ALMOST NORMAL SURFACES IN 3-MANIFOLDS

MICHELLE STOCKING

ABSTRACT. J. H. Rubinstein introduced the theory of almost normal surfaces
to solve several homeomorphism problems for 3-manifolds. A. Thompson sim-
plified Rubinstein's algorithm for recognizing the 3-sphere by using almost
normal surface theory and thin position. This paper discusses higher genus
analogues to A. Thompson's work.

1. INTRODUCTION

In 1961 Haken [5] developed the theory of normal surfaces to find an algorithm
that decides if a knot is the unknot. Since then, normal surface theory has been
used to prove several different results and theories. In 1982 Jaco and Oertel [11]
found an algorithm to determine if a manifold is Haken. In 1992 Rubinstein [15]
described algorithms to determine if a triangulated 3-manifold is the 3-sphere and to
solve the homeomorphism problem for triangulated, closed, orientable, irreducible
3-manifolds with Heegaard genus at most two, using PL minimal surface theory
along with almost normal surface theory. Thompson [17] simplified the algorithm
for the 3-sphere recognition problem and used thin position, almost normal surface
theory, and standard 3-manifold techniques to prove that the algorithm works.
Matveev [13] modified Thompson's proof on S^3 using handle decompositions instead
of triangulations. Prior to Rubinstein's algorithm, Birman and Hilden [1] and
Volodin, Kusnezov, and Fomenko [18] solved the recognition problem for S^3 for
3-manifolds with genus 2 Heegaard splittings.

Rubinstein's results and Thompson's techniques are the key influences on this
paper. The main goal of this paper is to prove the following theorem using thin
position and almost normal surface theory:

Theorem 1. *If M is a triangulated, irreducible 3-manifold with a non-trivial genus
g strongly irreducible Heegaard splitting, then M contains an almost normal genus
g surface isotopic to the Heegaard surface.*

Rubinstein announced a similar result in [15], based on a minimax sweepout of
the 3-manifold with surfaces. A sweepout resembles a foliation except that the
leaves are allowed to intersect each other. A minimax sweepout is a sweepout that
has the smallest maximum complexity of one of the surfaces in the sweepout. The
complexity depends on the number of intersections of the surface with the edges
of the triangulation as well as the length of the graph where the surface meets the

2-skeleton. The 2-skeleton of the triangulation is given a metric by gluing together hyperbolic metrics on triangles for each 2-simplex so that the angles are all some very small number.

In this paper we will use Gabai's thin position of the 1-skeleton with respect to a foliation of the 3-manifold. Gabai introduced thin position to solve the Property R conjecture, [4]. Other important ideas in this paper come from Pitts and Rubinstein [14] and Jaco and Rubinstein [12]. Pitts and Rubinstein introduced techniques to study unstable Heegaard surfaces. Jaco and Rubinstein developed the idea of least weight normal surfaces to obtain PL minimal surfaces in 3-manifolds.

Note that throughout this paper we will be assuming that all arcs are in general position.

2. Normal surfaces and almost normal surfaces

2.1. **The definition of an almost normal surface.** Let M be an orientable 3-manifold with triangulation Δ. This triangulation need not be combinatorial. See Hempel, [9], for details on triangulations. A properly embedded surface in M, S, is *almost normal* if S intersects each tetrahedron in Δ in a collection of normal triangles and quadrilaterals and in one of the tetrahedra there is exactly one exceptional piece. This exceptional piece is either a disk whose boundary is a normal curve of length eight (sometimes called an octagon; see Figure 1), or it is an annulus which consists of two normal disk types with a tube between them that is parallel to an edge of the one-skeleton (see Figure 2).

Note that the intersection of S and the tetrahedron that contains the exceptional piece may consist of normal triangles along with the exceptional piece.

2.2. **Lengths of normal curves.**

Claim 1. *The only odd length normal curve is of length three.*

Proof of Claim 1. Any odd length normal curve on a tetrahedron must separate one vertex from the other three. See Figure 3.

The normal curve cuts the boundary of the tetrahedron into two disks. Consider the disk containing one vertex. The only intersections of the edges with this disk
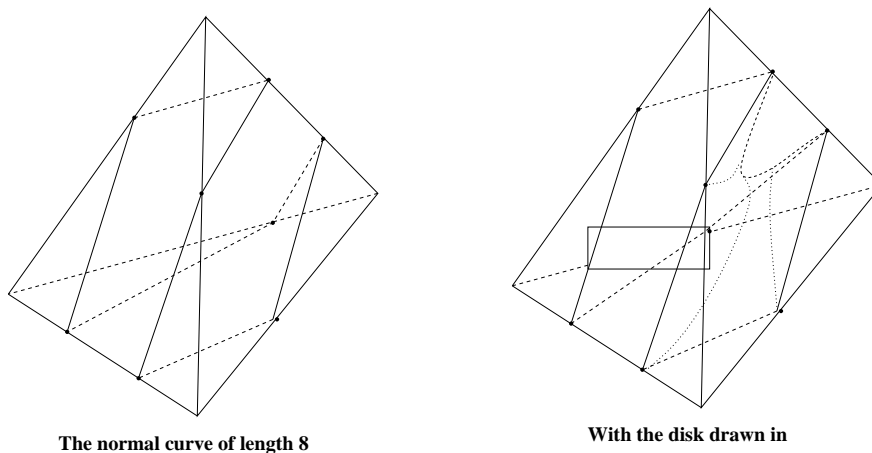

The normal curve of length 8   **With the disk drawn in**
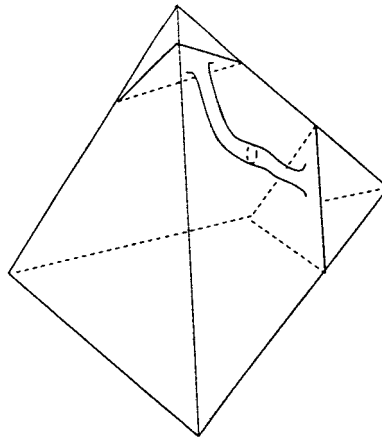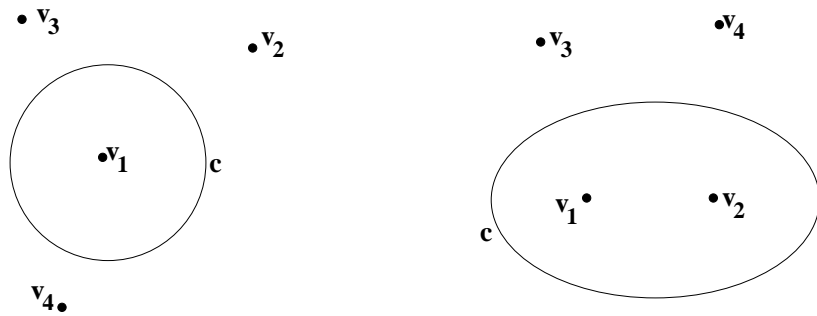
FIGURE 1. Normal octagon

FIGURE 2. An annulus

Of the 6 edges, three must intersect c an odd number of times, and the other 3 an even number of times, totaling an odd number for the length of c.

Of the 6 edges, four must intersect c an odd number of times, the other two an even number of times, totaling an even number for the length of c.

FIGURE 3. Odd length normal curves separate one vertex from the others

are arcs connecting the vertex to a point on the curve, since the curve is normal. The vertices are valence three, so the normal curve has length three. See Figure 4.

$\square$

Claim 2. *There are no normal curves of length six.*

Proof of Claim 2. Let $D \subset \partial H$, where H is a tetrahedron, be a disk bounded by c (a normal curve of even length) containing the vertices v_1 and v_2. There are only two ways that the edges can intersect this disk. Either they separate the vertices, or they connect a vertex to an edge. See Figure 5.

Suppose an edge of H lies entirely inside D. Then the edge connecting the two vertices outside of D lies completely outside of D. Thus the length of the normal curve is four.
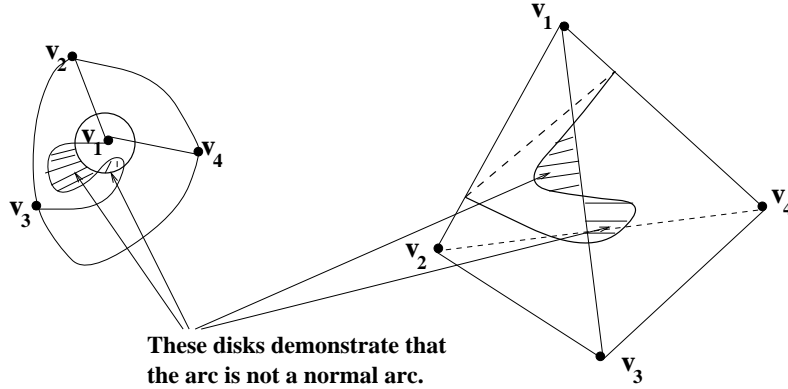
These disks demonstrate that the arc is not a normal arc.

FIGURE 4. All odd length normal curves have length three

c

v_1

v_2

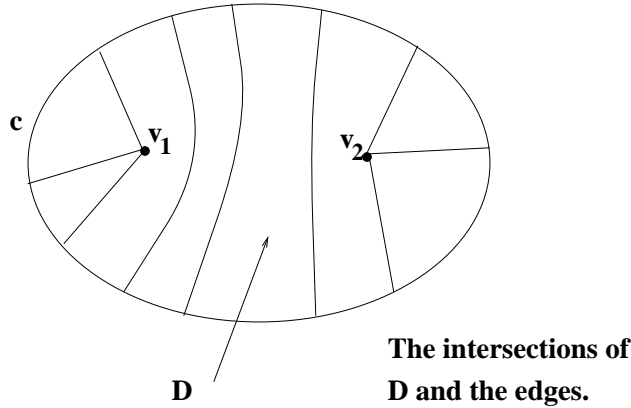**The intersections of D and the edges.**

D

FIGURE 5. The edges intersect D in only two ways

Suppose no edge of H lies entirely inside D. Then consider $D' = \partial H - D$. No edge lies entirely inside D'. So a counting argument tells you that the length of c is at least eight (four odds plus two non-zero evens $\geq$ eight.) $\qquad\square$

For more details on normal surfaces see Hemion, [8].

2.3. **An example.** We will list all possible normal surfaces and almost normal surfaces for the lens space $L(3,1)$ with the triangulation depicted in Figure 6. Note that this triangulation is not combinatorial.

The two triangles around vertex v_1, one in each tetrahedron, have their boundaries identified to create an embedded normal 2-sphere. The six triangles that are around vertex v_2 glue together to form a normal 2-sphere. Starting with a quadrilateral in one of the tetrahedra we get another normal 2-sphere. There are three normal 2-spheres obtained this way. See Figure 7. There are no other normal surfaces that are not parallel to these normal 2-spheres.

There is no almost normal surface with exceptional piece an octagon. So, the only possible almost normal surfaces have exceptional piece an annulus. The normal surfaces are all 2-spheres. So the only possible almost normal surfaces are either one
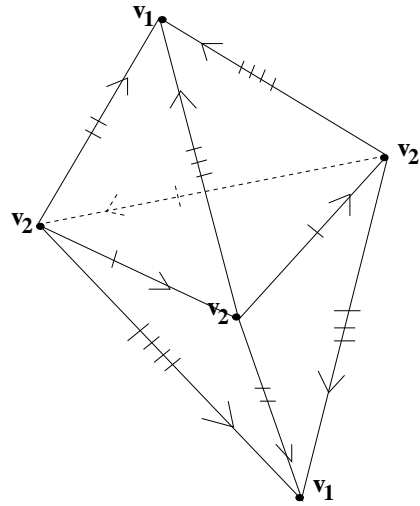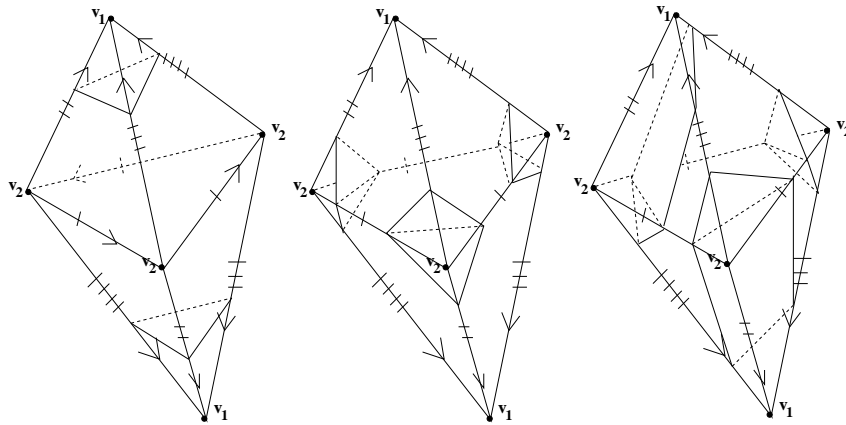
FIGURE 6. A triangulation of $L(3, 1)$

FIGURE 7. Normal 2-spheres

normal 2-sphere with a tube to itself, or two normal 2-spheres tubed together. The former gives an almost normal torus. The latter gives an almost normal 2-sphere.

