LOCALISING DEHN’S LEMMA AND THE LOOP THEOREM IN 3-MANIFOLDS

I. R AITCHISON AND J. HYAM RUBINSTEIN

Abstract. In this paper, we give a new proof of Dehn’s lemma and the loop theorem. This is a fundamental tool in the topology of 3-manifolds. Dehn’s lemma was originally formulated by Dehn in [?], where an incorrect proof was given. A proof was finally given by Papakyriakopolous in his famous 1957 paper [?] where the fundamental idea of towers of coverings was introduced. This was later extended to the loop theorem [?], and the version used most frequently was given by Stallings [?].

In [?] we showed that hierarchies for Haken 3-manifolds could be understood by a ‘local version’ of Dehn’s lemma and the loop theorem. Developing the idea further enables us here to give a new proof of the classical theorems of Papakyriakopolous, which do not use towers of coverings. A similar result was obtained by Johannson [?], with the added assumption that the 3-manifold in question was Haken. Our approach means that no extra hypotheses are necessary. Our method uses the concept of boundary patterns of hierarchies, as developed by Johannson in [?]. Marc Lackenby has independently produced a very similar proof in his lecture notes in [?].

Note that the more difficult sphere theorem ([?], [?], [?]) can then be deduced using Dehn’s lemma and the loop theorem, plus the PL theory of minimal surfaces in [?]. Other applications like the important result of [?] that a covering of an irreducible 3-manifold is irreducible, then follow also by PL minimal surface theory.

1. Introduction

The key idea in this paper is to introduce a ‘geometric’ version of the hierarchies of Haken and Waldhausen ([?], [?], [?], [?], [?]). We will work with a ‘very short hierarchy’, which is related to the short hierarchies discussed in [?]. These are useful for algorithms in the study of recognition and classical decomposition theory of 3-manifolds [?]. David Gabai has independently noted the existence of such hierarchies. We show that all orientable Haken 3-manifolds and non-orientable 3-manifolds without 2-sided Klein bottles contain such hierarchies.

Our aim is to give a new proof of Dehn’s lemma and the loop theorem, which uses two basic tools of geometric topology of 3-manifolds, but no covering space theory. One technique is the method of normal surfaces, developed by Haken [?] after the earlier work

Date: February 23, 2004.

1991 Mathematics Subject Classification. Primary 57M25, 57N10.

Key words and phrases. Hierarchies, Dehn’s lemma, loop theorem.

We would like to thank Klaus Johannson and Peter Scott for helpful comments on early versions of this paper. The authors’ research was partially supported by the Australian Research Council.
of Kneser [?] and the other method uses the concept of a simple hierarchy, introduced by Waldhausen [?] in his solution of the word problem in the fundamental groups of Haken 3-manifolds. We acknowledge the influence of a beautiful paper of Johannson [?], which explores this direction in the context of Haken 3-manifolds only. Johannson developed the theory of hierarchies and boundary patterns in [?]. We also note that in [?], these ideas are studied in the context of a theory of Haken 4-manifolds.

2. Geometric hierarchies and very short hierarchies

Hierarchies give fundamental ways of decomposing a large class of 3-manifolds into 3-cells. They are also used in Thurston’s beautiful theory of hyperbolic and geometric structures on 3-manifolds [?]. Our approach here is to start with embedded surfaces which cannot be ‘geometrically compressed’ i.e do not have embedded compressing disks. All surfaces and 3-manifolds will be assumed connected.

Definition 2.1. Assume that $M$ is a compact 3-manifold and $S$ is a properly embedded compact 2-sided surface in $M$ which is not a disk, 2-sphere or projective plane. We will say that $S$ is geometrically incompressible in $M$ if given any embedded disk $D$ in $M$ with $D \cap S = \partial D$, then $\partial D$ is a contractible loop in $S$.

As usual, we also need a relative version of geometrical incompressibility.

Definition 2.2. Assume that $M$ is a compact 3-manifold with non-empty geometrically incompressible boundary components. Suppose $S$ is a properly embedded compact 2-sided surface in $M$ which is not a disk and has non-empty boundary in $\partial M$. We will say that $S$ is boundary geometrically incompressible in $M$ if given any embedded disk $D$ in $M$ with $\partial D = \alpha \cup \beta$, where $\alpha \subset S$ and $\beta \subset \partial M$, with $\text{int}D$ disjoint from $S$, then there is an arc $\tau$ in $\partial S$ with the same endpoints as $\alpha$ and $\beta$ so that there is a subdisk $D_0$ of $S$ with $\partial D_0 = \tau \cup \alpha$.

