Notes on Basic 3-Manifold Topology

Allen Hatcher

These notes, originally written in the 1980's, were intended as the beginning of a book on 3-manifolds, but unfortunately that project has not progressed very far since then. A few small revisions have been made in 1999 and 2000, but much more remains to be done, both in improving the existing sections and in adding more topics. The next topic to be added will probably be Haken manifolds in §3.2. For any subsequent updates which may be written, the interested reader should check my webpage:

http://www.math.cornell.edu/˜hatcher

The three chapters here are to a certain extent independent of each other. The main exceptions are that the beginning of Chapter 1 is a prerequisite for almost everything else, while some of the later parts of Chapter 1 are used in Chapter 2.

Chapter 1. Canonical Decomposition

This chapter begins with the first general result on 3-manifolds, Kneser's theorem that every compact orientable 3-manifold M decomposes uniquely as a connected sum $M = P_1 \sharp \cdots \sharp P_n$ of 3-manifolds P_i which are prime in the sense that they can be decomposed as connected sums only in the trivial way $P_i = P_i \sharp S^3$.

After the prime decomposition, we turn in the second section to the canonical torus decomposition due to Jaco-Shalen and Johannson.

We shall work in the C^∞ category throughout. All 3-manifolds in this chapter are assumed to be connected, orientable, and compact, possibly with boundary, unless otherwise stated or constructed.

1. Prime Decomposition

Implicit in the prime decomposition theorem is the fact that S^3 is prime, otherwise one could only hope for a prime decomposition modulo invertible elements, as in algebra. This is implied by Alexander's theorem, our first topic.

Alexander's Theorem

This quite fundamental result was one of the earliest theorems in the subject:

Theorem 1.1. *Every embedded* 2*-sphere in* $\mathbb{R}^3$ *bounds an embedded* 3*-ball.*

Proof: Let $S \subset \mathbb{R}^3$ be an embedded closed surface, with $h : S \to \mathbb{R}$ the height function given by the z-coordinate. After a small isotopy of S we may assume h is a morse function with all its critical points in distinct levels. Namely, there is a small homotopy of h to such a map. Keeping the same x and y coordinates for S, this gives a small homotopy of S in $\mathbb{R}^3$. But embeddings are open in the space of all maps, so if this homotopy is chosen small enough, it will be an isotopy.

Let $a_1 < \cdots < a_n$ be noncritical values of h such that each interval $(-\infty, a_1)$, (a_1, a_2), $\cdots$, (a_n, ∞) contains just one critical value. For each i, $h^{-1}(a_i)$ consists of a number of disjoint circles in the level $z = a_i$. By the two-dimensional Schoenflies Theorem (which can be proved by the same method we are using here) each circle of $h^{-1}(a_i)$ bounds a disk in the plane $z = a_i$. Let C be an innermost circle of $h^{-1}(a_i)$, in the sense that the disk D it bounds in $z = a_i$ is disjoint from all the other circles of $h^{-1}(a_i)$. We can use D to **surger** S along C. This means that for some small $\varepsilon > 0$ we first remove from S the open annulus A consisting of points near C between the two planes $z = a_i \pm \varepsilon$, then we cap off the resulting pair of boundary circles of $S - A$ by adding to $S - A$ the disks in $z = a_i \pm \varepsilon$ which these circles bound. The result of this surgery is thus a new embedded surface, with perhaps one more component than S, if C separated S.

This surgery process can now be iterated, taking at each stage an innermost remaining circle of $h^{-1}(a_i)$, and choosing ε small enough so that the newly introduced horizontal cap disks intersect the previously constructed surface only in their boundaries. See Figure 1.1. After surgering all the circles of $h^{-1}(a_i)$ for all i, the original surface S becomes a disjoint union of closed surfaces S_j, each consisting of a number of horizontal caps together with a connected subsurface S_j' of S containing at most one critical point of h.

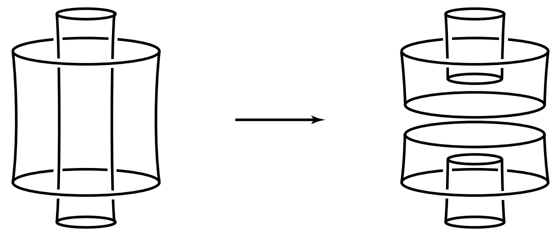


Figure 1.1

Lemma 1.2. *Each S_j is isotopic to one of seven models: the four shown in Figure 1.2 plus three more obtained by turning these upside down. Hence each S_j bounds a ball.*

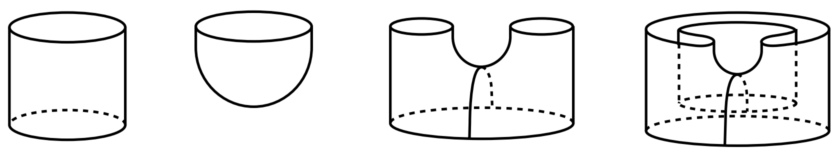


Figure 1.2

Proof: Consider the case that S_j has a saddle, say in the level $z = a$. First isotope S_j in a neighborhood of this level $z = a$ so that for some $\delta > 0$ the subsurface S_j^δ of S_j lying in $a - \delta \le z \le a + \delta$ is vertical, i.e., a union of vertical line segments, except in a neighborhood $N \subset \mathrm{int}(S_j^\delta)$ of the saddle, where S_j has the standard form of the saddles in the models. Next, isotope S_j so that its subsurface S_j' (the complement of the horizontal caps) lies in S_j^δ. This is done by pushing its horizontal caps, innermost ones first, to lie near $z = a$, as in Figure 1.3, keeping the caps horizontal throughout the deformation.

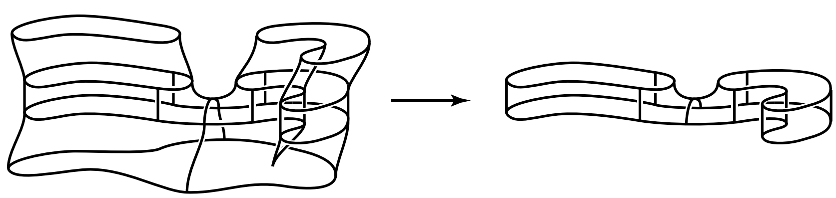


Figure 1.3

After this move S_j is entirely vertical except for the standard saddle and the horizontal caps. Viewed from above, S_j minus its horizontal caps then looks like two smooth circles, possibly nested, joined by a 1-handle, as in Figure 1.4.

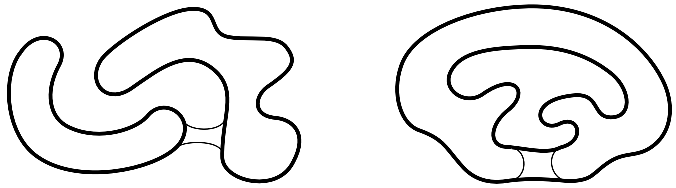


Figure 1.4

Since these circles bound disks, they can be isotoped to the standard position of one of the models, yielding an isotopy of S_j to one of the models.

The remaining cases, when S'_j has a local maximum or minimum, or no critical points, are similar but simpler, so we leave them as exercises. □

Now we assume the given surface S is a sphere. Each surgery then splits one sphere into two spheres. Reversing the sequence of surgeries, we start with a collection of spheres S_j bounding balls. The inductive assertion is that at each stage of the reversed surgery process we have a collection of spheres each bounding a ball. For the inductive step we have two balls A and B bounded by the spheres ∂A and ∂B resulting from a surgery. Letting the ε for the surgery go to 0 isotopes A and B so that $\partial A \cap \partial B$ equals the horizontal surgery disk D. There are two cases, up to changes in notation:

(i) $A \cap B = D$, with pre-surgery sphere denoted $\partial(A + B)$

(ii) $B \subset A$, with pre-surgery sphere denoted $\partial(A - B)$.

Since B is a ball, the lemma below implies that A and $A \pm B$ are diffeomorphic. Since A is a ball, so is $A \pm B$, and the inductive step is completed. □

Lemma 1.3. *Given an n-manifold M and a ball $B^{n-1} \subset \partial M$, let the manifold N be obtained from M by attaching a ball B^n via an identification of a ball $B^{n-1} \subset \partial B^n$ with the ball $B^{n-1} \subset \partial M$. Then M and N are diffeomorphic.*

Proof: Any two codimension-zero balls in a connected manifold are isotopic. Applying this fact to the given inclusion $B^{n-1} \subset \partial B^n$ and using isotopy extension, we conclude that the pair (B^n, B^{n-1}) is diffeomorphic to the standard pair. So there is an isotopy of ∂N to ∂M in N, fixed outside B^n, pushing $\partial N - \partial M$ across B^n to $\partial M - \partial N$. By isotopy extension, M and N are then diffeomorphic. □

Existence and Uniqueness of Prime Decompositions

Let M be a 3-manifold and $S \subset M$ a surface which is properly embedded, i.e., $S \cap \partial M = \partial S$, a transverse intersection. We do not assume S is connected. Deleting a small open tubular neighborhood $N(S)$ of S from M, we obtain a 3-manifold $M|S$ which we say is obtained from M by **splitting** along S. The neighborhood $N(S)$ is an interval-bundle over S, so if M is orientable, $N(S)$ is a product $S\times(-\varepsilon,\varepsilon)$ iff S is orientable.

Now suppose that M is connected and S is a sphere such that $M|S$ has two components, M_1' and M_2'. Let M_i be obtained from M_i' by filling in its boundary sphere corresponding to S with a ball. In this situation we say M is the **connected sum** $M_1 \sharp M_2$. We remark that M_i is uniquely determined by M_i' since any two ways of filling in a ball B^3 differ by a diffeomorphism of ∂B^3, and any diffeomorphism of ∂B^3 extends to a diffeomorphism of B^3. This last fact follows from the stronger assertion that any diffeomorphism of S^2 is isotopic to either the identity or a reflection (orientation-reversing), and each of these two diffeomorphisms extends over a ball. See [Cerf].

The connected sum operation is commutative by definition and has S^3 as an identity since a decomposition $M = M \sharp S^3$ is obtained by choosing the sphere S to bound a ball in M. The connected sum operation is also associative, since in a sequence of connected sum decompositions, e.g., $M_1 \sharp (M_2 \sharp M_3)$, the later splitting spheres can be pushed off the balls filling in earlier splitting spheres, so one may assume all the splitting spheres are disjointly embedded in the original manifold M. Thus $M = M_1 \sharp \cdots \sharp M_n$ means there is a collection S consisting of $n-1$ disjoint spheres such that $M|S$ has n components M_i', with M_i obtained from M_i' by filling in with balls its boundary spheres corresponding to spheres of S.

A connected 3-manifold M is called **prime** if $M = P\sharp Q$ implies $P = S^3$ or $Q = S^3$. For example, Alexander's theorem implies that S^3 is prime, since every 2-sphere in S^3 bounds a 3-ball. The latter condition, stronger than primeness, is called irreducibility: M is **irreducible** if every 2-sphere $S^2 \subset M$ bounds a ball $B^3 \subset M$. The two conditions are in fact very nearly equivalent:

Proposition 1.4. *The only orientable prime 3-manifold which is not irreducible is* $S^1\times S^2$.

Proof: If M is prime, every 2-sphere in M which separates M into two components bounds a ball. So if M is prime but not irreducible there must exist a nonseparating sphere in M. For a nonseparating sphere S in an orientable manifold M the union of a product neighborhood $S\times I$ of S with a tubular neighborhood of an arc joining $S\times\{0\}$ to $S\times\{1\}$ in the complement of $S\times I$ is a manifold diffeomorphic to $S^1\times S^2$ minus a ball. Thus M has $S^1\times S^2$ as a connected summand. Assuming M is prime, then $M = S^1\times S^2$.

It remains to show that $S^1\times S^2$ is prime. Let $S \subset S^1\times S^2$ be a separating sphere, so $S^1\times S^2|S$ consists of two compact 3-manifolds V and W each with boundary a 2-sphere. We have $\mathbb{Z} = \pi_1(S^1\times S^2) \approx \pi_1 V * \pi_1 W$, so either V or W must be simply-connected, say V is simply-connected. The universal cover of $S^1\times S^2$ can be identified with $\mathbb{R}^3 - \{0\}$, and V lifts to a diffeomorphic copy $\widetilde{V}$ of itself in $\mathbb{R}^3 - \{0\}$. The sphere

$\partial\widetilde{V}$ bounds a ball in $\mathbb{R}^3$ by Alexander's theorem. Since $\partial\widetilde{V}$ also bounds $\widetilde{V}$ in $\mathbb{R}^3$ we conclude that $\widetilde{V}$ is a ball, hence also V. Thus every separating sphere in $S^1 \times S^2$ bounds a ball, so $S^1 \times S^2$ is prime. □

Theorem 1.5. *Let M be compact, connected, and orientable. Then there is a decomposition $M = P_1 \sharp \cdots \sharp P_n$ with each P_i prime, and this decomposition is unique up to insertion or deletion of S^3's.*

Proof: The existence of prime decompositions is harder, and we tackle this first. If M contains a nonseparating S^2, this gives a decomposition $M = N \sharp S^1 \times S^2$, as we saw in the proof of Proposition 1.4. We can repeat this step of splitting off an $S^1 \times S^2$ summand as long as we have nonseparating spheres, but the process cannot be repeated indefinitely since each $S^1 \times S^2$ summand gives a $\mathbb{Z}$ summand of $H_1(M)$, which is a finitely generated abelian group since M is compact. Thus we are reduced to proving existence of prime decompositions in the case that each 2-sphere in M separates. Each 2-sphere component of ∂M corresponds to a B^3 summand of M, so we may also assume ∂M contains no 2-spheres.

We shall prove the following assertion, which clearly implies the existence of prime decompositions:

There is a bound on the number of spheres in a system S of disjoint spheres satisfying:

$(*)$ No component of $M \,|\, S$ is a punctured 3-sphere, i.e., a compact manifold obtained from S^3 by deleting finitely many open balls with disjoint closures.

Before proving this we make a preliminary observation: If S satisfies $(*)$ and we do surgery on a sphere S_i of S using a disk $D \subset M$ with $D \cap S = \partial D \subset S_i$, then at least one of the systems S', S'' obtained by replacing S_i with the spheres S_i' and S_i'' resulting from the surgery satisfies $(*)$. To see this, first perturb S_i' and S_i'' to be disjoint from S_i and each other, so that S_i, S_i', and S_i'' together bound a 3-punctured sphere P.

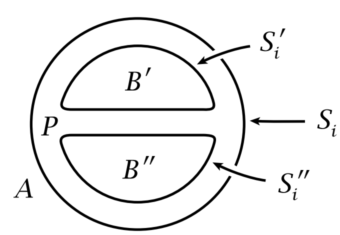


Figure 1.5

On the other side of S_i from P we have a component A of $M \,|\, S$, while the spheres S_i' and S_i'' split the component of $M \,|\, S$ containing P into pieces B', B'', and P. If both B' and B'' were punctured spheres, then $B' \cup B'' \cup P$, a component of $M \,|\, S$, would be a punctured sphere, contrary to hypothesis. So one of B' and B'', say B', is not a punctured sphere. If $A \cup P \cup B''$ were a punctured sphere, this would force A to be a punctured sphere, by Alexander's theorem. This is also contrary to hypothesis. So we conclude that neither component of $M\,|\,S'$ adjacent to S_i' is a punctured sphere, hence the sphere system S' satisfies $(*)$.

Now we prove the assertion that the number of spheres in a system S satisfying $(*)$ is bounded. Choose a smooth triangulation $\mathcal{T}$ of M. This has only finitely many simplices since M is compact. The given system S can be perturbed to be transverse to all the simplices of $\mathcal{T}$. This perturbation can be done inductively over the skeleta of $\mathcal{T}$: First make S disjoint from vertices, then transverse to edges, meeting them in finitely many points, then transverse to 2-simplices, meeting them in finitely many arcs and circles.

For a 3-simplex τ of $\mathcal{T}$, we can make the components of $S \cap \tau$ all disks, as follows. Such a component must meet $\partial\tau$ by Alexander's theorem and condition $(*)$. Consider a circle C of $S \cap \partial\tau$ which is innermost in $\partial\tau$. If C bounds a disk component of $S \cap \tau$ we may isotope this disk to lie near $\partial\tau$ and then proceed to a remaining innermost circle C. If an innermost remaining C does not bound a disk component of $S \cap \tau$ we may surger S along C using a disk D lying near $\partial\tau$ with $D \cap S = \partial D = C$, replacing S by a new system S' satisfying $(*)$, in which either C does bound a disk component of $S' \cap \tau$ or C is eliminated from $S' \cap \tau$. After finitely many such steps we arrive at a system S with $S \cap \tau$ consisting of disks, for each τ. In particular, note that no component of the intersection of S with a 2-simplex of $\mathcal{T}$ can be a circle, since this would bound disks in both adjacent 3-simplices, forming a sphere of S bounding a ball in the union of these two 3-simplices, contrary to $(*)$.

Next, for each 2-simplex σ we eliminate arcs α of $S \cap \sigma$ having both endpoints on the same edge of σ. Such an α cuts off from σ a disk D which meets only one edge of σ. We may choose α to be 'edgemost,' so that D contains no other arcs of $S \cap \sigma$, and hence $D \cap S = \alpha$ since circles of $S \cap \sigma$ have been eliminated in the previous step. By an isotopy of S supported near α we then push the intersection arc α across D, eliminating α from $S \cap \sigma$ and decreasing by two the number of points of intersection of S with the 1-skeleton of $\mathcal{T}$.

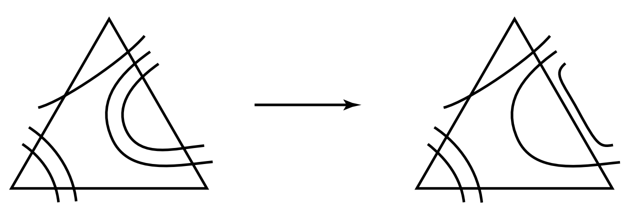


Figure 1.6

After such an isotopy decreasing the number of points of intersection of S with the 1-skeleton of $\mathcal{T}$ we repeat the first step of making S intersect all 3-simplices in disks. This does not increase the number of intersections with the 1-skeleton, so after finitely many steps, we arrive at the situation where S meets each 2-simplex only in arcs connecting adjacent sides, and S meets 3-simplices only in disks.

Now consider the intersection of S with a 2-simplex σ. With at most four exceptions the complementary regions of $S \cap \sigma$ in σ are rectangles with two opposite sides on $\partial\sigma$ and the other two opposite sides arcs of $S \cap \sigma$, as in Figure 1.7. Thus if $\mathcal{T}$ has t 2-simplices, then all but at most $4t$ of the components of $M|S$ meet all the 2-simplices of $\mathcal{T}$ only in such rectangles.

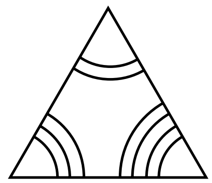


Figure 1.7

Let R be a component of $M|S$ meeting all 2-simplices only in rectangles. For a 3-simplex τ, each component of $R \cap \partial\tau$ is an annulus A which is a union of rectangles. The two circles of ∂A bound disks in τ, and A together with these two disks is a sphere bounding a ball in τ, a component of $R \cap \tau$ which can be written as $D^2 \times I$ with $\partial D^2 \times I = A$. The I-fiberings of all such products $D^2 \times I$ may be assumed to agree on their common intersections, the rectangles, to give R the structure of an I-bundle. Since ∂R consists of sphere components of S, R is either the product $S^2 \times I$ or the twisted I-bundle over $\mathbb{R}\mathrm{P}^2$. (R is the mapping cylinder of the associated ∂I-subbundle, a union of spheres which is a two-sheeted covering space of a connected base surface.) The possibility $R = S^2 \times I$ is excluded by $(*)$. Each I-bundle R is thus the mapping cylinder of the covering space $S^2 \to \mathbb{R}\mathrm{P}^2$. This is just $\mathbb{R}\mathrm{P}^3$ minus a ball, so each I-bundle R gives a connected summand $\mathbb{R}\mathrm{P}^3$ of M, hence a $\mathbb{Z}_2$ direct summand of $H_1(M)$. Thus the number of such components R of $M|S$ is bounded. Since the number of other components was bounded by $4t$, the number of components of $M|S$ is bounded. Since every 2-sphere in M separates, the number of components of $M|S$ is one more than the number of spheres in S. This finishes the proof of the existence of prime decompositions.

For uniqueness, suppose the nonprime M has two prime decompositions $M = P_1 \sharp \cdots \sharp P_k \sharp \ell(S^1 \times S^2)$ and $M = Q_1 \sharp \cdots \sharp Q_m \sharp n(S^1 \times S^2)$ where the P_i's and Q_i's are irreducible and not S^3. Let S be a disjoint union of 2-spheres in M reducing M to the P_i's, i.e., the components of $M|S$ are the manifolds $P_1, \cdots, P_k$ with punctures, plus possibly some punctured S^3's. Such a system S exists: Take for example a collection of spheres defining the given prime decomposition $M = P_1 \sharp \cdots \sharp P_k \sharp \ell(S^1 \times S^2)$ together with a nonseparating S^2 in each $S^1 \times S^2$. Note that if S reduces M to the P_i's, so does any system S' containing S.

Similarly, let T be a system of spheres reducing M to the Q_i's. If $S \cap T \neq \emptyset$, we may assume this is a transverse intersection, and consider a circle of $S \cap T$ which is innermost in T, bounding a disk $D \subset T$ with $D \cap S = \partial D$. Using D, surger the sphere S_j of S containing ∂D to produce two spheres S_j' and S_j'', which we may take to be disjoint from S_j, so that S_j, S_j', and S_j'' together bound a 3-punctured 3-sphere P. By an earlier remark, the enlarged system $S \cup S_j' \cup S_j''$ reduces M to the P_i's. Deleting S_j from this enlarged system still gives a system reducing M to the P_i's since this affects only one component of $M \,|\, S \cup S_j' \cup S_j''$, by attaching P to one of its boundary spheres, which has the net effect of simply adding one more puncture to this component.

The new system S' meets T in one fewer circle, so after finitely many steps of this type we produce a system S disjoint from T and reducing M to the P_i's. Then $S \cup T$ is a system reducing M both to the P_i's and to the Q_i's. Hence $k = m$ and the P_i's are just a permutation of the Q_i's.

Finally, to show $\ell = n$, we have $M = N \sharp \ell(S^1 \times S^2) = N \sharp n(S^1 \times S^2)$, so $H_1(M) = H_1(N) \oplus \mathbb{Z}^\ell = H_1(N) \oplus \mathbb{Z}^n$, hence $\ell = n$. □

The proof of the Prime Decomposition Theorem applies equally well to manifolds which are not just orientable, but oriented. The advantage of working with oriented manifolds is that the operation of forming $M_1 \sharp M_2$ from M_1 and M_2 is well-defined: Remove an open ball from M_1 and M_2 and then identify the two resulting boundary spheres by an orientation-reversing diffeomorphism, so the orientations of M_1 and M_2 fit together to give a coherent orientation of $M_1 \sharp M_2$. The gluing map $S^2 \to S^2$ is then uniquely determined up to isotopy, as we remarked earlier.

Thus to classify oriented compact 3-manifolds it suffices to classify the irreducible ones. In particular, one must determine whether each orientable irreducible 3-manifold possesses an orientation-reversing self-diffeomorphism.

