

# From angled triangulations to hyperbolic structures

David Futer and François Guéritaud

ABSTRACT. This survey paper contains an elementary exposition of Casson and Rivin’s technique for finding the hyperbolic metric on a 3–manifold  $M$  with toroidal boundary. We also survey a number of applications of this technique.

The method involves subdividing  $M$  into ideal tetrahedra and solving a system of gluing equations to find hyperbolic shapes for the tetrahedra. The gluing equations decompose into a linear and non-linear part. The solutions to the linear equations form a convex polytope  $\mathcal{A}$ . The solution to the non-linear part (unique if it exists) is a critical point of a certain *volume functional* on this polytope. The main contribution of this paper is an elementary proof of Rivin’s theorem that a critical point of the volume functional on  $\mathcal{A}$  produces a complete hyperbolic structure on  $M$ .

## 1. Introduction

Around 1980, William Thurston showed that “almost every” 3–manifold  $M$ , whether closed or bounded, admits a complete hyperbolic metric [31]. When the boundary of  $M$  consists of tori, this metric is unique up to isometry [20, 24]. Thurston introduced a method for finding this unique metric. The idea was to subdivide the interior of  $M$  into *ideal tetrahedra* (tetrahedra whose vertices are removed), and then give those tetrahedra hyperbolic shapes that glue up coherently in  $M$ . The shape of a hyperbolic ideal tetrahedron can be completely described by a single complex number, namely the cross–ratio of its four vertices on the sphere at infinity. Thurston wrote down a system of *gluing equations* in those complex parameters, whose solution corresponds to the complete hyperbolic metric on the interior of  $M$  [30].

The difficulty with Thurston’s approach is that this non-linear system of equations is very difficult to solve in practice. Even proving the existence of a positively oriented solution, for a given triangulation, often turns out to be a daunting task.

In the 1990s, Andrew Casson and Igor Rivin discovered a powerful technique for solving Thurston’s gluing equations. Their main idea (which builds on a result of Colin de Verdière [3]) was to separate the system into a *linear part* and a *non-linear part*. The linear part of the equations corresponds geometrically to a study

---

2010 *Mathematics Subject Classification*. 57M50.

Futer is supported in part by NSF Grant No. DMS–1007221.

Guéritaud is supported in part by the ANR program ETTT (ANR-09-BLAN-0116-01).

of dihedral angles of the individual tetrahedra. In order for the tetrahedra to fit together in  $M$ , the dihedral angles at each edge of  $M$  must sum to  $2\pi$ ; in order for the tetrahedra to be positively oriented, all the angles must be positive. For a given triangulation  $\tau$ , the space of solutions to this linear system of equations and inequalities is a convex polytope  $\mathcal{A}(\tau)$ , with compact closure  $\overline{\mathcal{A}(\tau)}$ . A point of  $\mathcal{A}(\tau)$ , corresponding to an assignment of tetrahedron shapes that satisfies the angle conditions, is called an *angle structure* on  $\tau$ . See Definition 3.1 for details.

An angle structure on  $\tau$  is weaker than a hyperbolic metric, because the linear equations that define  $\mathcal{A}(\tau)$  impose a strictly weaker condition on tetrahedra than do Thurston's gluing equations. Nevertheless, Casson proved

**THEOREM 1.1.** *Let  $M$  be an orientable 3-manifold with boundary consisting of tori, and let  $\tau$  be an ideal triangulation of  $M$ . If  $\mathcal{A}(\tau) \neq \emptyset$ , then  $M$  admits a complete hyperbolic metric.*

In other words, the existence of an angle structure implies the existence of a hyperbolic structure. Finding this hyperbolic structure requires solving the non-linear part of the gluing equations. To do this, Casson and Rivin introduced a *volume functional*  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ , which assigns to every angle structure on  $\tau$  the sum of the hyperbolic volumes of the tetrahedra in this structure. See Definition 5.2 for a precise definition. As we shall see in Section 5, the function  $\mathcal{V}$  has a number of pleasant properties: it is smooth and strictly concave down on  $\mathcal{A}(\tau)$ , and extends continuously to  $\overline{\mathcal{A}(\tau)}$ . Casson and Rivin independently proved

**THEOREM 1.2.** *Let  $M$  be an orientable 3-manifold with boundary consisting of tori, and let  $\tau$  be an ideal triangulation of  $M$ . Then a point  $p \in \mathcal{A}(\tau)$  corresponds to a complete hyperbolic metric on the interior of  $M$  if and only if  $p$  is a critical point of the functional  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ .*

All told, the Casson–Rivin program can be summarized as follows. Solving the linear part of Thurston's gluing equations produces the convex polytope  $\mathcal{A}(\tau)$ . Solving this linear system is straightforward and algorithmic; by Theorem 1.1, the existence of a solution implies that a hyperbolic structure also exists. By Theorem 1.2, solving the non-linear part of the gluing equations amounts to finding a critical point (necessarily a global maximum) of the functional  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ . In practice, the search for this maximum point can be accomplished by gradient-flow algorithms. It is worth emphasizing that a critical point of  $\mathcal{V}$  does not always exist: the maximum of  $\mathcal{V}$  over the compact closure  $\overline{\mathcal{A}(\tau)}$  will occur at a critical point in  $\mathcal{A}(\tau)$  if and only all the tetrahedra of  $\tau$  are positively oriented in the hyperbolic metric on  $M$ .

**1.1. Where to find proofs.** Proofs of Theorem 1.1 are readily available in the literature. A nice account appears in Lackenby [13, Corollary 4.6], and a slightly more general result appears in Futer and Guéritaud [5, Theorem 1.1]. Here is a brief summary of the argument. By Thurston's hyperbolization theorem [31], the existence of a hyperbolic structure is equivalent to the *non*-existence of essential spheres, disks, tori, and annuli in  $M$ . Any such essential surface  $S$  can be moved into a *normal form* relative to the triangulation  $\tau$ ; this means that every component of intersection between  $S$  and a tetrahedron is a disk in one of several combinatorial types. The dihedral angles that come with an angle structure permit a natural

measure of complexity for normal disks, which mimics the area of hyperbolic polygons. As a result, one can show that every surface  $S$  with non-negative Euler characteristic must have non-positive area, which means it is inessential. Hence  $M$  is hyperbolic.

By contrast, direct proofs of Theorem 1.2 are harder to find. The standard reference for this result is Rivin [26]. However, the focus of Rivin’s paper is somewhat different: he mainly studies the situation where  $M$  is an ideal polyhedron, subdivided into ideal tetrahedra that meet at a single vertex. The notion of an angle space  $\mathcal{A}(\tau)$  and a volume functional  $\mathcal{V}$  still makes sense in this context, and Rivin’s proof of the analogous result [26, Lemma 6.12 and Theorem 6.16] extends (with some effort by the reader) to manifolds with torus boundary.

One reference that contains the above formulation of Theorem 1.2, together with a direct proof, is Chan’s undergraduate honors thesis [1, Theorem 5.1]. However, this thesis is not widely available. In addition, Chan’s argument relies on a certain symplectic pairing introduced by Neumann and Zagier [22], a layer of complexity that is not actually necessary.

The central goal of this paper is to write down an elementary, self-contained proof of Theorem 1.2. The argument is organized as follows. In Section 2, we recall the definition of Thurston’s gluing equations. In Section 3, we give a rigorous definition of the polytope of angle structures  $\mathcal{A}(\tau)$  and compute its dimension. In Section 4, we introduce a natural set of tangent vectors, called *leading–trailing deformations*, which span the tangent space  $T_p\mathcal{A}(\tau)$ . In Section 5, we focus on the volume functional  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ . We will show that the gluing equations for  $\tau$  are in 1–1 correspondence with the leading–trailing deformations; in particular, the non-linear part of a gluing equation is satisfied if and only if the derivative of  $\mathcal{V}$  vanishes along the corresponding deformation. It will follow that the full system of gluing equations is satisfied at  $p \in \mathcal{A}(\tau)$  if and only if  $p$  is a critical point of  $\mathcal{V}$ .

**1.2. Extending and applying the method.** Section 6, at the end of the paper, surveys some of the ways in which the Casson–Rivin program has been generalized and applied. Among the generalizations are versions of Theorem 1.2 for manifolds with polyhedral boundary (Theorem 6.2) and Dehn fillings of cusped manifolds (Theorem 6.1). Another generalization of the method considers angles modulo  $2\pi$ : we discuss this briefly while referring to Luo’s paper in this volume [14] for a much more thorough treatment.

Among the applications are explicit constructions of the hyperbolic metric on several families of 3–manifolds, including punctured torus bundles [11], certain link complements [9], and “generic” Dehn fillings of multi-cusped manifolds [12]. For each of these families of manifolds, the volume associated to *any* angle structure gives useful lower bounds on the hyperbolic volume of the manifold: see Theorem 6.3. For each of these families of manifolds, certain angle inequalities obtained while proving that  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$  has a critical point also imply that the combinatorially natural triangulation  $\tau$  is in fact the *geometrically canonical* way to subdivide  $M$ . See Section 6.5 for details.

Finally, the behavior of the volume functional  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$  gives a surprising and short proof of Weil’s local rigidity theorem for hyperbolic metrics [32]. Because  $\mathcal{V}$  is strictly concave down (Lemma 5.3), any critical point of  $\mathcal{V}$  on  $\mathcal{A}(\tau)$  must be unique. As a result, there is only one complete metric on  $M$  in which all the

tetrahedra of  $\tau$  are positively oriented. This turns out to imply Weil’s local rigidity theorem, as we shall show in Theorem 6.4.

**1.3. A reader’s guide.** For a relative novice to this subject, we recommend the following strategy. Skim over Section 2, as needed, to recall the crucial definitions of gluing equations and holonomy. Read carefully the definitions and opening discussion of Sections 3 and 4, but skip the proofs in those sections entirely. Then, proceed directly to Section 5, where the computation of Proposition 5.4 is particularly important. This computation almost immediately implies the forward direction of Theorem 1.2. On the other hand, the linear–algebra results in Sections 3 and 4 involve completely different ideas compared to the rest of the paper, and are only needed for the reverse direction of Theorem 1.2 (complete metric implies critical point of  $\mathcal{V}$ ) and for various applications in Section 6.

For a practical example of the method, the reader may also study the proof of [11, Lemma 6.2], which proves Theorem 1.2 in the special case of punctured–torus bundles.

**1.4. Acknowledgements.** The proof of Theorem 1.2 was developed while the two authors were preparing to give a mini–course at Osaka University in January 2006. This argument was then tested and refined in front of three workshop audiences containing many graduate students: at Osaka University in 2006, at Zhejiang University in 2007, and at Columbia University in 2009. We thank all the participants of those workshops for their suggestions. We are also grateful to the organizers of the Columbia workshop, “Interactions Between Hyperbolic Geometry, Quantum Topology and Number Theory,” for their encouragement to write up this material for the workshop proceedings. Finally, this paper benefited from a number of conversations with Marc Culler, Feng Luo, Igor Rivin, Makoto Sakuma, and Louis Theran.

## 2. Gluing equations

For the length of this paper,  $M$  will typically denote a compact, connected, orientable 3–manifold whose boundary consists of tori. For example, such a manifold can be obtained from  $S^3$  by removing an open tubular neighborhood of a knot or link. An *ideal triangulation*  $\tau$  is a subdivision of  $M \setminus \partial M$  into ideal tetrahedra, glued in pairs along their faces. One way to recover the compact manifold  $M$  is to truncate all the vertices of the tetrahedra: then, the triangles created by the truncation will fit together to tile  $\partial M$ .

