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.

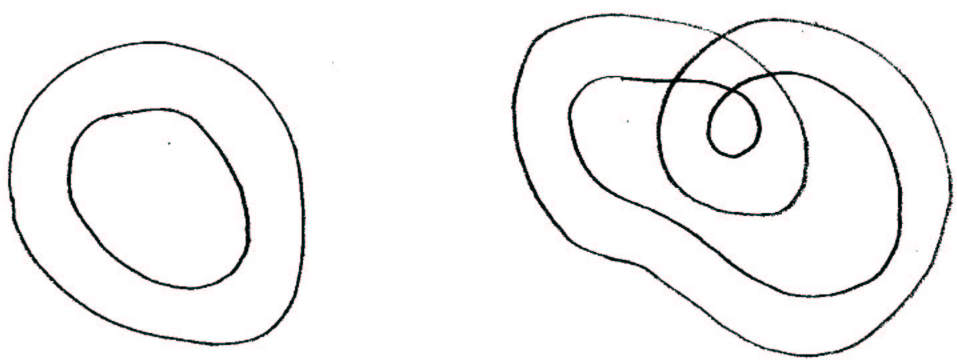
Thurston's *Three-Dimensional Geometry and Topology*, Vol. 1 (Princeton University Press, 1997) is a considerable expansion of the first few chapters of these notes. Later chapters have not yet appeared in book form.

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

CHAPTER 5

Flexibility and rigidity of geometric structures

In this chapter we will consider deformations of hyperbolic structures and of geometric structures in general. By a geometric structure on M, we mean, as usual, a local modelling of M on a space X acted on by a Lie group G. Suppose M is compact, possibly with boundary. In the case where the boundary is non-empty we do not make special restrictions on the boundary behavior. If M is modelled on (X, G) then the developing map $\tilde{M} \xrightarrow{D} X$ defines the holonomy representation $H : \pi_1 M \longrightarrow G$. In general, H does not determine the structure on M. For example, the two immersions of an annulus shown below define Euclidean structures on the annulus which both have trivial holonomy but are not equivalent in any reasonable sense.



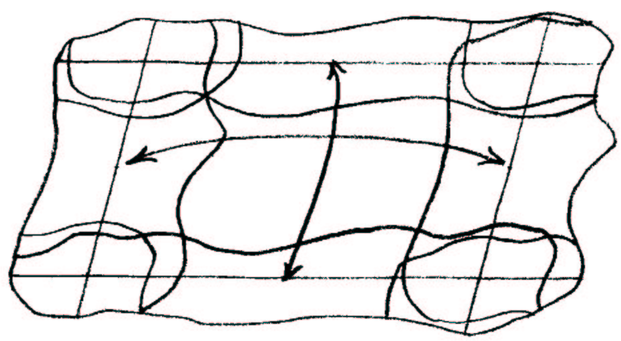
However, the holonomy is a complete invariant for (G, X)-structures on M *near* a given structure M_0, up to an appropriate equivalence relation: two structures M_1 and M_2 near M_0 are equivalent deformations of M_0 if there are submanifolds M_1' and 5.2
M_2', containing all but small neighborhoods of the boundary of M_1 and M_2, with a (G, X) homeomorphism between them which is near the identity.

Let M_0 denote a fixed structure on M, with holonomy H_0.

Proposition 5.1. *Geometric structures on M near M_0 are determined up to equivalency by holonomy representations of $\pi_1 M$ in G which are near H_0, up to conjugacy by small elements of G.*

Proof. Any manifold M can be represented as the image of a disk D with reasonably nice overlapping near ∂D. Any structure on M is obtained from the structure induced on D, by gluing via the holonomy of certain elements of $\pi_1(M)$.

Any representation of $\pi_1 M$ near H_0 gives a new structure, by perturbing the identifications on D. The new identifications are still finite-to-one giving a new manifold homeomorphic to M_0.



5.3

If two structures near M_0 have holonomy conjugate by a small element of G, one can make a small change of coordinates so that the holonomy is identical. The two structures then yield nearby immersions of D into X, with the same identifications; restricting to slightly smaller disks gives the desired (G, X)-homeomorphism. $\square$

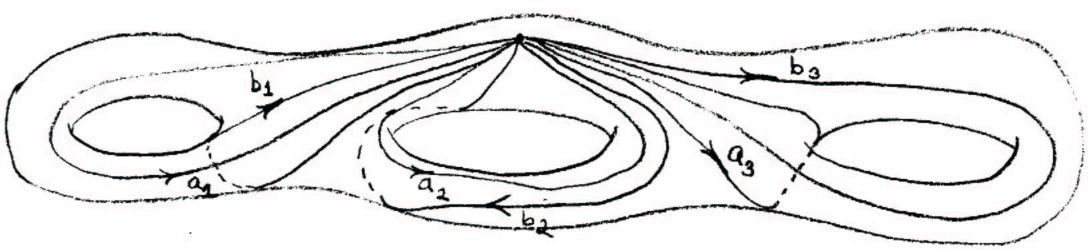
5.2

As a first approximation to the understanding of small deformations we can describe $\pi_1 M$ in terms of a set of generators $\mathcal{G} = \{g_1, \dots, g_n\}$ and a set of relators $\mathcal{R} = \{r_1, \dots, r_l\}$. [Each r_i is a word in the g_i's which equals 1 in $\pi_1 M$.] Any representation $\rho : \pi_1 M \to G$ assigns each generator g_i an element in G, $\rho(g_i)$. This embeds the space of representations R in $G^{\mathcal{G}}$. Since any representation of $\pi_1 M$ must respect the relations in $\pi_1 M$, the image under ρ of a relator r_j must be the identity in G. If $p : G^{\mathcal{G}} \to G^{\mathcal{R}}$ sends a set of elements in G to the $|\mathcal{R}|$ relators written with these elements, then D is just $p^{-1}(1, \dots, 1)$. If p is generic near H_0, (i.e., if the derivative dp is surjective), the implicit function theorem implies that $\mathcal{R}$ is just a manifold of dimension $(|\mathcal{G}| - |\mathcal{R}|) \cdot (\dim G)$. One might reasonably expect this to be the case, provided the generators and relations are chosen in an efficient way. If the action of G on itself by conjugation is effective (as for the group of isometries of hyperbolic space) then generally one would also expect that the action of G on $G^{\mathcal{G}}$ by conjugation, near
H_0, has orbits of the same dimension as G. If so, then the space of deformations of 5.4
M_0 would be a manifold of dimension

$$\dim G \cdot (|\mathcal{G}| - |\mathcal{R}| - 1).$$

5.2

Example. Let's apply the above analysis to the case of hyperbolic structures on closed, oriented two-manifolds of genus at least two. G in this case can be taken to be $\mathrm{PSL}(2, \mathbb{R})$ acting on the upper half-plane by linear fractional transformations. $\pi_1(M_g)$ can be presented with $2g$ generators $a_1, b_1, \dots a_g, b_g$ (see below) together with the single relator $\prod_{i=1}[a_i, b_i]$.



Since $\mathrm{PSL}(2, \mathbb{R})$ is a real three-dimensional Lie group the expected dimension of the deformation space is $3(2g - 1 - 1) = 6g - 6$. This can be made rigorous by showing directly that the derivative of the map $p : G^{\mathcal{G}} \to G^{\mathcal{R}}$ is surjective, but since we will have need for more global information about the deformation space, we won't make the computation here.

5.5

Example. The initial calculation for hyperbolic structures on an oriented three-manifold is less satisfactory. The group of isometries on H^3 preserves planes which, in the upper half-space model, are hemispheres perpendicular to $\mathbb{C} \cup \infty$ (denoted $\hat{\mathbb{C}}$). Thus the group G can be identified with the group of circle preserving maps of $\hat{\mathbb{C}}$. This is the group of all linear fractional transformations with complex coefficients $\mathrm{PSL}(2, \mathbb{C})$. (All transformations are assumed to be orientation preserving). $\mathrm{PSL}(2, \mathbb{C})$, is a *complex* Lie group with real dimensions 6. M^3 can be built from one zero-cell, a number of one- and two-cells, and (if M is closed), one 3-cell.

If M is closed, then $\chi(M) = 0$, so the number k of one-cells equals the number of two-cells. This gives us a presentation of $\pi_1 M$ with k generators and k relators. The expected (real) dimension of the deformation space is $6(k - k - 1) = -6$.

If $\partial M \neq \emptyset$, with all boundary components of positive genus, this estimate of the dimension gives

5.2.1. $$6 \cdot (-\chi(M)) = 3(-\chi(\partial M)).$$

This calculation would tend to indicate that the existence of *any* hyperbolic structure on a closed three-manifold would be unusual. However, subgroups of $\mathrm{PSL}(2, \mathbb{C})$ have many special algebraic properties, so that certain relations can follow from other relations in ways which do not follow in a general group. 5.6

The crude estimate 5.2.1 actually gives some substantive information when $\chi(M) < 0$.

PROPOSITION 5.2.2. *If M^3 possesses a hyperbolic structure M_0, then the space of small deformations of M_0 has dimension at least $6 \cdot (-\chi(M))$.*

PROOF. $\mathrm{PSL}(2,\mathbb{C})^{\mathcal{G}}$ is a complex algebraic variety, and the map

$$p : \mathrm{PSL}(2,\mathbb{C})^{\mathcal{G}} \to \mathrm{PSL}(2,\mathbb{C})^{\mathcal{R}}$$

is a polynomial map (defined by matrix multiplication). Hence the dimension of the subvariety $p = (1, \ldots, 1)$ is at least as great as the number of variables minus the number of defining equations. □

We will later give an improved version of 5.2.2 whenever M has boundary components which are tori.

5.3

In this section we will derive some information about the global structure of the space of hyperbolic structures on a closed, oriented surface M. This space is called the *Teichmüller space* of M and is defined to be the set of hyperbolic structures on M where two are equivalent if there is an isometry homotopic to the identity between them. In order to understand hyperbolic structures on a surface we will cut the surface up into simple pieces, analyze structures on these pieces, and study the ways they can be put together. Before doing this we need some information about closed geodesics in M.

PROPOSITION 5.3.1. *On any closed hyperbolic n-manifold M there is a unique, closed geodesic in any non-trivial free homotopy class.*

PROOF. For any $\alpha \in \pi_1 M$ consider the covering transformation T_α on the universal cover H^n of M. It is an isometry of H^n. Therefore it either fixes some interior point of H^n (elliptic), fixes a point at infinity (parabolic) or acts as a translation on some unique geodesic (hyperbolic). That all isometries of H^2 are of one of these types was proved in Proposition 4.9.3; the proof for H^n is similar.

NOTE. A distinction is often made between "loxodromic" and "hyperbolic" transformations in dimension 3. In this usage a loxodromic transformation means an isometry which is a pure translation along a geodesic followed by a non-trivial twist, while a hyperbolic transformation means a pure translation. This is usually not a useful distinction from the point of view of geometry and topology, so *we will use the term "hyperbolic" to cover either case.*

Since T_α is a covering translation it can't have an interior fixed point so it can't be elliptic. For any parabolic transformation there are points moved arbitrarily small distances. This would imply that there are non-trivial simple closed curves of arbitrarily small length in M. Since M is closed this is impossible. Therefore T_α translates a unique geodesic, which projects to a closed geodesic in M. Two geodesics corresponding to the translations T_α and T'_α project to the same geodesic in M if 5.8
and only if there is a covering translation taking one to the other. In other words, $\alpha' = g\alpha g^{-1}$ for some $g \in \pi_1 M$, or equivalently, α is free homotopic to α. □

Proposition 5.3.2. *Two distinct geodesics in the universal cover H^n of M which are invariant by two covering translations have distinct endpoints at ∞.*

Proof. If two such geodesics had the same endpoint, they would be arbitrarily close near the common endpoint. This would imply the distance between the two closed geodesics is uniformly $\leq \epsilon$ for all ϵ, a contradiction. □

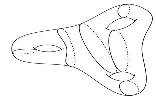
Proposition 5.3.3. *In a hyperbolic two-manifold M^2 a collection of homotopically distinct and disjoint nontrivial simple closed curves is represented by disjoint, simple closed geodesics.*

Proof. Suppose the geodesics corresponding to two disjoint curves intersect. Then there are lifts of the geodesics in the universal cover H^2 which intersect. Since the endpoints are distinct, the pairs of endpoints for the two geodesics must link each other on the circle at infinity. Consider any homotopy of the closed geodesics in M^2. It lifts to a homotopy of the geodesics in H^2. However, no homotopy of the geodesics moving points only a finite hyperbolic distance can move their endpoints; thus the images of the geodesics under such a homotopy will still intersect, and this intersection projects to one in M^2. 5.9

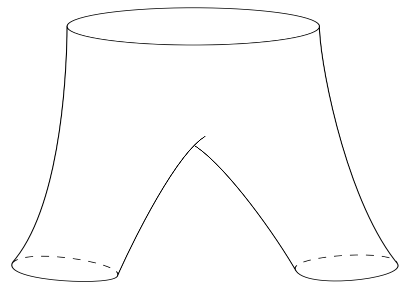
The proof that the closed geodesic corresponding to a simple closed curve is simple is similar. The argument above is applied to two different lifts of the same geodesic. □

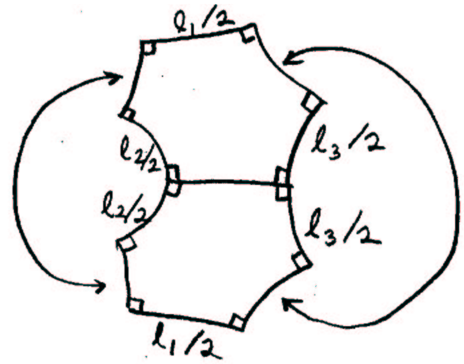
Now we are in a position to describe the Teichmüller space for a closed surface. The coordinates given below are due to Nielsen and Fenchel.

