CLASSICAL GEOMETRY AND LOW–DIMENSIONAL TOPOLOGY — LECTURE NOTES

DANNY CALEGARI

1. DEHN INVARIANTS AND EQUIDECOMPOSABILITY

1.1. **Scissors congruence in $\mathbb{E}^2$.** Consider a pair of polygons P, Q in the Euclidean plane $\mathbb{E}^2$. When can P be decomposed into a finite collection of pieces which can then be rearranged to form Q? There are two ambiguities to be resolved before we can answer this question:

1. What kind of decomposition is allowed? Is a piece of P allowed to consist of *any* subset? In this section we will restrict ourselves to *scissors congruence*: the decompositions of P that can be obtained with a pair of straight–edged scissors.
2. What kind of rearranging is allowed? When we move the pieces of P around, what kind of transformations can we use? In general we can restrict ourselves to arbitrary subgroups of the group of bijections of the plane to itself, but to get interesting mathematics, we usually restrict our motions to finite dimensional Lie groups of diffeomorphisms of our space, for instance, the group of similarities, or the group of rigid motions, or the group of affine transformations, or the group of translations, etc.

Definition 1.1. *Scissors Congruence data* consists of a given geometric space X, a collection of subsets of that space

$$S_\alpha \subset X$$

a collection of allowable finite decompositions of subsets

$$S_\alpha = \bigcup_i S_{\alpha_i}$$

and a group of allowable transformations of the pieces

$$G : \{S_\alpha\} \to \{S_\alpha\}$$

The *Scissors Congruence group* $\mathcal{P}(X)$ is the Abelian group with generators given by the S_α, and relations given by

$$[S_\alpha] = \sum_i [S_{\alpha_i}]$$

and

$$[S_\alpha] = [g(S_\alpha)]$$

Then we say S, R are *stably scissors congruent* if $[S] = [R]$ in this group.

If we consider $X = \mathbb{E}^2$, the S_α the set of (possibly disjoint collections of) polygons, and G the group of isometries of the plane, then the scissors congruence group is isomorphic to $\mathbb{R}$, and the natural isomorphism $\mathcal{P}(\mathbb{E}^2) \to \mathbb{R}$ is given by the area function. In fact, this is still true if we restrict to the subgroup of isometries of the plane generated by translations and by rotation (through any fixed point) by π.

Theorem 1.2. *Any two polygons P, Q in $\mathbb{E}^2$ with the same area are equidecomposable via translations and rotation through π.*

Proof: First subdivide P into a collection of triangles. By further subdivision, we can assume that each triangle has two sides parallel to the x and y axes. Each such triangle can be decomposed into a rectangle with sides parallel to the x and y axes. If R_1, R_2 are two such rectangles, translate them in the plane so that they have a common interior vertex. By a trick, any two such rectangles with equal area are equidecomposable. The idea of the trick is illustrated in figure 1. □

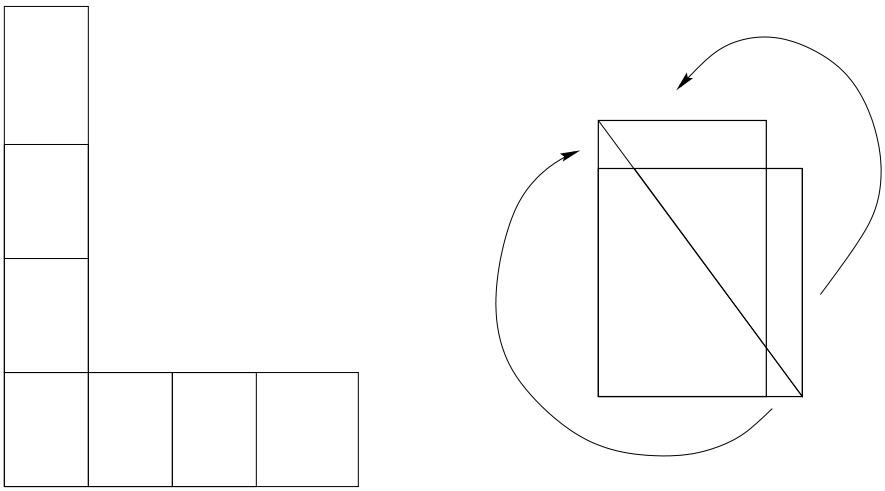

FIGURE 1. Two rectangles of the same area with parallel sides are equidecomposable via translations alone.

In particular, any two polygons P, Q of the same area which have a rotational symmetry of order 2 are equidecomposable by translation alone.

1.1.1. *Zylev's theorem.* There is a slight subtlety in the definition of $\mathcal{P}(X)$ for a G–space X which is the difference between scissors congruence and stable scissors congruence. Two polyhedra P, Q are stably scissors congruent if $P \coprod A \sim Q \coprod A$ for some A; this is the equivalence relation imposed by the equality of $[P]$ and $[Q]$ in $\mathcal{P}(X)$. Fortunately, for reasonable X, G stable scissors congruence agrees with the usual scissors congruence. That is, we have:

Theorem 1.3 (Zylev)**.** *If $X = \mathbb{S}^n, \mathbb{E}^n, \mathbb{H}^n$ and $G = Isom(X)$, then for two polytopes P, Q we have $P \sim Q$ iff $P \coprod A \sim Q \coprod A$ for some A.*

Proof: The relation $P \coprod A \sim Q \coprod A$ determines a decomposition of A into pieces. Further subdivide A into pieces A_i, for $i \in \{1, \ldots, n\}$ which are sufficiently small.

Now, look at the image $f(A_1)$ of A_1 in $Q \coprod A$. We may suppose, if the A_i are small enough, that A_1 is equidecomposable with a subset B_1 of $Q/f(A_1)$. We interchange $f(A_1)$ with B_1, further decomposing pieces which intersect B_1. Then we interchange B_1 with $A_1 \subset A$, further decomposing pieces which intersect A_1. Thus we have moved A_1 to the correct place, at the cost of further decomposing pieces which intersected the subsets B_1 and A_1 of $Q \coprod A$. Now look at the image of A_2. It may have been further subdivided

by our actions up to this point, but that is irrelevant; we can look at $f'(A_2) \subset Q \coprod A$, which is equidecomposable with A_2. This subset is equidecomposable with some $B_2 \subset Q/f'(A_2)$ so we can interchange them, further decomposing pieces which intersect B_2. Then exchange B_2 with $A_2 \subset A$, decomposing pieces which intersect A_2. We continue inductively. At the ith stage, we have an equidecomposition of $P \coprod A$ with $Q \coprod A$ so that the pieces in $A_1 \cup A_2 \cup \cdots \cup A_i$ do not move. Thus after n steps, we have an equidecomposition of P with Q, as required. □

1.2. **Dimension three and the Dehn invariant.** It is natural to ask the analogous question about scissors congruence in dimension 3, in particular, is the map vol : $\mathcal{P}(\mathbb{E}^3) \to \mathbb{R}$ an isomorphism? This was the third problem posed by Hilbert in his famous 1900 Congress address, and it was clear that he expected the answer "no". This was actually the first of Hilbert's problems to be solved, and in fact it was solved by Max Dehn in 1900 before Hilbert's problems had even appeared in print!

Dehn's negative solution to the problem consisted of defining a new invariant of scissors congruence, which has come to be known as the *Dehn invariant*, which can be defined as a homomorphism $\delta : \mathcal{P}(\mathbb{E}^3) \to \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}/\pi\mathbb{Q}$.

Note that the $\mathbb{Z}$–module $\mathbb{R} \otimes_{\mathbb{Z}} \mathbb{R}/2\pi\mathbb{Z}$ is isomorphic as a $\mathbb{Z}$–module (i.e. Abelian group) to $\mathbb{R} \otimes_{\mathbb{Z}} \mathbb{R}/\pi\mathbb{Q}$. This Abelian group is uniquely divisible, so can be enhanced in a unique way to a $\mathbb{Q}$–module (i.e. $\mathbb{Q}$–vector space) which is naturally isomorphic to $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}/\pi\mathbb{Q}$. Thus it does not matter whether we take $\mathbb{Z}$ or $\mathbb{Q}$ coefficients.

Definition 1.4. If E is an edge of a polyhedron P, we denote by $l(E)$ and $\theta(E)$ the length and the dihedral angle respectively of E. Then the *Dehn invariant* of P is defined as

$$\delta(P) = \sum_E l(E) \otimes \theta(E)$$

where the sum is taken over all edges E of P.

It can be seen that δ extends to a homomorphism on $\mathcal{P}(\mathbb{E}^3)$. One can easily check that it is well–defined; G–invariance is obvious, since G is the group of Euclidean motions, which preserves length and dihedral angle. Moreover, subdivision has the following potential effects:

1. If a cut is transverse to an edge E, that edge is split into a pair of edges E_1, E_2 with the same dihedral angles, and
$$l(E) \otimes \theta(E) = l(E_1) \otimes \theta(E) + l(E_2) \otimes \theta(E)$$
2. If a cut goes through an edge E, that edge is split into a pair of edges with the same length and with dihedral angles θ_1, θ_2 adding up to θ, so that
$$l(E) \otimes \theta(E) = l(E) \otimes \theta(E_1) + l(E) \otimes \theta(E_2)$$
3. When a cut is transverse to a face, two new edges are created with the same edge length, and with dihedral angles summing to π, so that
$$0 = l(E) \otimes \theta(E) + l(E) \otimes (\pi - \theta(E))$$

On the other hand, it is very easy to find examples of polyhedra where the Dehn invariant is non–zero. In particular, the regular tetrahedron of volume 1 is *not* scissors congruent to the regular cube of volume 1.

In this context, we have the following:

Theorem 1.5 (Sydler, 1965)**.** *The map*

$$vol \oplus \delta : \mathcal{P}(\mathbb{E}^3) \to \mathbb{R} \oplus \mathbb{R} \otimes_{\mathbb{Z}} \mathbb{R}/\pi\mathbb{Q}$$

is injective.

For a proof see Jessen's paper [7] or Sydler's original paper [14]. Jessen simplified Sydler's argument and proved an analogous result in dimension 4. The idea of Jessen's proof is to decompose a polyhedron into simplices of a particular kind — those of the kind that a plane cuts off from an octant. Such a polyhedron is determined by three parameters: the lengths of the sides not passing through the orthonormal vertex. Injectivity of the Dehn invariant follows if we can show that such objects with the same Dehn invariant are equidecomposable modulo prisms, since prisms have Dehn invariant zero and can take arbitrary volume.

Notice that the Dehn invariant is insensitive to orientation. The completeness of Dehn invariant therefore implies that polyhedra are equidecomposable with their mirror images. Actually, this is true in every dimension for the following elementary reason:

Theorem 1.6. *If P denotes a Euclidean polyhedron, and $\bar{P}$ denotes its mirror image, then $[P] = [\bar{P}]$ in $\mathcal{P}(\mathbb{E}^n)$.*

Proof: Since polyhedra can be decomposed into geodesic simplices, it suffices to prove the result for simplices. Let Δ be a Euclidean simplex with vertices $x_0, \ldots, x_n$ and let z denote the incenter of Δ — that is, the center of the largest inscribed sphere in Δ. Let z_i be the foot of the perpendicular from z to the codimension 1 face opposite x_0. Let Z_{ij} be the convex hull of z, z_i, z_j and the codimension 2 face with vertices $x_0, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_n$. Then each Z_{ij} is mirror–symmetric. Hence Δ is scissors congruent to its mirror image. □

This construction is very old and goes back at least to Lagrange.

1.3. **Tessellations, periodic and otherwise.** A useful property of the Dehn invariant is that it provides an obstruction for a polyhedron P to tile $\mathbb{E}^3$. In particular, we show that

Theorem 1.7. *Let P tile $\mathbb{E}^3$, possibly non–periodically. Then $\delta(P) = 0$.*

Proof: Notice that the group of values of the Dehn invariant is a $\mathbb{Q}$–vector space, so that it has many homomorphisms to $\mathbb{R}$. Pick any such homomorphism to $\mathbb{R}$. Such a homomorphism gives us a map γ from polyhedra in $\mathbb{E}^3$ to $\mathbb{R}$ which is additive, and which vanishes on edges with dihedral angle rational multiples of π. Let B be a large ball tiled with copies of P. Then if the radius of B is r, there are $O(r^3)$ copies of P in B, so the value of this homomorphism on the collection of copies of P is approximately $\gamma(P)r^3$. However, by examining dihedral angles of the polyhedron Q which is the union of the copies of P, we see that it has only $O(r^2)$ free edges, each of which is isomorphic to an edge of P. Hence $\gamma(Q) \sim \gamma(P)r^2$. It follows that $\gamma(P) = 0$. Since γ was arbitrary, $\delta(P) = 0$. □

The key property of Euclidean spaces used here is that surface area of balls grows slower than volume of balls. For Hyperbolic space, surface area grows at roughly the *same* rate as volume, and in fact, there are hyperbolic polyhedra which tile $\mathbb{H}^3$ (non–periodically) but which have non–zero Dehn invariant.

The following example is due to Doug Zare: in the upper half–space model of H^3, let P be the polyhedron cut out by four vertical planes which intersect the plane at infinity in

the lines $x = 0, x = 1, y = 0, y = 1$, and the five hemispheres with centers at

$$(1/2, 1/2), (1/4, 1/4), (1/4, 3/4), (3/4, 1/4), (3/4, 3/4)$$

with (Euclidean) radii x, $\frac{x}{2}$, $\frac{x}{2}$, $\frac{x}{2}$, $\frac{x}{2}$. Then copies of P can be stacked on top of each other, each copy standing on 4 copies, in a 4–adic tree. It is possible that $\delta(P) = 0$ for some x. But one can deform P by a little "wiggle" on the top and the bottom 4 faces, and after doing so, $\delta(P') = \delta(P) + 3\delta(\text{wiggle})$. In particular, $\delta(P')$ can be made non-zero.

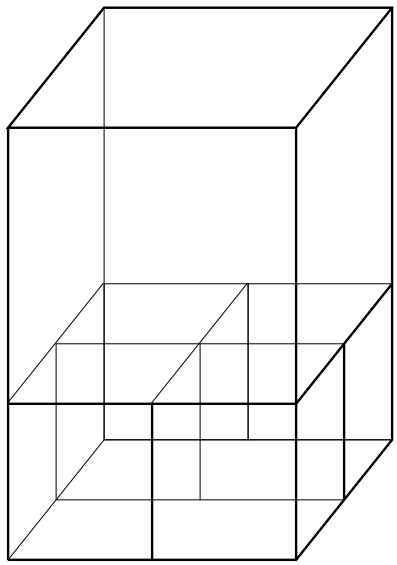

FIGURE 2. Zare's polyhedron tiles hyperbolic space (aperiodically) and has nonzero Dehn invariant.

In general, however, if a polyhedron tiles any of the homogeneous geometric spaces $\mathbb{S}^n, \mathbb{E}^n, \mathbb{H}^n$ *periodically*, its Dehn invariant vanishes. For, if P denotes the polyhedron, and Γ the discrete group of isometries which give the translates of P, then Γ has a finite index subgroup Γ_0 which acts properly discontinuously. Then the quotient of the geometric space by Γ_0 is a manifold M which has a decomposition into pieces isometric to P. Since around every edge, the sum of dihedral angles of the pieces is 2π, it follows that $\delta(M) = 0$. Since $[M] = n[P]$, we also have $\delta(P) = 0$.

1.4. **Miscellaneous Dissections.** We describe here some interesting dissections in dimension 3.