To obtain a prime decomposition theorem for nonorientable manifolds requires very little more work. In Proposition 1.4 there are now two prime non-irreducible manifolds, $S^1 \times S^2$ and $S^1 \widetilde{\times} S^2$, the nonorientable S^2 bundle over S^1, which can also arise from a nonseparating 2-sphere. Existence of prime decompositions then works as in the orientable case. For uniqueness, one observes that $N \sharp S^1 \times S^2 = N \sharp S^1 \widetilde{\times} S^2$ if N is nonorientable. This is similar to the well-known fact in one lower dimension that connected sum of a nonorientable surface with the torus and with the Klein bottle give the same result. Uniqueness of prime decomposition can then be restored by replacing all the $S^1 \times S^2$ summands in nonorientable manifolds with $S^1 \widetilde{\times} S^2$'s.

A useful criterion for recognizing irreducible 3-manifolds is the following:

Proposition 1.6. *If $p : \widetilde{M} \to M$ is a covering space and $\widetilde{M}$ is irreducible, then so is M.*

Proof: A sphere $S \subset M$ lifts to spheres $\widetilde{S} \subset \widetilde{M}$. Each of these lifts bounds a ball in $\widetilde{M}$ since $\widetilde{M}$ is irreducible. Choose a lift $\widetilde{S}$ bounding a ball B in $\widetilde{M}$ such that no other lifts of S lie in B, i.e., $\widetilde{S}$ is an innermost lift. We claim that $p : B \to p(B)$ is a covering space. To verify the covering space property, consider first a point $x \in p(B) - S$, with

U a small ball neighborhood of x disjoint from S. Then $p^{-1}(U)$ is a disjoint union of balls in $\widetilde{M} - p^{-1}(S)$, and the ones of these in B provide a uniform covering of U. On the other hand, if $x \in S$, choose a small ball neighborhood U of x meeting S in a disk. Again $p^{-1}(U)$ is a disjoint union of balls, only one of which, $\widetilde{U}$ say, meets B since we chose $\widetilde{S}$ innermost and p is one-to-one on $\widetilde{S}$. Therefore p restricts to a homeomorphism of $\widetilde{U} \cap B$ onto a neighborhood of x in $p(B)$, and the verification that $p : B \to p(B)$ is a covering space is complete. This covering space is single-sheeted on $\widetilde{S}$, hence on all of B, so $p : B \to p(B)$ is a homeomorphism with image a ball bounded by S. $\square$

The converse of Proposition 1.6 will be proved in §3.1.

By the proposition, manifolds with universal cover S^3 are irreducible. This includes $\mathbb{R}P^3$, and more generally each 3-dimensional lens space $L_{p/q}$, which is the quotient space of S^3 under the free $\mathbb{Z}_q$ action generated by the rotation $(z_1, z_2) \mapsto (e^{2\pi i/q} z_1, e^{2p\pi i/q} z_2)$, where S^3 is viewed as the unit sphere in $\mathbb{C}^2$.

For a product $M = S^1 \times F^2$, or more generally any surface bundle $F^2 \to M \to S^1$, with F^2 a compact connected surface other than S^2 or $\mathbb{R}P^2$, the universal cover of $M - \partial M$ is $\mathbb{R}^3$, so such an M is irreducible.

Curiously, the analogous covering space assertion with 'irreducible' replaced by 'prime' is false, since there is a 2-sheeted covering $S^1 \times S^2 \to \mathbb{R}P^3 \sharp \mathbb{R}P^3$. Namely, $\mathbb{R}P^3 \sharp \mathbb{R}P^3$ is the quotient of $S^1 \times S^2$ under the identification $(x, y) \sim (\rho(x), -y)$ with ρ a reflection of the circle. This quotient can also be described as the quotient of $I \times S^2$ where (x, y) is identified with $(x, -y)$ for $x \in \partial I$. In this description the 2-sphere giving the decomposition $\mathbb{R}P^3 \sharp \mathbb{R}P^3$ is $\{1/2\} \times S^2$.

Exercises

1. Prove the (smooth) Schoenflies theorem in $\mathbb{R}^2$: An embedded circle in $\mathbb{R}^2$ bounds an embedded disk.
2. Show that for compact M^3 there is a bound on the number of 2-spheres S_i which can be embedded in M disjointly, with no S_i bounding a ball and no two S_i's bounding a product $S^2 \times I$.
3. Use the method of proof of Alexander's theorem to show that every torus $T \subset S^3$ bounds a solid torus $S^1 \times D^2 \subset S^3$ on one side or the other. (This result is also due to Alexander.)
4. Develop an analog of the prime decomposition theorem for splitting a compact irreducible 3-manifolds along disks rather than spheres. In a similar vein, study the operation of splitting nonorientable manifolds along $\mathbb{R}P^2$'s with trivial normal bundles.
5. Show: If $M^3 \subset \mathbb{R}^3$ is a compact submanifold with $H_1(M) = 0$, then $\pi_1(M) = 0$.

2. Torus Decomposition

Beyond the prime decomposition, there is a further canonical decomposition of irreducible compact orientable 3-manifolds, splitting along tori rather than spheres. This was discovered only in the mid 1970's, by Johannson and Jaco-Shalen, though in the simplified geometric version given here it could well have been proved in the 1930's. (A 1967 paper of Waldhausen comes very close to this geometric version.) Perhaps the explanation for this late discovery lies in the subtlety of the uniqueness statement. There are counterexamples to a naive uniqueness statement, involving a class of manifolds studied extensively by Seifert in the 1930's. The crucial observation, not made until the 1970's, was that these Seifert manifolds give rise to the only counterexamples.

Existence of Torus Decompositions

A properly embedded surface $S \subset M^3$ is called **2-sided** or **1-sided** according to whether its normal I-bundle is trivial or not. A 2-sided surface without S^2 or D^2 components is called **incompressible** if for each disk $D \subset M$ with $D \cap S = \partial D$ there is a disk $D' \subset S$ with $\partial D' = \partial D$. See Figure 1.8. Thus, surgery on S cannot produce a simpler surface, but only splits off an S^2 from S.

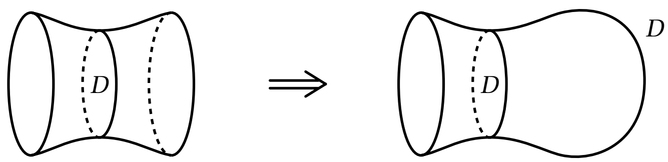


Figure 1.8

A disk D with $D \cap S = \partial D$ will sometimes be called a **compressing disk** for S, whether or not a disk $D' \subset S$ with $\partial D' = \partial D$ exists.

As a matter of convenience, we extend the definition of incompressibility to allow S to have disk components which are not isotopic rel boundary to disks in ∂M.

Here are some preliminary facts about incompressible surfaces:

(1) A surface is incompressible iff each of its components is incompressible. The 'if' half of this is easy. For the 'only if,' let $D \subset M$ be a compressing disk for one component S_i of S. If D meets other components of S, surger D along circles of $D \cap S$, innermost first as usual, to produce a new D with the same boundary circle, and meeting S only in this boundary circle. Then apply incompressibility of S.

(2) A connected 2-sided surface S which is not a sphere or disk is incompressible if the map $\pi_1(S) \to \pi_1(M)$ induced by inclusion is injective. For if $D \subset M$ is a compressing disk, then ∂D is nullhomotopic in M, hence also in S by the π_1 assumption. Then it is a standard fact that a nullhomotopic embedded circle in a surface bounds a disk in the surface. Note that it is enough for the two inclusions of S into $M | S$ to be injective on π_1.

The converse is also true, as is shown in Corollary 3.3. For 1-sided surfaces these two conditions for incompressibility are no longer equivalent, π_1-injectivity being strictly stronger in general; see the exercises. We emphasize, however, that in these notes we use the term 'incompressible' only in reference to 2-sided surfaces.

(3) There are no incompressible surfaces in $\mathbb{R}^3$, or equivalently in S^3. This is immediate from the converse to (2), but can also be proved directly, as follows. As we saw in the proof of Alexander's theorem, there is a sequence of surgeries on S along horizontal disks in $\mathbb{R}^3$ converting S into a disjoint union of spheres. Going through this sequence of surgeries in turn, either a surgery disk exhibits S as compressible, or it splits S into two surfaces one of which is a sphere. This sphere bounds balls on each side in S^3, and we can use one of these balls to isotope S to the other surface produced by the surgery. At the end of the sequence of surgeries we have isotoped S to a collection of spheres, but the definition of incompressibility does not allow spheres.

(4) A 2-sided torus T in an irreducible M is compressible iff either T bounds a solid torus $S^1 \times D^2 \subset M$ or T lies in a ball in M. For if T is compressible there is a surgery of T along some disk D which turns T into a sphere. This sphere bounds a ball $B \subset M$ by the assumption that M is irreducible. There are now two cases: If $B \cap D = \varnothing$ then reversing the surgery glues together two disks in ∂B to create a solid torus bounded by T. The other possibility is that $D \subset B$, and then $T \subset B$. Note that if $M = S^3$ the ball B can be chosen disjoint from D, so the alternative $D \subset B$ is not needed. Thus using statement (3) above we obtain the result, due to Alexander, that a torus in S^3 bounds a solid torus on one side or the other.

(5) If $S \subset M$ is incompressible, then M is irreducible iff $M|S$ is irreducible. For suppose M is irreducible. Then a 2-sphere in $M|S$ bounds a ball in M, which must be disjoint from S by statement (3) above, so $M|S$ is irreducible. Conversely, given an $S^2 \subset M$, consider a circle of $S \cap S^2$ which is innermost in S^2, bounding a disk $D \subset S^2$ with $D \cap S = \partial D$. By incompressibility of S, ∂D bounds a disk $D' \subset S$. The sphere $D \cup D'$ bounds a ball $B \subset M$ if $M|S$ is irreducible. We must have $B \cap S = D'$, otherwise the component of S containing D' would be contained in B, contrary to statement (3). Isotoping S^2 by pushing D across B to D', and slightly beyond, eliminates the circle ∂D from $S \cap S^2$, plus any other circles which happen to lie in D'. Repeating this step, we eventually get $S^2 \subset M|S$, so S^2 bounds a ball and M is irreducible.

Proposition 1.7. *For a compact irreducible M there is a bound on the number of components in a system $S = S_1 \cup \cdots \cup S_n$ of disjoint closed connected incompressible surfaces $S_i \subset M$ such that no component of $M|S$ is a product $T \times I$ with T a closed surface.*

Proof: This follows the scheme of the proof of existence of prime decompositions. First, perturb S to be transverse to a triangulation of M and perform the following

two steps to simplify the intersections of S with 2-simplices σ^2 and 3-simplices σ^3:
(1) Make all components of $S \cap \sigma^3$ disks. In the proof of prime decomposition, this was done by surgery, but now the surgeries can be achieved by isotopy. Namely, given a surgery disk $D \subset M$ with $D \cap S = \partial D$, incompressibility gives a disk $D' \subset S$ with $\partial D' = \partial D$. The sphere $D \cup D'$ bounds a ball $B \subset M$ since M is irreducible. We have $B \cap S = D'$, otherwise a component of S would lie in B. Then isotoping S by pushing D' across B to D and a little beyond replaces S by one of the two surfaces produced by the surgery.

Note that Step (1) eliminate circles of $S \cap \sigma^2$, since such a circle would bound disks in both adjacent σ^3's, producing a sphere component of S.
(2) Eliminate arcs of $S \cap \sigma^2$ with both endpoints on the same edge of σ^2, by isotopy of S.

After these simplifications, components of $M \,|\, S$ meeting 2-simplices only in rectangles are I-bundles (disjoint from ∂M), as before. Trivial I-bundles are ruled out by hypothesis, assuming $n \geq 2$ so that M is not a fiber bundle with fiber $S = S_1$. Nontrivial I-bundles are tubular neighborhoods of 1-sided (hence nonseparating) surfaces $T_1, \cdots, T_m$, say. We may assume M is connected, and then reorder the components S_i of S so that $M \,|\, (S_1 \cup \cdots \cup S_k)$ is connected and each of $S_{k+1}, \cdots, S_n$ separates $M \,|\, (S_1 \cup \cdots \cup S_k)$. The surfaces $S_1, \cdots, S_k, T_1, \cdots, T_m$ give linearly independent elements of $H_2(M;\mathbb{Z}_2)$, for a linear relation among them with $\mathbb{Z}_2$ coefficients would imply that some subcollection of them forms the boundary of a 3-dimensional submanifold of M. (Consider simplicial homology with respect to a triangulation of M in which all these surfaces are subcomplexes.) This is impossible since the complement of the collection is connected, hence also the complement of any subcollection.

Thus $k + m$ is bounded by the dimension of $H_2(M;\mathbb{Z}_2)$. The number of components of $M \,|\, S$, which is $n - k + 1$, is bounded by $m + 4t$, t being the number of 2-simplices in the given triangulation. Combining these bounds, we obtain the inequalities $n + 1 \leq k + m + 4t \leq 4t + \dim H_2(M;\mathbb{Z}_2)$. This gives a bound on n, the number of S_i's. □

A properly embedded surface $S \subset M$ is called ∂-**parallel** if it is isotopic, fixing ∂S, to a subsurface of ∂M. By isotopy extension this is equivalent to saying that S splits off a product $S \times [0,1]$ from M with $S = S \times \{0\}$. An irreducible manifold M is called **atoroidal** if every incompressible torus in M is ∂-parallel.

Corollary 1.8. *In a compact connected irreducible M there exists a finite collection T of disjoint incompressible tori such that each component of $M \,|\, T$ is atoroidal.*

Proof: Construct inductively a sequence of disjoint incompressible tori $T_1, T_2, \cdots$ in M by letting T_i be an incompressible torus in the manifold $M_i = M \,|\, (T_1 \cup \cdots \cup T_{i-1})$ which is not ∂-parallel in M_i, if M_i is not atoroidal. The claim is that this procedure must stop at some finite stage. In view of how M_i is constructed from M_{i-1} by splitting along a torus which is not ∂-parallel, there are just two ways that some M_i can have a component which is a product $S\times I$ with S a closed surface: Either $i = 1$ and $M_1 = M = S\times I$, or $i = 2$ and $M_2 = S\times I$ with S a torus. In the latter case M is a torus bundle with T_1 as a fiber. Thus if the process of constructing T_i's does not terminate, we obtain collections $T_1 \cup \cdots \cup T_i$ satisfying the conditions of the theorem but with arbitrarily large i, a contradiction. $\square$

Now we describe an example of an irreducible M where this torus decomposition into atoroidal pieces is not unique, the components of $M|T$ for the two splittings being in fact non-homeomorphic.

Example. For $i = 1,2,3,4$, let M_i be a solid torus with ∂M_i decomposed as the union of two annuli A_i and A_i' each winding $q_i > 1$ times around the S^1 factor of M_i. The union of these four solid tori, with each A_i' glued to A_{i+1} (subscripts mod 4), is the manifold M. This contains two tori $T_1 = A_1 \cup A_3$ and $T_2 = A_2 \cup A_4$. The components of $M|T_1$ are $M_1 \cup M_2$ and $M_3 \cup M_4$, and the components of $M|T_2$ are $M_2 \cup M_3$ and $M_4 \cup M_1$. The fundamental group of $M_i \cup M_{i+1}$ has presentation $\langle x_i, x_{i+1} \mid x_i^{q_i} = x_{i+1}^{q_{i+1}}\rangle$. The center of this amalgamated free product is cyclic, generated by the element $x_i^{q_i} = x_{i+1}^{q_{i+1}}$. Factoring out the center gives quotient $\mathbb{Z}_{q_i} * \mathbb{Z}_{q_{i+1}}$, with abelianization $\mathbb{Z}_{q_i} \oplus \mathbb{Z}_{q_{i+1}}$. Thus if the q_i's are for example distinct primes, then no two of the manifolds $M_i \cup M_{i+1}$ are homeomorphic.

Results from later in this section will imply that M is irreducible, T_1 and T_2 are incompressible, and the four manifolds $M_i \cup M_{i+1}$ are atoroidal. So the splittings $M|T_1$ and $M|T_2$, though quite different, both satisfy the conclusions of the Corollary.

Manifolds like this M which are obtained by gluing together solid tori along noncontractible annuli in their boundaries belong to a very special class of manifolds called Seifert manifolds, which we now define. A **model Seifert fibering** of $S^1\times D^2$ is a decomposition of $S^1\times D^2$ into disjoint circles, called **fibers**, constructed as follows. Starting with $[0,1]\times D^2$ decomposed into the segments $[0,1]\times\{x\}$, identify the disks $\{0\}\times D^2$ and $\{1\}\times D^2$ via a $2\pi p/q$ rotation, for $p/q \in \mathbb{Q}$ with p and q relatively prime. The segment $[0,1]\times\{0\}$ then becomes a fiber $S^1\times\{0\}$, while every other fiber in $S^1\times D^2$ is made from q segments $[0,1]\times\{x\}$. A **Seifert fibering** of a 3-manifold M is a decomposition of M into disjoint circles, the **fibers**, such that each fiber has a neighborhood diffeomorphic, preserving fibers, to a neighborhood of a fiber in some model Seifert fibering of $S^1\times D^2$. A **Seifert manifold** is one which possesses a Seifert fibering.

Each fiber circle C in a Seifert fibering of a 3-manifold M has a well-defined **multiplicity**, the number of times a small disk transverse to C meets each nearby fiber. For example, in the model Seifert fibering of $S^1\times D^2$ with $2\pi p/q$ twist, the fiber $S^1\times\{0\}$ has multiplicity q while all other fibers have multiplicity 1. Fibers of multiplicity 1 are **regular** fibers, and the other fibers are **multiple** (or **singular**, or **exceptional**). The multiple fibers are isolated and lie in the interior of M. The quotient space B of M obtained by identifying each fiber to a point is a surface, compact if M is compact, as is clear from the model Seifert fiberings. The projection $\pi : M \to B$ is an ordinary fiber bundle on the complement of the multiple fibers. In particular, $\pi : \partial M \to \partial B$ is a circle bundle, so ∂M consists of tori and Klein bottles, or just tori if M is orientable.

The somewhat surprising fact is that Seifert manifolds account for *all* the non-uniqueness in torus splittings, according to the following theorem, which is the main result of this section.

Theorem 1.9. *For a compact irreducible orientable 3-manifold M there exists a collection $T \subset M$ of disjoint incompressible tori such that each component of $M|T$ is either atoroidal or a Seifert manifold, and a minimal such collection T is unique up to isotopy.*

Here 'minimal' means minimal with respect to inclusions of such collections. Note the strength of the uniqueness: up to isotopy, not just up to homeomorphism of M, for example. The orientability assumption can be dropped if splittings along incompressible Klein bottles are also allowed, and the definition of 'atoroidal' is modified accordingly. For simplicity we shall stick to the orientable case, however.

Before proving the uniqueness statement we need to study Seifert manifolds a little more. This is done in the following subsection.

Incompressible Surfaces in Seifert Manifolds

There is a relative form of incompressibility which is often very useful: A surface $S \subset M$ is ∂-**incompressible** if for each disk $D \subset M$ such that ∂D decomposes as the union of two arcs α and β meeting only at their common endpoints, with $D \cap S = \alpha$ and $D \cap \partial M = \beta$ (such a D is called a ∂-**compressing disk** for S) there is a disk $D' \subset S$ with $\alpha \subset \partial D'$ and $\partial D' - \alpha \subset \partial S$. See Figure 1.9.

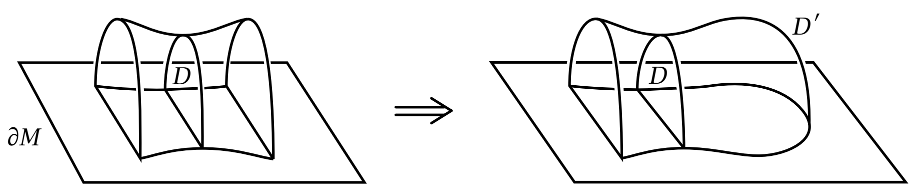


Figure 1.9

A surface which is both incompressible and ∂-incompressible we shall call **essential**. We leave it as an exercise to show that a surface S is essential iff each component of S is essential. Also, as in the absolute case, S is ∂-incompressible if $\pi_1(S, \partial S) \to \pi_1(M, \partial M)$ is injective for all choices of basepoint in ∂S.

Example. Let us show that the only essential surfaces in the manifold $M = S^1 \times D^2$ are disks isotopic to meridian disks $\{x\} \times D^2$. For let S be a connected essential surface in M. We may isotope S so that all the circles of ∂S are either meridian circles $\{x\} \times \partial D^2$ or are transverse to all meridian circles. By a small perturbation S can also be made transverse to a fixed meridian disk D_0. Circles of $S \cap D_0$ can be eliminated, innermost first, by isotopy of S using incompressibility of S and irreducibility of M. After this has been done, consider an edgemost arc α of $S \cap D_0$. This cuts off a ∂-compressing disk D from D_0, so α also cuts off a disk D' from S, meeting ∂M in an arc γ. The existence of D' implies that the two ends of γ lie on the same side of the meridian arc $\beta = D \cap \partial M$ in ∂M. But this is impossible since γ is transverse to all meridians and therefore proceeds monotonically through the meridian circles of ∂M. Thus we must have S disjoint from D_0, so ∂S consists of meridian circles. Moreover, S is incompressible in $M|D_0$, a 3-ball, so S must be a disk since each of its boundary circles bounds a disk in the boundary of the ball, and pushing such a disk slightly into the interior of the ball yields a compressing disk for S. It follows from Alexander's theorem that any two disks in a ball having the same boundary are isotopic fixing the boundary, so S is isotopic to a meridian disk in M.

Lemma 1.10. *Let S be a connected incompressible surface in the irreducible 3-manifold M, with ∂S contained in torus boundary components of M. Then either S is essential or it is a ∂-parallel annulus.*

Proof: Suppose S is ∂-compressible, with a ∂-compressing disk D meeting S in an arc α which does not cut off a disk from S. Let β be the arc $D \cap \partial M$, lying in a torus component T of ∂M. The circles of $S \cap T$ do not bound disks in T, otherwise incompressibility of S would imply S was a disk, but disks are ∂-incompressible. Thus β lies in an annulus component A of $T|\partial S$. If β were trivial in A, cutting off a disk D', incompressibility applied to the disk $D \cup D'$ would imply that α cuts off a disk from S, contrary to assumption; see Figure 1.10(a).

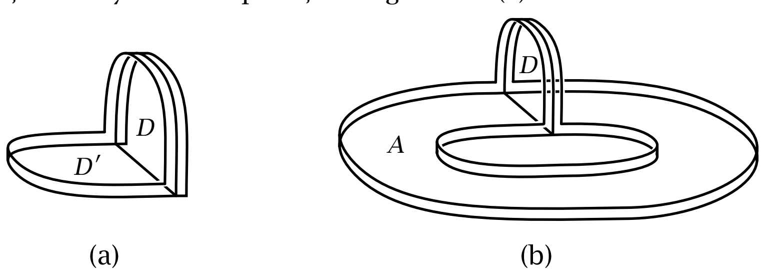


Figure 1.10

So β joins the two components of ∂A. If both of these components are the same circle of ∂S, i.e., if $S \cap T$ consists of a single circle, then S would be 1-sided. For consider the normals to S pointing into D along α. At the two points of $\partial\alpha$ these normals point into β, hence point to opposite sides of the circle $S \cap T$.

Thus the endpoints of β must lie in different circles of ∂S, and we have the configuration in Figure 1.10(b). Let N be a neighborhood of $\partial A \cup \alpha$ in S, a 3-punctured sphere. The circle $\partial N - \partial S$ bounds an obvious disk in the complement of S, lying near $D \cup A$, so since S is incompressible this boundary circle also bounds a disk in S. Thus S is an annulus. Surgering the torus $S \cup A$ via D yields a sphere, which bounds a ball in M since M is irreducible. Hence $S \cup A$ bounds a solid torus and S is ∂-parallel, being isotopic to A rel ∂S. □

Proposition 1.11. *If M is a connected compact irreducible Seifert-fibered manifold, then any essential surface in M is isotopic to a surface which is either* **vertical**, *i.e., a union of regular fibers, or* **horizontal**, *i.e., transverse to all fibers.*

Proof: Let $C_1, \cdots, C_n$ be fibers of the Seifert fibering which include all the multiple fibers together with at least one regular fiber if there are no multiple fibers. Let M_0 be M with small fibered open tubular neighborhoods of all the C_i's deleted. Thus M_0 is a circle bundle $M_0 \to B_0$ over a compact connected surface B_0 with nonempty boundary. Choose disjoint arcs in B_0 whose union splits B_0 into a disk, and let A be the pre-image in M_0 of this collection of arcs, a union of disjoint vertical annuli $A_1, \cdots, A_m$ in M_0 such that the manifold $M_1 = M_0 | A$ is a solid torus.

