Stabilization of Heegaard splittings

JOEL HASS
ABIGAIL THOMPSON
WILLIAM THURSTON

For each \( g \geq 2 \) there is a 3–manifold with two genus–\( g \) Heegaard splittings that require \( g \) stabilizations to become equivalent. Previously known examples required at most one stabilization before becoming equivalent. Control of families of Heegaard surfaces is obtained through a deformation to harmonic maps.

57M25; 53C43

1 Introduction

While minimal surfaces have proved to be a powerful tool in the study of the topology of 3–dimensional manifolds, harmonic maps of surfaces to 3–manifolds have not been as widely applied in this setting, due to several limitations. A single homotopy class of surfaces can give rise to a large number of harmonic maps. A harmonic map of a surface need not minimize self-intersections, so that an embedding of a surface may give rise to a homotopic harmonic map that is not embedded and may even fail to be immersed. In negatively curved 3–manifolds, however, there is a unique harmonic map in a homotopy class for each conformal class of metrics on the domain; see Hartman [7]. Moreover continuous families of maps and metrics give rise to continuous families of harmonic maps; see Eells and Lemaire [4] and Sampson [17]. In addition, the area of a harmonic map is bounded by a constant depending only on the genus and the sectional curvatures of the 3–manifold. In this paper we use families of harmonic surfaces to study Heegaard splittings of 3–manifolds.

A genus–\( g \) Heegaard splitting of a 3–manifold \( M \) is a decomposition of \( M \) into two genus–\( g \) handlebodies with a common boundary. It is described by an ordered triple \((H_1, H_2, S)\) where each of \( H_1, H_2 \) is a handlebody and the two handlebodies intersect along their common boundary \( S \), called a Heegaard surface. An orientation on \( S \) is determined by \( \partial H_1 \), and an equivalent definition of a Heegaard splitting is given by an oriented surface \( S \) in \( M \) whose complement consists of two handlebodies.

Two Heegaard splittings \((H_1, H_2, S)\) and \((H_1', H_2', S')\) of \( M \) are said to be equivalent if an ambient isotopy of \( M \) carries \((H_1, H_2, S)\) to \((H_1', H_2', S')\). Every 3–manifold
has a Heegaard splitting (see Moise [11]), and Heegaard splittings are one of the basic structures used to analyze and understand 3–manifolds.

Corresponding to a Heegaard splitting is a Heegaard sweepout, a family of surfaces that sweep out the manifold, starting with a spine of one handlebody and ending at a spine of the second. We will see that this family can be geometrically controlled by deforming it to a family of harmonic maps. When the manifold is negatively curved, harmonic maps of genus–$g$ surfaces have uniformly bounded area. In the manifolds we consider, the geometry forces small area surfaces to line up with small area cross sections of the manifold. As a result we obtain obstructions to the equivalence of distinct Heegaard splittings.

A stabilization of a genus–$g$ Heegaard surface is a surface of genus $g + 1$ obtained by adding a trivial 1–handle, a handle whose core is parallel to the surface. Such a surface splits the manifold into two genus $g + 1$ handlebodies, and thus gives a new Heegaard splitting. Any two Heegaard splittings become equivalent after a finite sequence of stabilizations; see Singer [21]. An upper bound on the number of stabilizations needed to make two splittings equivalent is known in some cases. If $G_p$ and $G_q$ are splittings of genus $p$ and $q$ with $p \leq q$, and $M$ is non-Haken, then Rubinstein and Scharlemann obtained an upper bound of $5p + 8q - 9$ for the genus of a common stabilization [15]. Less is known about lower bounds. In all previously known examples of manifolds with distinct Heegaard splittings, the splittings become equivalent after a single stabilization of the larger genus Heegaard surface. The question of whether a single stabilization always suffices is sometimes called the stabilization conjecture; see Kirby [9, Problem 3.89], Scharlemann [18], Schultens [19] and Sedgwick [20]. In Section 7 we show that this conjecture does not hold. There are pairs of genus–$g$ splittings of a 3–manifold that become equivalent after no fewer than $g$ stabilizations.

Theorem 1.1 For each $g > 1$ there is a 3–manifold $M_g$ with two genus–$g$ Heegaard splittings that require $g$ stabilizations to become equivalent.

We note that there is an alternate way of defining a Heegaard splitting in which the orientation of $S$, or the ordering of the handlebodies, is not considered. One can also consider equivalence under homeomorphism rather than under isotopy. Our results do not apply directly with these alternate definitions.

We outline the proof of Theorem 1.1 in Section 2. In Section 3 we describe the construction of the negatively curved, but not hyperbolic 3–manifolds $M_g$. We derive isoperimetric inequalities used in the proof in Section 4 and discuss deformations of surfaces to harmonic maps in Section 5. In Section 6 we show how a Heegaard splitting gives rise to a family of Heegaard surfaces that sweep out the 3–manifold. The proof
of Theorem 1.1 is given in Section 7. Finally in Section 8 we show how a somewhat weaker result can be obtained for an easily constructed class of hyperbolic manifolds. In these manifolds \( g - 4 \) stabilizations are required.

This result was presented at the American Institute of Mathematics Conference on Triangulations, Heegaard Splittings and Hyperbolic Geometry held in December 2007. At this conference D Bachman announced, using different methods, examples giving a lower bound of \( g - 3 \) for the number of required stabilizations.

The research of all three authors was partly supported by the National Science Foundation.

2 Outline of the argument

Let \( M_\phi \) be a hyperbolic 3–manifold that fibers over \( S^1 \) with monodromy \( \phi \) and let \( \tilde{M}_\phi \) denote its infinite cyclic cover. A pictorial representation of \( M_\phi \) is given in Figure 1. Cutting open \( M_\phi \) along a fiber gives a fundamental domain \( B \) of the \( \mathbb{Z} \)–action on the infinite cyclic cover, which we call a block. Blocks are homeomorphic, but not isometric, to the product of a surface and an interval. They are foliated by fibers of \( M_\phi \), which in \( B \) we call slices. By cutting open a cyclic cover of \( M_\phi \) we obtain a hyperbolic 3–manifold with as many adjacent blocks as we wish.

![Figure 1: \( M_\phi \) and one block \( B \), a fundamental domain of the \( \mathbb{Z} \)–action on its infinite cyclic cover. \( B \) is homeomorphic to \( F_g \times I \).](image)

The manifolds \( M_g \) used in our main result have sectional curvatures between \(-3/2\) and \(-1/2\). They contain two handlebodies \( H_L \) and \( H_R \) with fixed Riemannian metrics, separated by a region homeomorphic to the product of a surface of genus \( g \) with an interval. This intermediate piece is hyperbolic and isometric to \( 2n \) adjacent blocks. The first \( n \) blocks form a submanifold called \( L \) and the next \( n \) form a submanifold called \( R \), as in Figure 2. The value of \( n \) can be chosen as large as desired without
changing the geometry of $H_L$ and $H_R$, or the areas of slices filling $L$ and $R$. A more
detailed construction of $M_g$ is given in Section 3.

$M_g$ has two obvious Heegaard splittings $E_0 = (H_L \cup L, H_R \cup R, S)$ and $E_1 =
(H_R \cup R, H_L \cup L, -S)$, where $S$ is a surface of genus $g$ separating $H_L \cup L$ and
$H_R \cup R$ and $-S$ indicates $S$ with reversed orientation. We will show that these
splittings are not equivalent with fewer than $g$ stabilizations.

