

A CAT(0)-APPROACH TO THE MARKED LENGTH SPECTRAL RIGIDITY OF SINAI BILLIARDS

DOUGLAS FINAMORE AND MARTIN LEGUIL

ABSTRACT. We study the (marked) length spectral rigidity problem for Sinai billiards with finite horizon on the two-torus: to what extent do the lengths of (marked) periodic trajectories determine the geometry of the table? Since the billiard flow is neither smooth nor structurally stable, the natural marking is provided by the surface K obtained by gluing two copies of the table along their boundary. We introduce the *enriched marked length spectrum* $\mathbf{EL}$, which assigns to each free homotopy class of K the length of its shortest representative for the intrinsic metric of the table; it records the billiard periodic orbits together with the “gliding” periodic trajectories that slide along the scatterers. Our main result is that two finite-horizon Sinai billiards with the same enriched marked length spectrum are isometric. Following Kourganoff, the billiard flow is approximated by the Anosov geodesic flows of a family of Riemannian metrics on K degenerating along the gluing locus. The limit is a CAT(0) space on which we adapt Otal’s argument for negatively curved surfaces, extending marked length spectral rigidity beyond the Riemannian setting.

CONTENTS

1. Introduction	1
2. Preliminaries on Sinai billiards	6
3. Geodesic approximations	11
4. Spectral Rigidity	40
Appendix A. Generic properties of periodic orbits of Sinai billiards	62
Appendix B. Billiard cycles and the enriched length functional	69
References	73

1. INTRODUCTION

We consider dynamical systems with hyperbolic behaviour arising from geometry, such as geodesic flows on negatively curved Riemannian manifolds or dispersing billiards. A central question is how to parametrise the moduli space of such systems. Specifically, given a Riemannian manifold equipped with a negatively curved metric, one seeks to describe the set of all such metrics up to the action of the diffeomorphism group. Similarly, for billiards, the goal is to understand the set of isometries within a fixed diffeomorphism class of dispersing planar billiards. These systems typically exhibit a dense set of periodic orbits (restricted to the non-wandering set in the non-compact case), making periodic data a natural choice for moduli. The information recorded for periodic orbits can be categorised as follows:

1991 *Mathematics Subject Classification.* 37D40, 37C83, 51F30.

Key words and phrases. spectral rigidity, CAT(0) spaces, hyperbolic billiards.

- **Topological or symbolic data.** On negatively curved manifolds, and more generally, for metrics with Anosov geodesic flow, it is well-known (see e.g. [35]) that there is a unique closed geodesic in each free homotopy class. For dispersing billiards, such as open dispersing billiards or Sinai billiards, the symbolic data naturally assigned to each periodic orbit is the sequence of scatterers encountered by the trajectory.
- **Length data.** The *length spectrum* is the collection of all lengths (or periods) of closed orbits.
- **Lyapunov exponents.** Each periodic orbit in a hyperbolic system is associated with exponents measuring the exponential divergence rate of nearby orbits. For 3-dimensional volume preserving hyperbolic flows, this reduces to a single exponent.

A natural question that arises is: to what extent do periodic orbits characterise the system? A dynamical analogue of Mark Kac’s famous question, “Can one hear the shape of a drum?”, is the following inverse problem: does the length spectrum encode the system’s geometry? Unfortunately, this problem is typically intractable due to the lack of structure in the length spectrum. Even if two systems share the same length spectrum, one cannot *a priori* rule out the possibility that a given length corresponds to periodic orbits with very different behaviour in each system; we refer e.g. to [44], where examples of isospectral yet non-isometric manifolds are constructed. Matching data between such systems is therefore generally a hard task. To impose more structure on the spectrum, one usually augments it with symbolic data, yielding a *marked length spectrum*. Informally, this can be viewed as a (set-valued) function

$$\text{symbolic data } \omega \mapsto \{\text{lengths of all periodic orbits with symbol } \omega\}.$$

In the hyperbolic setting, the expansiveness of the dynamics typically reduces the set on the right-hand side to a single positive number. For certain systems, such as geodesic flows on negatively curved Riemannian manifolds or open dispersing billiards satisfying the no-eclipse condition (see e.g. [38]), the dynamics is structurally stable. Consequently, the symbolic data does not depend on the choice of negatively curved metric for a given Riemannian manifold, nor on the specific geometry of the open dispersing billiard. This defines a natural marked length spectrum, allowing us to rephrase the previous question as follows: does the marked length spectrum fully determine the system up to isometry?

The question of marked length spectral determination for negatively curved metrics on surfaces was answered affirmatively and simultaneously by Otal [39] and Croke [18] in their celebrated work. Significant progress has recently been made in generalising these results in several directions. In higher dimensions, Guillarmou and Lefeuvre [28] proved local marked length spectral determination for metrics of non-positive curvature with Anosov geodesic flow. In the same work, they also provided new stability estimates, quantifying the extent to which the marked length spectrum controls the distance between metrics. Even more recently, for surfaces of genus at least 2, Guillarmou, Lefeuvre, and Paternain [29] were able to drop the negative curvature assumption and establish global marked length spectral determination. More precisely, they showed that two metrics with Anosov geodesic flows on a given surface that share the same marked length spectrum are isometric via an isometry isotopic to the identity.

A related framework is that of flat metrics, more precisely, non-positively curved Euclidean cone metrics, where curvature is concentrated at conical singularities. In this context, Bankovic and Leininger [7] showed that two flat metrics on a closed,

oriented surface that assign the same lengths to all closed curves differ by an isometry isotopic to the identity. It is also worth noting that the question of marked length spectral rigidity has been investigated in different contexts. For instance, recent work by Abbondandolo and Mazzucchelli [1] establishes that S^1 -symmetric metrics on the 2-sphere with a single equator and the same marked length spectrum have conjugate geodesic flows, and coincide under an additional $\mathbb{Z}_2$ -symmetry; here the dynamics is integrable, which sets it apart from the hyperbolic geodesic flows discussed earlier.

Spectral rigidity of billiards has been studied for various classes of domains. De Simoi, Kaloshin and Wei [21] showed that convex domains with $\mathbb{Z}_2$ -symmetry and that are sufficiently close to a circle exhibit dynamical (deformational) spectral rigidity. In [20], it is shown that open dispersing billiards with analytic boundaries satisfying the no-eclipse condition and with $\mathbb{Z}_2 \times \mathbb{Z}_2$ -symmetry can be recovered from their marked length spectral data. The dynamics of such billiards is hyperbolic, aligning them more closely with Anosov geodesic flows. These billiards are structurally stable and admit natural symbolic dynamics [38]. However, the non-wandering dynamics occurs only on a Cantor set, and analyticity is required to globally determine the domain.

In this paper, we investigate the case of Sinai billiards, obtained by removing from the 2-torus $\mathbb{T}^2$ a finite number of strongly convex scatterers. While the dynamics of Sinai billiards shares many features with Anosov geodesic flows on surfaces (such as hyperbolicity, ergodicity, and exponential mixing, see e.g. [15, 4]), it also introduces several additional challenges:

- the billiard map has a dense set of singularities, and the billiard flow is not Anosov;
- the system is not structurally stable: perturbing the geometry of the boundary can create or destroy periodic orbits.

A symbolic coding for Sinai billiards does arise naturally: by considering a $\mathbb{Z}^2$ -periodic lift of the billiard table to $\mathbb{R}^2$ and labelling each scatterer, each orbit can be assigned a symbolic sequence consisting of the labels of the obstacles it successively encounters. However, due to the lack of structural stability, the question of marked length spectral determination for Sinai billiards may seem less canonical than for Anosov metrics on surfaces. The marking is not *a priori* stable under perturbation, and the symbolic data itself must encode substantial geometric information. In particular, if we discard the length data and collect all symbolic data corresponding to sequences realised by (periodic) billiard orbits (the realisable sequences of Section 2) into a *symbolic spectrum*, we can even pose the following stronger question:

Question 1.1. Does the symbolic spectrum determine a Sinai billiard up to isometry?

While we are unable to answer this question for Sinai billiards, we note that several works address similar questions for flat billiards (see e.g. [22, 14]). In the present work, rather than focusing on the marked length spectrum corresponding to periodic orbits of a Sinai billiard, we take a different approach. Specifically, we build on a construction suggested by Birkhoff and Arnold and made explicit in the work of Kourganoff [36]; it demonstrates that the billiard flow on a Sinai billiard can, in a certain sense, be approximated by a one-parameter family of Anosov geodesic flows on a surface lifting the table. This allows us to define an *enriched length spectrum*, obtained as the “limit” of the marked length spectra of the approximating geodesic flows. Although this construction introduces periodic data that do not correspond to actual billiard orbits, it renders the marking more natural by leveraging the symbolic data shared across the family of approximating geodesic flows.

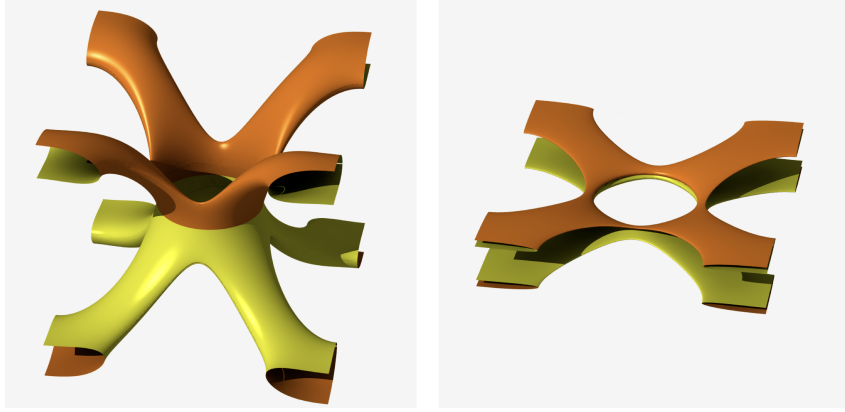


FIGURE 1. A surface K lifting the table $\mathcal{D}$ and a deformation K_ε (Credits: Mickaël Kourganoff, Jos Leys).

Let us quickly recall the construction in [36]. Fix a finite-horizon Sinai billiard $\mathcal{D} \subset \mathbb{T}^2$. Let $E = \mathbb{T}^2 \times \mathbb{R}$, and denote by $p: E \rightarrow \mathbb{T}^2$ the canonical projection onto the first two coordinates. We consider a smooth closed immersed surface K in E such that $p(K) = \mathcal{D}$, and such that $K^\circ := K \setminus p^{-1}(\partial\mathcal{D})$ has two connected components each of which is diffeomorphically mapped to $\text{int}(\mathcal{D})$ under p . We interpret K as a surface obtained by gluing two different copies of $\mathcal{D}$ along their boundary components. Let us denote by $\mathcal{C} := \text{Conj}(\pi_1(K))$ the set of free homotopy classes on K . Note that $\mathcal{C}$ is independent of the choice of surface K lifting $\mathcal{D}$ that satisfies our requirements.

For $\varepsilon \in (0, 1]$, we let α_ε be the contraction given by $(x, y, z) \mapsto (x, y, \varepsilon z)$ and write g_ε to denote the metric induced on $K_\varepsilon := \alpha_\varepsilon(K)$ by the Euclidean metric. Pulling back via α_ε , we will consider $\{g_\varepsilon\}$ as a family of metrics on the surface K . We denote by $\mathbf{g}_\varepsilon$ the geodesic flow of the surface (K, g_ε) . By Kourganoff [36], for $\varepsilon_0 > 0$ small enough, the geodesic flows $\{\mathbf{g}_\varepsilon\}_{0 < \varepsilon < \varepsilon_0}$ associated to the family of metrics $\{g_\varepsilon\}_{0 < \varepsilon < \varepsilon_0}$ are all Anosov. Therefore, for $0 < \varepsilon < \varepsilon_0$, there exists a unique closed geodesic $\gamma_\varepsilon^\alpha$ of (K, g_ε) in each free homotopy class $\alpha \in \mathcal{C}$. We will see in Lemma 3.39 that the family $\{p(\gamma_\varepsilon^\alpha)\}_{0 < \varepsilon < \varepsilon_0}$ converges in the Hausdorff topology to some closed curve γ_0^α in $\mathcal{D}$ as $\varepsilon \rightarrow 0$, and their respective lengths $\{l_\varepsilon(\gamma_\varepsilon^\alpha)\}_{0 < \varepsilon < \varepsilon_0}$ also converge to a limiting length $l_0(\gamma_0^\alpha)$. We can then define the *enriched marked length spectrum* of $\mathcal{D}$ as the function

$$\begin{aligned} \mathbf{EL}_\mathcal{D}: \mathcal{C} &\rightarrow \mathbb{R} \\ \alpha &\mapsto l_0(\gamma_0^\alpha). \end{aligned}$$

Remark 1.2. Every periodic billiard orbit of $\mathcal{D}$ is the projection of γ_0^α for some $\alpha \in \mathcal{C}$: its lift to K , which changes sheet at each collision, is a closed geodesic of K_0 (after two periods if the number of collisions is odd), hence the unique closed geodesic of its class (Lemma 3.38); orbits with only regular collisions give rise to exactly two classes, exchanged by the sheet involution (Remark 3.27). Consequently, $\mathbf{EL}_\mathcal{D}$ can be interpreted as a natural completion of the marked length spectrum of $\mathcal{D}$, which is the restriction of $\mathbf{EL}_\mathcal{D}$ to those $\alpha \in \mathcal{C}$ associated with genuine billiard orbits.

Remark 1.3. As we shall see (e.g. in Section 3), there are multiple ways to rigorously define the “limit” of the family of Anosov geodesic flows $\{\mathbf{g}_\varepsilon\}_{0 < \varepsilon < \varepsilon_0}$ as $\varepsilon \rightarrow 0$, all of which are equivalent. Although the resulting limit object is no longer an Anosov flow, it retains several key features inherited from the limiting CAT(0) spaces. We also emphasise that the enriched spectrum $\mathbf{EL}_\mathcal{D}$ is independent of the specific choice of the surface K lifting $\mathcal{D}$, hence is intrinsic to the billiard table $\mathcal{D}$ (see e.g. Appendix B).

Let $\mathcal{D}, \mathcal{D}' \subset \mathbb{T}^2$ be two Sinai billiards with finite horizon and C^r boundaries, $r \geq 3$ (see Section 2), that are diffeomorphic to each other. We can choose surfaces K, K' lifting $\mathcal{D}, \mathcal{D}'$ which are diffeomorphic to each other; a diffeomorphism $K \rightarrow K'$ identifies the sets of free homotopy classes of K and K' (in other words, it is a *marking*), and the equality of spectra in the statement below is understood with respect to such an identification, fixed once and for all (see Section 4.3). Our main result is:

Theorem A. *Let $\mathcal{D}, \mathcal{D}' \subset \mathbb{T}^2$ be two Sinai billiards with finite horizon and C^r boundaries, $r \geq 3$, that are diffeomorphic to each other. If the billiards $\mathcal{D}, \mathcal{D}'$ share the same enriched marked length spectrum, i.e., $\mathbf{EL}_{\mathcal{D}} = \mathbf{EL}_{\mathcal{D}'}$, then $\mathcal{D}$ and $\mathcal{D}'$ are actually isometric. In other words, there is an isometry φ of $\mathbb{T}^2$ with $\varphi(\mathcal{D}) = \mathcal{D}'$.*

Outline of the proof. The strategy is to reduce the marked length spectral rigidity of the billiard to a boundary rigidity statement for a limiting CAT(0) space, where an argument in the spirit of Otal [39] can be applied. It proceeds in four main steps.

From billiards to a degenerating family of Anosov surfaces. Following Kourganoff [36], each Sinai billiard $\mathcal{D}$ is lifted to a surface K whose geodesic flow, after a vertical contraction by a factor ε , is Anosov and shadows in a certain sense the billiard flow (Section 3, Theorems 3.1 and 3.30). As $\varepsilon \rightarrow 0$, the Riemannian metrics g_ε degenerate along the gluing locus $B(K)$ which lifts the boundary of the billiard $\mathcal{D}$; the central technical difficulty, addressed throughout Section 3, is that no estimate uniform in ε survives at $B(K)$. We control the degeneration through a boundary-layer normal form (Lemma 3.13) and a uniform Gromov-hyperbolicity statement for the family of geodesic flows (Lemma 3.22).

The limiting CAT(0) space. The distances d_ε decrease monotonically to a limit length metric d_0 (Theorem 3.6), which realises K_0 as two copies of the table $(\mathcal{D}, d_{\mathcal{L}})$ glued along their boundaries (Theorem 3.12). The universal cover $\widetilde{K}_0$ is a Hadamard space; the dispersing billiard dynamics forbids flat strips (Lemma 3.31), so $\widetilde{K}_0$ is Gromov-hyperbolic (Corollary 3.32) and its geodesics are determined by their endpoints at infinity (Lemma 3.33). The comparison of lengths (Theorem 3.16) shows that the g_ε -geodesics are quasi-geodesics of $\widetilde{K}_1$ with constants independent of ε , which is what transfers the coarse geometry across the limit (Section 3.2).

The Liouville current and the correspondence of geodesics. On $\widetilde{K}_0$ there is no Liouville form; instead, $\widetilde{K}_0$ carries an intrinsic Liouville geodesic current $\widehat{m}_0$, defined by the kinematic measure of its flat pieces (Definition 4.3), whose support is exactly the set of admissible geodesics (Proposition 4.14) and whose values on rectangles of endpoint pairs are the logarithms of the cross-ratios cr_0 at infinity (Proposition 4.15). We show that the enriched marked length spectrum $\mathbf{EL}_{\mathcal{D}} = \mathbf{EL}_{\mathcal{D}'}$ determines all cross-ratios (Proposition 4.25), hence the current. The identification of the boundaries at infinity induced by the marking then gives a homeomorphism h between the spaces of (unparametrised) geodesics of $\widetilde{K}_0$ and $\widetilde{K}'_0$ which preserves the currents (Lemma 4.27). At this step, the map h is dynamical and not geometric yet, i.e., it acts on geodesics, not on points.

From angle preservation to isometry. An averaging argument on the angle distortion of h , in the spirit of Otal [39] and Wilkinson [46], shows that h is angle-preserving (Theorem 4.35). There, the product structure of the Liouville measure on pairs of crossing geodesics replaces the use of ergodicity. Since the curvature of $\widetilde{K}_0$ is concentrated on the gluing locus $B(K)$, angle preservation is exploited along the walls: it forces the images of the geodesics through a wall point to be concurrent,

which defines an isometric map between the gluing loci. The current then controls the crossing positions along every geodesic, so that h is induced by an isometry $\widetilde{K}_0 \rightarrow \widetilde{K}'_0$ (Theorem 4.37); this isometry descends to the tables and extends to an isometry of $\mathbb{T}^2$.

Natural questions then arise.

Question 1.4. Does partial information from the enriched marked length spectrum suffice to recover this conclusion? In particular, to what extent does the set of periodic billiard orbits encode the geometry of the billiard $\mathcal{D}$? Given our main result Theorem A, this amounts to asking whether the billiard marked length spectrum encodes the enriched spectrum $\mathbf{EL}_{\mathcal{D}}$.

Question 1.5. It is natural to ask whether the enriched spectrum $\mathbf{EL}_{\mathcal{D}}$ can be detected by an analytic spectrum. The relevant operator is naturally attached to the singular surface K_0 obtained by gluing two copies of $\overline{\mathcal{D}}$ along the walls: the Laplace spectrum of the singular surface K_0 is the union of the Dirichlet and Neumann spectra of the table $\mathcal{D}$. For the table itself, the Poisson relation of Andersson–Melrose [3] (see also [41]) locates the singularities of the wave trace inside the length spectrum of generalised geodesics in the sense of Melrose–Sjöstrand. Since dispersing walls are everywhere diffractive, the generalised geodesics of a Sinai table contain no gliding arcs, and generically (no grazing periodic rays, simple non-degenerate lengths [41], see also Proposition 2.7) the wave trace of $\mathcal{D}$ detects exactly the (unmarked) billiard length spectrum. On the other hand, classes whose limit representatives γ_0^α contain gliding arcs exist for every finite-horizon table (see Corollary 3.40 and Lemma 3.25). These classes are not seen by the billiard’s wave trace: the enrichment is precisely the part of $\mathbf{EL}_{\mathcal{D}}$ that the classical Poisson relation cannot see. We expect the situation to change on K_0 : the seam of K_0 is a closed geodesic carrying concentrated negative curvature, and the gliding representatives are genuine closed geodesics of K_0 . By analogy with the diffractive trace formulas at conical singularities [31], it seems possible that the wave trace of K_0 is singular at the gliding lengths as well, but we do not pursue it here. Combined with Theorem A (and with the local constancy of the marking along deformations), such a trace formula would yield spectral deformation-rigidity of generic Sinai billiards for the pair of Dirichlet and Neumann boundary conditions.

2. PRELIMINARIES ON SINAI BILLIARDS

We begin by introducing some established facts and notation on Sinai billiards. Most of the standard facts recalled in this section can be found in [15]. Let $\mathcal{D} \subset \mathbb{T}^2$ be a Sinai billiard table with finite horizon whose boundary $\partial\mathcal{D}$ is of class C^r , $r \in \mathbb{N}_{\geq 3}$. More precisely, we have

$$\partial\mathcal{D} = \Gamma_1 \cup \Gamma_2 \cup \cdots \cup \Gamma_m, \quad m \geq 2,$$

where each Γ_i is a C^r curve bounding a strongly convex open set $\mathcal{O}_i \subset \mathbb{T}^2$, called an obstacle, and $\mathcal{D} := \mathbb{T}^2 \setminus \bigcup_{i=1}^m \mathcal{O}_i$.

For each $i \in \{1, \dots, m\}$, we denote by $\ell_i > 0$ the perimeter of $\mathcal{O}_i$, we let $\mathbb{T}_i := \mathbb{R}/\ell_i\mathbb{Z}$, and let $\gamma_i: \mathbb{T}_i \rightarrow \mathbb{T}^2$ be an anticlockwise parametrisation of Γ_i by arc-length. We set $\mathbb{T}_{\mathcal{D}} := \bigcup_{i=1}^m \mathbb{T}_i$, and let $\gamma: \mathbb{T}_{\mathcal{D}} \rightarrow \mathbb{T}^2$, $\mathbb{T}_i \ni s \mapsto \gamma_i(s)$. For each $s \in \mathbb{T}_{\mathcal{D}}$, we denote by n_s the unit outward normal vector at the point $\gamma(s) \in \partial\mathcal{D}$, and by $\mathcal{K}(s)$ the curvature at this point; following the convention in [15], we assume that $\mathcal{K}(s) > 0$.

We let $\mathcal{M} := \mathbb{T}_{\mathcal{D}} \times [-\frac{\pi}{2}, \frac{\pi}{2}]$, and denote by $f: \mathcal{M} \rightarrow \mathcal{M}$, $(s, \phi) \mapsto (s', \phi')$ the billiard map within $\mathcal{D}$, where s corresponds to the position on the boundary, and ϕ is the oriented angle from the inward normal vector $-n_s$ to the velocity vector at $\gamma(s)$.

Moreover, $\mathcal{S}_0 := \partial\mathcal{M} = \{(s, \phi); |\phi| = \pi/2\}$ is the set of *grazing* collisions of f ; a collision which is not grazing is called *regular*. Set

$$\mathcal{S}_{n+1} := \mathcal{S}_n \cup f^{-1}(\mathcal{S}_n) \text{ and } \mathcal{S}_{-(n+1)} := \mathcal{S}_{-n} \cup f(\mathcal{S}_{-n}).$$

Then, for every $n \in \mathbb{N}$, $\mathcal{S}_{n+1}$, resp. $\mathcal{S}_{-(n+1)}$ is the singularity set of the map f^{n+1} , resp. $f^{-(n+1)}$, while these maps are of class C^{r-1} at $\mathcal{M} \setminus (\mathcal{S}_{n+1} \cup \mathcal{S}_{-(n+1)})$. The points in $\mathcal{S}_{\mathcal{M}} := \cup_{i \in \mathbb{Z}} \mathcal{S}_i$ are called the *singular points* of the billiard map, and $\mathcal{M}^{\text{reg}} := \mathcal{M} \setminus \mathcal{S}_{\mathcal{M}}$ is the *regular set* of f . The singular set has empty interior, but is dense in $\mathcal{M}$, while the regular set has full Lebesgue measure, but is badly disconnected. A periodic orbit of f is called *singular* if it has a grazing collision, i.e., if it meets $\mathcal{S}_0$ (equivalently, $\mathcal{S}_{\mathcal{M}}$), and *regular* otherwise.

Finally, for each $(s, \phi) \in \mathcal{M}$, its free-flight time is the function

$$\tau(s, \phi) = \|\gamma(\text{Pr}_1(f(s, \phi))) - \gamma(s)\|,$$

where we denote by $\|\cdot\|$ the Euclidean distance in $\mathbb{T}^2$ and by $\text{Pr}_1: \mathcal{M} \rightarrow \partial\mathcal{D}$ the projection onto the first coordinate $s \simeq \gamma(s)$. The *finite horizon* hypothesis means that no trajectory of $\mathcal{D}$ avoids the obstacles forever; equivalently, every line of $\mathbb{R}^2$ meets a lifted obstacle. By compactness, there is then a uniform upper bound $\tau_{\max}$ on the time it takes for any trajectory, regardless of its starting point, to experience a collision. Since the obstacles are pairwise disjoint and closed, there is also a uniform lower bound $\tau_{\min} > 0$ on the free flight between two consecutive collisions, so that $\tau_{\min} \leq \tau(x) \leq \tau_{\max}$ for every $x \in \mathcal{M}$. The suspension flow of f with roof function τ is called the billiard flow Ψ of $\mathcal{D}$. Its phase space is denoted by Ω , and the usual coordinates are (x, y, ω) , where (x, y) are coordinates on $\mathcal{D}$ and ω is the angle between the velocity of the trajectory and the x -axis. The saturations of $\mathcal{S}_{\mathcal{M}}$ and $\mathcal{M}^{\text{reg}}$ under the flow form the singular and regular set of Ω , denoted by $\mathcal{S}_{\Omega}$ and Ω^{reg} . It is known that the billiard flow preserves the Lebesgue measure $dx dy d\omega$ and the contact form $\lambda = \cos \omega dx + \sin \omega dy$ [15]. Indeed, the flow is generated by the Reeb field of such form.

Neither the coordinates nor the hyperbolic structure recalled below are used in the sequel; we record them for the reader's convenience. The Jacobi coordinates (η, ξ, ω) , introduced in [47], are defined as

$$\begin{aligned} \eta &= x \cos \omega + y \sin \omega, \\ \xi &= x \sin \omega - y \cos \omega. \end{aligned}$$

In particular, while in the usual coordinates the action of the flow between collisions is $\Psi^t(x, y, \omega) = (x + t \cos \omega, y + t \sin \omega, \omega)$, in Jacobi coordinates this action is simply translation on the first coordinate:

$$\Psi^t(\eta, \xi, \omega) = (\eta + t, \xi, \omega).$$

Inverting the change of coordinates gives $x = \eta \cos \omega + \xi \sin \omega$ and $y = \eta \sin \omega - \xi \cos \omega$. Hence the forms

$$d\eta + \xi d\omega = \cos \omega dx + \sin \omega dy = \lambda$$

and

$$\lambda \wedge d\lambda = d\eta \wedge d\xi \wedge d\omega = -dx \wedge dy \wedge d\omega$$

are both preserved by the flow. In Jacobi coordinates, the Reeb field of λ is exactly the coordinate field ∂_η , which corresponds to the displacement of a particle in the direction of the flow.

The map f and the flow Ψ are (uniformly) hyperbolic: at every regular point there is an invariant splitting into stable and unstable directions (together with the flow direction ∂_η in the case of the flow), with uniform contraction and expansion rates; the associated Lyapunov exponents are defined almost everywhere by Oseledets' theorem. This gives a decomposition

$$T_x\Omega = T_x^c\Omega \oplus T_x^\perp\Omega = \mathbb{R}\partial_\eta \oplus E^s(x) \oplus E^u(x),$$

of the tangent space of every regular point, where $E^s(x), E^u(x)$ are the stable and unstable bundles, respectively. As usual, $E^s(x)$, resp. $E^u(x)$, contracts, resp. expands exponentially under the flow as $t \rightarrow \infty$, and is locally integrable to C^{r-1} curves, called the strong stable (resp. unstable) local manifolds $\mathcal{W}_{\text{loc}}^s(x)$ (resp. $\mathcal{W}_{\text{loc}}^u(x)$). The end points of the manifolds $\mathcal{W}_{\text{loc}}^s(x)$ (not included in the manifolds) are always singular. These manifolds are not uniform in length, and get smaller as one approaches the singular set. In particular, every open set of Ω contains points whose unstable manifolds are arbitrarily small, and therefore points in Ω do not have product neighbourhoods. The saturations of $\mathcal{W}_{\text{loc}}^s(x)$ and $\mathcal{W}_{\text{loc}}^u(x)$ with respect to the flow are the weak, or central, stable and unstable manifolds, respectively, denoted by $\mathcal{W}_{\text{loc}}^{cs}(x)$ and $\mathcal{W}_{\text{loc}}^{cu}(x)$.

2.1. Liftings and the symbolic space. We want to associate with each orbit of f a sequence of symbols that represents the scatterers it encounters along the way. However, this approach is problematic on the torus because an orbit could hit the same scatterer several times in succession. To avoid this issue, let $\tilde{\mathcal{D}} \subset \mathbb{R}^2$ be the $\mathbb{Z}^2$ -periodic billiard table such that $\pi(\tilde{\mathcal{D}}) = \mathcal{D}$, where $\pi: [0, 1]^2 \rightarrow \mathbb{T}^2$ denotes the canonical projection. Abusing notation, we denote the billiard map and flow on $\tilde{\mathcal{D}}$ by f and Ψ as well.

Let $\{B_1, \dots, B_k\}$ be the set of connected components of the lift $\pi^{-1}(\partial\mathcal{D}) \subset [0, 1]^2$. We index all the scatterers in the lifted table on $\mathbb{R}^2$ using the alphabet $\mathcal{A} := \mathbb{Z}^2 \times \{1, \dots, k\}$, as follows:

$$B_{(i,j;l)} := B_l + (i, j),$$

i.e., $B_{(i,j;l)}$ is the unique scatterer obtained by translating B_l by $(i, j) \in \mathbb{Z}^2$. Let Σ denote the space $\mathcal{A}^{\mathbb{Z}}$ of two-sided sequences of elements in $\mathcal{A}$. This is a complete metric space with the usual distance

$$d_\Sigma(\alpha, \beta) = \sum_{i \in \mathbb{Z}} 2^{-|i|} (1 - \delta_{a_i b_i})$$

between two sequences $\alpha = \{a_i\}_{i \in \mathbb{Z}}$ and $\beta = \{b_i\}_{i \in \mathbb{Z}}$. In what follows, we will work with the equivalent metric ρ_Σ given by

$$\rho_\Sigma(\alpha, \beta) := 2^{-\max\{n \in \mathbb{N}; a_i = b_i \text{ for all } |i| \leq n\}}.$$

We associate to the points in $\mathcal{M}$ the sequence of indices of scatterers where its lifted trajectory bounces.

Definition 2.1. The map $\iota: \mathcal{M} \rightarrow \Sigma$ is given by

$$\iota(s, \phi) = \{a_i\}_{i \in \mathbb{Z}},$$

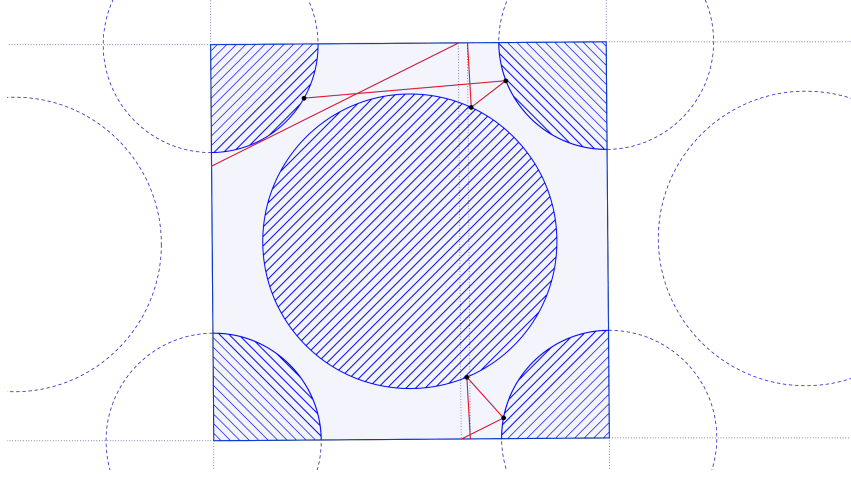


FIGURE 2. Lifted table.

where $B_{a_0} \subset \pi^{-1}(\partial\mathcal{D})$, and for all $i \in \mathbb{Z}$, a_i is the unique element of $\mathcal{A}$ such that $\text{Pr}_1(f^i(\pi^{-1}(s), \phi)) \in B_{a_i}$. We call the set $\Sigma_{\mathcal{M}} := \iota(\mathcal{M})$ the *space of realisable sequences*.

The space Σ admits a natural $\mathbb{Z}^2$ -action, defined by $\{(i_p, j_p; l_p)\}_{p \in \mathbb{Z}} \mapsto \{(i_p + i, j_p + j; l_p)\}_{p \in \mathbb{Z}}$, for all $(i, j) \in \mathbb{Z}^2$. We denote by $\hat{\Sigma}$ the quotient of Σ under this action, and by $\hat{\iota}: \mathcal{M} \rightarrow \hat{\Sigma}$ the associated map, identifying $\hat{\Sigma}_{\mathcal{M}} := \hat{\iota}(\mathcal{M})$ with $\Sigma_{\mathcal{M}}$.

Remark 2.2. The left shift $\sigma: \Sigma \rightarrow \Sigma$, $\{a_i\}_{i \in \mathbb{Z}} \mapsto \{a_{i+1}\}_{i \in \mathbb{Z}}$ induces a map $\hat{\sigma}: \hat{\Sigma} \rightarrow \hat{\Sigma}$. Moreover, by definition, the space $\hat{\Sigma}_{\mathcal{M}}$ is invariant under $\hat{\sigma}$, and $\hat{\iota} \circ f = \hat{\sigma} \circ \hat{\iota}$.

Proposition 2.3. *The mapping $\iota: \mathcal{M} \rightarrow \Sigma_{\mathcal{M}}$ is a bijection satisfying the following properties:*

- (i) *the restriction of ι to the regular set $\mathcal{M}^{\text{reg}}$ is continuous;*
- (ii) *the inverse mapping $(\iota|_{\mathcal{M}^{\text{reg}}})^{-1}$ is Hölder continuous. In particular,*

$$\iota|_{\mathcal{M}^{\text{reg}}}: \mathcal{M}^{\text{reg}} \rightarrow \Sigma_{\mathcal{M}^{\text{reg}}} := \iota(\mathcal{M}^{\text{reg}})$$

is a homeomorphism;

- (iii) *the discontinuity set of ι is precisely the singular set $\mathcal{S}_{\mathcal{M}}$.*

Proof. For $\alpha \in \Sigma$ and $n \in \mathbb{N}$, consider the preimage $C(\alpha; n) := \iota^{-1}(B(\alpha, 2^{-n}))$, where $B(\alpha, 2^{-n})$ is the open ball centred at α with radius 2^{-n} for the metric ρ_{Σ} . It consists of points $x \in \mathcal{M}$ whose first $n+1$ iterates under f , both in the past and future, are encoded by the n -cylinder obtained by truncating α , i.e.,

$$C(\alpha; n) = \{x \in \mathcal{M} \mid f^i(x) \in B_{a_i} \text{ for all } |i| \leq n+1\}.$$

The set H_n^s of points of $\mathcal{M}$ sharing the forward itinerary $(a_0, \dots, a_{n+1})$ is a closed set each of whose components is a “stable strip”, bounded by singularity curves of f^{n+1} , on which f^{n+1} is continuous. Unstable curves of $\mathcal{M}$ have length bounded by a constant D depending only on the table (they are graphs of functions with bounded slope over the compact set $\mathcal{M}$), while unstable curves inside H_n^s are expanded by f^{n+1} at a uniform exponential rate: there are constants $K, \mathcal{B} > 0$, depending only on the table (on the curvature function $\mathcal{K}$ and the free-flight bounds $\tau_{\min}, \tau_{\max}$), such that unstable vectors are expanded by at least $K^{-1}e^{\mathcal{B}n}$ (see the growth estimates for dispersing billiards in [15, Chapters 4–5]). Here an obstacle B that is not in the sequence $B_{a_1}, \dots, B_{a_n}$ may be eclipsed by the dispersing wave fronts generated by

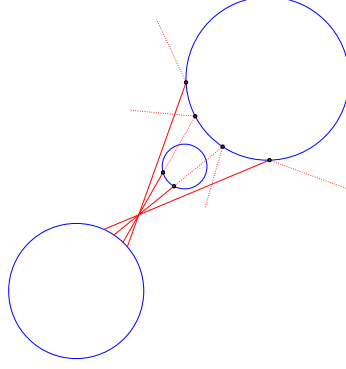


FIGURE 3. A smaller obstacle eclipsing a section of the interior of a wave front.

the unstable curves of H_n^s , and can then safely be ignored, see Figure 3. Hence every unstable curve inside H_n^s has length at most KDe^{-Bn} , i.e., the unstable width of H_n^s is at most KDe^{-Bn} . Symmetrically, the set H_n^u of points sharing the backward itinerary $(a_{-(n+1)}, \dots, a_0)$ is a strip of stable width at most KDe^{-Bn} . Since the stable and unstable cone fields are uniformly transverse, the cylinder $C(\alpha; n) \subset H_n^s \cap H_n^u$ satisfies

$$\text{diam } C(\alpha; n) \leq c(\mathcal{D}) e^{-Bn} \quad (1)$$

for a constant $c(\mathcal{D}) > 0$ depending only on the table. In particular, if $\iota(x) = \iota(y)$, then $x, y \in C(\alpha; n)$ for every n , hence $x = y$, and ι is indeed a bijection onto its image.

Now, let us denote by $j := \iota|_{\mathcal{M}^{\text{reg}}}$ the restriction of ι to the regular set. We claim that $j^{-1}(B(\alpha, 2^{-n})) = C(\alpha; n) \cap \mathcal{M}^{\text{reg}}$ is open. Indeed, since each f^i is continuous on $\mathcal{M}^{\text{reg}}$, there is a neighbourhood of any x in such preimage where every point bounces in exactly the same obstacles as x for $|i| < n + 1$. Hence every point in $C(\alpha; n) \cap \mathcal{M}^{\text{reg}}$ is interior and the map j is continuous, proving (i).

As for (ii), note that if $\rho_\Sigma(\alpha, \beta) = 2^{-n}$, then $\iota^{-1}(\alpha), \iota^{-1}(\beta) \in C(\alpha; n - 1)$, and it follows from the inequality in (1) (applied with $n - 1$ in place of n) that,

$$\|\iota^{-1}(\alpha) - \iota^{-1}(\beta)\| \leq \text{diam } C(\alpha; n - 1) \leq c(\mathcal{D}) e^{-B(n-1)} \leq c'(\mathcal{D}) \rho_\Sigma(\alpha, \beta)^{\mathcal{B}(\log 2)^{-1}},$$

hence ι^{-1} is $\mathcal{B}(\log 2)^{-1}$ -Hölder continuous.