As discussed in the Introduction, our eventual goal is to find a complete hyperbolic metric on  $M$  by gluing together metric tetrahedra. The metric models for the tetrahedra in  $M$  come from ideal tetrahedra in  $\mathbb{H}^3$ :

**DEFINITION 2.1.** A *hyperbolic ideal tetrahedron*  $T$  is the convex hull in  $\mathbb{H}^3$  of four distinct points on  $\partial\mathbb{H}^3$ . The four points on  $\partial\mathbb{H}^3$  are called *ideal vertices* of  $T$ , and are not part of the tetrahedron. The tetrahedron  $T$  is called *degenerate* if it lies in a single plane, and *non-degenerate* otherwise.

Recall that an isometry of  $\mathbb{H}^3$  is completely determined by its action on three points on  $\partial\mathbb{H}^3$ . Thus we may assume, for concreteness, that three vertices of  $T$  lie at  $0, 1, \infty$  in the upper half-space model. If  $T$  is non-degenerate, we may also

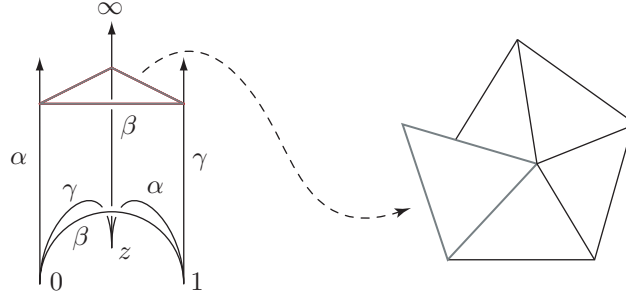


FIGURE 1. Left: the shape of an ideal tetrahedron is defined by three dihedral angles. Right: the gluing of tetrahedra, seen from an ideal vertex.

assume that the fourth vertex lies at  $z \in \mathbb{C}$ , with  $\text{Im}(z) > 0$ . This number  $z \in \mathbb{C}$  determines  $T$  up to isometry.

Suppose we move  $T$  by an orientation-preserving isometry of  $\mathbb{H}^3$ , so that three vertices again lie at  $0, 1, \infty$ , the fourth vertex again lies in the upper half-plane, and edge  $e$  is mapped to the edge  $0\infty$ . Then the fourth vertex of  $T$  will be sent to one of

$$(2.1) \quad z, \quad z' = \frac{z-1}{z} \quad \text{or} \quad z'' = \frac{1}{1-z}.$$

The corresponding number  $z, z'$ , or  $z''$  is called the *shape parameter* of  $e$ . Notice that opposite edges of  $T$  have the same shape parameter, and that  $zz'z'' = -1$ . The arguments

$$\arg z > 0, \quad \arg z' > 0, \quad \arg z'' > 0$$

represent the dihedral angles of  $T$ , or equivalently the angles of the Euclidean triangle with vertices  $0, 1, z$ . If one of the vertices of  $T$  is truncated by a horosphere, the intersection will be a Euclidean triangle in precisely this similarity class. This Euclidean triangle of intersection between  $T$  and a horosphere is called a *boundary triangle* of  $T$ . See Figure 1, left.

The tiling of  $\partial M$  by boundary triangles of the tetrahedra provides a way to understand what conditions are required to glue the tetrahedra coherently.

**DEFINITION 2.2.** Let  $C$  be an oriented surface with a specified triangulation (for example, a boundary torus of  $M$  with the tessellation by boundary triangles). A *segment* in  $C$  is an embedded arc in one triangle, which is disjoint from the vertices of  $C$ , and whose endpoints lie in distinct edges of  $C$ . A *normal closed curve*  $\sigma \subset C$  is an immersed closed curve that is transverse to the edges of  $C$ , such that the intersection between  $\sigma$  and a triangle is a union of segments.

**DEFINITION 2.3.** Let  $\tau$  be an ideal triangulation of  $M$ , and let  $C$  be one torus component of  $\partial M$ . Then  $\tau$  induces a tessellation of  $C$  by boundary triangles. If each tetrahedron of  $\tau$  is assigned a hyperbolic shape, then every corner of each boundary triangle can be labeled with the corresponding shape parameter.

Let  $\sigma$  be an oriented normal closed curve in  $C$ . Then every segment of  $\sigma$  in a triangle of  $C$  cuts off a single vertex of a triangle, labeled with a single shape

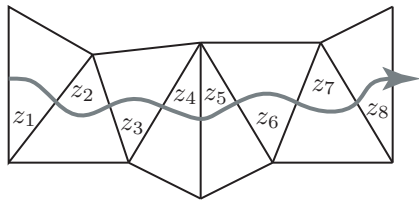


FIGURE 2. A normal path  $\sigma$  passing through the tessellation of a torus by boundary triangles.

parameter (see Figure 2). Let  $z_1, \dots, z_k$  be the sequence of shape parameters corresponding to the segments of  $\sigma$ . Then the *holonomy* of  $\sigma$  is defined to be

$$(2.2) \quad H(\sigma) = \sum_{i=1}^k \epsilon_i \log(z_i),$$

where  $\epsilon_i = 1$  for corners of triangles to the left of  $\sigma$  and  $\epsilon_i = -1$  for corners of triangles to the right of  $\sigma$ . We always choose the branch of the log where  $0 < \arg(z_i) < \pi$  for  $\text{Im}(z_i) > 0$ .

DEFINITION 2.4. Thurston's *gluing equations* for a triangulation  $\tau$  require, in brief, that the holonomy of every curve  $\sigma \subset \partial M$  should be trivial. More precisely, the system consists of *edge equations* and *completeness equations*. The edge equation for an edge  $e$  of  $\tau$  requires that all the shape parameters adjacent to  $e$  satisfy

$$(2.3) \quad \sum_{i=1}^k \log(z_i) = 2\pi i.$$

The *completeness equations* require that, if  $\partial M$  consists of  $k$  tori, there should be a collection of simple closed normal curves  $\sigma_1, \dots, \sigma_{2k}$  spanning  $H_1(\partial M)$ , such that

$$(2.4) \quad H(\sigma_j) = 0 \quad \forall j.$$

As the name suggests, the completeness equations ensure we obtain a complete metric.

PROPOSITION 2.5. *Let  $\tau$  be an ideal triangulation of  $M$ . Suppose that each ideal tetrahedron of  $\tau$  is assigned a non-degenerate hyperbolic shape with positive imaginary part, as above. If the shape parameters of the tetrahedra satisfy the edge gluing equations, these metric tetrahedra can be glued together to obtain a (possibly incomplete) hyperbolic metric on  $M \setminus \partial M$ . This metric will be complete if and only if the completeness equations are satisfied.*

PROOF. Because all hyperbolic ideal triangles are isometric, there is no obstruction to gluing the tetrahedra isometrically along the interiors of their faces. If the edge equation is satisfied for every edge of  $\tau$ , then this isometric gluing along faces of  $\tau$  extends continuously across the edges. Thus we obtain a (possibly incomplete) hyperbolic metric on  $M \setminus \partial M$  and a developing map  $D : \widetilde{M} \rightarrow \mathbb{H}^3$ .

If the metric on  $M$  is complete, then  $D : \widetilde{M} \rightarrow \mathbb{H}^3$  is an isometry. It is well-known that the deck transformations corresponding to a component of  $\partial M$  must be parabolic isometries of  $\mathbb{H}^3$ . In particular, if a cusp of  $M$  is lifted to  $\infty$  in the upper

half-space model of  $\mathbb{H}^3$ , the parabolic isometries preserving that cusp are Euclidean translations of  $\mathbb{C}$ . But a simple closed curve  $\sigma \subset \partial M$  realizes a translation of  $\mathbb{C}$  precisely when its holonomy is  $H(\sigma) = 0$ . Thus equation (2.4) holds for all cusps of  $M$ .

Conversely, suppose that the completeness equations (2.4) are satisfied for a basis of  $H_1(\partial M)$ . Then, for every cusp torus  $C \subset \partial M$ , the boundary triangles of  $C$  can be developed to tile a horosphere in  $\mathbb{H}^3$ . In other words, the developing image of a collar neighborhood of  $C$  is a horoball in  $\mathbb{H}^3$ , which is complete. But if we remove the collar neighborhoods of every boundary torus in  $M$ , the remaining set is compact, hence also complete. Thus the metric on  $M$  is complete.  $\square$

### 3. The polytope of angle structures

As we have seen in the last section, the shape of a hyperbolic ideal tetrahedron is completely determined by its dihedral angles. The notion of an angle structure is that these dihedral angles should fit together coherently.

DEFINITION 3.1. An *angle structure* on an ideal triangulation  $\tau$  is an assignment of an (internal) dihedral angle to each edge of each tetrahedron, such that opposite edges carry the same dihedral angle, and such that

- (1) all angles lie in the range  $(0, \pi)$ ,
- (2) around each ideal vertex of a tetrahedron, the dihedral angles sum to  $\pi$ ,
- (3) around each edge of  $M$ , the dihedral angles sum to  $2\pi$ .

The set of all angle structures on  $\tau$  is denoted  $\mathcal{A}(\tau)$ .

Conditions (1) and (2) in the definition above are precisely what is needed to specify a non-degenerate hyperbolic ideal tetrahedron up to isometry. For concreteness, if the three dihedral angles meeting at a vertex of  $T$  are labeled  $\alpha, \beta, \gamma$  in clockwise order, then the shape parameter corresponding to  $\alpha$  is

$$(3.1) \quad z(\alpha) = \frac{\sin \gamma}{\sin \beta} e^{i\alpha},$$

by the law of sines. Meanwhile, condition (3) in the definition is nothing other than the imaginary part of the edge equation (2.3).

The notion of an angle structure can be summarized by saying that the tetrahedra of  $\tau$  carry genuine hyperbolic shapes, but the conditions these shapes must satisfy are much weaker than the gluing equations. The completeness equations are discarded entirely<sup>1</sup>, and the real part of the edge equations is also discarded. If we attempt to glue the metric tetrahedra coming from an angle structure, we can encounter *shearing singularities* at the edges of  $\tau$ , as in Figure 1.

One may separate the gluing equations of Definition 2.4 into a real part and an imaginary part. The real part of the equations is non-linear, because the shape parameters within one tetrahedron are related in a non-linear way in equation (2.1). On the other hand, the imaginary part of the equations is linear in the dihedral angles, and the set of angle structures  $\mathcal{A}(\tau)$  is defined by a system of linear equations and strict linear inequalities. This leads to

---

<sup>1</sup>There is an alternative version of an angle structure, in which the angles must also satisfy the imaginary part of the completeness equations. This corresponds to taking a linear slice of the polytope  $\mathcal{A}(\tau)$ . We will not need this version for the main part of the paper — but see Section 6.1 for variations on this theme.

PROPOSITION 3.2. *Let  $\tau$  be an ideal triangulation of  $M$ , containing  $n$  tetrahedra. The set of all ways to assign a real number to each pair of opposite edges of each tetrahedron is naturally identified with  $\mathbb{R}^{3n}$ . Then  $\mathcal{A}(\tau)$  is a convex, finite-sided, bounded polytope in  $\mathbb{R}^{3n}$ . If  $\mathcal{A}(\tau) \neq \emptyset$ , its dimension is*

$$\dim \mathcal{A}(\tau) = |\tau| + |\partial M|,$$

where  $|\tau| = n$  is the number of tetrahedra and  $|\partial M|$  is the number of boundary tori in  $M$ .