Let $G_0$ be the Heegaard splitting obtained by stabilizing $E_0$ $(g - 1)$ times and $G_1$ be
the Heegaard splitting obtained by stabilizing $E_1$ $(g - 1)$ times. If $G_0$ and $G_1$ are
equivalent, then there is an isotopy $\{I_s: M_g \rightarrow M_g, 0 \leq s \leq 1\}$ with $I_0$ the identity
and $I_1$ a diffeomorphism of $M_g$ carrying $G_0$ to $G_1$. A family of Heegaard splittings
$\{G_s = I_s(G_0), 0 \leq s \leq 1\}$ interpolates between $G_0$ and $G_1$.

Associated to each Heegaard splitting $G_s$ is a family of surfaces $F_{s,t}$ sweeping out $M_g$
from one spine to the other. In Section 5 we show that such a family of surfaces can
be deformed to a family of harmonic, or energy minimizing, maps. These harmonic
surfaces have area uniformly bounded by a constant $A_0$ that is independent of $n$. This
area bound restricts the way that a surface divides the volume of $M_g$. We will show
that a surface of area less than $A_0$ and genus less than $2g$ cannot simultaneously split
in half the volumes of $L$ and $R$. In Section 7 we use this to show that surfaces in such
bounded area families cannot interpolate between $G_0$ and $G_1$. We conclude that $E_0$
and $E_1$ require at least $g$ stabilizations to become equivalent.

**Remark** The two Heegaard splittings $E_0$ and $E_1$ of $M_g$ do become equivalent
after $g$ stabilizations. To see this, consider the graph formed by connecting the two
spines with an arc that intersects once each surface in the family of Heegaard surface
between them. A regular neighborhood of this graph gives a handlebody of genus $2g$
with handlebody complement. Its boundary surface is obtained by adding $g$ trivial
1–handles to either of the two Heegaard surfaces $S$ and $-S$, and so gives a common
stabilization of each of them.

This genus $2g$ Heegaard surface can be isotoped to the surface formed by connecting
with a tube a slice that bisects the volume of $L$ and a slice that bisects the volume

Figure 2: $M_2$, constructed with 10 blocks, 5 in each of $L$ and $R$
of $R$. Such a surface arises in a family of bounded area surfaces interpolating between stabilizations of $E_0$ and $E_1$ having genus $2g$. As we will see, this property cannot be achieved in a family of smaller genus surfaces.

3 Construction of the manifolds

In this section we describe a construction that gives negatively curved 3–manifolds $M_g$ of Heegaard genus $g$, based on the work of Namazi and Souto [12]. Each of these manifolds has two genus–$g$ splittings that require $g$ stabilizations to become equivalent. In Section 8 we give a simpler construction of a hyperbolic manifold, containing two genus–$g$ splittings that require $g–4$ stabilizations to become equivalent.

We begin with a hyperbolic manifold $M_\phi$ that fibers over the circle, with fiber a genus–$g$ surface and pseudo-Anosov monodromy $\phi$. Fix a fibration of $M_\phi$ with fibers \{${S_t, 0 \leq t \leq 1}$\} satisfying $\phi(S_0) = S_1$. Define a block $B$ to be the manifold obtained by cutting open $M_\phi$ along $S_0$ and a block manifold $B_n$ to be a union of $n$ blocks, formed by placing $n$ copies of $B$ end to end. The manifold $B_n$ is topologically, though not geometrically, the product of a surface of genus $g$ and an interval. Its geometry can be obtained by cutting open the $n$–fold cover of $M_\phi$ along a lift of $S_0$. We call the fibers of $B_n$ slices and label them by $S_t, 0 \leq t \leq n$.

The manifold $M_g$ is a union of four pieces, as in Figure 2. Two pieces are genus–$g$ handlebodies, $H_L$ and $H_R$, and the other two $L$ and $R$ are each homeomorphic, though not isometric, to a product $F_g \times [-1, 1]$, where $F_g$ is a surface of genus $g$. $L$ and $R$ are each hyperbolic, and isometric to a block manifold $B_n$. The sectional curvatures of $M_g$ are pinched between $-1 – \epsilon$ and $-1 + \epsilon_0$, where $\epsilon_0 > 0$ can be chosen arbitrarily small. For our purpose we take $\epsilon_0 = 1/2$.

Namazi and Souto produced manifolds very similar to $M_g$ [12]. The manifolds they construct are suitable for the constructions we give, but the argument is somewhat simpler if we modify the metric slightly so that the middle part $L \cup R$ of $M_g$ is precisely, rather than approximately, hyperbolic. This can be arranged by beginning with the manifold $M$ in Theorem 5.1 of [12], which has sectional curvatures pinched between $-1 – \epsilon$ and $-1 + \epsilon$ where $\epsilon > 0$ can be chosen arbitrarily small. The manifold $M$ contains a region separating two handlebodies whose Riemannian metric is $\epsilon$ close (in the $C^2$–metric) to the Riemannian metric of a block. A small $C^2$–perturbation gives a new Riemannian metric on $M$ with sectional curvatures between $-3/2$ and $-1/2$ and isometric to a block on a region separating the two handlebodies. Now cut open the manifold along a slice in this hyperbolic block and insert a copy of $B_{2n}$ to obtain $M_g$, where $n$ can be chosen as large as desired. For any choice of $n$, the geometry of

Geometry & Topology, Volume 13 (2009)
the two complementary handlebodies of $B_{2n}$, which we call $H_L$ and $H_R$, remains unchanged. We split $B_{2n}$ along a slice into two halves $L$ and $R$, each consisting of $n$ blocks. We will call the resulting manifold $M_R$, omitting the index $n$.

## 4 Isoperimetric inequalities

In this section we develop some isoperimetric inequalities for piecewise-smooth curves in surfaces and for piecewise-smooth surfaces in 3–manifolds. The first isoperimetric inequality holds for a curve in hyperbolic space $H^2$ [2]. A proof can be obtained by using symmetry to show that a round circle bounds at least as much area as any other curve, using no more length. Explicit formulas for length and area in hyperbolic space then imply this Lemma. The curve does not need to be embedded or connected.

**Lemma 4.1** Let $c$ be a closed curve in $H^2$ with length $L$. Then $c$ is the boundary of a disk $f: D^2 \to H^2$ of area $A$, with $A < L$.

We now consider more general isoperimetric inequalities for curves in surfaces and surfaces in 3–manifolds. In these settings we consider the areas and volumes of 2–chains and 3–chains with boundary.

To simplify the next calculations, it is convenient to work with chains having $\mathbb{Z}_2$–coefficients. With this choice we do not need to pay attention to orientations, signs, or multiplicity other than even or odd. A 2–chain spanning a general position curve in a surface is obtained by 2–coloring the complement of the curve, dividing the surface into black and white regions separated by the curve. The area of a spanning 2–chain is the area of either the white or the black subsurface, and has a value between zero and the area of the entire surface. The two choices of spanning 2–chain have areas that sum to the area of the entire surface. A similar statement applies for 3–chains spanning a surface in a 3–manifold. While many of these results hold equally well for $\mathbb{Z}$–coefficients, we will assume $\mathbb{Z}_2$–coefficients below.

We first give an isoperimetric inequality for curves in a compact family of Riemannian metrics on a surface. We consider the area of a 2–chain bounding a smooth curve on a surface.