Finally, we prove (iii). Property (i) amounts to saying that the discontinuity set of ι is contained in the singular set $\mathcal{S}_{\mathcal{M}}$, so all that remains to show is the converse inclusion. Let $x_0 \in \mathcal{S}_{\mathcal{M}}$; replacing x_0 by an iterate, and ι by its composition with a power of the shift, which does not affect continuity, we assume that $f(x_0) \in \mathcal{S}_0$. Let $\text{Pr}_1(x_0) \in \Gamma_i$ and $\text{Pr}_1(f(x_0)) \in \Gamma_j$. Then, due to the discontinuity of f at x_0 , we can find points x arbitrarily close to x_0 and such that $\text{Pr}_1(f(x)) \notin \Gamma_j$, as well as points arbitrarily close and such that $\text{Pr}_1(f(x)) \in \Gamma_j$. Thus, there are sequences $x_n \rightarrow x_0$ in $\mathcal{M}$ such that $\rho_\Sigma(\iota(x_n), \iota(x_{n+1})) = 1$ for every $n \in \mathbb{N}$, and x_0 is therefore a discontinuity point for ι . $\square$

2.2. Periodic orbits. Although it is not needed for the main part of the argument, we collect in this section general properties of periodic orbits and their lengths. By Proposition 2.3, we have:

Corollary 2.4. *For any $T > 0$, the number of periodic orbits of period at most T is finite.*

Proof. Fix $T > 0$. Fix a periodic point (s, ϕ) whose orbit under the billiard flow has length at most T . Let $q > 0$ be its period under the billiard map, $f^q(s, \phi) = (s, \phi)$. In particular, we have $q \leq q_0$, with $q_0 := \lfloor \frac{T}{\tau_{\min}} \rfloor + 1$. Let $\iota(s, \phi) =: (\rho_k)_k \in \Sigma$. Since $f^q(s, \phi) = (s, \phi)$, we have $\rho_q = \rho_0 + (\tilde{i}, \tilde{j}; 0)$, for some $(\tilde{i}, \tilde{j}) \in \mathbb{Z}^2$, and

$$\rho_p = \rho_{p \bmod q} + \left\lfloor \frac{p}{q} \right\rfloor (\tilde{i}, \tilde{j}; 0), \quad \forall p \in \mathbb{Z}.$$

Moreover, by the finite horizon assumption, and the $\mathbb{Z}^2$ -periodicity of $\tilde{\mathcal{D}}$, there exists a uniform upper bound $K \geq 1$ on the difference $|i_{p+1} - i_p| + |j_{p+1} - j_p| \leq K$ between the labels (i_p, j_p) and (i_{p+1}, j_{p+1}) of the cells at steps p and $p + 1$, for each $p \in \mathbb{Z}$. In particular, we have $|\tilde{i}| + |\tilde{j}| \leq Kq$. We conclude that periodic points of period at most T are completely encoded by finite words $(\rho_0, \dots, \rho_q)$, $q \leq q_0$. Moreover, given $(\rho_0, \dots, \rho_\ell)$, $1 \leq \ell \leq q - 1$, the label $(i_{\ell+1}, j_{\ell+1})$ satisfies $|i_{\ell+1} - i_\ell| + |j_{\ell+1} - j_\ell| \leq K$, so there are at most $(2K^2 + 2K + 1)k$ choices for the symbol $\rho_{\ell+1}$. Therefore, the number of (infinite) words $\iota(s, \phi)$ associated to periodic points with period at most T is finite. By the injectivity of ι shown in Proposition 2.3, we deduce that the set of periodic orbits with period less than T for the billiard flow is finite. $\square$

Fix an integer $r \geq 3$. Given an integer $m \geq 2$, we denote by $\mathfrak{D}^m$ the space of all Sinai billiard tables in $\mathbb{T}^2$ with m scatterers and C^r boundary, that have finite horizon. Each table $\mathcal{D}$ is defined by its boundary map $\gamma: \mathbb{T}_{\mathcal{D}} \rightarrow \mathbb{T}^2$, allowing us to identify $\mathfrak{D}^m$ with an open subset of the space of embeddings $C_{\text{emb}}^r(\mathbb{T}_{\mathcal{D}}, \mathbb{T}^2)$, endowed with the C^r topology (strict convexity of the components and finite horizon are open conditions). In particular, it is a Baire space.

We want to show that for a generic choice of table in $\mathfrak{D}^m$, the set of periodic orbits has simple length spectrum and contains no singular orbits.

Definition 2.5. For each integer $n \geq 1$, we define $\mathcal{R}_n$ as the subset of $\mathfrak{D}^m$ consisting of tables without any singular periodic orbit of period less than or equal to n .

Definition 2.6. For each integer $n \geq 1$, we define $\mathfrak{S}_n$ as the subset of $\mathfrak{D}^m$ consisting of tables such that no distinct periodic orbits of period less than or equal to n have the same length.

Proposition 2.7. *For each integer $n \geq 1$, the set $\mathcal{R}_n \cap \mathfrak{S}_n$ is open and dense in $\mathfrak{D}^m$. Consequently, the set*

$$\mathfrak{G} := \bigcap_n (\mathcal{R}_n \cap \mathfrak{S}_n)$$

of tables with no grazing periodic orbit and simple length spectrum is residual in the C^r topology.

This result was already obtained by Petkov–Stoyanov, see e.g. [41]; for completeness, we give a proof of Proposition 2.7 in Appendix A, where further properties of periodic orbits are investigated.

3. GEODESIC APPROXIMATIONS

3.1. Kourganoff surfaces. Recall that Ψ is the billiard flow of the table $\mathcal{D}$, and let Ω denote its phase space. It is the suspension space of the billiard map f with roof function τ . We consider $E = \mathbb{T}^2 \times \mathbb{R}$ with coordinates (x, y, z) and a canonical

global frame $(\partial_x, \partial_y, \partial_z)$. We denote by $p: E \rightarrow \mathbb{T}^2$ the canonical projection on the first two coordinates.

Consider a smooth closed immersed surface K in E such that $p(K) = \mathcal{D}$. In what follows, we abuse the notation slightly and write

$$\begin{aligned} K^\circ &:= K \cap p^{-1}(\text{int}(\mathcal{D})), \\ B(K) &:= K \cap p^{-1}(\partial\mathcal{D}). \end{aligned}$$

Moreover, we assume K satisfies

- (K0) K° has exactly two connected components, each mapped diffeomorphically onto $\text{int}(\mathcal{D})$ by p .
- (K1) The vector ∂_z is not tangent to K at any point $x \in K^\circ$, i.e. $\partial_z \notin T_x K$.
- (K2) For every point $x \in B(K)$, the curve defined by the intersection of K and the affine plane $\{x\} + (\mathbb{R}\partial_z \oplus T_x^\perp K)$ has non-vanishing curvature around x .

By (K1), the restriction $p|_{K^\circ}$ is a local diffeomorphism. By (K0), K° consists of two sheets, each a copy of $\text{int}(\mathcal{D})$, so that every point of $\text{int}(\mathcal{D})$ has exactly two preimages. Near a point of $B(K)$, condition (K2) says that $p|_K$ is a fold, along which two local sheets meet; since interior points have only two preimages, there is exactly one fold curve above each scatterer $\Gamma_i \subset \partial\mathcal{D}$: each connected component of $B(K)$ is projected diffeomorphically onto a single Γ_i , and every point of $\partial\mathcal{D}$ has a unique preimage under p . In particular, $p|_K$ fails to be a local diffeomorphism precisely along $B(K)$. Since the kernel of dp is the vertical direction $\mathbb{R}\partial_z$, we obtain

$$\ker(dp_x|_{T_x K}) = T_x K \cap \mathbb{R}\partial_z = \mathbb{R}\partial_z, \quad x \in B(K),$$

so the vertical direction is tangent to K exactly at the boundary.

We thus think of K as being obtained by gluing two copies of $\mathcal{D}$ along their corresponding boundary components. We will refine this statement in Lemma 3.11 below. We refer the reader to Figure 1 for a picture of the construction.

For $\varepsilon \in (0, 1]$, we let α_ε be the contraction given by $(x, y, z) \mapsto (x, y, \varepsilon z)$ and denote by g_ε the metric induced on $K_\varepsilon := \alpha_\varepsilon(K)$ by the Euclidean metric. Pulling back via α_ε , we obtain a family of metrics $\{g_\varepsilon\}$ on the surface K . We denote by $\mathbf{g}_\varepsilon$ the geodesic flow of the surface (K, g_ε) . Then p induces a smooth mapping

$$\begin{aligned} \pi_K: T^1 K^\circ &\rightarrow \Omega \\ (x, v) &\mapsto \left(p(x), \frac{dp_x(v)}{\|dp_x(v)\|} \right) \end{aligned} \tag{2}$$

(note that (K1) ensures we are not dividing by 0 in the definition). Finally, we denote by A_0 the subset of $\mathbb{R} \times T^1 K^\circ$ consisting of points $(t; x, v)$ such that the orbit segment $\{\Psi^s(\pi_K(x, v)); s \in [0, t]\}$ has endpoints in $\text{int}(\mathcal{D})$ and does not contain grazing collisions. The set A_0 is clearly open; moreover, it contains

$$\{(t; x, v) \in \mathbb{R} \times \pi_K^{-1}(\Omega^{\text{reg}}); \Psi^0(\pi_K(x, v)), \Psi^t(\pi_K(x, v)) \notin \partial\mathcal{D}\},$$

which is dense in $\mathbb{R} \times T^1 K^\circ$. We define mappings $F_\varepsilon: A_0 \rightarrow \Omega$, $\varepsilon \in [0, 1]$ by

$$\begin{aligned} F_0: (t; x, v) &\mapsto \Psi^t(\pi_K(x, v)), \\ F_\varepsilon: (t; x, v) &\mapsto \pi_K(\mathbf{g}_\varepsilon^t(x, v)), \text{ for } \varepsilon > 0. \end{aligned}$$

Theorem 3.1 (Theorem 5 in [36]). *The family of maps F_ε converges uniformly to F_0 , as $\varepsilon \rightarrow 0$, on every compact subset of A_0 . Moreover, there is $\varepsilon_0 > 0$ such that $\mathbf{g}_\varepsilon$ is an Anosov flow for every $0 < \varepsilon < \varepsilon_0$.*

In particular, up to rescaling, without loss of generality, we can assume that

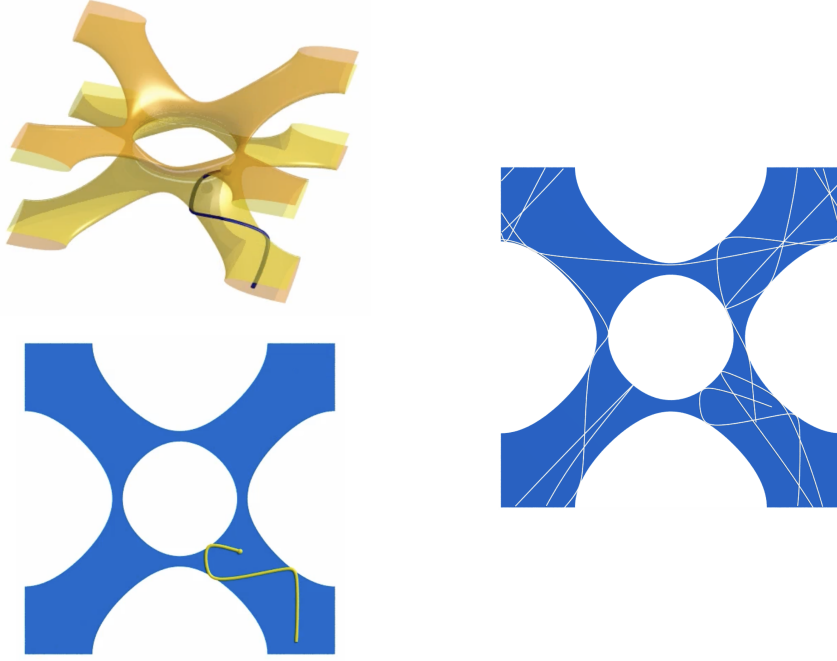


FIGURE 4. *Left:* A geodesic arc on some surface K_ε and its projection on the billiard table $\mathcal{D}$. *Right:* for $\varepsilon > 0$ small and $(t; x, v) \in A_0$, the projection $\{F_\varepsilon(s; x, v)\}_{s \in [0, t]}$ closely shadows the billiard orbit segment $\{F_0(s; x, v)\}_{s \in [0, t]}$ (Credits for pictures: Mickaël Kourganoff, Jos Leys).

(K3) The family $\{\mathbf{g}_\varepsilon\}_{\varepsilon \in (0, 1]}$ is a one-parameter family of Anosov flows in K .

We will denote by $\mathcal{K}(\mathcal{D})$ the collection of all such surfaces:

$$\mathcal{K}(\mathcal{D}) := \{K \subset E; p(K) = \mathcal{D} \text{ and } K \text{ satisfies (K0), (K1), (K2) and (K3)}\}.$$

3.1.1. Metric convergence of Kourganoff surfaces. Fix a Kourganoff surface $K \in \mathcal{K}(\mathcal{D})$. Given a patch $\varphi(u, v) = (x(u, v), y(u, v), z(u, v))$ for K , we consider the associated patch $\alpha_\varepsilon \circ \varphi = (x(u, v), y(u, v), \varepsilon z(u, v))$ for K_ε . Comparing the First Fundamental Forms $I = E du^2 + 2F du dv + G dv^2$ and $I_\varepsilon = E_\varepsilon du^2 + 2F_\varepsilon du dv + G_\varepsilon dv^2$ for K and K_ε in these coordinates, we obtain

$$\begin{aligned} E_\varepsilon &= E + (\varepsilon^2 - 1)(\partial_u z)^2, \\ F_\varepsilon &= F + (\varepsilon^2 - 1)\partial_u z \partial_v z, \\ G_\varepsilon &= G + (\varepsilon^2 - 1)(\partial_v z)^2. \end{aligned} \tag{3}$$

Hence, for any $\varepsilon, \delta \in (0, 1]$, the difference between the coefficients of the metrics g_ε and g_δ , as well as the differences between their derivatives, are of order $|\varepsilon^2 - \delta^2|$. Thus, the C^r norm of their difference, as a tensor field on K , can be bounded by

$$\|g_\varepsilon - g_\delta\|_{C^r(K)} \leq C|\varepsilon^2 - \delta^2| \leq C, \tag{4}$$

where C is a constant depending only on K (controlled by the C^r norm of the height function z on the compact surface K). Hence, given a curve γ , denoting by $l_\varepsilon(\gamma)$ its length with respect to the metric g_ε , the order of length variation of γ for two

different metrics g_ε and g_δ is

$$|l_\varepsilon(\gamma) - l_\delta(\gamma)| \leq \int_\gamma \left| \sqrt{I_\varepsilon} - \sqrt{I_\delta} \right| \leq C_K |\varepsilon - \delta| l_1(\gamma). \quad (5)$$

A first consequence of the equation above is that a curve γ^1 is rectifiable (i.e., has finite length) with respect to g_ε if and only if it is rectifiable with respect to the reference metric g_1 , thus the class of rectifiable curves is the same for all the metrics. Given $x, y \in K$, we will denote by $[x, y]_\varepsilon$ the minimising geodesic for g_ε joining them. Note that $\{[x, y]_\varepsilon\}$ is a collection of curves in K with endpoints x, y , all of which are rectifiable with respect to every metric in the family $\{g_\varepsilon\}$.

Moreover, by comparing the fundamental forms in (3), for $v \in TK$, we have $g_\varepsilon(v, v) = |dp(v)|^2 + \varepsilon^2 |dz(v)|^2$, which is non-decreasing in ε . Hence, for every fixed curve γ , the length $\varepsilon \mapsto l_\varepsilon(\gamma)$ is non-decreasing, and therefore so is $\varepsilon \mapsto d_\varepsilon(x, y)$; being bounded below by 0, it converges as $\varepsilon \rightarrow 0$. We may thus define

$$d_0(x, y) := \lim_{\varepsilon \rightarrow 0} d_\varepsilon(x, y) = \lim_{\varepsilon \rightarrow 0} l_\varepsilon([x, y]_\varepsilon). \quad (6)$$

A priori d_0 is only a pseudo-distance on K ; Theorem 3.6 below shows that it is a distance, that $d_\varepsilon \rightarrow d_0$ uniformly on $K \times K$, and describes the resulting length space $K_0 := (K, d_0)$.

Away from the seam $B(K)$, Equation (4) controls the geometry. No such control holds near $B(K)$, where $\det g_\varepsilon \rightarrow 0$ and $\|g_\varepsilon^{-1}\| \rightarrow \infty$, and the sprays are not uniformly Lipschitz as $\varepsilon \rightarrow 0$. Taking the non-degenerate limit spray f_0 as the fixed reference field and by the continuous dependence of ODEs on the field (see e.g. [45, Chapter 3.12]), we obtain the following statement on the local uniform convergence of sprays.

Lemma 3.2. *Let $Q \subset K^\circ$ be relatively compact and $J \subset \mathbb{R}$ compact, and suppose there are a compact $Q' \Subset K^\circ$ and $\varepsilon_* > 0$ with $\mathbf{g}_\varepsilon^t(X) \in T^1 Q'$ for all $X \in T^1 Q$, all $t \in J$, and all $0 < \varepsilon \leq \varepsilon_*$. Here $T^1 Q, T^1 Q'$ denote the g_0 -unit tangent bundles, the flows $\mathbf{g}_\varepsilon$ being compared on TK through their sprays, and d_G is a fixed Riemannian distance on $TK|_{Q'}$. Writing $\mathbf{g}_0$ for the geodesic flow of the non-degenerate metric $g_0|_{Q'}$, we have*

$$\sup_{X \in T^1 Q} \sup_{t \in J} d_G(\mathbf{g}_\varepsilon^t X, \mathbf{g}_0^t X) \xrightarrow{\varepsilon \rightarrow 0} 0,$$

and the convergence is C^1 in X .

Proof. On $Q' \Subset K^\circ$ we have $\det g_0 \geq c(Q') > 0$, so $g_\varepsilon \rightarrow g_0$ in $C^r(Q')$ gives $g_\varepsilon^{-1} \rightarrow g_0^{-1}$ in $C^r(Q')$; hence the Christoffel symbols and sprays converge, $f_\varepsilon \rightarrow f_0$ in $C^{r-1}(T^1 Q')$, and the sectional curvatures of g_ε are uniformly bounded on Q' for $\varepsilon \leq \varepsilon_*$. Take f_0 as the fixed reference field: it is Lipschitz on $T^1 Q'$ with a constant $L = L(Q')$ independent of ε . By Grönwall's lemma applied to the orbit segments (which remain in $T^1 Q'$) with reference f_0 , perturbation f_ε , $\delta_\varepsilon := \|f_\varepsilon - f_0\|_{C^0(T^1 Q')} \rightarrow 0$, and equal initial data, we have

$$\sup_{t \in J} d_G(\mathbf{g}_\varepsilon^t X, \mathbf{g}_0^t X) \leq \delta_\varepsilon |J| e^{L|J|} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

uniformly in $X \in T^1 Q$. By differentiating the geodesic equation and repeating the estimate for the Jacobi equation, whose coefficients also converge in $C^0(T^1 Q')$, we upgrade this to C^1 convergence in X . $\square$

¹ Please remark that through this section we frequently abuse the terminology and refer to both the maps $[a, b] \rightarrow K$ and their images as curves.

Lemma 3.3. *There is a uniform lower bound $R_0 > 0$ on the injectivity radius R_ε of the metrics g_ε .*

Proof. As the geodesic flows $\mathbf{g}_\varepsilon$ are all Anosov, none of the surfaces K_ε have conjugated points (see e.g. [35]). Therefore, according to Klingenberg's lemma (see [2, Lemma 1.8]) the injectivity radius of K_ε is

$$\text{inj}(g_\varepsilon) = \frac{1}{2}l_\varepsilon(\sigma_\varepsilon)$$

where σ_ε is the closed geodesic of minimal g_ε -length in K . We claim that

$$\inf_{0 < \varepsilon \leq 1} l_\varepsilon(\sigma_\varepsilon) \geq s_0 := \min\{\text{sys}(\mathbb{T}^2), \min_i l_i, 2\tau_{\min}\} > 0,$$

where $\text{sys}(\mathbb{T}^2)$ is the length of the shortest closed flat geodesic of the torus, l_i the length of the scatterer $\Gamma_i \subset \mathcal{D}$, and $\tau_{\min}$ the minimal distance between two distinct lifted obstacles in $\mathbb{R}^2$ (lattice translates of the same obstacle included), i.e., the minimal flight time of the table.

Since the projection $p: E \rightarrow \mathbb{T}^2$ is 1-Lipschitz for every g_ε , we have $l_\varepsilon(\sigma_\varepsilon) \geq l_E(\gamma_\varepsilon)$ for the projected loop $\gamma_\varepsilon := p(\sigma_\varepsilon)$, which is a loop in the closed table $\overline{\mathcal{D}}$ (it avoids the open obstacles); therefore, it suffices to bound $l_E(\gamma_\varepsilon)$ below by s_0 for the projection of any loop σ which is non-contractible in K . Note that non-contractibility of σ in K does not imply non-contractibility of $\gamma = p(\sigma)$ in $\mathcal{D}$: the kernel of p_* contains for instance all the classes of billiard trajectories bouncing back and forth between two scatterers, whose projections are segments traversed twice. We argue according to the crossings of σ with the fold locus $B(K)$, using that each transversal crossing switches the sheet, so that the number of fold crossings is even.

Case 1: σ stays in one closed sheet. In that case, p restricts to a homeomorphism of the sheet onto $\overline{\mathcal{D}}$, so γ is non-contractible in $\overline{\mathcal{D}}$. By minimising the Euclidean length in the free homotopy class of γ over the compact $\overline{\mathcal{D}}$, we obtain a closed Euclidean geodesic of $\overline{\mathcal{D}}$: a concatenation of straight chords, with endpoints on the walls, and of wall arcs. If this minimiser is essential in $\mathbb{T}^2$, its length is at least $\text{sys}(\mathbb{T}^2)$. If it is trivial in $\mathbb{T}^2$, we lift it to a closed loop in $\mathbb{R}^2$: a chord leaving a strictly convex lifted obstacle never returns to the same lift, so

- either the lifted loop touches a single obstacle lift and contains no chord; in that case, the class is a non-trivial power of a single scatterer class Γ_i , and any representative winding $m \neq 0$ times around the convex obstacle O_i has length at least $|m|l_i \geq l_i$;
- or it contains at least two chords, each joining two distinct obstacle lifts, hence each of length $\geq \tau_{\min}$, for a total $\geq 2\tau_{\min}$.

Case 2: σ crosses the fold locus. The crossings decompose σ into an even number ≥ 2 of arcs alternating between the two sheets, with endpoints on the wall circles. We write the class of σ as the corresponding alternating word in the amalgamated product $\pi_1(K) = \pi_1(\overline{\mathcal{D}}) *_{\pi_1(B(K))} \pi_1(\overline{\mathcal{D}})$. If some projected arc is homotopic (relative to its endpoints and within $\overline{\mathcal{D}}$) into the wall containing them, the word reduces; indeed, we can push the arc into the wall and merge it with its neighbours. After reduction, either the class lies in a vertex group (a one-sheet class, handled by Case 1, with the same length bound; indeed, replacing an arc by the wall arc it is homotopic to does not increase the Euclidean length) or there remain at least two irreducible arcs. Let us lift an irreducible arc to $\mathbb{R}^2$, with endpoints on lifted walls. If the endpoints lie on two distinct obstacle lifts, then the arc has length at least $\tau_{\min}$. If they lie on the same lift $\tilde{\Gamma}$, irreducibility means that the arc is not homotopic into

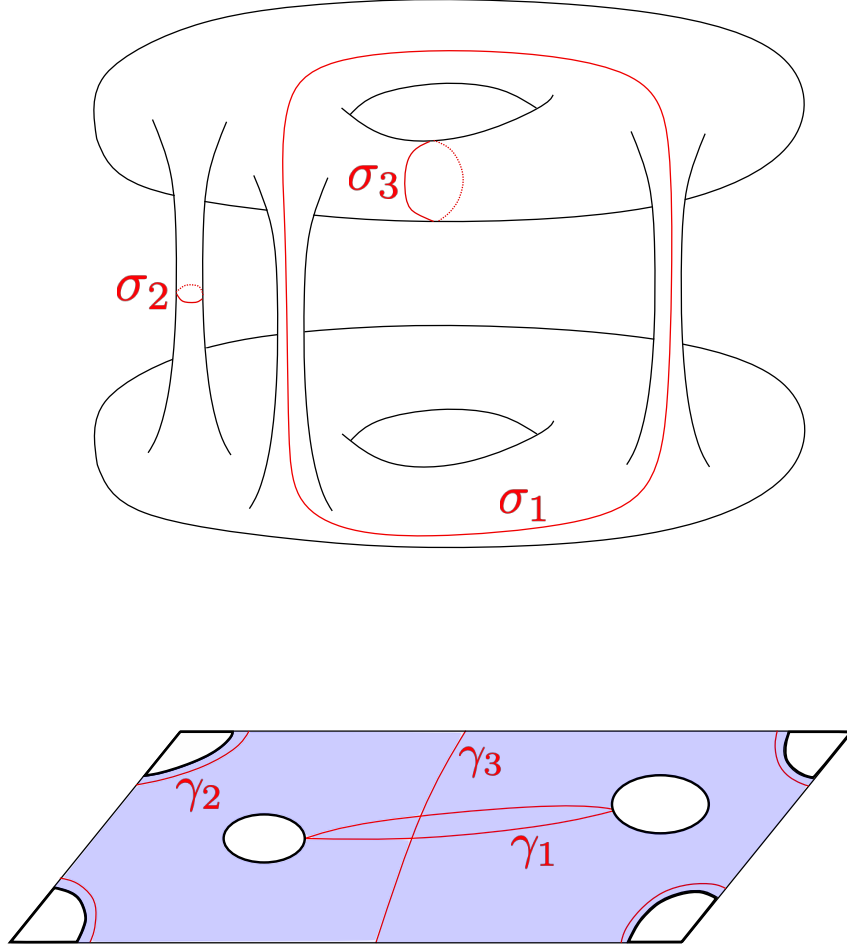


FIGURE 5. Some loops $\{\sigma_i\}_{i=1}^3$ on the surface K , and their projections $\{\gamma_i = p(\sigma_i)\}_{i=1}^3$ on the billiard table $\mathcal{D}$.

$\tilde{\Gamma}$ relative to its endpoints within the complement of the open obstacles; then the Jordan region bounded by the arc and a boundary arc of $\tilde{\Gamma}$ contains another obstacle lift Ω , and the arc passes through a point z on the far side of Ω : moving from any $x \in \Omega$ along the ray opposite to the nearest-point direction to $\tilde{\Gamma}$, the distance to the convex set $\tilde{\Gamma}$ is convex and increasing, so $d(z, \tilde{\Gamma}) \geq d(x, \tilde{\Gamma}) \geq \tau_{\min}$, and the arc, whose endpoints lie on $\tilde{\Gamma}$, has length $\geq 2d(z, \tilde{\Gamma}) \geq 2\tau_{\min}$. In all cases the total Euclidean length is at least $2\tau_{\min}$ when it remains at least two irreducible arcs (see Figure 5).

We conclude that $l_\varepsilon(\sigma_\varepsilon) \geq l_E(\gamma_\varepsilon) \geq s_0$, and $R_0 := \frac{1}{2}s_0$ is the desired uniform lower bound. $\square$

Lemma 3.4. *Let $x \in K^\circ$, and let U be the connected component containing x of $p^{-1}(B_{\mathcal{D}}(p(x), \delta))$, where $\delta > 0$ is small enough such that $B_{\mathcal{D}}(p(x), \delta) \subset \text{int}(\mathcal{D})$ contains no collision point of any billiard trajectory, and $\bar{U} \subset B_{g_1}(x, R_0)$, so that $\bar{U}$ lies in a normal neighbourhood of x for every metric g_ε (by Lemma 3.3, since $d_\varepsilon \leq d_1$). For every $y \in U$, the geodesics $[x, y]_\varepsilon$ converge uniformly to a C^1 curve $[x, y]_0 \subset U$, the lift of the Euclidean segment $[p(x), p(y)]$; in particular $d_0(x, y) = |p(x) - p(y)|$.*

Proof. Let $\tilde{x} \in K^\circ$, $x = p(\tilde{x}) \in \text{int}(\mathcal{D})$, and consider a neighbourhood U of $\tilde{x}$, which we choose sufficiently small so that $\bar{U}$ is a normal neighbourhood for all of the metrics g_ε (recall Lemma 3.3), and is moreover mapped under p into an open ball

$B := B_{\mathcal{D}}(x, \delta)$ with sufficiently small radius δ , ensuring that the billiard trajectories starting at x in any direction do not have any collision inside B .

Moreover, for a fixed $\tilde{y} \in U$, distinct from $\tilde{x}$, we write $l_\varepsilon = d_\varepsilon(\tilde{x}, \tilde{y})$ and let $\tilde{v}_\varepsilon$ be the velocity of the curve $[\tilde{x}, \tilde{y}]_\varepsilon$ at the starting point $\tilde{x}$. Finally, we set $y = p(\tilde{y})$, let $T = |x - y|$, $v = T^{-1}(y - x)$, and consider the geodesics γ_ε in K_ε with initial condition $(\tilde{x}, \tilde{v})$, where $\pi_K(\tilde{x}, \tilde{v}) = (x, v)$.

By construction, the point $(T; \tilde{x}, \tilde{v})$ lies in A_0 , and Theorem 3.1 guarantees that $\pi_K(\mathbf{g}_\varepsilon^t(\tilde{x}, \tilde{v}))$ converges uniformly to $\Psi^t(x, v)$ in the interval $[0, T]$ as $\varepsilon \rightarrow 0$. Thus, for ε sufficiently small we have $\gamma_\varepsilon(T)$ arbitrarily close to $\tilde{y}$. Since the exponential map $\exp_x^\varepsilon$ is a diffeomorphism onto a neighbourhood of $\bar{U}$, and since $\exp_x^\varepsilon(l_\varepsilon \tilde{v}_\varepsilon) = \tilde{y}$ while $\exp_x^\varepsilon(T\tilde{v}) = \gamma_\varepsilon(T)$, we can estimate

$$\|l_\varepsilon \tilde{v}_\varepsilon - T\tilde{v}\|_\varepsilon \leq \left\| d((\exp_x^\varepsilon)^{-1})|_{\bar{U}} \right\|_\varepsilon d_\varepsilon(\tilde{y}, \gamma_\varepsilon(T)).$$

Around $\tilde{x}$ the norm of the derivative is close to 1. Moreover, as pointed out in [36], as the surface K flattens out, the curvature away from $B(K)$ tends to 0, while in a neighbourhood of $B(K)$ it becomes strictly negative. Therefore, for ε sufficiently small, the sectional curvatures of K_ε on $\bar{U}$ are bounded in absolute value by some $1/R^2$, and we can estimate the derivative of $\exp_x^\varepsilon$ in terms of the norms of Jacobi fields (using [37, Theorem 11.2], for instance), which are bounded by $R \sinh(r/R)$ at radius r , hence bounded on $\bar{U}$ uniformly in ε . We may then find a constant L , independent of ε , such that

$$\|l_\varepsilon \tilde{v}_\varepsilon - T\tilde{v}\|_\varepsilon \leq L d_\varepsilon(\tilde{y}, \gamma_\varepsilon(T)) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

On the other hand, $l_\varepsilon = d_\varepsilon(\tilde{x}, \tilde{y}) \geq |x - y| = T$ because p is 1-Lipschitz, and $l_\varepsilon \leq l_\varepsilon(\sigma)$, where $\sigma \subset U$ is the lift of the segment $[x, y]$, whose g_ε -length decreases to T as $\varepsilon \rightarrow 0$; hence $l_\varepsilon \rightarrow T$, and therefore $\tilde{v}_\varepsilon \rightarrow \tilde{v}$. Since $\bar{U} \Subset K^\circ$ is collision-free and the geodesics $[\tilde{x}, \tilde{y}]_\varepsilon$ stay in $\bar{U}$ (being minimising in a normal ball), we may apply Lemma 3.2 with $Q = U$ and $Q' = \bar{U}$: the g_ε -geodesics with initial data $(\tilde{x}, \tilde{v}_\varepsilon)$ converge, uniformly on $[0, T]$ and in C^1 , to the g_0 -geodesic with initial data $(\tilde{x}, \tilde{v})$, which is the lift σ of the segment $[x, y]$. Hence $[\tilde{x}, \tilde{y}]_\varepsilon \rightarrow [\tilde{x}, \tilde{y}]_0 := \sigma$, and $d_0(\tilde{x}, \tilde{y}) = \lim_\varepsilon l_\varepsilon = T = |x - y|$. $\square$

Essentially the same argument shows that we may fix the tangent vectors v_ε and let the endpoint y vary instead, which gives the following corollary.

Corollary 3.5. *Given $(x, v) \in T^1 K^\circ$, let $\gamma_{\varepsilon;v}$ denote the unique geodesic of (K, g_ε) with initial condition (x, v) . If U is any connected neighbourhood of x not intersecting $B(K)$, then there is a unique differentiable curve $\gamma_{0;v}$, starting at x , and whose tangent space at x is spanned by v .*

Theorem 3.6. *The function d_0 defined in (6) is a distance on K : $d_0(x, y) > 0$ for all distinct $x, y \in K$. Moreover, $K_0 := (K, d_0)$ is a complete geodesic length space, whose topology is that of K , and $d_\varepsilon \rightarrow d_0$ uniformly on $K \times K$. Every g_1 -rectifiable curve is d_0 -rectifiable, and the length function is*

$$l_0: \text{Rec}(K, g_1) \rightarrow [0, \infty) \\ \gamma \mapsto \lim_{\varepsilon \rightarrow 0} l_\varepsilon(\gamma) = l_E(p \circ \gamma),$$

where l_E denotes the Euclidean length. Every locally uniform limit of geodesics of K_{ε_n} , $\varepsilon_n \rightarrow 0$, is a geodesic² of K_0 ; conversely, every billiard trajectory with only

² We mean geodesic in the sense of length spaces, i.e., a locally minimising path with respect to d_0 . In general, the tensors g_ε degenerate as $\varepsilon \rightarrow 0$ and the metric d_0 is not associated to any Riemannian structure.

transversal collisions is the projection of such a limit. They are C^1 on $\gamma^{-1}(K \setminus B(K))$, with well-defined one-sided directions in $T\mathcal{D}$ at each collision; the convergence is C^1 away from $B(K)$, but at a transversal crossing the limit may not be differentiable.

Proof. Symmetry, the triangle inequality and $d_0(x, x) = 0$ follow from the definition (6). Moreover, since p is 1-Lipschitz from (K, g_ε) to the Euclidean torus for every ε , we have $d_0(x, y) \geq |p(x) - p(y)|$; hence $d_0(x, y) > 0$ whenever $p(x) \neq p(y)$. If $p(x) = p(y)$ and $x \neq y$, then $p(x) \in \text{int}(\mathcal{D})$ (boundary points have a unique preimage) and x, y lie on distinct sheets, so every rectifiable path γ from x to y crosses the seam $B(K)$: its projection travels from $p(x)$ to $\partial\mathcal{D}$ and back, so that $l_\varepsilon(\gamma) \geq l_E(p \circ \gamma) \geq 2d_{\mathcal{L}}(p(x), \partial\mathcal{D})$ for every ε , and $d_0(x, y) \geq 2d_{\mathcal{L}}(p(x), \partial\mathcal{D}) > 0$. Thus d_0 is a metric.

Since $d_0 \leq d_1$, we have $|d_0(x, y) - d_0(x', y')| \leq d_1(x, x') + d_1(y, y')$, so d_0 is continuous on $K \times K$ for the manifold topology. Hence $d_\varepsilon \rightarrow d_0$ uniformly (by Dini's theorem), and the identity $(K, d_1) \rightarrow (K, d_0)$ is a homeomorphism. In particular K_0 is compact, hence complete, and it is the Gromov–Hausdorff limit of the length spaces K_ε ; therefore it is a length space [13, Theorem 7.5.1], and, being compact, a geodesic space.

Let γ be g_1 -rectifiable. By monotone convergence,

$$\lim_{\varepsilon \rightarrow 0} l_\varepsilon(\gamma) = \lim_{\varepsilon \rightarrow 0} \int \sqrt{|dp(\dot{\gamma})|^2 + \varepsilon^2 |dz(\dot{\gamma})|^2} = \int |dp(\dot{\gamma})| = l_E(p \circ \gamma).$$

The length $L_0(\gamma)$ induced by the metric d_0 satisfies $L_0(\gamma) \leq \lim_\varepsilon l_\varepsilon(\gamma)$, because $d_0 \leq l_\varepsilon$ on every sub-arc, and $L_0(\gamma) \geq l_E(p \circ \gamma)$, because $d_0 \geq |p(\cdot) - p(\cdot)|$. Hence $l_0 := L_0 = l_E \circ p = \lim_\varepsilon l_\varepsilon$ on g_1 -rectifiable curves (by the gluing description of Theorem 3.12 below, we will justify why every d_0 -rectifiable curve has rectifiable projection, with the same formula).

Let us now study limits of geodesics. Let γ_n be unit-speed geodesics of K_{ε_n} , $\varepsilon_n \rightarrow 0$, converging locally uniformly to a curve γ (the γ_n are 1-Lipschitz for $d_0 \leq d_{\varepsilon_n}$, so such limits exist along subsequences by Arzelà–Ascoli). Fix $a_0 < b_0$ with $b_0 - a_0 < R_0$, the uniform injectivity radius of Lemma 3.3; then $\gamma_n|_{[a_0, b_0]}$ is minimising, $d_{\varepsilon_n}(\gamma_n(a_0), \gamma_n(b_0)) = b_0 - a_0$, and by uniform convergence of d_{ε_n} to d_0 and continuity of d_0 we get $d_0(\gamma(a_0), \gamma(b_0)) = b_0 - a_0$. On the other hand,

$$L_0(\gamma|_{[a_0, b_0]}) = \sup \sum_i d_0(\gamma(t_i), \gamma(t_{i+1})) = \sup \sum_i \lim_n d_{\varepsilon_n}(\gamma_n(t_i), \gamma_n(t_{i+1})) \leq b_0 - a_0,$$

where the suprema is over partitions of $[a_0, b_0]$. Hence γ is locally minimising and parametrised by d_0 -arc length, i.e., it is a unit-speed geodesic of K_0 . Conversely, if $(t; x, v) \in A_0$ for every t , i.e., if the billiard trajectory of $\pi_K(x, v)$ has only transversal collisions, Theorem 3.1 shows that the g_ε -geodesics with initial data (x, v) converge locally uniformly to the lift of that trajectory starting at x (the lift being determined by the alternation of sheets at transversal crossings), which is a geodesic of K_0 by what precedes.

The regularity statements away from $B(K)$ follow from Lemma 3.4. We will investigate later (see the normal form of Lemma 3.13 below) the behavior of the height coordinate $z \circ \gamma$ at a transversal crossing and see it has a square-root cusp, so the limit may not be differentiable there. $\square$

Remark 3.7. The results above, in particular Theorem 3.6, do not guarantee uniqueness of the geodesic when the initial condition lies in $T^1B(K)$. In this case, if v is transverse to $B(K)$ (i.e., if $\pi_K(x, v)$ corresponds to a regular collision on Ω), then Kourganoff's approximation theorem guarantees $\gamma_{0;v}$ is unique. Otherwise, if v is tangent to $B(K)$, then there are three geodesics in K_0 starting at x with velocity

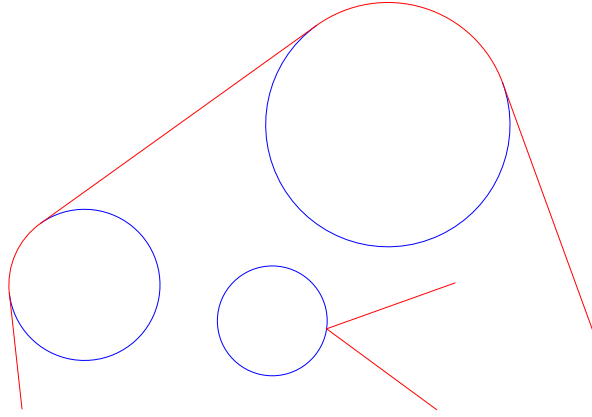


FIGURE 6. The longer curve following the obstacles is a $d_{\mathcal{L}}$ -geodesic, while the billiard trajectory segment is merely piecewise geodesic.

v : the component of $B(K)$ containing x , and the straight continuation of the line $t \mapsto p(x) + tv$ in either of the two sheets (see e.g. Lemma 3.48(iv)).