PROOF. Condition (1) of Definition 3.1 is a system of strict inequalities that constrains the coordinates of  $\mathcal{A}(\tau)$  to the open cube  $(0, \pi)^{3n}$ . Meanwhile, conditions (2) and (3) impose a system of linear equations, whose solution set is an affine subspace of  $\mathbb{R}^{3n}$ . The intersection between this affine subspace and  $(0, \pi)^{3n}$  will be a bounded, convex, finite-sided polytope.

We claim that the number of edges in the triangulation  $\tau$  is  $n$ , the same as the tetrahedra. To see this, observe that each tetrahedron gives rise to 4 boundary triangles that lie in  $\partial M$ . Thus  $\partial M$  is subdivided into  $4n$  boundary triangles. These triangles have a total of  $12n$  sides, glued in pairs; thus  $\partial M$  has  $6n$  edges. Since every component of  $\partial M$  is a torus,

$$\chi(\partial M) = 0 = 4n - 6n + 2n,$$

hence there are  $2n$  vertices on  $\partial M$ . Since every edge of  $\tau$  accounts for two vertices on  $\partial M$ , the number of edges in  $\tau$  is  $n$ , as claimed. It will be convenient to think of the tetrahedra as numbered 1 to  $n$ , and the edges as numbered  $n + 1$  to  $2n$ .

Definition 3.1 involves  $2n$  equations and  $3n$  unknowns. This system can be encoded in a  $(2n \times 3n)$  matrix  $A$ , as follows. Each column of  $A$  corresponds to a pair of opposite edges of one tetrahedron. For  $1 \leq i \leq n$ , the entry  $a_{ij}$  records whether or not the  $i^{\text{th}}$  tetrahedron contains the  $j^{\text{th}}$  edge pair. In other words,  $a_{ij}$  will be 1 when  $3i - 2 \leq j \leq 3i$ , and 0 otherwise. For  $n + 1 \leq i \leq 2n$ , the entry  $a_{ij}$  records how many edges out of the  $j^{\text{th}}$  edge pair become identified to the  $i^{\text{th}}$  edge of the glued-up manifold. Thus the entries in the bottom half of  $A$  can be 0, 1, or 2. With this setup, the system of equations is

$$A \mathbf{v} = [\pi, \dots, \pi, 2\pi, \dots, 2\pi]^T,$$

where  $\mathbf{v} \in \mathbb{R}^{3n}$  is the vector of angles. Then  $\dim \mathcal{A}(\tau) = 3n - \text{rank}(A)$ .

LEMMA 3.3.  $\text{rank}(A) = 2n - |\partial M|$ .

This result is due to Neumann [21]. Our proof is adapted from Choi [2, Theorem 3.7].

PROOF. Since the matrix  $A$  has  $2n$  rows, it suffices to show that the row null space has dimension  $|\partial M|$ . We will do this by constructing an explicit basis for the row null space, with basis vectors in 1 – 1 correspondence with the cusps of  $M$ .

Let  $c$  be a cusp of  $M$ . Associated to  $c$ , we construct a row vector  $\mathbf{r}_c \in \mathbb{R}^{2n}$ . For  $1 \leq i \leq n$ , the  $i^{\text{th}}$  entry of  $\mathbf{r}_c$  is *minus* the number of ideal vertices that the  $i^{\text{th}}$  tetrahedron has at cusp  $c$ . For  $n + 1 \leq i \leq 2n$ , the  $i^{\text{th}}$  entry of  $\mathbf{r}_c$  is *plus* the number of endpoints that the  $i^{\text{th}}$  edge has at cusp  $c$ . Thus  $\mathbf{r}_c$  records incidence, with tetrahedra counted negatively and edges counted positively.

Claim: For every cusp  $c$ ,  $\mathbf{r}_c A = 0$ .



Let  $\mathbf{v}_j$  be a column of  $A$ , and recall that  $\mathbf{v}_j$  corresponds to one pair of opposite edges in one tetrahedron. Then the dot product  $\mathbf{r}_c \cdot \mathbf{v}_j$  counts *minus* the number of times that one of these edges has an endpoint at  $c$ , *plus* the number of times that one of these edges is identified to an edge of  $M$  that has an endpoint at  $c$ . Naturally the sum is 0.

*Claim:* The collection of vectors  $\mathbf{r}_c$  is linearly independent.

Suppose that  $\sum \lambda_c \mathbf{r}_c = 0$ , and let  $e_i$  be the  $i^{\text{th}}$  edge in  $M$ . The  $i^{\text{th}}$  entry of  $\mathbf{r}_c$  will be 0, except when  $c$  is one of the endpoints of  $e_i$ . Thus, if the edge  $e_i$  has endpoints at cusps  $a$  and  $b$ , and the  $i^{\text{th}}$  entry of  $\sum \lambda_c \mathbf{r}_c$  is 0, we must have  $\lambda_a + \lambda_b = 0$ .

Now, let  $\Delta$  be an ideal triangle of  $\tau$ , whose ideal vertices are at cusps  $a$ ,  $b$ , and  $c$ . The argument above, applied to the three edges of  $\Delta$ , gives

$$\lambda_a + \lambda_b = 0, \quad \lambda_a + \lambda_c = 0, \quad \lambda_b + \lambda_c = 0.$$

The only way that these three equalities can hold simultaneously is if  $\lambda_a = \lambda_b = \lambda_c = 0$ . Thus all coefficients  $\lambda_c$  must vanish.

*Claim:* The collection of vectors  $\mathbf{r}_c$  spans the row null space of  $A$ .

Let  $\mathbf{q} = [q_1, \dots, q_{2n}]$  be a vector such that  $\mathbf{q}A = 0$ . Our goal is to show that  $\mathbf{q} = \sum \lambda_k \mathbf{r}_k$ . To that end, let  $T_i$ , ( $1 \leq i \leq n$ ), be one tetrahedron of  $\tau$ . Let  $a, b, c, d$  be the cusps of  $M$  corresponding to the four ideal vertices of  $T_i$ . (Some of these cusps may coincide.) A pair of letters from the collection  $\{a, b, c, d\}$  determines an edge of  $T_i$ , which is identified to some edge of  $M$ . For notational convenience, suppose that edge  $ab$  from tetrahedron  $T_i$  is identified to the edge  $i(ab)$  in  $M$ , and similarly for the other letters.

Now, let  $\mathbf{v}_j$  be the  $j^{\text{th}}$  column vector of  $A$ . This column of  $A$  corresponds to a pair of opposite edges in one tetrahedron, which will be tetrahedron  $T_i$  if and only if  $3i - 2 \leq j \leq 3i$ . After relabeling  $\{a, b, c, d\}$ , we may assume that  $j = 3i - 2$  corresponds to opposite edges  $i(ab)$  and  $i(cd)$ ,  $j = 3i - 1$  corresponds to  $i(ac)$  and  $i(bd)$ , and  $j = 3i$  corresponds to  $i(ad)$  and  $i(bc)$ . By the definition of the matrix  $A$ , the only non-zero entries of  $\mathbf{v}_{3i-2}$  are in row  $i$  (corresponding to the tetrahedron  $T_i$ ) and rows  $i(ab)$ ,  $i(cd)$  (corresponding to edges of  $M$ ). The analogous statement holds for  $j = 3i - 1$  and  $j = 3i$ .

Since  $\mathbf{q}A = 0$ , we must have  $\mathbf{q} \cdot \mathbf{v}_j = 0$  for every  $j$ . Applying this to  $3i - 2 \leq j \leq 3i$  gives

$$\begin{aligned} \mathbf{q} \cdot \mathbf{v}_{3i-2} &= q_i + q_{i(ab)} + q_{i(cd)} = 0, \\ \mathbf{q} \cdot \mathbf{v}_{3i-1} &= q_i + q_{i(ac)} + q_{i(bd)} = 0, \\ \mathbf{q} \cdot \mathbf{v}_{3i} &= q_i + q_{i(ad)} + q_{i(bc)} = 0, \end{aligned}$$

which implies

$$(3.2) \quad q_{i(ab)} + q_{i(cd)} = q_{i(ac)} + q_{i(bd)} = q_{i(ad)} + q_{i(bc)} = -q_i.$$

We are now ready to find coefficients  $\lambda_k$  such that  $\mathbf{q} = \sum \lambda_k \mathbf{r}_k$ . Let  $\Delta$  be an ideal triangle in  $T_i$ , whose ideal vertices are at cusps  $a, b, c$  and whose sides are at edges  $i(ab)$ ,  $i(ac)$ ,  $i(bc)$ . Define

$$(3.3) \quad \lambda_a = \frac{q_{i(ab)} + q_{i(ac)} - q_{i(bc)}}{2}, \quad \lambda_b = \frac{q_{i(ab)} + q_{i(bc)} - q_{i(ac)}}{2}, \quad \lambda_c = \frac{q_{i(ac)} + q_{i(bc)} - q_{i(ab)}}{2}.$$

A priori, this definition depends on the ideal triangle  $\Delta$ . To check this is well-defined, let  $\Delta'$  be another ideal triangle of  $T_i$ , for example the triangle with ideal

vertices at  $b, c, d$ . Note that equation (3.2) implies that

$$q_{i(cd)} - q_{i(bd)} = q_{i(ac)} - q_{i(ab)}.$$

Thus the definition of  $\lambda_c$  coming from  $\Delta'$  will be

$$\lambda_c = \frac{q_{i(bc)} + q_{i(cd)} - q_{i(bd)}}{2} = \frac{q_{i(bc)} + q_{i(ac)} - q_{i(ab)}}{2}.$$

Therefore, triangles  $\Delta$  and  $\Delta'$  that belong to the same tetrahedron and share a vertex at cusp  $c$  induce the same definition of  $\lambda_c$ . Since any pair of ideal triangles that meet the cusp  $c$  are connected by a sequence of tetrahedra, it follows that  $\lambda_c$  is well-defined.

It remains to check that this definition of  $\lambda_c$  satisfies  $\mathbf{q} = \sum \lambda_k \mathbf{r}_k$ . Recall that for  $n+1 \leq i \leq 2n$ , the  $i^{\text{th}}$  entry of  $\mathbf{r}_k$  is the number of endpoints that the  $i^{\text{th}}$  edge has at cusp  $k$ . If the ends of this edge are at cusps  $a$  and  $b$ , i.e. if this is the edge we have referred to as  $i(ab)$ , then the  $i^{\text{th}}$  entry of  $\sum \lambda_k \mathbf{r}_k$  is

$$(3.4) \quad \lambda_a + \lambda_b = q_{i(ab)},$$

using equation (3.3). This is exactly the  $i^{\text{th}}$  entry of  $\mathbf{q}$ . Similarly, for  $1 \leq i \leq n$ , the  $i^{\text{th}}$  entry of  $\mathbf{r}_k$  is *minus* the number of ideal vertices that  $T_i$  has at cusp  $k$ . Thus, if  $T_i$  has ideal vertices at cusps  $a, b, c, d$  (where some of these may coincide), the  $i^{\text{th}}$  entry of  $\sum \lambda_k \mathbf{r}_k$  is

$$-(\lambda_a + \lambda_b + \lambda_c + \lambda_d) = -q_{i(ab)} - q_{i(cd)} = q_i,$$

using equations (3.4) and (3.2). Thus  $\mathbf{q} = \sum \lambda_k \mathbf{r}_k$ , as desired.