Pick $3g-3$ disjoint, non-parallel simple closed curves on M^2. (This is the maximum number of such curves on a surface of genus g.) Take the corresponding geodesics and cut along them. This divides M^2 into $2g-2$ surfaces homeomorphic to S^2—three disks (called "pairs of pants" from now on) with geodesic boundary.



5.9a





5.10

On each pair of pants P there is a unique arc connecting each pair of boundary components, perpendicular to both. To see this, note that there is a unique homotopy class for each connecting arc. Now double P along the boundary geodesics to form a surface of genus two. The union of the two copies of the arcs connecting a pair of boundary components in P defines a simple closed curve in the closed surface. There is a unique geodesic in its free homotopy class and it is invariant under the reflection which interchanges the two copies of P. Hence it must be perpendicular to the geodesics which were in the boundary of P.

This information leads to an easy computation of the Teichmüller space of P.

PROPOSITION 5.3.4. *$\mathcal{T}(P)$ is homeomorphic to $\mathbb{R}^3$ with coordinates*

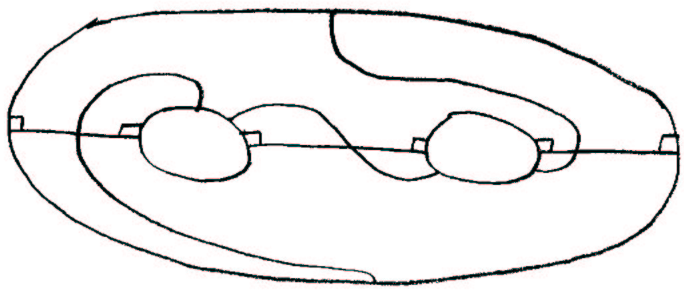
$$(\log l_1, \log l_2, \log l_3),$$

where l_i = length of the i-th boundary component.

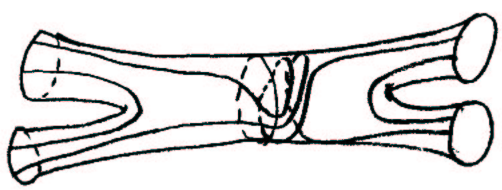
PROOF. The perpendicular arcs between boundary components divide P into two right-angled hexagons. The hyperbolic structure of an all-right hexagon is determined by the lengths of three alternating sides. (See page 2.19.) The lengths of the connecting arcs therefore determine both hexagons so the two hexagons are isometric. Reflection in these arcs is an isometry of the hexagons and shows that the boundary curves are divided in half. The lengths $l_i/2$ determine the hexagons; hence they also determine P. Any positive real values for the l_i are possible so we are done. $\square$ 5.11

In order to determine the hyperbolic structure of the closed two-manifold from that of the pairs of pants, some measurement of the twist with which the boundary geodesics are attached is necessary. Find $3g-3$ more curves in the closed manifold which, together with the first set of curves, divides the surface into hexagons.

In the pairs of pants the geodesics corresponding to these curves are arcs connecting the boundary components. However, they may wrap around the components. In P it is possible to isotope these arcs to the perpendicular connecting arcs discussed above. Let $2d_i$ denote the total number of degrees which this isotopy moves the feet of arcs which lie on the i-th boundary component of p.



5.12



Of course there is another copy of this curve in another pair of pants which has a twisting coefficient d_i'. When the two copies of the geodesic are glued together they cannot be twisted independently by an isotopy of the *closed* surface. Therefore $(d_i - d_i') = \tau_i$ is an isotopy invariant.

Remark. If a hyperbolic surface is cut along a closed geodesic and glued back together with a twist of $2\pi n$ degrees (n an integer), then the resulting surface is isometric to the original one. However, the isometry is not isotopic to the identity so the two surfaces represent distinct points in Teichmüller space. Another way to say this is that they are isometric as surfaces but not as *marked* surfaces. It follows that τ_i is a well-defined real number, not just defined up to integral multiples of 2π.

Theorem 5.3.5. *The Teichmüller space $\mathcal{T}(M)$ of a closed surface of genus g is homeomorphic to $\mathbb{R}^{6g-6}$. There are explicit coordinates for $\mathcal{T}(M)$, namely*

$$(\log l_1, \tau_1, \log l_2, \tau_2, \ldots, \log l_{3g-3}, \tau_{3g-3}),$$

where l_i is the length and τ_i the twist coefficient for a system of $3g-3$ simple closed 5.13
geodesics.

In order to see that it takes precisely $3g-3$ simple closed curves to cut a surface of genus g into pairs of pants P_i notice that $\chi(P_i) = -1$. Therefore the number of P_i's is equal to $-\chi(M_g) = 2g-2$. Each P_i has three curves, but each curve appears in two P_i's. Therefore the number of curves is $\frac{3}{2}(2g-2) = 3g-3$. We can rephrase Theorem 5.3.5 as

$$\mathcal{T}(M) \approx \mathbb{R}^{-3\chi(M)}.$$

It is in this form that the theorem extends to a surface with boundary.

The *Fricke space* $\mathcal{F}(M)$ of a surface M with boundary is defined to be the space of hyperbolic structures on M such that the boundary curves are geodesics, modulo isometries isotopic to the identity. A surface with boundary can also be cut into pairs of pants with geodesic boundary. In this case the curves that were boundary curves in M have no twist parameter. On the other hand these curves appear in only one pair of pants. The following theorem is then immediate from the gluing procedures above.

THEOREM 5.3.6. *$\mathcal{F}(M)$ is homeomorphic to $\mathbb{R}^{-3\chi(M)}$.*

5.14

The definition of Teichmüller space can be extended to general surfaces as the space of all metrics of constant curvature up to isotopy and change of scale. In the case of the torus T^2, this space is the set of all Euclidean structures (i.e., metrics with constant curvature zero) on T^2 with area one. There is still a closed geodesic in each free homotopy class although it is not unique. Take some simple, closed geodesic on T^2 and cut along it. The Euclidean structure on the resulting annulus is completely determined by the length of its boundary geodesic. Again there is a real twist parameter that determines how the annulus is glued to get T^2. Therefore there are two real parameters which determine the flat structures on T^2, the length l of a simple, closed geodesic in a fixed free homotopy class and a twist parameter τ along that geodesic.

THEOREM 5.3.7. *The Teichmüller space of the torus is homeomorphic to $\mathbb{R}^2$ with coordinates $(\log l, \tau)$, where l, τ are as above.*

5.4. Special algebraic properties of groups of isometries of H^3.

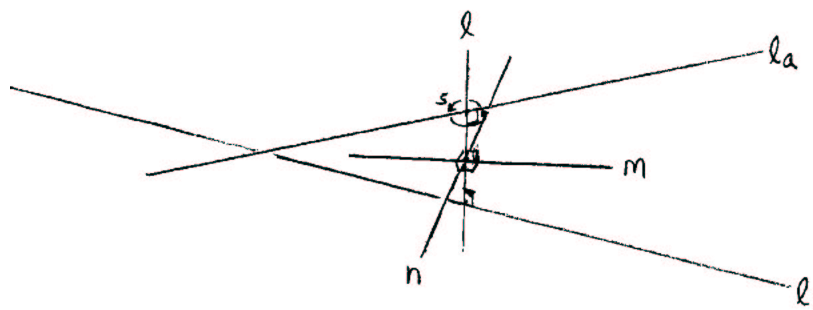
On large open subsets of $\mathrm{PSL}(2,\mathbb{C})^{\mathcal{G}}$, the space of representations of a generating set $\mathcal{G}$ into $\mathrm{PSL}(2,\mathbb{C})$, certain relations imply other relations. This fact was anticipated in the previous section from the computation of the expected dimension of small deformations of hyperbolic structures on closed three manifolds. The phenomenon that dp is not surjective (see 5.3) suggests that, to determine the structure of $\pi_1 M^3$ 5.15
as a discrete subgroup of $\mathrm{PSL}(2,\mathbb{C})$, not all the relations in $\pi_1 M^3$ as an abstract group are needed. Below are some examples.

gJørgensen

Proposition 5.4.1 (Jørgensen). *Let a, b be two isometries of H^3 with no common fixed point at infinity. If $w(a, b)$ is any word such that $w(a, b) = 1$ then $w(a^{-1}, b^{-1}) = 1$. If a and b are conjugate (i.e., if* $\mathrm{Trace}(a) = \pm \mathrm{Trace}(b)$ *in* $\mathrm{PSL}(2, \mathbb{C})$ *) then also $w(b, a) = 1$.*

Proof. If a and b are hyperbolic or elliptic, let l be the unique common perpendicular for the invariant geodesics l_a, l_b of a and b. (If the geodesics intersect in a point x, l is taken to be the geodesic through x perpendicular to the plane spanned by l_a and l_b). If one of a and b is parabolic, (say b is) l should be perpendicular to l_a and pass through b's fixed point at ∞. If both are parabolic, l should connect the two fixed points at infinity. In all cases rotation by 180° in l takes a to a^{-1} and b and b^{-1}, hence the first assertion.

If a and b are conjugate hyperbolic elements of $\mathrm{PSL}(2, \mathbb{C})$ with invariant geodesics l_a and l_b, take the two lines m and n which are perpendicular to l and to each other and which intersect l at the midpoint between g_b and l_a. Also, if g_b is at an angle of θ to l_b along l then m should be at an angle of $\theta/2$ and n at an angle of $\theta + \pi/2$. 5.16



Rotations of 180° through m or n take l_a to l_b and vice versa. Since a and b are conjugate they act the same with respect to their respective fixed geodesics. It follows that the rotations about m and n conjugate a to b (and b to a) or a to b^{-1} (and b to a^{-1}).

If one of a and b is parabolic then they both are, since they are conjugate. In this case take m and n to be perpendicular to l and to each other and to pass through the unique point x on l such that $d(x, ax) = d(x, bx)$. Again rotation by 180° in m and n takes a to b or a to b^{-1}. □

Remarks. 1. This theorem fails when a and b are allowed to have a common fixed point. For example, consider

$$a = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix},$$

where $\lambda \in \mathbb{C}^*$. Then

$$(b^{-k}ab^k)^l = b^{-k}a^lb^k = \begin{bmatrix} 1 & l\lambda^{2k} \\ 0 & 1 \end{bmatrix}.$$

5.17

If λ is chosen so that λ^2 is a root of a polynomial over $\mathbb{Z}$, say $1 + 2\lambda^2 = 0$, then a relation is obtained: in this case

$$w(a,b) = (a)(bab^{-1})^2 = I.$$

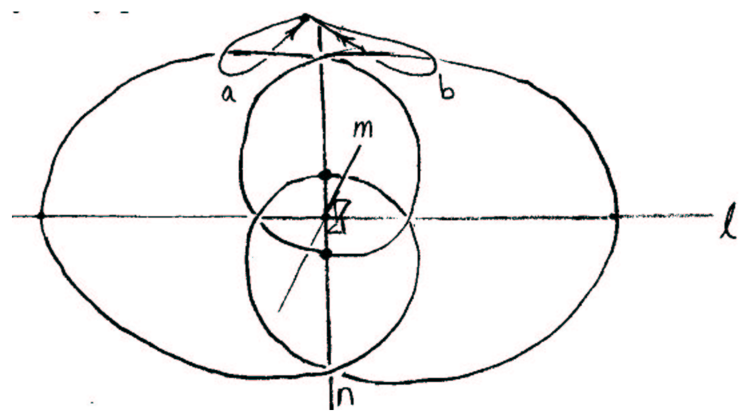
However, $w(a^{-1}, b^{-1}) = I$ only if λ^{-2} is a root of the same polynomial. This is not the case in the current example.

2. The geometric condition that a and b have a common fixed point at infinity implies the algebraic condition that a and b generate a solvable group. (In fact, the commutator subgroup is abelian.)

Geometric Corollary 5.4.2. *Any complete hyperbolic manifold M^3 whose fundamental group is generated by two elements a and b admits an involution s (an isometry of order 2) which takes a to a^{-1} and b to b^{-1}. If the generators are conjugate, there is a $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ action on M generated by s together with an involution t which interchanges a and b unless a and b have a common fixed point at infinity.*

Proof. Apply the rotation of 180° about l to the universal cover H^3. This conjugates the group to itself so it induces an isometry on the quotient space M^3. The same is true for rotation around m and n in the case when a and b are conjugate. It can happen that a and b have a common fixed point x at infinity, but since the 5.18
group is discrete they must both be parabolic. A 180° rotation about any line through x sends a to a^{-1} and b to b^{-1}. There is not generally a symmetry group of order four in this case. □

As an example, the complete hyperbolic structure on the complement of the figure-eight knot has symmetry implied by this corollary. (In fact the group of symmetries extends to S^3 itself, since for homological reasons such a symmetry preserves the meridian direction.)



Here is another illustration of how certain relations in subgroups of $\mathrm{PSL}(2,\mathbb{C})$ can imply others:

Proposition 5.4.3. *Suppose a and b are not elliptic. If $a^n = b^m$ for some $n, m \neq 0$, then a and b commute.*

Proof. If $a^n = b^m$ is hyperbolic, then so are a and b. In fact they fix the same geodesic, acting as translations (perhaps with twists) so they commute. If $a^n = b^m$ 5.19
is parabolic then so are a and b. They must fix the same point at infinity so they act as Euclidean transformations of any horosphere based there. It follows that a and b commute. □

Proposition 5.4.3. *If a is hyperbolic and a^k is conjugate to a^l then $k = \pm l$.*

Proof. Since translation distance along the fixed line is a conjugacy invariant and $\rho(a^k) = \pm k\rho(a)$ (where $\rho(\ )$ denotes translation distance), the proposition is easy to see. □

Finally, along the same vein, it is sometimes possible to derive some nontrivial topological information about a hyperbolic three-manifold from its fundamental group.

