

You can hear the shape of a billiard table: Symbolic dynamics and rigidity for flat surfaces

Moon Duchin,^{*} Viveka Erlandsson,^{**}
Christopher J. Leininger[‡] and Chandrika Sadanand[§]

Abstract. We give a complete characterization of the relationship between the shape of a Euclidean polygon and the symbolic dynamics of its billiard flow. We prove that the only pairs of tables that can have the same bounce spectrum are right-angled tables that differ by an affine map. The main tool is a new theorem that establishes that a flat cone metric is completely determined by the *support* of its Liouville current.

Mathematics Subject Classification (2020). 37B10, 57M50, 58J50.

Keywords. Billiards, surfaces, Euclidean, cone metric, Liouville current, symbolic dynamics.

1. Introduction

There is a well-established line of *spectral rigidity* problems in geometry, where one tries to show that various discrete invariants of an object determine its geometry. In 1966, Kac memorably asked if one could *hear the shape of a drum*; or more precisely whether the shape of a planar domain is determined by the spectrum of its Laplacian [28]. In this paper we define a “bounce spectrum” for polygons, recording the symbolic dynamics of the billiard flow, and we explore the question of how precisely this determines the polygon, and thus whether one can “hear” the shape of the polygon from the bounce spectrum.

Every Euclidean polygon P supports a *billiard flow*, in which a point mass in the interior travels in a straight line until it hits a wall, then bounces off the wall with *optical reflection*: angle of incidence equals angle of reflection. Trajectories hitting

^{*}This work was initiated with funding from M.D.’s NSF CAREER award DMS-1255442.

^{**}V.E. was partially supported by Academy of Finland project #297258.

[‡]C.J.L. is partially supported by NSF grant DMS-1510034 and also acknowledges support from NSF grants DMS 1107452, 1107263, 1107367 “RNMS: GEometric structures And Representation varieties” (the GEAR Network).

[§]C.S. acknowledges support from a Postdoctoral Fellowship at the Einstein Institute of Mathematics, Hebrew University.

corners are regarded as singular and discarded. A polygon together with this billiard flow constitutes a dynamical system called a billiard table (or simply *table*) below.

Suppose the sides of a simply connected (finite-sided, not necessarily convex) polygon P are labeled in cyclic order by letters from an ordered alphabet $\mathcal{A}$. Then for a nonsingular billiard path, $\gamma: \mathbb{R} \rightarrow P$, we can record the labels of the sides of P in the order that they are encountered along γ , forming a bi-infinite sequence of labels called the *bounce sequence*, an element of $\mathcal{A}^{\mathbb{Z}}$. Here the zeroth term of the sequence corresponds to the first side encountered on $\gamma([0, \infty))$, and so starting sooner or later along a trajectory amounts to shifting the corresponding bounce sequence. The set of all bounce sequences that can occur on P is denoted $B(P)$ and called the *bounce spectrum* of the billiard table. This closely resembles the construction of *cutting sequences*, a fundamental concept in symbolic dynamics. If there exist cyclic labelings of the sides of polygons P_1 and P_2 with letters from $\mathcal{A}$ which induce a bijection of bounce spectra, then we will identify the spectra and write $B(P_1) = B(P_2)$.

It is immediately clear that applying a *similarity* (any combination of dilation, reflection, and rotation) to a polygonal table does not change its bounce spectrum. Upon investigation, one quickly observes that the coordinate-affine map given by $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ preserves optical reflection in horizontal and vertical edges. It follows that any two rectangles R_1, R_2 have the same bounce spectrum $B(R_1) = B(R_2)$, and more generally that $B(P_1) = B(P_2)$ for any two right-angled tables (all angles in $\frac{\pi}{2}\mathbb{N}$) that are related by an affine map. Outside of this exception, we get the strongest possible rigidity result: the bounce spectrum determines not only the precise angles in the polygon, but also the proportions of side lengths.

Bounce theorem. *If two simply connected Euclidean polygons P_1, P_2 have $B(P_1) = B(P_2)$, then either P_1, P_2 are right-angled and affinely equivalent, or they are similar polygons.*

Remark. Pursuing the analogy between rigidity for the bounce spectrum and the length spectrum, we can regard the cyclic labeling of edges as a way to identify sides in one polygon with sides in the other; this is like the *marked* length spectral problem for manifolds, in which a correspondence of curves is specified. It is natural to wonder about the unmarked problem for billiard tables. In a new preprint, Calderon–Coles–Davis–Lanier–Oliveira [8] prove that for an arbitrary labeling of the sides of a polygon, one can “hear” edge adjacency (i.e., detect which labels correspond to incident edges) from the bounce sequences. Combining this result with ours, we obtain bounce-spectral rigidity in the unmarked case as well.

This theorem also shows that the Laplace spectrum and the bounce spectrum contain different information. It is well known that the Laplace spectrum determines the area and perimeter of a plane domain, from which it follows that rectangles can be distinguished by their Laplace spectra, while all having the same bounce spectrum.

On the other hand, the “propeller pair” constructions of Gordon–Webb–Wolpert [20] have the same Laplace spectrum but, by our main theorem, different bounce spectra.¹

In the special case of rational tables (all angles in $\pi\mathbb{Q}$), there is a well-established toolkit for studying the billiard flow centered on the “unfolding” to a closed *translation surface* and all the machinery that comes with that. A recent sequence of papers of Bobok–Troubetzkoy [2–4] culminates in a proof of symbolic rigidity (as in the Bounce theorem above) when one of the tables is assumed to be rational. In fact, their results are stronger in that setting: they show that only part of the bounce spectrum (the bounce sequence of a single generic trajectory, or the subset of periodic sequences) suffices to determine the table.

Our proof works for general tables, and we note that this setting is quite different; for instance, rational tables have periodic trajectories in a dense set of directions, while it is unknown if all irrational tables (or even all triangular ones!) have even a single periodic trajectory.²

Our proof utilizes essentially the entire bounce spectrum in a crucial way. On the other hand, any set that determines the bounce spectrum clearly also determines the shape of the table by the Bounce theorem. For example, we have the following corollary. A *generalized diagonal* in a Euclidean polygon P is a billiard trajectory $\gamma: [a, b] \rightarrow P$ that starts and ends at a vertex of the polygon. Given a cyclic labeling of the sides of P , a generalized diagonal induces a *finite* bounce sequence, and we let $B_\Delta(P)$ denote the *countable* set of bounce sequences of generalized diagonals.

Corollary 1. *If two simply connected Euclidean polygons P_1, P_2 have $B_\Delta(P_1) = B_\Delta(P_2)$, then either P_1, P_2 are right-angled and affinely equivalent, or they are similar polygons.*

The Bounce theorem is a consequence of a result about geodesic currents associated to flat metrics. To state this result, we require a few definitions; see Section 2 below for fuller background. The space of equivalence classes of *flat metrics* (nonpositively curved Euclidean cone metrics) on a closed, oriented surface S is denoted $\text{Flat}(S)$. Associated to each metric $\varphi \in \text{Flat}(S)$ is a geodesic current L_φ called the *Liouville current*, which is formally a measure on the double boundary of the universal cover of S . We affirmatively answer an open question of Banković–Leininger showing that the support of L_φ determines φ , up to affine deformation; see [1, Section 6]. That is, a flat metric is not only determined by its geodesic current but even by the support alone, which is quite different from the hyperbolic case where currents have full support. This Support Rigidity theorem yields the Bounce theorem via a new unfolding technique we introduce below.

¹See also [8, 31], in which the authors investigate what geometric properties of a polygon can be reconstructed from just a finite part of the bounce spectrum, Laplace spectrum, or length spectrum.

²Indeed, in [38], Schwartz proves that there is a sequence of triangles converging to the $(30, 60, 90)$ triangle for which the combinatorial length of the shortest periodic billiard trajectory tends to infinity. See also Hooper [24] for more on instability of periodic trajectories.

Support Rigidity theorem. *Suppose φ_1, φ_2 are two unit-area flat metrics whose Liouville currents have the same support, $\text{supp}(L_{\varphi_1}) = \text{supp}(L_{\varphi_2})$. Then φ_1, φ_2 differ by an affine deformation, up to isotopy. If either metric has holonomy of order greater than 2 (i.e., is not induced by a quadratic differential, or does not support a foliation by straight lines), then equal support implies that φ_1 and φ_2 differ by isometry, isotopic to the identity.*

We briefly describe the steps involved in the proof of the Support Rigidity theorem. A key observation for carrying out this program is that the support of L_φ consists precisely of the closure of the set of nonsingular $\tilde{\varphi}$ -geodesics in $\tilde{S}$; see Proposition 7 in Section 2.5. We then use a technical device called chains (see Section 2.5) to fix an identification of the cone points, and from these an identification of the saddle connections and the nonsingular geodesics, between φ_1 and φ_2 . This lets us build a careful correspondence of triangulations between φ_1 and φ_2 , and even of directions of travel as trajectories cross the edges of a triangulation. This can be used to show that the holonomy group (the rotations observed when transporting a tangent vector around a loop) is the same for the two metrics.

Our goal is to construct an affine map $(S, \varphi_1) \rightarrow (S, \varphi_2)$ isotopic to the identity. We carry this out by picking a geometric triangulation of S with respect to φ_1 and then proving that the map to the corresponding triangulation in φ_2 can be adjusted by an isotopy first on the vertices, then on the edges, and finally triangles, to produce a map which is affine on each of the triangles. Next we must analyze the piecewise affine maps and show that they are globally affine, and in fact isometries, up to the controlled exceptions identified in the theorem statement.

To prove the Bounce theorem from the Support Rigidity theorem, we consider arbitrary nonpositively curved “unfoldings” X of a polygon P and observe that, as with the unfoldings to translation surfaces in the rational case, the nonsingular geodesics on X correspond to billiard trajectories in P . We have seen that the set $\text{supp}(L_\varphi)$ coarsely encodes (the closure of) the set of nonsingular φ -geodesics. The bounce spectrum of P records how geodesics cut through edges of a triangulation and can thus be viewed as providing the same coarse information. From two polygons P_1, P_2 with $B(P_1) = B(P_2)$, we can find a pair of nonpositively curved unfoldings X_1, X_2 with a common underlying topological surface S such that their nonsingular geodesics can be identified. Appealing to the Support Rigidity theorem, we conclude that X_1, X_2 differ by an affine map (generically an isometry, or similarity if we do not normalize the areas). With some care, this induces a suitable affine map between P_1 and P_2 . For Corollary 1, we observe that generalized diagonals correspond to saddle connections in the unfolding and use limits of codings to understand the nonsingular geodesic as limits of saddle connections.

The ideas used to deduce the Bounce theorem from the Support Rigidity theorem can be applied to other settings as well: any data that coarsely determines the nonsingular geodesics of a flat surface can be seen to determine the flat metric, up to affine equivalence. For example, in Section 4 we sketch a similar result for *cutting sequences*.

1.1. Connections to symbolic dynamics literature. The results of Bobok and Troubetzkoy mentioned above [2–4] are directly relevant to our work, and the results of Calderon et al. [8] give constructive results on the exact questions treated here. These are part of a substantial literature relating symbolic dynamics to the geometry of billiards and associated flat surfaces. Variants on bounce sequences include cutting sequences and Sturmian sequences. Characterizing the sequences occurring in Veech surfaces (such as regular polygons and square-tiled surfaces) appears in the work of Morse and Hedlund from the 1930s and 40s [35, 36] and includes very recent work of Smillie–Ulcigrai [39, 40], Davis [14, 15], Davis–Pasquinelli–Ulcigrai [16], and Johnson [27]. Beyond the Veech case, there is a long string of papers of Lopez–Narbel developing a language-theoretic formulation of generalized Sturmian sequences for interval exchange transformations, beginning with [30].

Complexity of billiards has been studied via symbolic coding by Katok [29], Hubert [25], Troubetzkoy [42], and Hubert–Vuillon [26]. From the point of view of determining geometric information from a bounce sequence, Galperin, Krüger, and Troubetzkoy [19], for example, prove sharp relationships between periodic trajectories and periodic bounce sequences, which in turn influenced the Flat Strip theorem of Hassell–Hillairet–Marzuola [21]; see Theorem 4. We refer the reader to these works and their references for further discussion of the many connections between symbolic coding and geometry.

Acknowledgements. The authors would like to thank all the participants and visitors to the Polygonal Billiards Research Cluster held at Tufts University in Summer 2017. In particular, we thank Pat Hooper, Rich Schwartz, Caglar Uyanik, and the authors of [8] for illuminating conversations on billiards and symbolic dynamics. Erlandsson and Leininger would like to thank the School of Mathematics at Fudan University and the Mathematics Research Centre and the University of Warwick, respectively, for their hospitality while this project was being completed.

2. Preliminaries

In this section we discuss background, establish notation, and give preliminary results we will use in this paper. Throughout, S will denote a closed, oriented surface of genus at least 2 and $p: \tilde{S} \rightarrow S$ will denote the universal covering map. We also fix an action of $\pi_1 S$ on $\tilde{S}$ by covering transformations.

2.1. Spaces of geodesics. Fix once and for all an arbitrary hyperbolic metric ρ on S , and let $\tilde{\rho} = p^*(\rho)$ be the pullback to $\tilde{S}$, thus specifying an identification of $\tilde{S}$ with the hyperbolic plane. Let S_∞^1 denote the circle at infinity bounding $\tilde{\rho}$, and equip it with the action of $\pi_1 S$ obtained by extending the action on $\tilde{S}$. Given any other geodesic metric m on S , with the pullback $\tilde{m} = p^*(m)$ on $\tilde{S}$, the identity on $\tilde{S}$

is a $\pi_1 S$ -equivariant quasi-isometry to $\tilde{\rho}$, and hence the Gromov boundary of $\tilde{m}$ is identified with S_∞^1 ; see [7, Chapter III.H.3]. In particular, we view S_∞^1 as the Gromov boundary of the pullback of *any* metric on S .

Let $\mathcal{G}(\tilde{S})$ denote the set of unordered pairs of distinct points in S_∞^1 , that is,

$$\mathcal{G}(\tilde{S}) = \{\{x, y\} \mid x, y \in S_\infty^1, x \neq y\} = (S_\infty^1 \times S_\infty^1 \setminus \Delta) / \sim,$$

where Δ is the diagonal and $\sim$ is the equivalence relation $(x, y) \sim (y, x)$. The action of $\pi_1 S$ on S_∞^1 determines an action on $\mathcal{G}(\tilde{S})$.

For any geodesic metric m on S an $\tilde{m}$ -geodesic on $\tilde{S}$ is a map $\tilde{\gamma}: I \rightarrow \tilde{S}$ which is an isometric embedding from a (finite, infinite, or bi-infinite) interval $I \subset \mathbb{R}$ to $(\tilde{S}, \tilde{m})$. Below, we will use the term *geodesic* for the map and its image interchangeably. If we need to distinguish them, we will refer to parametrized and unparametrized geodesics. Downstairs on S , an m -geodesic γ is the composition of an $\tilde{m}$ -geodesic with the covering projection; that is, $\gamma = p \circ \tilde{\gamma}: I \rightarrow S$. These are locally isometric embeddings, but for arbitrary metrics, not all locally isometric embeddings to S are obtained in this way. However, for nonpositively curved metrics (e.g., the hyperbolic metric ρ or flat metrics), every locally isometric embedding $\gamma: I \rightarrow S$ lifts to an isometric embedding since $\tilde{S}$ is a CAT(0) space; see [7, Proposition II.1.4].

We let $\mathcal{G}(\tilde{m})$ denote the space of all unparametrized bi-infinite $\tilde{m}$ -geodesics on $\tilde{S}$ given the Chabauty–Fell topology (see [10, 18] and [9]), or equivalently the quotient of the compact–open topology on parametrized geodesics, equipped with the action of $\pi_1 S$ induced by its action on $\tilde{S}$. A sequence $\{\tilde{\gamma}_n\}_{n=1}^\infty$ in $\mathcal{G}(\tilde{m})$ converges to a geodesic $\tilde{\gamma}$ if and only if the geodesics can be parametrized so that the sequence of parametrized geodesics converges uniformly on compact sets to the parametrization of $\tilde{\gamma}$. There is a continuous, closed, proper, surjective map

$$\partial_{\tilde{m}}: \mathcal{G}(\tilde{m}) \rightarrow \mathcal{G}(\tilde{S})$$

defined by setting $\partial_{\tilde{m}}(\tilde{\gamma})$ to be the endpoints at infinity of the geodesic. When $\tilde{m}$ has negative curvature, $\partial_{\tilde{m}}$ is a homeomorphism – in particular, $\partial_{\tilde{\rho}}$ is a homeomorphism. Below, we will build up to a better understanding of $\partial_{\tilde{\varphi}}$ for the case of flat metrics φ .

We say that pairs $\{x, y\}, \{x', y'\} \in \mathcal{G}(\tilde{S})$ *link* if the two points x' and y' are in different components of $S_\infty^1 \setminus \{x, y\}$. A pair of $\tilde{m}$ -geodesics with linking endpoints necessarily intersect each other, though even when $\tilde{m}$ is CAT(0), the intersection may not be transverse; see the left-hand side of Figure 1. If $\{x, y\}, \{x', y'\} \in \mathcal{G}(\tilde{S})$ do not link, then we say that $\{x'', y''\}$ is *between* $\{x, y\}$ and $\{x', y'\}$ if, with respect to some choice of ordering of each of the three pairs, the six points appear cyclically as

$$x \leq x'' \leq x' \leq y' \leq y'' \leq y$$

as in the right-hand side of Figure 1. The terminology is suggestive of the behavior of geodesics with those endpoints (although for general metrics, the actual intersection patterns of geodesics can be more complicated). We write $[\{x, y\}, \{x', y'\}]$ for the set of all pairs $\{x'', y''\}$ between $\{x, y\}$ and $\{x', y'\}$.

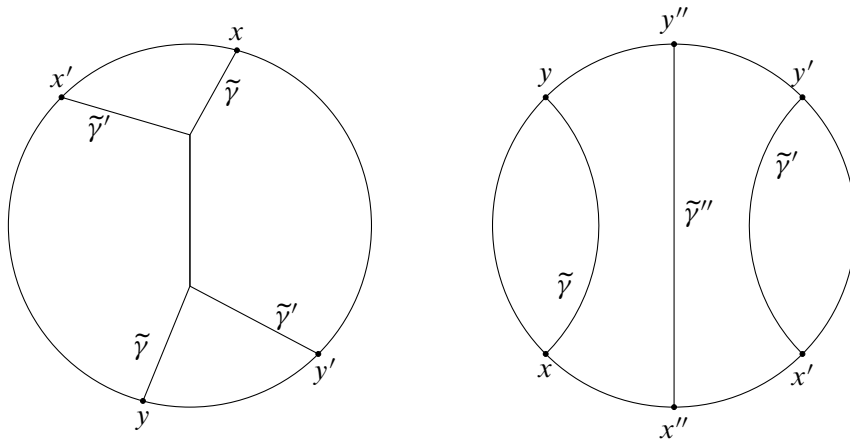


Figure 1. Left: Two geodesics with linking endpoints that do not meet transversely. Right: A geodesic $\tilde{\gamma}''$ between geodesics $\tilde{\gamma}$ and $\tilde{\gamma}'$, illustrating the notion of betweenness of their endpoints, $\{x'', y''\} \in [\{x, y\}, \{x', y'\}]$ in $\mathcal{G}(\tilde{S})$.

