§8. Hierarchies

In this section, we consider not just a single incompressible surface, but a whole sequence of them.

Terminology. Let M be a 3-manifold, containing an incompressible surface S. Then $M_S = M - \text{int}(\mathcal{N}(S))$ is the result of *cutting M along S*.

Definition. A *partial hierarchy* for a Haken 3-manifold M_1 is a sequence of 3-manifolds $M_1, \dots, M_n$, where M_{i+1} is obtained from M_i by cutting along an orientable incompressible properly embedded surface in M_i, no component of which is a 2-sphere. This is a *hierarchy* if, in addition, M_n is a collection of 3-balls. We denote (partial) hierarchies as follows:

$$M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} M_n.$$

Example. The following is a hierarchy for $S^1 \times S^1 \times S^1$:

$$S^1 \times S^1 \times S^1 \overset{S^1 \times S^1 \times \{*\}}{\longrightarrow} S^1 \times S^1 \times I \overset{S^1 \times \{*\} \times I}{\longrightarrow} S^1 \times I \times I \overset{\{*\} \times I \times I}{\longrightarrow} I \times I \times I.$$

Example. An example of hierarchy for a knot exterior is given below.

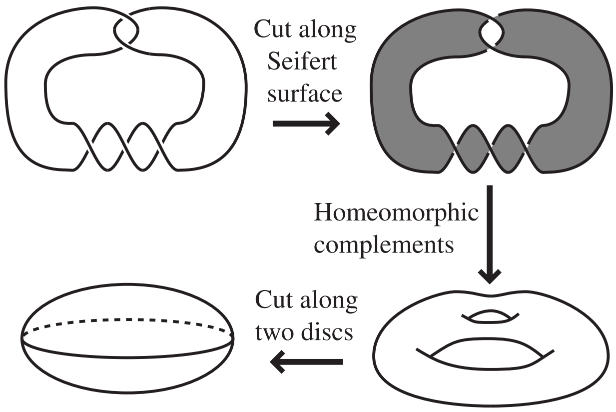

Figure 25.

Non-example. Let M be any 3-manifold with non-empty boundary. Let D be a disc in ∂M. Let D' be D with its interior pushed a little into the interior of M. Then decomposing M along D' gives a copy of M and a 3-ball. Hence, we may repeat this process indefinitely.

Non-example. Let S be the genus one orientable surface with one boundary component. Then $S \times I$ is homeomorphic to a genus two handlebody. Pick a simple closed non-separating curve C in the interior of S. Then $C \times I$ is a properly embedded annulus that is π_1-injective and hence incompressible. Cutting S along C gives a pair of pants $F_{0,3}$, and $F_{0,3} \times I$ is again a genus two handlebody. Hence, we may cut along a similar surface again, and repeat indefinitely.

Lemma 8.1. *Let M be a compact orientable irreducible 3-manifold. Let S be a properly embedded incompressible surface, no component of which is a 2-sphere. Then M_S is irreducible, and hence Haken since $\partial M_S \neq \emptyset$.*

Proof. Let S^2 be a 2-sphere in M_S. As M is irreducible, it bounds a 3-ball in M. If this 3-ball contained any component of S, then S would be compressible, by Theorem 3.8. Hence, S is disjoint from the 3-ball, and so the 3-ball lies in M_S. □

Despite the 'non-examples' above, the following theorem is in fact true.

Theorem 8.2. *Every Haken 3-manifold has a hierarchy.*

Theorem 8.2 will be proved in §11, but first, we show why hierarchies are useful.

9. Boundary patterns and the Loop Theorem

Definition. A *boundary pattern* P in a 3-manifold M is a (possibly empty) collection of disjoint simple closed curves and trivalent graphs in ∂M, such that no simple closed curve in ∂M intersects P transversely in a single point.

If S is a 2-sided surface properly embedded in a compact 3-manifold M, with ∂S intersecting P transversely (and missing the vertices of P), then the manifold M_S obtained by cutting along S inherits a boundary pattern, as follows. Note that ∂M_S is the union of subsurfaces, one of which is $\partial M \cap \partial M_S$, the other of which is $\partial \mathcal{N}(S) \cap \partial M_S$, which is two copies S_1 and S_2 of S. Then, M_S inherits a boundary pattern $(P \cap \partial M_S) \cup \partial S_1 \cup \partial S_2$.

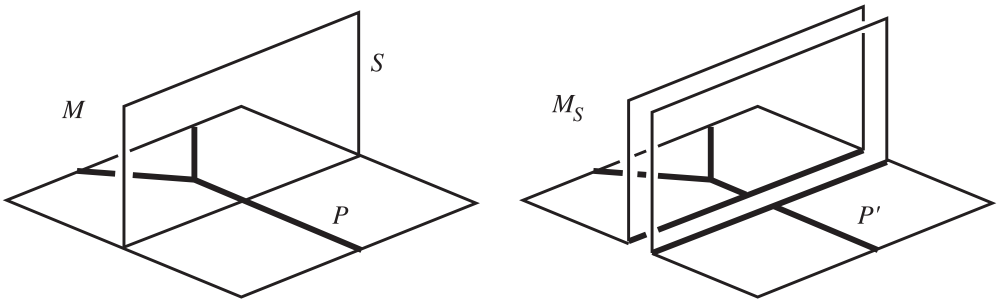

Figure 26.

The motivation for defining boundary patterns is as follows. If

$$M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} M_n$$

is a partial hierarchy for a 3-manifold M_1, then ∂M_n is a union of subsurfaces, which come from bits of ∂M_1 and $S_1, \dots, S_{n-1}$. The union of the boundaries of these bits of surface forms a boundary pattern for M_n.

Definition. A boundary pattern P for M is *essential* if, for each disc D properly embedded in M with $\partial D \cap P$ at most three points, there is a disc $D' \subset \partial M$ with $\partial D' = \partial D$, and D' containing at most one vertex of P and no simple closed curves of P.

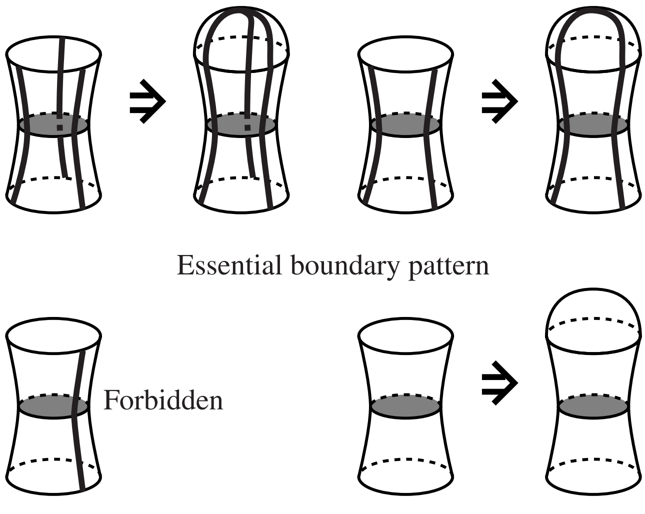

Figure 27.

Definition. A boundary pattern P is *homotopically essential* if, for each map of a disc $(D, \partial D) \to (M, \partial M)$ with $\partial D \cap P$ at most three points (which are disjoint), there is a homotopy (keeping $\partial D \cap P$ fixed, introducing no new points of $\partial D \cap P$, and keeping ∂D in ∂M) to an embedding of D into ∂M so that the image of D contains at most one vertex of P and no simple closed curves of P.