A compact 3-manifold $M$ will be called irreducible if any embedded 2-sphere bounds a 3-ball and $P^2$-irreducible if in addition, there are no embedded 2-sided projective planes. A compressing disk for $\partial M$, where $M$ is a compact 3-manifold with non-empty boundary, is a properly embedded disk $D$ with $\partial D$ being a non-contractible loop in $\partial M$.

We can now define the key concept in our technique, i.e a geometrical hierarchy for a 3-manifold.

Definition 2.3. Assume that $M$ is $P^2$-irreducible and either closed or compact with geometrically incompressible boundary components. A geometrical hierarchy $\mathcal{H}$ for $M$ is a collection of embedded 2-sided surfaces $S_1, S_2, ..., S_k$ with the property that each $S_j$ is either closed or has boundary in $S_1 \cup S_2 \cup ... \cup S_{j-1}$. Moreover if $M_{j-1}$ denotes the result of cutting
DEHN’S LEMMA

$M$ open along $S_1 \cup S_2 \cup \ldots \cup S_{j-1}$ then either $S_j$ is boundary geometrically incompressible in $M_{j-1}$ or $S_j$ is a compressing disk for $\partial M_{j-1}$. Finally $M_k$ is a collection of 3-cells.

A hierarchy can be chosen so that all the closed surfaces occur first, then the surfaces with boundary contained in the closed surfaces and so on. We can ‘group’ the surfaces of the hierarchy together into maximal sets of disjoint surfaces. The length of the hierarchy is then the number of sets of surfaces making up such a group. In [?], it is shown that a hierarchy can always be chosen with length 4 for a Haken 3-manifold. We will prove below that in fact a length 3 hierarchy can always be constructed for such 3-manifolds, so long as there are no (geometrically) incompressible 2-sided Klein bottles. (We call this a very short hierarchy). Since Dehn’s lemma and the loop theorem are needed to define a Haken 3-manifold, the proof that such hierarchies exist is embedded in our proof of Dehn’s lemma and the loop theorem. Assuming these facts, the construction of such hierarchies is quite straightforward. Our hierarchies are ‘only’ incompressible in the geometrical sense - the main part of the proof will be first the existence of such geometrically incompressible hierarchies and second that they are actually incompressible in the classical sense.

**Definition 2.4.** Assume that $M$ is a compact 3-manifold and $S$ is a properly embedded compact 2-sided surface in $M$ which is not a disk, 2-sphere or projective plane. We will say that $S$ is algebraically incompressible in $M$ if the induced map $\pi_1(S) \to \pi_1(M)$ is one-to-one.

**Definition 2.5.** Assume that $M$ is a compact 3-manifold with non-empty algebraically incompressible boundary components. Suppose $S$ is a properly embedded compact 2-sided surface in $M$ which is not a disk and has non-empty boundary in $\partial M$. We will say that $S$ is boundary algebraically incompressible in $M$ if $S$ is algebraically incompressible and the induced map $\pi_1(S, \partial S) \to \pi_1(M, \partial M)$ is one-to-one.

Note that the usual notions of incompressibility and boundary incompressibility are what we are calling algebraically incompressible and boundary algebraically incompressible. We do this to emphasize the fact that there are two notions of these concepts, namely whether one deals with embedded or singular disks. Then Dehn’s lemma and the loop theorem can be viewed as showing that the two concepts are in fact identical. An ordinary hierarchy is then a sequence of surfaces $\mathcal{H}$ which are all algebraically incompressible and boundary algebraically incompressible in the manifold obtained by cutting open along the previous surface in the sequence. Also cutting open along $\mathcal{H}$ gives a collection of 3-cells. Finally $M$ is then a Haken 3-manifold if it is $P^2$-irreducible and has such a hierarchy. Note that if $M$ is compact with non-empty boundary, then the components of the boundary must be algebraically incompressible and can be chosen as part of the hierarchy. If $M$ is closed, then it must contain some closed 2-sided embedded algebraically incompressible surfaces to be Haken.
3. Very short hierarchies

Our idea is to show first that a geometrically incompressible boundary component of a $P^2$-irreducible 3-manifold $M$ can be extended to a geometrically incompressible hierarchy for $M$. It is then easy to deal with the case of a general non-$P^2$-irreducible 3-manifold in the standard way by first splitting $M$ up along a maximal disjoint non-parallel collection of embedded essential 2-spheres (not bounding 3-balls) and embedded 2-sided projective planes.

**Theorem 3.1.** Suppose that $M$ is a compact $P^2$-irreducible 3-manifold with non-empty $\partial M$. Also assume that each component of $\partial M$ is geometrically incompressible in $M$ and that $M$ has no embedded geometrically incompressible 2-sided Klein bottles. Then a very short hierarchy (in the geometrically incompressible sense) can be found for $M$, which contains all the components of $\partial M$ amongst its surfaces. Also if $M$ has geometrically incompressible 2-sided Klein bottles, then $M$ has a short geometrically incompressible hierarchy.

