Topology and its Applications 13 (1982) 263–282
North-Holland Publishing Company

INCOMPRESSIBLE SURFACES IN PUNCTURED-TORUS BUNDLES

W. FLOYD* and A. HATCHER*

Department of Mathematics, University of California, Los Angeles, CA 90024, USA

Received 29 September 1980
Revised 1 June 1981

We derive in this paper the classification up to isotopy of the incompressible surfaces in hyperbolic 3-manifolds which fiber over the circle with fiber a once-punctured torus. From this classification it follows that most of the 3-manifolds obtained by compactifying these bundles via a circle at infinity are closed hyperbolic 3-manifolds which contain no incompressible surfaces, i.e., are not Haken manifolds.

AMS (MOS) Subj. Class.: 55A25, 55A99

incompressible surface   Haken manifold

In spite of the major role which incompressible surfaces play in 3-manifold topology, there have been rather few 3-manifolds for which the isotopy classification of incompressible surfaces has been worked out completely. Before Thurston's analysis in 1977 of the figure-eight knot complement [5], it appears that the only cases known were Seifert-fibered manifolds and, more generally, graph manifolds [7, 8]. (Also, the fact that in a torus bundle with hyperbolic monodromy the fiber is the only incompressible surface was surely known.) Generalizing the figure-eight knot complement, the next examples treated were 2-bridge knot complements [3]. In the present paper, we generalize the figure-eight knot complement in another direction, to the punctured-torus bundles with hyperbolic monodromy. This has also been done independently by Culler–Jaco–Rubinstein [2].

Our techniques are much the same as in [3], though the details are quite a bit simpler in the present case, reflecting the fact that the number of incompressible surfaces is always finite for the manifolds considered here. Thurston's original approach for the figure-eight complement was somewhat different: to analyse surfaces via their intersections with a nice spine for the knot complement. In an Appendix we pursue this point of view far enough to give Jørgensen's elegant construction of spines for punctured-torus bundles and to describe all the incompressible surfaces via their intersections with these spines.

* Supported in part by NSF Grant MCS 78-04378.

1. Statement of results

A punctured-torus bundle M_φ over S^1 is determined by a monodromy matrix $\varphi \in \mathrm{SL}(2, \mathbb{Z})$. Namely, φ determines a homeomorphism of the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ to itself, fixing the "origin" $0 \in T^2$, hence also a homeomorphism $\varphi : T^2 - \{0\} \to T^2 - \{0\}$, whose mapping torus is M_φ. For notational convenience we shall not distinguish between the open manifold M_φ and its natural compactification obtained by adding a boundary torus; context should make clear which is meant. To avoid special cases, we shall restrict ourselves to the case that φ is hyperbolic (has distinct real eigenvalues), though our method also works, with some modifications, when φ is elliptic or parabolic. We remark in passing that according to Jørgensen [4] and Thurston [6], the open manifold M_φ has a complete hyperbolic structure of finite volume exactly when φ is hyperbolic. For example, the figure-eight knot complement occurs as M_φ for

$$\varphi = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

the simplest hyperbolic φ.

In both stating and proving our results we shall make heavy use, as in [3], of a certain very classical "Diagram of PSL(2, $\mathbb{Z}$)" (see Fig. 1). If we view the hyperbolic plane H as the upper half plane of $\mathbb{C}$, bounded by the circle $\partial H = \mathbb{R} \cup \{\infty\}$, then the Diagram of PSL(2, $\mathbb{Z}$) has as its vertices the points of $\mathbb{Q} \cup \{1/0\} \subset \partial H$ and its edges the geodesics in H with endpoints the pairs a/b, c/d such that $ad - bc = \pm 1$. It is an elementary fact that no two of these geodesics cross each other, and that they determine an ideal triangulation of H. (An *ideal triangle* in H has its vertices on ∂H.) SL(2, $\mathbb{Z}$) acts on $H \cup \partial H$ via its quotient $\mathrm{PSL}(2, \mathbb{Z}) = \mathrm{SL}(2, \mathbb{Z})/\pm I$ as Möbius transformations, which are hyperbolic isometries of H. This action takes the Diagram (Fig. 1) to itself, since the determinant condition $ad - bc = \pm 1$ is preserved under multiplication by elements of SL(2, $\mathbb{Z}$). The action is transitive on edges ($\left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$ takes the edge joining $0/1$ and $1/0$ to the edge joining b/a and d/c), and in fact, PSL(2, $\mathbb{Z}$) is the group of all combinatorial symmetries of the Diagram which preserve orientation of H. The action of PSL(2, $\mathbb{Z}$) extends naturally to an action of $\mathrm{PGL}(2, \mathbb{Z}) = \mathrm{GL}(2, \mathbb{Z})/\pm I$ on the Diagram as hyperbolic isometries, including orientation reversing ones. This gives PGL(2, $\mathbb{Z}$) as the full combinatorial symmetry group of the Diagram.

Briefly, the Diagram arises when analyzing surfaces in M_φ because its points with rational barycentric coordinates correspond, with relatively small ambiguity, to the systems of disjoint (non-trivial) curves on a punctured torus, such as occur as the transverse intersections of a surface $S \subset M_\varphi$ with the fibers of M_φ. Lifting to the covering $(T^2 - \{0\}) \times \mathbb{R}$ and letting the fiber $(T^2 - \{0\}) \times t$ vary from $t = -\infty$ to $t = +\infty$, we obtain from S a bi-infinite sequence of curve systems on $T^2 - \{0\}$, i.e., points in the Diagram. This sequence of points is by construction invariant under φ. Joining successive points in the sequence by geodesic segments, we obtain a

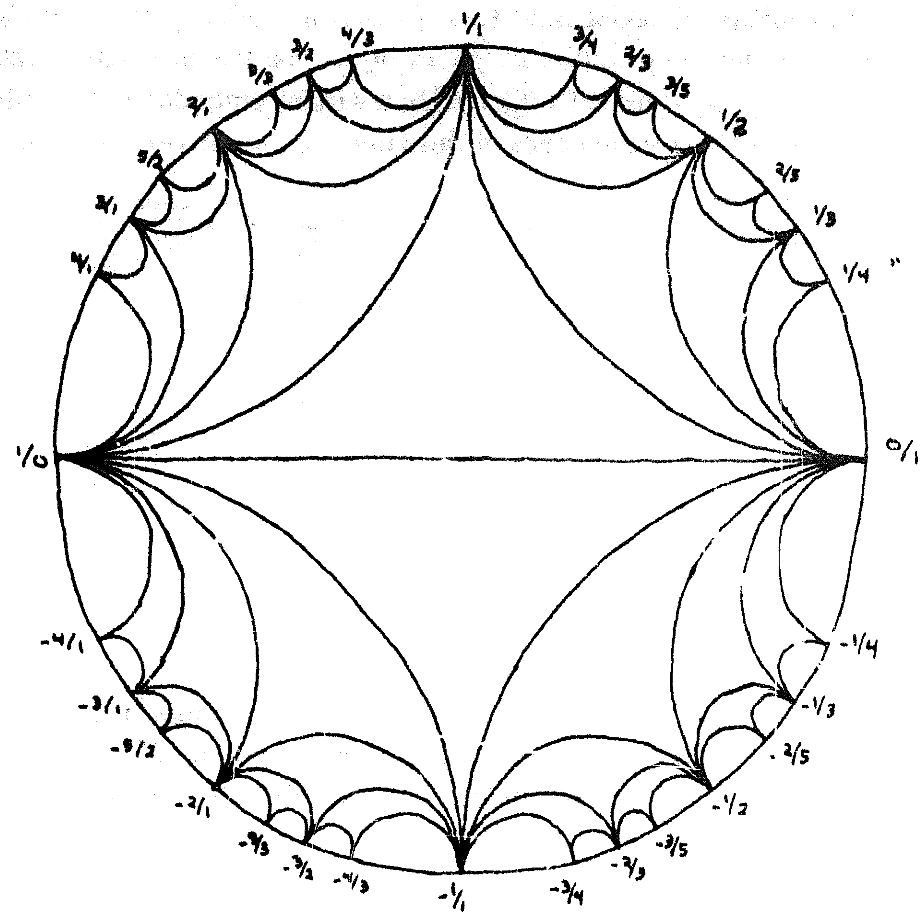


Fig. 1. Diagram of PSL(2, Z).

φ-invariant polygonal path in the Diagram. Incompressibility then can be translated into very restrictive conditions on this path; for example, it lies in the 1-skeleton of the Diagram.

Here is the main result. (The smooth category is understood; surfaces $S \subset M_\varphi$ are always properly embedded, i.e., $S \cap \partial M_\varphi = \partial S$.)