2.4. **Isotoping almost normal surfaces to normal surfaces.**

Lemma 1. *Let S be an almost normal surface in M, an irreducible 3-manifold. Let S be incompressible to one side. Then S is isotopic to a normal surface that does not intersect S and that does not contain S to the incompressible side.*

Proof of Lemma 1. **Case 1.** *The exceptional piece of the almost normal surface, S, is an annulus, such that the meridian disk of the annulus is not to the incompressible side of S.*

The annulus is parallel to an arc of the one-skeleton, and can be isotoped so that it surrounds the arc. See Figure 8.
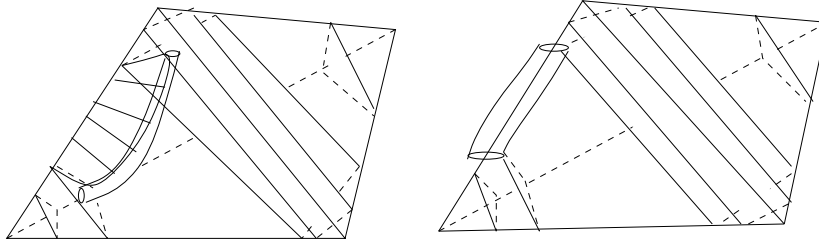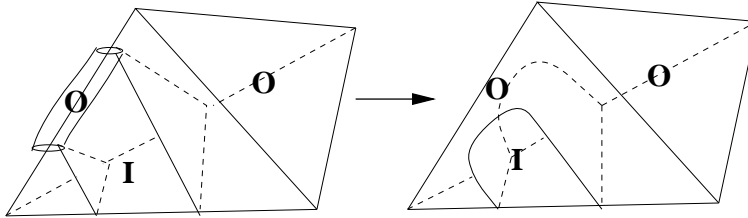
FIGURE 8. Sliding the annulus around an edge

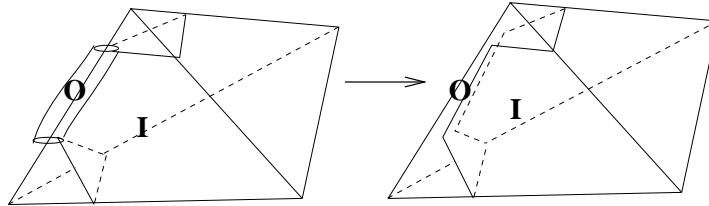FIGURE 9. Parallel normal triangles

FIGURE 10. Non-parallel normal triangles

Call the incompressible side of the almost normal surface the I side, and the other side the O side. When we slide the annulus around an edge, two points of the surface intersecting the edge will be isotoped together. The two points bound a subarc of the edge that is contained to the I side. In each tetrahedron that has this edge on it we will see the two points getting "pinched" together. There are four possibilities for the type of normal disks these two points lie on.

1. The points lie on two normal triangles that are parallel. See Figure 9.
2. The points lie on two normal triangles that are not parallel. See Figure 10.
3. The points lie on a normal triangle and a normal quadrilateral. See Figure 11.
4. The points lie on two parallel normal quadrilaterals. See Figure 12.

After we isotope the annulus around the edge, the intersection of the surface with the two-skeleton will no longer consist only of normal arcs. The only tetrahedra that have non-normal arcs of intersection of the faces with the surface that we have at this point are the ones resulting from the four possibilities above. All other tetrahedra contain only normal triangles and normal quadrilaterals.

Define the weight of the surface in M_i to be the number of points where the surface intersects the one-skeleton.
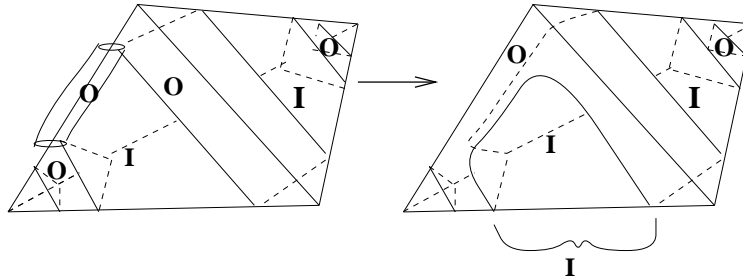
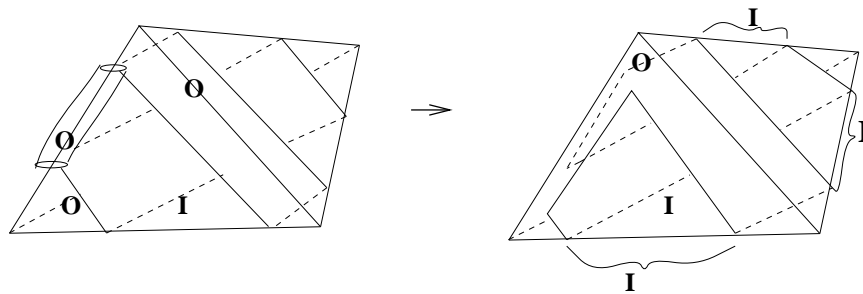FIGURE 11. A normal triangle and a normal quadrilateral

FIGURE 12. Parallel normal quadrilaterals

Note that when we isotope the annulus around the edge we are decreasing the weight.

We are going to isotope this new surface to a normal surface. During this isotopy we will always be pushing to the same side and we will always be reducing the weight.

Just after pushing the annulus across an edge if we look at any tetrahedron, the surface in that tetrahedron consists of disks. The number of times a disk intersects the one-skeleton will be called the length of the disk. The disk resulting from pinching together the two points on the two parallel normal triangles will be called the length four non-normal disk. The disk resulting from pinching together the two points on a normal triangle and a normal quadrilateral will be called the length five non-normal disk. The disk resulting from pinching together two points on two quadrilaterals will be called the length six non-normal disk. In these cases a non-normal arc on a face connects two points on an edge, and the subarc of the edge that the two points bound is contained on the I side of the surface.

Pick a non-normal arc in which the surface meets the faces of the tetrahedra. The non-normal disk it cobounds has length 4, 5, or 6. The non-normal arc and a subarc of the edge bound a disk in the face. We can isotope the non-normal arc through this disk to the subarc on the edge and then push it off the edge slightly. See Figures 13–15. The only resulting disk that we have not seen already is the length two non-normal disk in Figure 13. Any tetrahedron that shares the edge we isotoped over will have the two points on the edge pinch together. The two points that are pinched together bound a subarc of the edge to the I side of the surface.
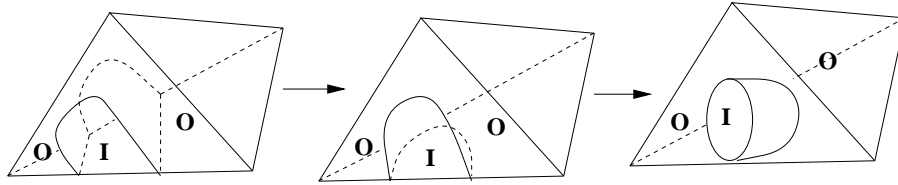
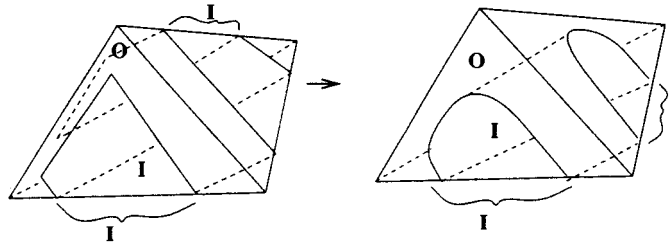FIGURE 13. Isotoping the length four and two non-normal disks

FIGURE 14. Isotoping the length six non-normal disk
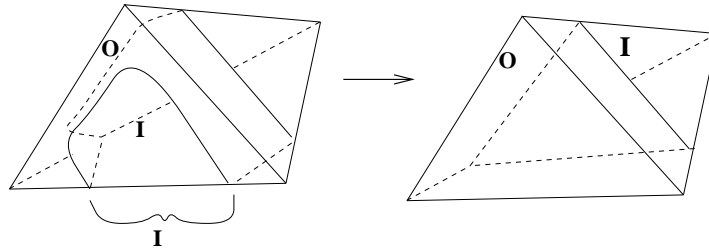
FIGURE 15. Isotoping the length five non-normal disk

These two points are either on two different normal disks, on one non-normal disk, on a non-normal disk and a normal disk, or on two different non-normal disks.

If these two points are on two different normal disks, then when they are pinched together we will be creating a disk from the previous four possibilities that we had when we isotoped the annulus over an edge.

If these two points are on one non-normal disk, then the two points that are pinched together either are connected by a non-normal arc on a face or are not. When the two points are connected by a non-normal arc on a face, then they pinch off a circle of intersection of the surface with a face. The resulting surface is the surface that would have been obtained by isotoping the non-normal arc across the edge and attaching an annulus to a face to give the circle of intersection. The only case where the two points do not lie on a non-normal arc on a face is when we have the non-normal arc of length six resulting from pinching two parallel normal quadrilaterals together. The result of pinching together these two points is two length two non-normal disks with an annulus between them. See Figures 16–18. In all of these cases we have the same disks that we obtained before, possibly with an
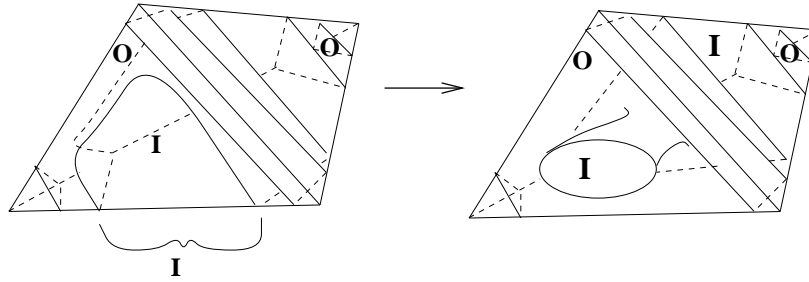
FIGURE 16. Length five pinched
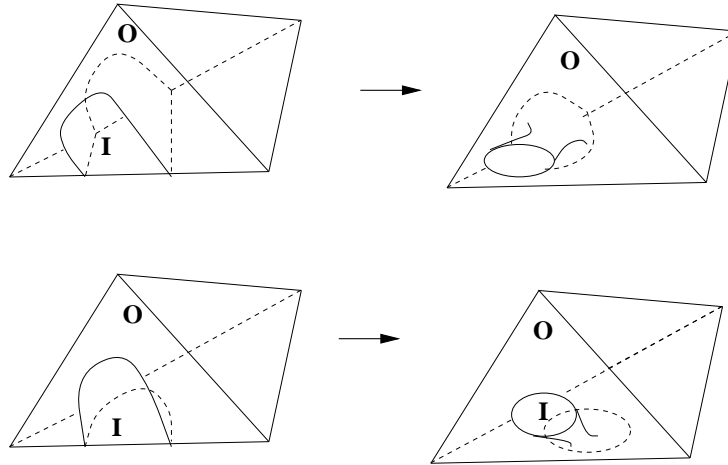
FIGURE 17. Length four and two pinched

annulus running from the disk to a face or between two disks.

If these two points are on a non-normal disk and a normal disk, then the only possibility for the non-normal disk is the one of length five. A length five non-normal disk is the only case where a normal disk could be to the I side of it. In all other cases the edges of the tetrahedron have either two or no points of intersection with the non-normal disk, and the two points bound a subarc contained to the I side. Note that in the case where we have a length five non-normal disk and a normal disk, the normal disk would have to be a normal triangle. When we pinch the two points together, the resulting disk is the length six disk we had in a previous case or possibly a new length six non-normal disk (see Figure 19).

It is not possible for the points to be on two different non-normal disks since the only non-normal disk that intersects an edge only once is the disk of length five. The only normal disks that can be on the I side of the non-normal disk of length five are triangles. The only way to create a non-normal disk of length five is to pinch together a normal triangle and a normal quadrilateral. Since no normal quadrilateral can exist on the I side of the non-normal disk, there cannot be another length five non-normal disk on the I side of the first length five disk.
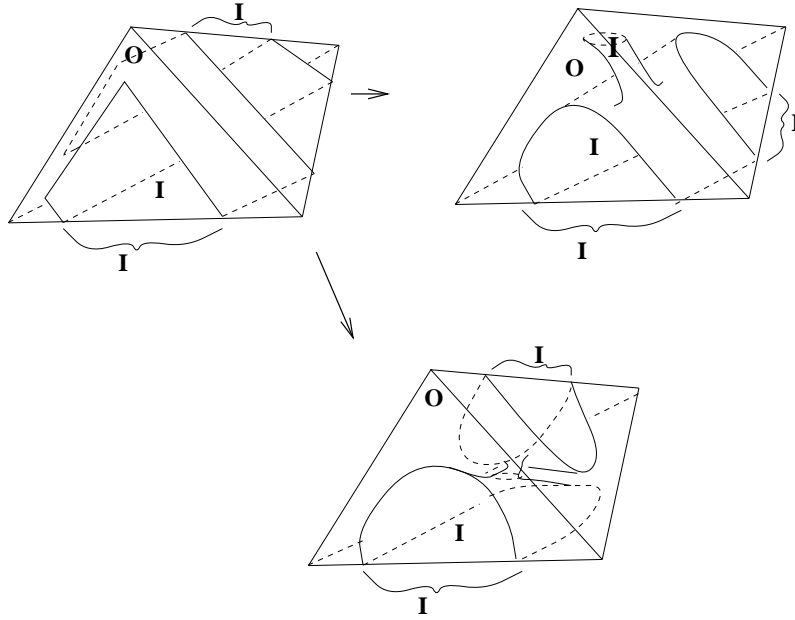
FIGURE 18. Length six pinched
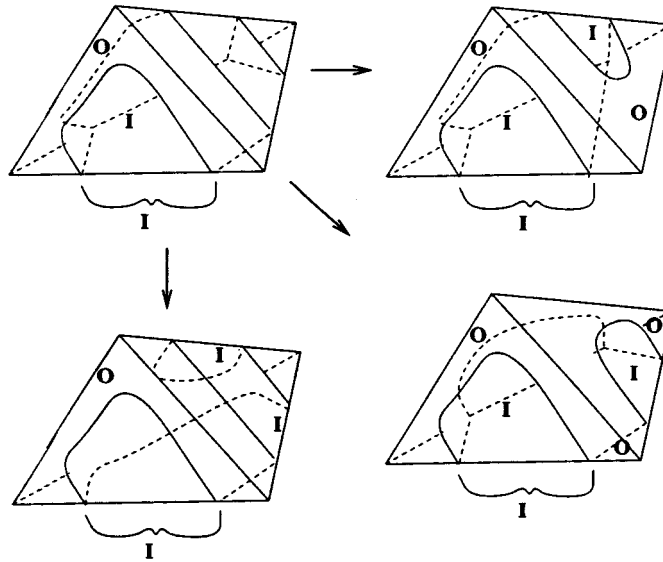
FIGURE 19. Length five and normal disk pinched

