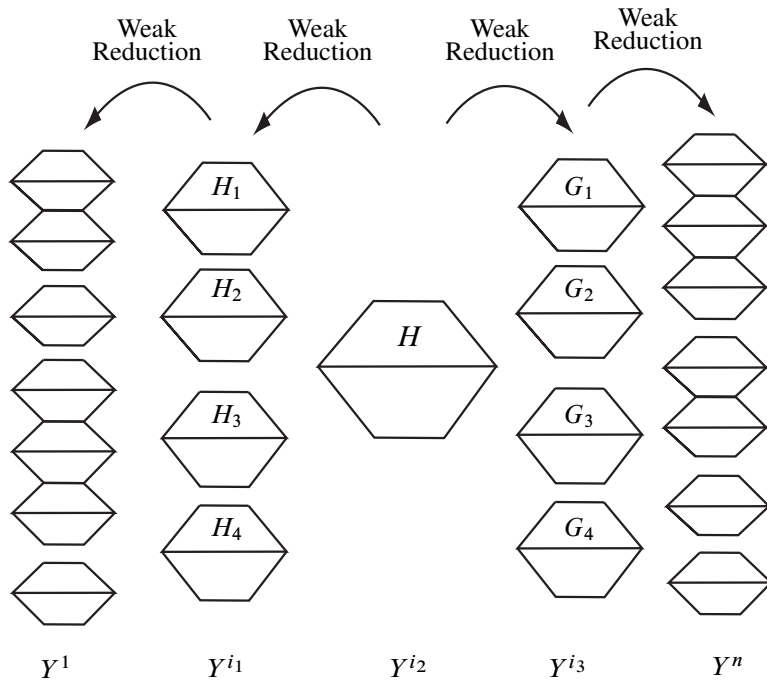


Figure 19

Let \mathbf{Y} be the SOG defined by Figure 19. In the figure, Y^1 and Y^n are strongly irreducible GHSs, obtained from Y^{i_1} and Y^{i_3} , by maximal sequence of weak reductions.

Since, by assumption, Y^{i_2} is unstabilized and every other GHS in \mathbf{Y} is obtained from Y^{i_2} by weak reductions, it follows from Corollary 7.9 that the irreducible genus never changes in \mathbf{Y} .

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

Since the irreducible genus never changes in \mathbf{Y} , it follows from Lemma 8.12 that \mathbf{X} is not initially increasing. Let k be the largest number such that for all $l \leq k$, $\text{genus}^-(X^l)$ is the same. As X^l is a GHS of an irreducible manifold for all l , it follows from Corollary 7.5 that $\mathcal{A}(X^l)$ is isotopic for all $l \leq k$. In particular, $\mathcal{A}(X^k) \sim \mathcal{A}(X^1)$.

As Y^{i_1} contains a single thick level in each component of the manifold it is a GHS of, it follows that $\text{Thick}(Y^{i_1}) = \mathcal{A}(Y^{i_1})$. As this manifold is irreducible, it follows from Corollary 7.5 that $\mathcal{A}(Y^1) \sim \mathcal{A}(Y^{i_1})$. Finally, as $X^1 = Y^1$, we have $\mathcal{A}(X^1) = \mathcal{A}(Y^1)$.

Putting all of this together, we have:

$$(5) \quad \text{Thick}(Y^{i_1}) = \mathcal{A}(Y^{i_1}) \sim \mathcal{A}(Y^1) = \mathcal{A}(X^1) \sim \mathcal{A}(X^k).$$

In particular, as Y^{i_1} is unstabilized, so is $\mathcal{A}(X^k)$. It follows that X^k is also unstabilized. If $k < m$ it now follows from Corollary 7.9 that

$$\text{genus}^- X^k \leq \text{genus}^- X^{k+1}.$$

But we know \mathbf{X} is not initially increasing, so we conclude

$$\text{genus}^- X^k = \text{genus}^- X^{k+1}.$$

This, however, violates the maximality of our choice of k . We conclude $k = m$ and thus (5) implies $\text{Thick}(Y^{i_1}) \sim \mathcal{A}(X^m)$.