Theorem 1.1. *If φ is hyperbolic, then a connected, orientable, incompressible, ∂-incompressible surface in M_φ is isotopic to exactly one of:*

(a) *the peripheral torus ∂M_φ,*

(b) *the fiber $T^2 - \{0\}$,*

(c) *a finite number (≥ 2) of non-closed surfaces S_γ indexed by edge-paths γ in the Diagram of* PSL(2, $\mathbb{Z}$), *which are invariant by φ and minimal in the sense that no two successive edges of γ lie in the same triangle.*

Corollary 1.2. *Let M be a closed 3-manifold containing a knot K such that $M - K$ is homeomorphic to M_φ for some hyperbolic $\varphi \in$* SL(2, $\mathbb{Z}$). *Then all but finitely many Dehn surgeries on K yield non-Haken hyperbolic (hence irreducible but not Seifert-fibered) 3-manifolds.*

To define the surface S_γ associated to a φ-invariant edge-path γ (minimal or not) in the Diagram, let $\cdots a_{-1}/b_{-1}, a_0/b_0, a_1/b_1, \ldots$ be the successive vertices of γ, with $\varphi(a_i/b_i) = a_{i+k}/b_{i+k}$ for all i. Consider a saddle embedded in a cube (see Fig. 2). Deleting the four vertical edges of this cube and identifying opposite lateral

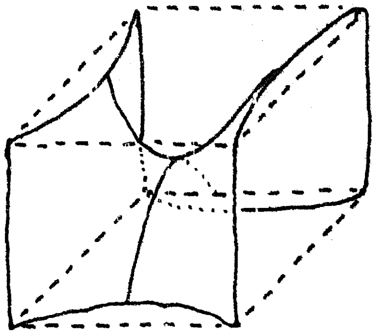


Fig. 2.

faces of the cube yields $(T^2 - \{0\}) \times I$, containing this same saddle surface. The two bottom edges of the surface are two parallel lines of slope 0 in $(T^2 = \{0\}) \times 0$, while the top edges are of slope ∞. Applying in each level $(T^2 - \{0\}) \times t$ the linear homeomorphism

$$\begin{bmatrix} b_i & b_{i+1} \\ a_i & a_{i+1} \end{bmatrix}$$

of $T^2 - \{0\}$ yields a saddle S_i whose bottom edges have slope a_i/b_i and whose top edges have slope a_{i+1}/b_{i+1}. Let $\tilde{S}_\gamma \subset (T^2 - \{0\}) \times \mathbb{R}$ be the surface which is S_i in the slab $(T^2 - \{0\}) \times [i/k, (i+1)/k]$, the two arcs of $\tilde{S}_\gamma \cap [(T \times \{0\}) \times i/k]$ having slope a_i/b_i for each i. By construction $\tilde{S}_\gamma$ is invariant under the map $\Phi(x, t) = (\varphi(x), t+1)$ of $(T^2 - \{0\}) \times \mathbb{R}$, and thus determines a surface $\tilde{S}_\gamma/\Phi$ in $(T^2 - \{0\}) \times \mathbb{R}/\Phi = M_\varphi$. One can easily see that $\tilde{S}_\gamma$ is orientable, while $\tilde{S}_\gamma/\Phi$ is orientable if and only if k is even. We set

$$S_\gamma = \begin{cases} \tilde{S}_\gamma/\Phi & \text{if } k \text{ even}, \\ \partial N(\tilde{S}_\gamma/\Phi) & \text{if } k \text{ odd} \end{cases}$$

where $N(\tilde{S}_\gamma/\Phi)$ is a tubular neighborhood of $\tilde{S}_\gamma/\Phi$, so that S_γ is the orientable double cover of $\tilde{S}_\gamma/\Phi$ if k is odd. As we shall show, S_γ is incompressible if and only if γ is minimal.

Next we describe the behavior of φ-invariant minimal edge-paths in the Diagram. Since φ is hyperbolic, it acts on H as translation along the geodesic l whose endpoints in ∂H are the slopes of the two eigenvectors of φ. The strip Σ_φ consisting of triangles in the Diagram which meet l is also translated along itself, so the triangulation of Σ_φ is periodic, and has the form shown in Fig. 3. Here the numbers $a_i \geq 1$ indicate the number of smaller triangles in each larger triangle. The cycle

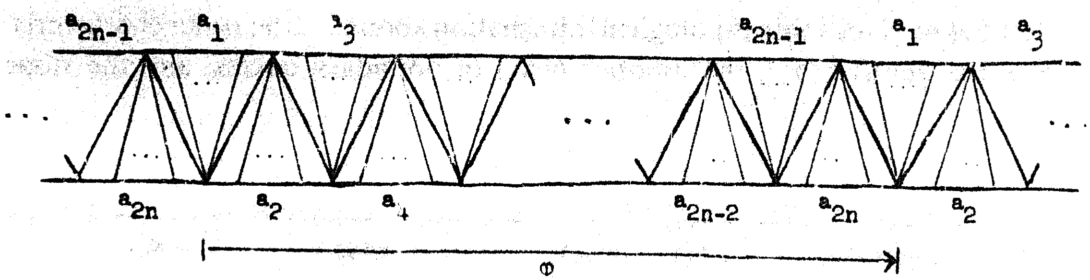


Fig. 3.

$(a_1, \ldots, a_{2n})$ is clearly an invariant of the conjugacy class of φ in PGL(2, $\mathbb{Z}$), in fact a complete invariant, since PGL(2, $\mathbb{Z}$) is the full symmetry group of the Diagram, as mentioned earlier. As a representative of the conjugacy class corresponding to the cycle $(a_1, \ldots, a_{2n})$ one can choose

$$\varphi = \pm\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}^{a_1}\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}^{a_2}\cdots\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}^{a_{2n}} \quad (a_i \geq 1),$$

the sign being the sign of the eigenvalues of φ. (This sign is precisely what is lost in passing from GL to PGL.)

Lemma 1.3. *A minimal edge-path invariant under φ must lie in the strip Σ_φ.*

This is fairly evident, but we shall give a proof later. Obviously, φ-invariant minimal edge-paths in Σ_φ can involve only the heavily shaded edges which form the large triangles, and they must proceed monotonically from left to right. The number of such edge-paths is clearly finite. For example, for $\varphi = [\begin{smallmatrix}2 & 1\\1 & 1\end{smallmatrix}]$, Σ_φ has the form shown in Fig. 4. The only invariant minimal edge-paths here are the two

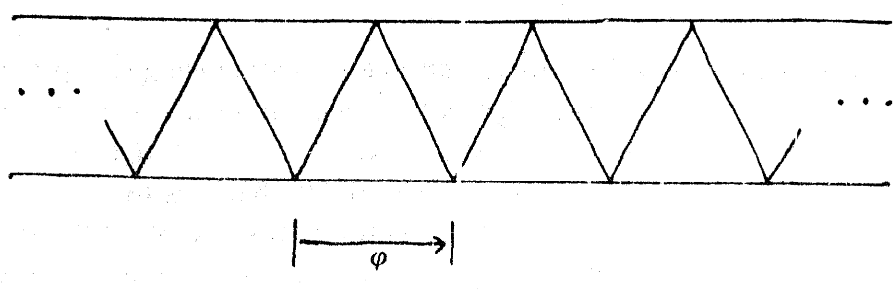


Fig. 4.

which every Σ_φ has: the two borders of Σ_φ. So according to Theorem 1.1, the figure-eight knot complement has just four connected, orientable, incompressible, ∂-incompressible surfaces, as was shown in [5].

Table 1 below gives some topological information about S_γ: the Euler characteristic $\chi(S_\gamma)$, the genus $g(S_\gamma)$, the number $b(S_\gamma)$ of boundary circles, and the slope

Table 1

$k(=L+R)$	eigenvalues of φ	$\chi(S_\gamma)$	$b(S_\gamma)$	$m(S_\gamma)$	$g(S_\gamma)$
odd	+	$-2k$	2	$\frac{1}{4}(L-R)$	k
odd	−	$-2k$	2	$\frac{1}{4}(L-R+2)$	k
even	+	$-k$	$gcd(L-R, 4)$	$\frac{1}{4}(L-R)$	$\frac{1}{2}k$ if $b=2$
even	−	$-k$	$gcd(L-R+2, 4)$	$\frac{1}{4}(L-R+2)$	$\frac{1}{2}k-1$ if $b=4$

$m(S_\gamma)$ of these boundary circles with respect to some choice of coordinates on the torus ∂M_φ in which the fibers have slope ∞. (A different choice of coordinates can change these slopes all by an integer or by a sign.) The numbers L and R are the numbers of vertices of γ, in one full period, at which γ turns left or right, respectively; equivalently, these are the numbers of vertices of γ on each of the two borders of Σ_φ, in each period.