**Lemma 4.2** Let $(S, g_r)$ be a closed 2–dimensional surface with Riemannian metric $g_r$, $r \in U$, where $g_r$ is a family of smooth metrics continuously parameterized by a compact set $U \subset \mathbb{R}^n$. There is a constant $K$ such that for any surface $S$ in the family $(S, g_r)$:

*Geometry & Topology, Volume 13 (2009)*
(1) A null-homotopic curve \( c \) in \( S \) is the boundary of a disk \( f: D \to S \) with 
\[
\text{Area}(f(D)) \leq K \cdot \text{Length}(c).
\]

(2) A collection of curves \( c = \bigcup c_i \), with each \( c_i \) null-homotopic in \( S \), is the boundary of a collection of disks \( f_i: D_i \to S \) with 
\[
\sum \text{Area}(f_i(D_i)) \leq K \cdot \text{Length}(c).
\]

(3) A null-homologous curve \( c \) on \( S \) bounds a two-chain \( X_2 \) in \( S \) with 
\[
\text{Area}(X_2) \leq K \cdot \text{Length}(c).
\]

**Proof** First consider the case where \( c \) is a null-homotopic curve. If it lies in a hyperbolic surface \( S \), \( c \) lifts to the universal cover \( \mathbb{H}^2 \) where it bounds a mapped in disk of area less than \( \text{Length}(c) \) by Lemma 4.1. The surface \( (S, g_r) \) is conformally equivalent to a hyperbolic surface, so there is a hyperbolic metric \( (S, g) \) and a diffeomorphism from \( (S, g) \) to \( (S, g_r) \) that stretches or compresses any vector in the tangent space of \( (S, g) \) by a factor of at most \( \lambda \). Since all quantities change continuously with the metric, a single choice of \( \lambda \) can be made for all the surfaces in the compact family \( (S, g_r) \). Lengths of curves measured in \( (S, g) \) differ from lengths in \( (S, g_r) \) by a factor of at most \( \lambda \) and the area of a region in \( (S, g) \) differs from the area in \( (S, g_r) \) by a factor of at most \( \lambda^2 \). By comparing the area of a spanning disk and \( \text{Length}(c) \) in \( (S, g) \), we see that in \( (S, g_r) \)
\[
\text{Area}(f(D)) \leq \lambda^3 \cdot \text{Length}(c).
\]
The same argument applies for a collection of null-homotopic components.

There is a constant \( \varepsilon \) that gives a lower bound for the injectivity radius for the family of metrics \( g_r \). If the length of \( c \) is less than \( \varepsilon \) then each of the components of \( c \) is null-homotopic and the above bound holds for the area of a collection of spanning disks.

Now consider a general null-homologous curve \( c \) with length greater than \( \varepsilon \). A small perturbation and cut and paste transforms \( c \) into a union of disjoint embedded components that is homologous to \( c \). A homology between this embedded curve and \( c \) has arbitrarily small area if the perturbation is small, so we can assume that \( c \) is a collection of embedded disjoint curves. Each side of \( c \) gives a 2–chain with boundary \( c \), and so \( \text{Area}(S) \) gives an upper bound to the area of a 2–chain \( X_2 \) spanning \( c \). Therefore
\[
\text{Area}(X_2) \leq \text{Area}(S) < \frac{\text{Length}(c)}{\varepsilon} \cdot \text{Area}(S) = \frac{\text{Area}(S)}{\varepsilon} \cdot \text{Length}(c).
\]
The Lemma follows with \( K = \lambda^3 + \text{Area}(S)/\varepsilon \). \( \square \)

*Geometry & Topology, Volume 13 (2009)*
We now analyze the geometry of surfaces in block manifolds. The metrics on the slices $S_t$ of a block manifold, up to isometry, are parameterized by a closed interval, so Lemma 4.2 applies. Therefore there is a constant $K_1$ independent of $n$, such that a null-homologous curve $c$ on a slice $S_t$ bounds a 2–chain $X_2$ in that slice with

$$\text{Area}(X_2) \leq K_1 \cdot \text{Length}(c).$$

Each slice is an incompressible surface in $B_n$, so its induced metric has injectivity radius greater or equal to that of $M$. Set $\epsilon$ to be the constant

$$\epsilon = \min_{t \in [0,1]} \{\text{injectivity radius}(M), \text{Area}(S_t)/2K_1\}$$

Now consider a proper separating surface $(F, \partial F)$ in $(B_n, \partial B_n)$. Since $F$ is null-homologous $F \cap S_t$ separates $S_t$ into two regions. In slices $S_t$ where $\text{Length}(F \cap S_t) \leq \epsilon$, each curve of $F \cap S_t$ is null-homotopic. Lemma 4.2 implies that there is a collection of disks spanning the components of $F \cap S_t$ whose total area is less than $\text{Area}(S_t)/2$. Call the subsurface of $S_t$ missing these disks the large side of $F \cap S_t$. This subsurface contains the connected complementary component of $F \cap S_t$ in $S_t$ of largest area.

**Lemma 4.3** Suppose $(F, \partial F)$ is a proper separating surface in $(B_n, \partial B_n)$ and $F$ intersects slices $S_{t_1}, S_{t_2}$ transversely with $\text{Length}(F \cap S_{t_1}) \leq \epsilon$ and $\text{Length}(F \cap S_{t_2}) \leq \epsilon$. If the genus of $F$ is less than $g$ then an arc $\alpha$ from the large side of $F \cap S_{t_1}$ to the large side of $F \cap S_{t_2}$ intersects $F$ algebraically zero times.

**Proof** Cap off the curves in $F \cap S_{t_1}, F \cap S_{t_2}$, each of which is shorter than the injectivity radius, by adding disks missing the large side of each surface. These disks do not intersect the arc $\alpha$. The resulting closed surface in the region between $S_{t_1}$ and $S_{t_2}$ has genus less than $g$. Such a surface is null-homologous in the region between these two slices, which is homeomorphic to the product of a genus–$g$ surface and an interval, and therefore cannot algebraically separate them.

The next lemma gives a 3–dimensional isoperimetric inequality for surfaces in block manifolds.

**Lemma 4.4** Let $(F, \partial F) \subset (B_n, \partial B_n)$ be a null-homologous proper surface in a block manifold, having genus less than $g$ and area less than a constant $A_0$. Then there is a constant $C_0$, independent of $n$, such that $F$ bounds a 3–chain $X$ in $(B_n, \partial B_n)$ with $\text{Volume}(X) < C_0$. As $n$ grows,

$$\lim_{n \to \infty} \frac{\text{Volume}(X)}{\text{Volume}(B_n)} = 0.$$
Proof} Since $F$ is null-homologous, we can 2–color its complement. This produces two 3–chains with boundary $F$, and we take $X$ to be the one of smaller volume.

A formula from differential geometry called the \textit{coarea formula} states that the volume of a region in a three dimensional manifold can be estimated by an integral of the area of the intersection of the region with surfaces that sweep across it [6; 8]. The coarea formula implies that the volume of $X$ can be estimated by the integral of its cross-sectional areas in $S_t$. There is a constant $C_1$ such that

\begin{equation}
\frac{\text{Volume}(X)}{C_1} \leq \int_0^n \text{Area}(X \cap S_t) \, dt \leq C_1 \cdot \text{Volume}(X).
\end{equation}

The coarea formula also gives an inequality between the area of a surface $F$ in $B_n$ and the lengths of the curves of intersection of $F$ with slices of $B_n$. There is a constant $C_2$ such that

\begin{equation}
\int_0^n \text{Length}(F \cap S_t) \, dt \leq C_2 \cdot \text{Area}(F).
\end{equation}

The constants $C_1$ and $C_2$ are determined by the geometry of a single block $B$, and do not depend on $n$. In a product metric $C_1 = C_2 = 1$, and equality occurs in Equation (2) when $F$ is itself a product surface, perpendicular to each slice.