1. If O denotes the regular Euclidean octahedron of side length x, and T the regular Euclidean tetrahedron of side length x, then $\delta([O] + 2[T]) = 0$. In fact, the polyhedron P obtained from gluing two tetrahedra to opposite faces of an octahedron is a parallelepiped which tiles $\mathbb{E}^3$. To see this, observe that the regular tetrahedron T has a description as the convex hull of four vertices of a cube. The regular octahedron also has a description as the convex hull of 8 lattice points of $\mathbb{Z}^3 \subset \mathbb{E}^3$. With this description, P is the parallelepiped with sides given by the vectors $(1, 1, 0), (1, 0, 1), (0, 1, 1)$.
2. If C denotes the regular ideal hyperbolic cube, and T denotes the regular ideal hyperbolic tetrahedron, then $[C] = 5[T]$ in $\mathcal{P}(\mathbb{H}^3)$. One of the copies of T is found in

C by taking the convex hull of an alternating set of vertices, as above. The complement of T in C consists of 4 ideal tetrahedra, each of which has an order 3 symmetry inherited from rotations of C about its long diagonals, which implies that they are regular.

3. If D denotes the regular hyperbolic dodecahedron with dihedral angles $\frac{\pi}{2}$ and C denotes the regular hyperbolic cube with dihedral angles $\frac{2\pi}{5}$ then $2[D] = 5[C]$. Consider the tessellation of $\mathbb{H}^3$ by cubes isometric to C. Then the dual tessellation contains totally right–angled geodesic pentagons which meet each other at right angles; in particular, these must bound regular right angled dodecahedra. After subdividing the cubes into octants, one sees that one can rearrange 20 such octants to make a copy of D.

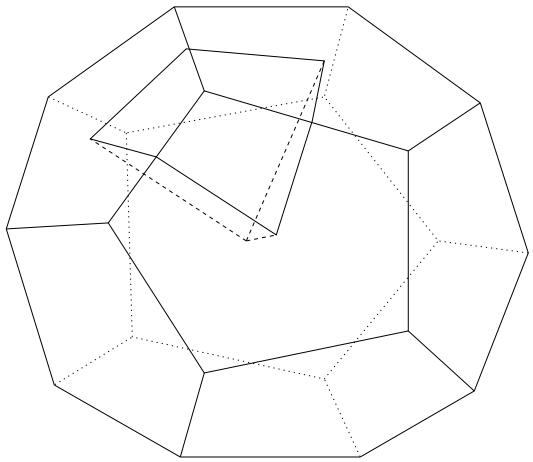

FIGURE 3. A regular right–angled hyperbolic dodecahedron is proportional in $\mathcal{P}(\mathbb{H}^3)$ to the regular hyperbolic cube with dihedral angles $\frac{2\pi}{5}$.

1.5. **Ideal Scissors Congruence.** In hyperbolic space there are a special class of polyhedra that one might want to consider — namely, the *ideal* polyhedra. An ideal polyhedron is one whose vertices all lie at infinity. There is a special scissors congruence group in $\mathbb{H}^n$ which is generated by the ideal simplices, and subject to the relation that the only decompositions allowed are those decompositions which are "boundaries of higher dimensional simplices". This cryptic remark is meant to be illustrated by the following examples:

1. Lay an ideal 3–simplex flat in $\mathbb{H}^2$ so that its vertices go to 4 distinct points in $\partial\mathbb{H}^2$. Then the ideal quadrilateral bounded by these vertices can be decomposed into a pair of ideal triangles in two distinct ways.
2. Lay an ideal 4–simplex flat in $\mathbb{H}^3$ so that its vertices go to 5 distinct points in $\partial\mathbb{H}^3$. Then the ideal polygon bounded by these points can be decomposed into a pair, or a triple of ideal tetrahedra. This is known as the 2–3 move, or the *cocycle* move.

There are analogous constructions in all dimensions.

Let $\widehat{\mathcal{P}}(\mathbb{H}^3)$ denote the group with these generators and relations in dimension 3. The Dehn invariant can be defined for ideal tetrahedra just as for regular tetrahedra, by truncating cusps with horoballs. In fact, there is a stronger *complex Dehn invariant* for this restricted scissors congruence group, which takes the value $z \wedge (1-z)$ on an ideal simplex with cross–ratio z. We are interested in the kernel of the Dehn invariant restricted to

$\widehat{\mathcal{P}}(\mathbb{H}^3)$. Denote this group by $\mathcal{B}(\mathbb{C})$. This is known as the *Bloch group* in dimension 3, and was actually invented for number theoretic reasons. More generally, if we consider the ideal scissors congruence group generated by ideal simplices with cross–ratio in a number field k, one can form the Bloch group $\mathcal{B}(k)$.

Let M be a non–compact, finite volume hyperbolic 3–manifold. Then a theorem of Thurston implies that it can be dissected into a finite collection of ideal tetrahedra. These tetrahedra determine an element of $\mathcal{B}(\mathbb{C})$, (and even $\mathcal{B}(k)$ for some number field k determined by M) since polyhedra which tile hyperbolic space periodically have zero Dehn invariant, as above. In fact, the element determined does *not* depend of the triangulation of M. This can be seen in a number of ways — directly, by a theorem of Piergallini which states that any two ideal triangulations of a 3–manifold are related by a series of 2–3 moves, or indirectly, by identifying $\mathcal{B}(k)$ with a quotient of $H_3(PSL(2,\mathbb{C})^\delta,\mathbb{Z})$ and seeing that M determines an element in $H_3(\Gamma,\mathbb{Z})$ where $M = \mathbb{H}^3/\Gamma$. More explicitly, an ideal tetrahedron determines a 3–chain in $C_*(PSL(2,\mathbb{C})^\delta,\mathbb{Z})$ by looking at the transformations in $PSL(2,\mathbb{C})$ which permute a set of flags, one for each vertex. Roughly, the condition that the complex Dehn invariant vanish implies that this chain is closed, and the 2–3 move quotients out by boundaries of 4–chains.

There is a refined, orientation sensitive invariant of $\mathcal{B}(\mathbb{C}) \to \mathbb{R}/\mathbb{Q}$ which, when evaluated on $[M]$, is just the Chern–Simons invariant. This is the real part of the Rogers dilogarithm function, evaluated on the cross–ratios of a collection of simplices determining the element of $\mathcal{B}(\mathbb{C})$. The ambiguity in which branch of the dilogarithm to choose leads to the indeterminacy of $\mathbb{R}/\mathbb{Q}$.

1.6. **More general dissections, and the Banach–Tarski paradox.** If we allow the pieces in a decomposition of a space to be arbitrary, more exotic equivalences result. To be precise, we fix a space X with a G–action for G some group.

Definition 1.8. Two subsets $A, B \subset X$ are *G–equidecomposable* (or just *equidecomposable* if there is no ambiguity) if there are finite partitions

$$A = \coprod_{i=1}^{n} A_i, B = \coprod_{i=1}^{n} B_i$$

and elements $\alpha_i \in G$ such that $\alpha_i(A_i) = B_i$. We write $A \sim_G B$ or, if there is no ambiguity, $A \sim B$.

If A is equidecomposable with a subset of B, we write, $A \prec B$. We prove a couple of basic lemmas.

Lemma 1.9. *If $A \prec B$ and $B \prec A$ then $A \sim B$.*

Proof: Let $f : A \to B$ take A to the subset of B with which it is equidecomposable, and let $g : B \to A$ take B to the subset of A with which it is equidecomposable. Define $A_1 = A/g(B)$ and inductively, define $A_n = gf(A_{n-1})$. Then all the A_i are disjoint, and we can define $A' = \bigcup_i A_i$. Similarly, define $B_1 = B/f(A)$ and $B_n = fg(B_{n-1})$ and $B' = \bigcup_i B_i$. Then A/A' is in the image of g, so we can apply g^{-1} to it. Clearly, $f(A') = B/B'$ and $g^{-1}(A/A') = B'$. So $A \sim B$. □

Lemma 1.10. *If X contains disjoint subsets A, B such that $A \sim X \sim B$, then $X/A \sim X$.*

Proof: We have $A \sim B \subset X/A$, so $A \prec X/A$. Also, $X/A \subset X \sim A$ so $X/A \prec A$. So $X/A \sim A \sim X$. □

Definition 1.11. Say that a space X has a *paradoxical decomposition* if there is a partition $X = A \coprod B$ such that $A \sim X \sim B$.

The content of the preceding lemma is that if a space contains two disjoint subsets which are equidecomposable with it, then that space has a paradoxical decomposition. If we need to specify the group G, we say that the space is *G–paradoxical*.

Example 1.12. If F_2 denotes the free group on two generators, then the set of elements of F_2 has a paradoxical decomposition with respect to the underlying action of the group on itself. For, say $F_2 = \langle a, b | \rangle$. Then we can let A^+ (resp. A^-) denote the set of reduced words starting with a (resp. a^{-1}) and define similarly $B^{\pm}$. Let $A = A^+ \cup A^-$ and $B = B^+ \cup B^-$. Then $A \sim F_2 \sim B$, since $F = A^+ \cup aA^- = B^+ \cup bB^-$.

This paradoxical quality of the free group can be carried over to any space on which this group acts freely.

Lemma 1.13. *Suppose F_2 acts on a set X freely (i.e. no nontrivial $w \in F_2$ has a fixed point). Then X has an F_2–paradoxical decomposition.*

Proof: The space X can be written as a disjoint union of orbits. For each orbit, choose a representative, and let $P \subset X$ be the set of representatives (Note that defining P requires the axiom of choice). Then for $T \in \{A^+, A^-, B^+, B^-\}$ we can set

$$X_T = TP = \{t(p) \text{ such that } t \in T, p \in P\}$$

Then $X_{A^+}, X_{A^-}, X_{B^+}, X_{B^-}$ are disjoint, because F_2 acts freely. Moreover,

$$X = X_{A^+} \cup aX_{A^-} = X_{B^+} \cup bX_{B^-}$$

□

We use at this point the following fact, which is proved for instance in [16].

Fact 1.14. *The group $SO(3, \mathbb{R})$ of 3×3 matrices A with real entries and determinant 1 satisfying $A^t A = id$ contains a subgroup isomorphic to F_2.*

Suppose G is such a subgroup of isometries of the sphere. Then each nontrivial $g \in G$ is a rotation, and therefore fixes exactly two points. Let $D \subset S^2$ be the union of the fixed points of nontrivial elements of G. Then D is a countable set.

Moreover, lemma 1.13 implies that $S^2 - D$ is paradoxical for F_2, and therefore for $SO(3, \mathbb{R})$.

Now, let l be a line passing through the origin in 3–space which avoids the set D. Then since D is countable, there is a rotation r with axis l with the property that $r^n(D)$ is disjoint from D for all $n \neq 0$. Set

$$\overline{D} = \bigcup_{n=0}^{\infty} r^n(D)$$

Then $r(\overline{D}) = \overline{D} - D$. It follows that $\overline{D} \sim \overline{D} - D$ and therefore we have

$$S^2 = (S^2 - \overline{D}) \cup \overline{D} \sim (S^2 - \overline{D}) \cup (\overline{D} - D) = (S^2 - D)$$

Since $S^2 - D$ is $SO(3, \mathbb{R})$–paradoxical, it follows from lemma 1.10 that S^2 is $SO(3, \mathbb{R})$–paradoxical; that is: *the sphere can be decomposed into finitely many pieces which can be rearranged by rotations to form two spheres of equal size*. This statement is often known as the *Banach–Tarski paradox*, although evidently there is nothing paradoxical about it.

By contrast, subsets of the Euclidean plane behave much better under equidecomposition with respect to $\mathrm{Isom}(\mathbb{E}^2)$. In particular, there is the following theorem of Banach, improving a result of Hilbert:

Theorem 1.15 (Banach)**.** *For $n = 1, 2$ Lebesgue measure can be extended to* all *subsets of $\mathbb{R}^n$ as a* finitely *additive and isometry–invariant measure.*

In particular, any two equidecomposable measurable subsets of $\mathbb{R}^1$ or $\mathbb{R}^2$ have the same measure, even if the pieces into which they are decomposed are not measurable. The reason for the distinct relationships between equidecomposability and measure in the plane and the sphere have to do with a property of $\mathrm{Isom}(\mathbb{E}^2)$ known as *amenability* which is not satisfied for $\mathrm{Isom}(\mathbb{S}^2)$.

Definition 1.16. A group G is said to be *amenable* if there is a finitely additive measure μ defined on all subsets of G such that $\mu(G) = 1$ and $\mu(gA) = \mu(A)$ for every $A \subset G$ and $g \in G$.

The group $\mathrm{Isom}(\mathbb{E}^2)$ is amenable, whereas the group $\mathrm{Isom}(\mathbb{S}^2)$ is not.

Exercise 1.17. *Show that $Isom(\mathbb{E}^2)$ does not contain a subgroup isomorphic to F_2. Hint: show $Isom^+(\mathbb{E}^2)$ admits a homomorphism to an Abelian group with Abelian kernel, and therefore the same is true of any subgroup.*

It is clear that a group which is paradoxical with respect to its action on itself is not amenable. For instance, F_2 is not amenable, nor is any group which contains it as a subgroup. There is a converse, proved by Tarski:

Theorem 1.18 (Tarski)**.** *Suppose G acts on X and $E \subset X$. Then there is a G–invariant finitely additive measure defined on every subset of X such that $\mu(E) = 1$ if and only if E is not G–paradoxical.*

We will study the amenability property in more detail in the sequel, in the context of groups of discrete isometries of $\mathbb{S}^n, \mathbb{E}^n, \mathbb{H}^n$ for $n = 2, 3$.

2. 3–DIMENSIONAL HYPERBOLIC GEOMETRY

2.1. **Spaces, structures, and stiffenings.** A space X can have more or less structure on it. Here is a very incomplete list of the kinds of structures a space can have:

1. A space can have a *topology* — a notion of closeness or of convergence of sequences in the space. Formally, a topology on a space X consists of picking out a collection of distinguished subsets of X satisfying certain properties, called the *open sets*. The open sets must have the property that every union of open sets is open, and every finite intersection of open sets is open; also, the empty set and the whole space are required to be open. For the spaces we will encounter, a topology on a space is tantamount to defining what maps into and out of the space are continuous. A map between topological spaces is *continuous* in general if the inverse image of every open set is open. It is a *homeomorphism* if it is continuous and has a continuous inverse. For more details on basic topology, see [6].

   The spaces we will encounter in the sequel will typically be *manifolds*. A topological space X is a *manifold* if it is *Hausdorff* — that is, if every two distinct points are contained in disjoint open sets — and if every point is contained in some open set which is homeomorphic to an open ball in $\mathbb{R}^n$ for some (typically constant) n. The index n is known as the *dimension* of the manifold, and is constant on connected components. We will be very interested in 3–manifolds, spaces locally modelled on $\mathbb{R}^3$.