Clearly, if a boundary pattern is homotopically essential, then it is essential. (A proof of this requires the fact from surface topology that if two properly embedded arcs in a surface are homotopic keeping their endpoints fixed, then they are ambient isotopic keeping their endpoints fixed.) The main technical result that we will prove is that the converse holds.

Theorem 9.1. *An essential boundary pattern for a compact orientable irreducible 3-manifold is homotopically essential.*

The Loop Theorem is a corollary of this result. This remarkable result is one of the most important theorems in 3-manifold theory. In this course, we will give a new proof of it, using hierarchies.

Theorem 9.2. (The Loop Theorem) *Let M be a compact orientable irreducible 3-manifold. Then ∂M is incompressible if and only if $\pi_1(F) \rightarrow \pi_1(M)$ is injective for each component F of ∂M.*

Proof of 9.2 from 9.1. A standard fact from surface topology gives that a simple closed curve in ∂M is homotopically trivial in ∂M if and only if it bounds a disc in ∂M. Hence, if a component F of ∂M is compressible, then $\pi_1(F) \rightarrow \pi_1(M)$ is not injective.

To prove the converse, suppose that ∂M is incompressible. Let P be the empty boundary pattern in ∂M. This is then essential. By Theorem 9.1, P is homotopically essential. Hence, if ℓ is any loop in ∂M that is homotopically trivial in M, then ℓ is homotopically trivial in ∂M. □

We can in fact prove the following slightly stronger version of the Loop Theorem.

Theorem 9.3. *Let M be a compact orientable irreducible 3-manifold, and let F be a connected surface in ∂M. If $\pi_1(F) \rightarrow \pi_1(M)$ is not injective, then F is compressible.*

Proof of 9.3 from 9.1. Suppose that F is incompressible. Let ∂F be the boundary pattern of M. If this is not essential, then there is a compressing disc for ∂M that intersects ∂F at most twice. Decompose M along this disc to give a new 3-manifold M'. Let F' be $M' \cap F$. Then $\pi_1(F) \to \pi_1(M)$ is injective if and only if each component of F' is π_1-injective in M'. Also, F' is incompressible in M'. Repeat this process if necessary. At each stage, we reduce the complexity of ∂M. Hence, we may assume that the boundary pattern ∂F is essential in M. By Theorem 9.1, it is homotopically essential, and therefore $\pi_1(F) \to \pi_1(M)$ is injective. □

This stronger version of Theorem 9.3 allows us to prove Theorem 3.3.

Theorem 3.3. *Let S be a connected compact orientable surface properly embedded in a compact orientable irreducible 3-manifold M. Then S is incompressible if and only if the map $\pi_1(S) \to \pi_1(M)$ induced by inclusion is an injection.*

Proof. Suppose that $\pi_1(S) \to \pi_1(M)$ is not injective. There is then a map $h\colon (D, \partial D) \to (M, S)$ of a disc D such that $h(\partial D)$ is homotopically non-trivial in S. Using an argument similar to that in Lemma 7.8, we may perform a homotopy of D (keeping ∂D fixed) so that $h^{-1}(S)$ is a collection of simple closed curves in D. Pick one innermost in D. If this is sent to a curve that is homotopically trivial in S, we may modify h and remove this curve. Hence, we may assume that there is a map $h\colon D \to M$ so that $h^{-1}(S) = \partial D$ and so that $h(\partial D)$ is homotopically non-trivial in S. We may also assume that $h|_{\mathcal{N}(\partial D)}$ respects the product structure on $\mathcal{N}(S)$. Hence, h restricts to a trivialising homotopy for some loop in one of the two copies of S in M_S. Applying Theorem 9.3 to this copy F of S gives that F is compressible. Extending the compression disc using the product structure $\mathcal{N}(S) \cong S \times I$ gives a compression disc for S. □

Remark. This argument fails (and the result need not be true) when S is non-orientable: since $\mathcal{N}(S)$ is not a product, a compression disc for the ∂I-bundle of $\mathcal{N}(S)$ does not necessarily extend to a compression for S.

Theorem 9.3, together with the existence of hierarchies, also allows us to prove the following.

Theorem 9.4. *Let M be a compact orientable Haken 3-manifold. Then $\pi_k(M) = 0$ for all $k \geq 2$.*

Proof. Pick a hierarchy

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} M_n.$$

Consider a map $h\colon S^2 \to M$, and let S_i be the first surface to intersect $h(S^2)$. We may homotope h so that $h^{-1}(S_i)$ is a collection of simple closed curves. Let C be one innermost in S^2 bounding a disc D. Then $h(C)$ is homotopically trivial in M_{i+1}. Hence, by the argument of Theorem 9.3, we may homotope D into S_i. There is then a a further homotopy removing C from $h^{-1}(S_i)$. We may therefore assume that $h(S^2) \subset M_{i+1}$. Repeating this as far as M_n gives that $h(S^2) \subset M_n$. Since $\pi_2(M_n)$ is trivial, h represents a trivial element of $\pi_2(M)$. Therefore $\pi_2(M) = 0$.

If M is closed, then $\pi_1(M)$ contains the fundamental group of a closed orientable surface other than a 2-sphere, and hence $\pi_1(M)$ is infinite. If M has non-empty boundary, then (providing it is not a 3-ball), $H_1(M)$ is infinite, by Theorem 7.5, and so $\pi_1(M)$ is infinite. Therefore the universal cover $\tilde{M}$ of M is non-compact. Hence, $H_k(\tilde{M}) = 0$ for all $k \geq 3$. Now, $\pi_k(\tilde{M}) \cong \pi_k(M)$ for all $k \geq 2$. Therefore, $\pi_2(\tilde{M}) = 0$. Hence, by the Hurewicz theorem, $\pi_k(\tilde{M}) \cong H_k(\tilde{M}) \cong 0$ for all $k \geq 3$. This proves the theorem. □

Remark. It is possible to show (using rather different methods) that $\pi_2(M) = 0$ for all irreducible orientable 3-manifolds M. Hence, if in addition $\pi_1(M)$ is infinite, $\pi_k(M) = 0$ for all $k \geq 3$.

10. Special hierarchies

Definition. Let S be a surface properly embedded in a 3-manifold M with boundary pattern P. Then a *pattern-compression disc* for S is a disc D embedded in M such that

- $D \cap S$ is an arc α in ∂D,
- $\partial D - \mathrm{int}(\alpha) = D \cap \partial M$ intersects P at most once, and
- α does not separate off a disc from S intersecting P at most once.

If no such pattern-compression disc exists, then S is *pattern-incompressible*.

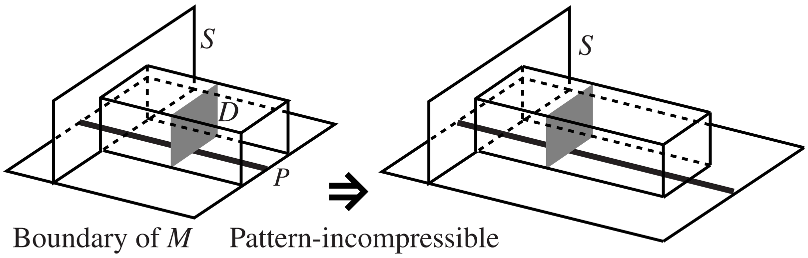

Figure 28.