**Proof.** We modify the procedure of Jaco [?] for finding a ‘short’ hierarchy. Choose a maximal family of disjoint embedded 2-sided closed surfaces $S_1, S_2, \ldots, S_m$ in $M$, which are not parallel and are all geometrically incompressible, including all the components of $\partial M$. (The trivial case when $M$ is a product $S \times I$ is an exception).

Notice that to show such a family can be found, uses Kneser’s technique ([?]) as extended by Haken [?]. Normal surface theory only requires that the surfaces $S_i$ are geometrically incompressible, not algebraically incompressible. We will prove that each $S_i$ can be modified to be normal and the number of disjoint non-parallel embedded surfaces is bounded by counting the rank of $H_1(M, \mathbb{Z}_2)$ and adding in the maximum number of non product regions in a fixed triangulation of $M$, which can occur in the complement of all the surfaces. (This is essentially Kneser’s theorem [?]). So we conclude that such a maximal family can be constructed.

We now give some details of the modification needed of the Kneser-Haken method applied to any geometrically incompressible embedded surface $S$. First of all, we can assume that $S$ meets the faces of the triangulation in arcs only. This follows since loops of intersection can be removed by a simple disk swap and isotopy, using geometrical incompressibility of $S$ to observe that such an innermost loop bounds a disk on $S$. (This also uses the assumption that $M$ is irreducible). Moreover arcs with both ends on the same edge can be eliminated by an isotopy of $S$, which reduces the number of intersections of $S$ with the edges. Next, consider the intersection of $S$ with some tetrahedron. Choose an innermost disk on the boundary of the tetrahedron, which is bounded by a loop on $S$. This is a compression disk for $S$ unless either the corresponding component of intersection of $S$ with the tetrahedron is a disk, or there is a component of $S$ outside the tetrahedron which is a disk. In the latter
case, we can perform a disk swap and isotopy to reduce the number of intersections of $S$ with the triangulation.

Next we have to eliminate the possibility of a disk $D$ of intersection of $S$ with the tetrahedron, which has more than four spanning arcs in the faces. In this case, clearly there is at least one face with at least two spanning arcs. But then a path between these arcs has common endpoints with an arc on $S$ and these arcs cobound an embedded 2-gon on the union of $D$ and the boundary of the tetrahedron. This 2-gon can be pushed off the union of $D$ and the boundary of the tetrahedron and we can use it to isotope $S$ to decrease the number of intersections with the edges, exactly as in the Kneser-Haken method.

Finally we need to prove that minimising the intersection of a geometrically incompressible surface with the triangulation as above yields a normal surface. The only other possibility is that the surface is isotoped inside a single tetrahedron, i.e a ball. But in this case, we can slice the surface by a family of disks, foliating the tetrahedron, as in the classical argument of Alexander [?]. We may assume that the disks are in general position relative to the surface, so that there are finitely many critical levels, where the disk has a single point of tangency with the surface. It is easy to see that if the surface is not a 2-sphere or projective plane, then at some transverse levels, the disks will have non-contractible loops of intersection with the surface, which yield non-trivial compressing disks. This completes the discussion of why the Kneser finiteness result applies in our context.

The main point to emphasize is that none of the above steps requires Dehn’s lemma and the loop theorem.

Suppose next that $M$ is cut open along all the surfaces $S_1, S_2, \ldots, S_m$. Let $R_1, R_2, \ldots, R_n$ be the complementary regions formed. For each such a region $R_i$, choose a collection of geometrically incompressible and boundary incompressible ‘spanning’ surfaces. We want these surfaces in $R_i$ to have the properties that every component of $\partial R_i$ is intersected by one of these surfaces. This can be done by carefully applying the procedure of Stallings [?] by mapping each region $R_i$ to a circle, pulling back a transverse point and surgering the map. Notice that again the geometric versions of incompressibility and boundary incompressibility suffice, so that we end up with proper surfaces $S^i_q, q = 1, 2, \ldots, n_i$ in $R_i$, which have maximal Euler characteristic and so have no embedded compressing or boundary compressing disks.

