William P. Thurston

# The Geometry and Topology of Three-Manifolds

Electronic version 1.1 - March 2002

http://www.msri.org/publications/books/gt3m/

This is an electronic edition of the 1980 notes distributed by Princeton University. The text was typed in TEX by Sheila Newbery, who also scanned the figures. Typos have been corrected (and probably others introduced), but otherwise no attempt has been made to update the contents. Genevieve Walsh compiled the index.

Numbers on the right margin correspond to the original edition's page numbers.

Thurston's *Three-Dimensional Geometry and Topology*, Vol. 1 (Princeton University Press, 1997) is a considerable expansion of the first few chapters of these notes. Later chapters have not yet appeared in book form.

Please send corrections to Silvio Levy at levy@msri.org.

CHAPTER 4

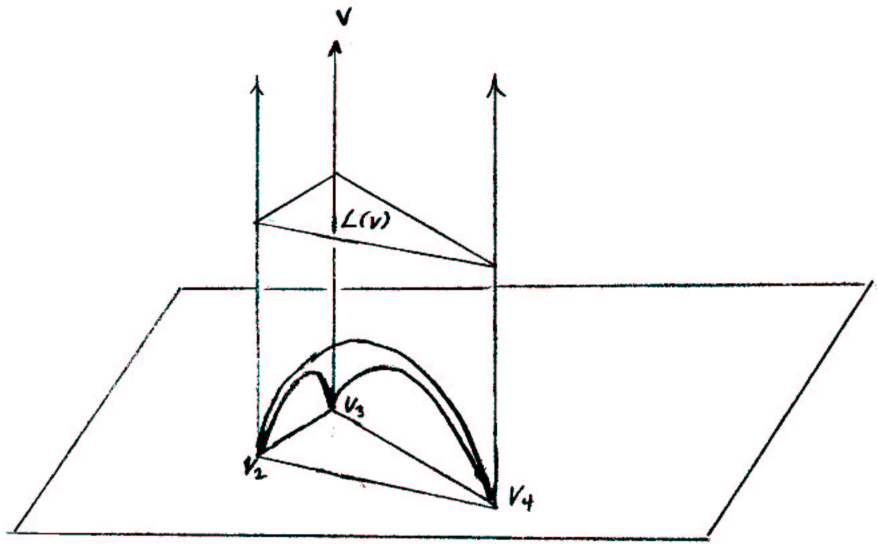
# Hyperbolic Dehn surgery

A hyperbolic structure for the complement of the figure-eight knot was constructed in 3.1. This structure was in fact chosen to be complete. The reader may wish to verify this by constructing a horospherical realization of the torus which is the link of the ideal vertex. Similarly, the hyperbolic structure for the Whitehead link complement and the Borromean rings complement constructed in 3.3 and 3.4 are complete.

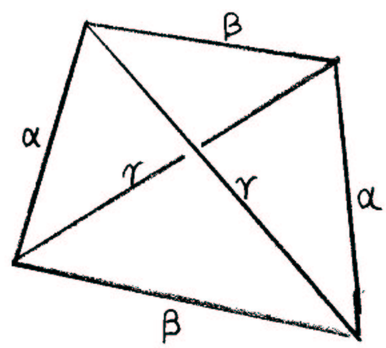
There is actually a good deal of freedom in the construction of hyperbolic structures for such manifolds, although most of the resulting structures are not complete. We shall first analyze the figure-eight knot complement. To do this, we need an understanding of the possible shapes of ideal tetrahedra.

## 4.1. Ideal tetrahedra in $H^3$.

The link $L(v)$ of an ideal vertex $v$ of an oriented ideal tetrahedron $T$ (which by definition is the set of rays in the tetrahedron through that vertex) is a Euclidean triangle, well-defined up to orientation-preserving similarity. It is concretely realized as the intersection with $T$ of a horosphere about $v$. The triangle $L(v)$ actually determines $T$ up to congruence. To see this, picture $T$ in the upper half-space model and arrange it so that $v$ is the point at infinity. The other three vertices of $T$ form a triangle in $E^2$ which is in the same similarity class as $L(v)$. Consequently, if two 4.2
tetrahedra $T$ and $T'$ have vertices $v$ and $v'$ with $L(v)$ similar to $L(v')$, then $T'$ can be transformed to $T$ by a Euclidean similarity which preserves the plane bounding upper half-space. Such a similarity is a hyperbolic isometry.

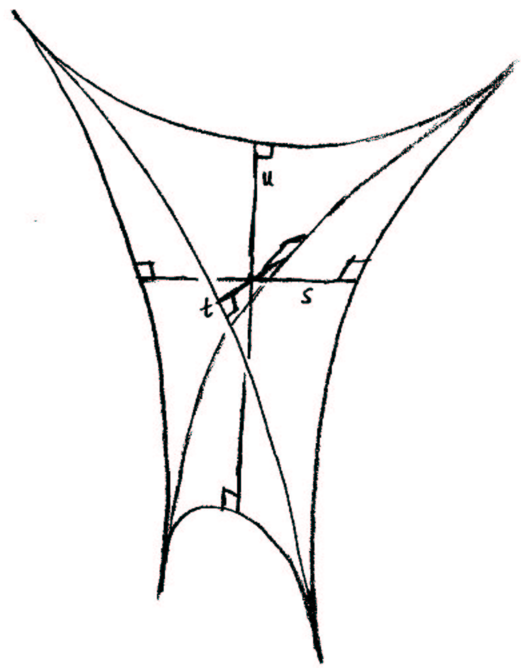



It follows that $T$ is determined by the three dihedral angles $\alpha, \beta$ and $\gamma$ of edges incident to the ideal vertex $v$, and that $\alpha + \beta + \gamma = \pi$. Using similar relations among angles coming from the other three vertices, we can determine the other three dihedral angles:




4.3

Thus, dihedral angles of opposite edges are equal, and the oriented similarity class of $L(v)$ does not depend on the choice of a vertex $v$! A geometric explanation of this phenomenon can be given as follows. Any two non-intersecting and non-parallel lines in $H^3$ admit a unique common perpendicular. Construct the three common perpendiculars $s, t$ and $u$ to pairs of opposite edges of $T$. Rotation of $\pi$ about $s$, for instance, preserves the edges orthogonal to $s$, hence preserves the four ideal vertices of $T$, so it preserves the entire figure. It follows that $s$, $t$ and $u$ meet in a point and that they are pairwise orthogonal. The rotations of $\pi$ about these three axes are the three non-trivial elements of $z_2 \oplus z_2$ acting as a group of symmetries of $T$.

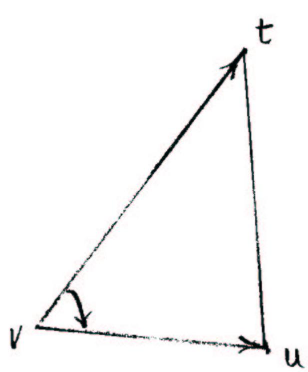


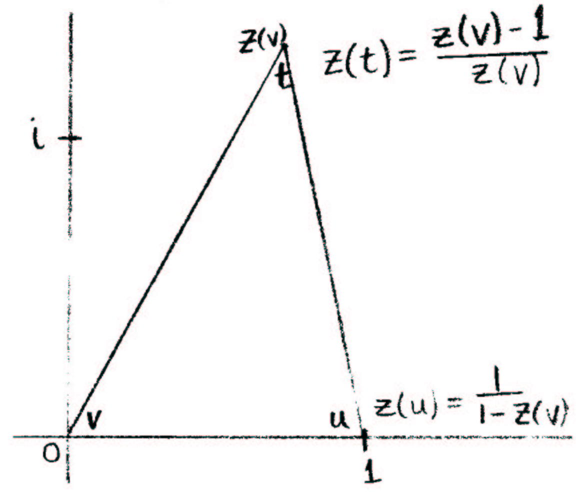

4.4

In order to parametrize Euclidean triangles up to similarity, it is convenient to regard $E^2$ as $\mathbb{C}$. To each vertex $v$ of a triangle $\Delta(t,u,v)$ we associate the ratio

$$\frac{(t-v)}{(u-v)} = z(v)$$

of the sides adjacent to $v$. The vertices must be labelled in a clockwise order, so that



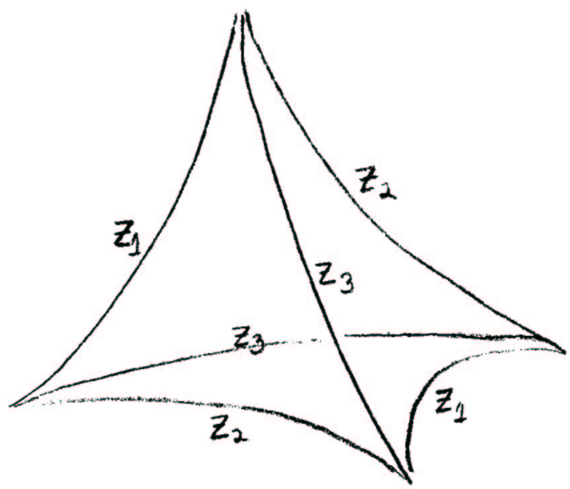
$\mathrm{Im}\big(z(v)\big) > 0$. Alternatively, arrange the triangle by a similarity so that $v$ is at 0 and $u$ at 1; then $t$ is at $z(v)$. The other two vertices have invariants

4.1.1.
$$\begin{aligned} z(t) &= \tfrac{z(v)-1}{z(v)} \\ z(u) &= \tfrac{1}{1-z(v)}. \end{aligned}$$

Denoting the three invariants $z_1, z_2, z_3$ in clockwise order, with any starting point, we have the identities 4.5

4.1.2.
$$\begin{gathered} z_1\, z_2\, z_3 = -1 \\ 1 - z_1 + z_1 z_2 = 0 \end{gathered}$$

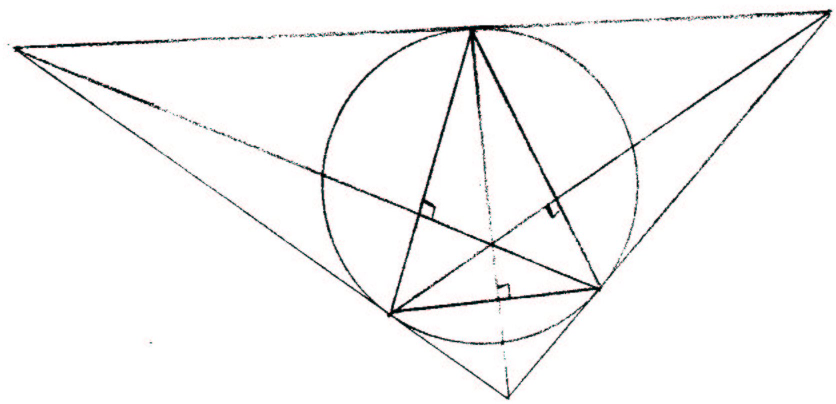
We can now translate this to a parametrization of ideal tetrahedra. Each edge $e$ is labelled with a complex number $z(e)$, opposite edges have the same label, and the three distinct invariants satisfy 4.1.2 (provided the ordering is correct.) Any $z_i$ determines the other two, via 4.1.2.



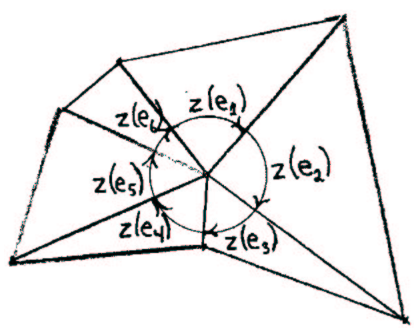

## 4.2. Gluing consistency conditions.

Suppose that $M$ is a three-manifold obtained by gluing tetrahedra $T_i, \ldots, T_j$ and then deleting the vertices, and let $K$ be the complex which includes the vertices. 4.6

Any realization of $T_1, \ldots, T_j$ as ideal hyperbolic tetrahedra determines a hyperbolic structure on $(M - (1 - \text{skeleton}))$, since any two ideal triangles are congruent. Such a congruence is uniquely determined by the correspondence between the vertices. (This fact may be visualized concretely from the subdivision of an ideal triangle by its altitudes.)



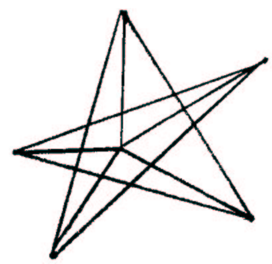
The condition for the hyperbolic structure on $(M - (1 - \text{skeleton}))$ to give a hyperbolic structure on $M$ itself is that its developing map, in a neighborhood of each edge, should come from a local homeomorphism of $M$ itself. In particular, the sum of the dihedral angles of the edges $e_1, \dots, e_k$ must be $2\pi$. Even when this condition is satisfied, though, the holonomy going around an edge of $M$ might be a non-trivial translation along the edge. To pin down the precise condition, note that for each ideal vertex $v$, the hyperbolic structure on $M - (1 - \text{skeleton})$ gives a similarity structure to $L(v) - (0 - \text{skeleton})$. The hyperbolic structure extends over an edge $e$ of $M$ if and 4.7
only if the similarity structure extends over the corresponding point in $L(v)$, where $v$ is an endpoint of $e$. Equivalently, the similarity classes of triangles determined by $z(e_1), \dots, z(e_k)$ must have representatives which can be arranged neatly around a point in the plane:




The algebraic condition is

4.2.1. $$z(e_1) \cdot z(e_2) \cdot \dots \cdot z(e_k) = 1.$$

This equation should actually be interpreted to be an equation in the universal cover $\tilde{\mathbb{C}}^*$, so that solutions such as 4.8