Definition. Two surfaces S_0 and S_1 embedded in a 3-manifold M are *parallel* if there is an embedding of $S \times [0,1]$ in M such that $S_0 = S \times \{0\}$ and $S_1 = S \times \{1\}$. If $\partial(S \times [0,1]) - S_0 \subset \partial M$, we say that S_0 is *boundary-parallel*.

Definition. A *special hierarchy* for a compact orientable irreducible 3-manifold M with boundary pattern P is a hierarchy for M of properly embedded connected pattern-incompressible incompressible surfaces, none of which is a 2-sphere or boundary-parallel disc. (At each stage, the cut-open 3-manifold inherits its boundary pattern from the previous one.) We write the manifolds and boundary patterns as:

$$(M,P) = (M_1,P_1) \xrightarrow{S_1} (M_2,P_2) \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} (M_n,P_n).$$

We now give an overview of the proof of Theorem 9.1. It proceeds in four main steps:

1. Show that any compact connected orientable irreducible 3-manifold M with essential boundary pattern P and non-empty boundary has a special hierarchy

$$(M,P) = (M_1,P_1) \xrightarrow{S_1} (M_2,P_2) \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} (M_n,P_n).$$

2. Show that (M_i, P_i) is essential if and only if (M_{i+1}, P_{i+1}) is.
3. Show, using simple properties of the 3-ball, that (M_n, P_n) being essential implies that it is homotopically essential.
4. Show that if (M_{i+1}, P_{i+1}) is homotopically essential, then so is (M_i, P_i).

We will save step 1 until §11. We now embark on steps 2, 3 and 4.

Lemma 10.1. *Let M be a compact orientable irreducible 3-manifold with essential boundary pattern P. Let S be a connected pattern-incompressible incom-*

pressible surface in M, which is not a boundary-parallel disc. Then the 3-manifold M_S obtained by cutting along S inherits an essential boundary pattern P'.

Proof. Let D be a disc properly embedded in M_S with $\partial D \cap P'$ at most three points. The curve ∂D may run through parts of ∂M_S coming from ∂M and parts coming from S. Note however the points where it swaps must be points of $\partial D \cap P'$, and that at most one side of any point of $\partial D \cap P'$ lies in S. Hence, at most one arc of $\partial D - P'$ lies in S.

Case 1. ∂D is disjoint from S.

Then $\partial D \subset \partial M$. Since P is essential, ∂D bounds a disc D' in ∂M containing at most one vertex of P and no simple closed curves. If S intersects D', then pick a simple closed curve of $S \cap D'$ innermost in D'. The disc this bounds cannot be a compression disc for S. Hence, S must be a disc. Since M is irreducible, it is parallel to a disc in ∂M, contrary to assumption. Hence, D' is disjoint from S, and therefore lies in ∂M_S. This verifies that D does not violate the essentiality of P'.

Case 2. ∂D intersects S.

Then $\partial D - S$ intersects P at most once. Since D is not a pattern-compressing disc for S, $D \cap S$ separates off a disc D_1 of S intersecting P in at most one point. Then, $D \cup D_1$ is a disc properly embedded in M, intersecting P in at most two points. There is therefore a disc D_2 in ∂M with $\partial D_2 = \partial(D \cup D_1)$, containing at most one vertex of P and no simple closed curves, since P is essential. Since $D \cup D_1$ intersects P in at most two points, D_2 cannot therefore contain any vertex of P. Therefore, $D_1 \cup D_2$ is a disc in ∂M_S containing at most one vertex of P' and no simple closed curves. This gives that P' is essential. □

Lemma 10.2. *Suppose that M is a 3-ball with essential boundary pattern P. Then P is homotopically essential.*

Proof. Consider a map $(D, \partial D) \to (M, \partial M)$ with $\partial D \cap P$ at most three points. Since P is essential, each component of $\partial M - P$ is a disc. We may therefore homotope each arc of $\partial D - P$ so that it is embedded. The arcs $\partial D - P$ lie in different components of $\partial M - P$, since P is a boundary pattern. Hence, we have homotoped ∂D so that it is embedded. It therefore bounds an embedded disc D'

in ∂M. Since P is essential, D' contains at most one vertex of P and no simple closed curves. As the 3-ball has trivial π_2, there is a homotopy taking D to D' keeping ∂D fixed. □

Lemma 10.3. *Let M be a compact orientable 3-manifold with boundary pattern P. Let S be an orientable incompressible pattern-incompressible surface properly embedded in M. Let P' be the boundary pattern inherited by M_S. If P' is homotopically essential, then so is P.*

Proof. Consider a map $h\colon (D, \partial D) \to (M, \partial M)$ such that ∂D intersects P in at most three points. We may perform a small homotopy so that $h^{-1}(S)$ is a collection of properly embedded arcs and circles in D.

Suppose that there is some simple closed curve of $h^{-1}(S)$. Pick one C innermost in D, bounding a disc D'. Since P' is homotopically essential, we may homotope D' to an embedded disc in S. Perform a further small homotopy to reduce $|h^{-1}(S)|$.

Hence, we may assume that there are no simple closed curves of $h^{-1}(S)$. If there is more than one arc, at least two are extrememost in D. They separate off discs D_1 and D_2 from D. Similarly, if there is only one arc of $h^{-1}(S)$, it divides D into two discs D_1 and D_2. There are only three points of $h^{-1}(P)$, and so D_1, say, contains at most one of these points. Hence, $h(\partial D_1)$ intersects P' in at most three points. Since P' is homotopically essential, we may homotope D_1 to an embedded disc D' in ∂M_S containing at most one vertex of P' and no simple closed curves. Replace D_1 with D', and perform a homotopy to reduce $|h^{-1}(S)|$.

Repeat this process until $h^{-1}(S) = \emptyset$. Then, use that P' is homotopically essential to construct the desired homotopy of D to an embedded disc in ∂M containing at most one vertex of P and no simple closed curves. □

This completes steps 2, 3 and 4. A similar argument to that of Lemma 10.3 gives the following.

Lemma 10.4. *Let M be a compact orientable 3-manifold with boundary pattern P. Let S be an orientable incompressible pattern-incompressible surface properly embedded in M. Let P' be the boundary pattern inherited by M_S. If P' is essential, then so is P.*

All that is now required in the proof of the Loop Theorem is to establish the existence of special hierarchies. For this, we need extra machinery.

11. Normal surfaces

Definition. A *triangle* (respectively, *square*) in a 3-simplex Δ^3 is a properly embedded disc D such that ∂D intersects precisely three (respectively, four) 1-simplices transversely in a single point, and is disjoint from the remaining 1-simplices and all the vertices.

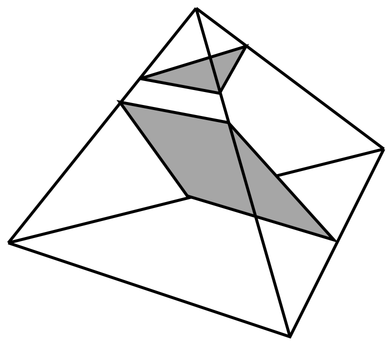

Figure 29.

Fix a triangulation T of the 3-manifold M.

Definition. A properly embedded surface in M is *in normal form* with respect to T if it intersects each 3-simplex in a finite (possibly empty) collection of disjoint triangles and squares.

Theorem 11.1. *Let M be a compact irreducible 3-manifold. Let S be a properly embedded closed incompressible surface in M, with no component of S a 2-sphere. Then, for any triangulation T of M, S may ambient isotoped into normal form.*

Proof. First, a small ambient isotopy makes S transverse to the 2-skeleton of the triangulation. Then S intersects each 2-simplex in a collection of arcs and simple closed curves. We may assume that it misses the vertices of T. Let the *weight* $w(S)$ of S be the number of intersections between S and the 1-simplices.