Now we claim that by a judicious choice of the map of $R_i$ to a circle, the desired property of these surfaces can be ensured. Such a map corresponds to a choice of a cohomology class $\alpha$ in $H^1(R_i, \mathbb{Z})$. Now for every component $T_j$ of $\partial R_i$ we can find such a cohomology class $\alpha_j$ with the property that $\alpha_j([C_j]) = 1$ for some non-separating simple closed curve $C_j$ on $T_j$, by Alexander duality, in case that $M$ and $R_i$ are orientable. It is simple to choose a linear combination $\Sigma n_r \alpha_r$ with the property that $(\Sigma n_r \alpha_r)[C_j] \neq 0$ for all $j$. Following through the Stallings argument then shows that the resulting surfaces $S^i_q, q = 1, 2, \ldots, n_i$ meet all the components $T_j$ of $\partial R_i$ as required.
If $M$ is non-orientable, an additional argument is needed, assuming that there are no geometrically incompressible 2-sided Klein bottles. Let $R$ denote one of the regions obtained by splitting $M$ open along a maximal collection of closed 2-sided geometrically incompressible non-parallel surfaces in $M$ and let $\tilde{R}$ denote the orientable double covering of $R$. If $T$ is some component of $\partial R$, then by assumption $T$ is not a Klein bottle, 2-sphere or projective plane. We claim there is still a cohomology class $\alpha$ with $\alpha([C]) = 1$, where $C$ is some non-separating 2-sided simple closed curve on $T$, as for the orientable case.

Let $g : \tilde{R} \to \tilde{R}$ be the covering involution. There is an induced action $g_*$ on $H_1(\tilde{T}, \mathbb{Q})$, which preserves the kernel $K$ of the map to $H_1(\tilde{R}, \mathbb{Q})$ and also preserves the intersection numbers of homology classes on the surface $\tilde{T}$. Note that $\tilde{T}$ has a single component if $T$ is non-orientable or two components if $T$ is orientable, where $\tilde{T}$ denotes the preimage of $T$ in $\tilde{R}$. In the first case, since $T$ has non-orientable genus $g \geq 3$, $\tilde{T}$ has orientable genus $g - 1 \geq 2$. By Alexander duality in $\tilde{R}$, it follows that the rank of $K$ is at most $g - 1$ and there cannot be classes in $K$ with non-zero intersection number. Moreover, notice that $g_*$ splits $H_1(\tilde{T}, \mathbb{Q})$ into subspaces of equal rank, which are the $+1$ and $-1$ eigenspaces. The $+1$ eigenspace cannot be contained in $K$ or we would violate the property that $K$ has no pair of classes with non-zero intersection number. Projecting to $H_1(R, \mathbb{Q})$, we conclude that there are classes of infinite order in the image of $H_1(T, \mathbb{Z})$ in $H_1(R, \mathbb{Z})$. Hence it is elementary to find the desired cohomology class $\alpha$ of $H^1(R, \mathbb{Z})$, so that $\alpha[C] = 1$ for a non-separating 2-sided loop $C$ in $T$. (Notice that if $T$ was a Klein bottle, this argument might fail, since the $+1$ eigenspace might be contained in $K$, which has rank 1. So there is no violation of the intersection number property.)

If $T$ is orientable, then there are two components of $\tilde{T}$ in $\tilde{R}$. We can use a very similar technique to the previous paragraph. If $K$ again denotes the kernel of the map from $H_1(\tilde{T}, \mathbb{Q})$ to $H_1(\tilde{R}, \mathbb{Q})$, it follows that the rank of $K$ is at most $2g$, where $g$ is the genus of $T$ and there cannot be classes in $K$ with non-zero intersection number. As before, the covering involution $g$, which interchanges the two components of $\tilde{T}$, has equal rank $\pm 1$ eigenspaces, in the induced action on first homology. We deduce that the $+1$ eigenspace cannot be contained in $K$, as this would violate the property that there cannot be classes in $K$ with non-zero intersection number. Projecting to $T$ in $R$, we find the required cohomology class exactly as before.

Finally we cut open each region $R_i$ along the proper surfaces $S^i_q$, $q = 1, 2, ... n_i$, to give a (possibly disconnected) region $R^i_j$. Exactly as in [?], it follows that each component of $R^i_j$ is a handlebody (which could be orientable or non-orientable). The reason is that performing a maximal collection of compressions along embedded disks must cut each such component into a collection of 3-cells, since any closed surface produced which is not a 2-sphere or projective plane, would be 2-sided and geometrically incompressible, but not parallel to one of the original family of such surfaces $S_1, S_2, ..., S_m$. This follows by the property that there
are surfaces of $S^q_i, q = 1, 2, \ldots, n_i$ meeting every component of $\partial R_i$ so that such a new surface being parallel into the boundary of $R_i$ is impossible. So, finding a new boundary surface, which is not a 2-sphere or projective plane, contradicts our choice of the original family as being maximal.

Consequently the hierarchy can be completed by adding compressing disks. Therefore cutting $M$ open along all the surfaces gives a collection of 3-cells. This finishes the construction of the very short geometrically incompressible and boundary incompressible hierarchy containing all the surfaces of $\partial M$. In the presence of 2-sided Klein bottles, the argument of [?] gives a short hierarchy. □

For convenience, we need to extend this construction to 3-manifolds which are not necessarily $P^2$-irreducible.

