An alternative method for computing the canonical cell decomposition of a hyperbolic link complement

Preliminary version (no figures yet)

Phillip Andreae, Laura Mansfield,
Kristen Mazur, Morwen Thistlethwaite, William Thistlethwaite

The canonical cell decomposition of a cusped 3-manifold was introduced in [E-P], and was rendered as a practical method of distinguishing such manifolds in [SnapPea]. In this article we describe a different method for computing the canonical cell decomposition for hyperbolic link complements.

1. Introduction

A *classical knot* is a smooth or piecewise linear simple closed curve in 3-dimensional Euclidean space $\mathbb{R}^3$, or in the 3-sphere $S^3 = \mathbb{R}^3 \cup \infty$. Two knots are *equivalent* if there is a continuous deformation of the ambient space ($\mathbb{R}^3$ or S^3) taking one knot to the other.

A *knot invariant* is a function defined on a set of knots, which assigns the same value to equivalent knots. The values of the invariant can simply be numbers, or they might be algebraic objects such as polynomials or groups; they might also be combinatorial structures. If we apply the invariant to two knots and obtain values that are recognizably different, then we'll have proved that the knots are inequivalent. Typically, knot invariants are not 1–1, and therefore will sometimes fail to distinguish pairs of inequivalent knots. An invariant that has the special property of being 1–1 is called *complete*.

This paper concerns a certain complete invariant of *hyperbolic knots*, the *canonical cell decomposition* of the complement of the hyperbolic knot in S^3. Although this invariant is only applicable to hyperbolic knots, this is not a serious disadvantage as almost all knots in the tables are hyperbolic; for example, only 32 of the 1,701,935 prime knots of up to 16 crossings fail to be hyperbolic.

Recall that *Mostow-Prasad rigidity* [Thu] asserts that if the complement of a knot admits a hyperbolic structure, then there is a structure, unique up to isometry, that is *complete*, meaning that the associated metric on the knot complement is a complete metric (*i.e.* Cauchy sequences converge).

The canonical cell decomposition was first described in [E-P], and a method for computing it is implemented in Jeff Weeks's software application *SnapPea* [SnapPea]. The implementation in SnapPea almost always works well; in particular it is extremely fast. However, very rarely (i) it can produce incorrect results owing to round-off error, and (ii) it can fail to arrive at an answer for geometric reasons. For knots, these problems begin to appear at the 16-crossing level. We shall seek to overcome these problems by means of an alternative algorithm, and by employing multiprecision arithmetic. We expect that our method, when implemented as a computer program, will be somewhat slower than that of *SnapPea*.

2. Geometric structures

The first step of our method for computing the canonical cell decomposition is to determine the complete hyperbolic structure on the knot complement, $S^3 - K$. We can then identify the fundamental group of $S^3 - K$ with a group G consisting of orientation-preserving isometries of $\mathbb{H}^3$. We shall use these isometries to compute a sufficient supply of *horoballs* for us to obtain the canonical cell decomposition.

A good introduction to geometric structures is provided by Euclidean structures on the torus $S^1 \times S^1$, arising from actions on the Euclidean plane by discrete groups of Euclidean isometries.

Consider, for example, the following two Euclidean translations of $\mathbb{R}^2$: $\sigma : (x, y) \mapsto (x + 1, y)$, $\tau : (x, y) \mapsto (x, y + 1)$. These generate a group Γ of isometries isomorphic to $\mathbb{Z} \times \mathbb{Z}$; two points (x_1, y_1), (x_2, y_2) of $\mathbb{R}^2$ belong to the same orbit of the action of Γ if and only if $x_1 - x_2 \in \mathbb{Z}$ and $y_1 - y_2 \in \mathbb{Z}$. We may identify the space of orbits $\mathbb{R}^2/\Gamma$ with the torus $S^1 \times S^1$, and the natural map $\mathbb{R}^2 \to \mathbb{R}^2/\Gamma$ with the familiar (universal) covering map

$$
\begin{aligned}
&p : \mathbb{R}^2 \to \mathbb{S}^1 \times \mathbb{S}^1 \quad , \\
&p(x, y) = (e^{2\pi i x}, e^{2\pi i y}) \quad .
\end{aligned}
$$

The covering map p not only gives us a *topology* on the torus, but also gives us a *geometry* on the torus. For example, the torus inherits notions of line, angle, area from corresponding notions in the plane. In particular, the area of the torus is equal to the area of a unit square in the plane, namely 1. (Technically the torus has inherited a *Riemannian metric* from the Euclidean metric on the plane.)