3.1.2. K_0 as a metric gluing. There is yet another way to construct K_0 . According to the proof of Lemma 3.4, given a connected component S of K° , every $x \in S$ has a neighbourhood U such that, for every $y \in U$, the curve $[x, y]_0$ satisfies

$$l_0([x, y]_0) = \lim_{\varepsilon \rightarrow 0} l_\varepsilon([x, y]_\varepsilon) = |p(x) - p(y)|.$$

We can then restate the main ingredient of the proof as

Lemma 3.8. *Let S be a connected component of K° and $|\cdot|$ the Euclidean distance in $\mathcal{D}$. Then $p|_S: (S, d_0) \rightarrow (\mathcal{D}, |\cdot|)$ is a local isometry.*

The argument is highly dependent on the approximation by geodesic flows, which fails if the points whose distance we are measuring are eclipsed by a scatterer, or after the billiard trajectory experiences a grazing collision. To deal with that, we consider a different metric on $\mathcal{D}$, making it into a length space.

Definition 3.9 (Length metric of $\mathcal{D}$). The table $\mathcal{D}$ has a length structure $\mathcal{L}$ consisting of all Euclidean rectifiable paths $\gamma: I \rightarrow \mathbb{R}^2$ whose image is completely contained in $\mathcal{D}$ together with the usual Euclidean length l_E induced by $|\cdot|$. We call the *length metric* of $\mathcal{D}$, denoted by $d_{\mathcal{L}}$, the metric induced by $\mathcal{L}$, i.e.,

$$d_{\mathcal{L}}(x, y) = \inf \{l_E(\gamma); \gamma \in \mathcal{L}, \gamma(0) = x, \gamma(1) = y\}.$$

The geodesics of $d_{\mathcal{L}}$ are straight lines in $\mathcal{D}$, segments of boundary, and paths consisting of concatenations of these two types: such curves are straight lines that meet a scatterer at a tangential angle, follow along the boundary for a while, and leave into a straight trajectory again, with another grazing angle (see Figure 6).

Remark 3.10. Similarly to K_0 , the space $(\mathcal{D}, d_{\mathcal{L}})$ is not uniquely geodesic. If $x \in \Gamma_i \subset \partial\mathcal{D}$ is a boundary point, and v is the tangent vector to Γ_i at x , then there are two geodesics starting at x in the direction of v : the straight line $x + tv$ into the interior of the table, and the curve Γ_i itself. Similarly, the points of $B(K)$ are exactly those at which d_0 fails to be uniquely geodesic.

Lemma 3.11. *For any connected component $S \subset K^\circ$, $p|_{\bar{S}}: (\bar{S}, d_0) \rightarrow (\mathcal{D}, d_{\mathcal{L}})$ is an isometry.*

Proof. The projection p is Lipschitz with respect to the metric g_1 on K and the Euclidean norm $|\cdot|$ on $\mathbb{T}^2$, hence the image $p(\gamma)$ of every rectifiable curve in K is rectifiable with respect to the Euclidean metric, and therefore $p(\text{Rec}(K)) \subset \mathcal{L}$.

According to Lemma 3.8 the map p is a local isometry on S , hence it preserves the length of every rectifiable curve γ mapped into $\text{int}(\mathcal{D})$, that is, $l_0(\gamma) = l_E(p(\gamma))$. More generally, $l_0(\gamma) = l_E(p \circ \gamma)$ for every g_1 -rectifiable curve γ in K , by Theorem 3.6.

Let $x, y \in \bar{S}$. Every rectifiable path in K from x to y projects to a path in $\bar{\mathcal{D}}$ from $p(x)$ to $p(y)$ of the same length, so $d_0(x, y) \geq d_{\mathcal{L}}(p(x), p(y))$: the map $p|_{\bar{S}}$ is non-expanding. Conversely, by (K0) the map $p|_{\bar{S}}$ is a homeomorphism onto $\bar{\mathcal{D}}$ (in fact, it is a diffeomorphism on S , extending continuously and bijectively to the closure, the fold being one-to-one over $\partial\mathcal{D}$). A $d_{\mathcal{L}}$ -geodesic from $p(x)$ to $p(y)$, which exists by compactness of $\bar{\mathcal{D}}$, lifts to a path in $\bar{S}$ from x to y of l_0 -length $d_{\mathcal{L}}(p(x), p(y))$, so that $d_0(x, y) \leq d_{\mathcal{L}}(p(x), p(y))$. Hence $p|_{\bar{S}}$ is an isometry. $\square$

Theorem 3.12. *The space K_0 is isometric to the metric quotient of the disjoint union $\coprod_{i=1}^2(\mathcal{D}, d_{\mathcal{L}})$ of two copies of the table, obtained by identifying corresponding boundary points.*

Proof. Let $\hat{d}$ be the metric on the disjoint union $X := \coprod_{i=1}^2(\mathcal{D}, d_{\mathcal{L}})$. It is defined as

$$\hat{d}(x, y) := \begin{cases} d_{\mathcal{L}}(x, y), & \text{if } x, y \text{ belong to the same copy of } \mathcal{D}; \\ \infty, & \text{otherwise.} \end{cases}$$

In general, for an equivalence relation R on the disjoint union, the semi-metric d_R “identifying” the R -related points (by assigning distance zero to them) is given by

$$d_R(x, y) := \inf \left\{ \sum_{i=0}^q \hat{d}(x_i, y_i); \ x_0 = x, y_q = y \text{ and } y_i R x_{i+1} \text{ for } i = 1, \dots, q-1 \right\}.$$

In our situation such a semi-metric has the equivalent, simpler description

$$d_{\sim}(x, y) = \inf \{ d_{\mathcal{L}}(x, m_1) + d_{\mathcal{L}}(m_2, y); \ m_1 \sim m_2 \},$$

since it is possible to change from one copy of the table to the other by “passing” through the boundary of $\mathcal{D}$ only once, and a path switching copies more than once can be shortened to one switching once, the identification being the identity on $\partial\mathcal{D}$. We also note that $d_{\sim}(x, y) = 0$ if and only if $x \sim y$, so the only pairs of distinct points (x, x') whose distance is zero are those where x and x' correspond to the same boundary point. In particular, since X is compact the quotients $K \approx X/\sim$ and $X/d_{\sim}$ are homeomorphic³.

We need to show that $\text{id}: (K, d_{\sim}) \rightarrow (K, d_0)$ is an isometry. Due to Lemma 3.11, we already know that the metrics coincide for any two points in the same copy of $\mathcal{D}$, so it is enough to consider $\tilde{x}, \tilde{y}$ lying in distinct connected components of X . Let x, y be the corresponding projected points onto $\mathcal{D}$, and let $[x, y]_{\mathcal{L}}$ be the geodesic of $(\mathcal{D}, d_{\mathcal{L}})$ joining them. There are two cases to analyse.

Case 1: $[x, y]_{\mathcal{L}}$ meets the boundary $\partial\mathcal{D}$. In this situation we have $d_{\sim}(\tilde{x}, \tilde{y}) = l_E([x, y]_{\mathcal{L}}) = d_{\mathcal{L}}(x, y)$. Indeed, given any point $\tilde{m}$ in the boundary $\partial\mathcal{D}$, we have

$$d_{\mathcal{L}}(\tilde{x}, \tilde{m}) + d_{\mathcal{L}}(\tilde{m}, \tilde{y}) = d_{\mathcal{L}}(x, m) + d_{\mathcal{L}}(m, y) \geq d_{\mathcal{L}}(x, y),$$

³ For a semi-metric d on X , we denote by X/d the set of equivalence classes of the relation $d(x, y) = 0$. For a general relation R on X , it is not always true that the quotients X/R and X/d_R are the same. Even when they coincide as sets, in general one can only guarantee that the topology on X/d_R is coarser, i.e., $\text{id}: X/R \rightarrow X/d_R$ is continuous [13].

with equality holding exactly when m is a point in the segment $[x, y]_{\mathcal{L}}$. Thus the infimum $d_{\sim}(\tilde{x}, \tilde{y}) = d_{\mathcal{L}}(x, y)$ is achieved by the assumption that $[x, y]_{\mathcal{L}} \cap \partial\mathcal{D} \neq \emptyset$.

On the other hand, given any $m \in [x, y]_{\mathcal{L}} \cap \partial\mathcal{D}$, we may lift the geodesic segments $[x, m]_{\mathcal{L}}$ and $[m, y]_{\mathcal{L}}$ isometrically to geodesics $[\tilde{x}, m]_0$ and $[m, \tilde{y}]_0$ of K_0 . Hence the path of minimal length $[\tilde{x}, \tilde{y}]_0$ has length at most $l_0([\tilde{x}, m]_0) + l_0([m, \tilde{y}]_0)$. But the projection map is non-expanding, that is, $d_{\mathcal{L}}(x, y) \leq d_0(\tilde{x}, \tilde{y})$, thus $d_0(\tilde{x}, \tilde{y}) = d_{\mathcal{L}}(x, y) = d_{\sim}(\tilde{x}, \tilde{y})$, as we wanted.

Case 2: $[x, y]_{\mathcal{L}} \cap \partial\mathcal{D} = \emptyset$. If the intersection is empty, then the geodesic $[x, y]_{\mathcal{L}}$ is a straight line contained entirely in $\text{int}(\mathcal{D})$. Then no $d_{\mathcal{L}}$ -geodesic from x to y meets $\partial\mathcal{D}$, and we argue directly. For any $m \in \partial\mathcal{D}$, lifting $[x, m]_{\mathcal{L}}$ to the sheet of $\tilde{x}$ and $[m, y]_{\mathcal{L}}$ to the sheet of $\tilde{y}$ (Lemma 3.11) gives a path in K from $\tilde{x}$ to $\tilde{y}$ through the seam point over m , of l_0 -length $d_{\mathcal{L}}(x, m) + d_{\mathcal{L}}(m, y)$; taking the infimum over m gives $d_0(\tilde{x}, \tilde{y}) \leq d_{\sim}(\tilde{x}, \tilde{y})$. Conversely, every rectifiable path γ in K from $\tilde{x}$ to $\tilde{y}$ crosses the seam at some point over some $m \in \partial\mathcal{D}$, so that, by Theorem 3.6, $l_0(\gamma) = l_E(p \circ \gamma) \geq d_{\mathcal{L}}(x, m) + d_{\mathcal{L}}(m, y) \geq d_{\sim}(\tilde{x}, \tilde{y})$; hence $d_0(\tilde{x}, \tilde{y}) \geq d_{\sim}(\tilde{x}, \tilde{y})$, which concludes (the same two inequalities also prove Case 1. In Case 2 the minimising paths are the billiard paths $[x, m]_{\mathcal{L}} * [m, y]_{\mathcal{L}}$ through a point m where the entry and exit angles agree, or paths through a tangency point). $\square$

3.1.3. Comparison of the lengths for d_0 and d_{ε} , $\varepsilon \in [0, 1]$. The goal of this part is to show certain estimates on the lengths for $\{d_{\varepsilon}\}_{\varepsilon \in [0, 1]}$ which will be useful in the following (see e.g. Lemma 3.22). We start by deriving a boundary-layer normal form.

Lemma 3.13. *There exist $0 < \delta_1 < \tau_{\min}/2$ and $C_1 > 0$ such that, for every scatterer Γ_i , there are functions $z_i \in C^r(\Gamma_i)$ and $R_i \in C^{r-2}(\Gamma_i \times (0, \delta_1])$ with the following property. Given a connected component S of K° and its height function $\phi_S := z \circ (p|_S)^{-1}$, if x lies in the collar $\{d_i(x) := d(x, \mathcal{O}_i) \leq \delta_1\}$, then*

$$\phi_S(x) = z_i(u(x)) + s_{S,i} R_i(u(x), d_i(x)), \quad s_{S,i} \in \{-1, +1\},$$

where $u(x) \in \Gamma_i$ is the point of Γ_i nearest to x , and R_i satisfies, uniformly in u and $0 < d \leq \delta_1$,

$$R_i(u, 0) = 0, \quad 0 < \partial_d R_i(u, d) \leq \frac{C_1}{\sqrt{d}}, \quad |\partial_u R_i(u, d)| \leq C_1 \sqrt{d}, \quad R_i(u, d) \leq C_1 \sqrt{d}.$$

Moreover, away from these collars the functions ϕ_S have uniformly bounded gradient.

Proof. Fix a boundary point $x_0 \in B(K)$, projecting to $u_0 \in \Gamma_i$. Let σ be the normal section obtained by intersecting K with the affine plane $x_0 + (\mathbb{R}\partial_z \oplus T_{x_0}^\perp K)$, and set $\zeta := z - z_i$, where $z_i := z \circ (p|_{B(K)})^{-1} \in C^r(\Gamma_i)$, where the inverse branch is chosen according to the sheet S . We write $\sigma = \sigma(n(\zeta), \zeta)$, where n is the inward normal coordinate to Γ_i . By the computation of the kernel of dp along $B(K)$ (see the discussion after (K2)) $n'(0) = 0$ and (K2) gives $n''(0) \neq 0$; by Taylor, we then have

$$n = \frac{a(u_0)}{2} \zeta^2 + O(\zeta^3), \quad a(u_0) > 0,$$

and, by compactness of $B(K)$, $a(u_0) \geq a_0 > 0$ uniformly. Solving for ζ , we get

$$\zeta = \pm \sqrt{\frac{2n}{a(u_0)}} (1 + O(\sqrt{n})),$$

and the expansion is uniform along $B(K)$. Shrinking the tubular neighbourhood so that the collars are disjoint and of width $< \tau_{\min}/2$, and using that the nearest-point projection u onto a strictly convex obstacle is well-defined and 1-Lipschitz with $n = d_i$, we obtain

$$\phi_S(x) = z_i(u) + s_{S,i}(\psi_i(u)\sqrt{d_i} + d_i E_i(u, d_i)), \quad \psi_i := \sqrt{\frac{2}{a}},$$

with E_i bounded. Setting $R_i(u, d) := \psi_i(u)\sqrt{d} + d E_i(u, d)$, all the stated estimates follow (for the derivative bound, note that E_i is a smooth function of $(u, \sqrt{d})$, so $\partial_d(d E_i)$ is bounded while $\partial_d(\psi_i\sqrt{d}) = \psi_i/(2\sqrt{d})$), after possibly shrinking δ_1 and enlarging C_1 . Away from the collars, K° is a C^r graph over $\mathcal{D}$ by (K1), so $\nabla\phi_S$ is uniformly bounded. $\square$

Let us now study the variation of ϕ_S along a chord:

Lemma 3.14. *There is $C_2 = C_2(K) > 0$ such that for every straight segment $J \subset \mathcal{D}$ of Euclidean length $\leq \tau_{\max}$ and every sheet S , we have $\text{Var}(\phi_S|_J) \leq C_2$.*

Proof. Away from the collars, Lemma 3.13 gives a uniform bound $\text{Var}(\phi_S|_{J \cap \{d_i > \delta_1\}}) \leq M|J| \leq M\tau_{\max}$; it remains to bound the contribution of a single collar visit. Lift to $\mathbb{R}^2$ and let $\tilde{\mathcal{O}}$ be a $\mathbb{Z}^2$ -translate of $\mathcal{O}_i$ whose δ_1 -collar meets J ; since $|J| \leq \tau_{\max}$, the number of such translates is bounded by $N = N(K)$. On the collar write $d(s) = \text{dist}(J(s), \tilde{\mathcal{O}})$; as $\tilde{\mathcal{O}}$ is convex and J is straight, d is convex, hence has at most two monotone pieces, so $\text{Var}(\sqrt{d}|_J) \leq 2\sqrt{\delta_1}$. Writing $\phi_S = z_i(u) + s_{S,i}R_i(u, d_i)$ on the collar, for two points $(u_1, d_1), (u_2, d_2)$ we split

$$\begin{aligned} |\phi_S(x_2) - \phi_S(x_1)| &\leq |R_i(u_2, d_2) - R_i(u_2, d_1)| \\ &\quad + |z_i(u_2) - z_i(u_1) + s_{S,i}(R_i(u_2, d_1) - R_i(u_1, d_1))|. \end{aligned}$$

Since $0 < \partial_d R_i \leq C_1/\sqrt{d}$, the first term is bounded by $C_1 |\sqrt{d_2} - \sqrt{d_1}|$ uniformly in u_2 ; summing over a partition of J and using that $\sqrt{d}$ has at most two monotone pieces, the total contribution is at most $C_1 \text{Var}(\sqrt{d}|_J) \leq 2C_1\sqrt{\delta_1}$. For the second, $|\partial_u R_i| \leq C_1\sqrt{\delta_1}$ and $z_i \in C^1$, and u is 1-Lipschitz, so this term is bounded by $(\text{Lip}(z_i) + C_1\sqrt{\delta_1})\tau_{\max}$. We conclude by summing over the at most N visits and the two regimes (collar/exterior). $\square$

In the next lemma, we study the vertical variation of d_0 -geodesics.

Lemma 3.15. *There is $C_3 = C_3(K) > 0$ such that every unit-speed d_0 -geodesic $\gamma: [0, L] \rightarrow K$ satisfies $\text{Var}(z \circ \gamma) \leq C_3(L + 1)$.*

Proof. By Theorem 3.12, γ decomposes into free arcs (maximal subarcs in a single sheet, projecting to straight chords of $\mathcal{D}$) and sliding arcs (maximal subarcs in $B(K)$), separated by isolated transversal crossings. Each free arc projects to a straight segment of $\mathcal{D}$; since a tangential contact does not change sheet, it is a chord of Euclidean length at most $\tau_{\max}$ or the union of two such chords through a tangency point. In both cases Lemma 3.14 bounds its vertical variation by $2C_2$. Along a sliding arc, z is C^r on the boundary curve, so $z \circ \gamma$ has derivative bounded by some C_4 , contributing at most $C_4 \times (\text{its } d_0\text{-length})$. By finite horizon every complete free flight has length at least $\tau_{\min}$, so among the free arcs meeting $\gamma([0, L])$ all but the first and last have d_0 -length at least $\tau_{\min}$; their number is less than $L/\tau_{\min} + 2$. The total d_0 -length of sliding arcs is at most L . Hence

$$\text{Var}(z \circ \gamma) \leq 2C_2\left(\frac{L}{\tau_{\min}} + 2\right) + C_4L \leq C_3(L + 1). \quad \square$$

We can finally compare the lengths for d_0 and d_ε , $\varepsilon \in (0, 1]$:

Theorem 3.16. *There is $C = C(K) > 0$ such that, for every g_1 -rectifiable $\gamma: [a, b] \rightarrow K$ parametrised by arc length and every $\varepsilon \in (0, 1]$,*

$$l_0(\gamma) \leq l_\varepsilon(\gamma) \leq l_0(\gamma) + C \text{Var}(z \circ \gamma) \varepsilon.$$

Therefore, for all $\varepsilon \in (0, 1]$ and all $x, y \in K$,

$$d_0(x, y) \leq d_\varepsilon(x, y) \leq (1 + C\varepsilon) d_0(x, y) + C\varepsilon,$$

and $d_\varepsilon \downarrow d_0$ pointwise and monotonically. Moreover, if γ is a unit-speed d_ε -geodesic on $\widetilde{K}_\varepsilon$, then:

- (a) $\frac{|t-s| - C\varepsilon}{1 + C\varepsilon} \leq d_0(\gamma(t), \gamma(s)) \leq |t-s|$; in particular γ is a $(1 + C\varepsilon, C\varepsilon)$ -quasi-geodesic of $(\widetilde{K}, d_0)$;
- (b) $|t-s| \leq d_1(\gamma(t), \gamma(s)) \leq (1 + C)|t-s| + C$; in particular γ is a $(1 + C, C)$ -quasi-geodesic of $(\widetilde{K}, d_1)$.

Proof. The lower bound follows from $g_\varepsilon \geq g_0$. Writing $v_H(t) = |dp(\gamma'(t))|$, $v_V(t) = |dz(\gamma'(t))|$, we have $l_\varepsilon(\gamma) = \int_a^b \sqrt{v_H^2 + \varepsilon^2 v_V^2} dt$. Noting that $\sqrt{v_H^2 + \varepsilon^2 v_V^2} \leq v_H + \varepsilon v_V$, and since $z \circ \gamma$ is absolutely continuous, we have

$$l_\varepsilon(\gamma) \leq l_0(\gamma) + \varepsilon \int_a^b v_V dt = l_0(\gamma) + \varepsilon \text{Var}(z \circ \gamma).$$

Taking $\gamma = [x, y]_0$, and by Lemma 3.15, we obtain the distance bounds. On $\widetilde{K}$ the height $z_{\widetilde{K}} := z \circ \Pi$ (with $\Pi: \widetilde{K} \rightarrow K$ the covering) is $\pi_1(K)$ -invariant and hence bounded, so the same estimates hold on each $(\widetilde{K}_\varepsilon, d_\varepsilon)$. Since $\widetilde{K}_\varepsilon$ has no conjugate points for every $\varepsilon \in (0, 1]$ (by (K3) and Klingenberg's theorem, as in the proof of Lemma 3.3), geodesics are globally minimising, and evaluating the distance bounds on $(\gamma(t), \gamma(s))$ gives (a) (from $d_\varepsilon(\gamma(t), \gamma(s)) = |t-s|$) and (b) (from $d_1 \geq d_\varepsilon$ and $d_1 \leq (1 + C)d_0 + C \leq (1 + C)d_\varepsilon + C$). $\square$

3.2. Coarse geometry of $K_0 = (K, d_0)$. Recall that, by (K0) and Theorem 3.12, $K_0 = (K, d_0)$ consists of two copies of $(\mathcal{D}, d_\mathcal{L})$ glued along the boundary.

Let $\widetilde{K}_0$ be the universal covering of K . For simplicity, when $\varepsilon > 0$, we denote the lift of the metric d_ε to $\widetilde{K}_0$ by d_ε as well, and the same for the lift of the Riemannian tensor g_ε . We denote by $\widetilde{K}_\varepsilon$ the space $(\widetilde{K}, d_\varepsilon)$. The canonical projection $\widetilde{K}_\varepsilon \rightarrow K_\varepsilon$ is then a smooth covering and a local isometry.

Let us recall the following notion:

Definition 3.17. A metric space is called *non-positively curved*, or *locally CAT(0)*, if every point has a neighbourhood on which every geodesic triangle satisfies the CAT(0) inequality.

Proposition 3.18. *The metric space $K_0 = (K, d_0)$ is non-positively curved. The universal covering $\widetilde{K}_0$ is a contractible Hadamard space, that is, a contractible and complete (globally) CAT(0) geodesic space.*

Proof. Any point of $K_0 \setminus B(K)$ has a neighbourhood isometric to an open subset of $\mathbb{R}^2$, where the CAT(0) inequality holds with equality. Let then $p \in B(K)$. A neighbourhood of p in K_0 is isometric to the double, across its boundary curve, of a plane region

$$U = \{(x, y) \in B_r(0) : y \leq -g(x)\},$$

where g is convex and C^2 (the local graph of the scatterer); indeed the two sheets are two copies of the table, glued along the wall.

Let us first pretend that g is piecewise affine and convex. The double of U is then a flat surface with isolated conical singularities, one for each vertex of the graph of g ; at a vertex with exterior turning angle $\tau \geq 0$ (by convexity), the interior angle of U is $\pi + \tau$, so the cone angle of the double is $2\pi + 2\tau \geq 2\pi$. A Euclidean cone

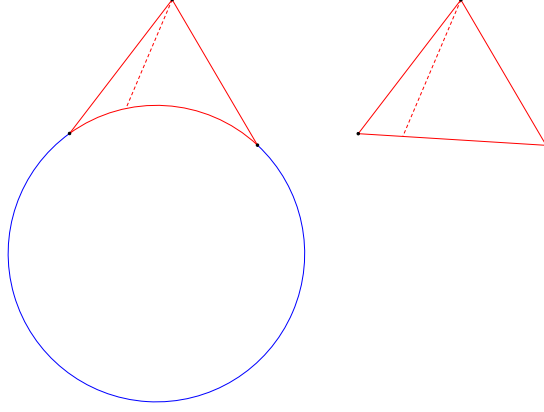


FIGURE 7. The local model at the seam: the double of the region below a convex graph. In the piecewise affine approximation, each vertex contributes a cone point of angle $2\pi + 2\tau \geq 2\pi$.

of angle $\geq 2\pi$ is CAT(0), and a locally finite flat surface with cone angles $\geq 2\pi$ is locally CAT(0) (see [11, Chapter II.5]); the smooth points of the seam are flat in the double, the two adjacent angles being π and π (see Figure 7).

For a smooth convex function g describing a piece of boundary, we approximate g uniformly by convex piecewise affine functions g_n (e.g. the maxima of finitely many tangent lines). The doubles of the corresponding regions U_n converge to the double of U : parametrisising both boundaries by arc length and extending sheetwise, one obtains homeomorphisms which distort lengths of curves, hence distances, by a factor $1 + o(1)$. The CAT(0) comparison inequality is a closed condition under such uniform convergence of metrics [11, II.3.10], so the double of U is CAT(0) for small r , and K_0 is non-positively curved. Alternatively, it follows from Reshetnyak's gluing theorem for two-dimensional manifolds of bounded curvature: two flat pieces are glued along a curve whose geodesic curvatures relative to the two sides sum to $-2K \leq 0$; see [42].

The Cartan-Hadamard theorem for complete locally CAT(0) spaces (see e.g. [11, Theorem II.4.1]) then gives the assertion about the universal covering.

Finally, we note that each closed sheet of K_0 is locally convex: through a smooth seam point the seam is a local geodesic of K_0 (both adjacent angles equal π), and a d_0 -geodesic between two nearby points of one closed sheet cannot enter the other sheet, since it would have to cross the seam twice, contradicting local uniqueness of geodesics between its crossing points. Consequently $(\mathcal{D}, d_{\mathcal{L}})$, which is isometric to a closed sheet, is itself locally CAT(0), being a locally convex subspace of the locally CAT(0) space K_0 . $\square$

Let us now recall some standard definitions on Gromov hyperbolicity.

Definition 3.19 (Gromov hyperbolicity). Let (X, d) be a geodesic metric space and $\delta \geq 0$. A geodesic triangle in X is δ -thin if each of its sides is contained in the closed δ -neighbourhood of the union of the other two. The space X is δ -hyperbolic if all its geodesic triangles are δ -thin; it is called *Gromov-hyperbolic* if it is δ -hyperbolic for some δ .

Recall that hyperbolicity is invariant under quasi-isometries between geodesic spaces: if X is δ -hyperbolic and $f: X' \rightarrow X$ is a (λ, c) -quasi-isometry from a geodesic

space, then X' is δ' -hyperbolic with δ' depending only on (λ, c, δ) [11, Theorem III.H.1.9]. In the following, we use two sources of hyperbolicity:

- a proper cocompact CAT(0) space containing no isometrically embedded flat plane is Gromov-hyperbolic (Flat Plane Theorem, [11, Theorem III.H.1.5]);
- the universal cover of a closed Riemannian manifold with Anosov geodesic flow is Gromov-hyperbolic [27, Appendix B].

Definition 3.20 (Gromov boundary). Let (X, d) be a proper geodesic δ -hyperbolic space. Two geodesic rays $r, r': [0, \infty) \rightarrow X$ are *asymptotic* if $\sup_{t \geq 0} d(r(t), r'(t)) < \infty$. The *boundary at infinity* (or *Gromov boundary*) $\partial_\infty X$ is the set of asymptotic classes of geodesic rays; the class of r is denoted $r(\infty)$, and for a bi-infinite geodesic γ we write $\gamma(\pm\infty)$ for the classes of its two half-rays. Finally we denote by $\partial_\infty^2 X := (\partial_\infty X \times \partial_\infty X) \setminus \Delta$ the space of pairs of distinct boundary points.

Let (X, d) be a proper geodesic δ -hyperbolic space. For every $x \in X$ and $\xi \in \partial_\infty X$ there is a geodesic ray issued from x in the class ξ [11, Lemma III.H.3.1]; when X is CAT(0) this ray is unique. The union $\bar{X} := X \cup \partial_\infty X$ carries a natural topology in which a sequence $x_n \in X$ converges to $\xi \in \partial_\infty X$ if and only if, for one (equivalently, every) base point x , the geodesic segments $[x, x_n]$ converge uniformly on compact sets to a ray issued from x in the class ξ ; for CAT(0) spaces this is the cone topology of [11, II.8]. With this topology $\bar{X}$ is compact, X is open and dense, and $\partial_\infty X$ is closed [11, Proposition III.H.3.7]; moreover every quasi-isometry between proper geodesic hyperbolic spaces extends to a homeomorphism of their boundaries [11, Theorem III.H.3.9].

Finally, we recall the notion of Gromov product and some of its basic properties.

Definition 3.21 (Gromov products). Let (X, d) be a metric space. The *Gromov product* of $y, z \in X$ at $x \in X$ is

$$(y \mid z)_x := \frac{1}{2} [d(x, y) + d(x, z) - d(y, z)] \geq 0.$$

Below, we will use the following standard facts:

- (i) $(y \mid z)_x \leq \min\{d(x, y), d(x, z)\}$, with $(y \mid z)_x = 0$ if and only if, in a geodesic space, x lies on a geodesic from y to z .
- (ii) Isometry invariance: $(\psi y \mid \psi z)_{\psi x} = (y \mid z)_x$ for every isometry ψ .
- (iii) In any geodesic space $(y \mid z)_x \leq d(x, [y, z])$ for every geodesic segment $[y, z]$, and in a δ -hyperbolic geodesic space the product computes this distance up to an additive constant:

$$(y \mid z)_x \leq d(x, [y, z]) \leq (y \mid z)_x + 4\delta.$$

- (iv) Let X be proper, geodesic and δ -hyperbolic, with boundary $\partial_\infty X$ as in Definition 3.20. The product controls convergence at infinity: a sequence (x_n) converges to a point of $\partial_\infty X$ if and only if $(x_n \mid x_m)_x \rightarrow \infty$ as $n, m \rightarrow \infty$, and two such sequences have the same limit iff $(x_n \mid y_n)_x \rightarrow \infty$ [17, Ch. 7]. Thus, for $\xi \neq \eta \in \partial_\infty X$ and sequences $y_n \rightarrow \xi$, $z_n \rightarrow \eta$ in $\bar{X}$, the products $(y_n \mid z_n)_x$ remain bounded, and any two limit values differ by at most 2δ .

In the following lemma, we show that the spaces $\widetilde{K}_\varepsilon$, $\varepsilon \in [0, 1]$, are uniformly Gromov hyperbolic.

Lemma 3.22. *There exists $\delta = \delta(K) > 0$ such that the spaces $\widetilde{K}_\varepsilon$, $\varepsilon \in [0, 1]$, are δ -hyperbolic.*

Proof. Since $g_\varepsilon \leq g_{\varepsilon'}$ pointwise for $\varepsilon \leq \varepsilon'$, the distances are monotone: $d_0 \leq d_\varepsilon \leq d_1$ on $\widetilde{K}$ for every $\varepsilon \in [0, 1]$. By Theorem 3.16 at $\varepsilon = 1$, $d_1 \leq (1 + C)d_0 + C$, hence

$$d_\varepsilon \leq d_1 \leq (1 + C)d_\varepsilon + C, \quad \forall \varepsilon \in [0, 1],$$

so that the identity map $(\widetilde{K}, d_\varepsilon) \rightarrow (\widetilde{K}, d_1)$ is a $(1 + C, C)$ -quasi-isometry, with constants independent of ε . All the spaces involved are geodesic: the $(\widetilde{K}, d_\varepsilon)$, $\varepsilon \in (0, 1]$, are complete Riemannian manifolds, while for $(\widetilde{K}, d_0)$ it follows from Theorem 3.6. Finally $(\widetilde{K}, d_1)$ is Gromov-hyperbolic, the geodesic flow of g_1 being Anosov by (K3) (Definition 3.19). Since hyperbolicity is stable under quasi-isometries between geodesic spaces, with a constant depending only on the quasi-isometry constants and on $\delta(g_1)$ (Definition 3.19), the spaces $\widetilde{K}_\varepsilon$ are δ -hyperbolic, for some $\delta = \delta(K) > 0$. $\square$

The space of (parametrised) geodesics of $\widetilde{K}_\varepsilon$, $\varepsilon \in [0, 1]$, is the subset $\widehat{\mathcal{G}}_\varepsilon \subset C(\mathbb{R}, \widetilde{K})$ consisting of unit speed parametrisations⁴ of geodesics $\gamma: \mathbb{R} \rightarrow \widetilde{K}_\varepsilon$. We define the equivalence relation $\sim$ on $\widehat{\mathcal{G}}_\varepsilon$ by declaring $\gamma_1 \sim \gamma_2$ if and only if there is $T \in \mathbb{R}$ such that $\gamma_1(\cdot) = \gamma_2(\cdot + T)$, and let $\mathcal{G}_\varepsilon$ be the space of unparametrised geodesics

$$\mathcal{G}_\varepsilon := \widehat{\mathcal{G}}_\varepsilon / \sim = \{[\gamma]; d_\varepsilon(\gamma(t), \gamma(s)) = |t - s|, t, s \in \mathbb{R}\},$$

equipped with the compact-open topology inherited from $C(\mathbb{R}, \widetilde{K})$. In the Riemannian cases (i.e., for $\varepsilon > 0$), there is an identification between $\widehat{\mathcal{G}}_\varepsilon$ and $T^1\widetilde{K}_\varepsilon$ given by

$$\gamma \mapsto (\gamma(0), \dot{\gamma}(0)),$$

and in this case the space $\mathcal{G}_\varepsilon$ is simply the space of orbits of the geodesic flow acting on $T^1\widetilde{K}_\varepsilon$. While there is no unit tangent bundle of $\widetilde{K}_0$ and therefore no such identification, we can still define an abstract geodesic flow on $\widetilde{K}_0$ as the $\mathbb{R}$ -action on $\widehat{\mathcal{G}}_0$ by precomposition with translations:

$$\begin{aligned} \mathbf{g}_0: \widehat{\mathcal{G}}_0 \times \mathbb{R} &\rightarrow \widehat{\mathcal{G}}_0 \\ (\gamma, t) &\mapsto \gamma(t + \cdot). \end{aligned}$$

By Lemma 3.22, every $(\widetilde{K}, d_\varepsilon)$, $\varepsilon \in [0, 1]$, is a proper geodesic Gromov-hyperbolic space. Therefore, it has a boundary at infinity $\partial_\infty \widetilde{K}_\varepsilon$ as in Definition 3.20, which for the CAT(0) spaces is the boundary of asymptotic geodesic rays with the cone topology, and $\widetilde{K}_\varepsilon \cup \partial_\infty \widetilde{K}_\varepsilon$ is compact. Since $\widetilde{K}_\varepsilon$ is proper, it is a visibility space [11, Lemma III.H.3.2]: any two distinct points $\alpha, \omega \in \partial_\infty \widetilde{K}_\varepsilon$ are joined by a geodesic, which is unique for $\varepsilon > 0$ since a Hadamard manifold with Anosov geodesic flow contains no flat strips (Flat Strip Theorem, [11, II.2.13]); the same argument is carried out for $\varepsilon = 0$ in Lemma 3.31 below. Therefore, for any $\varepsilon \in [0, 1]$, we have

$$\mathcal{G}_\varepsilon \approx \partial_\infty^2 \widetilde{K}_\varepsilon = (\partial_\infty \widetilde{K}_\varepsilon \times \partial_\infty \widetilde{K}_\varepsilon) \setminus \Delta.$$

3.2.1. The space $\mathcal{G}_0$ and the abstract geodesic flow. Based on the relation between geodesics of $\widetilde{K}_0$, K_0 and $(\mathcal{D}, d_{\mathcal{L}})$, we classify geodesics in terms of their projections on the table $\mathcal{D}$. Let us recall that in the gluing construction, the seam is the subset $B(K) \subset K_0$ corresponding to the boundary points where the identification happens. In other words, $B(K)$ consists of closed curves on K whose image under the projection $p: E \rightarrow \mathbb{T}^2$ are the scatterers on $\mathcal{D}$.

First, we associate to a geodesic a set of collision points; these are the points at which the corresponding trajectory segment on $\mathcal{D}$ bounces off a scatterer.

⁴ In a general metric space like $\widetilde{K}_0$, a parametrisation γ is unit-speed if $d(\gamma(t), \gamma(s)) = |t - s|$ for every $t, s \in \mathbb{R}$.

Definition 3.23. For a geodesic γ in K_0 , the set of *collision points* of γ is defined as

$$B(\gamma) := \partial(\gamma(\mathbb{R}) \cap B(K)).$$

The intersection of a segment of γ with $B(K)$ could be a single point (corresponding under p to a bounce of the associated trajectory on $\mathcal{D}$) or an entire segment of a scatterer on the table, in which case the geodesic does not correspond to any billiard trajectory. This motivates the following terminology.

Definition 3.24 (Geodesic trichotomy). We will say a collision point $x \in K_0$ is *degenerate* if the connected component of $\gamma(\mathbb{R}) \cap B(K)$ to which it belongs has positive length, i.e., is not an isolated point.

A non-degenerate point is said to be *regular* or *grazing* if the corresponding collision of $p(\gamma)$ is regular or grazing, respectively.

Similarly, given a geodesic $\tilde{\gamma}$ on $\tilde{K}_0$, let γ be its projection on K_0 . Then the set $B(\tilde{\gamma})$ of *collision points* of $\tilde{\gamma}$ is defined as the pre-image of $B(\gamma)$ under the canonical projection. A collision point of $\tilde{\gamma}$ will be called regular, grazing or degenerate when its projection on K_0 has the corresponding property.

A geodesic is *degenerate* if it contains a degenerate collision point. A non-degenerate geodesic is *grazing* if it contains a grazing point. Otherwise it is *regular*. Let

$$\mathcal{G}^{\text{reg}} := \{[\gamma] \in \mathcal{G}_0; \gamma \text{ is regular}\},$$

and write $\mathcal{G}^{\text{graz}}$ and $\mathcal{G}^{\text{deg}}$ for the sets of grazing and of degenerate geodesics, respectively. We denote by $\mathcal{G}(\mathcal{D})$ the closure of $\mathcal{G}^{\text{reg}}$ in $\mathcal{G}_0$ with respect to the compact-open topology. We call $\mathcal{G}(\mathcal{D})$ the space of *admissible* geodesics.

The non-degenerate geodesics are those which project to genuine billiard trajectories on $\mathcal{D}$, while the degenerate geodesics project to curves which are tangent to a scatterer Γ_i for some positive time interval $[t, t + \varepsilon]$. In particular, degenerate geodesics cannot be the limit of regular ones, so that admissible geodesics are non-degenerate, as we make explicit in the following statement.

Lemma 3.25. *The set $\mathcal{G}^{\text{deg}}$ of degenerate geodesics is open in $\mathcal{G}_0$; in particular, every admissible geodesic is non-degenerate. Hence*

$$\mathcal{G}^{\text{reg}} \subseteq \mathcal{G}(\mathcal{D}) \subseteq \mathcal{G}^{\text{nd}} := \mathcal{G}^{\text{reg}} \sqcup \mathcal{G}^{\text{graz}},$$

i.e., admissible geodesics are regular or grazing.

Proof. We start with a quantitative estimate on chords that hug a scatterer. Since the scatterers are compact and strictly convex, their boundary curves have curvature bounded below by some $\mathcal{K}_{\min} > 0$. We claim there is a constant $c_0 > 0$, depending only on $\mathcal{D}$, such that any straight segment of length ℓ lying entirely within distance η of a curve of curvature at least $\mathcal{K}_{\min}$ satisfies $\ell \leq c_0 \sqrt{\eta}$. Indeed, let J be such a segment and u a point of the curve nearest to the midpoint of J . If J is not nearly tangent to the curve at u , it leaves the η -neighbourhood within length $O(\eta)$ and the estimate is trivial. Assume therefore J is nearly tangent at u . In the tangent frame at u the curve is the graph of a function h with $h(0) = h'(0) = 0$ and $h(s) \geq \frac{\mathcal{K}_{\min}}{2} s^2 (1 + O(s))$, while J is the graph of an affine function a . Evaluating $h - a$ at the two endpoints and the midpoint of J , the affine part cancels in the second difference; since each of the three values is at most $C\eta$, we obtain

$$\frac{\mathcal{K}_{\min}}{4} \ell^2 (1 + O(\ell)) \leq 4C\eta,$$

which gives $\ell \leq c_0 \sqrt{\eta}$ for η small, hence for all η after enlarging c_0 .