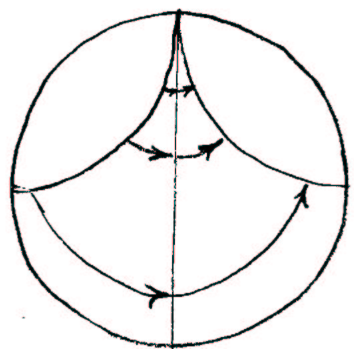
Proposition 5.4.4. *If M^3 is a complete, hyperbolic three-manifold, $a, b \in \pi_1 M^3$ and $[a, b] = 1$, then either*

(i) *a and b belong to an infinite cyclic subgroup generated by x and $x^l = a$, $x^k = b$, or*
(ii) *M has an end, E, homeomorphic to $T^2 \times [0, \infty)$ such that the group generated by a and b is conjugate in $\pi_1 M^3$ to a subgroup of finite index in $\pi_1 E$.*

Proof. If a and b are hyperbolic then they translate the same geodesic. Since $\pi_1 M^3$ acts as a discrete group on H^3, a and b must act discretely on the fixed geodesic. 5.20

Thus, (i) holds.

If a and b are not both hyperbolic, they must both be parabolic, since they commute. Therefore they can be thought of as Euclidean transformations on a set of horospheres. If the translation vectors are not linearly independent, a and b generate a group of translations of $\mathbb{R}$ and (i) is again true. If the vectors are linearly independent, a and b generate a lattice group $L_{a,b}$ on $\mathbb{R}^2$. Moreover as one approaches the fixed point at infinity, the hyperbolic distance a point x is moved by a and b goes to zero.



Recall that the subgroup $G_\epsilon(X)$ of $\pi_1 M^3$ generated by transformations that moves a point x less than ϵ is abelian. (See pages 4.34-4.35). Therefore all the elements of $G_\epsilon(X)$ commute with a and b and fix the same point p at infinity. By discreteness $G_\epsilon(X)$ acts as a lattice group on the horosphere through x and contains $L_{a,b}$ as a subgroup of finite index. 5.21

Consider a fundamental domain of $G_\epsilon(X)$ acting on the set of horocycles at p which are "contained" in the horocycle H_x through x. It is homeomorphic to the product of a fundamental domain of the lattice group acting on H_x with $[0,\infty)$ and is moved away from itself by all elements in $\pi_1 M^3$ not in $G_\epsilon(X)$. Therefore it is projected down into M^3 as an end homeomorphic to $T^2 \times [0,1]$. This is case (ii). □

5.22

5.5. The dimension of the deformation space of a hyperbolic three-manifold.

Consider a hyperbolic structure M_0 on $T^2 \times I$. Let α and β be generators for $\mathbb{Z} \oplus \mathbb{Z} = \pi_1(T^2 \times I)$; they satisfy the relation $[\alpha, \beta] = 1$, or equivalently $\alpha\beta = \beta\alpha$. The representation space for $\mathbb{Z} \oplus \mathbb{Z}$ is defined by the equation

$$H(\alpha)\, H(\beta) = H(\beta)\, H(\alpha),$$

where $H(\alpha), H(\beta) \in \mathrm{PSL}(2,\mathbb{C})$. But we have the identity

$$\mathrm{Tr}(H(\alpha)\, H(\beta)) = \mathrm{Tr}(H(\beta)\, H(\alpha)),$$

as well as $\det(H(\alpha)\,H(\beta)) = \det(H(\beta)\,H(\alpha)) = 1$, so this matrix equation is equivalent to two ordinary equations, at least in a neighborhood of a particular non-trivial solution. Consequently, the solution space has a complex dimension four, and the deformation space of M_0 has complex dimension two. This can easily be seen directly: $H(\alpha)$ has one complex degree of freedom to conjugacy, and given $H(\alpha) \neq \mathrm{id}$, there is a one complex-parameter family of transformations $H(\beta)$ commuting with it. This example shows that 5.2.2 is not sharp. More generally, we will improve 5.2.2 for any compact oriented hyperbolic three-manifold M_0 whose boundary contains toruses, under a mild nondegeneracy condition on the holonomy of M_0:

THEOREM 5.6. *Let M_0 be a compact oriented hyperbolic three-manifold whose holonomy satisfies*

(a) *the holonomy around any component of ∂M homeomorphic with T^2 is not trivial, and*
(b) *the holonomy has no fixed point on the sphere at ∞.*

Under these hypotheses, the space of small deformations of M_0 has dimension at least as great as the total dimension of the Teichmüller space of ∂M, that is,

$$\dim_{\mathbb{C}}(\mathrm{Def}(M)) \geq \sum_i \begin{cases} +3\big|\chi((\partial M)_i)\big| & \text{if } \chi((\partial M)_i) < 0, \\ 1 & \text{if } \chi((\partial M)_i) = 0, \\ 0 & \text{if } \chi((\partial M)_i) > 0. \end{cases}$$

REMARK. Condition (b) is equivalent to the statement that the holonomy representation in $\mathrm{PSL}(2,\mathbb{C})$ is irreducible. It is also equivalent to the condition that the holonomy group (the image of the holonomy) be solvable.

EXAMPLES. If N is any surface with nonempty boundary then, by the immersion theorem [Hirsch] there is an immersion ϕ of $N \times S^1$ in $N \times I$ so that ϕ sends $\pi_1(N)$ to $\pi_1(N \times I) = \pi_1(N)$ by the identity map. Any hyperbolic structure on $N \times I$ has a $-6\chi(N)$ complex parameter family of deformations. This induces a $(-6\chi(N))$-parameter family of deformations of hyperbolic structures on $N \times S^1$, showing that the inequality of 5.6 is not sharp in general.

Another example is supplied by the complement M_k of k unknotted unlinked solid tori in S^3. Since $\pi_1(M_k)$ is a free group on k generators, every hyperbolic structure on M_k has at least $3k-3$ degrees of freedom, while 5.6 guarantees only k degrees of freedom. Other examples are obtained on more interesting manifolds by considering hyperbolic structures whose holonomy factors through a free group.

PROOF OF 5.6. We will actually prove that for any compact oriented manifold M, the complex dimension of the representation space of $\pi_1 M$, near a representation satisfying (a) and (b), is at least 3 greater than the number given in 5.6; this suffices, by 5.1. For this stronger assertion, we need only consider manifolds which have no boundary component homeomorphic to a sphere, since any three-manifold M has the same fundamental group as the manifold $\hat{M}$ obtained by gluing a copy of D^3 to each spherical boundary component of M.

Remark. Actually, it can be shown that when $\partial M \neq 0$, a representation

$$\rho : \pi_1 M \to \mathrm{PSL}(2, \mathbb{C})$$

is the holonomy of some hyperbolic structure for M if and only if it lifts to a representation in $\mathrm{SL}(2, \mathbb{C})$. (The obstruction to lifting is the second Stiefel–Whitney class ω_2 of the associated H^3-bundle over M.) It follows that if H_0 is the holonomy of a hyperbolic structure on M, it is also the holonomy of a hyperbolic structure on $\hat{M}$, provided $\partial \hat{M} \neq \emptyset$. Since we are mainly concerned with structures which have more geometric significance, we will not discuss this further.

Let H_0 denote any representation of $\pi_1 M$ satisfying (a) and (b) of 5.6. Let $T_1, \ldots, T_k$ be the components of ∂M which are toruses.

Lemma 5.6.1. *For each i, $1 \leq i \leq k$, there is an element $\alpha_i \in \pi_1(M)$ such that the group generated by $H_0(\alpha_i)$ and $H_0(\pi_1(T_i))$ has no fixed point at ∞. One can choose α_i so $H_0(\alpha_i)$ is not parabolic.*

Proof of 5.6.1. If $H_0(\pi_1 T_i)$ is parabolic, it has a unique fixed point x at ∞ and the existence of an α_i' not fixing x is immediate from condition (b). If $H_0(\pi_1 T_i)$ has two fixed points x_1 and x_2, there is $H_0(\beta_1)$ not fixing x_1 and $H_0(\beta_2)$ not fixing x_2. If $H_0(\beta_1)$ and $H_0(\beta_2)$ each have common fixed points with $H_0(\pi_1 T_i)$, $\alpha_1' = \beta_1 \beta_2$ does not.

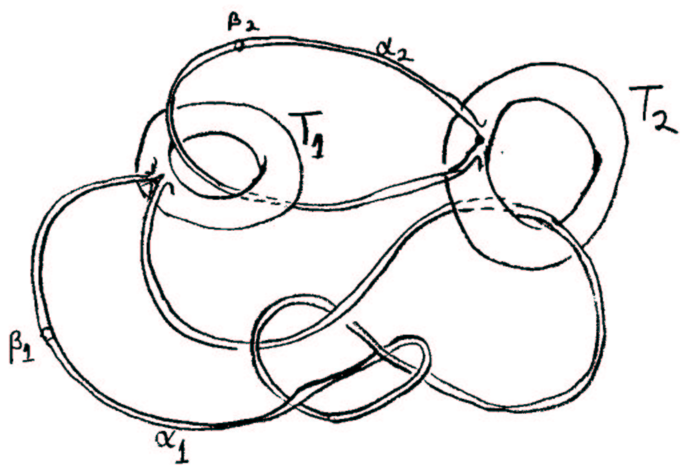
If $H_0(\alpha_i')$ is parabolic, consider the commutators $\gamma_n = [\alpha_i'^n, \beta]$ where $\beta \in \pi_1 T_i$ is some element such that $H_0(\beta) \neq 1$. If $H_0[\alpha_i'^n, \beta]$ has a common fixed point x with $H_0(\beta)$ then also $\alpha_i'^n \beta \alpha_i'^{-n}$ fixes x so β fixes $\alpha_i'^{-n} x$; this happens for at most three values of n. We can, after conjugation, take $H_0(\alpha_i') = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$. Write

$$H_0(\beta \alpha_i'^{-1} \beta^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where $a + d = 2$ and $c \neq 0$ since $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ is not an eigenvector of β. We compute $\mathrm{Tr}(\gamma_n) = 2 + n^2 c$; it follows that γ_n can be parabolic ($\Leftrightarrow \mathrm{Tr}(\gamma_n) = \pm 2$) for at most 3 values of n. This concludes the proof of Lemma 5.6.1. □

Let $\{\alpha_i, 1 \leq i \leq k\}$ be a collection of simple disjoint curves based on T_i and 5.26
representing the homotopy classes of the same names. Let $N \subset M$ be the manifold obtained by hollowing out nice neighborhoods of the α_i. Each boundary component

of N is a surface of genus ≥ 2, and M is obtained by attaching k two-handles along non-separating curves on genus two surfaces $S_1, \ldots, S_k \subset \partial N$.



Let α_i also be represented by a curve of the same name on S_i, and let β_i be a curve on S_i describing the attaching map for the i-th two-handle. Generators γ_i, δ_i can be chosen for $\pi_1 T_i$ so that $\alpha_i, \beta_i, \gamma_i$, and δ_i generate $\pi_1 B_i$ and $[\alpha_i, \beta_i] \cdot [\gamma_i, \delta_i] = 1$. $\pi_1 M$ is obtained from $\pi_1 M$ by adding the relations $\beta_i = 1$.

LEMMA 5.6.2. *A representation ρ of $\pi_1 N$ near H_0 gives a representation of $\pi_1 M$ if and only if the equations* 5.27

$$\begin{aligned} \operatorname{Tr}\left(\rho\left(\beta_i\right)\right) &= 2 \\ \textit{and } \operatorname{Tr}\left(\rho\left[\alpha_i, \beta_i\right]\right) &= 2 \end{aligned}$$

are satisfied.

PROOF OF 5.6.2. Certainly if ρ gives a representation of $\pi_1 M$, then $\rho(\beta_i)$ and $\rho[\alpha_i, \beta_i]$ are the identity, so they have trace 2.

To prove the converse, consider the equation

$$\operatorname{Tr}[A, B] = 2$$

in $\mathrm{SL}(2, \mathbb{C})$. If A is diagonalizable, conjugate so that

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix}.$$

Write

$$BA^{-1}B^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We have the equations

$$a + d = \lambda + \lambda^{-1}$$

$$\mathrm{Tr}\,[A,B] = \lambda a + \lambda^{-1} d = 2$$

which imply that

$$a = \lambda^{-1},\ d = \lambda.$$

Since $ad - bc = 1$ we have $bc = 0$. This means B has at least one common eigenvector $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ or $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ with A; if $[A,B] \neq 1$, this common eigenvector is the unique eigenvector 5.28
of $[A,B]$ (up to scalars). As in the proof of 5.6.1, a similar statement holds if A is parabolic. (Observe that $[A,B] = [-A,B]$, so the sign of $\mathrm{Tr}\,A$ is not important).

It follows that if $\mathrm{Tr}\ \rho[\alpha_i,\beta_i] = 2$, then since $[\gamma_i,\delta_i] = [\alpha_i,\beta_i]$, either $\rho\,(\alpha_i)$, $\rho\,(\beta_i)$, $\rho\,(\gamma_i)$ and $\rho\,(\delta_i)$ all have a common fixed point on the sphere at infinity, or $\rho[\alpha_i,\beta_i] = 1$.

By construction $H_0, \pi_1 S_i$ has no fixed point at infinity, so for ρ near $H_0 \rho\, \pi_1 S_i$ cannot have a fixed point either; hence $\rho[\alpha_i,\beta_i] = 1$.