2. A manifold M can have a *smooth structure*. This amounts to a particular choice of open sets called *charts* U_i and distinguished homeomorphisms $\phi_i : U_i \to \mathbb{R}^n$ from the charts to subsets of $\mathbb{R}^n$, satisfying the properties that the union of the charts should be all of M, and that if $V \subset U_i \cap U_j$, the map $\phi_i \circ \phi_j^{-1} : \phi_j(V) \to \phi_i(V)$ is *smooth* (i.e. infinitely differentiable). Notice that we can add to the collection U_i any new set U and map $\phi : U \to \mathbb{R}^n$ which is compatible with the other charts in the sense above. A smooth structure allows us to do Calculus on M: if $f : M \to \mathbb{R}$ is a function, we know what it should mean for f to be a smooth function. In particular, for any i, the composition $f \circ \phi_i^{-1} : \phi_i(U_i) \to \mathbb{R}$ should be smooth. For every $p \in M$, there is a well–defined tangent space, which is defined as $T_{\phi_i(p)}\mathbb{R}^n$ for any i. For distinct i, j with $p \in U_i \cap U_j$, the map
$$d(\phi_j \circ \phi_i^{-1}) : T_{\phi_i(p)}\mathbb{R}^n \to T_{\phi_j(p)}\mathbb{R}^n$$
is an isomorphism, so the vector space structure is well–defined. For more details on basic differential topology, see [17].
3. A manifold M can have a *Riemannian metric*. This is a path–metric in the sense of metric space. A Riemannian metric consists of a choice of inner–product structure $\langle \cdot, \cdot \rangle_p$ on each tangent space T_pM which varies smoothly with p. Recall that an inner–product structure on a vector space V is a bilinear form $\phi : V \times V \to \mathbb{R}$ satisfying $\phi(v, w) = \phi(w, v)$, and $\phi(v, v) > 0$ for all nonzero v. The *length* or *norm* of a vector v is $\phi(v, v)^{1/2}$, and is denoted $|v|$. Then for X, Y smooth vector fields, $\langle X(p), Y(p) \rangle_p$ should be a smooth function of p. Then if $\gamma : I \to M$ is a smooth path, the *length* of γ is defined to be
$$\text{length}(\gamma) = \int_0^1 |\gamma'(t)| dt$$
where $\gamma'(t)$ denotes the tangent vector to γ at the point t. It is an easy exercise to see that this is independent of the parameterization of γ. Then the distance between p and q is defined to be
$$d(p, q) = \inf_\gamma \text{length}(\gamma)$$
where the infimum is taken over all smooth paths from p to q.
4. Notice that a Riemannian structure tells us more than just the length of (piecewise) smooth curves. For $u, v \in T_pM$, the *angle* between these two vectors is defined to be
$$\angle(u, v) = \cos^{-1} \left\langle \frac{u}{|u|}, \frac{v}{|v|} \right\rangle$$
If $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle'$ are *proportional* — that is, if there is a $\eta > 0$ such that
$$\langle u, v \rangle = \eta \langle u, v \rangle'$$
for all u, v then the angle between u and v is the same as measured with respect to $\langle \cdot, \cdot, \rangle$ as with $\langle \cdot, \cdot \rangle'$.

   A *conformal structure* is an equivalence class of Riemannian metrics, where $\langle \cdot, \cdot \rangle' \sim \langle \cdot, \cdot \rangle$ if there is a smooth function $\eta(p)$ such that
$$\langle u, v \rangle = \eta(p) \langle u, v \rangle'$$
for all $u, v \in T_pM$. Lengths are not well–defined on a manifold with a conformal structure, but angles are well–defined. This may not sound like a very rigid structure,

and in fact in dimension two conformal structures are very flexible. But in dimensions 3 and greater, conformal structures are very rigid, and it is saying a great deal when two conformal manifolds M, N are (locally) conformally isomorphic.

There is a natural partial ordering on structures. If a structure s can be obtained from a structure s' by forgetting some information, we say s is a *stiffening* of s'. For the examples we have worked out, there is the following ordering:

topological structure $\to$ smooth structure $\to$ conformal structure $\to$ Riemannian metric

For a manifold M with a structure s, the group of *automorphisms* $\mathrm{Aut}(M, s)$ of M, s is the group of transformations of M as a set which preserve the structure s. So for M a topological manifold, this is the group of self–homeomorphisms; for M a manifold with a conformal structure, this is the set of smooth self–homeomorphisms (also called *diffeomorphisms*) which preserve angles. If s is a stiffening of s', the group $\mathrm{Aut}(M, s)$ is naturally a subgroup of $\mathrm{Aut}(M, s')$. Structures can be defined implicitly, simply by specifying the group $\mathrm{Aut}(M, s)$. Of course for $\mathrm{Aut}(M, s)$ some arbitrary group G, it might be hard to think of a geometric interpretation of the structure s.

A *geometry* will consist of a space X together with a structure s, and the group $\mathrm{Aut}(X, s)$ of automorphisms of this structure. Equivalently, we could just give X and a group G acting on X and define the structure implicitly. If a structure is understood, we denote the group of automorphisms as $\mathrm{Aut}(X)$.

Definition 2.1. A geometric space is *homogeneous* if the group $\mathrm{Aut}(X)$ is transitive.

Definition 2.2. A Riemannian manifold M is *isotropic* if the group $\mathrm{Aut}(X)$ is transitive on orthonormal frames. i.e. if $v_1, v_2 \dots v_n$ is an orthonormal frame at $T_p(M)$ and $w_1, w_2 \dots w_n$ is an orthonormal frame at $T_q(M)$, there is a $\phi \in \mathrm{Aut}(X)$ such that $d\phi(v_i) = w_i$. Informally, a space is isotropic if it looks the same from any point in any direction.

Theorem 2.3. *Let M be a homogeneous 3–dimensional Riemannian manifold, and suppose Aut(M) is transitive on UTM, the set of unit tangent vectors. Then M is transitive on oriented orthonormal frames. If it contains an orientation–reversing automorphism, it is transitive on orthonormal frames — that is, M is isotropic.*

Proof: Let $p \in M$ be arbitrary. Since $\mathrm{Aut}(M)$ is transitive on UTM, the subgroup $\mathrm{stab}(p) \subset O(T_pM)$ is transitive on $UT_p(M)$. In particular, it is at least two–dimensional. Since $\mathrm{stab}(p)$ is a subgroup, it passes through $e \in O(T_pM)$, the identity. Let $X, Y \in T_e\mathrm{stab}(p)$ be two independent vectors, tangent to rotations r_X, r_Y in axes l_X, l_Y. Let $r_X(t)$ denote the rotation in the same axis as r_X through angle t, so that there are 1–parameter families of rotations $r_X(t), r_Y(t)$ in axes very close to l_X, l_Y for small t. Then $Z(t) = r_X(t)r_Y(t)r_X(-t)r_Y(-t)$ is a rotation in an axis $l_Z(t)$. But $l_Z(t)$ is asymptotically orthogonal to l_X, l_Y as $t \to 0$. In particular, the dimension of $\mathrm{stab}(p)$ at e is at least 3. Since $O(T_pM)$ is 3–dimensional, $\mathrm{stab}(p)$ is either $O(T_pM)$ or $SO(T_pM)$. □

2.2. **The geometry of $\mathbb{E}^3$.** Since $\mathbb{E}^3$ is isotropic, the stabilizer $\mathrm{stab}(p) \cong O(3, \mathbb{R})$. An algebraic feature of the group of isometries of $\mathbb{E}^3$ which distinguishes Euclidean geometry from spherical or hyperbolic geometry is the existence of a homomorphism

$$\rho : Isom(\mathbb{E}^3) \to O(3, \mathbb{R})$$

where $\rho(\alpha)$ can be thought of as the *rotational part* of α. The homomorphism ρ is defined as follows: let (e_1, e_2, e_3) be an orthonormal frame at $T_p\mathbb{E}^3$ for some point $p \in \mathbb{E}^3$. Let t_α be the unique translation of $\mathbb{E}^3$ taking $\alpha(p)$ to p. Then $t_\alpha\alpha \in \mathrm{stab}(p)$. One can check that

if α, β are arbitrary, then $t_{\alpha\beta}\alpha\beta = t_\alpha \alpha t_\beta \beta$. The kernel of ρ is exactly the subgroup T of translations. So there is a short exact sequence

$$0 \to T \to \mathrm{Isom}(\mathbb{E}^3) \to O(3,\mathbb{R}) \to 0$$

This short exact sequence is *split*. That is, there is a section $s : O(3,\mathbb{R}) \to \mathrm{Isom}(\mathbb{E}^3)$ with the property that $s\rho : O(3,\mathbb{R}) \to O(3,\mathbb{R})$ is the identity.

Another way of saying this is that $\mathrm{Isom}(\mathbb{E}^3)$ is a semi–direct product

$$\mathrm{Isom}(\mathbb{E}^3) \cong T \rtimes O(3,\mathbb{R})$$

That is, the group consists of the pairs (t, r) with $t \in T, r \in O(3,\mathbb{R})$ and multiplication rule

$$(t_1, r_1)(t_2, r_2) = (t_1 r_1 t_2 r_1^{-1}, r_1 r_2)$$

Lemma 2.4. *Let $\alpha \in Isom^+(\mathbb{E}^3)$ with $\rho(\alpha) \neq e$. Then either α fixes some point and is a rotation centered at that point, or it is the composition of a rotation r and a translation t such that the axis of r is parallel to the axis of t.*

Proof: We know from the structure of $\mathrm{Isom}^+(\mathbb{E}^3)$ that $\alpha = tr$ for some translation t and rotation r. We show that we can choose t, r with the required properties.

If $t = \mathrm{id}$ there is nothing to prove.

Otherwise, the axis l of r is well–defined. Let π be a plane perpendicular to l. Then r fixes π and acts as a rotation of it. Let $c \in \pi$ be the center of this rotation. Then we can write $t = t_1 t_2$ where t_1 is a translation in the direction of l and t_2 fixes π. So $\alpha = tr = t_1 t_2 r$. The composition $t_2 r$ fixes π. It has a nontrivial rotational part, so it is a rotation r' of π, fixing c' say. Since r' fixes π, its axis is perpendicular to π, so is parallel to l. Then the expression $\alpha = t_1 r'$ has the required properties. □

A transformation α with $\rho(\alpha)$ nontrivial which is not itself a rotation is called a *screw motion*.

2.3. **Conformal transformations.** Recall that a conformal structure is an equivalence class of Riemannian metrics, where two metrics are equivalent if they are proportional on every T_pM. Isometries of $\mathbb{E}^n$ are conformal, and so are dilations. Another important class of conformal transformations, defined now only on complements of points in $\mathbb{E}^n$, are *inversions*.

Definition 2.5. If S is a sphere of radius r centered at $p \in \mathbb{E}^n$ the inversion in S, denoted i_S, is a conformal transformation

$$i_S : \mathbb{E}^n - p \to \mathbb{E}^n - p$$

defined as follows. For $x \in \mathbb{E}^n - p$ there is a unique ray r_x from p which passes through x. Then $i_S(x)$ is the unique point on r_x with the property that $d(x,p)d(i_S(x),p) = r^2$.

It follows from the definition that i_S fixes S pointwise and interchanges the interior and exterior. It is easy to see that i_S takes spheres and planes to spheres and planes. In particular, if π is a plane, $i_S(\pi)$ is a sphere passing through p, and *vice versa*. Notice that S is not the only sphere taken to itself by i_S: if T is a sphere perpendicular to S, then the sphere $i_S(T)$ is also perpendicular to S (because i_S is conformal) and intersects S in the same circle (because i_S fixes every point in S). So $i_S(T) = T$.

In two dimensions, conformal structures are very flexible. If U, V are any two open domains in $\mathbb{C}$ bounded by topological circles, there is a conformal map from U and V. However in higher dimensions, conformal structures are much more rigid, even locally. In this context, there is the following theorem of Liouville.

Theorem 2.6 (Liouville)**.** *Let U, V be open sets in $\mathbb{E}^n$ with $n \geq 3$ and suppose $\phi : U \to V$ is conformal. Then ϕ is of the form*

$$\phi : x \to \lambda A i(x) + t$$

where $\lambda > 0$ is a dilation, t is a translation, A is orthogonal, and i is either an inversion in a sphere or the identity.

Sketch of proof: We say that ϕ is of *type (a)* if it is of the form $\lambda A + b$ and of *type (b)* if it is of the form $\lambda A i + b$ for i an inversion.

By the definition of a conformal map, there is a positive valued function μ on U such that

$$|df(v)| = \mu(x)|v|$$

for all $x \in U$ and $v \in T_x\mathbb{E}^n$. Define $\rho(x) = \frac{1}{\mu(x)}$.

It is easy to see that ϕ is determined, up to composition with an isometry of $\mathbb{E}^n$, by the function ρ. Furthermore, it is pretty easy to see that ϕ is of type (a) iff ρ is constant, and ϕ is of type (b) iff $\rho = \eta|x - x_0|^2$ for some non–zero constant η and basepoint x_0.

We introduce some notation as follows. For each $p \in U$,

$$d\phi : T_pU \to T_{\phi(p)}V$$

denotes the induced map on tangent spaces. For $u \in T_pU$, we have

$$d(d\phi) : T_uT_pU \to T_{d\phi(u)}T_{\phi(p)}V$$

Since T_pU is a vector space, there is a canonical isomorphism $T_uT_pU \cong T_pU$ for each u. So there is a well–defined map

$$d^2\phi : T_pU \times T_pU \to T_{\phi(p)}V$$

Since partial derivatives commute, $d^2\phi$ is *symmetric* in its variables.

Suppose $u \perp w$. We show that $d^2\rho(u, w) = 0$. Let v be another vector orthogonal to both u and w. We extend u, v, w to vector fields locally which are constant in local co–ordinates. Now, we have the identity

$$\langle d_\phi(u), d_\phi(v)\rangle = 0$$

Taking the partial derivative of both sides with respect to w, we get

$$\langle d^2\phi(u, w), d\phi(v)\rangle = -\langle d\phi(u), d^2\phi(v, w)\rangle$$

Note that the left side is symmetric in u, w and the right side is symmetric in v, w. On the other hand the equality implies both sides are antisymmetric in u, v. So both sides must be identically zero.

Since v is arbitrary, it follows that there are functions α, β such that

$$d^2\phi(u, w) = \alpha(x)d\phi(u) + \beta(x)d\phi(w)$$

By definition, we have

$$\langle d\phi(u), d\phi(u)\rangle = \mu(x)^2\langle u, u\rangle$$

Taking the partial derivatives of both sides with respect to w, we get

$$\langle d^2\phi(u, w), d\phi(u)\rangle = \mu(x)d\mu(w)|u|^2$$

Notice that $|u|$ is constant as a function of x by our assumption.

It follows that

$$\alpha(x) = \frac{d\mu(w)}{\mu(x)}$$

and

$$\beta(x) = \frac{d\mu(u)}{\mu(x)}$$

Since $d\rho(z) = \frac{d\mu(z)}{\mu(x)^2}$ we have

$$\rho(x)d^2\phi(u,w) + d\rho(w)d\phi(u) + d\rho(u)d\phi(w) = 0$$

Taking partial derivatives with respect to v we get

$$d\rho(v)d^2\phi(u,w) + \rho(x)d^3\phi(u,w,v) + d^2\rho(w,v)d\phi(u) +$$

$$d\rho(w)d^2\phi(u,v) + d^2\rho(u,v)d\phi(w) + d\rho(u)d^2\phi(w,v) = 0$$

Now, terms 2,4,5 are symmetric in u, v. Moreover the sum of terms 1 and 6 is also symmetric in u, v. It follows that term 3 is symmetric in u, v. That is,

$$d^2\rho(w,v)d\phi(u) = d^2\rho(w,u)d\phi(v)$$

Since $d\phi(u)$ and $d\phi(v)$ are orthogonal, $d^2\rho(w,u) = 0$.

It follows that at each point, $d^2\rho$ is a symmetric bilinear form, and if e_i is the canonical basis for $\mathbb{E}^n$, $d^2\rho(e_i, e_j) = k_i\delta^i_j$ for some constants k_i. Moreover

$$0 = d^2\rho(e_i + e_j, e_i - e_j) = k_i - k_j$$

so these constants are equal, and $d^2\rho$ is k times the usual Euclidean inner product, and we can write

$$d^2\rho(u,w) = \eta(x)\langle u, w\rangle$$

for some $\eta(x)$.

Taking the partial derivative of this expression with respect to v, we get

$$d^3\rho(w,u,v) = d\eta(v)\langle w, u\rangle$$

where now w, u, v are arbitrary. Since d^3 is symmetric, we have

$$\langle d\eta(v)w - d\eta(w)v, u\rangle = 0$$

which implies that $d\eta(v)w = d\eta(w)v$ and therefore $d\eta = 0$; i.e. η is a constant.