In particular, $b(S_\gamma)$ is always 2 or 4. So the only orientable incompressible, ∂-incompressible surface with a single circle as boundary is the fiber $T^2-\{0\}$. It follows by well-known arguments that the group $\pi_0\,\mathrm{diff}(M_\varphi)$ of diffeomorphisms of M_φ modulo isotopy can be computed by restricting attention to diffeomorphisms which are fiber-preserving and linear on each fiber. Thus $\pi_0\,\mathrm{diff}(M_\varphi)$ is $N\langle\varphi\rangle/\langle\varphi\rangle$, where $\langle\varphi\rangle$ is the subgroup of $\mathrm{GL}(2,\mathbb{Z})$ generated by φ and $N\langle\varphi\rangle$ is the normalizer of $\langle\varphi\rangle$ in $\mathrm{GL}(2,\mathbb{Z})$. The $\mathbb{Z}_2$-quotient $N\langle\varphi\rangle/\langle\varphi, \pm I\rangle$ can be described neatly in terms of the Diagram of $\mathrm{PSL}(2,\mathbb{Z})$: it is the symmetry group of the triangulated cylinder Σ_φ/φ. (This symmetry group contains a cyclic or dihedral subroup of index at most two.) Also, one can easily see which surfaces S_γ are equivalent under diffeomorphisms of M_φ, namely, those whose γ's are equivalent under $N\langle\varphi\rangle$.

Proof of Lemma 1.3. To conclude this introductory section we give a proof of the lemma, that a φ-invariant minimal edge-path γ must lie in Σ_φ. If $\gamma\not\subset\Sigma_\varphi$, then for some edge e in the border of Σ_φ, γ has an edge in the component C of $H-\Sigma_\varphi$ bounded by e. It cannot be the case that some infinite half-segment of γ (i.e., all edges after, or before, some edge of γ) lies entirely in C, since all the translates of C by powers of φ are disjoint, and γ is φ-invariant. Hence some finite segment σ of γ enters and leaves C at endpoints of e. We may filter C by finite subcomplexes C_i, where C_0 is the triangle containing e, C_1 is obtained by adding the two triangles of C bordering C_0, C_2 is obtained by adding the four triangles bordering C_1, and so on. Let C_n be the smallest C_i containing σ. Since each triangle of C_n-C_{n-1} meets C_{n-1} in only one edge, we see that σ must have two successive edges in one triangle of C_n-C_{n-1}, making γ non-minimal. $\square$

2. Surface bundles in general

Let $\varphi : F \to F$ be a diffeomorphism of the compact connected surface F, and let $F \to M_\varphi \xrightarrow{p} S^1$ be the associated bundle. We suppose $F \neq S^2, P^2, D^2$. Then $M_\varphi - \partial M_\varphi$ has $\mathbb{R}^3$ as universal cover, making M_φ irreducible. Let $S \subset M_\varphi$ be an incompressible surface, in the geometric sense that every circle on S bounding a disk in $M_\varphi - S$ also bounds a disk in S. We assume that no component of S is peripheral—isotopic to a subsurface of ∂M_φ. After an isotopy of S we may assume:

(*) Each component of ∂S either is transverse to all the fibers of M_φ, or lies entirely in a fiber of M_φ.

By well-known techniques we shall prove:

Proposition 2.1. *S can be isotoped, preserving (*), so that either all critical points of $p|S$ are saddles in $S - \partial S$, or S is a union of fibers of M_φ.*

Proof. We may suppose S has been isotoped to minimize the number of components of the transverse intersection of S with a fiber F_0 of M_φ (preserving (*), always). If $S \cap F_0 = \emptyset$, then S is isotopic to a union of fibers of M_φ, by the following lemma.

Lemma 2.2. *An incompressible surface in a product $F \times I$, disjoint from $F \times \partial I$ and having no components which are peripheral annuli, is isotopic to a union of fibers $F \times \{x\}$.*

With orientability hypotheses (which are not really necessary) this follows from Lemma 3.1 of [8]; in any case, the proof is an elementary exercise.

So we may assume $S \cap F_0 \neq \emptyset$. If we split open (M_φ, S) along F_0, we obtain a pair $(F \times I, S')$ with S' incompressible in $F \times I$ (otherwise there would be contractible circles in $S \cap F_0$, which could be eliminated by isotopy of S). The projection $p' : S' \to I$ we may take to be a morse function with all critical points in $S' - \partial S'$, in distinct levels. Assuming there are local maxima or minima on S', we shall give a procedure for isotoping S' rel $S' \cap (F \times \partial I)$ to decrease the total number of these local maxima and minima.

Suppose without loss that S' has local maxima, and let $m \in S'$ be the one of these at the lowest level. In a level $F_t = F \times \{t\}$ just below m, there is a disk $X_t \subset F_t$ such that

(i) $X_t \cap S' = \partial X_t$,

(ii) ∂X_t cuts off a subsurface S'_t from S', $m \in S'_t$, and $X_t \cup S'_t$ bounds a product cobordism in $F \times I$ (pinched at $\partial X_t = \partial S'_t$).

The idea now is to let t decrease further, assuming inductively that we have a connected subsurface $X_t \subset F_t$ satisfying (i) and (ii). Everything varies only by isotopy unless one of the following three situations occurs.

(1) X_t encounters a saddle s of S', say in the level F_{t_0}. We regard s in the familiar way as a 1-handle h attached at its two "ends" to $S' \cap F_{t_0+\varepsilon}$, with one or both ends on $\partial X_{t_0+\varepsilon}$. There are three possibilities:

(a) h projects inside $X_{t_0+\varepsilon}$. Then the following steps produce an isotopy of S' decreasing the number of its local maxima and minima. First, isotope $S'_{t_0+\varepsilon}$ rel its boundary down to $X_{t_0+\varepsilon}$. Then the core arc of h projects up to an arc $\alpha \subset S'_{t_0+\varepsilon}$, and there is a disc $D \subset F \times I$, $D \cap S' = \partial D$, with ∂D the union of α and the core arc of h. Hence there is a disc $D' \subset S'$ with $\partial D' = \partial D$. Isotope S' by moving D' to D. Now the part of the resulting S' contained in the original $X_{t_0+\varepsilon}$ can be pushed back up above the level $F_{t_0+\varepsilon}$ so as to make the projection to I a morse function with only one local maximum on it and no local minima. This process has decreased the total number of local maxima and minima on S' since D' had to contain a local minimum below $F_{t_0+\varepsilon}$ (see Fig. 5).

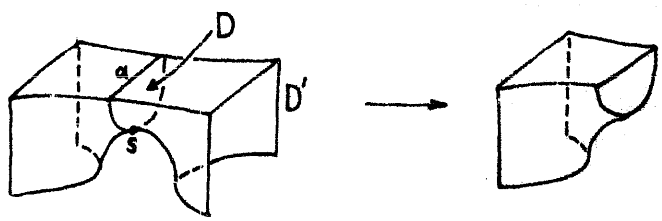


Fig. 5.

(b) h projects outside $X_{t_0+\varepsilon}$ and has only one end on $\partial X_{t_0+\varepsilon}$. Then we can push $S'_{t_0+\varepsilon}$ down to lie just above $F_{t_0+\varepsilon}$ and cancel s and m in the usual way (see Fig. 6).

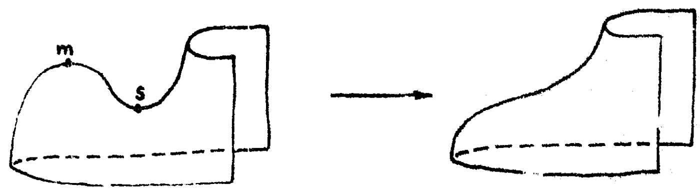


Fig. 6.

(c) h projects outside $X_{t_0+\varepsilon}$ and has both ends on $\partial X_{t_0+\varepsilon}$. Then clearly there exists a subsurface $X_{t_0-\varepsilon} \subset F_{t_0-\varepsilon}$ satisfying (i) and (ii). Topologically, $X_{t_0-\varepsilon}$ is $X_{t_0+\varepsilon}$ with h attached.

(2) X_t encounters a local minimum of S', in a level F_{t_0}. Regarding this as a 2-handle h^2 attaching to $\partial X_{t_0+\varepsilon}$, we see that h^2 must project outside $X_{t_0+\varepsilon}$, otherwise its projection into $F_{t_0+\varepsilon}$ would be $X_{t_0+\varepsilon}$ and $S'_{t_0+\varepsilon} \cup h^2$ would be a 2-sphere component of S'. Since h^2 projects outside $X_{t_0+\varepsilon}$, we can just add it to $X_{t_0+\varepsilon}$ to form $X_{t_0-\varepsilon}$ satisfying (i) and (ii).