2.2. Geodesic currents. A *geodesic current* on S is a $\pi_1 S$ -invariant Radon measure on $\mathcal{G}(\tilde{S})$. Let $C(S)$ denote the set of all geodesic currents on S with the weak* topology. Any essential closed curve γ on S determines a geodesic current of the same name which is the counting measure on the set of pairs of endpoints on S_∞^1 of $p^{-1}(\gamma)$. In fact, Bonahon proved that currents are a completion of the set of closed curves, in the sense that the set of real multiples of associated currents

$$\{t \cdot \gamma \mid t \in \mathbb{R}_+, \gamma \subset S \text{ closed curve}\}$$

forms a dense subset of $C(S)$. The geometric intersection number between two closed curves γ and δ is the minimal number of transverse double-points of intersection among all curves γ' and δ' homotopic to γ and δ , respectively. This is realized by the ρ -geodesic representatives, and Bonahon proved that the geometric intersection number has a continuous extension to the full space of currents.

Theorem 2 ([5, 6]). *There is a continuous, bilinear function*

$$\iota: C(S) \times C(S) \rightarrow \mathbb{R}$$

such that for every pair of closed curves γ, δ , with associated geodesic currents of the same name, $\iota(\gamma, \delta)$ recovers the geometric intersection number.

Another important class of geodesic currents is the *Liouville currents* L_m associated to certain types of metrics m . Liouville currents have the geometricity property that the intersection form recovers lengths: for every essential closed curve γ on S ,

$$\iota(\gamma, L_m) = \ell_m(\gamma), \quad (2.1)$$

where $\ell_m(\gamma)$ is the length of the m -geodesic representative of the homotopy class of γ . Such currents exist for a wide range of metrics, generalizing the classical Liouville measure on geodesics in the hyperbolic plane given in terms of cross-ratios; see [1, 6, 11–13, 17, 23, 37]. The construction of the Liouville current for a flat metric and an investigation of its various properties was the focus of [17] and [1]. We describe the key properties in Section 2.5, setting the stage for the more detailed analysis we will carry out in Section 3.

2.3. Flat metrics, geodesics, and holonomy. A *flat metric* on S will mean a non-positively curved Euclidean cone metric φ on S . This is a singular Riemannian metric, locally isometric to $\mathbb{R}^2$ away from finitely many cone singularities $\Sigma = \Sigma(\varphi) \subset S$, each with cone angle greater than 2π . Since φ is nonpositively curved, any two points of $\tilde{S}$ are connected by a unique $\tilde{\varphi}$ -geodesic. In the complement of the cone points, geodesics are Euclidean geodesics (straight lines, rays, or segments), and when a geodesic meets a cone point, it makes angle at least π on both sides; see Figure 2.

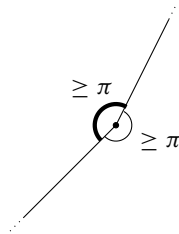


Figure 2. Local picture of a geodesic through a cone point.

Geodesics between a pair of cone points which meet no other cone points are called *saddle connections*. A geodesic segment, ray, or line containing no cone points is called *nonsingular*, and is called *singular* otherwise. We write $\mathcal{G}^*(\tilde{\varphi})$ for the closure in $\mathcal{G}(\tilde{\varphi})$ of the set of nonsingular geodesics, and we will call these the *basic geodesics* in the metric $\tilde{\varphi}$. The following is proved in [1, Section 2.4].

Proposition 3 (Structure of basic geodesics). *The geodesics in $\mathcal{G}^*(\tilde{\varphi})$ are precisely the following $\tilde{\varphi}$ -geodesics: (i) the nonsingular ones; (ii) the ones that meet a single cone point and make cone angle exactly π on one side; and (iii) those that meet two or more cone points in such a way that they always make cone angle π on one side and that side switches from right to left or from left to right at most once along their entire length. Furthermore, there are only countably many of type (iii), and we will denote that set by $\mathcal{G}^2(\tilde{\varphi})$.*

Suppose $\tilde{\gamma}_0, \tilde{\gamma}_1 \in \mathcal{G}^*(\tilde{\varphi})$ are asymptotic in one direction, meaning that the Hausdorff distance in one direction is finite. Then either we can pass to subrays $\tilde{\gamma}_0^+ = \tilde{\gamma}_1^+$ that coincide, or else there is an isometric embedding $[0, \infty) \times [0, W] \rightarrow \tilde{S}$ for some $W > 0$ such that $[0, \infty) \times \{0\}$ maps to $\tilde{\gamma}_0^+$ and $[0, \infty) \times \{W\}$ maps to $\tilde{\gamma}_1^+$;

see [7, Theorem II.2.13]. In the former case, there is a maximal such common subray, necessarily emanating from a cone point $\tilde{\zeta} \in \Sigma(\tilde{\varphi})$, and we say that $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ are *cone-point asymptotic*. In the latter case, the embedding of $[0, \infty) \times [0, W]$ is called a *flat half-strip*.

A *flat strip* is an isometric embedding $\mathbb{R} \times [0, W] \rightarrow \tilde{S}$, for some $W > 0$, or the composition of such an isometric embedding with the projection down to S . Flat strips in S naturally arise from Euclidean *cylinders* in S , which are locally isometric immersions $S_r^1 \times [0, W] \rightarrow S$ where $W > 0$ and S_r^1 is a circle of length $r > 0$. The next result of Hassell–Hillairet–Marzuola [21] (generalizing work of Galperin–Krüger–Troubetzkoy [19] for billiards) says that all flat strips arise in this way. The proofs in [19, 21] are easily modified to handle flat half-strips as well.

Theorem 4 (Flat Strip theorem [19, 21]). *With respect to the metric $\varphi \in \text{Flat}(S)$, any flat strip or flat half-strip in S is contained in a cylinder.*

A $\tilde{\varphi}$ -Euclidean triangle is a $\tilde{\varphi}$ -geodesic triangle in $\tilde{S}$ with vertices at the cone points which is isometric to a geodesic triangle in $\mathbb{R}^2$; equivalently, a $\tilde{\varphi}$ -Euclidean triangle is a triangle whose sides are $\tilde{\varphi}$ -saddle connections and which bounds a 2-simplex with no cone points in its interior. (Note: we may refer to either the 2-simplex or its boundary as a triangle.) A φ -Euclidean triangle T is the image in S of a $\tilde{\varphi}$ -Euclidean triangle $\tilde{T}$ in $\tilde{S}$ for which the restriction of $p: \tilde{S} \rightarrow S$ to $\tilde{T}$ is injective on the interior. In this case, there is a map from T to a Euclidean triangle in $\mathbb{R}^2$ (though it may not be defined at the vertices) which is an isometry for the induced path metric on T .

A φ -triangulation of S is a triangulation (more precisely, a Δ -complex structure in the sense of [22]) such that every triangle is a φ -Euclidean triangle. A φ -triangulation lifts to a $\tilde{\varphi}$ -triangulation of $\tilde{S}$. It was shown in [33] that φ -triangulations exist for quadratic differential metrics, and we note that the same proof is valid for arbitrary flat metrics.

For a flat metric φ , any homotopy class of a curve γ on S has a φ -geodesic representative; such a representative can be obtained as a uniform limit of constant-speed parametrizations of homotopic curves whose lengths limit to the infimum over all representatives. The geodesic representative is either unique (up to parametrization) or else it is part of a Euclidean *cylinder*, $S_r^1 \times [0, W] \rightarrow S$, as above. In this case the set of geodesic representatives of γ is precisely the set of *core curves* (i.e., the images of $S_r^1 \times \{t\}$ for $t \in [0, W]$). If a curve γ is homotopic to a core curve of a cylinder, then we say that γ is a cylinder curve. The set of all homotopy classes of cylinder curves for φ will be denoted $\text{cyl}(\varphi)$. In general, *closed curves* will refer to closed essential curves on S up to homotopy.

Suppose φ is a flat metric on S . Since φ has zero curvature on $S \setminus \Sigma(\varphi)$, we have a well-defined *holonomy homomorphism* $P_\varphi: \pi_1(S \setminus \Sigma(\varphi), \zeta) \rightarrow SO(2)$ given by parallel transport around loops based at a point $\zeta \in S \setminus \Sigma(\varphi)$. We will often refer to the image of P_φ as *the holonomy of φ* .

2.4. The space $\text{Flat}(S)$ and affine equivalence. We say that two flat metrics φ_1, φ_2 are equivalent if there exists an isometry $f: (S, \varphi_1) \rightarrow (S, \varphi_2)$ isotopic to the identity on S . We then define $\text{Flat}(S)$ to be the set of equivalence classes of unit-area flat metrics. We will use the notation φ for either a particular flat metric or its equivalence class.

We will also need the following coarser equivalence relation on flat metrics. We say that two (unit area) flat metrics φ_1, φ_2 are *affine-equivalent* if there is a map $f: (S, \varphi_1) \rightarrow (S, \varphi_2)$, isotopic to the identity, which is real-affine in isometric local coordinates: for all $\zeta \in S \setminus \Sigma(\varphi_1)$ and orientation preserving, isometric local coordinates $\xi_1: U_1 \rightarrow \mathbb{R}^2$ and $\xi_2: U_2 \rightarrow \mathbb{R}^2$ about ζ and $f(\zeta)$, respectively, with $f(U_1) = U_2$, we have $\xi_2 \circ f|_{U_1} = A \circ \xi_1 + b$, for some $A \in \text{SL}(2, \mathbb{R})$ and $b \in \mathbb{R}^2$. If f is not an isometry, then at each nonsingular point there is a direction of maximal stretch. Using the local coordinates we see that the direction of maximal stretch defines a parallel line field on $S \setminus \Sigma(\varphi_1)$. It follows from the existence of such a parallel line field that the holonomy of φ_1 is contained in $\{\pm I\} < \text{SO}(2)$. Consequently, φ_1 and φ_2 are defined by quadratic differentials differing by an element of $\text{SL}(2, \mathbb{R})$ from the usual action of $\text{SL}(2, \mathbb{R})$ on the space of quadratic differentials; see [34], for example.

Proposition 5. *If φ_1, φ_2 are two affine-equivalent flat metrics which are not equivalent, then φ_1, φ_2 are defined by quadratic differentials in the same $\text{SL}(2, \mathbb{R})$ -orbit.* $\square$

In [17], $\text{Flat}(S)$ denoted only the space of flat metrics coming from quadratic differentials. In the current paper (as well as [1]), the flat metrics coming from quadratic differentials, which we denote $\text{Flat}_2(S) \subset \text{Flat}(S)$ is a (fairly small) subspace (characterized, for example, by Proposition 5 as those admitting affine deformations). The following theorem, proved in [17, Lemma 19], gives a useful way of deciding when quadratic differential metrics are affine-equivalent.

Theorem 6. *If φ_1 and φ_2 are quadratic differential metrics on any surface of finite type, then they are affine-equivalent if and only if their cylinder curves agree.*

Remark. Note that this theorem holds for finite-type surfaces, which allows a finite number of punctures. We will apply it to a closed surface minus the cone points of a flat metric in the proof of the Support Rigidity theorem in Section 3.5.

When φ_1 has holonomy with order at least 3, then by appealing to Proposition 5, affine equivalence implies the stronger conclusion that $\varphi_1 = \varphi_2$ in $\text{Flat}(S)$.

2.5. Currents and chains for flat metrics. The fact that $\text{SL}(2, \mathbb{R})$ orbits in $\text{Flat}_2(S)$ are determined by cylinder sets (Theorem 6 above) was used in [17] as a step in proving that marked simple closed curves are spectrally rigid over $\text{Flat}_2(S)$; that is, the lengths of those curves entirely determine a quadratic differential metric. This rigidity result was generalized in [1] to all of $\text{Flat}(S)$, although in this setting all closed curves are needed, not just simple ones. In both [17] and [1], a key object is

the Liouville current for a flat metric $\varphi \in \text{Flat}(S)$. In this section we recall some of the key properties of the current, particularly relating to its support.

In [17], the Liouville current L_φ associated to a flat metric φ is defined as a kind of average of the measured foliations in all directions and which satisfies the geometricity property $\iota(\alpha, L_\varphi) = \ell_\varphi(\alpha)$ for all closed curves α . For general $\varphi \in \text{Flat}(S)$, the Liouville current L_φ is defined in [1] as the push-forward via the “endpoint map”

$$\partial_{\tilde{\varphi}}: \mathcal{G}(\tilde{\varphi}) \rightarrow \mathcal{G}(\tilde{S})$$

of a $\pi_1 S$ -invariant measure on $\mathcal{G}(\tilde{\varphi})$ obtained from Riemannian geometry. It is also shown there to have the same geometricity property as in [17]. One of the key facts we will need here, proved in [1, Proposition 3.4 and Corollary 3.5], involves the support of L_φ . Recall from Section 2.3 that the set of basic geodesics $\mathcal{G}^*(\tilde{\varphi}) \subset \mathcal{G}(\tilde{\varphi})$ is the closure of the set of nonsingular geodesics, and that they are completely described in Proposition 3. We will write

$$\mathcal{G}_{\tilde{\varphi}}^* = \partial_{\tilde{\varphi}}(\mathcal{G}^*(\tilde{\varphi})) \subset \mathcal{G}(\tilde{S})$$

for the set of endpoints on S_∞^1 of the basic geodesics.

Proposition 7. *For any $\varphi \in \text{Flat}(S)$, the support of its Liouville current is precisely given by endpoints of basic geodesics:*

$$\text{supp}(L_\varphi) = \mathcal{G}_{\tilde{\varphi}}^* \subset \mathcal{G}(\tilde{S}).$$

Next we introduce a key tool from [1] used to study Liouville currents for flat metrics, the set of *chains*, a technical device to encode information about cone points in terms of boundary points. The definition is somewhat technical, and involves an auxiliary countable subset $\Omega \subset \mathcal{G}_{\tilde{\varphi}}^*$.

Definition 8. Fix a flat metric φ and countable subset $\Omega \subset \mathcal{G}_{\tilde{\varphi}}^*$. Then a $(G_{\tilde{\varphi}}^*, \Omega)$ -chain is a bi-infinite sequence of boundary points $\mathbf{x} = (\dots, x_0, x_1, \dots) \subset S_\infty^1$ such that

- (i) $\{x_i, x_{i+1}\} \in \mathcal{G}_{\tilde{\varphi}}^* \setminus \Omega$,
 - (ii) x_i, x_{i+1}, x_{i+2} is a counterclockwise-ordered triple of distinct points, and
 - (iii) $[\{x_i, x_{i+1}\}, \{x_{i+1}, x_{i+2}\}] \cap \mathcal{G}_{\tilde{\varphi}}^* = \{\{x_i, x_{i+1}\}, \{x_{i+1}, x_{i+2}\}\}$
- for all i . Let the set of $(\mathcal{G}_{\tilde{\varphi}}^*, \Omega)$ -chains be denoted $\text{Chain}(\mathcal{G}_{\tilde{\varphi}}^*, \Omega)$.

Remark. This definition is a more concise version of the one given in [1] which involved first defining chains using $\mathcal{G}_{\tilde{\varphi}}^*$ and then introducing the countable set later, and additionally allowed for finite or half-infinite chains. Furthermore, the notation differs from [1], where these were called (L_φ, Ω) -chains. Since the definition depends only on $\text{supp}(L_\varphi) = \mathcal{G}_{\tilde{\varphi}}^*$, this notation is more descriptive and will be useful below. We note that chains may be periodic, and hence may contain only a finite number of boundary points, repeated infinitely often.

A chain arises naturally from a cone point $\zeta \in \Sigma(\tilde{\varphi})$ via families of basic geodesics meeting only ζ and winding around in a manner such that two successive ones are cone-point asymptotic. In [1, Proposition 4.1] it was shown that essentially all chains arise in this way.

The set Ω is used to discard geodesics meeting more than one cone point. More precisely, recall that the set $\mathcal{G}^2(\tilde{\varphi}) \subset \mathcal{G}^*(\tilde{\varphi})$ of basic geodesics meeting at least two cone points is countable, by Proposition 3 (Structure of basic geodesics). We let

$$\mathcal{G}_{\tilde{\varphi}}^2 = \partial_{\tilde{\varphi}}(\mathcal{G}^2(\tilde{\varphi})) \subset \mathcal{G}_{\tilde{\varphi}}^*$$

denote the countable set of endpoints of such geodesics, and then require our countable set Ω to contain $\mathcal{G}_{\tilde{\varphi}}^2$.

Proposition 9. *Suppose $\Omega \subset \mathcal{G}_{\tilde{\varphi}}^*$ is any countable set containing $\mathcal{G}_{\tilde{\varphi}}^2$ and $\mathbf{x} \in \text{Chain}(\mathcal{G}_{\tilde{\varphi}}^*, \Omega)$ is any chain. Then there exists a unique cone point ζ and sequence $\boldsymbol{\gamma} = (\dots, \tilde{\gamma}_0, \tilde{\gamma}_1, \dots) \subset \mathcal{G}^*(\tilde{\varphi})$ of geodesics meeting ζ , and no other cone points, such that $\partial_{\tilde{\varphi}}(\boldsymbol{\gamma}) = \mathbf{x}$, i.e., $\partial_{\tilde{\varphi}}(\tilde{\gamma}_i) = \{x_i, x_{i+1}\}$ for each i . Moreover, every cone point is related in this way to some chain.*

From this we obtain a well-defined, surjective map from chains to the set of cone points $\Sigma(\tilde{\varphi})$:

$$\partial_{\tilde{\varphi}}^{\#}: \text{Chain}(\mathcal{G}_{\tilde{\varphi}}^*, \Omega) \rightarrow \Sigma(\tilde{\varphi}); \quad (2.2)$$

see [1, Lemma 4.2]. Because $\partial_{\tilde{\varphi}}$ is π_1 -equivariant, so is $\partial_{\tilde{\varphi}}^{\#}$.

The upshot of this is that when we have two metrics $\varphi_1, \varphi_2 \in \text{Flat}(S)$ whose Liouville currents have the same support $\mathcal{G}_{\tilde{\varphi}_1}^* = \mathcal{G}_{\tilde{\varphi}_2}^*$, we can take Ω to be a countable subset containing $\mathcal{G}_{\tilde{\varphi}_1}^2 \cup \mathcal{G}_{\tilde{\varphi}_2}^2$ so that the resulting chains are the same, producing an identification of the cone points of the two metrics. The following is essentially a consequence of [1, Lemma 4.4] and the proof of [1, Theorem 5.1].

Proposition 10. *Suppose that $\varphi_1, \varphi_2 \in \text{Flat}(S)$ are two metrics whose Liouville currents have the same support, which we denote*

$$\mathcal{G}^* = \mathcal{G}_{\tilde{\varphi}_1}^* = \mathcal{G}_{\tilde{\varphi}_2}^*.$$

Then for any countable set $\Omega \subset \mathcal{G}^$ containing $\mathcal{G}_{\tilde{\varphi}_1}^2 \cup \mathcal{G}_{\tilde{\varphi}_2}^2$, we have*

$$\text{Chain}(\mathcal{G}^*, \Omega) = \text{Chain}(\mathcal{G}_{\tilde{\varphi}_1}^*, \Omega) = \text{Chain}(\mathcal{G}_{\tilde{\varphi}_2}^*, \Omega),$$

and there exists a $\pi_1 S$ -equivariant bijection

$$\tilde{f}^0: \Sigma(\tilde{\varphi}_1) \rightarrow \Sigma(\tilde{\varphi}_2)$$

so that $\tilde{f}^0 \circ \partial_{\tilde{\varphi}_1}^{\#} = \partial_{\tilde{\varphi}_2}^{\#}$.