We use this estimate to keep non-degenerate geodesics away from sliding arcs. Fix $s > 0$ and let $\eta_0 = \eta_0(s)$ be small enough that $c_0\sqrt{\eta_0} < \min(\tau_{\min}, s/4)$. We claim that no non-degenerate geodesic can remain within distance $\eta < \eta_0$ of a sliding arc of length s throughout a parameter window of length s . Indeed, on such a window a non-degenerate geodesic is a concatenation of straight chords, of which at most two meet the endpoints of the window while the remaining ones are complete free flights of length at least $\tau_{\min}$ (due to our finite horizon assumption). If the window contains at least two collisions, one such complete chord is present; if it contains exactly one collision, the two truncated chords have total length s , so one of them has length at least $s/2$; and if it contains none, a single chord spans the whole window and has length at least s . In every case the window carries a chord of length at least $\min(\tau_{\min}, s/2)$. Since d_0 dominates the Euclidean distance on each sheet, that chord projects to a straight table segment lying within η of the boundary arc Γ_i carrying the sliding arc. But then it would have length at most $c_0\sqrt{\eta} < \min(\tau_{\min}, s/4)$, which is impossible.

Now let $\{\gamma_n\}_n$ be a sequence of non-degenerate geodesics, and suppose $\gamma_n \rightarrow \gamma$ in the compact-open topology, with γ degenerate, i.e., γ slides along a seam arc over a window of positive length. Shrinking it we may take a sliding arc of length s over a window of length s . For n large, γ_n is within $\eta < \eta_0(s)$ of γ uniformly on that window, hence within η of the sliding arc throughout it, contradicting the previous paragraph. Therefore the non-degenerate geodesics form a closed set, $\mathcal{G}^{\deg}$ is open, and since $\mathcal{G}^{\text{reg}}$ is contained in this closed set, $\mathcal{G}(\mathcal{D}) = \overline{\mathcal{G}^{\text{reg}}}$ consists of non-degenerate geodesics.

The inclusion $\mathcal{G}^{\text{reg}} \subseteq \mathcal{G}(\mathcal{D})$ holds by definition, and the trichotomy gives the last statement. $\square$

Remark 3.26. We note that, in K_0 , branching of geodesics occurs only at grazing collisions: at $(x, v) \in T^1K^\circ$ the geodesics $\gamma_{\varepsilon;v}$ converge locally uniformly, on collision-free windows, to a well-defined geodesic $\gamma_{0;v}$, by Corollary 3.5; moreover, transversal crossings have a unique continuation, by Lemma 3.48(iii).

The following is immediate from the above definitions.

Remark 3.27. The metric space K_0 carries a canonical isometric involution $\iota: K_0 \rightarrow K_0$ exchanging the two sheets and fixing the seam $B(K)$ pointwise. A non-degenerate geodesic changes sheet at every transversal collision, and may or may not change sheet at a grazing one; consequently, every billiard trajectory of $\mathcal{D}$ with only transversal collisions admits exactly two non-degenerate lifts to K_0 , interchanged by ι , while a trajectory with $k < \infty$ grazing collisions admits 2^{k+1} non-degenerate lifts.

Lemma 3.28. *The projection induces a continuous surjection $\mathcal{G}^{\text{nd}} \rightarrow \Omega/\Psi$ from the space of non-degenerate geodesics onto the orbit space of the billiard flow. Over a trajectory with only transversal collisions it is exactly two-to-one, the fibre being an orbit of the involution ι ; in particular it induces a bijection $\mathcal{G}^{\text{reg}}/\iota \rightarrow \Omega^{\text{reg}}/\Psi$. Over a trajectory with $k < \infty$ grazing collisions the fibre has 2^{k+1} elements.*

Proof. A non-degenerate geodesic of K_0 is a concatenation of chords, reflected at transversal collisions (Lemma 3.48(iii)) and continued straight at tangential ones, so it projects to a billiard trajectory; conversely, a billiard trajectory lifts to a non-degenerate geodesic once a starting sheet and, at each grazing collision, a continuation sheet have been chosen. Continuity follows from the continuity of π_K on T^1K° . $\square$

The following result will be useful later to derive that two Sinai billiards sharing the same enriched length spectrum are isometric:

Proposition 3.29. *Let $K_0 = \coprod_{i=1}^2 (\mathcal{D}, d_{\mathcal{L}}) / \sim$ and $K'_0 = \coprod_{i=1}^2 (\mathcal{D}', d'_{\mathcal{L}}) / \sim$ be metric gluings of Sinai billiard tables. If $F: (K_0, d_0) \rightarrow (K'_0, d'_0)$ is an isometry, then the tables $(\mathcal{D}, d_{\mathcal{L}})$ and $(\mathcal{D}', d'_{\mathcal{L}})$ are also isometric; moreover, there is a Euclidean isometry φ of $\mathbb{T}^2$ with $\varphi(\mathcal{D}) = \mathcal{D}'$.*

Proof. An isometry preserves distances and consequently must map points where geodesic bifurcation happens to points with the same property⁵. As we have observed above, the points where geodesics bifurcate are exactly those in the seam, hence F must send the set $B(K)$ exactly onto $B(K')$. But $K_0 \setminus B(K)$ is the disjoint union of two open sheets, each isometric to $(\mathcal{D}, d_{\mathcal{L}})$. Since F carries sheets to sheets, it induces on each one a distance-preserving bijection with the corresponding sheet in K'_0 , i.e., an isometry between $\mathcal{D}$ and $\mathcal{D}'$. Finally, an isometry $f: (\mathcal{D}, d_{\mathcal{L}}) \rightarrow (\mathcal{D}', d'_{\mathcal{L}})$ restricts to a local isometry between the flat surfaces $\text{int}(\mathcal{D})$ and $\text{int}(\mathcal{D}')$ mapping boundary components to boundary components; thus, it lifts to a local isometry between the $\mathbb{Z}^2$ -covers $\mathbb{R}^2 \setminus \bigcup_i \tilde{\mathcal{O}}_i$ and $\mathbb{R}^2 \setminus \bigcup_i \tilde{\mathcal{O}}'_i$, i.e., to a map between connected open subsets of the plane whose differential is everywhere orthogonal and locally constant, hence to the restriction of a Euclidean motion $\tilde{\varphi}$ of $\mathbb{R}^2$. Since $\tilde{\varphi}$ is equivariant with respect to the deck translations, its linear part preserves the lattice $\mathbb{Z}^2$, and $\tilde{\varphi}$ descends to an isometry φ of $\mathbb{T}^2$ with $\varphi(\mathcal{D}) = \mathcal{D}'$. $\square$

Let us lift $\mathcal{G}(\mathcal{D})$ and consider the space $\hat{\mathcal{G}}(\mathcal{D})$ of *parametrised admissible geodesics*. As noted above, in K_0 , branching of geodesics can occur only at grazing collisions. Therefore, as in the Anosov case, at least at the set $\mathcal{R} := \pi_K^{-1}(\Omega^{\text{reg}}) \subset T^1 K$ there is an identification with a subset of geodesics $\hat{\mathcal{G}}_0$ given by

$$\mathcal{R} \ni (x, v) \mapsto \gamma_{0;v} \in \hat{\mathcal{G}}^{\text{reg}} \subset \hat{\mathcal{G}}_0.$$

Moreover, using this representation, the restriction of the abstract geodesic flow on $\hat{\mathcal{G}}_0$ to the set of regular geodesics has the more familiar form

$$\begin{aligned} \mathbf{g}_0^t: \mathcal{R} &\rightarrow \mathcal{R} \\ (x, v) &\mapsto (\gamma_{0;v}(t), \dot{\gamma}_{0;v}(t)) = \lim_{\varepsilon \rightarrow 0} (\gamma_{\varepsilon;v}(t), \dot{\gamma}_{\varepsilon;v}(t)), \end{aligned} \tag{7}$$

which is well defined for $t \in \mathbb{R}$. Finally, Kourganoff's theorem [36] implies that such flow maps are uniformly continuous on compact sets. These flow maps do not extend to a flow on $T^1 \tilde{K}$; indeed, at a vector tangent to the seam, three germs of geodesics share the same initial data (see e.g. Lemma 3.48(iv)), and a grazing geodesic may or may not change sheet at a tangential contact. Rather, writing $\hat{\mathcal{G}}^{\text{nd}} \supseteq \hat{\mathcal{G}}(\mathcal{D})$ for the parametrised non-degenerate geodesics, the map $\hat{\mathcal{G}}^{\text{nd}} \rightarrow T^1 \tilde{K}$, $\gamma \mapsto (\gamma(0), \dot{\gamma}(0^+))$, is a continuous surjection which is bijective over $\mathcal{R}$, with inverse $(x, v) \mapsto \gamma_{0;v}$, and whose fibre over a vector tangent to $B(\tilde{K})$ consists of the non-degenerate geodesics with that initial germ. We may then restate Kourganoff's result as

Theorem 3.30 (Kourganoff). *Let $\mathcal{D}$ be a Sinai billiard table and $K \in \mathcal{K}(\mathcal{D})$. Then the abstract geodesic flow on $\hat{\mathcal{G}}_0$ preserves the set $\hat{\mathcal{G}}^{\text{nd}}$ of non-degenerate geodesics and its closed subset $\hat{\mathcal{G}}(\mathcal{D})$ of admissible geodesics, and the map*

$$\pi_K: \hat{\mathcal{G}}^{\text{nd}} \rightarrow \Omega, \quad \gamma \mapsto \left(p(\gamma(0)), \frac{dp(\dot{\gamma}(0^+))}{|dp(\dot{\gamma}(0^+))|} \right),$$

⁵ In other words, if for any direction there is a single geodesic emanating from x , then the same must be true for $F(x)$.

(which agrees with (2) on geodesics through T^1K°) is a continuous surjection semi-conjugating the abstract flow to the billiard flow Ψ :

$$\pi_K \circ \mathbf{g}_0^t = \Psi^t \circ \pi_K \quad \text{on } \widehat{\mathcal{G}}^{\text{nd}}, \quad t \in \mathbb{R}.$$

Over the trajectories with only transversal collisions, π_K is 2-to-1, its fibres being the orbits of the sheet involution of Remark 3.27; over a trajectory with $k < \infty$ grazing collisions its fibre has 2^{k+1} elements.

3.2.2. The boundary $\partial_\infty \widetilde{K}_0$. Since $\widetilde{K}_0$ is a Hadamard space, it has a well-defined circle at infinity $\partial_\infty \widetilde{K}_0$, defined as the set of equivalence classes of asymptotic rays. More precisely, $\gamma, \gamma': [0, \infty) \rightarrow \widetilde{K}_0$ are in the same class if $\sup\{d_0(\gamma(t), \gamma'(t)); t \in [0, \infty)\} < \infty$. We denote the class of the ray γ by $\gamma(\infty)$. More generally, if $\gamma: \mathbb{R} \rightarrow \widetilde{K}_0$ is a geodesic, we set its endpoints at infinity to be

$$\gamma(\infty) := \gamma^+(\infty), \quad \gamma(-\infty) := \gamma^-(\infty),$$

where $\gamma^\pm: [0, \infty)$ is the geodesic ray defined by $\gamma^\pm(t) = \gamma(\pm t)$.

The compactification $\widetilde{K}_0 \cup \partial_\infty \widetilde{K}_0$ of Definition 3.20 is homeomorphic to a closed disc, because $\widetilde{K}_0$ is a Hadamard space homeomorphic to the plane [11, II.8.11]; the isometric action of $\pi_1(K)$ on $\widetilde{K}_0$ extends continuously to $\partial_\infty \widetilde{K}_0$ [11, II.8.9].

There is a well-defined π_1 -equivariant map

$$\begin{aligned} \partial: \mathcal{G}_0 &\rightarrow \partial_\infty^2 \widetilde{K}_0, \\ [\gamma] &\mapsto (\gamma(-\infty), \gamma(\infty)). \end{aligned}$$

Lemma 3.31. *The space $\widetilde{K}_0$ contains no isometrically embedded flat strip $\mathbb{R} \times [0, \varepsilon]$; a fortiori, no isometrically embedded flat plane.*

Proof. Suppose $\widetilde{K}_0$ contained an isometric copy of $\mathbb{R} \times [0, \varepsilon]$, and consider the family $\{\mathbb{R} \times \{s\}\}_{s \in [0, \varepsilon]}$ of parallel geodesics. We first note that no geodesic in the interior of the strip can be degenerate: if some η_{s_0} , $s_0 \in (0, \varepsilon)$, contained a sliding arc along (a lift of) a scatterer Γ , then, on a sub-window of the sliding arc, the nearby parallels η_s would coincide with equidistant curves of the strictly convex arc Γ inside a flat sheet; such curves have non-vanishing geodesic curvature, contradicting the fact that the η_s are geodesics. We may then assume that the interior geodesics of the strip project to genuine billiard trajectories. These project, via the composition of the covering projection $\widetilde{K}_0 \rightarrow K_0$ with $p: E \rightarrow \mathbb{T}^2$, onto a continuous family of billiard trajectories $\{\eta_s\}$ on $\mathcal{D}$; by the finite-horizon hypothesis, each η_s meets a scatterer within time $\tau_{\max}$, tangentially or transversally. We claim that some interior η_s has a regular (transversal) collision. Indeed, assume by contradiction it is not true: each interior η_s , having no deflections and no sliding arcs, would project to a complete straight line ℓ_s of $\mathbb{T}^2$, tangent to every scatterer it meets. Pick a tangency point q of ℓ_{s_0} at a scatterer Γ , for some interior s_0 . The strip contains parallels on both sides of η_{s_0} , and those on the side of Γ project, near q , to curves at constant small distance δ from ℓ_{s_0} on the side of the obstacle; but the straight line parallel to ℓ_{s_0} at distance δ on that side crosses the strictly convex Γ , while the projections must stay in $\mathcal{D}$, so these projections are not straight lines, a contradiction. Thus some interior η_s collides regularly with some scatterer Γ , and for s' sufficiently close to s the trajectory $\eta_{s'}$ collides transversally with Γ as well. Since the trajectories are parallel and Γ is strictly convex, they are dispersed after the collision, contradicting the fact that the transverse distance is constant in a flat strip. $\square$

As a consequence of Lemma 3.31, we can give an intrinsic proof of the Gromov-hyperbolicity of $\widetilde{K}_0$, independent from that of Lemma 3.22.

Corollary 3.32. *The space $\widetilde{K}_0$ is Gromov-hyperbolic: there is $\delta > 0$ such that every geodesic triangle in $\widetilde{K}_0$ has each side contained in the δ -neighbourhood of the union of the other two.*

Proof. The space $\widetilde{K}_0$ is a proper cocompact CAT(0) space (it is the universal cover of the compact space K_0), and by Lemma 3.31 it contains no isometrically embedded flat plane. By the Flat Plane Theorem (see Definition 3.19), it is Gromov-hyperbolic. $\square$

Lemma 3.33. *The space $\widetilde{K}_0$ is a visibility space, i.e., $\partial: \mathcal{G}_0 \rightarrow \partial_\infty^2 \widetilde{K}_0$ is a homeomorphism. Moreover, any two geodesics in $\mathcal{G}^{\text{reg}}$ intersect at most once, and transversely.*

Proof. Let us first show the existence, i.e., the surjectivity of ∂ . Since $\widetilde{K}_0$ is proper, cocompact, CAT(0) and Gromov-hyperbolic (Corollary 3.32), any two distinct points $\xi \neq \eta$ of $\partial_\infty \widetilde{K}_0$ are joined by a bi-infinite geodesic [11, Lemma III.H.3.2].

Let us now prove uniqueness. If two distinct geodesic lines had the same endpoints at infinity, they would remain at bounded Hausdorff distance and, by the Flat Strip Theorem [11, Theorem II.2.13], would bound a flat strip, contradicting Lemma 3.31; this shows the injectivity of ∂ . Continuity of ∂ and of its inverse for the compact-open and cone topologies then follows from Arzelà–Ascoli together with uniqueness (see [11, Chapter II.8]), hence ∂ is a homeomorphism.

Let us now prove the point on transversality. The Hadamard space $\widetilde{K}_0$ is uniquely geodesic; hence if two geodesic lines meet at two points $p \neq p'$, they coincide along the segment $[p, p']$ and can only separate at p or p' by branching. Since branching occurs only at grazing collision points, two distinct geodesics in $\mathcal{G}^{\text{reg}}$ meet in at most one point; and a tangential intersection at such a point would again be a branching point, which is excluded for regular geodesics. Hence the intersection is transverse. $\square$

The geodesic flows $\{\mathbf{g}_\varepsilon\}_{\varepsilon \in (0,1]}$ are all Anosov; in particular, by structural stability, they are all orbit equivalent to each other. In other words, for any $\varepsilon > 0$, there is a Hölder continuous homeomorphism $H_\varepsilon: T^1 \widetilde{K}_\varepsilon \rightarrow T^1 \widetilde{K}_1$ which is an orbit equivalence between the geodesic flows $\mathbf{g}_\varepsilon$ and $\mathbf{g}_1$ (see [26, 24]): given a geodesic $\gamma \in \mathcal{G}_\varepsilon$, there is a unique geodesic $H_\varepsilon \gamma \in \mathcal{G}_1$ corresponding to γ , and such correspondence is equivariant with respect to the natural action of $\pi_1(K)$ on the space of geodesics, induced by deck transformations.

Lemma 3.34. *For every $\varepsilon \in [0, 1]$ there is a homeomorphism*

$$\rho_\varepsilon: \partial_\infty \widetilde{K}_\varepsilon \rightarrow \partial_\infty \widetilde{K}_1, \quad \rho_1 = \text{id},$$

which is equivariant with respect to the isometric actions of $\pi_1(K)$, i.e.,

$$\rho_\varepsilon(\alpha \cdot \xi) = \alpha \cdot \rho_\varepsilon(\xi) \quad \forall \xi \in \partial_\infty \widetilde{K}_\varepsilon, \forall \alpha \in \pi_1(K).$$

In particular, $\rho_\varepsilon^{-1} \circ \rho_0: \partial_\infty \widetilde{K}_0 \rightarrow \partial_\infty \widetilde{K}_\varepsilon$ identifies, equivariantly, the boundary of the limit space with the boundary of each approximating surface.

Proof. Given a geodesic $\gamma \in \mathcal{G}_\varepsilon$, the geodesic ray $\gamma^+: [0, \infty) \rightarrow K_\varepsilon$ defines a unique boundary point $\gamma(\infty)$. For $\varepsilon \in (0, 1]$, we define ρ_ε by mapping the forward point at infinity of γ to the forward point at infinity of $H_\varepsilon \gamma$:

$$\begin{aligned} \rho_\varepsilon: \partial_\infty \widetilde{K}_\varepsilon &\rightarrow \partial_\infty \widetilde{K}_1, \\ \gamma(\infty) &\mapsto (H_\varepsilon \gamma)^+(\infty). \end{aligned}$$

This extension is well-defined because γ' is another geodesic ray in the class $\gamma(\infty)$ if and only if $(\gamma', \dot{\gamma}')$ lies on the weak (centre-)stable manifold $\mathcal{W}_\gamma^{\text{cs}}$ of $\mathbf{g}_\varepsilon$, the

boundary point being defined by bounded, not vanishing, distance. Since H_ε is an orbit equivalence, it carries weak-stable leaves to weak-stable leaves, hence $H_\varepsilon\gamma'(\infty) = H_\varepsilon\gamma(\infty)$.

As for the continuity of the map ρ_ε on the boundary, we note that the cone topology on $\partial_\infty\widetilde{K}_\varepsilon$ coincides with the compact-open topology on the space of geodesic rays [11, Chapter II.8]. Moreover, if σ is another geodesic sufficiently close to γ along a compact subset of $[0, \infty)$, then $H_\varepsilon\sigma$ will be close to $H_\varepsilon\gamma$ along a corresponding compact, implying that such a map is continuous in the compact-open topology of the space of geodesics (and, in particular, geodesic rays) as we wanted.

Reversing the roles of the indices ε and 1 and applying the exact same construction to H_ε^{-1} provides a continuous inverse, so ρ_ε is a homeomorphism; it is equivariant because H_ε is.

For the case $\varepsilon = 0$, Theorem 3.16(b) shows that every geodesic $\gamma_\varepsilon \in \mathcal{G}_\varepsilon$, seen as a curve in $\widetilde{K}_1$, is a $(1 + C, C)$ -quasi-geodesic, with constants independent of ε :

$$|t - s| \leq d_1(\gamma_\varepsilon(t), \gamma_\varepsilon(s)) \leq (1 + C)|t - s| + C, \quad t, s \in \mathbb{R}.$$

We may now apply the Morse Lemma on stability of quasi-geodesics (see e.g. [11, Theorem III.H.1.7]): there is $R = R(g_1, C) > 0$ such that any geodesic in $\mathcal{G}_\varepsilon$, $\varepsilon \in (0, 1]$ (seen as a curve in $\widetilde{K}_1$) is in a R -neighbourhood of a geodesic in $\mathcal{G}_1$. In other words, given $\gamma_\varepsilon \in \mathcal{G}_\varepsilon$, there is $\gamma_1 \in \mathcal{G}_1$ such that

$$\gamma_\varepsilon(\mathbb{R}) \subset B_1(\gamma_1; R) := \coprod_{t \in \mathbb{R}} \{x \in \widetilde{K}_1; d_1(x, \gamma_1(t)) < R\}.$$

Moreover, since the geodesic flow of g_1 is Anosov, distinct g_1 -geodesics diverge (there are no conjugate points and $\widetilde{K}_1$ is a visibility space), so at most one g_1 -geodesic can stay at bounded distance from $\gamma_\varepsilon(\mathbb{R})$.

For $\varepsilon = 0$ we argue directly. Recall that (see e.g. Theorem 3.16)

$$d_0 \leq d_1 \leq (1 + C)d_0 + C, \quad (8)$$

hence every d_0 -geodesic γ satisfies $|t - s| \leq d_1(\gamma(t), \gamma(s)) \leq (1 + C)|t - s| + C$, so it is a $(1 + C, C)$ -quasi-geodesic of $\widetilde{K}_1$. By the Morse Lemma and the divergence of g_1 -geodesics, it lies within distance R of a unique $\gamma_1 \in \mathcal{G}_1$. A quasi-geodesic ray $c: [0, \infty) \rightarrow \widetilde{K}_1$ defines a unique point at infinity $\xi(c) \in \partial_\infty\widetilde{K}_1$ [11, Lemma III.H.3.1], and we set

$$\begin{aligned} \rho_0: \partial_\infty\widetilde{K}_0 &\rightarrow \partial_\infty\widetilde{K}_1 \\ \gamma(\infty) &\mapsto \xi(\gamma^+). \end{aligned}$$

This is well defined and injective; indeed, two d_0 -rays define the same point of $\partial_\infty\widetilde{K}_1$ if and only if they stay at bounded d_1 -distance, if and only if (by (8)) they stay at bounded d_0 -distance, if and only if they are asymptotic in $\widetilde{K}_0$. It is continuous, because the Morse Lemma is uniform: if $\gamma_n \rightarrow \gamma$ on compact sets, the g_1 -geodesics shadowing γ_n^+ converge on compact sets to the one shadowing γ^+ , hence their ideal points converge. Moreover, it is surjective: given $\xi \in \partial_\infty\widetilde{K}_1$, let γ be a g_1 -geodesic with $\gamma(\infty) = \xi$; the curves $H_\varepsilon^{-1}\gamma$, $\varepsilon \in (0, 1]$, parametrised by d_ε -arc length, are 1-Lipschitz for d_ε , hence for d_0 , and they stay within d_1 -distance R of γ by the Morse Lemma applied in $\widetilde{K}_1$. By Arzelà–Ascoli, some sequence $H_{\varepsilon_n}^{-1}\gamma$ converges on compact sets to a curve γ_0 , which is a d_0 -geodesic by Theorem 3.6 and stays within distance R of γ , so that $\rho_0(\gamma_0(\infty)) = \xi$. Finally, ρ_0 is a continuous bijection between compact Hausdorff spaces, hence a homeomorphism, and it is $\pi_1(K)$ -equivariant by construction. $\square$

Let us recall some basic facts about Busemann cocycles which will be useful in the following.

Definition 3.35 (Busemann cocycles). Let (X, d) be a proper CAT(0) space and $\xi \in \partial_\infty X$. The *Busemann cocycle* of ξ is the function

$$B_\xi: X \times X \rightarrow \mathbb{R}, \quad B_\xi(x, y) := \lim_{t \rightarrow \infty} [d(x, \gamma(t)) - d(y, \gamma(t))],$$

where γ is any geodesic ray with $\gamma(\infty) = \xi$; the limit exists (the bracket is bounded and monotone by the triangle inequality and convexity) and does not depend on the chosen ray [11, II.8]. We will use the following standard properties:

- (i) B_ξ is a cocycle: $B_\xi(x, z) = B_\xi(x, y) + B_\xi(y, z)$;
- (ii) B_ξ is 1-Lipschitz in each variable;
- (iii) $B_{\psi\xi}(\psi x, \psi y) = B_\xi(x, y)$ for every isometry ψ of X ;
- (iv) along a ray towards ξ , we have $B_\xi(\gamma(t), \gamma(0)) = -t$;
- (v) *horofunction continuity*: if $z_n \rightarrow \xi$ in $\overline{X} = X \cup \partial_\infty X$, then $d(x, z_n) - d(y, z_n) \rightarrow B_\xi(x, y)$, uniformly for (x, y) in compact sets [6, Ch. II].

Remark 3.36. The orbit equivalences H_ε above are essentially unique, due to the Anosov Shadowing Theorem (more precisely, they are *transversally* unique, see [24, Theorem 5.4.1]). The choice of homeomorphism does not change the equivalence between orbits, just the reparametrisation function τ appearing in the relation

$$H_\varepsilon \mathbf{g}_\varepsilon^t(z) = \mathbf{g}_1^{\tau(z,t)}(H_\varepsilon(z)).$$

One possibility is to choose H_ε is the following construction, from [27, Appendix B, Theorem B.1]. Given $v \in T^1 \widetilde{K}_\varepsilon$, we write

$$(v_\varepsilon^-, v_\varepsilon^+) = (\gamma_{\varepsilon;v}(-\infty), \gamma_{\varepsilon;v}(\infty)) \in \partial_\infty^2 \widetilde{K}.$$

The homeomorphism $H_\varepsilon: T^1 \widetilde{K}_\varepsilon \rightarrow T^1 \widetilde{K}_1$ is then obtained by setting $H_\varepsilon(v) = w$, where $w \in T^1 \widetilde{K}_1$ is the unique vector such that

$$(v_\varepsilon^-, v_\varepsilon^+) = (w_1^-, w_1^+) \quad \text{and} \quad B_{v_\varepsilon^+}^1(\pi(v), \pi(w)) = 0.$$

Here $\pi: T^1 \widetilde{K} \rightarrow \widetilde{K}$ is the canonical projection and $B_{v_\varepsilon^+}^1$ is the Busemann cocycle (Definition 3.35) of the boundary point v_ε^+ with respect to the metric g_1 . The condition $B_{v_\varepsilon^+}^1(\pi(v), \pi(w)) = 0$ selects, on the g_1 -geodesic with the prescribed endpoints, the unique point lying on the horosphere of v_ε^+ through $\pi(v)$.

With such choice, the family $\{H_\varepsilon^{-1}\}$ becomes a uniformly equicontinuous deformation of the g_1 -geodesics on $\widetilde{K}$. Indeed, if that were not the case there would be $\eta > 0$ and sequences $(v_n, u_n) \in T^1 K_1, \varepsilon_n \rightarrow 0$ such that

$$d_1(v_n, u_n) < \frac{1}{n} \quad \text{and} \quad d_\varepsilon(H_\varepsilon^{-1}(v_n), H_\varepsilon^{-1}(u_n)) > \eta.$$

The common limit $\lim_n v_n = \lim_n u_n = z \in T^1 K_1$ defines a unique geodesic $\gamma \in \widehat{\mathcal{G}}_0$. Moreover, by construction, the geodesics defined by the initial conditions $H_\varepsilon^{-1}(v_n)$ and $H_\varepsilon^{-1}(u_n)$ have the same endpoints at infinity as v_n and u_n , and thus must both converge to γ on every compact subset as $\varepsilon \rightarrow 0$, a contradiction.

3.2.3. Closed geodesics in the space K_0 . We first recall the classification of isometries of CAT(0) spaces and the relevant properties of the action of $\pi_1(K)$ on $\widetilde{K}_0$; we refer to [11, Chapter II.6] for more details.

Definition 3.37 (Isometries of CAT(0) spaces). Let (X, d) be a complete CAT(0) space and g an isometry of X . The *displacement function* of g is $d_g: X \rightarrow [0, \infty)$, $d_g(x) := d(x, gx)$, and its *translation length* is

$$\ell(g) := \inf_{x \in X} d(x, gx).$$

Write $\text{Min}(g) := \{x \in X : d(x, gx) = \ell(g)\}$. The isometry g is called

- (i) *elliptic* if it has a fixed point (equivalently, $\ell(g) = 0$ and $\text{Min}(g) \neq \emptyset$);
- (ii) *hyperbolic* (or *axial*) if $\text{Min}(g) \neq \emptyset$ and $\ell(g) > 0$;
- (iii) *parabolic* if $\text{Min}(g) = \emptyset$, i.e., the infimum is not attained.

Isometries which are elliptic or hyperbolic are called *semi-simple*.

We will use the following standard facts about a hyperbolic isometry g of a proper CAT(0) space X [11, II.6.8]:

- (a) a geodesic line $A \subset X$ is an *axis* of g if g acts on A as a translation by $\ell(g)$; the set $\text{Min}(g)$ is the union of the axes of g , and any two axes are parallel, i.e., at bounded Hausdorff distance (hence they bound a flat strip, by the Flat Strip Theorem);
- (b) all the axes of g have the same endpoints at infinity, denoted $g^- := A(-\infty)$ and $g^+ := A(+\infty)$ and called the *repelling* and *attracting* fixed points of g ; these are fixed by the action of g on $\partial_\infty X$, and $g^n x \rightarrow g^\pm$ as $n \rightarrow \pm\infty$ for every $x \in X$;
- (c) $\ell(g^n) = |n| \ell(g)$ and $\text{Min}(g^n) = \text{Min}(g)$ for every $n \neq 0$, and $\ell(hgh^{-1}) = \ell(g)$ for every isometry h , with $(hgh^{-1})^\pm = h g^\pm$; in particular ℓ is a class function on any group of isometries;
- (d) if X is moreover Gromov-hyperbolic, then $g^- \neq g^+$ are the only fixed points of g on $\partial_\infty X$, and g has *North-South dynamics*: for every neighbourhood $U^\pm$ of $g^\pm$ in $\bar{X}$, one has $g^n(\bar{X} \setminus U^\mp) \subset U^\pm$ for all n large enough (i.e., $g^{\pm n} \rightarrow g^\pm$ uniformly on compact subsets of $\bar{X} \setminus \{g^\mp\}$); in Gromov's terminology, g is *loxodromic*.

Here, the word “hyperbolic” refers here to the classification of isometries, and not to any property of the dynamics; in the Gromov-hyperbolic CAT(0) spaces considered below, hyperbolic isometries in the CAT(0) sense and loxodromic isometries in Gromov's sense coincide.

Lemma 3.38. *The group $\pi_1(K)$ acts on the Hadamard space $\widetilde{K}_0$ by deck transformations, which are isometries; the action is free, properly discontinuous and cocompact. Moreover:*

- (i) $\pi_1(K)$ is the fundamental group of a closed orientable surface of genus $m+1 \geq 2$, where m is the number of scatterers; in particular it is torsion-free and non-elementary, and $\partial_\infty \widetilde{K}_0 \approx \mathbb{S}^1$ is its Gromov boundary;
- (ii) every $g \in \pi_1(K) \setminus \{1\}$ is a hyperbolic isometry of $\widetilde{K}_0$;
- (iii) every $g \in \pi_1(K) \setminus \{1\}$ has a unique axis A_g , namely the geodesic line (g^-, g^+) of Lemma 3.33; the projection of A_g to K_0 is a closed geodesic of length $\ell(g)$ in the free homotopy class of g , and $\ell(g) \geq 2R_0 > 0$ with R_0 as in Lemma 3.3;
- (iv) two elements $g, h \in \pi_1(K) \setminus \{1\}$ either have the same pair of fixed points $\{g^\pm\} = \{h^\pm\}$, in which case they have a common axis and generate a cyclic group up to finite index, or they have four distinct fixed points.

Proof. The covering $\widetilde{K}_0 \rightarrow K_0$ is a local isometry for the lifted length metric, so deck transformations are isometries; the action is free and properly discontinuous (as for any covering) and cocompact, since K_0 is compact.

(i) The surface K is the double of the table $\overline{D}$, which is a torus with m discs removed, along its m boundary circles; hence K is a closed orientable surface with

$\chi(K) = 2\chi(\overline{D}) = -2m$, i.e., of genus $m+1 \geq 2$. Its fundamental group is torsion-free and non-elementary (it contains free subgroups of rank two), and, being quasi-isometric to $\widetilde{K}_0$ by the Švarc–Milnor lemma, its Gromov boundary is $\partial_\infty \widetilde{K}_0$, which is a circle by [11, II.8.11] (see also the beginning of this subsection).

(ii) Since the action is cocompact, every element of $\pi_1(K)$ is a semi-simple isometry [11, II.6.10(2)]; since the action is free, no non-trivial element is elliptic. Hence every $g \neq 1$ is hyperbolic.

(iii) Two distinct axes of g would be parallel geodesic lines, bounding a flat strip, which is excluded by Lemma 3.31; so the axis is unique, and it is the unique geodesic line joining g^- to g^+ (Lemma 3.33). Its projection to K_0 is a closed curve of length $\ell(g)$ in the free homotopy class of g , which is a local geodesic since A_g is a geodesic line; its length is at least the length of the shortest non-contractible loop of K for d_0 , which is at least $2R_0$ by the proof of Lemma 3.3 (the bound s_0 obtained there is for the Euclidean length of projections, which is dominated by the d_0 -length).

(iv) If g and h share one fixed point $\xi \in \partial_\infty \widetilde{K}_0$, then the stabiliser of ξ in $\pi_1(K)$ contains the two loxodromic elements g, h ; in a group acting properly on a proper Gromov-hyperbolic space, the stabiliser of a point fixed by a loxodromic element is virtually cyclic and fixes both its fixed points [17, Chapter 8], so $\{g^\pm\} = \{h^\pm\}$, the axes coincide by (iii), and $\langle g, h \rangle$ is virtually cyclic. $\square$

With this terminology, the length spectrum of K_0 is the function ℓ on the set $\mathcal{C}(K) := \text{Conj}(\pi_1(K))$ of conjugacy classes, and the next lemma identifies it with the lengths of the closed d_0 -geodesics of K_0 .

Lemma 3.39. *Any element $\alpha \in \mathcal{C}(K) := \text{Conj}(\pi_1(K))$ contains a single closed geodesic curve γ_0^α , which is the curve of minimal d_0 -length in α ; write $\ell_0(\alpha) := \ell_0(\gamma_0^\alpha)$. Moreover, $\gamma_\varepsilon^\alpha \rightarrow \gamma_0^\alpha$ in the C^0 topology and $\ell_\varepsilon(\gamma_\varepsilon^\alpha) \rightarrow \ell_0(\alpha)$ as $\varepsilon \rightarrow 0$.*

Proof. *Step 1: existence, minimality and uniqueness of γ_0^α .* Fix a representative $g \in \pi_1(K)$ of α . By Lemma 3.38, g is a hyperbolic isometry of $\widetilde{K}_0$ with a unique axis A , on which it translates by $\ell(g) = \inf_{x \in \widetilde{K}_0} d_0(x, gx) > 0$. Set $\gamma_0^\alpha := p(A)$, a closed d_0 -geodesic in α , and $\ell_0(\alpha) := \ell(g)$. Any closed curve $c \in \alpha$ lifts to a path from some x to gx , so $\ell_0(c) \geq d_0(x, gx) \geq \ell(g) = \ell_0(\gamma_0^\alpha)$; thus γ_0^α is length-minimising in α . If α contained a second closed d_0 -geodesic, its lift would be a second axis of g , contradicting the uniqueness of the axis (Lemma 3.38(iii)); thus γ_0^α is the unique closed geodesic in α .

Step 2: convergence $\gamma_\varepsilon^\alpha \rightarrow \gamma_0^\alpha$. Each K_ε is Anosov, so α contains a unique closed g_ε -geodesic $\gamma_\varepsilon^\alpha$ [35], of length $\ell_\varepsilon(\alpha) := \inf_x d_\varepsilon(x, g_\varepsilon x)$. By Theorem 3.16, $\ell_0(\alpha) \leq \ell_\varepsilon(\alpha)$, and testing against γ_0^α ,

$$\ell_\varepsilon(\alpha) \leq \ell_\varepsilon(\gamma_0^\alpha) \leq \ell_0(\gamma_0^\alpha) + C \text{Var}(z \circ \gamma_0^\alpha) \varepsilon = \ell_0(\alpha) + O(\varepsilon),$$

so $\ell_\varepsilon(\alpha) \rightarrow \ell_0(\alpha)$ as $\varepsilon \rightarrow 0$. (*)

Parametrise each $\gamma_\varepsilon^\alpha$ on $\mathbb{S}^1$ proportionally to d_ε -arclength; since $d_0 \leq d_\varepsilon$, the resulting maps are $\ell_1(\alpha)$ -Lipschitz into (K, d_0) , so $\{\gamma_\varepsilon^\alpha\}$ is precompact in $C^0(\mathbb{S}^1, K)$ by Arzelà–Ascoli. Let γ_* be an accumulation point, $\gamma_{\varepsilon_n}^\alpha \rightarrow \gamma_*$ uniformly.

As (K, d_0) and (K, d_1) are compact with the same topology, the identity is a uniform homeomorphism, so uniform d_0 -convergence is uniform d_1 -convergence; hence for large n , $(\gamma_{\varepsilon_n}^\alpha(s), \gamma_*(s)) \in D_{R_0} := \{(x, y) : d_1(x, y) < R_0\}$ for all s , where R_0 is the uniform injectivity radius (Lemma 3.3). Since D_{R_0} deformation-retracts to the diagonal via $(x, y) \mapsto (x, \exp_x^1 y)$, the loops $\gamma_{\varepsilon_n}^\alpha$ and γ_* are freely homotopic, i.e., $\gamma_* \in \alpha$.

Let us now show that $l_0(\gamma_*) = \ell_0(\alpha)$. By lower semicontinuity of length under uniform convergence [13] and $l_0 \leq l_{\varepsilon_n}$,

$$l_0(\gamma_*) \leq \liminf_n l_0(\gamma_{\varepsilon_n}^\alpha) \leq \liminf_n l_{\varepsilon_n}(\gamma_{\varepsilon_n}^\alpha) = \liminf_n l_{\varepsilon_n}(\alpha) \stackrel{(*)}{=} \ell_0(\alpha),$$

while $l_0(\gamma_*) \geq \ell_0(\alpha)$ by Step 1. Hence γ_* is a minimal closed d_0 -geodesic in α , so $\gamma_* = \gamma_0^\alpha$. Every accumulation point equals γ_0^α and the family is precompact, so $\gamma_\varepsilon^\alpha \rightarrow \gamma_0^\alpha$ in C^0 , and $l_\varepsilon(\gamma_\varepsilon^\alpha) = \ell_\varepsilon(\alpha) \rightarrow \ell_0(\alpha)$ by $(*)$. $\square$

Corollary 3.40. *The pairs of endpoints (g^-, g^+) of the axes of the elements $g \neq 1$ of $\pi_1(K)$ (i.e., of the lifts of closed geodesics of K_0) are dense in $\partial_\infty^2 \widetilde{K}_0$. In particular, lifts of closed geodesics are dense in $\mathcal{G}_0$, and the fixed points of the elements of $\pi_1(K)$ are dense in $\partial_\infty \widetilde{K}_0$.*

Proof. Since the geodesic flow $\mathbf{g}_1$ is Anosov, hence transitive with dense periodic orbits, the pairs of endpoints of the axes of the elements of $\pi_1(K)$ are dense in $\partial_\infty^2 \widetilde{K}_1$. The equivariant homeomorphism $\rho_0: \partial_\infty \widetilde{K}_0 \rightarrow \partial_\infty \widetilde{K}_1$ of Lemma 3.34 maps the fixed points of each g to the fixed points of the same g , so $(\rho_0 \times \rho_0)^{-1}$ carries this dense set onto the set of pole pairs in $\partial_\infty^2 \widetilde{K}_0$, which is therefore dense; the remaining assertions follow from Lemma 3.33. $\square$

3.2.4. Cross-ratios on the boundary $\partial_\infty \widetilde{K}_0$. Fix the homeomorphisms $H_\varepsilon: \mathcal{G}_0 \rightarrow \mathcal{G}_\varepsilon$ induced by the boundary identifications ρ_ε of Lemma 3.34, and identify all the boundaries $\partial_\infty \widetilde{K}_\varepsilon$ with $\partial_\infty \widetilde{K}_0 \approx \mathbb{S}^1$. Given $\gamma_0 \in \mathcal{G}_0$, we write $\gamma_\varepsilon := H_\varepsilon \gamma_0$, and assume, without loss of generality, that $\gamma_\varepsilon(\infty) = \gamma_0(\infty) \in \mathbb{S}^1$ for all $\varepsilon \in [0, 1]$. In particular, γ_0 is a quasi-geodesic on every $\widetilde{K}_\varepsilon$ and thus defines a unique point at infinity in $\partial_\infty \widetilde{K}_\varepsilon$, which, by our previous identifications, is just $\gamma_0(\infty)$.