It follows that $d^2\rho(u,w) = \eta\langle u, w\rangle$ for some constant η, and therefore ρ is a quadratic function. By examining the possibilities, it is easy to show ϕ is either of type (a) or of type (b). □

Remark 2.7. For $U \subset M$ a Riemannian manifold with an arbitrary metric, the derivative

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

where $\nabla_X Y$ is the *covariant derivative* of the vector field Y in the direction X. In general this expression is *not* symmetric in X and Y unless $[X, Y] = 0$. This condition cannot necessarily be satisfied simultaneously with the condition $X \perp Y$.

2.4. **The geometry of $\mathbb{H}^3$.** One can define 3–dimensional hyperbolic geometry implicitly as $(\mathbb{D}^3, G)$ where G is the group of conformal transformations of the unit ball $\mathbb{D}^3$ to itself. By Liouville's theorem, one can show that G consists of transformations of the form Ai_S where $A \in O(3,\mathbb{R})$ and i_S is either the identity or inversion in a sphere perpendicular to the unit sphere.

First, observe that G is transitive. We investigate the point stabilizers.

Lemma 2.8. *Let $p \in \mathbb{D}^3$ be arbitrary. Then stab$(p) \subset G$ is isomorphic to $O(3,\mathbb{R})$.*

Proof: Since G is transitive, we can assume without loss of generality that $p = 0$. Now, if $\phi \in \text{stab}(p)$ we write $\phi(x) = \lambda Ai(x) + t$. Since ϕ takes the center of S to the center of S, it follows that i is the identity. For, if S, T are arbitrary spheres with distinct centers, the inversion in T takes S to a new sphere $i_T(S)$ and if c is the center of S, $i_T(c)$ is *not* the center of $i_T(S)$. Let d denote the center of T. One way to see this is to look at the effect of i_T on the ray r_c from d through c. Then the ray r_c intersects S at points p, q with $d(p,c) = d(c,q)$. However, we have

$$d(d, i_T(p)) = \frac{r^2}{d(d,p)}, d(d, i_T(q)) = \frac{r^2}{d(d,q)}, d(d, i_T(c)) = \frac{r^2}{d(d,c)}$$

so in particular $d(i_T(p), i_T(c)) \neq d(i_T(q), i_T(c))$. By definition, ϕ preserves the unit sphere and its center, so i is the identity.

Thus $\phi(t) = \lambda A(x) + t$. Since it preserves the unit sphere, $\lambda = 1$ and $t = 0$. Thus $\phi \in O(3,\mathbb{R})$. The lemma follows. □

We define the hyperbolic metric at the origin by declaring it to be equal to 2 times the Euclidean metric there. That is,

$$\langle u, v\rangle^{\text{hyp}}_{T_0\mathbb{D}^3} = 4\langle u, v\rangle^{\text{Euc}}_{T_0\mathbb{D}^3}$$

Now, for any $p \in \mathbb{D}^3$ we define the hyperbolic inner product on $T_p\mathbb{D}^3$ by

$$\langle u, v\rangle^{\text{hyp}}_{T_p\mathbb{D}^3} = \langle d\alpha(u), d\alpha(v)\rangle^{\text{hyp}}_{T_0\mathbb{D}^3}$$

where $\alpha \in G$ takes p to 0.

To check that this is well–defined, suppose β is some other element with $\beta(p) = 0$. Then $\alpha\beta^{-1} \in \text{stab}(0)$ and therefore the action of $d(\alpha\beta^{-1}) = d\alpha d(\beta^{-1})$ on $T_0\mathbb{D}^3$ preserves lengths of vectors in the Euclidean metric there, and therefore also in the hyperbolic metric. So

$$\langle d\alpha(u), d\alpha(v)\rangle^{\text{hyp}}_{T_0\mathbb{D}^3} = \langle d\beta d(\alpha^{-1})d\alpha(u), d\beta d(\alpha^{-1})d\alpha(v)\rangle^{\text{hyp}}_{T_0\mathbb{D}^3} = \langle d\beta(u), d\beta(v)\rangle^{\text{hyp}}_{T_0\mathbb{D}^3}$$

and the inner product at $T_p\mathbb{D}^3$ is well–defined. In the sequel we will drop the subscripts and superscripts from our inner products if it is clear where and how they are defined.

A similar calculation shows that this metric is invariant under the action of G. That is, if $u, v \in T_p\mathbb{D}^3$ and $\alpha \in G$ is arbitrary,

$$\langle u, v\rangle = \langle d\alpha(u), d\alpha(v)\rangle$$

Exercise 2.9. *Verify the claim made in the last paragraph.*

Since stab$(0) = O(3,\mathbb{R})$ preserves the Euclidean length of vectors, and acts transitively on the set of points at Euclidean distance r from 0, there are the following consequences:

1. The hyperbolic distance from 0 to p only depends on the Euclidean distance r from 0 to p.

2. The ratio of the Euclidean to the hyperbolic metric on $T_p\mathbb{D}^3$ only depends on the Euclidean distance r from 0 to p.

Exercise 2.10. *Show that*

$$\langle u, v\rangle^{hyp} = \frac{4}{(1-r^2)^2}\langle u, v\rangle^{Euc}$$

for $u, v \in T_p\mathbb{D}^3$ where the Euclidean distance from 0 to p is r. In particular,

$$|u|^{hyp} = \frac{2}{1-r^2}|u|^{Euc}$$

2.4.1. *Geodesics and totally geodesic planes.*

Definition 2.11. In a Riemannian manifold, a curve $\alpha : I \to M$ is a *geodesic* if it is *locally distance minimizing*. That is, there is a fine mesh $p_i \in I$ so that

$$d_M(\alpha(p_i), \alpha(p_{i+1})) = \text{length}(\alpha([p_i, p_{i+1}]))$$

Note that any (globally) distance–minimizing curve, if it exists, is a geodesic.

We have the following basic theorem of Riemannian geometry:

Theorem 2.12. *Let M be a Riemannian manifold. For any $p \in M$ there exists a neighborhood U of p such that for all $q \in U$, there is a unique distance–minimizing geodesic from p to q.*

Theorem 2.13. *Let $v \in T_pM$ be arbitrary. Then there is a unique geodesic emanating from p which is tangent to v at p. If M is* complete *as a metric space, any geodesic can be continued indefinitely in either direction.*

For a proof, see [2].

Theorem 2.14. *Suppose M is a Riemannian manifold and $i : M \to M$ is an isometry. Then the fixed point set of i consists of totally geodesic components.*

Sketch of proof: If some fixed points of i are isolated, they are totally geodesic 0–dimensional components. Suppose α is a 1–dimensional component, and p, q are two close points on α. Then if γ is the unique distance minimizing geodesic between p and q, the image of γ under i is another curve of the same distance between p and q, so it must necessarily be γ. Moreover, if $r \in \gamma$ is such that $d(p, r) = t$, the image $i(r) \in \gamma$ must also satisfy $d(p, i(r)) = t$ since $d(p, i(r)) = d(i(p), i(r)) = d(p, r)$ since i is an isometry. Hence $r = i(r)$ and γ is fixed by i. Since the fixed set is 1–dimensional near p, γ is a segment of α. So α is a geodesic. A similar argument works in higher dimensions. □

Let S be a sphere perpendicular to the unit sphere in $\mathbb{E}^3$. Then i_S is a conformal transformation of the unit ball to itself, and therefore an isometry of the hyperbolic metric. The fixed point set of i_S is exactly the intersection $S \cap \mathbb{D}^3$. If S_1, S_2 are two spheres perpendicular to the unit sphere and transverse to each other, their intersection is a circle or straight line perpendicular to the unit sphere. It follows that the arcs of circles and straight line segments in $\mathbb{D}^3$ perpendicular to the boundary are geodesics. By the uniqueness of geodesics passing through a point with a given tangent vector, one sees that these are exactly the set of all geodesics. That is

Theorem 2.15. *The geodesics in $\mathbb{D}^3$ with the hyperbolic metric are the arcs of circle and straight line segments perpendicular to the boundary.*

2.5. **Matrix models of Isom$^+(\mathbb{H}^3)$.**

2.5.1. *The conformal models.* In the conformal upper half–space model of $\mathbb{H}^3$, the sphere at infinity S^2_∞ can be identified with the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \infty$. The isometries of $\mathbb{H}^3$ are the restriction of conformal transformations of $\mathbb{R}^3 \cup \infty$ which fix the upper half–space, and therefore they restrict to conformal transformations of $\widehat{\mathbb{C}}$. A conformal transformation ϕ from $\widehat{\mathbb{C}}$ to itself is necesarily analytic. One can easily check that the group of analytic isomorphisms from $\widehat{\mathbb{C}}$ to itself is exactly the group $PSL(2,\mathbb{C})$ consisting of 2×2 complex matrices of determinant 1 modulo the subgroup $\pm$Id, and acting on $\widehat{\mathbb{C}}$ by transformations

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} : z \to \frac{\alpha z + \beta}{\gamma z + \delta}$$

Since every complex number has a complex square root, there is an isomorphism between the groups $PSL(2,\mathbb{C}) = PGL(2,\mathbb{C})$ where $PGL(2,\mathbb{C})$ denotes the group of 2×2 complex matrices of nonzero determinant modulo the subgroup of matrices which are scalar multiples of the identity; that is, matrices of the form $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$.

With this notation, we can think of $\widehat{\mathbb{C}}$ as the *complex projective line* $\mathbb{CP}^1$. That is, we can think of it as the equivalence space $(\mathbb{C}^2 - 0)/\mathbb{C}^*$, the space of ratios of pairs of complex numbers which are not both zero. Geometrically, this is the space of complex lines in $\mathbb{C}^2$. Two points (z_1, w_1) and (z_2, w_2) both not the origin are on the same complex line iff there is some $\lambda \in \mathbb{C}^*$ with $z_2 = \lambda z_1$ and $w_2 = \lambda w_1$. So the two ratios $(z_1 : w_1)$ and $(z_2 : w_2)$ are equal. The action of $PGL(2,\mathbb{C})$ on $\mathbb{CP}^1$ is the projectivization of the usual action of $GL(2,\mathbb{C})$ on $\mathbb{C}^2$.

These transformations extend naturally to conformal transformations of the upper half–space, and we see therefore that this group is isomorphic to $\mathrm{Isom}^+(\mathbb{H}^3)$. Thinking of the upper half space as the subspace

$$H = \{x + iy + jz \in \text{quaternions} \mid z > 0\}$$

we can extend the action of $A \in PSL(2,\mathbb{C})$ as follows:

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} : q \to (\alpha q + \beta)(\gamma q + \delta)^{-1}$$

Note that the order of multiplication of the quaternionic factors is important.

Definition 2.16. Let A be a 2×2 complex matrix with determinant 1 representing an element of $PSL(2,\mathbb{C})$. We say

- A is *elliptic* if $\mathrm{tr}(A)$ is real and contained in $(-2, 2)$.
- A is *parabolic* if $\mathrm{tr}(A)$ is real and equal to ± 2.
- A is *hyperbolic* if $\mathrm{tr}(A)$ is real and has absolute value > 2.
- A is *loxodromic* if $\mathrm{tr}(A)$ is not real.

Theorem 2.17. *The notation above characterizes the nature of the action on $\mathbb{H}^3$ of an element of $\mathrm{Isom}^+(\mathbb{H}^3)$. A transformation is* elliptic *if it is a rotation around some axis. It is* parabolic *if it fixes a unique point in S^2_∞ and no points in the interior of $\mathbb{H}^3$. It is* hyperbolic *if it acts as a translation along some geodesic axis, and* loxodromic *if it acts as a "screw motion" along some geodesic axis.*

Proof: The conditions on the trace are invariant under $A \to -A$ so the type of a transformation is well–defined independent of the representative in $SL(2,\mathbb{C})$. Moreover, trace is invariant under conjugation. Since the equation $\alpha z + \beta = \gamma z^2 + \delta z$ always has a solution in $\widehat{\mathbb{C}}$, we can assume after conjugation that ∞ is a fixed point — that is, that our

matrix A is of the form $\begin{bmatrix} \alpha & \beta \\ 0 & \alpha^{-1} \end{bmatrix}$. Such a transformation acts on $\mathbb{C}$ by $z \to \alpha^2 z + \alpha\beta$, which is to say, by an arbitrary similarity of the standard Euclidean structure on $\mathbb{C}$. One sees that the transformation is parabolic iff $\alpha = \pm 1$, in which case it acts on $\mathbb{C}$ as a translation. It is elliptic iff $|\alpha| = 1$ and $\alpha \neq \pm 1$ in which case it acts on $\mathbb{C}$ as a rotation about some point p, and on $\mathbb{H}^3$ as a rotation about the axis joining p to ∞. It is hyperbolic iff α is real and not ± 1 in which case it acts on $\mathbb{C}$ as a dilation about some point p, and on $\mathbb{H}^3$ as a translation along the axis joining p to ∞. It is loxodromic iff it is a product of an elliptic and a hyperbolic element with the same axis, in which case it acts on $\mathbb{H}^3$ as a screw motion, translating along an axis and rotating about it simultaneously. □

The other conformal model of $\mathbb{H}^3$ is the unit ball. Define

$$B = \{x + iy + jz \in \text{quaternions} \mid x^2 + y^2 + z^2 < 1\}$$

For a quaternion $q = q_0 + q_1 i + q_2 j + q_3 k$ denote $q' = q_0 - q_1 i - q_2 j + q_3 k$ and $\overline{q} = q_0 - q_1 i - q_2 j - q_3 k$. Then define $SB(2, H)$ to be the group of 2×2 matrices with quaternionic entries of the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $d = a', b = c'$ and $a\overline{a} - c\overline{c} = 1$. The natural action of $SB(2, H)$ on B by fractional linear transformations as above has kernel exactly $\pm$Id and gives an isomorphism of $SB(2, H)/\pm \text{Id}$ with the group of conformal transformations of B — that is, with the group $Isom^+(\mathbb{H}^3)$ for the conformal unit ball model.

2.5.2. *The hyperboloid model.* Let $\langle \cdot, \cdot \rangle$ denote the symmetric bilinear form on $\mathbb{R}^{n+1}$ which in terms of a basis $e_1 \dots e_{n+1}$ is given by

$$\left\langle \sum_{i=1}^{n+1} r_i e_i, \sum_{i=1}^{n+1} s_i e_i \right\rangle = \sum_{i=1}^{n} r_i s_i - r_{n+1} s_{n+1}$$

For a vector v, define the *Lorentz norm* $\|v\| = \langle v, v \rangle^{\frac{1}{2}}$.

Let J denote the matrix

$$J = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{bmatrix}$$

The group of linear transformations of $\mathbb{R}^{n+1}$ which preserve the Lorentz norm is exactly the matrices A such that $A^t J A = J$, and is denoted $O(n, 1)$. The subgroup of determinant 1 is $SO(n, 1)$.

The set of vectors of norm i is a two–sheeted hyperboloid. Denote the positive sheet (i.e. for which the last co–ordinate is positive) by H^+, and the subgroup preserving this sheet as $SO^+(n, 1)$. Then $(H^+, SO^+(n, 1))$ is a model for hyperbolic n space $\mathbb{H}^n$.

