PERGAMON

Topology 39 (2000) 619–641

TOPOLOGY

www.elsevier.com/locate/top

Heegaard structures of manifolds in the Dehn filling space

Yo'av Rieck*

Mathematics Department, University of California, Santa Barbara, CA 93106, USA

Abstract

We prove that after Dehn filling an incompressible torus in the boundary of an a-cylindrical 3-manifold the Heegaard genus degenerates by at most one for all but finitely many fillings. We do so by proving that for all but finitely many fillings the core of the attached solid torus can be isotoped into the minimal Heegaard surface of the filled manifold, we say that these manifolds are "good". For these fillings, after stabilizing the Heegaard surface once, it becomes a Heegaard surface of the original manifold. We show that any two Heegaard surfaces in different fillings, into which the core is not isotopic, can be isotoped to intersect essentially. Using this, a bound on the distance between fillings containing such surfaces is given in terms of the genera of the Heegaard surfaces.

Keywords: 3-manifolds; Heegaard Surfaces; Cerf theory; Dehn filling

1. Introduction

Heegaard structures for 3-manifolds have been around for a long time ([7]). Much progress has been made understanding Heegaard structures of manifolds with certain geometric structure (Seifert Fiber Spaces, graph manifolds), but little is known about manifolds that are given as a Heegaard structure (unless the genus is 2 or less).

Ever since Thurston used Dehn surgery to construct hyperbolic 3-manifolds ([13]) viewing 3-manifolds as originating from Dehn surgery has proven very profitable. In this paper we show that the Heegaard structures of a manifold behave quite nicely under Dehn filling: if M is obtained from X by a Dehn filling, it is clear that any Heegaard surface for X is a Heegaard surface for M (see Section 4 for details). In Sections 5 and 6 we will prove theorems needed to show that the converse is usually true (in a sense that we describe in the sequel).

* Corresponding author. Oklahoma State University, Department of Mathematics, 401 Mathematical Sciences, Stillwater, OK 74078, USA Tel.: + 1-405-744-6296; fax: + 1-405-744-8275.

E-mail address: yoav@math.okstate.edu (Y. Rieck)

0040-9383/00/$ - see front matter
PII: S0040-9383(99)00026-9

In [9] Moriah and Rubinstein studied the Heegaard structure of negatively curved manifolds using analytical techniques. They proved the finiteness result we obtain in Corollary 4.2 (compare with Theorem 0.1 of [9]). The main intent of this paper is to prove this result using purely topological techniques. We do that using Theorems 5.2 and 6.2. We quote Corollary 4.2 here, rephrased to avoid terminology that has not yet been introduced. It is given as a finiteness result; in the body of the paper we will give a concrete upper bound (in terms of the genus).

Corollary 4.2. *Let X be an a-cylindrical 3-manifold, $T \subset \partial X$ an incompressible torus. Fix g and pick Heegaard surfaces (one or more) of genus at most g for each of the manifolds obtained by filling T.*

Then for all but finitely many of the filled manifolds the core of the attached solid torus is isotopic into all the chosen Heegaard surfaces.

Considering filled manifolds with pre-chosen Heegaard surfaces for them, our main concept is that of a *good* Dehn filling, this being one where the core of the attached solid torus (assuming for now we filled just one boundary component) is isotopic into all the given Heegaard surfaces. (Such curves play a role similar to that of geodesics in the study of hyperbolic 3-manifolds.) We show that most fillings are good, and for these fillings one can "push" the Heegaard surface back into X after at most one stabilization, thus increasing the genus by at most 1.

Using this and Corollary 4.2 we obtain Corollary 4.4, where $g(\cdot)$ denotes Heegaard genus.

Corollary 4.4. *For all but finitely many manifolds M resulting from the filling of an incompressible boundary component T of an a-cylindrical manifold X:*

$$g(X) - 1 \leqslant g(M) \leqslant g(X).$$

We discuss this and other geometric interpretations of our work in Sections 3 and 4.

The main theorem which enables us to achieve these goals is Theorem 6.2. This theorem is proven in Section 6 independently of the work done in the rest of this paper. We quote it here, noting to the reader less familiar with this type of technique that essential intersection is defined in Section 2 and is slightly non-standard, as we ignore simple closed curves of intersection entirely, concentrating only on arcs of intersection.

Let (X,T) be a manifold with $T \subset \partial X$ a collection of c tori. Pick T_{j_0} a component of T. Let M be a result of filling X along T_{j_0} (or simply *a filling of* (X,T_{j_0})). Let Σ be a Heegaard surface for M. We say that M is "T_{j_0}-bad" w.r.t Σ if the core of the solid torus attached to T_{j_0} is not isotopic into Σ.

Theorem 6.2. *Let X be a 3-manifold, $T = \bigcup_{j=1}^{c} T_j \subset \partial X$ a collection of c tori. Let $\lambda_{i,j}$ $(i = 1,2,\ j = 1,\ldots,c)$ be slopes on T_j and $\lambda_i = \bigcup_{j=1}^{c} \lambda_{i,j}$. Let $M_i = X(\lambda_i)$ be fillings of (X,T), both T_{j_0}-bad with respect to Σ_i, $i = 1,2$.*

Then the surfaces Σ_i can be isotoped so that $\Sigma_i^* \subset X$ intersect essentially while meeting the boundary component T_{j_0}.

The structure of the paper is as follows: In Section 2 we give the necessary definitions and standardize the notation. In Section 3 we go over the well-known construction of 3-manifolds via Dehn surgery (see, e.g. [11, pp. 275–277]). Our goal is to show that the procedure of Dehn surgery

can be done (using our notion of "good links") while respecting a given Heegaard structure. The main theorems and their corollaries are stated and proven in Section 4. The proofs of the technical theorems are in Section 5 (the combinatorial theorem) and 6 (the Cerf-theoretic argument, which poses the main technical difficulty).

2. Preliminary notations

Except where otherwise stated, all surfaces and manifolds considered are orientable. Let Σ_g be the surface of genus g. For a submanifold A, let $N(A)$ denote a normal neighborhood of A. A manifold pair (A,B) is a 3-manifold A and a surface B embedded in A (possibly in the interior of A, possibly in the boundary of A). A manifold pair (N,Σ_g) is called a *compression body* if it is obtained by attaching two and three handles to $(\Sigma_g \times \mathbf{I}, \Sigma_g \times \{1\})$, $\mathbf{I} = [-1,1]$, where no attachment is performed along $\Sigma_g \times \{1\}$. The boundary of a compression body[1] is partitioned into two parts, $\partial_+ N \cup \partial_- N$. The surface $\partial_+ N = \Sigma_g \times \{1\}$ is called the attaching region (the reason for this terminology is that we use this region to attach the compression body to another compression body to construct manifolds). A compression body for which $\partial N = \partial_+ N$ is a handlebody – view it up-side-down. A closed surface Σ embedded in M is a *Heegaard surface* if M cut open along Σ consists of two compression bodies with the two copies of Σ as their attaching regions.

Every compression body (N, Σ) contains an embedded graph Γ (not unique) so that $N - (\Gamma \cup \partial_-(N))$ is a product $\Sigma \times (-\infty, 0]$. (Existence of such a Γ is easily proven by induction on the number of two and three handles.) $\Gamma \cup \partial_-(N)$ is called a *spine*. Let M be a 3-manifold and Σ a Heegaard surface for it. A *spine* for (M,Σ) is defined as a union of spines of the two complementary compression bodies. Note two things:

- the complement of the spine is $\Sigma \times \mathbb{R}$.
- the spine is ∂M union of an one-dimensional set, so other one-dimensional objects in the interior of M will miss it by transversality.

An important notion for many of our corollaries is stabilization (e.g. [15]). Let Σ be a Heegaard surface for M. In Definition 2.1 we use the product structure $\Sigma \times I$ in a neighborhood of Σ in order to discuss a small straight tube. A tube (i.e. a homeomorph of $\mathbb{S}^1 \times I$) is said to be *small and straight* if $\mathbb{S}^1 \times \{-1\}$ bounds a disk in Σ, and for every point $p \in \mathbb{S}^1$ the arc $\{p\} \times I$ is monotonic w.r.t. the projection of $\Sigma \times I$ into I.

Definition 2.1. Let γ be a (closed) curve that cobounds an annulus with Σ, a Heegaard surface of M. *Stabilizing* Σ along γ to get $S_\gamma(\Sigma) = S(\Sigma)$ means attaching the torus $T = \partial N(\gamma)$ to Σ by a small straight tube[2] as shown in Fig. 1.

[1] We allow ∂N to have $\mathbb{S}^2$ components.

[2] Traditionally, stabilization is defined without respect to a curve, in which case the procedure is unique (up to isotopy of the stabilized surface). The particulars of the construction given in Figure 1 are important for our work in $M - N(\gamma)$. We will *never* stabilize without specifying a curve along which we stabilize (although the subscript γ will often be omitted). In order for the resulting surface to be a Heegaard surface this curve must be parallel to Σ.