(3) A boundary circle of X_t approaches a boundary circle C of S' in a level F_{t_0}. In this case we can eliminate m (and any other local maxima or minima in S'_{t_0}) by first pushing S'_{t_0} down to X_{t_0}, then pushing it back up slightly so that $p'|S'_{t_0}$ is a morse function $(S'_{t_0}, \partial S'_{t_0} - C, C) \to ([0, \varepsilon], 0, \varepsilon)$ having only saddles.

If iteration of the preceding steps does not decrease the number of local maxima and minima of S', then we reach the level $t=0$ with X_0 satisfying (i) and (ii). If $X_0 = F$, then $S' \cap F_0 = \emptyset$, a case previously considered. If $X_0 \neq F$, then we could decrease the number of components of $S \cap F_0$ by isotoping S'_0 to X_0 and a little beyond, thereby removing $\partial S'_0$ from $S \cap F_0$, contrary to hypothesis. □

Lemma 2.3. *Suppose the fibering $M_\varphi \to S^1$ restricted to ∂M_φ is trivial. Let $S \subset M_\varphi$ be an incompressible surface without peripheral components and with the property that for some fiber F_0 transverse to S, $S \cap F_0$ consists entirely of circles which are either contractible in F_0 or isotopic in F_0 to components of ∂F_0. Then S is isotopic to a union of fibers of M_φ.*

Proof. First S can be isotoped in the usual way to eliminate all trivial circles of $S \cap F_0$. Then we can perform a sequence of "peripheral surgeries" on S to eliminate any remaining circles of $S \cap F_0$ one by one (see Fig. 7). The result is a sequence

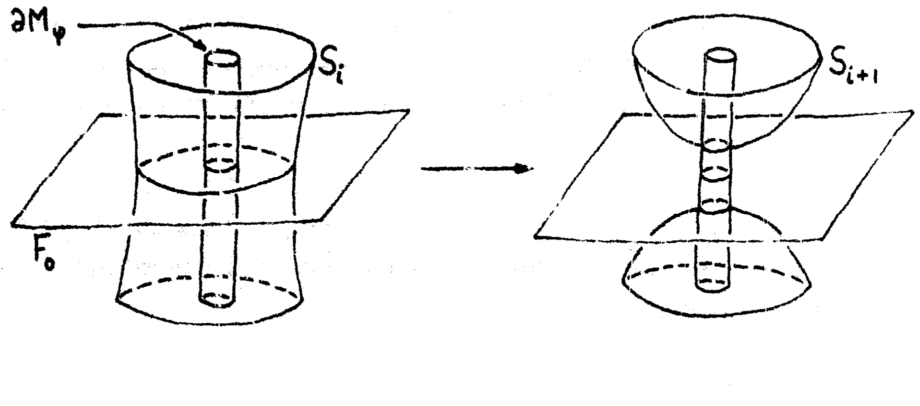


Fig. 7.

of surfaces $S_0 = S, S_1, \ldots, S_n$, with $S_n \cap F_0 = \emptyset$. It is easy to see that S_i incompressible implies S_{i+1} incompressible, since a candidate for a compressing disc for S_{i+1} can be pushed away from the region where the surgery was performed. By Lemma 2.2, S_n consists of fibers of M_φ, up to isotopy, plus possibly some peripheral annuli. Let S_k be the last S_i in the sequence which has this form, of fibers and peripheral annuli. If $k > 0$, then S_{k-1} is obtained from S_k by "peripherally tubing" two distinct fiber components of S_k together. It follows that S_{k-1} can be isotoped to be disjoint from some fiber of M_φ, so the component of S_{k-1} resulting from tubing the two fibers together would violate Lemma 2.2. Hence $k = 0$. □

3. Punctured-torus bundles

Let F be a torus with an open disk removed. We first recall some elementary facts about curve systems on F (compact 1-manifolds $C \subset F$ with $C \cap \partial F = \partial C$). A circle on F has one of three types:
- trivial: bounding a disk in F
- peripheral: isotopic to ∂F
- essential: representing a non-zero class in $H_1(F)$.

If an essential circle represents $\pm(a, b) \in H_1(F)$, with respect to a basis for $H_1(F)$ chosen once and for all, we call $b/a \in \mathbf{Q} \cup \{1/0\}$ the *slope* of the circle. (This does not depend on a choice of orientation for the circle.) Similarly, an arc on F (with ends on ∂F) can be of two types:
- peripheral: isotopic to an arc in ∂F (keeping ends on ∂F, of course)
- essential: representing a non-zero class in $H_1(F, \partial F) \approx H_1(F)$.

Again, an essential arc has a well-defined slope in $\mathbf{Q} \cup \{1/0\}$. An essential circle or arc is determined up to isotopy by its slope.

If a finite system of disjoint essential circles and arcs on F contains any circles, these circles must all have the same slope, and any arcs in the system must also have this same slope. If the curve system has no circles, just essential arcs, then these arcs can have at most three different slopes, which must form the vertices of a triangle (or edge, in the case of just two slopes) in the Diagram of PSL(2, $\mathbf{Z}$). (To see this, one can first by a homeomorphism of F take all the arcs of the first slope to parallel meridians on F, (slope $1/0$) then by a further homeomorphism, fixing these meridians, take all the arcs of a second slope to parallel longitudes (slope $0/1$). Any essential arcs disjoint from all these meridians and longitudes must then have slope ± 1.)

In any case, a system of disjoint essential circles and arcs is determined up to isotopy by the numbers of circles and arcs of each slope. To a curve system on F having n_i curves (circles or arcs) of slope a_i/b_i we associate the point in the Diagram having barycentric coordinates

$$\sum_i n_i(a_i/b_i) \Big/ \sum_i n_i$$

where (a_i/b_i) denotes the vertex labelled a_i/b_i. This definition applies also for curve systems containing inessential curves, simply by ignoring any such curves, except of course if the system consists entirely of inessential curves.

Proof of Theorem 1.1. Let $S \subset M_\varphi$ be an orientable, incompressible, ∂-incompressible surface, not necessarily connected. We suppose no components of S are tori isotopic to ∂M_φ (peripheral annuli are ruled out by ∂-incompressibility) or punctured tori isotopic to fibers of M_φ. We may isotope S so that condition (*) of section 2 holds, that is, ∂S either lies in fibers of M_φ or is transverse to all fibers of M_φ. In the latter case, no transverse intersection of S with a fiber can contain arcs which are peripheral in that fiber, since S is ∂-incompressible.

Let $\tilde{S}$ be the preimage of S in the cover $F\times\mathbb{R}$ of M_φ. Specifying a product structure on $F\times\mathbb{R}$ means that each transverse intersection $\tilde{S}\cap F_t$, where $F_t = F\times t\subset F\times\mathbb{R}$, determines a point $\lambda(F_t)$ in the Diagram of PSL(2, $\mathbb{Z}$). (The possibility that $\tilde{S}\cap F_t$ contains no essential curves is excluded by Lemma 2.3 and our hypotheses on S.) Perturbing S slightly, we may suppose the projection $\tilde{S}\to\mathbb{R}$ is a morse function having all its critical points in distinct fibers. The point $\lambda(F_t)$ can then change only at isolated t values, namely the saddles of $\tilde{S}\to\mathbb{R}$. Letting t go from $-\infty$ to $+\infty$, the succession of points $\lambda(F_t)$ forms the *λ-sequence* of S, $\{\lambda_i \mid i\in\mathbb{Z}\}$, where by definition $\lambda_{i+1}\neq\lambda_i$ for all i. Since φ is hyperbolic, it leaves no finite subset of the Diagram invariant, hence the λ-sequence must be infinite.

We can now easily dispose of the case that the components of ∂S lie in fibers of M_φ (including the case $\partial S=\emptyset$). Under this hypothesis each transverse intersection $\tilde{S}\cap F_t$ contains only circles. As F_t varies through a fiber T_{t_0} containing a saddle of $\tilde{S}$, the curve system $\tilde{S}\cap F_{t_0-\varepsilon}$ changes to another curve system $\tilde{S}\cap F_{t_0+\varepsilon}$ whose projection into $F_{t_0-\varepsilon}$ can be isotoped to be disjoint from $\tilde{S}\cap F_{t_0-\varepsilon}$. (Near the saddle such a disjunction isotopy clearly exists; to extend to the rest of $\tilde{S}\cap F_{t_0+\varepsilon}$ all one needs is that $\tilde{S}$ be 2-sided, to determine a well-defined direction in which to push the projection of $\tilde{S}\cap F_{t_0+\varepsilon}$ off $\tilde{S}\cap F_{t_0-\varepsilon}$.) Since disjoint essential circles on the punctured torus must have the same slope, we conclude that the λ-sequence of S would be constant, a contradiction.