For $\varepsilon \in [0, 1]$ and $x, y, z, w \in \widetilde{K}_\varepsilon$, write

$$E_\varepsilon(x, y, z, w) := d_\varepsilon(x, y) + d_\varepsilon(z, w) - d_\varepsilon(x, w) - d_\varepsilon(z, y).$$

As in [40, 10], given distinct points $\{\zeta_i\}_{i=1}^4 \in \partial_\infty \widetilde{K}_0^4$, we define their cross-ratio as

$$\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) = \lim_{t \rightarrow \infty} e^{\frac{1}{2} E_0(\gamma_1(t), \gamma_2(t), \gamma_3(t), \gamma_4(t))} \quad (9)$$

where the γ_i are unit-speed geodesic rays, issued from arbitrary base points, such that $\gamma_i(\infty) = \zeta_i$, for $i = 1, 2, 3, 4$. Lemma 3.41 shows that the limit exists and depends neither on the base points nor on any other feature of the chosen rays.

Lemma 3.41. *For distinct $\zeta_1, \dots, \zeta_4 \in \partial_\infty \widetilde{K}_0$ and unit-speed rays γ_i with $\gamma_i(\infty) = \zeta_i$, the limit defining cr_0 in (9) exists and is independent of the chosen rays.*

Proof. Let $f_{ij}(t) := d_0(\gamma_i(t), \gamma_j(t)) - 2t$. By convexity of $t \mapsto d_0(\gamma_i(t), \gamma_j(t))$ in the CAT(0) space $\widetilde{K}_0$ [11, II.2.2] and its 2-Lipschitz property, f_{ij} is non-increasing; by δ -hyperbolicity (Corollary 3.32) and $\zeta_i \neq \zeta_j$, it is bounded below. Hence each f_{ij} converges, and so does $E_0(\gamma_1(t), \gamma_2(t), \gamma_3(t), \gamma_4(t)) = f_{12} + f_{34} - f_{14} - f_{32}$. Independence of the rays follows from a CAT(0) argument: for two forward-asymptotic rays γ, γ' in $\widetilde{K}_0$, the function $t \mapsto d_0(\gamma(t), \gamma'(t))$ is convex and bounded, hence non-increasing, and converges to some $c \geq 0$; if $c > 0$, the rays would bound a flat half-strip and, translating it by a suitable sequence of deck transformations and passing to a limit, we would obtain a complete flat strip, contradicting Lemma 3.31 (see [6, Chapter II] for more details). Hence $c = 0$, and the limits defining cr_0 do not depend on the chosen rays. $\square$

Lemma 3.42. *Let $\xi \neq \eta \in \partial_\infty \widetilde{K}_0$ and let $u_k, v_k \in \widetilde{K}_0$ with $u_k \rightarrow \xi$, $v_k \rightarrow \eta$. Then the geodesics $[u_k, v_k]$ (segments, rays or lines) converge, after reparametrisation, uniformly on compact sets to the line (ξ, η) . Moreover:*

- (i) $d(p, [u_k, v_k]) \rightarrow 0$ for every $p \in (\xi, \eta)$;
- (ii) if $x_k \in [u_k, v_k]$ and $d(o, x_k) \rightarrow \infty$, then the accumulation points of (x_k) in $\overline{X}$ are contained in $\{\xi, \eta\}$.

Proof. By δ -hyperbolicity every point of $[u_k, v_k]$ lies within 4δ of $[u_k, o] \cup [o, v_k]$; since u_k, v_k converge to distinct boundary points, the Gromov products $(u_k | v_k)_o$ are bounded (Definition 3.21(iv)), hence so is $d(o, [u_k, v_k]) \leq (u_k | v_k)_o + 4\delta$ (Definition 3.21(iii)): there is a ball $B(o, C)$ met by all $[u_k, v_k]$. Reparametrise so that $[u_k, v_k](0) \in B(o, C)$; the maps are 1-Lipschitz, so Arzelà–Ascoli and properness give subconvergence, uniform on compact sets, to a complete geodesic (the domains exhaust $\mathbb{R}$ since u_k, v_k leave every ball). Its endpoints are ξ, η because far-away points of $[u_k, v_k]$ lie 4δ -close to the rays $[o, u_k]$, $[o, v_k]$, whose ideal points converge; this concludes the proof of point (ii). By uniqueness of the geodesic joining ξ to η (by Lemma 3.33) every sublimit is (ξ, η) , hence we have full convergence and (i). $\square$

Corollary 3.43. *The function cr_0 is continuous on the set of quadruples of pairwise distinct points of $\partial_\infty \widetilde{K}_0$.*

Proof. Let (ξ, η) be one of the pairs (ζ_i, ζ_j) appearing in E_0 , with rays u_t, v_s issued from base points c, c' towards ξ, η , and let p be any point of the line (ξ, η) (Lemma 3.33). We claim that

$$\lim_{t,s \rightarrow \infty} [t + s - d_0(u_t, v_s)] = -[B_\xi(p, c) + B_\eta(p, c')],$$

with the Busemann cocycles of Definition 3.35. Indeed $B_\xi(p, c) = \lim_t [d_0(p, u_t) - t]$ and $B_\eta(p, c') = \lim_s [d_0(p, v_s) - s]$, so the difference between the two sides is $\lim [d_0(p, u_t) + d_0(p, v_s) - d_0(u_t, v_s)]$, which is non-negative by the triangle inequality and at most $2d_0(p, [u_t, v_s])$; the latter tends to 0 because the segments $[u_t, v_s]$ converge to the line (ξ, η) (by Lemma 3.42(i)). The right-hand side depends continuously on $(\xi, \eta) \in \partial_\infty^2 \widetilde{K}_0$: the line (ξ, η) and a point p on it can be chosen continuously by Lemma 3.42, and Busemann cocycles are jointly continuous in (ξ, x, y) on a proper CAT(0) space; indeed, $\xi \mapsto B_\xi(\cdot, o)$ is the horofunction map, which is a homeomorphism onto its image for the cone topology [11, II.8.13]. As $\lim_t E_0(\gamma_1(t), \dots, \gamma_4(t))$ is a signed sum of four such limits (with $s = t$), cr_0 is continuous. $\square$

Remark 3.44. Since $d_0 = \lim_{\varepsilon \rightarrow 0} d_\varepsilon$ monotonically, one may ask whether cr_0 is consistent with the ε -cross-ratios, i.e.,

$$\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) = \lim_{\varepsilon \rightarrow 0} \text{cr}_\varepsilon(\zeta_1, \zeta_2, \zeta_3, \zeta_4), \quad (10)$$

where cr_ε is defined as in (9) using E_ε , the same geodesics $\gamma_i \in \mathcal{G}_0$ being viewed as uniform quasi-geodesics of $\widetilde{K}_\varepsilon$ with the same endpoints at infinity (Theorem 3.16). The interchange of the limits in t and ε can be justified by Moore–Osgood, using the uniform shadowing of Theorem 3.16 together with a uniform exponential convergence of asymptotic g_ε -geodesics. Since we do not need this fact in the present paper (the constructions of Section 4 are intrinsic to $(\widetilde{K}_0, d_0)$), we do not elaborate on this point.

It is easy to check the following properties from the definition:

Proposition 3.45. *The mapping $\text{cr}_0: (\partial_\infty \widetilde{K}_0)^4 \rightarrow \mathbb{R}$ satisfies*

- (i) cr_0 is positive and invariant under the action of $\pi_1(K)$ on $(\partial_\infty \widetilde{K}_0)^4$;

- (ii) $\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) = \text{cr}_0(\zeta_1, \zeta_4, \zeta_3, \zeta_2)^{-1}$;
- (iii) $\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) = \text{cr}_0(\zeta_3, \zeta_4, \zeta_1, \zeta_2)$;
- (iv) $\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) \text{cr}_0(\zeta_1, \zeta_4, \zeta_3, \zeta_5) = \text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_5)$;
- (v) $\text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) \text{cr}_0(\zeta_2, \zeta_3, \zeta_1, \zeta_4) \text{cr}_0(\zeta_3, \zeta_1, \zeta_2, \zeta_4) = 1$.

3.2.5. Geodesic facts on the limit space. Below we abbreviate $X := \widetilde{K}_0$ and $B := B(\widetilde{K})$ for the seam (the union of the lifted walls); $\overline{X} = X \cup \partial_\infty X$ carries the cone topology. Let us recall that the minimal free path $\tau_{\min} > 0$ of the table, and the minimal distance between distinct lifted walls, are positive.

We first recall the notion of angle. We recall that, in a CAT(0) space (X, d) , the angle $\angle_x(\gamma, \gamma') \in [0, \pi]$ between two geodesics $\gamma, \gamma': [0, 1] \rightarrow X$ intersecting at the point $x = \gamma(0) = \gamma'(0)$ is given by

$$\angle_x(\gamma, \gamma') = \lim_{t \rightarrow 0} 2 \arcsin \left(\frac{1}{2t} d(\gamma(t), \gamma'(t)) \right).$$

In particular, for any geodesic triangle $\triangle(x, y, z)$ the following standard comparison estimate holds:

$$\angle_x(y, z) + \angle_y(z, x) + \angle_z(x, y) \leq \pi. \quad (11)$$

When (X, d) is a Riemannian manifold with the Riemannian metric, this coincides with the Riemannian angle (see e.g. [11, Corollary II.1A.7]). In our situation, at a flat point $x \in X \setminus B$ a neighbourhood of x is isometric to a Euclidean disc (see e.g. Lemma 3.8), so the angle between two geodesics through x is the Euclidean angle of their chords; at a seam point the CAT(0) angle is used directly, and it is described in Lemma 3.48 below.

Remark 3.46. Note that angles are always well defined on $\widetilde{K}_0$, even though they do not specify the geodesic, at least not at degenerate points. In fact, two geodesics bifurcate at a point x if and only if their angle at x is zero, and we may characterise the boundary set $B(\widetilde{K})$ exactly as the locus of points where geodesic bifurcation happens, that is, points from which emanate a pair of geodesics with angle zero.

The following elementary lemmas will be used repeatedly in the following.

Lemma 3.47. *For every $x \in X$ and $\zeta \in \partial_\infty X$ there is a unique ray $r_{x, \zeta}$ from x to ζ , and $(x, \zeta) \mapsto r_{x, \zeta}$ is continuous (uniformly on compact time intervals).*

Proof. If r, r' are rays from x to ζ , the convex function $t \mapsto d(r(t), r'(t))$ vanishes at 0 and is bounded (by asymptoticity); a bounded convex function on $[0, \infty)$ vanishing at the origin is identically zero (since $d(t)/t$ is non-decreasing and tends to 0). Let us now show continuity. If $x_k \rightarrow x$ and $\zeta_k \rightarrow \zeta$, the rays $r_k := r_{x_k, \zeta_k}$ satisfy $d(r_k(t), r_{o, \zeta_k}(t)) \leq d(x_k, o)$ for all t (a convex function bounded by its value at 0, by the same argument), so the r_k stay in a tube around rays from o ; Arzelà–Ascoli gives sublimits, which are rays from x asymptotic to ζ (their ideal points are $\lim \zeta_k = \zeta$, by Lemma 3.42(ii) applied to the rays r_{o, ζ_k}); by uniqueness, all sublimits are equal. $\square$

Lemma 3.48. *At every $x \in X$ the space of directions Σ_x , with the Alexandrov angle, is a circle of length 2π , and:*

- (i) *the two directions of a geodesic through x cut Σ_x into two arcs of length π ;*
- (ii) *for directions u, v, w in this cyclic order on an arc of length $\leq \pi$: $\angle(u, w) = \angle(u, v) + \angle(v, w)$;*
- (iii) *(reflection law) a concatenation of two segments crossing B transversally at x is geodesic iff the two incidence angles with the seam, measured in the respective sheets, are equal;*

- (iv) (branching) *two distinct geodesics containing a common ray or segment ending at x split at x with angle 0; this occurs only if the common direction at x is tangent to B . In particular a concatenation of a segment and a seam arc is geodesic iff the segment arrives tangentially.*

Proof. For x in the flat interior everything is Euclidean. Let us consider $x \in B$; the tangent cone of X at x is the pointed Gromov–Hausdorff limit of the rescalings $(X, x, \lambda d)$, $\lambda \rightarrow \infty$. A neighbourhood of x is the gluing of two flat half-discs along the wall through x , which is a C^2 curve; under rescaling each half-disc converges to a Euclidean half-plane and the wall to its tangent line, so the tangent cone is the gluing of two half-planes along their boundary lines, i.e., the Euclidean plane $\mathbb{E}^2$. For a CAT(0) space, the tangent cone in the sense of [11, II.3.19] is the Euclidean cone over Σ_x with the angle metric; here it coincides with the Gromov–Hausdorff limit above, since the Alexandrov angle between two germs at x is the limit of their comparison angles, which are computed in the rescaled spaces and converge to the Euclidean angle of the limit configuration (the two half-discs being C^2 -close to half-planes). Hence Σ_x is the round circle of length 2π ; the two sheets contribute the two semicircles, glued at the two seam directions $\pm\tau$. Angles are then arc lengths, which gives (ii), and a concatenation at x is geodesic if and only if the two directions make angle π [11, II.1.4 and II.9], i.e., iff both complementary arcs measure exactly π ; (i) follows.

For (iii): with incidence angles ϕ^- (incoming, in sheet 1) and ϕ^+ (outgoing, in sheet 2), the two arcs measure $\pi - \phi^- + \phi^+$ and $\pi + \phi^- - \phi^+$, both $\geq \pi$ iff $\phi^- = \phi^+$.

For (iv): two geodesics splitting at x with a common incoming direction u must both continue with the direction antipodal to u in Σ_x ; if u is not tangent to B , the antipode is the direction of the unique straight continuation of the incoming chord inside its sheet (or the transversally reflected chord, by (iii)), and the germ of the continuation is unique. If $u = \pm\tau$, the antipode is realised by three germs: the seam and one chord in each sheet, all tangent to B . It corresponds to the branching configuration, which gives the last statement (an arc and then a segment junction with incidence ϕ^- has arcs $\pi \mp \phi^-$; it is geodesic iff $\phi^- = 0$). $\square$

Let us conclude this section with a remark on the contacts with the seam:

Lemma 3.49. *For every geodesic γ of X , consecutive connected components of the contact set $C_\gamma := \{t : \gamma(t) \in B\}$ are separated by open intervals of length $\geq \tau_{\min}$. Consequently, on every compact window, C_γ is a finite union of closed intervals (seam arcs) and isolated points (transversal crossings or tangential contacts), and $|C_\gamma \setminus \text{int } C_\gamma|$ has zero Lebesgue measure.*

Proof. A maximal open interval (t, t') of non-contact times is a straight chord of one flat sheet, i.e. (developing a neighbourhood of the chord isometrically into the plane), a Euclidean segment in a lift of the table, from a wall point to a wall point, meeting no wall in between. A segment leaving a point of a strictly convex wall into the table never meets the same obstacle lift again; so the chord joins two distinct obstacle lifts (possibly two lifts of the same wall of K) and its length is at least the minimal free path $\tau_{\min} > 0$ of the table. The remaining statements follow: on a compact window there are finitely many components; non-interval components are single times, at which γ either crosses B transversally or touches it tangentially (see Lemma 3.48); and the complement of the interior of C_γ in C_γ consists of finitely many points per window. $\square$

4. SPECTRAL RIGIDITY

In what follows, we use the notation of Section 3, i.e., $K_0 = (K, g_0)$ is the Kourganoff surface obtained by gluing two copies of $(\mathcal{D}, d_{\mathcal{L}})$ at the boundary, with $K^\circ = K \setminus B(K) = \mathcal{D}^{\text{up}} \amalg \mathcal{D}^{\text{down}}$. As previously, we consider a family of metrics g_ε on K converging to the gluing metric when ε approaches 0. In other words, we have metric convergence $(K, g_\varepsilon) =: K_\varepsilon \xrightarrow{GH} K_0$ in the Gromov–Hausdorff distance, as $\varepsilon \rightarrow 0$. Below we abbreviate $X := \widetilde{K}_0$, with $\overline{X} = X \cup \partial_\infty X$, and we denote by $B := B(\widetilde{K})$ the seam (the union of the lifted walls).

The strategy of this section is that of Otal [39], carried out directly on the limit space $\widetilde{K}_0$; here, the enriched spectrum is the marked length spectrum of the CAT(0) space K_0 . We show that the space carries an intrinsic Liouville geodesic current, defined by the kinematic measure of its flat pieces (§4.1). The spectrum determines the cross-ratio at infinity (§4.2), hence the current, and therefore a homeomorphism h between the spaces of unparametrised geodesics which preserves the currents. An average-angle argument in the spirit of Otal [39] then shows that h preserves angles (§4.3), and the dispersing walls force h to be induced by an isometry (§4.4). The proofs do not use its ergodicity; in particular the rigidity theorem does not rely on Sinai’s theorem about the ergodicity of finite horizon Sinai billiards.

Recall that $\mathcal{C}(K) := \text{Conj}(\pi_1(K))$ is the set of free homotopy classes on K .

Definition 4.1 (Enriched marked length spectrum). Let $\mathcal{D}$ be a Sinai billiard and K a surface in $\mathcal{K}(\mathcal{D})$. The function

$$\begin{aligned} \mathbf{EL}_{\mathcal{D};K} : \mathcal{C}(K) &\rightarrow \mathbb{R} \\ \alpha &\mapsto l_0(\gamma^\alpha) \end{aligned}$$

is called the enriched marked length spectrum of $\mathcal{D}$.

Remark 4.2. By Lemma 3.39, this definition agrees with the one given in the Introduction; moreover, it does not depend on the choice of $K \in \mathcal{K}(\mathcal{D})$, since by Theorem B.2 and the results of Appendix B the values $l_0(\gamma^\alpha)$ are computed by the intrinsic length functionals $\mathbf{EL}_\rho^I$ of the table. We therefore drop K from the notation and write $\mathbf{EL}_{\mathcal{D}}$. We have $\mathbf{EL}_{\mathcal{D}}(\alpha) = \ell_0(\alpha) = \inf_x d_0(x, gx)$ for any representative $g \in \pi_1(K)$ of α : as we noted earlier, the enriched marked length spectrum of $\mathcal{D}$ is thus the marked length spectrum of the CAT(0) space K_0 .

4.1. The intrinsic Liouville current of the limit space. In this part, we define an intrinsic Liouville current on the limit space. We then prove an exact Crofton formula (Lemma 4.9) and show that the values of the current on rectangles of boundary pairs are precisely $|\ln \text{cr}_0|$ (Proposition 4.15), so that the cross-ratios above determine the current completely.

Let us start by quickly recalling some facts about geodesic currents. For $\varepsilon > 0$, the surface $K_\varepsilon \in \mathcal{K}(\mathcal{D})$ carries a smooth Riemannian metric g_ε . This structure induces a canonical contact form λ_ε on the unit tangent bundle $T^1 K_\varepsilon$, obtained by pulling back the canonical symplectic form from the cotangent bundle. The Reeb vector field of λ_ε coincides with the geodesic spray of g_ε , that is, the vector field on $T^1 K_\varepsilon$ generating the geodesic flow. The associated Liouville measure $\int \lambda_\varepsilon \wedge d\lambda_\varepsilon$ on $T^1 K_\varepsilon$ is a finite measure, invariant under the geodesic flow, and consequently under its time reversal.

All such structures are lifted to $T^1 \widetilde{K}_\varepsilon$ via the local isomorphism defined by the covering map $\widetilde{K}_\varepsilon \rightarrow K_\varepsilon$, so that the space $T^1 \widetilde{K}_\varepsilon$ has the corresponding contact and

volume forms $\tilde{\lambda}_\varepsilon$ and $\int \tilde{\lambda}_\varepsilon \wedge d\tilde{\lambda}_\varepsilon$. Given a rectifiable curve $\gamma \subset \widetilde{K}_\varepsilon$, its unit tangent lift $\dot{\gamma} \subset T^1\widetilde{K}_\varepsilon$ defines the *Liouville current* m_ε on the space of geodesics $\mathcal{G}_\varepsilon$ by integration against $d\tilde{\lambda}_\varepsilon$, assigning mass to Borel sets of geodesics intersecting γ transversely [46].

More generally, a geodesic current for K_ε is any Radon measure on the space $\partial_\infty^2 \widetilde{K}_\varepsilon$ that is invariant under both the diagonal action of $\pi_1(K_\varepsilon)$ and the flip transformation $(\zeta_1, \zeta_2) \mapsto (\zeta_2, \zeta_1)$. There is an affine one-to-one correspondence between the cones of finite measures on $T^1 K_\varepsilon$ invariant under the geodesic flow and the space of geodesic currents with the weak-* topology (see [8], [25]).

In particular, the contact structure λ_ε , the cross-ratio cr_ε , and the Liouville current are related in the following way. For measurable rectangles in the space of geodesics determined by intervals $[\zeta_1, \zeta_2]$ and $[\zeta_3, \zeta_4]$ in the boundary $\partial \widetilde{K}_\varepsilon$, we have:

$$m_\varepsilon([\zeta_1, \zeta_2] \times [\zeta_3, \zeta_4]) = |\ln \text{cr}_\varepsilon(\zeta_1, \zeta_2, \zeta_3, \zeta_4)| = \int_\Sigma d\lambda_\varepsilon,$$

where $\Sigma \subset T^1 \widetilde{K}_\varepsilon$ is a transversal to the geodesic flow intersecting exactly the geodesics joining $[\zeta_1, \zeta_2]$ to $[\zeta_3, \zeta_4]$, oriented accordingly.

Such a construction cannot be replicated directly on $\widetilde{K}_0$, since the Riemannian tensors degenerate as $\varepsilon \rightarrow 0$ and the metric on $\widetilde{K}_0$ is not induced by any Riemannian metric. Below, we will nevertheless construct a Liouville current on $\mathcal{G}_0$ intrinsically. For that, we use an atlas consisting of two kinds of charts.

For a geodesic σ meeting B only transversally, the *flat chart* $\mathcal{U}_\sigma^{\text{reg}}$ is the set of geodesics crossing σ transversally at a flat point from the right, i.e., with oriented crossing angle $\theta \in (0, \pi)$, an orientation of X being fixed. Its coordinates are $(t, \theta) \in \mathbb{R} \times (0, \pi)$, the crossing time on σ and the oriented crossing angle. The geodesics crossing σ from the left form the flat chart of the reversed geodesic $\bar{\sigma}$, so that every transversal crosser of σ lies in exactly one of $\mathcal{U}_\sigma^{\text{reg}}, \mathcal{U}_{\bar{\sigma}}^{\text{reg}}$.

For a lifted wall w , parametrised by arc length r , the two *wall charts* $\mathcal{W}_w^\pm$ are the sets of geodesics crossing w transversally from a fixed side, $\pm$ referring to the two sheets adjacent to w , with coordinates $(r, \phi) \in \mathbb{R} \times (0, \pi)$, where ϕ is the incidence angle measured in the incoming sheet; by the reflection law (Lemma 3.48(iii)) the outgoing sheet gives the same coordinates, so the charts are well defined. A geodesic crossing w from one side is the reverse of a geodesic crossing it from the other side, so $\mathcal{W}_w^-$ is the image of $\mathcal{W}_w^+$ under the flip, which reads $(r, \phi) \mapsto (r, \pi - \phi)$ in these coordinates; as for flat charts, each chart is one-sided. We write $\mathcal{W}_w$ for either of them when the side is irrelevant. We can now give the intrinsic definition of the Liouville current.

Definition 4.3. The *Liouville current* $\widehat{m}_0$ is the unique Borel measure on $\widehat{\mathcal{G}}_0/\mathbf{g}_0 \approx \partial_\infty^2 \widetilde{K}_0$ whose restriction to every chart is given by

$$d\widehat{m}_0 = \frac{1}{4} \sin \theta \, d\theta \, dt \quad \text{on } \mathcal{U}_\sigma^{\text{reg}}, \quad d\widehat{m}_0 = \frac{1}{4} \sin \phi \, d\phi \, dr \quad \text{on } \mathcal{W}_w. \quad (12)$$

The next results show that this is well defined.

Lemma 4.4. *Let c_1, c_2 be embedded C^1 curves in the Euclidean plane, parametrised by arc length. On the set of lines crossing both transversally (crossing parameters separated locally), the correspondence $(t_1, \theta_1) \mapsto (t_2, \theta_2)$ between crossing data is a local C^1 -diffeomorphism and satisfies $\sin \theta_2 \, dt_2 \, d\theta_2 = \sin \theta_1 \, dt_1 \, d\theta_1$.*

Proof. Parametrise non-oriented lines by $(p, \psi) \in \mathbb{R} \times [0, \pi)$, $L_{p, \psi} = \{x : \langle x, n_\psi \rangle = p\}$, $n_\psi = (-\sin \psi, \cos \psi)$. If $L_{p, \psi}$ crosses c_i at $c_i(t_i)$ with angle θ_i , then $\psi = \alpha_i(t_i) + \theta_i \bmod \pi$ (α_i the direction of $\dot{c}_i$) and $p = \langle c_i(t_i), n_\psi \rangle$, so $(t_i, \theta_i) \mapsto (p, \psi)$ is C^1 with,

at fixed ψ , $\partial p / \partial t_i = \langle \dot{c}_i, n_\psi \rangle = \sin \theta_i \neq 0$. So $dp d\psi = \sin \theta_i dt_i d\theta_i$, and chart-to-line maps are local C^1 -diffeomorphisms [43]; the claim follows. $\square$

Lemma 4.5. *Fix a flat or wall chart. Each of the following sets of chart parameters is contained in a countable union of C^1 curves, hence is null for the planar Lebesgue measure of the chart:*

- (a) *parameters whose geodesic has a tangential contact with B ;*
- (b) *parameters over which the chart fibre contains more than one geodesic;*
- (c) *parameters whose geodesic passes through a fixed point $p \in X$.*

Proof. (a) By Lemma 3.49, along any bounded time window a geodesic has finitely many transversal crossings before a possible tangential contact; so the parameters in (a) are the countable union, over finite itineraries $I = (w_1, \dots, w_k)$ of lifted walls crossed transversally, over target walls w , over bounded windows, and over the two time directions, of the sets

$$T_{I,w} := \{\text{parameters whose continuation crosses } w_1, \dots, w_k \text{ transversally,} \\ \text{then meets } w \text{ tangentially}\}.$$

Fix I and let O_I be the open set of parameters whose continuation realises the itinerary I transversally and then enters the flat sheet adjacent to w along a chord. On O_I , let $\Lambda: (\text{chart data}) \rightarrow (p, \psi)$ assign the line data of that final chord in a fixed local development of its sheet. Then Λ is a composition of the chart-to-line map of the initial piece, and, at each w_i , the map (line data) $\rightarrow$ (wall data (r_i, ϕ_i)) followed by (reflected wall data) $\rightarrow$ (line data of the next developed sheet). By Lemma 4.4 applied with c_1 a straight developed transversal or the C^2 curve w_i , every factor is a local C^1 -diffeomorphism (the Jacobians are $\sin \theta$, $\sin \phi_i \neq 0$ at transversal data while the reflection itself is the identity in wall data). Hence Λ is a local C^1 -diffeomorphism onto open sets of line data. The lines tangent to the developed image of w form an immersed C^1 curve in (p, ψ) : parametrised by the tangency foot s , the direction $\psi(s)$ has $\psi'(s) = \pm \mathcal{K}(s) \neq 0$ by strict convexity of the walls. Its preimage under Λ is a countable union of C^1 curves; and $T_{I,w} \subset \Lambda^{-1}(\text{tangent lines to } w)$.

(b) Two geodesics over the same chart datum share the crossing point and direction, hence coincide until they branch; by Lemma 3.48(iv) branching requires a tangential contact with B , so (b) $\subset$ (a).

(c) If p is a flat point, the lines through the developed image of p form a C^1 curve in (p, ψ) ; argue as in (a). If $p = w(r_0)$ is a seam point, the geodesics through p that cross w transversally form $\{r = r_0\}$ in the wall chart of w , and we transfer to the given chart through Λ ; the remaining geodesics through p (tangent to w at p , or running along w) lie in (a). $\square$

Proposition 4.6. *The formulas (12) are compatible on all chart overlaps and define a unique Radon measure $\hat{m}_0$ on $\partial_\infty^2 X$, invariant under $\pi_1(K)$ and the flip, for which tangential-contact configurations, point pencils, and degenerate geodesics are null.*

Proof. Let two charts (flat or wall) overlap at a geodesic γ crossing their transversals at $\gamma(s)$ and $\gamma(s')$, $s < s'$, and remove the null set of geodesics described in Lemma 4.5 with a tangential contact on the connecting arc, or belonging to a multiple fibre. For the remaining γ , the arc $\gamma([s, s'])$ crosses B transversally, finitely many times (Lemma 3.49), at walls $w_1, \dots, w_k$, and contains no seam arc. On a neighbourhood of γ in the first chart, the transition map factors as

$$(\text{chart}) \rightarrow (r_1, \phi_1) \rightarrow (r_1, \phi_1)^{\text{out}} \rightarrow (r_2, \phi_2) \rightarrow \dots \rightarrow (\text{chart}'),$$

where consecutive arrows are the planar transitions of Lemma 4.4 inside one developed sheet (with straight or C^2 -curved transversals), or the reflection, which is the identity in wall data. Every factor preserves the density $\frac{1}{4} \sin(\cdot) d(\cdot) d(\cdot)$, hence so does the transition, almost everywhere on the overlap, which shows the desired consistency. The charts cover the complement of the set $\mathcal{T}$ of complete wall lifts, which is countable. Indeed, let η be a geodesic which is not a complete wall lift; then η contains a chord c , whose interior points are flat. In any wall chart $\mathcal{W}_w$, the set of geodesics crossing c transversally is open and non-empty, hence of positive measure, while the geodesics with a tangential contact form a null set by Lemma 4.5(a); so some tangency-free σ crosses c transversally at a flat point, and $\eta \in \mathcal{U}_\sigma^{\text{reg}} \cup \mathcal{U}_\sigma^{\text{reg}}$. By setting $\hat{m}_0(\mathcal{T}) = 0$ and gluing over a countable subatlas, this defines a Radon measure. Isometries of X permute the charts preserving the densities, which gives $\pi_1(K)$ -invariance; the flip acts by $(t, \theta) \mapsto (-t, \pi - \theta)$ and $(r, \phi) \mapsto (r, \pi - \phi)$ on wall charts, both preserving the densities. Nullity of tangential configurations and pencils comes from Lemma 4.5; degenerate geodesics contain a tangential junction (Lemma 3.48(iv)), hence are null. $\square$

The normalisation is chosen so that, counting both sides, the geodesics crossing a segment have mass equal to its length (Lemma 4.9). In particular, by Lemma 4.4 and the consistency shown above, formula (12) holds along every straight segment of a flat sheet and every wall arc, whether or not the segment extends to a tangency-free geodesic: the geodesics crossing such a segment J in a sector (θ_1, θ_2) have mass $|J| \int_{\theta_1}^{\theta_2} \frac{1}{2} \sin \theta d\theta$.

Let us summarise what has been proved so far, together with two consequences that will be used repeatedly.

Lemma 4.7. *Formula (12) defines a unique Radon measure $\hat{m}_0$ on $\partial_\infty^2 \widetilde{K}_0$. It is invariant under $\pi_1(K)$ and under the flip; moreover,*

- (i) *two geodesics intersect transversally at a single point if and only if their endpoint pairs are linked in $\partial_\infty \widetilde{K}_0$ and they do not share a sub-arc; on the other hand, pairs sharing a sub-arc (which are pairs of geodesics with a common tangential contact) form a $\hat{m}_0 \otimes \hat{m}_0$ -null set;*
- (ii) *(Null tangencies) the set of geodesics tangent to a fixed C^2 curve, and the set of geodesics through a fixed point, are $\hat{m}_0$ -null;*
- (iii) *(Null degenerates) the set of degenerate geodesics is $\hat{m}_0$ -null.*

Proof. (i) Linked endpoints force an intersection by a Jordan-curve argument in the closed disc $\widetilde{K}_0 \cup \partial_\infty \widetilde{K}_0$; if two geodesics had two common points they would share the segment between them (uniqueness of CAT(0) segments), whose endpoints are branch points, hence tangential contacts with the seam by Lemma 3.48(iv); in particular a shared sub-arc forces both geodesics into the tangential-contact family, which is $\hat{m}_0 \otimes \hat{m}_0$ -null by Proposition 4.6. Points (ii) and (iii) are contained in Lemma 4.5 and Proposition 4.6, because every degenerate geodesic meets the seam tangentially at the endpoints of each of its seam arcs (see Lemma 3.48(iv)). $\square$

Let $\hat{\mathcal{G}}_0^\flat \subset \hat{\mathcal{G}}_0$ be the set of parametrised geodesics γ such that $\gamma(0)$ is a flat point. The map

$$\hat{\mathcal{G}}_0^\flat \rightarrow T^1 \widetilde{K}^\circ, \quad \gamma \mapsto (\gamma(0), \dot{\gamma}(0)),$$

is well defined and continuous. It is injective on the set of regular geodesics (see Remark 3.26), and $\hat{\mathcal{G}}_0^\flat$ has full $\hat{m}_0 \otimes dt$ -measure, since the times at which a geodesic

lies on the seam form a Lebesgue-null set for regular geodesics (by Lemma 3.49) and non-regular geodesics are $\widehat{m}_0$ -null (by Lemma 4.7).

We use this map to identify, up to null sets, the flow-invariant measure $\widehat{m}_0 \otimes dt$ on $\widehat{\mathcal{G}}_0$ with a measure on $T^1\widetilde{K}^\circ$. In the coordinates (x, φ) of a developed sheet, φ being the direction angle, it is defined as follows.

Corollary 4.8. *Under the above identification, $d(\widehat{m}_0 \otimes dt) = \frac{1}{4} dx d\varphi$ on $T^1\widetilde{K}^\circ$. Consequently, the quotient $\mu := (\widehat{m}_0 \otimes dt)/\pi_1(K)$, a measure on $\widehat{\mathcal{G}}_0/\pi_1(K)$ which we call the Liouville measure of K_0 , is finite, of total mass $\pi \text{Area}(\mathcal{D})$, and invariant under the flow, the flip, and each rotation $R_{\theta_1}: (x, \varphi) \mapsto (x, \varphi + \theta_1)$. Until further notice, μ is not normalised.*

Proof. In a flat chart with the position coordinate s along the crossing geodesic, $d(\widehat{m}_0 \otimes dt) = \frac{1}{4} \sin \theta d\theta dt ds$; developing the sheet and writing $x = \sigma(t) + se_\varphi$, $\varphi = \alpha + \theta$, the Jacobian of $(t, s) \mapsto x$ at fixed θ is $|\det(e_\alpha, e_\varphi)| = \sin \theta$, so $dx d\varphi = \sin \theta d\theta dt ds$. The two flat charts $\mathcal{U}_\sigma, \mathcal{U}_{\bar{\sigma}}$ cover the two half-circles $\varphi \in (\alpha, \alpha + \pi)$ and $(\alpha + \pi, \alpha + 2\pi)$, so the formula holds for all φ . The identification makes sense a.e., since times on the seam form a Lebesgue-null set for every geodesic (by Lemma 3.49 for regular ones, while degenerate ones are $\widehat{m}_0$ -null). The invariances and the mass computation follow, R_{θ_1} clearly preserving $\frac{1}{4} dx d\varphi$ and mapping the conull regular set to itself. $\square$

Lemma 4.9 (Crofton formula). *For every geodesic segment $\gamma|_{[a,b]}$ of X , we have $\widehat{m}_0(\{\eta : \eta \cap \gamma|_{[a,b]} \neq \emptyset\}) = b - a$.*

Proof. First, off a null set, a geodesic η crosses γ at most once; indeed, two common points give a common segment (by uniqueness of CAT(0) segments), whose endpoints are branch points, hence tangential contacts of η with B (Lemma 3.48(iv)); such η are null. By Lemma 3.49, $[a, b]$ decomposes into finitely many maximal open flat intervals I_j (chords), finitely many maximal seam intervals W_i (wall arcs), and a finite set Z of times. Modulo null sets, the crossing set of $\gamma|_{[a,b]}$ is then the disjoint union of the crossing sets of the pieces; indeed, disjointness follows from the at-most-once property; the geodesics crossing exactly at a time of Z form finitely many point pencils, which are null by Lemma 4.5(c). Each chord piece is a tangency-free transversal segment; its crossers from the right form the box $\overline{I_j} \times (0, \pi)$ of the flat chart along γ restricted to I_j , of mass $|I_j| \int_0^\pi \frac{1}{4} \sin = \frac{1}{2} |I_j|$, and its crossers from the left form the corresponding box of the chart along $\bar{\gamma}$, of the same mass; together, $|I_j|$. Similarly, for each wall piece $\gamma(W_i) \subset w$, the geodesics crossing it from either side form two boxes $W_i \times (0, \pi)$ of the two wall charts of w , of total mass $|W_i|$, by (12). By summing, the total mass is then $\sum |I_j| + \sum |W_i| = (b - a) - |Z| = b - a$. $\square$

Corollary 4.10. *For every compact $Q \subset X$, $\widehat{m}_0(\{\eta : \eta \cap Q \neq \emptyset\}) \leq \frac{\pi}{4} \text{Area}(N_1(Q)) < \infty$, where $N_1(Q)$ is the closed 1-neighbourhood.*

Proof. If η meets Q , the sub-segment of η of length 2 centred at a point of $\eta \cap Q$ lies in $N_1(Q)$, so $\text{length}(\eta \cap N_1(Q)) \geq 2$. Integrating and using Corollary 4.8, we get

$$\begin{aligned} 2\widehat{m}_0(\{\eta \cap Q \neq \emptyset\}) &\leq \int \text{length}(\eta \cap N_1(Q)) d\widehat{m}_0(\eta) \\ &= (\widehat{m}_0 \otimes dt)(\{(\eta, t) : \eta(t) \in N_1(Q)\}) = \frac{\pi}{2} \text{Area}(N_1(Q)). \quad \square \end{aligned}$$

Corollary 4.11. *If $A, B \subset \partial_\infty X$ are disjoint closed arcs, the geodesics with one endpoint in A and the other in B all meet a common compact set; in particular their $\widehat{m}_0$ -mass is finite.*

Proof. Otherwise there are such geodesics η_k with $\eta_k \cap B(o, k) = \emptyset$ for all $k \geq 0$; up to passing to a subsequence, their endpoint pairs converge to $(p, q) \in A \times B$ with $p \neq q$, so by Lemma 3.42 the η_k meet a fixed ball, a contradiction. We conclude by Corollary 4.10. $\square$

Lemma 4.12. *For any $\zeta \in \partial_\infty X$, the set $\{\eta : \zeta \in \{\eta(-\infty), \eta(\infty)\}\}$ is $\widehat{m}_0$ -null.*

Proof. We work in a flat chart along σ . A geodesic with forward endpoint ζ crossing σ at $\sigma(t)$ contains the ray $r_{\sigma(t), \zeta}$, which is unique by Lemma 3.47; off the multiple-fibre null set, its angle coordinate is the Alexandrov angle $\theta_\zeta(t) := \angle_{\sigma(t)}(\sigma, r_{\sigma(t), \zeta})$. The function $t \mapsto \theta_\zeta(t)$ is Borel: the comparison angles $\overline{\angle}(\sigma(t+s), \sigma(t), r_{\sigma(t), \zeta}(s'))$ are continuous in t for fixed rational $s, s' > 0$ (Lemma 3.47), and the Alexandrov angle is their infimum over s, s' [11, II.1.7]. Hence the set $\{\eta : \zeta \in \{\eta(-\infty), \eta(\infty)\}\}$ meets the chart in a subset of the union of the graphs of two Borel functions and a null set; graphs of Borel functions are Lebesgue-null by Fubini. Covering with countably many charts, this concludes the proof. $\square$

Remark 4.13. In the seam chart used in the proof of Lemma 4.9, the restriction of $\widehat{m}_0$ to the geodesics crossing a wall Γ_i from a fixed side is $\frac{1}{4} \sin \phi d\phi dr$, i.e., it is the classical invariant collision measure of the billiard map on $\mathcal{D}$; the current $\widehat{m}_0$ is thus the geodesic-current analogue of the billiard Liouville measure (for each side; the classical collision measure $\frac{1}{2} \sin \phi d\phi dr$ of the billiard map counts each collision once, i.e., both sides of the seam together).