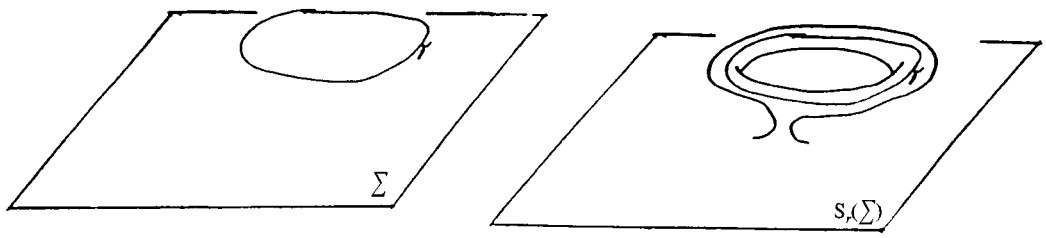


Fig. 1. Stabilization of Σ along γ.

Let Σ_1^* and Σ_2^* be surfaces with boundary properly embedded in some manifold. An arc of intersection α is called *boundary parallel* in Σ_i^* (for either $i = 1$ or 2) if there exists a disk from Σ_i^* whose boundary consists of two arcs meeting at their endpoints, one being α and the other from $\partial\Sigma_i^*$. No requirements about the intersection of that disk with Σ_{3-i}^* are made.

The following is a technical definition which is essential for everything we do, and is therefore singled out here.

Definition 2.2. Let $(\Sigma_i^*,\partial\Sigma_i^*) \subset (X,T)$, $i = 1,2$ be properly embedded surfaces, so that

- $\partial\Sigma_i^* \neq \emptyset$;
- each component of $\partial\Sigma_1^*$ meets each component of $\partial\Sigma_2^*$ minimally;
- no arc of $\partial\Sigma_1^* \cap \partial\Sigma_2^*$ is boundary parallel on either surface.

Then Σ_1^* and Σ_2^* are said to intersect essentially.

Let T be a torus boundary component of X. M is called *Dehn filling* of (X,T) if it is the result of attaching a solid torus V to X along T. The diffeomorphism class of M is determined by the isotopy class of λ, the meridian of V, in T. The set of all manifolds obtained by this procedure is called the Dehn Surgery Space of X with respect to T, denoted $D(X,T)$.[3] After choosing a basis for $H_1(T)$ (or a *framing*) $D(X)$ is parameterized by $\bar{\mathbb{Q}} = \mathbb{Q} \cup \{1/0\}$. More generally, if T is a union of c tori, after choosing a basis for $H_1(T)$, $D(X,T)$ is parameterized by $\bar{\mathbb{Q}}^c$.

The distance between slopes p/q and p'/q' in $H_1(T)$ (or between the corresponding manifolds in $D(X)$) is defined to be $|\Delta(p/q, p'/q')| = |pq' - p'q|$, where $\Delta(\ ,\ )$ is the algebraic intersection form on $H_1(T)$. It is well known that a bounded set in $D(X)$ is finite.

A subset of $D(X)$ is called a *line* if it is of the form $\{n\alpha + \beta | n \in \mathbb{Z}\}$, for some α, $\beta \in H_1(T)$ with $|\Delta(\alpha,\beta)| = 1$. This line is in fact all curves on T that intersect α once. We say that this line is *determined by* α.

Let γ be a knot (or a link component) in some 3-manifold M that can be isotoped into a surface Σ. Isotope γ to be parallel[4] to Σ. For convenience we will use the following parameterization of

[3] Often, T will be suppressed from notation $D(X)$.

[4] Let $\Sigma \times \mathbf{I}$ be a product neighborhood of $\Sigma = \Sigma \times \{0\}$. We will name a knot "parallel" to a Σ, denoted $\gamma \| \Sigma$, if it lies on a level $\Sigma \times \{\varepsilon\}$ for some non-zero ε.

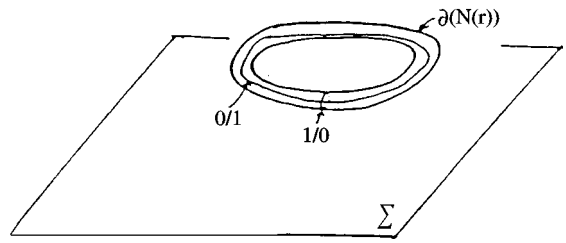


Fig. 2. Parameterization of $N(\gamma)$.

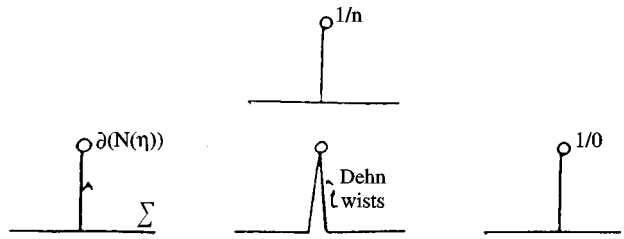


Fig. 3. Good Heegaard surfaces survive $1/n$ surgery.

$H_1(\partial N(\gamma))$ (see Fig. 2): $1/0$ is the meridian in M; 0 is the surface framing (i.e. the curve that cobounds an annulus with the surface).[5] Thus the curves that meet the surface framing exactly once are called $1/n$ (for some n) with respect to Σ. These curves form the line determined by $0/1$ (or by the surface framing).

Let (M,Σ) be a manifold and a Heegaard surface for it. Let γ be a curve parallel to Σ. Let A be an annulus with one boundary component γ, the other on Σ, so that $int(A)\cap\Sigma = \emptyset$. Our reason for picking the framing described above is that performing a $1/n$ surgery on γ (with respect to Σ) is equivalent to the following procedure (see [11] for details), see Fig. 3: cut M open along Σ to obtain N_+, N_-. Say $\gamma \subset N_+$. Cut N_+ open along the annulus A. Glue A after n Dehn Twists, and fill the manifold along the $1/0$ curve, as γ was before. Glue Σ back as before. The resulting manifold is the same as performing a $1/n$ surgery on γ. This new manifold can also be obtained by deforming the map between the attaching regions of the compression bodies by n Dehn twists along the curve on Σ parallel to γ.

The main point to remember from the paragraph above is that Σ will remain a Heegaard surface for the manifold obtained by a $1/n$ surgery on γ.

Let A be a surface or a 3-manifold with boundary. An arc α properly embedded in A is called *boundary parallel* if there exists a disk embedded in A whose boundary consists of two arcs intersecting at their endpoints, one arc being α and the other on ∂A.

By an *essential surface* in X we mean a surface F properly embedded in X with the following two properties: $\pi_1(F)$ injects into $\pi_1(X)$ under the homomorphism induced by inclusion, and every arc properly embedded in F that is boundary parallel in X is boundary parallel in F.

[5] Note that the 0 framing depends not only on the γ and the Σ but also on the particular embedding of γ into Σ chosen.

A manifold pair (X,T) is called *a-cylindrical* if there does not exist an essential annulus $(S^1 \times \mathbf{I},\ S^1 \times S^0)$ embedded in (X,T).

A *virtually finite* set will be defined as follows: let $G = (V,E)$ be an oriented infinite tree so all edges are pointing away from the root. Assume the number of edges pointing away from each vertex is infinite (except the leaves where it is zero) and the number of edges pointing towards a vertex is one (except the root where it is zero). Choose $E_0 \subset E$, a set of edges, so at each vertex v we pick only finitely many edges, *all pointing away from v*. Let V_0 be the set of all vertices that are successors of some edge in E_0, where by succesor we mean a vertex that can be reached from E_0 by a path that is always oriented forward. Any subset of V_0 is called *virtually finite*. Note that the complement of a virtually finite set is always infinite.

3. Families of manifolds

Manifold pairs (M_1,Σ_1) and (M_2,Σ_2) (where Σ_i is a Heegaard surface for M_i) are called *sisters* (or *cousins of the first degree*) if there exist knots γ_i parallel to Σ_i, $i = 1,2$ so that $M_1 - N(\gamma_1) \cong M_2 - N(\gamma_2) = X$, and (in X) $S_{\gamma_1}(\Sigma_1)$ is isotopic to $S_{\gamma_2}(\Sigma_2)$. (Recall that by $S_\gamma(\Sigma)$ we mean the stabilization of Σ along γ, as described in Fig. 1, Section 2.) X is called the *parent manifold*. Note that the condition $\gamma_i \| \Sigma_i$ is equivalent to both M_1 and M_2 being good fillings of X. In this section we will describe the tree obtained by filling multiple boundary components, and see how Heegaard surfaces behave with respect to that structure.

Generally, (M_1,Σ_1) and (M_2,Σ_2) are said to be *cousins of degree c* if there exist knots $\gamma_i \| \Sigma_i$, $i = 1,2$ so that $(M - N(\gamma_i),\ S_{\gamma_i}(\Sigma_i))$ are cousins of degree $c-1$. We will show that *any* two closed manifolds with *any* two Heegaard splittings for them are cousins (of some degree). (If the Heegaard surfaces do not have the same genus stabilize the surface with the lower genus – along any set of curves parallel to it—until the genera are equal.)

So our next goal is to compare two Heegaard surfaces of the same genus, in either the same or different manifolds. The basic techniques we employ in this section have been used in [8] to show that any two manifolds differ by Dehn surgery on a link. We extend them here (using the idea of good links) to show that this can be performed while respecting the Heegaard structures. See, e.g, [11] for more details. First we define one of our main concepts:

Definition 3.1. A link in a manifold M is called *good with respect to* Σ (where Σ is a Heegaard surface for M) if all of its components are isotopic into Σ (perhaps not simultaneously).