So now we may assume ∂S is non-empty and consists of circles transverse to the fibers of M_φ. There can be no peripheral circles in any transverse intersection $\tilde{S}\cap F_t$, since this would force peripheral arcs, previously ruled out. By Proposition 2.1, we may isotope S to eliminate all non-saddle critical points of the projection $\tilde{S}\to\mathbb{R}$. It follows that no transverse intersection $\tilde{S}\cap F_t$ can contain a trivial circle, since such a circle would bound a disk on $\tilde{S}$ which would have to have a local minimum or maximum in its interior. Nor can there be any essential circles of a transverse intersection $\tilde{S}\cap F_t$. For if there were, some saddle (1-handle) would have to have at least one end attached to this essential circle (the λ-sequence is not constant). Examining the small number of possibilities for such a saddle, one sees that after passing through the fiber containing the saddle, $\tilde{S}\cap F_t$ would have to contain a trivial circle or peripheral arc (see Fig. 8).

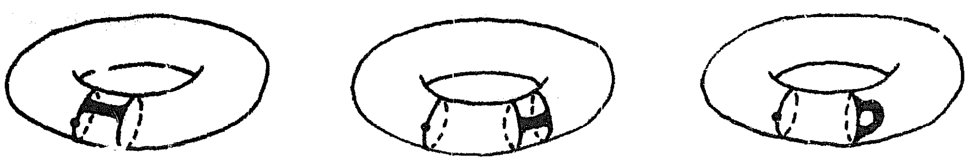


Fig. 8.

Thus each transverse intersection $\tilde{S}\cap F_t$ consists entirely of essential arcs. In any such $\tilde{S}\cap F_t$ there cannot be arcs of three distinct slopes, for if there were, then each component of $F_t-\tilde{S}$ would be a planar region bounded by 2 or 3 arcs of $\tilde{S}\cap F_t$

(plus arcs of ∂F_t), and the next saddle would yield either a peripheral arc (if the saddle joined two different arcs of $\mathring{S} \cap F_t$) or a trivial circle (if the saddle joined an arc of $\mathring{S} \cap F_t$ to itself).

In fact there is just one admissible position for a saddle, up to fiber-preserving isotopy: it must join two arcs of the same slope a/b and be disjoint from any other arcs of slope a/b or second slope c/d which might be present (see Fig. 9). For, a saddle joining an arc to itself would yield either a trivial circle (if both ends of the saddle attached to the same side of the arc) or a nonorientable surface (in the opposite case); while a saddle joining two arcs of different slopes would yield a peripheral arc.

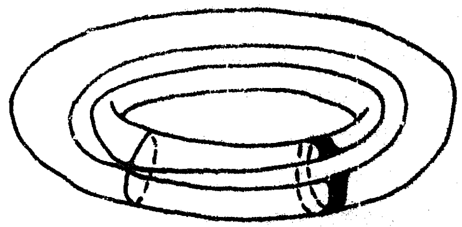


Fig. 9.

Thus in passing a saddle, $\lambda(F_t)$ changes from the point $[m(a/b)+n(c/d)]/m+n$ to $[(m-2)(a/b)+(n+2)(c/d)]/m+n$. Both m and n must be even, for otherwise, after passing more saddles we would eventually reach the impossible case $m=1$. As a result, we see that $\mathring{S}$ consists of some number of parallel copies of the surface $\mathring{S}_\gamma$ defined in the Introduction, for some φ-invariant edge-path γ. The surface S then lies in a tubular neighborhood $N(\mathring{S}_\gamma/\Phi)$ of $\mathring{S}_\gamma/\Phi$ such that the projection $N(\mathring{S}_\gamma/\Phi) \to \mathring{S}_\gamma/\Phi$ restricts to a covering map $S \to \mathring{S}_\gamma/\Phi$. If S is connected this covering must be either 1–1 or 2–1 depending on whether $N(\mathring{S}_\gamma/\Phi)$ is a product or not. Hence S is isotopic to the surface S_γ.

If there were back-tracking in the λ-sequence, i.e., $\lambda_{i+2}=\lambda_i$ $(\neq \lambda_{i+1})$, then the two saddles S_i and S_{i+1} changing λ_i to λ_{i+1} and λ_{i+2} could be isotoped to lie in the same fiber (see Fig. 10). But then if s_{i+1} is moved before s_i in the sequence, it becomes a saddle not of the admissible type. Similarly, if λ_i and λ_{i+2} lay on two different edges of one triangle of the Diagram, with λ_{i+1} at the common vertex of

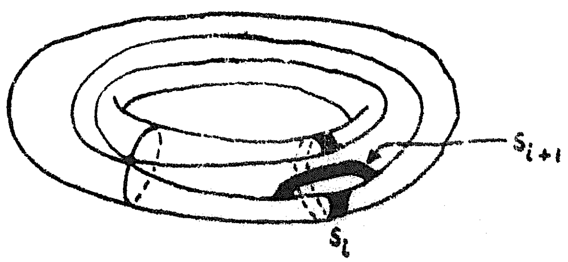


Fig. 10.

these two edges, we could put the two relevant saddles s and s_{i+1} into one fiber (see Fig. 11). Here, s_{i-1} becomes an inadmissible saddle if it is put before s_i.

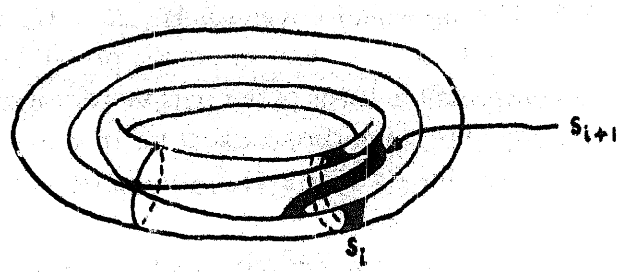


Fig. 11.

Thus we have shown that S can be isotoped so that its components are parallel copies of the surface S_γ, for some φ-invariant minimal edge-path γ. Note that this is enough already to deduce Corollary 1.2.

It remains to show that if γ is minimal, then S_γ is incompressible and ∂-incompressible, and γ is invariant under isotopy of S_γ. Consider the following properties of the surface S_γ:

(i) The intersection of $\tilde{S}_\gamma$ with each non-critical fiber F_t contains no peripheral arcs.

(ii) Each circle of intersection of $\tilde{S}_\gamma$ with a non-critical fiber F_t bounds a disc on $\tilde{S}_\gamma$.

(iii) The λ-sequence of $\tilde{S}_\gamma$ traces out the vertex sequence of the given minimal edge-path γ.

We claim that (i)–(iii) are preserved by any isotopy of $\tilde{S}_\gamma$ (rel $\partial\tilde{S}_\gamma$). If this is so, the rest of Theorem 1.1 follows. For suppose S_γ were compressible. A compressing disc would lift to a disc $D \subset F \times \mathbb{R}$ with $D \cap \tilde{S}_\gamma = \partial D$. A small subdisc D' can be isotoped to lie in a fiber F_t. Shrinking D to D' extends to an isotopy of $\tilde{S}_\gamma$ to $\tilde{S}'_\gamma$. Condition (ii) implies that $\partial D'$ bounds a disc on $\tilde{S}'_\gamma$, so ∂D bounds a disc on $\tilde{S}_\gamma$, and the image of ∂D would bound a disc on S_γ. Thus S_γ is incompressible. Since S_γ is not an annulus, it is also ∂-incompressible. (In an irreducible 3-manifold with torus boundary components, an incompressible surface is either ∂-incompressible or a ∂-parallel annulus [7].)

To prove the claim, consider a generic isotopy of $\tilde{S}_\gamma$. At each time during this isotopy, the projection to $\mathbb{R}$ will be a morse function, except for the following isolated phenomena (see, e.g., [1]):

(A) A saddle and a local maximum (or minimum) are introduced or cancelled in a region $(T^2-\{0\})\times(a, b)$ containing no other critical points.

(B) A pair of critical points interchange levels.