Later we will see that by enlarging Ω further, we can also make the nonsingular geodesics match up for a pair of metrics φ_1, φ_2 as we did here for chains and cone points.

2.6. Identifying geodesics. For two flat metrics whose Liouville currents have the same support, we combine facts we know about basic geodesics and flat strips to provide a geometrically useful identification of “most of” the basic geodesics for the two metrics. Recall that having the same support precisely means that $\mathcal{G}_{\tilde{\varphi}_1}^* = \mathcal{G}_{\tilde{\varphi}_2}^*$, or in other words the basic geodesics have the same endpoints. Below, we will simply write

$$\mathcal{G}^* = \mathcal{G}_{\tilde{\varphi}_1}^* = \mathcal{G}_{\tilde{\varphi}_2}^*,$$

for the mutual support of the pair of Liouville currents.

Metrics in $\text{Flat}(S)$ are only defined up to isotopy. In the previous Section 2.5 where chains are defined we recorded the fact that for any two metrics $\varphi_1, \varphi_2 \in \text{Flat}(S)$ whose Liouville currents have the same support, we can take any countable set Ω containing the boundary pairs of “multiple-cone-point geodesics”

$$\mathcal{G}_{\varphi_1}^2 \cup \mathcal{G}_{\varphi_2}^2 \subset \mathcal{G}^*$$

in either metric, and using the chains, produce a $\pi_1 S$ -equivariant bijection between the cone points of the two metrics; see Proposition 10. From arbitrary representatives of the metrics, we can therefore perform “point-pushing” isotopies on the cone points so that the two metrics have exactly the same set of cone points $\Sigma(\tilde{\varphi}_1) = \Sigma(\tilde{\varphi}_2)$ in a way that is compatible with the chains (i.e., so that $\partial_{\tilde{\varphi}_1}^\# = \partial_{\tilde{\varphi}_2}^\#$).

Second, for the same Ω , we can observe that every pair $\{x, y\} \in \mathcal{G}^* \setminus \Omega$ has exactly one geodesic in its preimage by both $\partial_{\tilde{\varphi}_1}$ and $\partial_{\tilde{\varphi}_2}$,

$$|\partial_{\tilde{\varphi}_1}^{-1}(\{x, y\})| = 1 = |\partial_{\tilde{\varphi}_2}^{-1}(\{x, y\})|.$$

To see this, note that the only way $\partial_{\tilde{\varphi}_i}^{-1}(\{x, y\})$ can contain more than one geodesic is if it consists of a (closed) flat strip of parallel geodesics; see [7, Theorem II.2.13]. By the Flat Strip theorem 4, such a strip covers a maximal cylinder. Since the boundary of such a cylinder is a concatenation of saddle connections in S , the boundary of the strip itself is a concatenation of saddle connections in $\tilde{S}$, and hence $\{x, y\} \in \mathcal{G}_{\tilde{\varphi}_i}^2$, which is a contradiction.

We can then enlarge the countable set Ω so that for any $\{x, y\} \in \mathcal{G}^* \setminus \Omega$, either x, y can be ordered to become part of a bi-infinite (possibly periodic) chain, or else it is the image of nonsingular geodesics in both $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$. This is possible because, for each $i = 1, 2$ and for each of the countably many cone points $\zeta \in \Sigma(\tilde{\varphi}_i)$, there are only countably many nonsingular rays emanating from ζ whose endpoints on S_∞^1 fail to be extendable to a bi-infinite chain. Thus, there are countably many geodesics $\tilde{\gamma} \in \mathcal{G}_{\tilde{\varphi}_i}^*$ containing a single cone point whose endpoints are not part of a bi-infinite chain. So by taking Ω to be the union of the $\partial_{\tilde{\varphi}_i}$ -images of these, together with $\mathcal{G}_{\varphi_1}^2 \cup \mathcal{G}_{\varphi_2}^2$, we obtain the desired property.

Therefore, setting $G_i = \partial_{\tilde{\varphi}_i}^{-1}(\mathcal{G}^* \setminus \Omega)$, for $i = 1, 2$, we can define $g: G_1 \rightarrow G_2$ to be the bijection

$$g = \partial_{\tilde{\varphi}_2}^{-1} \circ \partial_{\tilde{\varphi}_1} |_{\mathcal{G}^* \setminus \Omega}. \quad (2.3)$$

That is, g is the $\tilde{\varphi}_2$ -straightening of each $\tilde{\varphi}_1$ -geodesic in G_1 , and it sends the set of singular geodesics to the set of singular geodesics. Furthermore, observe that any singular geodesic $\tilde{\gamma} \in G_i$ containing a cone point ζ has endpoints in a chain whose $\partial_{\tilde{\varphi}_i}^\#$ -image is in a chain determining ζ . It follows that ζ is the unique cone point in a singular geodesic $\tilde{\gamma} \in G_1$ if and only if it is the unique cone point in $g(\tilde{\gamma}) \in G_2$.

Proposition 11. *For any pair of flat metrics with the same support $\mathcal{G}^*$, after discarding a suitable countable set $\Omega \subset \mathcal{G}^*$ the bijection $g: G_1 \rightarrow G_2$ between the basic geodesics in the two metrics is a homeomorphism. Furthermore, $g(\tilde{\gamma})$ is singular if and only if $\tilde{\gamma}$ is. In the singular case, the unique cone point of $\tilde{\gamma}$ is also contained in $g(\tilde{\gamma})$.*

Proof. Suppose $\{\tilde{\gamma}_n\}_{n=1}^\infty \subset G_1$ is a sequence such that $\tilde{\gamma}_n \rightarrow \tilde{\gamma} \in G_1$. To prove $g(\tilde{\gamma}_n) \rightarrow g(\tilde{\gamma})$, it suffices to prove that every subsequence of $\{g(\tilde{\gamma}_n)\}$ has a subsequence that converges to $g(\tilde{\gamma})$. For this, observe that $\{\partial_{\tilde{\varphi}_1}(\tilde{\gamma}_n)\}_{n=1}^\infty \cup \{\partial_{\tilde{\varphi}_1}(\tilde{\gamma})\}$ is compact subset of $\mathcal{G}(\tilde{S})$. Therefore, $\{g(\tilde{\gamma}_n)\}_{n=1}^\infty \cup \{g(\tilde{\gamma})\}$ is precompact in $\mathcal{G}(\tilde{\varphi}_2)$, and since $\mathcal{G}^*(\tilde{\varphi}_2)$ is a closed subspace containing G_2 , our subset is precompact in $\mathcal{G}^*(\tilde{\varphi}_2)$. For any subsequence of $\{g(\tilde{\gamma}_n)\}$ we may choose a convergent subsequence $\{g(\tilde{\gamma}_{n_k})\}_{k=1}^\infty$ with $g(\tilde{\gamma}_{n_k}) \rightarrow \tilde{\gamma}' \in \mathcal{G}^*(\tilde{\varphi}_2)$. Then, by definition of g (and the fact that limits in the Hausdorff space $\mathcal{G}(\tilde{S})$ are unique) we have

$$\partial_{\tilde{\varphi}_2}(\tilde{\gamma}') = \lim_{k \rightarrow \infty} \partial_{\tilde{\varphi}_2}(g(\tilde{\gamma}_{n_k})) = \lim_{k \rightarrow \infty} \partial_{\tilde{\varphi}_1}(\tilde{\gamma}_{n_k}) = \partial_{\tilde{\varphi}_1}(\tilde{\gamma}).$$

Thus, $\tilde{\gamma}' = g(\tilde{\gamma})$ since $\partial_{\tilde{\varphi}_2}$ is one-to-one on G_2 . Therefore, g is continuous. Reversing the roles of G_1 and G_2 , we see that g^{-1} is also continuous, and hence g is a homeomorphism. $\square$

We now adopt the machinery needed to define this homeomorphism as a convention for the remainder of the paper. For two flat metrics φ_1, φ_2 with equal support, we will be able to unambiguously write $\mathcal{G}^*$ for their mutual support as pairs of endpoints; $\Omega \subset \mathcal{G}^*$ for a countable set as described above; $G_i = \partial_{\tilde{\varphi}_i}(\mathcal{G}^* \setminus \Omega)$ for the remaining basic geodesics; $\tilde{\Sigma} = \Sigma(\tilde{\varphi}_1) = \Sigma(\tilde{\varphi}_2)$ for the cone points; and $g: G_1 \rightarrow G_2$ for the homeomorphism identifying the basic geodesics.

3. Support Rigidity theorem

In this section we will prove our main theorem about Liouville currents of flat metrics.

Support Rigidity theorem. *Suppose φ_1, φ_2 are two unit-area flat metrics whose Liouville currents have the same support, $\text{supp}(L_{\varphi_1}) = \text{supp}(L_{\varphi_2})$. Then φ_1, φ_2 differ by an affine deformation, up to isotopy. If either metric has holonomy of order greater than 2 (i.e., is not induced by a quadratic differential, or does not support a foliation by straight lines), then equal support implies that φ_1 and φ_2 differ by isometry, isotopic to the identity.*

For this entire section, we assume that φ_1 and φ_2 are flat metrics with equal support, and we follow the conventions specified above to identify their cone points and their basic geodesics.

3.1. Cone point partitions. For each $i = 1, 2$, any geodesic $\tilde{\gamma} \in G_i$ divides $\tilde{S}$ into two *half-planes*, $\mathcal{H}_i^\pm(\tilde{\gamma})$, which are the closures of the connected components of the complement of $\tilde{\gamma}$. When $\tilde{\gamma}$ is a nonsingular geodesic, each cone point lies in exactly one of these two half-planes, and thus $\tilde{\gamma}$ determines a partition of $\tilde{S}$ into two disjoint subsets depending on which side of $\tilde{\gamma}$ the point lies. When $\tilde{\gamma}$ is singular, it contains exactly one cone point, and makes cone angle π on one side. We then declare the cone point to lie on the side opposite the one in which it makes an angle π , and so $\tilde{\gamma}$ also determines a partition of $\tilde{S}$ into two disjoint subsets.

Lemma 12 (Cone point partitions are well-defined). *For any $\tilde{\gamma} \in G_1$, the geodesic $\tilde{\gamma}$ defines the same partition as its image $g(\tilde{\gamma})$ under the straightening homeomorphism.*

The straightening homeomorphism is defined above in (2.3).

Proof. First, suppose that $\tilde{\gamma} \in G_1$ is nonsingular, and fix a $\zeta \in \tilde{S}$. After orienting $\tilde{\gamma}$, let us suppose $\mathcal{H}_1^+(\tilde{\gamma})$ is the right half-plane and $\mathcal{H}_1^-(\tilde{\gamma})$ is the left half-plane. The orientation of $\tilde{\gamma}$ induces an orientation of $g(\tilde{\gamma})$ and hence a choice of right and left half-planes $\mathcal{H}_2^+(g(\tilde{\gamma}))$ and $\mathcal{H}_2^-(g(\tilde{\gamma}))$. The endpoints $\partial_{\tilde{\varphi}_1}(\tilde{\gamma}) = \partial_{\tilde{\varphi}_2}(g(\tilde{\gamma}))$ divide S_∞^1 into two components, $S_\infty^{1\pm}(\tilde{\gamma})$, which are the boundaries at infinity of $\mathcal{H}_1^\pm(\tilde{\gamma})$, as well as that of $\mathcal{H}_2^\pm(g(\tilde{\gamma}))$.

We will show $\zeta \in \mathcal{H}_1^+(\tilde{\gamma}) \implies \zeta \in \mathcal{H}_2^+(g(\tilde{\gamma}))$. To prove this, observe that there is a singular geodesic $\tilde{\delta} \in G_1$ containing ζ and contained in the interior of $\mathcal{H}_1^+(\tilde{\gamma})$. Such a geodesic can be obtained by first parallel translating the direction of $\tilde{\gamma}$ to ζ along a geodesic segment connecting a point of $\tilde{\gamma}$ to ζ (making appropriate choices at any cone points encountered along this geodesic segment), then perturbing slightly to avoid the countably many geodesics through ζ not in G_1 . Then $\zeta \in g(\tilde{\delta})$. On the other hand, $\tilde{\gamma}$ and $g(\tilde{\gamma})$ have the same endpoints $\partial_{\tilde{\varphi}_1}(\tilde{\delta}) = \partial_{\tilde{\varphi}_2}(g(\tilde{\delta}))$, and these lie in $S_\infty^{1+}(\tilde{\gamma})$. If $\zeta \in \mathcal{H}_2^-(g(\tilde{\gamma}))$, then $g(\tilde{\delta})$ would have to cross $g(\tilde{\gamma})$ twice, creating a bigon; see Figure 3. This contradiction shows that $\zeta \in \mathcal{H}_2^+(g(\tilde{\gamma}))$, as required.

Next, observe that if $\zeta \in \tilde{S}$ is on the left-hand side of a sequence of $\tilde{\varphi}_i$ -geodesics $\{\tilde{\gamma}_n\} \subset G_i$ (with respect to a choice of orientations) and $\tilde{\gamma}_n \rightarrow \tilde{\gamma} \in G_i$ (limiting as oriented geodesics), then ζ is also on the left-hand side of $\tilde{\gamma}$, for $i = 1, 2$. Now suppose that $\zeta \in \tilde{S}$ is any cone point and $\tilde{\gamma} \in G_1$ is a geodesic for which ζ is on the left-hand side. Choose a sequence of nonsingular geodesics $\tilde{\gamma}_n \in G_1$ limiting to $\tilde{\gamma}$, such that ζ is on the left-hand side of $\tilde{\gamma}_n$ for all n : this is easy to do if ζ is not on $\tilde{\gamma}$ (then any sequence will have a tail that has this property). If, on the other hand, $\zeta \in \tilde{\gamma}$, then it makes angle π on the right, and any sequence $\{\tilde{\gamma}_n\} \subset G_1$ of nonsingular geodesics limiting to $\tilde{\gamma}$ from the right near ζ will have ζ on the left-hand side (to find such a sequence, take any sequence of points approaching ζ in the interior of $\mathcal{H}_1^+(\tilde{\gamma})$ whose

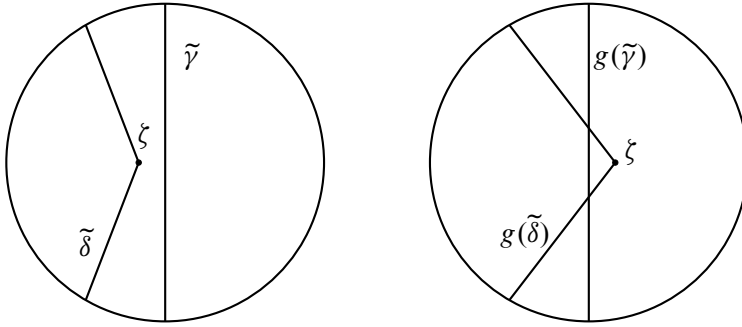


Figure 3. Switching sides creates a bigon.

tangent vectors limit to the direction of $\tilde{\gamma}$ and so that there are nonsingular geodesics through those points which lie in G_1). Now since g is a homeomorphism, it follows that $g(\tilde{\gamma}_n) \rightarrow g(\tilde{\gamma})$. Since $\tilde{\gamma}_n$ are all nonsingular, ζ is on the left-hand side of $g(\tilde{\gamma}_n)$ by the first part of the proof, and hence also on the left-hand side of $g(\tilde{\gamma})$. $\square$

3.2. Identifying triangulations. Recall that given $\varphi \in \text{Flat}(S)$, a $\tilde{\varphi}$ -Euclidean triangle is a triangle with cone point vertices, $\tilde{\varphi}$ -saddle connection sides, and no cone points in its interior.

Lemma 13. *Every pair $x, y \in \tilde{\Sigma}$ that determines a saddle connection for $\tilde{\varphi}_1$ also determines a saddle connection for $\tilde{\varphi}_2$; this defines a bijection $g_{sc}: G_{1,sc} \rightarrow G_{2,sc}$ between all $\tilde{\varphi}_1$ -saddle connections and all $\tilde{\varphi}_2$ -saddle connections. Furthermore, $x, y, z \in \tilde{\Sigma}$ are vertices of a positively oriented $\tilde{\varphi}_1$ -Euclidean triangle if and only if they are vertices of a positively oriented $\tilde{\varphi}_2$ -Euclidean triangle.*

Proof. To prove the first statement, suppose x, y determines a saddle connection δ for $\tilde{\varphi}_1$. This means that there is a $\tilde{\varphi}_1$ -geodesic segment connecting x and y that does not pass through any other cone points. Now consider the unique $\tilde{\varphi}_2$ -geodesic path between x and y . Suppose there is a cone point z on this path. We claim that it is possible to separate z from x and y by a $\tilde{\varphi}_1$ -geodesic, $\tilde{\gamma} \in G_1$ as in Lemma 12. Note that this is impossible in the $\tilde{\varphi}_2$ metric, since this would create a geodesic bigon; see Figure 4.

To create the desired separating geodesic, consider the $\tilde{\varphi}_1$ -geodesic segment from z to x and the $\tilde{\varphi}_1$ -geodesic segment from z to y . There must be sub-segments of each of these that do not intersect δ . This is because in the $\tilde{\varphi}_1$ metric, z is not on the saddle connection δ . We denote these sub-segments as α and β in Figure 4, below. Any geodesic not passing through x, y or z and transversally intersecting both α and β will separate z from x and y . It is possible to find the required geodesic $\tilde{\gamma} \in G_1$ by perturbing an arbitrary such $\tilde{\varphi}_1$ -geodesic to one in G_1 . This completes the proof of the first statement.

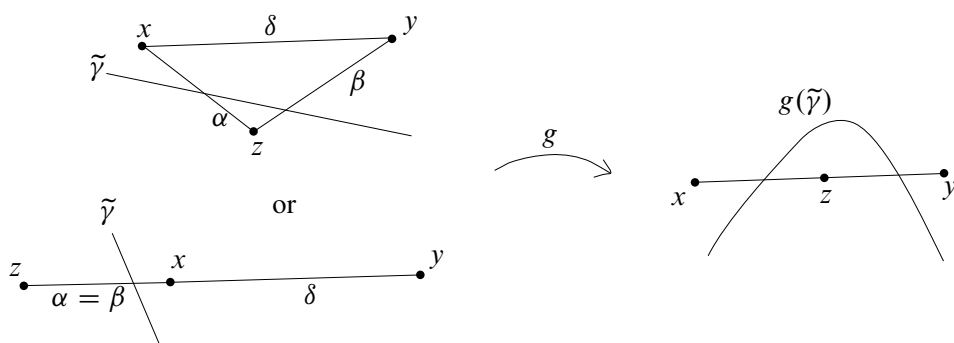


Figure 4. Failure to preserve saddle connections creates a bigon.

To prove the second statement, suppose that x, y, z are the vertices of a $\tilde{\varphi}_1$ -Euclidean triangle $\tilde{T}$. From the above, there are $\tilde{\varphi}_2$ -saddle connections forming a complete graph on x, y, z . Further, these saddle connections bound a connected component, which we call $\tilde{T}'$. We need to show that $\tilde{T}'$ does not contain any cone points, hence x, y, z form a $\tilde{\varphi}_2$ -Euclidean triangle, and that this triangle $\tilde{T}'$ has the same orientation as $\tilde{T}$.