We could vary either the magnitudes of σ, τ or the angle between them, thus obtaining different Euclidean structures on the torus. In general, the group of isometries corresponds to a tiling of the plane by fundamental regions that are parallelograms. In similar fashion we can find geometric structures on surfaces of higher genus, but these cannot be obtained from Euclidean geometry on the plane; instead we would be forced to use a discrete group of isometries of the *hyperbolic plane*.

An analogous situation exists in three dimensions: given a hyperbolic knot K in the 3-sphere, we can identify the complement $S^3 - K$ with the quotient of 3-dimensional hyperbolic space $\mathbb{H}^3$ under the action of a discrete group of isometries.

For the figure-eight knot complement, the isometry group is generated by the following matrices:

$$
\begin{bmatrix} 1 & \omega \\ 0 & 1 \end{bmatrix} , \begin{bmatrix} 1 & 0 \\ \omega & 1 \end{bmatrix} \quad ,
$$

where $\omega = \frac{1}{2}(-1+\sqrt{3}i) = e^{2\pi i/3}$ is a primitive cube root of 1 (we could equally well have used the other primitive cube root $\frac{1}{2}(-1-\sqrt{3}i)$). Later, we shall explain in detail how to derive the isometry group from a specific labelling of the knot diagram, where complex numbers are assigned to crossings and edges.

Let $p : \mathbb{H}^3 \to S^3 - K$ be the associated covering map. We are particularly interested in the preimage of a peripheral torus with respect to p. From the above discussion concerning geometric structures on the torus, perhaps it isn't too surprising that Euclidean planes are involved. Indeed, it's possible to choose this peripheral torus so that the preimage is the union of a disjoint collection of *horospheres* in $\mathbb{H}^3$.

Recall that a horosphere has the geometry of a Euclidean plane, and that in the upper half space model of $\mathbb{H}^3$ a horosphere manifests itself either as a round (punctured) sphere touching the finite part of the boundary, or as a horizontal plane. The point of contact with the boundary is the *center* of the horosphere ("horizontal" horospheres are centered at infinity). For hyperbolic knot complements it's known that the horosphere centers form a dense subset of the boundary of $\mathbb{H}^3$.

By choosing coordinates appropriately, we may assume that one of the horospheres in the preimage of the peripheral torus is the horizontal plane at height 1. If we imagine standing on this horosphere and looking down at all the others, we'll see a fractal arrangement of horospheres, called the "horoball pattern" of the knot.

It'll be explained later how to use these horospheres to get a special decomposition of $\mathbb{H}^3$ into ideal polyhedra. The vertices of the polyhedra are precisely the centers of the horospheres, and the edges of the polyhedra are geodesics joining certain horosphere centers. These ideal polyhedra are mapped by the covering map q to finitely many ideal polyhedra in the knot complement, giving us the so-called *canonical cell decomposition* of the knot complement. The combinatorial structure of this decomposition, *i.e.* information telling us how to glue the ideal polyhedra together to construct the knot complement, is a *complete invariant* of the hyperbolic knot, for the following reasons:

- the (complete) hyperbolic structure on the knot complement is unique [Thu];
- we can construct the knot complement from the glueing information;
- knots are determined by their complements [G-L].

3. Labelling knot diagrams

Suppose that we're given an oriented diagram with n crossings of a hyperbolic knot K. Our method of obtaining the hyperbolic structure on $S^3 - K$ begins with a certain labelling of the crossings and edges of the diagram. We choose a checkerboard shading of the regions of the diagram, and then assign complex numbers $w_1, \ldots, w_n$ to the crossings, and complex numbers $u_1, \ldots u_{2n}$ and $v_1, \ldots v_{2n}$ to the $2n$ edges. More precisely, we think of each label u_i as being attached to the *shaded* side of the ith edge, and v_i as being attached to the unshaded side.