The only new disks resulting from pinching two points together and isotoping a non-normal arc are the length two non-normal disk and the new length six non-normal disk. We only need to investigate what happens to each of these when we isotope a non-normal arc across an edge or when we pinch together two points. Both of these disks intersect the edges in pairs of points that bound a subarc of
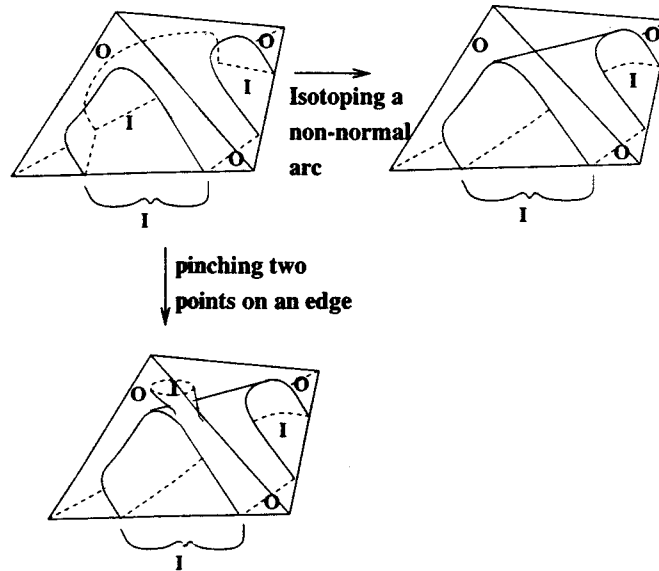
FIGURE 20. The other length six non-normal disk

the edge that is to the I side. When we isotope a non-normal arc of the length two disk across an edge, we get a disk in the tetrahedron that intersects a face in a simple closed curve. See Figure 13. When we pinch the two points together on the length two disk, we get an annulus that intersects two faces in simple closed curves. See Figure 17. When a non-normal arc of the new length six non-normal disk intersected with the 2-skeleton is isotoped across an edge, we obtain the same non-normal length four disk as before. When two points on the new length six non-normal disk are pinched together, we obtain the length four non-normal disk with an annulus attached. See Figure 20.

Each time we isotope a non-normal arc across an edge we are isotoping to the I side and we are reducing the weight. Continue to isotope until the intersection of the surface with the two-skeleton consists of normal arcs and simple closed curves in a face. All of the possibilities for the disks (possibly with annuli attached) are of length six and under. This implies that when we compress the simple closed curves in the faces we get a collection of normal surfaces. Since the simple closed curves all bound a disk in a face that is to the I side (incompressible side) of the surface, when we compress the curve we are only cutting off a 2-sphere. Since the manifold is irreducible, the 2-spheres all bound 3-balls. This implies that the remaining normal surface is isotopic to the almost normal surface. The almost normal surface is to the O side of the normal surface obtained, since we are always isotoping to the I side.

Case 2. *The exceptional piece of S is an annulus with meridian disk to the incompressible side of S.*

One of the normal surfaces that the annulus connects must be a 2-sphere. It is not difficult to see that the 3-ball that the normal 2-sphere bounds does not contain the other normal surface. Isotope S by isotoping the normal 2-sphere through the
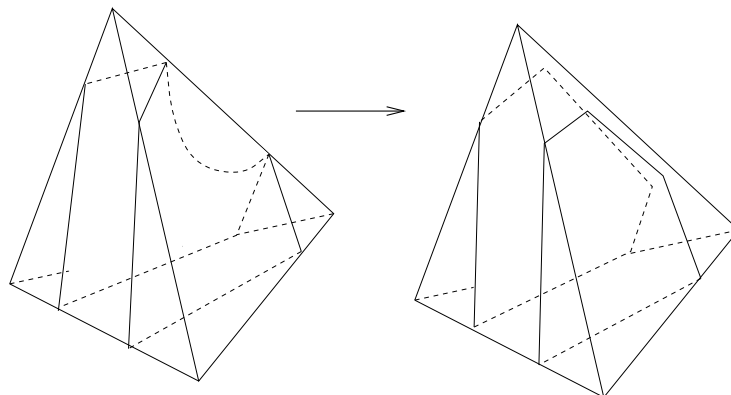
FIGURE 21. Isotoping the octagon

3-ball to the meridian disk of the annulus. This isotopy is to the incompressible side of S.

Case 3. *The exceptional piece of the almost normal surface is an octagon.*

Isotope the octagon of the almost normal surface to the incompressible side. When we do this, we are pinching two points on the edge together. This leads to the same set of pictures as in Case 1. See Figure 21.   □

3. THIN POSITION

3.1. **Definitions.** Thin position was first introduced by Gabai, [4].

Let (W, W') be a Heegaard splitting for a 3-manifold M. Note that the compression bodies in the Heegaard splittings we will consider can have 2-sphere boundary components. We will define a *standard singular foliation* of (M, W, W') to be a collection of embedded surfaces, Q_t (which we will call *leaves* of the foliation), where $0 < t < 1$, along with two singular leaves Q_0 and Q_1 such that Q_t for $0 \leq t \leq 1$ are disjoint. A singular leaf is the collection of boundary surfaces of M in one of the compression bodies along with the core of the compression body attached to the surfaces. It is possible for a singular leaf to consist only of the core if one of the compression bodies is a handlebody or a point if one of the compression bodies is a 3-ball. The leaves are chosen such that the union of Q_t over t from 0 to 1 is M and such that $Q_{1/2}$ is the Heegaard surface, Q_t as t goes from $0 + \epsilon$ to $1/2$ are isotopic copies of the Heegaard surface that fill in the compression body from Q_0 out to $Q_{1/2}$, and Q_t as t goes from $1/2$ to $1 - \epsilon$ are isotopic copies of the Heegaard surface that fill in the other compression body from $Q_{1/2}$ to Q_1. See Figure 22.

Note that there is a height function associated to a standard singular foliation. We will call Q_0 the top leaf. Then the height function will be the map h from M onto the unit interval such that all points on a leaf Q_t will be sent to t.

Define an *isotopy of a foliation* to be the image of the foliation under an isotopy of the 3-manifold.

Let (M, W, W') be a 3-manifold with a Heegaard splitting, and let F be a standard singular foliation of M. Let Δ_1 be a collection of arcs properly embedded in M, such that all but a finite number of the leaves of F intersect Δ_1 transversely,
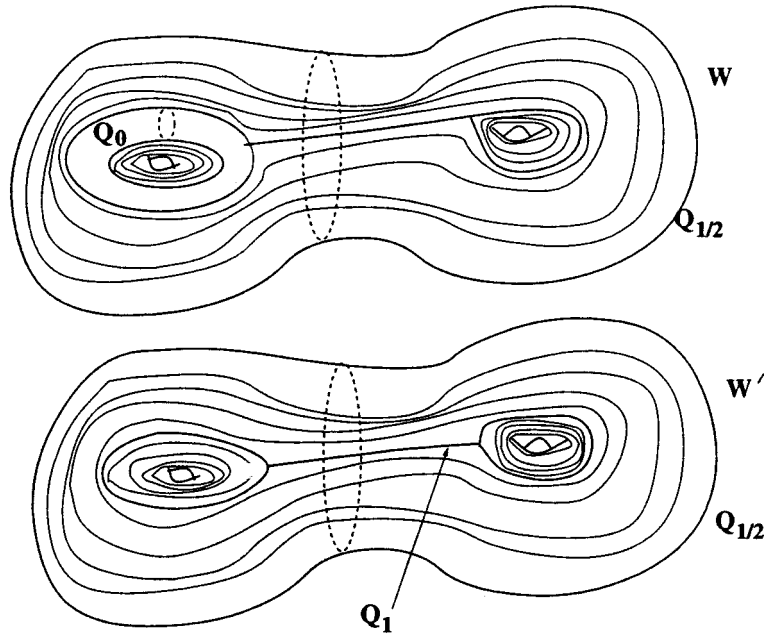
FIGURE 22. A singular foliation

every leaf of F has no more than one point of tangency with Δ_1, and Δ_1 is disjoint from the arcs of singularities on the singular leaves. (Note that we will be considering the one-skeleton of a triangulation, Δ, as the collection of arcs Δ_1 in the proof of Theorem 1.) The leaves of F with a point of tangency to Δ_1 are called *tangent leaves* and the others are *transverse leaves*. Label the tangent leaves, $S_1, S_2, ..., S_n$. Pick a transverse leaf of F between each adjacent pair of tangent leaves, S_i and S_{i+1}; call it L_i. Pick a transverse leave of F between the top singular leaf and S_1; call it L_0. Pick a transverse leaf between S_n and the bottom singular leaf; call it L_n. Define the *width of Δ_1 with respect to F* to be the sum over i, $i = 0, 1, ..., n$, of the number of times Δ_1 intersects L_i. Let $\mathcal{F}$ be the collection of all standard singular foliations of (M, W, W'). Define the *width of Δ_1 with respect to $\mathcal{F}$* to be the minimum width of Δ_1 with respect to F over all foliations F in $\mathcal{F}$. If a foliation, F, realizes the width of Δ_1 with respect to $\mathcal{F}$, then Δ_1 is in *thin position* relative to $\mathcal{F}$.

Let F be a foliation of M that realizes the width of Δ_1 with respect to $\mathcal{F}$. The arcs will form sequences of maxima and minima with respect to F. A leaf of F in the region where there is a shift as you move downward from maxima to minima is called a *thick leaf* and the region is called a *thick region*. See Figure 23.

An *upper (lower) disk* for a leaf, L, of F is a disk K in $\text{int}(M) - \Delta_1$ such that $\partial K = \alpha \cup \beta$, where α is an arc embedded in L, $\partial\alpha = \partial\beta$, β is a subarc of Δ_1, $K - \alpha$ intersects L transversely, and a small product neighborhood of α lies above (below) L. K describes an isotopy of Δ_1 in which the arc β can be replaced by α, and then pushed slightly off L. A *strict upper (lower) disk* is an upper (lower) disk such that $\text{int}(\beta)$ lies entirely above (below) L.
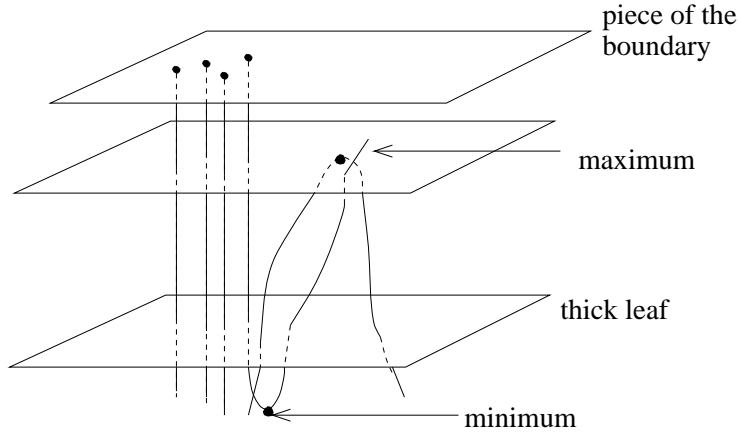
FIGURE 23. A maximum with respect to Δ_1

3.2. Isotoping upper (lower) disks.

Lemma 2. *If there exists an upper (lower) disk for a leaf that is not strict, then Δ_1 is not in thin position with respect to the foliation type.*

The proof of Lemma 2 is in [4]. See Scharlemann and Thompson, [16], for details.

Lemma 3. *If there is a leaf with an upper and a lower disk that are either disjoint or intersect at one point on Δ_1, then Δ_1 is not in thin position.*

The proof of Lemma 3 is in [4]. See [16] for details.

4. RECOGNITION ALGORITHMS

4.1. $S^2 \times S^1$.

Theorem 2. *Given M, a closed, orientable, triangulated, 3-manifold, there exists an algorithm to decide if M is $S^2 \times S^1$.*

Proof of Theorem 2. Let M be a closed, orientable, triangulated 3-manifold. If there is a non-separating 2-sphere in M, then we can find a non-separating normal 2-sphere by isotoping the non-normal arcs of intersection of the 2-sphere with the two-skeleton and compressing the simple closed curves of intersection of the 2-sphere with the two-skeleton. Thus, if there aren't any non-separating normal 2-spheres, then M is not $S^2 \times S^1$.