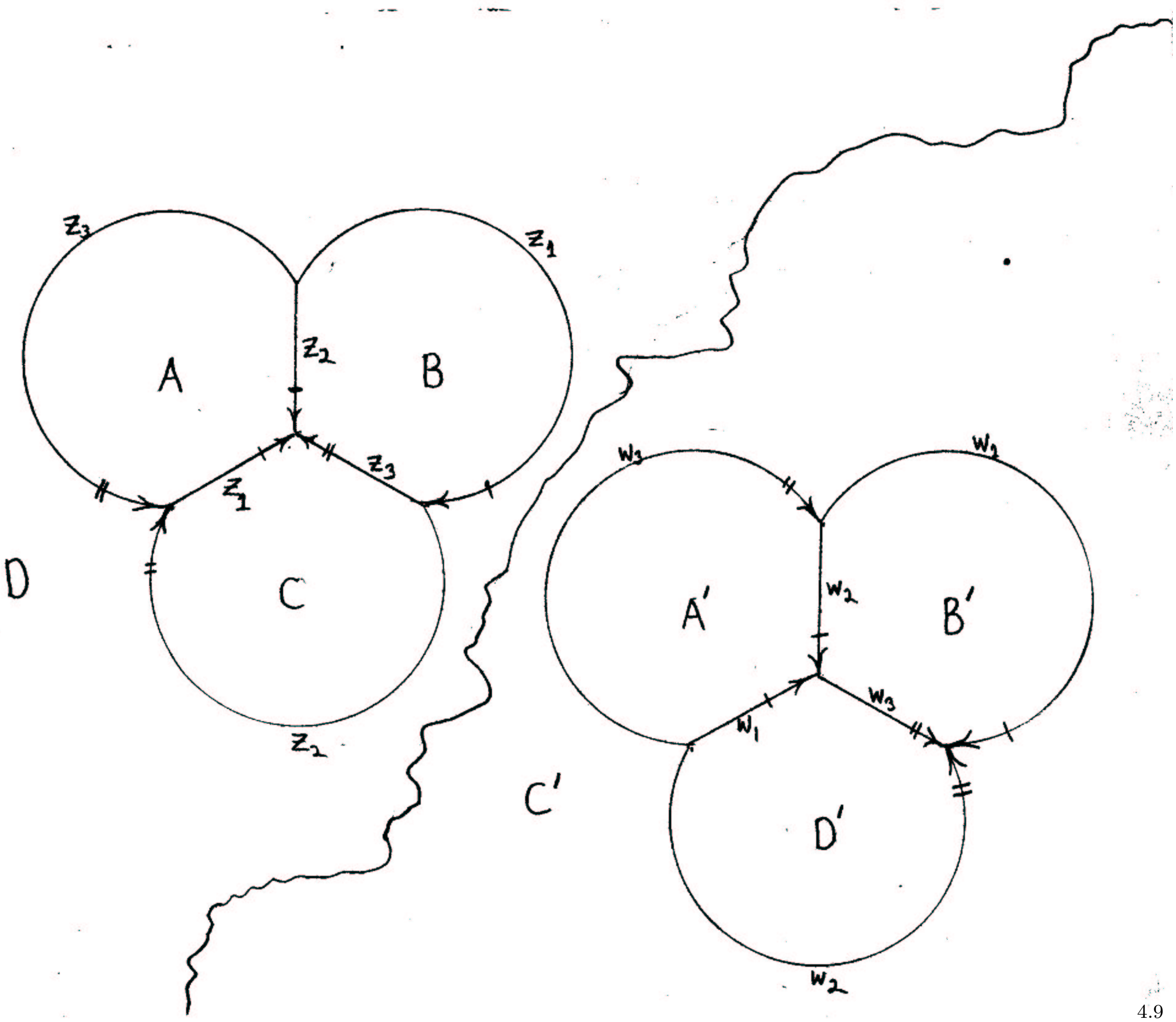
are ruled out. In other words, the auxiliary condition

4.2.2. $$\arg z_1 + \cdots + \arg z_k = 2\pi$$

must also be satisfied, where $0 < \arg z_i \leq \pi$.

## 4.3. Hyperbolic structure on the figure-eight knot complement.

Consider two hyperbolic tetrahedra to be identified to give the figure eight knot complement:




4.9

We read off the gluing consistency conditions for the two edges:

$$(\not\longrightarrow) z_1^2 z_2 w_1^2 w_2 = 1 \qquad (\not\!\not\longrightarrow) z_3^2 z_2 w_3^2 w_2 = 1.$$

From 4.1.2, note that the product of these two equations,

$$(z_1 z_2 z_3)^2 (w_1 w_2 w_3)^2 = 1$$

is automatically satisfied. Writing $z = z_1$, and $w = w_1$, and substituting the expressions from 4.1.1 into $(\not\longrightarrow)$, we obtain the equivalent gluing condition,

4.3.1. $$z(z-1)\, w(w-1) = 1.$$

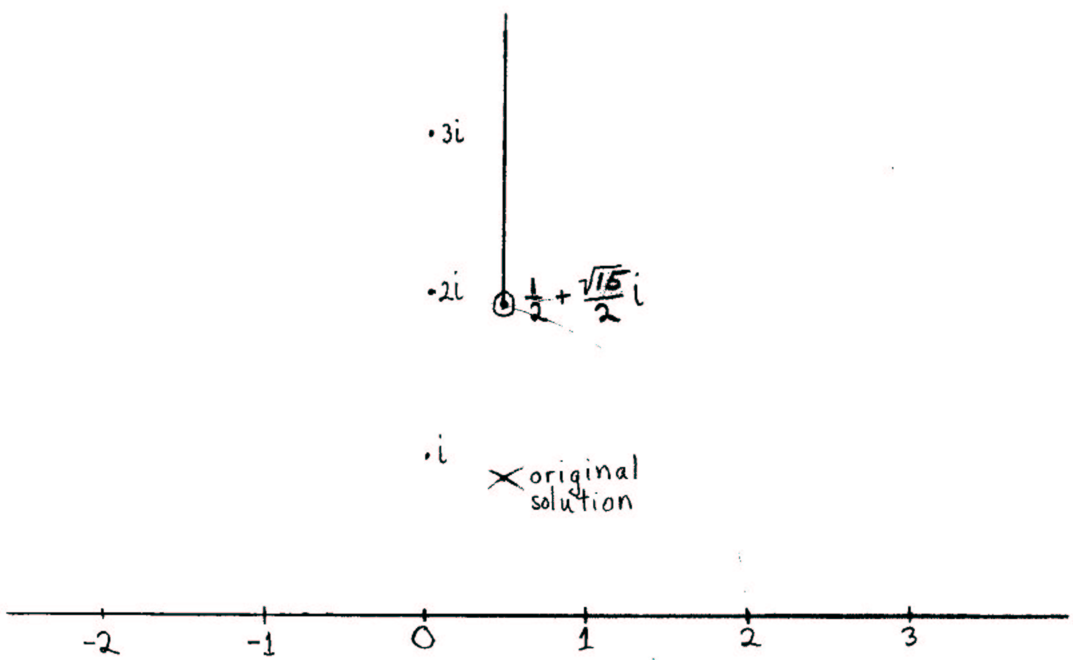
We may solve for $z$ in terms of $w$ by using the quadratic formula.

4.3.2. $$z = \frac{1 \pm \sqrt{1 + 4/w(w-1)}}{2}$$

We are searching only for geometric solutions

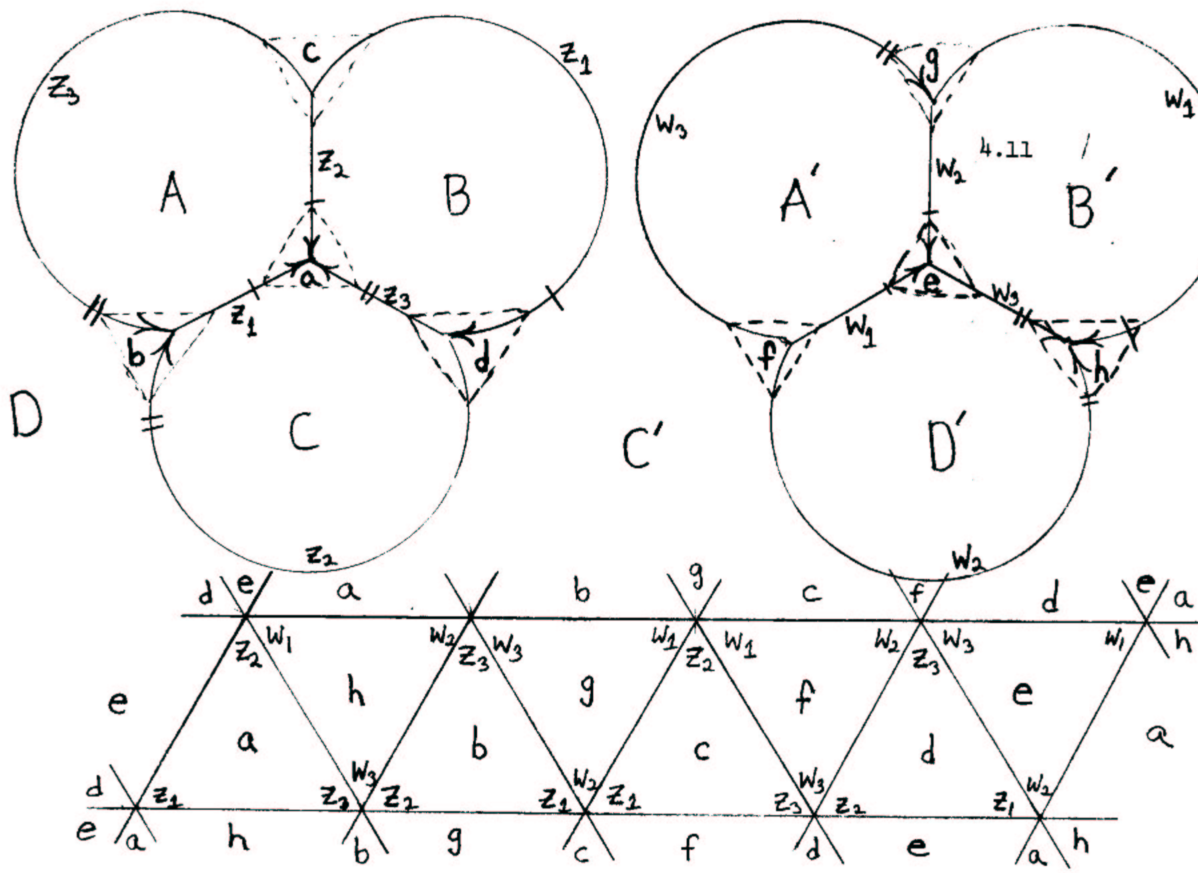
$$\mathrm{Im}(z) > 0 \quad \mathrm{Im}(w) > 0$$

so that the two tetrahedra are non-degenerate and positively oriented. For each value of $w$, there is at most one solution for $z$ with $\mathrm{Im}(z) > 0$. Such a solution exists provided that the discriminant $1 + 4/w(w-1)$ is not positive real. Solutions are therefore parametrized by $w$ in this region of $\mathbb{C}$: 4.10

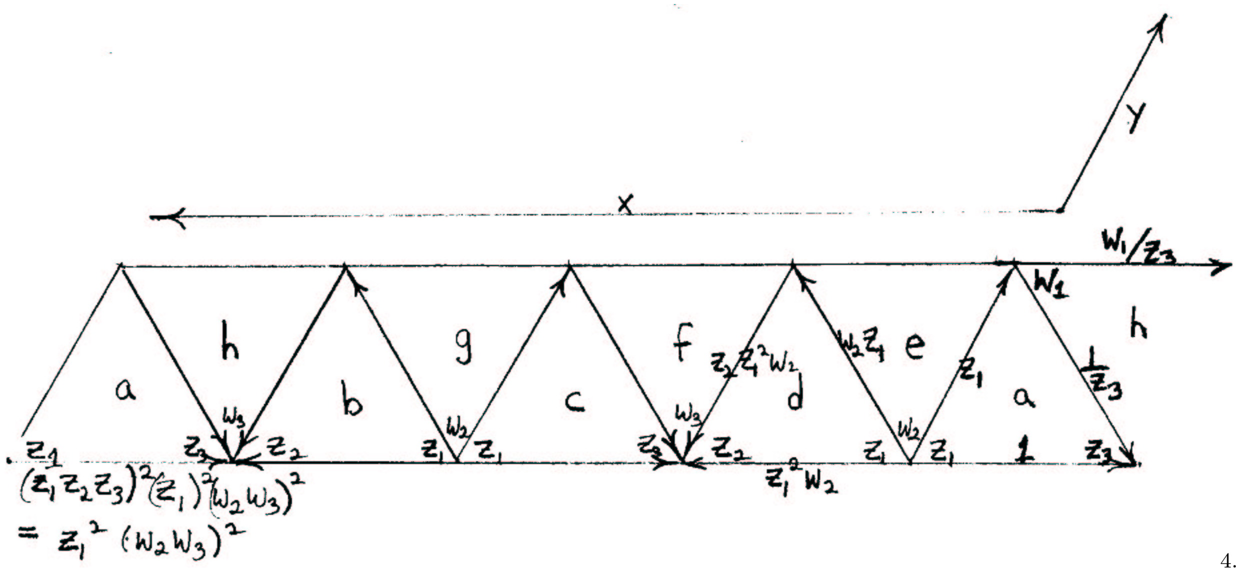



Note that the original solution given in 3.1 corresponds to $w = z = \sqrt[3]{-1} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$.