Recall that $\widehat{\mathcal{G}}(\mathcal{D})$ is the admissible class of Section 3 (see Definition 3.24); in particular, by Lemma 3.25 the open set of degenerate geodesics is disjoint from $\widehat{\mathcal{G}}(\mathcal{D})$.

Proposition 4.14. *The support of the current $\widehat{m}_0$ is equal to*

$$\text{supp } \widehat{m}_0 = \overline{\widehat{\mathcal{G}}^{\text{reg}}} = \widehat{\mathcal{G}}(\mathcal{D}).$$

The same holds for μ on the quotient. In particular $\widehat{m}_0$ does not have full support on the space of all geodesics.

Proof. ($\subseteq$) Non-regular geodesics, i.e., degenerate or grazing ones, are $\widehat{m}_0$ -null, since they have a tangential contact with B (see Proposition 4.6, Lemma 4.5(a)), so every set of positive mass contains regular geodesics; hence every point of the support is a limit of regular geodesics. ($\supseteq$) Let η_0 be regular and $O \ni \eta_0$ a basic C^0 -neighbourhood over a window $[-R, R]$. On that window η_0 has finitely many seam contacts (Lemma 3.49), all of them transversal since η_0 is regular. Fix a transversal segment τ inside a chord of η_0 . Transversality of finitely many crossings is an open condition on the chart data along τ : there is an open box $U \times (\theta_1, \theta_2)$ of data whose geodesics realise the same itinerary on $[-R, R]$ and depend continuously on the data there, hence lie in O for a small enough box. By (12) the box has positive mass, so $\widehat{m}_0(O) > 0$ and $\eta_0 \in \text{supp } \widehat{m}_0$; the support being closed, $\widehat{\mathcal{G}}^{\text{reg}} \subset \text{supp } \widehat{m}_0$. Finally, degenerate geodesics are not limits of regular ones (see Lemma 3.25). $\square$

Proposition 4.15 (Rectangle formula). *Let $\zeta_1, \zeta_2, \zeta_3, \zeta_4 \in \partial_\infty X$ be pairwise distinct, in the cyclic order $\zeta_1, \zeta_3, \zeta_2, \zeta_4$ (i.e., the pairs (ζ_1, ζ_2) and (ζ_3, ζ_4) are linked). Let*

$$R := \{\eta : \text{one endpoint of } \eta \text{ lies in the arc } (\zeta_1 \zeta_3), \text{ the other in the arc } (\zeta_2 \zeta_4)\},$$

the arcs being those not containing the remaining two points. Then

$$\widehat{m}_0(R) = \ln \text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) \geq 0.$$

Proof. Let $L_{41} := (\zeta_4, \zeta_1)$ and $L_{32} := (\zeta_3, \zeta_2)$ be the two “opposite sides” of the ideal quadrilateral (they exist and are unique by Lemma 3.33, and are disjoint). Choose the rays in the definition (9) on the sides: $x_1(T), x_4(T)$ the points of L_{41} at parameter $\pm T$ converging to ζ_1, ζ_4 respectively, and $x_3(T), x_2(T)$ the points of L_{32} converging to ζ_3, ζ_2 ; this is legitimate by the ray-independence in Lemma 3.41.

Let G^* be the set of geodesics η which

- are not tangent to and share no arc with L_{41}, L_{32} (so η meets each of them in at most one point, transversally, by the branching criterion of Lemma 3.48(iv) and uniqueness of segments);
- have no endpoint in $\{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$;
- pass through none of the countably many points $x_i(T)$, $T \in \mathbb{Q}$.

By Lemmas 4.5 and 4.12, G^* is conull. For $\eta \in G^*$ and any T with $x_i(T) \notin \eta$ (all but at most four values of T , namely, the crossing parameters), each $x_i(T)$ has a well-defined side of η (the components of $X \setminus \eta$; η is embedded and closed, and X is a plane).

Fix T with $x_i(T) \notin \eta$ for the four indices, write $x_i = x_i(T)$, and set $[x_i|x_j] \in \{0, 1\}$ for the indicator of “opposite sides”. Off a null set, η crosses the segment $[x_i, x_j]$ if and only if $[x_i|x_j] = 1$ (indeed, a crossing separates; conversely, separated endpoints force an odd, hence exactly one, transversal crossing, while tangencies and shared arcs with the segments are null configurations). By the Crofton formula (Lemma 4.9), $d(x_i, x_j) = \hat{m}_0(\{[x_i|x_j] = 1\})$, hence we have

$$E_0(x_1, x_2, x_3, x_4) = \int s_T(\eta) d\hat{m}_0(\eta), \quad s_T := [x_1|x_2] + [x_3|x_4] - [x_1|x_4] - [x_3|x_2].$$

Note that the integrand is integrable since each crossing set has finite mass. Enumerating the partitions of $\{x_1, x_2, x_3, x_4\}$ into sides shows that $s_T = 0$ unless the partition is $\{x_1, x_4\} \mid \{x_2, x_3\}$, where $s_T = 2$, or $\{x_1, x_2\} \mid \{x_3, x_4\}$, where $s_T = -2$. Let us write $S_+(T)$, $S_-(T)$ for the corresponding sets of geodesics; we have

$$E_0(x_1(T), x_2(T), x_3(T), x_4(T)) = 2\hat{m}_0(S_+(T)) - 2\hat{m}_0(S_-(T)). \quad (13)$$

For $\eta \in G^*$: if two points of L_{41} lie on the same side of η , then the whole segment of L_{41} between them lies on that side; indeed η crosses L_{41} at most once, and a crossing of the sub-segment would put its endpoints on opposite sides. Thus, for $T < T'$ (avoiding the finitely many exceptional parameters of η), we have

- $\eta \in S_+(T') \Rightarrow \eta \in S_+(T)$: the pairs $\{x_1(T'), x_4(T')\}$ and $\{x_3(T'), x_2(T')\}$ lie on their respective sides, hence so do the sub-segments $[x_4(T'), x_1(T')] \supset [x_4(T), x_1(T)]$ and $[x_3(T'), x_2(T')] \supset [x_3(T), x_2(T)]$.
- $\eta \in S_+(T)$, $T \geq 1$, implies that η separates $x_1(1)$ from $x_3(1)$ (same sub-segment argument), i.e., η crosses the fixed compact segment $[x_1(1), x_3(1)]$; by Crofton, the sets $S_+(T)$, $T \geq 1$, share the finite-mass dominator $D_+ := \{\eta \cap [x_1(1), x_3(1)]\}$.
- $\eta \in S_-(T)$ implies that η separates $x_1(T)$ from $x_4(T)$, both on L_{41} , hence crosses L_{41} ; in the same way, it crosses L_{32} . A geodesic crossing both L_{41} and L_{32} has its endpoints linked with both (ζ_4, ζ_1) and (ζ_3, ζ_2) , which forces (up to the null endpoint fans, see Lemma 4.12) one endpoint in the closed arc $[\zeta_4\zeta_1]$ and the other in $[\zeta_3\zeta_2]$, which are two disjoint closed arcs. By Corollary 4.11, the sets $S_-(T)$ share a finite-mass dominator D_- .

We claim that, for $\eta \in G^*$, as $T \rightarrow \infty$,

$$\mathbf{1}_{S_+(T)}(\eta) \longrightarrow \mathbf{1}_R(\eta), \quad \mathbf{1}_{S_-(T)}(\eta) \longrightarrow 0.$$

Let us show the first convergence: if $\eta \in R$, its endpoints are interior to the arcs $(\zeta_1\zeta_3)$, $(\zeta_2\zeta_4)$; then η 's endpoint pair is unlinked with (ζ_4, ζ_1) and with (ζ_3, ζ_2) , so η crosses neither side; each side, being connected and disjoint from η , lies in one component of $X \setminus \eta$, and the component is the one whose ideal arc contains its endpoints: L_{41} on the side of the arc containing ζ_4, ζ_1 , and L_{32} on the other side (the arcs of η split $\{\zeta_4, \zeta_1\}$ from $\{\zeta_3, \zeta_2\}$ since $\eta \in R$). Hence $\eta \in S_+(T)$ for every non-exceptional T . Conversely, if $\eta \in S_+(T_k)$ for integers $T_k \rightarrow \infty$: after passing to a subsequence, the assignment of pairs to the two (fixed) sides of η is constant; since $x_i(T_k) \rightarrow \zeta_i$ in $\bar{X}$ and the closures of the two components of $X \setminus \eta$ meet $\partial_\infty X$ in the two closed arcs bounded by the endpoints of η , we get ζ_1, ζ_4 in one closed arc and ζ_2, ζ_3 in the other; as $\eta \in G^*$ has no endpoint among the ζ_i , the points lie in the open arcs, and reading off the cyclic order, we get one endpoint of η in $(\zeta_1\zeta_3)$, the other in $(\zeta_2\zeta_4)$, i.e., $\eta \in R$.

For the second one: if $\eta \in S_-(T_k)$ along a subsequence with constant side assignment, the same argument puts ζ_1, ζ_2 in one closed arc of η and ζ_3, ζ_4 in the other; but in the cyclic order $\zeta_1, \zeta_3, \zeta_2, \zeta_4$ any arc containing ζ_1, ζ_2 and avoiding ζ_3, ζ_4 is impossible unless one of the ζ_i is an endpoint of η , which is excluded on G^* .

By dominated convergence, with dominators $D_\pm$ from (b), (c), we conclude that

$$\widehat{m}_0(S_+(T)) \rightarrow \widehat{m}_0(R), \quad \text{and} \quad \widehat{m}_0(S_-(T)) \rightarrow 0$$

along integers; by (13) and the definition (9),

$$2 \ln \text{cr}_0(\zeta_1, \zeta_2, \zeta_3, \zeta_4) = \lim_{T \rightarrow \infty} E_0(x_1(T), x_2(T), x_3(T), x_4(T)) = 2 \widehat{m}_0(R). \quad \square$$

4.2. The spectrum determines the cross-ratio. In this part, we use the notions introduced in Section 3, namely, the classification of isometries of CAT(0) spaces, Busemann cocycles B_ξ , Gromov products and the cross-ratio cr_0 of (9), which is well defined (Lemma 3.41) and continuous (Corollary 3.43). By Lemma 3.38, every non-trivial $g \in \pi_1(K)$ is a hyperbolic isometry of X , with a unique axis $A_g = (g^-, g^+)$ on which it translates by $\ell(g) = \mathbf{EL}([g])$, and with North–South dynamics on $\partial_\infty X$. Throughout, $g, h \in \pi_1(K)$ are two such elements with four pairwise distinct fixed points $g^\pm, h^\pm$ (equivalently, by Lemma 3.38(iv), with distinct axes), and $a \in A_g, b \in A_h$ are base points on the axes. The goal is to express $\text{cr}_0(g^-, g^+, h^-, h^+)$ through the translation lengths $\ell(g^n h^n), \ell(g), \ell(h)$.

Lemma 4.16. *For any hyperbolic isometry w and any $x \in X$, we have $\ell(w) = B_{w^+}(x, wx)$.*

Proof. By naturality of Busemann cocycles (Definition 3.35(iii)), with $\psi = w, \xi = w^+$: $B_{w^+}(wu, wv) = B_{w^+}(u, v)$. For $a \in A_w$, wa lies on the ray from a to w^+ at distance $\ell(w)$, so $B_{w^+}(a, wa) = \ell(w)$; for general x , the cocycle relation gives $B_{w^+}(x, wx) = B_{w^+}(x, a) + \ell(w) + B_{w^+}(wa, wx) = B_{w^+}(x, a) + \ell(w) + B_{w^+}(a, x) = \ell(w)$. $\square$

Lemma 4.17. *For $y, z \in X$, we have $d(y, g^n z) = n\ell(g) + B_{g^+}(y, a) + B_{g^-}(z, a) + o(1)$ as $n \rightarrow \infty$, locally uniformly in (y, z) .*

Proof. Write $d(y, g^n z) - n\ell(g) = [d(y, g^n z) - d(a, g^n z)] + [d(a, g^n z) - d(a, g^n a)]$. Since $d(g^n z, g^n a) = d(z, a)$ is bounded and $g^n a \rightarrow g^+$ radially, $g^n z \rightarrow g^+$ in the cone topology; by horofunction continuity (Definition 3.35(v)), the first bracket converges to $B_{g^+}(y, a)$, uniformly for y in compact sets. Moreover, it is 2-Lipschitz in z with a z -independent limit, so the convergence is locally uniform in (y, z) (by equi-Lipschitzness). The second bracket equals $d(g^{-n} a, z) - d(g^{-n} a, a) \rightarrow B_{g^-}(z, a)$, locally uniformly in z , by the same fact. $\square$

Lemma 4.18. *For $\xi \neq \eta \in \partial_\infty X$ and base points $c, c' \in X$, we consider the gap*

$$G(\xi, \eta; c, c') := B_\xi(p, c) + B_\eta(p, c'),$$

where p is any point on the line (ξ, η) . Note that $G(\xi, \eta; c, c') = G(\eta, \xi; c', c)$. Then

$$\inf_{y \in X} [B_\xi(y, c) + B_\eta(y, c')] = G(\xi, \eta; c, c'),$$

the infimum being attained on the line (ξ, η) . Moreover, if a moves along the axis A_g and b along A_h , the sum $G(g^+, h^-; a, b) + G(g^-, h^+; a, b)$ is unchanged.

Proof. Along (ξ, η) , parametrised towards ξ , $B_\xi(\cdot, c)$ decreases and $B_\eta(\cdot, c')$ increases at unit speed, so the sum is constant on the line; the swap symmetry is clear from the definition. For arbitrary y and $p = \gamma(0)$ on the line $\gamma = (\xi, \eta)$, $\gamma(+\infty) = \xi$:

$$\begin{aligned} [B_\xi(y, c) + B_\eta(y, c')] - G(\xi, \eta; c, c') &= B_\xi(y, p) + B_\eta(y, p) \\ &= \lim_{t, s \rightarrow \infty} [d(y, \gamma(t)) + d(y, \gamma(-s)) - (t + s)] \geq 0 \end{aligned}$$

by the triangle inequality, since $t + s = d(\gamma(t), \gamma(-s))$. Finally, replacing a by a point at parameter t_0 further on A_g changes $B_{g^+}(\cdot, a)$ by $+t_0$ and $B_{g^-}(\cdot, a)$ by $-t_0$, leaving the sum unchanged; we argue similarly for b . $\square$

Lemma 4.19. *Write $w_n := g^n h^n$ and fix $p^* \in (h^-, g^+)$. Then:*

- (i) $w_n^+ \rightarrow g^+$ and $w_n^- \rightarrow h^-$;
- (ii) let $y_n \in A_{w_n}$ be nearest to p^* ; then the sequence (y_n) is bounded;
- (iii) let $\zeta_n \in [y_n, w_n y_n] \subset A_{w_n}$ be nearest to $g^n p^*$; the points $z_n := g^{-n} \zeta_n$ form a bounded sequence.

Proof. (i) Hyperbolic isometries of a proper δ -hyperbolic space act on the boundary with North–South dynamics (by fact (d) following Definition 3.37; see also [17, Ch. 10]). Choose disjoint closed neighbourhoods $U^\pm \ni g^\pm$, $V^\pm \ni h^\pm$ in $\bar{X}$, with $V^+ \cap U^- = \emptyset$. For large n , $h^n(\bar{X} \setminus V^-) \subset V^+$ and $g^n(\bar{X} \setminus U^-) \subset U^+$, so $w_n(\bar{X} \setminus V^-) \subset U^+$; iterating, $w_n^k(\bar{X} \setminus V^-) \subset U^+$ for all $k \geq 1$, and since $w_n^k x \rightarrow w_n^+$ as $k \rightarrow \infty$ for every $x \in X \cap U^+$, while $w_n^k(U^+) \subset U^+$ with U^+ closed, we have $w_n^+ \in U^+$. Applying this to $w_n^{-1} = h^{-n} g^{-n}$ gives $w_n^- \in V^-$. By shrinking the neighbourhoods, we obtain (i).

(ii) The axes $A_{w_n} = (w_n^-, w_n^+)$ have endpoints converging to (h^-, g^+) ; by Lemma 3.42 they converge to $(h^-, g^+) \ni p^*$ uniformly on compacts, so $d(p^*, A_{w_n})$ is bounded.

(iii) In a δ -hyperbolic space (recall Corollary 3.32), we have $d(m, [u, v]) \leq (u | v)_m + 4\delta$ for any geodesic segment [17, Ch. 3]. With $m = g^n p^*$, $u = y_n$, $v = w_n y_n$, and using invariance of Gromov products (Definition 3.21) under the isometry g^{-n} , we obtain

$$(y_n | w_n y_n)_{g^n p^*} = (g^{-n} y_n | h^n y_n)_{p^*}.$$

Since the y_n are bounded, $g^{-n} y_n \rightarrow g^-$ and $h^n y_n \rightarrow h^+$, which are two distinct boundary points (Definition 3.21(iv)), hence the Gromov products are bounded, and $d(z_n, p^*) = d(\zeta_n, g^n p^*) = d(g^n p^*, [y_n, w_n y_n]) \leq C$. $\square$

Theorem 4.20. *We have the following length asymptotics:*

$$\lim_{n \rightarrow \infty} [\ell(g^n h^n) - n\ell(g) - n\ell(h)] = G(g^+, h^-; a, b) + G(g^-, h^+; a, b).$$

Proof. Let us start by showing the upper bound. Fix $y \in (g^+, h^-)$ and $z \in (g^-, h^+)$. For every n , we have

$$\ell(w_n) \leq d(y, w_n y) \leq d(y, g^n z) + d(g^n z, g^n h^n y) = d(y, g^n z) + d(z, h^n y),$$

and by Lemma 4.17 (for g , and for h with base $b \in A_h$),

$$\begin{aligned} \limsup_n [\ell(w_n) - n\ell(g) - n\ell(h)] &\leq [B_{g^+}(y, a) + B_{h^-}(y, b)] + [B_{g^-}(z, a) + B_{h^+}(z, b)] \\ &= G(g^+, h^-; a, b) + G(g^-, h^+; a, b), \end{aligned}$$

where the equalities hold because y, z lie on the respective lines (Lemma 4.18).

Let us now show the lower bound. With y_n, ζ_n, z_n as in Lemma 4.19: since $\zeta_n = g^n z_n$ lies on the segment $[y_n, w_n y_n]$, whose length is $\ell(w_n)$, we have

$$\ell(w_n) = d(y_n, g^n z_n) + d(g^n z_n, w_n y_n) = d(y_n, g^n z_n) + d(z_n, h^n y_n).$$

The sequences $(y_n), (z_n)$ are bounded, so Lemma 4.17 applies locally uniformly, and each of the two resulting brackets is bounded below by the corresponding gap (by Lemma 4.18 at y_n, z_n); hence $\liminf_n \geq G(g^+, h^-; a, b) + G(g^-, h^+; a, b)$. $\square$

Gathering the previous results, we can conclude that the spectrum determines cross-ratios:

Corollary 4.21. *We have*

$$\ln \text{cr}_0(g^-, g^+, h^-, h^+) = -\frac{1}{2} \lim_{n \rightarrow \infty} [\ell(g^n h^n) - n\ell(g) - n\ell(h)]. \quad (14)$$

Proof. For $\xi, \eta \in \partial_\infty X$ distinct and base points $c, c' \in X$, let u, v be the rays from c, c' to ξ, η , and set $P(\xi, \eta; c, c') := \lim_t [2t - d(u_t, v_t)]$; the diagonal limit agrees with the double limit $\lim_{t,s} [t + s - d(u_t, v_s)]$, both expressions being bounded and monotone in each variable by first variation (as in Lemma 3.41). By the identity established in the proof of Corollary 3.43, for every p on the line (ξ, η) ,

$$P(\xi, \eta; c, c') = -[B_\xi(p, c) + B_\eta(p, c')] = -G(\xi, \eta; c, c'). \quad (15)$$

Let us now evaluate (9) with the rays $\gamma_1(t) = \sigma_1(-t)$, $\gamma_2(t) = \sigma_1(t)$, $\gamma_3(t) = \sigma_2(-t)$, $\gamma_4(t) = \sigma_2(t)$, where $\sigma_1 = A_g$ is parametrised towards g^+ with $\sigma_1(0) = a$, and $\sigma_2 = A_h$ towards h^+ with $\sigma_2(0) = b$ (ray-independence, Lemma 3.41). With $f_{ij}(t) = d(\gamma_i(t), \gamma_j(t)) - 2t$: $f_{12} = f_{34} \equiv 0$, $\lim f_{14} = -P(g^-, h^+; a, b)$, $\lim f_{32} = -P(h^-, g^+; b, a)$, so

$$\begin{aligned} \ln \text{cr}_0(g^-, g^+, h^-, h^+) &= \frac{1}{2} \lim_t E_0 = \frac{1}{2} [P(g^-, h^+; a, b) + P(h^-, g^+; b, a)] \\ &= -\frac{1}{2} [G(g^-, h^+; a, b) + G(h^-, g^+; b, a)]. \end{aligned}$$

By Lemma 4.18, we have $G(h^-, g^+; b, a) = G(g^+, h^-; a, b)$, and by Theorem 4.20, this concludes the proof. $\square$

Remark 4.22. The principle that cross-ratios can be recovered from length data (as in Theorem 4.20 and Corollary 4.21) goes back to Otal's proof that the marked length spectrum of a negatively curved surface determines its Liouville cross-ratio [39, Th. 2.2]. It was extended to CAT(-1) spaces by Bourdon [9, 10], whose formulation of the cross-ratio through Busemann functions is the one we use here, to CAT(-1) spaces with cone singularities by Hersensky and Paulin [30], and to symmetric spaces by Dal'bo and Kim [19, 33, 34]. Otal's global scheme was carried over to non-positively curved surfaces with cone singularities by Constantine [16] (see also [23, 7] for flat cone metrics). Theorem 4.20 shows that the limit is exact in a CAT(0) space with flat pieces, using only Gromov hyperbolicity and the absence of flat strips.

For hyperbolic billiards, in particular, open dispersing billiards, similar asymptotic expansions have already been considered in several e.g. in [5, 20]. In this purely hyperbolic setting, the goal was to recover periodic multipliers from length data through exponentially small error terms in such expansions. Here, we do not claim

any estimate on the speed of convergence in (14); instead, we focus directly on the limit of the expansions.

4.3. Preservation of angles. In the following, we suppose that $\mathcal{D}$ and $\mathcal{D}'$ are two distinct billiard tables in $\mathfrak{D}^m \approx C_{\text{emb}}^r(\mathbb{T}_{\mathcal{D}}, \mathbb{T}^2)$. We consider the latter as a metric space with the C^r metric. Remark that the tables $\mathcal{D}$ and $\mathcal{D}'$ are diffeomorphic. Indeed, the space of configurations of m pairwise disjoint closed discs in $\mathbb{T}^2$ is connected (we can shrink the discs to points, use the connectedness of the configuration space of m distinct points of $\mathbb{T}^2$, and enlarge the points back to discs); since the obstacles of $\mathcal{D}$ and $\mathcal{D}'$ are isotopic to discs through strictly convex domains, there is an isotopy of embeddings of $\mathbb{T}_{\mathcal{D}} = \bigsqcup_{i=1}^m \mathbb{S}^1$ carrying $\partial\mathcal{D}$ onto $\partial\mathcal{D}'$, and by the isotopy extension theorem (see e.g. [32, Chapter 8, Theorem 1.3]) it extends to a diffeotopy of $\mathbb{T}^2$, whose time-one map sends $\mathcal{D}$ onto $\mathcal{D}'$.

We will consider surfaces $K \subset \mathcal{K}(\mathcal{D})$, $K' \subset \mathcal{K}(\mathcal{D}')$, obtained by gluing two copies of $\mathcal{D}$, respectively $\mathcal{D}'$, that is

$$K \setminus B(K) = \mathcal{D}^{\text{up}} \coprod \mathcal{D}^{\text{down}}, \quad K' \setminus B(K') = (\mathcal{D}')^{\text{up}} \coprod (\mathcal{D}')^{\text{down}}.$$

The surfaces K and K' are also diffeomorphic. Indeed, $p(\mathcal{D}^{\text{up}}) \approx \mathcal{D} \approx \mathcal{D}' = p((\mathcal{D}')^{\text{up}})$, and similarly for $\mathcal{D}^{\text{down}}$. As $\partial\mathcal{D} \approx \partial\mathcal{D}'$, we obtain a diffeomorphism $h: K \rightarrow K'$, as wanted. We fix such a diffeomorphism once and for all; it induces an isomorphism $\phi := h_*: \pi_1(K) \rightarrow \pi_1(K')$, the *marking*, well defined up to the action of the mapping class group.

In what follows we assume that $\mathcal{D}$ and $\mathcal{D}'$ have the same enriched length spectrum:

$$\mathbf{EL}_{\mathcal{D}} = \mathbf{EL} = \mathbf{EL}' = \mathbf{EL}_{\mathcal{D}'}.$$

Let us stress that the hypothesis $\mathbf{EL} = \mathbf{EL}'$ is understood with respect to the marking ϕ , i.e., $\mathbf{EL}'(\phi(\alpha)) = \mathbf{EL}(\alpha)$ for every conjugacy class $\alpha \in \mathcal{C}(K)$. The conclusion of the main theorem, i.e., the existence of an isometry of $\mathbb{T}^2$, does not depend on the marking, but the hypothesis does, and the isometry obtained need not induce ϕ .

In what follows, we will fix K as our reference surface. As in Section 3, we will consider families of metrics $g_\varepsilon, g'_\varepsilon$ on K converging to the gluing metric when ε approaches 0. More specifically, $g = g_1$ is our reference metric on K , induced by the Euclidean metric of $E = \mathbb{T}^2 \times \mathbb{R}$. For each of the tables, the metrics $g_\varepsilon, g'_\varepsilon$ are defined as the pullback to K , under the mapping $\alpha_\varepsilon: (x, y, z) \mapsto (x, y, \varepsilon z)$, of the metric induced on $\alpha_\varepsilon(K), \alpha_\varepsilon(K')$, respectively. The g'_ε are seen as metrics in K by pulling back with the diffeomorphism between the surfaces.

We have metric convergence $(K, g_\varepsilon) =: K_\varepsilon \xrightarrow{GH} K_0 := (K, g_0)$, and similarly $(K, g'_\varepsilon) =: K'_\varepsilon \xrightarrow{GH} K'_0 := (K, g'_0)$, in the Gromov–Hausdorff distance, as $\varepsilon \rightarrow 0$. The metric space K_0 (respectively K'_0) consists of two copies of $(\mathcal{D}, d_{\mathcal{D}})$ (respectively $(\mathcal{D}', d_{\mathcal{D}'})$) glued at the boundary. In the following, we abbreviate $X = \widetilde{K}_0$, $X' = \widetilde{K}'_0$. We denote by $\mathcal{G}_0, \mathcal{G}'_0$ the spaces of unparametrised oriented geodesics, and by $\mathcal{G}(\mathcal{D}), \mathcal{G}(\mathcal{D}')$ the subsets of admissible ones.

Lemma 4.23 (Boundary identification). *The marking isomorphism $\phi: \pi_1(K) \rightarrow \pi_1(K')$ induces a homeomorphism*

$$\partial\phi: \partial_\infty X \longrightarrow \partial_\infty X',$$

which is equivariant with respect to the actions of $\pi_1(K)$ (on the right via ϕ), and mapping, for every hyperbolic $g \in \pi_1(K)$, the fixed pair (g^-, g^+) to $(\phi(g)^-, \phi(g)^+)$.

Proof. Both actions are free, properly discontinuous and cocompact by isometries on the proper geodesic spaces X, X' , which are Gromov-hyperbolic (Lemma 3.22

and its analogue for $\mathcal{D}'$). By the Švarc–Milnor lemma [11, Proposition I.8.19] the orbit maps are quasi-isometries, so ϕ induces an equivariant quasi-isometry $X \rightarrow X'$, and every quasi-isometry between hyperbolic proper geodesic spaces extends to a homeomorphism of the boundaries [11, Theorem III.H.3.9]. Equivariance gives the statement about fixed pairs, since $g^\pm$ are the unique fixed points of g on the boundary. Finally, $\partial\phi$ does not depend on the choice of orbit maps: two equivariant quasi-isometries induced by ϕ are at bounded distance from each other, hence induce the same boundary map. $\square$

Remark 4.24. One could try to identify the boundaries $\partial_\infty X$ and $\partial_\infty X'$ through the Anosov approximations instead: indeed, an isotopy $(\mathcal{D}^t)_{t \in [0,1]}$ of tables from $\mathcal{D}$ to $\mathcal{D}'$ induces an isotopy $(K^t)_{t \in [0,1]}$ of surfaces, and if, for some fixed $\varepsilon > 0$, the geodesic flows of all the (K^t, g_ε) are Anosov, they are orbit equivalent by structural stability, so that their boundaries at infinity, hence (by Lemma 3.34) those of the limit spaces, are identified; by uniqueness of equivariant boundary maps, this identification coincides with $\partial\phi$. Two points would however need to be verified: that the isotopy can be chosen through finite-horizon tables with strictly convex obstacles (finite horizon is an open but not a convex condition), and that Kourganoff’s threshold ε_0 can be taken uniform along the isotopy. The use of Švarc–Milnor lemma avoids both issues, and uses only the marking ϕ .

The next proposition says that the enriched spectrum determines the cross-ratio and the current:

Proposition 4.25. *If $\mathbf{EL} = \mathbf{EL}'$, then we have $\text{cr}'_0 \circ (\partial\phi)^{\times 4} = \text{cr}_0$ on quadruples of pairwise distinct boundary points. Therefore, $(\partial\phi \times \partial\phi)_* \hat{m}_0 = \hat{m}'_0$.*

Proof. By Remark 4.2, the hypothesis gives $\ell(w) = \ell'(\phi(w))$ for every $w \in \pi_1(K)$. Corollary 4.21, applied to (g, h) in X and to $(\phi(g), \phi(h))$ in X' , together with $\partial\phi(g^\pm) = \phi(g)^\pm$ (Lemma 4.23), gives the equality on all quadruples (g^-, g^+, h^-, h^+) . Given an ordered quadruple $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)$ of pairwise distinct points, choose, by density of pole pairs (Corollary 3.40), hyperbolic elements g_k, h_k with $(g_k^-, g_k^+) \rightarrow (\zeta_1, \zeta_2)$ and $(h_k^-, h_k^+) \rightarrow (\zeta_3, \zeta_4)$; both cross-ratios are continuous on pairwise distinct quadruples (Corollary 3.43) and $\partial\phi$ is a homeomorphism, so the identity passes to the limit.

Let us now show the second statement. We claim that a flip-invariant Radon measure on $\partial_\infty^2 X$ giving zero mass to endpoint fans is determined by its values on the rectangles R of Proposition 4.15. Indeed, ordered products $A \times B$ of open arcs with disjoint closures form a π -system generating the Borel σ -algebra of $\partial^2 \setminus \Delta$, on which our measures are supported and σ -finite (Corollary 4.11). The unordered rectangles determine the ordered products by flip invariance, and arcs with common endpoints are handled by monotone approximation from inside, the endpoint fans being null by Lemma 4.12. Uniqueness follows from Dynkin’s lemma.

Since $\partial\phi$ transports cross-ratios, by Proposition 4.15 applied on both sides, the two measures of the statement agree on all rectangles; and $(\partial\phi \times \partial\phi)_* \hat{m}_0$ gives zero mass to endpoint fans since $\partial\phi$ is a bijection. $\square$

Below, we let $\phi: \pi_1(K) \rightarrow \pi_1(K')$ as in Lemma 4.23.

Definition 4.26 (Unparametrised conjugacy). For every geodesic $\gamma \in \mathcal{G}_0$, let $h\gamma \in \mathcal{G}'_0$ be the unique geodesic of X' with endpoints

$$(h\gamma)(-\infty) = \partial\phi(\gamma(-\infty)), \quad (h\gamma)(+\infty) = \partial\phi(\gamma(+\infty)),$$

which exists and is unique by Lemma 3.33 applied to X' .

Lemma 4.27. *The map $h: \mathcal{G}_0 \rightarrow \mathcal{G}'_0$ is a homeomorphism of the spaces of unparametrised geodesics; it is $\pi_1(K)$ -equivariant over ϕ , it commutes with the flip, and preserves the Liouville currents: $h_*\hat{m}_0 = \hat{m}'_0$. In particular, $h(\hat{\mathcal{G}}(\mathcal{D})) = \hat{\mathcal{G}}(\mathcal{D}')$.*

Proof. Under the identifications $\mathcal{G}_0 \approx \partial_\infty^2 X$ and $\mathcal{G}'_0 \approx \partial_\infty^2 X'$ (Lemma 3.33 and Lemma 3.42), h is $\partial\phi \times \partial\phi$, which is a homeomorphism; equivariance and the commutation with the flip $(\xi, \eta) \mapsto (\eta, \xi)$ are inherited from $\partial\phi$ (Lemma 4.23). Moreover, $(\partial\phi \times \partial\phi)_*\hat{m}_0 = \hat{m}'_0$ follows from Proposition 4.25. The homeomorphism h carries the support of the measure $\hat{m}_0$ to the support of its push-forward $\hat{m}'_0$; by Proposition 4.14, we deduce that h maps $\text{supp } \hat{m}_0 = \hat{\mathcal{G}}(\mathcal{D})$ onto $\text{supp } \hat{m}'_0 = \hat{\mathcal{G}}(\mathcal{D}')$. $\square$

Definition 4.28. We call a geodesic *clean* if it meets the seam only transversally.

Clean geodesics are neither degenerate nor tangent to a wall, they share no arc with any other geodesic (indeed, a shared compact arc has branch points at its ends, which are tangential contacts by Lemma 3.48(iv); moreover, a shared ray is excluded for pairs with distinct endpoints), and they cross any geodesic they are linked with in exactly one point (as linked pairs intersect, and two intersection points would give a shared segment).

Next lemma shows the stability of crossing points.

Lemma 4.29. *We work in X or X' . Let $\rho_n \rightarrow \rho$ uniformly on compacts with $(\rho_n(-\infty), \rho_n(\infty)) \rightarrow (\rho(-\infty), \rho(\infty))$, and let τ be a geodesic with endpoints disjoint from those of ρ , linked with ρ_n for all large n , and meeting ρ in exactly one point x^* . If $x_n \in \tau \cap \rho_n$, then $x_n \rightarrow x^*$.*

Proof. Assume $d(x^*, x_{n_k}) \rightarrow \infty$. Since $x_{n_k} \in \tau$, the accumulation points of (x_{n_k}) in $\bar{X}$ lie in $\{\tau(\pm\infty)\}$; since $x_{n_k} \in \rho_{n_k}$ with the ρ_{n_k} as in Lemma 3.42, they also lie in $\{\rho(\pm\infty)\}$ by part (ii) of that lemma, a contradiction, the two pairs being disjoint. Moreover, any sublimit $\tilde{x}$ of the bounded sequence lies on τ and on ρ (by uniform convergence on the compact containing the tail), hence $\tilde{x} = x^*$; we conclude that $x_n \rightarrow x^*$. $\square$

Recall the CAT(0) angle from §3.2.5, and recall the measure μ of Corollary 4.8, which is supported on the admissible class (Proposition 4.14). A point of $\hat{\mathcal{G}}_0/\pi_1(K)$ is a *pointed geodesic* (γ, t) , i.e., a geodesic with a marked point $\gamma(t)$; for μ -a.e. (γ, t) the point $\gamma(t)$ is flat and, for a.e. $\theta \in (0, \pi)$, there is a unique geodesic $\eta_{\gamma, t, \theta}$ crossing γ at $\gamma(t)$ with oriented angle θ (Lemma 4.5); we write $\eta_{\gamma, \theta}$ when $t = 0$.

In the following, we say that a geodesic η *crosses γ from the right* if η and γ are linked and the oriented crossing angle from γ to η lies in $(0, \pi)$; equivalently, the endpoints occur in the cyclic order $\gamma^-, \eta^-, \gamma^+, \eta^+$ (the other cyclic order of a linked pair, $\gamma^-, \eta^+, \gamma^+, \eta^-$, corresponds to crossing from the left). Let

$$\mathcal{P}^+ := \{(\gamma, \eta) \in \hat{\mathcal{G}}_0 \times \hat{\mathcal{G}}_0 : \eta \text{ crosses } \gamma \text{ from the right}\},$$

with the diagonal action of $\pi_1(K)$, and define $\mathcal{P}^-$, $\mathcal{P}'^\pm$ likewise. For μ -a.e. pointed geodesic (γ, t) and every $\theta \in (0, \pi)$ there is a unique $\eta_{\gamma, t, \theta}$ crossing γ at $\gamma(t)$ from the right with angle θ , and every right crosser of γ arises this way for a unique (t, θ) ; thus

$$(\gamma, t, \theta) \longmapsto (\gamma, \eta_{\gamma, t, \theta}) \tag{16}$$

is an a.e. bijection from $\hat{\mathcal{G}}_0 \times (0, \pi)$ onto $\mathcal{P}^+$, equivariant for $\pi_1(K)$, which acts trivially on θ .

Let us now define a product measure on crossing pairs:

- Lemma 4.30.** (i) In the coordinates (16), the restriction of $\widehat{m}_0 \otimes \widehat{m}_0$ to $\mathcal{P}^+$ is $(\widehat{m}_0 \otimes dt) \otimes \frac{1}{4} \sin \theta d\theta$; hence the quotient of $(\widehat{m}_0 \otimes \widehat{m}_0)|_{\mathcal{P}^+}$ by $\pi_1(K)$ is $\mu \otimes \frac{1}{4} \sin \theta d\theta$. The same holds for $\mathcal{P}^-$, $\mathcal{P}'^+$ and $\mathcal{P}'^-$, with μ' .
- (ii) The map $h \times h$ sends $\mathcal{P}^+$ onto $\mathcal{P}'^+$ if $\partial\phi$ preserves the cyclic orientation of $\partial_\infty X$, and onto $\mathcal{P}'^-$ if it reverses it; in both cases it carries $(\widehat{m}_0 \otimes \widehat{m}_0)|_{\mathcal{P}^+}$ to the restricted product measure of the image family, and descends to the quotients.
- (iii) Consequently $\mu(\widehat{\mathcal{G}}_0/\pi_1(K)) = \mu'(\widehat{\mathcal{G}}'_0/\pi_1(K'))$, i.e. $\text{Area}(\mathcal{D}) = \text{Area}(\mathcal{D}')$. From now on μ and μ' are normalised to probability measures.

Proof. (i) By Fubini, $(\widehat{m}_0 \otimes \widehat{m}_0)(A) = \int \widehat{m}_0(\{\eta : (\gamma, \eta) \in A\}) d\widehat{m}_0(\gamma)$ for $A \subset \mathcal{P}^+$. For a.e. fixed γ , the right crossers of γ are parametrised by (t, θ) , and by the flux formula (12) along γ (valid along every straight segment and wall arc, by the remark following Proposition 4.6) the measure $\widehat{m}_0$ restricted to them is $\frac{1}{4} \sin \theta d\theta dt$. This shows the first claim; the quotient statement follows since $\pi_1(K)$ acts on the (γ, t) factor only and $(\widehat{m}_0 \otimes dt)/\pi_1(K) = \mu$. The other three families are treated in the same way (for left crossers, we use the chart along the reversed geodesic, and the density is the same).