The circles of ∂S are nontrivial in ∂M since S is incompressible and M is irreducible. Hence S can be isotoped so that the circles of ∂S are either vertical or horizontal in each component torus or Klein bottle of ∂M. Vertical circles of S may be perturbed to be disjoint from A. We may assume S meets the fibers C_i transversely, and hence meets the neighborhoods of these fibers in disks transverse to fibers. So the surface $S_0 = S \cap M_0$ also has each its boundary circles horizontal or vertical.

Circles of $S \cap A$ bounding disks in A can be eliminated by isotopy of S in the familiar way, using incompressibility of S and irreducibility of M. Arcs of $S \cap A$ with both endpoints on the same component of ∂A can be eliminated as follows. An edgemost such arc α cuts off a disk D from A. If the two endpoints of α lie in a component of $\partial M_0 - \partial M$, then S can be isotoped across D to eliminate two intersection points with a fiber C_i. The other possibility, that the two endpoints of α lie in ∂M, cannot actually occur, for if it did, the disk D would be a ∂-compressing disk for S in M, a configuration ruled out by the monotonicity argument in the Example preceding Lemma 1.10, with the role of meridians in that argument now played by vertical circles.

So we may assume the components of $S \cap A$ are either vertical circles or horizontal arcs. If we let $S_1 = S_0 | A$ in $M_0 | A = M_1$, it follows that ∂S_1 consists entirely of horizontal or vertical circles in the torus ∂M_1. We may assume S_1 is incompressible in M_1. For let $D \subset M_1$ be a compressing disk for S_1. Since S is incompressible, ∂D bounds a disk $D' \subset S$. If this does not lie in S_1, we can isotope S by pushing D'

across the ball bounded by $D \cup D'$, thereby eliminating some components of $S \cap A$.

Since S_1 is incompressible, its components are either ∂-parallel annuli or are essential in the solid torus M_1, hence are isotopic to meridian disks by the Example before Lemma 1.10. If S_1 contains a ∂-parallel annulus with horizontal boundary, then this annulus has a ∂-compressing disk D with $D \cap \partial M_1$ a vertical arc in ∂M_0. As in the earlier step when we eliminated arcs of $S \cap A$ with endpoints on the same component of ∂A, this leads to an isotopy of S removing intersection points with a fiber C_i. So we may assume all components of S_1 are either ∂-parallel annuli with vertical boundary or disks with horizontal boundary.

Since vertical circles in ∂M_1 cannot be disjoint from horizontal circles, S_1 is either a union of ∂-parallel annuli with vertical boundary, or a union of disks with horizontal boundary. In the former case S_1 can be isotoped to be vertical, staying fixed on ∂S_1 where it is already vertical. This isotopy gives an isotopy of S to a vertical surface. In the opposite case that S_1 consists of disks with horizontal boundary, isotopic to meridian disks in M_1, we can isotope S_1 to be horizontal fixing ∂S_1, and this gives an isotopy of S to a horizontal surface. ◻

Vertical surfaces are easy to understand: They are circle bundles since they are disjoint from multiple fibers by definition, hence they are either annuli, tori, or Klein bottles.

Horizontal surfaces are somewhat more subtle. For a horizontal surface S the projection $\pi : S \to B$ onto the base surface of the Seifert fibering is a branched covering, with a branch point of multiplicity q for each intersection of S with a singular fiber of multiplicity q. (To see this, look in a neighborhood of a fiber, where the map $S \to B$ is equivalent to the projection of a number of meridian disks onto B, clearly a branched covering.) For this branched covering $\pi : S \to B$ there is a useful formula relating the Euler characteristics of S and B,

$$\chi(B) - \chi(S)/n = \sum_i (1 - 1/q_i)$$

where n is the number of sheets in the branched cover and the multiple fibers of M have multiplicities $q_1, \cdots, q_m$. To verify this formula, triangulate B so that the images of the multiple fibers are vertices, then lift this to a triangulation of S. Counting simplices would then yield the usual formula $\chi(S) = n\chi(B)$ for an n-sheeted unbranched cover. In the present case, however, a vertex in B which is the image of a fiber of multiplicity q_i has n/q_i pre-images in S, rather than n. This yields a modified formula $\chi(S) = n\chi(B) + \sum_i(-n + n/q_i)$, which is equivalent to the one above.

There is further structure associated to a horizontal surface S in a Seifert-fibered manifold M. Assume S is connected and 2-sided. (If S is 1-sided, it has an I-bundle neighborhood whose boundary is a horizontal 2-sided surface.) Since $S \to B$ is onto,

S meets all fibers of M, and $M|S$ is an I-bundle. The local triviality of this I-bundle is clear if one looks in a model-fibered neighborhood of a fiber. The associated ∂I-subbundle consists of two copies of S, so the I-bundle is the mapping cylinder of a 2-sheeted covering projection $S \amalg S \to T$ for some surface T. There are two cases, according to whether S separates M or not:

(1) If $M|S$ is connected, so is T, and $S \amalg S \to T$ is the trivial covering $S \amalg S \to S$, so $M|S = S \times I$ and hence M is a bundle over S^1 with fiber S. The surface fibers of this bundle are all horizontal surfaces in the Seifert fibering.

(2) If $M|S$ has two components, each is a twisted I-bundle over a component T_i of T, the mapping cylinder of a nontrivial 2-sheeted covering $S \to T_i$, $i = 1, 2$. The parallel copies of S in these mapping cylinders, together with T_1 and T_2, are the leaves of a foliation of M. These leaves are the 'fibers' of a natural projection $p : M \to I$, with T_1 and T_2 the pre-images of the endpoints of I. This 'fiber' structure on M is not exactly a fiber bundle, so let us give it a new name: a **semi-bundle**. Thus a semi-bundle $p : M \to I$ is the union of two twisted I-bundles $p^{-1}[0, {}^1\!/_2]$ and $p^{-1}[{}^1\!/_2, 1]$ glued together by a homeomorphism of the fiber $p^{-1}({}^1\!/_2)$. For example, in one lower dimension, the Klein bottle is a semi-bundle with fibers S^1, since it splits as the union of two Möbius bands. More generally, one could define semi-bundles with base any manifold with boundary.

The techniques we have been using can also be applied to determine which Seifert manifolds are irreducible:

Proposition 1.12. *A compact connected Seifert-fibered manifold M is irreducible unless it is $S^1 \times S^2$, $S^1 \tilde{\times} S^2$, or $\mathbb{R}P^3 \sharp \mathbb{R}P^3$.*

Proof: We begin by observing that if M is reducible then there is a horizontal sphere in M not bounding a ball. This is proved by imitating the argument of the preceding proposition, with S now a sphere not bounding a ball in M. The only difference is that when incompressibility was used before, e.g., to eliminate trivial circles of $S \cap A$, we must now use surgery rather than isotopy. Such surgery replaces S with a pair of spheres S' and S''. If both S' and S'' bounded balls, so would S, as we saw in the proof of Alexander's theorem, so we may replace S by one of S', S'' not bounding a ball. With these modifications in the proof, we eventually get a sphere which is either horizontal or vertical, but the latter cannot occur since S^2 is not a circle bundle.

If S is a horizontal sphere in M, then as we have seen, M is either a sphere bundle or a sphere semi-bundle. The only two sphere bundles are $S^1 \times S^2$ and $S^1 \tilde{\times} S^2$. A sphere semi-bundle is two copies of the twisted I-bundle over $\mathbb{R}P^2$ glued together via a diffeomorphism of S^2. Such a diffeomorphism is isotopic to either the identity or the antipodal map. The antipodal map extends to a diffeomorphism of the I-bundle $\mathbb{R}P^2 \tilde{\times} I$, so both gluings produce the same manifold, $\mathbb{R}P^3 \sharp \mathbb{R}P^3$. $\square$

Note that the three manifolds $S^1 \times S^2$, $S^1 \tilde{\times} S^2$, and $\mathbb{R}P^3 \sharp \mathbb{R}P^3$ do have Seifert fiberings. Namely, $S^1 \tilde{\times} S^2$ is $S^2 \times I$ with the two ends identified via the antipodal map, so the I-bundle structure on $S^2 \times I$ gives $S^1 \tilde{\times} S^2$ a circle bundle structure; and the I-bundle structures on the two halves $\mathbb{R}P^2 \tilde{\times} I$ of $\mathbb{R}P^3 \sharp \mathbb{R}P^3$, which are glued together by the identity, give it a circle bundle structure.

Now we can give a converse to Proposition 1.11:

Proposition 1.13. *Let M be a compact irreducible Seifert-fibered* 3*-manifold. Then every* 2*-sided horizontal surface $S \subset M$ is essential. The same is true of every connected* 2*-sided vertical surface except:*

(a) *a torus bounding a solid torus with a model Seifert fibering, containing at most one multiple fiber, or*

(b) *an annulus cutting off from M a solid torus with the product fibering.*

Proof: For a 2-sided horizontal surface S we have noted that the Seifert fibering induces an I-bundle structure on $M\,|\,S$, so $M\,|\,S$ is the mapping cylinder of a 2-sheeted covering $S \amalg S \to T$ for some surface T. Being a covering space projection, this map is injective on π_1, so the inclusion of the ∂I-subbundle into the I-bundle is also injective on π_1. Therefore S is incompressible. (In case S is a disk, $M\,|\,S$ is $D^2 \times I$, so S is clearly not ∂-parallel.) Similarly, ∂-incompressibility follows from injectivity of relative π_1's.

Now suppose S is a compressible 2-sided vertical surface, with a compressing disk D which does not cut off a disk from S. Then D is incompressible in $M\,|\,S$, and can therefore be isotoped to be horizontal. The Euler characteristic formula in the component of $M\,|\,S$ containing D takes the form $\chi(B) - 1/n = \sum_i (1 - 1/q_i)$. The right-hand side is non-negative and $\partial B \neq \emptyset$, so $\chi(B) = 1$ and B is a disk. Each term $1 - 1/q_i$ is at least $1/2$, so there can be at most one such term, and so at most one multiple fiber. Therefore this component of $M\,|\,S$ is a solid torus with a model Seifert fibering and S is its torus boundary. (If S were a vertical annulus in its boundary, S would be incompressible in this solid torus.)

Similarly, if S is a ∂-compressible vertical annulus there is a ∂-compressing disk D with horizontal boundary, and D may be isotoped to be horizontal in its interior as well. Again D must be a meridian disk in a solid torus component of $M\,|\,S$ with a model Seifert fibering. In this case there can be no multiple fiber in this solid torus since ∂D meets S in a single arc. ◻

Note that the argument just given shows that the only Seifert fiberings of $S^1 \times D^2$ are the model Seifert fiberings.

Uniqueness of Torus Decompositions

We need three preliminary lemmas:

Lemma 1.14. *An incompressible, ∂-incompressible annulus in a compact connected Seifert-fibered M can be isotoped to be vertical, after possibly changing the Seifert fibering if $M = S^1 \times S^1 \times I$, $S^1 \times S^1 \tilde{\times} I$ (the twisted I-bundle over the torus), $S^1 \tilde{\times} S^1 \times I$ (the Klein bottle cross I), or $S^1 \tilde{\times} S^1 \tilde{\times} I$ (the twisted I-bundle over the Klein bottle).*

Proof: Suppose S is a horizontal annulus in M. If S does not separate M then $M|S$ is the product $S \times I$, and so M is a bundle over S^1 with fiber S, the mapping torus $M \times I/\{(x,0) \sim (\varphi(x),1)\}$ of a diffeomorphism $\varphi : S \to S$. There are only four isotopy classes of diffeomorphisms of $S^1 \times I$, obtained as the composition of either the identity or a reflection in each factor, so φ may be taken to preserve the S^1 fibers of $S = S^1 \times I$. This S^1-fibering of S then induces a circle bundle structure on M in which S is vertical. The four choices of φ give the four exceptional manifolds listed.

If S is separating, $M|S$ is two twisted I-bundles over a Möbius band, each obtained from a cube by identifying a pair of opposite faces by a 180 degree twist. Each twisted I-bundle is thus a model Seifert fibering with a multiplicity 2 singular fiber. All four possible gluings of these two twisted I-bundles yield the same manifold M, with a Seifert fibering over D^2 having two singular fibers of multiplicity 2, with S vertical. This manifold is easily seen to be $S^1 \tilde{\times} S^1 \tilde{\times} I$. $\square$

Lemma 1.15. *Let M be a compact connected Seifert manifold with ∂M orientable. Then the restrictions to ∂M of any two Seifert fiberings of M are isotopic unless M is $S^1 \times D^2$ or one of the four exceptional manifolds in Lemma 1.14.*

Proof: Let M be Seifert-fibered, with $\partial M \neq \emptyset$. First we note that M contains an incompressible, ∂-incompressible vertical annulus A unless $M = S^1 \times D^2$. Namely, take $A = \pi^{-1}(\alpha)$ where α is an arc in the base surface B which is either nonseparating (if $B \neq D^2$) or separates the images of multiple fibers (if $B = D^2$ and there are at least two multiple fibers). This guarantees incompressibility and ∂-incompressibility of A by Proposition 1.13. Excluding the exceptional cases in Lemma 1.14, A is then isotopic to a vertical annulus in any other Seifert fibering of M, so the two Seifert fiberings can be isotoped to agree on ∂A, hence on the components of ∂M containing ∂A. Since α could be chosen to meet any component of ∂B, the result follows. $\square$

Lemma 1.16. *If M is compact, connected, orientable, irreducible, and atoroidal, and M contains an incompressible, ∂-incompressible annulus meeting only torus components of ∂M, then M is a Seifert manifold.*

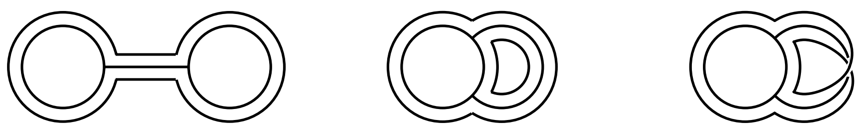
Proof: Let A be an annulus as in the hypothesis. There are three possibilities, indicated in Figure 1.11 below:

(a) A meets two different tori T_1 and T_2 in ∂M, and $A \cup T_1 \cup T_2$ has a neighborhood N which is a product of a 2-punctured disk with S^1.

(b) A meets only one torus T_1 in ∂M, the union of A with either annulus of $T_1 \,|\, \partial A$ is a torus, and $A \cup T_1$ has a neighborhood N which is a product of a 2-punctured disk with S^1.

(c) A meets only one torus T_1 in ∂M, the union of A with either annulus of $T_1 \,|\, \partial A$ is a Klein bottle, and $A \cup T_1$ has a neighborhood N which is an S^1 bundle over a punctured Möbius band.

In all three cases N has the structure of a circle bundle $N \to B$ with A vertical.



(a) (b) (c)

Figure 1.11

By hypothesis, the tori of $\partial N - \partial M$ must either be compressible or ∂-parallel in M. Suppose D is a nontrivial compressing disk for $\partial N - \partial M$ in M, with ∂D a nontrivial loop in a component torus T of $\partial N - \partial M$. If $D \subset N$, then N would be a solid torus $S^1 \times D^2$ by Proposition 1.13, which is impossible since N has more than one boundary torus. So $D \cap N = \partial D$. Surgering T along D yields a 2-sphere bounding a ball $B^3 \subset M$. This B^3 lies on the opposite side of T from N, otherwise we would have $N \subset B^3$ with T the only boundary component of N. Reversing the surgery, B^3 becomes a solid torus outside N, bounded by T.

The other possibility for a component T of $\partial N - \partial M$ is that it is ∂-parallel in M, cutting off a product $T \times I$ from M. This $T \times I$ cannot be N since $\pi_1 N$ is nonabelian, the map $\pi_1 N \to \pi_1 B$ induced by the circle bundle $N \to B$ being a surjection to a free group on two generators. So $T \times I$ is an external collar on N, and hence can be absorbed into N.

Thus M is N with solid tori perhaps attached to one or two tori of $\partial N - \partial M$. The meridian circles $\{x\} \times \partial D^2$ in such attached $S^1 \times D^2$'s are not isotopic in ∂N to circle fibers of N, otherwise A would be compressible in M (recall that A is vertical in N). Thus the circle fibers wind around the attached $S^1 \times D^2$'s a non-zero number of times in the S^1 direction. Hence the circle bundle structure on N extends to model Seifert fiberings of these $S^1 \times D^2$'s, and so M is Seifert-fibered. □

Proof of Theorem 1.9: Only the uniqueness statement remains to be proved. So let $T = T_1 \cup \cdots \cup T_m$ and $T' = T_1' \cup \cdots \cup T_n'$ be two minimal collections of disjoint incompressible tori splitting M into manifolds M_j and M_j', respectively, which are either atoroidal or Seifert-fibered. We may suppose T and T' are nonempty, otherwise the theorem is trivial.

Having perturbed T to meets T' transversely, we isotope T and T' to eliminate circles of $T \cap T'$ which bound disks in either T or T', by the usual argument using incompressibility and irreducibility.

For each M_j the components of $T' \cap M_j$ are then tori or annuli. The annulus components are incompressible in M_j since they are noncontractible in T' and T' is incompressible. Annuli of $T' \cap M_j$ which are ∂-compressible are then ∂-parallel, by Lemma 1.10, so they can be eliminated by isotopy of T'.

A circle C of $T \cap T'$ lies in the boundary of annulus components A_j of $T' \cap M_j$ and A_k of $T' \cap M_k$ (possibly $A_j = A_k$ or $M_j = M_k$). By Lemma 1.16 M_j and M_k are Seifert-fibered. If $M_j \neq M_k$ Lemma 1.14 implies that we can isotope Seifert fiberings of M_j and M_k so that A_j and A_k are vertical. In particular, the two fiberings of the torus component T_i of T containing C induced from the Seifert fiberings of M_j and M_k have a common fiber C. Therefore these two fiberings can be isotoped to agree on T_i, and so the collection T is not minimal since T_i can be deleted from it.

Essentially the same argument works if $M_j = M_k$: If we are not in the exceptional cases in Lemma 1.14, then the circle C is isotopic in T_i to fibers of each of the two induced fiberings of T_i, so these two fiberings are isotopic, and T_i can be deleted from T. In the exceptional case $M_j = S^1 \times S^1 \times I$, if we have to rechoose the Seifert fibering to make A_j vertical, then as we saw in the proof of Lemma 1.14, the new fibering is simply the trivial circle bundle over $S^1 \times I$. The annulus A_j, being vertical, incompressible, and ∂-incompressible, must then join the two boundary tori of M_j, since its projection to the base surface $S^1 \times I$ must be an arc joining the two boundary components of $S^1 \times I$. The two boundary circles of A_j in T_i either coincide or are disjoint, hence isotopic, so once again the two induced fiberings of T_i are isotopic and T_i can be deleted from T. The other exceptional cases in Lemma 1.14 cannot arise since M_j has at least two boundary tori.

Thus $T \cap T' = \varnothing$. If any component T_i of T lies in an atoroidal M'_j it must be isotopic to a component T'_i of T'. After an isotopy we then have $T_i = T'_i$ and M can be split along this common torus of T and T', and we would be done by induction. Thus we may assume each T_i lies in a Seifert-fibered M'_j, and similarly, each T'_i lies in a Seifert-fibered M_j. These Seifert-fibered manifolds all have nonempty boundary, so they contain no horizontal tori. Thus we may assume all the tori $T_i \subset M'_j$ and $T'_i \subset M_j$ are vertical.

Consider T'_1, contained in M_j and abutting M'_k and M'_ℓ (possibly $M'_k = M'_\ell$). If some T_i is contained in M'_k then M'_k is Seifert-fibered, by the preceding paragraph. If no T_i is contained in M'_k then $M'_k \subset M_j$, and M'_k inherits a Seifert fibering from M_j since $\partial M'_k$ is vertical in M_j. Thus in any case $M_j \cap M'_k$ has two Seifert fiberings: as a subset of M_j and as a subset of M'_k. By Lemma 1.15 these two fiberings can be isotoped to agree on T'_1, apart from the following exceptional cases:

— $M_j \cap M'_k = S^1 \times D^2$. This would have T'_1 as its compressible boundary, so this case cannot occur.

— $M_j \cap M'_k = S^1 \times S^1 \times I$. One boundary component of this is T'_1 and the other must be a T_i. (If it were a T'_i, then T' would not be minimal unless $T'_i = T'_1$, in which case $T = \varnothing$, contrary to hypothesis.) Then T'_1 can be isotoped to T_i and we would be done by induction.

— $M_j \cap M'_k = S^1 \tilde{\times} S^1 \tilde{\times} I$. This has only one boundary component, so $M'_k \subset M_j$ and we can change the Seifert fibering of M'_k to be the restriction of the Seifert fibering of M_j.

Thus we may assume the fibering of T'_1 coming from M'_k agrees with the one coming from M_j. The same argument applies with M'_ℓ in place of M'_k. So the Seifert fiberings of M'_k and M'_ℓ agree on T'_1, and T'_1 can be deleted from T'. □

Exercises

1. Show: If $S \subset M$ is a 1-sided connected surface, then $\pi_1 S \to \pi_1 M$ is injective iff $\partial N(S)$ is incompressible, where $N(S)$ is a twisted I-bundle neighborhood of S in M.

2. Call a 1-sided surface $S \subset M$ *geometrically incompressible* if for each disk $D \subset M$ with $D \cap S = \partial D$ there is a disk $D' \subset S$ with $\partial D' = \partial D$. Show that if $H_2 M = 0$ but $H_2(M;\mathbb{Z}_2) \neq 0$ then M contains a 1-sided geometrically incompressible surface which is nonzero in $H_2(M;\mathbb{Z}_2)$. This applies for example if M is a lens space $L_{p/2q}$. Note that if $q > 1$, the resulting geometrically incompressible surface $S \subset L_{p/2q}$ cannot be S^2 or $\mathbb{R}P^2$, so the map $\pi_1 S \to \pi_1 L_{p/2q}$ is not injective. (See [Frohman] for a study of geometrically incompressible surfaces in Seifert manifolds.)

3. Develop a canonical torus and Klein bottle decomposition of irreducible nonorientable 3-manifolds.

Chapter 2. Special Classes of 3-Manifolds

In this chapter we study prime 3-manifolds whose topology is dominated, in one way or another, by embedded tori. This can be regarded as refining the results of the preceding chapter on the canonical torus decomposition

1. Seifert Manifolds

Seifert manifolds, introduced in the last chapter where they play a special role in the torus decomposition, are among the best-understood 3-manifolds. In this section our goal is the classification of orientable Seifert manifolds up to diffeomorphism. We begin with the classification up to fiber-preserving diffeomorphism, which is fairly straightforward. Then we show that in most cases the Seifert fibering is unique up to isotopy, so in these cases the diffeomorphism and fiber-preserving diffeomorphism classifications coincide. But there are a few smaller Seifert manifolds, including some with non-unique fiberings, which must be treated by special techniques. Among these are the lens spaces, which we classify later in this section.

The most troublesome Seifert fiberings are those with base surface S^2 and three multiple fibers. These manifolds are too large for the lens space method to work and too small for the techniques of the general case. They have been classified by a study of the algebra of their fundamental groups, but a good geometric classification has yet to be found, so we shall not prove the classification theorem for these Seifert manifolds.