Claim 3. *If M is $S^2 \times S^1$, then there is a non-separating, normal 2-sphere in the collection of fundamental surfaces for M.*

Proof of Claim 3. This is similar to the proof for showing that it is a finite process to search for a normal 2-sphere, [17]. Let S be a non-separating normal 2-sphere. Assume that S is not a fundamental surface. Then $S = J + K$ where J and K are connected, normal surfaces. Since S is non-separating, at least one of J or K is also non-separating. Suppose J is non-separating. The Euler characteristic of S is equal to the Euler characteristic of J plus the Euler characteristic of K. Since the Euler characteristic of S is two, either both J and K have Euler characteristic one or one of them has Euler characteristic two and the other zero.
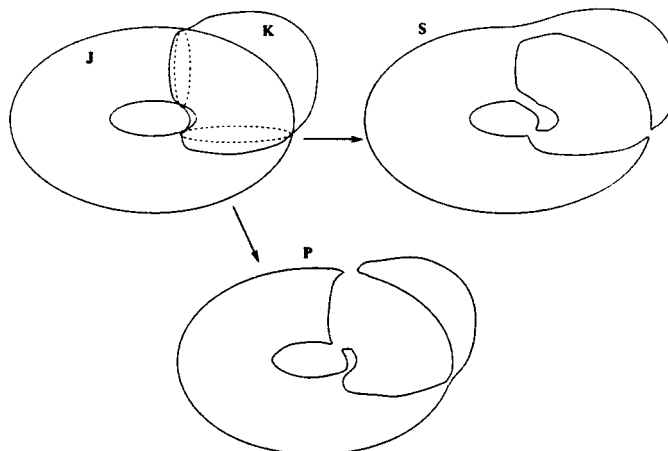
FIGURE 24. Alterations on $J + K$

Assume that S is a normal non-separating 2-sphere that is composed of as few normal disks as possible. Then J cannot be a 2-sphere, since this would be a normal non-separating 2-sphere composed of fewer normal disks. Suppose J has Euler characteristic one. Then J and K must be projective planes. Since M is orientable and J is not, J is one-sided. Thus, the boundary of a regular neighborhood of J is a 2-sphere. This implies that M has an RP^3 as a connect summand. See [9], Lemma 6.3. In this case M is not $S^2 \times S^1$.

This implies that J has Euler characteristic zero. So, J is a normal non-separating torus, or a normal non-separating Klein bottle, and K is a normal separating 2-sphere. When we perform a Haken sum, there are two choices for how to glue the surfaces together. One way yields the normal 2-sphere, S; the other yields a non-normal 2-sphere (call it P). See Figure 24. We can isotope the non-normal arcs of intersection of the surface P with the two-skeleton into the adjacent tetrahedra; when we do this we are reducing the number of disk intersections of the surface with the tetrahedra. We continue to isotope all non-normal arcs and we compress any simple closed curves in the faces. The resulting collection of 2-spheres either are contained in a 3-ball or are normal. At least one of these 2-spheres is non-separating since S is non-separating. All of these 2-spheres are composed of fewer normal disks than S. This contradicts S being a non-separating normal 2-sphere composed of the fewest normal disks. Therefore, S must be a fundamental surface. □

Count the number of boundary components of the only component of M minus the non-separating normal 2-sphere. If there are not exactly two boundary components, then M is not $S^2 \times S^1$. Decide if M minus the non-separating 2-sphere, S, is $S^2 \times I$ or not. This can be done by using the algorithm for S^3, [17]. That is, we cut $M - S$ open along a maximal collection of non-parallel normal 2-spheres, Σ. If there is a non-separating normal 2-sphere in $M - S$, then $M - S$ is not $S^2 \times I$, and therefore, M is not $S^2 \times S^1$. By [17] the components of $M - S - \Sigma$ are either 3-balls, punctured 3-balls, or have only one boundary component. If a component has only one boundary 2-sphere and is not a 3-ball, then $M - S$ is not $S^2 \times I$.

A component of this type is a 3-ball if and only if it contains an almost normal 2-sphere with an octagon (see [17]). Therefore, we search these components for an almost normal 2-sphere with an octagon to decide if $M - S$ is $S^2 \times I$ or not. Since M is orientable, $M - S$ is $S^2 \times I$ if and only if M is $S^2 \times S^1$. □

4.2. **Lens spaces.** If we are given a lens space, by definition it has a genus one Heegaard splitting, and if it is not S^3 or $S^2 \times S^1$, then the Heegaard splitting is strongly irreducible. Certainly, if we can find an almost normal torus in a closed 3-manifold M that bounds a solid torus to each side, then M must be a lens space. So, with Theorem 1 it is enough to search a given closed 3-manifold for an almost normal torus that bounds a solid torus to each side, in order to solve the recognition problem for lens spaces. If we can say that no such thing exists, then we can say that M is not a lens space. Arguments for showing that it is a finite search for such a torus were developed by Rubinstein, [15]. Dave Bachman filled in the details.

In the example in section 2.4 we noticed that the only almost normal surfaces are tori consisting of an annulus attached to one normal 2-sphere or 2-spheres consisting of two normal 2-spheres with an annulus between them. Thus, there is only one possibility for the almost normal torus satisfying Theorem 1 for this example. It must be a normal 2-sphere with an annulus.

5. Proof of Theorem 1 for closed 3-manifolds

Let (M, W, W') be a closed 3-manifold, M, with a non-trivial, strongly irreducible, genus g Heegaard splitting of M, (W, W'). Let G be the Heegaard surface. Let M be triangulated, and call the triangulation of M, Δ. Note that the triangulation need not be combinatorial.

We can find a maximal collection of separating, normal, non-parallel 2-spheres, Σ in M, using an algorithm due to Haken. Components of $M - \Sigma$ with more than one 2-sphere boundary component are punctured 3-balls. Components of $M - \Sigma$ with just one boundary component are either 3-balls around a vertex or contain no vertices and are 3-balls if and only if they contain an almost normal 2-sphere with exceptional piece an octagon, [17]. If all components with a single boundary component are 3-balls, then M is a 3-sphere. It is a result of Waldhausen that the only irreducible Heegaard splitting of S^3 is of genus zero. (Note that we do not consider the standard genus one splitting of S^3 to be strongly irreducible.) If M is S^3, then take two normal 2-spheres and an annulus between them parallel to an edge as the almost normal 2-sphere. There is always at least one normal 2-sphere since there is at least one vertex, and the two normal 2-spheres could be parallel. Thus, M has an almost normal 2-sphere, as required by Theorem 1.

We can assume M is not S^3, and so there is at least one component of $M - \Sigma$, call it M_0, with only one boundary 2-sphere, containing no vertices and no almost normal 2-spheres with exceptional piece an octagon. In fact, there will only be one of these components of $M - \Sigma$, since M is irreducible. M_0 has a Heegaard splitting where W_0 and W'_0 are the compression bodies obtained from W and W' after first isotoping G so that the normal 2-sphere lies to one side of the splitting surface. Call this new splitting surface G_0. Then cut out the 3-ball that the normal 2-sphere bounds.

We can foliate (M_0, W_0, W'_0) with a standard singular foliation. Consider the normal 2-sphere plus the core of the compression body it lies in to be the top singular leaf of the foliation. See Figure 25.
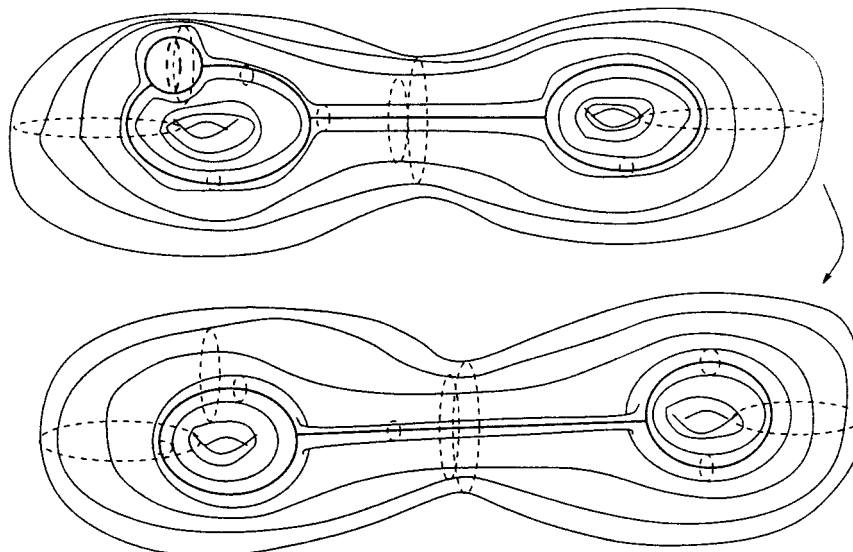
FIGURE 25. A singular foliation of M_0

Call the collection of all standard singular foliations of (M_0, W_0, W'_0), $\mathcal{F}_{G_0}$. Call the one-skeleton of M that is in M_0, Δ_1. Put Δ_1 into thin position with respect to $\mathcal{F}_{G_0}$. Let F_0 be a foliation in $\mathcal{F}_{G_0}$ that realizes the width of Δ_1. If the core of the compression body intersects Δ_1, then isotope Δ_1 slightly off of the core. We can find F_0 since the width is a non-negative integer.

We define the quintuple $(M_i, \Delta_1, W_i, W'_i, F_i)$ as follows. M_i is a bounded, submanifold of M with one-skeleton, Δ_1. M_i has a non-trivial, strongly irreducible Heegaard splitting (W_i, W'_i). F_i is a standard singular foliation of (M_i, W_i, W'_i). Let $(M_i, \Delta_1, W_i, W'_i, F_i)$ have the following properties:

1. ∂M_i consists of normal surfaces that are incompressible in M_i and possibly one normal 2-sphere.
2. M_i does not contain an almost normal 2-sphere with exceptional piece an octagon and any normal 2-sphere it contains is parallel to the boundary normal 2-sphere.
3. M_i does not contain any vertices of the triangulation of M.
4. Let $\mathcal{F}_{G_i}$ denote the set of all standard singular foliations of (M_i, W_i, W'_i), where G_i is the Heegaard surface for the Heegaard splitting (W_i, W'_i). Δ_1 is in thin position with respect to $\mathcal{F}_{G_i}$. F_i is a foliation in $\mathcal{F}_{G_i}$ that realizes the width of Δ_1.

Notice that $(M_0, \Delta_1, W_0, W'_0, F_0)$ satisfies the above properties, so we can apply the following lemma to $(M_0, \Delta_1, W_0, W'_0, F_0)$:

Lemma 4. *If there is no thick region in $(M_i, \Delta_1, W_i, W'_i, F_i)$, then there is an almost normal surface in M_i that is isotopic to a surface obtained from a leaf of F_i by compressing the leaf to one side.*

Claim 4. *A closed, orientable, normal surface, S, in a 3-manifold, M, with triangulation, Δ, is incompressible in the complement of the one-skeleton.*

The proof of the claim can be seen by an innermost disk/outermost arc argument, [17].

If we assume there is no thick region, then the width only decreases. Call the leaf just above the first minimum L. The first leaf of the foliation is a singular leaf that consists of a collection of boundary surfaces with the core of the compression body attached. Since there is no change in the width between the first leaf and L, L consists of normal surfaces parallel to the boundary surfaces with the boundary of a closed neighborhood of the core attached. L is just before the first minimum so it has a lower disk. Call this disk E.

Note that the arcs of the one-skeleton intersect L only on the normal surfaces. Consider the compression body that contains the first leaf of the foliation. Let t be a thickened copy of the core of this compression body such that the one-skeleton does not intersect t and L is a copy of the normal surfaces with the boundary of t attached to it. We can find t since the core and Δ_1 are assumed to be disjoint and in general position. We will isotope t (and in turn we will be isotoping the boundary of t) by extending an isotopy of the graph associated to t that allows the edges to slide over the graph. Define an *edge of t* to be a thickened edge of the graph associated to t. Define a *tube* to be the boundary of an edge of t. The vertices of the graph associated to t initially are the ends of the graph that are on the normal surfaces and the points of the core that branch. We will slide the ends of the graph over the graph, creating more multivalent vertices. Consider a tetrahedron that contains a multivalent vertex. A thickened subarc of the graph that runs from the vertex directly to the two-skeleton without passing through any vertices will be called a *leg*.

Define the complexity of L to be (a, b) where a is the number of intersections t makes with the two-skeleton and b is the number of intersections the lower disk, E, makes with the two-skeleton, and (a, b) is lexicographically ordered. E runs once over an arc of the one-skeleton; call the arc β. Call the arc of intersection of E and L, α.

E will run over the side of the boundary of t that does not contain the core. We will say that E runs over the "outside" of t. Let's assume that E runs over all of the tubes of t. If it doesn't, then we can compress the meridian disks of the tubes of t that E doesn't run over and define t to be the remaining thickened graph. E runs over at least one tube of t. Otherwise, we have a ∂-compressing disk in the complement of the one-skeleton for a normal surface. This contradicts Claim 4.

Proof of Lemma 4. Let H be a tetrahedron in Δ. Consider $t \cap H$ as a union of connected components; call them t_i. Note that if t_i contains an endpoint of the graph associated to t, then we can slide the legs of t_i off of t and onto a normal disk without increasing the complexity. Assume we have done this for all tetrahedra after each of the following cases. See Figure 26.

Consider the intersections of E and the two-skeleton, Δ_2, on E. See Figure 27. By considering outermost arcs and innermost circles of $E \cap \Delta_2$ we will reduce the complexity to $(0, 0)$.