**Definition 3.2.** Let $M$ be a compact or closed 3-manifold, where $M$ is not necessarily $P^2$-irreducible and all the components of $\partial M$ are neither 2-spheres nor projective planes and are geometrically incompressible. An extended geometrical hierarchy for $M$ is a collection $H_i, 1 \leq i \leq k$ of sets of surfaces, satisfying all the conditions of a geometrical hierarchy, except that the closed surfaces can include essential 2-spheres and 2-sided projective planes. Also cutting $M$ open along all the surfaces produces punctured 3-cells, punctured real or fake copies of $RP^2 \times I$ and punctured orientable 3-manifolds which are closed connected summands of $M$ without any embedded geometrically incompressible surfaces.

**Remarks** Notice that any non orientable closed connected summand of $M$ must have embedded geometrically incompressible surfaces which are non separating, since as is well known, such a manifold has infinite first homology. So the only ‘part’ of $M$ which cannot be decomposed by such a hierarchy are closed orientable connected summands without incompressible surfaces. A fake copy of $RP^2 \times I$ is a manifold with two boundary components which are projective planes and which is homotopy equivalent to $RP^2 \times I$.

**Theorem 3.3.** Assume that $M$ is compact and its non-empty boundary has every component which is neither a 2-sphere nor a projective plane and is geometrically incompressible. Then there is an extended geometrical hierarchy $H_i, 1 \leq i \leq k$ of length 3 for $M$, assuming that $M$ has no 2-sided Klein bottles. If there are such Klein bottles, then there is an extended geometrical hierarchy of length 4 for $M$.

**Proof.** The idea is very similar to the proof of the preceding theorem. The main point is that we can first decompose $M$ along a maximal collection of disjoint non-parallel essential 2-spheres (i.e. which do not bound 3-cells) and 2-sided projective planes, by Kneser’s theorem [?]. This splits $M$ into a number of pieces. Each such piece is compact with boundary consisting of some components of $\partial M$ and some 2-sided projective planes and 2-spheres. We are only interested in the pieces with at least one component of $\partial M$ in their boundary or
which have embedded geometrically incompressible surfaces. For the latter, we can further split them along geometrically incompressible surfaces into smaller pieces. To summarise, we will denote by $M_j$, $1 \leq j \leq k$, all the pieces found with some geometrically incompressible surfaces in their boundaries. Any pieces without incompressible surfaces can be left alone.

For such a 3-manifold piece $M_j$, following the procedure of the previous result gives a geometrically incompressible set of surfaces of length 3 or 4 cutting up $M_j$ into punctured 3-cells and punctured real or fake copies of $RP^2 \times I$. (Any punctured fake 3-cell will be included as a connected summand of $M$ with no geometrically incompressible surfaces). To see this, note that choosing a cohomology class giving properly embedded boundary geometrically incompressible surfaces in $M_j$ meeting all the components of $\partial M$ in $\partial M_j$, means that after cutting $M_j$ open along these surfaces, we must have punctured handlebodies. Note in the non-orientable case, a punctured handlebody can have boundary surfaces which are 2-spheres, projective planes and a single closed non-orientable surface which is not a projective plane. (We can start with two copies of $RP^2 \times I$ and attach a 1-handle between these, as an example). Consequently, cutting along compressing disks gives punctured 3-cells and punctured real or fake copies of $RP^2 \times I$. This completes the proof. □

4. Dehn’s lemma and the loop theorem

We start with the case of a closed 3-manifold and indicate the changes needed to deal with general compact 3-manifolds later.

**Theorem 4.1.** Suppose that $M$ is a closed 3-manifold and $S$ is an embedded 2-sided closed surface in $M$. Assume that $S$ is not a 2-sphere nor a real projective plane and the induced map from $\pi_1(S)$ to $\pi_1(M)$ is not one-to-one. Then there is an embedded disk $D$ in $M$ with $D \cap S = C$, where $C = \partial D$ is non-contractible in $S$.

**Proof.** Suppose that no such embedded disk $D$ can be found. The first step is to extend $S$ to a geometrically incompressible hierarchy for $M$. We will establish a property which will play a similar role to the ‘simplicity’ condition of Waldhausen [?], for our hierarchy. Rather than following the ideas in [?] or [?], we only use the closed and spanning surfaces of the hierarchy and establish that the compressing disks for the handlebodies have to meet the boundary pattern at least 4 times, whether the disks are embedded or singular. We will go through the details, since the existence of a very short hierarchy gives an interesting different and quite simple argument.