All 3-manifolds in this chapter are assumed to be orientable, compact, and connected. The nonorientable case is similar, but as usual we restrict to orientable manifolds for simplicity.

Classification of Seifert Fiberings

We begin with an explicit construction of Seifert fiberings. Let B be a compact connected surface, not necessarily orientable. Choose disjoint disks $D_1, \cdots, D_k$ in the interior of B, and let B' be B with the interiors of these disks deleted. Let $M' \to B'$ be the circle bundle with M' orientable. Thus if B' is orientable M' is the product $B' \times S^1$, and if B' is nonorientable, M' is the twisted product in which circles in B' are covered either by tori or Klein bottles in M' according to whether these circles are orientation-preserving or orientation-reversing in B'. Explicitly, B' can be constructed by identifying pairs of disjoint arcs a_i and b_i in the boundary of a disk D^2, and then we can form M' from $D^2 \times S^1$ by identifying $a_i \times S^1$ with $b_i \times S^1$ via the product of the given identification of a_i and b_i with either the identity or a reflection in the S^1 factor, whichever makes M' orientable.

Let $s: B' \to M'$ be a cross section of $M' \to B'$. For example, we can regard M' as the double of an I-bundle, that is, two copies of the I-bundle with their sub ∂I-bundles identified by the identity map, and then we can choose a cross section in one of the I-bundles. The cross section s together with a choice of orientation for the manifold M' allows us to speak unambiguously of slopes of nontrivial circles in the tori of $\partial M'$. Namely, we can choose a diffeomorphism φ of each component of $\partial M'$ with $S^1 \times S^1$ taking the cross section to $S^1 \times \{y\}$ (slope 0) and a fiber to $\{x\} \times S^1$ (slope ∞). An orientation of M' induces an orientation of $\partial M'$, and this determines φ up to simultaneous reflections of the two S^1 factors, which doesn't affect slopes. We are assuming the standard fact that each nontrivial circle in $S^1 \times S^1$ is isotopic to a unique 'linear' circle which lifts to the line $y = (p/q)x$ of slope p/q in the universal cover $\mathbb{R}^2$.

From M' we construct a manifold M by attaching k solid tori $D^2 \times S^1$ to the torus components T_i of $\partial M'$ lying over $\partial D_i \subset \partial B'$, attaching by diffeomorphisms taking a meridian circle $\partial D^2 \times \{y\}$ of $\partial D^2 \times S^1$ to a circle of some finite slope $\alpha_i/\beta_i \in \mathbb{Q}$ in T_i. The k slopes α_i/β_i determine M uniquely, since once the meridian disk $D^2 \times \{y\}$ is attached to M' there is only one way to fill in a ball to complete the attaching of $D^2 \times S^1$. The circle fibering of M' extends naturally to a Seifert fibering of M via a model Seifert fibering on each attached $D^2 \times S^1$, since the fibers of M' in T_i are not isotopic to meridian circles of the attached $D^2 \times S^1$. Namely, fibers have slope ∞, meridians have slope $\alpha_i/\beta_i \neq \infty$. Note that the singular fiber in the i^{th} $D^2 \times S^1$ has multiplicity β_i since the meridian disk of $D^2 \times S^1$ is attached to a slope α_i/β_i circle and hence meets each fiber of $\partial M'$ β_i times. Recall that the multiplicity of a singular fiber is the number of times a transverse disk meets nearby regular fibers.

We use the notation $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$ for this Seifert-fibered manifold M, where g is the genus of B, with the sign $+$ if B is orientable and $-$ if B is nonorientable, and b is the number of boundary components of B. Here 'genus' for nonorientable surfaces means the number of $\mathbb{R}\mathrm{P}^2$ connected summands. Reversing the orientation of $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$ has the effect of changing it to $M(\pm g, b; -\alpha_1/\beta_1, \cdots, -\alpha_k/\beta_k)$.

We say two Seifert fiberings are **isomorphic** if there is a diffeomorphism carrying fibers of the first fibering to fibers of the second fibering.

Proposition 2.1. *Every orientable Seifert-fibered manifold is isomorphic to one of the models* $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$. *Seifert fiberings* $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$ *and* $M(\pm g, b; \alpha_1'/\beta_1', \cdots, \alpha_k'/\beta_k')$ *are isomorphic by an orientation-preserving diffeomorphism iff, after possibly permuting indices,* $\alpha_i/\beta_i \equiv \alpha_i'/\beta_i' \bmod 1$ *for each* i *and, if* $b = 0$, $\sum_i \alpha_i/\beta_i = \sum_i \alpha_i'/\beta_i'$.

This gives the complete isomorphism classification of Seifert fiberings since the numbers $\pm g$ and b are determined by the isomorphism class of a fibering, which determines the base surface B, and the Seifert fiberings $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$ and $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k, 0)$ are the same.

Proof: Given an oriented Seifert-fibered manifold M, let M' be the complement of open solid torus model-fibered neighborhoods of fibers $C_1, \cdots, C_k$ including all multiple fibers. Choose a section s of the circle bundle $M' \to B'$. As before, this determines slopes for circles in $\partial M'$, and we see that M has the form $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$. It remains to see the effect on the α_i/β_i's of choosing a different section s.

Let a be an arc in B' with endpoints in $\partial B'$. Above this in M' lies an annulus A. We can rechoose s near A so that instead of simply crossing A transversely, it winds m times around A as it crosses. See Figure 2.1. The effect of this change in s is to add m to all slopes in the boundary torus of M' at one edge of A and subtract m from all slopes in the boundary torus at the other edge of A. In particular, if both ends of A lie in the same boundary torus there is no change in boundary slopes.

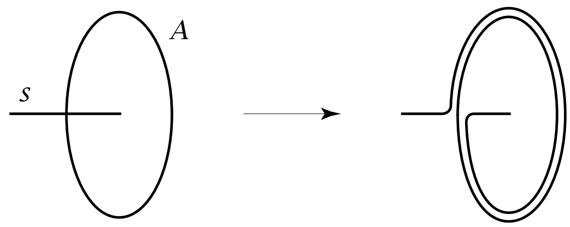


Figure 2.1

Thus if $b \neq 0$ we can choose A connecting the boundary torus near the fiber C_i with a torus in ∂M, and then change α_i/β_i by any integer, keeping all other α_j/β_j's fixed. Similarly, if $b = 0$ we can add and subtract an integer m from any two α_i/β_i's, so we can change the α_i/β_i's to any fractions which are congruent $\bmod 1$, subject only to the constraint that $\sum_i \alpha_i/\beta_i$ stays constant.

We claim that any two choices of the section s are related by a sequence of 'twist' modifications near vertical annuli A as above, together with homotopies through sections, which have no effect on slopes. To see this, take disjoint vertical annuli A_j splitting M' into a solid torus. Any two sections can be homotoped, through sections, to coincide outside a neighborhood of the A_j's. Then it is clear that near the A_j's the two sections can be homotoped to coincide, after inserting the appropriate number of twists. □

In the case of Seifert fiberings $M(\pm g, 0; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$ of closed manifolds, the invariant $\sum_i \alpha_i/\beta_i$ is called the **Euler number** of the fibering. When there are no multiple fibers we can take $k = 1$ and then the Euler number, which is an integer, is the obstruction to the existence of a section $B \to M$, i.e., the Euler number vanishes iff such a section exists. (Exercise.) More generally:

Proposition 2.2. *Let M be an orientable Seifert-fibered manifold.*

(a) *If $\partial M \neq \emptyset$, horizontal surfaces exist in M.*

(b) *If $\partial M = \emptyset$, horizontal surfaces exist iff the Euler number of the fibering is zero.*

Proof: In (a), view M as a circle bundle M_0 with model-fibered solid tori M_i attached, each M_i attaching along an annulus in its boundary. Namely, these annuli in M project to arcs in the base surface cutting off disks each containing the image of one multiple fiber. Given a positive integer n there is a horizontal surface $S_0 \subset M_0$ meeting each fiber in n points. To see this, we can regard M_0 as a quotient of a trivially fibered solid torus $S^1 \times D^2$ in which certain vertical annuli in $S^1 \times \partial D^2$ are identified. Each identification can be chosen to be either the identity or a fixed reflection in the S^1 factor. Taking n points $x_j \in S^1$ which are invariant (as a set) under the reflection, the n meridian disks $\{x_j\} \times D^2$ in $S^1 \times D^2$ give the desired surface S_0 in the quotient M_0.

Now let n be a common multiple of the multiplicities q_i of the multiple fibers in the solid tori M_i attached to M_0. In M_i let S_i be the union of n/q_i meridian disks, so S_i meets each regular fiber in n points. We can isotope S_i through horizontal surfaces so that its n arcs of intersection with the vertical annulus $M_0 \cap M_i$ match up with the n arcs of S_0 in $M_0 \cap M_i$. Then the union of S_0 with the S_i's is a horizontal surface in M.

For (b), let $M = M(\pm g, 0; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$, with section $s : B' \to M'$ as before. Let M_0 be M with a fibered solid torus neighborhood of a regular fiber in M' deleted. Suppose S_0 is a horizontal surface in M_0.

Claim. The circles of ∂S_0 in ∂M_0 have slope equal to $e(M)$, the Euler number of M.

This easily implies (b): By (a), such a surface S_0 exists. If $e(M) = 0$, S_0 extends via meridian disks in $M - M_0$ to a horizontal surface $S \subset M$. Conversely, if a horizontal surface $S \subset M$ exists, the surface $S_0 = S \cap M_0$ must have its boundary circles of slope 0 since these circles bound disks in $M - M_0$. So $e(M) = 0$.

To prove the Claim, let $M_0' = M_0 \cap M'$ and $S_0' = S_0 \cap M'$. The circles of $\partial S_0'$ in $\partial M'$ have slopes α_i/β_i, and we must check that the circles of $\partial S_0'$ in ∂M_0 have slope $\sum_i \alpha_i/\beta_i$. This we do by counting intersections of these circles with fibers and with the section s. Since S_0' is horizontal, it meets all fibers in the same number of points, say n. Intersections with s we count with signs, according to whether the slope of $\partial S_0'$ near such an intersection point is positive or negative. The total number of intersections of $\partial S_0'$ with s is zero because the two intersection points at the end of an arc of $s \cap S_0'$ have opposite sign. Thus the number of intersections of $\partial S_0'$ with s in ∂M_0 equals the number of intersections in $\partial M'$. The latter number is $\sum_i n\alpha_i/\beta_i$ since the slope of $\partial S_0'$ near the i^{th} deleted fiber is α_i/β_i, which must equal the ratio of intersection number with s to intersection number with fiber; the denominator of this ratio is n, so the numerator must be $n\alpha_i/\beta_i$.

Thus the slope of ∂S_0 is $(\sum_i n\alpha_i/\beta_i)/n = \sum_i \alpha_i/\beta_i = e(M)$. $\square$

Classification of Seifert Manifolds

Here is a statement of the main result:

Theorem 2.3. *Seifert fiberings of orientable Seifert manifolds are unique up to isomorphism, with the exception of the following fiberings:*

(a) $M(0,1;\alpha/\beta)$*, the various model Seifert fiberings of* $S^1\times D^2$.

(b) $M(0,1;{}^1\!/_2,{}^1\!/_2) = M(-1,1;)$*, two fiberings of* $S^1 \tilde{\times} S^1 \tilde{\times} I$.

(c) $M(0,0;\alpha_1/\beta_1,\alpha_2/\beta_2)$*, various fiberings of* S^3, $S^1\times S^2$*, and lens spaces.*

(d) $M(0,0;{}^1\!/_2,{}^{-1}\!/_2,\alpha/\beta) = M(-1,0;\beta/\alpha)$ *for* $\alpha,\beta \neq 0$.

(e) $M(0,0;{}^1\!/_2,{}^1\!/_2,{}^{-1}\!/_2,{}^{-1}\!/_2) = M(-2,0;)$*, two fiberings of* $S^1 \tilde{\times} S^1 \tilde{\times} S^1$.

The two Seifert fiberings of $S^1 \tilde{\times} S^1 \tilde{\times} I$ in (b) are easy to see if we view $S^1 \tilde{\times} S^1 \tilde{\times} I$ as obtained from $S^1\times I\times I$ by identifying $S^1\times I\times\{0\}$ with $S^1\times I\times\{1\}$ via the diffeomorphism φ which reflects both the S^1 and I factors. One fibering of $S^1 \tilde{\times} S^1 \tilde{\times} I$ then comes from the fibers $S^1\times\{y\}\times\{z\}$ of $S^1\times I\times I$ and the other comes from the fibers $\{x\}\times\{y\}\times I$; in the latter case the two fixed points of φ give multiplicity-two fibers.

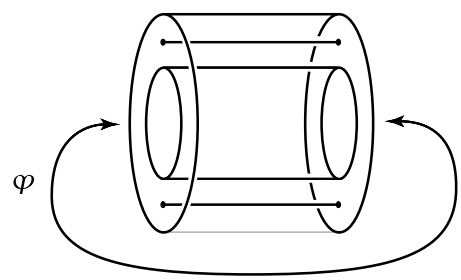


Figure 2.2

Note that the examples in (a) and (b) generate the remaining examples: The fiberings in (c) are obtained by gluing together two model fiberings from (a); (d) is obtained by gluing a model fibering into each of the two fiberings in (b); and (e) is simply the double of (b).

In most cases the Theorem is a consequence of the following:

Proposition 2.4. *If* M_1 *and* M_2 *are irreducible orientable Seifert-fibered manifolds which are diffeomorphic, then there is a fiber-preserving diffeomorphism provided that* M_1 *contains vertical incompressible,* ∂*-incompressible annuli or tori, and* M_2 *contains no horizontal incompressible,* ∂*-incompressible annuli or tori.*

Proof: First we do the case of closed manifolds. In the base surface B_1 of M_1 choose two transversely intersecting systems C and C' of disjoint 2-sided circles not passing through singular points (projections of singular fibers), such that:

(1) No circle of C or C' bounds a disk containing at most one singular point.

(2) The components of $B_1 \,|\, (C\cup C')$ are disks containing at most one singular point.

(3) No component of $B_1 \,|\, (C \cup C')$ is a disk bounded by a single arc of C and a single arc of C', and containing no singular point.

For example, one can choose C to be a single circle, then construct C' from suitably chosen arcs in $B_1 \,|\, C$, matching their ends across C. (Details left as an exercise.) Let T and T' be the collections of incompressible vertical tori in M_1 lying over C and C'.

Let $f : M_1 \to M_2$ be a diffeomorphism. Since M_2 contains no incompressible horizontal tori, we may isotope f to make $f(T)$ vertical in M_2 by Proposition 1.11. The circles of $T \cap T'$ are nontrivial in T, so we may isotope f to make the circles of $f(T \cap T')$ vertical or horizontal in each torus of $f(T)$. By condition (3) the annuli of $T' \,|\, (T \cap T')$ are incompressible and ∂-incompressible in $M_1 \,|\, T$, so we may isotope f, staying fixed on T, to make the annuli of $f(T') \,|\, f(T \cap T')$ vertical or horizontal in M_2. If any of these were horizontal, they would be part of horizontal tori in M_2, so we now have $f(T \cup T')$ vertical in M_2. Since $f(T \cap T')$ is vertical, we can can isotope f to be fiber-preserving on $T \cup T'$, and then make f fiber-preserving in a neighborhood M_1' of $T \cup T'$.

By condition (2) the components of $M_1 - M_1'$ are solid tori, so the same is true for $M_2 - M_2'$, where $M_2' = f(M_1')$. Choose an orientation for M_1 and a section for $M_1' \to B_1'$. Via f these choices determine an orientation for M_2 and a section for $M_2' \to B_2'$. Note that f induces a diffeomorphism of B_1' onto B_2', so the closed surfaces B_1 and B_2 are diffeomorphic. The fractions α_i / β_i for corresponding solid tori of $M_1 - M_1'$ and $M_2 = M_2'$ must be the same since these are the slopes of boundary circles of meridian disks, and f takes meridian disks to meridian disks (up to isotopy). Thus M_1 and M_2 have the same form $M(\pm g, 0; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$, completing the proof for closed manifolds.

For non-empty boundary the proof is similar but easier. Let T be a collection of incompressible, ∂-incompressible vertical annuli splitting M_1 into solid tori. Isotope f first to make $f(T)$ vertical, then to make f fiber-preserving in a neighborhood of $T \cup \partial M_1$. The rest of the argument now proceeds as in the closed case. □

Let us see how close Proposition 2.4 comes to proving the Theorem. Consider first the case of irreducible manifolds with non-empty boundary. Then vertical incompressible, ∂-incompressible annuli exist except in the model fiberings $M(0,1;\alpha/\beta)$. To see when horizontal annuli exist we apply the Euler characteristic formula $\chi(B) - \chi(S)/n = \sum_i (1 - 1/\beta_i)$. In the present case S is an annulus, so we have $\chi(B) = \sum_i (1 - 1/\beta_i) \ge 0$, so B is a disk, annulus, or Möbius band. If $B = D^2$, we have $1 = \sum_i (1 - 1/\beta_i)$, a sum of terms $1/2, 2/3, 3/4, \cdots$, so the only possibility is $1 = 1/2 + 1/2$ and the fibering $M(0,1;1/2,1/2)$. If $B = S^1 \times I$, we have $0 = \sum_i (1 - 1/\beta_i)$ so there are no multiple fibers and we have a product fibering $M(0,2;)$ of $S^1 \times S^1 \times I$. Similarly, if $B = S^1 \tilde{\times} I$ we have the fibering $M(-1,1;)$ of $S^1 \tilde{\times} S^1 \tilde{\times} I$. The manifolds $S^1 \times S^1 \times I$ and $S^1 \tilde{\times} S^1 \tilde{\times} I$ are not diffeomorphic since they deformation retract onto a torus and a Klein bottle, respectively, and so are of distinct homotopy types. Thus the Theorem follows in the case of irreducible manifolds with non-empty boundary.

For closed irreducible manifolds the analysis is similar but more complicated. Incompressible vertical tori exist unless the base surface B is S^2 and there are at most three multiple fibers, or B is $\mathbb{R}P^2$ and there is at most one multiple fiber. If horizontal tori exist we must have $\chi(B) \geq 0$, so there are the following cases:

— $B = S^2$ and $2 = \sum_i (1 - 1/\beta_i)$. Since $1/2 \leq 1 - 1/\beta_i < 1$ there must be either three or four multiple fibers. With four multiple fibers, the multiplicities must all be 2, so we have the fibering $M(0,0;1/2,1/2,-1/2,-1/2)$ since the Euler number must be zero. With three multiple fibers we have $1/\beta_1 + 1/\beta_2 + 1/\beta_3 = 1$ so $(\beta_1,\beta_2,\beta_3) = (2,3,6)$, $(2,4,4)$, or $(3,3,3)$, up to permutations. We leave it for the reader to check that the fibering must be isomorphic to $M(0,0;1/2,-1/3,-1/6)$, $M(0,0;1/2,-1/4,-1/4)$, or $M(0,0;2/3,-1/3,-1/3)$, allowing orientation-reversing isomorphisms.

— $B = \mathbb{R}P^2$ and $1 = \sum_i (1 - 1/\beta_i)$, implying just two multiple fibers, of multiplicity 2. Since the Euler number is zero the fibering must be $M(-1,0;1/2,-1/2)$.

— $B = S^1 \times S^1$ or $S^1 \widetilde{\times} S^1$ with no multiple fibers and Euler number zero, the fiberings $M(1,0;)$ and $M(-2,0;)$.

Thus we have the following six manifolds:

$$
\begin{aligned}
M_1 &= M(1,0;) = S^1 \times S^1 \times S^1 \\
M_2 &= M(-2,0;) = M(0,0;1/2,1/2,-1/2,-1/2) = S^1 \widetilde{\times} S^1 \widetilde{\times} S^1 \\
M_3 &= M(0,0;2/3,-1/3,-1/3) \\
M_4 &= M(0,0;1/2,-1/4,-1/4) \\
M_5 &= M(0,0;1/2,-1/3,-1/6) \\
M_6 &= M(-1,0;1/2,-1/2)
\end{aligned}
$$

Each of these seven fiberings has Euler number zero, hence does in fact contain a horizontal surface S. By the Euler characteristic formula, $\chi(S) = 0$, so S is either a torus or Klein bottle. In the latter case S is one-sided, and the boundary of a tubular neighborhood of S is a horizontal torus. So horizontal tori exist in all seven Seifert fiberings. As we saw in §1.2 following Proposition 1.11, this implies the manifolds M_1-M_6 are torus bundles or torus semi-bundles.

There remains the case of reducible Seifert manifolds. As shown in the proof of Proposition 1.13 there must exist a horizontal sphere in this case, which implies that the manifold is closed and the Euler number of the Seifert fibering is zero. By the Euler characteristic formula the base surface B must be S^2 or $\mathbb{R}P^2$, with at most three multiple fibers in the first case and at most one multiple fiber in the second case. In the latter case there is in fact no multiple fiber since the Euler number is zero, so we have the Seifert fibering $M(-1,0;)$, and it is not hard to see that this manifold is $\mathbb{R}\mathrm{P}^3 \sharp \mathbb{R}\mathrm{P}^3$. In the case $B = S^2$ it is an exercise with fractions to rule out the possibility of three multiple fibers, using the Euler characteristic formula and the fact that the Euler number is zero. Then if there are at most two multiple fibers, we are in the exceptional case (c) of the classification theorem, with a Seifert fibering $M(0,0;\alpha/\beta,-\alpha/\beta)$ of $S^1\times S^2$.

To complete the main classification theorem, there remain three things to do:

(1) Show the manifolds M_1-M_6 above are all distinct, and have only the seven Seifert fiberings listed.
(2) Classify the Seifert manifolds which possess fiberings over S^2 with at most two multiple fibers.
(3) Show that the different Seifert fiberings over S^2 with three multiple fibers are all distinct manifolds, distinct also from the manifolds in (2).

We shall do (2) in the following subsection, and in §2.2 we shall do (1) as part of a general classification of torus bundles and torus semi-bundles. As for (3), the only way which seems to be known for doing this is to show these manifolds are all distinguished by their fundamental groups, apart from a few special cases where geometric techniques are available. We shall not cover this in the present version of these notes; see [Orlik] for a proof.

Classification of Lens Spaces

Let $L_{p/q}$ be the manifold obtained by attaching two solid tori $S^1\times D^2$ together by a diffeomorphism $\varphi: S^1\times\partial D^2 \to S^1\times\partial D^2$ sending a meridian $\{x\}\times\partial D^2$ to a circle of slope p/q, where we use the convention that a meridian has slope ∞ and a longitude $S^1\times\{y\}$ has slope 0. The fraction p/q determines $L_{p/q}$ completely since circles in $S^1\times\partial D^2$ are determined up to isotopy by their slopes, and once the meridian disk of the second $S^1\times D^2$ has been attached to the first $S^1\times D^2$ there is only one way to attach the remaining 3-ball. Note that $L_{1/0} = S^1\times S^2$, and $L_{0/1} = S^3$ since $S^3 = \partial(D^2\times D^2) = \partial D^2\times D^2 \cup D^2\times\partial D^2$.

The manifolds $L_{p/q}$ are Seifert-fibered in many ways, by taking model fiberings on the two $S^1\times D^2$'s whose restrictions to $S^1\times\partial D^2$ correspond under $\varphi\in SL_2(\mathbb{Z})$; a model fibering on the second $S^1\times D^2$ with boundary fibers of slope u/v extends to a model fibering on the first $S^1\times D^2$ provided $\varphi(u/v)\neq 1/0$. These Seifert fiberings have base surface S^2 and at most two multiple fibers. Conversely, such a Seifert-fibered manifold is the union of two solid tori with boundaries identified, hence is an $L_{p/q}$.