If v is a vector, $v^\perp$ denotes the subspace of vectors w such that $\langle v, w \rangle = 0$. If v has imaginary norm, every vector in w has real (positive) norm, and the restriction of $\langle \cdot, \cdot \rangle$ to $v^\perp$ gives a standard Euclidean inner product. For $v \in H^+$ the tangent space $T_v H^+$ is a translate of $v^\perp$. To see this, observe that w is tangent to H^+ at v iff

$$\frac{d}{dt} \langle v + tw, v + tw \rangle = 0$$

Giving $T_v H^+$ the inner product $\langle \cdot, \cdot \rangle$ gives H^+ the structure of a Riemannian manifold. Since everything is invariant under the action of $SO^+(n,1)$, this must be (a scaled multiple of) the standard hyperbolic metric on H^+. It is easy to check by a direct computation that this metric on the space is given by

$$\cosh(d(v,w)) = -\langle v, w \rangle$$

for vectors v, w of norm 1.

Example 2.18. In $\mathbb{R}^{25}$ consider the "standard" form g of signature $(24,1)$. The vector

$$\lambda = (3, 5, 7, \ldots, 47, 51, 145)$$

satisfies $|\lambda|^2 = -1$, so the restriction of the form g to $L = \lambda^\perp$ is positive definite. The intersection of the cubical lattice I with L is a lattice in L called the *Leech lattice*. See [3] for more details, and many more constructions of the Leech and other lattices.

2.6. **Horospherical geometry.** In one of the conformal models of hyperbolic geometry, a Euclidean sphere whose interior is contained in $\mathbb{H}^3$ and is tangent to S^2_∞ at some point is called a *horosphere*. In the upper half–space model, a sphere tangent to ∞ is a horizontal plane. The stabilizer of a horosphere can be seen to be isomorphic to the group of isometries of $\mathbb{E}^2$, that is, to $\mathbb{R}^2 \rtimes O(2)$. The orientation–preserving subgroup preserving ∞ in the upper half–space model is exactly the group of matrices of the form $\begin{bmatrix} \theta & \alpha \\ 0 & \theta^{-1} \end{bmatrix}$, where θ is a complex number of norm 1, and $\alpha \in \mathbb{C}$ is arbitrary. The intrinsic geometry $(H, \mathrm{stab}(H))$ is thus identical to $(\mathbb{E}^2, \mathrm{Isom}(\mathbb{E}^2))$. That is, the intrinsic geometry of a horosphere is Euclidean.

A horosphere is a limit of spheres which are contained entirely in the interior of $\mathbb{H}^3$. These spheres are isometric to Euclidean spheres of various sizes; a horosphere can be thought of as a "sphere of infinite radius", and exhibits Euclidean geometry as a limit of spherical geometry in the small.

Definition 2.19. A *Busemann function* is a function $f : \mathbb{H}^3 \to \mathbb{R}$ defined as a limit

$$f(p) = \lim_{i \to \infty} \mathrm{dist}(p, p_i) - \mathrm{dist}(q, p_i)$$

for some sequence $p_i \to p_\infty \in S^2_\infty$ and fixed $q \in \mathbb{H}^3$.

A Busemann function measures the relative distance to a point $p_\infty \in S^2_\infty$. The level sets of a Busemann function are the horospheres tangent to p_∞. Thus p_∞ may be thought of as the "center" of these horospheres, and the horospheres may be thought of as the locus of points equidistant to an ideal point. One must be careful not to abuse this metaphor, but it is a useful and suggestive one.

2.7. **Volumes of ideal hyperbolic simplices.** Let Δ be an ideal tetrahedron — that is, a geodesic tetrahedron with all four vertices contained in S^2_∞. After translating Δ by an element of $PSL(2,\mathbb{C})$, we can assume that the vertices are at $0, 1, z, \infty$ for some z. For different choices of vertices to be sent to $0, 1, \infty$ the parameter z is replaced by $\frac{1}{1-z}$ or $\frac{z-1}{z}$ for an even permutation, or $\frac{1}{z}$, $\frac{1-z}{1}$, $\frac{z}{z-1}$ for an odd permutation. We say an ordering of vertices is *positively oriented* if $\mathrm{Im}(z) > 0$. The space of (possibly degenerate) oriented ideal tetrahedra is isomorphic to the quotient of $\mathbb{C} - \{0,1\}$ by the group G generated by $z \to \frac{1}{1-z}$ and $z \to \frac{1}{z}$. This group has order 6, and the quotient of the thrice punctured sphere by this group is the $(2,3,\infty)$ orbifold, which is exactly the moduli space of flat structures on a torus. The correspondence is given as follows: if $T = \mathbb{C}/\Lambda$ is a flat torus, we can form the quotient $T/\langle i \rangle$ where i is the involution $z \to -z$. The quotient is an

orbifold S which is a sphere with four orbifold points of order 2. The underlying sphere is conformally equivalent to $\mathbb{CP}^1$, and the four orbifold points pick out 4 distinct points in $\mathbb{CP}^1$.

Example 2.20. With the notation as above, what is the locus of flat structures on the torus which correspond to real simplex parameters?

If Δ is an ideal simplex with one vertex at infinity, the three faces incident to this vertex are subsets of vertical Euclidean planes, in the upper half–space model. A horosphere S centered at infinity is a horizontal plane, so the intersection of Δ with S is a Euclidean triangle. If γ is an edge of Δ incident to ∞, represented by a vertical line in this picture, a hyperbolic plane perpendicular to γ at some point $p = S \cap \gamma$ is tangent to S at that point. It follows that the hyperbolic dihedral angle between the planes incident at γ is equal to the Euclidean angle at p in the triangle $S \cap \Delta$. It follows that for three edges $\gamma_1, \gamma_2, \gamma_3$ incident to a vertex of an ideal simplex, the sum of the dihedral angles at those edges is π.

If the six dihedral angles of Δ are labelled $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ where α and α' are the angles on "opposite" edges, and so on for β, γ then we have identities

$$\alpha + \beta + \gamma = \alpha + \beta' + \gamma' = \alpha' + \beta + \gamma' = \alpha' + \beta' + \gamma = \alpha' + \beta' + \gamma' = \pi$$

It follows that $\alpha = \alpha', \beta = \beta', \gamma = \gamma'$. That is, we have proved

Lemma 2.21. *Dihedral angles of opposite edges of an ideal hyperbolic simplex are equal.*

In fact, it is easy to see from this that a permutation of the vertices of Δ by $(12)(34)$ or its conjugates can be realized as a hyperbolic isometry. There are 3 such conjugates; these elements together with an isometry form a normal subgroup of S_4, the group of permutations of the vertices of a simplex, called the *Klein* 4*–group* K. The quotient group S_4/K is the group G discussed above.

Definition 2.22. The *Lobachevsky function* $\Lambda(\theta)$ is defined by the following integral formula

$$\Lambda(\theta) = -\int_0^\theta \log|2\sin t|dt$$

for $\theta \in \mathbb{R}$.

We establish some basic properties of the function Λ.

Theorem 2.23. *The function $\Lambda(\theta)$ satisfies the following properties:*

1. *Λ is continuous*
2. *Λ is odd*
3. *$\Lambda'(\theta)$ exists for all $\theta \neq k\pi$ for $k \in \mathbb{Z}$*
4. *Λ is periodic with period π*
5. *in the interval $[0, \pi]$ the function Λ has a unique maximum at $\frac{\pi}{6}$*
6. *for $m \in \mathbb{Z}$ we have*

$$\Lambda(m\theta) = m\sum_{k=0}^{m-1} \Lambda\Big(\theta + \frac{k\pi}{m}\Big)$$

Proof: To prove 1 it is obvious that Λ is continuous for θ not a multiple of π. For θ close to 0, $2\sin(\theta) \sim 2\theta$, and $\int_0^\theta \log t dt = t\log t - t|_0^\theta$ which is certainly finite, since t goes to 0 much faster than $\log t$ blows up.

2 is obvious, since the integrand $\log|2\sin t|$ is even.

3 is immediate from the existence of the integral formula.

4 will follow if we can show $\int_0^\pi \log|2\sin t|dt = 0$. We use the identity

$$\sin(2t) = 2\sin t\cos t = 2\sin t\sin(t+\pi/2)$$

which gives us

$$\begin{aligned}\frac{\Lambda(2\theta)}{2} &= -\frac{1}{2}\int_0^{2\theta}\log|2\sin t|dt = -\int_0^{\theta}\log|2\sin(2t)|dt\\ &= -\int_0^{\theta}\log|2\sin t|dt - \int_0^{\theta}\log|2\sin(t+\pi/2)|dt\\ = \Lambda(\theta) - \int_{\pi/2}^{\pi/2+\theta}&\log|2\sin t|dt = \Lambda(\theta)+\Lambda(\theta+\pi/2)-\Lambda(\pi/2)\end{aligned}$$

So if $\theta = \pi/2$ this formula shows $\Lambda(\pi) = 0$.

To prove 5, observe that $\log|2\sin t|$ vanishes when $2\sin t = 1$, so $t = \pi/6$ or $5\pi/6$. The derivative of $\Lambda(\theta)$ is negative in the interval $(\pi/6, 5\pi/6)$ so the unique maximum is at $\pi/6$.

Finally, to prove 6, we need the following identity for an integer m:

$$2\sin(mt) = \prod_{k=0}^{m-1} 2\sin(t+k\pi/m)$$

Denoting $e^{it} = z$ and $e^{ik\pi/m} = \omega$, the left hand side of this expression is equal to

$$-i\Big(z^m - \frac{1}{z^m}\Big)$$

and the right hand side is equal to

$$(-i)^m\Big(z-\frac{1}{z}\Big)\Big(z\omega - \frac{1}{z\omega}\Big)\Big(z\omega^2 - \frac{1}{z\omega^2}\Big)\dots\Big(z\omega^{m-1} - \frac{1}{z\omega^{m-1}}\Big)$$

Evidently, the product of the $z\omega^k$ terms gives $-iz^m$. The product of the $\frac{1}{z\omega^k}$ terms gives iz^{-m}. Each cross term has $\sum_{j=0}^{m-1}\omega^j$ as a factor, but this expression is equal to zero, so there are no other terms, and the identity is proved.

It follows that

$$\begin{aligned}\frac{\Lambda(m\theta)}{m} &= -\frac{1}{m}\int_0^{m\theta}\log|2\sin t|dt = -\int_0^{\theta}\log|2\sin(mt)|dt\\ &= -\sum_{k=0}^{m-1}\int_0^{\theta}\log|2\sin(t+k\pi/m)|dt\\ = -\sum_{k=0}^{m-1}&\left(\int_0^{\theta+k\pi/m}\log|2\sin t|dt - \int_0^{k\pi/m}\log|2\sin t|dt\right)\end{aligned}$$

The second term doesn't involve θ, and therefore gives a constant $C(m)$ depending only on m.

$$= \sum_{k=0}^{m-1}\Lambda(\theta+k\pi/m) + C(m)$$

Since Λ is odd and π–periodic, $\int_0^\pi \frac{\Lambda(m\theta)}{m}d\theta = 0$ and $\int_0^\pi \sum_{k=0}^{m-1}\Lambda(\theta+k\pi/m)d\theta = 0$. It follows that $C(m) = 0$ and the theorem is proved. □

The volume of an ideal hyperbolic simplex has a very simple description in terms of the Lobachevsky function.

Theorem 2.24 (Lobachevsky)**.** *If σ is an ideal simplex with dihedral angles α, β, γ,*

$$\textit{volume}(\sigma) = \Lambda(\alpha) + \Lambda(\beta) + \Lambda(\gamma)$$

Proof: Place one of the vertices at ∞. Then σ is bounded by four hyperbolic planes, three of which can be realized as vertical planes in the upper half–space model. These vertical planes intersect $\mathbb{C}$ in three lines cutting out a Euclidean triangle T with angles α, β, γ. By a hyperbolic isometry, we can arrange for the vertices of T to lie on the unit circle in $\mathbb{C}$. We cut T into six triangles, such that each triangle has vertices consisting of a vertex of T, a midpoint of an edge of T incident to that vertex, and 0. These triangles are right angled, with the angle incident to 0 one of α, β or γ. Up to isometry, there are two regions K_α, and two each of K_β, K_γ. We calculate the volume of the region in σ lying above one of these triangles.

$$\begin{aligned}
\text{volume}(K_\alpha) &= \int_{x=0}^{\cos\alpha} dx \int_{y=0}^{x\tan\alpha} dy \int_{z=\sqrt{1-x^2-y^2}}^{\infty} \frac{dz}{z^3} \\
&= \frac{1}{2}\int_0^{\cos\alpha} dx \int_0^{x\tan\alpha} \frac{dy}{1-x^2-y^2} \\
&= \frac{1}{4}\int_0^{\cos\alpha} \log\frac{\sqrt{1-x^2}\cos\alpha + x\sin\alpha}{\sqrt{1-x^2}\cos\alpha - x\sin\alpha}\frac{dx}{\sqrt{1-x^2}} \\
&= -\frac{1}{4}\int_{\pi/2}^{\alpha} \log\frac{\sin t\cos\alpha + \cos t\sin\alpha}{\sin t\cos\alpha - \cos t\sin\alpha}dt \\
&= \frac{1}{4}\int_{\alpha}^{\pi/2} \log\frac{\sin(t+\alpha)}{\sin(t-\alpha)}dt \\
&= \frac{1}{4}\int_{\alpha}^{\pi/2} \log|2\sin(t+\alpha)|dt - \frac{1}{4}\int_{\alpha}^{\pi/2} \log|2\sin(t-\alpha)|dt \\
&= \frac{1}{4}\int_{2\alpha}^{\alpha+\pi/2} \log|2\sin t|dt - \frac{1}{4}\int_0^{\pi/2-\alpha} \log|2\sin t|dt \\
&= \frac{1}{4}(-\Lambda(\alpha+\pi/2) + \Lambda(2\alpha) + \Lambda(\pi/2-\alpha)) = \frac{\Lambda(\alpha)}{2}
\end{aligned}$$

and the theorem follows. □

2.7.1. *The Lobachevsky function and the Dilogarithm.* The Lobachevsky function is related to the *dilogarithm function*

Definition 2.25. The *dilogarithm function* is defined by the power series

$$\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}$$

Lemma 2.26. *The dilogarithm function is defined by the following integral formula*

$$\textit{Li}_2(z) = \int_0^z \frac{-\log(1-x)}{x}dx$$

Proof: This follows from the following computation

$$\begin{aligned}
-\log(1-x) &= \int_0^x \frac{1}{1-s}ds = \int_0^x 1 + s + s^2 + \dots\, ds \\
&= x + \frac{x^2}{2} + \frac{x^3}{3} + \dots
\end{aligned}$$

Dividing by x and integrating termwise, the identity follows. □

Lemma 2.27. *The dilogarithm and Lobachevsky functions are related by the identity*

$$Li_2(e^{2i\theta}) = \frac{\pi^2}{6} - \theta(\pi - \theta) + 2i\Lambda(\theta)$$

Proof: We can easily calculate

$$\frac{d}{d\theta}\mathrm{Li}_2(e^{2i\theta}) = 2\theta - \pi - 2i\log(2\sin\theta)$$

which proves the identity up to an additive constant. Evaluating either side at $z \to 0$ and applying Euler's formula $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ the formula follows. □

Taking imaginary parts of both sides, we get a Fourier series expansion for $\Lambda(\theta)$:

$$2\Lambda(\theta) = \sum \frac{\sin(2n\theta)}{n^2}$$

For our purposes, the most significant relation between the dilogarithm and the Lobachevsky function is the following theorem of Kummer:

Theorem 2.28 (Kummer)**.**

$$Im(Li_2(z)) = \Lambda(\arg z) + \Lambda\left(\arg\frac{z-1}{z}\right) + \Lambda\left(\arg\frac{1}{1-z}\right) - \log|z|\arg(1-z)$$

Defining $D(z) = \mathrm{Im}(\mathrm{Li}_2(z)) + \log|z|\arg(1-z)$ we see that the volume of an ideal simplex with parameter z is $D(z)$. It follows that $D(z)$ satisfies the following identities:

$$D(z) = D(1 - \frac{1}{z}) = D(\frac{1}{1-z})$$

and

$$D(z) = -D(-z)$$

For $\omega = e^{i\pi/3}$, one can calculate

$$D(\omega) = 1.0149\ldots$$

the volume of a regular ideal tetrahedron.