As Y^{i_3} contains a single thick level in each component of the manifold it is a GHS of, it follows that $\text{Thick}(Y^{i_3}) = \mathcal{A}(Y^n)$. As this manifold is irreducible, it follows from Corollary 7.5 that $\mathcal{A}(Y^n) \sim \mathcal{A}(Y^{i_3})$. Finally, as $X^m = Y^n$, we have $\mathcal{A}(X^m) = \mathcal{A}(Y^n)$.

Putting everything together, we now have:

$$\{H_i\} = \text{Thick}(Y^{i_1}) \sim \mathcal{A}(X^m) = \mathcal{A}(Y^n) \sim \mathcal{A}(Y^{i_3}) = \text{Thick}(Y^{i_3}) = \{G_i\}.$$

Thus, $H_i \sim G_i$ for all i . □

References

- [1] **D Bachman**, *Critical Heegaard surfaces*, Trans. Amer. Math. Soc. 354 (2002) 4015–4042 MR1926863
- [2] **J S Birman**, *Braids, links, and mapping class groups*, Annals of Mathematics Studies 82, Princeton University Press (1974) MR0375281
- [3] **F Bonahon, J-P Otal**, *Scindements de Heegaard des espaces lenticulaires*, Ann. Sci. École Norm. Sup. (4) 16 (1983) 451–466 (1984) MR740078
- [4] **A J Casson, C M Gordon**, *Reducing Heegaard splittings*, Topology Appl. 27 (1987) 275–283 MR918537
- [5] **J Cerf**, *Sur les difféomorphismes de la sphère de dimension trois ($\Gamma_4 = 0$)*, Lecture Notes in Mathematics, No. 53, Springer, Berlin (1968) MR0229250
- [6] **R Engmann**, *Nicht-homöomorphe Heegaard-Zerlegungen vom Geschlecht 2 der zusammenhängenden Summe zweier Linsenräume*, Abh. Math. Sem. Univ. Hamburg 35 (1970) 33–38 MR0283803
- [7] **D Gabai**, *Foliations and the topology of 3-manifolds. III*, J. Differential Geom. 26 (1987) 479–536 MR910018

- [8] **W Haken**, *Some results on surfaces in 3-manifolds*, from: “Studies in Modern Topology”, Math. Assoc. Amer. (distributed by Prentice–Hall, Englewood Cliffs, N.J.) (1968) 39–98 MR0224071
- [9] **R Kirby, editor**, *Problems in low-dimensional topology*, from: “Geometric topology (Athens, GA, 1993)”, AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc. (1997) 35–473 MR1470751
- [10] **H A Masur, Y N Minsky**, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. 10 (2000) 902–974 MR1791145
- [11] **Y Moriah, J Schultens**, *Irreducible Heegaard splittings of Seifert fibered spaces are either vertical or horizontal*, Topology 37 (1998) 1089–1112 MR1650355
- [12] **R Qiu**, *Stabilizations of Reducible Heegaard Splittings* arXiv:0409497
- [13] **H Rubinstein, M Scharlemann**, *Comparing Heegaard splittings of non-Haken 3-manifolds*, Topology 35 (1996) 1005–1026 MR1404921
- [14] **M Scharlemann**, *Reconfiguring Qiu’s proof of the Gordon Conjecture* arXiv:0801.4581
- [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] **J Schultens**, *The classification of Heegaard splittings for (compact orientable surface) $\times S^1$* , Proc. London Math. Soc. (3) 67 (1993) 425–448 MR1226608
- [17] **J Schultens**, *The stabilization problem for Heegaard splittings of Seifert fibered spaces*, Topology Appl. 73 (1996) 133–139 MR1416756
- [18] **F Waldhausen**, *Heegaard–Zerlegungen der 3–Sphäre*, Topology 7 (1968) 195–203 MR0227992
- [19] **R Weidmann**, *Generating tuples of free products*, Bull. Lond. Math. Soc. 39 (2007) 393–403 MR2331566

Pitzer College, 1050 N Mills Ave
Claremont, CA 91711, USA

bachman@pitzer.edu

Proposed: David Gabai
Seconded: Colin Rourke, Joan Birman

Received: 23 November 2004
Revised: 3 May 2005