Let α be the saddle connection between x and z . There is a sequence of nonsingular geodesics $\{\tilde{\alpha}_n\}$ in G_1 , intersecting $\tilde{T}'$ which limits to α in $\tilde{T}$. Let β and δ be the saddle connections between y and z and between x and z , respectively. Define similarly sequences $\{\tilde{\beta}_n\}$ and $\{\tilde{\delta}_n\}$. For large enough N , $\tilde{\alpha}_N$ defines a partition that separates z from x and $\tilde{\delta}_N$ a partition that separates z from y . Consider the intersection of corresponding two half-planes containing z . The boundary of this region determines an interval I_x on S_∞^1 . Similarly $\tilde{\alpha}_N, \tilde{\beta}_N, z$ and $\tilde{\beta}_N, \tilde{\delta}_N, y$ determine intervals I_y and I_z on S_1 , respectively; see the left-hand side of Figure 5. Give the triangle bounded by $\tilde{\alpha}_N, \tilde{\beta}_N$, and γ_N the same orientation as $\tilde{T}'$. Note that the orientation of the intervals I_x, I_y, I_z along S_∞^1 determine the orientation of the triangle (although we note that the intersection pattern of the geodesics in the $\tilde{\varphi}_2$ -metric is not *a priori* the same; see the right-hand side of Figure 5). Moreover, the 6 intervals determined by I_x, I_y, I_z and their complements partition S_∞^1 in such a way that the $\tilde{\varphi}_1$ -geodesics obtained by straightening $\tilde{\alpha}_N$ and $\tilde{\beta}_N$ separate vertex x from the other vertices of $\tilde{T}$. Similarly the straightening of $\tilde{\alpha}_N$ and $\tilde{\delta}_N$ separate vertex z from y and x and the straightening of $\tilde{\beta}_N$ and $\tilde{\delta}_N$ separate vertex y from z and x . This results in a triangle whose orientation must agree with both $\tilde{T}'$ and $\tilde{T}$ and hence the orientation of $\tilde{T}$ and $\tilde{T}'$ must agree.

Finally, if w is a cone point inside $\tilde{T}'$, we similarly approximate the side of $\tilde{T}'$ by geodesics $\tilde{\alpha}_N, \tilde{\beta}_N, \tilde{\delta}_N \in G_2$ which together separate this cone point from the cone points x, y, z in the $\tilde{\varphi}_2$ -metric, and hence, since $g^{-1}: G_2 \rightarrow G_1$ is a homeomorphism, induces such a separation in the $\tilde{\varphi}_1$ -metric as well, contradicting that $\tilde{T}$ is a $\tilde{\varphi}$ -Euclidean triangle. $\square$

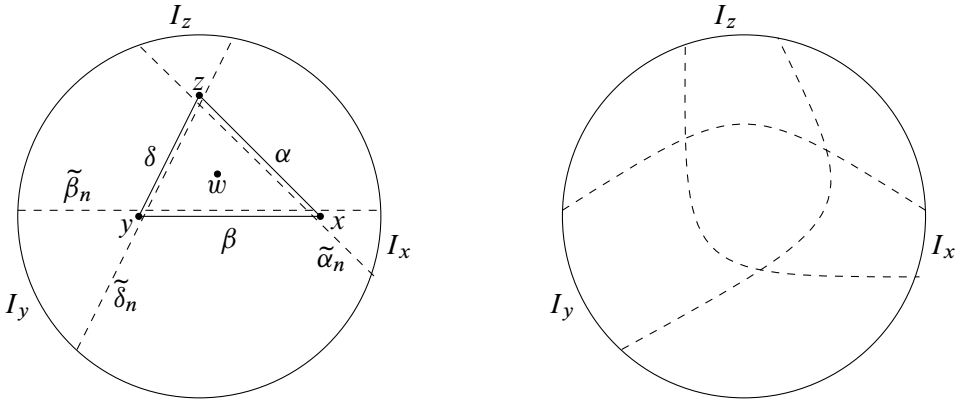


Figure 5. Left-hand side: Approximating the sides α, β, γ of $\tilde{T}'$ with nonsingular geodesics and the resulting intervals I_x, I_y, I_z on S_1 . Right-hand side: Potential change of intersection pattern.

In fact, g_{sc} also defines a bijection between all $\tilde{\varphi}_1$ -geodesics that are concatenation of saddle connections and the set of all such $\tilde{\varphi}_2$ -geodesic concatenations:

Proposition 14. *Consider a $\tilde{\varphi}_1$ -geodesic $\tilde{\gamma}_1 = \cdots \delta_0 \delta_1 \cdots$ which is a finite, infinite, or bi-infinite concatenation of saddle connections. Then $\tilde{\gamma}_2 = \cdots g_{sc}(\delta_1) g_{sc}(\delta_2) \cdots$ is a $\tilde{\varphi}_2$ -geodesic. We write $\tilde{\gamma}_2 = g(\tilde{\gamma}_1)$.*

Proof. Recall that a concatenation of saddle connections is a geodesic segment if and only if the angle between any two consecutive saddle connections measure at least π on either side. Let

$$\tilde{\gamma}_1 = \cdots \delta_0 \delta_1 \cdots$$

be a $\tilde{\varphi}_1$ -geodesic segment which is a concatenation of saddle connections. Then by Lemma 13,

$$\tilde{\gamma}_2 = \cdots g_{sc}(\delta_1) g_{sc}(\delta_2) \cdots$$

is a concatenation of $\tilde{\varphi}_2$ -saddle connections. Suppose $g_{sc}(\delta_i)$ is a saddle connection between cone points y and x and $g_{sc}(\delta_{i+1})$ a saddle connection between x and z and that they meet at x in an angle less than π on one side. Then, similar to the proof of Lemma 13, there is a nonsingular $\tilde{\varphi}_2$ -geodesic that induces a partition separating x from y and z ; see Figure 6. But since δ_i and δ_{i+1} make up a geodesic segment in the $\tilde{\varphi}_1$ -metric, this gives us a $\tilde{\varphi}_1$ -geodesic bigon, a contradiction. Hence the angle between two consecutive saddle connections must be at least π on both sides, and hence $\tilde{\gamma}_2$ is a $\tilde{\varphi}_2$ -geodesic segment. $\square$

Consider now any φ_1 -triangulation $\mathcal{T}$ of S . Recall that this means that the vertex set is precisely Σ and each triangle is a φ_1 -Euclidean triangle. We lift to a $\pi_1 S$ -invariant triangulation $\tilde{\mathcal{T}}$ of $\tilde{S}$. By Lemma 13, we can extend the identity from $\tilde{\Sigma}$ to

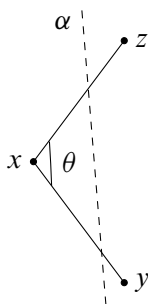


Figure 6. If the angle θ between two consecutive saddle connections measures less than π there is a nonsingular geodesic α separating x from y and z .

itself to a $\pi_1 S$ -equivariant map

$$\tilde{f}_{\mathcal{T}}: \tilde{S} \rightarrow \tilde{S}$$

which sends each triangle in $\tilde{\mathcal{T}}$ to a Euclidean $\tilde{\varphi}_2$ -triangle, and is orientation preserving and affine on the corresponding 2-simplices. This map descends to a map $f_{\mathcal{T}}: S \rightarrow S$ which is also affine on each 2-simplex in $\mathcal{T}$. Since the orientation on each triangle is preserved by $f_{\mathcal{T}}$, and $f_{\mathcal{T}}$ has degree 1, it follows that $f_{\mathcal{T}}$ is a homeomorphism. Our goal is to show that, when φ_1 has holonomy greater than 2, this map is in fact an isometry. Since the areas of φ_1 and φ_2 are both 1, if we prove that the interior angles of the triangles in $\mathcal{T}$ are preserved by $f_{\mathcal{T}}$, then $f_{\mathcal{T}}$ must be an isometry.

Adjusting φ_2 by an isotopy, we can assume that the φ_1 -saddle connections which are the edges of $\mathcal{T}$ are also φ_2 -saddle connections. In particular, $\mathcal{T}$ is also a φ_2 -triangulation. Moreover, through a further isotopy, we can assume that $f_{\mathcal{T}}$ and $\tilde{f}_{\mathcal{T}}$ are the identity on each of the edges in the triangulation. This is convenient to assume throughout the following, and so we make the following additional assumption. In addition to the conventions adopted at the end of Section 2.6 to identify G_1 and G_2 homeomorphically, we further assume that we have fixed a φ_1 -triangulation $\mathcal{T}$ which is also a φ_2 -triangulation, and let $\tilde{\mathcal{T}}$ be its lift to $\tilde{S}$. We also assume that the identity on $(S, \varphi_1) \rightarrow (S, \varphi_2)$ is affine on each triangle.

For each triangle T of $\mathcal{T}$ and each $i = 1, 2$, choose an isometry from T (with its induced path metric) to a Euclidean triangle with respect to φ_i (as noted in Section 2.3, this may not be well-defined on the vertices). Using this, we identify the unit tangent space at each point of each triangle (other than the vertices) with S^1 . Given $\theta \in S^1$ and a triangle T of $\mathcal{T}$, we can then unambiguously refer to the φ_i -direction θ of T . We let $S_i^1(T)$ denote the φ_i -directions of T .

We do the same for triangles of $\tilde{\mathcal{T}}$ in $\tilde{S}$, choosing isometries from triangles in $\tilde{\mathcal{T}}$ to Euclidean triangles by composing the covering map p with the isometries of the corresponding triangles in $\mathcal{T}$. Observe that by construction, for any $\tilde{T}$ in $\tilde{\mathcal{T}}$ and $x \in \tilde{T}$,

the derivative dp_x sends $\tilde{\varphi}_i$ -direction θ to φ_i -direction θ in $T = p(\tilde{T})$. Thus, for any triangle $\tilde{T}$ in $\tilde{\mathcal{T}}$ with $T = p(\tilde{T})$, we use this to identify $S_i^1(T) = S_i^1(\tilde{T})$.

With this convention, we are now able to talk about common triangulations and associated directions in a consistent way, for example in proving the following corollary to Proposition 14.

Corollary 15. *Two currents with the same support have the same cylinder curves: $\text{cyl}(\varphi_1) = \text{cyl}(\varphi_2)$. Moreover, we can homotope any φ_1 -cylinder curve to a φ_2 -cylinder curve in $S \setminus \Sigma$.*

Proof. Lift any φ_1 -cylinder to an infinite strip in $\tilde{S}$. The boundary is a pair of bi-infinite geodesics, $\tilde{\gamma}^\pm$ which are each concatenations of $\tilde{\varphi}_1$ -saddle connections. Moreover, these remain a bounded Hausdorff distance apart (Hausdorff distance equal to the width of the strip, in fact). Now observe that $g(\tilde{\gamma}^\pm)$ is a pair of asymptotic $\tilde{\varphi}_2$ -geodesics which are concatenations of $\tilde{\varphi}_2$ -saddle connections. These must bound an infinite strip, invariant under a cyclic group (since the same is true of the strip bounded by $\tilde{\gamma}^\pm$), and hence the quotient is a φ_2 -cylinder whose core curve is homotopic to the core curve of the φ_1 -cylinder. Thus $\text{cyl}(\varphi_1) \subset \text{cyl}(\varphi_2)$. A symmetric argument proves the reverse containment.

To see that the cylinder curves are homotopic in $S \setminus \Sigma$ we fix a triangulation $\mathcal{T}$ of S . Let $\tilde{\gamma}$ be a $\tilde{\varphi}_1$ -geodesic contained in some $\tilde{\varphi}_1$ -strip. Note that any $\tilde{\varphi}_2$ -geodesic in the corresponding $\tilde{\varphi}_2$ -strip defines the same partition of $\tilde{\Sigma}$ as $\tilde{\varphi}_1$. This follows by approximating the $\tilde{\varphi}_1$ -geodesic by nonsingular geodesics $\tilde{\gamma}_n$ in G_1 and applying g to these. Each $g(\tilde{\gamma}_n)$ define the same partition as $\tilde{\varphi}_1$. By passing to a subsequence, $g(\tilde{\gamma}_n)$ converges to a $\tilde{\varphi}_2$ geodesic in the $\tilde{\varphi}_2$ -strip. However, all $\tilde{\varphi}_2$ geodesics in the $\tilde{\varphi}_2$ -strip determine the same partition. Now, take a nonsingular φ_1 -representative γ_1 of a cylinder curve γ . The ordered and oriented edges of $\mathcal{T}$ crossed by γ determine the homotopy class of γ in $S \setminus \Sigma$. However, a nonsingular φ_2 -geodesic representative γ_2 of the homotopy class of γ must cross the same set of ordered and oriented edges, since their lifts $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ to $\tilde{S}$ determine the same partition of $\tilde{\Sigma}$ in $\tilde{S}$. Hence, γ_1 and γ_2 must be homotopic in $S \setminus \Sigma$. $\square$

3.3. Parallelism preserved. Suppose $\tilde{T}$ is any triangle in $\tilde{\mathcal{T}}$, $i \in \{1, 2\}$, and $\tilde{\gamma} \in G_i$ is a geodesic nontrivially intersecting the interior of $\tilde{T}$. Define $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ to be the closure of the set of geodesics in $\mathcal{G}^*(\tilde{\varphi}_i)$ intersecting $\tilde{T}$ in an arc parallel to the arc of intersection $\tilde{\gamma} \cap \tilde{T}$. We say that $\tilde{\gamma}$ is $\tilde{\varphi}_i$ -generic for $\tilde{T}$ if $\mathcal{G}^*(\tilde{T}, \tilde{\gamma}) \subset G_i$.

Before stating the main technical result of this section, we describe the structure of the set $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ for the triangle $\tilde{T}$, classifying the possibilities for how these geodesics parallel to $\tilde{\gamma}$ hit a particular side δ of each triangle, and allowing us to essentially parametrize the leaves of this foliation by where it hits δ .

Lemma 16. *Fix $i \in \{1, 2\}$, a triangle $\tilde{T}$ in $\tilde{\mathcal{T}}$, and a $\tilde{\varphi}_i$ -generic geodesic $\tilde{\gamma} \in G_i$. Then the intersection of the geodesics of $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ with $\tilde{T}$ defines a foliation by*

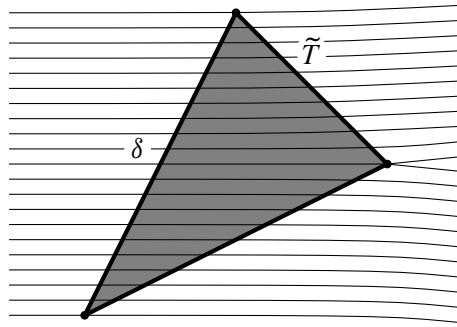


Figure 7. The foliation of a triangle $\tilde{T}$ from a $\tilde{\varphi}_i$ -generic geodesic $\tilde{\gamma}$. The geodesics on the top and bottom of this figure are the unique geodesics in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ which do not meet the interior of $\tilde{T}$. The side δ meets every geodesic in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$.

parallel geodesic segments meeting each side of $\tilde{T}$ transversely. One side δ of $\tilde{T}$ nontrivially intersects every geodesic in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$, defining a function

$$h: \mathcal{G}^*(\tilde{T}, \tilde{\gamma}) \rightarrow \delta$$

that sends each geodesic to its point of intersection with δ . For any $x \in \delta$, the preimage $h^{-1}(x)$ consists of either (1) one nonsingular geodesic, (2) one singular geodesic through one of the two endpoints of δ , or (3) two cone-point asymptotic geodesics meeting $\tilde{T}$ in the same geodesic arc.

Proof. The geodesics of $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ intersect $\tilde{T}$ in the leaves of a parallel geodesic foliation by definition. To see that this foliation meets the sides transversely, we observe that otherwise one of the sides is part of a leaf. Since such a leaf is a geodesic that contains both endpoints of that side, hence two cone points, it is not in G_i , contradicting the $\tilde{\varphi}_i$ -genericity assumption.

To find the unique side δ of $\tilde{T}$ intersecting every geodesic in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$, we consider an isometry from $\tilde{T}$ (with the $\tilde{\varphi}_i$ -metric) to a triangle in $\mathbb{R}^2$ for which the foliation is horizontal. The unique side connecting the highest and lowest vertices is (the image of) δ ; see Figure 7.

Finally, recall that all geodesics in G_i are either nonsingular and uniquely determined by any arc contained in it, or singular and contain exactly one cone point. In the latter case, any arc β in a singular geodesic $\tilde{\gamma}' \in G_i$ not containing the (unique) cone point of $\tilde{\gamma}'$ is contained in exactly one other geodesic $\tilde{\gamma}'' \in G_1$, cone-point asymptotic to $\tilde{\gamma}'$. These two statements imply the claim about the fibers of h . $\square$

With this basic structure in hand, the goal of this section is to prove that the map g preserves the parallel structure described above.

Proposition 17 (Parallelism is preserved). *Suppose $\tilde{T}$ is a triangle in $\tilde{\mathcal{T}}$ and $\tilde{\gamma} \in G_1$ is a $\tilde{\varphi}_1$ -generic geodesic meeting $\tilde{T}$. Then $g(\tilde{\gamma})$ is a $\tilde{\varphi}_2$ -generic geodesic and*

$$\mathcal{G}^*(\tilde{T}, g(\tilde{\gamma})) = g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma})).$$

We will need the following construction in the proof of this proposition. Suppose $\tilde{\gamma}_0, \tilde{\gamma}_1 \in G_i$ is a pair of cone-point asymptotic geodesics. For each $j = 0, 1$, we write $\tilde{\gamma}_j = \tilde{\gamma}_j^- \cup \tilde{\gamma}_j^+$ as a union of rays based at the (unique) cone point ζ in $\tilde{\gamma}_j$, so that $\tilde{\gamma}_0^+ = \tilde{\gamma}_1^+$. The union of the negative rays $\tilde{\gamma}_0^- \cup \tilde{\gamma}_1^-$ bounds a closed *slit space* containing $\tilde{\gamma}_0^+ = \tilde{\gamma}_1^+$ which we denote $\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1)$; see Figure 8. The boundary of $\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1)$ is precisely $\tilde{\gamma}_0^- \cup \tilde{\gamma}_1^-$ and contains the single cone point ζ , making an interior angle of 2π .

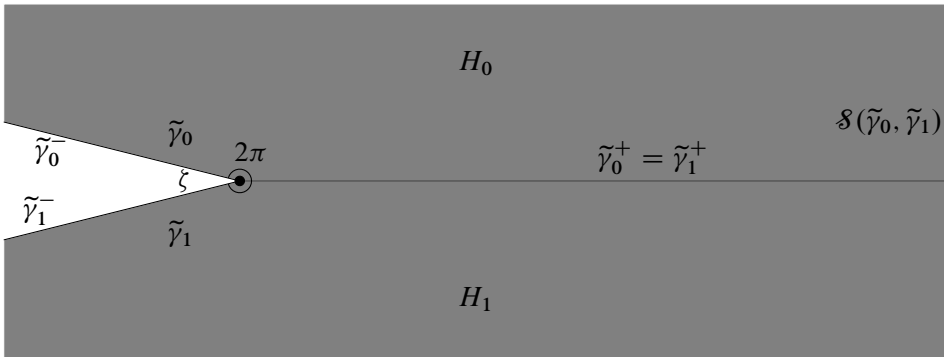


Figure 8. The “slit space” $\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1)$ is a union of half-planes H_0 and H_1 bounded by geodesics $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$, respectively, along their maximal common sub-ray $\tilde{\gamma}_0^+ = \tilde{\gamma}_1^+$ based at the cone point ζ .