This completes the proof that the vectors  $\mathbf{r}_c$ , indexed by the cusps of  $M$ , form a basis for the row null space of  $A$ . Therefore,  $\text{rank}(A) = 2n - |\partial M|$   $\square$

We conclude that  $\dim \mathcal{A}(\tau) = 3n - \text{rank}(A) = n + |\partial M|$ , completing the proof of Proposition 3.2.  $\square$

#### 4. Leading–trailing deformations

Given a point  $p$  of the angle polytope  $\mathcal{A}(\tau)$ , let  $T_p \mathcal{A}(\tau)$  be the tangent space to  $\mathcal{A}(\tau)$  at the point  $p$ . There is a particularly convenient choice of spanning vectors for  $T_p \mathcal{A}(\tau)$ . These *leading–trailing deformations*, which were probably folklore knowledge to experts for several years, are the main innovation in our direct proof of Theorem 1.2. In a certain sense, they serve as a more concrete reformulation of Neumann and Zagier’s symplectic pairing [22]. To the best of the authors’ knowledge, this is the first place where they are described in print.

**DEFINITION 4.1.** Let  $C$  be a cusp torus of  $M$ , with a tessellation by boundary triangles coming from  $\tau$ . Let  $\sigma \subset C$  be an oriented normal closed curve (see Definition 2.2), consisting of segments  $\sigma_1, \dots, \sigma_k$ . Every oriented segment  $\sigma_i$  lies in a boundary triangle  $\Delta_i$ . We define the *leading* corner of  $\Delta_i$  to be the corner opposite the side where  $\sigma_i$  enters  $\Delta_i$ , and mark it with a  $+$ . We define the *trailing* corner of  $\Delta_i$  to be the corner opposite the side where  $\sigma_i$  leaves  $\Delta_i$ , and mark it with a  $\ominus$ . See Figure 3.

For every oriented segment  $\sigma_i$ , we define a vector  $\mathbf{w}(\sigma_i) \in \mathbb{R}^{3n}$ , where as above each coordinate of  $\mathbb{R}^{3n}$  corresponds to one pair of opposite edges in one of the  $n$  tetrahedra. Every corner of  $\Delta_i$  corresponds to one such edge pair in a tetrahedron. The vector  $\mathbf{w}(\sigma_i)$  will have a 1 in the coordinate corresponding to the leading corner

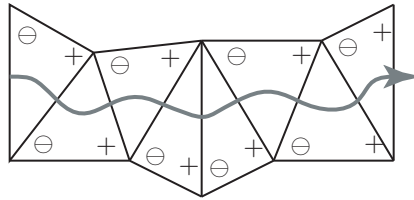


FIGURE 3. The leading–trailing deformation along  $\sigma$  increases the angles marked  $+$  and decreases the angles marked  $\ominus$ .

of  $\Delta_i$ , a  $-1$  in the coordinate corresponding to the trailing corner of  $\Delta_i$ , and  $0$ 's otherwise.

Finally, the *leading-trailing deformation* corresponding to  $\sigma$  is the vector  $\mathbf{w}(\sigma) = \sum_i \mathbf{w}(\sigma_i)$ .

EXAMPLE 4.2. Let  $e$  be an edge of  $M$ , and suppose for simplicity that all tetrahedra adjacent to  $e$  are distinct. Let  $\sigma$  be a simple closed curve on  $\partial M$ , running counterclockwise about one endpoint of  $e$ . The boundary triangles intersected by  $\sigma$  have angles  $\alpha_i, \beta_i, \gamma_i$ , labeled clockwise with  $\alpha_i$  inside  $\sigma$ . Then the leading–trailing deformation  $\mathbf{w}(\sigma)$  adds 1 to every  $\beta_i$  and subtracts 1 from every  $\gamma_i$ , keeping the angle sum in each tetrahedron equal to  $\pi$ . In addition, the dihedral angle marked  $\beta_i$  is adjacent to the same edge of  $M$  as the dihedral angle marked  $\gamma_{i+1}$ . Thus the angle sum at each edge is unchanged. (See Figure 4.) Observe that  $\mathbf{w}(\sigma)$  has no effect at all on the dihedral angles  $\alpha_i$  adjacent to  $e$ .

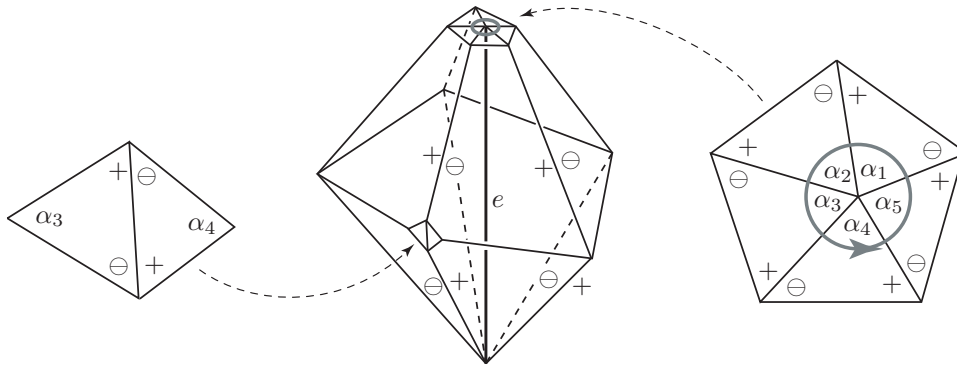


FIGURE 4. The leading–trailing deformation about a single edge  $e$  of  $M$ . Every increase to an angle in a tetrahedron is canceled out by a decrease to an adjacent angle.

In the more general setting, we will prove in Lemma 4.5 that leading–trailing deformations are always tangent to  $\mathcal{A}(\tau)$ . To do that, we need a better understanding of the interaction between different deformations.

DEFINITION 4.3. Let  $\tau$  be an ideal triangulation of  $M$ , and let  $\rho, \sigma$  be oriented normal closed curves on  $\partial M$  that intersect transversely (if at all). Define the *signed intersection number*  $\iota(\rho, \sigma)$  to be the number of times that  $\sigma$  crosses  $\rho$  from right to

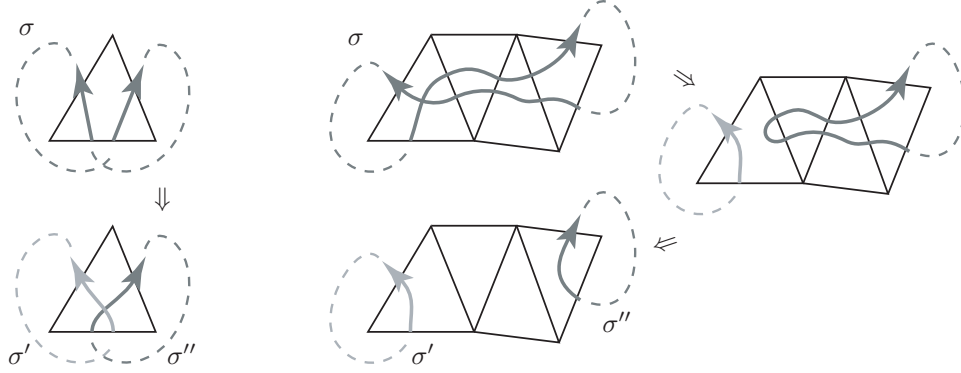


FIGURE 5. Decomposing  $\sigma$  into smaller curves. Left: when two segments of  $\sigma$  pass through the same boundary triangle, we may cut and rejoin to form two normal curves  $\sigma'$ ,  $\sigma''$ . Right: a local isotopy in the complement of the vertices may be required to ensure that  $\sigma'$ ,  $\sigma''$  are again normal curves.

left, minus the number of times that  $\sigma$  crosses  $\rho$  from left to right. This definition has a few immediate properties:

- It is anti-symmetric:  $\iota(\rho, \sigma) = -\iota(\sigma, \rho)$ .
- It depends only on the homology classes of  $\sigma$  and  $\rho$  in  $H_1(\partial M)$ .
- By considering a transverse pushoff of  $\sigma$ , one obtains  $\iota(\sigma, \sigma) = 0$ .

LEMMA 4.4. *Let  $\rho, \sigma$  be oriented normal closed curves on  $\partial M$  that intersect transversely, if at all. Then*

$$\frac{\partial}{\partial \mathbf{w}(\sigma)} \text{Im}(H(\rho)) = 2\iota(\rho, \sigma).$$

Recall that  $\text{Im}(H(\rho))$  is the linear, angled part of the holonomy in Definition 2.3.

PROOF. The proof involves three steps.

*Step 1* introduces several simplifying assumptions with no loss of generality. First, it will help to assume that every tetrahedron  $T$  is embedded in  $M$  (that is,  $T$  does not meet itself along an edge or face). This assumption can always be met by passing to a finite-sheeted cover of  $M$ , since  $\pi_1(M)$  is residually finite by Selberg's lemma. Note that the tetrahedra, angles, boundary curves, holonomies, and leading-trailing deformations all lift naturally to covers. This assumption also implies that boundary triangles  $\Delta_i, \Delta_j$  meet in either at most one edge or at most one vertex.

Next, no generality is lost by assuming that  $\sigma$  passes through each boundary triangle at most once. For, if two segments  $\sigma_i, \sigma_j$  run through the same boundary triangle  $\Delta$ , we may cut and rejoin the curve  $\sigma$  into a pair of shorter normal closed curves  $\sigma', \sigma''$ . See Figure 5. (To ensure that  $\sigma'$  and  $\sigma''$  are *normal* curves, it may be necessary to move them by isotopy in the complement of the vertices, as in the right panel of the figure.) Note that this operation is natural and topologically additive: we have  $\mathbf{w}(\sigma) = \mathbf{w}(\sigma') + \mathbf{w}(\sigma'')$ , and  $\iota(\rho, \sigma) = \iota(\rho, \sigma') + \iota(\rho, \sigma'')$ . Thus, if

we prove the lemma for each of  $\sigma'$  and  $\sigma''$ , the result will follow for  $\sigma$ . In particular, we may now assume that  $\sigma$  is embedded.

Similarly, we may assume that  $\rho$  also passes through each boundary triangle at most once. For, if  $\rho_i, \rho_j$  run through the same boundary triangle  $\Delta$ , we may decompose  $\rho$  into a pair of curves  $\rho', \rho''$  just as above. This natural operation ensures  $H(\rho) = H(\rho') + H(\rho'')$ , and  $\iota(\rho, \sigma) = \iota(\rho', \sigma) + \iota(\rho'', \sigma)$ . Thus it suffices to consider each of  $\rho'$  and  $\rho''$  in place of  $\rho$ .

Given these simplifying assumptions, suppose that the segments  $\sigma_1, \dots, \sigma_k$  of  $\sigma$  are contained in distinct boundary triangles  $\Delta_1, \dots, \Delta_k$ . A priori, there are two ways in which  $\mathbf{w}(\sigma)$  could affect the angular component of  $H(\rho)$ :

- (1) A segment of  $\rho$  lies in the same boundary triangle  $\Delta_i$  that contains  $\sigma_i$ . Then,  $\mathbf{w}(\sigma_i)$  changes the angles of  $\Delta_i$ , thereby changing  $\text{Im}(H(\rho))$ .
- (2) A segment of  $\rho$  lies in a boundary triangle  $\Delta'_i$ , where  $\Delta_i \neq \Delta'_i$  are two truncated vertices of the same ideal tetrahedron. By changing dihedral angles in the ambient tetrahedron,  $\mathbf{w}(\sigma_i)$  also changes the angles of  $\Delta'_i$ .

We consider the effect of (2) in Step 2 and the effect of (1) in Step 3.

Step 2 is to show that the effect of  $\mathbf{w}(\sigma)$  in the boundary triangles “on the other side” of the ambient tetrahedra, as in (2) above, will never change  $\text{Im}(H(\rho))$ . To see why this is true, it helps to group together boundary triangles that share a common vertex  $v$ .