The link $L$ of the single ideal vertex has a triangulation which can be calculated from the gluing diagram: 4.11




Now let us compute the derivative of the holonomy of the similarity structure on $L$. To do this, regard directed edges of the triangulation as vectors. The ratio of any two vectors in the same triangle is known in terms of $z$ or $w$. Multiplying appropriate ratios, we obtain the derivative of the holonomy:



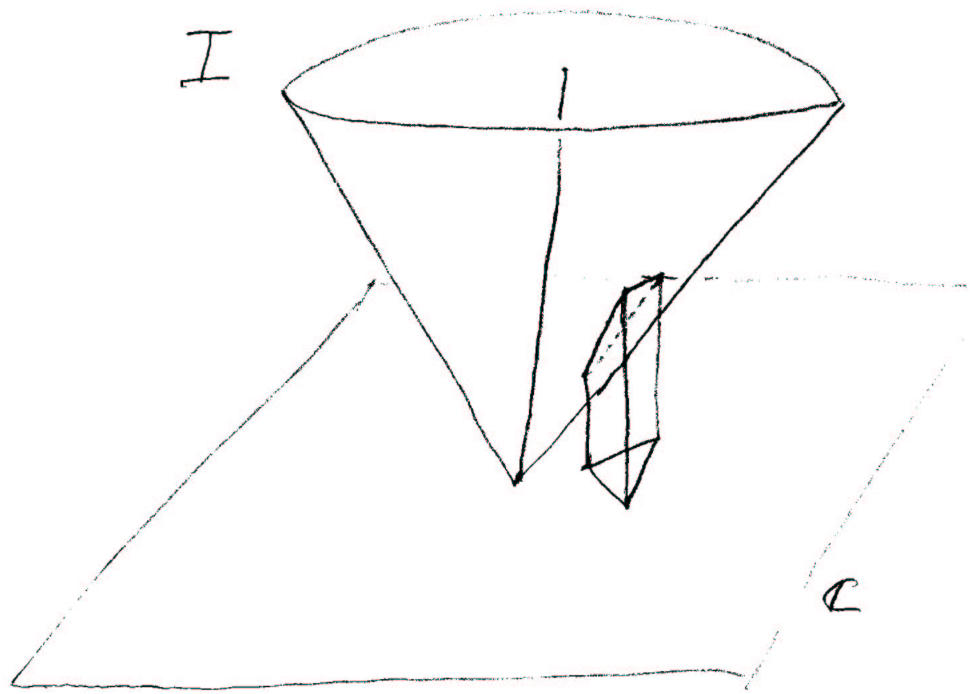

4.12

$$\begin{aligned} H'(x) &= z_1^2(w_2w_3)^2 = (\tfrac{z}{w})^2 \\ H'(y) &= \tfrac{w_1}{z_3} = w(1-z). \end{aligned}$$

Observe that if $M$ is to be complete, then $H'(x) = H'(y) = 1$, so $z = w$. From 4.3.1, $(z(z-1))^2 = 1$. Since $z(z-1) < 0$, this means $z(z-1) = -1$, so that the only possibility is the original solution $w = z = \sqrt[3]{-1}$.

## 4.4. The completion of hyperbolic three-manifolds obtained from ideal polyhedra.

Let $M$ be any hyperbolic manifold obtained by gluing polyhedra with some vertices at infinity, and let $K$ be the complex obtained by including the ideal vertices. The completion $\bar{M}$ is obtained by completing a deleted neighborhood $\mathcal{N}(v)$ of each ideal vertex $v$ in $k$, and gluing these completed neighborhoods $\bar{\mathcal{N}}(v)$ to $M$. The developing map for the hyperbolic structure on $\mathcal{N}(v)$ may be readily understood in terms of the developing map for the similarity structure on $L(v)$. To do this, choose coordinates so that $v$ is the point at infinity in the upper half-space model. The developing images of corners of polyhedra near $v$ are "chimneys" above some polygon in the developing image of $L(v)$ on $\mathbb{C}$ (where $\mathbb{C}$ is regarded as the boundary of upper half-space.) If $M$ is not complete near $v$, we change coordinates if necessary by a translation of $\mathbb{R}^3$ so that the developing image of $L(v)$ is $\mathbb{C} - 0$. The holonomy for $\mathcal{N}(v)$ now consists of similarities of $\mathbb{R}^3$ which leave invariant the $z$-axis and the $x - y$ plane ($\mathbb{C}$). Replacing $\mathcal{N}(v)$ by a smaller neighborhood, we may assume that the developing image $I$ of $\mathcal{N}(v)$ is a solid cone, minus the $z$-axis. 4.13




The completion of $I$ is clearly the solid cone, obtained by adjoining the $z$-axis to $I$. It follows easily that the completion of

$$\widetilde{\mathcal{N}(v)} = \tilde{I}$$

is also obtained by adjoining a single copy of the $z$-axis.

The projection

$$p : \widetilde{\mathcal{N}(v)} \to \mathcal{N}(v)$$

extends to a surjective map $\bar{p}$ between the completions. [$\bar{p}$ exists because $p$ does not increase distances. $\bar{p}$ is surjective because a Cauchy sequence can be replaced by a Cauchy path, which lifts to $\mathcal{N}\tilde{(v)}$.] Every orbit of the holonomy action of $\pi_1\big(\mathcal{N}(v)\big)$ on the $z$-axis is identified to a single point. This action is given by

$$H(\alpha) : (0,0,z) \mapsto |H'(\alpha)| \cdot (0,0,z)$$

where the first $H(\alpha)$ is the hyperbolic holonomy and the second is the holonomy of 4.14
$L(v)$. There are two cases:

*Case* 1. The group of moduli $\{|H'(\alpha)|\}$ is dense in $\mathbb{R}_+$. Then the completion of $\mathcal{N}(v)$ is the one-point compactification.

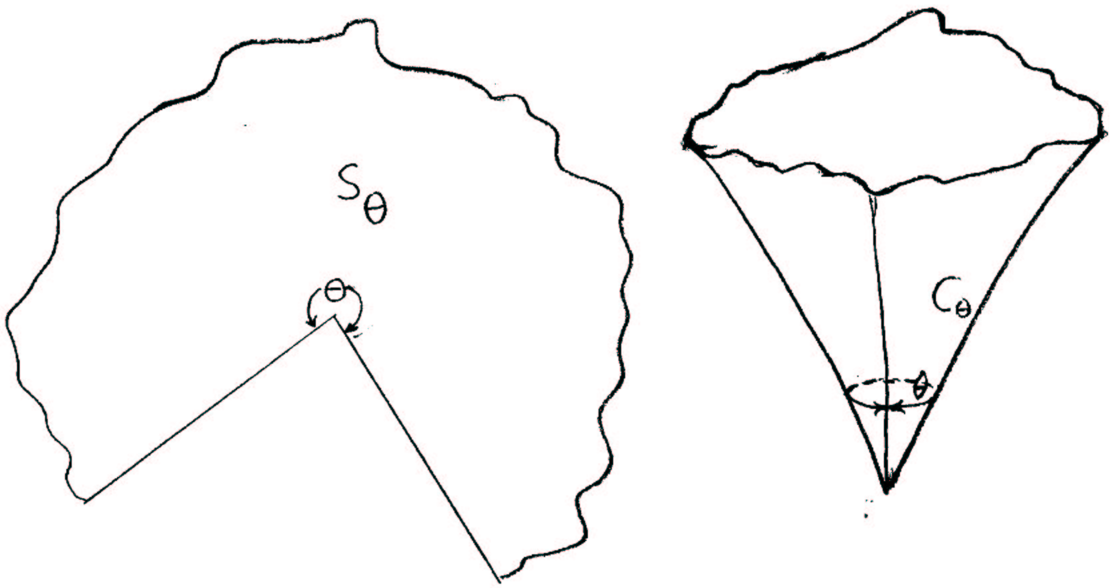
*Case* 2. The group of moduli $\{|H'(\alpha)|\}$ is a cyclic group. Then the completion

$$\overline{\mathcal{N}(v)}$$

is *topologically* a manifold which is the quotient space $\widetilde{(\overline{\mathcal{N}})}/H$, and it is obtained by adjoining a circle to $\mathcal{N}(v)$. Let $\alpha_1 \in \pi_1(L(v))$ be a generator for the kernel of $\alpha \mapsto |H'(\alpha)|$ and let $1 < |H'(\alpha_2)|$ generate the image, so that $\alpha_1$ and $\alpha_2$ generate $\pi_1(L(v)) = \mathbb{Z} \oplus \mathbb{Z}$. Then the added circle in

$$\overline{\mathcal{N}(v)}$$

has length $\log |H'(\alpha_2)|$. A cross-section of $\bar{\mathcal{N}}(v)$ perpendicular to the added circle is a cone $C_\theta$, obtained by taking a two-dimensional hyperbolic sector $S_\theta$ of angle $\theta$, $[0 < \theta < \infty]$ and identifying the two bounding rays:




4.15

It is easy to make sense of this even when $\theta > 2\pi$. The cone angle $\theta$ is the argument of the element $\tilde{H}'(\alpha_2) \in \tilde{\mathbb{C}}^*$. In the special case $\theta = 2\pi$, $C_\theta$ is non-singular, so

$$\overline{\mathcal{N}(v)}$$

is a hyperbolic manifold. $\overline{\mathcal{N}(v)}$ may be seen directly in this special case, as the solid cone $I \cup (z - \text{axis})$ modulo $H$.

## 4.5. The generalized Dehn surgery invariant.

Consider any three-manifold $M$ which is the interior of a compact manifold $\hat{M}$ whose boundary components $P_1, \ldots, P_k$ are tori. For each $i$, choose generators $a_i, b_i$ for $\pi_1(P_i)$. If $M$ is identified with the complement of an open tubular neighborhood of a link $L$ in $S^3$, there is a standard way to do this, so that $a_i$ is a *meridian* (it bounds a disk in the solid torus around the corresponding component of $L$) and $b_i$ is a *longitude* (it is homologous to zero in the complement of this solid torus in $S^3$). In this case we will call the generators $m_i$ and $l_i$.

We will use the notation $M_{(\alpha_1,\beta_1),\dots,(\alpha_k,\beta_k)}$ to denote the manifold obtained by gluing solid tori to $M$ so that a meridian in the $i$-th solid torus goes to $\alpha_i, a_i + \beta_i b_i$. If an ordered pair $(\alpha_i, \beta_i)$ is replaced by the symbol $\infty$, this means that nothing is glued in near the $i$-th torus. Thus, $M = M_{\infty,\dots,\infty}$.

These notions can be refined and extended in the case $M$ has a hyperbolic structure whose completion $\bar{M}$ is of the type described in 4.4. (In other words, if $M$ is not complete near $P_i$, the developing map for some deleted neighborhood $\mathcal{N}_i$ of $P_i$ should be a covering of the deleted neighborhood $I$ of radius $r$ about a line in $H^3$.) The developing map $D$ of $\mathcal{N}_i$ can be lifted to $\tilde{I}$, with holonomy $\tilde{H}$. The group of isometries of $\tilde{I}$ is $\mathbb{R} \oplus \mathbb{R}$, parametrized by (translation distance, angle of rotation); this parametrization is well-defined up to sign. 4.16

Definition 4.5.1. The generalized Dehn surgery invariants $(\alpha_i, \beta_i)$ for $\bar{M}$ are solutions to the equations

$$\alpha_i \tilde{H}(a_i) + \beta_i \tilde{H}(b_i) = (\text{rotation by } \pm 2\pi),$$

(or, $(\alpha_i, \beta_i) = \infty$ if $M$ is complete near $P_i$).

Note that $(\alpha_i, \beta_i)$ is unique, up to multiplication by $-1$, since when $M$ is not complete near $P_i$, the holonomy $\tilde{H} : \pi_1(\mathcal{N}_i) \to \mathbb{R} \oplus \mathbb{R}$ is injective. We will say that $\bar{M}$ is a hyperbolic structure for

$$M_{(\alpha_1,\beta_1),\dots,(\alpha_k,\beta_k)}.$$

If all $(\alpha_i, \beta_i)$ happen to be primitive elements of $\mathbb{Z} \oplus \mathbb{Z}$, then $\bar{M}$ is the topological manifold $M_{(\alpha_1,\beta_1),\dots,(\alpha_k,\beta_k)}$ with a non-singular hyperbolic structure, so that our extended definition is compatible with the original. If each ratio $\alpha_i/\beta_i$ is the rational number $p_i/q_i$ in lowest terms, then $\bar{M}$ is topologically the manifold $M_{(p_1,q_1),\dots,(p_k,q_k)}$. The hyperbolic structure, however, has singularities at the component circles of $\bar{M} - M$ with cone angles of $2\pi(p_i/\alpha_i)$ [since the holonomy $\tilde{H}$ of the primitive element $p_i a_i + q_i b_i$ in $\pi_1(P_i)$ is a pure rotation of angle $2\pi(p_i/\alpha_i)$]. 4.17

There is also a topological interpretation in case the $(\alpha_i, \beta_i) \in \mathbb{Z} \oplus \mathbb{Z}$, although they may not be primitive. In this case, all the cone angles are of the form $2\pi/n_i$, where each $n_i$ is an integer. Any branched cover of $\bar{M}$ which has branching index $n_i$ around the $i$-th circle of $\bar{M} - M$ has a non-singular hyperbolic structure induced from $\bar{M}$.

### 4.6. Dehn surgery on the figure-eight knot.

For each value of $w$ in the region $R$ of $\mathbb{C}$ shown on p. 4.10, the associated hyperbolic structure on $S^3 - K$, where $K$ is the figure-eight knot, has some Dehn surgery invariant $d(w) = \pm(\mu(w), \lambda(w))$. The function $d$ is a continuous map from $R$ to the one-point compactification $\mathbb{R}^2/\pm 1$ of $\mathbb{R}^2$ with vectors $v$ identified to $-v$. Every primitive element $(p, q)$ of $\mathbb{Z} \oplus \mathbb{Z}$ which lies in the image $d(R)$ describes a closed manifold $(S^3 - K)_{(p,q)}$ which possesses a hyperbolic structure.

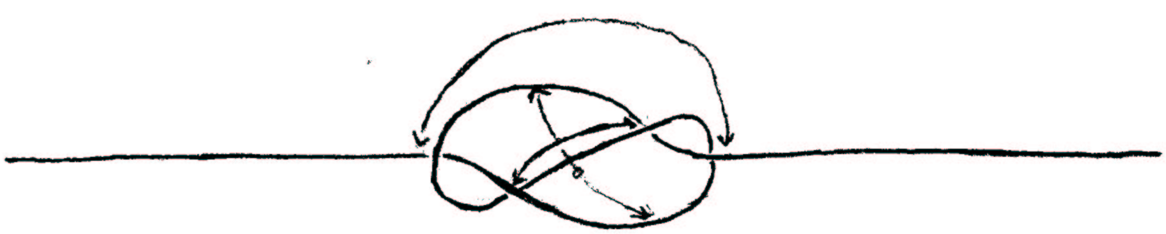
Actually, the map $d$ can be lifted to a map $\tilde{d} : R \to \hat{\mathbb{R}}^2$, by using the fact that the sign of a rotation of

$$(H^3 \widetilde{- z\text{-axis}})$$

is well-defined. (See §4.4. The extra information actually comes from the orientation of the $z$-axis determined by the direction in which the corners of tetrahedra wrap around it). $\tilde{d}$ is defined by the equation $\tilde{d}(w) = (\mu, \lambda)$ where 4.18

$$\mu \tilde{H}(m) + \lambda \tilde{H}^2(l) = (\text{a rotation by } +2\pi)$$

In order to compute the image $\tilde{d}(R)$, we need first to express the generators $l$ and $m$ for $\pi_1(P)$ in terms of the previous generators $x$ and $y$ on p. 4.11. Referring to page 6, we see that a meridian which only intersects two two-cells can be constructed in a small neighborhood of $K$. The only generator of $\pi_1(L(v))$ (see p. 4.11) which intersects only two one-cells is $\pm y$, so we may choose $m = y$. Here is a cheap way to see what $l$ is. The figure-eight knot can be arranged (passing through the point at infinity) so that it is invariant by the map $v \mapsto -v$ of $\hat{\mathbb{R}}^3 = S^3$.