Remark 2.29. The polylogarithms are defined as power series by the formulae

$$\mathrm{Li}_m(z) = \sum \frac{z^n}{n^m}$$

Volumes of $(2m-1)$–dimensional ideal hyperbolic simplices are related to values of Li_m.

3. Topological 3–manifolds

A *manifold* is a Hausdorff, second countable topological space X with the property that every point $x \in X$ has a neighborhood (i.e. is contained in an open set) homeomorphic to $\mathbb{R}^n$ for some n. If X is connected, n is contstant, and is referred to as the *dimension* of the manifold. As discussed earlier, the topological structure on manifolds can sometimes be stiffened to smooth or piecewise–linear (PL) structures. The fundamental theorem pertinent to topological 3–manifolds is a theorem of Moise:

Theorem 3.1 (Moise)**.** *Let M be a compact 3–manifold (possibly with boundary). Then $(M, \partial M)$ has a unique smooth or PL structure.*

For a proof, see [9].

3.1. **Building 3–manifolds from tetrahedra.** Informally speaking, a PL structure on a n manifold amounts to a description of the manifold as being built up from n–dimensional simplices in a combinatorial way, such that the local combinatorial pictures are "standard".

Definition 3.2. A *combinatorial* structure on a 3–manifold is a description of the manifold as the total space $|\Sigma|$ of a simplicial complex Σ satisfying the following conditions:

- Σ is made by gluing faces of a collection Δ_i of tetrahedra (3–simplices) in pairs
- There are only finitely many simplices glued together around each edge
- The link of every vertex is a sphere

Such conditions are certainly sufficient for $|\Sigma|$ to be a 3–manifold, and it is not too hard to verify that they are also necessary.

There is a useful combinatorial criterion which verifies that the link of every vertex is a sphere.

Lemma 3.3. *Let Σ be obtained from finitely many tetrahedra by glueing the faces in pairs. $|\Sigma|$ is a 3–manifold iff the Euler characteristic $\chi(\Sigma) = 0$.*

Proof: Recall that the Euler characteristic $\chi(\Sigma)$ is defined to be the alternating sum

$$\chi(\Sigma) = v - e + f - t$$

where v is the number of vertices, e edges, f faces and t tetrahedra in the glued up complex.

Since faces are identified in pairs, $f = 2t$. Moreover, each edge contributes 2 vertices, each face 3 edges, each tetrahedron 4 faces to the union of links of vertices. It follows that

$$\sum_{v_i} \chi(\text{link}(v_i)) = 2e - 3f + 4t$$

Note that this sum is equal to $2v$ iff all links of vertices are spheres, since $\chi(S) \leq 2$ for every surface S, with equality iff S is a sphere. Hence $2v - 2e + 3f - 4t = 0$ iff all links are spheres. Equivalently, $v - e + 1\frac{1}{2}f - 2t = 0$. Since $f = 2t$, this sum is equal to $v - e + f - t = \chi(\Sigma)$, and the lemma is proved. □

If M is a 3–manifold with boundary ∂M a (possibly disconnected) surface other than a sphere, one can often think of the interior of M as given by a simplicial complex minus some (or all) of its vertices. The links of removed vertices correspond to removed boundary components of M, and are thus not required to be spheres. If all boundary components of M are tori, a simplicial complex Σ with all vertices removed and all links of vertices tori, whose underlying topological space is homeomorphic to the interior of M, is called an *ideal triangulation* of M.

3.2. **Heegaard decompositions.** If one thinks of the standard embedding $\Sigma_g \to \mathbb{R}^3$ of the surface of genus g in $\mathbb{R}^3$, the bounded component of $\mathbb{R}^3 - \Sigma_g$ is a standard 3–manifold called the *handlebody of genus g*. If Γ is a finite graph embedded in some 3–manifold M, then a thickened neighborhood $N(\Gamma)$ is homeomorphic to a handlebody of some genus.

Definition 3.4. A finite graph $\Gamma \subset H$ is called a *standard graph* if $H - \Gamma$ is homeomorphic to $\partial H \times [0, 1)$.

A handlebody admits many combinatorially distinct standard graphs.

Definition 3.5. A system $\mathcal{D}$ of pairwise disjoint discs in a handlebody H with $\mathcal{D} \cap \partial H = \partial\mathcal{D}$ is called a *meridian system* in H if $H - (\partial H \cup \mathcal{D})$ is homeomorphic to an (open) 3–ball.

To any standard graph $\Gamma \subset H$ is associated (up to topological equivalence) a meridian system $\mathcal{D}_\Gamma$. The system is obtained by first fixing a spanning tree T for Γ and then choosing the disk system dual to the edges of $\Gamma - T$. Conversely, a disk system gives rise to a standard graph by taking the midpoints of the disks and coning them to the midpoint of the complementary ball.

Definition 3.6. Let H_1, H_2 be two handlebodies of genus g. A description of a 3–manifold $M = H_1 \cup H_2$ where $H_1 \cap H_2 = \partial H_1 = \partial H_2$ is called a *Heegaard decomposition* of M. i.e. a Heegaard decomposition of M is a decomposition of the manifold into two handlebodies (necessarily of the same genus).

Theorem 3.7 (Heegaard)**.** *Let M be a closed, oriented 3–manifold. Then M admits a Heegaard decomposition.*

Proof: Triangulate $M = \cup_i \Delta_i$. Let H_1 be a thickened neighborhood of the 1–skeleton. Then H_1 is a handlebody. The complement is a thickened neigborhood of the 1–skeleton of the dual cell decomposition, and therefore also a handlebody. □

We can think of a handlebody as being composed of neighborhoods of the disks in a meridian system, together with the complementary 3–ball. Gluing H_2 to H_1 to make a 3–manifold M can be done in two steps — gluing on the disks, and then attaching the 3–ball. If H_2 is of genus g, a meridian system $\mathcal{D}$ contains g disks. The boundaries of the disks $D_i \in \mathcal{D}$ are g circles $\gamma_i \subset \partial H_2$. Specifying the image of these g circles in ∂H_1 determines how the D_i are attached to H_1 in M.

Topologically, a set $\{\gamma_i\}$ of circles in ∂H_1 is the image of the boundaries of a meridian system of an attaching handlebody H_2 iff the circles satisfy the following conditions:

- the circles are pairwise disjoint and embedded
- the number of circles in the collection is equal to the genus of the handlebody
- the complement of the circles is connected

These conditions ensure that if we build a 3–manifold

$$M' = H_1 \cup_i N(D_i)$$

where the D_i are glued along γ_i, and thickened slightly, then $\partial M' = S^2$. It follows that we can attach the remaining 3–ball to this boundary sphere to make a closed 3–manifold M; manifestly, M has a handlebody decomposition into H_1 and its complement. To see that M is well–defined it suffices to show that there is a unique way to attach an n–disk D^n along its boundary; this is sometimes known as *Alexander's trick*.

Lemma 3.8 (Alexander's trick)**.** *Let X be a topological space, and $S \subset X$ an embedded sphere of dimension $n-1$. Let ϕ_1, ϕ_2 be two homeomorphisms from ∂D^n to S. Then the result of the glueings $X \cup_{\phi_1} D^n$ and $X \cup_{\phi_2} D^n$ are homeomorphic.*

Proof: Let

$$\psi = \phi_2 \phi_1^{-1} : \partial D^n \to \partial D^n$$

Then ψ can be extended to a homeomorphism $\Psi : D^n \to D^n$ by the formula $\Psi(r, \theta) = (r, \psi(\theta))$ where we use radial co–ordinates (r, θ) on D^n, where $\theta \in S^{n-1}$.

Then we have a homeomorphism

$$\mathrm{id} \cup \Psi : X \cup_{\phi_1} D^n \to X \cup_{\phi_2} D^n$$

□

With $n = 2$, this shows that the disks are attached in an essentially unique way, given their image. With $n = 3$, this shows that the ball is at the end is attached uniquely. It follows that any closed oriented 3–manifold M can be described uniquely by a collection of g embedded circles in the boundary of a genus g handlebody H whose complement in ∂H is connected. Such a presentation is called a *Heegaard diagram* for M.

3.3. **Knots, Links and Surgery.** The idea of surgery — building up a manifold by balls of the appropriate dimension, attached along canonical pieces of their boundary — is the foundation of the theory of high dimensional manifolds.

Definition 3.9. An *n–dimensional k–handle* is an n–ball D^n thought of as a product $D^k \times D^{n-k}$. It's boundary has a natural decomposition

$$\partial D^k \times D^{n-k} = S^{k-1} \times D^{n-k} \cup D^k \times S^{n-k-1}$$

The subset $S^{k-1} \times D^{n-k}$ is called the *attaching boundary* of the k–handle, and the sphere $S^{k-1} \times 0$ is called the *core* of the attaching boundary.

If W is an n–manifold with boundary, and $S \subset \partial W$ is a $n-k-1$–sphere, we can attach a k–handle to ∂W so that the core of the attaching boundary of the handle goes along S.

With this terminology, we can think of a Heegaard decomposition as giving a decomposition of a 3–manifold into 3 dimensional handles: the first handlebody H_1 can be built from a 0–handle (a solid ball) and g 1–handles. Then g 2–handles are attached, corresponding to a meridian system for H_2, and finally the result is "capped off" with a 3–handle.

3.3.1. *Framings.* In general, we need to know how to identify the attaching boundary $S^{k-1} \times D^{n-k}$ with the neighborhood of S in ∂W; such an identification is called a *framing* of S. The different choices of framings are given by the different trivializations of this neighborhood.

For $k = 1$ and $n = 4$, we are attaching a $D^2 \times D^2$ along a solid torus $S^1 \times D^2$ whose core goes around a knot K in the 3–manifold ∂W. A framing of K amounts to specifying how many times the circle $S^1 \times p$ in the attaching boundary, for some $p \neq 0$ wraps around the core $S^1 \times 0$ when it is identified with a neighborhood $N(K) \subset \partial W$.

The boundary $\partial N(K)$ is a torus. There is a natural choice of meridian for this torus — a small circle which bounds a disk D intersecting the knot in a single point. A *longitude* for $\partial N(K)$ is any curve l which intersects the meridian m transversely in a single point. It is easy to see that there are $\mathbb{Z}$ choices for the longitude: if l is some choice of longitude, $l+nm$ is another choice of longitude; here $l+nm$ denotes the circle which follows l around once, then winds n times around the meridian m before joining up. It is the image $t_m^n(l)$ of l under the nth power of a Dehn twist in the curve m. We usually orient the meridian and longitude so that at the crossing c, the tangent vectors (m', l') are a positively oriented basis for the tangent space $T_c \partial N(K)$.

In this language, a framing for a knot amounts to a choice of a longitude for $\partial N(K)$.

3.3.2. *Dehn surgery.* If W is a 4–manifold with boundary M, and K is a framed knot in M, the effect on M of adding a 2–handle to W along K is to remove a solid torus neighborhood of K and to glue in a new solid torus, where the meridian of the new solid torus wraps once around the longitude of K and some number of times around the meridian. But if we are only interested in 3–manifolds, there is nothing to stop us from gluing the solid torus back in an arbitrary way, one that has nothing to do with W. This procedure — removing a solid torus from a 3–manifold and then gluing it back in a different way — is called *Dehn surgery*.

To specify a Dehn surgery on a manifold M, we need to specify firstly a knot K in M. We remove a neighborhood $N(K)$ of K from M to get N, a manifold with torus boundary. We glue in a solid torus along ∂N. Think of a solid torus as a (3–dimensional) 2–handle together with a 3–handle: once we specify where the 2–handle goes, the 3–handle can be glued in uniquely, by Alexander's trick. So to specify a Dehn surgery, we just need to specify the isotopy class of the circle in ∂N where the 2–handle will be attached. Based on our experience of surgery on 4–manifolds, we might worry that we need to specify a framing of this circle; but remember a framing is the same thing as a choice of trivialization of a neighborhood of the knot. For a knot in a 2–manifold, a neighborhood is equivalent to a product iff going around the knot preserves orientation. In this case, there is a *unique* (up to isotopy) trivialization of a neighborhood of the knot. The isotopy classes of essential circles in the torus ∂N correspond to *primitive* elements of $\pi_1(\partial N) = \mathbb{Z} \oplus \mathbb{Z}$; after we choose a basis for $\mathbb{Z} \oplus \mathbb{Z}$ as $\langle m \rangle \oplus \langle l \rangle$, such a loop corresponds to an element $pm + ql$ where p, q are coprime integers. Equivalently, after choosing a basis for $\pi_1(\partial N(K))$, Dehn surgery on K can be specified by a choice of an element of $\mathbb{Q} \cup \infty$.

3.3.3. *Knot and Link projections.* A *knot* in a 3–manifold is an embedded circle. Usually we are interested in *tame knots* — i.e. those embedded circles which have a neighborhood homeomorphic to an embedded solid torus. A tame knot can also be thought of as a polygonal embedded circle with respect to the unique PL structure on the manifold. A knot in $\mathbb{S}^3$ can be thought of as sitting in $\mathbb{R}^3$, by removing a point from the complement of the knot. Since the complement is connected, any point will do. Imagine a far away light source, whose rays are almost parallel when they reach the knot: the (thickened) knot casts a shadow on a horizontal plane in $\mathbb{R}^3$. If the knot is in general position, this shadow is an immersed circle in $\mathbb{R}^2$ whose singularities occur when the circle crosses itself transversely at finitely many places. If we record at each such crossing the relative position of the two knot segments in $\mathbb{R}^3$ — i.e. which one is closer, and which further away from the light source — we can reconstruct the embedding of the knot in $\mathbb{R}^3$ from this picture. Such a picture — an immersed circle in $\mathbb{R}^2$ with resolutions of crossings into over and under — is called a *knot diagram*. Two such knot diagrams describe equivalent knots iff they are related by a sequence of *Reidemeister moves*, which describe the topological changes in the picture when the knot moves from one general position to another by a generic isotopy. These moves are illustrated below.

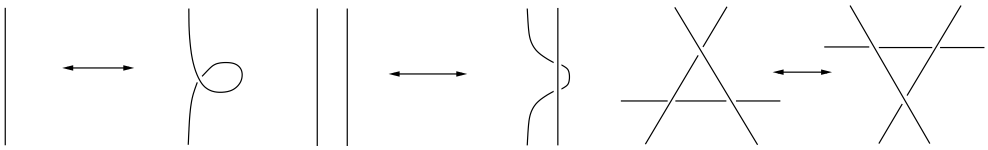

FIGURE 4. Two knot diagrams represent the same topological knot iff they are related by a sequence of Reidemeister moves.

For more details and a proof, consult [13].

If $M^3 = \partial W^4$ and $K \subset M$ is a knot, adding a 2–handle to W along $N(K) \subset M$ has the effect on M of removing a solid torus neighborhood of K, and gluing in a new solid torus along the resulting boundary. The meridian of the glued in solid torus goes once around the longitude of $\partial N(K)$ and some number of times around the meridian of $\partial N(K)$. The exact number of times it winds around the meridian depends exactly on the choice of framing of K.

3.3.4. *Seifert surfaces, linking number and natural framings.* It is convenient to have a canonical choice of "base" framing for a knot, so that other framings can be calculated relative to it. Such a base framing is provided by a geometric object called a *Seifert surface* for a knot.