The equation $\mathrm{Tr}\ \rho\,(\beta_i) = 2$ implies $\rho\,(\beta_i)$ is parabolic; but it commutes with $\rho\,(\beta_i)$, which is hyperbolic for ρ near H_0. Hence $\rho\,(\beta_i) = 1$. This concludes the proof of Lemma 5.6.2. □

To conclude the proof of 5.6, we consider a handle structure for N with one zero-handle, m one-handles, p two-handles and no three-handles (provided $\partial M \neq \emptyset$). This gives a presentation for $\pi_1 N$ with m generators and p relations, where

$$1 - m + p = \chi(N) = \chi(M) - k.$$

The representation space $R \subset \mathrm{PSL}(2,\mathbb{C})^m$ for $\pi_1 M$, in a neighborhood of H_0, is defined by the p matrix equations

$$r_i = 1, \quad (1 \le i \le p),$$

where the r_i are products representing the relators, together with $2k$ equations 5.29

$$\mathrm{Tr}\,\rho(\beta_i) = 2$$

$$\mathrm{Tr}\,\rho([\alpha_i,\beta_i]) = 2 \quad [1 \le i \le k]$$

The number of equations minus the number of unknowns (where a matrix variable is counted as three complex variables) is

$$3m - 3p - 2k = -3\chi(M) + k + 3.$$

□

Remark. If M is a closed hyperbolic manifold, this proof gives the estimate of 0 for $\dim_{\mathbb{C}} \mathrm{def}(M)$: simply remove a non-trivial solid torus from M, apply 5.6, and fill in the solid torus by an equation $\mathrm{Tr}(\gamma) = 2$.

5.7

There is a remarkable, precise description for the global deformation space of hyperbolic structures on closed manifolds in dimensions bigger than two:

THEOREM 5.7.1 (Mostow's Theorem [algebraic version]). *Suppose* Γ_1 *and* Γ_2 *are two discrete subgroups of the group of isometries of* H^n, $n \geq 3$ *such that* H^n/Γ_i *has finite volume and suppose* $\phi : \Gamma_1 \to \Gamma_2$ *is a group isomorphism. Then* Γ_1 *and* Γ_2 *are conjugate subgroups.*

This theorem can be restated in terms of hyperbolic manifolds since a hyperbolic manifold has universal cover H^n with fundamental group acting as a discrete group of isometries. 5.30

THEOREM 5.7.2 (Mostow's Theorem [geometric version]). *If* M_1^n *and* M_2^n *are complete hyperbolic manifolds with finite total volume, any isomorphism of fundamental groups* $\phi : \pi_1 M_1 \to \pi_1 M_2$ *is realized by a unique isometry.*

REMARK. Multiplication by an element in either fundamental group induces the identity map on the manifolds themselves so that ϕ needs only to be defined up to composition with inner automorphisms to determine the isometry from M_1 to M_2.

Since the universal cover of a hyperbolic manifold is H^n, it is a $K(\pi, 1)$. Two such manifolds are homotopy equivalent if and only if there is an isomorphism between their fundamental groups.

COROLLARY 5.7.3. *If* M_1 *and* M_2 *are hyperbolic manifolds which are complete with finite volume, then they are homeomorphic if and only if they are homotopy equivalent.* (*The case of dimension two is well known.*)

For any manifold M, there is a homomorphism $\operatorname{Diff} M \to \operatorname{Out}(\pi_1 M)$, where $\operatorname{Out}(\pi_1 M) = \operatorname{Aut}(\pi_1 M)/\operatorname{Inn}(\pi_1 M)$ is the group of outer automorphisms. Mostow's Theorem implies this homomorphism splits, if M is a hyperbolic manifold of dimension $n \geq 3$. It is unknown whether the homomorphism splits when M is a surface. When $n = 2$ the kernel $\operatorname{Diff}_0(M)$ is contractible, provided $\chi(M) \leq 0$. If M is a Haken 5.31 three-manifold which is not a Seifert fiber space, Hatcher has shown that $\operatorname{Diff}_0 M$ is contractible.

COROLLARY 5.7.4. *If* M^n *is hyperbolic* (*complete, with finite total volume*) *and* $n \geq 3$, *then* $\operatorname{Out}(\pi_1 M)$ *is a finite group, isomorphic to the group of isometries of* M^n.

PROOF. By Mostow's Theorem any automorphism of $\pi_1 M$ induces a unique isometry of M. Since any inner automorphism induces the identity on M, it follows that the group of isometries is isomorphic to $\operatorname{Out}(\pi_1 M)$. That $\operatorname{Out}(\pi_1 M)$ is finite is immediate from the fact that the group of isometries, $\operatorname{Isom}(M^n)$, is finite.

To see that $\mathrm{Isom}(M^n)$ is finite, choose a base point and frame at that point and suppose first that M is compact. Any isometry is completely determined by the image of this frame (essentially by "analytic continuation"). If there were an infinite sequence of isometries there would exist two image frames close to each other. Since M is compact, the isometries, ϕ_1, ϕ_2, corresponding to these frames would be close on all of M. Therefore ϕ, is homotopic to ϕ_2. Since the isometry $\phi_2^{-1}\phi_1$ induces the trivial outer automorphism on $\pi_1 M$, it is the identity; i.e., $\phi_2 = \phi_1$.

If M is not compact, consider the submanifold $M_\epsilon \subset M$ which consists of points which are contained in an embedded hyperbolic disk of radius ϵ. Since M has finite total volume, M_ϵ is compact. Moreover, it is taken to itself under any isometry. The 5.32
argument above applied to M_ϵ implies that the group of isometries of M is finite even in the non-compact case. □

Remark. This result contrasts with the case $n = 2$ where $\mathrm{Out}(\pi_1 M^2)$ is infinite and quite interesting.

The proof of Mostow's Theorem in the case that H^n/Γ is not compact was completed by Prasad. Otherwise, 5.7.1 and 5.7.2 (as well as generalizations to other homogeneous spaces) are proved in Mostow. We shall discuss Mostow's proof of this theorem in 5.10, giving details as far as they can be made geometric. Later, we will give another proof due to Gromov, valid at least for $n = 3$.

5.8. Generalized Dehn surgery and hyperbolic structures.

Let M be a non-compact, hyperbolic three-manifold, and suppose that M has a finite number of ends $E_1, \ldots, E_k$, each homeomorphic to $T^2 \times [0, \infty)$ and isometric to the quotient space of the region in H^3 (in the upper half-space model) above an interior Euclidean plane by a group generated by two parabolic transformations which fix the point at infinity. Topologically M is the interior of a compact manifold $\bar{M}$ whose boundary is a union of $T_1, \ldots, T_k$ tori.

Recall the operation of generalized Dehn surgery on M (§4.5); it is parametrized 5.33
by an ordered pair of real numbers (a_i, b_i) for each end which describes how to glue a solid torus to each boundary component. If nothing is glued in, this is denoted by ∞ so that the parameters can be thought of as belonging to S^2 (i.e., the one point compactification of $\mathbb{R}^2 \approx H_1(T^2, \mathbb{R})$). The resulting space is denoted by $M_{d_1,\ldots,d_k}$ where $d_i = (a_i, b_i)$ or ∞.

In this section we see that the new spaces often admit hyperbolic structures. Since $M_{d_1,\ldots,d_k}$ is a closed manifold when $d_i = (a_i, b_i)$ are primitive elements of $H_1(T^2, \mathbb{Z})$, this produces many closed hyperbolic manifolds. First it is necessary to see that small deformations of the complete structure on M induce a hyperbolic structure on *some* space $M_{d_1,\ldots,d_k}$.

LEMMA 5.8.1. *Any small deformation of a "standard" hyperbolic structure on* $T^2 \times [0,1]$ *extends to some* $(D^2 \times S^1)_d$. $d = (a,b)$ *is determined up to sign by the traces of the matrices representing generators* α, β *of* $\pi_1 T^2$.

PROOF. A "standard" structure on $T^2 \times [0,1]$ means a structure as described on an end of M truncated by a Euclidean plane. The universal cover of $T^2 \times [0,1]$ is the region between two horizontal Euclidean planes (or horospheres), modulo a group of translations. If the structure is deformed slightly the holonomy determines the new structure and the images of α and β under the holonomy map H are slightly 5.34
perturbed.

If $H(\alpha)$ is still parabolic then so is $H(\beta)$ and the structure is equivalent to the standard one. Otherwise $H(\alpha)$ and $H(\beta)$ have a common axis l in H^3. Moreover since $H(\alpha)$ and $H(\beta)$ are close to the original parabolic elements, the endpoints of l are near the common fixed point of the parabolic elements. If $T^2 \times [0,1]$ is thought to be embedded in the end, $T^2 \times [0,\infty)$, this means that the line lies far out towards ∞ and does not intersect $T^2 \times [0,1]$. Thus the developing image of $T^2 \times [0,1]$ in H^3 for new structure misses l and can be lifted to the universal cover

$$\widetilde{H^3 - l}$$

of $H^3 - l$.

This is the geometric situation necessary for generalized Dehn surgery. The extension to $(D^2 \times S^1)_d$ is just the completion of

$$\widetilde{H^3 - l}/\{\tilde{H}(\alpha), \tilde{H}(\beta)\}$$

where $\tilde{H}$ is the lift of H to the cover

$$\widetilde{H^3 - l}.$$

Recall that the completion depends only on the behavior of $\widetilde{H(\alpha)}$ and $\widetilde{H(\beta)}$ along l. In particular, if $\tilde{H}(\ )$ denotes the complex number determined by the pair (translation distance along l, rotation about l), then the Dehn surgery coefficients $d = (a,b)$ are determined by the formula:

$$a\,\tilde{H}(\alpha) + b\,\tilde{H}(\beta) = \pm 2\pi i.$$

The translation distance and amount of rotation of an isometry along its fixed line is determined by the trace of its matrix in $\mathrm{PSL}(2,\mathbb{C})$. This is easy to see since trace is a conjugacy invariant and the fact is clearly true for a diagonal matrix. In 5.35
particular the complex number corresponding to the holonomy of a matrix acting on H^3 is $\log \lambda$ where $\lambda + \lambda^{-1}$ is its trace. □

The main result concerning deformations of M is

THEOREM 5.8.2. *If* $M = M_{\infty,\dots,\infty}$ *admits a hyperbolic structure then there is a neighborhood* U *of* $(\infty,\dots,\infty)$ *in* $S^2 \times S^2 \times \dots \times S^2$ *such that for all* $(d_1,\dots,dk) \in U, M_{d_1,\dots,d_k}$ *admits a hyperbolic structure.*

PROOF. Consider the compact submanifold $M_0 \subset M$ gotten by truncating each end. M_0 has boundary a union of k tori and is homeomorphic to the manifold $\bar{M}$ such that $M =$ interior $\bar{M}$. By theorem 5.6, M_0 has a k complex parameter family of non-trivial deformations, one for each torus. From the lemma above, each small deformation gives a hyperbolic structure on some $M_{d_1,\dots,d_k}$. It remains to show that the d_i vary over a neighborhood of $(\infty,\dots,\infty)$.

Consider the function

$$\mathrm{Tr} : \mathrm{Def}(M) \to (\mathrm{Tr}(H(\alpha_1)),\dots,\mathrm{Tr}(H(\alpha_k)))$$

which sends a point in the deformation space to the k-tuple of traces of the holonomy of $\alpha_1,\alpha_2,\dots,\alpha_k$, where α_i,β_i generate the fundamental group of the i-th torus. Tr is a holomorphic (in fact, algebraic) function on the algebraic variety 5.36 $\mathrm{Def}(M)$. $\mathrm{Tr}(M_{\infty,\dots,\infty}) = (\pm 2,\dots,\pm 2)$ for some fixed choice of signs. Note that $\mathrm{Tr}(H(\alpha_i)) = \pm 2$ if and only if $H(\alpha_i)$ is parabolic and $H(\alpha_i)$ is parabolic if and only if the i-th surgery coefficient d_i equals ∞. By Mostow's Theorem the hyperbolic structure on $M_{\infty,\dots,\infty}$ is unique. Therefore $d_i = \infty$ for $i = l,\dots,k$ only in the original case and $\mathrm{Tr}^{-1}(\pm 2,\dots,\pm 2)$ consists of exactly one point. Since $\dim(\mathrm{Def}(M)) \geq k$ it follows from [ ] that the image under Tr of a small open neighborhood of $M_{\infty,\dots,\infty}$ is an open neighborhood of $(\pm 2,\dots,\pm 2)$.