Clearly, (i)–(iii) could be affected only when two saddles interchange levels in (B). Up to fiber-preserving diffeomorphism, there is a rather small finite number of possibilities for such a pair of saddles. Without being fancy, one can just check each of these possibilities in turn, to see that (i), (ii), and (iii) (together) are preserved. We leave this tedious task to the reader. (A slightly more general argument at this point is given in [3].) □

Proof of Corollary 1.2. The universal cover of $M-K=M_\varphi$ is $\mathbb{R}^3$, so $M-K$ is irreducible. By Theorem 1.1, $M-K$ has no incompressible, ∂-incompressible annuli, and the only incompressible torus is the peripheral torus. So by Thurston's theorem [6], M_φ has a complete hyperbolic structure of finite volume. Hence by Theorem 5.9 of [5], all but finitely many Dehn surgeries on K give closed 3-manifolds having hyperbolic structures.

On the other hand, since M_φ has only finitely many connected orientable incompressible, ∂-incompressible surfaces, up to isotopy, then only finitely many Dehn surgeries on K could yield closed manifolds containing (closed) orientable incompressible surfaces. This follows by a well-known general argument – see [3] or Chapter 4 of [5]. □

Appendix

A different view of M_φ and its incompressible surfaces

In [5], the incompressible surfaces in M_φ, $\varphi=\left[\begin{smallmatrix}2&1\\1&1\end{smallmatrix}\right]$, were analyzed in the following steps:

(1) "Triangulate" the open manifold M_φ into ideal tetrahedra (tetrahedra with their vertices deleted).

(2) Represent M_φ as the mapping cylinder of a map $T^2\to T^2/\tau$, where τ identifies pairwise the 2-cells in a certain cell division of $T^2=\partial M_\varphi$. This cell division is dual to the triangulation of ∂M_φ induced by the ideal triangulation of M_φ in step (1).

(3) Describe the incompressible surfaces in M_φ by describing their intersection with the spine T^2/τ or M_φ.

In this Appendix we shall sketch steps (1) and (2) for a general hyperbolic φ, following Jørgensen, and then fit the surfaces S_γ into the picture by describing their intersection with the spine T^2/τ of M_φ. Presumably, one could also prove the Theorem using this point of view.

Step (1). Let $\ldots, T_0, T_1, T_2, \ldots$ be the sequence of triangles meeting the line l in the Diagram of PSL(2, $\mathbb{Z}$), where $\varphi(T_i)=T_{i+m}$ for each i. Each T_i determines a "triangulation" K_i of $T^2=\mathbb{R}^2/\mathbb{Z}^2$ into two triangles, whose three edges have slopes given by the three vertices of T_i. Let $\tilde{K}_i$ be the lift of K_i to a triangulation of $\mathbb{R}^2$, with vertices $\mathbb{Z}^2$. $\tilde{K}_i$ is invariant under translations in $\mathbb{Z}^2$, and also under 180° rotations about points of $\mathbb{Z}^2$. Since T_i and T_{i+1} have two vertices in common, K_i and K_{i+1} have two edges in common. If we superimpose $\tilde{K}_{i+1}$ upon $\tilde{K}_i$, we see an array of tetrahedra, whose bottom faces are the triangles of $\tilde{K}_i$ and whose top faces are the triangles of $\tilde{K}_{i+1}$ (see Fig. 12). All these layers of tetrahedra can be stacked up in order, allowing some distortion of distances in the direction perpendicular

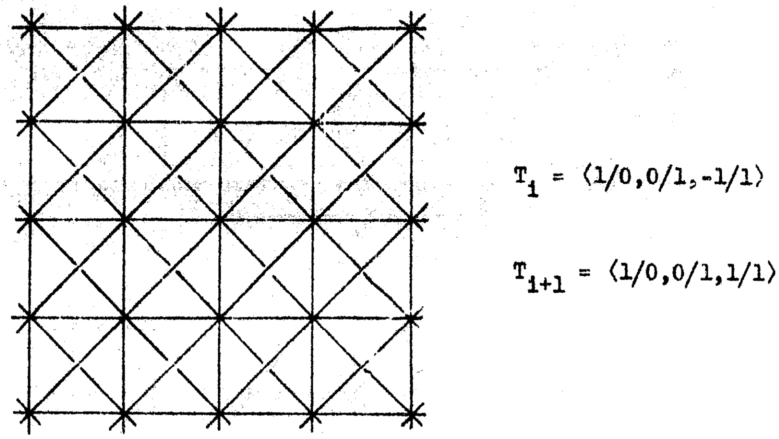


Fig. 12.

to the layers. Deleting the vertices ($\mathbb{Z}^2$), we have a decomposition $\tilde{\tilde{K}}$ of $(\mathbb{R}^2-\mathbb{Z}^2)\times\mathbb{R}$ into ideal tetrahedra because no vertex is common to all the triangles T_i, hence no edge is common to all the triangulations K_i. $\tilde{\tilde{K}}$ is invariant by the translations of $\mathbb{Z}^2$ in the first factor, so $\tilde{\tilde{K}}$ lifts a decomposition $\tilde{K}$ of $\tilde{M}_\varphi=(T^2-\{0\})\times\mathbb{R}$ into ideal tetrahedra. $\tilde{K}$ in turn is invariant by Φ (acting by φ in $T^2-\{0\}$ and by translation by m layers of tetrahedra in the $\mathbb{R}$ direction), so we obtain a decomposition K of $M_\varphi=\tilde{M}_\varphi/\Phi$ into ideal tetrahedra.

Step (2). If we truncate the tetrahedra of K by chopping off their corners in a uniform way, we obtain a decomposition of the compact manifold M_φ into truncated tetrahedra, with an induced triangulation L of the torus ∂M_φ. L lifts to triangulations $\tilde{L}$ of the infinite cyclic cover $\partial\tilde{M}_\varphi$ and $\tilde{\tilde{L}}$ of the universal cover $\partial M_\varphi=\mathbb{R}^2$. These triangulations can be described very neatly, as follows. Let $\mathcal{T}_j\subset\mathbb{R}^2$ be the horizontal strip $\mathbb{R}\times[j,j+1]$. We triangulate $\mathcal{T}_0$ by the triangles T_i, in the same way that the T_i's triangulate the strip $\bigcup_i T_i$ in the diagram of PSL(2, $\mathbb{Z}$), and we relable these triangles T_{i0}. This triangulation of $\mathcal{T}_0$ determines triangulations of all the other T_j's by successive reflections across the edges of these strips. The triangle of $\mathcal{T}_j$ corresponding to T_{i0} we label T_{ij}. This triangulation of $\mathbb{R}^2$ is periodic under the vertical translation $V(x,y)=(x,y+2)$ and the horizontal translation H which takes T_{ij} to $T_{i,j+m}$.

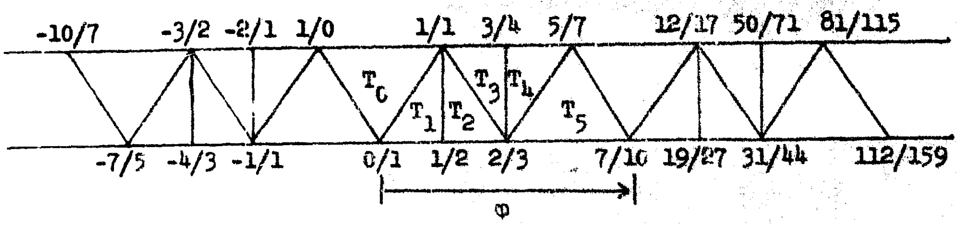
Example A1. Let

$$\varphi=\begin{bmatrix}10&7\\7&5\end{bmatrix}=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}1&0\\1&1\end{bmatrix}^2\begin{bmatrix}1&1\\0&1\end{bmatrix}^2\begin{bmatrix}1&0\\1&1\end{bmatrix}.$$

See Fig. 13.

Assertion. $\tilde{\tilde{L}}$ is the triangulation $\{T_{ij}\}$, $\tilde{L}$ is $\tilde{\tilde{L}}/V^2$, and

$$L=\begin{cases}\tilde{L}/H & \text{if } \varphi \text{ has positive eigenvalues,}\\ \tilde{L}/HV & \text{if } \varphi \text{ has negative eigenvalues.}\end{cases}$$



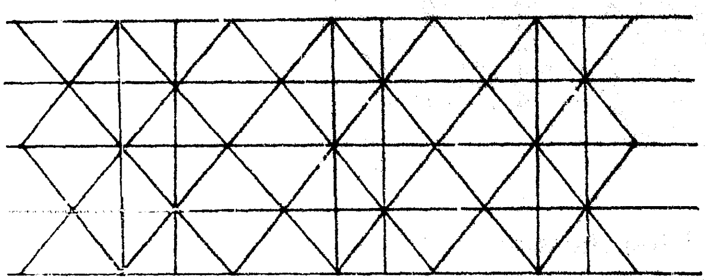


Fig. 13.