Definition 3.10. Let $K \subset \mathbb{S}^3$ be a knot. A *Seifert surface* for K is an orientable surface Σ properly embedded in $M - K$ such that $\partial\Sigma = K$.

A Seifert surface is thus a surface in the knot complement whose boundary is the knot. For instance, if K is the unknot, a Seifert surface is a spanning disk for K.

Definition 3.11. Let K be a knot. The *natural longitude* on $\partial N(K)$ is the (isotopy class of) intersection $\Sigma \cap \partial N(K)$ where Σ is a Seifert surface for K.

This definition should raise many questions. Why does a Seifert surface exist for a knot? Why should two different Seifert surfaces give rise to topologically equivalent longitudes? The reasons for these facts are related to the topology of $\mathbb{S}^3$; for a knot in an arbitrary 3–manifold M, there are no "naturally" defined longitudes.

Definition 3.12. A *link* in a 3–manifold is a disjoint collection of knots.

We usually require our links to be tame. A link gives rise to a link diagram, much like a knot diagram. Two link diagrams represent the same topological link iff they are related by a sequence of Reidemeister moves.

If L is a link with two components K_1, K_2, we can define the *linking number* of K_1 with K_2. This definition requires a choice of orientation on each of K_1, K_2, and is defined as follows

Definition 3.13. Let v be a non–zero vector at some point in $\mathbb{R}^3$. The unit length vector parallel to v centered at the origin can be identified with an element of the unit sphere, $g(v) \in S^2$. This defines the *Gauss map*

$$g : T\mathbb{R}^3 - 0 \to S^2$$

Note that 0 denotes the subspace of $T\mathbb{R}^3$ consisting of zero vectors based at any point.

Let $T = S^1 \times S^1$ be parameterized by points on the knots K_1, K_2. Define $g : T \to S^2$ by sending the point (θ, ϕ) to the image of the Gauss map of the vector from the point $\theta \in K_1$ to $\phi \in K_2$.

The *linking number* of K_1 with K_2 is defined to be the degree of the map g; that is, the number of points (counted with sign) in the preimage of a generic point $s \in S^2$.

An easy way to count this number is by looking at a link projection. If s is the south pole of S^2, the preimages of s correspond to the pairs of points $p \in K_1, q \in K_2$ such that p is vertically above q — that is, such that the knot K_1 crosses over the knot K_2 in the projection. So we can define the linking number to be the sum over the undercrossings of K_2 under K_1 in a diagram, of either ± 1, where the count is determined by the following figure:

Now we may define the *natural longitude* of a knot K as the choice of longitude on $\partial N(K)$ which has linking number 0 with K.

3.3.5. *Surgery diagrams.*

Definition 3.14. A *surgery diagram* is a link $L = K_1 \cup K_2 \cup \cdots \cup K_n$ in $\mathbb{S}^3$, where each component K_i may or may not be decorated with an element r_i of $\mathbb{Q} \cup \infty$. Such a diagram represents a 3–manifold, possibly with some number of torus boundary components, defined by surgering the decorated components by (p_i, q_i) surgery with respect to a natural

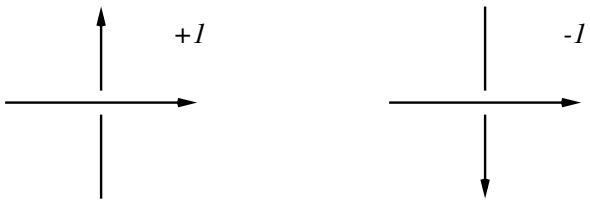

FIGURE 5. The linking number of K_1 and K_2 is the sum over undercrossings of K_2 under K_1 of the sign of the crossing, defined as in this figure.

meridian and longitude, where $r_i = p_i/q_i$ is a representation of the rational number r_i as a ratio of coprime integers, and $(p_i, q_i) = (1, 0)$ if $r_i = \infty$.

Surgery diagrams are certainly not unique: a fixed 3–manifold admits infinitely many representations by a surgery diagram. Moreover, there is no known algorithm which will decide if two diagrams represent the same manifold; in fact, this seems to be a difficult question to decide in practice, as well as in theory.

There are some standard "moves" on diagrams which generate equivalent manifolds; these basic moves always suffice to move between two different diagrams of the same manifold. These moves are as follows:

1. *Twist an unknotted component.*
   If $K \subset L$ is an unknot, it bounds a disk D which can be chosen to be transverse to $L - K$. The manifold obtained from $\mathbb{S}^3 - K$ by cutting along D, twisting through 2π in the anticlockwise direction and then regluing is obviously topologically still $\mathbb{S}^3 - K$, but in terms of the diagram, we have "twisted" the components of L passing through D through one full twist. This has the effect of changing the framings on these components as follows: if K' is a component that passes through D, orient K' and D, and let u, d be the number of times K' passes through D is the positive (resp. negative) direction. Then the framing on K' changes by $(u - d)^2$. This framing change is obviously independent of the orientation of K' and D. Moreover, this twist changes the framing of the boundary of a neighborhood of K. In particular, the natural choice of meridian and longitude for K change by $(m, l) \to (m + l, l)$, so if K is decorated with a rational number, this number must be suitably modified to reflect the change of basis.
2. *Add or remove a knot with framing* ∞
   Doing ∞ surgery on a knot corresponds to removing a solid torus neighborhood of a knot, and replacing it so that the meridian of the added torus goes to the meridian of the knot. In other words, this operation does nothing to the manifold.

These moves suffice to generate all equivalences between manifolds described by surgery diagrams.

Example 3.15. 1 surgery on the figure 8 knot complement is equivalent to -1 surgery on the right–handed trefoil:

As explained above, *integral* surgery on a component of a link in a 3–manifold M corresponds to adding a 2–handle to a 4 manifold W with boundary M. If K_1, K_2 with framings n_1, n_2 are the cores of the attaching handles, we can imagine attaching 2–handles in some particular order — first along K_1, and then along K_2. Let W_1 be the result of adding a 2–handle to W along K_1. Isotopic attaching maps give rise to homeomorphic

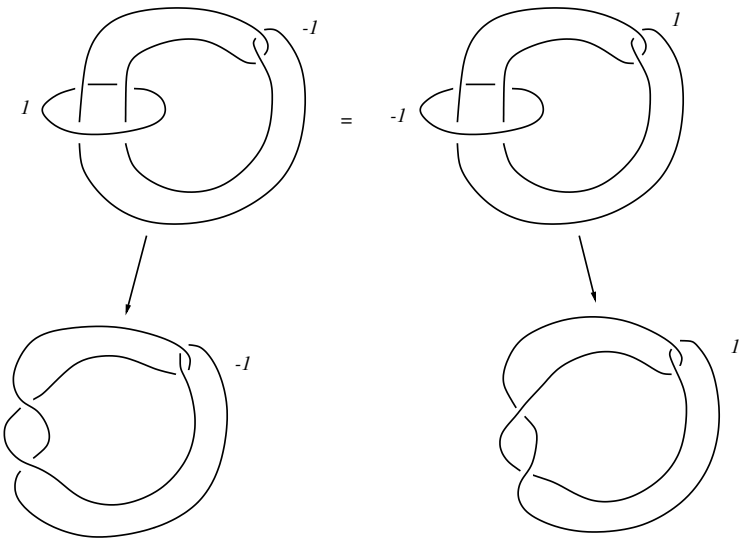

FIGURE 6. This sequence of moves demonstrates the equivalence of the two manifolds obtained by 1 and -1 surgery on the figure 8 knot and right–handed trefoil respectively

manifolds, so we can "slide" K_2 around in $M_1 = \partial W_1$ without changing the end result of surgery. If we slide K_2 over the 2–handle used to attach K_1, the resulting knot K_2' will not be isotopic *in* $M - K_1$ (except under very special circumstances) to K_2. The effect of such a slide is illustrated below.

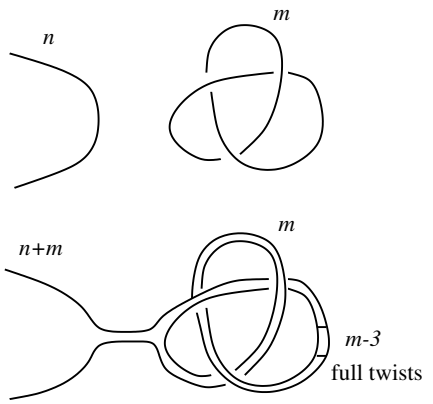

FIGURE 7. The result of sliding K_2 with framing n over K_1 with framing m, both integers

Surgery diagrams with integer framings can equally well represent 3–manifolds or the 4–manifolds that they bound. In the 4–dimensional context, the rules for equivalence of diagrams is called *Kirby Calculus*. A brief exposition can be found in [8]. A much more detailed and thorough account is found in [5].

4. Discrete Groups and Hyperbolic 3–Manifolds

4.1. **Discrete Groups.**

Definition 4.1. Let Γ be a group of homeomorphisms of X. We say the action is:

- *free if no point of X is fixed by any nontrivial $\gamma \in \Gamma$.*
- *discrete if Γ is a discrete subgroup of Homeo(X), thought of as a topological group.*
- *has discrete orbits if every $x \in X$ has a neighborhood U such that the set of $\gamma \in \Gamma$ with $\gamma(x) \in U$ is finite.*
- *wandering if every x has a neighborhood U such that the set of $\gamma \in \Gamma$ with $\gamma(U)$ intersecting U is finite.*
- *properly discontinuous if for all compact sets $K \subset X$ the set of $\gamma \in \Gamma$ for which $\gamma(K)$ intersects K is finite.*

Theorem 4.2. *If X is a manifold, the projection $X \to X/\Gamma$ is a covering map iff Γ acts freely and properly discontinuously.*

If Γ merely acts properly discontinuously, the quotient space naturally has the structure of an *orbifold*.

For a proof of this theorem, and a description of the various fine distinctions between the quality of different kinds of group actions, see [15].

Definition 4.3. A *hyperbolic* 3*–manifold* is the quotient $\mathbb{H}^3/\Gamma$ where Γ is a free, properly discontinuous subgroup of Isom$(\mathbb{H}^3)$.

Definition 4.4. Let M be a 3–manifold. A *hyperbolic structure* on M is an equivalence class of homeomorphisms $\widetilde{d} : \widetilde{M} \to \mathbb{H}^3$ such that i induces an *injective* homomorphism

$$\widetilde{d}_* : \pi_1(M) \to \text{Isom}(\mathbb{H}^3)$$

from the deck group of the covering $p : \widetilde{M} \to M$ to Isom$(\mathbb{H}^3)$, and where the image of i_* is free and properly discontinuous. A hyperbolic structure induces a homeomorphism $d : M \to \mathbb{H}^3/\widetilde{d}_*(\pi_1(M))$. Two such hyperbolic structures $\widetilde{d_1}, \widetilde{d_2} : \widetilde{M} \to \mathbb{H}^3$ are *equivalent* if there is an isometry $i : d_1(M) \to d_2(M)$ such that the composition

$$d_2^{-1} \circ i \circ d_1 : M \to M$$

is *isotopic to the identity*. That is, there is a 1–parameter family of *homeomorphisms* from M to itself, interpolating between $d_2^{-1} \circ i \circ d_1$ and id.

One of the most important theorems in the study of hyperbolic 3–manifolds is the following *rigidity theorem*:

Theorem 4.5 (Mostow, Gabai–Meyerhoff–N.Thurston)**.** *If M is a closed* 3*–manifold, a hyperbolic structure on M, if it exists, is unique up to equivalence.*

Remark 4.6. Actually, we have used an unusually strong defininion of equivalence of hyperbolic structures. It is much more usual to say that two such structures are unique if the composition $d_2^{-1} \circ i \circ d_1$ is merely *homotopic* to the identity. The uniqueness of hyperbolic structures on n manifolds, for all $n \geq 3$ with the usual definition of equivalence of structures, was established by Mostow in [11]. The improvement of homotopy to isotopy for $n = 3$ was achieved only very recently, in [4].

Conversely, we can go from a group Γ acting freely and properly discontinuously on $\mathbb{H}^3$ to a topological manifold $\mathbb{H}^3/\Gamma$ which comes equipped with a hyperbolic structure.

Where do such groups come from? What about if we just ask for properly discontinuous actions? Before the work of Thurston, not many examples of properly discontinuous subgroups of $\mathrm{Isom}(\mathbb{H}^3)$ were known, and most of them fell into the following relatively special classes. Note that there is considerable overlap between these classes.

1. Reflection groups
2. Groups derived from "regular" tilings
3. Arithmetic groups
4. Fuchsian surface groups and their quasifuchsian deformations

We will examine these examples in turn.

4.1.1. *Reflection groups, regular tilings, and Poincaré's polyhedron theorem.* Suppose P is a hyperbolic polyhedron, with all sides totally geodesic, and all dihedral angles of the form π/n_i for various integers $n_1 \geq 2$. Embed P isometrically in $\mathbb{H}^3$. For each face f of P, let r_f denote the reflection of $\mathbb{H}^3$ in the geodesic plane through f. Let Γ_P be the subgroup of $\mathrm{Isom}(\mathbb{H}^3)$ generated by all the r_f. Notice that Γ_P has a subgroup of index 2 consisting of orientation–preserving isometries.

We claim that Γ_P is a properly discontinuous group. The reason is that the condition on dihedral angles implies that $\mathbb{H}^3$ can be tiled with copies of the polyhedron P which alternate in orientation. How can we justify this claim? Poincaré's polyhedron theorem describes very general conditions under which copies of a set of polyhedral "tiles" can be glued together in a periodic manner to give a discrete tiling of one of the standard "geometries" $\mathbb{S}^n, \mathbb{E}^n, \mathbb{H}^n$ for some n.

Theorem 4.7 (Poincaré's polyhedron theorem)**.** *Given the following setup:*

1. *A "model space" $\mathbb{X}$ which is one of $\mathbb{S}^n, \mathbb{E}^n, \mathbb{H}^n$*
2. *A collection of finite–sided polyhedra P_i modelled on $\mathbb{X}$, with totally geodesic faces*
3. *A pairing p of the codimension one faces of P_i, so that the pairing maps are the restrictions of isometries of $\mathbb{X}^n$.*
4. *For each codimension two face e of P_i, let $f_1(e), f_2(e)$ be the two codimension one faces of P_i incident to e. Then $p(f_1(e))$ is a codimension one face of some P_j, and there is some codimension two face e' which is mapped to e by this pairing. The other codimension one face of P_j incident to e' is glued to some other P_k by the face pairings, and determines a new codimension two face, etc. After finitely many steps, this sequence of faces should return to $f_1(e)$, and the composition of the isometries corresponding to this chain should restrict to some isometry ϕ of $\mathbb{X}^n$ on $f_1(e)$, which has some finite power which is the identity: $\phi^m = id$.*
5. *The sum of the dihedral angles around the codimension two edges determined in the previous condition should be $2\pi/m$ (a priori the previous condition just ensures that this sum is of the form $2\pi/m + 2\pi l$ for some integer l).*
6. *Each polyhedron P_i is* compact.

then the geometric complex Σ obtained from the P_i by gluing the faces by the pairing p is an $\mathbb{X}$–orbifold *— that is, it is isometric to the quotient $\mathbb{X}/\Gamma$ for Γ a properly discontinuous group of isometries.*

Remark 4.8. The statement of the theorem above is not as general as it might be. For instance, the polyhedra P_i do not really need to be finite–sided, merely *locally* finite sided. Moreover, they do not really need to be compact, or even to have finite volume. However, to guarantee the desired result in this more general context, we need to add more conditions. We will revisit this subtlety later, when we study 3–manifolds obtained by gluing *ideal polyhedra*.