Theorem 3.2. *Let* (M_1,Σ_1), (M_2,Σ_2) *be manifold pairs with Heegaard surfaces of the same genus g. Denote the compression bodies of* M_i *cut open along* Σ_i *by* (N_i^+,Σ_i) *and* (N_i^-,Σ_i), $i = 1,2$.

If $(N_1^+,\Sigma_1) \cong (N_2^+,\Sigma_2)$ *and* $(N_1^-,\Sigma_2) \cong (N_2^-,\Sigma_2)$ *then there exist good links* $L_i = \bigcup_{j=1}^{c} \gamma_{i,j} \subset M_i (i = 1,2)$ *so that*

$$(M_1 - N(L_1),\ S^c_{L_1}(\Sigma_1)) \cong (M_2 - N(L_2),\ S^c_{L_2}(\Sigma_2)).$$

Moreover, the filling of each component γ_j *is* $1/n_{i,j}$ *(for some* $n_{i,j} \in \mathbb{Z}$*) with respect to* $S^j(\Sigma)$.

Remark. Although the notion of stabilization was defined along knots (and not links), in the proof we will show that it extends naturally.

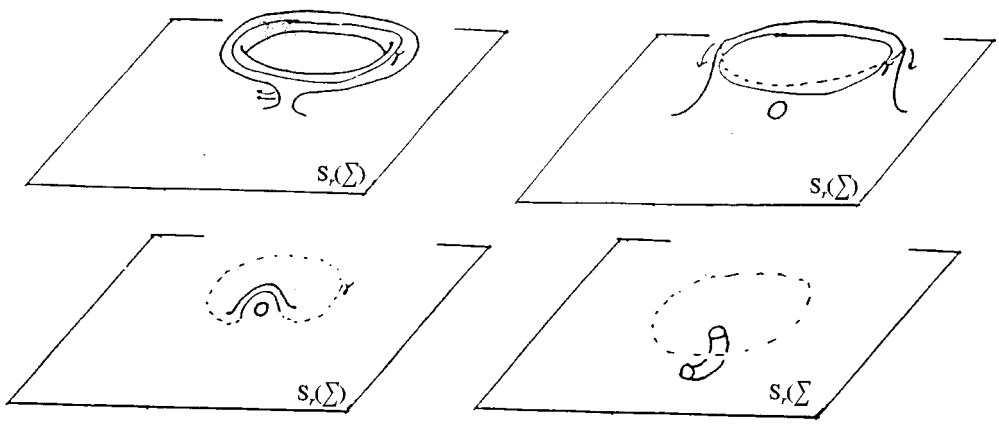


Fig. 4. Pushing the stabilization under the Heegaard surface.

Proof. Let (M_1, Σ_1) and (M_2, Σ_2) be two manifolds and Heegaard surfaces of the same genus – say g – for them. Assume $N_1^+ \cong N_2^+$ and $N_1^- \cong N_2^-$ (of course, this assumption always holds for closed manifolds).

We follow [8] (where it was shown that any two closed manifolds differ by a Dehn surgery).

Consider $N^+ \cong N_1^+$ and $N^- \cong N_1^-$ as two abstract manifolds. Each of these is embedded in both M_1 and M_2. The embedding of the attaching region of N^+ into M_1 followed by the inverse of the embedding of the attaching region of M^- into M_1 yields a homeomorphism f_1 of the surface of genus g. Similarly we obtain f_2.

As with any homeomorphism of a surface, $f_2 f_1^{-1}$ is a composition of Dehn twists (see [2,1]). Name the set of curves along which we twist $\{\zeta_j\}_{j=1}^c$ (in order). Pick L_1 to be the link whose components $\gamma_{1,j}$ lie one above the other, all above Σ_1, and $\gamma_{1,j} \| \zeta_j$. (For the notion of height use the structure of $\Sigma_i \times \mathbb{R}$ in a neighborhood of Σ_i.)

Note that (as mentioned in Section 2) performing a $1/n_j$ Dehn surgery on L_1 will deform f_1 into f_2.

As shown in Fig. 1, Section 2 stabilize Σ_1 along γ_1. Call the surface thus obtained $S_{\gamma_1}(\Sigma_1) = \Sigma_{1,1}$. Clearly, γ_1 is a core of a handle of a compression body of M cut open along $\Sigma_{1,1}$, so $\Sigma_{1,1}$ is a Heegaard surface for $M_{1,1} = M_1 - N(\gamma_1)$.

By induction, having obtained $(M_{1,j-1}, \Sigma_{1,j-1})$ with $\gamma_j \| \Sigma_{1,j-1}$, we stabilize $\Sigma_{1,j-1}$ along γ_j to obtain $\Sigma_{1,j}$. We then drill out γ_j. Finally we obtain $(X, \Sigma) \cong (M_1 - \cup_{j=1}^c \gamma_j,\ S_{L_1}^c(\Sigma_1))$.

To justify the inductive step we need to show that γ_j can be isotoped to lie on $\Sigma_{1,j-1}$. To see that, we perform as isotopy described in Fig. 4. After performing this isotopy $j-1$ times, once per stabilization, the surface $\Sigma_{1,j-1}$ will look (from above) exactly like Σ_1, with $2(j-1)$ punctures in it. Having γ_j (which "hovers" over $\Sigma_{1,j-1}$) miss these punctures is now a generic condition.

So we have ascended in the tree of the Dehn filling space from M_1 to X. Next we need to descend down to M_2. Note that the filling needed for that is $1/n_j$ (the numbers n_j given by the description of $f_2 f_1^{-1}$ as a composition of Dehn twists above) on each component of L. After each filling a destabilization[6] is performed by cutting the stabilized surface along D_j a meridional disk of the

[6] Destabilization is the inverse operation to stabilization. It can be performed whenever we find two disks meeting at one point. We need to specify the disk we cut open along, see Fig. 5.

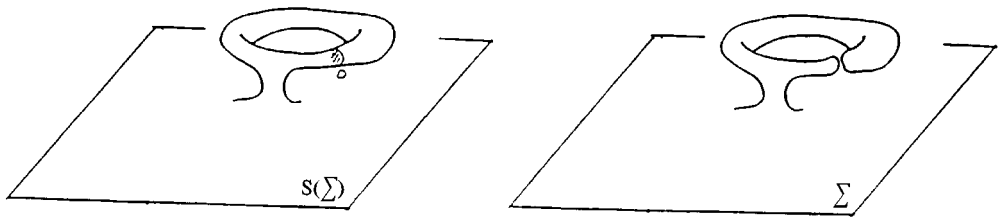


Fig. 5. Destabilization along D_j.

attached solid torus (see Fig. 5). Thus we end up with a procedure leading us to (M_2, Σ_2) which is the exact opposite of the one used when ascending the tree. □

Remark. We make the following remark to emphasize the analogy between geodesics in hyperbolic manifolds and good links. In the realm of complete hyperbolic manifolds of finite volume, define a Dehn drilling (resp. filling) to be *hyperbolic* if the link drilled (resp. filled) consist entirely of geodesics (all manifold involved being hyperbolic). A hyperbolic Dehn surgery is a hyperbolic drilling followed by hyperbolic filling. By analogy to the work done in this section, we can define hyperbolic manifolds to be *hyperbolic cousins* if they are related by a hyperbolic drilling followed by a hyperbolic filling. The set of all hyperbolic manifolds of volume bounded by V_0 forms a *single* extended family. In this family every pair of manifolds are hyperbolic cousins. Of course, any finite collection of hyperbolic manifolds has a bound on their volume and hence are all cousins. See Jørgensen's Theorem, Theorem 5.11.2 in [13].

4. Main corollaries of Theorem 6.2.

Fix (X,T) an a-cylindrical manifold pair with T a collection of c incompressible tori in its boundary. For each $M \in D(X,T)$ pick Heegaard surfaces Σ. Everything we do from now on is done with respect to these surfaces.[7]

The objects of our desire are "*good*" fillings. We define that notion from two points of view, one in (M,Σ) and the other in X. The reader will readily see the connection to good links (see Definition 3.1).

Definition 4.1. Let $T \subset \partial X$ be a collection of tori and (M,Σ) be a filling of (X,T). (M,Σ) is called *good* (or M is *good with respect to* Σ) if the cores of the attached solid tori are isotopic into Σ (although not necessarily simultaneously).

[7] Results in this section about finiteness of slopes yielding "bad" Dehn fillings require discussing Heegaard surfaces of bounded genus. We may pick surfaces for all the manifolds in $D(X)$ to be of bounded genus: the Heegaard genus of all these manifolds is bounded above by $g(X)$. However, we never restrict the number of surfaces picked for any manifold and do not even assume finiteness.

Remark.

- Equivalent to Definition 4.1 is the following condition: after pushing the cores off Σ and drilling them out, for each T_j (a component of T) we can find two parallel curves, one on T_j and the other on Σ (i.e. the two cobound an annulus in X), with the meridian intersecting the curve on T_j once exactly. The annulus may be punctured by other components of T, and the punctures have meridional slope.
- A filling (M,Σ) which is not good is called bad.
- The assumptions that X is a-cylindrical and T incompressible are used only in the combinatorial work (Section 5), and not in the Cerf-Theoretic work (Section 6). The precise assumption we need is that there exists no essential Möbius band embedded in (X,T) with its boundary essential in T. In Section 5 we prove Lemma 5.1, where we show that if X is a-cylindrical and T incompressible then T supports no essential Möbius bands.