The Riemannian metric on $B_n$ is not a product metric and the magnitude of $C_1$ and $C_2$ gives a measure of how far this metric differs from a product metric in a neighborhood of a slice. Since the local geometry near a slice in $B_n$ is determined by the local geometry of the slices in a single block, the constants $C_1$ and $C_2$ do not depend on the number of blocks $n$.

To estimate the volume of $X$, we integrate the area of its cross sections. To do so, we need to consistently specify for each slice $S_t$ which of the two complementary subsurfaces of $F \cap S_t$ lies in $X$. If Length$(F \cap S_t) < \epsilon$ then we take the side of $F$ containing the complementary subsurface of smaller area in $S_t$. This subsurface has area less than $K_1 \epsilon$ by Lemma 4.2 and misses the large side of $F \cap S_t$. Each such subsurface lies on the same side of $F$ (mod 2) by Lemma 4.3. For slices in which Length$(F \cap S_t) \geq \epsilon$ the construction given in Lemma 4.2 gives a bound $K_1 \cdot$ Length$(F \cap S_t)$ for the area of each of the two sides in $S_t$ of $F \cap S_t$, and in particular for the side lying in $X$. By Equation (1), $X$ has volume bounded above by

$$\text{Volume}(X) \leq C_1 \int_0^n \text{Area}(X \cap S_t) \, dt \leq C_1 K_1 \int_0^n \text{Length}(F \cap S_t) \, dt \leq C_1 K_1 C_2 A_0 = C_0.$$ 

This bound is independent of $n$, so as $n$ grows, $\text{Volume}(X)/\text{Volume}(B_n)$ tends to 0. \quad $\Box$

\textit{Geometry \& Topology, Volume 13 (2009)}
Note that Lemma 4.4 is false for surfaces of genus \( k \geq g \). A slice of genus \( g \) splits the volume of \( B_n \) in half and the volume of each half is unbounded with \( n \), showing that one cannot hope to bound the volume of a spanning \( 3 \)-chain by a function of the area of a boundary surface when the surface has genus \( g \).

Finally we give an isoperimetric inequality that states that small area surfaces in a Riemannian \( 3 \)-manifold bound regions of small volume. A more general result is proved in Federer [6, Theorem 4.4.2]. The result is valid for both \( \mathbb{Z} \) and \( \mathbb{Z}_2 \) coefficients, though we need only the latter.

**Lemma 4.5** Let \( M \) be a compact, connected Riemannian \( 3 \)-manifold and \( F \) a proper surface in \( M \), not necessarily embedded or connected. Given \( \epsilon_0 > 0 \) there is a \( \delta_0 > 0 \) such that if \( \text{Area}(F) < \delta_0 \) then \( F \) bounds a \( 3 \)-chain \( X \) in \( M \) with \( \text{Volume}(X) < \epsilon_0 \).

## 5 Harmonic and bounded area maps

Eells and Sampson showed that a map from a Riemannian surface \( F \) with metric \( h \) to a negatively curved manifold \( M \) is homotopic to a harmonic map. This harmonic map is unique unless its image is a point or a closed geodesic, cases that we will not need to consider [5; 7]. A map from a surface to a \( 3 \)-manifold is *elementary* if it induces a homomorphism of fundamental groups whose image is trivial or cyclic, and we will be considering nonelementary maps. The harmonic map is obtained by deforming an initial map of a surface with domain having a fixed Riemannian metric (or complex structure) in the direction of fastest decrease for the energy, along a vector field called the *tension field*. The resulting harmonic map depends continuously on the metrics of both the domain and the image, as shown by Eells–Lemaire [4, Theorem 3.1] and also by Sampson [17].

**Lemma 5.1** [4] Let \( \{g_u : u \in U\} \) be a family of metrics on a closed surface, parametrized by an compact set \( U \subset \mathbb{R}^n \) and let \( M \) be a closed manifold with negative sectional curvature. Then a continuous family of nonelementary smooth maps \( \{f_u : (F, g_u) \to M\} \) is homotopic to a family of harmonic maps \( \{h_u : (F, g_u) \to M\} \) by a homotopy along which the energy of each map is nonincreasing.

Since harmonic maps satisfy a maximal principle, their induced metric is more negatively curved than the ambient manifold wherever they are immersed [17]. The Gauss–Bonnet theorem then implies an area bound proportional to the genus. The immersion assumption can be removed via an approximation of a general map by immersions. A detailed argument is given in Theorem 3.2 of [10].
Lemma 5.2  A harmonic map $f : F \to M$ from a Riemannian genus–$g$ surface to a hyperbolic 3–manifold $M$ has area bounded above by $4\pi (g-1)$. If $M$ has sectional curvatures bounded above by $-r^2$ then the area of the surface is bounded above by $4\pi (g-1)/r^2$.

For any Riemannian metric on $F$, a basic inequality relates the energy and area of a map $f$ [3]:

$$\text{Energy}(f) \geq 2 \text{Area}(f).$$

Equality occurs precisely when $f$ is almost conformal, i.e., conformal except possibly at finitely many singular points with zero derivative. In particular, equality holds for an isometric immersion, an immersion of a surface into a Riemannian manifold in which the domain has the induced metric.

Lemma 5.3  A continuous family of immersions $\{f_{u,0} : F \to M, u \in U\}$ from a surface $F$ to a closed negatively curved 3–manifold $M$, parameterized by a compact set $U \subset \mathbb{R}^n$, is homotopic to a family of harmonic maps $\{f_{u,1} : (F, g_u) \to M\}$ through maps $\{f_{u,s}, 0 \leq s \leq 1\}$ satisfying

$$\text{Area}(f_{u,s}) \leq \text{Area}(f_{u,0}).$$

Proof  We begin by taking the induced family of metrics $\{g_u\}$ on the surface $F$ obtained from the family of immersions $f_{u,0}$. This gives a family of isometric immersions of $F$, for each of which the energy equals twice the area. The Eells–Sampson process of deforming along the tension field gives an energy decreasing homotopy of each surface in this family that converges to a harmonic map [5]. Since the energy is nonincreasing during this flow, and the area is bounded above by twice the energy, the area of each surface in this flow is bounded above by its initial value.

6  Sweepouts

Consider a 1–parameter family of maps $k : F \times (-1, 1) \to M$ such that

$$\lim_{t \to \pm 1} \text{Area}(k(F, t)) = 0$$

and let $F_t$ denote the surface $k(F, t)$. By Lemma 4.5 we know that there is a 3–chain bounded by $F_t$ whose volume approaches 0 as $t \to \pm 1$. It follows that there is a constant $\beta > 0$ such that one side of $F_t$ has volume less than $\text{Volume}(M)/2$ if $t \in (-1, -1 + \beta] \cup [1 - \beta, 1)$. In particular $F_{-1+\beta}$ bounds a 3–chain $C_{-1+\beta}$ of volume less than $\text{Volume}(M)/2$ and similarly $F_{1-\beta}$ bounds such a chain $C_{1-\beta}$. Consider the...
3–cycle $Z \in H_3(M; \mathbb{Z}_2)$ formed by the union of the 3–chains $k(F \times [-1 + \beta, 1 - \beta])$, $C_{-1 + \beta}$ and $C_{1 - \beta}$. We say the family of maps $k$ has degree one if $Z$ represents the fundamental homology class of $M$ (with $\mathbb{Z}_2$–coefficients) and in that case we say $k$ is a sweepout of $M$. Note that changing the value of $\beta$ leads to a continuous change of the volume of $Z$, while a change in the homology class of $Z$ would cause the volume to change by a multiple of the volume of $M$. Thus the choice of $\beta$ does not affect whether $k$ has degree one.