Since the surgery coefficients of the i-th torus depend on the trace of both $H(\alpha_i)$ and $H(\beta_i)$, it is necessary to estimate $H(\beta_i)$ in terms of $H(\alpha_i)$ in order to see how the surgery coefficients vary. Restrict attention to one torus T and conjugate the original developing image of $M_{\infty,\dots,\infty}$ so that the parabolic fixed point of the holonomy, $H_0,(\pi_1 T)$, is the point at infinity. By further conjugation it is possible to put the holonomy matrices of the generators α,β of $\pi_1 T$ in the following form:

$$H_0(\alpha) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad H_0(\beta) = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}.$$

Note that since $H_0(\alpha), H_0(\beta)$ act on the horospheres about ∞ as a two-dimensional lattice of Euclidean translations, c and l are linearly independent over $\mathbb{R}$. Since 5.37 $H_0(\alpha), H_0(\beta)$ have $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ as an eigenvector, the perturbed holonomy matrices

$$H(\alpha), H(\beta)$$

will have common eigenvectors near $\begin{bmatrix}1\\0\end{bmatrix}$, say $\begin{bmatrix}1\\\epsilon_1\end{bmatrix}$ and $\begin{bmatrix}1\\\epsilon_2\end{bmatrix}$. Let the eigenvalues of $H(\alpha)$ and $H(\beta)$ be (λ, λ^{-1}) and (μ, μ^{-1}) respectively. Since $H(\alpha)$ is near $H_0(\alpha)$,

$$H(\alpha)\begin{bmatrix}0\\1\end{bmatrix} \approx \begin{bmatrix}1\\1\end{bmatrix}.$$

However

$$H(\alpha)\begin{bmatrix}0\\1\end{bmatrix} = \frac{1}{\epsilon_1-\epsilon_2}\,H(\alpha)\left(\begin{bmatrix}1\\\epsilon_1\end{bmatrix} - \begin{bmatrix}1\\\epsilon_2\end{bmatrix}\right) = \frac{1}{\epsilon_1-\epsilon_2}\left(\begin{bmatrix}\lambda\\\lambda\epsilon_1\end{bmatrix} - \begin{bmatrix}\lambda^{-1}\\\lambda^{-1}\epsilon_2\end{bmatrix}\right).$$

Therefore

$$\frac{\lambda-\lambda^{-1}}{\epsilon_1-\epsilon_2} \approx 1.$$

Similarly,

$$\frac{\mu-\mu^{-1}}{\epsilon_1-\epsilon_2} \approx c.$$

For λ, μ near l,

$$\frac{\log(\lambda)}{\log(\mu)} \approx \frac{\lambda-1}{\mu-1} \approx \frac{\lambda-\lambda^{-1}}{\mu-\mu^{-1}} \approx \frac{1}{c}.$$

Since $\tilde{H}(\alpha) = \log\lambda$ and $\tilde{H}(\beta) = \log\mu$ this is the desired relationship between $\tilde{H}(\alpha)$ and $\tilde{H}(\beta)$.

The surgery coefficients (a, b) are determined by the formula 5.38

$$a\tilde{H}(\alpha) + b\tilde{H}(\beta) = \pm 2\pi i.$$

From the above estimates this implies that

$$(a+bc) \approx \frac{\pm 2\pi i}{\log\lambda}.$$

(Note that the choice of sign corresponds to a choice of λ or λ^{-1}.) Since 1 and c are linearly independent over $\mathbb{R}$, the values of (a, b) vary over an open neighborhood of ∞ as λ varies over a neighborhood of 1. Since $\mathrm{Tr}(H(\alpha)) = \lambda + \lambda^{-1}$ varies over a neighborhood of 2 (up to sign) in the image of $\mathrm{Tr} : \mathrm{Def}(M) \to \mathbb{C}^k$, we have shown that the surgery coefficients for the $M_{d_1,\dots,d_k}$ possessing hyperbolic structures vary over an open neighborhood of ∞ in each component. □

Example. The complement of the Borromean rings has a complete hyperbolic structure. However, if the trivial surgery with coefficients $(1, 0)$ is performed on one component, the others are unlinked. (In other words, $M_{(1,0),\infty,\infty}$ is S^3 minus two unlinked circles.) The manifold $M_{(1,0),x,y}$ (where M is S^3 minus the Borromean rings) is then a connected sum of lens spaces if x, y are primitive elements of $H_1(T_i^2, \mathbb{Z})$ so it cannot have a hyperbolic structure. Thus it may often happen that an infinite number of non-hyperbolic manifolds can be obtained by surgery from a hyperbolic one. However, the theorem does imply that if a finite number of integral pairs of coefficients is excluded from *each boundary component*, then all remaining three-manifolds obtained by Dehn surgery on M are also hyperbolic.

5.39

5.9. A Proof of Mostow's Theorem.

This section is devoted to a proof of Mostow's Theorem for closed hyperbolic n-manifolds, $n \geq 3$. The proof will be sketchy where it seems to require analysis. With a knowledge of the structure of the ends in the noncompact, complete case, this proof extends to the case of a manifold of finite total volume; we omit details. The outline of this proof is Mostow's.

Given two closed hyperbolic manifolds M_1 and M_2, together with an isomorphism of their fundamental groups, there is a homotopy equivalence inducing the isomorphism since M_1 and M_2 are $K(\pi, 1)$'s. In other words, there are maps $f_1 : M_1 \to M_2$ and $f_2 : M_2 \to M_1$ such that $f_1 \circ f_2$ and $f_2 \circ f_1$ are homotopic to the identity. Denote lifts of f_1, f_2 to the universal cover H^n by $\tilde{f}_1, \tilde{f}_2$ and assume $\tilde{f}_1 \circ \tilde{f}_2$ and $\tilde{f}_2 \circ \tilde{f}_1$ are equivariantly homotopic to the identity.

The first step in the proof is to construct a correspondence between the spheres at infinity of H^n which extends $\tilde{f}_1$ and $\tilde{f}_2$.

Definition. A map $g : X \to Y$ between metric spaces is a *pseudo-isometry* if there are constants c_1, c_2 such that $c_1^{-1} d(x_1, x_2) - c_2 \leq d(gx_1, gx_2) \leq c_1 d(x_1, x_2)$ for all $x_1, x_2 \in X$.

Lemma 5.9.1. *$\tilde{f}_1, \tilde{f}_2$ can be chosen to be pseudo-isometries.*

Proof. Make f_1 and f_2 simplicial. Then since M_1 and M_2 are compact, f_1 and 5.40
f_2 are Lipschitz and lift to $\tilde{f}_1$ and $\tilde{f}_2$ which are Lipschitz with the same coefficient. It follows immediately that there is a constant c_1 so that $d(\tilde{f}_i x_1, \tilde{f}_i x_2) \leq c_1 d(x_1, x_2)$ for $i = 1, 2$ and all $x_1, x_2 \in H^n$.

If $x_i = \tilde{f}_1 y_i$, then this inequality implies that

$$d(\tilde{f}_2 \circ \tilde{f}_1(y_1),\, \tilde{f}_2 \circ \tilde{f}_1(y_2)) \leq c_1 d(\tilde{f}_1 y_1, \tilde{f}_1 y_2).$$

However, since M_1 is compact, $\tilde{f}_2 \circ \tilde{f}_1$ is homotopic to the identity by a homotopy that moves every point a distance less than some constant b. It follows that

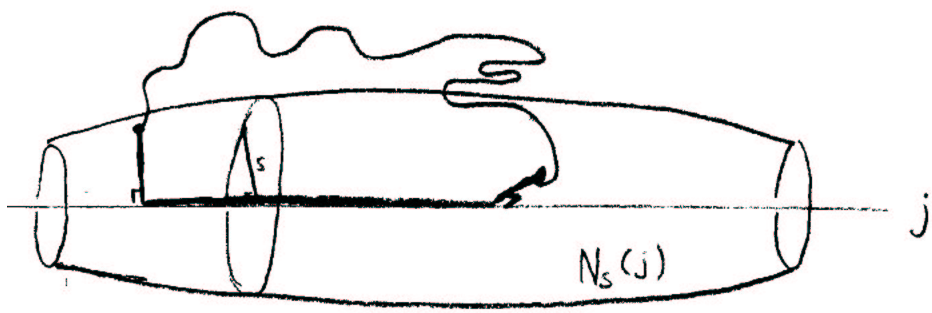
$$d(y_1, y_2) - 2b \leq d(\tilde{f}_2 \circ \tilde{f}_1 y_1, \tilde{f}_2 \circ \tilde{f}_1 y_2),$$

from which the lower bound $c_1^{-1} d(y_1, y_2) - c_2 \leq d(\tilde{f}_1 y_1, \tilde{f}_1 y_2)$ follows. □

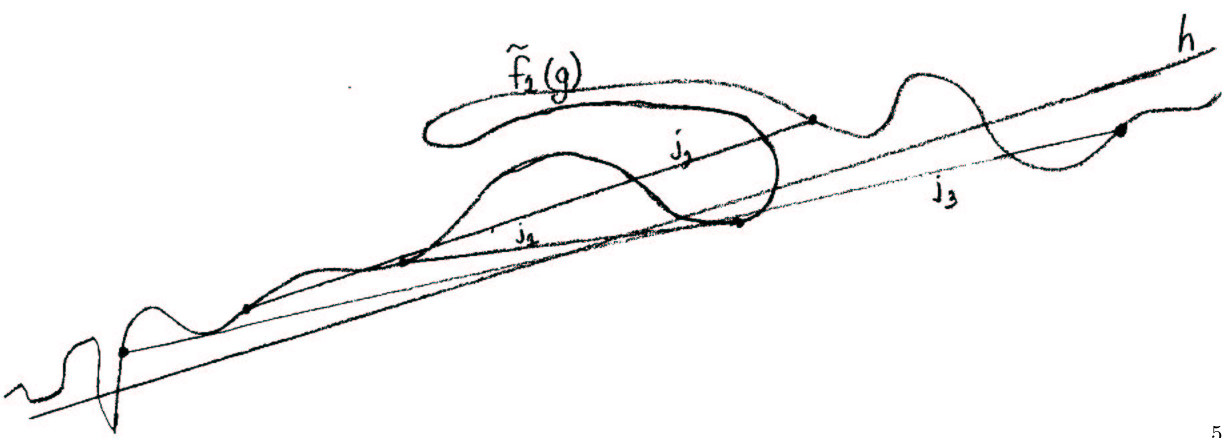
Using this lemma it is possible to associate a unique geodesic with the image of a geodesic.

Proposition 5.9.2. *For any geodesic $g \subset H^n$ there is a unique geodesic h such that $f_1(g)$ stays in a bounded neighborhood of h.*

Proof. If j is any geodesic in H^n, let $N_s(j)$ be the neighborhood of radius s about j. We will see first that if s is large enough there is an upper bound to the length of any bounded component of $g - \left(\tilde{f}_1^{-1}(N_s(j))\right)$, for any j. In fact, the perpendicular projection from $H^n - N_s(j)$ to j decreases every distance by at least a factor of $1/\cosh s$, so any long path in $H^n - N_s(j)$ with endpoints on $\partial N_s(j)$ can be replaced by a much shorter path consisting of two segments perpendicular to j, together with a segment of j. 5.41



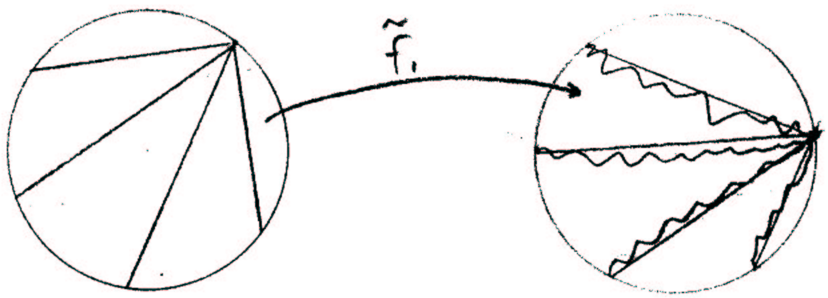
When this fact is applied to a line j joining distant points p_1 and p_2 on $\tilde{f}_1(g)$, it follows that the segment of g between p_1 and p_2 must intersect each plane perpendicular to j a bounded distance from j. It follows immediately that there is a limit line h to such lines j as p_1 and p_2 go to $+\infty$ and $-\infty$ on $\tilde{f}_1(g)$, and that $\tilde{f}_1(g)$ remains a bounded distance from h. Since no two lines in H^n remain a bounded distance apart, h is unique. □

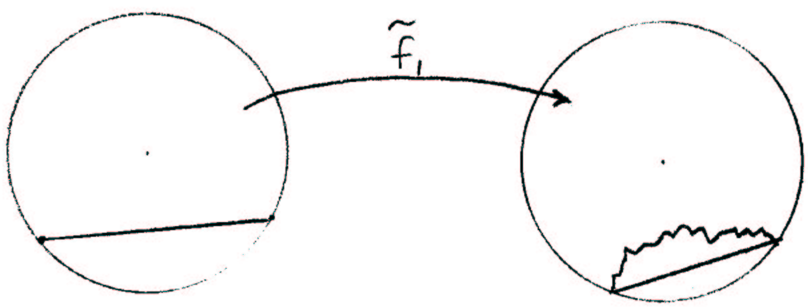


5.42

Corollary 5.9.3. *$\tilde{f}_1 : H^n \to H^n$ induces a one-to-one correspondence between the spheres at infinity.*

Proof. There is a one-to-one correspondence between points on the sphere at infinity and equivalence classes of directed geodesics, two geodesics being equivalent if they are parallel, or asymptotic in positive time. The correspondence of 5.9.2 between geodesics in $\tilde{M}_1$ and geodesics in $\tilde{M}_2$ obviously preserves this relation of parallelism, so it induces a map on the sphere at infinity. This map is one-to-one since any two distinct points in the sphere at infinity are joined by a geodesic, hence must be taken to the two ends of a geodesic. □





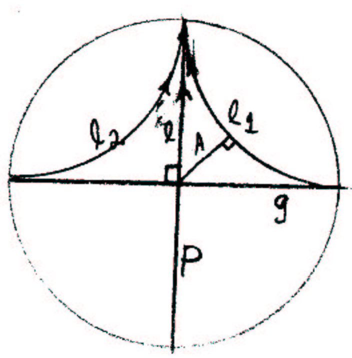
5.43

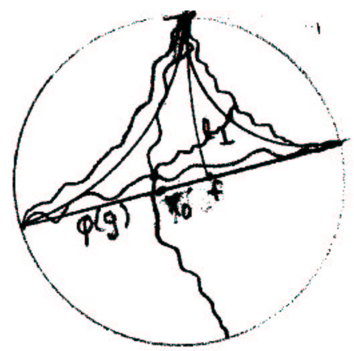
The next step in the proof of Mostow's Theorem is to show that the extension of $\tilde{f}_1$ to the sphere at infinity S_∞^{n-1} is continuous. One way to prove this is by citing Brouwer's Theorem that every function is continuous. Since this way of thinking is not universally accepted (though it is valid in the current situation), we will give another proof, which will also show that f is quasi-conformal at S_∞^{n-1}. A basis for the neighborhoods of a point $x \in S_\infty^{n-1}$ is the set of disks with center x. The boundaries of the disks are $(n-2)$-spheres which correspond to hyperplanes in H^2 (i.e., to $(n-1)$-spheres perpendicular to S_∞^{n-1} whose intersections with S_∞^{n-1} are the $(n-2)$-spheres).