To see this, consider the ith layer of ideal tetrahedra in $(\mathbb{R}^2-\mathbb{Z}^2)\times\mathbb{R}$. Truncation yields four boundary triangles around each ideal vertex (see Fig. 14). The vertices of these boundary triangles correspond to edges of the ideal tetrahedra of $\tilde{K}$, so

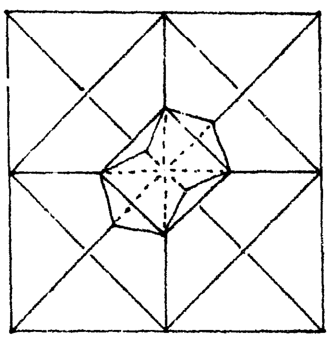


Fig. 14.

each vertex has associated to it a slope in $\mathbb{Q}\cup\{1/0\}$. The four triangles of this cycle have as their slopes the slopes of T_i and T_{i+1}, alternately. Lifting to the universal cover $\widetilde{\partial M_\varphi}$ we see an infinite cycle of alternating triangles (see Fig. 15). The lower faces of the given layer of tetrahedra meet the cycle of four triangles in a cycle $\tilde{C}_i\subset\partial\tilde{M}_\varphi$ of the six edges. The six vertices of $\tilde{C}_i$ are labelled by the slopes of the vertices of T_i, traversed twice around. Similarly, the upper faces of this layer of tetrahedra determine a cycle $\tilde{C}_{i+1}$ of six edges, so that $\tilde{C}_i\cup\tilde{C}_{i+1}$ is the boundary of the cycle of four triangles. The next cycle of four triangles, bounded by $\tilde{C}_{i+1}\cup\tilde{C}_{i+2}$, attaches to this cycle along the common edge-cycle $\tilde{C}_{i+1}$, in a way determined by which edge of T_{i+1} is also an edge of T_{i+2} (see Fig. 16). Lifting to the universal

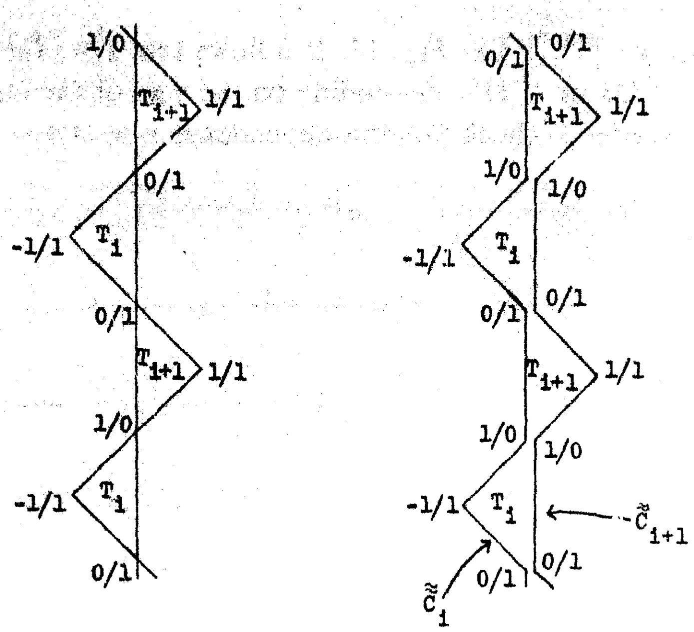


Fig. 15.

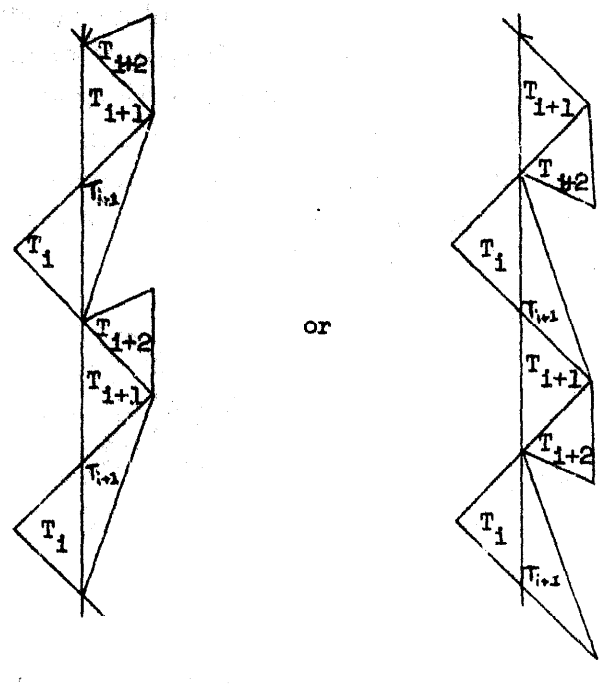
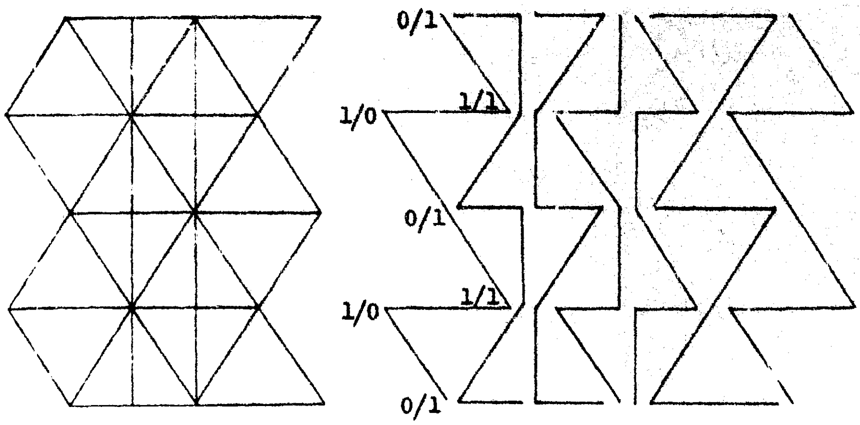


Fig. 16.

cover $\widetilde{\partial M_\varphi}$, we see the triangulation is determined by the infinite edge-cycles $\tilde{\tilde{C}}_i$ lifting $\tilde{C}_i$. The triangulation $\{T_{ij}\}$ also has this structure: its 1-skeleton is the union of (non-crossing) infinite vertical "sawtooth" cycles $\tilde{\tilde{C}}_i$, the successive vertices of $\tilde{\tilde{C}}_i$ having slopes the infinitely repeating 3-cycle of slopes of the vertices of T_i.

Example A2. Let $\varphi = \left[\begin{smallmatrix}10 & 7\\ 7 & 5\end{smallmatrix}\right]$. See Fig. 17. It follows that $\tilde{L} = \{T_{ij}\}$ and $\hat{L} = \tilde{L}/V^2$. Also, L is clearly $\hat{L}/H$ or $\hat{L}/HV$, depending on the sign of the eigenvalues of φ; we leave it to the reader to check that this dependence is as stated.



The triangulation L of ∂M_φ The cycles $C_0, \ldots, C_5$, $C_6 = C_0$

Fig. 17.

To obtain a spine for M_φ we can simultaneously collapse all the truncated tetrahedra to their "spines", where by the "spine" of a truncated tetrahedran we mean the subcomplex of the first barycentric subdivision of the tetrahedron, consisting of those simplices not containing any vertices of the original tetrahedron. The collapse of the truncated tetrahedron to its "spine" is by projection from the original vertices (see Fig. 18). Each triangle of L is thus subdivided into three quadrilaterals by adding the three edges joining its barycenter to the midpoints of its edges. These new edges form the dual cell division L^* of L.

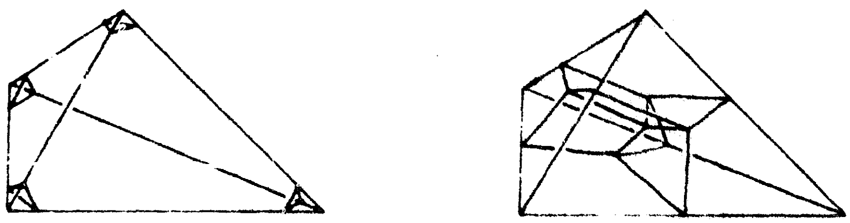


Fig. 18.

Example A3. Let $\varphi = \left[\begin{smallmatrix}10 & 7\\ 7 & 5\end{smallmatrix}\right]$. See Fig. 19. The spine of M_φ is formed from ∂M_φ by identifying these quadrilateral cells in pairs, the two cells of each pair being contained in the same tetrahedron. The careful reader can check that this identification is given by an involution τ on the 2-cells on L^* which sends a 2-cell e of L^* to $V(e)$ reflected across the diameter which joins opposite vertices of $V(e)$ and is as nearly horizontal as possible, of positive or negative slope in alternating horizontal rows of cells of L^*. (In Fig. 19 these diameters are indicated by dashed lines.) Note that the edge-cycles C_i are invariant by τ.