Suppose first that there is a simple closed curve of intersection between S and the interior of some 2-simplex. Pick one C innermost in the 2-simplex, bounding a disc D in the 2-simplex. Then C bounds a disc D' in S, since D is not a compression disc for S. Since M is irreducible, we may ambient isotope D' onto D. This does not increase $w(S)$. Hence, we may assume that S intersects each 2-simplex in a (possibly empty) collection of arcs.

If $w(S)$ is zero, then each component of S lies in a 3-simplex. By Theorem 3.8, any such component is 2-sphere, contrary to assumption. We will perform a sequence of ambient isotopies to the surface, which will reduce $w(S)$ and hence are guaranteed to terminate.

Let Δ^3 be a 3-simplex of M. Suppose first that S intersects Δ^3 in something other than a collection of discs. If there is a non-disc component of $S \cap \Delta^3$ with non-empty boundary, then pick a curve of $S \cap \partial\Delta^3$ innermost in $\partial\Delta^3$ among all curves not bounding discs of $S \cap \Delta^3$. This bounds a compression disc D for $S \cap \Delta^3$. Since S is incompressible in M, ∂D bounds a disc in S. Ambient isotope this disc onto D to decrease $w(S)$. If every component of $S \cap \Delta^3$ with non-empty boundary is disc, then any closed component of $S \cap \Delta^3$ lies in the complement of these discs, which is a 3-ball. Hence, it is a 2-sphere by Theorem 3.8. Thus, we may assume that each component of $S \cap \Delta^3$ is a disc.

Now suppose that some disc D of $S \cap \Delta^3$ intersects a 1-simplex σ more than once, as in Figure 30. We claim that we can find such a disc D, and two points of $D \cap \sigma$, so that no other points of $S \cap \sigma$ lie between them on σ. First pick two points of $D \cap \sigma$ having no points of $D \cap \sigma$ between them on σ. Let β be the arc of σ between them. Note that ∂D separates $\partial\Delta^3$ into two discs and that β is properly embedded in one of these. Hence, if D' is any other disc of $S \cap \Delta^3$, it intersects β in an even number of points. Hence, we may find a disc D of $S \cap \Delta^3$ intersecting σ in adjacent points on σ. Let β be the arc of σ between them, and let α be some arc properly embedded in D joining these two points. Note that $S \cap \Delta^3$ separates Δ^3 into 3-balls and that $\alpha \cup \beta$ lies in the boundary of one of these balls. Hence, there is a disc D' embedded in Δ^3 with $D' \cap (S \cup \partial\Delta^3) = \alpha \cup \beta$. Then we may use the disc D' to ambient isotope S, reducing $w(S)$, as in Figure 30.

Hence, we may assume that each disc of $S \cap \Delta^3$ intersects each 1-simplex at most once. It is then a triangle or square. Hence, S is now normal. □

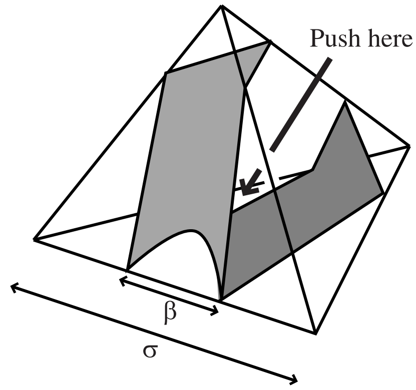

Figure 30.

Theorem 11.2. *Let M be a compact orientable irreducible 3-manifold. Then there is some integer $n(M)$ with the following property. If S is a closed properly embedded incompressible surface in M with more than $n(M)$ components, none of which is a 2-sphere, then at least two components of S are parallel (with no component of S in the product region between them).*

Proof. We let $n(M) = 2\beta_1(M;\mathbb{Z}_2) + 6t$, where t is the number 3-simplices in some triangulation of M. Let S have components $S_1, \dots, S_k$, with $k > n(M)$. Then, by Theorem 11.1, S may be ambient isotoped into normal form. Note that M_S has more than $\beta_1(M;\mathbb{Z}_2) + 6t$ components. Also, for each 3-simplex Δ^3, all but at most six components of $\Delta^3 - S$ is a product region, lying between adjacent triangles or squares. Therefore, more than $\beta_1(M;\mathbb{Z}_2)$ components of M_S are composed entirely of product regions. Each such component X of M_S is an I-bundle. If X is not a product I-bundle, then it is an I-bundle over a non-orientable surface. Then we can calculate that $H_1(\partial X;\mathbb{Z}_2) \to H_1(X;\mathbb{Z}_2)$ is not surjective. Hence, there is a non-trivial summand of $H_1(M;\mathbb{Z}_2)$ for each such component X of M. So, at most $\beta_1(M;\mathbb{Z}_2)$ are of this form. Hence, there is at least one product I-bundle of M_S. Its two boundary components are parallel in M. □

Lemma 11.3. *Let M be a compact orientable 3-manifold, and let*

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} M_n$$

be a partial hierarchy. Let $X = \mathcal{N}(\partial M \cup S_1 \cup \dots \cup S_{n-1})$. Then $\partial X - \partial M$ is incompressible in X.

Proof. Consider a compression disc D for $\partial X - \partial M$ in X. Let S_i be the first surface in the hierarchy it intersects. Then we may assume $D \cap S_i$ is a collection of simple closed curves in the interior of D. Pick one innermost in D, bounding a disc D_1. This cannot be a compression disc for S_i, and so it bounds a disc D_2 in S_i. Remove D_1 from D, replace it with D_2, and perform a small isotopy to reduce $|D \cap S_i|$. This does not introduce any new intersections with $S_1 \cup \ldots \cup S_{i-1}$. Thus, we may assume that D is disjoint from S_i, and, repeating, from all of the surfaces in the partial hierarchy. It therefore lies in the the space X, with the interior of a small regular neighbourhood of $S_1 \cup \ldots \cup S_{n-1}$ removed. This is a copy of $F \times I$, for a closed orientable surface F, with $F \times \{1\}$ identified with $\partial X - \partial M$. But the boundary of $F \times I$ is π_1-injective, and hence incompressible, which is a contradiction. □

Theorem 11.4. *Let M be a compact orientable irreducible 3-manifold with non-empty boundary and an essential boundary pattern P. Then M has a special hierarchy. Furthermore, if M has non-empty boundary, we may assume that no surface in this hierarchy is closed.*

Proof. Suppose first that ∂M is compressible. Let D be a compression disc. If there is a pattern-compression disc for D, then 'compressing' D along this disc decomposes D into two discs. Both of these discs have fewer intersections with P, and at least one of these is a compression disc for ∂M. Focus on this disc, and repeat until we have a pattern-incompressible compression disc for M. Decompose M along this disc. By Lemma 10.1, the resulting manifold M_2 inherits an essential boundary pattern. If its boundary is compressible, cut again along a pattern-incompressible compression disc. Repeat, giving a partial special hierarchy

$$M = M_1 \xrightarrow{S_1} M_2 \xrightarrow{S_2} \ldots \xrightarrow{S_{i-1}} M_i,$$

where ∂M_i is incompressible in M_i. We must reach such an M_i, since the complexity of ∂M_2 is less than that of ∂M_1, and so on. Push ∂M_i a little into M, giving a closed properly embedded surface F_1.