Let  $v$  be a vertex of  $\partial M$ , and let  $\sigma_v = \sigma_1 \cup \dots \cup \sigma_j$  be a maximal union of consecutive segments in  $\sigma$ , such that the ambient boundary triangles  $\Delta_1, \dots, \Delta_j$  are all adjacent to  $v$ . By our embeddedness assumption, each  $\Delta_i$  is adjacent to  $v$  in only one corner. Recall that  $v$  is one endpoint of an edge  $e \subset M$ ; we will investigate the effect of  $\mathbf{w}(\sigma_v)$  on boundary triangles  $\Delta'_1, \dots, \Delta'_j$  at the other end of edge  $e$ . The arc  $\sigma_v$  can take one of three forms:

- (i)  $\sigma_v = \sigma$ , and forms a closed loop about  $v$ .
- (ii)  $\sigma_v$  takes a right turn in  $\Delta_1$ , followed by a sequence of left turns about  $v$ , followed by a right turn in  $\Delta_j$ .
- (iii)  $\sigma_v$  takes a left turn in  $\Delta_1$ , followed by a sequence of right turns about  $v$ , followed by a left turn in  $\Delta_j$ .

Scenario (i) is depicted in the right panel of Figure 4. Here,  $v$  is in the interior of a polygon  $P_v = \Delta_1 \cup \dots \cup \Delta_j$ . Notice that none of the external angles along  $\partial P_v$  actually change, even though the shapes of constituent triangles are changing. The same will be true in the boundary polygon  $P'_v = \Delta'_1 \cup \dots \cup \Delta'_j$  at the other end of  $e$ . Thus, if  $\rho$  passes through  $P'_v$ , its holonomy will be unaffected by this change.

In Scenario (ii), we once again construct a polygon  $P_v$  by gluing  $\Delta_1$  to  $\Delta_2$  along their (unique) shared edge, and so on up to  $\Delta_j$ . Let  $P'_v = \Delta'_1 \cup \dots \cup \Delta'_j$  be the polygon formed by the boundary triangles in the same tetrahedra, on the other end of  $e$ . These two polygons are depicted in Figure 6. Notice that although  $\Delta_i$  and  $\Delta'_i$  are in the same oriented similarity class (both have the same angles  $\alpha_i, \beta_i, \gamma_i$ , in clockwise order), the triangles in these similarity classes are rearranged to form  $P'_v$ . In particular,  $\mathbf{w}(\sigma_v)$  does not change any of the external angles along  $\partial P'_v$ , even though the shapes of constituent triangles  $\Delta'_i$  are changing. Thus, if  $\rho$  passes through  $P'_v$ , its holonomy will be unaffected.

Scenario (iii) is the mirror image of (ii). Once again,  $\mathbf{w}(\sigma_v)$  does not change any of the external angles along  $\partial P'_v$ .

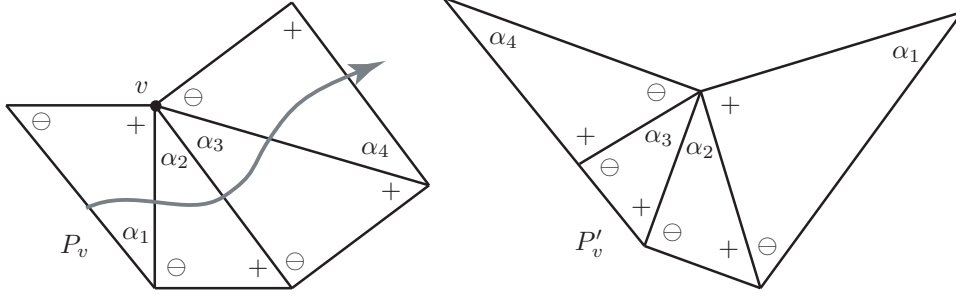


FIGURE 6. Left: the boundary polygon  $P_v$  comprised of boundary triangles  $\Delta_1, \dots, \Delta_j$  adjacent to  $v$ . In each  $\Delta_i$ , the smallest angle is  $\alpha_i$ . Right: in the boundary polygon  $P'_v = \Delta'_1 \cup \dots \cup \Delta'_j$ , at the other end of the edge that starts at  $v$ , none of the total angles along  $\partial P'_v$  change.

We conclude that for every vertex  $v$  in a boundary triangle visited by  $\sigma$ , the polygon  $P'_v$  on the other end of the same edge will have all its external angles unaffected by  $\mathbf{w}(\sigma_v)$ . Since each segment  $\sigma_i$  of  $\sigma$  belongs to three different maximal arcs  $\sigma_v$  (corresponding to the three distinct vertices of  $\Delta_i$ ), we have  $\sum_v \mathbf{w}(\sigma_v) = 3\mathbf{w}(\sigma)$ . Thus, since each  $\mathbf{w}(\sigma_v)$  does not affect the holonomy of  $\rho$  “on the other side” of the ambient tetrahedra, neither does  $\mathbf{w}(\sigma)$ . This completes Step 2.

*Step 3* considers the holonomy of  $\rho$  in the same boundary triangles that are also visited by  $\sigma$ . Consider a maximal consecutive string of segments  $\sigma_1, \dots, \sigma_j$ , contained in  $\Delta_1, \dots, \Delta_j$ , such that consecutive segments  $\rho_1, \dots, \rho_j$  run through the same boundary triangles. Depending on orientations, we can have either  $\rho_1 \subset \Delta_1$  or  $\rho_1 \subset \Delta_j$ .

In the special case where  $\Delta_1 \cup \dots \cup \Delta_j$  contains all of  $\sigma$  and all of  $\rho$ , the two curves cut off exactly the same corners of boundary triangles (possibly in opposite cyclic order). Then, since  $\partial M$  is orientable, we must have  $\iota(\rho, \sigma) = 0$ . In this case, the angles that go into computing the angular holonomy  $\text{Im}(H(\rho))$  are exactly the angles that are unaffected by  $\mathbf{w}(\sigma)$ . Thus, in this special case,  $\partial \text{Im}(H(\rho)) / \partial \mathbf{w}(\sigma) = 0 = 2\iota(\rho, \sigma)$ .

In the general case, we may assume that  $\sigma$  and  $\rho$  do not run in parallel through  $\Delta_1$  or through  $\Delta_j$  (otherwise, the sequence  $1, \dots, j$  is not maximal). As above, it helps to construct a polygon  $P$  by gluing  $\Delta_1$  to  $\Delta_2$  along their (unique) shared edge, and so on up to  $\Delta_j$ . By Definition 4.1, each  $\Delta_i$  has one mark of  $+$  in the leading corner relative to  $\sigma_i$ , and one mark of  $\ominus$  in the trailing corner. Altogether, the polygon  $P$  contains  $2j$  markers, with  $j$  pluses and  $j$  minuses. We consider how many of these markers are to the left and right of  $\rho$ .

If  $\sigma$  enters polygon  $P$  to the right of  $\rho$  and leaves to the left of  $\rho$ , the interior of  $P$  contributes  $+1$  to  $\iota(\rho, \sigma)$ . Also, the part of  $P$  to the left of  $\rho$  will have a surplus of two  $+$ 's. Thus  $\mathbf{w}(\sigma_1 \cup \dots \cup \sigma_j)$  increases the angular component of  $H(\rho)$  by 2. See Figure 7.

If  $\sigma$  enters polygon  $P$  to the left of  $\rho$  and leaves to the right of  $\rho$ , the interior of  $P$  contributes  $-1$  to  $\iota(\rho, \sigma)$ . Also, the part of  $P$  to the left of  $\rho$  will have a surplus of two  $\ominus$ 's. Thus  $\mathbf{w}(\sigma_1 \cup \dots \cup \sigma_j)$  increases the angular component of  $H(\rho)$  by  $-2$ .

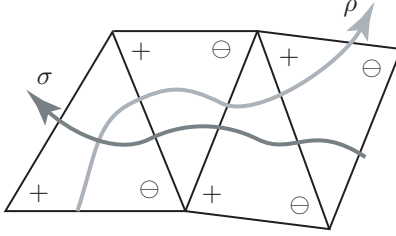


FIGURE 7. When  $\sigma$  crosses  $\rho$  from right to left, the part of the polygon  $P$  to the left of  $\rho$  contains an excess of two  $+$ 's.

Finally, if  $\sigma$  enters and leaves  $P$  on the same side of  $\rho$ , the interior of  $P$  contributes 0 to  $\iota(\rho, \sigma)$ . Also, the part of  $P$  to the left of  $\rho$  will have the same number of  $+$ 's and  $\ominus$ 's, hence  $\mathbf{w}(\sigma_1 \cup \dots \cup \sigma_j)$  does not affect the angular component of  $H(\rho)$ .

Summing these contributions over the (disjoint) polygons of intersection, we conclude that  $\partial \text{Im}(H(\rho)) / \partial \mathbf{w}(\sigma) = 2 \iota(\rho, \sigma)$ , as desired.  $\square$

We can now show that every leading–trailing deformation  $\mathbf{w}(\sigma)$  is tangent to  $\mathcal{A}(\tau)$ .

LEMMA 4.5. *Let  $p \in \mathcal{A}(\tau)$  be an angle structure, and let  $\sigma$  be an oriented normal curve on a cusp of  $M$ . Then the vector  $\mathbf{w}(\sigma)$  is tangent to  $\mathcal{A}(\tau)$ . In other words, for all sufficiently small  $\varepsilon > 0$ ,  $p + \varepsilon \mathbf{w}(\sigma) \in \mathcal{A}(\tau)$ .*

PROOF. Let us check the conditions of Definition 3.1. By condition (1) of the definition,  $p$  lies in the open set  $(0, \pi)^{3n}$ . Thus, for sufficiently small  $\varepsilon$ , we have  $p + \varepsilon \mathbf{w}(\sigma) \in (0, \pi)^{3n}$  also.

Next, observe that in Definition 4.1, every vector  $\mathbf{w}(\sigma_i)$  only affects a single tetrahedron. In this tetrahedron,  $\varepsilon \mathbf{w}(\sigma_i)$  adds  $\varepsilon$  to the dihedral angle on one (leading) pair of opposite edges, and adds  $-\varepsilon$  to the dihedral angle on another (trailing) pair of opposite edges. Thus the deformation of angles coming from  $\mathbf{w}(\sigma_i)$  keeps the angle sum equal to  $\pi$  in every tetrahedron. Clearly, the same is true for  $\mathbf{w}(\sigma) = \sum_i \mathbf{w}(\sigma_i)$ .

Finally, we check condition (3) of Definition 3.1. Let  $e$  be an edge of  $M$ , and let  $\rho$  be a normal closed curve on  $\partial M$  that encircles one endpoint of  $e$ . Because  $\rho$  is homotopically trivial, we have  $\iota(\rho, \sigma) = 0$ . Thus, by Lemma 4.4, deforming the dihedral angles along  $\mathbf{w}(\sigma)$  does not change the imaginary part of  $H(\rho)$ . But  $\text{Im}(H(\rho))$  is nothing other than the sum of dihedral angles about edge  $e$ ; this sum stays constant, equal to  $2\pi$ .  $\square$

There is a convenient choice of leading–trailing deformations that will span  $T_p \mathcal{A}(\tau)$ .

PROPOSITION 4.6. *For every edge  $e_i$  of  $M$ , where  $1 \leq i \leq n$ , choose a normal closed curve  $\rho_i$  about one endpoint of  $e_i$ . In addition, if  $M$  has  $k$  cusps, choose simple closed normal curves  $\sigma_1, \dots, \sigma_{2k}$  that will span  $H_1(\partial M)$ . Then the vectors  $\mathbf{w}(\rho_i)$  and  $\mathbf{w}(\sigma_j)$  span the tangent space  $T_p \mathcal{A}(\tau)$ .*

Notice that the curves  $\rho_1, \dots, \rho_n, \sigma_1, \dots, \sigma_{2k}$  are exactly the ones whose holonomy is being considered in Definition 2.4.