Given two bad fillings (M_i, Σ_i) $(i = 1, 2)$ Theorem 6.2 demonstrates that the two Heegaard surfaces Σ_i can be isotoped so that the punctured surfaces $\Sigma_i^* = \Sigma_i \cap X$ intersect essentially while meeting T.

Theorem 5.2 tells us that when $|T| = 1$ the distance between curves bounding (non-closed) surfaces that intersect essentially is bounded in terms the genera of those surface by a quadratic polynomial $f(g_1, g_2)$, where g_i is the genus of Σ_i. Hence the distance between two *bad* fillings (M_1, Σ_1) and (M_2, Σ_2) is bounded by $f(g_1, g_2)$. f is given explicitly in Eq. (5), Section 5. Hence there are only finitely many slopes yielding bad surgeries of bounded genus.

We get:

Corollary 4.2. *Let (X,T) be an a-cylindrical 3-manifold, $T \subset \partial X$ an incompressible torus. Then the distance between bad fillings of genera g_1 and g_2 is bounded by $f(g_1, g_2)$. In particular, there are only finitely many bad surgeries of bounded genus.*

It may, of course, happen that a good Σ is a Heegaard surface for X, but that need not be the case: consider a knot lying on a Heegaard torus in a Lens space (or even in $\mathbb{S}^3$). However, we have seen before that after stabilizing Σ along γ we obtain $S_\gamma(\Sigma)$ for which γ is a core of a handle (recall Definition 2.1). This means $S_\gamma(\Sigma)$ can be pushed back into $X = M - N(\gamma)$: after drilling γ, $S_\gamma(\Sigma)$ is still a Heegaard surface!

Thus we get:

Corollary 4.3. *Let (M, Σ) be a good filling of a manifold X. Then $S_\gamma(\Sigma)$ is a Heegaard surface for X.*

By picking a minimal genus Heegaard splittings for each manifold in $D(X)$, and combining Corollaries 4.2 and 4.3 we get (recall that the genus of the filled manifold never exceeds the genus of the original manifold):

Corollary 4.4. *Let X be an a-cylindrical manifold and $T \subset \partial X$ an incompressible torus. Then for all but finitely many manifolds $M \in D(X)$ the following holds*:

$$g(X) - 1 \leqslant g(M) \leqslant g(X). \tag{1}$$

Moreover, the distance between fillings for which the genus drop by two or more is bounded by $f(g(X)-2, g(X)-2)$.

For bad fillings the genus may, of course, drop down to zero: we may be getting $\mathbb{S}^3$.

Let (M, Σ) be a good filling with Σ not a Heegaard surface for X. As discussed in Section 2, Σ will be a Heegaard surface for all manifolds obtained by $1/n$ surgery on γ (which is isotopic into Σ by assumption). Thus fillings containing such Heegaard surfaces are arranged on lines. Since these Heegaard surfaces are the result of destabilizing a Heegaard surface for X these lines are sometimes call *destabilization lines*.

To illustrate our point, we pick the minimal genus Heegaard splitting for the manifolds in $D(X)$ and get:

Corollary 4.5. *Let* (X,T) *be an a-cylindrical manifold pair with* $T \subset \partial X$ *an incompressible torus.*

If the Heegaard genus drops for two fillings of distance $f(g(X)-1, g(X)-1)$ *or more, it will drop for all the manifolds on at least one line.*

Example 4.6. *Let X be the trefoil knot exterior.*[8] It is of genus 2. Consider the manifolds in $D(X)$ with their minimal genus Heegaard splittings. Most are of genus 2 as well. On one line the manifolds are of genus 1 or less (lens spaces) and are all good but one. That manifold – S^3 – is of genus zero and corresponds to a bad filling. So we get a finite set (with one element) of bad fillings. If you view S^3 as a lens space with the genus one splitting (or with any other positive genus splitting) the filling will be good. Doubling the Seifert surface for the trefoil gives the standard genus-2 surface in S^3 with surface framing given by the longitude, i.e. the standard framing. All the $1/n$ surgeries with respect to that framing will have an extra Heegaard surface of genus 2. (These surfaces are usually called *horizontal* Heegaard surfaces; we make no claim as to whether or not the horizontal Heegaard surfaces are isotopic to the "regular" ones or genuinely new.)

Our final remarks for this section is that all we have done can (virtually) be done when filling more than one boundary component. The construction (below) is done by filling all boundary components but one, and then considering the last filling as a filling of a single torus. If $T = \{T_j\}_{j=1}^c$, we view the Dehn Fillings Space $D(X,T)$ as a tree (see Fig. 6), with X as its root. Oriented edges originating from the root correspond to the different slopes on T_1. These edges terminate on vertices that form $D(X,T_1)$. Edges originating on $D(X,T_1)$ correspond to slopes on T_2, where we see each slope appearing once exactly from each vertex. These edges terminate on the set $D(X,T_1 \cup T_2)$, and so on until we reach the leafs that correspond to $D(X,T) = D(X, \bigcup_{j=1}^c T_j)$.

With each drilling we will need to exclude a finite set of fillings (hence our definition of a virtually finite set, see Section 2). As before, we use minimal genus Heegaard surfaces to demonstrate our techniques, which are in fact valid for any collection of Heegaard surfaces for $D(X,T)$ with bounded genus.

[8] The trefoil knot exterior is cylindrical, however the conclusions of all the theorems hold.

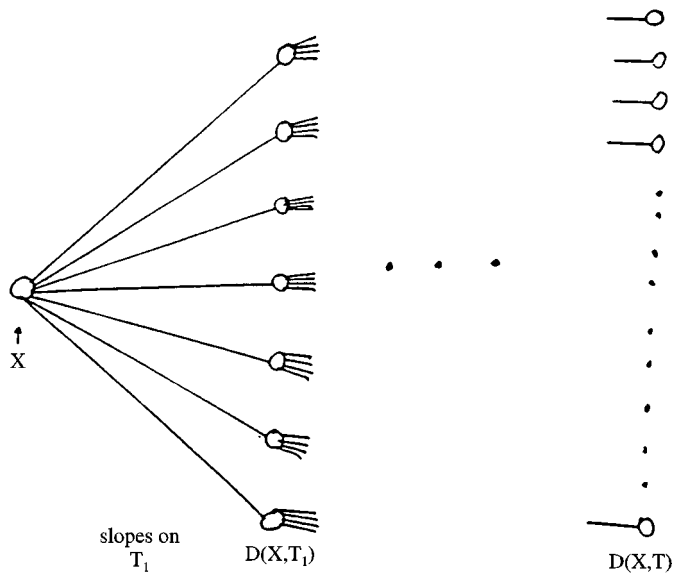


Fig. 6. Tree structure for $D(X,T)$.

Corollary 4.7. *Let (X,T) be a-cylindrical manifold pair with T a collection of c incompressible tori in ∂X. Then for all but virtually finitely many manifolds in $D(X,T)$ we get*

$$g(X) - c \leqslant g(M) \leqslant g(X). \tag{2}$$

Proof. Order the components of T as $\{T_j\}_{j=1}^{c}$. Pick a minimal genus Heegaard surface Σ for each manifold in $D(X,T)$. Then by our discussion above γ_c, the core of the solid torus attached to T_c, is isotopic into the chosen Heegaard surface in each manifold in $D(X,\bigcup_{j=1}^{c-1} T_j)$ — unless it belongs to a bounded (and hence finite) set of bad fillings. Exclude that set. For the good fillings, stabilize Σ along γ_c to obtain $S_{\gamma_c}(\Sigma)$ which is a Heegaard surface for $M - N(\gamma_c) \in D(X,\bigcup_{j=1}^{c-1} T_j)$. Now drill out γ_{c-1} and proceed in reverse order, the structure of the set we exclude is that of a virtually finite set.[9] □

As there are good fillings for which the genus degenerates, (see, e.g., Example 4.6) the inequality in Eq. (1) is the best possible. Similarly, by taking a manifold (M,Σ) of genus g and drilling out c curves parallel to Σ but not isotopic to a core in the complementary compression bodies one can

[9] The set of bad slopes we obtain when filling T_k $(k \neq 1)$ is finite for each manifold in $D(X,\bigcup_{j=1}^{k-1} T_j)$, but we cannot conclude that union of these sets is finite. At the last step, when filling T_1, even though we exclude a finite set of slopes from T_1, not all fillings of other slopes are good all the way down to every manifold that succeeds them in $D(X)$: we may start with a good filling of T_1 and continue to a bad filling of another component. Hence, unlike Thurston's Hyperbolic Dehn Surgery Theorem, we cannot conclude that avoiding a finite set of each cusp will guarantee that the filling is good. This conclusion was obtained by Moriah and Rubinstein in [9] using analytic techniques, and it will be interesting to find a combinatorial technique that will retrieve that result.

easily construct an examples of good fillings where the genus drops by c for any c, thus showing that the inequality 2 is the best possible as well.

5. The combinatorial bound

We start by proving the following lemma:

Lemma 5.1. *Let (X,T) be an a-cylindrical manifold pair with $T \subset \partial X$ an incompressible torus. Then there does not exists a Möbius band embedded in X with its boundary essential in T.*

Proof. Assume F is an essential Möbius band in (X,T), we will show that this implies that T compresses, thus contradicting our hypothesis.