For any geodesic g in $\tilde{M}_1$, let $\phi(g)$ be the geodesic in $\tilde{M}_2$ which remains a bounded distance from $\tilde{f}_1(g)$.

LEMMA 5.9.4. *There is a constant c such that, for any hyperplane P in H^n and any geodesic g perpendicular to P, the projection of $\tilde{f}_1(P)$ onto $\phi(g)$ has diameter $\leq c$.*

PROOF. Let x be the point of intersection of P and g and let l be a geodesic ray based at x. Then there is a unique geodesic l_1 which is parallel to l in one direction and to a fixed end of g in the other. Let A denote the shortest arc between x and l_1. It has length d, where d is a fixed contrast ($= \operatorname{arc\,cosh} \sqrt{2}$). 5.44

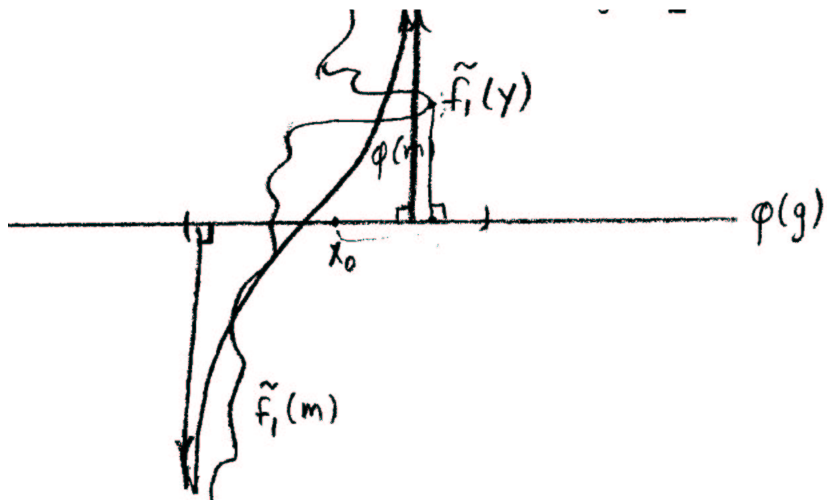




The idea of the proof is to consider the image of this picture under $\tilde{f}_l$. Let $\phi(l), \phi(l_1), \phi(g)$ denote the geodesics that remain a bounded distance from l, l_1 and g respectively. Since ϕ preserves parallelism $\phi(l)$ and $\phi(l_1)$ are parallel. Let $l^\perp$ denote the geodesic from the endpoint on S^{n-1}_∞ of $\phi(l)$ which is perpendicular to $\phi(g)$. Also let x_0 be the point on $\phi(g)$ nearest to $\tilde{f}_l(x)$.

Since $\tilde{f}_l(x)$ is a pseudo-isometry the length of $\tilde{f}_l(A)$ is at most $c_1 d$ where c_1 is a fixed constant. Since $\phi(l_1)$ and $\phi(g)$ are less than distance s (for a fixed constant s) from $\tilde{f}_l(l_1)$ and $\tilde{f}_l(g)$ respectively, it follows that x_0 is distance less than $C_1 d + 2s = \bar{d}$ from $\phi(l_1)$. This implies that the foot of $l^\perp$ (i.e., $l^\perp \cap \phi(g)$) lies distance less than $\bar{d}$ to one side of x_0. By considering the geodesic l_2 which is parallel to l and to the other end of g, it follows that f lies a distance less than $\bar{d}$ from x_0. 5.45

Now consider any point $y \in P$. Let m be any line through y. The endpoints of $\phi(m)$ project to points on $\phi(g)$ within a distance $\bar{d}$ of x_0; since $\tilde{f}_l(y)$ is within a distance s of $\phi(m)$, it follows that y projects to a point not farther than $\bar{d} + s$ from x_0. □



Corollary 5.9.5. *The extension of* $\tilde{f}_l$ *to* S_∞^{n-1} *is continuous.*

Proof. For any point $y \in S_\infty^{n-1}$, consider a directed geodesic g bending toward y, and define $\tilde{f}_l(y)$ to be the endpoint of $\phi(g)$. The half-spaces bounded by hyperplanes perpendicular to $\phi(g)$ form a neighborhood basis for $\tilde{f}_l(y)$. For any such half-space H, there is a point $x \in g$ such that the projection of $\tilde{f}_l(x)$ to $\phi(g)$ is a distance $> C$ from ∂H. Then the neighborhood of y bounded by the hyperplane through x perpendicular to g is mapped within H. □

Below it will be necessary to use the concept of quasi-conformality. If f is a 5.46
homeomorphism of a metric space X to itself, f is K-*quasi-conformal* if and only if for all $z \in X$

$$\lim_{r \to 0} \frac{\sup_{x,y \in S_r(z)} d\left(f(x), f(y)\right)}{\inf_{x,y \in S_r(z)} d\left(f(x), f(y)\right)} \leq K$$

where $S_r(Z)$ is the sphere of radius r around Z, and x and y are diametrically opposite. K measures the deviation of f from conformality, is equal to 1 if f is conformal, and is unchanged under composition with a conformal map. f is called *quasi-conformal* if it is K-quasi-conformal for some K.

Corollary 5.9.6. $\tilde{f}_l$ *is quasi-conformal at* S_∞^{n-1}.

Proof. Use the upper half-space model for H^n since it is conformally equivalent to the ball model and suppose x and $\tilde{f}_l x$ are the origin since translation to the origin is also conformal. Then consider any hyperplane P perpendicular to the geodesic g from 0 to the point at infinity. By Lemma 5.9.4 there is a bound, depending only on $\tilde{f}_l$, to the diameter of the projection of $\tilde{f}_l(P)$ onto $\phi(g) = g$. Therefore, there are hyperplanes P_1, P_2 perpendicular to g contained in and containing $\tilde{f}_l(P)$ and the distance (along g) between P_1 and P_2 is uniformly bounded for all planes P.

But this distance equals $\log r, r > 1$, where r is the ratio of the radii of the $n-2$ spheres

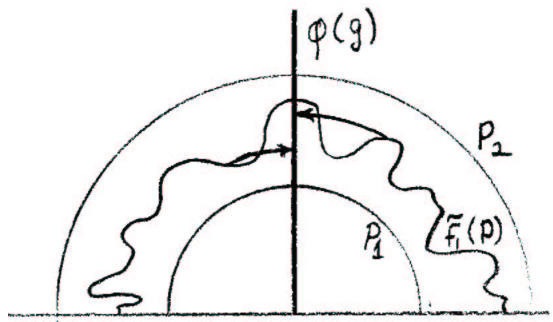
$$S^{n-2}_{p_1}, \quad S^{n-2}_{p_2}$$

in S^{n-1}_∞ corresponding to P_1 and P_2. The image of the $n-2$ sphere S^{n-2}_P corresponding 5.47
to P lies between $S^{n-2}_{p_2}$ and $S^{n-2}_{p_1}$ so that r is an upper bound for the ratio of maximum to minimum distances on

$$\tilde{f}_l(S^{n-2}_p).$$

Since $\log r$ is uniformly bounded above, so is r and $\tilde{f}_l$ is quasi-conformal. □

Corollary 5.9.6 was first proved by Gehring for dimension $n = 3$, and generalized to higher dimensions by Mostow.



At this point, it is necessary to invoke a theorem from analysis (see Bers).

Theorem 5.9.7. *A quasi-conformal map of an $n-1$-manifold, $n > 2$, has a derivative almost everywhere (= a.e.).*

Remark. It is at this stage that the proof of Mostow's Theorem fails for $n = 2$. 5.48
The proof works to show that $\tilde{f}_l$ extends quasi-conformally to the sphere at infinity, S^1_∞, but for a one-manifold this does not imply much.

Consider $\tilde{f}_l : S^{n-1}_\infty \to S^{n-1}_\infty$; by theorem 5.9.7 $d\tilde{f}_l$ exists a.e. At any point x where the derivative exists, the linear map $d\tilde{f}_l(x)$ takes a sphere around the origin to an ellipsoid. Let $\lambda_1, \ldots, \lambda_{n-1}$ be the lengths of the axes of the ellipsoid. If we normalize so that $\lambda_1 \cdot \lambda_2 \cdots \lambda_{n-1} = 1$, then the λ_i are conformal invariants. In particular denote the maximum ratio of the λ_i's at x by $e(x)$, the eccentricity of $\tilde{f}_l$ at x. Note that if $\tilde{f}_l$ is K-quasi-conformal, the supremum of $e(x)$, $x \in S^{n-1}_\infty$, is K. Since $\pi_1 M_1$ acts on S^{n-1}_∞ conformally and e is invariant under conformal maps, e is a measurable, $\pi_1 M_1$ invariant function on S^{n-1}_∞. However, such functions are very simple because of the following theorem:

Theorem 5.9.8. *For a closed, hyperbolic n-manifold M, $\pi_1 M$ acts ergodically on S^{n-1}_∞, i.e., every measurable, invariant set has zero measure or full measure.*

Corollary 5.9.9. *e is constant a.e.*

Proof. Any level set of e is a measurable, invariant set so precisely one has full measure. □

In fact more is true: 5.49

Theorem 5.9.10. $\pi_1(M)$ *acts ergodically on* $S^{n-1}_\infty \times S^{n-1}_\infty$.

Remark. This theorem is equivalent to the fact that the geodesic flow of M is ergodic since pairs of distinct points on S^{n-1}_∞ are in a one-to-one correspondence to geodesics in H^n (whose endpoints are those points).

From Corollary 5.9.9 e is equal a.e. to a constant K, and if the derivative of $\tilde{f}_l$ is not conformal, $K \neq 1$.

Consider the case $n = 3$. The direction of maximum "stretch" of df defines a measurable line field l on S^2_∞. Then for any two points $x, y \in S^2_\infty$ it is possible to parallel translate the line $l(x)$ along the geodesic between x and y to y and compute the angle between the translation of $l(x)$ and $l(y)$. This defines a measurable $\pi_1 M$-invariant function on $S^2_\infty \times S^2_\infty$. By theorem 5.9.10 it must be constant a.e. In other words l is determined by its "value" at one point. It is not hard to see that this is impossible.

For example, the line field determined by a line at x agrees with the line field below a.e. However, any line field determined by its "value" at y will have the same form and will be incompatible. 5.50

The precise argument is easy, but slightly more subtle, since l is defined only a.e.

The case $n > 3$ is similar.

Now one must again invoke the theorem, from analysis, that a quasi-conformal map whose derivative is conformal a.e. is conformal in the usual sense; it is a sphere-preserving map of S^{n-1}_∞, so it extends to an isometry I of H^n. The isometry I conjugates the action of $\pi_1 M_1$ to the action of $\pi_1 M_2$, completing the proof of Mostow's Theorem. □

5.10. A decomposition of complete hyperbolic manifolds.

5.51

Let M be any complete hyperbolic manifold (possibly with infinite volume). For $\epsilon > 0$, we will study the decomposition $M = M_{(0,\epsilon]} \cup M_{[\epsilon,\infty)}$, where $M_{(0,\epsilon]}$ consists of those points in M through which there is a non-trivial closed loop of length $\leq \epsilon$, and $M_{[\epsilon,\infty)}$ consists of those points through which every non-trivial loop has length $\geq \epsilon$.

In order to understand the geometry of $M_{(0,\epsilon]}$, we pass to the universal cover $\tilde{M} = H^n$. For any discrete group Γ of isometries of H^n and any $x \in H^n$ let $\Gamma_\epsilon(x)$ be the subgroup generated by all elements of Γ which move x a distance $\leq \epsilon$, and let

$\Gamma'_\epsilon(x) \subset \Gamma_\epsilon(x)$ be the subgroup consisting of elements whose derivative is also ϵ-close to the identity.

LEMMA 5.10.1 (The Margulis Lemma). *For every dimension n there is an $\epsilon > 0$ such that for every discrete group Γ of isometries of H^n and for every $x \in H^n$, $\Gamma'_\epsilon(x)$ is abelian and $\Gamma_\epsilon(x)$ has an abelian subgroup of finite index.*

REMARK. This proposition is much more general than stated; if "abelian" is replaced by "nilpotent," it applies in general to discrete groups of isometries of Riemannian manifolds with bounded curvature. The proof of the general statement is essentially the same.

PROOF. In any Lie group G, since the commutator map $[\ ,\ ] : G \times G \to G$ has derivative 0 at $(1,1)$, it follows that the size of the commutator of two small elements 5.52
is bounded above by some constant times the product of their sizes. Hence, if Γ'_ϵ is any discrete subgroup of G generated by small elements, it follows immediately that the lower central series $\Gamma'_\epsilon \supset [\Gamma'_\epsilon, \Gamma'_\epsilon] \supset [\Gamma'_\epsilon, [\Gamma'_\epsilon, \Gamma'_\epsilon]], \dots$ is finite (since there is a lower bound to the size of elements of Γ'_ϵ). In other words, Γ'_ϵ is nilpotent. When G is the group of isometries of hyperbolic space, it is not hard to see (by considering, for instance, the geometric classification of isometries) that this implies Γ'_ϵ is actually abelian.