Case 1. *There is a circle of intersection of the two-skeleton with E.*

An innermost circle of intersection of the two-skeleton with E bounds a disk on E and a disk on a face of the two-skeleton. The disk on E is contained in a tetrahedron. The two disks together form a 2-sphere, which bounds a 3-ball in the
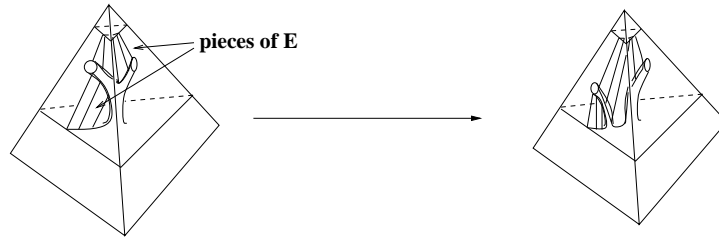
FIGURE 26. Sliding off the legs of t
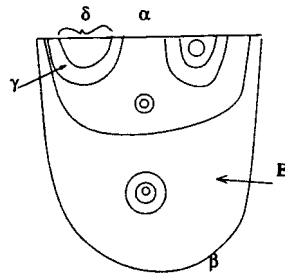
FIGURE 27. The lower disk, E
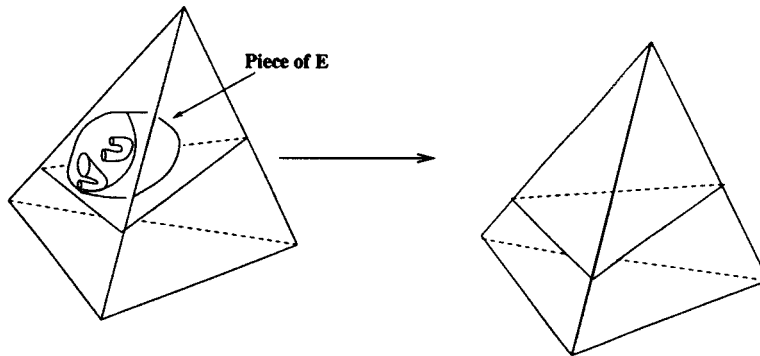
FIGURE 28. Case 1: circles of intersection

tetrahedron. We can isotope E through the 3-ball to reduce the complexity. See Figure 28.

We can eliminate all circles of intersection of the two-skeleton and E by Case 1. For the remaining cases we will be considering an outermost arc of intersection, γ. Let δ be an arc on α, such that $\gamma \cup \delta$ is the boundary of an outermost disk. γ and δ together bound a subdisk of E which is contained in a tetrahedron. Call this subdisk an outermost disk for γ. δ will alternate running over t and running on a normal disk.

Case 2. *δ runs over a normal disk only.*
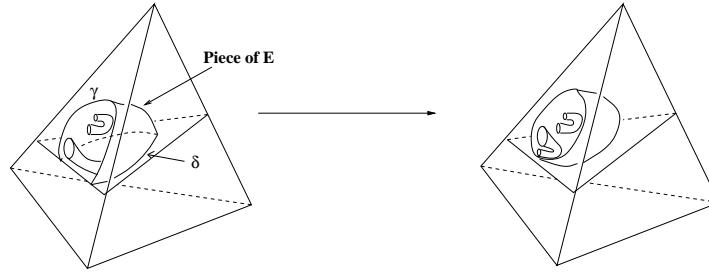
FIGURE 29. Case 2: δ runs over a normal disk only

γ has endpoints on the same normal disk. See Figure 29.

Consider the 3-ball bounded by the outermost disk for γ, a subdisk of the normal disk and a subdisk of the face. We will isotope the outermost disk through the 3-ball and into the adjacent tetrahedron, by first isotoping δ along the normal disk to the face. If t attached to the normal disk in the subdisk we are isotoping across, then we will isotope it along too. Then we will isotope E across the 3-ball to the circle of intersection we have just created with the two-skeleton. If t intersects the normal disk, then t must be attaching to the side of the normal disk that the subdisk of E is on. If t attached to the other side, then E could not run on the outside of the tube, since the leaf is separating. So, when we isotope δ, we may be increasing the number of intersections of t with the two-skeleton, but when we isotope through the 3-ball we are reducing the number of intersections of t with the two-skeleton by the same or more than what we just increased it by. So, in the end we have not increased the number of t intersections with the two-skeleton, and we have definitely reduced the number of intersections of E with the two-skeleton by at least one.

Case 3. *δ runs along only t and no normal disks.*

Either γ connects to the same circle of intersection of t with the two-skeleton or to two different circles.

Define δ "doubles back" to mean that δ is a trivial loop on t inside the tetrahedron.

If δ connects to the same circle of intersection of $t \cap \Delta_2$ and is non-trivial on t, then there exists a loop of the graph in the tetrahedron. This implies that there is a handle of the leaf contained in the tetrahedron. So, if γ connects to the same circle, then δ either doubles back or runs over a handle.

If it doubles back, then we can isotope δ out of the tetrahedron, pushing along pieces of E or t as necessary. Then we will be in Case 1 where we have a circle of intersection which can be eliminated by isotoping through the 3-ball. Whatever the complexity has increased by in isotoping δ it will decrease by when we isotope through the 3-ball. We will also have at least decreased b by one, so the complexity has been reduced. See Figure 30.

If δ runs over a handle, then we have a handle contained in a 3-ball. The 3-ball is the tetrahedron that the handle is in. The boundary of this tetrahedron is a 2-sphere. Since the manifold is not S^3, the 2-sphere does not bound a 3-ball in M to the other side. See Figure 31. See Frohman's paper, [3]. Drill out a solid torus in the handle.
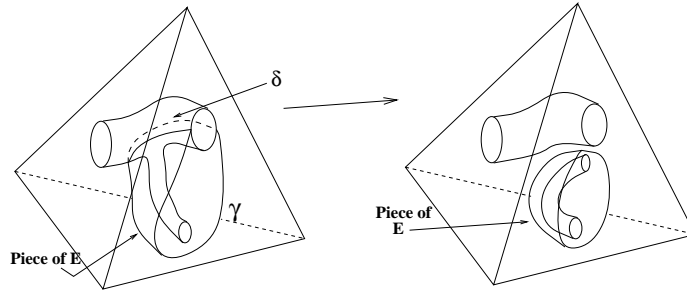
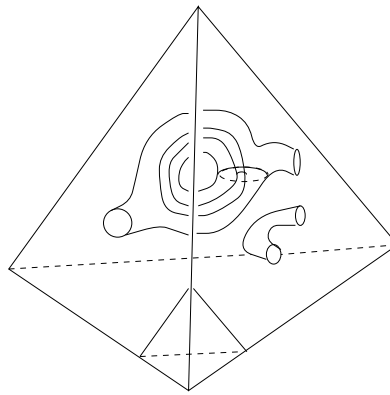FIGURE 30. Case 3: γ goes to the same circle

FIGURE 31. A handle contained in a tetrahedron

The leaf, L, is isotopic to a Heegaard splitting surface for compression bodies or comes from one by compressions all made to the same side. After drilling out the torus the 2-sphere becomes essential. By Haken's lemma [2] (see Casson and Gordon for a generalization of Haken's lemma [6]; see Jaco for a nice proof of Haken's lemma [10]), we know that there is an essential 2-sphere that intersects the Heegaard splitting surface in only one circle, and this circle is essential on the Heegaard surface. So we have a compressing disk on one side that has the same boundary as the compressing disk on the other. This tells us that the Heegaard surface is reducible, which is a contradiction.

If γ connects two different circles of $t \cap \Delta_2$, then either the piece of t δ stays on doesn't have any other edges attached to it in that tetrahedron or it does. In the first case, we have a disk to isotope the piece of t out of the tetrahedron which reduces the complexity. In the second case, it may not work to isotope t into the next tetrahedron because we may be pulling along other pieces of t and we may not decrease the number of intersections of t with the two-skeleton. So, in this case we will slide one of the legs of t across this disk to eliminate a circle of intersection of the two-skeleton with the t. We can take the thickened subarc of t that has boundary component equal to one of the circles of intersection on the face, as the leg we slide. Therefore, we can reduce the complexity. See Figure 32.

We have covered the only cases where δ doesn't run over both t and normal disks. In the rest of the cases, δ running over both implies that the pieces of t that
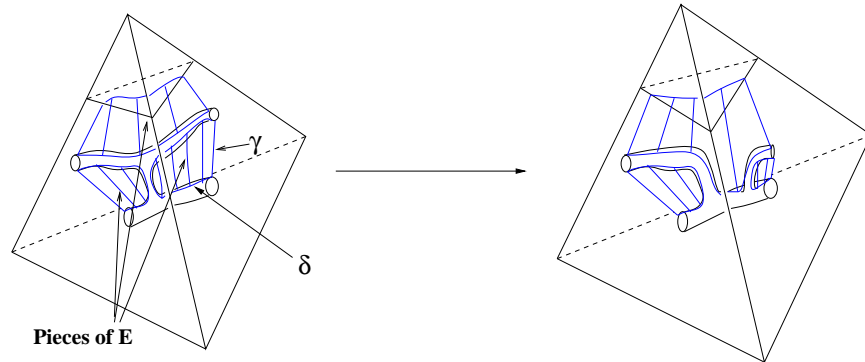
FIGURE 32. Case 3: γ connects different circles
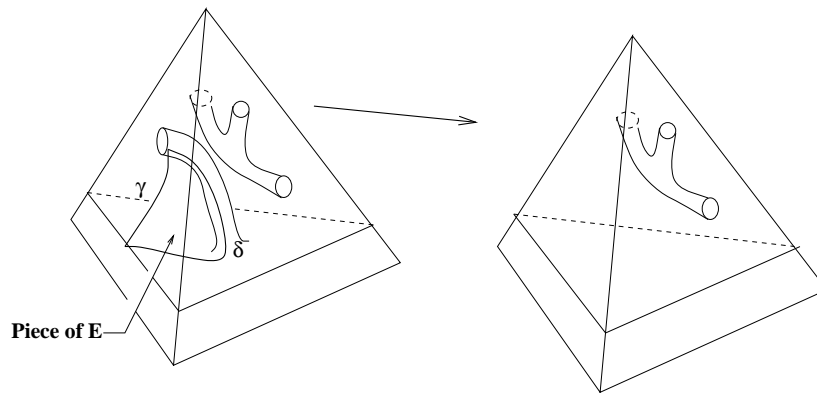
FIGURE 33. Case 4: δ runs along t and a normal disk

δ runs over will not have any legs. We assumed in the beginning that we would be isotoping the legs off of t when we can do this without increasing the complexity. And in the following cases there is an endpoint of t which allows this isotopy for the pieces of t on which δ travels.

If δ doubles back on t, then there is a disk on t bounded by a subarc of δ and a subarc of t intersect the normal surface through which we can isotope δ off of t. We may have pushed along pieces of E, but since this takes place inside the tetrahedron, we have not increased the complexity. Thus, we may assume that δ doesn't double back on t.

Case 4. δ *runs along* t *and then a piece of a normal disk.*

We can isotope the piece of t on which δ runs, through the outermost disk into the adjacent tetrahedron. See Figure 33. The complexity will be reduced, since a is reduced (b may increase).

Case 5. δ *runs over a normal disk, then* t, *then a normal disk.*

If γ has endpoints on the same normal arc, then a handle of the Heegaard splitting surface is contained in a 3-ball. See Figure 34.
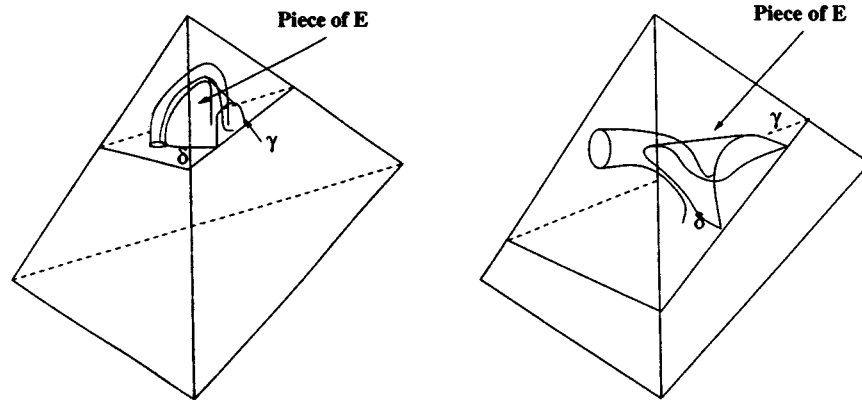
FIGURE 34. Case 5: γ goes to the same normal arc
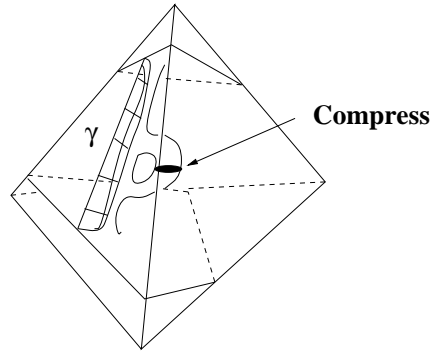
FIGURE 35. Case 5: γ connects different normal arcs

As before, the second possibility leads to a contradiction since the Heegaard splitting is irreducible.

If γ has endpoints on two different normal arcs, then we compress all of the tubes except the one δ runs on. The tube that δ runs over is parallel to an arc in the one-skeleton since γ is parallel to an arc of the one-skeleton. See Figure 35. So, we will get an almost normal surface (consisting of one or two normal surfaces with an annulus parallel to an edge). This surface was obtained by isotoping and compressing (to the same side) a leaf of the foliation.

Case 6. *δ runs along t, then a normal disk, then t again.*

If γ connects to two different circles of intersection of t with the two-skeleton, then isotope a leg of t through the disk and into the adjacent tetrahedron. We have reduced the complexity. See Figure 36.