We shall exclude the special case $L_{1/0} = S^1\times S^2$ from now on.

The manifold $L_{p/q}$ can also be obtained as the quotient space $S^3/\mathbb{Z}_q$ where $\mathbb{Z}_q$ is the group of rotations of S^3 generated by $\rho(z_1,z_2) = (e^{2\pi i/q}z_1, e^{2\pi ip/q}z_2)$, where S^3 is regarded as the unit sphere in $\mathbb{C}^2$. This action can be pictured in $\mathbb{R}^3$, with the point at infinity added to get $S^3 = S^1\times D^2 \cup D^2\times S^1$, as a $1/q$ rotation of the first factor and a p/q rotation of the second factor.

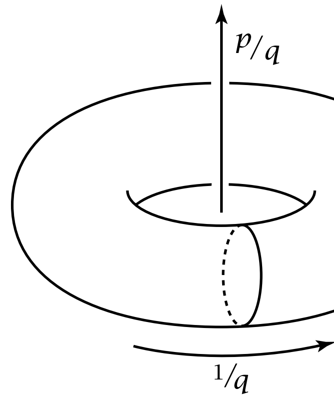


Figure 2.3

Each quotient $S^1\times D^2/\mathbb{Z}_q$ and $D^2\times S^1/\mathbb{Z}_q$ is a solid torus, so the quotient $S^3/\mathbb{Z}_q$ has the form $L_{p'/q'}$. To see that $p'/q' = p/q$, note that slope p circles on the common boundary torus $S^1\times S^1$ are invariant under ρ, hence become longitudes in $S^1\times D^2/\mathbb{Z}_q$. The boundary circle of a meridian disk of $D^2\times S^1/\mathbb{Z}_q$ then intersects the longitude of $S^1\times D^2/\mathbb{Z}_q$ p times and the meridian q times, hence has slope p/q, up to a sign determined by orientations.

In particular $\pi_1 L_{p/q} \approx \mathbb{Z}_q$. This can also be seen directly from the definition of $L_{p/q}$ since the meridian disk of the second solid torus is attached to $S^1\times D^2$ along a circle which wraps around the S^1 factor q times, and the subsequent attaching of a 3-ball has no effect on π_1. Since $\pi_1 L_{p/q} \approx \mathbb{Z}_q$, the number q is uniquely determined by $L_{p/q}$. (We take q non-negative always.)

For p there is some ambiguity, however. First, we can rechoose the longitude in $S^1\times D^2$ by adding any number of twists around the meridian. This changes slopes by an arbitrary integer, so $L_{p'/q} = L_{p/q}$ if $p' \equiv p \bmod q$. Also, reversing the orientation of $S^1\times D^2$ changes the sign of slopes, so $L_{-p/q} = L_{p/q}$. (We are not specifying an orientation of $L_{p/q}$.) Finally, the roles of the two $S^1\times D^2$ halves of $L_{p/q}$ can be reversed, replacing φ by φ^{-1}. Taking $\varphi = \left(\begin{smallmatrix} r & q \\ s & p \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, then $\varphi^{-1} = \left(\begin{smallmatrix} p & -q \\ -s & p \end{smallmatrix}\right)$, so we obtain $L_{-r/q} = L_{p/q}$, where $r \equiv p^{-1} \bmod q$ since $pr - qs = 1$. Summarizing, $L_{p'/q} = L_{p/q}$ if $p' \equiv \pm p^{\pm 1} \bmod q$.

Theorem 2.5. *$L_{p'/q}$ is diffeomorphic to $L_{p/q}$ iff $p' \equiv \pm p^{\pm 1} \bmod q$.*

Proof: We shall show that there is, up to isotopy, only one torus T in $L_{p/q}$ bounding a solid torus on each side. This implies the theorem since such a T yields a decomposition $L_{p/q} = S^1\times D^2 \cup_\varphi S^1\times D^2$, and the only ambiguities in the slope p'/q' defining such a representation of $L_{p/q}$ are the ones considered in the paragraph preceding the theorem: choice of longitude and orientation in the boundary of the first solid torus, and switching the two solid tori.

Let T be the standard torus bounding solid tori on each side, from the definition of $L_{p/q}$. Let Σ be the core circle $S^1\times\{0\}$ of the first solid torus $S^1\times D^2$ bounded by T. The union of a meridian disk D of the second $S^1\times D^2$ with the radial segments in slices $\{x\}\times D^2$ of the first $S^1\times D^2$ joining points of ∂D to Σ is a 2-complex $\Delta\subset L_{p/q}$. Abstractly, Δ is obtained from the circle Σ by attaching a disk D^2 by a q-to-1 covering map $\partial D^2\to\Sigma$.

Let T' be another torus in $L_{p/q}$ bounding solid tori on both sides. The concentric copies of T' in these two solid tori, together with the two core circles, define a singular foliation $\mathcal{F}$ of $L_{p/q}$, whose leaves can also be described as the level sets of a function $f:L_{p/q}\to[0,1]$.

The first thing we do is put Δ in good position with respect to $\mathcal{F}$, as follows. Perturb Σ to be disjoint from the two singular leaves (circles) of $\mathcal{F}$, and so that $f\,|\,\Sigma$ is a morse function with all critical points in distinct levels. Near local maxima of $f\,|\,\Sigma$, perturb the q sheets of Δ to lie 'above' Σ, as in Figure 2.4. Likewise, at local minima of $f\,|\,\Sigma$ make Δ lie 'below' Σ.

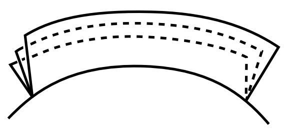


Figure 2.4

Now, keeping Δ fixed in a small neighborhood of Σ, perturb Δ to be transverse to the two singular leaves of $\mathcal{F}$ and so that away from these two leaves f is a morse function on $\Delta-\Sigma$ with all saddles in levels distinct from each other and from the levels of the critical points of $f\,|\,\Sigma$.

The foliation $\mathcal{F}$ then induces a singular foliation on Δ, hence also on D^2 via the quotient map $D^2\to\Delta$. The singularities of the foliation of D^2 are of the following types:

— 'centers' in the interior of D^2, where $f\,|\,(\Delta-\Sigma)$ has local maxima or minima, in particular where Δ meets the two singular leaves of $\mathcal{F}$.
— saddles in the interior of D^2.
— 'half-saddles' in ∂D^2, at local maxima or minima of $f\,|\,\Sigma$.

Figure 2.5 shows the various possible configurations for a singular leaf in D^2 containing a saddle or half-saddle.

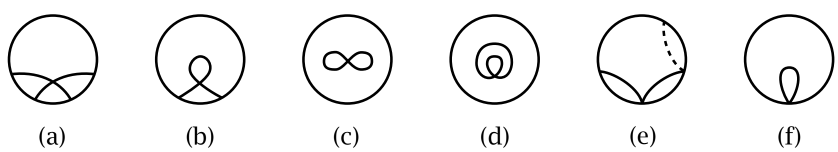


Figure 2.5

Recall that all saddles and half-saddles on Δ lie in distinct levels. However, each local maximum or minimum of $f\,|\,\Sigma$ gives q half-saddles in D^2 in the same level, and some of these may be contained in the same singular leaf, as indicated by the dashed line in case (e). The saddles and half-saddles in cases (a) and (e) are called *essential*, all others *inessential*. In (a) the singular leaf divides D^2 into four *quarter disks*, and in (e) the singular leaf cuts off two or more *half disks* from D^2.

Let D be a quarter or half disk containing no other quarter or half disks, and let α be the arc $D \cap \partial D^2$. The two endpoints of α are part of the same singular leaf, and as we move into the interior of α we have pairs of points joined by a nonsingular leaf in D^2. Continuing into the interior of α, this pairing of points of α by nonsingular leaves in D^2 continues until we reach an inessential saddle of type (b) or an inessential half-saddle, since D was chosen as a smallest quarter or half disk, and the singular leaves in cases (c) and (d) do not meet ∂D^2. An inessential saddle of type (b) presents no obstacle to continuing the pairing of points of α, which can therefore continue until we reach an inessential half-saddle. At this point, however, the two paired points of α coalesce into one point, the half-saddle point itself. We conclude from this that D contains exactly one inessential half-saddle.

In particular, note that the projection $\alpha \to \Sigma$ must be an embedding on the interior of α since $f\,|\,\Sigma$ has at least two critical points. The two endpoints of α might have the same image in Σ, however. If this happens, D must be a half disk and the two endpoints of α must be half-saddles at the same critical point of $f\,|\,\Sigma$. In this case isotoping α rel $\partial\alpha$ across D to $\partial D - \alpha$ induces an isotopy of Σ so that it lies in a leaf of $\mathcal{F}$. This is a desirable outcome, as we shall see below, and our next goal is to reduce to this case.

So suppose the disk D is a half disk and α has a half-saddle at only one of its endpoints. The configuration is as shown in Figure 2.6. Here we can isotope Σ, and hence Δ, via isotopy extension, by dragging α across D to a new position, shown in dashed lines, with $f\,|\,\Sigma$ having two fewer critical points.

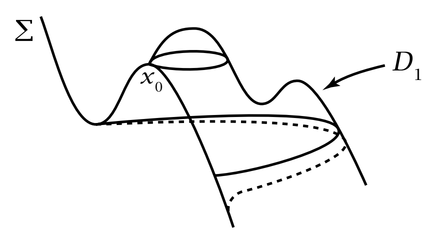


Figure 2.6

The other possibility is that D is a quarter disk, with the configuration in Figure 2.7 below. In this case we again isotope Σ by pushing α across D, dragging the other $q-1$ sheets of Δ along behind, enlarging them by parallel copies of D. This does not change the number of critical points of $f\,|\,\Sigma$, while the number of essential saddles decreases, since the essential saddle in ∂D becomes q half-saddles and no new essential saddles are created. (The inessential half-saddle in α becomes $q-1$

inessential saddles. Inessential saddles in the interior of D are also replicated. In addition, other essential saddles in $D^2 - D$ may become inessential.)

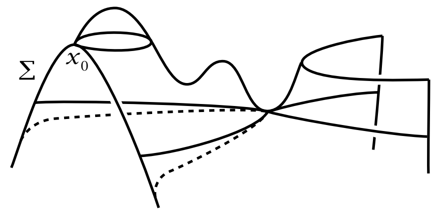


Figure 2.7

Thus after finitely many steps we are reduced to the earlier case when Σ is isotoped into a leaf of $\mathcal{F}$, which we may assume is the torus T'. The circle Σ is nontrivial in T' since it generates $\pi_1 L_{p/q} \approx \mathbb{Z}_q$ and we may assume $q > 1$. The torus T is the boundary of a solid torus neighborhood of Σ. Shrinking this neighborhood if necessary, we may assume $T \cap T'$ consists of two circles, parallel copies of Σ in T', so T intersects each of the solid tori X' and Y' bounded by T' in an annulus, and likewise T' intersects each of the solid tori X and Y bounded by T in an annulus.

We now appeal to the following fact: An annulus $A \subset S^1 \times D^2$ with ∂A consisting of nontrivial, nonmeridional circles in $S^1 \times \partial D^2$ must be ∂-parallel. For such an A must be incompressible since $\pi_1 A \to \pi_1(S^1 \times D^2)$ is injective, hence must be either ∂-incompressible or ∂-parallel. In the former case A could be isotoped to be vertical in the product Seifert fibering of $S^1 \times D^2$, but then A would again be ∂-parallel.

This applies to each of the four annuli of $T - T'$ and $T' - T$ Note that meridional boundary circles cannot occur since Σ is not nullhomotopic in $L_{p/q}$. We conclude that the two annuli of $T - T'$ can be isotoped to the two annuli of $T' - T$, as indicated in Figure 2.8, so T is isotopic to T'. □

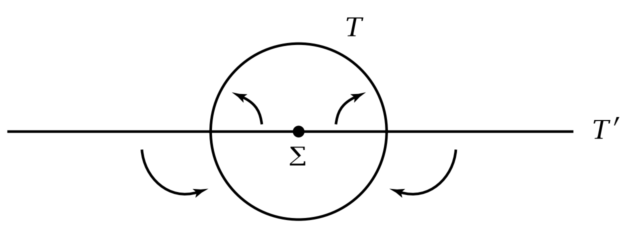


Figure 2.8

Exercises

1. What happens if $\alpha_i/\beta_i = 1/0$ in the manifold $M(\pm g, b; \alpha_1/\beta_1, \cdots, \alpha_k/\beta_k)$?

2. Torus Bundles and Semi-Bundles

In the preceding section we encountered a few Seifert manifolds which are bundles over the circle with torus fibers. Most torus bundles are not Seifert-fibered, however, though they share one essential feature with most Seifert manifolds: Their topology is dominated by incompressible tori. This makes them easy to classify, and Theorem 2.6 below reduces this classification to understanding conjugacy classes in $GL_2(\mathbb{Z})$, an algebraic problem which has a nice geometric solution described later in this section. After the case of torus bundles we move on to a closely-related class of manifolds, torus semi-bundles, which always have a torus bundle as a 2-sheeted covering space.

The torus bundles and torus semi-bundles which are not Seifert-fibered have a special interest because they are the manifolds in one of Thurston's eight geometries, called solvgeometry.

The Classification of Torus Bundles

A torus bundle has the form of a mapping torus, the quotient $M_\varphi = T\times I/(x,0) \sim (\varphi(x),1)$ for some diffeomorphism $\varphi \in GL_2(\mathbb{Z})$ of the torus T. This is because every diffeomorphism $f:T\to T$ is isotopic to a linear diffeomorphism $\varphi \in GL_2(\mathbb{Z})$. Moreover, φ is uniquely determined by f since it is essentially the map on $H_1(T) \approx \mathbb{Z}^2$ induced by f. By Proposition 1.6, M_φ is irreducible since its universal cover is $\mathbb{R}^3$.

Once again we restrict attention to orientable 3-manifolds for simplicity. For M_φ this means restricting φ to be in $SL_2(\mathbb{Z})$.

Theorem 2.6. *M_φ is diffeomorphic to M_ψ iff φ is conjugate to $\psi^{\pm 1}$ in $GL_2(\mathbb{Z})$.*

Proof: The 'if' half is easy: Applying a diffeomorphism $f:T\to T$ in each slice $T\times\{y\}$ of $T\times I$ has the effect of conjugating φ by f when we form the quotient $M_\varphi = T\times I/(x,0) \sim (\varphi(x),1)$. And by switching the two ends of $T\times I$ we see that $M_{\varphi^{-1}}$ is diffeomorphic to M_φ.

The main step in the converse is the following:

Lemma 2.7. *An incompressible surface in M_φ is isotopic to a union of torus fibers, unless φ is conjugate to $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$.*

Proof: Let $S\subset M_\varphi$ be incompressible. We may suppose S has transverse intersection with a fiber T_0 of M_φ. After we eliminate trivial circles of $S\cap T_0$ by isotopy of S in the usual way, the surface $S' = S\,|\,T_0$ is incompressible in $M_\varphi\,|\,T_0 = T\times I$. (Given a compressing disk D for S' in $T\times I$, then ∂D bounds a disk in S, which must lie in S' since any intersections with T_0 would be trivial circles of $S\cap T_0$.) If S' is ∂-compressible, it must have ∂-parallel annulus components. These can be eliminated by pushing them across T_0, decreasing the number of circles of $S\cap T_0$. So we may assume S' is incompressible and ∂-incompressible in $T\times I$.

Thinking of $T\times I$ as a trivial circle bundle, S' can be isotoped to be either horizontal or vertical, hence to consist either of vertical tori parallel to T_0, i.e., fibers of M_φ, or annuli whose boundary circles have the same slope in both ends $T\times\{0\}$ and $T\times\{1\}$. In the latter case, φ must preserve this slope in order for S' to glue together to form the original surface S. This means φ has an eigenvector in $\mathbb{Z}^2$, so is conjugate to $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$. $\square$

Continuing the proof of the theorem, suppose $f:M_\varphi\to M_\psi$ is a diffeomorphism. The lemma implies that if ψ is not conjugate to $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$, we can isotope f to take a fiber T of M_φ to a fiber of M_ψ. We may assume that T is the image of $T\times\partial I$ in the quotient mapping torus $T\times I/(x,0)\sim(\varphi(x),1)$, and similarly that $f(T)$ is the image of $T\times\partial I$ in $T\times I/(x,0)\sim(\psi(x),1)$. Then f is the quotient of a diffeomorphism $F:T\times I\to T\times I$. We may assume F takes each of $T\times\{0\}$ and $T\times\{1\}$ to themselves, rather than interchanging them, by replacing ψ by ψ^{-1} if necessary, switching the ends of $T\times I$. Let f_0 and f_1 be the restrictions of F to $T\times\{0\}$ and $T\times\{1\}$. Since $(x,0)$ is identified with $(\varphi(x),1)$, it must be that $(f_0(x),0)$ is identified with $(f_1\varphi(x),1)$. But $(f_0(x),0)$ is identified with $(\psi f_0(x),1)$, so we conclude that $f_1\varphi=\psi f_0$. This equation implies that φ and ψ are conjugate in $GL_2(\mathbb{Z})$ since f_0 and f_1 are isotopic to the same diffeomorphism in $GL_2(\mathbb{Z})$, as they induce the same map on $H_1(T)$, namely the map induced by F on $H_1(T\times I)$.

There remains the case that ψ is conjugate, hence we may assume equal, to $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$. Since ψ preserves slope ∞, M_ψ is a circle bundle with fibers the slope ∞ circles in the torus fibers of M_ψ. It is not hard to see that, with a suitable choice of orientation for M_ψ, this circle bundle, as a Seifert manifold, is $M(\varepsilon,0;n)$ with ε equal to either $+1$ or -2, the sign of ε being the same as the sign of $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$. The bundle projection $M_\psi\to S^1$ factors as a composition of two circle bundles $M_\psi\to B\to S^1$, where B is the base surface of the Seifert fibering.

We may isotope a diffeomorphism $f:M_\varphi\to M_\psi$ so that the incompressible torus $f(T)$, with T a fiber of M_φ, is either horizontal or vertical in M_ψ. If it is vertical, we may rechoose the torus fibering of M_ψ so that $f(T)$ is a fiber, by rechoosing the fibering $B\to S^1$ so that the image circle of $f(T)$ is a fiber. (In the case that B is a Klein bottle, this circle is 2-sided and nonseparating in B, so its complement must be an annulus, not a Möbius band.) In the new fibering, the new ψ is still $\pm\left(\begin{smallmatrix}1&0\\n&1\end{smallmatrix}\right)$ since this matrix was determined by the Seifert fibering $M(\varepsilon,0;n)$, as we saw. Now we have f taking T to a torus fiber of M_ψ, and the argument can be completed as before.

If $f(T)$ is horizontal we must have the Euler number n equal to zero, so $\psi=\pm I$. We cannot have $\psi=-I$ since then the base surface of $M(-2,0;0)$ is nonorientable, so the circle fibers of $M(-2,0;0)$ cannot be coherently oriented, which means that the horizontal surface $f(T)$ is separating, splitting $M(-2,0;0)$ into two twisted I-bundles. If $\psi=I$, then M_ψ is the 3-torus, and after composing f with a diffeomorphism in $GL_3(\mathbb{Z})$ we may assume the map on π_1 induced by f takes the subgroup $\pi_1 T$ into the $\mathbb{Z}^2$ subgroup corresponding to the torus fiber of M_ψ, with cyclic image in $\pi_1 B$. Then when we isotope f to make $f(T)$ horizontal or vertical, $f(T)$ cannot be horizontal since for a horizontal surface the projection to B is a finite-sheeted covering space, with π_1-image of finite index in $\pi_1 B$, hence non-cyclic. This reduces us to the previous case that $f(T)$ was vertical. ◻

We now determine which Seifert-fibered manifolds are torus bundles M_φ. If the fiber torus T is isotopic to a vertical surface, then since its complement is $T \times I$, which has only the product Seifert fibering, up to isomorphism, the Seifert fibering of M_φ must be of the form $M(1,0;n)$ or $M(-2,0;n)$, with φ conjugate to $\pm\begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$. If T is isotopic to a horizontal surface, M_φ must be one of the manifolds M_1-M_5 from §2.1:

$$M_1 = M(1,0;\,) = S^1 \times S^1 \times S^1, \qquad \text{with } \varphi = I$$
$$M_2 = M(0,0;\,1/2, 1/2, -1/2, -1/2) = S^1 \tilde{\times} S^1 \tilde{\times} S^1, \qquad \text{with } \varphi = -I$$
$$M_3 = M(0,0;\,2/3, -1/3, -1/3), \qquad \text{with } \varphi = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$$
$$M_4 = M(0,0;\,1/2, -1/4, -1/4), \qquad \text{with } \varphi = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$M_5 = M(0,0;\,1/2, -1/3, -1/6), \qquad \text{with } \varphi = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

For M_2 and M_4, φ is rotation by 180 and 90 degrees, respectively. For M_3 and M_5, if we view $\mathbb{Z}^2$ not as the usual square lattice in $\mathbb{R}^2$ but as the regular hexagonal lattice, φ is rotation by 120 and 60 degrees, respectively. So for M_1-M_5, φ is periodic, of period $1, 2, 3, 4$, and 6, respectively. We leave it for the reader to verify that with these φ's, M_φ is the given Seifert fibering, with circle fibers the images of the segments $\{x\} \times I \subset T \times I$ in the mapping torus M_φ. This just amounts to checking that the multiple fibers have the multiplicities given, since the Euler number is necessarily zero; see §2.1.

In particular, the theorem implies that none of these Seifert manifolds M_1-M_5 are diffeomorphic to each other or to the manifolds $M(1,0;n)$ or $M(-2,0;n)$ with $n \neq 0$, since their φ's are not conjugate, being of different orders. This completes the classification of Seifert-fibered torus bundles, left over from §2.1.

The Classification of Torus Semi-bundles

Let $N = S^1 \tilde{\times} S^1 \tilde{\times} I$, the twisted I-bundle over the Klein bottle. Two copies of N glued together by a diffeomorphism φ of the torus ∂N form a closed orientable manifold N_φ, which we called a torus semi-bundle in §2.1: N_φ is foliated by tori parallel to ∂N, together with a Klein bottle at the core of each copy of N. We can identify φ, or rather its isotopy class, with an element of $GL_2(\mathbb{Z})$ via a choice of coordinates in ∂N, which we can get by thinking of N as the quotient of $S^1 \times S^1 \times I$ obtained by identifying (x, y, z) with $(-x, \rho(y), \rho(z))$, where the ρ's are reflections.

Then the x and y coordinates define 'horizontal' and 'vertical' in ∂N, and an arbitrary choice of orientations of these directions completes the choice of coordinates for ∂N.

If we compose φ on the left or the right with a diffeomorphism of ∂N which extends to a diffeomorphism of N, then we get the same manifold N_φ. For example, we can compose φ with diffeomorphisms reflecting either the horizontal or vertical directions in ∂N, with matrices $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$, since the reflection ρ acting in either S^1 factor of $S^1 \times S^1 \times I$ passes down to a diffeomorphism of the quotient N. As we shall see below, these four diffeomorphisms are, up to isotopy, the only diffeomorphisms of ∂N which extend over N.