This map can be made an isometry of the complete hyperbolic structure constructed for $S^3 - K$. (This can be seen directly; it also follows immediately from Mostow's Theorem, ...). This hyperbolic isometry induces an isometry of the Euclidean structure on $L(v)$ which takes $m$ to $m$ and $l$ to $-l$. Hence, a geodesic representing $l$ must be orthogonal to a geodesic representing $m$, so from the diagram on the bottom of 4.19
p. 4.11 we deduce that the curve $l = +x + 2y$ is a longitude. (Alternatively, it is not hard to compute $m$ and $l$ directly).

From p. 4.12, we have

4.6.1.
$$\begin{aligned} H(m) &= w(1-z) \\ H(l) &= z^2(1-z)^2 \end{aligned}$$

The behavior of the map $\tilde{d}$ near the boundary of $R$ is not hard to determine. For instance, when $w$ is near the ray $\mathrm{Im}(w) = 0$, $\mathrm{Re}(w) > 1$, then $z$ is near the ray $\mathrm{Im}(z) = 0$, $\mathrm{Re}(z) < 0$. The arguments of $\tilde{H}(m)$ and $\tilde{H}(l)$ are easily computed by analytic continuation from the complete case $w = z = \sqrt[3]{-1}$ (when the arguments are 0) to be

$$\arg \tilde{H}(m) = 0 \quad \arg \tilde{H}(l) \approx \pm 2\pi.$$

Consequently, $(\mu, \lambda)$ is near the line $\lambda = +1$. As $w \to 1$ we see from the equation

$$z(1-z)\, w(1-w) = 1$$

that

$$|z|^2 \cdot |w| \to 1$$

so $(\mu, \lambda)$ must approach the line $\mu + 4\lambda = 0$. Similarly, as $w \to +\infty$, then $|z|\,|w|^2 \to 1$, so $(\mu, \lambda)$ must approach the line $\mu - 4\lambda = 0$. Then the map $\tilde{d}$ extends continuously to send the line segment

$$\overline{1, +\infty}$$

to the line segment 4.20

$$\overline{(-4, +1), (+4, +1)}.$$

There is an involution $\tau$ of the region $R$ obtained by interchanging the solutions $z$ and $w$ of the equation $z(l-z)\, w(l-w) = 1$. Note that this involution takes $H(m)$ to $1/H(m) = H(-m)$ and $H(l)$ to $H(-l)$. Therefore $\tilde{d}(\tau w) = -\tilde{d}(w)$. It follows that $\tilde{d}$ extends continuously to send the line segment

$$\overline{-\infty, 0}$$

to the line segment

$$\overline{(+4, -1), (-4, -1)}.$$

When $|w|$ is large and $0 < \arg(w) < \pi/2$, then $|z|$ is small and

$$\arg(z) \approx \pi - 2\arg(w).$$

Thus $\arg \tilde{H}(m) \approx \arg w$, $\arg \tilde{H}(l) \approx 2\pi - 4 \arg w$ so $\mu \arg w + \lambda(2\pi - 4\arg w) = 2\pi$. By considering $|H(m)|$ and $|H(l)|$, we have also $\mu - 4\lambda \approx 0$, so $(\mu, \lambda) \approx (4, 1)$.

There is another involution $\sigma$ of $R$ which takes $w$ to

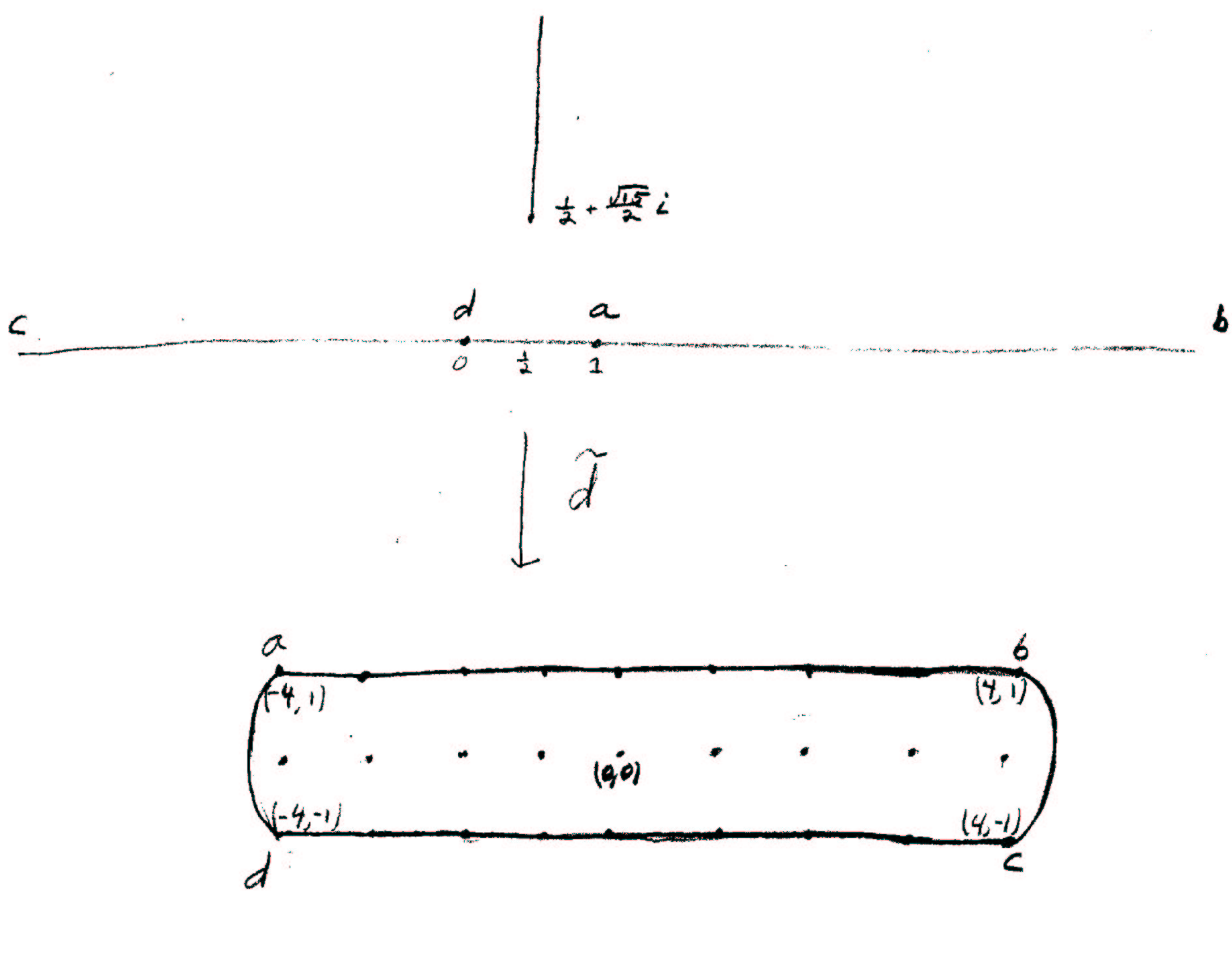
$$\overline{1 - w}$$

(and $z$ to $\overline{1-z}$). From 4.6.1 we conclude that if $\tilde{d}(w) = (\mu, \lambda)$, then $\tilde{d}(\sigma w) = (\mu, -\lambda)$. With this information, we know the boundary behavior of $\tilde{d}$ except when $w$ or $\tau w$ is near the ray $r$ described by

$$\operatorname{Re}(w) = \tfrac{1}{2}, \quad \operatorname{Im}(w) \geq \tfrac{\sqrt{15}}{2}\, i.$$

The image of the two sides of this ray is not so neatly described, but it does not represent a true boundary for the family of hyperbolic structures on $S^3 - K$, as $w$ crosses $r$ from right to left, for instance, $z$ crosses the real line in the interval $(0, \frac{1}{2})$. For a while, a hyperbolic structure can be constructed from the positively oriented simplex determined by $w$ and the negatively oriented simplex determined by $z$, by cutting the $z$-simplex into pieces which are subtracted from the $w$-simplex to leave a polyhedron $P$. $P$ is then identified to give a hyperbolic structure for $S^3 - K$. 4.21

For this reason, we give only a rough sketch of the boundary behavior of $\tilde{d}$ near $r$ or $\tau(r)$:



Since the image of $\tilde{d}$ in $\hat{\mathbb{R}}^2$ does not contain the origin, and since $\tilde{d}$ sends a curve winding once around the boundary of $R$ to the curve $abcd$ in $\hat{\mathbb{R}}^2$, it follows that the image of $\tilde{d}(R)$ contains the exterior of this curve. 4.22

In particular

Theorem 4.7. *Every manifold obtained by Dehn surgery along the figure-eight knot $K$ has a hyperbolic structure, except the six manifolds:*

$$(S^3 - K)_{(\mu,\lambda)} = (S^3 - K)_{(\pm\mu,\pm\lambda)}$$

*where* $(\mu, \lambda)$ *is* $(1, 0)$, $(0, 1)$, $(1, 1)$, $(2, 1)$, $(3, 1)$ *or* $(4, 1)$.

The equation

$$(S^3 - K)_{(\alpha,\beta)} = (S^3 - K_{(-\alpha,\beta)}$$

follows from the existence of an orientation reversing homeomorphism of $S^3 - K$.

I first became interested in studying these examples by the beautiful ideas of Jørgensen (compare Jørgensen, "Compact three-manifolds of constant negative curvature fibering over the circle," *Annals* 106 (1977) 61-72). He found the hyperbolic structures corresponding to the ray $\mu = 0$, $\lambda > 1$, and in particular, the integer and half-integer (!) points along this ray, which determine discrete groups.

The statement of the theorem is meant to suggest, but not imply, the true fact that the six exceptions do *not* have hyperbolic structures. Note that at least

$$S^3 = (S^3 - K)_{(1,0)}$$

does not admit a hyperbolic structure (since $\pi_1(S^3)$ is finite). We shall arrive at an understanding of the other five exceptions by studying the way the hyperbolic structures are degenerating as $(\mu, \lambda)$ tends to the line segment

$$\overline{(-4, 1), (4, 1)}.$$

## 4.8. Degeneration of hyperbolic structures.

Definition 4.8.1. A codimension-$k$ foliation of an $n$-manifold $M$ is a $\mathcal{G}$-structure, on $M$, where $\mathcal{G}$ is the pseudogroup of local homeomorphisms of $\mathbb{R}^{n-k} \times \mathbb{R}^k$ which 4.23
have the local form

$$\phi(x, y) = (f(x, y),\ g(y)).$$

In other words, $\mathcal{G}$ takes horizontal $(n - k)$-planes to horizontal $(n - k)$-planes. These horizontal planes piece together in $M$ as $(n - k)$-submanifolds, called the *leaves* of the foliation. $M$, like a book without its cover, is a disjoint union of its leaves.

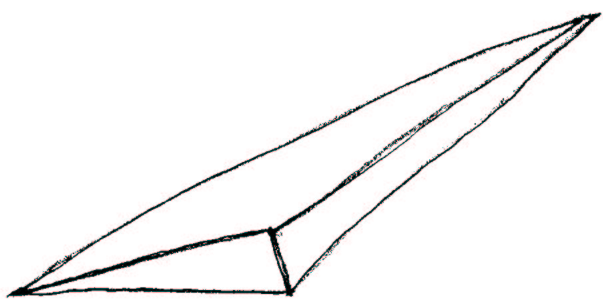
For any pseudogroup $\mathcal{H}$ of local homeomorphisms of some $k$-manifold $N^k$, the notion of a codimension-$k$ foliation can be refined:

DEFINITION 4.8.2. An $\mathcal{H}$-*foliation* of a manifold $M^n$ is a $\mathcal{G}$-structure for $M^n$, where $\mathcal{G}$ is the pseudogroup of local homeomorphisms of $\mathbb{R}^{n-k} \times N^k$ which have the local form

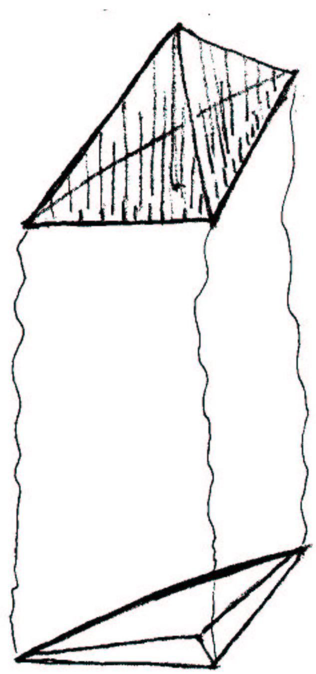
$$\phi(x, y) = (f(x, y),\, g(y))$$

with $g \in \mathcal{H}$. If $\mathcal{H}$ is the pseudogroup of local isometries of hyperbolic $k$-space, then an $\mathcal{H}$-foliation shall, naturally, be called a codimension-$k$ hyperbolic foliation. A hyperbolic foliation determines a hyperbolic structure for each $k$-manifold transverse to its leaves.

When $w$ tends in the region $R \subset \mathbb{C}$ to a point $\mathbb{R} - [0, 1]$, the $w$-simplex and the $z$-simplex are both flattening out, and in the limit they are flat: 4.24



If we regard these flat simplices as projections of nondegenerate simplices $A$ and $B$ (with vertices deleted), this determines codimension-2 foliations on $A$ and $B$, whose leaves are preimages of points in the flat simplices:



4.25

$A$ and $B$ glue together (in a unique way, given the combinatorial pattern) to yield a hyperbolic foliation on $S^3 - K$. You should satisfy yourself that the gluing consistency conditions for the hyperbolic foliation near an edge result as the limiting case of the gluing conditions for the family of squashing hyperbolic structures.