A neighborhood of F in X is a solid torus N (twisted **I** bundle over the Möbius band), which meets T in an annulus. Let $\tilde{F}$ be the complementary annulus $cl(\partial N - T)$. The boundary of this annulus defines the same slope as ∂F, and hence is essential in T. As (X,T) is a-cylindrical by assumption, $\tilde{F}$ must compress or boundary compresses. IF $\tilde{F}$ copresses so does T, contradicting our assumption. Hence $\tilde{F}$ boundary compreses. Let D be the disk defining this boundary compression. Since the boundary of $\tilde{F}$ is homotopic to a square of the generator of the solid torus N, we may assume $int(D) \cap N = \emptyset$. Let $\bar{N} = N(D) \cup N(F)$ a neighborhood of $D \cup F$. $\bar{N}$ is still a solid torus, now meeting the boundary in a punctured torus. Hence a collar neighborhood of the boundary union $\bar{N}$ is a punctured solid torus, and T compresses, contradiction. □

The rest of this section is devoted to proving the following theorem:

Theorem 5.2. *Let (X,T) be an a-cylindrical 3-manifold pair, $T \subset \partial X$ an incompressible torus. Let $(\Sigma_i^*, \partial\Sigma_i^*)$, $i = 1,2$, be two surfaces of genus g_i (w.l.o.g. $g_1 \geqslant g_2$), both with non-empty boundary, properly embedded in (X,T). Assume that each boundary component of Σ_i^* defines a non-trivial slope λ_i on T.*

If Σ_1^* intersects Σ_2^* essentially then the intersection number of λ_1 and λ_2 satisfies:

$$\Delta(\lambda_1, \lambda_2) < f(g_1, g_2), \tag{3}$$

where

$$f(g_1, g_2) = f_1(g_1, g_2) = 36g_1g_2 + 36g_1 + 18g_2 + 18. \tag{4}$$

If, in addition, we know that $|\partial\Sigma_i^*| \geqslant 2$ for both surfaces, $f(g_1, g_2)$ can be taken to be

$$f(g_1, g_2) = f_2(g_1, g_2) = 18g_1g_2 + 18g_1 + 18g_2 + 18. \tag{5}$$

Remark.

- For our main purposes we do know that $|\partial\Sigma_i^*| \geqslant 2$ for both surfaces as Heegaard surfaces separate, so the stronger, symmetric bound given by Eq. (5) holds.

- Recall (Definition 2.2) that surfaces are said to intersect essentially if each component of $\partial\Sigma_1^*$ meets each component of $\partial\Sigma_2^*$ minimally, and their intersection contains no arc that is boundary parallel in either of them. Recall also that an arc properly embedded in Σ_i^* is called boundary parallel in Σ_i^* if it cobounds a disk with $\partial\Sigma_i^*$. We do *not* require that this disk misses Σ_{3-i}^* in its interior.
- In [14], Torisu, using a lemma of Gordon and Litherland (see [5]), proved a similar bound for boundary slopes of embedded incompressible, boundary incompressible surfaces.

Proof. Recall our assumption that all surfaces are orientable.

Endow T with an (arbitrary) flat metric and pull all components of $\partial\Sigma_1^*$ and $\partial\Sigma_2^*$ tight to be straight lines. More precisely, each boundary component lifts to a straight line in the universal cover $\mathbb{E}^2$. This can be done without changing the topology of the surfaces and their intersection. We will therefore assume the intersection remains essential.

Pick a maximal set of non-parallel, non-boundary parallel arcs on each surface. Denote the number of arcs in each family by e_i and the number of boundary components of each surface by p_i. Our first goal is to bound e_i. Note that Σ_i^* is not a Möbius band by assumption of orientability, nor a disk by assumption that T is incompressible. If Σ_i^* is an annulus $e_i = 1$. This easily yields a much stronger bound than the case $\chi(\Sigma_i^*) < 0$. We will not show that here, assuming from now on that $\chi(\Sigma_i^*) < 0$.

After pinching $\partial\Sigma_i^*$, so the punctures become vertices in the closed surface that we call Σ_i, we obtain the following from an Euler characteristic count (with f_i denoting the number of faces):

1. $2e_i = 3f_i$ (maximality of e_i, and $\chi(\Sigma_i^*) < 0$
2. $p_i - e_i + f_i = \chi(\Sigma_i)$.

Combining the two and solving for e_i we obtain that the *maximal* number of non-parallel non-boundary parallel arcs is

$$e_i = 6g_i + 3p_i - 6.$$

Thus every collection of $k(6g_i + 3p_i - 6) + 1$ arcs on Σ_i^* contain a family of $k + 1$ parallel arcs.

Assume then that the total number of arcs of intersection is at least

$$(6g_1 + 3p_1 - 6)(6g_2 + 3p_2 - 6) + 1.$$

Then on Σ_1^* we have a set of $(6g_2 + 3p_2 - 6) + 1$ parallel arcs. Viewing these as arcs on Σ_2^*, we get that at least two of these are parallel there. Name those arcs a_1 and a_2. The arcs a_1 and a_2 cobound disks D_i on Σ_i^*. Pasting these two discs together we obtain a surface F which is either a Möbius band or an annulus. By passing to an innermost annulus/Möbius band, and deleting simple closed curves via disk swaps, we can obtain an annulus/Möbuis band F that is embedded. (Since the arcs a_1 and a_2 cut F to disks, no essential simple closed curve of self intersection can exist in F.)

We shall need the following "parity rule" (see [5]), which is an orientation argument. Choose an orientation for the λ_i's and orient the surfaces Σ_i^*. Each boundary component of each Σ_i^* inherits an orientation from Σ_i^*; we call it "clockwise" if it agrees with the orientation of λ_i, counter-clockwise otherwise. The parity rule states that the arcs a_1 and a_2 will connect two boundary components

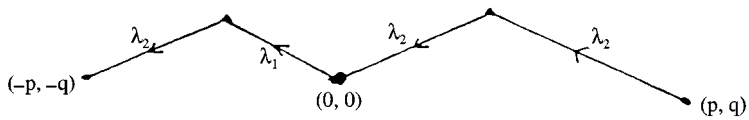


Fig. 7. Lifts of ∂F (F an annulus).

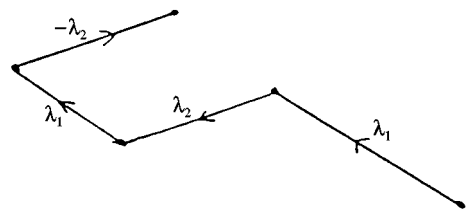


Fig. 8. ∂F is essential (F an Möbius strip.)

with the same orientation – say clockwise – on one surface and two with opposite orientation on the other.

We first rule out the case where F is an annulus. This is characterized by $|\partial F| = 2$. Note that each boundary component consists of two arcs, one in direction $\pm\lambda_1$ and the other in direction $\pm\lambda_2$. Since the two boundary components are parallel their liftings to $\mathbb{E}^2$ define the same slope (see Fig. 7). Hence if the lifting of the first (after shifting it to start at (0,0)) ends at some point (p,q) and the second (after shifting it to start at (0,0) as well) will terminate at either (p,q) or $(-p,-q)$. Each boundary component corresponds to a linear combination of the vectors λ_1 and λ_2, the first yielding (p,q) and the second $\pm(p,q)$, say $(-p,-q)$. As λ_1 and λ_2 are linearly independent (equivalent to the slopes being different), the coefficient of λ_1 in the first combination has opposite sign to that in the second. So a_1 connects boundary components of opposite signs. A similar argument for the coefficients of λ_2 shows that a_2 connects components of opposite signs as well, contradicting the parity rule. In case the lifting of the second boundary component terminates at (q,p) a similar argument demonstrates that a_1 connects boundary components of the same sign on both surfaces, again contradicting the parity rule.

Next, we show that if F is a Möbius band then its boundary is essential in T (hence in M). If F were a Möbius band its boundary would have lifted to four segments. The first and the third in direction $\pm\lambda_1$, the second and fourth in $\pm\lambda_2$. By the parity rule one pair-say the first and third segments – will have the same sign, while the other pair opposite signs. The reader will easily verify that such curve cannot be closed. See Fig. 8.

Lemma 5.1 excludes such Möbius bands.

We conclude that on Σ_1^* we cannot have $(6g_1 + 3p_1 - 6)(6g_2 + 3p_2 - 6) + 1$ arcs. As the total number of arcs is $\dfrac{|\Delta(\lambda_1,\lambda_2)|p_1p_2}{2}$, we get

$$\frac{|\Delta(\lambda_1,\lambda_2)|p_1p_2}{2} < (6g_1 + 3p_1 - 6)(6g_2 + 3p_2 - 6) + 1.$$

After distributing the right-hand side, we may use the term $-18(p_1+p_2)$ to cancel $+37$ (since $(p_1+p_2)\geqslant 3$ for homological reasons) and remove negative terms to get:

$$|\Delta(\lambda_1,\lambda_2)|<\frac{72g_1g_2}{p_1p_2}+\frac{36(p_2-2)}{p_1p_2}g_1+\frac{36(p_1-2)}{p_1p_2}g_2+18.$$

As $p_i\geqslant 2$ for some i (say $i=2$, however to side with caution we must assume $g_1\geqslant g_2$) so we plug $p_1p_2\geqslant 2$ and get (we also round the number down to an integer):

$$|\Delta(\lambda_1,\lambda_2)|\leqslant 36g_1g_2+36g_1+18g_2+18. \tag{6}$$

If, in addition, we know that each surface has at least two boundary components we can use $p_1p_2\geqslant 4$ and get the improved and symmetric bound:

$$|\Delta(\lambda_1,\lambda_2)|\leqslant 18g_1g_2+18g_1+18g_2+18. \qquad \square \tag{7}$$

6. The Cerf theory

Recall that all manifolds and surfaces are assumed to be orientable.