Replacing φ by φ^{-1} obviously produces the same manifold N_φ, by interchanging the roles of the two copies of N in N_φ. Thus we have proved the easier 'if' half of the following:

Theorem 2.8. *N_φ is diffeomorphic to N_ψ iff $\varphi = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \psi^{\pm 1} \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$ in $GL_2(\mathbb{Z})$, with independent choices of signs understood.*

Proof: First, consider an incompressible, ∂-incompressible surface $S \subset N$. Give N the Seifert fibering $M(-1,1;)$ as a circle bundle over the Möbius band, and isotope S to be either horizontal or vertical. If S is vertical, it consists either of tori parallel to ∂N, or nonseparating annuli with boundaries of slope ∞ in ∂N, with respect to the coordinates in ∂N chosen earlier. If S is horizontal it consists of separating annuli with boundaries of slope 0. This can be seen as follows. Orient ∂S by the positive direction of the x-coordinate in N. Let s be a section of the circle bundle N, with ∂s of slope 0. The arcs of $S \cap s$ pair off the points of $\partial S \cap \partial s$ into pairs of opposite sign, so S and s have algebraic intersection number 0. Since all the components of ∂S have parallel orientations, each component must have intersection number 0 with ∂s, hence must have slope 0.

Now let S be an incompressible surface in N_ψ. As in the proof of Lemma 2.7, we may isotope S so that for each of the two copies N_1 and N_2 of N whose union is N_ψ, $S \cap N_i$ is an incompressible, ∂-incompressible surface S_i in N_i. By the preceding paragraph, either S consists of parallel copies of the torus T splitting N_ψ into N_1 and N_2, or ψ takes slope 0 or ∞ to slope 0 or ∞. We shall see later that this means N_ψ is Seifert-fibered.

Suppose $f : N_\varphi \to N_\psi$ is a diffeomorphism, and let T be the torus in N_φ splitting N_φ into two copies of N. Applying the preceding remarks to $S = f(T)$, there are two possibilities:

(1) We can isotope f so that $f(T)$ is the torus splitting N_ψ into two copies of N. Then φ must be obtained from ψ by composing on the left and right by diffeomorphisms of ∂N which extend to diffeomorphisms of N. Such diffeomorphisms must preserve slope 0 and slope ∞, since these are the unique slopes of the boundaries of incompressible, ∂-incompressible separating and nonseparating annuli in N, respectively,

as noted earlier. The theorem is thus proved in this case.

(2) We can isotope f so that $f(T)$ meets each of the two copies of N in incompressible, ∂-incompressible annuli. These must be separating, since T is separating, so they must have boundaries of slope 0. Thus ψ preserves slope 0, having, we may assume, the form $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. In this case if we take all the slope 0 circles in the torus and Klein bottle fibers of the torus semi-bundle N_ψ we obtain the Seifert fibering $M(0,0;{}^1\!/_2,{}^1\!/_2,{}^{-1}\!/_2,{}^{-1}\!/_2,n)$, with Euler number n, as the reader can verify by inspection. The torus $f(T)$ can be isotoped to be vertical in this Seifert fibering, since horizontal surfaces in $M(0,0;{}^1\!/_2,{}^1\!/_2,{}^{-1}\!/_2,{}^{-1}\!/_2,n)$ are nonseparating (because the base surface of this Seifert fibering is orientable, hence the fiber circles are coherently orientable, making any horizontal surface the fiber of a surface bundle). We can then rechoose the torus semi-bundle structure on N_ψ so that $f(T)$ becomes a fiber torus. (The other torus and Klein bottle fibers are also unions of circle fibers of $M(0,0;{}^1\!/_2,{}^1\!/_2,{}^{-1}\!/_2,{}^{-1}\!/_2,n)$.) The ψ for the new torus semi-bundle structure is clearly the same as the old one, namely $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. So this reduces us to case (1). $\square$

Which Seifert fiberings are torus semi-bundles N_φ? If the torus fiber is vertical in the Seifert fibering, it splits N_φ into two Seifert-fibered copies of N. Since N has just the two Seifert fiberings $M(-1,1;\,)$ and $M(0,1;{}^1\!/_2,{}^1\!/_2)$, the possibilities are:

(1) $M(-2,0;n)$, with $\varphi = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$

(2) $M(-1,0;{}^1\!/_2,{}^{-1}\!/_2,n)$, with $\varphi = \begin{pmatrix} 0 & 1 \\ 1 & n \end{pmatrix}$, or equivalently $\begin{pmatrix} n & 1 \\ 1 & 0 \end{pmatrix}$

(3) $M(0,0;{}^1\!/_2,{}^1\!/_2,{}^{-1}\!/_2,{}^{-1}\!/_2,n)$, with $\varphi = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$

On the other hand, if the torus fiber of N_φ is horizontal, then as we saw in §2.1, the Seifert fibering is either the one in (1) or (2) with $n = 0$.

The theorem implies that all the manifolds listed in (1)-(3) are distinct, modulo changing the sign of n and the coincidence of (1) and (3) when $n = 0$. In most cases this was already proved in §2.1. The one exception is the manifold $M(-1,0;{}^1\!/_2,{}^{-1}\!/_2)$, labelled M_6 in §2.1, which we did not then distinguish from the other manifolds in (1)-(3).

For each N_φ there is a torus bundle M_ψ which is a natural 2-sheeted covering space of N_φ. Over each of the two copies of N in N_φ, M_ψ consists of the double cover $S^1\times S^1\times I \to N$ by which N was defined. The first copy of $S^1\times S^1\times I$ is glued to the second by φ along $S^1\times S^1\times\{0\}$ and by $\tau\varphi\tau$ along $S^1\times S^1\times\{1\}$, where $\tau(x,y) = (-x,\rho(y))$, as in the definition of N. Thus $\psi = \tau\varphi^{-1}\tau\varphi$. Since the 2-sheeted covering $M_\psi \to N_\varphi$ takes fibers to fibers, it fits into a commutative diagram as at the right, where the two vertical maps are the bundle and semi-bundle projections, and the map $S^1 \to I$ is the 'folding' map which factors out by the reflection ρ.

$$\begin{array}{ccc} M_\psi & \longrightarrow & N_\varphi \\ \downarrow & & \downarrow \\ S^1 & \longrightarrow & I \end{array}$$

The bundle $M_\psi \to S^1$ can be regarded as the pullback of the semi-bundle $N_\varphi \to I$ via the folding map $S^1 \to I$. There are similar

n-to-1 folding maps $I \to I$, folding up I like an n-segmented carpenter's ruler, and $N_\varphi \to I$ can be pulled back to other torus semi-bundles via these folding maps.

The Geometry of $SL_2(\mathbb{Z})$

A matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ gives rise to a linear fractional transformation $z \mapsto (az+b)/(cz+d)$ of the Riemann sphere $\mathbb{C} \cup \{\infty\}$. Since this has real coefficients it takes $\mathbb{R} \cup \{\infty\}$ to itself, preserving orientation since the determinant is positive. Hence it takes the upper half plane to itself since all linear fractional transformations preserve orientation of $\mathbb{C} \cup \{\infty\}$. Let $\mathcal{H}$ denote the upper half plane, including its boundary points $\mathbb{R} \cup \{\infty\}$. The action of $SL_2(\mathbb{Z})$ on $\mathcal{H}$ extends to an action of $GL_2(\mathbb{Z})$ on $\mathcal{H}$ if we identify points in the upper and lower half planes via complex conjugation, since a linear fractional transformation with real coefficients takes conjugate points to conjugate points. Then an element of $GL_2(\mathbb{Z})$ preserves or reverses orientation of $\mathcal{H}$ according to whether its determinant is $+1$ or -1.

For a matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2(\mathbb{Z})$ let ℓ_A be the semicircle in $\mathcal{H}$ orthogonal to $\partial\mathcal{H}$ with endpoints at the rational numbers c/a and d/b in $\partial\mathcal{H}$. We orient ℓ_A from c/a to d/b. In case a or b is zero, ℓ_A is a vertical line in $\mathcal{H}$ with lower endpoint at d/b or c/a, respectively, and other 'endpoint' at $\infty = \pm 1/0$. Note that there are exactly four A's with the same oriented ℓ_A, obtained by multiplying either column of A by -1. When the interior of $\mathcal{H}$ is taken as a model for the hyperbolic plane, the ℓ_A's are geodesics, so we shall refer to them as 'lines.'

The collection of all the lines ℓ_A is invariant under the action of $GL_2(\mathbb{Z})$ on $\mathcal{H}$. Namely, if $A, B \in GL_2(\mathbb{Z})$ then $A(\ell_B) = \ell_{AB}$, as the reader can readily verify using the fact that linear fractional transformations preserve angles and take circles and lines in $\mathbb{C}$ to circles or lines. Note that this action is transitive: $\ell_B = BA^{-1}(\ell_A)$.

Proposition 2.9. *The collection of all lines ℓ_A forms a tessellation $\mathcal{T}$ of $\mathcal{H}$ by ideal triangles: triangles with their vertices on $\partial\mathcal{H}$.*

See Figure 2.9 for two pictures of this tesselation $\mathcal{T}$, the second picture being a redrawing of the first to make $\mathcal{H}$ a disk.

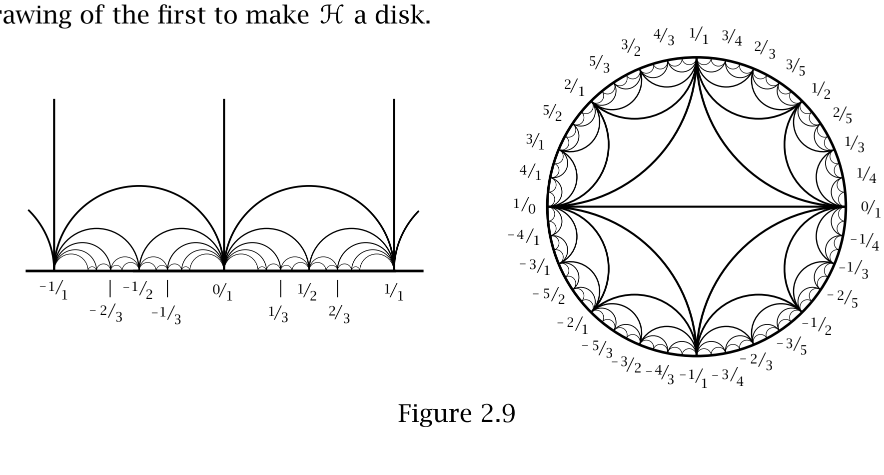


Figure 2.9

Proof: First we check that no two lines ℓ_A cross. Since the action of $GL_2(\mathbb{Z})$ on the ℓ_A's is transitive, it suffices to consider an ℓ_A crossing ℓ_I. If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and ℓ_A is oriented from left to right (which can be achieved by switching the columns of A if necessary), then $c/a < 0 < d/b$. We may assume c is negative and a, b, and d are positive. But then $ad - bc > 1$, which is impossible since elements of $GL_2(\mathbb{Z})$ have determinant ± 1.

Next we observe that any two rational numbers are joined by an edgepath consisting of finitely many edges ℓ_A, each two successive edges in the edgepath having a common vertex. For given $c/a \in \mathbb{Q}$ with $(a,c) = 1$, there exist integers b,d with $ad - bc = \pm 1$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2(\mathbb{Z})$. Thus each rational is the endpoint of a line ℓ_A. To get an edgepath joining c/a to $0/1$, use the fact that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ can be reduced to a diagonal matrix by a sequence of elementary column operations, each operation consisting of adding an integer multiple of one column to the other column. (This is essentially the Euclidean algorithm in $\mathbb{Z}$.) This produces a finite sequence of matrices in $GL_2(\mathbb{Z})$, each pair of successive matrices having a column in common, so the corresponding sequence of ℓ_A's forms an edgepath. The final edge has endpoints $0 = 0/\pm 1$ and $\infty = \pm 1/0$, since the diagonal matrix has ± 1's down the diagonal.

The three ℓ_A's with endpoints $0/1$, $1/0$, and $1/1$ form one triangle in $\mathcal{H}$. Let U be the union of all the images of this triangle under the action of $GL_2(\mathbb{Z})$, with all vertices in $\partial\mathcal{H}$ deleted. To finish the proof we need to show that $U = \mathrm{int}(\mathcal{H})$. We do this by showing that U is both open and closed in $\mathrm{int}(\mathcal{H})$. Openness is clear, since $GL_2(\mathbb{Z})$ acts transitively on the ℓ_A's, and the edge ℓ_I is shared by the two triangles whose third vertices are $+1$ and -1. If U were not closed, there would be an infinite sequence $z_i \in U$ converging to a point $z \in \mathrm{int}(\mathcal{H}) - U$. We may assume the z_i's lie in distinct lines ℓ_{A_i}. Since no ℓ_{A_i}'s cross, there would have to be a limit line ℓ containing z. Since $z \notin U$, ℓ is not an ℓ_A. Neither endpoint of ℓ can be rational since every line through a rational point crosses lines ℓ_A (true for $1/0$, hence true for all rationals by transitivity) and if ℓ crossed an ℓ_A so would some ℓ_{A_i}. On the other hand, if both endpoints of ℓ are irrational, then again ℓ cannot cross any ℓ_A's, and in this case it would not be possible to join rationals on one side of ℓ to rationals on the other side by edgepaths consisting of ℓ_A's, contradicting an earlier observation. Thus U must be closed in $\mathrm{int}(\mathcal{H})$. ◻

For the action of $GL_2(\mathbb{Z})$ on $\mathcal{H}$, only $\pm I$ act as the identity. So we have an injection of the quotient $PGL_2(\mathbb{Z}) = GL_2(\mathbb{Z})/\pm I$ into the symmetry group of the tesselation $\mathcal{T}$. This is in fact an isomorphism, since $PGL_2(\mathbb{Z})$ acts transitively on oriented edges, and also contains orientation-reversing maps of $\mathcal{H}$. The index-two subgroup $PSL_2(\mathbb{Z})$ is the group of orientation-preserving symmetries of $\mathcal{T}$.

An element $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ not equal to $\pm I$ has one or two fixed points in $\mathcal{H}$, the roots of the quadratic $z = (az+b)/(cz+d)$. There are three possibilities:

(1) There are two distinct real irrational roots r_1 and r_2. Then the line (semicircle) in $\mathcal{H}$ joining r_1 and r_2 is invariant under A. The infinite sequence of triangles of $\mathcal{T}$ meeting this line is arranged in a periodic pattern, invariant under A, which acts as a translation of this strip, as in Figure 2.10, with a_i triangles having a common vertex on one side of the strip, followed by a_{i+1} triangles with a common vertex on the other side, etc.

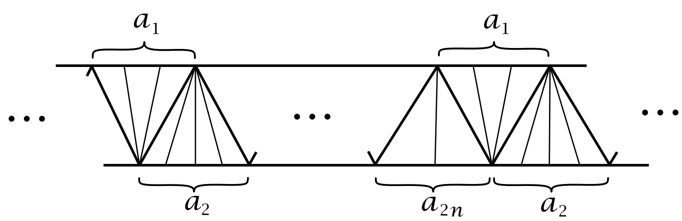


Figure 2.10

The cycle $(a_1, \cdots, a_{2n})$ is a complete invariant of the conjugacy class of A in $PGL_2(\mathbb{Z})$, by our earlier comment that $PGL_2(\mathbb{Z})$ is the full symmetry group of $\mathcal{T}$. The fixed points r_1 and r_2 are slopes of eigenvectors of A. The two corresponding eigenvalues have the same sign since $A \in SL_2(\mathbb{Z})$, and this sign together with the cycle $(a_1, \cdots, a_{2n})$ is a complete invariant of the conjugacy class of A in $GL_2(\mathbb{Z})$. A specific representative of this conjugacy class is

$$\pm\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix}\begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ a_3 & 1 \end{pmatrix}\cdots\begin{pmatrix} 1 & a_{2n} \\ 0 & 1 \end{pmatrix}$$

The conjugacy class of A^{-1} is represented by the inverse cycle $(a_{2n}, \cdots, a_1)$ with the same sign of the eigenvalues. Geometrically, we can think of the conjugacy class of the pair $\{A, A^{-1}\}$ as the triangulated cylinder obtained from the strip in Figure 5.2 by factoring out by the translation A, together with a sign. Diffeomorphism classes of torus bundles M_φ with φ having distinct real eigenvalues correspond bijectively with such triangulated, signed cylinders, by Theorem 2.6.

(2) There is only one rational root r. In this case the triangles having r as a vertex form a degenerate infinite strip, with quotient a cone divided into some number n of triangles. The number n together with the sign of the eigenvalue ± 1 of A is a complete invariant of the conjugacy class of $\{A, A^{-1}\}$ in $GL_2(\mathbb{Z})$. A representative of this conjugacy class is $\pm\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$.

(3) There are two complex conjugate roots. The root in $\operatorname{int}(\mathcal{H})$ gives a fixed point of A in $\operatorname{int}(\mathcal{H})$, so A acts either by rotating a triangle 120 degrees about its center or by rotating an edge 180 degrees about its center. Since $PGL_2(\mathbb{Z})$ acts transitively on triangles and edges, these rotations of $\mathcal{H}$ fall into two conjugacy classes in $PGL_2(\mathbb{Z})$, according to the angle of rotation. In $GL_2(\mathbb{Z})$ we thus have just the conjugacy classes of the matrices of orders $2, 3, 4$, and 6 discussed earlier.

Now let us use the tesselation $\mathcal{T}$ to make explicit the equivalence relation on elements of $GL_2(\mathbb{Z})$ corresponding to diffeomorphism of torus semi-bundles, described in Theorem 2.8. An element $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z})$ determines an oriented edge ℓ_A in $\mathcal{T}$, as above. The other three A's with the same oriented ℓ_A differ by multiplication on the right by $\begin{pmatrix} \pm1 & 0 \\ 0 & \pm1 \end{pmatrix}$, so give the same torus semi-bundle N_A. Multiplication on the left by $\begin{pmatrix} \pm1 & 0 \\ 0 & \pm1 \end{pmatrix}$ corresponds either to the identity of $\mathcal{H}$ or a reflection of $\mathcal{H}$ across the edge ℓ_I, so we may normalize A by choosing it so that ℓ_A lies in the right half of $\mathcal{H}$. Let $T(A)$ be the smallest connected union of triangles of $\mathcal{T}$ containing ℓ_I and ℓ_A. This is a finite triangulated strip with two distinguished oriented edges in its boundary, ℓ_I and ℓ_A, as in Figure 2.11.

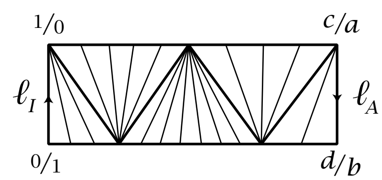


Figure 2.11

In degenerate cases the top or bottom edges of the strip can reduce to a single point. We leave it as an exercise for the reader to check that $T(A^{-1})$ is obtained from $T(A)$ by simply interchanging the labelling of the two oriented edges ℓ_I and ℓ_A. Thus diffeomorphism classes of torus semi-bundles correspond bijectively with isomorphism classes of triangulated strips with oriented left and right edges.

As another exercise, the reader can show that the torus bundle double covering N_A, as described earlier in this section, corresponds to the cylinder (or cone) obtained by gluing two copies of $T(A)$ together by the identity map of their boundary, with the sign $+1$ if the two ends of $T(A)$ have parallel orientation and the sign -1 in the opposite case; this assumes we are not in the degenerate case that $T(A)$ is just a single edge, with no triangles.

Chapter 3. Homotopy Properties

1. The Loop and Sphere Theorems

These are the basic results relating homotopy theory, specifically π_1 and π_2, to more geometric properties of 3-manifolds.

The Loop Theorem

Let M be a 3-manifold-with-boundary, not necessarily compact or orientable.

Theorem 3.1. *If there is a map $f:(D^2,\partial D^2)\to(M,\partial M)$ with $f\,|\,\partial D^2$ not nullhomotopic in ∂M, then there is an embedding with the same property.*

Proof: The first half of the proof consists of covering space arguments which reduce the problem to finitely many applications of the relatively simple special case that f is at most two-to-one. The second half of the proof then takes care of this special case. It is a curious feature of the proof that at the end of the first step one discards the original f completely and starts afresh with a new and nicer f.

It seems more convenient to carry out the first half of the proof in the piecewise linear category. Choose a triangulation of M and apply the simplicial approximation theorem (for maps of pairs) to homotope the given f to a map f_0 which is simplicial in some triangulation of the domain D^2 and which still satisfies the same hypotheses as f. We now construct a diagram as at the right. To obtain the bottom row, let D_0 be $f_0(D^2)$, a finite subcomplex of M, and let V_0 be a neighborhood of D_0 in M that is a compact 3-manifold deformation retracting onto D_0. The classical construction of such a V_0 is to take the union of all simplices in the second barycentric subdivision of M that meet D_0.

$$\begin{tikzcd}
 & D_n \arrow[r, hook] & V_n \arrow[r, hook] & M_n \arrow[dl, "p_n"] \\
 & & \vdots & {} \arrow[dl, "p_2"] \\
 & D_1 \arrow[r, hook] & V_1 \arrow[r, hook] & M_1 \arrow[dl, "p_1"] \\
D^2 \arrow[r, "f_0"] \arrow[ur, "f_1"] \arrow[uuur, "f_n"] & D_0 \arrow[r, hook] & V_0 \arrow[r, hook] & M
\end{tikzcd}$$

In particular, V_0 is connected since D_0 is connected. If V_0 has a connected 2-sheeted cover $p_1:M_1\to V_0$, then f_0 lifts to a map $f_1:D^2\to M_1$ since D^2 is simply-connected. Let $D_1=f_1(D^2)$ and let V_1 be a neighborhood of D_1 in M_1 chosen as before. This gives the second row of the diagram. If V_1 has a connected 2-sheeted cover we can repeat the process to construct a third row, and so on up the tower.

To see that we must eventually reach a stage where the tower cannot be continued further, consider the covering $p_i^{-1}(D_{i-1})\to D_{i-1}$. Both these spaces have natural simplicial structures with simplices lifting the simplices of D_0. The nontrivial deck transformation τ_i of the covering space $p_i^{-1}(D_{i-1})\to D_{i-1}$ is a simplicial homeomorphism. Since M_i is connected, so is $p_i^{-1}(D_{i-1})$ since M_i deformation retracts to $p_i^{-1}(D_{i-1})$ by lifting a deformation retraction of V_{i-1} to D_{i-1}. The set $p_i^{-1}(D_{i-1})=D_i\cup\tau_i(D_i)$ is connected, so $D_i\cap\tau_i(D_i)$ must be nonempty. This means that τ_i must take some simplex of D_i to another simplex of D_i, distinct from the first simplex since τ_i has no fixed points. These two simplices of D_i are then identified in D_{i-1}, so D_{i-1} is a quotient of D_i with fewer simplices than D_i. Since the number of simplices of D_i is bounded by the number of simplices in the original triangulation of the source D^2, it follows that the height of the tower is bounded.

Thus we arrive at a V_n having no connected 2-sheeted cover. This says that $\pi_1 V_n$ has no subgroup of index two, so there is no nontrivial homomorphism $\pi_1 V_n \to \mathbb{Z}_2$, and hence $H^1(V_n;\mathbb{Z}_2) = \mathrm{Hom}(H_1(V_n),\mathbb{Z}_2) = 0$. Consider now the following three terms of the long exact sequence of homology groups for the pair $(V_n, \partial V_n)$:

$$\begin{array}{ccccc} H_2(V_n,\partial V_n;\mathbb{Z}_2) & \longrightarrow & H_1(\partial V_n;\mathbb{Z}_2) & \longrightarrow & H_1(V_n;\mathbb{Z}_2) \\ \wr\wr & & & & \wr\wr \\ H^1(V_n;\mathbb{Z}_2) & & & & H^1(V_n;\mathbb{Z}_2) \end{array}$$

The isomorphism on the left comes from Poincaré duality and the one on the right comes from the universal coefficient theorem. Since the two outer terms of the exact sequence vanish, so does $H_1(\partial V_n;\mathbb{Z}_2)$. This implies that all the components of the compact surface ∂V_n are 2-spheres.