At the ith edge e_i, the labels u_i, v_i are related as follows:

$$v_i = \begin{cases} u_i + 1 & \text{if } e_i \text{ descends from an overpass to an underpass,} \\ u_i & \text{if } e_i \text{ is level, and} \\ u_i - 1 & \text{if } e_i \text{ ascends from an underpass to an overpass.} \end{cases}$$

The remaining equations relating the labels are given by each region of D: a certain product of Möbius transformations determined by the labels incident to that region is declared to equal the identity transformation; this yields three polynomial equations in the labels.

Specifically, suppose that we have a shaded region in the diagram with at least three sides, with crossing labels $w_1, \ldots, w_k$ in cyclic order, and with edge labels $u_1, \ldots, u_k$, where it is assumed that u_i, u_{i+1} are incident to w_i (suffixes modulo k). We define new complex numbers

$$\zeta_i = \begin{cases} -\frac{w_i}{u_i u_{i+1}} & \text{if } e_i, e_{i+1} \text{ are both directed towards or both directed away from } w_i, \\ \frac{w_i}{u_i u_{i+1}} & \text{if just one of } e_i, e_{i+1} \text{ is directed towards } w_i. \end{cases}$$

(Unshaded regions are dealt with similarly, with v_i in place of u_i.)

Then, representing Möbius transformations in the usual way by 2×2 matrices, we have the matrix identity

$$\begin{bmatrix} 0 & -\zeta_1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -\zeta_2 \\ 1 & -1 \end{bmatrix} \cdots \begin{bmatrix} 0 & -\zeta_k \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

For example, for a 3-sided shaded region, it is readily verified that the “region” identity yields $w_i = -u_i u_{i+1}$ if the edges incident to w_i are both directed towards or both away from w_i, and the identity $w_i = u_i u_{i+1}$ if one incident edge is directed away from w_i and the other towards w_i.

For the anomalous case of a 2-sided shaded region, we have $w_1 = w_2$ and $u_1 = u_2 = 0$ ($v_1 = v_2 = 0$ for an unshaded region).

The geometric interpretation for the labels and the above identities is as follows. We choose (complex) affine coordinates on each horosphere, so that a lift of a meridional curve around the peripheral torus travels unit distance in the real direction. For each region of the diagram, consider the disk Δ spanning that region, that lies on the projection 2-sphere except near its boundary, and whose boundary is the union of an alternating sequence of arcs of two kinds, one kind travelling along the peripheral torus near an edge of the diagram, and the other kind travelling “vertically” from an overpass to the corresponding underpass at a crossing (or from underpass to overpass). A lift of the boundary of Δ to $\mathbb{H}^3$ is the union of an alternating sequence of arcs α_i travelling along horospheres and geodesic arcs β_i travelling from one horosphere to another.

For each edge e_i incident to the region, the associated complex number u_i or v_i represents the parabolic isometry that fixes the center of the relevant horosphere and that takes the initial point of α_i to its terminal point. If we assume that the center of this horosphere is at infinity, then the parabolic isometry in question is simply

$$\begin{bmatrix} 1 & u_i \\ 0 & 1 \end{bmatrix} \text{ (shaded), or } \begin{bmatrix} 1 & v_i \\ 0 & 1 \end{bmatrix} \text{ (unshaded)} .$$

The equations relating the labels u_i with the labels v_i arise from the simple observation that the associated arcs on the peripheral torus differ by a meridian, the negative of a meridian or 0, respectively.

The complex number w_i at the ith crossing expresses the relative positions of the two horospheres incident to the lift β_i of the "vertical" arc at that crossing: the distance between the two horospheres, as measured along the geodesic, is $-\log(w_i)$, and the argument of w_i is the angle between the affine coordinate systems of the two horospheres, measured by parallel transport along the geodesic.

Thus, informally, the edge complex numbers tell us how to travel along a horosphere from one crossing geodesic to the next, and the crossing complex numbers tell us how far away the next horosphere is, and how much we have to turn our heads in order to adapt to the new coordinate system.