First of all, we can split $M$ along $S$ into one or two compact 3-manifolds with boundary. Each piece then satisfies the hypotheses of the preceding section on hierarchies. Namely, there is non-empty boundary with components being geometrically incompressible surfaces. Hence, there is an extended hierarchy for each piece, with the copies of $S$ included. Gluing back along the two copies of $S$ yields an extended hierarchy $\mathcal{H}$ for $M$ including $S$. Notice
DEHN’S LEMMA

that this extended hierarchy still has length 3 or 4, i.e. is very short or just short, if there are 2-sided Klein bottles. For the next part of the argument, we will assume \( M \) is orientable, as it turns out to be simple to deal with the non orientable case last.

To complete the proof of Dehn’s lemma and the loop theorem in the closed orientable case, we observe an elementary property (*) of the extended hierarchy \( \mathcal{H} \) constructed above, which plays a similar role to Waldhausen’s simplicity condition. Recall that the collection of surfaces in the extended hierarchy consists of disjoint closed geometrically incompressible surfaces and possibly essential 2-spheres \( S_1, S_2, \ldots, S_k \), followed by boundary geometrically incompressible spanning surfaces \( S_{k+1}, S_{k+2}, \ldots S_m \) with non-empty boundaries on the closed surfaces, then finally compressing disks \( S_{m+1}, S_{m+2}, \ldots S_n \). Moreover, the result of splitting \( M \) open along all these surfaces has components which are punctured 3-cells and connected summands of \( M \) without geometrically incompressible surfaces.

Consider the possibly punctured handlebody components, denoted \( H_1, H_2, \ldots, H_q \) obtained by splitting \( M \) open along \( S_1, S_2, \ldots, S_m \). The boundary pattern on these components consists of a collection \( \mathcal{C} \) of disjoint simple closed curves, which are copies of the boundaries of the spanning surfaces \( S_{k+1}, S_{k+2}, \ldots S_m \). The property (*) that we require is that there is no singular compressing disk \( D \) for any handlebody \( H_i \) which meets the boundary pattern \( \mathcal{C} \) in exactly two points. Notice that any simple closed curve in \( \partial H_i \) intersects \( \mathcal{C} \) an even number of times, assuming a transverse intersection, since \( \mathcal{C} \) separates \( \partial H_i \) into regions which are copies of the closed surfaces \( S_1, S_2, \ldots, S_k \) and copies of the spanning surfaces \( S_{k+1}, S_{k+2}, \ldots S_m \).

Note that if there was an embedded compressing disk \( D \) for some \( H_i \) meeting \( \mathcal{C} \) in exactly two points, then \( \partial D = \alpha \cup \beta \) would consist of two arcs, \( \alpha \) in some spanning surface and \( \beta \) in some closed incompressible surface. (Observe that since no spanning surface meets any essential 2-sphere, it follows that there are no loops of \( \mathcal{C} \) in such a 2-sphere). But this would contradict our construction that the spanning surfaces are geometrically boundary incompressible.

To prove property (*), assume there is a singular compressing disk \( D' \) for some \( H_i \) with boundary meeting \( \mathcal{C} \) in exactly two points. Choose a family of embedded compressing disks \( D_1, D_2, \ldots, D_r \) for \( H_i \) with the two properties that firstly the disks cut \( H_i \) up into a single (possibly punctured) 3-cell \( B \) and that secondly the disks are chosen with minimal possible intersection numbers with \( \mathcal{C} \) amongst disks with the first property. It is easy to arrange that the intersection of \( D' \) with all these disks consists of arcs only, since any loops can be removed by a standard disk exchange argument. Also we can suppose the number of such arcs is minimal, amongst choices of \( D' \). Choose an outermost arc \( \gamma \) in the domain of \( D' \), which maps to some arc in a disk \( D_1 \) of the family. Let \( \delta \) be the subarc in the domain of \( \partial D' \) with the same endpoints as \( \gamma \), so that when \( \delta \) is mapped into \( H_i \) it has no interior intersections with the compressing disks \( D_1, D_2, \ldots, D_r \). By assumption, the arc \( \delta \) cannot be
homotoped into $D_i$ keeping its ends fixed, since otherwise the number of arcs of intersection of $D'$ and the compressing disks could be reduced. Moreover, since there are assumed to be exactly two intersection points of $\partial D'$ and $C$, we may suppose that $\gamma$ and $\delta$ are chosen so that the image of $\delta$ crosses $C$ at most once. But then we can replace this singular arc by an embedded arc $\delta'$ with the same endpoints as $\delta$ and which meets $C$ the same number of times and is not homotopic keeping its ends fixed into $D_i$. The reason is that the image of $\delta$ lies in the boundary of $B$, which consists of 2-spheres. The 2-sphere component containing this singular arc can be cut open by the loops of $\partial D_1, ..., \partial D_r$ and the arcs of $C$ into two copies of each of the disks $D_1, D_2, ..., D_r$ and some even-sided polygons. (There cannot be any regions which are not disks topologically, since then the boundary loops of such a region would give embedded compressing disks for $H_i$ disjoint from $C$). If the image of $\delta$ misses $C$, then it lies in one of these polygonal regions and has ends on two different boundary edges lying on the same copy of $\partial D_j$. Hence the required embedded arc can be found in this polygonal region with the same endpoints as the singular arc. Similarly, if the image of $\delta$ crosses $C$ exactly once, then there must be two polygonal regions with a common boundary arc so that their union has two boundary edges on the same copy of $\partial D_j$, with the singular arc running between these edges. As before, we can find the desired embedded arc with the same endpoints and running between the same edges. In fact, in both cases we see that the singular and embedded arcs are homotopic keeping their ends fixed.