Next we observe that these slit spaces hit the cone points in the same way as their g -images.

Lemma 18. *If $\tilde{\gamma}_0, \tilde{\gamma}_1 \in G_1$ are cone-point asymptotic geodesics, then so are $g(\tilde{\gamma}_0)$, $g(\tilde{\gamma}_1)$ and*

$$\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1) \cap \tilde{\Sigma} = \mathcal{S}(g(\tilde{\gamma}_0), g(\tilde{\gamma}_1)) \cap \tilde{\Sigma}.$$

Proof. As already noted $\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1)$ is a union of closed half-planes H_0 and H_1 bounded by $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$, respectively; see Figure 8. We have

$$\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1) \cap \tilde{\Sigma} = (H_0 \cap \tilde{\Sigma}) \cup (H_1 \cap \tilde{\Sigma}).$$

Similarly,

$$\mathcal{S}(g(\tilde{\gamma}_0), g(\tilde{\gamma}_1)) = H'_0 \cup H'_1,$$

where H'_0, H'_1 are closed half-planes bounded by $g(\tilde{\gamma}_0)$ and $g(\tilde{\gamma}_1)$, respectively. Since cone-point partitions are preserved by g (Lemma 12), we have

$$H'_0 \cap \tilde{\Sigma} = H_0 \cap \tilde{\Sigma} \quad \text{and} \quad H'_1 \cap \tilde{\Sigma} = H_1 \cap \tilde{\Sigma},$$

and the lemma follows. $\square$

Proof of Proposition 17. From our basic observations about how the foliation meets each triangle (Lemma 16), there is a unique side δ of the triangle $\tilde{T}$ that meets all the geodesics in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$. For all $\tilde{\gamma}' \in \mathcal{G}^*(\tilde{T}, \tilde{\gamma})$, the two endpoints of δ (which are cone points) lie in different subsets of the partition of $\tilde{\Sigma}$ determined by $\tilde{\gamma}'$, so since the partition is preserved (Lemma 12), the same is true for $g(\tilde{\gamma}')$. Since δ is also a $\tilde{\varphi}_2$ -geodesic (by our convention identifying geodesics), this property of the partitions implies $g(\tilde{\gamma}') \cap \delta \neq \emptyset$.

Let $\tilde{\gamma}_\alpha, \tilde{\gamma}_\omega$ denote the geodesics of $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ intersecting the endpoints of δ (i.e., the unique pair of geodesics in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ that are disjoint from the interior of $\tilde{T}$) and set

$$\mathcal{G}_0^*(\tilde{T}, \tilde{\gamma}) = \mathcal{G}^*(\tilde{T}, \tilde{\gamma}) \setminus \{\tilde{\gamma}_\alpha, \tilde{\gamma}_\omega\}.$$

Note that $g(\mathcal{G}_0^*(\tilde{T}, \tilde{\gamma}))$ consists of geodesics intersecting the interior of $\tilde{T}$ (and so also the interior of δ).

Claim. *For all $\tilde{\gamma}' \in \mathcal{G}_0^*(\tilde{T}, \tilde{\gamma})$, the images $g(\tilde{\gamma})$ and $g(\tilde{\gamma}')$ intersect $\tilde{T}$ in parallel arcs.*

Let us first establish that the proposition follows from this claim. Observe that the claim implies $g(\mathcal{G}_0^*(\tilde{T}, \tilde{\gamma})) \subset \mathcal{G}^*(\tilde{T}, g(\tilde{\gamma}))$, and since g is continuous, we can extend this to both $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\omega$, and therefore

$$g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma})) \subset \mathcal{G}^*(\tilde{T}, g(\tilde{\gamma})).$$

Now suppose $\tilde{\gamma}'' \in \mathcal{G}^*(\tilde{T}, g(\tilde{\gamma}))$ is any geodesic and consider its $\tilde{\varphi}_1$ -straightening. Observe that $\tilde{\gamma}''$ has no transverse intersections with any geodesic in $g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma}))$ and lies between $g(\tilde{\gamma}_\alpha)$ and $g(\tilde{\gamma}_\omega)$. Considering the endpoints on the circle at infinity, we see that a $\tilde{\varphi}_1$ -straightening of $\tilde{\gamma}''$ lies between $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\omega$ and has no transverse intersections with any geodesic from $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$. But such a $\tilde{\varphi}_1$ -geodesic would have to intersect $\tilde{T}$ in a leaf of the foliation of $\tilde{T}$ coming from $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$, and would thus be a geodesic in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$. Therefore, $\tilde{\gamma}'' \in g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma}))$ proving that

$$\mathcal{G}^*(\tilde{T}, g(\tilde{\gamma})) \subset g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma})),$$

and hence $\mathcal{G}^*(\tilde{T}, g(\tilde{\gamma})) = g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma}))$, as required. Therefore, all that remains is to prove the claim.

Proof of Claim. For the remainder of this proof, set $\mathfrak{g} = \mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ and $\mathfrak{g}_0 = \mathcal{G}_0^*(\tilde{T}, \tilde{\gamma})$ to make the notation less cumbersome. We start by defining Z_1 to be the subsurface of $\tilde{S}$ foliated by geodesics from $\mathfrak{g}$. It has singular $\tilde{\varphi}_1$ -geodesic boundary. The proof strategy will be to define a corresponding Z_2 and to show that it can be “zipped up” to a nonsingular Euclidean subsurface with a well-defined notion of angle. This will let us conclude that parallelism is suitably maintained.

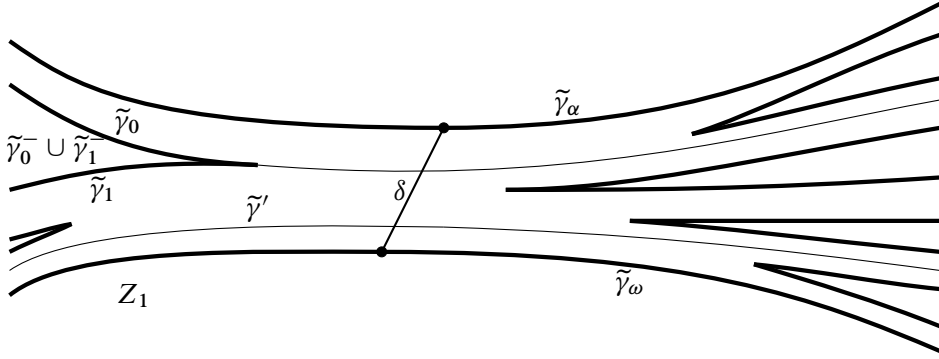


Figure 9. The surface Z_1 together with the geodesic segment δ intersected by all geodesics in $\mathfrak{g}$, the nonsingular boundary components $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\omega$, a pair $\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0$ defining a singular boundary component $\tilde{\gamma}_0^- \cup \tilde{\gamma}_1^-$, and another generic geodesic $\tilde{\gamma}' \in \mathfrak{g}_0$.

We will need an alternative description of Z_1 as follows. From Lemma 16 on the structure of the geodesics in $\mathfrak{g}$, all singular geodesics in $\mathfrak{g}_0$ occur in cone-point asymptotic pairs, and we consider the set of all such pairs:

$$P\mathfrak{g}_0 = \{\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \mid \tilde{\gamma}_0, \tilde{\gamma}_1 \in \mathfrak{g}_0, \text{ cone-point asymptotic}\},$$

which has an obvious map to cone points. Let $\mathcal{H}(\tilde{\gamma}_\alpha)$ and $\mathcal{H}(\tilde{\gamma}_\omega)$ denote the half-spaces bounded by $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\omega$, respectively, containing δ . Then, we have

$$Z_1 = \mathcal{H}(\tilde{\gamma}_\alpha) \cap \mathcal{H}(\tilde{\gamma}_\omega) \cap \bigcap_{\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0} \mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1),$$

where $\mathcal{S}(\tilde{\gamma}_0, \tilde{\gamma}_1)$ is the slit space bounded by the union of subrays $\tilde{\gamma}_0^-$ and $\tilde{\gamma}_1^-$ of $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$, respectively. That Z_1 is a subsurface follows from this description since the boundaries of the half planes and slit spaces in the intersection form a locally finite set. The boundary of Z_1 decomposes as the union of the boundaries of these half-planes and slit spaces, which are precisely $\tilde{\gamma}_\alpha$, $\tilde{\gamma}_\omega$, and the union of pairs of rays $\tilde{\gamma}_0^- \cup \tilde{\gamma}_1^-$, one for each $\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0$. Each boundary component of the form $\tilde{\gamma}_0^- \cup \tilde{\gamma}_1^-$ contains exactly one cone point and makes cone angle 2π on the interior. There are no cone points in the interior of Z_1 . Although there are cone points (the endpoints of δ) on the two boundary components, $\tilde{\gamma}_\alpha$ and $\tilde{\gamma}_\omega$, the interior cone angle in Z_1 is π ; see Figure 9 for an illustration of the various features of Z_1 .

From this description, we obtain a subsurface of $\tilde{S}$ for the $\tilde{\varphi}_2$ -metric in an exactly similar way:

$$Z_2 = \mathcal{H}(g(\tilde{\gamma}_\alpha)) \cap \mathcal{H}(g(\tilde{\gamma}_\omega)) \cap \bigcap_{\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0} \mathcal{S}(g(\tilde{\gamma}_0), g(\tilde{\gamma}_1)).$$

The boundary of Z_2 decomposes just like Z_1 . By Lemma 12 and Lemma 18, Z_2 has no cone points in the interior and exactly one cone point on each boundary

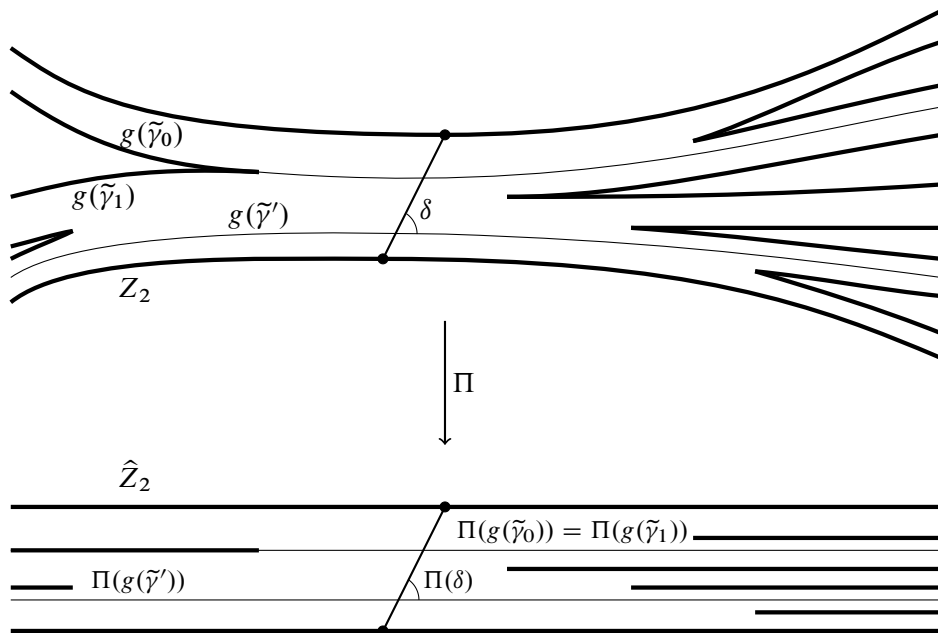


Figure 10. The surface Z_2 and nonsingular quotient $\Pi: Z_2 \rightarrow \hat{Z}_2$, together with: $g(\tilde{\gamma}') \in g(\mathfrak{g}_0)$, the cone-point asymptotic pair $\{g(\tilde{\gamma}_0), g(\tilde{\gamma}_1)\} \in g(P\mathfrak{g}_0)$, and δ , as well as their images under Π . The angle between δ and $g(\tilde{\gamma}')$ is equal to the angle between $\Pi(\delta)$ and $\Pi(g(\tilde{\gamma}'))$.

component. For each pair $\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0$, the corresponding boundary component of Z_2 is of the form

$$g(\tilde{\gamma}_0)^- \cup g(\tilde{\gamma}_1)^-$$

and has interior cone angle 2π at the unique cone point it contains, while $g(\tilde{\gamma}_\alpha)$ and $g(\tilde{\gamma}_\omega)$ have interior cone angle π at their cone point.

We now construct a quotient *nonsingular* Euclidean surface $\Pi: Z_2 \rightarrow \hat{Z}_2$ by “zipping up the slits”: we isometrically identify the pair of rays $g(\tilde{\gamma}_0)^-$ and $g(\tilde{\gamma}_1)^-$, for each $\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0$. Away from the cone points, Π is a local isometric embedding, and it is a local isometry away from ∂Z_2 . In particular, $\Pi(\delta)$ is a geodesic arc connecting the boundary components $\Pi(g(\tilde{\gamma}_\alpha))$ and $\Pi(g(\tilde{\gamma}_\omega))$, and for any geodesic $\tilde{\gamma}' \in \mathfrak{g}_0$, $\Pi(g(\tilde{\gamma}'))$ is a geodesic in $\hat{Z}_2$ intersecting $\Pi(\delta)$ in the same angle that $g(\tilde{\gamma}')$ intersects δ . Observe that for any pair $\{\tilde{\gamma}_0, \tilde{\gamma}_1\} \in P\mathfrak{g}_0$ we have $\Pi(g(\tilde{\gamma}_0)) = \Pi(g(\tilde{\gamma}_1))$; see Figure 10.

Now let $\tilde{\gamma}' \in \mathfrak{g}_0$ be any disjoint geodesic distinct from $\tilde{\gamma}$, and suppose that $g(\tilde{\gamma}')$ and $g(\tilde{\gamma})$ intersect δ in different angles. The same is then true of the images in $\hat{Z}_2$. Since $\hat{Z}_2$ is simply connected and complete with geodesic boundary, the developing map is globally defined and injective, and is thus an isometric embedding to a closed subset of $\mathbb{R}^2$. Therefore, the images of $\Pi(g(\tilde{\gamma}'))$ and $\Pi(g(\tilde{\gamma}))$, being bi-infinite

lines in $\mathbb{R}^2$ making different angles with the image of $\Pi(\delta)$ must intersect. On the other hand, $\tilde{\gamma}'$ and $\tilde{\gamma}$ are disjoint, and hence so are $g(\tilde{\gamma}')$ and $g(\tilde{\gamma})$. By construction of $\hat{Z}_2$, $\Pi(g(\tilde{\gamma}'))$ and $\Pi(g(\tilde{\gamma}))$ are also disjoint, and hence so are their images in $\mathbb{R}^2$. This is impossible, and so $g(\tilde{\gamma}')$ and $g(\tilde{\gamma})$ intersect δ in the same angle, proving the claim. $\square$

Since the claim proves the proposition, we are done. $\square$

3.4. Holonomy preserved. Our next goal is to use this preservation of parallelism to conclude that holonomy is also preserved.

Suppose that $\tilde{T}$ is a triangle in $\tilde{\mathcal{T}}$ and $\tilde{\gamma} \in G_i$ is a $\tilde{\varphi}_i$ -generic geodesic for $\tilde{T}$ (and so intersecting the interior of $\tilde{T}$). Observe that any other geodesic in $\mathcal{G}^*(\tilde{T}, \tilde{\gamma})$ intersecting the interior of $\tilde{T}$ is also $\tilde{\varphi}_i$ -generic for $\tilde{T}$. Therefore, if $\theta \in S_i^1(\tilde{T})$ is the direction in $\tilde{T}$ of $\tilde{\gamma}$, we can unambiguously refer to θ as a $\tilde{\varphi}_i$ -generic direction, and we write

$$\mathcal{G}^*(\tilde{T}, \theta) = \mathcal{G}^*(\tilde{T}, \tilde{\gamma}).$$

If T is the image $T = p(\tilde{T})$ in S , then we have an identification of the space of directions $S_i^1(\tilde{T}) = S_i^1(T)$, and we define a direction in $S_i^1(T)$ to be φ_i -generic if it is $\tilde{\varphi}_i$ -generic in $S_i^1(\tilde{T})$. Since G_1 is invariant under the action of $\pi_1 S$, this is independent of the choice of triangle $\tilde{T}$ in the preimage of T . Note that for any triangle T , there are only countably many non- φ_i -generic directions.

Let Γ denote the dual graph of the 1-skeleton $\mathcal{T}^{(1)}$, which we view as embedded transversely and minimally intersecting $\mathcal{T}^{(1)}$. An edge path $\delta: [0, 1] \rightarrow \Gamma$ has initial vertex $\delta(0)$ in some unique triangle T of $\mathcal{T}$ and terminal vertex in a unique triangle T' of $\mathcal{T}$. We say that δ is an *edge path of Γ from T to T'* , and we denote the initial and terminal vertices as $v_T = \delta(0)$ and $v_{T'} = \delta(1)$. Given such a path δ and direction $\theta \in S_i^1(T)$, we can φ_i -parallel translate θ along δ to obtain a φ_i -direction $P_{\delta, \varphi_i}(\theta)$ in T' . The parallel translate $P_{\delta, \varphi_i}(\theta)$ depends only on the homotopy class of δ , rel endpoints, in Γ (or equivalently in $S \setminus \Sigma$) since the metric is Euclidean away from the cone points. In particular, if $v_T \in \Gamma^{(0)}$ is the base point in T then this defines the φ_i -holonomy homomorphism

$$P_{\varphi_i}: \pi_1(\Gamma, v_T) \cong \pi_1(S \setminus \Sigma, v_T) \rightarrow SO(2),$$

by $P_{\varphi_i}([\delta]) \cdot \theta = P_{\delta, \varphi_i}(\theta)$ for any loop δ based at v_T . The isomorphism

$$\pi_1(\Gamma, v_T) \cong \pi_1(S \setminus \Sigma, v_T)$$

is induced by inclusion. We note that the map P_{φ_i} has no dependence on the base point. This is justified by the fact that the fundamental groups based at any two points differ by an isomorphism which is canonical, up to inner isomorphism, and

since $SO(2)$ is abelian, the resulting holonomy homomorphism differs from the one based at v_T in this prescribed way. In particular, we may write

$$P_{\varphi_i}: \pi_1(S \setminus \Sigma) \rightarrow SO(2)$$

without any confusion.

Given a triangle T of $\mathcal{T}$, we say that a direction $\theta \in S_i^1(T)$ is φ_i -stably generic if for every triangle T' in $\mathcal{T}$ and path δ from T to T' , $P_{\delta, \varphi_i}(\theta) \in S_i^1(T')$ is φ_i -generic. It follows that for any path δ in Γ from T to T' , the direction $\theta \in S_i^1(T)$ is φ_i -stably generic if and only if $P_{\delta, \varphi_i}(\theta) \in S_i^1(T')$ is φ_i -stably generic. Since there are only countably many homotopy classes of paths between any two vertices of Γ , and only countably many non- φ_i -generic directions in $S_i^1(T)$, for any T in $\mathcal{T}$, it follows that there are only countably many non- φ_i -stably generic directions in $S_i^1(T)$. We denote the complementary set of φ_i -stably generic directions in T by

$$\Delta(T, \varphi_i) = \{\theta \in S_i^1(T) \mid \theta \text{ is } \varphi_i\text{-stably generic in } T\}.$$

The next result is the key to determining angles of triangles when the holonomy is infinite.