If γ connects to the same circle, then δ bounds a disk on L (possibly ignoring some edges of t). We can isotope δ through the disk to an arc on t, pushing along any pieces of t or E that may be inside the disk. Now we apply Case 3 to reduce the complexity. See Figure 37.
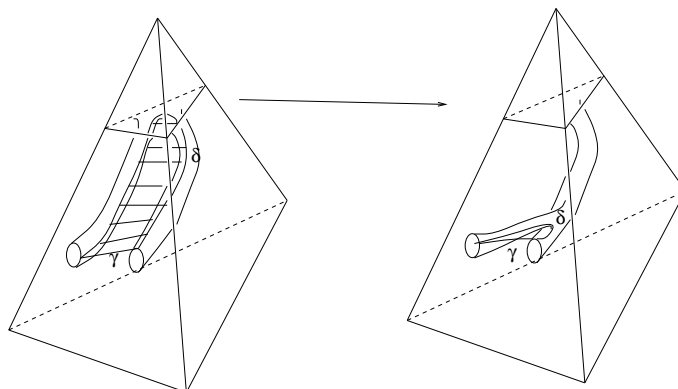
FIGURE 36. Case 6: γ connects different circles
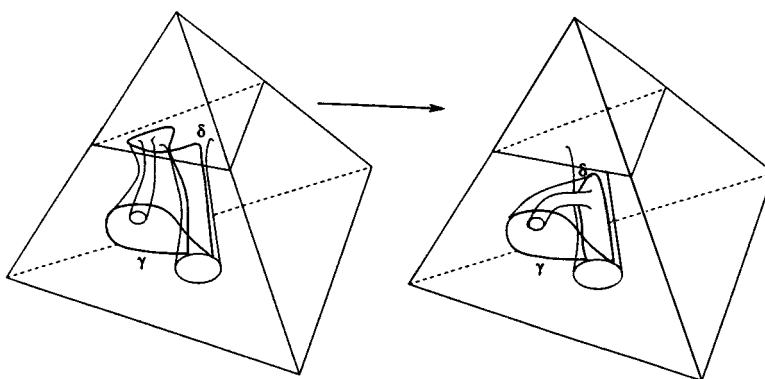
FIGURE 37. Case 6: γ goes to the same circle

Case 7. *δ runs on more than one tube of t, runs completely over at least one tube, and doesn't run over any tubes more than once.*

Since a handle of the surface cannot be contained in a 3-ball, the tube that δ runs entirely over must attach to two different normal disks and be contained in the tetrahedron. Attached to at least one of these normal disks there will be another tube on which δ runs. Isotope the corresponding edge of t over the edge of t that δ runs over entirely. This will not increase the complexity, since the isotopy takes place inside a tetrahedron. If we are not in one of the previous cases, then we can reapply Case 7. δ cannot run over infinitely many tubes, so eventually we will be in a previous case. See Figure 38.

Case 8. *δ runs entirely over at least one tube more than once.*

Given a tetrahedron with a collection of non-intersecting normal disks inside it, consider the components of the tetrahedron after cutting along the normal disks. Note that the largest number of normal disks that can bound a connected component of the tetrahedron is four. Since a handle cannot be contained in a tetrahedron, this implies that the largest number of tubes completely contained in a component
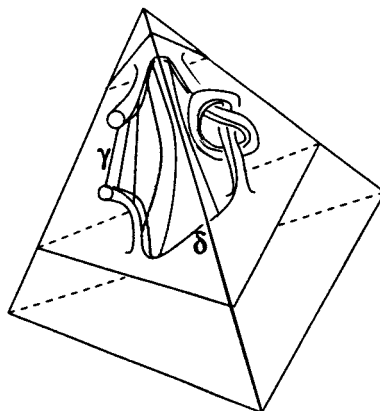
FIGURE 38. Case 7
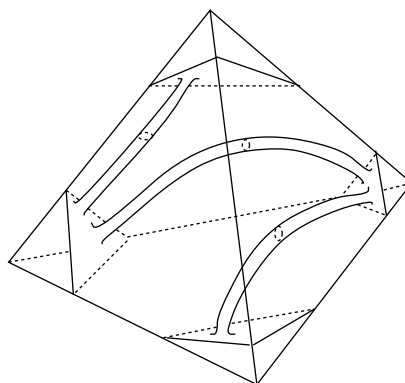
FIGURE 39. There are no more than three tubes

of the tetrahedron is three. See Figure 39. Since the tubes are all to the same side, a connected component of L intersect a tetrahedron has at most three tubes.

Consider the tubes and the normal disks that δ runs over. This surface will be a punctured disk (the number of punctures depends on the number of tubes δ runs over). See Figure 40. If it was not a planar surface, then there would be a handle contained in a 3-ball, which contradicts the Heegaard surface being irreducible.

We can represent the meridian curves for the tubes as circles around the boundary circles of the planar surface. δ will represent an arc connecting one or two boundary circles.

If δ connects two boundary circles, then δ can be isotoped to intersect each meridian curve at most once. So, we can reduce to a previous case without increasing the complexity, since this isotopy takes place inside a tetrahedron.

If δ connects a boundary circle to itself, then δ can be isotoped to intersect the meridian curves of the tubes at most twice. We can slide the edges of t over each other. This implies that we can isotope δ so that it doesn't intersect any meridian curves. So, again, we can reduce to a previous case without increasing the complexity. See Figure 41.
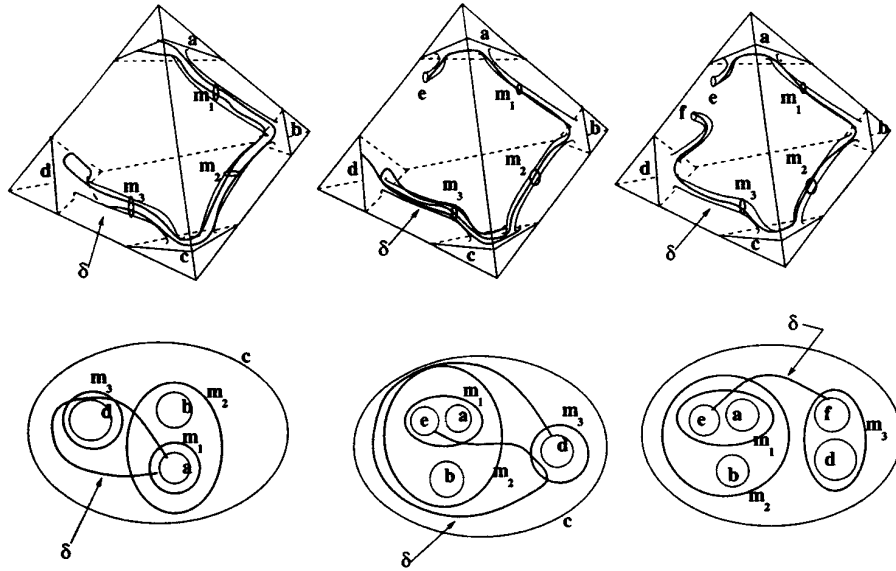
FIGURE 40. The tubes and normal disks form a planar region
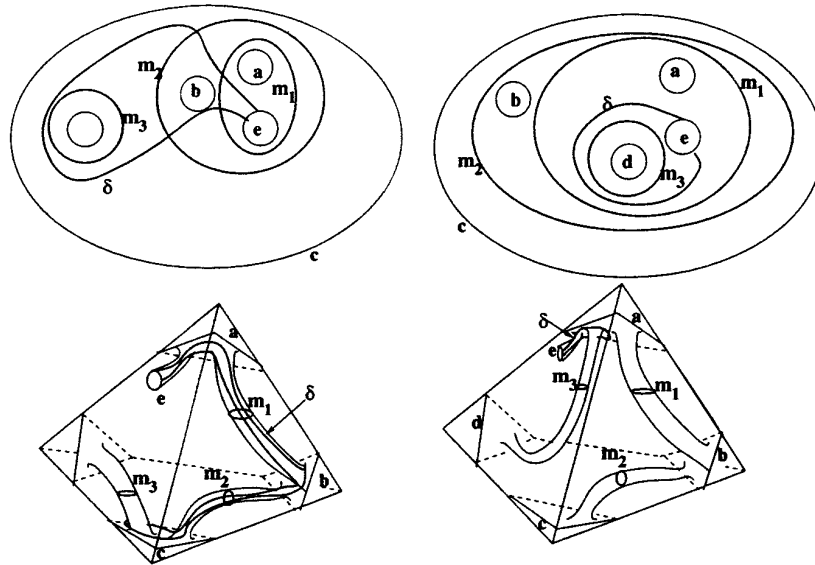
FIGURE 41. Shortening δ

All possible outermost disks and innermost circles are covered in the above cases. Case 1 eliminates all circles of intersection of the two-skeleton with E. Case 2 isotopes all outermost arcs where δ runs over a normal disk only. Case 3 isotopes all outermost arcs where δ runs along only t and no normal disks. Case 4 isotopes all outermost arcs where δ runs along t and then a piece of a normal disk. Case 5 isotopes all outermost arcs where δ runs over a normal disk, t, then a normal

Upper disk

L
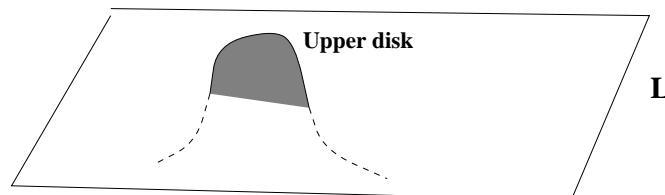
FIGURE 42. The upper disk for L

disk. Case 6 isotopes all outermost arcs where δ runs over t, a normal disk, then t. In the remaining cases there will be a subarc of δ that will run on t, a normal disk, then t, then a normal disk. Since we are assuming that δ doesn't double back, this implies that δ must run over an entire tube and that δ runs on more than one tube. Case 7 isotopes all outermost arcs where δ runs on more than one tube, runs completely over at least one tube, and doesn't run over any tubes more than once. Case 8 eliminates the possibilities where δ runs over at least one tube more than once by reducing to a previous case. So, we have covered all possible cases.

Since there are only a finite number of intersections, we can isotope the edges of t and E until the complexity is $(0,0)$. This tells us that the tubes are all contained in tetrahedra. If E runs over only k tubes, then ignore the tubes E doesn't run over. The tubes that E runs over do not connect to each other since the tubes lie in a tetrahedron and, therefore, can be isotoped off one another. If α runs over the tubes more than once, then consider the punctured disk consisting of the tubes and the normal disks they attach to. Recall that the graph associated to the tubes has no loops, since, if it did, this would give a handle in a tetrahedron. β runs between two different normal disks, so α connects two different boundary circles in the punctured disk. Therefore, we can isotope α so that it intersects the meridian circles of the tubes at most once. This implies that we can isotope E so that it runs over the tubes at most once. If E still runs over more than one tube, then we can isotope the edges of t over each other, through E, so that E runs over only one tube. This surface is one or two normal surfaces with an annulus that is parallel (by E) to an arc of the one-skeleton. Therefore, we have an almost normal surface.

Recall that L consists of the boundary normal surfaces of M_i with the boundary of t attached to them. This almost normal surface comes from isotoping the tubes of t and compressing the meridian disks. If the almost normal surface is isotopic to a leaf of the foliation, then it is the almost normal surface we are looking for. If the almost normal surface is not isotopic to a leaf of the foliation, then it is isotopic to a compression of a leaf. □

Let $(M_i, \Delta_1, W_i, W_i', F_i)$ have the same properties as before.

Lemma 5. *If there is a thick region in M_i, then there is a collection of normal or almost normal surfaces in M_i that is obtained from a leaf of the foliation by compressing to one side. At most one of the surfaces in the collection can be almost normal. Not all of the surfaces are boundary parallel. There is at most one normal 2-sphere.*

Proof of Lemma 5. The proofs of the following four claims are in [17]. Note that in this more general setting these proofs can be carried out because the boundary of M_i consists of normal surfaces.

Claim 5. *There exists a transverse leaf, L, of the foliation, F_i, such that $L \cap \partial H$ consists of normal and simple curves for each tetrahedron, H, in Δ.*

This leaf, L, of the foliation is found in a thick region. It has both an upper disk and a lower disk in M_i, since immediately above it there are only maxima and immediately below it there are only minima. By Lemmas 2 and 3, these disks must intersect other than at an endpoint.

For each tetrahedron H, $L \cap H$ is almost determined by $L \cap \partial H$. Picture $L \cap H$ as a collection of disks, $\mathcal{D}$, connected together by tubes (possibly very messy). $\partial \mathcal{D}$ is equal to the normal curves of $L \cap \partial H$

Claim 6. *$L \cap \partial H$ contains no curves of length greater than eight.*

Claim 7. *$L \cap \partial H$ does not contain two parallel curves of length eight.*

Claim 8. *For two distinct tetrahedra H_1 and H_2, $L \cap \partial H_1$ and $L \cap \partial H_2$ do not both contain a curve of length eight.*

With Claims 5– 8 we can find a leaf of F_i that is transverse and intersects each tetrahedron in simple and normal arcs. If there is a normal arc of length eight in a tetrahedron, H, then there are no other normal arcs of length eight, either in H or in any other tetrahedron. There are no normal curves of length greater than eight. The only other normal curves are of length three or four, by Claims 1 and 2. This tells us that L intersects ∂H for each tetrahedron H in the triangulation Δ in simple curves and normal curves of length three or four and possibly one of length eight in only one tetrahedron.

Compress the simple curves and the annuli contained in a tetrahedron to get a collection of normal surfaces and possibly one almost normal surface. Recall that there are no almost normal 2-spheres in M_i with exceptional piece an octagon, so this almost normal surface is not a 2-sphere. We can consider the leaf L to be this collection of surfaces with tubes attached to the surfaces. These tubes do not intersect the one-skeleton and the one-skeleton does not run through the tubes.

Claim 9. *L cannot be one normal 2-sphere with tubes attached to it.*

Proof of Claim 9. Recall that the only normal 2-spheres in M_i are parallel to the one boundary normal 2-sphere, if one exists. If there is no normal boundary 2-sphere, then there are no normal 2-spheres. Assume L is one normal 2-sphere with any number of tubes. Since L is the leaf that we found in Claim 5, L has an upper and a lower disk. Consider the upper disk (the disk to the side with the boundary normal 2-sphere). The boundary of this disk must run along L and the one-skeleton. See Figure 42.

The arcs of the one-skeleton above the normal 2-sphere are all parallel to a straight arc, α. Therefore, above the 2-sphere the one-skeleton goes from the boundary normal 2-sphere to L without intersecting the 2-sphere more than once. See Figure 43.