In the component of ∂V_i containing $f_i(\partial D^2)$, let $F_i = (p_1 \circ \cdots \circ p_i)^{-1}(\partial M)$, a compact surface if the neighborhoods V_i are chosen reasonably. Let $N_i \subset \pi_1 F_i$ be the kernel of $(p_1 \circ \cdots \circ p_i)_* : \pi_1 F_i \to \pi_1(\partial M)$. Note that $[f_i \,|\, \partial D^2] \notin N_i$ by the initial hypothesis on f, so N_i is a proper normal subgroup of $\pi_1 F_i$.

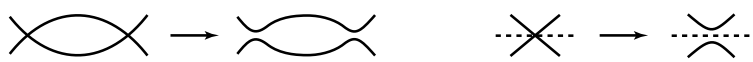
At the top of the tower the surface F_n is a planar surface, being a subsurface of a sphere. Hence $\pi_1 F_n$ is normally generated by the circles of ∂F_n. Since $N_n \neq \pi_1 F_n$, some circle of ∂F_n must represent an element of $\pi_1 F_n - N_n$. This circle bounds a disk in ∂V_n. Let $g_n : D^2 \to V_n$ be this embedding of a disk, with its interior pushed into V_n so as to be a proper embedding. We have $[g_n | \partial D^2] \notin N_n$. We can take g_n to be a smooth embedding and work from now on in the smooth category.

The rest of the proof consists of descending the tower step by step, producing embeddings $g_i : D^2 \to V_i$ with $[g_i | \partial D^2] \notin N_i$. In the end we will have an embedding $g_0 : D^2 \to V_0 \subset M$ with $[g_0 | \partial D^2] \notin N_0$, which says that $g_0 | \partial D^2$ is not nullhomotopic in ∂M, and the proof will be complete.

Consider the inductive step of producing g_{i-1} from g_i. Since p_i is a 2-sheeted cover, we can perturb g_i so that the immersion $p_i g_i$ has only simple double curves, either circles or arcs, where two sheets of $p_i g_i(D^2)$ cross transversely. Our aim is to modify g_i so as to eliminate each of these double curves in turn.

Suppose C is a double circle. A neighborhood $N(C)$ of C in $p_i g_i(D^2)$ is a bundle over C with fiber the letter X. Thus $N(C)$ is obtained from $X \times [0,1]$ by identifying $X \times \{0\}$ with $X \times \{1\}$ by some homeomorphism given by a symmetry of X. A 90 degree rotational symmetry is impossible since p_i is two-to-one. A 180 degree rotational symmetry is also impossible since it would force the domain D^2 to contain a Möbius band. For the same reason a symmetry which reflects the X across one of its crossbars is impossible. The two remaining possibilities are the identity symmetry, with $N(C) = X \times S^1$, and a reflection of X across a horizontal or vertical line. Note that a reflection can occur only if M is nonorientable.

In the case that $N(C) = X \times S^1$ there are two circles in D^2 mapping to C. These circles bound disks D_1 and D_2 in D^2. If D_1 and D_2 are nested, say $D_1 \subset D_2$, then we redefine g_i on D_2 to be $g_i | D_1$, smoothing the resulting corner along C. This has the effect of eliminating C from the self-intersections of $p_i g_i(D^2)$, along with any self-intersections of the immersed annulus $p_i g_i(D_2 - D_1)$. If D_1 and D_2 are disjoint we modify g_i by interchanging its values on the disks D_1 and D_2, then smoothing the resulting corners along C, as shown in Figure 3.1(a). Thus we eliminate the circle C from the self-intersections of $p_i g_i(D^2)$ without introducing any new self-intersections.

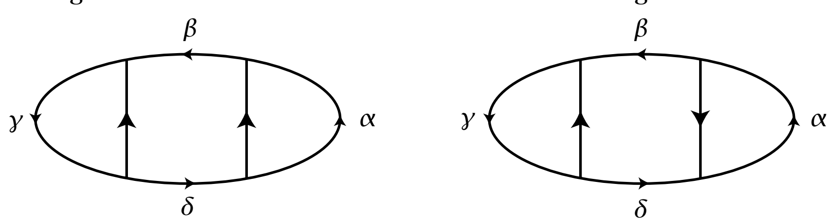


(a) (b)

Figure 3.1

In the opposite case that $N(C)$ is not $X \times S^1$ we modify the immersed disk $p_i g_i(D^2)$ in the way shown in Figure 3.1(b), replacing each cross-sectional X of $N(C)$ by two disjoint arcs. This replaces an immersed annulus by an embedded annulus with the same boundary, so the result is again an immersed disk with fewer double circles. Iterating these steps, all double circles can be eliminated without changing $g_i | \partial D^2$.

Now consider the case of a double arc C. The two possibilities for how the two corresponding arcs in D^2 are identified to C are shown in Figure 3.2.



(a) (b)

Figure 3.2

In either case, Figure 3.3 shows two ways to modify g_i to produce new immersions g_i' and g_i'' eliminating the double arc C.

g_i' g_i''

Figure 3.3

The claim is that either $[g_i' | \partial D^2] \notin N_i$ or $[g_i'' | \partial D^2] \notin N_i$. To see this, we break the loop $g_i | \partial D^2$ up into four pieces α, β, γ, δ as indicated in Figure 3.2. Then in case

(a) we have

$$\begin{array}{c} \alpha\beta\gamma\delta = \underbrace{(\alpha\gamma)}_{g_i'|\partial D^2}\delta^{-1}\underbrace{(\alpha\beta^{-1}\gamma\delta^{-1})^{-1}}_{g_i''|\partial D^2}\underbrace{(\alpha\gamma)}_{g_i'|\partial D^2}\delta \end{array}$$

and in case (b) we have

$$\alpha\beta\gamma\delta = \underbrace{(\alpha\gamma^{-1})}_{g_i'|\partial D^2}(\gamma\delta)^{-1}\underbrace{(\alpha\gamma^{-1})^{-1}}_{g_i'|\partial D^2}\underbrace{(\alpha\delta\gamma\beta)}_{g_i''|\partial D^2}(\gamma\delta)$$

So in either case, if $g_i'|\partial D^2$ and $g_i''|\partial D^2$ were both in the normal subgroup N_i, so would $g_i|\partial D^2$ be in N_i.

We choose the new g_i to be either g_i' or g_i'', whichever one preserves the condition $[g_i|\partial D^2] \notin N_i$. Repeating this step, we eventually get g_i with $p_i g_i$ an embedding g_{i-1}, with $[g_{i-1}|\partial D^2] \notin N_{i-1}$. □

Corollary 3.2 (Dehn's Lemma). *If an embedded circle in ∂M is nullhomotopic in M, it bounds a disk in M.*

Proof: First delete from ∂M everything but a neighborhood of the given circle C, producing a new manifold M' with $\partial M'$ an annulus. (If C had a Möbius band neighborhood in ∂M, it would be an orientation-reversing loop in M, which is impossible since it is nullhomotopic.) Apply the loop theorem to M' to obtain a disk $D^2 \subset M'$ with ∂D^2 nontrivial in $\partial M'$, hence isotopic to C. □

Corollary 3.3. *Let $S \subset M$ be a 2-sided surface. If the induced map $\pi_1 S \to \pi_1 M$ is not injective for some choice of basepoint in some component of S, then there is a disk $D^2 \subset M$ with $D^2 \cap S = \partial D^2$ a nontrivial circle in S.*

Proof: Let $f : D^2 \to M$ be a nullhomotopy of a nontrivial loop in S. We may assume f is transverse to S, so $f^{-1}(S)$ consists of ∂D^2 plus possibly some circles in $D^2 - \partial D^2$. If such a circle in $D^2 - \partial D^2$ gives a trivial loop in the component of S it maps to, then we can redefine f so as to eliminate this circle (and any other circles inside it) from $f^{-1}(S)$. So we may assume no circles of $f^{-1}(S)$ are trivial in S. Then the restriction of f to the disk bounded by an innermost circle of $f^{-1}(S)$ (this circle may be ∂D^2 itself) gives a nullhomotopy in $M - S$ of a nontrivial loop in S. Now apply the loop theorem to $(M - \partial M)\,|\,(S - \partial S)$. □

Here is another application of the loop theorem:

Proposition 3.4. *A compact connected orientable prime 3-manifold M with $\pi_1 M \approx \mathbb{Z}$ is either $S^1 \times S^2$ or $S^1 \times D^2$.*

Proof: First let us do the case $\partial M \neq \varnothing$. Then ∂M can contain no spheres, otherwise primeness would imply $M = B^3$, contradicting $\pi_1 M \approx \mathbb{Z}$. To restrict the possibilities for ∂M further we need the following basic result.

Lemma 3.5. *If M is a compact orientable 3-manifold then the image of the boundary map $H_2(M,\partial M)\to H_1(\partial M)$ has rank equal to one half the rank of $H_1(\partial M)$.*

Here 'rank' means the number of $\mathbb{Z}$ summands in a direct sum splitting into cyclic groups. Equivalently, we can use homology with $\mathbb{Q}$ coefficients and replace 'rank' by 'dimension.'

Proof: With $\mathbb{Q}$ coefficients understood, consider the commutative diagram

$$\begin{array}{ccccc} H_2(M,\partial M) & \xrightarrow{\partial} & H_1(\partial M) & \xrightarrow{i_*} & H_1(M) \\ \wr\wr & & \wr\wr & & \wr\wr \\ H^1(M) & \xrightarrow{i^*} & H^1(\partial M) & \xrightarrow{\delta} & H^2(M,\partial M) \end{array}$$

where the vertical isomorphisms are given by Poincaré duality. Since we have coefficients in a field, the map i^* is the dual of i_*, obtained by applying $\mathrm{Hom}(-,\mathbb{Q})$. Linear algebra then implies that $\dim\mathrm{Ker}\, i_* = \dim\mathrm{Coker}\, i^*$. Hence $\dim\mathrm{Im}\,\partial = \dim\mathrm{Ker}\, i_* = \dim\mathrm{Coker}\, i^* = \dim\mathrm{Coker}\,\partial$ which implies the result. $\square$

Continuing with the proof of the Proposition, $H_1(M)$ has rank one, hence so does $H^1(M)\approx H_2(M,\partial M)$ and we conclude that ∂M must be a single torus. The map $\pi_1\partial M\to\pi_1 M$ is not injective, so the loop theorem gives a disk $D^2\subset M$ with ∂D^2 nontrivial in ∂M. Splitting M along this D^2 gives a manifold with boundary S^2, and this leads to a connected sum decomposition $M = N \sharp S^1\times D^2$. Since M is prime, we conclude that $M = S^1\times D^2$.

When $\partial M=\emptyset$ we use another basic fact about 3-manifolds:

Lemma 3.6. *For a 3-manifold M, every class in $H_2(M)$ is represented by an embedded closed orientable surface $S\subset M$, and similarly every class in $H_2(M,\partial M)$ is represented by an embedded compact orientable surface $(S,\partial S)\subset(M,\partial M)$. If M is orientable, S may be chosen in both cases so that the inclusion of each component of S into M induces an injective map on π_1.*

Proof: Taking simplicial homology with respect to some triangulation of M, represent a class in $H_2(M)$ by a cycle z, a sum of the oriented 2-simplices in M with integer coefficients. Reversing orientations of some simplices, we may assume all coefficients are non-negative. To desingularize this 2-cycle z into a surface, first replace each 2-simplex σ with coefficient n_σ by n_σ nearby parallel copies of itself, all with the same boundary $\partial\sigma$. Next, look in a small disk transverse to a 1-simplex of M. The copies of 2-simplices incident to this 1-simplex have induced normal orientations, which we can think of as either clockwise or counterclockwise around the 1-simplex. Since z is a cycle, the number of incident 2-simplices with clockwise orientation equals the number with counterclockwise orientation. Thus there must be two consecutive 2-simplices with opposite orientation. The edges of these along the given 1-simplex can be joined together and pushed off the 1-simplex. Repeat this step until z is 'resolved' to a surface whose only singularities are at vertices of M. To eliminate these isolated singularities, consider a small ball about a vertex. The surface meets the boundary of this ball in disjoint embedded circles, so we can cap off these circles by disjoint embedded disks in the interior of the ball. Doing this for all vertices produces an embedded oriented surface S, with a natural map to the 2-skeleton of M taking the fundamental class for S to z. The case of relative homology is treated in similar fashion.

To obtain the final statement of the lemma, suppose the map $\pi_1 S \to \pi_1 M$ is not injective for some component of S. Then Corollary 3.3 gives a disk $D \subset M$ realizing this non-injectivity. Surgering S along D produces a new surface which is obviously in the same homology class. Since S is compact, surgery simplifies S, either splitting one component into two components of lower genus, in case ∂D separates S, or in the opposite case, reducing the genus of a component of S. So after finitely many steps we reach the situation that each component of S is π_1-injective. □

Returning to the proof of Proposition 3.4 in the case of closed M, we have $\mathbb{Z} \approx H_1(M) \approx H^1(M) \approx H_2(M)$. By the lemma we may represent a nontrivial class in $H_2(M)$ by a closed embedded surface S with each component π_1-injective. Since $\pi_1 M \approx \mathbb{Z}$, each component of S must then be a 2-sphere. The separating spheres are trivial in homology, so can be discarded, leaving at least one nonseparating sphere. As in Proposition 1.4 this gives a splitting $M = N \sharp S^1 \times S^2$, finishing the proof. □

Dropping the hypothesis that M is prime, the proof shows that $M = N \sharp S^1 \times S^2$ or $N \sharp S^1 \times D^2$. Then $\pi_1 M \approx \pi_1 N * \mathbb{Z}$, hence N is simply-connected, since a nontrivial free product is necessarily nonabelian. This implies that ∂N, hence also ∂M, consists of spheres, since as we showed in the proof of the loop theorem, $H_1(N; \mathbb{Z}_2) = 0$ is sufficient to obtain this conclusion. Let N' be obtained from N by capping off its boundary spheres by balls. Thus N' is a closed simply-connected 3-manifold. The next proposition describes the homotopy type of such an N'.

Proposition 3.7. *A closed simply-connected* 3*-manifold is a homotopy sphere, i.e., homotopy equivalent to* S^3. *A compact* 3*-manifold is contractible iff it is simply-connected and has boundary a* 2*-sphere.*

Proof: Let P be a closed simply-connected 3-manifold. It is orientable since $\pi_1 P = 0$. By Poincaré duality and the Hurewicz theorem we have $\pi_2 P \approx H_2(P) \approx H^1(P) \approx H_1(P) = 0$. Then $\pi_3 P \approx H_3(M) \approx \mathbb{Z}$, and we have a degree one map $S^3 \to P$. This induces an isomorphism on all homology groups. Since both S^3 and P are simply-connected, Whitehead's theorem implies that the map $S^3 \to P$ is a homotopy equivalence.

The second assertion can be proved in similar fashion. □

The Poincaré Conjecture (still unproved) is the assertion that S^3 is the only homotopy 3-sphere, or equivalently, B^3 is the only compact contractible 3-manifold.

The Sphere Theorem

Here is the general statement:

Theorem 3.8. *Let M be a connected 3-manifold. If $\pi_2 M \neq 0$, then either*
(a) *there is an embedded S^2 in M representing a nontrivial element in $\pi_2 M$, or*
(b) *there is an embedded 2-sided $\mathbb{R}\mathrm{P}^2$ in M such that the composition of the cover $S^2 \to \mathbb{R}\mathrm{P}^2$ with the inclusion $\mathbb{R}\mathrm{P}^2 \hookrightarrow M$ represents a nontrivial element of $\pi_2 M$.*

A 2-sided $\mathbb{R}\mathrm{P}^2 \subset M$ can exist only if M is nonorientable, so for orientable M the theorem asserts that there is an $S^2 \subset M$ which is nontrivial in $\pi_2 M$, if $\pi_2 M$ is nonzero. An example of a nonorientable manifold where $\mathbb{R}\mathrm{P}^2$'s are needed to represent a nontrivial element of π_2 is $S^1 \times \mathbb{R}\mathrm{P}^2$. This has $S^1 \times S^2$ as a double cover, so π_2 is $\mathbb{Z}$. However, $S^1 \times \mathbb{R}\mathrm{P}^2$ is irreducible, by Proposition 1.12, so any embedded S^2 in $S^1 \times \mathbb{R}\mathrm{P}^2$ is nullhomotopic.

We remark that it is not true in general for orientable M that every element of $\pi_2 M$ is represented by an embedded sphere. For example, $\pi_2(S^1 \times S^2) \approx \mathbb{Z}$, but as we saw in Proposition 1.4, a separating S^2 in $S^1 \times S^2$ bounds a ball since $S^1 \times S^2$ is prime, while a nonseparating S^2 is a slice $\{x\} \times S^2$ in a product structure $S^1 \times S^2$ hence represents a generator of π_2.

Corollary 3.9. *Let M be a compact connected orientable irreducible 3-manifold with universal cover $\widetilde{M}$.*
(a) *If $\pi_1 M$ is infinite, and in particular if $\partial M \neq \varnothing$, then M is a $K(\pi,1)$, i.e., $\pi_i M = 0$ for all $i > 1$, or equivalently, $\widetilde{M}$ is contractible.*
(b) *If $\pi_1 M$ is finite, then either $M = B^3$ or M is closed and $\widetilde{M}$ is a homotopy 3-sphere.*

Proof: Since M is irreducible, the Sphere Theorem implies that $\pi_2 M$ is zero. It follows that $\pi_3 M \approx \pi_3 \widetilde{M} \approx H_3 \widetilde{M}$ by the Hurewicz theorem. If $\pi_1 M$ is infinite, $\widetilde{M}$ is noncompact, so $H_3 \widetilde{M} = 0$. Thus $\pi_3 M = 0$. By Hurewicz again, $\pi_4 M \approx \pi_4 \widetilde{M} \approx H_4 \widetilde{M}$. Since H_4 of a 3-manifold is zero, we get $\pi_4 M = 0$. Similarly, all higher homotopy groups of M are zero. Now assume $\pi_1 M$ is finite. If $\partial M \neq \varnothing$ then by Lemma 3.5, ∂M must consist of spheres and irreducibility implies $M = B^3$. If $\partial M = \varnothing$, then $\widetilde{M}$ is also closed since the covering $\widetilde{M} \to M$ is finite-sheeted, and so by Proposition 3.7 $\widetilde{M}$ is a homotopy sphere. □

Proposition 3.10. *An embedded sphere $S^2 \subset M$ is zero in $\pi_2 M$ iff it bounds a compact contractible submanifold of M. A 2-sided $\mathbb{R}\mathrm{P}^2 \subset M$ is always nontrivial in $\pi_2 M$.*

Proof: Let $\widetilde{S}$ be any lift of the nullhomotopic sphere $S \subset M$ to the universal cover $\widetilde{M}$. Via the isomorphism $\pi_2 M \approx \pi_2 \widetilde{M}$, $\widetilde{S}$ is nullhomotopic in $\widetilde{M}$, hence it is also homologically trivial in $\widetilde{M}$, bounding a compact submanifold $\widetilde{N}$ of $\widetilde{M}$. Rechoosing $\widetilde{S}$ to be an innermost lift of S in $\widetilde{N}$, we may assume $\widetilde{N}$ contains no other lifts of S. Then the restriction of the covering projection $\widetilde{M} \to M$ to $\widetilde{N}$ is a covering space $\widetilde{N} \to N$, as in the proof of Proposition 1.6. This covering is 1-sheeted, i.e., a homeomorphism, since it is 1-sheeted over S. By the van Kampen theorem, $\widetilde{N}$ is simply-connected since $\widetilde{M}$ is. By Proposition 3.7, N is contractible.

For the case of a 2-sided $\mathbb{R}P^2 \subset M$, consider the orientable double cover $\widetilde{M} \to M$, corresponding to the subgroup of $\pi_1 M$ represented by orientation-preserving loops in M. The given $\mathbb{R}P^2$ lifts to a unique $S^2 \subset \widetilde{M}$. If this is nullhomotopic, the first half of the proposition implies that this S^2 bounds a contractible submanifold $\widetilde{N}$ of $\widetilde{M}$. The nontrivial deck transformation of $\widetilde{M}$ takes $\widetilde{N}$ to itself since the given $\mathbb{R}P^2$ is 2-sided in M. So the restriction of the covering projection to $\widetilde{N}$ is a double cover $\widetilde{N} \to N$, with $\partial N = \mathbb{R}P^2$. This is impossible by the following lemma. (Alternatively, the Lefschetz fixed point theorem implies that the nontrivial deck transformation $\widetilde{N} \to \widetilde{N}$ must have a fixed point since $\widetilde{N}$ is contractible.) □

Lemma 3.11. *If N is a compact 3-manifold, then the Euler characteristic $\chi(\partial N)$ is even.*

Proof: By the proof of Lemma 3.5 with $\mathbb{Z}_2$ coefficients now instead of $\mathbb{Q}$ coefficients, we see that $H_1(\partial N; \mathbb{Z}_2)$ is even-dimensional, from which the result follows. □

The converse statement, that a closed surface of even Euler characteristic is the boundary of some compact 3-manifold, is easily seen to be true also.

Here is another consequence of the Sphere Theorem:

Proposition 3.12. *In a compact orientable 3-manifold M there exists a finite collection of embedded 2-spheres generating $\pi_2 M$ as a $\pi_1 M$-module.*

Proof: By the prime decomposition there is a finite collection S of disjoint spheres in M such that each component of $M|S$ is an irreducible manifold with punctures. We claim that the spheres in S, together with all the spheres in ∂M, generate $\pi_2 M$ as a $\pi_1 M$-module. For let $f : S^2 \to M$ be given. Perturb f to be transverse to S, so that $f^{-1}(S)$ is a collection of circles in S^2. Since the components of S are simply-connected, we may homotope f, staying transverse to S, so that f takes each circle in $f^{-1}(S)$ to a point. This means that f is homotopic to a linear combination, with coefficients in $\pi_1 M$, of maps $f_j : S^2 \to M - S$. It suffices now to show that if P is irreducible-with-punctures then $\pi_2 P$ is generated as a $\pi_1 P$-module by the spheres in ∂P. This is equivalent to saying that if $\widetilde{P} \to P$ is the universal cover, then $\pi_2 \widetilde{P}$ is generated by the spheres in $\partial \widetilde{P}$. To show this, let $\widetilde{P}'$ be $\widetilde{P}$ with its boundary spheres filled in with balls B_ℓ. We have a Mayer-Vietoris sequence:

$$
\begin{array}{ccccc}
 & & 0 & & \\
 & & \parallel & & \\
H_2(\cup_\ell \partial B_\ell) & \longrightarrow & H_2(\cup_\ell B_\ell) \oplus H_2(\widetilde{P}) & \longrightarrow & H_2(\widetilde{P}') \\
\wr\wr & & \wr\wr & & \wr\wr \\
\pi_2(\cup_\ell \partial B_\ell) & \longrightarrow & \pi_2(\widetilde{P}) & \longrightarrow & \pi_2(\widetilde{P}')
\end{array}
$$

Now $\widetilde{P}'$ is the universal cover of the irreducible manifold P' obtained by filling in the boundary spheres of P, so $\pi_2\widetilde{P}' \approx \pi_2 P' = 0$ by the Sphere Theorem. Since $H_2(\bigcup_\ell B_\ell) = 0$, we conclude that $\pi_2(\bigcup_\ell \partial B_\ell) \to \pi_2(\widetilde{P})$ is surjective, i.e., $\pi_2\widetilde{P}$ is generated by the spheres in its boundary. $\square$

Proof of the Sphere Theorem: Let $p:\widetilde{M}\to M$ be the universal cover. By Lemma 3.6 we may represent a nonzero class in $\pi_2 M \approx \pi_2\widetilde{M} \approx H_2\widetilde{M}$ by an embedded closed orientable surface $S \subset \widetilde{M}$, each component of which is π_1-injective, hence a sphere since $\pi_1\widetilde{M} = 0$. Some component will be nontrivial in $H_2\widetilde{M}$, hence in $\pi_2\widetilde{M}$. Call this component S_0.