We can apply Theorem 5.2 in various situation. One is $-f(g_1,g_2)$ bounds the distance between slopes supporting (non-closed) incompressible, boundary incompressible surfaces of genera g_1 and g_2 (similar to [14]). Another application that follows immediately from Gabai's thin position argument (see [4]) is bounding the distance between boundary slopes of essential surfaces and slopes of bad surgeries.

Theorem 6.1. *Let (X,T) be an a-cylindrical 3-manifold, $T\subset\partial X$ an incompressible torus. Let λ_1 be a boundary slope of an incompressible, boundary incompressible surface (with non-empty boundary) of genus g_1. Let λ_2 be a slope yielding a manifold $(X(\lambda_2),\Sigma)$ with a Heegaard surface of genus g_2 so that the filling is bad.*

Then $|\Delta(\lambda_1\cdot\lambda_2)|<f_1(g_1,g_2)$.

This was proven in [10].

The applications we are after were given in Section 4. They all follow from the following theorem (for them we need $c=1$ only). This theorem concerns the surfaces Σ_i^*, which are the part of Σ_i which exists in X. These surfaces depend both on Σ_i and on the way it is embedded in M_i and are *not* invariant under isotopy of Σ_i in M_i. We shall be using techniques of *Cerf Theory* (see [3,6]).

Fix a boundary component $T_{j_0}\in T$. For Theorem 6.2 we need to assume that the fillings are T_{j_0}-bad, which simply means the core of the solid torus attached to the component T_{j_0} is not isotopic into the pre-chosen Heegaard surfaces. The only case that we use of this Theorem is when $c=1$. In that case, specifying a particular boundary component is, of course, meaningless.

Theorem 6.2. *Let X be a 3-manifold, $T=\bigcup_{j=1}^{c}T_j\subset\partial X$ a collection of c tori. Let $\lambda_{i,j}$ ($i=1,2$, $j=1,\ldots,c$) be slopes on T_j and $\lambda_i=\bigcup_{j=1}^{c}\lambda_{i,j}$. Let $M_i=X(\lambda_i)$ be fillings of (X,T), both T_{j_0}-bad with respect to Σ_i, $i=1,2$.*

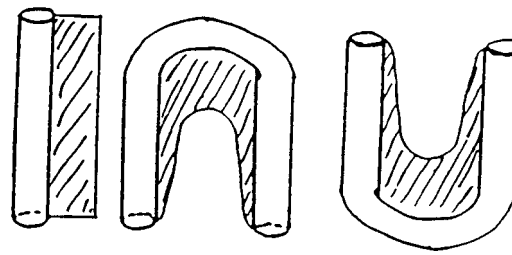


Fig. 9. Surface in the neighborhood of the boundary torus.

Then the surfaces Σ_i can be isotoped so that $\Sigma_i^* \subset X$ intersect essentially while meeting the boundary component T_{j_0}.

Proof. By our assumption, γ_{i,j_0} are both bad w.r.t. Σ_i.

We briefly define the notion of *thin position.* For more detail see [4]. Given γ a set of curves embedded in a 3-manifold M and a height function h on M, assume $h|\gamma$ (h restricted to γ) is Morse. Define the *width* of the embedding as follows: between every two consecutive critical points of $h|\gamma$ choose a regular level and count the number of times that level intersects γ (this is clearly independent of choice). The sum of these numbers is called the width.

Any embedding that minimizes the width is called *thin position.*[10]

Let Γ_i be a spine for (M_i, Σ_i). Recall that $M_i - \Gamma_i$ is foliated as $\Sigma_i \times \mathbb{R}$.

Let $\gamma_i = \bigcup_{j=1}^{c} \gamma_{i,j}$ be the set of cores of the attached solid tori in M_i. Isotope the foliations $\Sigma_i \times \mathbb{R}$ (fixing γ_i) so that both γ_i's are in thin position. Some components $\gamma_{1,j}$ may have width zero – if they were isotoped into a level – but not γ_{i,j_0}. Any such component will be ignored from now on.

A *fat layer* in M_i is a stack of the form $\Sigma_i \times [a,b]$, where $h_i|\gamma_i$ has a minimum at a and a maximum at b, and no other critical value a and b. Choose a fat layer meeting γ_{i,j_0} in each of the two manifolds. Parameterize the fat layers as $\Sigma_i \times \mathbf{I}$, $\mathbf{I} = [-1,1]$. Our levels are now parameterized by two parameters, say $s \in \mathbf{I}$ and $t \in \mathbf{I}$. The first picks the level $\Sigma_1 \times \{s\} = \Sigma_1(s)$ and the second $\Sigma_2(t)$.

View the two fat layers in X. Let h_i be restriction of the height function in M_i to X. The fat layer from M_i is bounded above by a maximum and below by a minimum of h_i. Near that maximum (minimum resp.) bend the fat layer from M_{3-i} downwards (upwards resp.), and away from these points "straighten" that fat layer, so locally the surfaces near the boundary are modeled on the three types seen in Fig. 9.

Perturb the foliations slightly so that h_2 is Morse on $\Sigma_1^*(s)$ for all but finitely many values of s. On the square $\mathbf{I} \times \mathbf{I}$ mark all points corresponding to surfaces $\Sigma_1^*(s)$ and $\Sigma_2^*(t)$ that do *not* intersect transversely. This set is called the Graphic. Equivalently, view the Graphic as the set: $\{(s,t) \in \mathbf{I} \times \mathbf{I} | \Sigma_1^*(s)$ has a critical point at height t (with respect to $h_2)\}$. So long as the parameter (s,t) varies within a region[11] (away from the Graphic), the height functions h_1 and h_2 are both Morse without critical points so the topology of the two surfaces and their complements will not change.

[10] To achieve thin position one needs to isotope γ or, equivalently, change the height function. We will take the second point of view, and will see the height function as defined by the level surfaces. Thus we will isotope the surfaces in M_i instead of γ. The advantage of this semantic change is that is allows us to view our work entirely in X.

[11] By a "region" we mean a connected component of the complement of the Graphic.

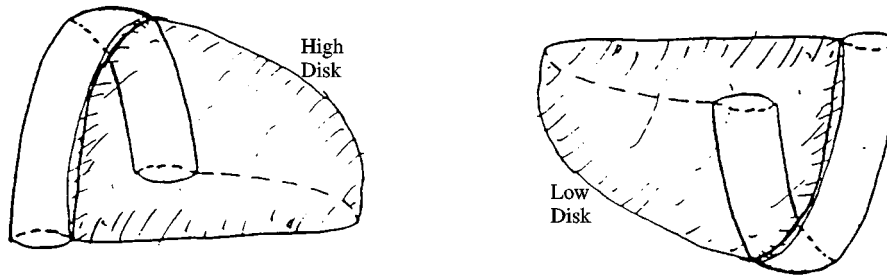


Fig. 10. High and low disks.

The concept required for connecting the assumption (thin position) with the desired outcome (essential intersection) is that of a high (resp. low) disk: pick two level surfaces by picking $(s_0,t_0)\in \mathbf{I}\times\mathbf{I}$. A disk $D\subset \Sigma_2^*(t_0)$ is called a *high* (resp. low) disk for $\Sigma_1^*(s_0)$ if the boundary of D consists of two arcs (meeting at their endpoints), one along $\Sigma_1^*(s_0)$ and the other along ${}_{j_0}T$ entirely *above* (resp. below) the level $\Sigma_1^*(s_0)$. We define a high (low) disk for $\Sigma_2^*(t_0)$ in a similar manner. See Fig. 10.

A boundary parallel arc bounds a disk that shall call *one sided.* An outermost one sided disk is either high or low. Therefore $\Sigma_1^*(s_0)$ and $\Sigma_2^*(t_0)$ intersect essentially if and only if there do not exist high and low disks for either surface, and the entire proof now boils down to showing existence of $(s_0,t_0)\in\mathbf{I}$ for which neither high nor low disk exists.[12]

We cannot have a high and a low disk at the same level (and for the same surface) else the width can be reduced (see Gabai's original argument in [4]). Hence the following labeling system is well defined: label each region by two letters, each being L,H or N, in the following way:

- L (H) stands for low (high) disk;
- N stands for neither a high nor a low disk exists;
- the first letter refers to a high/low disk for Σ_1^* from Σ_2^*, the second – for Σ_2^* from Σ_1^*;

In case a label is unknown or is not relevant, the label X will be used, so labeling a region LX, for instance, means there exists a low disk for Σ_1^* from Σ_2^*, but we do not know (or do not care about) the label for Σ_2^*.

Thus our goal is to prove the existence of a region marked NN.

Since we bent the surfaces near the boundary (recall Fig. 9), we know much about the labels around the perimeter. This information is given in Fig. 11.