We digress somewhat to point out that harmonic maps gives a new approach to constructing sweepouts of bounded area. Pitts and Rubinstein showed the existence of an unstable minimal surface in a 3–manifold using a minimax argument [13]. Starting with a strongly irreducible Heegaard splitting, they obtain a minimal surface that has maximal area in a 1–parameter family of surfaces obtained from a sweepout. Since a genus–$g$ minimal surface in a hyperbolic manifold has area less than $4\pi (g-1)$, this implies the same area bound for each surface in the sweepout. As noted by Rubinstein [16] and by Bachman, Cooper and White [1], the existence of a sweepout by bounded area surfaces has implications on the geometry and Heegaard genus of a 3–manifold. Bachman, Cooper and White constructed a sweepout using piecewise-geodesic surfaces with two vertices, which gives a somewhat weaker area bound.

Harmonic maps give an alternate way to obtain a sweepout by bounded area surfaces, with the same area bound implied by Pitts–Rubinstein, but without assuming that the Heegaard splitting is strongly irreducible. The harmonic sweepout has the drawback of allowing singular surfaces, like [1], but unlike [13], where a sweepout by embedded surfaces is implied.

**Theorem 6.1** If $M$ is a hyperbolic 3–manifold with a genus–$g$ Heegaard splitting then $M$ has a sweepout in which each surface has area bounded above by $4\pi (g-1)$.

**Proof** Putting the induced metric on each surface associated to a Heegaard splitting sweepout gives a continuous family of isometric embeddings

$$\{ f_{0,t}: (F, g_t) \to M, -1 < t < 1 \},$$

with $g_t$ the induced metric on $F$ under the map $f_{0,t}$. By Lemma 5.1, this family of maps is homotopic to a family of harmonic maps $\{ f_{1,t}: (F, g_t) \to M, -1 < t < 1 \}$. Lemma 5.3 implies that during the homotopy from $f_{0,t}$ to $f_{1,t}$, the area of $(f_{s,t}$ stays below the initial area of $f_{0,t}$. So for each $s$ we have

$$\lim_{t \to \pm 1} \text{Area}(f_{s,t}) = 0.$$
Stabilization of Heegaard splittings

Since $f_{s,t}$ changes continuously with $s$ the degree remains one and $f_{s,t}$ forms a sweepout for each fixed value of $s$. Each map $f_{1,t}$ in the final sweepout has area bounded above by $4\pi(g-1)$, by Lemma 5.2.

Given a sweepout of a 3–manifold $M$ and two subsets $L$ and $R$ with disjoint interiors, we can characterize the direction of the sweepout relative to $L$ and $R$, describing which of $L$ and $R$ is first engulfed. Let $\beta$ be a constant that satisfies the conditions of Lemma 4.5, so that for $-1 < s \leq -1 + \beta$, $F_s$ bounds a 3–chain of volume less than $\text{Volume}(L \cup R)/2$. If $-1 < t < -1 + \beta$ define $K_t$ to be the $\mathbb{Z}_2$ 3–chain $C_t$ with boundary $F_t$, given by the side of $F_t$ having smaller volume. If $t \geq -1 + \beta$ define $K_t$ to be the 3–chain with boundary $F_t$ obtained by adding the 3–chains $k(F \times [-1 + \beta, t])$ and $C_{-1 + \beta} \pmod{2}$. As $t$ increases on $(-1, 1)$ the volume of $K_t$ changes continuously, while

$$\lim_{t \to -1} \text{Volume}(K_t \cap (L \cup R)) = 0$$

and

$$\lim_{t \to 1} \text{Volume}(K_t \cap (L \cup R)) = \text{Volume}(L \cup R).$$

Note that we are using $\mathbb{Z}_2$–chains, so that the volume of $K_t$ need not grow monotonically with $t$.

Let $Y$ denote the set of points $t \in (-1, 1)$ where the volume of $K_t \cap (L \cup R)$ equals half the volume of $L \cup R$,

$$Y = \left\{ t : \frac{\text{Volume}(K_t \cap (L \cup R))}{\text{Volume}(L \cup R)} = \frac{1}{2} \right\}.$$

For a sweepout of disjoint embedded surfaces coming from a Heegaard splitting, $Y$ contains exactly one point. For a generic sweepout $Y$ is finite and contains an odd number of points.

**Definition** A point $t \in Y$ is an $L$ point if

$$\text{Volume}(K_t \cap L) > \text{Volume}(K_t \cap R)$$

and an $R$ point otherwise. A sweepout is an $LR$–sweepout if it has an odd number of $L$ points and an $RL$–sweepout if it has an odd number of $R$ points.

The Heegaard splitting $E_0$ gives rise to an $LR$–sweepout that begins with surfaces near a graph at the spine of $H_L$, sweeps out $L \cup R$ with embedded slices, and ends with surfaces that collapse to a graph at the spine of $H_R$. The stabilized Heegaard splitting $G_0$ gives rise to an $LR$–sweepout by embedded surfaces of genus $2g-1$. A stabilization adds a loop to each of the graphs forming the spines of $E_0$, the two loops...
linking once in \( M_g \), and each crossing the separating surface \( S \). Between the resulting graphs is a product region which can be filled with Heegaard surfaces of the stabilized splitting. See Figure 3, which shows two spines and a Heegaard surface that has been stabilized once. Note that for arbitrarily large \( n \) and any \( \delta > 0 \), the stabilizations can be chosen so that the added area in going from the surfaces of \( E_0 \) to those of \( G_0 \) is less than \( \delta \). When \( G_0 \) and \( G_1 \) are equivalent, composing the Heegaard sweepout of \( G_0 \) with the diffeomorphisms \( \{ I_s, 0 \leq s \leq 1 \} \) gives a family of genus \( 2g - 1 \) Heegaard sweepouts connecting \( G_0 \) and \( G_1 \).

![Figure 3: Stabilized spines](image)

**Lemma 6.1** Suppose that \( E_0 \) is equivalent to \( E_1 \) after \( (g - 1) \) stabilizations. Then there is a constant \( A_0 \), independent of the number of blocks \( n \), and a family of embedded surfaces 
\[
\{ f_{s,t}: F \to M_g, \ 0 \leq s \leq 1, \ -1 < t < 1 \}
\]
satisfying the following conditions:

1. \( F \) has genus \( 2g - 1 \).
2. For each \( s \), \( \{ f_{s,t}: F \to M_g, \ -1 < t < 1 \} \) is a sweepout.
3. \( \{ f_{0,t}(F), \ -1 < t < 1 \} \) is an LR–sweepout.
4. \( \{ f_{1,t}(F), \ -1 < t < 1 \} \) is an RL–sweepout.
5. Each surface in the two sweepouts \( f_{0,t}(F) \) and \( f_{1,t}(F) \) has area bounded above by \( A_0 \).

**Proof** \( M_g \) is formed from the union of a handlebody \( H_L \), \( n \)–blocks forming \( L \), \( n \)–blocks forming \( R \), and a second handlebody \( H_R \). The geometry of a single block, and of each of \( H_L, H_R \), does not depend on \( n \). Pick a spine for each of \( H_L, H_R \) and foliate the complement of this spine in each handlebody by embedded Heegaard surfaces, connecting the spine to the slice that forms the boundary of each handlebody. Then fill \( L \) and \( R \) with interpolating slices. This Heegaard sweepout is a foliation of the complement of the two spines in \( M_g \) by genus– \( g \) leaves \( \{ L_t, -1 < t < 1 \} \).