The notation of the developing map extends in a straightforward way to the case of an $\mathcal{H}$-foliation on a manifold $M$, when $\mathcal{H}$ is the set of restrictions of a group $J$ of real analytic diffeomorphisms of $N^k$; it is a map

$$D : \tilde{M}^n \to N^k.$$

Note that $D$ is not a local homeomorphism, but rather a local projection map, or a submersion. The holonomy

$$H : \pi_1(M) \to J$$

is defined, as before, by the equation

$$D \circ T_\alpha = H(\alpha) \circ D.$$

Here is the generalization of proposition 3.6 to $\mathcal{H}$-foliations. For simplicity, assume that the foliation is differentiable:

Proposition 4.8.1. *If $J$ is transitive and if the isotropy subgroups $J_x$ are compact, then the developing map for any $\mathcal{H}$-foliation $\mathcal{F}$ of a closed manifold $M$ is a fibration* labeled 4.8.1def 4.26

$$D : \tilde{M}^n \to N^k.$$

Proof. Choose a plane field $\tau^k$ transverse to $\mathcal{F}$ (so that $\tau$ is a complementary subspace to the tangent space to the leaves of $\mathcal{F}$, called $T\mathcal{F}$, at each point). Let $h$ be an invariant Riemannian metric on $N^k$ and let $g$ be any Riemannian metric on $M$. Note that there is an upper bound $K$ for the difference between the $g$-length of a nonzero vector in $\tau$ and the $k$-length of its local projection to $N^k$.

Define a horizontal path in $\tilde{M}$ to be any path whose tangent vector always lies in $\tau$. Let $\alpha : [0,1] \to N$ be any differentiable path, and let $\tilde{\alpha}_0$ be any point in the preimage $D^{-1}(\alpha_0)$. Consider the problem of lifting $\alpha$ to a horizontal path in $\tilde{M}$ beginning at $\tilde{\alpha}_0$. Whenever this has been done for a closed interval (such as $[0,0]$), it can be obviously extended to an open neighborhood. When it has been done for an open interval, the horizontal lift $\tilde{\alpha}$ is a Cauchy path in $\tilde{M}$, so it converges. Hence, by "topological induction", $\alpha$ has a (unique) global horizontal lift beginning at $\tilde{\alpha}_0$. Using horizontal lifts of the radii of disks in $N$, local trivializations for $D : \tilde{M} \to N$ are obtained, showing that $D$ is a fibration. □

Definition. An $\mathcal{H}$-foliation is *complete* if the developing map is a fibration.

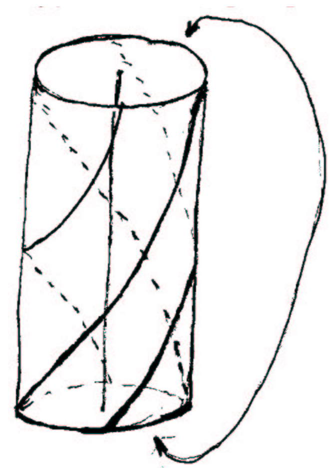
4.27

Any three-manifold with a complete codimension-2 hyperbolic foliation has universal cover $H^2 \times \mathbb{R}$, and covering transformations act as global isometries in the first coordinate. Because of this strong structure, we can give a complete classification of such manifolds. A *Seifert fibration* of a three-manifold $M$ is a projection $p : M \to B$ to some surface $B$, so that $p$ is a submersion and the preimages of points are circles in $M$. A Seifert fibration is a fibration except at a certain number of singular points $x_1, \ldots, x_k$. The model for the behavior in $p^{-1}(N_\epsilon(x_i))$ is a solid torus with a foliation having the core circle as one leaf, and with all other leaves winding $p$ times around the short way and $q$ times around the long way, where $1 < p < q$ and $(p,q) = 1$.

The projection of a meridian disk of the solid torus to its image in $B$ is $q$-to-one, 4.28
except at the center where it is one-to-one.

A group of isometries of a Riemannian manifold is *discrete* if for any $x$, the orbit of $x$ intersects a small neighborhood of $x$ only finitely often. A discrete group $\Gamma$ of orientation-preserving isometries of a surface $N$ has quotient $N/\Gamma$ a surface. The projection map $N \to N/\Gamma$ is a local homeomorphism except at points $x$ where the isotropy subgroup $\Gamma_x$ is nontrivial. In that case, $\Gamma_x$ is a cyclic group $\mathbb{Z}/q\mathbb{Z}$ for some $q > 1$, and the projection is similar to the projection of a meridian disk cutting across a singular fiber of a Seifert fibration.

Theorem 4.9. *Let $\mathcal{F}$ be a hyperbolic foliation of a closed three-manifold $M$. Then either*



A meridian disk of the solid torus wraps $q$ times around its image disk. Here $p = 1$ and $q = 2$.

(a) *The holonomy group $H(\pi_1 M)$ is a discrete group of isometries of $H^2$, and the developing map goes down to a Seifert fibration*

$$D_{/\pi_1 M} : M \to H^2/H(\pi_1 M),$$

*or*

(b) *The holonomy group is not discrete, and $M$ fibers over the circle with fiber a torus.*

The structure of $\mathcal{F}$ and $M$ in case (b) will develop in the course of the proof. 4.29

Proof. (a) If $H(\pi_1 M)$ is discrete, then $H^2/H(\pi_1 M)$ is a surface. Since $M$ is compact the fibers of the fibration $D : \tilde{M}^3 \to H^2$ are mapped to circles under the projection $\pi : \tilde{M}^3 \to M^3$. It follows that $D/H(\pi_1 M) : M^3 \to H^2/H(\pi_1 M)$ is a Seifert fibration.

(b) When $H(\pi_1 M)$ is not discrete, the proof is more involved. First, let us assume that the foliation is oriented (this means that the leaves of the foliation are oriented, or in other words, it is determined by a vector field). We choose a $\pi_1$ $M$-invariant Riemannian metric $g$ in $\tilde{M}^3$ and let $\tau$ be the plane field which is perpendicular to the fibers of $D : \tilde{M}^3 \to H^2$. We also insist that along $\tau$, $g$ be equal to the pullback of the hyperbolic metric on $H^2$.

By construction, $g$ defines a metric on $M^3$, and, since $M^3$ is compact, there is an infimum $I$ to the length of a nontrivial simple closed curve in $M^3$ when measured with respect to $g$. Given $g_1, g_2 \in \pi_1 M$, we say that they are *comparable* if there is a $y \in \tilde{M}^3$ such that

$$d\big(D(g_1(y)),\, D(g_2(y))\big) < I,$$

where $d(\ ,\ )$ denotes the hyperbolic distance in $H^2$. In this case, take the geodesic in $H^2$ from $D(g_1(y))$ to $D(g_2(y))$ and look at its horizontal lift at $g_2(y)$. Suppose its other endpoint $e$ where $g_1(y)$. Then the length of the lifted path would be equal to the length of the geodesic in $H^2$, which is less than $I$. Since $g_1 g_2^{-1}$ takes $g_2(y)$ to $g_1(y)$, the path represents a nontrivial element of $\pi_1 M$ and we have a contradiction. Now if we choose a trivialization of $H^2 \times \mathbb{R}$, we can decide whether or not $g_1(x)$ is greater than $e$. If it is greater than $e$ we say that $g_1$ is greater than $g_2$, and write 4.30
$g_1 > g_2$, otherwise we write $g_1 < g_2$. To see that this local ordering does not depend on our choice of $y$, we need to note that

$$U(g_1, g_2) = \{x \mid d\big(H(g_1(x)),\ H(g_2(x))\big) < I\}$$

is a connected (in fact convex) set. This follows from the following lemma, the proof of which we defer.

LEMMA 4.9.1. *$f_{g_1,g_2}(x) = d(g_1 x, g_2 x)$ is a a convex function on $H^2$.*

One useful property of the ordering is that it is invariant under left and right multiplication. In other words $g_1 < g_2$ if and only if, for all $g_3$, we have $g_3 g_1 < g_3 g_2$ and $g_1 g_3 < g_2 g_3$. To see that the property of comparability is equivalent for these three pairs, note that since $H(\pi_1 H^3)$ acts as isometries on $H^2$,

$$d(Dg_1 y, Dg_2 y) < I \quad \text{implies that} \quad d(Dg_3 g_1 y, Dg_3 g_2 y) < I.$$

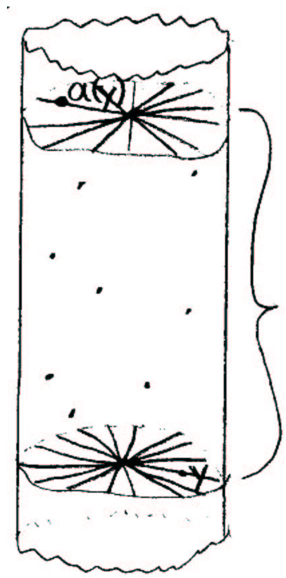
Also, if $d(Dg_1 y, Dg_2 y)$ then $d(Dg_3 g_1(g_3^{-1} y), Dg_3 g_2(g_3^{-1} y)) < I$, so that $g_1 g_3$ and $g_2 g_3$ are comparable. The invariance of the ordering easily follows (using the fact that $\pi_1 M$ preserves orientation of the $\mathbb{R}$ factors).

For a fixed $x \in H^2$ we let $G_\epsilon(X) \subset \pi_1 M$ be those elements whose holonomy acts on $x$ in a way $C^1 - \epsilon$-close to the identity. In other words, for $g \in G_\epsilon(x)$, $d(x, H_g(x)) < \epsilon$ and the derivative of $H_g(x)$ parallel translated back to $x$, is $\epsilon$-close to the identity.

PROPOSITION 4.9.2. *There is an $\epsilon_0$ so that for all $\epsilon < \epsilon_0\, [G_\epsilon, G_\epsilon] \subset G_\epsilon$.*

PROOF. For any Lie group the map $[*, *] : G \times G \to G$ has derivative zero at $(\mathrm{id}, \mathrm{id})$. Since for any $g \in G$, $(g, \mathrm{id}) \mapsto \mathrm{id}$ and $(\mathrm{id}, g) \mapsto 1$. The tangent spaces of $G \times \mathrm{id}$ and $\mathrm{id} \times G$ span the tangent space to $G \times G$ at $(\mathrm{id}, \mathrm{id})$. Apply this to the group of isometries of $H^2$. □

From now on we choose $\epsilon < I/8$ so that any two words of length four or less in $G_\epsilon$ are comparable. We claim that there is some $\beta \in G_\epsilon$ which is the "smallest" element in $G_\epsilon$ which is $>$ id. In other words, if $\mathrm{id} < \alpha \in G_\epsilon$, $a \neq \beta$, then $\alpha > \beta$. This can be seen as follows. Take an $\epsilon$-ball $B$ of $x \in H^2$ and look at its inverse image $\tilde{B}$ under $D$. Choose a point $y$ in $\tilde{B}$ and consider $y$ and $\alpha(y)$, where $\alpha \in G_\epsilon$. We can truncate $\tilde{B}$




There are only finitely many translates of $y$ in this region.

by the lifts of $B$ (using the horizontal lifts of the geodesics through $x$) through $y$ and $\alpha(y)$. Since this is a compact set there are only a finite number of images of $y$ under $\pi_1 M$ contained in it. Hence there is one $\beta(y)$ whose $\mathbb{R}$ coordinate is the closest to that of $y$ itself. $\beta$ is clearly our minimal element. 4.32

Now consider $\alpha > \beta > 1$, $\alpha \in G_\epsilon$. By invariance under left and right multiplication, $\alpha^2 > \beta_\alpha > \alpha$ and $\alpha > \alpha^{-1}\beta\alpha > 1$. Suppose $\alpha^{-1}\beta\alpha < \beta$. Then $\beta > \alpha^{-1}\beta\alpha > 1$ so that $1 > \alpha^{-1}\beta\alpha\beta^{-1} > \beta^{-1}$. Similarly if $\alpha^{-1}\beta\alpha > \beta > 1$ then $\beta > \alpha\beta\alpha^{-1} > 1$ so that $1 > \alpha\beta\alpha^{-1}\beta^{-1} > \beta^{-1}$. Note that by multiplicative invariance, if $g_1 > g_2$ then $g_2^{-1} = g_1^{-1}g_1g_2^{-1} > g_1^{-1}g_2g_2^{-1} = g_1^{-1}$. We have either $1 < \beta\alpha^{-1}\beta^{-1}\alpha < \beta$ or $1 < \beta\alpha\beta^{-1}\alpha^{-1} < \beta$ which contradicts the minimality of $\beta$ . Thus $\alpha^{-1}\beta\alpha = \beta$ for all $\alpha \in G_\epsilon$.

We need to digress here for a moment to classify the isometries of $H^2$. We will prove the following:

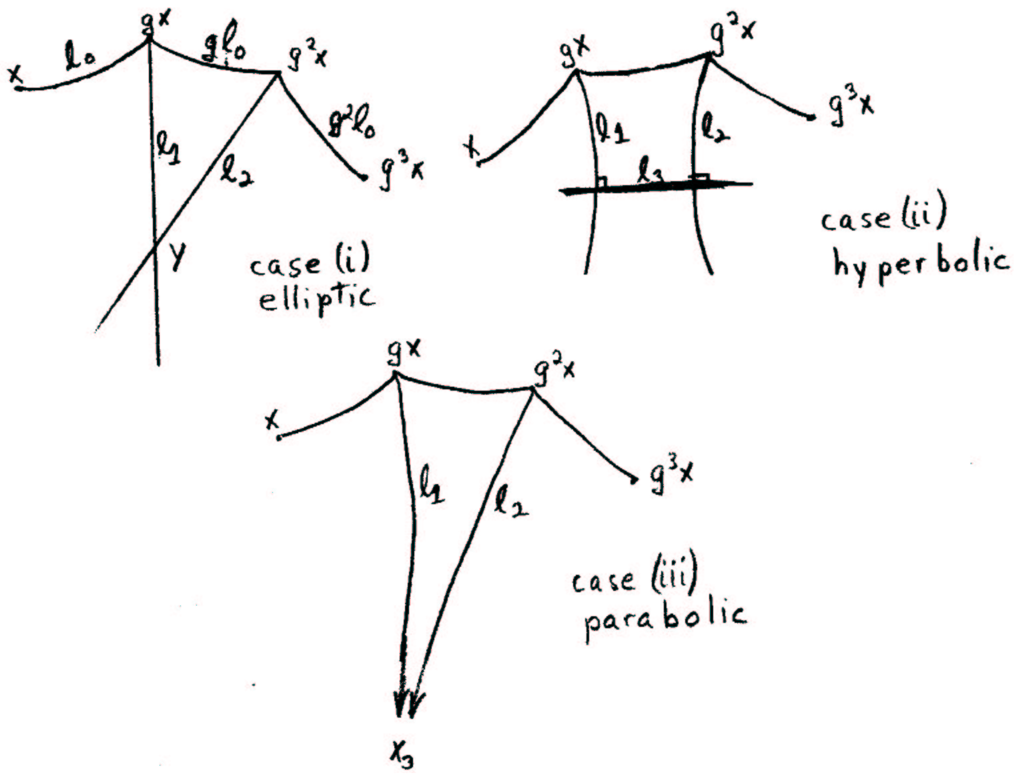
PROPOSITION 4.9.3. *If $g : H^2 \to H^2$ is a non-trivial isometry of $H^2$ which preserves orientation, then exactly one of the following cases occurs:*

(i) *$g$ has a unique interior fixed point or*
(ii) *$g$ leaves a unique invariant geodesic or*
(iii) *$g$ has a unique fixed point on the boundary of $H^2$.*

Case (i) is called *elliptic*, case (ii) *hyperbolic*, case (iii) *parabolic.*

Proof. *This follows easily from linear algebra, but we give a geometric proof.* Pick an interior point $x \in H^2$ and connect $x$ to $gx$ by a geodesic $l_0$. Draw the geodesics $l_1$, $l_2$ at $gx$ and $g^2x$ which bisect the angle made by $l_0$ and $gl_0$, $gl_0$ and $g^2 l_0$ respectively. There are three cases: 4.33

(i) $l_1$ and $l_2$ intersect in an interior point $y$
(ii) There is a geodesic $l_3$ perpendicular to $l_1$, $l_2$
(iii) $l_1, l_2$ are parallel, i.e., they intersect at a point at infinity $x_3$.