Claim. F_1 is incompressible in M.

The surface F_1 separates M into two components: M_i and $X = \mathcal{N}(\partial M \cup S_1 \cup \ldots \cup S_{i-1})$. By assumption, F_1 is incompressible in M_i. By Lemma 11.3, F_1 is incompressible in X. This proves the claim.

Each 2-sphere component of ∂M_i bounds a 3-ball. If every component of ∂M_i is a 3-ball, then we have constructed our special hierarchy as required. Suppose therefore that at least one component of ∂M_i is not a 2-sphere. By Theorem 7.6, M_i contains a properly embedded connected incompressible 2-sided non-separating surface S. If M_i has non-empty boundary, then we may assume that ∂S is non-empty. If S has a pattern-compression disc, then 'compress' S along this disc giving a surface S'. Then S' is incompressible and 2-sided, and at least one component S_1 of S' is non-separating. Then either $\chi(S_1) > \chi(S)$, or $\chi(S_1) = \chi(S)$ and $|S_1 \cap P| < |S \cap P|$. Hence, we may assume that S is pattern-incompressible. Cut along this surface to give M_{i+1}. If ∂M_{i+1} is compressible, then, as above, compress it as far as possible to give a closed incompressible surface F_2 in M. Note that F_1 and F_2 are disjoint. Continue this process. If we have not stopped by the time we have constructed $F_{n(M)+1}$, Theorem 11.2 implies that at some stage F_i and F_j are parallel for some $i < j$, with no F_k in the product region between them. Some S_p lies in this product region. The theorem is then proved by the following lemma. □

Lemma 11.5. *Let F be a compact orientable surface. Then there is no connected non-separating incompressible surface S properly embedded in $F \times [0,1]$ that is disjoint from $F \times \{1\}$.*

Proof. If F is closed, pick a simple closed curve C in F that does not bound a disc. Then $C \times [0,1]$ is an annulus A. A small ambient isotopy of S ensures that $S \cap A$ is a collection of arcs and simple closed curves. We may remove all simple closed curves of $S \cap A$ that bound discs in A. If there is an arc, it has both its endpoints in $C \times \{0\}$. We may find such an arc separating off a disc of A with interior disjoint from S. 'Compress' S along this disc to reduce $|S \cap A|$. The result is still an incompressible surface, and at least one component is non-separating. Hence, we may assume that $S \cap A$ contains only simple closed curves. By 'compressing' S along annuli in A, we may remove each of these. Hence, we may assume that S lies in $(F - C) \times [0,1]$. Therefore, we may assume that F has non-empty boundary. Pick a collection α of arcs properly embedded in F which cut F to a disc. Apply an argument as above to ensure that S is disjoint from $\alpha \times [0,1]$. It is then a disc properly embedded in $(F - \mathcal{N}(\alpha)) \times [0,1]$, which is a 3-ball. It is therefore separating. □

12. Topological rigidity

In this section, we will prove that homotopy equivalent closed Haken 3-manifolds are homeomorphic. The main ingredients are the existence of hierarchies and the loop theorem. A vital part of the argument is a version of topological rigidity for surfaces. Its proof is instructive, since it follows the same approach as the 3-manifold case.

Theorem 12.1. *Let F and G be connected compact surfaces with $\pi_1(F) \neq 0$. Let $f\colon (F, \partial F) \to (G, \partial G)$ be a map with $f_*\colon \pi_1(F) \to \pi_1(G)$ injective. Then, there is a homotopy through maps $f_t\colon (F, \partial F) \to (G, \partial G)$ with $f_0 = f$ and either*

(i) $f_1\colon F \to G$ is a covering map, or

(ii) F is an annulus or Möbius band and $f_1(F) \subset \partial G$.

If, for some components C of ∂F, $f|_C$ is a covering map, we can require that $f_t|_C = f|_C$ for all t.

Lemma 12.2. *Let $f\colon (F, \partial F) \to (G, \partial G)$ be a map between connected surfaces with non-empty boundary such that*

1. *$f|_{\partial F}$ is not injective, and its restriction to each component of ∂F is a cover,*
2. *$f_*\colon \pi_1(F) \to \pi_1(G)$ is an isomorphism,*
3. *$\pi_1(F) \neq 0$, and*
4. *F is compact.*

Then conclusion (ii) of Theorem 12.1 holds.

Proof. By (1), there are two points in ∂F mapping to the same point in ∂G, and there is a path $\gamma\colon I \to F$ joining them. Then $f \circ \gamma$ is a loop in G. By (2), there is a loop β in F based at $\gamma(0)$ such that $f_*([\beta]) = [f \circ \gamma]^{-1} \in \pi_1(G, f\gamma(0))$. Then $\alpha = \beta.\gamma$ is a path $(I, \partial I) \to (F, \partial F)$ such that $\alpha(0) \neq \alpha(1)$ and $f \circ \alpha$ is a homotopically trivial loop in G.

For $i = 0$ and 1, let J_i be the component of ∂F containing $\alpha(i)$. (Possibly, $J_0 = J_1$.) Orient J_i in some way, so that it is a loop based at $\alpha(i)$. Let K be the component of ∂G containing $f \circ \alpha(0) = f \circ \alpha(1)$. Then $f_*([J_0])$ and $f_*([\alpha.J_1.\alpha^{-1}])$

are both non-zero powers of $[K]$ in $\pi_1(G, f\alpha(0))$, by (1). Hence, by (2), some power of $[J_0]$ is some power of $[\alpha.J_1.\alpha^{-1}]$ in $\pi_1(F, \alpha(0))$. Let $x = \alpha(0)$. Let $p\colon (\tilde{F}, \tilde{x}) \to (F, x)$ be the covering of F such that $p_*\pi_1(\tilde{F}, \tilde{x}) = i_*\pi_1(J_0, x)$, where $i\colon J_0 \to F$ is the inclusion map. Lift α to a path $\tilde{\alpha}$ starting at $\tilde{x}$. Let $\tilde{J}_i$ be the component of $\partial\tilde{F}$ containing $\tilde{\alpha}(i)$. Since some power of $[\alpha.J_1.\alpha^{-1}]$ is some power of $[J_0] \in \pi_1(F, x)$, $\tilde{J}_1$ is compact.

Claim. $\tilde{J}_0 \neq \tilde{J}_1$.

Otherwise, since $\pi_1(\tilde{J}_0) \to \pi_1(\tilde{F})$ is an isomorphism, we may homotope $\tilde{\alpha}$ (keeping its endpoints fixed) to a path α_1 in $\tilde{J}_0$. But then $f \circ p \circ \alpha_1$ is a loop in K which lifts to a path under the covering $f|_{J_0}\colon J_0 \to f(J_0) \subset K$. Since $f \circ p \circ \alpha_1$ is null-homotopic in G, $\pi_1(K) \to \pi_1(G)$ is therefore not injective. Hence G is a disc and so, by (2), $\pi_1(F) = 0$. However, this contradicts (3) and so this proves the claim.

Claim. $\tilde{F}$ is compact.

We have the following exact sequence:

$$0 \to H_2(\tilde{F}, \tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2) \to H_1(\tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2) \to H_1(\tilde{F}; \mathbb{Z}_2).$$

The last of the above groups is isomorphic to $H_1(\tilde{J}_0; \mathbb{Z}_2) \cong \mathbb{Z}_2$. The middle group is $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. Hence, the first group must be non-trivial. Hence, $\tilde{F}$ is a compact surface.