The rest of the proof will be to show that the sphere S_0 can be chosen so that the union of all the images $\tau(S_0)$ under the deck transformations τ of $\widetilde{M}$ is a collection of disjoint embedded spheres. Then the restriction $p:S_0\to p(S_0)$ is a covering space, so its image $p(S_0)$ is an embedded S^2 or $\mathbb{R}\mathrm{P}^2$ in M representing a nontrivial element of $\pi_2 M$. If it is a 1-sided $\mathbb{R}\mathrm{P}^2$, then the boundary of an I-bundle neighborhood of this $\mathbb{R}\mathrm{P}^2$ is an S^2 with the inclusion $S^2 \hookrightarrow M$ homotopic to the composition $S_0 \to \mathbb{R}\mathrm{P}^2 \hookrightarrow M$, hence nontrivial in $\pi_2 M$ by the choice of S_0. Thus the proof will be finished.

We now introduce the key technical ideas which will be used. Let N be a triangulated 3-manifold, with triangulation $\mathcal{T}$. Then a surface $S \subset N$ is said to be in **normal form** with respect to the triangulation if it intersects each 3-simplex Δ^3 of $\mathcal{T}$ in a finite collection of disks, each disk being bounded by a triangle or square of one of the types shown in Figure 3.4.

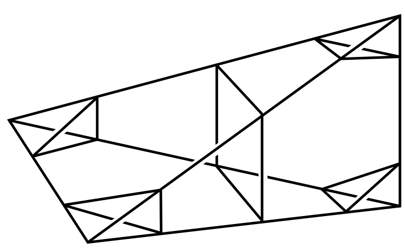


Figure 3.4

There are four types of triangles in each Δ^3 separating one vertex from the other three, and three types of squares separating opposite pairs of edges of Δ^3. Squares of different types necessarily intersect each other, so an embedded surface in normal form can contain squares of only one type in each Δ^3. The edges of the squares and triangles are assumed to be straight line segments, so 'triangles' are literally triangles, but ‘squares’ do not have to lie in a single plane or have all four sides of equal length. ‘Quadrilateral’ would thus be a more accurate term.

For a compact surface S which meets the 1-skeleton of the triangulation transversely we define the **weight** $w(S)$ to be the total number of points of intersection of S with all the edges of the triangulation.

Lemma 3.13. *If N contains an embedded sphere which is nontrivial in $\pi_2 N$, then it contains such a sphere in normal form, and we may take this sphere to have minimal weight among all spheres which are nontrivial in $\pi_2 N$.*

Proof: As in the proof of existence of prime decompositions, we begin by perturbing a nontrivial sphere $S \subset N$ to be transverse to all simplices of $\mathcal{T}$, and then we perform surgeries to produce an S meeting each 3-simplex only in disks. A surgery on S produces a pair of spheres S' and S'', and in $\pi_2 N$, S is the sum of S' and S'', with respect to suitably chosen orientations and paths to a basepoint. So at least one of S' and S'' must be nontrivial in $\pi_2 N$ if S is, and we replace S by this nontrivial S' or S''. Then we repeat the process until we obtain a nontrivial sphere S meeting all 3-simplices in disks. Note that replacing S by S' or S'' does not yield a sphere of larger weight since $w(S) = w(S') + w(S'')$.

Next, we isotope S so that for each 3-simplex Δ^3, the boundary ∂D of each disk component D of $S \cap \Delta^3$ meets each edge of Δ^3 at most once. For if such a circle ∂D meets an edge Δ^1 more than once, consider two points of $\partial D \cap \Delta^1$. After rechoosing these two points if necessary, we may assume the segment α of Δ^1 between these points meets D only in the two points of $\partial\alpha$. The arc α lies in one of the two disks of $\partial\Delta^3$ bounded by ∂D, and separates this disk into two subdisks. Let E be one of these subdisks. We may perturb E so that it lies entirely in the interior of Δ^3 except for the arc α in ∂E, which we do not move, and so that the rest of ∂E, an arc β, stays in D. Now we can perform an isotopy of S which removes the two points of $\partial\alpha = \partial\beta$ from the intersection of S with the edge Δ^1 by pushing β across the repositioned E. It may be that before this isotopy, E intersects S in other curves besides β, but this is not a real problem as we can just push these curves along ahead of the moving β. The key property of this isotopy is that it decreases the weight $w(S)$ by at least two, since it eliminates the two intersection points of $\partial\alpha$ without introducing any new intersections of S with edges of the triangulation.

Since this isotopy decreases $w(S)$, finitely many such isotopies, interspersed with the former surgery steps to make S meet each 3-simplex in disks, suffice to produce a sphere S meeting 3-simplices only in disks, with the boundaries of these disks meeting each 1-simplex at most once. We can then easily isotopy S to meet all 2-simplices in straight line segments, without changing the intersections with edges. We leave it as an exercise for the reader to check that the condition that the boundary circles of the disks of $S \cap \Delta^3$ meet each 1-simplex at most once implies that each of these boundary circles consists of either three or four segments, forming a triangle or square as in Figure 3.4.

Thus we produce a sphere S in normal form which is still nontrivial in $\pi_2 N$. Since the procedure for doing this did not increase the weight of S, it follows that the minimum weight of all homotopically nontrivial spheres in N is achieved by a normal form nontrivial sphere. □

Now we return to the proof of the Sphere Theorem. Choose a triangulation of M and lift this to a triangulation $\mathcal{T}$ of $\widetilde{M}$. By the preceding lemma we may take S_0 to be in normal form with respect to $\mathcal{T}$ and of minimum weight among all nontrivial spheres in $\widetilde{M}$. The immersed sphere $p(S_0) \subset M$ is then composed of triangles and squares in all the 3-simplices of M, and by small perturbations of the vertices of these triangles and squares we can put their boundary edges in general position with respect to each other. This means than any two edges of triangles or squares meet transversely in at most one point lying in the interior of a 2-simplex, and no three edges have a common point. We may also assume the interiors of the triangles and squares intersect each other transversely. Lifting back up to $\widetilde{M}$, this means that the collection Σ of all the translates $\tau(S_0)$ of S_0 under deck transformations $\tau \in \pi_1 M$ consists of minimum weight nontrivial spheres in normal form, having transverse intersections disjoint from the 1-skeleton of $\mathcal{T}$.

Define compact subsets $B_i \subset \widetilde{M}$ inductively, letting B_0 be any 3-simplex of $\mathcal{T}$ and then letting B_i be the union of B_{i-1} with all 3-simplices which meet B_{i-1}. So $B_0 \subset B_1 \subset \cdots \subset \bigcup_i B_i = \widetilde{M}$. Let Σ_i be the collection of all the spheres in Σ which meet B_i. This is a finite collection since only finitely many $\tau(S_0)$'s meet each 3-simplex of $\mathcal{T}$, and B_i is a finite union of 3-simplices.

We describe now a procedure for modifying the spheres in Σ_i to make them all disjoint. Suppose $S \cap S' \cap \mathcal{T}^2 \neq \emptyset$ for some $S, S' \in \Sigma_i$, where $\mathcal{T}^2$ denotes the 2-skeleton of $\mathcal{T}$. As a first step, circles of $S \cap S'$ in the interiors of 3-simplices of $\mathcal{T}$ can be eliminated by isotopy of S without affecting $S \cap \mathcal{T}^2$. Namely, since S and S' are in normal form, such a circle in a simplex Δ^3 bounds disks $D \subset S \cap \Delta^3$ and $D' \subset S' \cap \Delta^3$. We may choose D to be innermost, so that $D \cap S' = \partial D$. Then $D \cup D'$ is a sphere, which bounds a ball in Δ^3 by Alexander's theorem, and we can isotope S' by pushing D' across this ball to eliminate the circle $\partial D = \partial D'$ from $S \cap S'$.

Next we will appeal to the following crucial lemma:

Lemma 3.14. *Let S and S' be minimum weight nontrivial spheres in normal form, intersecting transversely and with $S \cap S'$ disjoint from 1-simplices. Assume that $S \cap S' \neq \emptyset$, and that in each 3-simplex, $S \cap S'$ consists only of arcs, no circles. Then there is a circle C of $S \cap S'$ which is innermost in both S and S', i.e., C bounds disks $D \subset S$ and $D' \subset S'$ with $D \cap S' = \partial D$ and $D' \cap S = \partial D'$. Furthermore we may choose D and D' so that the sphere $D \cup D'$ is trivial in $\pi_2 M$.*

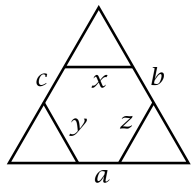
Let us postpone the proof of this lemma, which is rather lengthy, and continue with the proof of the Sphere Theorem. Having a circle C as in this lemma, we replace D by D' in S to obtain a new sphere S, and similarly modify S' by replacing D' by D. Since $D \cup D'$ is trivial, the new spheres S and S' are homotopic to the old ones, hence are nontrivial. They are also still of minimum weight since the total weight $w(S) + w(S')$ is unchanged, hence both terms are unchanged since S and S' each have minimal weight. With a small perturbation of S and S' we can eliminate the circle C from $S \cap S'$. The new spheres S and S' are not in normal form, but can easily be made so: First straighten their arcs of intersection with 2-simplices without changing the endpoints of these arcs; note that these arcs must join distinct edges of each 2-simplex, otherwise S or S' would not be of minimum weight. Having straightened the intersection of S and S' with 2-simplices, the intersections of S and S' with 3-simplices must still be disks, otherwise the procedure for putting them into normal form would decrease their weights, which is not possible. The new S and S' together have no more points of intersection in 2-simplices with other spheres in Σ_i than the old S and S' did, since these other spheres are in normal form, so when we straightened the intersections of the new S and S' with 2-simplices this will not produce new intersections with the other spheres.

Thus by repeating this process for the other circles of $S \cap S'$ and then for all pairs S, S', in Σ_i we eventually obtain a collection Σ_i of spheres which are disjoint in all 2-simplices. They can easily be made to be disjoint in the interior of each 3-simplex Δ^3 as well. For the new spheres of Σ_i meet $\partial\Delta^3$ in a collection of disjoint circles, so we need only choose disjoint disks in Δ^3 bounded by these circles, then isotope the disks of $\Sigma_i \cap \Delta^3$ to these disjoint disks, rel $\Sigma_i \cap \partial\Delta^3$.

Suppose we perform this disjunction procedure on the spheres in $\Sigma_0, \Sigma_1, \Sigma_2, \cdots$ in turn. We claim that no sphere is modified infinitely often. To see this, observe first that each sphere of Σ_i is contained in B_{i+w} where w is the common weight of all the spheres of Σ. This is because the cell structure on each sphere S in Σ consisting of the triangles and squares given by normal form has w vertices, so a maximal tree in its 1-skeleton has $w - 1$ edges and any two vertices can be joined by an edgepath with $w - 1$ edges, so if S meets B_i, S must be contained in B_{i+w}. So a sphere in Σ_i will not be modified at any stage after the spheres which meet B_{i+w} are modified, i.e., the spheres in Σ_{i+w}.

So the infinite sequence of modifications produces a collection Σ of disjoint nontrivial spheres. Let S denote the union of these spheres. This is a surface in normal form, perhaps noncompact. We claim that the intersection of S with the 2-skeleton $\mathcal{T}^2$ is invariant under all the deck transformation. This is certainly true for the intersection with the 1-skeleton $\mathcal{T}^1$, which was unchanged during the whole modification procedure. Observe that for a surface T in normal form, $T \cap \mathcal{T}^1$ determines $T \cap \mathcal{T}^2$ uniquely, since for each simplex Δ^2, the weights a, b, c of T at the three edges of

Δ^2 determine the numbers x, y, z of arcs of $T \cap \Delta^2$ joining each pair of edges, as indicated in Figure 3.8.



$$\begin{aligned} x + y &= c & \qquad x &= \tfrac{1}{2}(b + c - a) \\ x + z &= b & y &= \tfrac{1}{2}(a + c - b) \\ y + z &= a & z &= \tfrac{1}{2}(a + b - c) \end{aligned}$$

Figure 3.8

So for each deck transformation τ, the normal form surfaces S and $\tau(S)$ have the same intersection with $\mathcal{T}^1$ hence also with $\mathcal{T}^2$, as claimed.

Finally, to make S itself invariant under deck transformations we need only choose the disks of $S - \mathcal{T}^2$ to be invariant. This can be done by fixing a choice of these disks in a particular 3-simplex of $\mathcal{T}$ lying over each 3-simplex in M, then translating these disks to all other 3-simplices of $\mathcal{T}$ via the deck transformations. □

Proof of Lemma 3.14: Consider disks $D \subset S$ and $D' \subset S'$ bounded by circles of $S \cap S'$. Such disks may or may not be innermost, i.e., meeting $S \cap S'$ only in their boundary. Here are some facts about such disks.

(1) The minimum weight of disks in S and S' is the same. For consider a disk of minimum weight in both S and S', say $D \subset S$. We may assume D is innermost. Let D' and E' be the disks in S' bounded by ∂D. At least one of the spheres $D \cup D'$, $D \cup E'$ is non-trivial, say $D \cup E'$. Then $w(D \cup E') \geq w(S') = w(D' \cup E')$ since S' has minimal weight. Weights being additive, we get $w(D) \geq w(D')$. This implies $w(D) = w(D')$ by the choice of D, so D' also achieves the minimum weight of disks in S and S'.

(2) If $D \subset S$ is an innermost minimum weight disk, separating S' into disks D' and E' with $D \cup E'$ nontrivial, then D' is a minimum weight disk and $D \cup E'$ is a minimum weight nontrivial sphere. This was shown in (1) except for the observation that $D \cup E'$ is minimum weight, which holds because $w(D \cup E') = w(D' \cup E')$ since $w(D) = w(D')$.

(3) Each disk D has $w(D) > 0$. To verify this we may assume D is innermost. The normal form sphere S is decomposed into triangles and squares by the 2-skeleton $\mathcal{T}^2$. If $w(D) = 0$, D is disjoint from the corners of these triangles and squares, and so meets them in regions of the following types:

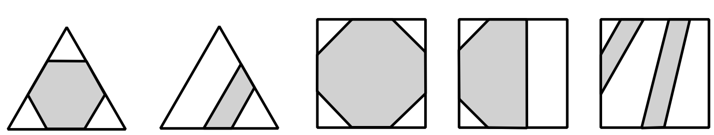


Figure 3.5

Construct a graph $G \subset D$ by placing a vertex in the interior of each of these regions, together with an edge joining this vertex to the midpoint of each of the arcs of intersection of the region with the sides of the triangle or square. Clearly D deformation retracts onto G, so G is a tree, but this contradicts the fact that each vertex of G is incident to at least two edges of G.

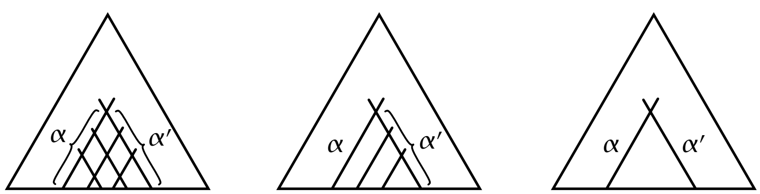
(4) By (3), a minimum weight disk D which is not innermost can contain only one innermost disk D_0, and the annuli in $D - D_0$ bounded by circles of $S \cap S'$ must have weight zero. By the Euler characteristic argument in (3), these annuli must consist entirely of the rectangular shaded regions in Figure 3.5.

(5) For a minimum weight disk D, there must exist in some 2-simplex Δ^2 of $\mathcal{T}$ an arc component α of $D \cap \Delta^2$ having one end on $\partial\Delta^2$ and the other in the interior of Δ^2. This is because D meets $\mathcal{T}^1$ (since $w(D) > 0$), and if a component of $D \cap \mathcal{T}^2$ containing a point of $D \cap \mathcal{T}^1$ consisted entirely of arc components of $D \cap \Delta^2$ with both endpoints on $\partial\Delta^2$, for each Δ^2, then D would be a closed surface.

(6) An arc α as in (5) is divided into segments by the points of $\alpha \cap S'$. Each such segment, except the one meeting $\partial\Delta^2$, must connect the two boundary components of an annulus of $D - S'$, as in (4). The segment of α meeting $\partial\Delta^2$ must lie in the unique innermost disk $D_0 \subset D$.

Having these various facts at our disposal, we now begin the proof proper.

Case I. Suppose we have a minimum weight disk D and an arc α of $D \cap \Delta^2$ as in (5) with the additional property that the endpoint of α in the interior of Δ^2 is the endpoint of an arc α' of $S' \cap \Delta^2$ whose other endpoint lies on the same edge of Δ^2 as the other endpoint of α; see Fig. 3.6a. The two arcs α and α' cut off a triangle T from Δ^2, and the following argument will involve replacing T by subtriangles until the situation of Fig. 3.6c is achieved, with $\alpha \cap S' = \alpha' \cap S = \alpha \cap \alpha'$.



(a) (b) (c)

Figure 3.6

After replacing D by its innermost subdisk we may assume by (6) that $\alpha \cap S' = \alpha \cap \alpha'$, as in Fig. 3.6b. Near $\alpha \cap \alpha'$, α' lies in a disk $D' \subset S'$. If all of α' is not contained in D', then the circle $\partial D = \partial D'$ meets α' at another point in the interior of α', and we can rechoose α to have this point as an endpoint. So we may assume $\alpha' \subset D'$. If $D \cup D'$ is nontrivial, relabel D' as E' in the notation of (1) and (2) above. But then the minimum weight nontrivial sphere $D \cup E'$ can obviously be isotoped to decrease its weight, a contradiction. So $D \cup D'$ is trivial, and by (1) D' is also a minimum weight disk. We can then repeat the preceding steps with the roles of D and D' reversed, obtaining in the end the desired situation of Fig. 3.6c, with both D and D' innermost and minimum weight, and $D \cup D'$ trivial. This finishes the proof in Case I.

Case II. Suppose we have a minimum weight disk D such that $D \cap \Delta^2$ has a component which is an arc β in the interior of Δ^2, for some 2-simplex Δ^2 of $\mathcal{T}$. The arc β is divided into segments by the points of $S' \cap D$. One of these segments must lie in the innermost disk D_0 in D, otherwise these segments would all be arcs crossing the annuli of $D - S'$ as we saw in (4), an impossibility since both ends of β are on ∂D. Replace D by its innermost disk, with the new β a subsegment of the old one. The two endpoints of β are joined to $\partial\Delta^2$ by arcs α_1 and α_2 of $S' \cap \Delta^2$, which may be chosen to have their outer endpoints on the same edge of Δ^2, as in Figure 3.7a.

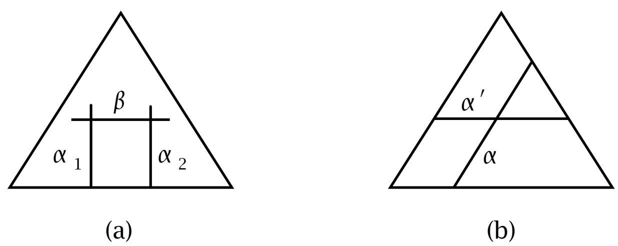


Figure 3.7

If there are other arcs of $D \cap \Delta^2$ meeting α_1 or α_2, these arcs cannot cross α_1 or α_2 since D is innermost. Furthermore, these arcs must lie on the same side of α_1 and α_2 as β since D lies on one side of S' near ∂D. Assuming we are not in Case I, these arcs then join α_1 to α_2, and we can replace β by an edgemost such arc joining α_1 to α_2. So we may assume α_1 and α_2 meet D only at $\partial\beta$. Then α_1 and α_2 lie in a disk D' of S'. If the sphere $D \cup D'$ is nontrivial, it must be minimum weight, by (2). But there is an obvious isotopy of $D \cup D'$ decreasing its weight by two. So $D \cup D'$ is trivial and D' is a minimum weight disk, by (2) again. If D' is innermost, we are done. Otherwise, replace D' by its innermost subdisk, α_1 and α_2 being replaced by subarcs which still meet $\partial\Delta^2$ by (6). If we are not in Case I, the endpoints of α_1 and α_2 in the interior of Δ^2 are then joined by a new arc β in some disk D of S with $\partial D = \partial D'$. By (2) the sphere $D \cup D'$ cannot be nontrivial, since there is an obvious isotopy decreasing its weight. So $D \cup D'$ is trivial and D is minimum weight. If D is not innermost we can go back to the beginning of Case II and repeat the argument with a smaller rectangle bounded by $\alpha_1 \cup \beta \cup \alpha_2$ in Δ^2. So we may assume D is innermost, finishing the proof in Case II.

Case III. Considering again an arc α as in (5), there remains the possibility that the component A' of $S' \cap \Delta^2$ containing the endpoint of α in the interior of Δ^2 has both its endpoints on edges of Δ^2 not containing the other endpoint of α. Thus we have the configuration shown in Figure 3.7b. The arc α lies in a minimum weight disk D

which we may assume is innermost. In particular this means that no arcs of $S' \cap \Delta^2$ below A' join the left and right edges of Δ^2. By (2) we have $\partial D = \partial D'$ for some minimum weight disk D' in S'. Assuming Cases I and II are not applicable, there is an arc α' in D' as shown in the figure, with one endpoint at the upper endpoint of α and the other endpoint in the left edge of Δ^2. If D' is not innermost we may replace it by an innermost disk, choosing a new α to the left of the current one. Thus we may assume both D and D' are innermost. If $D \cup D'$ is nontrivial then by (2) we can rechoose D' on the other side of α, but this configuration would be covered by Case I or II. So we may assume $D \cup D'$ is trivial, and Case III is finished. □

Now we can prove a converse to Proposition 1.6:

Theorem 3.15. *If $\widetilde{M} \to M$ is a covering space with M irreducible and orientable, then $\widetilde{M}$ is irreducible.*

Proof: This follows quite closely the scheme of proof of the Sphere Theorem. The main thing to observe is that Lemmas 3.13 and 3.14 remain valid if 'triviality' for embedded spheres is redefined to mean 'bounding a ball.' The two properties of trivial spheres which were used were: (1) Surgery on a nontrivial sphere cannot produce two trivial spheres, and (2) Triviality is invariant under isotopy.

Suppose $\widetilde{M}$ is reducible, hence contains a nontrivial sphere S_0. The proof of the Sphere Theorem gives a new nontrivial sphere $S \subset \widetilde{M}$ which covers either a sphere or an $\mathbb{R}P^2$ in M. In the latter case the $\mathbb{R}P^2$ is 1-sided since we assume M orientable, and an I-bundle neighborhood of this $\mathbb{R}P^2$ is $\mathbb{R}P^3$ minus a ball, so irreducibility of M implies $M = \mathbb{R}P^3$, hence $\widetilde{M} = S^3$, which is irreducible, a contradiction. In case the sphere $S \subset \widetilde{M}$ covers a sphere in M, the latter sphere bounds a ball in M, which lifts to a ball in $\widetilde{M}$ bounded by S, a contradiction again. □

Recall the example $S^1 \times S^2 \to S^1 \times \mathbb{R}P^2$ of a covering of an irreducible manifold by a reducible manifold, which shows the necessity of the orientability hypothesis in Theorem 3.15. To extend the theorem to the nonorientable case one needs to restrict M to be P^2-*irreducible*, that is, irreducible and containing no 2-sided $\mathbb{R}P^2$'s.

Exercises

1. Show that a closed connected 3-manifold with π_1 free is a connected sum of $S^1 \times S^2$'s, $S^1 \widetilde{\times} S^2$'s, and possibly a homotopy sphere.
2. Extend Proposition 3.4 to the nonorientable case.
3. Show that every closed surface of even Euler characteristic is the boundary of some compact 3-manifold.
4. Show that if M is simply-connected, then every circle in ∂M separates ∂M. In particular, components of ∂M which are closed must be spheres.