In case (i) the length of the arc $gx, y$ equals that of $g^2x, y$ since $\Delta(gx, g^2x, y)$ is an isoceles triangle. It follows that $y$ is fixed by $g$.

In case (ii) the distance from $gx$ to $l_3$ equals that from $g^2x$ to $l_3$. Since $l_3$ meets $l_1$ and $l_2$ in right angles it follows that $l_3$ is invariant by $g$. 4.34

Finally, in case (iii) $g$ takes $l_1$ and $l_2$, both of which hit the boundary of $H^2$ in the same point $x_3$. It follows that $g$ fixes $x_3$ since an isometry takes the boundary to itself.

Uniqueness is not hard to prove. □

Using the classification of isometries of $H^2$, it is easy to see that the centralizer of any non-trivial element $g$ in $\text{isom}(H^2)$ is abelian. (For instance, if $g$ is elliptic with fixed point $x_0$, then the centralizer of $g$ consists of elliptic elements with fixed point $x_0$). It follows that the centralizer of $\beta$ in $\pi_1(M)$ is abelian; let us call this group $N$.

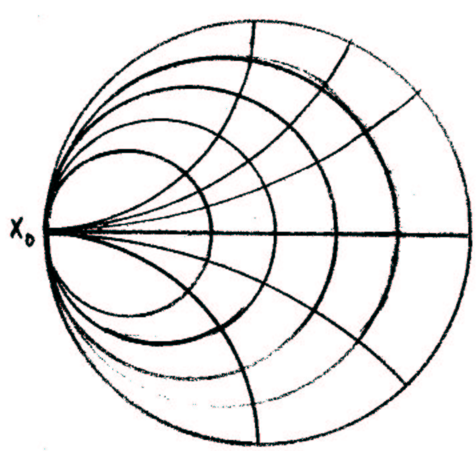
Although $G_\epsilon(x)$ depends on the point $x$, for any point $x' \in H^2$, if we choose $\epsilon'$ small enough, then $G_{\epsilon'}(x') \subset G_\epsilon(x)$. In particular if $x = H(g)x$, $g \in \pi_1 M$, then all elements of $G_{\epsilon'}(x')$ commute with $\beta$. It follows that $N$ is a normal subgroup of $\pi_1(M)$.

Consider now the possibility that $\beta$ is elliptic with fixed point $x_0$ and $n \in N$ fixes $x_0$ we see that all of $\pi_1 M$ must fix $x_0$. But the function $f_{x_0} : H^2 \to \mathbb{R}^+$ which measures the distance of a point in $H^2$ from $x_0$ is $H(\pi_1 M)$ invariant so that it lifts to a function $f$ on $M^3$. However, $M^3$ is compact and the image of $\tilde{f}$ is non-compact, which is impossible. Hence $\beta$ cannot be elliptic.

If $\beta$ were hyperbolic, the same reasoning would imply that $H(\pi_1 M)$ leaves invari- 4.35
ant the invariant geodesic of $\beta$. In this case we could define $f_l : H^2 \to \mathbb{R}$ to be the distance of a point from $l$. Again, the function lifts to a function on $M^3$ and we have a contradiction.

The case when $\beta$ is parabolic actually does occur. Let $x_0$ be the fixed point of $\beta$ on the circle at infinity. $N$ must also fix $x_0$. Using the upper half-plane model for $H^2$ with $x_0$ at $\infty$, $\beta$ acts as a translation of $\mathbb{R}^2$ and $N$ must act as a group of similarities; but since they commute with $\beta$, they are also translations. Since $N$ is normal, $\pi_1 M$ must act as a group of similarities of $\mathbb{R}^2$ (preserving the upper half-plane).

Clearly there is no function on $H^2$ measuring distance from the point $x_0$ at infinity. If we consider a family of finite points $x_\tau \to X$, and consider the functions $f_{x_\tau}$, even though $f_{x_\tau}$ blows up, its derivative, the closed 1-form $df_{x_\tau}$, converges to a closed 1 form $\omega$. Geometrically, $\omega$ vanishes on tangent vectors to horocycles about $x_0$ and takes the value 1 on unit tangents to geodesics emanating from $x_0$.




4.36

The non-singular closed 1-form $\omega$ on $H^2$ is invariant by $H(\pi_1 M)$, hence it defines a non-singular closed one-form $\bar{\omega}$ on $M$. The kernel of $\bar{\omega}$ is the tangent space to the leaves of a codimension one foliation $\mathcal{F}$ of $M$. The leaves of the corresponding foliation $\tilde{\mathcal{F}}$ on $\tilde{M}$ are the preimages under $D$ of the horocycles centered at $x_0$. The group of periods for $\omega$ must be discrete, for otherwise there would be a translate of the horocycle about $x_0$ through $x$ close to $x$, hence an element of $G_\epsilon$ which does not commute with $\beta$. Let $p_0$ be the smallest period. Then integration of $\omega$ defines a map from $M$ to $S^1 = \mathbb{R}/\langle p_0 \rangle$, which is a fibration, with fibers the leaves of $\mathcal{F}$. The fundamental group of each fiber is contained in $N$, which is abelian, so the fibers are toruses.

It remains to analyze the case that the hyperbolic foliation is not oriented. In this case, let $M'$ be the double cover of $M$ which orients the foliation. $M'$ fibers over $S^1$ with fibration defined by a closed one-form $\omega$. Since $\omega$ is determined by the unique fixed point at infinity of $H(\pi_1 M')$, $\omega$ projects to a non-singular closed one-form on $M$. This determines a fibration of $M$ with torus fibers. (Klein bottles cannot occur even if $M$ is not required to be orientable.) $\square$

We can construct a three-manifold of type (b) by considering a matrix

$$A \in SL(2, \mathbb{Z})$$

which is hyperbolic, i.e., it has two eigenvalues $\lambda_1, \lambda_2$ and two eigenvectors $V_1, V_2$. Then $AV_1 = \lambda_1 V_1, AV_2 = \lambda_2 V_2$ and $\lambda_2 = 1/\lambda_1$.

Since $A \in SL(2, \mathbb{Z})$ preserves $\mathbb{Z} \oplus \mathbb{Z}$ its action on the plane descends to an action on the torus $T^2$. Our three-manifold $M_A$ is the mapping torus of the action of $A$ 4.37
on $T^2$. Notice that the lines parallel to $V_1$ are preserved by $A$ so they give a one-dimensional foliation on $M_A$. Of course, the lines parallel to $V_2$ also define a foliation. The reader may verify that both these foliations are hyperbolic. When

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

then $M_A$ is the manifold $(S^3 - K)_{(D, \pm 1)}$ obtained by Dehn surgery on the figure-eight knot. The hyperbolic foliations corresponding to $(0, 1)$ and $(0, -1)$ are distinct, and they correspond to the two eigenvectors of

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

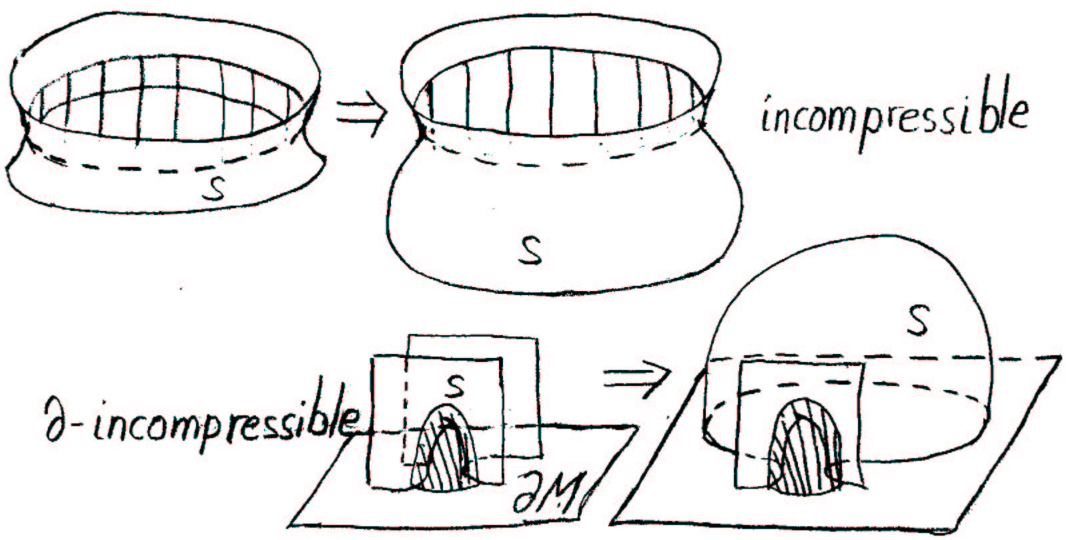
All codimension-2 hyperbolic foliations with leaves which are not closed are obtained by this construction. This follows easily from the observation that the hyperbolic foliation restricted to any fiber is given by a closed non-singular one-form, together with the fact that a closed non-singular one-form on $T^2$ is determined (up to isotopy) by its cohomology class.

The three manifolds $(S^3 - K)_{(1,1)}$, $(S^3 - K)_{(2,1)}$ and $(S^3 - K)_{(3,1)}$ also have codimension-2 hyperbolic foliations which arise as "limits" of hyperbolic structures.

Since they are rational homology spheres, they must be Seifert fiber spaces. A Seifert fiber space cannot be hyperbolic, since (after passing to a cover which orients the fibers) a general fiber is in the center of its fundamental group. On the other hand, the centralizer of an element in the fundamental group of a hyperbolic manifold is abelian. 4.38

### 4.10. Incompressible surfaces in the figure-eight knot complement.

Let $M^3$ be a manifold and $S \subset M^3$ a surface with $\partial S \subset \partial M$. Assume that $S \neq S^2$, $IP^2$, or a disk $D^2$ which can be pushed into $\partial M$. Then $S$ is *incompressible* if every loop (simple closed curve) on $S$ which bounds an (open) disk in $M - S$ also bounds a disk in $S$. Some people prefer the alternate, stronger definition that $S$ is (*strongly*) *incompressible* if $\pi_1(S)$ injects into $\pi_1(M)$. By the loop theorem of Papakyriakopoulos, these two definitions are equivalent if $S$ is two-sided. If $S$ has boundary, then $S$ is also $\partial$-*incompressible* if every arc $\alpha$ in $S$ (with $\partial(\alpha) \subset \partial S$) which is homotopic to $\partial M$ is homotopic in $S$ to $\partial S$.

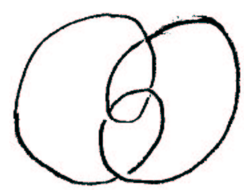



If $M$ is oriented and irreducible (every two-sphere bounds a ball), then $M$ is *sufficiently large* if it contains an incompressible and $\partial$-incompressible surface. A 4.39
compact, oriented, irreducible, sufficiently large three-manifold is also called a *Haken-manifold*. It has been hard to find examples of three-manifolds which are irreducible but can be shown not to be sufficiently large. The only previously known examples are certain Seifert fibered spaces over $S^2$ with three exceptional fibers. In what follows we give the first known examples of compact, irreducible three-manifolds which are not Haken-manifolds and are not Seifert fiber spaces.

NOTE. If $M$ is a compact oriented irreducible manifold $\neq D^3$, and either $\partial M \neq \emptyset$ or $H^1(M) \neq 0$, then $M$ is sufficiently large. In fact, $\partial M \neq 0 \Rightarrow H^1(M) \neq 0$. Think of a non-trivial cohomology class $\alpha$ as dual to an embedded surface; an easy argument using the loop theorem shows that the simplest such surface dual to $\alpha$ is incompressible and $\partial$-incompressible.

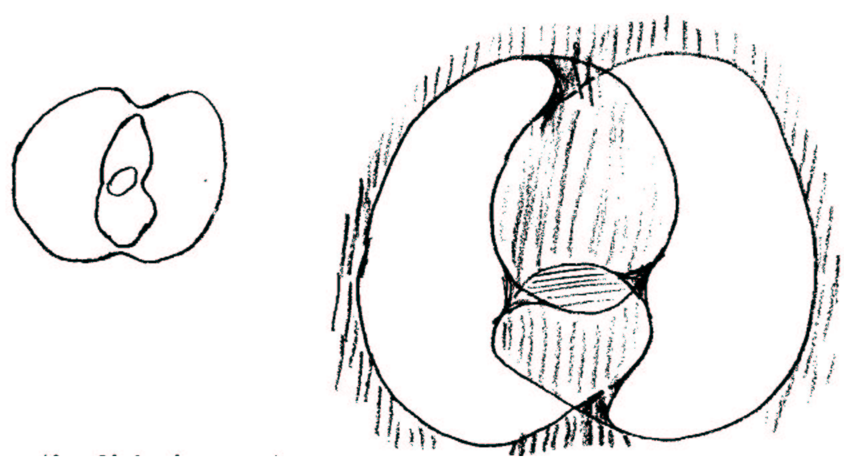
The concept of an incompressible surface was introduced by W. Haken (International Congress of Mathematicians, 1954), (*Acta. Math.* 105 (1961), *Math A.* 76 (1961), *Math Z* 80 (1962)). If one splits a Haken-manifold along an incompressible and $\partial$-incompressible surface, the resulting manifold is again a Haken-manifold. One can continue this process of splitting along incompressible surfaces, eventually arriving (after a bounded number of steps) at a union of disks. Haken used this to give algorithms to determine when a knot in a Haken-manifold was trivial, and when two knots were linked.