Sketch of proof: The proof of Poincaré's theorem in this context will rest on the following lemma:

Lemma 4.9. *Let M be a closed manifold with a Riemannian metric which is locally isometric to $\mathbb{X}^n$. Then M is isometric to $\mathbb{X}^n/\Gamma$ for some free, properly discontinuous group Γ in Isom$(\mathbb{X}^n)$.*

Proof: We form the universal cover $\widetilde{M}$ of M, and pull back the Riemannian structure on M to give a Riemannian metric on $\widetilde{M}$. With respect to this metric, the deck group $\pi_1(M)$ acts on $\widetilde{M}$ by isometries. Now, by hypothesis, $\widetilde{M}$ is locally isometric to $\mathbb{X}^n$. We construct a map from $\widetilde{M}$ to $\mathbb{X}^n$ called the *developing map*, defined as follows.

Every point $p \in \widetilde{M}$ is contained in an open neighborhood $U(p)$ which is isometric to a ball of radius $r(p)$ in $\mathbb{X}$. The maximal size of such a ball $r(p)$ depends only on the image of p in M under the covering projection. Since M is compact, there is a *uniform* minimum r for $r(p)$. Pick a basepoint $o \in \widetilde{M}$, and an isometry $i : B_r(o) \to B_r(0) \subset \mathbb{H}^3$. For every open ball V of radius $\leq r$ which intersects U, there is a unique isometry $i_V : V \to \mathbb{H}^3$ which agrees with i on the intersection $U \cap V$. We extend i to V by $i|_V = i_V$. Inductively, once we have defined i on some open set Q, we can extend it over a small open ball W intersecting Q nontrivially, by choosing $i|_W$ to be the unique isometry on W agreeing with i on $W \cap Q$. If $q \in \widetilde{M}$ is arbitrary, we can join q to o by a path α, and cover α with finitely many balls of radius $\leq r$ and define i inductively over α. To check that this is well–defined, observe that for two paths α, α' which only differ on a short subsegment, we can cover a homotopy from α to α' by small open balls, and define i inductively on this cover. Since $\widetilde{M}$ is simply connected, any two paths α, α' from o to q are homotopic. So $i : \widetilde{M} \to \mathbb{X}$ is well–defined, and a local isometry.

By construction, i is open. Moreover, it is closed. For, if $p_i \in \mathbb{X}$ is a sequence of points in the image of i converging to $p \in X$, and $i(q_i) \to p_i$, then if p is a limit of p_i, there is an i so that $d(p, p_i) < r$. But this implies that p is in the image of a ball of radius r contained around q_i. Hence the image of $\widetilde{M}$ is both open and closed, and therefore is all of $\mathbb{X}$. Since M is compact, $\widetilde{M}$ is complete as a metric space, and therefore i is a *covering map*. But $\mathbb{X}$ is simply connected, so i is an isometry.

It is clear that i is uniquely defined once we have fixed $i|_{B_r(o)}$. It follows that if $\alpha \in \pi_1(M)$ is a deck transformation, there is a unique $\rho(\alpha) \in \text{Isom}(\mathbb{X})$ such that

$$i \circ \alpha(p) = \rho(\alpha) \circ i(p)$$

for all $p \in \widetilde{M}$. So $\rho : \pi_1(M) \to \text{Isom}(\mathbb{X})$ is an injective homomorphism, and

$$M = \mathbb{X}/\rho(\pi_1(M))$$

□

Now if M is obtained by gluing polyhedra P_i according to p, M is certainly locally isometric to $\mathbb{X}$ on the interiors of the polyhedra P_i. On the interiors of codimension one faces, M looks locally like the union of two closed half–spaces in $\mathbb{X}$, so in particular the structure is locally isometric to $\mathbb{X}$. The conditions on codimension two faces in the statement of the theorem imply that the geometry in the interior of a codimension two face is locally isometric to that of a quotient $\mathbb{X}/\Gamma$ for some finite group Γ of rotations. There are some technical issues involved in constructing a *universal orbifold cover* of M which will be locally isometric to $\mathbb{X}$ on its codimension two faces, but assuming that this can be done,

we can apply the lemma to show that M is the quotient of $\mathbb{X}$ by a properly discontinuous group. So if M is 2–dimensional, we are done.

The proof now proceeds by induction. If M is 3–dimensional, we need only to be concerned about the geometry in a neighborhood of the vertices. Fix an ϵ, smaller than half the shortest length of an edge of any P_i. If P_i is one of the polyhedra, let Q_{ij} denote the union of the ϵ–neighborhoods of the vertices of the P_i. Each Q_{ij} is the cone on a spherical polygon, where the intrinsic geometry of this polygon is modelled on the sphere of radius ϵ $\mathbb{S}_\epsilon(\mathbb{X})$ in $\mathbb{X}$. The neighborhood of a vertex of M consists of a union of Q_{ij}. The geometry near this vertex is the cone over a 2–dimensional spherical cone manifold, modelled on $\mathbb{S}_\epsilon(\mathbb{X})$. To show that the neighborhood of a vertex is locally modelled on $\mathbb{X}$ is equivalent to showing that the link is isometric to $\mathbb{S}_\epsilon(\mathbb{X})$. But this follows from Poincaré's polyhedron theorem in dimension 2.

The proof for higher dimensions is similar. □

Notice that a key point in this proof is that $\widetilde{M}$ is complete as a metric space, since M is compact. If we obtain M by gluing ideal polyhedra, we have no guarantees that $\widetilde{M}$ will be complete without some extra conditions.

Returning to the geometry of reflection groups, observe that we can find a gluing satisfying the hypotheses of Poincaé's theorem by taking two copies of a fundamental polyhedron P with opposite orientation, and pairing their faces by the identity map.

We can construct some very beautiful tessellations of $\mathbb{H}^3$ by means of Poincaré's theorem. For instance, if P is a regular hyperbolic polyhedron with dihedral angles $2\pi/n$ for some n, we can tile $\mathbb{H}^3$ by copies of P, with n glued around every edge. We enumerate the possibilities:

- The regular ideal tetrahedron with dihedral angles $\pi/3$
- The regular ideal cube with dihedral angles $\pi/3$
- The regular cube with dihedral angles $2\pi/5$
- The regular ideal octahedron with dihedral angles $\pi/2$
- The regular ideal dodecahedron with dihedral angles $\pi/3$
- The regular dodecahedron with dihedral angles $2\pi/5$
- The regular dodecahedron with dihedral angles $\pi/2$
- The regular icosahedron with dihedral angles $2\pi/3$

The groups of symmetries of these tessellation are properly discontinuous groups of isometries of $\mathbb{H}^3$. Note that the tessellations by $\pi/2$ dodecahedra and $2\pi/5$ cubes are dual, and the symmmetry groups they give rise to are conjugate. Note too that a regular ideal cube can be decomposed into 5 regular ideal tetrahedra, so the symmetries of the tessellation by ideal cubes is of finite order in the symmetries of the tessellation by ideal tetrahedra.

4.1.2. *Arithmetic groups.* Let u be a positive real number. Define the symmetric quadratic form

$$\phi_u(x) = x_1^2 + \cdots + x_n^2 - ux_{n+1}^2$$

for $x \in \mathbb{R}^{n+1}$. The map $\Delta_u : x \to (x_1, \ldots, x_n, \sqrt{u}x_{n+1})$ takes the hyperboloid $x_1^2 + \ldots x_n^2 - x_{n+1}^2 = 1$ to the hyperboloid

$$I_u = \{x \in \mathbb{R}^{n+1} : \phi_u(x) = 1\}$$

Let I_u^+ denote the sheet of I_u with $x_{n+1} > 0$. Then Δ_u gives an identification of I_u^+ with $\mathbb{H}^n$.

Define $G(u)$ to be the subgroup of $GL(n+1,\mathbb{R})$ which preserves ϕ_u; that is, those matrices A such that $\phi_u(A(x)) = \phi_u(x)$ for all $x \in \mathbb{R}^{n+1}$. Let $G^+(u)$ be the subgroup preserving I_u^+. Then $G^+(u)$ is isomorphic as a group to $\mathrm{Isom}(\mathbb{H}^n)$. Let $\rho : G^+(u) \to \mathrm{Isom}(\mathbb{H}^n)$ denote some isomorphism

Define

$$\Gamma(u) = G^+(u) \cap GL(n+1, \mathbb{Z}[u])$$

where $\mathbb{Z}[u]$ denotes the ring obtained from $\mathbb{Z}$ by adjoining u.

Theorem 4.10. *Define $\Gamma(u)$ as above. Suppose u satisfies the following conditions:*

1. *u is an algebraic integer.*
2. *The only solution to $\phi_u(x) = 0$ for $x \in \mathbb{Z}^{n+1}$ is $x = 0$.*
3. *All roots of the minimal polynomial of u have multiplicity 1 and (except for u) are real and negative.*

Then $\Gamma(u)$ is a properly discontinuous subgroup of $G^+(u)$, and the corresponding quotient $\mathbb{H}^n/\rho(\Gamma(u))$ is compact.

Sketch of proof: Let f be the minimal polynomial of u of degree d, and $u = u_1, u_2, \ldots, u_d$ the Galois conjugates of u. These conjugates of u_1 determine d embeddings of the ring $\mathbb{Z}[u]$ in $\mathbb{R}$, by $e_i : \mathbb{Z}[u] \to \mathbb{Z}[u_i]$. Let

$$\Delta : GL(n+1, \mathbb{Z}[u]) \to GL(n+1, \mathbb{R}^d) = GL(n+1, \mathbb{R})^d$$

be the "diagonal" embedding associated with the embedding $\delta : \mathbb{Z}[u] \to \mathbb{R}^d$ given by $\delta(u) = (u_1, u_2, \ldots, u_d)$. Then the image of Δ is *discrete* in $GL(n+1,\mathbb{R})^d$. For example, if $A \in GL(n+1, \mathbb{Z}[u])$ we have

$$\mathrm{tr}(\Delta(A)) = \sum_i \mathrm{tr}(A_i) = \mathrm{tr}_{\mathbb{Q}(u)/\mathbb{Q}}(\mathrm{tr}(A)) \in \mathbb{Z}$$

Similarly, if σ is any symmetric function of the eigenvalues of matrix, $\sigma(\Delta(A)) \in \mathbb{Z}$. It is not hard to show therefore that the image of Δ cannot accumulate near id without being actually equal to id — that is, it is discrete. Moreover, the image of $\Gamma(u)$ is a subgroup of

$$G(u_1) \times G(u_2) \times \cdots \times G(u_d)$$

where $G(u_i)$ is the subgroup of $GL(n+1,\mathbb{R})$ preserving the form ϕ_{u_i}. For $i \neq 1$, the group $G(u_i)$ is isomorphic to $O(n+1,\mathbb{R})$; in particular, it is compact. So $\Delta(\Gamma(u))$ is a discrete subgroup of the product of $\mathrm{Isom}(\mathbb{H}^n)$ with a finite number of compact groups, so $\Gamma(u)$ is discrete in $\mathrm{Isom}(\mathbb{H}^n)$.

Since $\Delta(GL(n+1, \mathbb{Z}[u]))$ has finite covolume in $PGL(n+1, \mathbb{R}^d)$, to show that $\Gamma(u)$ has finite covolume in $\prod_i G(u_i)$ it suffices to show that there is no sequence $g_j \in \prod_i G(u_i)$ and $\gamma_j \in \Delta(\Gamma(u))$ such that $g_j\gamma_j^{-1} \to \mathrm{id}$ but is not actually equal to id. But if v is a vector with co–ordinates in $\delta(\mathbb{Z}[u])$, the same is true of $\gamma_j(v)$, so the sum $\sum_i \phi_{u_i}(\gamma_j(v))$ is integral. But $\sum_i \phi_{u_i}(g_j(v)) = \sum_i \phi_{u_i}(v) \in \mathbb{Z}$, so if $g_j\gamma_j^{-1} \to \mathrm{id}$, the values of these functions on basis vectors are eventually equal. In particular, $g_j\gamma_j^{-1}$ is eventually equal to id.

So $\Gamma(u)$ has finite covolume in $\prod_i G(u_i)$, and since all but one of the factors is compact, it has finite covolume in $G(u_1) = G(u)$. Since the stabilizer of the action of $G(u)$ on $\mathbb{H}^n$ is compact, the quotient $\mathbb{H}^n/\Gamma(u)$ has finite volume.

The condition that $\phi_u(x) = 0$ has no integer solutions ensures that the quotient is actually compact. □

The quotient is not necessarily a manifold, since $\Gamma(u)$ will often contain torsion elements. However, it is a general fact that there exists a finite index subgroup of $\Gamma(u)$ which is torsion–free. Such a subgroup can be obtained by examining the kernels of homomorphisms $GL(n+1,\mathbb{Z}[u]) \to GL(n+1,\mathbb{Z}[u]/p)$ for prime ideals p in $\mathbb{Z}[u]$. Hyperbolic manifolds constructed according to this procedure are known as *arithmetic manifolds*. It turns out that the arithmetic manifolds are quite sparsely distributed amongst the set of all hyperbolic manifolds in dimension 3.

The fact that $\Gamma(u)$ contains a torsion–free subgroup of finite index admits considerable generalization, and is known in this generality as *Selberg's lemma*.

Theorem 4.11 (Selberg's lemma)**.** *Let Γ be a finitely generated subgroup of $GL(n,\mathbb{C})$ for some n. Then there is a finite index subgroup $\Gamma' \subset \Gamma$ which is torsion–free.*

Sketch of proof:

□

4.1.3. *Quasifuchsian groups and "bending".*

4.2. **Ideal triangulations.**

4.2.1. *Gluing and cusp equations.*

4.2.2. *Incomplete structures.*

4.2.3. *Hyperbolic Dehn surgery.*

REFERENCES

[1] R. Benedetti and C. Petronio, *Lectures on Hyperbolic Geometry*, Springer–Verlag Universitext (1992)
[2] M. do Carmo, *Riemannian geometry*, Birkhauser (1992)
[3] J. Conway and N. Sloane, *Sphere packings, lattices and codes*, Springer–Verlag
[4] D. Gabai, R. Meyerhoff and N. Thurston, *Homotopy implies isotopy for hyperbolic manifolds*
[5] R. Gompf, *4–manifolds and Kirby Calculus??*
[6] J. Hocking and G. Young, *Topology*, Addison–Wesley (1961)
[7] B. Jessen, *The algebra of polyhedra and the Dehn–Sydler theorem*, Math. Scand. **22** (1968), pp. 241–256
[8] R. Kirby, *The topology of 4–manifolds*, Springer–Verlag LNM 1374 (1989)
[9] E. Moise, *Geometric Topology in Dimensions 2 and 3*, Springer–Verlag GTM **47** (1977)
[10] J. Montesinos, *Classical tessellations and three–manifolds*, Springer–Verlag (1987)
[11] D. Mostow, *Rigidity of rank 1 locally symmetric spaces*
[12] J. Ratcliffe, *Foundations of hyperbolic manifolds*, Springer–Verlag GTM **149** (1994)
[13] K. Reidemeister, *Knot Theory*, translated from the German; BCS associates (1983)
[14] J–P. Sydler, *Conditions nécessaires et suffisantes pour l'équivalence des polyèdres de l'espace euclidien à trois dimensions*, Comm. Math. Helv. **40** (1965), pp. 43–80
[15] W. Thurston, *Three–dimensional geometry and topology, vol. 1*, Princeton University Press, Princeton Math. Series 35 (1997)
[16] J. Tits, *Free subgroups in linear groups*, J. Algebra **20** (1972), pp. 250–270
[17] F. Warner, *Foundations of differentiable manifolds and Lie groups*, Scott, Foresman and co. (1971)

DEPARTMENT OF MATHEMATICS, HARVARD, CAMBRIDGE, MA 02138
E-mail address: dannyc@math.harvard.edu