Finally, let $\partial D_i = \sigma \cup \tau$, where the two arcs $\sigma, \tau$ have the same ends as $\delta'$. Then the loops $C_1 = \sigma \cup \delta'$ and $C_2 = \tau \cup \delta'$ both bound compressing disks for $H_i$, since they lie in $\partial B$ and neither can be contractible on $\partial H_i$, since $\delta'$ is chosen not homotopic into $D_i$. It is standard to observe that the disk $D_i$ can be replaced by one of these two disks in the family of compressing disks. However both disks have fewer intersections with $C$ than do $D_i$, for the arc $\delta'$ meets $C$ at most once. If say $\sigma$ meets $C$ at least twice, then obviously $\tau \cup \delta'$ meets $C$ fewer times than does $\partial D_i = \sigma \cup \tau$. On the other hand, if $\sigma$ intersects $C$ at most once, the disk bounded by $\sigma \cup \delta'$ meets $C$ at most twice, which contradicts our construction that there are no such embedded compressing disks. So this completes the discussion of the property (*)

But now the conclusion follows similarly to that of [?] (c.f also [?]). The argument shows that all the surfaces in the hierarchy are algebraically incompressible and boundary algebraically incompressible, i.e the maps $\pi_1(S_i) \to \pi_1(M_{i-1})$ and $\pi_1(S_i, \partial S_i) \to \pi_1(M_{i-1}, \partial M_{i-1})$ are all one-to-one. Here $M_{i-1}$ is obtained by cutting $M$ open along the first $i-1$ surfaces in the hierarchy. Note the method of [?] and [?] is to study the graph $\Gamma'$ formed by the preimage of the hierarchy, using the map of some singular compressing disk or boundary compressing disk for a surface such as $S$ in the hierarchy. We look instead at the graph $\Gamma''$ which is the preimage of the closed and spanning surfaces of $\mathcal{H}$ in the domain $D$ of the disk. The closed surfaces give simple closed curves and by an easy disk exchange, any such
loops which have contractible image can be removed. Moreover, if there is such a loop with non-contractible image, we can study the subdisk bounded by it instead of our original disk. In this way, it suffices to assume there are no interior points mapped to the closed surfaces of $\mathcal{H}$. Similarly, loops mapped into spanning surfaces can be removed if they have contractible image. An essential loop in a spanning surface again gives a subdisk which can be studied instead of the original disk. So we are reduced to the case where the inverse image of all the closed and spanning surfaces consists of arcs in $D$. Now an outermost arc $\gamma$ by property (*), must have image which is homotopic into the boundary of some spanning surface. Consequently we can homotop $D$ to remove such an arc and continue to remove all the arcs in the preimage of the spanning surfaces. So the disk can be homotoped to have interior disjoint from the spanning surfaces and closed surfaces. But we know by property (*) that there are no singular compressing disks for the handlebodies which miss $C$, so the disk is homotopic into $S$. This completes the proof for the case of orientable 3-manifolds.

One interesting feature of very short hierarchies, is that we only have to establish property (*), to deal with singular compressing disks. In the argument in [?], one has to order the complexity of vertices of the graph $\Gamma^*$. If the hierarchy is only short, as in the case where there are 2-sided geometrically incompressible Klein bottles, then the best approach is to use a very short hierarchy for the orientable double covering manifold $\tilde{M}$. So this is our approach for any non-orientable 3-manifold $M$. By the previous case, a lift $\tilde{S}$ to $\tilde{M}$ of a 2-sided geometrically incompressible surface $S$ in $M$ is algebraically incompressible in $\tilde{M}$, since it is geometrically incompressible. We just note that if there was an embedded compressing disk for $\tilde{S}$, by a simple exchange along arcs of intersection, we could find an embedded compressing disk for $S$. But it is very elementary to prove that $S$ is algebraically incompressible as well. For $\pi_1(\tilde{S})$ is a subgroup of index 2 in $\pi_1(S)$ and injects into $\pi_1(\tilde{M})$ which is a subgroup of index 2 in $\pi_1(M)$, assuming that $S$ is non orientable. (If $S$ is orientable then $\tilde{S}$ is homeomorphic to $S$ and this case is trivial). So if follows immediately that the map of $\pi_1(S)$ into $\pi_1(M)$ must be an injection also.