To guarantee that $\Gamma_\epsilon(x)$ has an abelian subgroup of finite index, the idea is first to find an ϵ_1 such that $\Gamma'_{\epsilon_1}(x)$ is always abelian, and then choose ϵ many times smaller than ϵ_1, so the product of generators of $\Gamma_\epsilon(x)$ will lie in $\Gamma'_{\epsilon_1}(x)$. Here is a precise recipe:

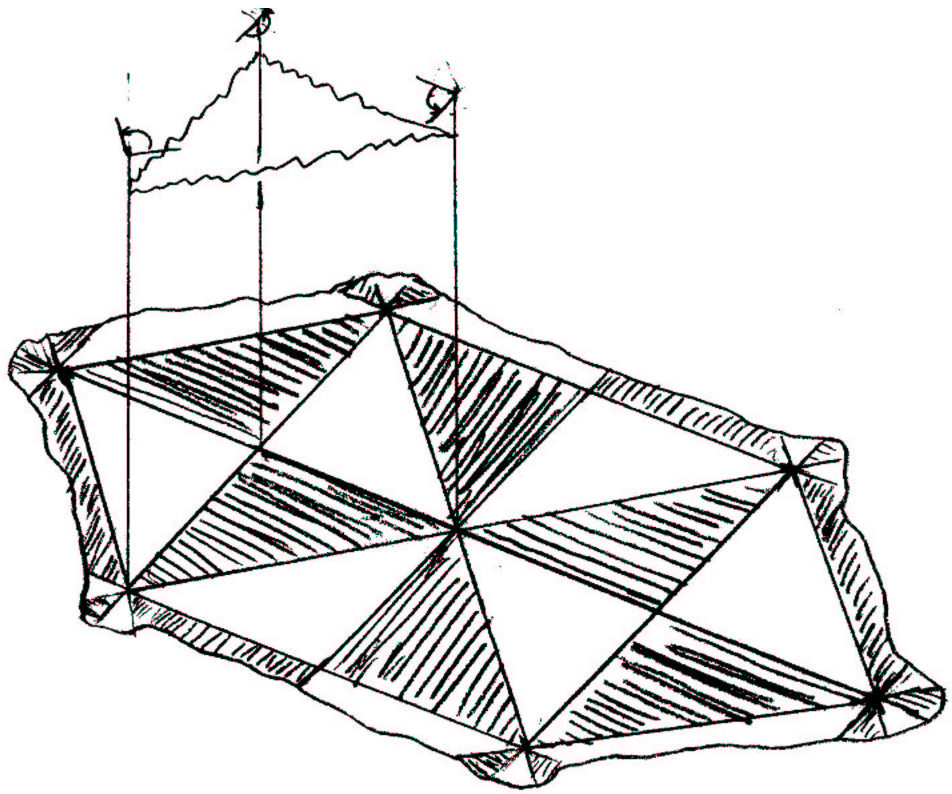
Let N be large enough that any collection of elements of $O(n)$ with cardinality $\geq N$ contains at least one pair separated by a distance not more than $\epsilon_{1/3}$.

Choose $\epsilon_2 \leq \epsilon_{1/3}$ so that for any pair of isometries ϕ_1 and ϕ_2 of H^n which translate a point x a distance $\leq \epsilon_2$, the derivative at x of $\phi_1 \circ \phi_2$ (parallel translated back to x) is estimated within $\epsilon_{1/6}$ by the product of the derivatives at x of ϕ_1 and ϕ_2 (parallel translated back to x).

Now let $\epsilon = \epsilon_{2/2N}$, so that a product of $2N$ isometries, each translating x a distance $\leq \epsilon$, translates x a distance $\leq \epsilon_2$. Let $g_1, \dots, g_k$ be the set of elements of Γ 5.53
which move x a distance $\leq \epsilon$; they generate $\Gamma_\epsilon(x)$. Consider the cosets $\gamma\, \Gamma'_{\epsilon_1}(x)$, where $\gamma \in \Gamma_\epsilon(x)$; the claim is that they are all represented by γ's which are words of length $< N$ in the generators $g_1, \dots, g_k$. In fact, if $\gamma = g_{i_1} \cdot \ldots \cdot g_{i_l}$ is any word of length $\geq N$ in the g_i's, it can be written $\gamma = \alpha \cdot \epsilon' \cdot \beta, (\alpha, \epsilon', \beta \neq 1)$ where $\epsilon' \cdot \beta$ has length $\leq N$, and the derivative of ϵ' is within $\epsilon_{1/3}$ of 1. It follows that $(\alpha\beta)^{-1} \cdot (\alpha\epsilon'\beta) = \beta^{-1}\epsilon'\beta$ is in $\Gamma'_{\epsilon_1}(x)$; hence the coset $\gamma\Gamma'_{\epsilon_1}(x) = (\alpha\beta)\Gamma'_{\epsilon_1}(x)$. By induction, the claim is verified. Thus, the abelian group $\Gamma'_{\epsilon_1}(x)$ has finite index in the group generated by $\Gamma_\epsilon(x)$ and $\Gamma'_{\epsilon_1}(x)$, so $\Gamma'_{\epsilon_1}(x) \cap \Gamma_\epsilon(x)$ with finite index. $\square$

Examples. When $n = 3$, the only possibilities for discrete abelian groups are $\mathbb{Z}$ (acting hyperbolically or parabolically), $\mathbb{Z} \times \mathbb{Z}$ (acting parabolically, conjugate to a group of Euclidean translations of the upper half-space model), $\mathbb{Z} \times \mathbb{Z}_n$ (acting as a group of translations and rotations of some axis), and $\mathbb{Z}_2 \times \mathbb{Z}_2$ (acting by 180° rotations about three orthogonal axes). The last example of course cannot occur as $\Gamma'_\epsilon(x)$. Similarly, when ϵ is small compared to $1/n$, $\mathbb{Z} \times \mathbb{Z}_n$ cannot occur as $\Gamma'_\epsilon(x)$.

Any discrete group Γ of isometries of Euclidean space E^{n-1} acts as a group of isometries of H^n, via the upper half-space model. 5.54

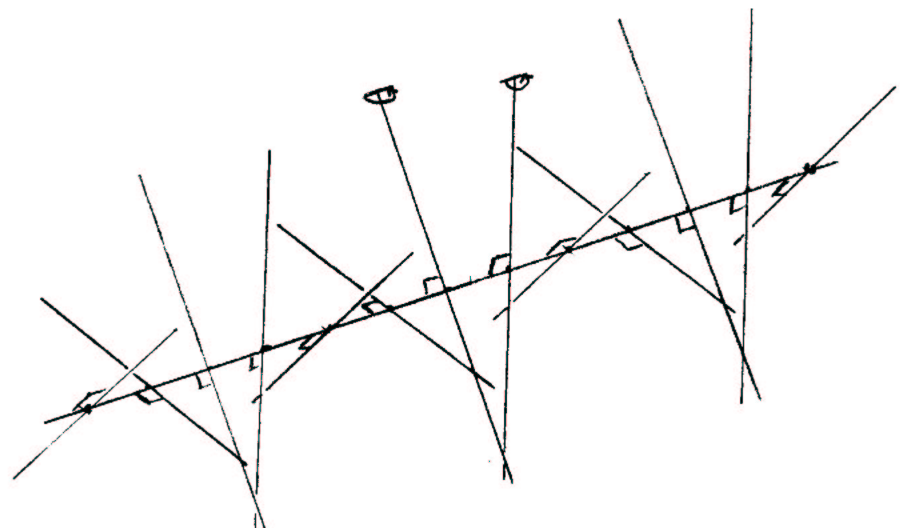


For any x sufficiently high (in the upper half space model), $\Gamma_\epsilon(x) = \Gamma$. Thus, 5.10.1 contains as a special case one of the Bieberbach theorems, that Γ contains an abelian subgroup of finite index. Conversely, when $\Gamma_\epsilon(x) \cap \Gamma'_{\epsilon_1}(x)$ is parabolic, $\Gamma_\epsilon(x)$ must be a Bieberbach group. To see this, note that if $\Gamma_\epsilon(x)$ contained any hyperbolic element γ, no power of γ could lie in $\Gamma'_{\epsilon_1}(x)$, a contradiction. Hence, $\Gamma_\epsilon(x)$ must consist of parabolic and elliptic elements with a common fixed point p at ∞, so it acts as a group of isometries on any horosphere centered at p.

If $\Gamma_\epsilon(x) \cap \Gamma'_{\epsilon_1}(x)$ is not parabolic, it must act as a group of translations and rotations of some axis a. Since it is discrete, it contains $\mathbb{Z}$ with finite index (provided $\Gamma_\epsilon(x)$ is infinite). It easily follows that $\Gamma_\epsilon(x)$ is either the product of some finite

FIGURE 1. The infinite dihedral group acting on H^3.

subgroup F of $O(n-1)$ (acting as rotations about a) with $\mathbb{Z}$, or it is the semidirect product of such an F with the infinite dihedral group, $\mathbb{Z}/2 * \mathbb{Z}/2$. 5.55



For any set $S \subset H^n$, let $B_r(S) = \{x \in H^n \,|\, d(x, S) \leq r\}$.

COROLLARY 5.10.2. *There is an $\epsilon > 0$ such that for any complete oriented hyperbolic three-manifold M, each component of $M_{(0,\epsilon]}$ is either*

(1) *a horoball modulo $\mathbb{Z}$ or $\mathbb{Z} \oplus \mathbb{Z}$, or*
(2) *$B_r(g)$ modulo $\mathbb{Z}$, where g is a geodesic.*

The degenerate case $r = 0$ may occur.

PROOF. Suppose $x \in M_{(0,\epsilon]}$. Let $\tilde{x} \in H^3$ be any point which projects to x. There is some covering translation γ which moves x a distance $\leq \epsilon$. If γ is hyperbolic, let a be its axis. All rotations around a, translations along a, and uniform contractions of hyperbolic space along orthogonals to a commute with γ. It follows that $\tilde{M}_{(0,\epsilon]}$ 5.56 contains $B_r(a)$, where $r = d(a, x)$, since γ moves any point in $B_r(a)$ a distance $\leq \epsilon$. Similarly, if γ is parabolic with fixed point p at ∞, $\tilde{M}_{(0,\epsilon]}$ contains a horoball about p passing through x. Hence $M_{(0,\epsilon]}$ is a union of horoballs and solid cylinders $B_r(a)$. Whenever two of these are not disjoint, they correspond to two covering transformations γ_1 and γ_2 which move some point x a distance $\leq \epsilon$; γ_1 and γ_2 must commute (using 5.10.1), so the corresponding horoballs or solid cylinders must be concentric, and 5.10.2 follows. $\square$

5.11. Complete hyperbolic manifolds with bounded volume.

It is easy now to describe the structure of a complete hyperbolic manifold with finite volume; for simplicity we stick to the case $n = 3$.

PROPOSITION 5.11.1. *A complete oriented hyperbolic three-manifold with finite volume is the union of a compact submanifold (bounded by tori) and a finite collection of horoballs modulo* $\mathbb{Z} \oplus \mathbb{Z}$ *actions.*

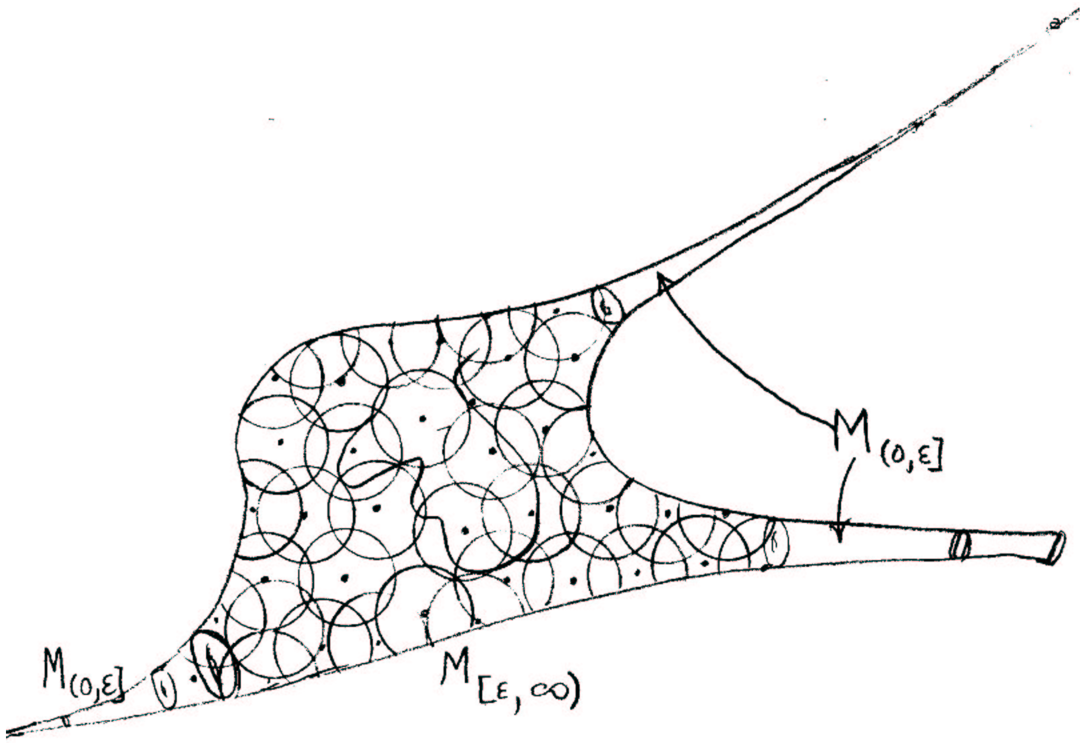
PROOF. $M_{[\epsilon,\infty)}$ must be compact, for otherwise there would be an infinite sequence of points in $M_{[\epsilon,\infty)}$ pairwise separated by at least ϵ. This would give a sequence of hyperbolic $\epsilon/2$ balls disjointly embedded in $M_{[\epsilon,\infty)}$, which has finite volume. $M_{(0,\epsilon]}$ must have finitely many components (since its boundary is compact). The proposition is obtained by lumping all compact components of $M_{(0,\epsilon]}$ with $M_{[\epsilon,\infty)}$. □