Proposition 19 (Defining angles). *For every triangle $T \in \mathcal{T}$ there is an orientation-preserving homeomorphism*

$$F_T: S_1^1(T) \rightarrow S_2^1(T),$$

which is $\pi_1(\Gamma, v_T)$ -equivariant with respect to the holonomy homomorphisms:

$$F_T(P_{\varphi_1}(\delta) \cdot \theta) = P_{\varphi_2}(\delta) \cdot F_T(\theta).$$

Furthermore, if $\tilde{T}$ is a triangle in $\tilde{\mathcal{T}}$ with $p(\tilde{T}) = T$ and $\theta \in \Delta(T, \varphi_1)$, then $F_T(\theta) \in \Delta(T, \varphi_2)$ and

$$g(\mathcal{G}^*(\tilde{T}, \theta)) = \mathcal{G}^*(\tilde{T}, F_T(\theta)).$$

By the last statement, F_T is determined by g via the stably generic geodesic foliations of $\tilde{T}$.

Proof. Fix any triangle $\tilde{T}$ in $\tilde{\mathcal{T}}$. According to Proposition 17, for any $\tilde{\varphi}_1$ -generic direction $\theta \in S_1^1(\tilde{T})$ and $\tilde{\varphi}_1$ -geodesic $\tilde{\gamma}$ intersecting $\tilde{T}$ in direction θ , we have that $g(\tilde{\gamma})$ is a $\tilde{\varphi}_2$ -generic geodesic, and we can uniquely define $F_{\tilde{T}}(\theta) \in S_2^1(\tilde{T})$ to be the $\tilde{\varphi}_2$ -generic direction of $g(\tilde{\gamma})$. This uniquely determines a bijection $F_{\tilde{T}}$ from the set of $\tilde{\varphi}_1$ -generic directions in $S_1^1(\tilde{T})$ to the $\tilde{\varphi}_2$ -generic directions in $S_2^1(\tilde{T})$ satisfying

$$g(\mathcal{G}^*(\tilde{T}, \theta)) = g(\mathcal{G}^*(\tilde{T}, \tilde{\gamma})) = \mathcal{G}^*(\tilde{T}, g(\tilde{\gamma})) = \mathcal{G}^*(\tilde{T}, F_{\tilde{T}}(\theta)).$$

Consider any three directions $\theta_1, \theta_2, \theta_3$ appearing cyclically in this order around $S_1^1(\tilde{T})$ and let $\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3$ be $\tilde{\varphi}_1$ -geodesics through a single point x in the

interior of $\tilde{T}$ in each of these directions, respectively. The forward endpoints on the circle at infinity for each of these geodesics is the same as for their images by g , and hence $F_{\tilde{T}}(\theta_1)$, $F_{\tilde{T}}(\theta_2)$, $F_{\tilde{T}}(\theta_3)$ also appear cyclically in the same order around $S_2^1(\tilde{T})$. That is, $F_{\tilde{T}}$ preserves the cyclic ordering. It follows that for $T = p(\tilde{T})$, $F_{\tilde{T}}$ uniquely extends to a homeomorphism

$$F_T: S_1^1(T) \rightarrow S_2^1(T).$$

By construction, the last condition of the proposition is satisfied.

Next, suppose $\tilde{T}, \tilde{T}'$ are triangles that share an edge $\tilde{e}$, let $T = p(\tilde{T})$ and $T' = p(\tilde{T}')$ with $e = p(\tilde{e}) \in T \cap T'$ the image of the shared edge, and let δ be the edge of Γ dual to e with $\tilde{\delta}$ a lift of δ intersecting $\tilde{e}$. Suppose $\theta \in S_1^1(T)$ is any φ_1 -stably generic direction and $\tilde{\gamma}$ is a geodesic intersecting $\tilde{T}$ in the direction θ , and suppose that $\tilde{\gamma}$ crosses $\tilde{e}$. From the property of F_T we have already proved, we know that the direction of $g(\tilde{\gamma})$ in $\tilde{T}$ is $F_T(\theta)$. Since $\tilde{\gamma}$ and $g(\tilde{\gamma})$ cross $\tilde{e}$, both intersect $\tilde{T}'$. Parallel transport along an arc of $\tilde{\gamma}$ contained in the union $\tilde{T} \cup \tilde{T}'$ or along $\tilde{\delta}$ (or equivalently along δ via the identification from the covering p) both define the same maps

$$P_{\delta, \varphi_1}: S_1^1(T) \rightarrow S_1^1(T'),$$

and it follows that the direction of $\tilde{\gamma}$ in $\tilde{T}'$ is $P_{\delta, \varphi_1}(\theta)$. Similarly, the direction of $g(\tilde{\gamma})$ in $\tilde{T}'$ is $P_{\delta, \varphi_2}(F_T(\theta))$. On the other hand, applying what we already proved about $F_{T'}$, we know that the direction of $g(\tilde{\gamma})$ in $\tilde{T}'$ is $F_{T'}(P_{\delta, \varphi_1}(\theta))$. That is,

$$F_{T'} \circ P_{\delta, \varphi_1}(\theta) = P_{\delta, \varphi_2} \circ F_T(\theta). \quad (3.1)$$

Since this is true on a dense set in $S_1^1(T)$, it is true on all of $S_1^1(T)$, by continuity.

If $\delta = \delta_1 \delta_2 \cdots \delta_k$ is any edge path connecting vertices $v_{T_0}, v_{T_1}, \dots, v_{T_k}$ (so that δ_j is an edge from $v_{T_{j-1}}$ to v_{T_j}), then by repeatedly applying equation (3.1) we have

$$\begin{aligned} P_{\delta, \varphi_2} \circ F_{T_0} &= P_{\delta_2 \cdots \delta_k, \varphi_2} \circ F_{T_1} \circ P_{\delta_1, \varphi_1} \\ &= P_{\delta_3 \cdots \delta_k, \varphi_2} \circ F_{T_2} \circ P_{\delta_1 \delta_2, \varphi_1} = \cdots = F_{T_k} \circ P_{\delta, \varphi_1}. \end{aligned}$$

The fact that F_T is equivariant with respect to the holonomy homomorphisms P_{φ_1} and P_{φ_2} now follows by taking δ in this equation to be a loop based at a vertex of Γ . $\square$

Corollary 20. φ_1 and φ_2 induce the same holonomy homomorphism

$$P_{\varphi_1} = P_{\varphi_2}: \pi_1(S \setminus \Sigma) \rightarrow SO(2).$$

Proof. For every $\gamma \in \pi_1(\Gamma, v_T) = \pi_1(S \setminus \Sigma, v_T)$, Proposition 19 implies that F_T is an orientation-preserving topological conjugacy between $P_{\varphi_1}(\gamma)$ and $P_{\varphi_2}(\gamma)$. Any

rotation is determined by the cyclic ordering on any orbit and its action on this cyclically ordered set. Consequently, $P_{\varphi_1}(\gamma) = P_{\varphi_2}(\gamma)$, and since γ was arbitrary,

$$P_{\varphi_1} = P_{\varphi_2}. \quad \square$$

When the holonomy is infinite, we get the most information about the homeomorphism F_T .

Corollary 21. *If $P_{\varphi_1}(\pi_1(\Gamma))$ is infinite, then for each triangle T , the map F_T is an isometric identification of the circle of directions in the two metrics. In particular, F_T conjugates P_{φ_1} to P_{φ_2} inside $SO(2)$.*

To emphasize this last point, we note that F_T is *a priori* a topological conjugacy, but when the holonomy is infinite, it is in fact an *isometric* conjugacy.

Proof. Since $P_{\varphi_1}(\pi_1(\Gamma, v_T))$ is an infinite, finitely generated abelian group, there exists an element $[\delta] \in \pi_1(\Gamma, v_T)$ such that $P_{\varphi_1}([\delta])$ has infinite order. Then both $P_{\varphi_1}([\delta])$ and $P_{\varphi_2}([\delta])$ are irrational rotations, topologically conjugate by the homeomorphism F_T . It follows that their angles of rotation are equal, and F_T is an isometry on the orbit of any point by $\langle P_{\varphi_1}([\delta]) \rangle$. Since this orbit is dense, F_T is an isometry. $\square$

From this, we easily deduce the following, which essentially proves the Support Rigidity theorem when the holonomy has infinite order.

Corollary 22. *If $P_{\varphi_1}(\pi_1(\Gamma))$ is infinite, then the φ_1 -interior angle at any vertex of any triangle T of $\tilde{\mathcal{T}}$ is equal to the φ_2 -interior angle of the same vertex.*

Proof. Let $\tilde{T}$ be a triangle with $p(\tilde{T}) = T$ and for each $i = 1, 2$, let $\Theta_i(\tilde{T}) \subset S_i^1(\tilde{T})$ be the set of six $\tilde{\varphi}_i$ -directions parallel to the sides of $\tilde{T}$ (each side is considered with both orientations so appears twice). For every $\theta \in S_i^1(\tilde{T}) \setminus \Theta_i(\tilde{T})$, the foliation of $\tilde{T}$ determined by $\mathcal{G}^*(\tilde{T}, \theta)$ is transverse to each of the sides. Furthermore, for any such θ , the side which meets every leaf defines a locally constant function

$$h_i: S_i^1(\tilde{T}) \setminus \Theta_i(\tilde{T}) \rightarrow \text{sides}(\tilde{T}).$$

This function changes values at precisely each of the six directions in $\Theta_i(\tilde{T})$. In fact, h_i changes value from one side to another at the direction of the third side. Now note that for a $\tilde{\varphi}_1$ -generic direction θ , we have $h_1(\theta) = h_2(F_T(\theta))$, by Proposition 19. Consequently,

$$F_T(\Theta_1(\tilde{T})) = \Theta_2(\tilde{T}).$$

Therefore, since the previous result shows that F_T is an isometry, the $\tilde{\varphi}_1$ -directions of the three sides of $\tilde{T}$ differ from the $\tilde{\varphi}_2$ -directions by an isometry, and consequently, the $\tilde{\varphi}_1$ -angles and $\tilde{\varphi}_2$ -angles of $\tilde{T}$ agree. Pushing back down to T proves the corollary. $\square$

3.5. Proof of the Support Rigidity theorem. We are now ready to put together all the pieces and show that a flat metric is determined (up to affine equivalence) by its support.

Proof of the Support Rigidity theorem. We continue to assume, without loss of generality, that φ_1 and φ_2 are representatives chosen so that we can identify geodesics and triangulations as above. The proof of the Support Rigidity theorem divides into two cases, depending on whether the holonomy is finite or infinite.

Case 1. The holonomy of φ_1 is infinite.

For every triangle T of $\mathcal{T}$ and every vertex of T , the interior φ_1 -angle is equal to the interior φ_2 -angle (Corollary 22), and thus, the φ_1 -metric and φ_2 -metric on T are similar. Since the identity $(S, \varphi_1) \rightarrow (S, \varphi_2)$ is already affine on each triangle $T \in \mathcal{T}$, it is in fact a similarity, and so it scales distances by a constant factor c_T . When two triangles share an edge, their scaling factors must be equal. Since S is connected, $c_T = c_{T'}$ for any two triangles T, T' of $\mathcal{T}$, and hence φ_1 and φ_2 differ by some global scalar $c > 0$. Since both φ_1 and φ_2 have unit area, the scalar c must equal 1, and hence $\varphi_1 = \varphi_2$.

Case 2. The holonomy of φ_1 is finite.

Since the two metrics induce the same holonomy homomorphism (Corollary 20), we can let $\pi: S' \rightarrow S$ denote the branched cover corresponding to the kernel of this homomorphism P_{φ_i} . More precisely, S' is the metric completion of the cover of $S \setminus \Sigma$ corresponding to the kernel, and π is the extension of the covering map to the completion (which sends completion points Σ' of S' to appropriate points of Σ). Write $\varphi'_i = \pi^*(\varphi_i)$ for the pullback of φ_i to S' , for $i = 1, 2$. Since the cover was constructed from the kernel, both of these metrics have trivial holonomy.

Moreover, the cylinder sets are equal in the two metrics (Corollary 15), and the homotopy from a cylinder curve for φ_1 to a cylinder curve for φ_2 occurs in the complement of Σ . Now observe that the cylinders for φ'_i are precisely the preimages of the cylinders for φ_i under π , for $i = 1, 2$, and hence $\text{cyl}(\varphi'_1) = \text{cyl}(\varphi'_2)$ on $S' \setminus \Sigma'$. By passing to these covers we are now in Flat_1 and we may invoke Theorem 6 to conclude from the equality of cylinder sets that φ'_1 and φ'_2 are affine-equivalent on $S' \setminus \Sigma'$. Let f' be the affine map, isotopic to the identity on $S' \setminus \Sigma'$, which extends by the identity over Σ' .

Now every triangle $T \in \mathcal{T}$ is covered by a triangle $T' \subset \pi^{-1}(T)$, and the map

$$f'|_{T'}: (T', \varphi'_1) \rightarrow (f'(T'), \varphi'_2)$$

is affine. By our convention for triangulations, the identity $(T', \varphi'_1) \rightarrow (T', \varphi'_2)$ is also affine. So, since f' is isotopic to the identity rel Σ' , it follows that for each edge e of T' , $f'(e)$ is the unique φ'_2 -geodesic representative of the isotopy class (namely, the straight segment). Since e is already a φ'_2 -geodesic, these must be equal.

That is, $f'(T') = T'$. Since both the identity and f' are affine maps from (T', φ'_1) to (T', φ'_2) , they must be equal. Since T' was an arbitrary triangle in the preimage of an arbitrary triangle of $\mathcal{T}$, it follows that f' is the identity. Therefore, the identity $(S, \varphi_1) \rightarrow (S, \varphi_2)$ is affine, and hence φ_1 and φ_2 are affine equivalent.

If the holonomy has order greater than 2, then affine-equivalent flat metrics are actually equal (Proposition 5), so $\varphi_1 = \varphi_2$. This completes the proof. $\square$

4. Bounce theorem and related results

We are now ready to prove our main theorem on billiards. We begin by recalling the statement.

Bounce theorem. *If finite sided, simply connected Euclidean polygons P_1, P_2 have $B(P_1) = B(P_2)$, then either P_1, P_2 are right-angled and affinely equivalent, or they are similar polygons.*

In Section 4.2–4.4 we describe corollaries of this result, and other applications of the ideas in the proof.

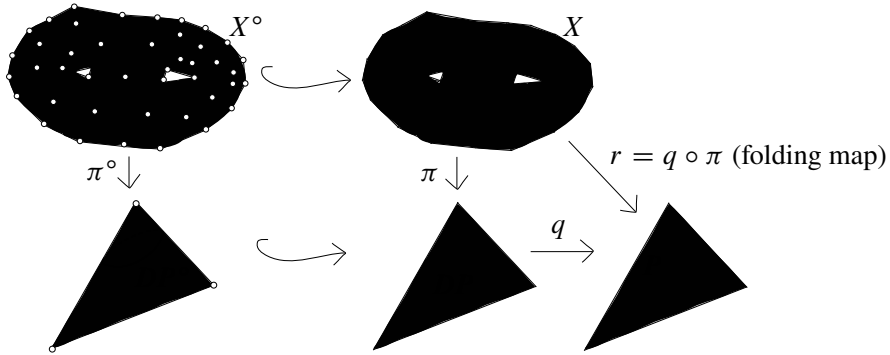
4.1. Proof of the Bounce theorem. We view billiard trajectories in P as unit speed piecewise geodesic paths $\tau: \mathbb{R} \rightarrow P$. Recall also that with respect to a cyclic labeling of the edges from an ordered alphabet $\mathcal{A}$, the bounce spectrum $B(P)$ is the set of all bounce sequences (sequences of letters from $\mathcal{A}$) that can occur along billiard trajectories.

In order to prove the Bounce theorem, we will heavily use the results of the last section. To do this, we relate billiard trajectories to nonsingular geodesics on flat surfaces.

Definition 23 (Unfolding). Suppose P is a simply connected Euclidean polygon with n vertices, X is an oriented surface of negative Euler characteristic, and G is a finite group acting faithfully on X with quotient P . Then the quotient map $r: X \rightarrow P$ is called a *folding map*. We pull back the flat metric on P by r to a metric on X . If the preimage of every vertex in P is a cone point with angle more than 2π , then X together with its pullback metric is called an *unfolding* of P .

We also give an equivalent constructive definition of an unfolding of P and introduce some notation. This highlights that unfoldings are tessellated by lifts of P and that each tile is a fundamental domain for the action of G . For the following description, refer to Figure 11 throughout.

Let DP be the double of P , obtained by gluing two copies of P along corresponding edges. There is an involution interchanging the two copies of P and a natural quotient map $q: DP \rightarrow P$ given by identifying all corresponding points. Let DP° be DP with the vertices of the copies of P removed. Note that DP° is a topological sphere with n punctures. Let $\pi^\circ: X^\circ \rightarrow DP^\circ$ be a finite-sheeted regular cover such that the involution on DP° interchanging the two copies of P lifts

Figure 11. A genus-2 unfolding of P .

to X° . Both X° and DP° have a flat metric induced by the metric on P . Let X be the metric completion of X° . As above, we assume that the cone angles are more than 2π . The covering map can be extended to a regular branched covering map $\pi: X \rightarrow DP$. Then X is an unfolding of P , and the map $r = q \circ \pi$ is a folding map. Furthermore, G is the group generated by the deck group of $\pi: X \rightarrow DP$ and a lift of the involution on DP . In particular, it follows that for any edge of the tessellation there is an element in G that locally acts as a reflection in this edge. The quotient of X by G is precisely P .

Remark. In contrast to the theory of rational billiards, we are considering all unfoldings, not just those that carry an induced translation structure. The angle condition ensures that X is nonpositively curved. We will use this construction to pass to simultaneous unfoldings of a pair of polygons P_1, P_2 to flat surfaces X_1, X_2 with the same underlying topology S .

Lemma 24. *For an unfolding as above, if $\gamma: \mathbb{R} \rightarrow X$ is a unit speed nonsingular geodesic, then $r \circ \gamma: \mathbb{R} \rightarrow P$ is a billiard trajectory on P . In fact, this defines a bijection between G -orbits of nonsingular geodesics in X and billiard trajectories on P .*

Viewing $r: X^{\circ} \rightarrow P^{\circ}$ as a Euclidean orbifold cover (and P as a reflector orbifold), this lemma follows from standard orbifold generalizations of covering space theory; see [41, Chapter 13]. However, the proof is very constructive, so we sketch it here.

Proof of Lemma 24. Fix a nonsingular geodesic γ on X . Since r is a local isometry away from the edges of the tessellation by copies of P , it follows that $r \circ \gamma$ is a concatenation of geodesic segments between the edges of P . Moreover, we know that for every edge of the tessellation, there is an element of G that acts locally as a reflection in that edge. Thus $r \circ \gamma$ has optical reflection at each point of intersection with an edge of P (see Figure 12), and $r \circ \gamma$ defines a billiard trajectory on P .

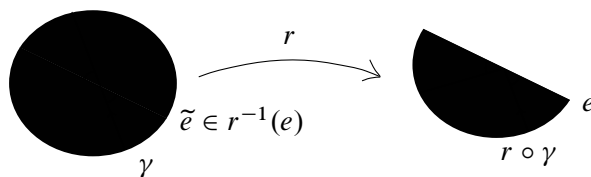


Figure 12. The image of a nonsingular geodesic γ on X under r has optical reflection at points of intersection with an edge e of P .