Let $K$ be a figure-eight knot, $M = S^3 - \mathcal{N}(K)$. $M$ is a Haken manifold by the above note [$M$ is irreducible, by Alexander's theorem that every differentiable two-sphere in $S^3$ bounds a disk (on each side)]. 4.40



Here is an enumeration of the incompressible and $\partial$-incompressible surfaces in $M$. There are six reasonably obvious choices to start with;

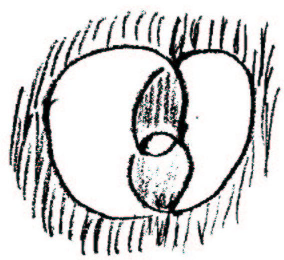
- $S_1$ is a torus parallel to $\partial M$,
- $S_2 = T^2$-disk = Seifert surface for $K$. To construct $S_2$, take 3 circles lying above the knot, and span each one by a disk. Join



the disks by a twist for each crossing at $K$ to get a surface $S_2$ with boundary the longitude $(0, \pm 1)$. $S_2$ is oriented and has Euler characteristic $= -1$, so it is $T^2$-disk.

- $S_3 =$ (Klein bottle-disk) is the unoriented surface pictured.

- $S_4 = \partial$ (tubular neighborhood of $S_3$) $= T^2 - 2$ disks. $\partial S_4 = (\pm 4, 1)$, (depending on the choice of orientation for the meridian).
- $S_5 =$ (Klein bottle-disk) is symmetric with $S_3$.



- $S_6 = \partial$ (tubular neighborhood of $S_5$ ) $= T^2 - 2$ disks. $\partial S_6 = (\pm 4, 1)$.

It remains to show that

Theorem 4.11. *Every incompressible and $\partial$-incompressible connected surface in $M$ is isotopic to one of $S_1$ through $S_6$.*

Corollary. *The Dehn surgery manifold $M_{(m,l)}$ is* irreducible*, and it is a Haken-manifold if and only if $(m, l) = (0, \pm 1)$ or $(\pm 4, \pm 1)$.*

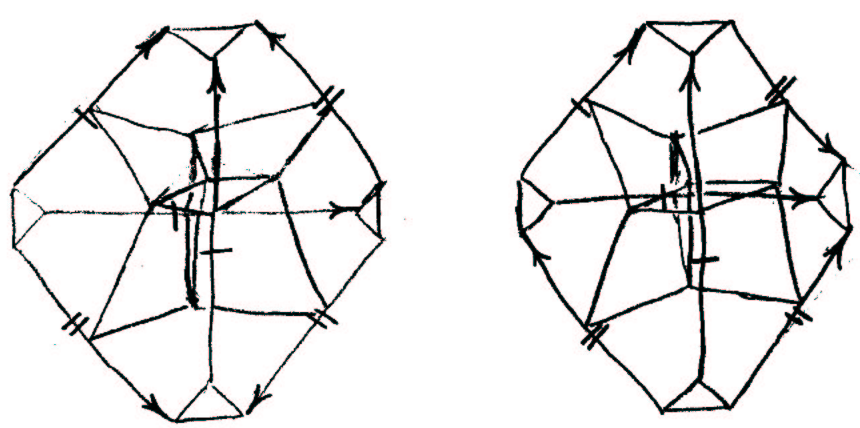
In particular, none of the hyperbolic manifolds obtained from $M$ by Dehn surgery is sufficiently large. (Compare 4.7.) Thus we have an infinite family of examples of oriented, irreducible, non-Haken-manifolds which are not Seifert fiber spaces. It seems likely that Dehn surgery along other knots and links would yield many more examples. 4.42

Proof of corollary from theorem. Think of $M_{(m,l)}$ as $M$ union a solid torus, $D^2 \times S^1$, the solid torus being a thickened core curve. To see that $M_{(m,l)}$ is irreducible, let $S$ be an embedded $S^2$ in $M_{(m,l)}$, transverse to the core curve $\alpha$ ($S$ intersects the solid torus in meridian disks). Now isotope $S$ to minimize its intersections with $\alpha$. If $S$ doesn't intersect $\alpha$ then it bounds a ball by the irreducibility

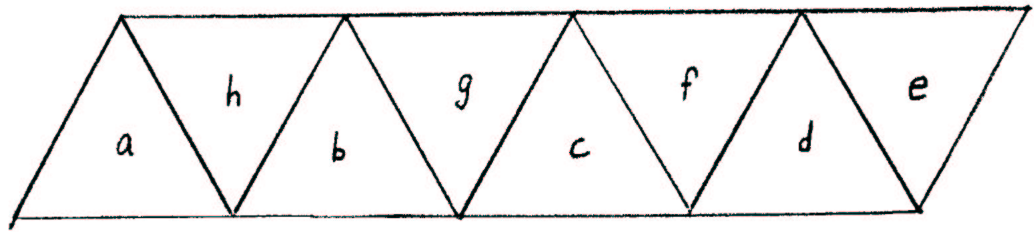
of $M$. If it does intersect $\alpha$ we may assume each component of intersection with the solid torus $D^2 \times S^1$ is of the form $D^2 \times x$. If $S \cap M$ is not incompressible, we may divide $S$ into two pieces, using a disk in $S \cap M$, each of which has fewer intersections with $\alpha$. If $S$ does not bound a ball, one of the pieces does not bound. If $S \cap M$ is $\partial$-incompressible, we can make an isotopy of $S$ to reduce the number of intersections with $\alpha$ by 2. Eventually we simplify $S$ so that if it does not bound a ball, $S \cap M$ 4.43
is incompressible and $\partial$-incompressible. Since none of the surfaces $S_1, \ldots, S_6$ is a submanifold of $S^2$, it follows from the theorem that $S$ in fact bounds a ball.

The proof that $M_{(m,l)}$ is not a Haken-manifold if $(m,l) \neq (0, \pm 1)$ or $(\pm 4, \pm 1)$ is similar. Suppose $S$ is an incompressible surface in $M_{(m,l)}$. Arrange the intersections with $D^2 \times S^1$ as before. If $S \cap M$ is not incompressible, let $D$ be a disk in $M$ with $\partial D \subset S \cap M$ not the boundary of a disk in $S \cap M$. Since $S$ in incompressible, $\partial D = \partial D'$ for some disk $D' \subset S$ which must intersect $\alpha$. The surface $S'$ obtained from $S$ by replacing $D'$ with $D$ is incompressible. (It is in fact isotopic to $S$, since $M$ is irreducible; but it is easy to see that $S'$ is incompressible without this.) $S'$ has fewer intersections with $\alpha$ than does $S$. If $S$ is not $\partial$-incompressible, an isotopy can be made as before to reduce the number of intersections with $\alpha$. Eventually we obtain an incompressible surface (which is isotopic to $S$) whose intersection with $M$ is incompressible and $\partial$-incompressible. $S$ cannot be $S_1$ (which is not incompressible in $M_(m,l)$), so the corollary follows from the theorem. $\square$

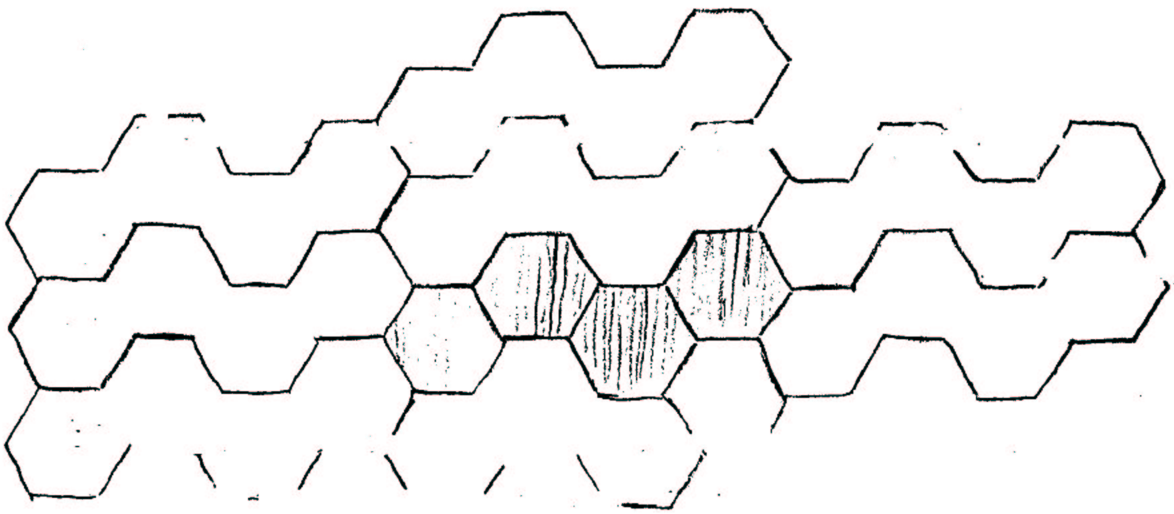
Proof of theorem 4.11. Recall that $M = S^3 - \mathcal{N}(K)$ is a union of two tetrahedra-without-vertices. To prove the theorem, it is convenient to use an alternate description of $M$ at $T^2 \times I$ with certain identifications on $T^2 \times \{1\}$ (compare Jørgensen, "Compact three-manifolds of constant negative curvature fibering over the 4.44
circle", *Annals of Mathematics* **106** (1977), 61–72, and R. Riley). One can obtain this from the description of $M$ as the union of two tetrahedra with corners as follows. Each tetrahedron $=$ (corners) $\times I$ with certain identifications on (corners) $\times \{1\}$.



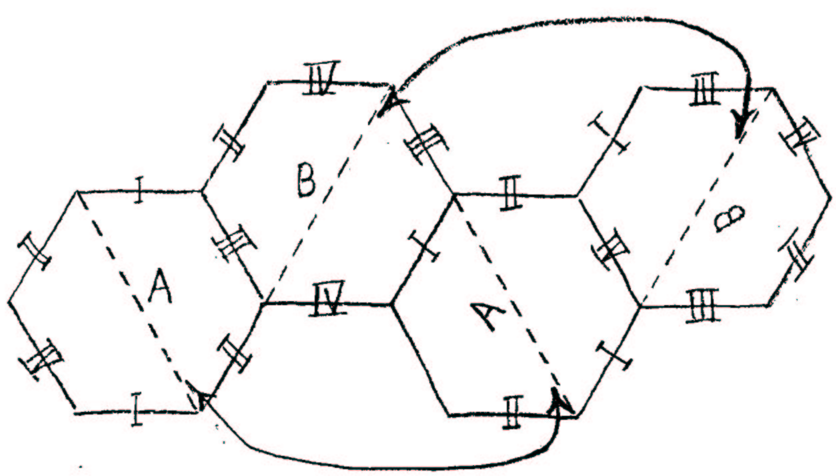
This “product” structure carries over to the union of the two tetrahedra. The boundary torus has the triangulation (p. 4.11)




4.45



$T^2 \times \{1\}$ has the dual subdivision, which gives $T^2$ as a union of four hexagons. The diligent reader can use the gluing patters of the tetrahedra to check that the identifications on $T^2 \times \{1\}$ are



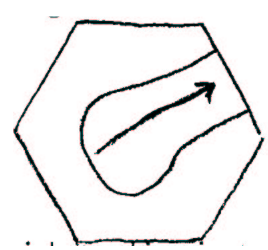
where we identify the hexagons by flipping through the dotted lines.

The complex $N = T^2 \times \{1\}/\text{identifications}$ is a spine for $N$. It has a cell subdivision with two vertices, four edges, and two hexagons. $N$ is embedded in $M$, and its complement is $T^2 \times [0, 1)$.

If $S$ is a connected, incompressible surface in $M$, the idea is to simplify it with respect to the spine $N$ (this approach is similar in spirit to Haken's). First isotope $S$ 4.46
so it is transverse to each cell of $N$. Next isotope $S$ so that it doesn't intersect any hexagon in a simple closed curve. Do this as follows.

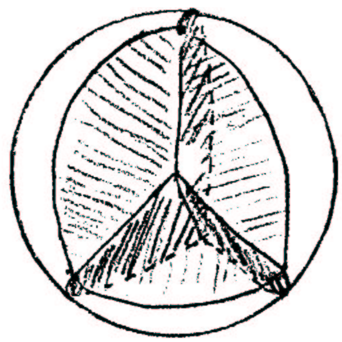
If $S \cap$ hexagon contains some loops, pick an innermost loop $\alpha$. Then $\alpha$ bounds an open disk in $M^2 - S$ (it bounds one in the hexagon), so by incompressibility it bounds a disk in $S$. By the irreducibility of $M$ we can push this disk across this hexagon to eliminate the intersection $\alpha$. One continues the process to eliminate all such loop intersections. This does not change the intersection with the one-skeleton $N_{(1)}$.

$S$ now intersects each hexagon in a collection of arcs. The next step is to isotope $S$ to minimize the number of intersections with $N_{(1)}$. Look at the preimage of $S \cap N$. We can eliminate any arc which enters and leaves a hexagon in the same edge by pushing the arc across the edge.