PROOF. By Proposition 3.2,  $\dim T_p \mathcal{A}(\tau) = n + k$ . Thus it suffices to show that  $\mathbf{w}(\rho_i), \mathbf{w}(\sigma_j)$  span a vector space of this dimension.

First, we claim that the vectors  $\mathbf{w}(\sigma_1), \dots, \mathbf{w}(\sigma_{2k})$  are linearly independent from one another, and from each of the  $\mathbf{w}(\rho_i)$ . Suppose, after renumbering, that  $\sigma_1$  and  $\sigma_2$  are homology basis curves on the same torus of  $\partial M$ . Then  $\iota(\sigma_1, \sigma_2) = \pm 1$ , while  $\iota(\sigma_j, \sigma_2) = 0$  for every  $j \neq 1$  (including  $j = 2$ ). Similarly, since each  $\rho_i$  is homotopically trivial,  $\iota(\rho_i, \sigma_2) = 0$  for every  $i$ . Thus, by Lemma 4.4,  $\mathbf{w}(\sigma_1)$  is the only deformation among the  $\rho_i$  and  $\sigma_j$  that affects  $\text{Im}(H(\sigma_2))$ . Therefore,  $\mathbf{w}(\sigma_1)$  is independent from all the other vectors in the collection. Similarly, each  $\mathbf{w}(\sigma_j)$  is independent from all the other vectors in the collection.

To complete the proof, it will suffice to show that  $\mathbf{w}(\rho_1), \dots, \mathbf{w}(\rho_n)$  span a vector space of dimension  $n - k$ . Let  $B$  be the  $(n \times 3n)$  matrix whose  $i^{\text{th}}$  row is  $\mathbf{w}(\rho_i)$ . Then, we claim that  $\text{rank}(B) = n - k$ . The proof of this claim is virtually identical to the proof of Lemma 3.3. For every cusp  $c$ , we define a row vector  $\mathbf{r}_c \in \mathbb{R}^n$ , whose  $i^{\text{th}}$  entry is the number of endpoints that the  $i^{\text{th}}$  edge  $e_i$  has at cusp  $c$ . (This is the second half of the vector  $\mathbf{r}_c$  from Lemma 3.3.) Then, by the same argument as in that lemma, one checks that the vectors  $\mathbf{r}_c$  form a basis for the row null space of  $B$ . Therefore,  $\text{rank}(B) = n - k$ , and the  $\mathbf{w}(\rho_i), \mathbf{w}(\sigma_j)$  span a vector space of dimension  $(n - k) + 2k$ , as required.  $\square$

## 5. Volume maximization

In this section, we show how volume considerations give a way to turn an angle structure into a genuine hyperbolic metric on  $M$ . To compute the volume of an ideal hyperbolic tetrahedron, recall the Lobachevsky function  $\mathbb{J} : \mathbb{R} \rightarrow \mathbb{R}$ . Its definition is

$$(5.1) \quad \mathbb{J}(x) = - \int_0^x \log |2 \sin t| dt.$$

LEMMA 5.1. *The Lobachevsky function  $\mathbb{J}(x)$  is well defined and continuous on  $\mathbb{R}$  (even though the defining integral is improper), and periodic with period  $\pi$ . Furthermore, if  $T$  is a hyperbolic ideal tetrahedron with dihedral angles  $\alpha, \beta, \gamma$ , its volume satisfies*

$$\text{vol}(T) = \mathbb{J}(\alpha) + \mathbb{J}(\beta) + \mathbb{J}(\gamma).$$

PROOF. See, for example, Milnor [19].  $\square$

Following Lemma 5.1, we may define the volume of an angle structure in a natural way.

DEFINITION 5.2. Let  $\tau$  be an ideal triangulation of  $M$ , containing  $n$  tetrahedra. Let  $\mathcal{A}(\tau) \subset \mathbb{R}^{3n}$  be the polytope of angle structures on  $M$ . Then we define a *volume functional*  $\mathcal{V} : \overline{\mathcal{A}(\tau)} \rightarrow \mathbb{R}$ , by assigning to a point  $p = (p_1, \dots, p_{3n})$  the real number

$$\mathcal{V}(p) = \mathbb{J}(p_1) + \dots + \mathbb{J}(p_{3n}).$$

By Lemma 5.1,  $\mathcal{V}(p)$  is equal to the sum of the volumes of the hyperbolic tetrahedra associated to the angle structure  $p$ .



LEMMA 5.3. *Let  $p = (p_1, \dots, p_{3n}) \in \mathcal{A}(\tau)$  be an angle structure on  $\tau$ , and let  $\mathbf{w} = (w_1, \dots, w_{3n}) \in T_p\mathcal{A}(\tau)$  be a nonzero tangent vector at  $p$ . Then the first two derivatives of  $\mathcal{V}(p)$  satisfy*

$$\frac{\partial \mathcal{V}}{\partial \mathbf{w}} = \sum_{i=1}^{3n} -w_i \log \sin p_i \quad \text{and} \quad \frac{\partial^2 \mathcal{V}}{\partial \mathbf{w}^2} < 0.$$

*In particular,  $\mathcal{V}$  is strictly concave down on  $\mathcal{A}(\tau)$ .*

PROOF. Since the definition of  $\mathcal{V}$  is linear over the tetrahedra in  $\tau$ , it suffices to consider the volume of one tetrahedron. Thus, suppose that a tetrahedron  $T$  has angles  $p_1, p_2, p_3 > 0$ , which are changing at rates  $w_1, w_2, w_3$ . Note that, since the tangent vector  $\mathbf{w}$  must preserve the angle sum in each tetrahedron, we have  $w_1 + w_2 + w_3 = 0$ . We may also assume that at least one (hence, at least two) of the  $w_i$  are nonzero. Then, by equation (5.1) and Lemma 5.1,

$$\frac{\partial \text{vol}(T)}{\partial \mathbf{w}} = \sum_{i=1}^3 -w_i \log |2 \sin p_i| = \sum_{i=1}^3 -w_i \log \sin p_i,$$

because all sines are positive and  $\sum w_i \log 2 = 0$ . This completes the computation of the first derivative.

To compute the second derivative, assume by symmetry that  $p_1, p_2 < \pi/2$ . Differentiating  $\text{vol}(T)$  a second time, we get

$$\begin{aligned} -\frac{\partial^2 \text{vol}(T)}{\partial \mathbf{w}^2} &= w_1^2 \cot p_1 + w_2^2 \cot p_2 + w_3^2 \cot p_3 \\ &= w_1^2 \cot p_1 + w_2^2 \cot p_2 + (w_1 + w_2)^2 \frac{1 - \cot p_1 \cot p_2}{\cot p_1 + \cot p_2} \\ &= \frac{(w_1 + w_2)^2 + (w_1 \cot p_1 - w_2 \cot p_2)^2}{\cot p_1 + \cot p_2} \\ &\geq 0. \end{aligned}$$

In fact, the numerator in the next-to-last line must be strictly positive. For the numerator to be 0, we must have  $w_1 = -w_2$ , hence  $\cot p_1 = -\cot p_2$ , which is impossible when  $p_1, p_2 \in (0, \pi/2)$ . Thus  $\partial^2 \text{vol}(T)/\partial \mathbf{w}^2 < 0$ , and the volume functional  $\mathcal{V}$  is also strictly concave down.  $\square$

By Lemma 5.3, the only potential critical point of  $\mathcal{V}$  is a global maximum. As the next proposition shows, derivatives of  $\mathcal{V}$  are closely connected to the holonomy of curves on  $\partial M$ .

PROPOSITION 5.4. *Let  $C$  be a cusp torus of  $M$ , with a tessellation by boundary triangles coming from  $\tau$ . Let  $\sigma \subset C$  be an oriented normal closed curve. Recall the holonomy  $H(\sigma)$  from Definition 2.3 and the tangent vector  $\mathbf{w}(\sigma) \in T_p\mathcal{A}(\tau)$  from Definition 4.1.*

*Then for every point  $p \in \mathcal{A}(\tau)$ , we have*

$$\frac{\partial \mathcal{V}}{\partial \mathbf{w}(\sigma)} = \text{Re}(H(\sigma)).$$

PROOF. Let  $\sigma_1, \dots, \sigma_k$  be the segments of  $\sigma$ , with  $\sigma_i$  contained in boundary triangle  $\Delta_i$ . For each  $i$ , label the angles of  $\Delta_i$  as  $\alpha_i, \beta_i, \gamma_i$ , in clockwise order, such

that  $\alpha_i$  is the angle cut off by  $\sigma_i$ . Recall that, in Definition 2.3, we defined

$$\epsilon_i = \begin{cases} 1, & \text{if } \alpha_i \text{ is to the left of } \sigma_i, \\ -1, & \text{if } \alpha_i \text{ is to the right of } \sigma_i. \end{cases}$$

Comparing this with Definition 4.1 and Figure 3, we see that the vector  $\mathbf{w}(\sigma_i)$  increases angle  $\beta_i$  at rate  $\epsilon_i$ , and increases  $\gamma_i$  at rate  $-\epsilon_i$ . By Lemma 5.3,

$$\begin{aligned} \frac{\partial \mathcal{V}}{\partial \mathbf{w}(\sigma)} &= \sum_{i=1}^k \frac{\partial \mathcal{V}}{\partial \mathbf{w}(\sigma_i)} \\ &= \sum_{i=1}^k (-\epsilon_i \log \sin \beta_i + \epsilon_i \log \sin \gamma_i) \\ &= \sum_{i=1}^k \epsilon_i \log \left( \frac{\sin \gamma_i}{\sin \beta_i} \right) \\ &= \sum_{i=1}^k \epsilon_i \log |z_i| \quad \text{by equation (3.1)} \\ &= \operatorname{Re} \sum_{i=1}^k \epsilon_i \log z_i, \end{aligned}$$

as desired.  $\square$

Recall, from Section 3, that an angle structure on  $\tau$  corresponds to solving the imaginary part of the edge gluing equations. By Proposition 5.4, solving the real part of each edge equation amounts to having vanishing derivative in the direction of the corresponding deformation. This turns out to be the crucial step in the proof of Theorem 1.2.

**THEOREM 1.2.** *Let  $M$  be an orientable 3-manifold with boundary consisting of tori, and let  $\tau$  be an ideal triangulation of  $M$ . Then a point  $p \in \mathcal{A}(\tau)$  corresponds to a complete hyperbolic metric on the interior of  $M$  if and only if  $p$  is a critical point of the functional  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ .*

**PROOF.** For one direction of the theorem, suppose that  $p \in \mathcal{A}(\tau)$  is a critical point of  $\mathcal{V}$ . This angle structure defines a shape parameter on each tetrahedron of  $\tau$ . By Proposition 2.5, proving that these shape parameters give a complete hyperbolic metric on  $M$  amounts to checking the edge and completeness equations of Definition 2.4.

First, consider the edge equation about an edge  $e$ . Note that the imaginary part of the gluing equation about edge  $e$  is automatically satisfied for any angle structure. To check the real part of the gluing equation, let  $\sigma \subset \partial M$  be a normal closed curve encircling one endpoint of  $e$ . Since  $p$  is a critical point of  $\mathcal{V}$ , Proposition 5.4 implies that  $\operatorname{Re}(H(\sigma)) = 0$ , as desired. Thus the edge gluing equations are satisfied.