(ii) $\partial\phi \times \partial\phi$ preserves linking, and it preserves or reverses all cyclic orders of $\partial_\infty X$ according to its orientation behaviour; by the characterisation of the side of crossing through cyclic order, $h \times h$ maps $\mathcal{P}^+$ onto $\mathcal{P}'^+$ or $\mathcal{P}'^-$. Since $h_*\widehat{m}_0 = \widehat{m}'_0$ (Lemma 4.27), $(h \times h)_*(\widehat{m}_0 \otimes \widehat{m}_0) = \widehat{m}'_0 \otimes \widehat{m}'_0$, and restrictions to corresponding sets correspond; equivariance over ϕ gives the quotient statement.

(iii) By (i) and (ii), the two quotient measures $\mu \otimes \frac{1}{4} \sin \theta d\theta$ and $\mu' \otimes \frac{1}{4} \sin \theta' d\theta'$ have the same total mass, i.e., $\frac{1}{2}\mu(\widehat{\mathcal{G}}_0/\pi_1(K)) = \frac{1}{2}\mu'(\widehat{\mathcal{G}}'_0/\pi_1(K'))$. By Corollary 4.8, these masses are equal to $\pi \text{Area}(\mathcal{D})$ and $\pi \text{Area}(\mathcal{D}')$. $\square$

Definition 4.31 (Angle transfer). For μ -a.e. (γ, t) and a.e. θ with $\eta_{\gamma, t, \theta}$ crossing γ from the right, the geodesics $h\gamma$ and $h\eta_{\gamma, t, \theta}$ are clean, linked, and cross at a unique point q (by Lemma 4.27 and Proposition 4.6). We let

$$\vartheta_{(\gamma, t)}(\theta) := \angle_q(h\gamma, h\eta_{\gamma, t, \theta}) \in (0, \pi)$$

be the angle between the forward directions. The same construction for h^{-1} defines $\vartheta'_{(\gamma', s)}$ for μ' -a.e. pointed geodesic of X' .

By flip equivariance of h , we have $\vartheta_{(\gamma, t)}(\pi - \theta) = \pi - \vartheta_{(\gamma, t)}(\theta)$.

The following lemma shows that the angle transfer is monotone and satisfies a superadditivity property.

Lemma 4.32. For μ -a.e. (γ, t) and a.e. $\theta_1, \theta_2 > 0$ with $\theta_1 + \theta_2 < \pi$, writing $\eta_1 = \eta_{\gamma, t, \theta_1}$ based at its crossing with γ ,

$$\vartheta_{(\gamma, t)}(\theta_1 + \theta_2) \geq \vartheta_{(\gamma, t)}(\theta_1) + \vartheta_{(\eta_1, 0)}(\theta_2).$$

In particular $\theta \mapsto \vartheta_{(\gamma, t)}(\theta)$ is non-decreasing, and $\vartheta_{(\gamma, t)}(\pi/2) = \pi/2$.

Proof. Let $\eta_2 = \eta_{\gamma, t, \theta_1 + \theta_2}$; at the flat point $x = \gamma(t)$ it makes oriented angle θ_2 with η_1 . The three image geodesics are pairwise linked and cross pairwise exactly once: $p_1 = h\gamma \cap h\eta_1$, $p_2 = h\gamma \cap h\eta_2$, $q = h\eta_1 \cap h\eta_2$. If $p_1 = p_2$ then $q = p_1$ and the claim is arc additivity in the circle of directions at that point (Lemma 3.48(ii)), with equality. Otherwise $p_1 p_2 q$ is a geodesic triangle, such that the sum of its interior angles is at most π [11, II.1.7]. The cyclic order $\gamma^+, \eta_1^+, \eta_2^+, \gamma^-, \eta_1^-, \eta_2^-$ of the endpoints is preserved by $\partial\phi$, so q lies on the side S of $h\gamma$ whose ideal arc contains η_1^+, η_2^+ , or on the opposite side.

If $q \in S$, the forward rays of $h\eta_i$ from q cannot recross $h\gamma$, so p_1, p_2 lie backward from q along both, and, since $h\eta_2$ passes from the left of $h\eta_1$ (at η_2^+) to its right, p_2 lies forward of p_1 on $h\gamma$. Writing u for the forward direction of $h\gamma$ and v_i for those of $h\eta_i$, the interior angles are $\angle_{p_1}(u, v_1) = \vartheta(\theta_1)$, $\angle_{p_2}(-u, v_2) = \pi - \vartheta(\theta_1 + \theta_2)$ (Lemma 3.48(i)), and $\angle_q(-v_1, -v_2) = \angle_q(v_1, v_2) = \vartheta_{(\eta_1, 0)}(\theta_2)$.

If $q \notin S$, then p_1, p_2 lie forward from q along both image geodesics, $h\eta_2$ passes from the right of $h\eta_1$ to its left, so p_2 lies backward of p_1 on $h\gamma$, and the interior angles are $\angle_{p_1}(-u, -v_1) = \vartheta(\theta_1)$, $\angle_{p_2}(u, -v_2) = \pi - \vartheta(\theta_1 + \theta_2)$ and $\angle_q(v_1, v_2) = \vartheta_{(\eta_1, 0)}(\theta_2)$: the same three values. In both cases the claim is proved. Monotonicity follows since angles are non-negative, and $\vartheta_{(\gamma, t)}(\pi/2) = \pi/2$ from the flip symmetry. $\square$

For μ -a.e. (γ, t) , the function $\theta \mapsto \vartheta_{(\gamma, t)}(\theta)$ is defined for a.e. $\theta \in (0, \pi)$ (Definition 4.31), takes values in $(0, \pi)$, is non-decreasing on its domain (Lemma 4.32), and satisfies $\vartheta(\pi - \theta) = \pi - \vartheta(\theta)$. A function which is non-decreasing on a conull subset of $(0, \pi)$ agrees a.e. with a non-decreasing function defined on all of $(0, \pi)$ (take $\theta \mapsto \sup\{\vartheta(\theta') : \theta' \leq \theta \text{ in the domain}\}$); since every quantity below is an integral in θ , we replace $\vartheta_{(\gamma, t)}$ by this extension, which is then non-decreasing everywhere. We also set $\vartheta(\pi/2) := \pi/2$, which is compatible with monotonicity since the symmetry gives $\vartheta(\theta) \leq \pi/2 \leq \vartheta(\theta')$ for $\theta < \pi/2 < \theta'$ in the domain. We have $0 \leq \vartheta(0^+)$ and $\vartheta(\pi^-) \leq \pi$ (a jump at the endpoints is not excluded, but is unimportant).

We define its *generalised inverse* as

$$\vartheta_{(\gamma, t)}^{-1}(b) := \sup\{\theta \in (0, \pi) : \vartheta_{(\gamma, t)}(\theta) < b\} \in [0, \pi], \quad b \in (0, \pi),$$

with $\sup \emptyset = 0$. On the conull set of clean configurations the crossing point $h\gamma \cap h\eta_{\gamma, t, \theta}$ depends continuously on (γ, t, θ) (Lemma 4.29) and the Alexandrov angle is the infimum of the comparison angles $\bar{Z}(\gamma(s), x, \gamma'(s'))$ over rational $s, s' > 0$, these being continuous in the configuration and non-decreasing in s, s' [11, II.3.1], so ϑ is Borel in (γ, t, θ) ; hence $\vartheta^{-1}(b) = \sup\{\theta \in \mathbb{Q} \cap (0, \pi) : \vartheta(\theta) < b\}$ is Borel in (γ, t) ; in particular, Fubini can be applied to the integrals below.

Lemma 4.33. *Let $f : (0, \pi) \rightarrow (0, \pi)$ be non-decreasing with $f(\pi/2) = \pi/2$, and let f^{-1} be its generalised inverse. Then:*

- (i) *for every $b \in (0, \pi)$, $\{\theta : f(\theta) < b\}$ is an interval with endpoints 0 and $f^{-1}(b)$, so that $\int_{\{f < b\}} \frac{1}{4} \sin \theta \, d\theta = \frac{1}{4}(1 - \cos f^{-1}(b))$;*
- (ii) *f^{-1} is non-decreasing, and $f^{-1}(b) \leq \pi/2$ for $b \leq \pi/2$;*
- (iii) *(Young's identity) $\int_0^{\pi/2} f(\theta) \, d\theta + \int_0^{\pi/2} f^{-1}(b) \, db = \left(\frac{\pi}{2}\right)^2$.*

Proof. (i) If $f(\theta) < b$ and $\theta' < \theta$ then $f(\theta') \leq f(\theta) < b$, so $\{f < b\}$ is an initial interval, whose supremum is $f^{-1}(b)$ by definition; whether the endpoint belongs to it is irrelevant for the integral.

(ii) Monotonicity of f^{-1} is clear from the definition; for $b \leq \pi/2$ and $\theta > \pi/2$ we have $f(\theta) \geq f(\pi/2) = \pi/2 \geq b$, so such θ do not enter the supremum.

(iii) For $\theta, b \in (0, \pi/2)$ one has $\theta < f^{-1}(b) \Rightarrow f(\theta) < b \Rightarrow \theta \leq f^{-1}(b)$ (the first once because some $\theta' > \theta$ has $f(\theta') < b$, the second one by definition), so the sets $\{(\theta, b) : b \leq f(\theta)\}$ and $\{(\theta, b) : \theta < f^{-1}(b)\}$ partition $(0, \pi/2)^2$ up to the null set $\{\theta = f^{-1}(b)\}$. Since $f \leq \pi/2$ on $(0, \pi/2]$ by (ii) applied to f itself ($f(\theta) \leq f(\pi/2) = \pi/2$), the first set has area $\int_0^{\pi/2} f$, the second $\int_0^{\pi/2} f^{-1}$, and their sum is the area of the square. $\square$

Lemma 4.34. *For every $b \in (0, \pi)$, we have*

$$\begin{aligned} \int \cos(\vartheta_{(\gamma,t)}^{-1}(b)) d\mu(\gamma, t) &= \cos b, \\ \int \cos(\vartheta_{(\gamma',s)}^{-1}(b)) d\mu'(\gamma', s) &= \cos b. \end{aligned}$$

Proof. Consider the set of crossing pairs whose image angle is $< b$, modulo π_1 . By Lemma 4.30 its mass can be computed in image coordinates: $\int_0^b \frac{1}{4} \sin \theta' d\theta' d\mu' = \frac{1}{4}(1 - \cos b)$. In source coordinates the same set is $\{(\gamma, t, \theta) : \vartheta_{(\gamma,t)}(\theta) < b\} = \{\theta < \vartheta_{(\gamma,t)}^{-1}(b)\}$ up to an endpoint, by Lemma 4.33(i), with mass $\int \frac{1}{4}(1 - \cos \vartheta_{(\gamma,t)}^{-1}(b)) d\mu$. This gives the first identity. The same argument for h^{-1} gives the second one. $\square$

Gathering, we can now conclude that h is angle preserving.

Theorem 4.35. *For μ -a.e. (γ, t) and every $\theta \in (0, \pi)$, we have*

$$\vartheta_{(\gamma,t)}(\theta) = \theta.$$

Therefore, by Lemma 4.29 and the continuity of adjacent angles (their sum being π at a transversal crossing), for every pair of admissible geodesics crossing at a point x , the images cross at a point q with $\angle_q(h\gamma, h\eta) = \angle_x(\gamma, \eta)$.

Proof. Following Otal [39], let us define the averaged angle

$$\Theta(\theta) := \int \vartheta_{(\gamma,t)}(\theta) d\mu.$$

By Lemma 4.32 each $\vartheta_{(\gamma,t)}$ is a monotone map of $(0, \pi/2]$ into itself fixing $\pi/2$. For a.e. $a \in (0, \pi/2)$,⁶ by integrating the superadditivity at $a + (\pi/2 - a)$,⁷ we have $\Theta(a) + \Theta(\pi/2 - a) \leq \Theta(\pi/2) = \pi/2$ for a.e. a , hence

$$\int_0^{\pi/2} (\Theta(a) - a) da \leq 0. \quad (17)$$

On the other hand, for $b \leq \pi/2$, the random variable $Z_b := \vartheta_{(\gamma,t)}^{-1}(b)$ takes values in $[0, \pi/2]$ (Lemma 4.33(ii)), where $\cos$ is strictly concave. Lemma 4.34 and Jensen's inequality give $\cos b = \int \cos Z_b d\mu \leq \cos(\int Z_b d\mu)$, i.e.,

$$\int \vartheta_{(\gamma,t)}^{-1}(b) d\mu \leq b, \quad \text{with equality if and only if } Z_b = b \text{ } \mu\text{-a.s.} \quad (18)$$

For each monotone map ϑ of $[0, \pi/2]$ fixing the endpoints, Young's identity holds (Lemma 4.33(iii)); integrating it against μ and using (18), we have

$$\int_0^{\pi/2} (\Theta(a) - a) da = \int_0^{\pi/2} \left(b - \int \vartheta_{(\gamma,t)}^{-1}(b) d\mu \right) db \geq 0.$$

With (17) both sides vanish, so the non-negative integrand is zero for a.e. b , and the equality case of (18) gives $\vartheta_{(\gamma,t)}^{-1}(b) = b$ for a.e. $b \in (0, \pi/2)$ and μ -a.e. (γ, t) . It remains to pass from $\vartheta_{(\gamma,t)}^{-1}(b) = b$ for a.e. $b \in (0, \pi/2)$ to $\vartheta_{(\gamma,t)}(\theta) = \theta$ for every θ . Fix such a (γ, t) , write $f = \vartheta_{(\gamma,t)}$, and let $B \subset (0, \pi/2)$ be the conull set on which $f^{-1}(b) = b$. Let $\theta \in (0, \pi/2)$. For $b \in B$ with $b > \theta$, the definition $f^{-1}(b) = \sup\{\theta'' : f(\theta'') < b\} = b > \theta$ provides some $\theta'' > \theta$ with $f(\theta'') < b$, hence $f(\theta) \leq f(\theta'') < b$ by monotonicity; letting $b \downarrow \theta$ within B gives $f(\theta) \leq \theta$. For $b \in B$

⁶By Fubini, the inequality of Lemma 4.32 at the pair $(a, \pi/2 - a)$ holds for μ -a.e. (γ, t) for a.e. such a .

⁷The base point of the second term is $R_a(\gamma, t)$ in the phase-space coordinates of Corollary 4.8, and μ is invariant by the rotation R_a .

with $b < \theta$, we have $f^{-1}(b) = b < \theta$, so θ does not belong to $\{f < b\}$, i.e., $f(\theta) \geq b$; letting $b \uparrow \theta$ within B gives $f(\theta) \geq \theta$. Thus $f(\theta) = \theta$ for every $\theta \in (0, \pi/2)$, and also at $\pi/2$; the flip symmetry $f(\pi - \theta) = \pi - f(\theta)$ then gives $f = \text{id}$ on $(0, \pi)$.

For the extension, let (γ, η) be admissible geodesics crossing at x , possibly a seam point; by Proposition 4.14 there are pairs (γ_n, η_n) of the conull set, crossing at flat points $x_n \rightarrow x$, with $\gamma_n \rightarrow \gamma$, $\eta_n \rightarrow \eta$. The images $h\gamma, h\eta$ are linked, hence meet, and the crossing points $q_n = h\gamma_n \cap h\eta_n$ stay in a bounded set (the crossing point of two linked geodesics lies within a distance of a base point controlled by the Gromov products of their endpoints, which converge). We first check that $h\gamma \cap h\eta$ is a single point. Otherwise, by uniqueness of CAT(0) segments, $h\gamma$ and $h\eta$ share a non-degenerate segment and branch at its endpoints, where both are tangent to the seam; at every point of this segment, including its endpoints, the two geodesics have the same or opposite directions (at a tangential contact, the two straight continuations have the same tangent vector, Lemma 3.48(iv)), so that the angle between them is 0 or π there. But any accumulation point q^* of (q_n) lies in $h\gamma \cap h\eta$, and the semicontinuity argument below, applied at q^* , gives $\angle_{q^*}(h\gamma, h\eta) = \angle_x(\gamma, \eta) \in (0, \pi)$, a contradiction. Hence $h\gamma \cap h\eta = \{q\}$, and $q_n \rightarrow q$ by Lemma 4.29, applied to $\rho_n = h\gamma_n$ and τ a fixed admissible transversal through q , and then to $h\eta_n$. At a transversal crossing the two adjacent angles sum to π (by Lemma 3.48(i)), and the Alexandrov angle is upper semicontinuous under such convergence; indeed, it is the infimum of the comparison angles, which are continuous in the configuration and non-decreasing in the parameters [11, II.3.1]. If α_n, β_n are the two adjacent angles at q_n and α, β those at q , then $\limsup \alpha_n \leq \alpha$, $\limsup \beta_n \leq \beta$ and $\alpha_n + \beta_n = \pi = \alpha + \beta$, so both sequences converge: $\alpha_n \rightarrow \alpha$, $\beta_n \rightarrow \beta$; the same holds at $x_n \rightarrow x$. Hence $\angle_q(h\gamma, h\eta) = \lim \angle_{q_n}(h\gamma_n, h\eta_n) = \lim \angle_{x_n}(\gamma_n, \eta_n) = \angle_x(\gamma, \eta)$. $\square$

Below, we also verify that h preserves incidence angles.

Proposition 4.36. *Let $\tilde{w}$ be a lifted wall, the axis of a peripheral element $p \in \pi_1(K)$, and let $\tilde{w}' = h\tilde{w}$ be the corresponding wall of X' , the axis of $\phi(p)$. For a.e. r and every $\phi_0 \in (0, \pi)$, the geodesic crossing $\tilde{w}$ at $\tilde{w}(r)$ with incidence angle ϕ_0 has image crossing $\tilde{w}'$ with incidence angle ϕ_0 ; by continuity, the same holds for every transversal crossing of $\tilde{w}$ by an admissible geodesic.*

Proof. Let $\vartheta_w(r, \phi)$ be the incidence angle of $h\eta$ with $\tilde{w}'$ at their unique crossing, for η crossing $\tilde{w}$ at $\tilde{w}(r)$ with incidence ϕ . The geodesics crossing $\tilde{w}$ from a fixed side modulo $\langle p \rangle$ carry the measure $\frac{1}{4} \sin \phi d\phi dr$ on $[0, \ell(p)) \times (0, \pi)$ (wall chart), h maps them to the geodesics crossing $\tilde{w}'$ from a fixed side modulo $\langle \phi(p) \rangle$, and $\ell(p) = \ell'(\phi(p))$ by **EL** = **EL'** (note that this also follows from the correspondence of the two families of crossing geodesics, by comparing total masses). Hence the identity of Lemma 4.34 holds for ϑ_w with μ replaced by $dr/\ell(p)$. The superadditivity argument of Lemma 4.32 applies verbatim with $h\gamma$ replaced by $\tilde{w}'$ (a geodesic of X'), and its second term is now the transfer of the pair (η_1, η_2) of *admissible* geodesics crossing at the wall point, which equals ϕ_2 by Theorem 4.35: $\vartheta_w(r, \phi_1 + \phi_2) \geq \vartheta_w(r, \phi_1) + \phi_2$. Thus, for a.e. r , the function $\phi \mapsto \vartheta_w(r, \phi) - \phi$ agrees a.e. with a non-decreasing function (the inequality holds for a.e. pair (ϕ_1, ϕ_2)) which vanishes at $\pi/2$, so $\vartheta_w(r, \cdot) \leq \text{id}$ on $(0, \pi/2]$, i.e., $\vartheta_w^{-1}(b) \geq b$. Together with the Jensen inequality $\int \vartheta_w^{-1}(b) dr/\ell(p) \leq b$ this forces equality, and the strict case gives $\vartheta_w^{-1}(b) = b$ a.e., hence $\vartheta_w = \text{id}$. $\square$

4.4. From angle preservation to isometry. In this part, we conclude that once h preserves angles, it is actually induced by an isometry, which completes the proof of Theorem A. If $\widetilde{K}_0$ were negatively curved, we could use the fact that the sum of the

angles on a geodesic triangle is always strictly smaller than π , except if the triangle degenerates into a single point. This is exactly how Otal originally argued in [39]. In our case, however, the space is only CAT(0), so in principle there's nothing to prevent the existence of triangles with sum of intern angles exactly equal to π . In fact, if h is “close to the identity”, a triple of geodesics intersecting at an interior point $x \in \widetilde{K}^\circ$ would be mapped into three geodesics of $\widetilde{K}'_0$ whose intersection determines a geodesic triangle $\triangle$ also inside $(\widetilde{K}')^\circ$. Since the negative curvature of the limit space $\widetilde{K}_0$ is concentrated at the boundary set $B(\widetilde{K})$, it would follow that $\triangle$ is a flat triangle, and the sum of its intern angles would be equal to π .

This indicates that, in order to use angle preservation to derive some sort of isometry, we must look at triangles on a neighbourhood of the boundary set. This is how we proceed to conclude the proof.

Theorem 4.37 (Enriched spectrum rigidity). *Suppose $\mathbf{EL} = \mathbf{EL}'$ (with respect to the marking ϕ fixed in §4.3). Then there is an isometry $\varphi: \mathbb{T}^2 \rightarrow \mathbb{T}^2$ such that $\mathcal{D}' = \varphi \circ \mathcal{D}$. In particular, the induced map*

$$\begin{aligned} \Phi: \Omega &\rightarrow \Omega' \\ (x, \omega) &\mapsto (\varphi(x), d\varphi(\omega)) \end{aligned}$$

is a flow conjugacy between billiard flows.

Throughout this subsection we assume $\mathbf{EL} = \mathbf{EL}'$, and $h: \mathcal{G}_0 \rightarrow \mathcal{G}'_0$ denotes the map of Definition 4.26. Let us recall what is known so far about h :

- h is a homeomorphism preserving linking and the currents;
- h maps admissible geodesics to admissible ones (Lemma 4.27);
- h maps the axis of every $g \in \pi_1(K)$ to the axis of $\phi(g)$, in particular each lifted wall $\tilde{w}$ to a lifted wall $\tilde{w}' := h\tilde{w}$ of $\mathcal{D}'$;
- h preserves the angles of admissible crossings (Theorem 4.35) and the incidence angles of admissible geodesics with the walls (Proposition 4.36).

Everything below is $\pi_1(K)$ -equivariant over ϕ , and we will not repeat it.

The proof of Theorem 4.37 consists in constructing a point map $\tilde{h}: X \rightarrow X'$ inducing h . We proceed in three steps: the images under h of the geodesics through a wall point are concurrent (Lemma 4.40), which defines $\tilde{h}$ on the seam, where it is isometric (Lemma 4.41); the current then controls the crossing positions along every regular geodesic, which gives concurrence at flat points (Lemma 4.42); finally $\tilde{h}$ is shown to be isometric on a dense set by a Crofton argument, and extended by continuity.

Definition 4.38 (Generic wall points). A wall point $x \in B$ is *generic* if the set of directions at x (in either adjacent sheet) whose geodesic is regular is conull in the circle of directions at x .

Lemma 4.39. *All but countably many points of each lifted wall are generic. In particular, generic points are dense in B .*

Proof. Let $\tilde{w}$ be a lifted wall. By Lemma 4.5(a), the parameters (r, ϕ) of a wall chart of $\tilde{w}$ whose geodesic has a tangential contact with B lie on countably many C^1 curves. A C^1 curve meets all but countably many vertical lines $\{r = r_0\}$ in a Lebesgue-null set (it can contain a vertical segment for at most countably many r_0 , and by Fubini a null set of the plane meets a.e. vertical line in a null set); hence, for all but countably many r_0 , the directions at $\tilde{w}(r_0)$ whose geodesic has a tangential contact form a null set of incidence angles, for each of the two charts. Since a geodesic through a

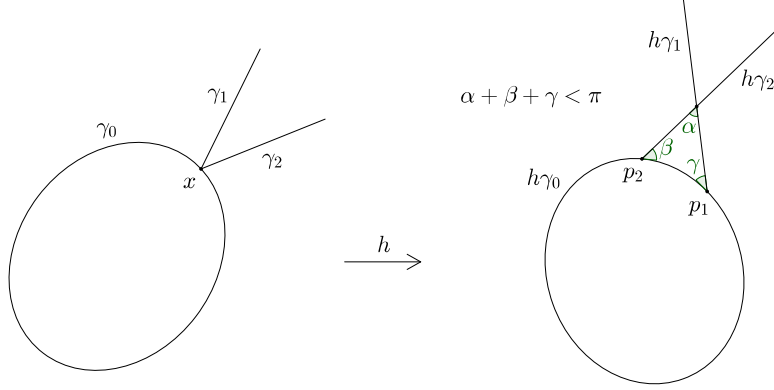


FIGURE 8. Three geodesics intersecting at a boundary point are mapped into a triangle of geodesics whose interior angles sum to π , which is impossible by the Flat Triangle Lemma since (p_1, p_2) is a wall arc.

wall point which crosses the wall transversally there and has no tangential contact is regular, $\tilde{w}(r_0)$ is generic. $\square$

We start the proof by showing that the images under h of the regular geodesics through a generic wall point are concurrent; in other words, h induces a map $\tilde{h}_0$ from the generic points of B to B' .

Lemma 4.40. *Let $x \in B$ be generic and $\tilde{w}$ the lifted wall through x . Then the images under h of all regular geodesics through x pass through a single point $\tilde{h}_0(x) \in \tilde{w}' = h\tilde{w}$.*

Proof. Let γ_1, γ_2 be regular geodesics through x ; they cross $\tilde{w}$ transversally at x , with incidence angles ϕ_1, ϕ_2 , and we label them so that γ_1 lies between $\tilde{w}$ and γ_2 , i.e., with $\psi := \angle_x(\gamma_1, \gamma_2)$, $\phi_1 + \psi + (\pi - \phi_2) = \pi$. The images $h\gamma_1, h\gamma_2$ are admissible, cross $\tilde{w}'$ at points p_1, p_2 and cross each other at a point q (they are linked, and by Theorem 4.35, applied to (γ_1, γ_2) crossing at x , they meet in a single point, at which their angle is ψ). Suppose that p_1, p_2, q are not all equal (see Figure 8). Then $p_1 p_2 q$ is a non-degenerate geodesic triangle of X' whose side $[p_1, p_2]$ is an arc of $\tilde{w}'$ (the unique geodesic between two points of a wall), and, by the proof of Lemma 4.32, its interior angles are ϕ_1 and $\pi - \phi_2$ at p_1, p_2 (Proposition 4.36) and ψ at q (Theorem 4.35); in particular, the sum of these angles is equal to π . Since each interior angle is bounded by the corresponding comparison angle and the comparison angles sum to π , all three are equal to their comparison angles, and by the Flat Triangle Lemma [11, II.2.9] the convex hull of $p_1 p_2 q$ is isometric to a Euclidean triangle, in which the side $[p_1, p_2]$ is straight. But $[p_1, p_2]$ is a wall arc of positive length, whose geodesic curvature with respect to the adjacent sheet is $-\mathcal{K} \neq 0$, whereas a side of a Euclidean triangle has zero geodesic curvature with respect to the triangle, and the geodesic curvature of a boundary arc of a flat region is intrinsic to the region (by Gauss–Bonnet), a contradiction. Hence $p_1 = p_2 = q$: the images of all regular geodesics through x meet $\tilde{w}'$ at one and the same point, which we call $\tilde{h}_0(x)$. $\square$

Next, we show that the map $\tilde{h}_0$ induces an isometry between the boundaries:

Lemma 4.41. *The map $\tilde{h}_0$ extends uniquely to an isometric bijection $\tilde{h}_0: (B, d_0) \rightarrow (B', d'_0)$, which maps each lifted wall $\tilde{w}$ onto $\tilde{w}' = h\tilde{w}$ and satisfies $\tilde{h}_0(x) \in h\gamma$ for every wall point x and every regular geodesic γ through x .*

Proof. Let us first assume that the points $x, y \in B$ are generic. By the Crofton formula (see Lemma 4.9), $d_0(x, y)$ is the $\hat{m}_0$ -mass of the geodesics η separating x from y . We claim that η separates x from y if and only if $h\eta$ separates $\tilde{h}_0(x)$ from $\tilde{h}_0(y)$, for every η through neither x nor y . Given such an η , let S be the side of η containing x . The directions at x whose geodesic is unlinked with η form an arc of length $\pi - \angle_x(\eta^-, \eta^+) > 0$, and regular directions are conull at x ; so there is a regular geodesic γ_x through x , unlinked with η , hence contained in $\bar{S}$. Then $h\gamma_x$ is unlinked with $h\eta$ and contained in the closure of the side S' of $h\eta$ corresponding to S (the sides correspond through $\partial\phi$), and $\tilde{h}_0(x) \in h\gamma_x \subset \bar{S}'$ by Lemma 4.40. The same holds for y , which proves the claim; since $h_*\hat{m}_0 = \hat{m}'_0$, we get $d'_0(\tilde{h}_0(x), \tilde{h}_0(y)) = d_0(x, y)$.

Now, since generic points are dense in B (Lemma 4.39), and $\tilde{h}_0$ is isometric on them, it extends uniquely to an isometric map $B \rightarrow B'$ which sends each lifted wall $\tilde{w}$ into $\tilde{w}'$. An isometric map from a geodesic line into a geodesic line is onto, so $\tilde{h}_0$ maps $\tilde{w}$ onto $\tilde{w}'$; since h maps the set of lifted walls (which are the axes of the peripheral elements) bijectively onto the set of lifted walls of X' , $\tilde{h}_0: B \rightarrow B'$ is bijective.

Let us now consider an arbitrary point $x \in B$ on the wall $\tilde{w}$, and let γ be regular through x . Choose generic points $x_n \rightarrow x$ on $\tilde{w}$, and regular geodesics γ_n through x_n converging to γ (since γ is regular, its collisions are transversal, so geodesics with nearby initial data converge to γ uniformly on compact sets, with converging endpoints, and the regular ones among them are dense). Then $\tilde{h}_0(x_n) = h\gamma_n \cap \tilde{w}' \rightarrow h\gamma \cap \tilde{w}'$ by Lemma 4.29 (with $\rho_n = h\gamma_n$ and $\tau = \tilde{w}'$), while $\tilde{h}_0(x_n) \rightarrow \tilde{h}_0(x)$ by continuity; hence $\tilde{h}_0(x) = h\gamma \cap \tilde{w}' \in h\gamma$. $\square$

Lemma 4.42. *Let γ be a regular geodesic, with wall points $x_i = \gamma(t_i)$, $i \in \mathbb{Z}$, $t_i < t_{i+1}$. Parametrise $h\gamma$ by arc length so that $(h\gamma)(t_1) = \tilde{h}_0(x_1)$ and $(h\gamma)(t_2) = \tilde{h}_0(x_2)$. Then:*

- (i) $(h\gamma)(t_i) = \tilde{h}_0(x_i)$ for every i ;
- (ii) for $\hat{m}_0$ -a.e. geodesic η crossing γ , at $\gamma(t)$ say, the image $h\eta$ crosses $h\gamma$ at $(h\gamma)(t)$;
- (iii) for every $t \in \mathbb{R}$ and every regular geodesic η through $\gamma(t)$, $h\eta$ passes through $(h\gamma)(t)$.

Proof. (i) By Lemma 4.41, the points $\tilde{h}_0(x_i)$ lie on $h\gamma$, and are at mutual distances $d'_0(\tilde{h}_0(x_i), \tilde{h}_0(x_j)) = d_0(x_i, x_j) = |t_i - t_j|$. A point of the line $h\gamma$ is determined by its distances to the two distinct points $(h\gamma)(t_1)$, $(h\gamma)(t_2)$; since $(h\gamma)(t_i)$ has the right distances, $\tilde{h}_0(x_i) = (h\gamma)(t_i)$.

(ii) It suffices to prove the statement on each chord $[x_1, x_2]$ of γ , i.e., for crossings at times $t \in (t_1, t_2)$, the crossers at the countably many wall points being null. For $t \in (t_1, t_2)$ and $\theta \in (0, \pi)$ write $\eta_{t,\theta} := \eta_{\gamma,t,\theta}$ for the geodesic crossing γ at $\gamma(t)$ from the right with angle θ (Lemma 4.30), and let $\sigma(\theta, t)$ be the position on $h\gamma$ of a crossing point of $h\gamma$ and $h\eta_{t,\theta}$, unique when $h\eta_{t,\theta}$ is clean. By Lemma 4.5(a) and Fubini in the flat chart along γ , there is a conull set $\Theta_0 \subset (0, \pi)$ such that, for $\theta \in \Theta_0$, the set T_θ of those $t \in (t_1, t_2)$ for which $\eta_{t,\theta}$ and $h\eta_{t,\theta}$ are clean is conull in (t_1, t_2) . We claim that $\sigma(\theta, t) = t$ for every $\theta \in \Theta_0$ and $t \in T_\theta$; by Lemma 4.30(i) this is the statement (ii) for the crossers from the right, and the crossers from the left are the reverses of the crossers from the right of $\bar{\gamma}$.

Let us now fix $\theta \in \Theta_0$ and $t < t'$ in T_θ . The geodesics $\eta_{t,\theta}$, $\eta_{t',\theta}$ are unlinked; indeed, if they were linked, they would cross at a point z , with a positive angle since both are clean, and the geodesic triangle with vertices $\gamma(t), \gamma(t'), z$ would have interior angles θ and $\pi - \theta$ at $\gamma(t), \gamma(t')$ (in one order or the other, according to the side of z) and a positive angle at z , contradicting the angle-sum inequality in the CAT(0) space X . Now, for three embedded arcs of the closed disc $\bar{X}$ with pairwise distinct endpoints, two of which cross the third exactly once and are unlinked with each other, the order in which the third arc meets the other two is determined by the cyclic order of the six endpoints on $\partial_\infty X$ (the two unlinked arcs are nested, and the one with the inner pair of endpoints is met first or last according to which end of the third arc its endpoints surround). Since h preserves linking, and $\partial\phi$ preserves or reverses all cyclic orders, the order of the crossing points of $h\eta_{t,\theta}$, $h\eta_{t',\theta}$ along $h\gamma$ is that of $\gamma(t), \gamma(t')$ along γ , or globally the opposite; with the orientation of $h\gamma$ fixed by (i), $t \mapsto \sigma(\theta, t)$ is non-decreasing on T_θ . Let F_θ be its non-decreasing extension to (t_1, t_2) , i.e., $F_\theta(t) := \sup\{\sigma(\theta, t') : t' \in T_\theta, t' \leq t\}$.

As $t \rightarrow t_2$ within T_θ , the crossers $\eta_{t,\theta}$ converge, with converging endpoints, to the crosser of γ at x_2 with angle θ , which for a.e. θ is admissible and crosses the wall at x_2 transversally; its image passes through $\hat{h}_0(x_2) = (h\gamma)(t_2)$ (Lemma 4.41 and (i)) and crosses $h\gamma$ only there, so Lemma 4.29 gives $F_\theta(t_2^-) = t_2$; likewise $F_\theta(t_1^+) = t_1$. Shrinking Θ_0 to a conull set, we assume this for all $\theta \in \Theta_0$.

For $t_1 < a < b < t_2$ and $0 < \theta_1 < \theta_2 < \pi$, the geodesics crossing $\gamma|_{[a,b]}$ from the right with angle in (θ_1, θ_2) have mass $(b-a) \int_{\theta_1}^{\theta_2} \frac{1}{4} \sin \theta d\theta$ by (12); their images are, by Theorem 4.35, geodesics crossing $h\gamma$ with angle in (θ_1, θ_2) , at the positions $\sigma(\theta, t)$. The images of the non-clean crossers form an $\hat{m}'_0$ -null set, since $h_*\hat{m}_0 = \hat{m}'_0$ and non-clean geodesics are $\hat{m}_0$ -null; hence, by (12) along $h\gamma$, the mass of the image set is $\int_{\theta_1}^{\theta_2} \frac{1}{4} \sin \theta \text{Leb}(\sigma(\theta, [a,b] \cap T_\theta)) d\theta$. Comparing the two masses for all rational $a < b$ and $\theta_1 < \theta_2$, we get $\text{Leb}(\sigma(\theta, [a,b] \cap T_\theta)) = b - a$ for all rational $a < b$ and all θ in a conull subset of Θ_0 , which we still call Θ_0 .

Fix $\theta \in \Theta_0$. For a non-decreasing function F on an interval and $a < b$ in it, $F([a,b]) \subset [F(a), F(b)]$, so $\text{Leb}(F([a,b])) \leq F(b) - F(a)$, with equality if and only if F is continuous on $[a,b]$ (a jump at a point of $[a,b]$ leaves a gap of positive length in $F([a,b])$); moreover $\sigma(\theta, [a,b] \cap T_\theta) \subset F_\theta([a,b])$. Summing over a partition $t_1 < a_0 < \dots < a_N < t_2$ by points in $\mathbb{Q}$ and letting $a_0 \rightarrow t_1$, $a_N \rightarrow t_2$, we get

$$\begin{aligned} t_2 - t_1 &= \lim \sum_i \text{Leb}(\sigma(\theta, [a_i, a_{i+1}] \cap T_\theta)) \leq \lim \sum_i (F_\theta(a_{i+1}) - F_\theta(a_i)) \\ &= F_\theta(t_2^-) - F_\theta(t_1^+) = t_2 - t_1. \end{aligned}$$

Hence the inequality is an equality for every partition, so that $\text{Leb}(F_\theta([a,b])) = F_\theta(b) - F_\theta(a) = b - a$ for all rational $t_1 < a < b < t_2$: F_θ is continuous, and $F_\theta(b) - F_\theta(a) = b - a$ for all rational $a < b$. Together with $F_\theta(t_1^+) = t_1$ this gives $F_\theta(t) = t$, i.e., $\sigma(\theta, t) = t$ for every $t \in T_\theta$, which proves the claim and (ii).

(iii) Let $t_0 \in (t_1, t_2)$ and let η be a regular geodesic through $\gamma(t_0)$, crossing γ there from the right (otherwise consider $\bar{\eta}$) with angle θ_0 . Choose $\theta_n \in \Theta_0$ and $t_n \in T_{\theta_n}$ with $(t_n, \theta_n) \rightarrow (t_0, \theta_0)$. Since η is regular, every collision of η is transversal, so geodesics with nearby initial data converge to η uniformly on compact sets, with converging endpoints: $\eta_{t_n, \theta_n} \rightarrow \eta$. By Lemma 4.29 and (ii), $h\gamma \cap h\eta = \lim h\gamma \cap h\eta_{t_n, \theta_n} = \lim (h\gamma)(t_n) = (h\gamma)(t_0)$. At a wall point $\gamma(t_i)$, (iii) is Lemma 4.41 together with (i). $\square$

We are now ready to conclude the proof of Theorem 4.37.

Proof of Theorem 4.37. We start by defining $\tilde{h}$ on a dense set. Let $P \subset X$ be the set of points lying on at least one regular geodesic. For $x \in P$, choose a regular $\gamma \ni x$, write $x = \gamma(t)$, and set $\tilde{h}(x) := (h\gamma)(t)$, with the parametrisation of Lemma 4.42. By Lemma 4.42(iii), $\tilde{h}(x)$ is the common point of the images of all regular geodesics through x , a description independent of the choice of γ ; hence $\tilde{h}$ is well defined on P , satisfies $\tilde{h}(\gamma) = h\gamma$ for every regular γ (its restriction to γ being the parametrisation of Lemma 4.42), and agrees with $\tilde{h}_0$ on $P \cap B$ by Lemma 4.42(i). The set P is dense in X : non-regular geodesics are $\hat{m}_0$ -null (Lemma 4.7), so the phase points lying on them form a $\hat{m}_0 \otimes dt$ -null, hence (Corollary 4.8) Lebesgue-null, subset of $T^1\tilde{K}^\circ$, and P contains the projection of its complement, which has full measure in X .