In order to see how the labels change as we pass from one region to another, we will study what happens as we cross the Graphic. Each singular curve in the Graphic is either a saddle or a minimum/maximum, henceforth a *center.* Passing through a center does not change the existence or non existence of high/low disks. However, passing across a saddle could change the labels, as we may be attaching a 1-handle to the high/low disk.

[12] In its interior the high (resp. low) disk may get below (resp. above) the given level, and so meet the other surface. Thus eliminating high and low disks corresponds to eliminating boundary parallel arcs *without* requiring that the one sided disks have no simple closed curves of intersection in their interior. It is exactly these arcs we must eliminate in order to use Theorem 5.2.

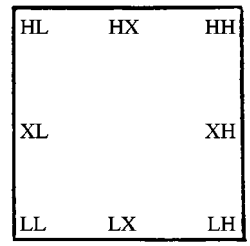
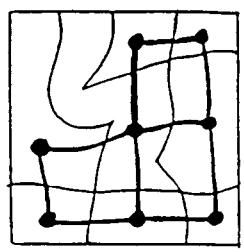


Fig. 11. The labels along the boundary.



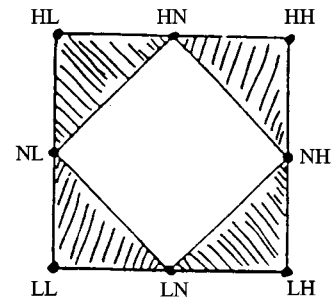


Fig. 12. The dual to the Graphic and the Range Space.

From Gabai's argument it follows that adjacent regions cannot be labeled LX and HX (resp. XL and XH) by thin position of γ_1 (resp. γ_2). Hence a region labeled LX, for instance, may change into NX but *never* to HX. However, we may have LN change into NH since the first L represents a low disk for Σ_1^*, and the second H represents a high disk for Σ_2^*.

We simplify the Graphic slightly by removing any component not connected to the boundary. (We do that in order to ensure that the dual hypergraph will embedded naturally in the disk.) The label given to the new region is the label we had near its boundary before. A *vertex* in the Graphic will be a point where two edges cross; all vertices have valence 4.

At this point (following the presentation of Rubinstein and Scharlemann in [12]), assume no region is labeled NN. We construct a map ϕ from the dual hypergraph to the Graphic to the cell complex "the Range", given in Fig. 12. To be specific, the dual hypergraph is a two-dimensional cell complex consisting of a vertex for each region in the Graphic, an edge for each edge and a face for each vertex. Label each vertex of the dual hypergraph by the label of the corresponding region in the Graphic. ϕ maps each vertex to the point in the Range with the same label.

It is clear that ϕ extends over all edges: the Range is path-connected. For each edge of the dual, we choose an extension that hits no vertices of the Range in its interior. However, it is not at all clear that this map extends over all faces.

Claim 6.3. *The map ϕ extends over all faces, and hence is defined globally.*

In order to prove this claim we only need to check how the labels change as we go around a vertex of the Graphic. Let us try and construct the only possible vertex around which the map

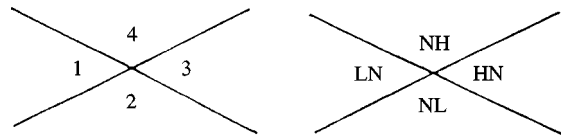


Fig. 13. The labels around (s_0,t_0).

Fig. 14. (a) "Almost disjoint" disks; (b) not "almost disjoint" disks.

ϕ does not extend. This vertex – say (s_0,t_0) – is the crossing of two singular points, both saddles. (We are motivating Fig. 13.) Number the regions of the parameter square adjacent to (s_0,t_0) $1,2,3,4$. Since the label NN does not exist by assumption, assume w.l.o.g. that the label in region 1 is LX. We can only have a problem continuing the map if its image is not simply connected, so the label HX must appear. Since crossing one curve of the Graphic cannot change LX to HX (by thin position) HX must be the label in region 3 while regions 2 and 4 are labeled NX.

Using again the fact that NN does not appear, region 2 is labeled NL or NH, say NL. If XH never appears, the regions 1,2,3,4 are mapped into a simply connected region, hence the map extends. So XH appears, and the only available region for it is 4. Thus region 4 is labeled NH. This information is summarized in Fig. 13.

We now describe the set up for the rest of our work. There are four disks involved, one in each of the regions 1–4: a high disk and a low disk for Σ_1^* and a high disk and a low disk for Σ_2^*. In this order, name them H_1 (in region 3), L_1 (in region 1), H_2 (in region 4) and L_2 (in region 2). The vertex in Fig. 13 corresponds to two saddles. Name them S_1, S_2. In Fig. 13 we see two ways to move out of each of the regions 1–4 by crossing an edge of the Graphic. Moving out of any region destroys the high/low disk that exists there, hence each of the two saddles is a point on the boundary of each of the four disks H_1, L_1, H_2, L_2.[13]

Let α_1 be the arc component of $H_1 \cap \Sigma_1^*(s_0)$, and β_1 be the arc component of $L_1 \cap \Sigma_1^*(s_0)$. Similarly define the arcs α_2 and β_2. As we move (s,t) across an edge of the Graphic the surfaces will cross one saddle, and as (s,t) crosses (s_0,t_0) the surfaces cross both saddles. When viewing the intersection pattern on the surfaces, crossing a saddle corresponds to replacing two little horizontal arcs by two little vertical arcs or vice versa.

Two disks are called "almost disjoint" if their boundaries meet in a way that allows common boundary arcs to be pushed away from each other by a small perturbation, e.g. Fig. 14(a). In many

[13] We refer to disks H_i and L_i (and in the next paragraph arcs α_i and β_i in their boundary) but in fact we are discussing families of disks and arcs. As we move within a region – say labeled HN – every point (s,t) in that region corresponds to surfaces $\Sigma_1^*(s)$ and $\Sigma_2^*(t)$ where a high disk for $\Sigma_1^*(s)$ exists. We will not distinguish between disks and arcs corresponding to different values for (s,t) within a region as the surfaces are diffeomorphic. When discussing the way different arcs meet, or saying "$S_j \in \alpha_i$", we are referring to the limit disks and arcs at the point (s_0,t_0). The limit disks are embedded in their interior but their boundary have exactly two double points.

cases our strategy is to show that either H_1 and L_1 are "almost disjoint" or H_2 and L_2 are. In the first case the width of γ_1 can be reduced, in the second that of γ_2: given two almost disjoint disks – say H_1 and L_1 – after perturbing one of them slightly (say H_1) near its boundary, we get two disjoint disks that can be used to reduce the width of γ_2 (see [4]), contradicting our assumption of thin position.

Remark. In the sequel we will provide figures of intersection between Σ_1^* and Σ_2^*. The following rules will be used: the curves studied are pictured in solid bold lines; dashed lines are used for parts of these curves where the only relevant information is which 2 arcs are connected by the dashed lines, not where the dashed line is on the surface. Dotted lines, when appear (in Fig. 18), indicate other curves of intersection we know of, and are given as a reference point only. Small circles indicate boundary component; we use them to indicate we know an arc ends at the boundary, whenever we know it is not a part of a simple closed curve of intersection.

We will now check all possible configurations for α_i and β_i, and reach a contradiction in each of them. Consider the level $(s_0,t_0+\varepsilon)$, slightly above t_0, for very small ε. Since for $\varepsilon=0$ we have S_1, $S_2 \in \alpha_1$, as $\varepsilon \mapsto 0$ there are two arcs from $\Sigma_1^*(s) \cap \Sigma_2^*(t)$ approaching α_1.

These two arcs may or may not be subarcs from α_1 itself. If they are we will say that α_1 "folds on itself". Assume an arc folds on itself. Orient the arc. The following Lemma will enable us to use orientation of X conveniently:

Lemma 6.4. *If an arc η folds on itself, the segments meeting at the saddle meet with opposite orientation.*

Proof. (See Fig. 15 *for illustration of this lemma*). Let $\eta \subset \Sigma_1^* \cap \Sigma_2^*$ be an arc on Σ_1^*. Consider the following two vector fields:

- Let $v_0(t)$ be a tangent vector to $\eta(t)$. By assumption, $v_0(0)$ is parallel to $v_0(1)$.
- Let $v_2(t)$ be a vector field along η which, together with $v_0(t)$, constitutes a basis for Σ_2^*. Note that as v_2 is normal to the Heegaard surface Σ_1, we may assume that $v_2(0)$ is parallel to $v_2(1)$, as seen in Fig. 15 (a).

Just above or below Σ_1^* lies S (say above), the saddle. In the neighborhood of S Σ_1^* intersects Σ_1^* as $\{z=x^2+y^2\}$ intersects $\{z=-\varepsilon\}$ (for small ε) near the origin in $\mathbb{R}^3$. Thus we can add a little "arc" from Σ_2^* to obtain Fig. 15 (b). The reader can now see that Σ_2^* contains an embedded Möbius band, contradicting our assumption of orientability. □

The arc α_1 can take various forms. There are eight different cases which respect Lemma 6.4, arranged according to whether the two arcs approach α_1 from the same side, whether α_1 folds on itself or not and the order in which we encounter the saddles as we travel along α_1. Cases 1–5 are given in Fig. 16. This is the local picture on Σ_1^*, in the neighborhood of α_1. Cases 6–8 are when α_1 folds on itself twice. They differ in the order that one encounters the saddles S_1 and S_2 as one travels along α_1. (Say we meet S_1 first.) When considering them keep Lemma 6.4 in mind.