Conversely, given a billiard trajectory, any lift to X is a piecewise geodesic path and optical reflection is precisely what is needed to ensure that this piecewise geodesic is in fact an actual geodesic. Hence every nonsingular geodesic is a lift of a billiard trajectory, and by definition of unfolding, any two lifts differ by an element in G . $\square$

Proof of the Bounce theorem. Suppose $B(P_1) = B(P_2)$. By counting observed labels, this implies that P_1 and P_2 have the same number of edges. This allows us to construct a homeomorphism $f: P_1 \rightarrow P_2$ respecting the cyclic labeling from the alphabet $\mathcal{A}$. We assume the restriction of f to each edge of P_1 is affine onto the image edge of P_2 .

The map $f: P_1 \rightarrow P_2$ lifts to a homeomorphism $Df: DP_1 \rightarrow DP_2$, and we denote the restriction to the complements of the vertices by $Df^\circ: DP_1^\circ \rightarrow DP_2^\circ$, which is also a homeomorphism. For each $i = 1, 2$, we choose compact, nonpositively curved unfoldings X_i and denote the folding map $r_i: X_i \rightarrow P_i$. We assume that the preimage of every vertex by r_i is a cone point of angle greater than 2π . The folding maps factor through branched coverings $X_i \rightarrow DP_i$ which restrict to (unbranched) covering maps $X_i^\circ \rightarrow DP_i^\circ$, and we can therefore identify the fundamental group $\pi_1 X_i^\circ$ as a subgroup $\pi_1 X_i^\circ < \pi_1 DP_i^\circ$. By passing to further finite-sheeted branched covers, we can assume that

$$Df_*(\pi_1 X_1^\circ) = \pi_1 X_2^\circ$$

and hence Df lifts to a homeomorphism $F: X_1 \rightarrow X_2$ fitting into a commutative diagram

$$\begin{array}{ccc} X_1 & \xrightarrow{F} & X_2 \\ \downarrow & & \downarrow \\ DP_1 & \xrightarrow{Df} & DP_2 \\ \downarrow & & \downarrow \\ P_1 & \xrightarrow{f} & P_2. \end{array}$$

Recall that for each $i = 1, 2$, X_i has a tessellation whose tiles are copies of P_i . The map F sends tiles of X_1 to tiles of X_2 by construction. Furthermore, the edges of the tessellation are labeled according to the labels on the original polygon, and F is affine on the edges and preserves the labeling.

Claim. *The map $F: X_1 \rightarrow X_2$ is isotopic to an affine map $F': X_1 \rightarrow X_2$ such that $F' = F$ on the edges of the tessellation.*

Proof. Suppose $\gamma_1: \mathbb{R} \rightarrow X_1$ is any bi-infinite nonsingular geodesic. By Lemma 24, the projection to P_1 is a billiard trajectory with bounce sequence $\mathbf{b}$. This sequence is precisely the sequence of labels on edges of the tessellation crossed by γ_1 . Since F preserves the labels on the edges of the tessellation, $F \circ \gamma_1$ crosses the edges creating the same sequence $\mathbf{b}$. Since $B(P_1) = B(P_2)$, this is also the bounce sequence of a billiard trajectory in P_2 , which lifts to a bi-infinite nonsingular geodesic $\gamma_2: \mathbb{R} \rightarrow X_2$ (again by Lemma 24) and produces the same sequence of labels according to the edges it crosses. By composing with an element of the group G associated to X_2 , we can assume that γ_2 crosses the same set of edges as $F \circ \gamma_1$, in the same direction. In particular, after reparametrizing if necessary, we can find a homotopy between $F \circ \gamma_1$ and γ_2 in X_2° with uniformly bounded length traces (that is, the homotopy $H_t: \mathbb{R} \rightarrow X_2^\circ$, $t \in [0, 1]$, from $F \circ \gamma_1$ to γ_2 has the property that the paths $t \mapsto H_t(x)$ have uniformly bounded length, independent of $x \in \mathbb{R}$; in fact, the length is bounded by the diameter of P_2 with its induced path metric).

Now let φ_1 be the given flat metric on $S = X_1$ and φ_2 the pullback of the flat metric on X_2 by the map F . Observe that by construction, φ_1 and φ_2 have the same set of cone points. The previous paragraph shows that any bi-infinite nonsingular $\tilde{\varphi}_1$ -geodesic in the universal cover $\tilde{\gamma}_1: \mathbb{R} \rightarrow \tilde{S}$ is uniformly close to a bi-infinite nonsingular $\tilde{\varphi}_2$ -geodesic $\tilde{\gamma}_2: \mathbb{R} \rightarrow \tilde{S}$. Moreover, these are homotopic in the complement of the (common) cone point set $\tilde{\Sigma}$. Therefore the metrics have the same set of endpoints of basic geodesics: $\mathcal{G}_{\tilde{\varphi}_1}^* = \mathcal{G}_{\tilde{\varphi}_2}^*$. Furthermore, since the homotopy between a nonsingular $\tilde{\varphi}_1$ -geodesic and a corresponding $\tilde{\varphi}_2$ -geodesic occurs within $\tilde{S} \setminus \tilde{\Sigma}$, it follows that these geodesics define the same partition of $\tilde{S}$.

Next, recall that our identification of chains (Proposition 10) gives us a $\pi_1 S$ -invariant bijection $\tilde{\Sigma} \rightarrow \tilde{\Sigma}$, and in the proof of the Support Rigidity theorem we extended this to a $\pi_1 S$ -equivariant affine map from $(\tilde{S}, \tilde{\varphi}_1) \rightarrow (\tilde{S}, \tilde{\varphi}_2)$, $\pi_1 S$ -equivariantly isotopic to the identity. In fact, since any nonsingular $\tilde{\varphi}_1$ -geodesic defines the same partition of $\tilde{\Sigma}$ as a corresponding nonsingular $\tilde{\varphi}_2$ -geodesic and since any cone point can be uniquely determined by finitely many such partitions, it follows from Proposition 12 (Cone point partitions are well-defined) that this bijection $\tilde{\Sigma} \rightarrow \tilde{\Sigma}$ is the identity.

Therefore, we have determined that the identity on $\tilde{S}$ is $\pi_1 S$ -equivariantly isotopic, rel $\tilde{\Sigma}$, to a $\pi_1 S$ -equivariant affine map $\tilde{\Phi}: (\tilde{S}, \tilde{\varphi}_1) \rightarrow (\tilde{S}, \tilde{\varphi}_2)$. Since each edge of the (lifted) tessellation of $\tilde{S}$ by copies of P_1 is both a $\tilde{\varphi}_1$ -saddle

connection as well as a $\tilde{\varphi}_2$ -saddle connection, and the identity is already an affine map, it follows that $\tilde{\Phi}$ is the identity on these edges. Equivalently, the induced map $\Phi: (S, \varphi_1) \rightarrow (S, \varphi_2)$ is an affine map, isotopic to the identity, which is the identity on the edges of the tessellation by copies of P_1 . The composition $F' = F \circ \Phi: X_1 \rightarrow X_2$ is an affine map, is isotopic to F , agreeing with F on the edges of the tessellation. $\square$

With the claim in hand, we finish the proof of the Bounce theorem. If P_1 is not a right-angled polygon, then there is a vertex with interior angle not in $\frac{\pi}{2}\mathbb{N}$; by an appropriate choice of unfolding X_1 of P_1 , we can assume that the holonomy of X_1 has order greater than 2 (specifically, we can force the holonomy around the cone point corresponding to the chosen vertex to have order greater than 2). By the Support Rigidity theorem, the affine map $F': X_1 \rightarrow X_2$ is a similarity (here we must allow a similarity, rather than isometry, since the two metrics have not been normalized to unit area). In this case, the map $F': X_1 \rightarrow X_2$ descends to a similarity $f': P_1 \rightarrow P_2$, preserving the labeling, as required.

If P_1 is a right-angled polygon, then the affine map $F': X_1 \rightarrow X_2$ still descends to a map $f': P_1 \rightarrow P_2$, but we also note that it descends to an affine map $Df': DP_1 \rightarrow DP_2$ on the doubles, sending each of the two copies of P_1 onto one of the copies of P_2 . By rotating if necessary, we may assume that P_1 has horizontal and vertical sides. Pick a horizontal side e_1 of P_1 , and rotate P_2 so that the corresponding side e_2 of P_2 is also horizontal. The derivative of $f': P_1 \rightarrow P_2$ therefore has the form $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ for some $a, b, c \in \mathbb{R}$. On the other hand, from Df' , we see that f' extends to an affine map from the reflection of P_1 over e_1 to the reflection of P_2 over e_2 , and hence its derivative commutes with the reflection in the x -axis. An easy computation shows that $b = 0$, and hence P_2 is also right-angled. $\square$

4.2. Generalized diagonals. Next, we prove Corollary 1, which states that the (finite) bounce sequences of all generalized diagonals – billiard trajectories starting and ending at a vertex – determine the shape of the polygon. For the proof, we will need some additional setup.

Given a polygon P with labels in $\mathcal{A} = \{1, \dots, n\}$, add a “dummy” label to form $\mathcal{A}_0 = \{0, 1, \dots, n\}$, and view $B(P)$ as a subset of $\mathcal{A}^{\mathbb{Z}} \subset \mathcal{A}_0^{\mathbb{Z}}$. A generalized diagonal $\tau: [a, b] \rightarrow P$ determines a finite bounce sequence, which we extend by 0 to a bi-infinite sequence in $\mathcal{A}_0^{\mathbb{Z}}$, so for example the sequence $(1, 2, 3)$ extends to $(\dots, 0, 0, 1, 2, 3, 0, 0, \dots)$. We are interested in certain types of elements $(b_n) \in \mathcal{A}_0^{\mathbb{Z}}$, characterized by where they are zero and nonzero. Specifically, let $I \subset \mathbb{Z}$ be such that $b_n \in \mathcal{A}$ for all $n \in I$ and $b_n = 0$ for $n \notin I$, then we say that (b_n) is *finite* if $I = [a, a']$, *forward infinite* if $I = [a, \infty)$, *backward infinite* if $I = (-\infty, a]$, and *bi-infinite* if $I = \mathbb{Z}$, for appropriate $a, a' \in \mathbb{Z}$. Thus, the bounce sequence of a generalized diagonal is a finite sequence in $\mathcal{A}_0^{\mathbb{Z}}$ in this terminology.

We topologize $\mathcal{A}_0$ with the discrete topology and $\mathcal{A}_0^{\mathbb{Z}}$ with the product topology, making it compact by Tychonoff’s theorem. We then view the full bounce spec-

trum $B(P)$ and the generalized diagonals $B_\Delta(P)$ as subspaces of $\mathcal{A}_0^{\mathbb{Z}}$. We consider their closures

$$\overline{B(P)} \subset \mathcal{A}^{\mathbb{Z}} \text{ and } \overline{B_\Delta(P)} \subset \mathcal{A}_0^{\mathbb{Z}}.$$

The closure on the left can be taken inside $\mathcal{A}_0^{\mathbb{Z}}$ instead, but since $\mathcal{A}^{\mathbb{Z}}$ is a closed subset, the closure is contained in this smaller set. In fact, these closures can be constructed in a more geometric way, as we now describe.

First take any unfolding $r: X \rightarrow P$ with symmetry group G (so that $X/G = P$) and consider the natural tiling of X by copies of P . As in the proof of the Bounce theorem, the nonsingular geodesics γ on X , up to the action of G , correspond bijectively to billiard trajectories; recording the edge labels encountered by γ on the copies of P in X produces its bounce sequence.

Next, we note that each basic geodesic also has an associated bounce sequence which records the labels from P crossed by γ in order. Here we say that γ *crosses a side e of a copy of P* if either (a) it intersects the interior of e transversely in a point, (b) it contains e as a subsegment and switches from making angle π on the left to angle π on the right (or vice versa) at the endpoints of e , or (c) it contains an endpoint of e and makes supplementary angles with that side. Condition (c) is equivalent to saying that γ contains the endpoint v of e , and that e lies on the side of γ which makes angle π at v . When $\gamma(0)$ is at a vertex, there are finitely many ways to decide which of the edge crossings (in the sense of (b) and (c)) is recorded in the index-0 position. This means that there are a uniformly bounded number of such sequences for a given γ (with the bound depending on the angles of the polygon), and they all differ by shifts.

Now consider any sequence of bounce sequences $\mathbf{b}_n$ converging in $\mathcal{A}^{\mathbb{Z}}$ to a sequence $\mathbf{b} \in \mathcal{A}^{\mathbb{Z}}$. For each $\mathbf{b}_n$ choose a bi-infinite nonsingular geodesic $\gamma_n: \mathbb{R} \rightarrow X$ having $\mathbf{b}_n$ as its associated bounce sequence and such that $\gamma_n(0)$ lies on an edge labeled by the zeroth term of $\mathbf{b}_n$. After passing to a subsequence if necessary, we can assume that γ_n converges to a basic geodesic γ in X (or more precisely, the projection to X of a basic geodesic in the universal cover). By the construction of bounce sequences for basic geodesics, $\mathbf{b}$ is one of the bounce sequences for γ (which is unique unless $\gamma(0)$ is at a vertex). Since every basic geodesic is a limit of nonsingular geodesics, it follows that $\overline{B(P)}$ consists precisely of the set of bounce sequences of basic geodesics.

One can carry out a similar construction for limits of generalized diagonals, recalling that a generalized diagonal on P is the image of a saddle connection on X . Suppose we have a sequence of bounce sequences $\mathbf{b}_n \in B_\Delta(P)$ limiting to $\mathbf{b} \in \mathcal{A}_0^{\mathbb{Z}}$. For each $\mathbf{b}_n$ consider an associated saddle connection $\gamma_n: [a, a'] \rightarrow X$, with $a \leq 0 \leq a'$, such that $\gamma(0)$ lies on an edge corresponding to the zeroth term of $\mathbf{b}_n$. Pass to a subsequence so that γ_n converges to a geodesic γ . If the lengths of γ_n are uniformly bounded, then γ_n must eventually be constant, and so $\mathbf{b}_n = \mathbf{b}$ for n sufficiently large. Since this case is trivial, we assume that the length of γ_n tends to infinity. In this case, γ is either a bi-infinite basic geodesic or a basic geodesic

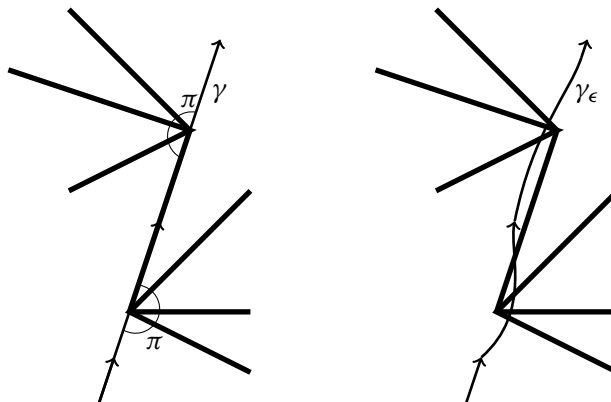


Figure 13. Replacing a singular basic geodesic γ with a perturbation γ_ϵ . The sides crossed by γ are easily read off from γ_ϵ .

ray (i.e., a subray of a basic geodesic). As above, we can associate (a finite set of) bounce sequences to γ and $\mathbf{b}$ is one of these. Thus, all sequences $\mathbf{b}$ in $\overline{\mathbf{B}_\Delta(P)}$ are bounce sequences of saddle connections, basic geodesics, or basic geodesic rays (i.e., subrays of basic geodesics emanating from a cone point).

We will need the following lemma that allows us to detect, via $\mathbf{B}_\Delta(P)$, the bounce sequences of basic geodesics that meet cone points. The only real difficulty in the proof comes from the awkward definition of the bounce sequence(s) for a basic geodesic described above, and the fact that periodic sequences arise as bounce sequences of both singular and nonsingular basic geodesics.

Lemma 25. *A sequence $\mathbf{b} = (b_n) \in \overline{\mathbf{B}(P)}$ is the bounce sequence of a singular basic geodesic if and only there exists a forward infinite (respectively, backward infinite) sequence $\mathbf{b}' = (b'_n) \in \overline{\mathbf{B}_\Delta(P)}$ and integer N such that, $b_n = b'_n$ for all $n > N$ (respectively, $n < N$).*

Proof. First choose an unfolding $r: X \rightarrow P$, and let $p: \tilde{X} \rightarrow X$ denote the universal covering. Consider the tessellation of $\tilde{X}$ by copies of P and let $\tilde{G}$ be the symmetry group, so that $\tilde{X}/\tilde{G} = P$.

Given any basic geodesic $\gamma: I \rightarrow \tilde{X}$ defined on an interval $I \subset \mathbb{R}$ and $\epsilon > 0$, we can perturb γ to a path $\gamma_\epsilon: I \rightarrow \tilde{X}$, so that $d(\gamma_\epsilon(t), \gamma(t)) < \epsilon$ and γ_ϵ transversely crosses the interiors of the edges of the tessellation of $\tilde{X}$ by copies of P according to the bounce sequence $\mathbf{b}$ of γ , and is disjoint from the cone points (if γ is nonsingular, $\gamma = \gamma_\epsilon$). A path γ_ϵ with this property is unique up to homotopy in the complement of the cone points through paths with this property; see Figure 13.

Next suppose γ with bounce sequence $\mathbf{b} = (b_n)$ and γ' with bounce sequence $\mathbf{b}' = (b'_n)$ are two basic geodesics and $b_n = b'_n$ for all $n > N$, for some integer N . Further suppose that γ is nonsingular and let γ'_ϵ be as above. As in the proof of

the Bounce theorem (see the proof of the claim), after applying an element of $\tilde{G}$, we can assume that subrays of γ and γ'_ϵ cross the same P -tiles in the same order. In particular, γ and γ'_ϵ are asymptotic in the forward direction. Since γ' and γ'_ϵ are ϵ apart, γ and γ' are forward asymptotic. Then either γ and γ' are cone point asymptotic, or there are subrays that bound a half-strip. The former cannot happen because γ contains no cone points. Therefore, γ and γ' contain subrays that bound a half-strip. In particular, their bounce sequences are necessarily periodic in the forward direction by Theorem 4. We can make a similar argument if $b_n = b'_n$ for all $n < N$.

Now we let γ be any basic geodesic with bounce sequence $\mathbf{b} = (b_n) \in \overline{\mathbf{B}(P)}$. If γ is singular, then let γ' be a basic geodesic subray of γ starting at a cone point. Its bounce sequence $\mathbf{b}' = (b'_n) \in \overline{\mathbf{B}_\Delta(P)}$ has $b_n = b'_n$ for all n sufficiently large. Next, suppose that there is a basic geodesic ray γ' emanating from a cone point with bounce sequence $\mathbf{b}' = (b'_n) \in \overline{\mathbf{B}_\Delta(P)}$ and $b_n = b'_n$ for all n sufficiently large (the argument for negative infinite sequences is similar). We suppose that γ is nonsingular, and prove that there is a singular geodesic with the same bounce sequence as γ . Extend γ' to a bi-infinite singular basic geodesic γ'' with bounce sequence $\mathbf{b}'' = (b''_n)$. Note that $b''_n = b'_n = b_n$ for all sufficiently large n . By the previous paragraph, since γ is nonsingular and γ'' is singular, it must be the case that there are subrays of each that bound a half-strip. Therefore, $\mathbf{b}$ (and hence $\mathbf{b}'$) are periodic in the forward direction by Theorem 4. This implies that γ is entirely contained in a flat strip (because it is nonsingular). The boundary geodesic γ''' of the flat strip has the same bounce sequence as γ , and is singular. This completes the proof of the lemma. $\square$

We are now ready for the proof of Corollary 1.