Now stabilize by adding \( (g - 1) \) loops to each spine and \( (g - 1) \) handles to each surface between the two spines, giving a Heegaard sweepout of \( G_0 \). Let \( a_0 > 0 \) and \( v_0 > 0 \)
be two constants. By adding thin handles in the stabilization, we can arrange that the area of any surface increases by less than $a_0$ and so that the volume bounded by the additional $(g - 1)$ 1–handles is less than $v_0$. For our purpose it suffices to take

$$v_0 = a_0 = \min\{1, \text{Volume}(B)/2\}.$$ 

Let $F$ be a surface of genus $2g - 1$ and construct maps $\{f_{0,t}: F \to M_g, -1 < t < 1\}$ that smoothly parametrize the stabilized surfaces. We will refer to $f_{0,t}(F)$ as $F_{0,t}$. Set the value of the constant $A_0$ to be the largest area of the surfaces $\{F_{0,t}, -1 < t < 1\}$. Note that our construction gives a value for $A_0$ that is determined by $g$, the area of the slices in a single block, and the geometry of $H_L$ and $H_R$. In particular $A_0$ is independent of the number of blocks $n$ used to construct $M_g$.

The Heegaard splitting $G_0 = (\tilde{H}_1, \tilde{H}_2, \tilde{S})$ obtained by stabilizing $E_0$ $(g - 1)$ times is equivalent to the reversed splitting $G_1 = (\tilde{H}_2, \tilde{H}_1, \tilde{-S})$, where $\tilde{S} = F_{0,0}$ and $\tilde{-S}$ indicates the orientation of $\tilde{S}$ has reversed. So an isotopy $I_s, 0 \leq s \leq 1$ from the identity map $I_0$ to a diffeomorphism $I_1$ carries $(\tilde{H}_1, \tilde{H}_2, \tilde{S})$ to $(\tilde{H}_2, \tilde{H}_1, \tilde{-S})$. Construct the family of surfaces $f_{s,t}: F \to M_g$ by defining

$$f_{s,t} = I_s \circ f_{0,t}: F \to M_g$$

and let $F_{s,t}$ denote $f_{s,t}(F)$.

For any constant $0 < \alpha < 1$, we can arrange, by stretching out a collar around the invariant surface $F_{0,0}$, that $I_1$ carries $F_{0,t}$ to $F_{0,-t}$ for each $t \in [-1 + \alpha, 1 - \alpha]$. For $\alpha$ sufficiently small the embedded surfaces

$$\{F_{s,t}: t \in (0, -1 + \alpha] \cup [1 - \alpha, 1)\}$$

lie in small neighborhoods of the images of the spines of $G_0$ under $I_s$, have area uniformly bounded above by $a_0$, and bound submanifolds having volume less than $v_0$.

The Heegaard splitting $E_0$ gives rise to a sweepout $\{L_t, -1 < t < 1\}$ of $M_g$ by genus–$g$ surfaces, starting near the spine of $H_L$ and ending near the spine of $H_R$. This sweepout foliates the complement of the two spines. The surface $L_t$ bounds a 3–chain $K_t$ that fills up the side containing $\{\bigcup L_{t'}: t' < t\}$. $K_0$ bisects the volume of $L \cup R$ and contains $L$ but not $R$. The surfaces $L_t$, when parametrized, give an $LR$–sweepout with $\text{Volume}(K_0 \cap L) = n\text{Volume}(B)$, while $\text{Volume}(K_0 \cap R) = 0$. The surface $F_{0,t}$, obtained by stabilizing $L_t$, bounds a 3–chain $K_{0,t}$ whose volume of intersection with $L$ and $R$ differs by less than $v_0$ from that of $K_t$. Therefore $F_{0,t}$ also gives an $LR$–sweepout. By the same argument applied to the stabilization of $E_1$, we have that $F_{1,t}$ gives an $RL$–sweepout. All properties now follow.  

$\square$

*Geometry & Topology, Volume 13 (2009)*
We now show that a path of sweepouts in \( M_g \) whose surfaces have area uniformly bounded by a constant \( A_0 \) cannot start with an \( LR \)-sweepout and end with \( RL \)-sweepout if \( n \) is sufficiently large. Define
\[
V_L = \text{Volume } (L) = \text{Volume } (R)
\]
so that \( 2V_L = \text{Volume } (L \cup R) \).

**Lemma 6.2** Given a constant \( A_0 \) there is a constant \( n_0 \) such that if \( n > n_0 \) and \( M_g \) is constructed with \( 2n \) blocks then there does not exist a smooth family of maps
\[
\{h_{s,t}: F \to M_g, \ 0 \leq s \leq 1, -1 < t < 1\}
\]
from a surface \( F \) to \( M_g \) satisfying the following conditions:

1. \( F \) has genus less than \( 2g \).
2. For each fixed \( s \), \( \{h_{s,t}: F \to M_g, -1 < t < 1\} \) is a sweepout.
3. Each surface \( h_{s,t}(F) \) has area bounded above by \( A_0 \).
4. \( \{h_{0,t}(F), -1 < t < 1\} \) is an \( LR \)-sweepout.
5. \( \{h_{1,t}(F), -1 < t < 1\} \) is an \( RL \)-sweepout.

**Proof** Suppose such a family exists for all \( n \) and let \( F_{s,t} \) denote \( h_{x,t}(F) \). For each \( s, t \) we now construct a 3–chain \( K_{s,t} \) with boundary \( F_{s,t} \). First pick \( \beta \) sufficiently small to satisfy the conditions of Lemma 4.5 for each \( 0 \leq s \leq 1 \). Then for each \( s \) and each \( t_0 \in (-1, -1+\beta] \) there is a 3–chain \( C_{s,t_0} \) with boundary \( F_{s,t_0} \) and volume less than \( V_L/2 \). If \( -1 < t < -1 + \beta \) take \( K_{s,t} \) to be equal to \( C_{s,t} \). If \( -1 + \beta < t < 1 \) take \( K_{s,t} \) to be the sum of \( C_{s,-1+\beta} \) and \( \bigcup_{-1+\beta \leq t' \leq t} F_{s,t'} \).

Note that the chains \( K_{s,t} \) and their volumes are defined with \( \mathbb{Z}_2 \)-coefficients. Thus even though for each \( s \) the chains \( K_{s,t} \) are images of nested sets that grow with \( t \), as \( \mathbb{Z}_2 \)-chains the volume of their image can both grow and shrink as \( t \to 1 \). Since they are null-homologous, the surfaces \( F_{s,t} \) have complement in \( M_g \) that can be two-colored into black and white regions, and the volume of \( K_{s,t} \) is equal to the volume of one of these regions. The volume of \( K_{s,t} \) varies continuously with \( s \) and \( t \), giving a continuous function from \([0, 1] \times (-1, 1) \to \mathbb{R} \).