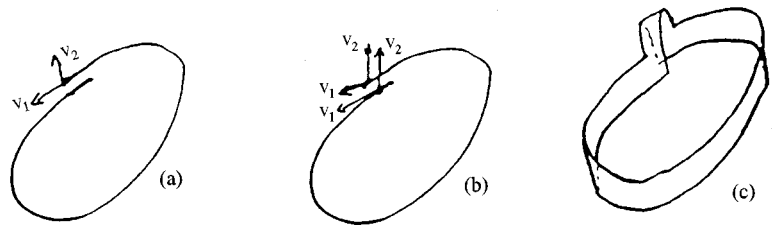


Fig. 15. Arc folding on itself with the orientation agreeing.

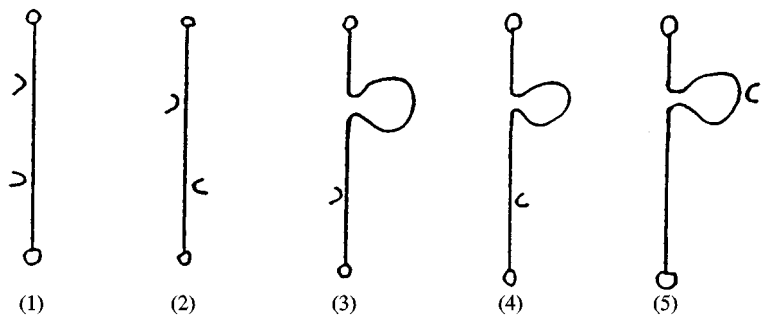


Fig. 16. α_1: the first five cases.

We demonstrate the different cases by order of reduction. (The crucial case seems to be case 2.) In each of these cases we will compare the four arcs (α_1, β_1, α_2, β_2). Recall that we get from one to another by crossing one or two of the saddles S_1, S_2. W.l.o.g., we call the top saddle in Fig. 16 S_1.

- *Case* 1: After crossing both saddles, we see that the arc β_1 is almost disjoint from α_1.
- *Case* 4: Crossing S_1 only: for α_2 or β_2 the picture is reduced to case 1.
- *Cases* 5: After crossing S_1 we see that there is no arc (candidate for α_2 or β_2) which involves both saddles, which we know there must be.
- *Case* 7: $S_1\ S_2\ S_2\ S_1$. Much like case 5, after crossing S_1 we see that either α_2 or β_2 does not involve S_2.
- *Case* 2: See below (reduced to cases 1 and 5).
- *Case* 3: Reduced to case 2 by crossing S_1 (again for α_2 or β_2).
- *Case* 6: $S_1\ S_1\ S_2\ S_2$. Cross both saddles: reduced to case 1 or 2 (for β_1).
- *Case* 8: $S_1\ S_2\ S_1\ S_2$. It is easy to see that the only two configurations possible are the two shown in Fig. 17 (recall Lemma 6.4). In subcase 8(i) crossing both saddles reduces to case 1 for α_2 or β_2, and in subcase 8(ii) crossing both saddles reduces to case 1 or 2 for β_1.

It now all boils down to case 2 in Fig. 16. Note that we may reduce this case to cases 1 and 5 as planned: those have been proven *without* using case 2.

After crossing S_1 or S_2 in case 2 (note the symmetry in this case) we get eight different possibilities β_2. These are summarized in Fig. 18. We now treat each of them. Note that subcase 2.vi forces us to

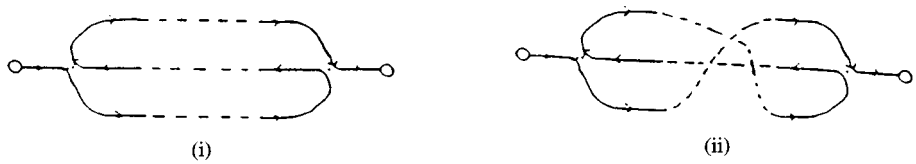


Fig. 17. The two possibilities for case 8.

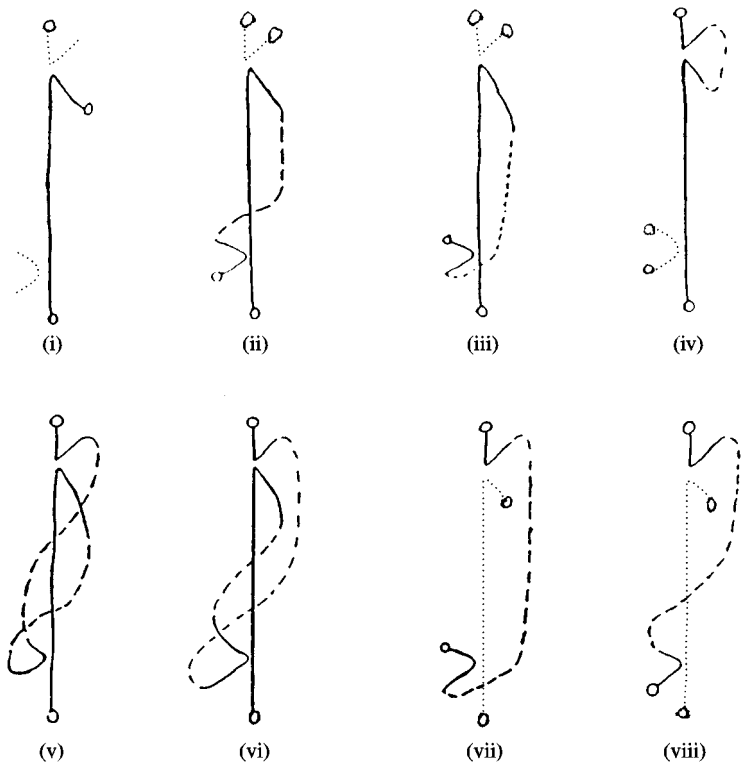


Fig. 18. β_2 in case 2.

view α_1 and β_1 again, but now we have more information about them, as we know how the arcs are connected.

- *Subcases* 2(i) and 2(vii) are case 1.
- *Subcase* 2(ii) is case 5.
- *Subcases* 2(iii) and 2(v) violate orientability (recall Lemma 6.4).
- *In subcase* 2(iv) the arc α_2 is as in case 1.
- *In subcase* 2(vi) α_1 and β_1 are almost disjoint.
- *In subcase* 2(viii) the arc α_2 is as in case 5.

This completes the proof of the claim. □

Claim 6.5. *ϕ cannot be defined globally.*

Proof. This follows from the fact that going around the boundary of the Graphic, counter-clockwise starting at the south-west corner, our image will first run along the bottom of the Range, then along the right side of it, the top, and back along the left side (see Figs. 11 and 12). Thus we have mapped a disk into the Range with its boundary mapping to a non null homotopic curve. Absurd. □

The observant reader will notice that Claims 6.3 and 6.5 imply a contradiction, hence we have proved the existence of a region labeled NN and correspondingly two surfaces which meet essentially, as desired. □

Acknowledgements

The author gratefully acknowledges the help of his dissertation advisor, Cameron Gordon, and thanks John Luecke and Eric Sedgwick for helpful conversations.

References

[1] J. Birman, Braids, Links and Mapping Class Groups, Princeton University Press, Princeton, 1974.
[2] A. Casson, S. Bleiler, Surface Automorphisms After Nielsen and Thurston, Cambridge University Press, Cambridge, 1988.
[3] J. Cerf, les diffeomorphismes de la sphere de dimension trois ($\Gamma(4) = 0$), Lecture Notes in Mathematics, vol. 53, Springer, Berlin, 1968.
[4] D. Gabai, Foliations and the topology of three-manifolds III, Journal of Differential Geometry 26 (1987) 179–536.
[5] C. McA. Gordon, R. Litherland, Incompressible Planar Surfaces in 3-Manifolds, Topology and its Applications 18 (1984) 121–144.
[6] C. McA. Gordon, J. Luecke, Knots are determined by their complements, Journal of American Mathematical Society 2 (1989) 371–415.
[7] P. Heegaard, Forstudier til en Topologiskteori for de Algebraiske Aladers Sammenhaeng, Ph.D. Thesis, 1899.
[8] W.B.R. Lickorish, A Representation of orientable combinatorial 3-manifolds, Annals of Mathematics 76 (1962) 531–540.
[9] Y. Moriah, H.J. Rubinstein, Heegaard structures of negatively curved manifolds, Communications in Geometrical Annals 5 (3) (1997) 375–412.
[10] Yo'av Rieck, Heegaard surfaces and Dehn fillings: $g(M) - 1 \leqslant t(X) \leqslant g(M)$, Ph.D. Thesis, University of Texas at Austin, 1997.
[11] D. Rolfsen, Knots and links, Publish or Perish, Berkeley, 1976.
[12] H. Rubinstein, M. Scharlemann, Compairing Heegaard splittings of non-Haken 3-manifolds, Topology 35 (1996) 1005–1026.
[13] W.P. Thurston, The Geometry and Topology of 3-mainfolds, Princeton University, Princeton, 1978.
[14] I. Torisu, Boundary slopes and knots, Osaka Journal of Mathematics 33 (1996) 47–55.
[15] F. Waldhausen, Gruppen mit Zentrum und 3-Dimensionale Mannigfaltigkeiten, Topology 6 (1967) 505–517.