To summarise, it follows that our original surface $S$ is algebraically incompressible and this completes the proof of Dehn’s lemma and the loop theorem for closed 3-manifolds. □

**Theorem 4.2.** Suppose that $M$ is a compact 3-manifold and $S$ is an embedded 2-sided properly embedded connected surface in $M$. Assume that $S$ is not a disk, 2-sphere or real projective plane and either the induced map from $\pi_1(S)$ to $\pi_1(M)$ is not one-to-one or the induced map from $\pi_1(S, \partial S)$ to $\pi_1(M, \partial M)$ is not one-to-one. Then there is an embedded disk $D$ in $M$ with $D \cap S = C$, where $C = \partial D$ is non-contractible in $S$ or there is an embedded disk $D$ in $M$ with $D \cap S = \alpha$ and $D \cap \partial M = \beta$, with $\partial D = \alpha \cup \beta$ and $\beta$ is not homotopic keeping its ends fixed into $\partial S$. 
Proof. The proof is very similar to the previous theorem. We first observe that if $S$ is closed, then we can compress the components of $\partial M$ which are not 2-spheres or projective planes. Namely, if there are any embedded compressing disks for such components, either the compressing disks can be cut-and-pasted to miss $S$, or else an innermost subdisk bounded by a loop of intersection with $S$ gives a compressing disk for $S$ and the theorem follows. So we may suppose without loss of generality that all the compressing disks are disjoint from $S$. Performing these compressions, i.e removing small open regular neighbourhoods of the compressing disks from $M$, we obtain a new 3-manifold $M'$ which has all boundary components which are either 2-spheres, projective planes or geometrically incompressible. We can then proceed exactly as in the previous theorem, including $S$ in an extended hierarchy for $M'$ and the case when $S$ is closed is now complete.

Next if $S$ has non-empty boundary in $\partial M$ and all the components of $\partial M$ are 2-spheres, projective planes or are geometrically incompressible in $M$, we can double $M$ along all the components of $\partial M$. So we get a new closed 3-manifold $2M$ with reflection involution $g$ between the two copies of $M$ and containing the closed surface $S$. If there is no embedded compressing disk or boundary compressing disk for $S$ in $M$, then it is easy to show that there is no embedded compressing disk for $2S$ in $2M$. For such a disk can be cut-and-pasted to eliminate all loops of intersection with the components of $\partial M$. If the disk misses $\partial M$ then it is an embedded compressing disk for one of the copies of $S$ and if it meets $\partial M$ in arcs, then either an innermost arc gives a boundary compressing disk for $S$, or the arc can be slid across $\partial S$ and the number of such arcs reduced. So in this case, the theorem is proved.

Finally, if $\partial M$ has some geometrically compressible components which are not 2-spheres or projective planes, we can proceed as in the first paragraph, to remove a small open neighbourhood of such compressing disks from $M$. Notice that if a compressing disk for a component of $\partial M$ meets $S$, then a standard argument as before shows that either $S$ has a compressing disk, boundary compressing disk, or the number of arcs and loops of intersection of $S$ and the compressing disk for $\partial M$ can be reduced. If all the compressing disks are disjoint from $S$ then removing a small regular neighbourhood gives a new 3-manifold $M'$ with geometrically incompressible boundary components, which are not 2-spheres or projective planes, and we are back in the previous case. This completes the general version of Dehn’s lemma and the loop theorem.

\[ \square \]

5. Remarks

- Notice that this proof of Dehn’s lemma and the loop theorem really only uses Kneser’s theorem on the finiteness of disjoint families of embedded normal surfaces (needed for the construction of hierarchies) and the concept of a hierarchy - which
DEHN'S LEMMA

follows from Waldhausen and ideas of Haken and Johannson. So the proof could have been found much before Papakyriakopoulos beautiful construction of towers of covering spaces, if the idea of hierarchies had been considered.

• It seems reasonable to call the procedure ‘localising’ since we show that cutting up a singular compressing disk into pieces, using the hierarchy, means that we only have to solve Dehn’s lemma and the loop theorem locally, i.e in regions which are (possibly punctured) handlebodies with simple closed curves as boundary pattern. In this case, the procedure is to cut the handlebody further up into a (possibly punctured) 3-cell and the disk into subdisks.

• The existence of very short hierarchies turns out to be rather useful for brief proofs of the torus and characteristic variety theorems for Haken 3-manifolds (see [7]). The above proof gives a procedure for constructing such hierarchies, which are also very valuable for various algorithmic problems (see [10]).

REFERENCES