Corollary 1. *If two simply connected Euclidean polygons P_1, P_2 have $\mathbf{B}_\Delta(P_1) = \mathbf{B}_\Delta(P_2)$, then either P_1, P_2 are right-angled and affinely equivalent, or they are similar polygons.*

Proof. By the Bounce theorem, it is enough to show that $\mathbf{B}_\Delta(P)$ uniquely determines $\mathbf{B}(P)$. First note that any bi-infinite limit of saddle connections is easily seen to be a limit of nonsingular geodesics. Consequently, we have

$$\overline{\mathbf{B}_\Delta(P)} \cap \mathcal{A}^{\mathbb{Z}} = \overline{\mathbf{B}(P)}. \quad (4.1)$$

Each $\mathbf{b} \in \overline{\mathbf{B}(P)}$ is the bounce sequence of a basic geodesic, so it suffices to determine which $\mathbf{b} \in \overline{\mathbf{B}(P)}$ are bounce sequences of nonsingular geodesics – this is exactly $\mathbf{B}(P)$.

For this, we apply Lemma 25 and note that we can decide if $\mathbf{b} \in \overline{\mathbf{B}(P)}$ is the bounce sequence of a singular geodesic from the set $\mathbf{B}_\Delta(P)$. Any $\mathbf{b}$ that is a bounce sequence for both a singular and nonsingular geodesic is a periodic bounce sequence, as shown in the proof of Lemma 25. By a result of Gal'perin–Krüger–Troubetzkoy [19], any periodic bounce sequence is the bounce sequence of a periodic billiard trajectory, and

hence of a nonsingular geodesic. Therefore, $B(P)$ is given by removing from $\overline{B(P)}$ all bounce sequences of singular geodesics (determined by $B_\Delta(P)$), then adding back in the periodic bounce sequences that were removed in the process. $\square$

4.3. The case of rational billiards. We note that we can also deduce the following stronger version of the Bounce theorem (first proved by Bobok-Troubetzkoy [4] using different methods) under the assumption that all angles of P_1 are rational multiples of π : *If two tables P_1 and P_2 , with sides cyclically labeled by $\mathcal{A}$ have the same periodic bounce spectra, and P_1 is rational, then P_1 and P_2 are either similar or they are right-angled and affinely equivalent.* Here the *periodic bounce spectrum* refers to the subset $B_p(P) \subset B(P)$ consisting of all periodic bounce sequences. In fact, a bounce sequence is periodic if and only if the corresponding billiard trajectory is periodic, as shown in [19].

To see how we can prove this result using the methods presented here, we first explain why P_1 being rational and having the same periodic bounce spectrum as P_2 implies that P_2 is also rational. If all angles of P_1 are rational multiples of π , then we can unfold to a *translation surface* $S = X_1$ with flat metric φ_1 (that is, φ_1 has trivial holonomy). We also choose a nonpositively curved unfolding X_2 of P_2 with the same underlying topological surface S and let φ_2 be the pullback of the metric on X_2 , as in the proof of the Bounce theorem. Since X_1 is a translation surface, a result of Vorobets [43] shows that tangent vectors to core geodesics of cylinders are dense in T^1X_1 ; see also [32]. Consequently, the endpoints of the lifts of cylinder curves are dense in the support of the Liouville current L_{φ_1} . Since $B_p(P_1)$ determines the cylinder curves in X_1 , it also determines $\text{supp}(L_{\varphi_1})$, and consequently all its chains. As pointed out in [1, Section 6], the chains also determine the angles of the cone points, and hence we know all the cone points of φ_1 as well as their cone angles. Although we do not know *a priori* that endpoints of lifts of cylinder curves for φ_2 are dense in $\text{supp}(L_{\varphi_2})$, we do know that the closure of these endpoints is contained in $\text{supp}(L_{\varphi_2})$. Since $B_p(P_1) = B_p(P_2)$, it follows that the two metrics have the same set of cylinder curves, and hence $\text{supp}(L_{\varphi_1}) \subset \text{supp}(L_{\varphi_2})$.

From this one can show that a generic³ chain for the φ_1 -metric is also a chain for the φ_2 -metric, giving a $\pi_1 S$ -equivariant *injection* from the cone points of $\tilde{\varphi}_1$ to those of $\tilde{\varphi}_2$ in $\tilde{S}$. Therefore, there is a cone-angle-preserving injection from the cone points of φ_1 to those of φ_2 . By Gauss–Bonnet, this accounts for all cone points, and thus the injection is actually a bijection. Therefore, the angles of P_2 must be the same as those of P_1 and, in particular, all angles of P_2 are rational multiples of π as well.

³Chains are defined in [1] in terms of pairs of geodesics in $\text{supp}(L_\varphi)$ which are asymptotic to each other in one direction, with no other geodesics from $\text{supp}(L_\varphi)$ between them. It is conceivable that two such geodesics in the set $\text{supp}(L_{\varphi_1}) \subset \text{supp}(L_{\varphi_2})$ may nonetheless have a geodesic in $\text{supp}(L_{\varphi_2})$ between them. This can happen only countably many times (i.e., *non-generically*) as a consequence of Theorem 4, and so we may ignore these.

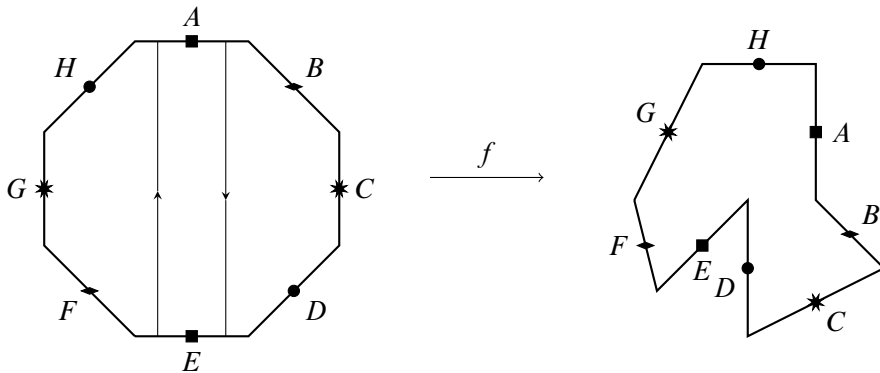


Figure 14. Two combinatorially equivalent octagons and the bijection f . There are two periodic, vertical trajectories shown in the left octagon: one has cutting sequence $\{\dots, A, A, A, \dots\}$ and the other has cutting sequence $\{\dots, E, E, E, \dots\}$.

Once we know both tables are rational, applying [43] again, we can deduce that $\text{supp}(L_{\varphi_1}) = \text{supp}(L_{\varphi_2})$. Now we can follow the proof of the Bounce theorem above verbatim to complete the argument.

4.4. Cutting sequences. We end by explaining how the application of the Support Rigidity theorem to the Bounce theorem can be extended to other settings, in particular to the context of *cutting sequences*. Recall that a cutting sequence is a symbolic coding of linear trajectories on translation surfaces, constructed by gluing polygons; see, for example, [16, 40]. More generally, suppose we have a Euclidean polygon P and we glue pairs of edges by isometries, obtaining an orientable surface S with cone metric φ . Given a cyclic labeling of the edges of P , the side pairing defines an equivalence relation on the labels. Consider a bi-infinite geodesic γ on S that avoids the vertices of P . Recording the labels crossed by γ determines a bi-infinite symbolic coding of the geodesic, which we call its cutting sequence. We note that each edge actually has two labels (since the edges are paired) and we record the one at which we enter, not the one from which we exit; see Figure 14. We note that our definition of a cutting sequence is different from the usual definition since this double labeling essentially records the *direction* in which a geodesic crosses an edge instead of just the edge it crosses.

Let P_1, P_2 be two cyclically labeled, oriented polygons with side pairings given by isometries as above. We say that P_1 and P_2 are *combinatorially equivalent* if they have the same set of labels, and if the side pairings induce the same equivalence relations; see Figure 14. We claim that if we have combinatorially equivalent polygons with the same set of cutting sequences, then the polygons can only differ by an affine deformation. To see this, note that combinatorial equivalence implies that there is a homeomorphism $P_1 \rightarrow P_2$ respecting the labelings and pairings, and hence a

homeomorphism of the resulting glued surfaces $f: S_1 \rightarrow S_2$. If the corresponding two metrics φ_1, φ_2 have all cone angles greater than 2π , then since the cutting sequences determine the bi-infinite geodesics, the Support Rigidity theorem applies as in the proof of the Bounce theorem above to produce an affine map from (S_1, φ_1) to (S_2, φ_2) that restricts to an affine homeomorphism $P_1 \rightarrow P_2$. If not, we can pass to finite-sheeted covers of S_1 and S_2 that are non-positively curved, and run the same argument there.

Theorem 26. *Let P_1, P_2 be two labeled, oriented, simply connected polygons with side pairings given by isometries. If P_1 and P_2 are combinatorially equivalent and have the same set of cutting sequences, then they are affinely equivalent.* $\square$

References

- [1] A. Bankovic and C. J. Leininger, Marked-length-spectral rigidity for flat metrics. *Trans. Amer. Math. Soc.* **370** (2018), no. 3, 1867–1884 [Zbl 1395.57020](#) [MR 3739194](#)
- [2] J. Bobok and S. Troubetzkoy, Does a billiard orbit determine its (polygonal) table? *Fund. Math.* **212** (2011), no. 2, 129–144 [Zbl 1219.37015](#) [MR 2784003](#)
- [3] J. Bobok and S. Troubetzkoy, Code and order in polygonal billiards. *Topology Appl.* **159** (2012), no. 1, 236–247 [Zbl 1259.37026](#) [MR 2852968](#)
- [4] J. Bobok and S. Troubetzkoy, Homotopical rigidity of polygonal billiards. *Topology Appl.* **173** (2014), 308–324 [Zbl 1348.37060](#) [MR 3227225](#)
- [5] F. Bonahon, Bouts des variétés hyperboliques de dimension 3. *Ann. of Math. (2)* **124** (1986), no. 1, 71–158 [Zbl 0671.57008](#) [MR 847953](#)
- [6] F. Bonahon, The geometry of Teichmüller space via geodesic currents. *Invent. Math.* **92** (1988), no. 1, 139–162 [Zbl 0653.32022](#) [MR 931208](#)
- [7] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Grundlehren Math. Wiss. 319, Springer, Berlin, 1999 [Zbl 0988.53001](#) [MR 1744486](#)
- [8] A. Calderon, S. Coles, D. Davis, J. Lanier, and A. Oliveira, How to hear the shape of a billiard table, 2018 [arXiv:1806.09644](#)
- [9] R. D. Canary, D. B. A. Epstein, and P. L. Green, Notes on notes of Thurston [[MR 0903850](#)]. In *Fundamentals of hyperbolic geometry: selected expositions*, pp. 1–115, London Math. Soc. Lecture Note Ser. 328, Cambridge Univ. Press, Cambridge, 2006 [MR 2235710](#)
- [10] C. Chabauty, Limite d’ensembles et géométrie des nombres. *Bull. Soc. Math. France* **78** (1950), 143–151 [Zbl 0039.04101](#) [MR 38983](#)
- [11] D. Constantine, Marked length spectrum rigidity in non-positive curvature with singularities. *Indiana Univ. Math. J.* **67** (2018), no. 6, 2337–2361 [Zbl 1423.53050](#) [MR 3900371](#)
- [12] C. Croke, A. Fathi, and J. Feldman, The marked length-spectrum of a surface of nonpositive curvature. *Topology* **31** (1992), no. 4, 847–855 [Zbl 0779.53025](#) [MR 1191384](#)
- [13] C. B. Croke, Rigidity for surfaces of nonpositive curvature. *Comment. Math. Helv.* **65** (1990), no. 1, 150–169 [Zbl 0704.53035](#) [MR 1036134](#)

- [14] D. Davis, Cutting sequences, regular polygons, and the Veech group. *Geom. Dedicata* **162** (2013), 231–261 [Zbl 1351.37147](#) [MR 3009542](#)
- [15] D. Davis, Cutting sequences on translation surfaces. *New York J. Math.* **20** (2014), 399–429 [Zbl 1347.37082](#) [MR 3193960](#)
- [16] D. Davis, I. Pasquinelli, and C. Ulcigrai, Cutting sequences on Bouw–Möller surfaces: An S -adic characterization, *Ann. Sci. Éc. Norm. Supér. (4)* **52** (2019), no. 4, 927–1023 [Zbl 1439.30067](#) [MR 4038456](#)
- [17] M. Duchin, C. J. Leininger, and K. Rafi, Length spectra and degeneration of flat metrics. *Invent. Math.* **182** (2010), no. 2, 231–277 [Zbl 1207.53052](#) [MR 2729268](#)
- [18] J. M. G. Fell, A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space. *Proc. Amer. Math. Soc.* **13** (1962), 472–476 [Zbl 0106.15801](#) [MR 139135](#)
- [19] G. Gal’perin, T. Krüger, and S. Troubetzkoy, Local instability of orbits in polygonal and polyhedral billiards. *Comm. Math. Phys.* **169** (1995), no. 3, 463–473 [Zbl 0924.58043](#) [MR 1328732](#)
- [20] C. Gordon, D. Webb, and S. Wolpert, Isospectral plane domains and surfaces via Riemannian orbifolds. *Invent. Math.* **110** (1992), no. 1, 1–22 [Zbl 0778.58068](#) [MR 1181812](#)
- [21] A. Hassell, L. Hillairet, and J. Marzuola, Eigenfunction concentration for polygonal billiards. *Comm. Partial Differential Equations* **34** (2009), no. 4–6, 475–485 [Zbl 1173.35620](#) [MR 2530705](#)
- [22] A. Hatcher, *Algebraic topology*. Cambridge University Press, Cambridge, 2002 [Zbl 1044.55001](#) [MR 1867354](#)
- [23] S. Hersonsky and F. Paulin, On the rigidity of discrete isometry groups of negatively curved spaces. *Comment. Math. Helv.* **72** (1997), no. 3, 349–388 [Zbl 0908.57009](#) [MR 1476054](#)
- [24] W. P. Hooper, Periodic billiard paths in right triangles are unstable. *Geom. Dedicata* **125** (2007), 39–46 [Zbl 1131.37038](#) [MR 2322537](#)
- [25] P. Hubert, Complexité de suites définies par des billards rationnels. *Bull. Soc. Math. France* **123** (1995), no. 2, 257–270 [Zbl 0836.58013](#) [MR 1340290](#)
- [26] P. Hubert and L. Vuillon, Complexity of cutting words on regular tilings. *European J. Combin.* **28** (2007), no. 1, 429–438 [Zbl 1111.68095](#) [MR 2261829](#)
- [27] C. C. Johnson, Cutting sequences on square-tiled surfaces. *Geom. Dedicata* **190** (2017), 53–80 [Zbl 1406.37038](#) [MR 3704812](#)
- [28] M. Kac, Can one hear the shape of a drum? *Amer. Math. Monthly* **73** (1966), no. 4, Part II, 1–23 [Zbl 0139.05603](#) [MR 201237](#)
- [29] A. Katok, The growth rate for the number of singular and periodic orbits for a polygonal billiard. *Comm. Math. Phys.* **111** (1987), no. 1, 151–160 [Zbl 0631.58020](#) [MR 896765](#)
- [30] L.-M. Lopez and P. Narbel, Generalized Sturmian languages. In *Automata, languages and programming (Szeged, 1995)*, pp. 336–347, Lecture Notes in Comput. Sci. 944, Springer, Berlin, 1995 [Zbl 1412.68144](#) [MR 1466480](#)
- [31] Z. Lu and J. Rowlett, The sound of symmetry. *Amer. Math. Monthly* **122** (2015), no. 9, 815–835 [Zbl 1344.58004](#) [MR 3418203](#)
- [32] H. Masur, Closed trajectories for quadratic differentials with an application to billiards. *Duke Math. J.* **53** (1986), no. 2, 307–314 [Zbl 0616.30044](#) [MR 850537](#)

- [33] H. Masur and J. Smillie, Hausdorff dimension of sets of nonergodic measured foliations. *Ann. of Math. (2)* **134** (1991), no. 3, 455–543 [Zbl 0774.58024](#) [MR 1135877](#)
- [34] H. Masur and S. Tabachnikov, Rational billiards and flat structures. In *Handbook of dynamical systems, Vol. 1A*, pp. 1015–1089, North-Holland, Amsterdam, 2002 [Zbl 1057.37034](#) [MR 1928530](#)
- [35] M. Morse and G. A. Hedlund, Symbolic Dynamics. *Amer. J. Math.* **60** (1938), no. 4, 815–866 [Zbl 64.0798.04](#) [MR 1507944](#)
- [36] M. Morse and G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62** (1940), 1–42 [Zbl 0022.34003](#) [MR 745](#)
- [37] J.-P. Otal, Le spectre marqué des longueurs des surfaces à courbure négative. *Ann. of Math. (2)* **131** (1990), no. 1, 151–162 [Zbl 0699.58018](#) [MR 1038361](#)
- [38] R. E. Schwartz, Obtuse triangular billiards. I. Near the $(2, 3, 6)$ triangle. *Experiment. Math.* **15** (2006), no. 2, 161–182 [Zbl 1112.37030](#) [MR 2253003](#)
- [39] J. Smillie and C. Ulcigrai, Geodesic flow on the Teichmüller disk of the regular octagon, cutting sequences and octagon continued fractions maps. In *Dynamical numbers – interplay between dynamical systems and number theory*, pp. 29–65, Contemp. Math. 532, Amer. Math. Soc., Providence, RI, 2010 [Zbl 1222.37012](#) [MR 2762132](#)
- [40] J. Smillie and C. Ulcigrai, Beyond Sturmian sequences: coding linear trajectories in the regular octagon. *Proc. Lond. Math. Soc. (3)* **102** (2011), no. 2, 291–340 [Zbl 1230.37021](#) [MR 2769116](#)
- [41] W. Thurston, Geometry and topology of 3-manifolds, Lecture Notes, Princeton University, 1986. Available at: <http://www.msri.org/publications/books/gt3m>
- [42] S. Troubetzkoy, Complexity lower bounds for polygonal billiards. *Chaos* **8** (1998), no. 1, 242–244 [Zbl 0986.37030](#) [MR 1609769](#)
- [43] Y. Vorobets, Periodic geodesics on generic translation surfaces. In *Algebraic and topological dynamics*, pp. 205–258, Contemp. Math. 385, Amer. Math. Soc., Providence, RI, 2005 [Zbl 1130.37015](#) [MR 2180238](#)

Received March 2, 2020

M. Duchin, Department of Mathematics, Tufts University,
503 Boston Avenue, Bromfield-Pearson, Medford, MA 02155, USA

E-mail: moon.duchin@tufts.edu

V. Erlandsson, School of Mathematics, University of Bristol,
Fry Building, Woodland Road, Bristol BS8 1UG, UK; and
Department of Mathematics and Statistics,

UiT The Arctic University of Norway, 9037 Tromsø, Norway

E-mail: v.erlandsson@bristol.ac.uk

C. J. Leininger, Department of Mathematics, Rice University,
6100 Main Street, Houston, TX 77005, USA

E-mail: cjl12@rice.edu

C. Sadanand, Department of Mathematics, Bowdoin College,
255 Maine Street, Brunswick, ME 04011, USA

E-mail: c.sadanand@bowdoin.edu