The "region identities" now follow easily from elementary hyperbolic geometry. For a two-sided region, the two "crossing" geodesics must coincide, as the two vertical crossing arcs in the knot complement are parallel; this explains why the edge complex numbers are zero. For regions of more than two sides, the individual Möbius transformations shift our viewpoint from one crossing to the next.

4. Extracting the knot group and generating horoballs

The next step is to find isometries that generate the group of covering translations corresponding to the fundamental group of the complement of our hyperbolic knot. We shall then be able to generate horospheres by applying these generating isometries recursively, starting with the horosphere centered at infinity and with height 1. Armed with sufficiently many horospheres, we shall be able to realize our ultimate aim of constructing the canonical cell decomposition of the knot complement.

The fundamental group of the knot complement, $\pi_1(S^3 - K, x_0)$, is an infinite group that is found by looking at loops in the knot complement that are based at an arbitrary point, x_0, situated above the knot (see figure 1). The elements of the group are equivalence classes of loops, where two loops are considered equivalent if they are path homotopic. Furthermore, the identity of the group is the "constant" loop that remains at x_0, and the inverse of a loop is determined by traveling in the opposite direction.

The *Wirtinger presentation* is a standard way of describing this group. A group presentation consists of a set of *generators* and a set of *relations*. In the Wirtinger presentation, the generators are so-called *meridians*.

Let us assume that an orientation (*i.e.* a direction) for the knot has been given. A meridian is represented by a path that travels from the basepoint down to a point near an overpass, then loops once around the overpass in the positive screw sense, and finally travels back along the original route to the basepoint (Figure 1). For example, the figure-eight knot has four overpasses; hence the fundamental group of the complement of knot will be generated by the four loops a, b, c, d of Figure 1, one for each overpass.

The relations of the Wirtinger presentation are obtained by looking at certain loops that are obviously trivial, and expressing these as products of generators. The trivial loops in question descend from the basepoint to a point near a crossing but slightly lower than the crossing, then travel around a horizontal circle under the crossing, and finally travel back along the original route to the basepoint. In this way we obtain a relation for each crossing of the knot diagram, although any one of these relations is dependent on the others.

Thus, the Wirtinger presentation the figure-eight knot group. obtained from its standard 4-crossing diagram, is as follows:

$$\pi_1(S^3 - K\,,\,x_0) = <a,\,b,\,c,\,d \mid a^{-1}cab^{-1} = 1,\,c^{-1}acd = 1,\,bdb^{-1}c^{-1} = 1> .$$

This group is crucial to knot theory because it is a very powerful knot invariant, and provides a strong link between geometry and algebra. However, the disadvantage of examining solely the fundamental groups of different knots is that they can quickly become complicated, thus making it difficult to distinguish groups from one another. Yet, the group of isometries of $\mathbb{H}^3$ that is isomorphic to $\pi_1(\mathbb{S}^3 - K, x_0)$ yields a faithful representation of $\pi_1(\mathbb{S}^3 - K, x_0)$ in the form of a subgroup of $SL(2, \mathbb{C})$, and these 2×2 matrices are more approachable.

The necessary isometries are found by using the edge and crossing labels to construct isometries based on Wirtinger generators for $\pi_1(S^3 - K)$. We place the basepoint at some convenient point on the peripheral torus, for example at the point where some crossing geodesic punctures the torus. We can then realize the Wirtinger generator corresponding to some overpass by travelling from the basepoint along the peripheral torus and crossing geodesics to the overpass in question, looping once around the overpass along the peripheral torus, and finally retracing our steps back to the basepoint. The isometry corresponding to this generator will then be expressed as a product $\alpha\beta\alpha^{-1}$, where α is a product (written from left to right) of isometries corresponding to the labels of the edges and crossings encountered on the way to the overpass, and β is the parabolic isometry $z \mapsto z + 1$. It is easily verified that the isometry corresponding to an edge label u_i is represented by the matrix

$$\begin{bmatrix} 1 & -u_i \\ 0 & 1 \end{bmatrix},$$

(similarly for a label v_i), and that the matrix for a crossing label w_i is

$$\begin{bmatrix} 0 & -w_i^{1/2} \\ w^{-1/2} & 0 \end{bmatrix}.$$

Equipped with generating matrices for the group of covering translations corresponding to $\pi_1(S^3 - K)$, we can apply these recursively to generate horospheres in the orbit of the "base" horosphere H_∞, centered at infinity and situated at height 1. Determining the centers and diameters of these horospheres is an easy exercise in elementary hyperbolic geometry.