There are no arcs inside the tubes since they represent compressions of simple curves in the faces of the tetrahedra or of tubes inside a tetrahedron. There is no arc of the one-skeleton above L that intersects L twice. This contradicts the fact that L has an upper disk.   □
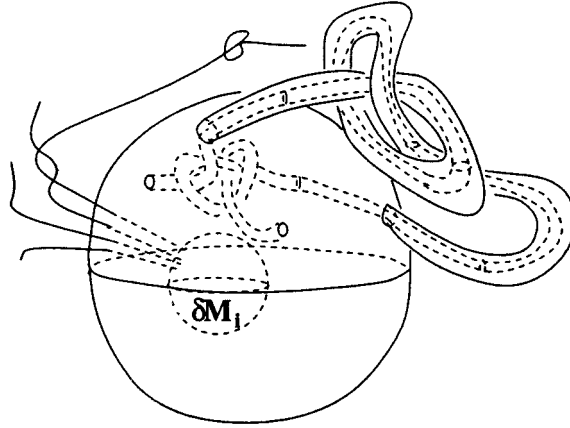
FIGURE 43. There is no upper disk for L

Claim 10. *L cannot consist of two normal 2-spheres, perhaps along with other normal surfaces, with tubes connecting them.*

Proof of Claim 10. Suppose L is thus constructed and call the outermost normal 2-spheres S_1 and S_2, where S_1 is closest to the boundary 2-sphere, ∂M_i. Here, S_1 and S_2 are parallel to ∂M_i. Since ∂M_i and S_2 cobound a 2-sphere cross I, all of the arcs in the one-skeleton are parallel in this region. Call one of these arcs α and the other arcs $\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n$. Let D_i be a parallel disk between α and α_i. D_i has boundary consisting of an arc on ∂M_i, an arc on S_2, call it δ_i, and the two arcs α and α_i. See Figure 44. We can pick the parallel disks, D_i, to be disjoint from each other. Thus, δ_i for $i = 1, 2, ..., n$ are disjoint from each other. Isotope α_i through D_i to a pushoff of the arc $\alpha \cup \delta_i$ such that the new arc has the same endpoint on S_2 as α_i. See Figure 44. Thus, the one-skeleton connecting to α_i inside S_2 is not affected by this isotopy.

Between ∂M_i and S_1, we will call the subarc of α, α'. The rest of α will be called α''.

Pick another path, β', that is homotopic rel endpoints to α' and is strictly decreasing; that is, it intersects each leaf of the foliation at most once.

How to pick this path: First we pick a path that has the same endpoints as α' and is strictly decreasing; call it ψ'. We can pick it so that close to ∂M_i it agrees with α'. $\psi'\alpha'^{-1}$ is homotopic to a path on L, call it ψ, since L is isotopic to the Heegaard surface. Let β' be equal to a push off of $\psi^{-1}\psi'$ so that the paths are intersecting each leaf of the foliation at most once. This adapts the strictly decreasing path, ψ', so that it is homotopic to α'.

There is an upper disk for L that runs along one arc of the one-skeleton and L. Label the intersection of this disk with L, γ. Note that the isotopy we have done so far involved only the parallel arcs in the 2-sphere cross I. This will not affect the upper disk since the upper disk cannot appear in that region.

We will find a path, call it θ, on L from the point where β' intersects L to the endpoint of α on L.. Furthermore, we will pick θ so that it doesn't intersect γ (except possibly at the endpoint) and such that θ is homotopic to α'' in M_i.
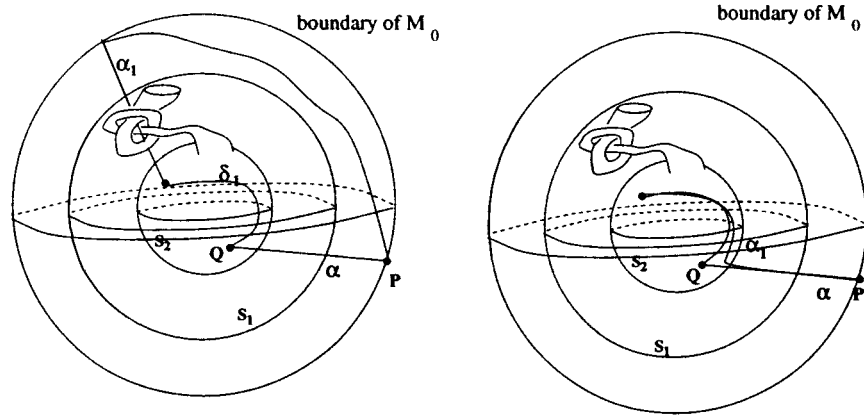
FIGURE 44. Isotoping α_i to α
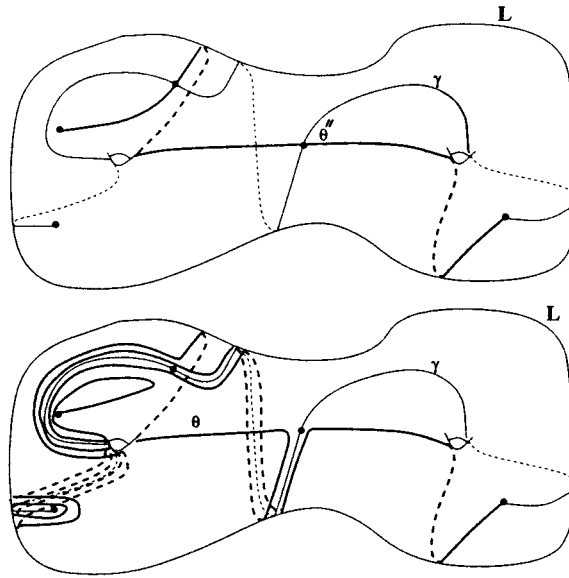
FIGURE 45. Finding θ

Note that α'' is contained in a compression body. Since L is isotopic to the Heegaard surface, there is a path, call it θ'', on L that is homotopic rel endpoints to α''. We can assume that θ'' has no self-intersections.

To build θ: Let θ coincide with θ'' until it approaches a point where γ and θ'' intersect; then θ will run parallel to γ, wind around the free endpoint of γ, and run parallel to γ back to a neighborhood of the point of intersection. Continue on θ'' repeating the above at each point of intersection. See Figure 45. θ is an embedded path avoiding γ that is homotopic rel endpoints to α''.

We can push θ above L slightly in the complement of the one-skeleton. This is possible since L is not tangent to the one-skeleton. We will do this so that the endpoints of θ are fixed. Call this new path, β''. Let β be the union of β' and β''.
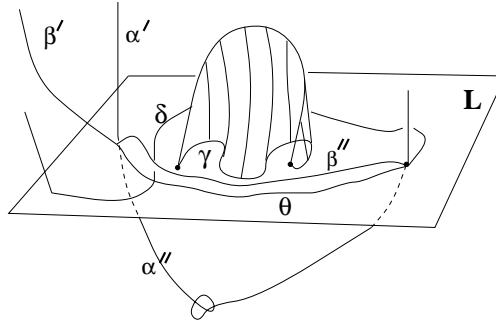
FIGURE 46. Creating β

At this point, β is a path from one endpoint of α to the other that agrees with α close to the endpoint of α on ∂M_i and that is homotopic rel endpoints to α. See Figure 46.

Let Γ consist of all arcs of the one-skeleton that are inside S_2 (inside means to the opposite side as ∂M_i) union the δ_i arcs that are on S_2. Now we can apply the unknotting lemma, [7], Proposition 7, to find a homeomorphism of M_i, isotopic to the identity (rel boundary) that fixes Γ and takes α to β.

A *product neighborhood* of a leaf of a foliation is a submanifold of M homeomorphic to $L \times [0, 1]$ such that the leaves of the foliation in this region are equal to $L \times \{\text{point}\}$ and $L = L \times \{1/2\}$.

β'' is in a product neighborhood of L. So we can isotope β'' slightly below L so that it has exactly one critical point. Note that δ_i for some i could possibly lie below β'' before we do this isotopy. If it does, then we will isotope it along, too. (Note: this is the reason we pushed θ above L first. We want to create a lower disk. So, we can't have δ_i puncturing the disk.) After the isotopy β has exactly one critical point below L. See Figure 47. α is taken to the new arc which we will also call β. α_i is taken to an arc close to $\beta \cup \delta_i$; call it β_i.

Isotope β_i so that it runs parallel to β and then along a pushoff of δ_i that lies slightly below L. We can do this so that it has exactly one critical point below L. See Figure 48.

We are considering the same foliation of M_i, but with the one-skeleton changed by an isotopy. Above L the width has not increased since the β arcs intersect each leaf at most once. Below L the α arcs must have at least one critical point since they intersect L twice. The β arcs have only one critical point. So we have not increased the number of critical points. We possibly have moved the critical points around. This could change the width. When a minimum moves above a maximum without creating any new critical points, then the width is decreased by use of Lemma 3. The β arcs are close to L. So, the first critical points we hit below L are this set of minima. The other arcs below L are the same as they were before the isotopy. The thinnest width is accomplished by putting the minima first. So, the width from the original position of the one-skeleton is either the same or thicker than the width of the position of the one-skeleton after this isotopy. If the width of the original position of the one-skeleton is bigger, then we have a contradiction to thin position, since we assumed the original foliation and position of the one-skeleton to be the
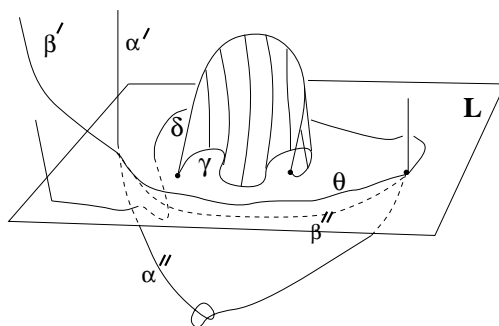
FIGURE 47. β is isotoped below L
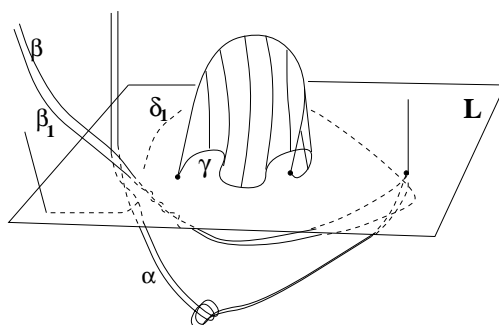
FIGURE 48. There is only one critical point on each β_i

thinnest. Assuming it is the same width, then we want to show that we can create a lower disk that is disjoint from an upper disk or intersects it in only one endpoint.

The only thing possibly preventing β from forming part of the boundary of a lower disk is if the parallel arcs β_i braid around β below L. Note that we picked β close enough to L so that the non-parallel arcs could not possibly wrap around β. If the parallel arcs wrap around β, then we can slide the braid up above L. The points of intersection with L are only fixed for the endpoints matching with the α arcs so we are free to move around the other endpoints to "unwind" the braiding below L and create braiding above L (i.e. we push the braid above L). See Figure 49. We can do this so that the arcs above L are still strictly decreasing.

Therefore, we can isotope β and its parallel arcs to arcs that bound disks below L. β'' is also picked so that the disk it cobounds does not intersect the upper disk, except possibly at one endpoint. Since the upper disk is disjoint from the lower disk (except possibly at one endpoint), we can isotope the one-skeleton across the upper and lower disks to get a thinner width. This contradicts thin position. □

When we compress the simple curves and annuli, we get a collection of normal surfaces and possibly one almost normal surface. There is at most one normal 2-sphere in the collection.

□
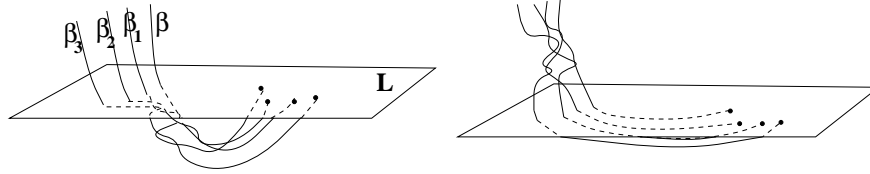
FIGURE 49. Isotoping the braiding of βs above L

Lemma 6. *Let S be the Heegaard surface for a strongly irreducible Heegaard splitting of a 3-manifold, and let S' be a surface (not a 2-sphere) obtained from S by compressing S to one side. Then S' is incompressible to the other side.*

See [2].

We are thinking of the leaf, L, as the collection of normal surfaces possibly along with an almost normal surface connected by tubes. If none of the normal surfaces is a 2-sphere, then the meridian disks of the tubes are all compressing disks for the leaf L. Since L (after one compression) is incompressible to the other side by Lemma 6, the tubes must all compress to the same side.

Case 1. *L is a collection of normal surfaces connected by tubes.*

Claim 11. *There must be a normal 2-sphere in the collection of normal surfaces, and the meridian disk of the tube connecting the normal 2-sphere to the rest of the normal surfaces must be to the opposite side of L as the meridian disks of the other tubes (there must be at least one other tube).*

Proof of Claim 11. Assume all tubes have meridian disks to the same side of L. Thus, the tubes do not run through each other. L is a thick leaf and has both an upper and a lower disk. Consider a meridian disk of a tube and the intersections of the upper (lower) disk with the meridian disk. Given an innermost circle of intersection, we can compress it. Continue to compress all innermost circles of intersection until there are none left. Given an outermost arc of intersection, we can take two copies of the subdisk of the meridian disk that it bounds, one on either side of the meridian disk. Eliminate a neighborhood of the outermost arc on the upper (lower) disk and glue in the two copies of the subdisk. See Figure 50. After doing this for all outermost arcs of the intersection of the meridian disk with the upper (lower) disk we may have cut the upper (lower) disk into several disks. None of these disks intersects the meridian disk. No arcs of the one-skeleton run through the tubes. Thus, at least one of these disks is still an upper (lower) disk, since the part of the disk that runs over an arc of the one-skeleton was not affected by these surgeries. We can pick another meridian disk for another tube and again find an upper (lower) disk that does not intersect this meridian disk. When we have done this for all meridian disks, then we have found an upper (lower) disk that doesn't run over any meridian disks. Since all tubes have meridian disks to the same side of L, this implies that the disk can be isotoped so that it doesn't run over any tubes. Thus we have a ∂-compressing disk in the complement of the one-skeleton for a normal surface. This contradicts Claim 4.