The only compact surface with the property that some power of one boundary component can be freely homotoped into one power of another boundary component is an annulus. Since $\chi(\tilde{F})$ is a multiple of $\chi(F)$, F is an annulus or Möbius band. Using that $f \circ \alpha$ is homotopically trivial, we can retract f into ∂G. So (ii) of Theorem 12.1 holds. □

Proof of Theorem 12.1. Let $p\colon \tilde{G} \to G$ be the cover where $p_*\pi_1(\tilde{G}) = f_*\pi_1(F)$. Construct a lift

$$\begin{array}{ccc} & & \tilde{G} \\ & \overset{\tilde{f}}{\nearrow} & \downarrow p \\ F & \xrightarrow{f} & G \end{array}$$

Then $\tilde{f}_*$ is an isomorphism. We will show that $\tilde{f}$ may homotoped so that either (i) or (ii) hold. This will prove the result.

Note that each boundary component of F is π_1-injective in F. Hence, if $\tilde{f}$ is not already a covering map on ∂F, we may homotope it to so that it is a cover. If $\tilde{f}|_{\partial F}$ is not a homeomorphism, then by Lemma 12.1, case (ii) of Theorem 12.1 holds for $\tilde{f}$ and hence f. So, we may assume that $\tilde{f}|_{\partial F}$ is a homeomorphism onto its image.

Claim. $\tilde{G}$ is compact.

If $\tilde{G}$ is non-compact, then $\pi_1(\tilde{G})$ is free. So, F is not a closed surface. Note that the following commutes

$$\begin{array}{ccc} H_2(F, \partial F; \mathbb{Z}_2) & \longrightarrow & H_1(\partial F; \mathbb{Z}_2) \\ \downarrow & & \downarrow \\ H_2(\tilde{G}, \partial \tilde{G}; \mathbb{Z}_2) & \longrightarrow & H_1(\partial \tilde{G}; \mathbb{Z}_2) \end{array}$$

Since the map along the top has non-zero image and the map on the right is injective, their composition is not the zero map. Hence, $H_2(\tilde{G}, \partial\tilde{G}; \mathbb{Z}_2)$ is non-trivial and so $\tilde{G}$ is compact.

By looking at $\tilde{f}$ instead of f, it therefore suffices to consider the case where f_* is an isomorphism and $f|_{\partial F}$ is a homeomorphism onto its image. Consider first the case where ∂G is non-empty. Pick a collection A of properly embedded arcs in G which cut it to a disc. We may homotope f (keeping it unchanged on ∂F) so that $f^{-1}(A)$ is a collection of properly embedded arcs and simple closed curves. If there is any simple closed curve, its image in G lies in an arc, and hence is homotopically trivial. Hence, each simple closed curve of $f^{-1}(A)$ bounds a disc. By repeatedly considering an innermost such curve, we may homotope f to remove all such simple closed curves.

Since $f|_{\partial F}$ is a homeomorphism, the endpoints of each arc of $f^{-1}(A)$ map to distinct points in G. Hence, we may homotope $f|_{\mathcal{N}(f^{-1}(A)\cup\partial F)}$ so that it is a homeomorphism. But the remainder $F-(f^{-1}(A)\cup\partial F)$ maps to a disc in G. Since f is π_1-injective, $F-(f^{-1}(A)\cup F)$ is a collection of discs. A map of a disc to a disc that is a homeomorphism from boundary to boundary may be homotoped to a homeomorphism. Hence, we have therefore homotoped f to a homeomorphism.

Now consider the case where G is closed. Pick a simple closed curve C in G that does not bound a disc. Homotope f so that $f^{-1}(C)$ is a collection of simple closed curves in F, none of which bounds discs. Then $f|_{F-\mathrm{int}(\mathcal{N}(f^{-1}(C)))} \colon F -$

$\text{int}(\mathcal{N}(f^{-1}(C))) \to G - \text{int}(\mathcal{N}(C))$ is π_1-injective. We have proved the theorem in the case of surfaces with non-empty boundary. Consider therefore a component of $F - \text{int}(\mathcal{N}(f^{-1}(C)))$. If it is an annulus or Möbius band that can be homotoped into C, then perform this homotopy. A further small homotopy reduces the number of components of $f^{-1}(C)$. Hence, we may assume that case (i) applies to each component of $F - \text{int}(\mathcal{N}(f^{-1}(C)))$. Then we have homotoped f to a cover. □

We can now tackle topological rigidity for Haken 3-manifolds. The full result is the following.

Theorem 12.3. *Let M and N be Haken 3-manifolds. Suppose that there is a map $f\colon (M, \partial M) \to (N, \partial N)$ such that $f_*\colon \pi_1(M) \to \pi_1(N)$ is injective, and such that for each component B of ∂M, $(f|_B)_*\colon \pi_1(B) \to \pi_1(B')$ is injective, where B' is the component of ∂N containing $f(B)$. Then there is a homotopy $f_t\colon (M, \partial M) \to (N, \partial N)$ such that $f_0 = f$ and either*

(i) $f_1\colon M \to N$ is a covering map,

(ii) M is an I-bundle over a closed surface and $f_1(M) \subset \partial N$, or

(iii) N and M are solid tori $D^2 \times S^1$ and

$$\begin{aligned} f_1\colon D^2 \times S^1 &\to D^2 \times S^1 \\ (r, \theta, \phi) &\mapsto (r, p\theta + q\phi, s\phi), \end{aligned}$$

where $p, s \in \mathbb{Z} - \{0\}$ and $q \in \mathbb{Z}$.

If, for any components B of ∂M, $f|_B$ is already a cover, then we may assume that $f_t|_B = f|_B$ for all t.

Corollary 12.4. *Let M and N be closed Haken 3-manifolds. Then a homotopy equivalence between them can be homotoped to a homeomorphism.*

In order to prove Theorem 12.3, we will need the following result. Its proof can be found in Chapter 10 of Hempel's book (Theorem 10.6).

Theorem 12.5. *Let M be a compact orientable irreducible 3-manifold, and suppose that $\pi_1(M)$ contains a finite index subgroup isomorphic to the fundamental group of a closed surface other than S^2 or $\mathbb{R}P^2$. Then M is an I-bundle over some closed surface.*

Lemma 12.6. *Suppose that $f\colon (M, \partial M) \to (N, \partial N)$ is a map between connected orientable irreducible 3-manifolds with non-empty boundary such that*

1. *$f|_{\partial M}$ is not injective, and its restriction to each component of ∂M is a cover,*
2. *$f_*\colon \pi_1(M) \to \pi_1(N)$ is an isomorphism, and*
3. *M is compact.*

Then either (ii) or (iii) of Theorem 12.3 holds.

Proof. This proof was omitted in the lectures. The argument is very similar to that of Lemma 12.2. By (1), there are two points in ∂M mapping to the same point in ∂N, and there is a path $\gamma\colon I \to M$ joining them. Then $f \circ \gamma$ is a loop in N. By (2), there is a loop β in M based at $\gamma(0)$ such that $f_*([\beta]) = [f \circ \gamma]^{-1} \in \pi_1(N, f\gamma(0))$. Then $\alpha = \beta.\gamma$ is a path $(I, \partial I) \to (M, \partial M)$ such that

($*$) $\alpha(0) \neq \alpha(1)$ and $f \circ \alpha$ is a homotopically trivial loop in N.