To apply a matrix to a horosphere center, we write this point in vector form as follows:

$$z \mapsto \begin{bmatrix} z \\ 1 \end{bmatrix} \quad \text{if } z \neq \infty \quad , \quad \infty \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix} .$$

If we apply an isometry $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}(2, \mathbb{C})$ to H_∞, the image horosphere will have Euclidean diameter $D = \frac{1}{|c|^2}$, and the distance between H_∞ and its image will be $-\log(D)$. A slightly more complicated formula gives the Euclidean diameter of the image of a horosphere centered at a finite point of the boundary.

5. Forming the canonical cell decomposition

The *canonical cell decomposition* of the knot complement is a special way of expressing the knot complement as a union of ideal polyhedra glued together along faces. Typically the polyhedra are all ideal tetrahedra, although in the presence of symmetry they can be more complicated.

The canonical cell decomposition is most easily described by considering its preimage in $\mathbb{H}^3$. It is sufficient to describe the vertices, edges and faces of these preimage tiles.

The vertices of the tiles are simply the horosphere centers. Two horosphere centers are joined by a geodesic edge of the tiling if and only if there exists some point in $\mathbb{H}^3$ equidistant from both horospheres and not closer to any other horosphere. Similarly, three or more horosphere centers form the vertices of a face if and only if there exists some point equidistant from them all that is not closer to any other horosphere.

Informally, we can envisage pumping air into all horospheres simultaneously, inflating them at a constant rate; when two horospheres collide, the collision points are thereafter stationary, as happens if one presses two softballs together. Eventually, all the space between the horospheres is exhausted, and we have a cellular decomposition of hyperbolic space, invariant under the action of the knot group, whose 2-skeleton is the *collision locus* resulting from this process. The (preimage of the) canonical cell decomposition is the geometric dual of this cellular decomposition.

Since the knot group acts transitively on the set of horospheres, the preimage tiling of $\mathbb{H}^3$ is determined by the cells (tiles) incident to the base horosphere H_∞. Furthermore, we may restrict our attention to the finitely many cells that are incident to a fundamental parallelogram of H_∞ with respect to its stabilizer in the knot group.

Our strategy is then as follows. The first step is to generate a "fair number" of horospheres (the number to be determined by experiment) by applying generators of the knot group recursively, starting from the base horosphere H_∞. We then attempt to construct the canonical cell decomposition from first principles, using the "nearness" criterion described above. This attempt might fail for two reasons: (i) we might not have sufficiently many horospheres available to construct a system of cells incident to our base parallelogram on H_∞, (ii) because of the absence of vitally important horospheres we might create unwarranted geodesic edges, thus creating a cell decomposition that is not the canonical one.

Fortunately, for (ii) there is a convenient test available, the so-called *tilt test*, due to Jeff Weeks [SnapPea]. To apply this test, we only need the side-lengths and angles of the polygons constituting the intersection of the ideal cells with the base parallelogram on H_∞, and this information is readily available. If our cell decomposition passes the test, all is well and good; if not, we generate some more horospheres and try again. This method must eventually succeed, as only finitely many horospheres are needed to construct the cells, and any particular horosphere is attainable in finite time.

References

[E-P] D.B.A. Epstein and R.C. Penner, *Euclidean decompositions of noncompact hyperbolic manifolds.* J. Diff. Geom. **27** (1988), no. 1, 67–80.

[G-L] C. Gordon and J. Luecke, *Knots Are Determined by Their Complements.* J. Amer. Math. Soc. **2** (1989), 371–415.

[SnapPea] Jeffrey Weeks, *Snappea.* A program for creating and studying hyperbolic 3-manifolds, available for download from http://www.geometrygames.org/snappea/, 1990.

[Thu] W.P. Thurston, Lecture Notes, available for download from http://www.msri.org.