To check completeness, let  $C$  be a boundary torus of  $M$ , and let  $\sigma_1, \sigma_2$  be a pair of simple closed normal curves that span  $\pi_1(C) = H_1(C)$ . Recall that  $C$  is tiled by boundary triangles that truncate the tips of ideal tetrahedra. The angle structure  $p$  gives each of these triangles a Euclidean shape, well-defined up to similarity. If  $\Delta$  is a triangle in which  $\sigma_1$  and  $\sigma_2$  intersect, we may place the corners of  $\Delta$  at  $0, 1, z \in \mathbb{C}$ , and develop the other triangles of  $C$  from there. Note that, since the edge gluing

equations are satisfied, the boundary triangles fit together correctly around every vertex of  $C$ .

Let  $d_1$  and  $d_2$  be the deck transformations of  $\tilde{C}$  corresponding to  $\sigma_1$  and  $\sigma_2$ . Consider the complex numbers

$$q = d_1(0), \quad r = d_2(0), \quad s = d_1(r) = d_2(q).$$

The four points  $0, q, r, s$  form four corners of a fundamental domain for  $C$ . By the definition of holonomy (Definition 2.3) and the solution to the edge equations,

$$H(\sigma_1) = \log \frac{s - q}{r - 0}, \quad H(\sigma_2) = \log \frac{s - r}{q - 0}.$$

By Proposition 5.4, the real part of each of these holonomies is 0. Thus we have  $|s - q| = |r|$  and  $|s - r| = |q|$ , hence the fundamental domain of  $C$  is a parallelogram. Therefore,  $H(\sigma_1) = H(\sigma_2) = 0$ , and the completeness equations are satisfied for cusp  $C$ .

For the converse implication of the theorem, suppose that  $p \in \mathcal{A}(\tau)$  defines a complete metric. Then the tetrahedron shapes corresponding to  $p$  must satisfy all the edge and completeness equations of Definition 2.4: we must have  $H(\sigma) = 2\pi i$  for every closed curve  $\sigma$  encircling an endpoint of an edge, and  $H(\sigma_j) = 0$  for a collection of simple closed curves  $\sigma_1, \dots, \sigma_{2k}$  that form a basis of  $H_1(\partial M)$ . In particular, the real part of each of these holonomies is 0. But, by Proposition 4.6, the leading-trailing deformations that correspond to these closed curves span  $T_p \mathcal{A}(\tau)$ . Thus  $p$  is a critical point of  $\mathcal{V}$ , as desired.  $\square$

## 6. Extensions, applications, generalizations

This final section of the paper surveys several ways in which the Casson–Rivin program has been generalized and extended, as well as several infinite families of manifolds to which the program has been successfully applied.

Theorem 1.2 concerns manifolds with non-empty boundary that consists of tori. It is natural to ask whether similar methods can be applied to treat manifolds with more general boundary, or closed manifolds that have no boundary at all. Indeed, there has been considerable progress in these areas.

**6.1. Closed manifolds via Dehn filling.** The most straightforward way to extend Theorem 1.2 to closed manifolds is via Dehn surgery. If  $C$  is a boundary torus of  $M$ , and  $\mu, \lambda$  are simple closed normal curves that form a basis for  $H_1(C)$ , then  $M(p/q)$  is the manifold obtained by attaching a solid torus to  $C$ , such that the boundary of the meridian disk is mapped to  $p\mu + q\lambda$ . In terms of the gluing equations of Definition 2.4, attaching a disk to the closed curve  $p\mu + q\lambda$  is equivalent to solving the holonomy equation

$$(6.1) \quad p H(\mu) + q H(\lambda) = 2\pi i.$$

Just as above, the imaginary part of equation (6.1) is linear in the angles of  $\tau$ . Imposing this linear equation corresponds to taking a codimension–1 linear slice of the angle space  $\mathcal{A}(\tau)$ .

**THEOREM 6.1.** *Let  $M$  be a manifold with boundary a single torus  $C$ , and let  $\tau$  be an ideal triangulation of  $M$ . Choose a pair  $(p, q)$  of relatively prime integers, and let  $\mathcal{A}_{p/q}(\tau) \subset \mathcal{A}(\tau)$  be the set of all angle structures that satisfy the imaginary*

part of equation (6.1). Then a critical point of the volume functional  $\mathcal{V}$  on  $\mathcal{A}_{p/q}(\tau)$  yields a complete hyperbolic structure on  $M(p/q)$ , the  $p/q$  Dehn filling of  $M$ .

The analogous statement holds for fillings along multiple boundary tori.

The proof follows the same outline as Theorem 1.2; see [1, Theorem 6.2] for more details.

**6.2. Manifolds with polyhedral boundary.** There is also an analogue of Theorem 1.2 for ideal triangulations where not every face of the tetrahedra is glued to another face. Given such a partial gluing of tetrahedra, one obtains a 3-manifold  $N$  whose boundary is subdivided into ideal triangles. If the tetrahedra carry dihedral angles, then every edge along  $\partial N$  will also carry a prescribed angle. The simplest case of this is when  $N$  is an ideal polyhedron, but one may also consider cases where  $N$  has more complicated topology.

It is worth asking exactly when a 3-manifold  $N$  with polyhedral boundary, and with a fixed assignment of convex dihedral angles, carries a complete hyperbolic metric with ideal vertices that realizes those angles. The following combinatorial condition was suggested by Rivin [25]:

- (\*) For every simple closed normal curve  $\sigma \subset \partial N$  that bounds a disk in  $N$ , the sum of exterior angles along  $\sigma$  is at least  $2\pi$ , with equality iff  $\sigma$  encircles an ideal vertex.

**THEOREM 6.2.** *Let  $N$  be a 3-manifold with polyhedral boundary and prescribed (convex) dihedral angles along every edge of  $\partial N$ . Suppose that  $N$  is irreducible and atoroidal, and furthermore that  $N$  is a 3-ball, or a solid torus, or has incompressible boundary. Then  $N$  carries a complete hyperbolic metric realizing the prescribed angles if and only if condition (\*) holds. Furthermore, any hyperbolic realization is unique up to isometry.*

The case when  $N$  is a polyhedron is due to Rivin [27, 28]; the case of solid tori, to Guéritaud [8]. In both cases, the argument works by first proving that  $N$  has an ideal triangulation with a non-empty angle space  $\mathcal{A}(\tau)$ , and then proving that the volume functional  $\mathcal{V}$  has a critical point in  $\mathcal{A}(\tau)$ . This critical point gives the hyperbolic realization of  $N$ . In the preprint [29], Schlenker gives a more analytic and general argument in the case of incompressible boundary. Conjecturally, the special hypotheses of all these papers are not needed: all that should be necessary is that  $N$  is irreducible and contains no incompressible tori, and that the angle assignments satisfy (\*). See [5, Conjecture 2.4].

**6.3. Generalized angle structures.** All of the angle structures discussed so far have involved strictly positive dihedral angles. One natural generalization of the definitions in Section 3 would be to allow negative angles, or more generally, to consider angles mod  $2\pi$ .

A *generalized angle structure* is an assignment of a real number to every pair of opposite edges in a tetrahedron, so that equations (2) and (3) of Definition 3.1 are satisfied, but the inequalities are discarded. The set of all such assignments is denoted  $\mathcal{GA}(\tau)$ . Luo and Tillmann showed that for any ideal triangulation  $\tau$  of a manifold with torus boundary,  $\mathcal{GA}(\tau)$  is always non-empty [18]; in other words, there is no analogue of Theorem 1.1. Furthermore, they establish a linear-algebraic duality between angle structures and normal surfaces (see also Rivin [28]). For

example, the obstruction to finding a non-empty positive polytope  $\mathcal{A}(\tau) \subset \mathcal{GA}(\tau)$  is a certain branched normal surface with non-negative Euler characteristic.

Generalizing further, an  $S^1$ -valued angle structure on a triangulation  $\tau$  is an assignment of a real number (mod  $2\pi$ ) to every pair of opposite edges in a tetrahedron, such that

- (1) Around each ideal vertex of a tetrahedron, the dihedral angles sum to  $\pi$  (mod  $2\pi$ ),
- (2) Around each edge of  $M$ , the dihedral angles sum to 0 (mod  $2\pi$ ).

The set of all  $S^1$ -valued angle structures on  $\tau$  is denoted  $\mathcal{SA}(\tau)$ .

Just as with real-valued solutions in  $\mathcal{GA}(\tau)$ , Luo showed that the existence of  $S^1$ -valued solutions is extremely general. For any triangulated closed pseudo-manifold (i.e., any cell complex obtained by gluing tetrahedra in pairs along all of their faces, whatever the link of a vertex), he showed that  $\mathcal{SA}(\tau) \neq \emptyset$ , and is a closed smooth manifold [15, Proposition 2.6].

Even though the existence of  $S^1$ -valued angle structures does not distinguish the class of hyperbolic manifolds, studying the volume of such a structure can still yield geometric information. Recall from Lemma 5.1 that the Lobachevsky function  $\mathbb{J}$  is  $\pi$ -periodic; as a result, Definition 5.2 of the volume functional  $\mathcal{V}$  extends in a natural way to  $\mathcal{SA}(\tau)$ . Because  $\mathcal{SA}(\tau) \neq \emptyset$  is a compact manifold,  $\mathcal{V}$  must achieve a maximum. The work of Luo [15] and Luo–Tillmann [17] uses this maximum point to either solve a generalized version of the gluing equations (which yields a representation from a double branched cover of  $M$  into  $PSL(2, \mathbb{C})$ ), or find certain highly restrictive normal surfaces in  $M$ . See Luo’s survey paper in this volume [14] for more details.

**6.4. Volume estimates.** Recall that, by Lemma 5.3, the volume functional  $\mathcal{V}$  is concave down on  $\mathcal{A}(\tau)$ . As a result, any critical point of  $\mathcal{V}$  must actually be the global maximum of the function over the compact closure  $\overline{\mathcal{A}(\tau)}$ . Thus Theorem 1.2 has the following corollary.

**THEOREM 6.3.** *Let  $M$  be an orientable 3-manifold with boundary consisting of tori, and let  $\tau$  be an ideal triangulation of  $M$ . Suppose that  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$  has a critical point at  $p \in \mathcal{A}(\tau)$ . Then, for any point  $q \in \overline{\mathcal{A}(\tau)}$ ,*

$$\mathcal{V}(q) \leq \text{vol}(M),$$

*with equality iff  $q = p$  (i.e., iff  $q$  gives the complete hyperbolic metric on  $M$ ).*

In fact, the analogous statement also holds in the settings of Section 6.1 (Dehn filling) or Section 6.2 (polyhedral boundary). The uniqueness of a critical point of  $\mathcal{V}$  turns out to be the key idea in Rivin’s proof of the uniqueness statement of Theorem 6.2.

Theorem 6.3 allows to compute effective, combinatorial volume estimates for hyperbolic 3-manifolds. In certain settings, the combinatorics of a 3-manifold naturally guides a choice of triangulation  $\tau$ , and the same combinatorial data provide a convenient point  $q \in \overline{\mathcal{A}(\tau)}$ . Then, Theorem 6.3 says that  $\mathcal{V}(q)$  is a lower bound on the volume of  $M$ . This approach is illustrated by the first family of manifolds to which the Casson–Rivin method was successfully applied [11].