For $i = 0$ and 1, let J_i be the component of ∂M containing $\alpha(i) = x_i$. (Possibly, $J_0 = J_1$.) Let K be the component of ∂N containing $y = f \circ \alpha(0) = f \circ \alpha(1)$. Let $p\colon (\tilde{M}, \tilde{x}_0) \to (M, x_0)$ be the covering of M such that $p_*\pi_1(\tilde{M}, \tilde{x}_0) = i_{0*}\pi_1(J_0, x_0)$, where $i_0\colon J_0 \to M$ is the inclusion map. Lift α to a path $\tilde{\alpha}$ starting at $\tilde{x}_0$ and ending at $\tilde{x}_1$, say. Let $\tilde{J}_i$ be the component of $\partial\tilde{M}$ containing $\tilde{\alpha}(i)$. There is a commutative diagram

$$\begin{array}{ccc}
\pi_1(J_0, x_0) & \xrightarrow{(f|_{J_0})_*} & \pi_1(K, y) \\
\Big\downarrow i_{0*} & & \Big\downarrow \\
\pi_1(M, x_0) & \xrightarrow{f_*} & \pi_1(N, y) \\
\Big\uparrow \psi_\alpha \circ i_{1*} & & \Big\uparrow \\
\pi_1(J_1, x_1) & \xrightarrow{(f|_{J_1})_*} & \pi_1(K, y)
\end{array}$$

where i_0 and i_1 are the relevant inclusion maps, and ψ_α is the 'change of basepoint map' $\pi_1(M, x_1) \to \pi_1(M, x_0)$ sending a loop ℓ based at x_1 to $\alpha.\ell.\alpha^{-1}$. Commutativity of the lower half of the diagram follows from the fact that $f \circ \alpha$ is homotopically trivial. Since $f|_{J_i}$ is a finite sheeted covering, we conclude that $\psi_\alpha i_{1*}\pi_1(J_1, x_1) \cap i_{0*}\pi_1(J_0, x_0)$ has finite index in each term. This intersection is $p_*\psi_\alpha\tilde{i}_{1*}\pi_1(\tilde{J}_1, \tilde{x}_1)$, where $\tilde{i}_1\colon J_1 \to \tilde{M}$ is the inclusion map. Hence, we conclude that $\tilde{J}_1$ is compact and that a nonzero power of each loop in $\tilde{J}_0$ is freely homotopic in $\tilde{M}$ to a loop in $\tilde{J}_1$. Note also that $p|_{\tilde{J}_0}\colon \tilde{J}_0 \to J_0$ is a homeomorphism.

Case 1. There is some path α satisfying $(*)$ which also satisfies

$(**)$ α is not homotopic (keeping $\partial\alpha$ fixed) to a path in ∂M.

Then $\tilde{J}_0 \neq \tilde{J}_1$. Otherwise, since $\pi_1(\tilde{J}_0) \to \pi_1(\tilde{M})$ is surjective, $\tilde{\alpha}$ would homotope into $\tilde{J}_0$ and projecting this homotopy would contradict $(**)$. In addition, we can conclude that $\tilde{J}_0$ is incompressible in $\tilde{M}$. If not, we could write $\pi_1(\tilde{M})$ as a free product, with $\pi_1(\tilde{J}_1)$ conjugate to a subgroup of one factor. This is not possible, since $\pi_1(\tilde{J}_1)$ maps to a subgroup of finite index in $\pi_1(\tilde{M})$. Thus, $\tilde{i}_{0*}\colon \pi_1(\tilde{J}_0) \to \pi_1(\tilde{M})$ is injective, and therefore an isomorphism. Hence, $\tilde{i}_0$ is a homotopy equivalence, as all the higher homotopy groups of $\tilde{J}_0$ and $\tilde{M}$ are trivial. We have the exact sequence

$$0 \to H_3(\tilde{M}, \tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2) \to H_2(\tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2) \to H_2(\tilde{M}; \mathbb{Z}_2).$$

Since $H_2(\tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$ and $H_2(\tilde{M}; \mathbb{Z}_2) \cong H_2(\tilde{J}_0; \mathbb{Z}_2) \cong \mathbb{Z}_2$, we deduce that $H_3(\tilde{M}, \tilde{J}_0 \cup \tilde{J}_1; \mathbb{Z}_2)$ is non-trivial, and hence $\tilde{M}$ is compact. Hence, $i_{0*}\pi_1(J_0)$ has finite index in $\pi_1(M)$. By Theorem 12.5, M is an I-bundle over a closed surface.

We now obtain a homotopy retracting M into ∂N. The map $i_{0*}\colon \pi_1(J_0) \to \pi_1(M)$ is a injection. For otherwise, J_0 is compressible and hence so is $\tilde{J}_0$, which we already know not be the case. This implies that $\pi_1(K) \to \pi_1(N)$ is an injection. For if some element of $\pi_1(K)$ were sent to the identity in $\pi_1(N)$, then some power of it would lie in the image of $\pi_1(J_0)$ and hence J_0 would not be π_1-injective. Consider the covering $q\colon \tilde{N} \to N$ corresponding to $f_*\pi_1(J_0)$. An appropriate lifting $\tilde{f}$ of f takes J_0 and J_1 into a component $\tilde{K}$ of $q^{-1}(K)$ (the same component since $[f \circ \alpha] = 1$). The map $\pi_1(\tilde{K}) \to \pi_1(\tilde{N})$ is necessarily surjective, and it is injective since $\pi_1(K) \to \pi_1(N)$ is injective. All higher homotopy groups of $\tilde{K}$ and $\tilde{N}$ are trivial, and so the inclusion of $\tilde{K}$ into $\tilde{N}$ is a homotopy equivalence. Hence, there is a deformation retract of $\tilde{N}$ onto $\tilde{K}$, by a homotopy $\rho_t\colon \tilde{N} \to \tilde{N}$. Then $f_t = q \circ \rho_t \circ \tilde{f}$ homotopes M into ∂N. Hence we have conclusion (ii) of Theorem 12.3.

Note that if F sends two different components of ∂M to the same component of ∂N, then we may find a path α satisfying $(*)$ and $(**)$. Hence, the theorem holds in this case. On the other hand, if F sends distinct components of ∂M to distinct components of ∂N, then the right-hand map in the following diagram is injective:

$$\begin{array}{ccc} H_3(M, \partial M) & \longrightarrow & H_2(\partial M) \\ \Big\downarrow \tilde{f}_* & & \Big\downarrow \tilde{f}_* \\ H_3(N, \partial N) & \longrightarrow & H_2(\partial N) \end{array}$$

Thus, $H_3(N, \partial N)$ is non-trivial, and therefore N is compact.

Case 2. No path α satisfies both $(*)$ and $(**)$.

Then every path α satisfying $(*)$ is homotopic (keeping its endpoints fixed) to a path α_1 in ∂M. Hence, $J_0 = J_1$. The loop $f \circ \alpha_1$ is not contractible in K, since $f|_{J_0}$ is a cover onto K. However $f \circ \alpha_1$ is homotopically trivial in N. Therefore, K is compressible in N. We wish to show that K is a torus and deduce (iii) of Theorem 12.3.

If f maps two distinct components of ∂M to the same component of ∂N then there is a path β joining these components such that $f \circ \beta$ is a loop. Since f_* is surjective, we may assume that $[f \circ \beta] = 1$, and hence β satisfies $(*)$ and $(**)$. Therefore, f takes distinct components of ∂M to distinct components of ∂N. Note that $f|_{J_0}$ is not injective, since α satisfies $(*)$.