If at any time a loop intersection is created with a hexagon, eliminate it before proceeding. 4.47

Next we concentrate on corner connections in hexagons, that is, arcs which connect two adjacent edges of a hexagon. Construct a small ball $\mathcal{B}$ about each vertex,

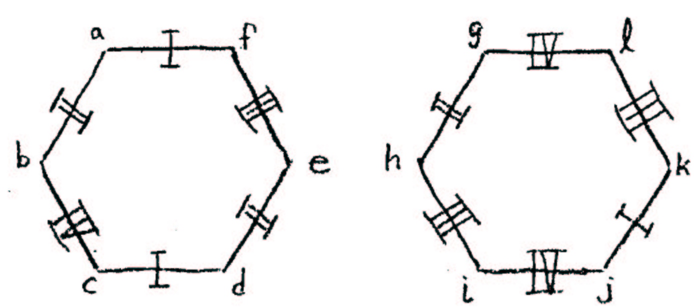


and push $S$ so that the corner connections are all contained in $\mathcal{B}$, and so that $S$ is transverse to $\partial\mathcal{B}$. $S$ intersects $\partial\mathcal{B}$ in a system of loops, and each component of intersection of S with $\mathcal{B}$ contains at least one corner connection, so it intersects $N_{(1)}$ at least twice. If any component of $S \cap \mathcal{B}$ is not a disk, there is some "innermost" such component $S_i$; then all of its boundary components bound disks in $\mathcal{B}$, hence in $S$. Since $S$ is not a sphere, one of these disks in $S$ contains $S_i$. Replace it by a disk in $\mathcal{B}$. This can be done without increasing the number of intersections with $N_{(1)}$, since every loop in $\partial\mathcal{B}$ bounds a disk in $\mathcal{B}$ meeting $N_{(1)}$ at most twice.

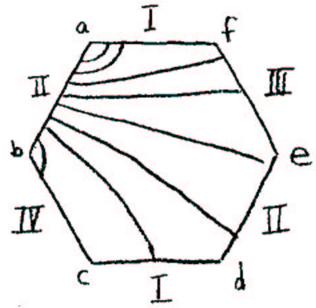
Now if there are any two corner connections in $\mathcal{B}$ which touch, then some component of $S \cap \mathcal{B}$ meets $N_{(1)}$ at least three times. Since this component is a disk, it can 4.48
be replaced by a disk which meets $N_{(1)}$ at most twice, thus simplifying $S$. (Therefore at most two corners can be connected at any vertex.)

Assume that $S$ now has the minimum number of intersections with $N_{(1)}$ in its isotopy class. Let I, II, III, and IV denote the number of intersections of $S$ with edges I, II, III, and IV, respectively (no confusion should result from this). It remains to analyze the possibilities case by case.

Suppose that none of I, II, III, and IV are zero. Then each hexagon has connections at two corners. Here are the possibilities for corner connections in hexagon A.




If the corner connections are at $a$ and $b$ then the picture in hexagon A is of the form



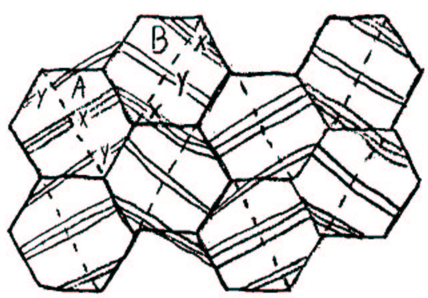

4.49

This implies that $\mathrm{II} = \mathrm{I} + \mathrm{III} + \mathrm{II} + \mathrm{I} + \mathrm{IV}$, which is impossible since all four numbers are positive in this case. A similar argument also rules out the possibilities c-d, d-e, a-f, b-f, and c-e in hexagon, and h-i, i-j, k-l, g-l, g-k and h-j in hexagon B.

The possibility a-c cannot occur since they are adjacent corners. For the same reason we can rule out a-e, b-d, d-f, g-i, i-k, h-l, and j-l.

Since each hexagon has at least two corner connections, at each vertex we must have connections at two opposite corners. This means that knowing any one corner connection also tells you another corner connection. Using this one can rule out all possible corner connections for hexagon A except for a-d.

If a-d occurs, then $\mathrm{I} + \mathrm{IV} + \mathrm{II} = \mathrm{I} + \mathrm{III} + \mathrm{II}$, or $\mathrm{III} = \mathrm{IV}$. By the requirement of opposite corners at the vertices, in hexagon B there are corner connections at i and l, which implies that $\mathrm{I} = \mathrm{II}$. Let $x = \mathrm{III}$ and $y = \mathrm{I}$. The picture is then

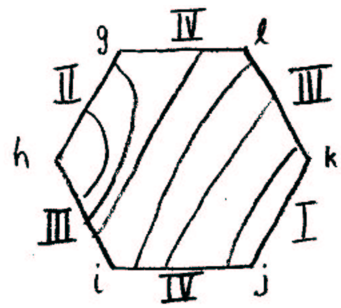



4.50

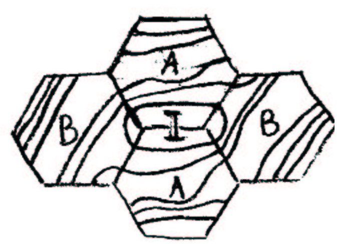
We may reconstruct the intersection of $S$ with a neighborhood of $N$, say $\mathcal{N}(N)$, from this picture, by gluing together $x + y$ annuli in the indicated pattern. This yields $x+y$ punctured tori. If an $x$-surface is pushed down across a vertex, it yields a $y$-surface, and similarly, a $y$-surface can be pushed down to give an $x$-surface. Thus, $S \cap \mathcal{N}(N)$ is $x+y$ parallel copies of a punctured torus, which we see is the fiber of a fibration of $\mathcal{N}(N) \approx M$ over $S^1$. We will discuss later what happens outside $\mathcal{N}(N)$. (Nothing.)

Now we pass on to the case that at least one of I, II, III, and IV are zero. The case $\mathrm{I} = 0$ is representative because of the great deal of symmetry in the picture.

First consider the subcase I $= 0$ and none of II, III, and IV are zero. If hexagon B had only one corner connection, at h, then we would have III + IV = II + IV + III,

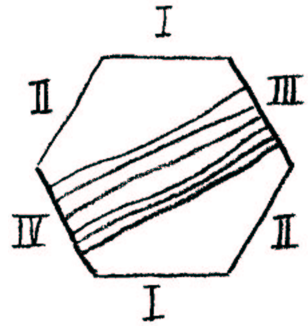



contradicting II $> 0$. By the same reasoning for all the other corners, we find that hexagon B needs at least two corner connections. At most one corner connection can occur in a neighborhood of each vertex in $N$, since no corner connection can involve the edge I. Thus, hexagon B must have exactly two corner connections, and hexagon A has no corner connections. By checking inequalities, we find the only possibility is corner connections at g-h. If we look at the picture in the pre-image $T^2 \times \{1\}$ near I we see that there is a loop around I. This loop bounds a disk in $S$ by incompressibility, 4.51

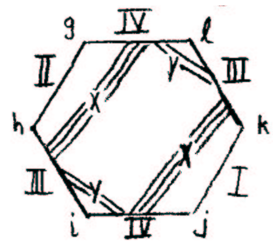

and pushing the disk across the hexagons reduces the number of intersections with $N_{(1)}$ by at least two (you lose the four intersections drawn in the picture, and gain possibly two intersections, above the plane of the paper). Since $S$ already has minimal intersection number with $N_{(1)}$ already, this subcase cannot happen.

Now consider the subcase I $= 0$ and II $= 0$. In hexagon A the picture is



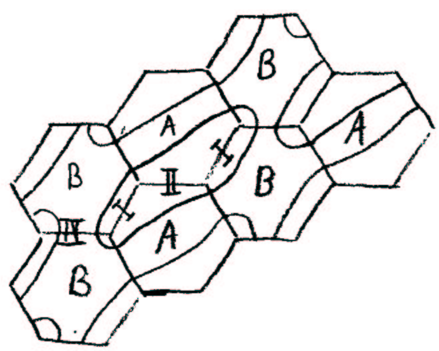

4.52

implying III = IV. The picture in hexagon B is



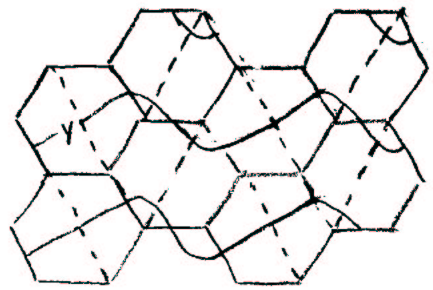

with $y$ the number of corner connections at corner $l$ and $x = \mathrm{IV} - y$. The three subcases to check are $x$ and $y$ both nonzero, $x = 0$, and $y = 0$.

If both $x$ and $y$ are nonzero, there is a loop in $S$ around

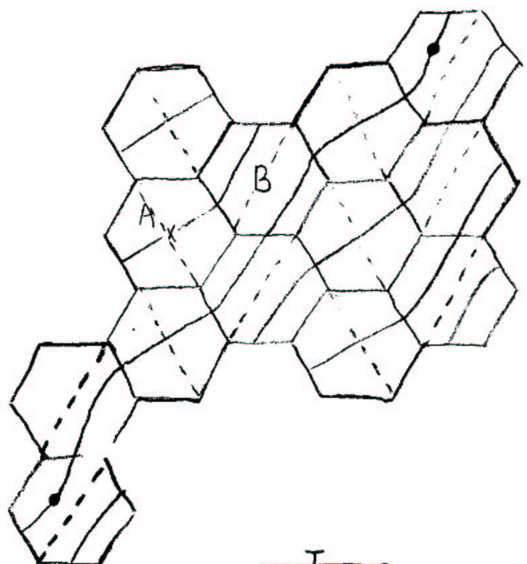


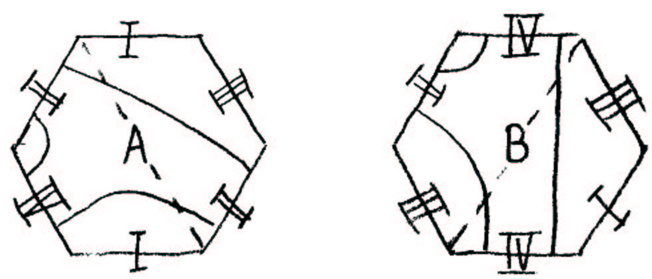

edges I and II. The loop bounds a disk in $S$, and pushing the disk across the hexagons reduces the number of intersections by at least two, contradicting minimality. So $x$ and $y$ cannot both be nonzero.

If I = II = 0 and $x = 0$, then $S \cap \mathcal{N}(N)$ is $y$ parallel copies of a punctured torus. 4.53

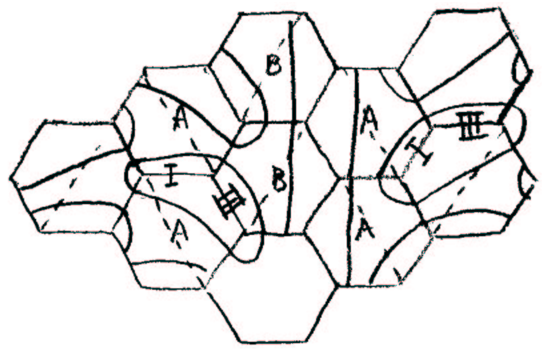



If I = II = 0 and $y = 0$, then $S \cap \mathcal{N}(N)$ consists of $\lfloor x/2 \rfloor$ copies of a twice punctured torus, together with one copy of a Klein bottle if $x$ is odd.

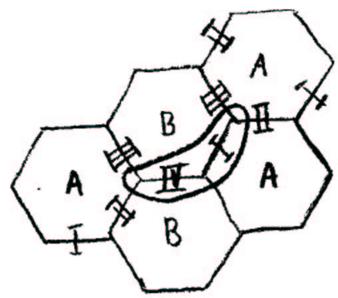


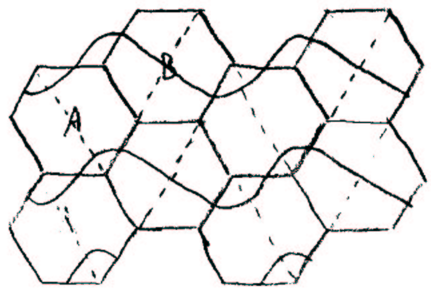

Now consider the subcase I = III = 0. If $S$ intersects the spine $N$, then II $\neq 0$ 4.54
because of hexagon A and IV $\neq 0$ because of hexagon B. But this means that there is a loop around edges I and III, and $S$ can be simplified further, contradicting minimality.

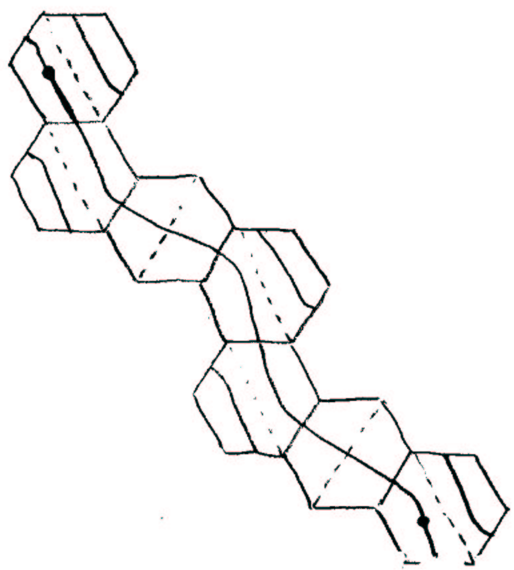



The subcase I = IV = 0 also cannot occur because of the minimality of the number of intersections of $S$ and $N_{(1)}$. Here is the picture.




By symmetric reasoning, we find that only one more case can occur, that III = IV = 0, with I = II. The pictures are symmetric with preceding ones: 4.55



To finish the proof of the theorem, it remains to understand the behavior of $S$ in $M - \mathcal{N}(N) = T^2 \times [0, .99]$. Clearly, $S \cap (T^2 \times [0, .99])$ must be incompressible. (Otherwise, for instance, the number of intersections of $S$ with $N_{(1)}$ could be reduced.) It is not hard to deduce that either $S$ is parallel to the boundary, or else a union of annuli. If one does not wish to assume $S$ is two-sided, this may be accomplished by studying the intersection of $S \cap (T^2 \times [0, .99])$ with a non-separating annulus. 4.56
If any annulus of $S \cap (T^2 \times [0, .99])$ has both boundary components in $T^2 \times .99$, then by studying the cases, we find that $S$ would not be incompressible. It follows that $S \cap (T^2 \times [0, .99])$ can be isotoped to the form (circles $\times$ $[0, .99]$). There are five possibilities (with $S$ connected). Careful comparisons lead to the descriptions of $S_2, \ldots, S_6$ given on pages 4.40 and 4.41. $\square$