A decade after Colin de Verdière, Casson, and Rivin developed the theory of volume maximization in the early 1990s, it was Guéritaud who first applied the method to find the hyperbolic metrics on an infinite family of manifolds, namely

punctured torus bundles [11]. In an appendix to the same paper, Futer extended the method to two-bridge links [11, Appendix]. For both of these families of manifolds, the existence of hyperbolic metrics was well-known, but the volume estimates coming from Theorem 6.3 were both new and sharp. Combined with Dehn surgery techniques, the volume estimates from [11] also give explicit, combinatorial bounds for the volume of several families of knot and link complements [6, 7], as well as of a number of closed manifolds [23].

It is worth asking whether the existence of a critical point of  $\mathcal{V}$  in  $\mathcal{A}(\tau)$  is actually necessary for the volume inequality of Theorem 6.3. Casson conjectured that this inequality holds for any angled triangulation  $\tau$ , whether or not the tetrahedra of this triangulation can be given positively oriented shapes in the hyperbolic metric on  $M$ . (Note that, by Theorem 1.1, the hyperbolic metric must exist.) If proved true, this conjecture would provide a practical tool for finding volume estimates on many more families of 3-manifolds.

**6.5. Canonical triangulations.** Nearly everything discussed thus far in this paper has depended on the choice of triangulation. As it turns out, many of the methods already discussed can show that a particular triangulation is *geometrically canonical* for  $M$ .

Given a hyperbolic 3-manifold  $M$  with  $k$  cusps, let  $H_1, \dots, H_k$  be disjoint horospherical neighborhoods of the cusps. Then the Ford-Voronoi domain  $\mathcal{F}$  is the set of all points in  $M$  that have a unique shortest path to the union of the  $H_i$ . This is an open set in  $M$ , whose complement  $L = M \setminus \mathcal{F}$  is a compact 2-complex, called the *cut locus*. The dual to  $L$  is an ideal polyhedral decomposition  $\mathcal{P}$  of  $M$ ; the  $n$ -cells of  $\mathcal{P}$  are in bijective correspondence with the  $(3-n)$ -cells of  $L$ . This is called the *canonical polyhedral decomposition* of  $M$ , relative to the cusp neighborhoods  $H_i$ .

The combinatorics of  $L$  —and therefore, of  $\mathcal{P}$ — depends only on the relative volumes of the  $H_i$ . In particular, if  $M$  has only one cusp, there are no choices whatsoever, and  $\mathcal{P}$  is completely determined by the hyperbolic metric. If the horoballs in  $\mathbb{H}^3$  obtained by lifting the cusp neighborhoods  $H_i$  are in “general position,” every vertex of  $\tilde{L}$  will have exactly four closest horoballs. Thus each vertex of  $L$  will meet four edges, and the dual polyhedral decomposition  $\mathcal{P}$  will generically be a triangulation.

If  $\tau$  is a given triangulation (or, more generally, a given polyhedral decomposition), proving that  $\tau$  is canonical amounts to verifying finitely many inequalities about nearest horoballs. Epstein and Penner found a way to translate these inequalities into convexity statements in the Minkowski space  $\mathbb{R}^{3+1}$ , where  $\mathbb{H}^3$  is modeled by a hyperboloid [4]. More recently, Guéritaud discovered that the convexity inequalities that imply canonicity (once translated to inequalities involving dihedral angles of the polyhedra) can be verified using information obtained in the course of showing that  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$  has a critical point [9].

To date, the method of angled triangulations has found both the hyperbolic metric and the canonical polyhedral decomposition of several families of 3-manifolds: punctured-torus bundles [9, Theorem 1.11.1], two-bridge links [9, Theorem 2.1.7], and certain special arborescent links [9, Theorem 2.3.1]. Guéritaud used the same ideas to find the canonical triangulations for convex cores of quasi-Fuchsian punctured-torus groups [10]. All of these families of manifolds actually have closely

related combinatorial features, with the structure of the triangulation effectively determined by the combinatorics of  $SL(2, \mathbb{Z})$  and continued fractions.

There is a recent result that brings together several themes from this section. Suppose that  $M$  is a hyperbolic 3-manifold with  $k$  cusps, and the canonical decomposition of  $M$  is indeed a triangulation. If we perform Dehn filling along one of these cusps, the result will (generically) be another hyperbolic manifold  $M(p/q)$ , with a new geometry. In [12], Guéritaud and Schleimer use the canonical triangulation of  $M$  to completely describe the canonical triangulation of its generic Dehn fillings. Their argument uses Theorems 6.1 and 6.2 to construct a triangulated solid torus with the right shape and glue it into  $M(p/q)$ .

**6.6. Weil rigidity.** We close this paper with an application whose statement has nothing to do with triangulations. The following rigidity theorem, due to Weil [32], precedes Mostow–Prasad rigidity by a dozen years.

**THEOREM 6.4.** *Let  $M$  be a 3-manifold with boundary consisting of tori. Then any complete hyperbolic metric on the interior of  $M$  is locally rigid: there is no local deformation of the metric through other complete hyperbolic metrics.*

We thank Marc Culler and Feng Luo for a fruitful discussion that produced the following extremely short proof.

**PROOF.** Suppose that  $M$  admits a complete hyperbolic metric. Choose horoball neighborhoods  $H_1, \dots, H_k$  about the cusps of  $M$ . Then, as in Section 6.5, this choice of cusp neighborhoods determines a decomposition  $\mathcal{P}$  of  $M$  into ideal polyhedra. Luo, Schleimer, and Tillmann showed that  $M$  has a finite-sheeted cover  $N$ , in which the lift of  $\mathcal{P}$  decomposes into positively oriented ideal tetrahedra [16]. Let  $\tau$  be this positively oriented ideal triangulation of  $N$ .

By Theorem 1.2, the complete hyperbolic metric on  $N$  (which was obtained by lifting the metric on  $M$ ) represents a critical point of  $\mathcal{V} : \mathcal{A}(\tau) \rightarrow \mathbb{R}$ . By Lemma 5.3, this critical point is unique. But any local deformation of the complete metric on  $M$  would lift to a deformation of the metric on  $N$ , which would violate Lemma 5.3.  $\square$

## References

- [1] Ken Chan, *Constructing hyperbolic 3-manifolds*, Undergraduate thesis with Craig Hodgson, University of Melbourne, 2002.
- [2] Young-Eun Choi, *Positively oriented ideal triangulations on hyperbolic three-manifolds*, *Topology* **43** (2004), no. 6, 1345–1371.
- [3] Yves Colin de Verdière, *Un principe variationnel pour les empilements de cercles*, *Invent. Math.* **104** (1991), no. 3, 655–669.
- [4] David B. A. Epstein and Robert C. Penner, *Euclidean decompositions of noncompact hyperbolic manifolds*, *J. Differential Geom.* **27** (1988), no. 1, 67–80.
- [5] David Futer and François Guéritaud, *Angled decompositions of arborescent link complements*, *Proc. London Math. Soc.* **98** (2009), no. 2, 325–364.
- [6] David Futer, Efstratia Kalfagianni, and Jessica S. Purcell, *On diagrammatic bounds of knot volumes and spectral invariants*, *Geom. Dedicata* **147** (2010), 115–130.
- [7] ———, *Cusp areas of Farey manifolds and applications to knot theory*, *Int. Math. Res. Notices* (to appear), doi: 10.1093/imrn/rnq037, arXiv:0808.2716.
- [8] François Guéritaud, *Deforming ideal solid tori*, arXiv:0911.3067.
- [9] ———, *Géométrie hyperbolique effective et triangulations idéales canoniques en dimension 3*, Ph.D. thesis, Université de Paris XI (Orsay), 2006.
- [10] ———, *Triangulated cores of punctured-torus groups*, *J. Differential Geom.* **81** (2009), no. 1, 91–142.

- [11] François Guéritaud and David Futer (appendix), *On canonical triangulations of once-punctured torus bundles and two-bridge link complements*, *Geom. Topol.* **10** (2006), 1239–1284.
- [12] François Guéritaud and Saul Schleimer, *Canonical triangulations of Dehn fillings*, *Geom. Topol.* **14** (2010), 193–242.
- [13] Marc Lackenby, *Word hyperbolic Dehn surgery*, *Invent. Math.* **140** (2000), no. 2, 243–282.
- [14] Feng Luo, *Triangulated 3-manifolds: from Haken’s normal surfaces to Thurston’s algebraic equations*, This volume.
- [15] ———, *Volume optimization, normal surfaces and Thurston’s equation on triangulated 3-manifolds*, arXiv:0903.1138.
- [16] Feng Luo, Saul Schleimer, and Stephan Tillmann, *Geodesic ideal triangulations exist virtually*, *Proc. Amer. Math. Soc.* **136** (2008), no. 7, 2625–2630.
- [17] Feng Luo and Stephan Tillmann, *Volume optimization on minimally triangulated 3-manifolds*, Preprint.
- [18] ———, *Angle structures and normal surfaces*, *Trans. Amer. Math. Soc.* **360** (2008), 2849–2866.
- [19] John Milnor, *Hyperbolic geometry: the first 150 years*, *Bull. Amer. Math. Soc. (N.S.)* **6** (1982), no. 1, 9–24.
- [20] George D. Mostow, *Strong rigidity of locally symmetric spaces*, Princeton University Press, Princeton, N.J., 1973, *Annals of Mathematics Studies*, No. 78.
- [21] Walter D. Neumann, *Combinatorics of triangulations and the Chern-Simons invariant for hyperbolic 3-manifolds*, *Topology ’90* (Columbus, OH, 1990), Ohio State Univ. Math. Res. Inst. Publ., vol. 1, de Gruyter, Berlin, 1992, pp. 243–271.
- [22] Walter D. Neumann and Don Zagier, *Volumes of hyperbolic three-manifolds*, *Topology* **24** (1985), no. 3, 307–332.
- [23] Carlo Petronio and Andrei Vesnin, *Two-sided bounds for the complexity of cyclic branched coverings of two-bridge links*, *Osaka J. Math.* **46** (2009), no. 4, 1077–1095.
- [24] Gopal Prasad, *Strong rigidity of  $\mathbf{Q}$ -rank 1 lattices*, *Invent. Math.* **21** (1973), 255–286.
- [25] Igor Rivin, *On geometry of convex ideal polyhedra in hyperbolic 3-space*, *Topology* **32** (1993), no. 1, 87–92.
- [26] ———, *Euclidean structures on simplicial surfaces and hyperbolic volume*, *Ann. of Math. (2)* **139** (1994), no. 3, 553–580.
- [27] ———, *A characterization of ideal polyhedra in hyperbolic 3-space*, *Ann. of Math. (2)* **143** (1996), no. 1, 51–70.
- [28] ———, *Combinatorial optimization in geometry*, *Adv. in Appl. Math.* **31** (2003), no. 1, 242–271.
- [29] Jean-Marc Schlenker, *Hyperbolic manifolds with polyhedral boundary*, arXiv:math/0111136.
- [30] William P. Thurston, *The geometry and topology of three-manifolds*, Princeton Univ. Math. Dept. Notes, 1979.
- [31] ———, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, *Bull. Amer. Math. Soc. (N.S.)* **6** (1982), no. 3, 357–381.
- [32] André Weil, *On discrete subgroups of Lie groups*, *Ann. of Math. (2)* **72** (1960), 369–384.

DEPARTMENT OF MATHEMATICS, TEMPLE UNIVERSITY, PHILADELPHIA, PA 19122, USA  
*E-mail address:* dfuter@temple.edu

LABORATOIRE PAUL PAINLEVÉ, CNRS UMR 8524, UNIVERSITÉ DE LILLE 1, 59650 VILLE-NEUVE D’ASCQ, FRANCE  
*E-mail address:* Francois.Gueritaud@math.univ-lille1.fr