Now f is a homotopy equivalence, and so

$$\frac{\chi(\partial M)}{2} = \chi(M) = \chi(N) = \frac{\chi(\partial N)}{2}.$$

(Here, we are using the assumption that N is compact.) Let ∂M have components $J_1, \dots, J_k$, and suppose that $f|_{J_i}$ is n_i-sheeted. Then

$$\sum n_i \chi(f(J_i)) = \sum \chi(J_i) = \chi(\partial M) = \chi(\partial N) = \sum \chi(f(J_i)).$$

So, $n_i = 1$ unless $\chi(f(J_i)) = 0$. Since $n_1 > 1$, $\chi(K) = 0$ and so K is a torus. We have already established that K is compressible. Thus N is a solid torus, since this is the only irreducible 3-manifold with a compressible torus boundary component. Also, J_0 is a torus and $\pi_1(J_0) \to \pi_1(M) \cong \pi_1(N) \cong \mathbb{Z}$. Therefore, J_0 is compressible and M is a solid torus. It is now straightforward to homotope f so that is in the form required by (iii) of Theorem 12.3. □

Proof of Theorem 12.3. Consider first the case where ∂N is non-empty. Let

$p: \tilde{N} \to N$ be the cover such that $p_* \pi_1(\tilde{N}) = f_* \pi_1(M)$. Consider the lift

$$\begin{array}{ccc} & & \tilde{N} \\ & \overset{\tilde{f}}{\nearrow} & \big\downarrow p \\ M & \xrightarrow{f} & N \end{array}$$

Then $\tilde{f}_*$ is an isomorphism. We will show that $\tilde{f}$ may homotoped so that either (i), (ii) or (iii) holds. This suffices to prove the theorem. For if (i) holds for $\tilde{f}$, then $p \circ \tilde{f}$ is a covering map. If (ii) holds for $\tilde{f}$, then composing the homotopy with p, we may homotope M into ∂N. Suppose that (iii) holds for $\tilde{f}$. In particular, $\tilde{N}$ is a solid torus. Then, N must have compressible boundary. Since it is irreducible, and has boundary a torus, it must be a solid torus. Therefore, p is a standard finite covering of the solid torus over itself. The composition of this with $\tilde{f}$ is a map as in (iii), as required.

We are assuming that the restriction of f to each boundary component of M is π_1-injective onto its image component of ∂M. Hence, by Theorem 12.1, we may homotope $f|_{\partial M}$ to a covering. So, $\tilde{f}|_{\partial M}$ is a cover. If $\tilde{f}|_{\partial M}$ sends two distinct components of ∂M to the same component of $\partial \tilde{N}$, then, by Lemma 12.6, (ii) or (iii) of 12.3 hold. So, we may assume that $\tilde{f}|_{\partial M}$ sends distinct components of ∂M to distinct components of $\partial \tilde{N}$. Hence, the right-hand map in the following diagram is injective.

$$\begin{array}{ccc} H_3(M, \partial M) & \longrightarrow & H_2(\partial M) \\ \big\downarrow \tilde{f}_* & & \big\downarrow \tilde{f}_* \\ H_3(\tilde{N}, \partial \tilde{N}) & \longrightarrow & H_2(\partial \tilde{N}) \end{array}$$

So, the fundamental class in $H_3(M, \partial M)$ has non-trivial image in $H_3(\tilde{N}, \partial \tilde{N})$ and hence $\tilde{N}$ is compact.

Hence, it suffices to consider the case where f_* is an isomorphism. By Lemma 12.6, we may assume that $f|_{\partial M}$ is a homeomorphism onto ∂N, for otherwise either (ii) or (iii) holds.

Let

$$N = N_1 \xrightarrow{S_1} N_2 \xrightarrow{S_2} \dots \xrightarrow{S_{n-1}} \dots N_n$$

be a hierarchy. By Theorem 11.4, we may assume that each surface has non-empty boundary. Let $F_1 = f^{-1}(S_1)$. After a homotopy of f (fixed on ∂M), we may assume that F_1 is a 2-sided incompressible surface, no component of which is a 2-sphere. We may also assume that f maps $\mathcal{N}(F_1)$ onto $\mathcal{N}(S_1)$ in way that sends fibres homeomorphically to fibres. The following diagram commutes.

$$\begin{array}{ccc} F_1 & \xrightarrow{f} & S_1 \\ \downarrow & & \downarrow \\ M & \xrightarrow{f} & N \end{array}$$

By Theorem 3.3, $\pi_1(F_1) \to \pi_1(M)$ is injective and hence $\pi_1(F_1) \to \pi_1(S_1)$ is injective. Note that the restriction of f to ∂F_1 is a homeomorphism. If any component of F_1 is a disc, so is S_1, and hence so is every component of F_1. We therefore homotope $f|_{F_1}$ keeping $f|_{\partial F_1}$ fixed, so that it is a homeomorphism on each component. If no component of F_1 is a disc, then we may apply Theorem 12.1. Note that (ii) of Theorem 12.1 cannot hold, since $f|_{\partial F_1}$ is a homeomorphism. So we may homotope $f|_{F_1}$ to a covering map, keeping $f|_{\partial F_1}$ fixed. This homotopy extends to M, so that f still sends fibres of $\mathcal{N}(F_1)$ onto fibres of $\mathcal{N}(S_1)$. The cover $f|_{F_1}$ is a homeomorphism on its boundary, and hence is a homeomorphism. Therefore, f restricts to a map $M_2 = M - \mathrm{int}(\mathcal{N}(F_1)) \to N_2$ that is a homeomorphism between the boundaries of these 3-manifolds. Applying an argument similar to that in Theorem 3.3, we get that $\pi_1(M_2) \to \pi_1(M_1)$ is injective. Hence, $M_2 \to N_2$ is π_1-injective.

Arguing inductively, we may assume that (i), (ii) or (iii) holds for $M_2 \to N_2$. However, neither (ii) nor (iii) holds, except possibly $|p| = |s| = 1$ in (iii), since $f|_{\partial M_2}$ is a homeomorphism. Thus, $f|_{M_2}$ is a cover. It is a homeomorphism near ∂M_2, and therefore f is a homeomorphism. This proves the inductive step.

The induction starts with $M_n \to N_n$, with N_n a collection of 3-balls. Since the restriction of this map to each component of ∂M_n is π_1-injective, each component of ∂M_n is a 2-sphere. But M_n is irreducible. Hence, it is a collection of 3-balls. The map may therefore be homotoped to a homeomorphism.

Suppose now that N is closed. Let S be an orientable incompressible surface in N, no component of which is a 2-sphere. Then we may homotope f so that $F = f^{-1}(S)$ is an orientable incompressible surface in M, no component of which is a 2-sphere. As above, the map $f|_F \colon F \to S$ is π_1-injective and may therefore be homotoped to a cover. Also, $f|_{M - \mathrm{int}(\mathcal{N}(F))} \colon M - \mathrm{int}(\mathcal{N}(F)) \to N - \mathrm{int}(\mathcal{N}(S))$ is

π_1-injective. Apply the theorem in the case of bounded 3-manifolds to this map. No component of $M - \text{int}(\mathcal{N}(F))$ satisfies (iii) of Theorem 12.3. If any component satisfies (ii), we may homotope f to reduce $|F|$. Therefore, we may assume that (i) holds for each component of $M - \text{int}(\mathcal{N}(F))$. We have therefore homotoped f to a cover. □