For each fixed \( s \) the surfaces \( F_{s,t} \) sweep out \( M_g \) with degree one, so
\[
\lim_{t \to -1} \text{Volume}(K_{s,t} \cap (L \cup R)) = 0
\]
and
\[
\lim_{t \to 1} \text{Volume}(K_{s,t} \cap (L \cup R)) = 2V_L.
\]
Set \[ Q = \left\{ (s, t) : \frac{\text{Volume}(K_{s,t} \cap (L \cup R))}{2V_L} = \frac{1}{2} \right\}. \]

If necessary, perturb the value 1/2 used to define \( Q \) to a regular value of the volume function evaluated on the rectangle. Then \( Q \) is a proper 1–manifold and any path from the edge \( t = -1 \) to the edge \( t = 1 \) of the \((s, t)\) rectangle must cross \( Q \). It follows that there is a path contained in \( Q \) connecting the edges \( s = 0 \) and \( s = 1 \) of the \((s, t)\) rectangle, as in Figure 4.

![Figure 4](image)

Figure 4: Each curve in \( Q \) consists entirely of \( L \) points or \( R \) points.

We now claim that each component of \( Q \) consists entirely of \( L \) points or entirely of \( R \) points. If a component of \( Q \) has both \( L \) points and \( R \) points then there is a point \((s, t) \in Q\) on that component where

\[ \text{Volume}(K_{s,t} \cap L) = \text{Volume}(K_{s,t} \cap R) \]

and since these two volumes sum to \( \text{Volume}(K_{s,t} \cap (L \cup R)) = V_L \), we have

\[ \frac{\text{Volume}(K_{s,t} \cap L)}{2V_L} = \frac{\text{Volume}(K_{s,t} \cap R)}{2V_L} = \frac{1}{4}. \]

The area of \( F_{s,t} \) is bounded above by \( A_0 \), so Lemma 4.4 implies that for \( n \) sufficiently large, \( F_{s,t} \) bounds a 3–chain \( X_{s,t} \) in \((L, \partial L)\) with volume less than \( V_L/3 \). Now \((K_{s,t} \cap L)\) and \( X_{s,t} \) have the same boundary in \((L, \partial L)\). Any two 3–chains in \((L, \partial L)\) with the same boundary sum to a 3–chain whose volume is a multiple of \( V_L \), so we have

\[ 0 \leq \frac{\text{Volume}(K_{s,t} \cap L)}{V_L} \leq \frac{1}{3} \]

or

\[ \frac{2}{3} \leq \frac{\text{Volume}(K_{s,t} \cap L)}{V_L} \leq 1. \]
As a fraction of $2V_L$

$$0 \leq \frac{\text{Volume}(K_{s,t} \cap L)}{2V_L} \leq \frac{1}{6}$$

or

$$\frac{1}{3} \leq \frac{\text{Volume}(K_{s,t} \cap L)}{2V_L} \leq \frac{1}{2}.$$ 

This contradicts Equation (3). We conclude that components of $Q$ consist either entirely of $L$ points or entirely of $R$ points.

A component curve in $Q$ either

1. meets each edge $s = 0$ and $s = 1$ once,
2. meets one edge of the $(s, t)$ rectangle twice, or
3. is a closed curve disjoint from both.

The parity of the number of $L$ points and the number of $R$ points on the two families $h_{0,t}(F)$ and $h_{1,t}(F)$ is the same for each edge $s = 0$ and $s = 1$, and thus either both are $LR$–sweepouts or both are $RL$–sweepouts.

We now show how to homotop a path of sweepouts whose initial and final sweepouts contain surfaces with area less than $A_0$ to a new path of sweepouts all of whose surfaces have area bounded above by $A_0$.

**Lemma 6.3** There is a constant $n_0$ such if $n > n_0$ such that if $M_g$ is constructed with $2n$ blocks and \{ $f_{s,t,0}$: $F \to M_g$, $0 \leq s \leq 1$, $-1 < t < 1$ \} is a family of surface maps such that

1. the genus of $F$ is less than $2g$,
2. for each $s$, \{ $f_{s,t,0}$: $F \to M_g$, $-1 < t < 1$ \} is a sweepout,
3. each surface in the two sweepouts $f_{0,t,0}(F)$ and $f_{1,t,0}(F)$ has area bounded above by a constant $A_0$,
4. $f_{0,t,0}(F)$, $-1 < t < 1$ is an $LR$–sweepout,
5. $f_{1,t,0}(F)$, $-1 < t < 1$ is an $RL$–sweepout,
6. each surface \{ $f_{s,t,0}$: $F \to M_g$, $0 \leq s \leq 1$, $-1 < t < 1$ \} is immersed,

then there is a homotopy

\{ $f_{s,t,u}$: $F \to M_g$, $0 \leq s \leq 1$, $-1 \leq t \leq 1$, $0 \leq u \leq 1$ \}

to a family of maps $f_{s,t,1}$ such that
(1) for each fixed \( s, u; \{ f_{s,t,u}: F \to M_g, -1 \leq t \leq 1 \} \) is a sweepout,
(2) each surface \( f_{s,1}(F) \) is harmonic, with area at most \( 8\pi(2g-1) \),
(3) \( f_{0,t,1}(F) \), \(-1 < t < 1\) is an LR–sweepout,
(4) \( f_{1,t,1}(F) \), \(-1 < t < 1\) is an RL–sweepout.

**Proof** Construct a family of Riemannian metrics \( h_{s,t} \) on \( F \) for each \( 0 \leq s \leq 1, -1 < t < 1 \) by taking the induced metric pulled back from \( M_g \) by the immersion \( f_{s,t,0} \).

![Figure 5: Maps \( \{ f_{s,t,u}: 0 \leq s \leq 1, -1 \leq t \leq 1, 0 \leq u \leq 1 \} \) are parametrized by points in a cube. The back face, where \( u = 0 \), represents the initial path of Heegaard surfaces that connects \( G_0 \) and \( G_1 \). The front face corresponds to harmonic maps giving a homotopic path of Heegaard sweepouts. The curves indicate parameter values on the boundary of the cube corresponding to surfaces that bisect \( L \cup R \).

Then each map \( f_{s,t,0} \) is conformal and has energy equal to twice its area. We flow the family \( f_{s,t,0} \) to a family of harmonic maps \( f_{s,t,1} \) as in Lemma 5.3. Each surface in the sweepout \( f_{0,t,0}(F) \) has area bounded above by \( A_0 \), and this area bound is maintained for each surface \( \{ f_{0,t,u}(F), 0 \leq u \leq 1 \} \) in the homotopy to harmonic maps, as in Lemma 5.2. Similarly the homotopy to harmonic maps of the sweepout \( f_{1,t,0}(F) \) given by \( \{ f_{1,t,u}(F), 0 \leq u \leq 1 \} \) has surfaces whose area is uniformly bounded above by \( A_0 \). Thus if \( n \) is sufficiently large, the sweepout \( \{ f_{0,t,1}(F) \} \) is an LR–sweepout, by Lemma 6.2 and similarly \( \{ f_{1,t,1}(F) \} \) is an RL–sweepout. \( \square \)
7 Inverting Heegaard surfaces

We now prove the main result:

**Proof of Theorem 1.1** We begin with the two genus–g Heegaard foliations, \(E_0\) and \(E_1\) of \(M_g\). We will show that if \(n\) is sufficiently large then these splittings are not equivalent with fewer than \(g\) stabilizations. It suffices to show that they are not equivalent after \(g - 1\) stabilizations.