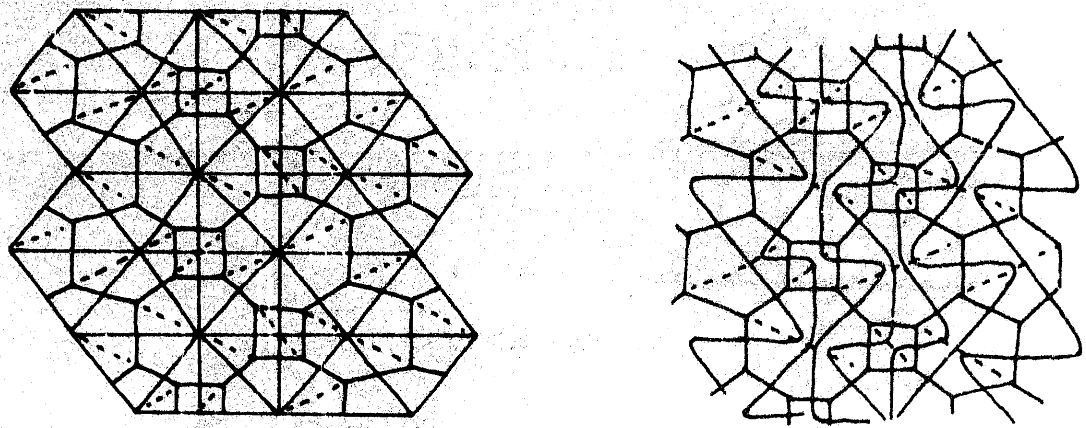


Fig. 19.

Step (3). We have expressed M_φ as the mapping cylinder of the projection $\partial M_\varphi \to \partial M_\varphi/\tau$. Lifting to the cover $\tilde{M}_\varphi = (T^2 \times \{0\}) \times \mathbb{R}$, this gives $\tilde{M}_\varphi$ as the mapping cylinder of the lifted projection $\partial\tilde{M}_\varphi \to \partial\tilde{M}_\varphi/\tilde{\tau}$. The incompressible surfaces $\tilde{S}_\gamma \subset \tilde{M}_\varphi$, as well as the fiber $T^2 - \{0\}$, have the form of restrictions of the mapping cylinder to certain 1-manifolds in $\partial\tilde{M}_\varphi$ invariant under $\tilde{\tau}$. For the fiber this is clear: it is the mapping cylinder on any of the edge-cycles $\tilde{C}_i$. For a minimal edge-path γ in one of the four horizontal infinite strips of the triangulation $\tilde{L}$, the surface $\tilde{S}_\gamma$ is the mapping cylinder on a curve system $\tilde{C}_\gamma \subset \partial\tilde{M}_\varphi$, obtained by the following two steps:

(i) Begin with the union of all the edge-cycles $\tilde{C}_{i_j}$ which contain the edges e_j of γ, perturbed slightly to be disjoint. This perturbation can be done equivariantly with respect to $\tilde{\tau}$, as in the right diagram of Fig. 19.

(ii) At each vertex v_j of γ and at the translate $V(v_j)$, switch the connections of the four incoming arcs of the edge-cycles in (i) (see Fig. 20).

This can also be done $\tilde{\tau}$-equivariantly.

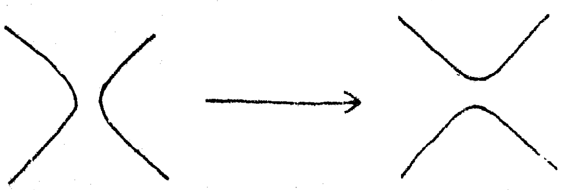


Fig. 20.

Example A4. Let $\varphi = [\begin{smallmatrix}10 & 7\\ 7 & 5\end{smallmatrix}]$. See Fig. 21. One can see that the mapping cylinder on $\tilde{C}_\gamma$ is $\tilde{S}_\gamma$ as follows. The edge-cycles $\tilde{C}_{i_j}$ in (i) have as their mapping cylinders certain fibers F_j of $\tilde{M}_\varphi$. The switching operations in (ii) at v_j and $V(v_j)$ effect surgeries connecting ∂F_j to ∂F_{j+1} twice. Forming the mapping cylinder, F_j and F_{j+1} are joined together by a sort of surgery operation, which is just the surgery shown in (ii) "cross I". The resulting surface intersects a fiber between F_j and F_{j+1} in two arcs, whose slope p_j/q_j is given by the label on the vertex v_j. At the next vertex v_{j+1} a similar thing is happening: F_{j+1} is "surgered" to connect it to F_{j+2}. The two surgeries on

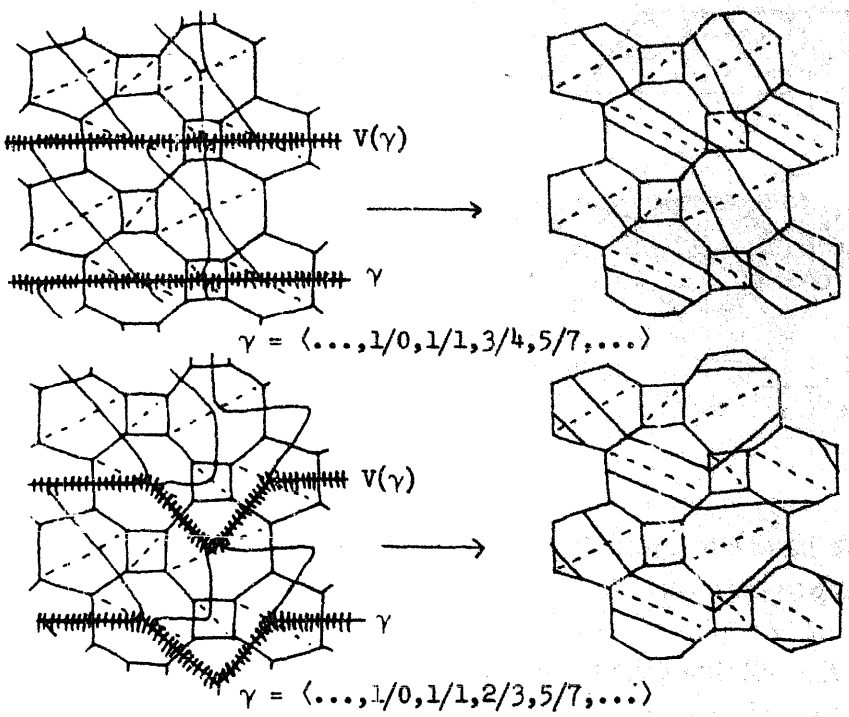


Fig. 21.

F_{j+1} connecting it to F_j and F_{j+2} are disjoint, and have the effect of making F_{j+1} into a saddle connecting two arcs of slope p_j/q_j in one fiber to two arcs of slope p_{j+1}/q_{j+1} in another fiber. This is just what $\tilde{S}_\gamma$ does.

The construction of $\tilde{C}_\gamma$ depends on an initial choice of one of the four horizontal strips of $\tilde{L}$ to contain γ. Choosing instead an adjacent strip may yield a curve system $\tilde{C}_\gamma$ intersecting $\tilde{L}^*$ in a different pattern. However, the associated mapping cylinder surfaces are isotopic, and in fact the curve systems $\tilde{C}_\gamma$ are $\tilde{\tau}$-equivariantly isotopic.

References

[1] J. Cerf, La stratification naturelle . . . , Publ. Math. I.H.E.S. 39 (1970) 5–173.

[2] M. Culler, W. Jaco and H. Rubinstein, Incompressible surfaces in once-punctured torus bundles, preprint.

[3] A. Hatcher and W. Thurston, Incompressible surfaces in 2-bridge knot complements, preprint.

[4] T. Jørgensen, On pairs of punctured tori, to appear.

[5] W. Thurston, Geometry and topology of 3-manifolds, xeroxed notes.

[6] W. Thurston, Hyperbolic structures on 3-manifolds II: surface groups and 3-manifolds which fiber over the circle, preprint.

[7] F. Waldhausen, Eine Klasse von 3-dimensionalen Mannigfaltigkeiten I, Invent. Math. 3 (1967) 308–333.

[8] F. Waldhausen, On irreducible 3-manifolds which are sufficiently large, Annals of Math. 87 (1968) 55–88.