So, there must be a tube with meridian disk on the opposite side of L as the other meridian disks. This tube must connect to a normal 2-sphere; otherwise we contradict strongly irreducibility. $\square$
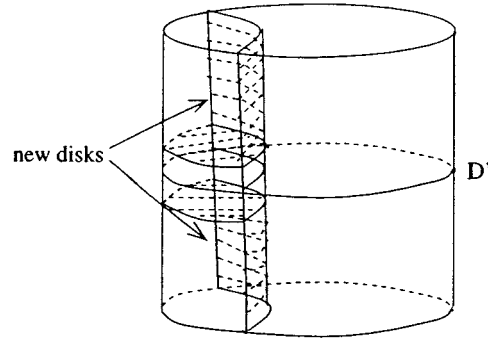
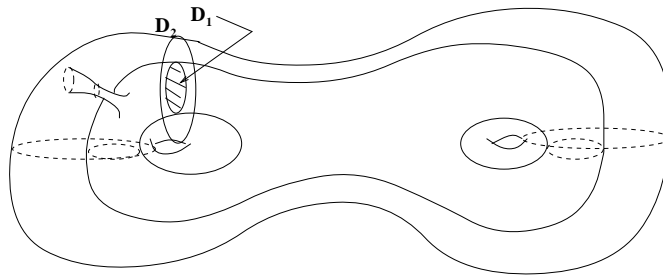FIGURE 50. Performing a one-surgery on the upper(lower) disk

FIGURE 51. This violates strongly irreducible

From Claim 10 there cannot be more than one normal 2-sphere in this collection. Note that L cannot be a normal genus g surface with a tube connecting it to a normal 2-sphere, because there would be only one tube.

A normal surface (not the 2-sphere), N_1, cannot be on the compressible side of another normal surface, N_2, by Lemma 6. See Figure 51.

Also, the tubes (except the one connecting the normal 2-sphere to the other normal surfaces) cannot be to the compressible side of any of the normal surfaces. This implies that the normal 2-sphere is to the compressible side of one of the normal surfaces, since the tube connecting it lies to the opposite side as the other tubes.

Cut M_i along the collection of normal surfaces (but not along the normal 2-sphere). Keep the connected component with boundary the normal surfaces we just cut along (possibly with additional boundary surfaces from before). Call this new submanifold of M M_{i+1}. The normal 2-sphere is not contained in M_{i+1}.

We will be reiterating Lemma 5 for M_{i+1}. We are guaranteed not to get Case 1 again, since there are no normal 2-spheres in M_{i+1}. To apply this lemma again we need to have a foliation of M_{i+1}. The normal surfaces we cut along come from compressing the leaf, which is isotopic to the Heegaard surface of M_i. So, the same Heegaard surface for M_i is a Heegaard surface for M_{i+1}. The compression body containing the normal surfaces we just cut along can be constructed by sewing in 2-handles to $L \times I$ where the compressions were made. We will consider the standard singular foliations of M_{i+1}. Consider the singular leaf that consists of the normal

surfaces we cut along to be the top leaf. Note that the boundary surfaces are all incompressible in M_{i+1} by Lemma 6.

Put the one-skeleton into thin position with respect to all foliations of this type. If there is no thick region, then apply Lemma 4. If there is a thick region, apply Lemma 5.

Case 2. *L is a collection of normal surfaces and one almost normal surface (not a 2-sphere) connected by tubes.*

If the almost normal surface is genus g, then we are done since this surface is isotopic to a (non-singular) leaf of the foliation which is isotopic to the Heegaard surface.

If the almost normal surface is not genus g, then continue. There is either a normal 2-sphere in the collection of normal surfaces or there isn't. By the same argument as above we know that the tubes and surfaces (not 2-spheres) cannot be contained in the compressible side of any of the other normal or almost normal surfaces, since the Heegaard splitting is strongly irreducible. Note that in the case where there is no thick leaf the resulting almost normal surface cannot be a 2-sphere, since there is at most one normal 2-sphere boundary of M_i.

Apply Lemma 1 to obtain a normal surface isotopic to the almost normal surface. Cut along the normal surfaces (possibly the normal 2-sphere, too) and keep the component that is to the incompressible side of the normal surfaces. Call this component M_{i+1}. Notice that the almost normal surface is to the compressible side of the normal surface obtained from it, since the isotopy to the normal surface was always to the incompressible side of the almost normal surface. Therefore, the almost normal surface is not in M_{i+1}. Refoliate M_{i+1} with a standard singular foliation. If the top leaf of M_i is contained in the part we cut off, then we will consider the leaf consisting of the normal surfaces that we cut along with the core of the compression body it lies in as the top leaf. After foliating, we can put the foliation in thin position with respect to the one-skeleton. Apply either Lemma 4 or Lemma 5. We cannot possibly get back L when we reiterate this since we have cut off the almost normal surface.

In both of the above cases we will obtain at least one more normal surface that is non-parallel to and disjoint from the boundary normal surfaces of M_i. The normal surface is non-parallel to the boundary surfaces since the normal surface either comes from a thick leaf or from an almost normal surface that was created by tubing together two boundary normal surfaces.

By an argument due to Haken, there are only a finite number of non-parallel, disjoint normal surfaces in M. This implies that we will only reapply the above lemmas a finite number of times, until we obtain the almost normal surface isotopic to the Heegaard surface.

6. Proof of Theorem 1 for bounded 3-manifolds

Let (M, W, W') be a 3-manifold M and a genus g Heegaard splitting of M into compression bodies W and W'. Let G be the Heegaard surface, ∂M be the boundary of M, and Δ be a triangulation of M. ∂M consists of faces of the tetrahedra in Δ, that are not identified to other faces. If there is a tetrahedron in Δ with more than one face in the boundary of M, then subdivide that tetrahedron until each tetrahedron in Δ has at most one face that is in ∂M. Push the boundary components slightly into M to create normal surfaces. Cut M along these normal

surfaces. We will keep the submanifold of M with boundary consisting of normal surfaces, where a collar of M is cut off. Call it M'.

As in the closed case, we can find a maximal collection of separating, normal, non-parallel 2-spheres, Σ, in M', using an algorithm due to Haken.

Since M is irreducible, the boundary components cannot be 2-spheres. There is a component of $M' - \Sigma$ with at least one boundary surface which is not a 2-sphere. We will call this component M_0.

Each 2-sphere boundary of M_0 bounds a 3-ball to the other side since M is irreducible. M_0 has a Heegaard splitting where W_0 and W'_0 are the compression bodies obtained from W and W' after first isotoping G so that the normal 2-sphere boundaries do not intersect the splitting surface. Call this new splitting surface G_0. The fact that the boundary surfaces are incompressible in M_0 follows from the strong irreducibility of the Heegaard splitting and from the assumption that the Heegaard splitting is non-trivial.

In Lemma 5, we need there to be no more than one normal 2-sphere boundary of M_0. Given any one of the boundary surfaces we can find an arc of the one-skeleton connecting it to another boundary surface. To see this, note that the one-skeleton is connected and that normal surfaces must intersect the one-skeleton. If we take a normal 2-sphere boundary of M_0 and find an arc of the one-skeleton connecting to another normal 2-sphere boundary, then we can use Lemma 2 in [17] to say that the component is a punctured 3-ball. If there is no such arc, then that means that there is an arc of the one-skeleton from the normal 2-sphere to a normal surface of genus n. Take an annulus that is parallel to this arc, connecting the two normal surfaces. This creates an almost normal surface of genus n. Since the boundary surfaces are incompressible in M, the normal genus n surface with a normal 2-sphere tubed on is incompressible to one side in M' and the meridian disk of the tube is on the opposite side of this almost normal surface. We can apply Lemma 1 to isotope the almost normal surface to a normal surface while always pushing away from the normal 2-sphere. Cut along this normal surface and keep the side not containing the normal 2-sphere we isotoped away from. The resulting boundary surface is isotopic to the first boundary surface. We continue to do this for all normal 2-spheres, until we no longer have any normal 2-spheres in the component. We will refer to this component as M_0.

There are no vertices in M_0 since there is always a normal 2-sphere around a vertex in the interior of a 3-manifold. There cannot be an almost normal 2-sphere in M_0 since there are no normal 2-spheres. If there was an almost normal 2-sphere with exceptional piece an octagon, then we can isotope the octagon as in Lemma 1 to get a normal 2-sphere.

We can foliate (M_0, W_0, W'_0) with a standard singular foliation. Call the collection of all standard singular foliations of M_0, $\mathcal{F}_{G_0}$. Put the one-skeleton, Δ_1, into thin position with respect to $\mathcal{F}_{G_0}$. Let F_0 be the foliation in $\mathcal{F}_{G_0}$ that realizes the width of Δ_1. $(M_0, \Delta_1, W_0, W'_0, F_0)$ has the necessary properties in order to apply Lemmas 4 and 5 to it.

By the same reasons as in the closed case, we cannot apply the lemmas an infinite number of times. Therefore, there is an almost normal genus g surface isotopic to the Heegaard surface.

Note that when M is a surface cross I, if the vertices of the triangulation are not all on the two boundary surfaces we can use the above ideas. Push the boundary to normal surfaces. Cut along a maximal collection of normal 2-spheres. Consider

M_0 to be a component that contains one of the normal genus g surfaces. Since there is a vertex in M', M_0 has a normal 2-sphere boundary as well. There must be an arc of the one-skeleton from the normal 2-sphere to one of the normal genus g surfaces. Attach an annulus parallel to this arc. The result is an almost normal genus g surface isotopic to the Heegaard splitting surface. However, if the vertices are all on the boundary, then it is possible that the boundary normal surfaces of M' are parallel. If they aren't, then there must be a thick region and we can use Lemma 5 to say that there must be an almost normal genus g surface isotopic to the Heegaard surface.

References

[1] J. Birman and H. Hilden, *The homeomorphism problem for the 3-sphere*, Bull. AMS **79** (1973), no. 5, 1006–1009. MR **47:**7726

[2] A. Casson and C. Gordon, *Reducing Heegaard splittings*, Topology and its Applications **27** (1987). MR **89c:**57020

[3] C. Frohman, *The topological uniqueness of triply periodic minimal surfaces in R^3*, J. Diff. Geom. **31** (1990), 277–283. MR **92e:**53008

[4] D. Gabai, *Foliations and the topology of 3-manifolds III*, J. Diff. Geom. **26** (1987). MR **89a:**57014b

[5] W. Haken, *Theorie der normalflachen*, Acta Math. **105** (1961), 245–375. MR **25:**4519a

[6] ———, *Some results on surfaces in 3-manifolds*, Studies in Modern Topology, M.A.A., Prentice-Hall (1968), 34–98. MR **36:**7118

[7] J. Hass and A. Thompson, *Neon bulbs and the unknotting of arcs in manifolds*, J. Knot Theory Ramifications **6** (1997), 235–242. MR **99c:**57043

[8] G. Hemion, *The classification of knots and 3-dimensional spaces*, Oxford University Press, 1992. MR **94g:**57015

[9] J. Hempel, *3-manifolds*, Princeton University Press, 1976. MR **54:**3702

[10] W. Jaco, *Lectures on three-manifold topology*, A.M.S., Regional conference series in math, no. 43, 1980. MR **81k:**57009

[11] W. Jaco and U. Oertel, *An algorithm to decide if a 3-manifold is a Haken manifold*, Topology **23** (1984), no. 2, 195–209. MR **85j:**57014

[12] W. Jaco and J.H. Rubinstein, *A piecewise linear theory of minimal surfaces in 3-manifolds*, J. Diff. Geom. **27** (1988), 493–524.

[13] S.V. Matveev, *Algorithms for the recognition of the three-dimensional sphere (after A. Thompson)*, Mat. Sb. **186** (1995), 69–84. MR **96g:**57016

[14] J. Pitts and J.H. Rubinstein, *Equivariant minimax and minimal surfaces in geometric three–manifolds*, Bull. Amer. Math. Soc. **19** (1988), 303–309. MR **90a:**53014

[15] J.H. Rubinstein, *Polyhedral minimal surfaces, Heegaard splittings and decision problems for 3-dimensional manifolds*, Proceedings of the Georgia Topology Conference, AMS/IP Stud. Adv. Math., vol. 21, Amer. Math. Soc., Providence, RI, 1997, pp. 1–20. MR **98f:**57030

[16] M. Scharlemann and A. Thompson, *Heegaard splittings of $(surface) \times I$ are standard*, Math. Ann. **295** (1993), 549–564. MR **94b:**57020

[17] A. Thompson, *Thin position and the recognition problem for the 3-sphere*, Math. Research Letter **1** (1994), 613–630. MR **95k:**57015

[18] I. Volodin, V. Kusnezov, and A. Fomenko, *On a problem of algorithmic recognition of the standard 3-sphere*, Usp. Mat. Nauk **29** (1974), no. 5, 71–168, Russian-English Translation in Russian Math. Survey. MR **53:**9219

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, DAVIS, CALIFORNIA 95616
Current address: Department of Mathematics, University of Texas, Austin, Texas 78712
E-mail address: `stocking@math.utexas.edu`