Assume to the contrary that for all \(n\), \(G_0\) and \(G_1\) are equivalent after \(g - 1\) stabilizations. Lemma 6.1 implies there is a constant \(A_0\) independent of \(n\) and a family of sweepouts \(\{f_{s,t,0}: F \to M_g, 0 \leq s \leq 1, -1 < t < 1\}\) of genus \(2g - 1\) with \(\{f_{0,t,0}(F), -1 < t < 1\}\) an LR–sweepout, \(\{f_{1,t,0}(F), -1 < t < 1\}\) an RL–sweepout with each surface in the two sweepouts \(f_{0,t,0}(F)\) and \(f_{1,t,0}(F)\) having area bounded above by \(A_0\). Lemma 6.3 implies that for \(n\) sufficiently large this family can be deformed to a new family of sweepouts \(\{f_{s,t,1}: F \to M_g, 0 \leq s \leq 1, -1 < t < 1\}\) in which \(\{f_{0,t,1}(F), -1 < t < 1\}\) is an LR–sweepout, \(\{f_{1,t,1}(F), -1 < t < 1\}\) is an RL–sweepout and each surface \(f_{s,t,1}(F)\) also has area bounded above by \(A_0\). Then \(\{f_{s,t,1}, 0 \leq s \leq 1\}\) is a path of sweepouts that connects an LR–sweepout \(f_{0,t,1}\) to a RL–sweepout \(f_{1,t,1}\), and all of the surfaces in this path of sweepouts have area less than \(A_0\). Lemma 6.2 states that no such path of sweepouts can exist, contradicting the assumption that fewer than \(g\) stabilizations can make \(E_0\) and \(E_1\) equivalent. 

8 Hyperbolic examples

The Riemannian manifolds \(M_g\) used in our construction were negatively curved, but not hyperbolic. We now show how to construct a family of hyperbolic manifolds that give somewhat weaker lower bounds on the number of stabilizations required to make two genus–g splittings equivalent. Note that block manifolds cannot be isometrically embedded into a hyperbolic 3–manifold in which their slices are Heegaard surfaces, because the fibers of a surface bundle lift to planes in the universal cover \(H^3\), while Heegaard surfaces lift to compressible surfaces. However there does exist an isometric embedding of block manifolds into a hyperbolic manifold in which slices are separating incompressible surfaces. These surfaces become Heegaard surfaces after two stabilizations.

Let \(N_0\) be a hyperbolic 3–manifold that is a union of two \(I\)–bundles over a nonorientable surface, glued along their common genus \(k\) boundary surface, where \(k\) is an
integer greater than one. Such hyperbolic manifolds are double covered by a hyperbolic surface bundle over $S^1$. Some explicit examples can be found in [14]. The boundary surface of each $I$–bundle has a neighborhood isometric to a neighborhood of a fiber in its double cover. We can cut open along this fiber and insert a block manifold with $n$ blocks to obtain a hyperbolic manifold $N$, still homeomorphic to a union of two $I$–bundles. We let $S$ denote a surface separating the two $I$–bundles in the center of the block manifold.

Removing a neighborhood of an interval fiber from a nontrivial $I$–bundle with a genus $k$ boundary surface results in a handlebody of genus $k + 1$. Thus $N$ contains a genus $g = k + 2$ Heegaard surface $S'$ obtained by adding a $1$–handle to each side of $S$, with the core of each $1$–handle an interval in each $I$–bundle. These two handlebodies give rise to two Heegaard splittings of $N$, and to corresponding Heegaard sweepouts that fill the two $I$–bundles in opposite order. The arguments applied to $M_g$ in Section 7 now apply to show that $N$ has two genus–$g$ splittings that require no fewer than $g - 4$ stabilizations to become equivalent.

References

[1] D Bachman, D Cooper, M E White, Large embedded balls and Heegaard genus in negative curvature, Algebr. Geom. Topol. 4 (2004) 31–47 MR2031911

[2] I Chavel, Isoperimetric inequalities. Differential geometric and analytic perspectives, Cambridge Tracts in Math. 145, Cambridge Univ. Press (2001) MR1849187

[3] J Eells, L Lemaire, A report on harmonic maps, Bull. London Math. Soc. 10 (1978) 1–68 MR495450

[4] J Eells, L Lemaire, Deformations of metrics and associated harmonic maps, Proc. Indian Acad. Sci. Math. Sci. 90 (1981) 33–45 MR653945

[5] J Eells, J H Sampson, Harmonic mappings of Riemannian manifolds, Amer. J. Math. 86 (1964) 109–160 MR0164306

[6] H Federer, Geometric measure theory, Grund. der math. Wissenschaften 153, Springer, New York (1969) MR0257325

[7] P Hartman, On homotopic harmonic maps, Canad. J. Math. 19 (1967) 673–687 MR0214004

[8] J Hass, P Scott, The existence of least area surfaces in 3–manifolds, Trans. Amer. Math. Soc. 310 (1988) 87–114 MR965747

[9] R Kirby, Problems in low dimensional manifold theory, from: “Algebraic and geometric topology (Proc. Sympos. Pure Math., Stanford Univ., 1976), Part 2”, (R J Milgram, editor), Proc. Sympos. Pure Math., XXXII, Amer. Math. Soc. (1978) 273–312 MR520548
[10] Y N Minsky, Harmonic maps into hyperbolic 3–manifolds, Trans. Amer. Math. Soc. 332 (1992) 607–632 MR1100698
[11] E E Moise, Geometric topology in dimensions 2 and 3, Graduate Texts in Math. 47, Springer, New York (1977) MR0488059
[12] H Namazi, J Souto, Heegaard splittings and pseudo-Anosov maps, preprint Available at http://www.math.princeton.edu/~hossein/contents/pseudo8.pdf
[13] J T Pitts, J H Rubinstein, Applications of minimax to minimal surfaces and the topology of 3–manifolds, from: “Miniconference on geometry and partial differential equations, 2 (Canberra, 1986)”, (J E Hutchinson, L M Simon, editors), Proc. Centre Math. Anal. Austral. Nat. Univ. 12, Austral. Nat. Univ., Canberra (1987) 137–170 MR924434
[14] A W Reid, A non-Haken hyperbolic 3–manifold covered by a surface bundle, Pacific J. Math. 167 (1995) 163–182 MR1318168
[15] H Rubinstein, M Scharlemann, Comparing Heegaard splittings of non-Haken 3–manifolds, Topology 35 (1996) 1005–1026 MR1404921
[16] J H Rubinstein, Minimal surfaces in geometric 3–manifolds, from: “Global theory of minimal surfaces”, (D Hoffman, editor), Clay Math. Proc. 2, Amer. Math. Soc. (2005) 725–746 MR2167286
[17] J H Sampson, Some properties and applications of harmonic mappings, Ann. Sci. École Norm. Sup. (4) 11 (1978) 211–228 MR510549
[18] M Scharlemann, Heegaard splittings of compact 3–manifolds, from: “Handbook of geometric topology”, (R J Daverman, R B Sher, editors), North-Holland, Amsterdam (2002) 921–953 MR1886684
[19] J Schultens, The stabilization problem for Heegaard splittings of Seifert fibered spaces, Topology Appl. 73 (1996) 133–139 MR1416756
[20] E Sedgwick, An infinite collection of Heegaard splittings that are equivalent after one stabilization, Math. Ann. 308 (1997) 65–72 MR1446199
[21] J Singer, Three-dimensional manifolds and their Heegaard diagrams, Trans. Amer. Math. Soc. 35 (1933) 88–111 MR1501673

JH, AT: Department of Mathematics, University of California Davis, California 95616, USA
WT: Department of Mathematics, Cornell University Ithaca, NY 14853, USA
hass@math.ucdavis.edu, thompson@math.ucdavis.edu, wpt@math.cornell.edu

Proposed: Joan Birman Received: 22 April 2008
Seconded: Jean-Pierre Otal, Ron Stern Revised: 9 February 2009

Geometry & Topology, Volume 13 (2009)