5.57

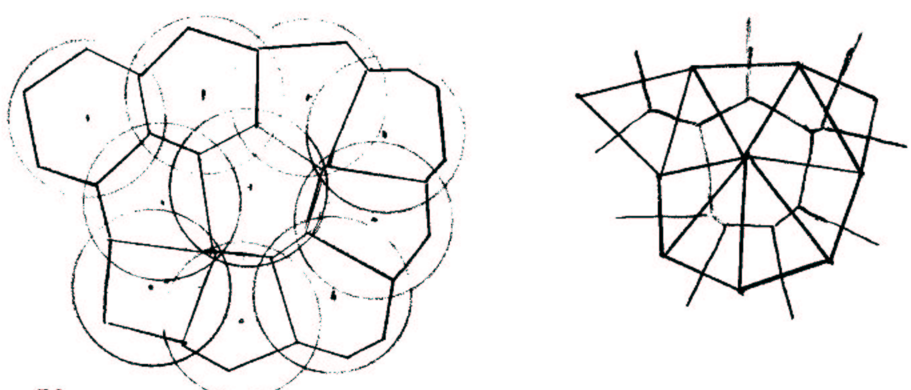
With somewhat more effort, we obtain Jørgensen's theorem, which beautifully describes the structure of the set of all complete hyperbolic three-manifolds with volume bounded by a constant C:

THEOREM 5.11.2 (Jørgensen's theorem [first version]). *Let $C > 0$ be any constant. Among all complete hyperbolic three-manifolds with volume $\leq C$, there are only finitely many homeomorphism types of $M_{[\epsilon,\infty)}$. In other words, there is a link L_c in S^3 such that every complete hyperbolic manifold with volume $\leq C$ is obtained by Dehn surgery along L_C. (The limiting case of deleting components of L_C to obtain a non-compact manifold is permitted.)*

PROOF. Let V be any maximal subset of $M_{[\epsilon,\infty)}$ having the property that no two elements of V have distance $\leq \epsilon/2$. The balls of radius $\epsilon/4$ about elements of V are embedded; since their total volume is $\leq C$, this gives an upper bound to the cardinality of V. The maximality of V is equivalent to the property that the balls of radius $\epsilon/2$ about V cover.



The combinatorial pattern of intersections of this set of $\epsilon/2$-balls determines $M_{[\epsilon,\infty)}$ up to diffeomorphism. There are only finitely many possibilities. (Alternatively a triangulation of $M_{[\epsilon,\infty)}$ with vertex set V can be constructed as follows. First, form a cell division of $M_{[\epsilon,\infty)}$ whose cells are indexed by V, associating to each $v \in V$ the subset of $M_{[\epsilon,\infty)}$ consisting of $x \in M_{[\epsilon,\infty)}$ such that $d(x,v) < d(x,v')$ for all $v' \in V$.



If

If V is in general position, faces of the cells meet at most four at a time. (The dual cell division is a triangulation.)

Any two hyperbolic manifolds M and N such that $M_{[\epsilon,\infty)} = N_{[\epsilon,\infty)}$ can be obtained from one another by Dehn surgery. All manifolds with volume $\leq C$ can therefore be obtained from a fixed finite set of manifolds by Dehn surgery on a fixed link in each manifold. Each member of this set can be obtained by Dehn surgery on some link in S^3, so all manifolds with volume $\leq C$ can be obtained from S^3 by Dehn surgery on the disjoint union of all the relevant links. $\square$ 5.59

The full version of Jørgensen's Theorem involves the geometry as well as the topology of hyperbolic manifolds. The geometry of the manifold $M_{[\epsilon,\infty)}$ completely determines the geometry and topology of M itself, so an interesting statement comparing the geometry of $M_{[\epsilon,\infty)}$'s must involve the approximate geometric structure. Thus, if M and N are complete hyperbolic manifolds of finite volume, Jørgensen defines M to be *geometrically near* N if for some small ϵ, there is a diffeomorphism which is approximately an isometry from the hyperbolic manifold $M_{[\epsilon,\infty)}$ to $N_{[\epsilon,\infty)}$. It would suffice to keep ϵ fixed in this definition, except for the exceptional cases when M and N have closed geodesics with lengths near ϵ. This notion of geometric nearness gives a topology to the set $\mathcal{H}$ of isometry classes of complete hyperbolic manifolds of finite volume. Note that neither coordinate systems nor systems of generators for the fundamental groups have been chosen for these hyperbolic manifolds; the homotopy class of an approximate isometry is arbitrary, in contrast with the definition for Teichmüller space. Mostow's Theorem implies that every closed manifold M in $\mathcal{H}$ is an isolated point, since $M_{[\epsilon,\infty)} = M$ when ϵ is small enough. On the other hand, a manifold in $\mathcal{H}$ with one end or *cusp* is a limit point, by the hyperbolic Dehn surgery theorem 5.9. A manifold with two ends is a limit point of limit points and a manifold with k ends is a k-fold limit point. 5.60

Mostow's Theorem implies more generally that the number of cusps of a geometric limit M of a sequence $\{M_i\}$ of manifolds distinct from M must strictly exceed the lim sup of the number of cusps of M_i. In fact, if ϵ is small enough, $M_{(0,\epsilon]}$ consists only of cusps. The cusps of M_i are contained in $M_{i_{(0,\epsilon]}}$; if all its components are cusps, and if $M_{i_{[\epsilon,\infty)}}$ is diffeomorphic with $M_{[\epsilon,\infty)}$ then M_i is diffeomorphic with M so M_i is isometric with M.

The volume of a hyperbolic manifold gives a function $v : \mathcal{H} \to \mathbb{R}_+$. If two manifolds M and N are geometrically near, then the volumes of $M_{[\epsilon,\infty)}$ and $N_{[\epsilon,\infty)}$ are approximately equal. The volume of a hyperbolic solid torus r_0 centered around a geodesic of length l may be computed as

$$\text{volume (solid torus)} = \int_0^{r_0} \int_0^{2\pi} \int_0^{l} \sinh r \cosh r \, dt \, d\theta \, dr = \pi l \sinh^2 r_0$$

while the area of its boundary is

$$\text{area (torus)} = 2\pi l \sinh r_0 \cosh r_0.$$

Thus we obtain the inequality

$$\frac{\text{area } (\partial \text{ solid torus})}{\text{volume (solid torus)}} = \frac{1}{2}\frac{\sinh r_0}{\cosh r_0} < \frac{1}{2}.$$

The limiting case as $r_0 \to \infty$ can be computed similarly; the ratio is $1/2$. Applying 5.61
this to M, we have

5.11.2. $$\text{volume } (M) \leq \text{ volume } (M_{[\epsilon,\infty)}) + \tfrac{1}{2} \text{ area } (\partial M_{[\epsilon,\infty)}).$$

It follows easily that v is a continuous function on $\mathcal{H}$. Changed this label to 5.11.2a.

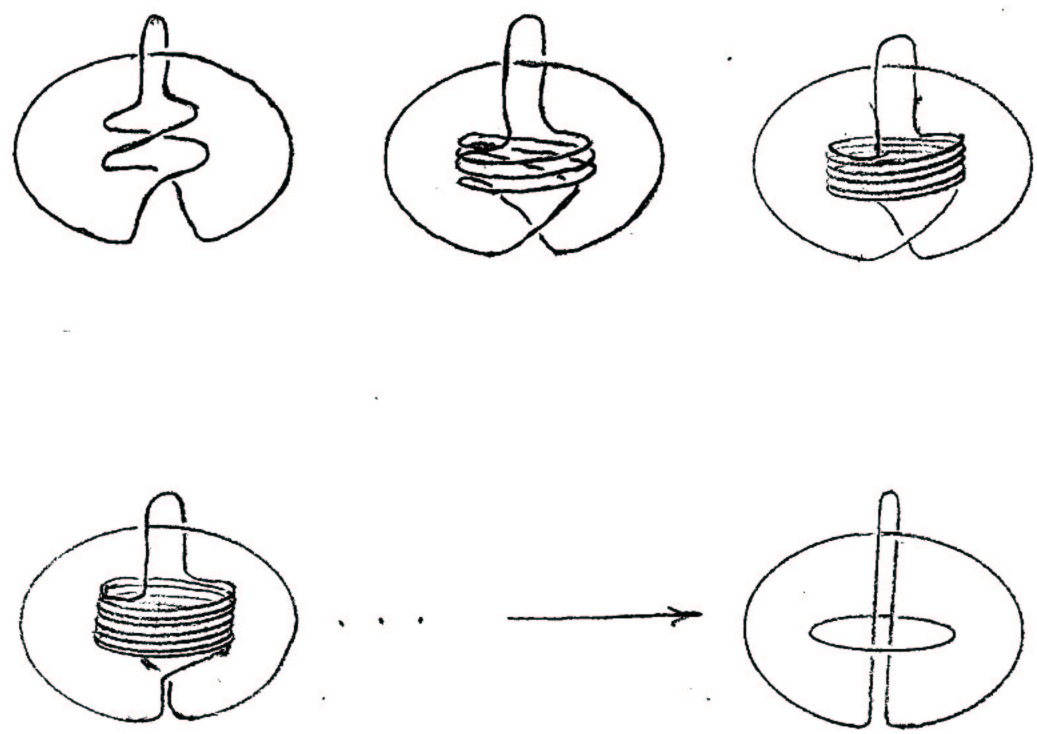
5.12. Jørgensen's Theorem.

Theorem 5.12.1. *The function $v : \mathcal{H} \to \mathbb{R}_+$ is proper. In other words, every sequence in $\mathcal{H}$ with bounded volume has a convergent subsequence. For every C, there is a finite set $M_1, \ldots, M_k$ of complete hyperbolic manifolds with volume $\leq C$ such that all other complete hyperbolic manifolds with volume $\leq C$ are obtained from this set by the process of hyperbolic Dehn surgery* (as in 5.9)*.*

Proof. Consider a maximal subset of V of $M_{[\epsilon,\infty)}$ having the property that no two elements of V have distance $\leq \epsilon/2$ (as in 5.11.1). Choose a set of isometries of the $\epsilon/2$ balls centered at elements of V with a standard $\epsilon/2$-ball in hyperbolic space. The set of possible gluing maps ranges over a compact subset of $\mathrm{Isom}(H^3)$, so any sequence of gluing maps (where the underlying sequence of manifolds has volume $\leq C$) has a convergent subsequence. It is clear that in the limit, the gluing maps still give a hyperbolic structure on $M_{[\epsilon,\infty)}$, approximately isometric to the limiting $M_{[\epsilon,\infty)}$'s. We must verify that $M_{[\epsilon,\infty)}$ extends to a complete hyperbolic manifold. To see this, note that whenever a complete hyperbolic manifold N has a geodesic which is very short compared to ϵ, the radius of the corresponding solid torus in $N_{(0,\epsilon]}$
becomes large. (Otherwise there would be a short non-trivial curve on $\partial N_{(0,\epsilon]}$—but 5.62
such a curve has length $\geq \epsilon$). Thus, when a sequence $\{M_{i_{[\epsilon,\infty)}}\}$ converges, there are approximate isometries between arbitrarily large balls $B_r(M_{i_{[\epsilon,\infty)}})$ for large i, so in the limit one obtains a complete hyperbolic manifold. This proves that v is a proper function. The rest of §5.12 is merely a restatement of this fact. □

Remark. Our discussion in §5.10, 5.11 and 5.12 has made no attempt to be numerically efficient. For instance, the proof that there is an ϵ such that $\Gamma_\epsilon(x)$ has an abelian subgroup of finite index gives the impression that ϵ is microscopic. In fact, ϵ can be rather large; see Jørgensen, for a more efficient approach. It would be extremely interesting to have a good estimate for the number of distinct $M_{[\epsilon,\infty)}$'s

Figure eight knot



Whitehead Link

where M has volume $\leq C$, and it would be quite exciting to find a practical way of computing them. The development in 5.10, 5.11, and 5.12 is completely inefficient in this regard. Jørgensen's approach is much more explicit and efficient.

Example. The sequence of knot complements below are all obtained by Dehn surgery on the Whitehead link, so 5.8.2 implies that all but a finite number possess complete hyperbolic structures. (A computation similar to that of Theorem 4.7 shows that in fact they all possess hyperbolic structures.) This sequence converges, in $\mathcal{H}$, to the Whitehead link complement: 5.63

Note. Gromov proved that in dimensions $n \neq 3$, there is only a finite number of complete hyperbolic manifolds with volume less than a given constant. He proved this more generally for negatively curved Riemannian manifolds with curvature varying between two negative constants. His basic method of analysis was to study the injectivity radius

$$\begin{aligned}\operatorname{inj}(x) &= \tfrac{1}{2}\inf\{\text{lengths of non-trivial closed loops through } x\} \\ &= \sup\{r \mid \text{the exponential map is injective on the ball of radius } r \text{ in } T(x)\}.\end{aligned}$$

Basically, in dimensions $n \neq 3$, little can happen in the region M^n_ϵ of M^n where $\text{inj}(x)$ is small. This was the motivation for the approach taken in 5.10, 5.11 and 5.12. Gromov also gave a weaker version of hyperbolic Dehn surgery, 5.8.2: he showed that many of the manifolds obtained by Dehn surgery can be given metrics of negative curvature close to -1.