Let us now show that $\tilde{h}$ is isometric on P . Let $x, y \in P$, on regular geodesics γ_x, γ_y . By the Crofton formula, $d_0(x, y)$ is the $\hat{m}_0$ -mass of the geodesics separating x from y , and $d'_0(\tilde{h}(x), \tilde{h}(y))$ that of the geodesics separating $\tilde{h}(x)$ from $\tilde{h}(y)$; since $h_*\hat{m}_0 = \hat{m}'_0$, it suffices to show that, for $\hat{m}_0$ -a.e. η , η separates x from y if and only if $h\eta$ separates $\tilde{h}(x)$ from $\tilde{h}(y)$. Let η be a geodesic through neither x nor y , not sharing a sub-arc with γ_x or γ_y (which is a conull condition by Lemma 4.7(i)). If η is unlinked with γ_x , then x lies on the side of η containing γ_x , and $\tilde{h}(x) \in h\gamma_x$ lies on the corresponding side of $h\eta$, the sides corresponding through $\partial\phi$. If η crosses γ_x , say at $\gamma_x(t)$, then for $\hat{m}_0$ -a.e. such η the image $h\eta$ crosses $h\gamma_x = \tilde{h}(\gamma_x)$ at $\tilde{h}(\gamma_x(t))$ (Lemma 4.42(ii)), and the side of $h\eta$ containing $\tilde{h}(x)$ is determined by the position of $\tilde{h}(x)$ on $h\gamma_x$ relative to this crossing point, which is that of x relative to $\gamma_x(t)$, the parametrisations of γ_x and $h\gamma_x$ corresponding; so again x and $\tilde{h}(x)$ lie on corresponding sides. The same holds for y , and the claim follows.

Since $\tilde{h}$ is isometric on the dense set P , it extends uniquely to an isometric map $\tilde{h}: (X, d_0) \rightarrow (X', d'_0)$, which is $\pi_1(K)$ -equivariant over ϕ . It is also surjective, since it is an injective continuous map between the topological planes X and X' , hence open by invariance of domain, so that its image is open and closed in X' . By continuity, $\tilde{h} = \tilde{h}_0$ on B (the two maps agree on the generic points, which lie on regular geodesics and are dense in B), so $\tilde{h}$ maps lifted walls onto lifted walls and B onto B' . Thus $\tilde{h}$ descends to an isometry $K_0 \rightarrow K'_0$ mapping $B(K)$ onto $B(K')$, and Proposition 3.29 provides a Euclidean isometry φ of $\mathbb{T}^2$ with $\varphi(\mathcal{D}) = \mathcal{D}'$. The last assertion of the theorem is immediate, the billiard flow being determined by the table. $\square$

The proof of Theorem A is complete.

Remark 4.43. The proof above can be read in two more geometric ways, which also explain where the difficulties lie.

Let us fix a wall point x and the sheet S on one side of it. For each direction at x into S , the geodesic leaves along a straight chord until its next wall point y ; let V_x be the union of these chords, the “visibility star” of x in S , which is a simply connected flat region. By Lemma 4.41 and Lemma 4.42(i), the image of such a chord is the chord of $h\gamma$ between $\tilde{h}_0(x)$ and $\tilde{h}_0(y)$, of the same length; by Proposition 4.36 and Theorem 4.35, the image chords make the same incidence angle with the wall and the same mutual angles. Consequently there is a Euclidean motion I_x of the developed sheet sending V_x onto the visibility star $V'_{\tilde{h}_0(x)}$ and agreeing with $\tilde{h}_0$ on $V_x \cap B$; in particular, the wall arcs visible from x and from $\tilde{h}_0(x)$ have the same shape.

Similarly, along one regular geodesic γ with collision points $x_i = \gamma(t_i)$, the image $h\gamma$ meets the image walls at $\tilde{h}_0(x_i)$, at the same mutual distances and with the same incidence angles; since $\partial\phi$ preserves or reverses all cyclic orders, the oriented reflection

angles coincide as well. Hence the billiard trajectories of $\mathcal{D}$ and $\mathcal{D}'$ corresponding to γ and $h\gamma$ have the same flight times and the same reflection angles, and their unfoldings to the plane $\mathbb{R}^2$ (the $\mathbb{Z}^2$ -cover of the torus) are polygonal paths with the same chord lengths and turning angles: they differ by a Euclidean motion M_γ (an anti-motion if $\partial\phi$ reverses the orientation), which maps the collision points of the first onto those of the second, together with the tangent lines of the obstacles there. Every regular trajectory of $\mathcal{D}$ thus “samples” the boundary of $\mathcal{D}'$, with the same local geometry.

In both descriptions, the geometry between collisions is flat and plays no role; what has to be shown is that the motions $I_x, I_{x'}$ (or $M_\gamma, M_{\gamma'}$) attached to different stars (or trajectories) fit together. This is not automatic, and it is hard to derive it from the data along a single trajectory, even a dense one. Indeed, the data of γ (flight times, reflection angles) is invariant under the deck transformations $g_n \in \pi_1(K)$ lifting the lattice translations, and ϕ acts trivially on $\mathbb{Z}^2$: so the unfolding of $g_n\gamma$ is the translate $\Gamma + n$ of the unfolding Γ of γ , and that of $h(g_n\gamma) = \phi(g_n)h\gamma$ is $\Gamma' + n$, whence $M_{g_n\gamma} = T_{n-L(n)} \circ M_\gamma$, where L is the linear part of M_γ and T_v the translation by v . The density of the projection of γ in $T^1\mathcal{D}$ only says that the collision points x_i are dense in $\partial\mathcal{D}$ modulo $\mathbb{Z}^2$, i.e., that they sample all the lifted obstacles $\tilde{\mathcal{O}} + n$ for infinitely many n ; but the sample taken on $\tilde{\mathcal{O}} + n$ is transported by M_γ to a point of $M_\gamma(\tilde{\mathcal{O}}) + L(n)$, and unless $L(n) \in \mathbb{Z}^2$ these samples land on different translates of $\partial\tilde{\mathcal{O}}'$ and say nothing about the shape of a single obstacle. The missing information, $L(n) = n$, i.e., $M_{g_n\gamma} = M_\gamma$, is exactly the statement that the motions attached to γ and to $g_n\gamma$ agree at a common point of these two geodesics: this is the concurrence of Lemma 4.42(iii). Once concurrence is known, the Crofton argument of the proof produces the isometry on a dense set at once; in particular, no dense orbit is used, and ergodicity of the billiard flow plays no role.

APPENDIX A. GENERIC PROPERTIES OF PERIODIC ORBITS OF SINAI BILLIARDS

A.1. Length functionals and periodic orbits. We use the notation introduced in Section 2. We let $\gamma_{(0,0)} := \pi^{-1} \circ \gamma$, and for each integer vector $(i, j) \in \mathbb{Z}^2$, we let $\gamma_{(i,j)} := \gamma_{(0,0)} + (i, j)$. Denote by

$$\begin{aligned} \tau_{(i,j)} : \mathbb{T}_{\mathcal{D}} \times \mathbb{T}_{\mathcal{D}} &\rightarrow \mathbb{R}_+ \\ (s, s') &\mapsto \tau_{(i,j)}(s, s') := \|\gamma_{(0,0)}(s) - \gamma_{(i,j)}(s')\| \end{aligned}$$

the Euclidean distance between the points of $\partial\tilde{\mathcal{D}}$ with respective parameters s, s' in the respective cells $(0, 0)$ and (i, j) . Recall that for any $(s, s') \in \mathbb{T}_{\mathcal{D}} \times \mathbb{T}_{\mathcal{D}}$, we have

$$\partial_1 \tau_{(i,j)}(s, s') = -\sin \phi_{(i,j)}, \quad \partial_2 \tau_{(i,j)}(s, s') = \sin \phi'_{(i,j)}, \quad (19)$$

where $\phi_{(i,j)}$ and $\phi'_{(i,j)}$ are the respective angles that the line segment between $\gamma_{(0,0)}(s)$ and $\gamma_{(i,j)}(s')$ makes with respect to the normal vector at these points. Moreover,

$$\begin{aligned} \partial_{11} \tau_{(i,j)}(s, s') &= \mathcal{K}(s) \cos \phi_{(i,j)} + \frac{\cos^2 \phi_{(i,j)}}{\tau_{(i,j)}(s, s')}, \\ \partial_{12} \tau_{(i,j)}(s, s') &= -\frac{\cos \phi_{(i,j)} \cos \phi'_{(i,j)}}{\tau_{(i,j)}(s, s')}, \\ \partial_{22} \tau_{(i,j)}(s, s') &= \mathcal{K}(s') \cos \phi'_{(i,j)} + \frac{\cos^2 \phi'_{(i,j)}}{\tau_{(i,j)}(s, s')}, \end{aligned} \quad (20)$$

where $\mathcal{K}(s), \mathcal{K}(s')$ are the respective curvatures at the points with parameters s and s' . Let $\sigma = (\bar{s}_0, \dots, \bar{s}_{q-1})$ denote the collision points of a periodic orbit of (least) period $q \geq 2$, with $\bar{s}_k \in \mathbb{T}_{\rho_k}$, $\rho_k \in \{1, \dots, m\}$, for $k \in \{0, \dots, q-1\}$. We let $\rho(\sigma) :=$

$(\rho_k)_{k=0,\dots,q-1}$. Let us consider the lift to $\tilde{\mathcal{D}}$ of this orbit starting in the cell $(0,0)$, and denote by $\{(\tilde{i}_k, \tilde{j}_k)\}_{k=0,\dots,q}$ the labels of the cells where the first $(q+1)$ bounces happen, with $(\tilde{i}_0, \tilde{j}_0) = (0,0)$. For $k \in \{1, \dots, q\}$, set $I_k := (\tilde{i}_k, \tilde{j}_k) - (\tilde{i}_{k-1}, \tilde{j}_{k-1})$. We denote $I(\sigma) := (I_k)_{k=1,\dots,q}$. As we assume that the table has finite horizon, there is a constant $M > 0$ such that $I_k \in [-M, M]^2$, for any $k \in \{1, \dots, q\}$.

The following result is essentially [20, Lemma A.2]; since the context is slightly different, we reproduce its proof for the reader's convenience.

Lemma A.1. *Let $\sigma = (\bar{s}_0, \dots, \bar{s}_{q-1})$ denote the collision points of a periodic orbit of (least) period $q \geq 2$. With the above notations, let $\rho := \rho(\sigma) = (\rho_k)_{k=0,\dots,q-1}$, and $I = I(\sigma) := (I_k)_{k=1,\dots,q}$. We define the length functional*

$$L_\rho^I: \begin{cases} \mathbb{T}_{\rho_0} \times \dots \times \mathbb{T}_{\rho_{q-1}} & \rightarrow \mathbb{R} \\ (s_0, \dots, s_{q-1}) & \mapsto \sum_{k=1}^{q-1} \tau_{I_k}(s_{k-1}, s_k) + \tau_{I_q}(s_{q-1}, s_0) \end{cases}.$$

Then σ is a critical point of L_ρ^I , and it holds:

$$L_\rho^I(s_0, \dots, s_{q-1}) - L_\rho^I(\bar{s}_0, \dots, \bar{s}_q) = Q_\rho^I(s_0 - \bar{s}_0, \dots, s_q - \bar{s}_q) + R_\rho^I,$$

where Q_ρ^I is a positive definite quadratic form and R_ρ^I is a remainder term that satisfies the estimate:

$$R_\rho^I = O(\|(s_0 - \bar{s}_0, \dots, s_q - \bar{s}_q)\|^3).$$

Proof. Let us abbreviate $L = L_\rho^I$. The fact that $\sigma = (\bar{s}_0, \dots, \bar{s}_{q-1})$ is a critical point for L among configurations $(s_0, \dots, s_{q-1})$ follows from (19), as σ is a closed billiard orbit. Hence the lemma follows from Taylor's formula, provided that we show that the Hessian of L at $(\bar{s}_0, \dots, \bar{s}_q)$ is positive definite. It is immediate from the definition of L that $\partial_{ij}L = 0$ if $|i - j| > 1 \pmod{q}$; the Hessian is thus a *cyclic* tridiagonal symmetric matrix, with corner entries $\partial_{0,q-1}L = \partial_{q-1,0}L \neq 0$ coming from the term $\tau_{I_q}(s_{q-1}, s_0)$. For notational convenience, let $\tau_k = \tau_{I_{k+1}}(\bar{s}_k, \bar{s}_{k+1})$; let $\bar{\phi}_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ denote the angle formed by the trajectory at the k -th collision point and the unit normal at $\bar{s}_k$; and let $\mathcal{K}_k > 0$ denote the curvature at $\bar{s}_k$. Using (20), a direct computation gives, for $u = (u_0, \dots, u_{q-1})$ (indices mod q),

$$D^2L(\bar{s})(u, u) = \sum_{k=0}^{q-1} \frac{1}{\tau_k} (\cos \bar{\phi}_k u_k - \cos \bar{\phi}_{k+1} u_{k+1})^2 + 2 \sum_{k=0}^{q-1} \mathcal{K}_k \cos \bar{\phi}_k u_k^2.$$

Since $\mathcal{K}_k > 0$ and $\cos \bar{\phi}_k > 0$ at a non-grazing orbit, both sums are non-negative and the second is positive unless $u = 0$; hence $D^2L(\bar{s})$ is positive definite. $\square$

Definition A.2. The spectrum of the table $\mathcal{D}$ is the metric space $(\text{Spec}(\mathcal{D}), d_H)$, where $\text{Spec}(\mathcal{D})$ is the set of all periodic orbits of the billiard flow of $\mathcal{D}$, and d_H is the Hausdorff metric on the phase space Ω .

It follows from Corollary 2.4 that $\text{Spec}(\mathcal{D})$ is a discrete totally bounded space.

Remark that a critical point of some length functional L as in Lemma A.1 is not necessarily a periodic orbit, as it may consist of a non-physical trajectory, that is, one that passes through one of the obstacles.

Let us introduce the following sets of orbits.

Definition A.3. Given a word $\omega = (\omega_0, \dots, \omega_n) \in \mathcal{A}^{n+1}$, the cylinder $C(\omega; n)$ is the set of points $x \in \mathcal{M}$ whose lifted orbits have the first $n+1$ bounces in the scatterers $B_{\omega_0}, \dots, B_{\omega_n}$, in this order.

Definition A.4. A *generalised periodic orbit* of $\mathcal{D}$ is the projection onto $\mathcal{M}$ of a critical point of one of the functionals L . A generalised orbit that is not a periodic orbit of the billiard $\mathcal{D}$ is called a *ghost* or *non-physical*.

Theorem A.5 (Theorem 2.2 in [12]). *Let $\omega = (\omega_0, \dots, \omega_n) \in \mathcal{A}^{n+1}$ be a word in $\Sigma_{\mathcal{M}}$ such that*

$$\begin{aligned} \pi(B_{\omega_i}) &\neq \pi(B_{\omega_{i+1}}), \text{ for } i = 0, \dots, n-1; \\ \pi(B_{\omega_n}) &= \pi(B_{\omega_0}). \end{aligned}$$

Then there is at most one generalised periodic orbit in the cylinder $C(\omega; n)$.

On the other hand, one periodic orbit in $\text{Spec}(\mathcal{D})$ is associated to several different words ω , corresponding to distinct liftings. Even when we restrict ourselves to liftings starting at the $(0, 0)$ cell, distinct starting points in the same orbit may lead to distinct words. We will therefore need a choice of point in each of the periodic orbits of $\text{Spec}(\mathcal{D})$.

Given $O \in \text{Spec}(\mathcal{D})$, we can see it as a union of line segments in the square $[0, 1)^2$. The finite set $(\cup_i B_i) \cap O$ corresponds to all the collision points of the orbit O . It is an ordered set, with the lexicographical ordering inherited from $[0, 1)^2 \times [-\pi/2, \pi/2]$. Set

$$\zeta(O) := \min\{O \cap (\cup_i B_i)\} \in \mathcal{M}.$$

Definition A.6. The *type* of the periodic orbit $O \in \text{Spec}(\mathcal{D})$ is the unique sequence $[O] := \iota(\zeta(O)) \in \Sigma_{\mathcal{M}}$.

A.2. Good tables. Recall that we denote by $\mathfrak{D}^m$ the space of all Sinai billiard tables in $\mathbb{T}^2$ with finite horizon, with m scatterers and C^r boundary, $r \geq 3$, identified with an open subset of $C_{\text{emb}}^r(\mathbb{T}_{\mathcal{D}}, \mathbb{T}^2)$ as in Section 2; recall also that a collision is *regular* if it is not grazing, and that a periodic orbit is *singular* if it has a grazing collision, and *regular* otherwise.

We want to show that for a generic choice of table in $\mathfrak{D}^m$, the set of periodic orbits has simple length spectrum and contains no singular orbits.

A.2.1. Table perturbations and periodic orbits. We begin by analysing how perturbations on the scatterers affect periodic orbits.

Definition A.7. Given $p = \gamma(s_0) \in \partial\mathcal{D}$, we say a table $\mathcal{D}' \in \mathfrak{D}^m$ is a *localised perturbation* of $\mathcal{D}$ at p if there is a neighbourhood U of $s_0 \in \mathbb{T}_{\mathcal{D}}$ such that the defining maps of $\mathcal{D}$ and $\mathcal{D}'$ coincide outside U .

Intuitively, a small perturbation around a single collision point of a periodic orbit should generate a new periodic orbit close to the original one. In order to see that this is in fact what happens, we investigate how a localised perturbation at a point $p = \gamma(s_0)$ alters the length functional of the table.

We begin by describing the perturbation in terms of a C^r function λ defined on $\mathbb{T}_{\mathcal{D}}$. For $\varepsilon > 0$, consider the ε -perturbation

$$\gamma^\lambda(s) = \gamma(s) + \varepsilon \lambda(s) n_s, \tag{21}$$

where n_s is the normal vector to $\gamma(s)$ at s . For sufficiently small ε , the perturbation γ^λ remains strictly convex and the resulting billiard table is still a Sinai billiard. The perturbation γ^λ is a localised perturbation at $p = \gamma(s_0)$ exactly when λ is supported in a small neighbourhood around s_0 .

As before, we look at liftings of such curves to the Euclidean plane and consider mappings

$$\tau_{(i,j)}^\lambda(s, s') := \|\gamma_{(0,0)}^\lambda(s) - \gamma_{(i,j)}^\lambda(s')\|,$$

whose first order approximation is

$$\tau_{(i,j)}^\lambda(s, s') = \tau_{(i,j)}(s, s') + \varepsilon(\lambda(s) \cos \phi_{(i,j)}(s, s') + \lambda(s') \cos \phi'_{(i,j)}(s, s')) + O(\varepsilon^2).$$

Now, we suppose $\sigma = (s_0, \dots, s_{q-1})$ corresponds to a sequence of bouncing points of the table $\mathcal{D}$. We want to study the relation between critical points for the length functionals of the original and perturbed tables. For notational convenience, we drop the indices (i, j) . Then the length functional L^λ can be written in terms of the original length functional L as

$$\begin{aligned} L^\lambda(\sigma) &= \sum_{i=0}^{q-1} \tau^\lambda(s_i, s_{i+1}) \\ &= \sum_{i=0}^{q-1} h(s_i, s_{i+1}) + \varepsilon \sum_{i=0}^{q-1} (\lambda(s_i) \cos \phi(s_i, s_{i+1}) + \lambda(s_{i+1}) \cos \phi'(s_i, s_{i+1})) + O(\varepsilon^2), \end{aligned}$$

or, more concisely, as

$$L^\lambda(\sigma) = L(\sigma) + \varepsilon P^\lambda(\sigma) + O(\varepsilon^2). \quad (22)$$

If σ is a periodic orbit of the unperturbed table $\mathcal{D}$, then consecutive angles ϕ and ϕ' have the same cosine, and hence P^λ becomes

$$P^\lambda(\sigma) = 2 \sum_{i=0}^{q-1} \lambda(s_i) \cos \phi(s_i, s_{i+1}), \quad (23)$$

with the usual convention that $s_q = s_0$. Writing $\sigma^\lambda := (s_0^\lambda, \dots, s_{q-1}^\lambda)$ to denote the generalised periodic orbit of the perturbed table with same type as σ , we have

$$s_i^\lambda = s_i + \varepsilon \psi_i + O(\varepsilon^2).$$

Thus, $\sigma^\lambda = \sigma + \varepsilon \psi$ satisfies

$$\begin{aligned} 0 &= DL^\lambda(\sigma^\lambda) \\ &= DL(\sigma^\lambda) + \varepsilon DP^\lambda(\sigma^\lambda) + O(\varepsilon^2) \\ &= DL(\sigma) + \varepsilon D^2L(\sigma)\psi + \varepsilon DP^\lambda(\sigma) + O(\varepsilon^2) \\ &= \varepsilon D^2L(\sigma)\psi + \varepsilon DP^\lambda(\sigma) + O(\varepsilon^2), \end{aligned} \quad (24)$$

where the last equality comes from the fact that σ is a critical point for L . Thus, the first order change ψ we see in the collision points for a critical point σ^λ is the solution of the linear system

$$D^2L(\sigma)\psi = -DP^\lambda(\sigma). \quad (25)$$

We already know that the Hessian of L is a non-degenerate positive definite matrix, so for a non-positive solution ψ to (25) to exist it is sufficient that the term $-DP^\lambda(\sigma)$ is non-zero. We conclude that, unless σ is a critical point of P^λ , the perturbed flow will have a new generalised periodic orbit arbitrarily close to the original one. In particular, for a generic choice of perturbations, such perturbed orbits always exist. Naturally, if σ is a critical point of P^λ , then $\psi = 0$ is the unique solution of (25) and σ remains a generalised periodic orbit after the perturbation.

Definition A.8. Given a point $p = \gamma(s_0)$ in the obstacle $\Gamma_i \subset \partial\mathcal{D}$, we say a localised perturbation is a *retraction of $\mathcal{O}_i$ at p* if it is of the form

$$\tilde{\gamma}(s) = \gamma(s) - \varepsilon \Psi(s) n_s,$$

where $\varepsilon > 0$ and Ψ is a positive bump function supported at a small interval around s_0 , with $\Psi(s_0) = 1$.

Remark A.9. If a periodic orbit σ has a grazing collision at p , then any retraction of the obstacle tangent to the orbit at p , no matter how small, removes the tangency without changing the original orbit, i.e. the perturbed orbit σ^λ has the same trace, but without the grazing collision at p .

On the other hand, a generic perturbation of σ at a regular collision point will generate a non-singular generalised orbit. Indeed, the law of reflection means that a change in the angle of one of the collisions will propagate through all the orbit, so that, generically, a change of the angle and position of one the vertices will either kill the periodic orbit σ by pushing it into an obstacle (i.e., the perturbed orbit σ^λ is a ghost orbit) or it will move the orbit away from the obstacles it previously touched tangentially (that is, σ^λ is regular).

A.2.2. Killing off singular orbits. As suggested in Remark A.9, localised perturbations can “destroy” singular periodic orbits by turning them into regular or ghost ones. We claim this procedure can be applied, in a controlled way, to remove any finite number of singular orbits without producing new ones.

Lemma A.10. *If $\mathcal{D}$ has finite horizon, so does every sufficiently close localised perturbation of $\mathcal{D}$.*

Proof. Let p be a collision point in $\partial\mathcal{D}$, and let L be the image of the orbit of p in the torus. If p is regular, then, for sufficiently small perturbations localised at p , the corresponding perturbed orbit L' is still regular and bounces at the perturbed obstacle $\partial\mathcal{D}'_i$. Thus, if a perturbation around p were to create a periodic orbit with no collisions at all, then the collision at p must be grazing. It must also be the only collision on the orbit L ; otherwise, a localised perturbation would not eliminate all collisions. Hence, the line L is closed, tangent to the obstacle Γ_i at p , and a positive distance away from all the other curves in $\partial\mathcal{D}$. In particular, this means the obstacle Γ_i lies entirely on one side of L . Now, on the side of L opposite the obstacle, any line parallel to L and sufficiently close to it would not meet any of the obstacles Γ_j , which contradicts our finite horizon hypothesis. It follows that localised perturbations of finite horizon tables cannot create tables of unbounded horizon, as we wanted. $\square$

Given $n \in \mathbb{N}$, there exists a positive constant M , depending on n and on the table $\mathcal{D}$ (more specifically, on the maximum free-flight time $\tau_{\max}$) with the following property: for every periodic point $p \in \mathcal{M}$ of order less or equal to n , the fundamental trace of its lift to $\tilde{\mathcal{D}}$ is completely contained in $[-M, M]^2$.

We consider the set of all possible symbolic sequences of period n orbits:

$$\Sigma_n := \left\{ (a_0, \dots, a_{n-1}) \left| \begin{array}{l} a_i \in ([-M, M]^2 \cap \mathbb{Z}^2) \times \{1, \dots, k\} \subset \mathcal{A}; \\ \pi(B_{a_i}) \neq \pi(B_{a_{i+1}}) \text{ for } i = 0, \dots, n-2; \\ \pi(B_{a_{n-1}}) = \pi(B_{a_0}). \end{array} \right. \right\}.$$

Now, according to Theorem A.5, to each element of the finite set Σ_n corresponds at most one periodic orbit, which can be non-physical. Whether or not a sequence has an associated periodic trajectory depends on the length functional of the table. The important point here is that localised perturbations may change the length

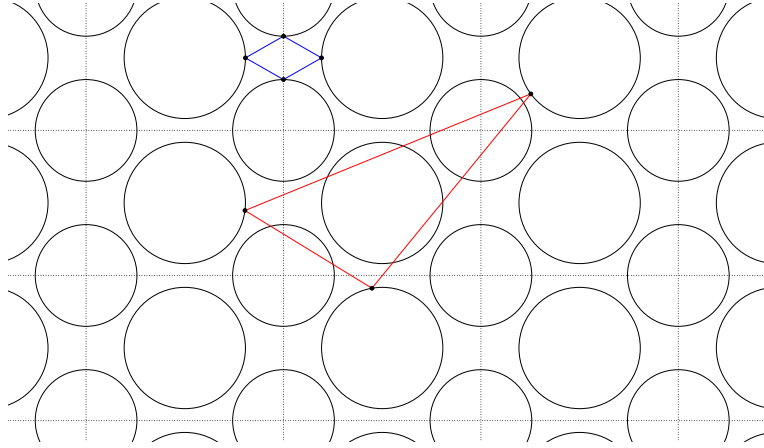


FIGURE 9. Physical and non-physical trajectories on the lifted table.

functional slightly, but they do not create or destroy scatterers, and hence do not change the set Σ_n . Moreover, as a consequence of Equation (25), ghost and regular trajectories are stable under perturbations, whereas singular orbits can become either regular or non-physical. Either way, the sequence associated to a generalised orbit will still be associated to the corresponding perturbed orbit.

Lemma A.11. *Let $p \in \partial\mathcal{D}$ and $n \in \mathbb{N}$. Then, for any sufficiently small neighbourhood $U \subset \partial\mathcal{D}$ of p , there exists localised perturbations of $\mathcal{D}$ at p with no singular points of order n or less inside U .*

Proof. As there are finitely many periodic orbits of order n , the result is straightforward for any point not lying on a periodic orbit. Assume $p \in \Gamma_a$, $a \in \mathbb{A}$, is a vertex of a periodic trajectory. Let

$$\mathcal{A} \subset \bigcup_{i=2}^n \Sigma_i$$

be the set of all sequences of n or less elements containing a . We further partition $\mathcal{A}$ into the sets $\mathcal{A}_{\text{emp}}$, of sequences corresponding to no periodic trajectory; and its complement $\mathcal{A}_{\text{orb}} := \mathcal{A} \setminus \mathcal{A}_{\text{emp}}$, of sequences corresponding to periodic trajectories (physical or not).

As there are finitely many closed orbits of order less than n , there exists a neighbourhood U of p which does not contain the vertices of any periodic orbits of order n or less other than ones reflecting exactly at p . Now, the generalised orbits reflecting at p will all persist under perturbation localised at U , provided it is sufficiently small. In other words, small localised perturbations do not decrease the size of the set $\mathcal{A}_{\text{orb}}$. They could however, create a new periodic orbit O associated to one the sequences previously in $\mathcal{A}_{\text{emp}}$. Since, however, such set is finite, there are only finitely many perturbations which could possibly create such new orbits.

Choose the perturbation to be a retraction $\tilde{\gamma}$ of Γ_a at p with support J such that $\tilde{\gamma}(J) \subset U$. If any of the orbits containing p were grazing at p , the grazing collision does not happen any more after the perturbation. The regular and ghost orbits reflecting at p will persist after the perturbation, provided it is small enough. Moreover, we can choose our perturbation in such a way that any singular orbit having a regular collision at p is either regular or non-physical after perturbation. Finally, if the perturbation creates any new periodic orbits, that is, if it increases the

size of $\mathcal{A}_{\text{orb}}$, then we repeat the procedure considering the new orbits and a suitable new neighbourhood $\tilde{U} \subset U$. Since $\#\mathcal{A}_{\text{emp}} < \infty$, the process terminates after finitely many steps. $\square$

Recall that $\mathcal{R}_n$ is the subset of $\mathfrak{D}^m$ consisting of tables without any singular periodic orbits of period less or equal to n .

Lemma A.12. *For every $n \geq 2$, the set $\mathcal{R}_n$ is open and dense in $\mathfrak{D}^m$.*

Proof. As there are only finitely many periodic orbits of order less than n , and regular orbits are persistent under sufficiently small perturbations, it follows that any table close enough to a table in $\mathcal{R}_n$ is also in $\mathcal{R}_n$. We need only show that $\mathcal{R}_n$ is dense. Let $\mathcal{D} \in \mathfrak{D}^m$. Using induction on the order of periodic orbits, we construct an element of $\mathcal{R}_n$ arbitrarily close to $\mathcal{D}$.

The base case $n = 2$. A singular periodic orbit of period 2 has two collision points x, y (on possibly distinct scatterers), at least one of which, say $p = x$, is grazing. Retract the obstacle at p by an arbitrarily small localised perturbation tangent to the orbit: by Remark A.9 this removes the grazing collision at p without creating a new period-2 singular orbit, since the other collision (if regular) persists and the reflection law at y is unchanged. Iterating over the finitely many period-2 singular orbits yields a table in $\mathcal{R}_2$ arbitrarily close to $\mathcal{D}$.

The general case. Assume, by induction, that the result holds for $\mathcal{R}_{n-1}$. We lift our initial table $\mathcal{D}$ and consider the set Σ_n of all possible encodings of periodic orbits of order n . Since there are only finitely many periodic orbits, the collection of all reflection points of singular orbits is also finite, and we can choose a minimal finite sub-collection $\{p_1, \dots, p_k\}$ containing at least one reflection point of each singular orbit.

Starting with p_1 , we can use Lemma A.11 to find an arbitrarily small perturbation $\mathcal{D}'$ of $\mathcal{D}$ around p_1 which removes all the singular orbits of order n reflecting at p_1 , without creating any new ones. Moreover, since the perturbation is localised away from singular orbits not reflecting at p_1 , such orbits comprise the full set of singular orbits of order n of $\mathcal{D}'$.

Now, according to Lemma A.10, $\mathcal{D}'$ still has finite horizon, so that we can once again repeat the construction done to $\mathcal{D}$ to encode its periodic orbits using the set Σ_n . Furthermore, the set $\{p_2, \dots, p_k\}$ is a minimal sub-collection of reflection points representing all the singular orbits of $\mathcal{D}'$, and we can repeat the procedure for p_2 .

After finitely many steps we obtain a perturbed table, which we still call $\mathcal{D}'$, all of whose periodic orbits of order n are regular. In particular, since regular orbits are persistent under perturbation, any table sufficiently close to $\mathcal{D}'$ also has this property. It follows from the induction hypothesis that $\mathcal{D}'$ can be approximated by a table with no singular orbits of order less than or equal to n . As such approximations can be made arbitrarily close, it follows that there are tables in $\mathcal{R}_n$ arbitrarily close to $\mathcal{D}$, as we wanted. $\square$

Recall that $\mathfrak{S}_n$ is the subset of $\mathfrak{D}^m$ consisting of tables such that no distinct periodic orbits of period less or equal to n have the same length.

Lemma A.13. *For every $n \geq 2$, the set $\mathfrak{S}_n$ is open and dense in $\mathfrak{D}^m$.*

Proof. From the equality in (24) we can see that the first order variation of perimeter from a perturbation of the form $\gamma^\lambda(s) = \gamma(s) + \varepsilon\lambda(s)n_s$ is

$$\Delta_\sigma(\psi) = P^\lambda(\sigma) + DP^\lambda(\sigma)(\psi), \quad (26)$$

where $\sigma = (s_1, \dots, s_n)$ are the reflection points of the original periodic orbit and the corresponding periodic orbit for the perturbation reflects at points $s_i^\lambda = s_i + \varepsilon\psi_i$. In particular, arbitrarily small perturbations induce arbitrarily small changes in length.

Given a table $\mathcal{D}$ in $\mathfrak{S}_n$, since there are only finitely many periodic orbits of order n or less, we have $|L(\sigma) - L(\sigma')| > \eta > 0$ for any two distinct periodic orbits σ, σ' of period at most n . Any perturbation of $\mathcal{D}$ small enough not to alter the length of such orbits by more than $\eta/2$ would still be an element of $\mathfrak{S}_n$, so $\mathfrak{S}_n$ is open.

As for density, let U be an open set in $\mathfrak{D}^m$, and $\mathcal{D}$ a table in U . Due to the existence of the minimum free-flight time for any table in $\mathfrak{D}^m$, it is sufficient to show that arbitrarily close to $\mathcal{D}$ there is a table where no two periodic orbits of the same order have the same length. If $\mathcal{D}$ itself has such a property there is nothing to show, so assume O, O' are two distinct closed trajectories in $\mathcal{D}$ with same length. Suppose first that O has a reflection point p which is not a collision point of O' . We choose the perturbation λ localised around p so that it changes the tangent vector of the obstacle at p , but does not move p itself. More explicitly, λ satisfies

$$\begin{cases} \lambda(s_0) = 0, \\ \lambda'(s_0) \neq 0. \end{cases} \quad (27)$$

With (27), every collision parameter of σ lying in the support of λ satisfies $\lambda(s_i) = 0$, so $P^\lambda(\sigma) = 0$ by (23) and the first-order variation of length vanishes. Expanding to second order and using (25),

$$\begin{aligned} L^\lambda(\sigma^\lambda) - L(\sigma) &= \varepsilon^2 \left(\frac{1}{2} D^2 L(\sigma)(\psi, \psi) + DP^\lambda(\sigma)(\psi) \right) + O(\varepsilon^3) \\ &= -\frac{\varepsilon^2}{2} D^2 L(\sigma)(\psi, \psi) + O(\varepsilon^3), \end{aligned}$$

which is strictly negative for ε small: indeed $\psi \neq 0$, because the component of $DP^\lambda(\sigma)$ at s_0 contains the non-vanishing term $2\lambda'(s_0) \cos \bar{\phi}_0$ (recall that the collision at p is regular, so $\cos \bar{\phi}_0 > 0$). Thus the length of the perturbed orbit $\tilde{O}$ differs from that of O' , and the two orbits now have distinct lengths. Thus, the increase in length of the perturbed orbit $\tilde{O}$ is non-zero, and now the two orbits have distinct lengths. As this process is localised at p , it does not change the perimeter of any other orbit not reflecting at p . Thus, the only situation where such a procedure could fail is when p is a collision point of more than one orbit (in particular, if the orbits O, O' share all their reflection points). If this happens, the orbits cannot have the same reflection angles, so the corresponding vectors $DP^\lambda(\sigma)$ and $DP^\lambda(\sigma')$ are distinct. In particular, after a preliminary perturbation of the type (27), the corresponding orbits now have different reflection points, and the procedure above can be applied.

Finally, as all the perturbations involved can be made arbitrarily small, it follows that $\mathcal{D}$ can be approximated arbitrarily well by a table in $\mathfrak{S}_n$, and this set is dense. $\square$

APPENDIX B. BILLIARD CYCLES AND THE ENRICHED LENGTH FUNCTIONAL

As seen in Section 3, every Kourganoff surface K , when compressed on its third coordinate, converges to a metric space K_0 , whose geometry is determined by the table $\mathcal{D}$. This suggests that the enriched marked length spectrum, while defined through the limiting surface K_0 , is in fact an object intrinsic to the table $\mathcal{D}$.

In what follows, we describe $\mathbf{EL}_{\mathcal{D};K}$ in terms of lengths on the table only. Recall that a geodesic of the length space $(\overline{\mathcal{D}}, d_{\mathcal{L}})$ is a locally shortest path; by Theorem 3.12 and the structure of the geodesics of K_0 , these are the concatenations of straight chords of $\mathcal{D}$ and of arcs of $\partial\mathcal{D}$, the junctions between a chord and an arc being tangential. We begin with the following definition.

Definition B.1 (Billiard cycle). A closed curve $\sigma: I \rightarrow \overline{\mathcal{D}}$ is a *billiard cycle* if there is a finite (possibly empty) set $t_0 < t_1 < \dots < t_{k-1}$ in I such that

- (i) $\sigma(t_i) \in \partial\mathcal{D}$ for $0 \leq i \leq k-1$;
- (ii) the segment $\sigma|_{[t_i, t_{i+1}]}$ is a geodesic of $d_{\mathcal{L}}$ for every $i = 0, \dots, k-1$, where $t_k = t_0$ (for $k = 0$, σ itself is a closed geodesic of $d_{\mathcal{L}}$);
- (iii) at each point $\sigma(t_i)$ the curve undergoes a specular reflection, that is:

$$\dot{\sigma}_+(t_i) = \dot{\sigma}_-(t_i) - 2\langle \dot{\sigma}_-(t_i), n \rangle n,$$

where $\dot{\sigma}_s(t_i) = \lim_{t \rightarrow t_i^s} \frac{1}{t-t_i}(\sigma(t) - \sigma(t_i))$, for $s \in \{+, -\}$, are the side limits of σ at t_i , n is the normal to $\partial\mathcal{D}$ at $\sigma(t_i)$ and $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product.

The points $\sigma(t_i)$ at which the reflection is not tangential ($\langle \dot{\sigma}_-(t_i), n \rangle \neq 0$) are the *regular collision points* of the cycle; since (iii) is automatic at a tangential contact, where the two adjacent $d_{\mathcal{L}}$ -geodesics concatenate into a single one, we may and shall always assume that the t_i are exactly the regular collision points.

It is clear that every closed orbit of the billiard flow is also a billiard cycle, as well as every closed geodesic of the length metric $d_{\mathcal{L}}$.

Theorem B.2. *The map $\alpha \mapsto p \circ \gamma_0^\alpha$ is a surjection from $\mathcal{C}(K)$ onto the set of billiard cycles of $\mathcal{D}$, up to parametrisation. More precisely, $p \circ \gamma_0^\alpha$ is a billiard cycle for every α , and every billiard cycle σ is of the form $p \circ \gamma_0^\alpha$, traversed once if σ has an even number of regular collision points and twice otherwise.*

Proof. We begin by showing that, for any $\alpha \in \mathcal{C}(K)$, the geodesic γ_0^α projects to a billiard cycle. We consider an arc-length parametrisation $\gamma_0^\alpha: [0, l_\alpha] \rightarrow K_0$. We need to construct a partition $\{t_0 < \dots < t_{k-1}\}$ of $[0, l_\alpha]$ satisfying properties (i), (ii) and (iii).

First, suppose $\gamma_0^\alpha(\mathbb{R}) \subset B(K)$. In this case, the curve γ_0^α coincides with one of the walls, hence projects to one of the Γ_i (possibly traversed several times), which is a closed geodesic of $d_{\mathcal{L}}$; we take $k = 0$.

Otherwise, the set $\gamma_0^\alpha \setminus B(K)$ is a non-empty open subset of γ_0^α , each of whose connected components is a chord of a flat sheet joining two points of $B(K)$. A chord leaving an obstacle tangentially does not meet that obstacle again (the obstacle is convex), so every chord has Euclidean length at least the distance between two distinct obstacles; hence there are finitely many of them, and

$$\gamma_0^\alpha \setminus B(K) = \bigcup_{j=0}^q \gamma_0^\alpha(I_j)$$

for finitely many open subintervals $I_j \subset [0, l_\alpha]$, labelled so that $\sup I_j < \inf I_{j+1}$. Let $t_j := \sup I_j$ and $s_j := \inf I_{j+1}$ (indices modulo $q+1$). Each $\gamma_0^\alpha|_{I_j}$ lies in a single sheet, hence projects to a straight segment of $\mathcal{D}$ (Lemma 3.8). At the contact $\gamma_0^\alpha(t_j)$ with the seam, three cases may occur (Lemma 3.48):

- the contact is *transversal*: then $s_j = t_j$, the geodesic changes sheet, and by Lemma 3.48(iii) the projections of $\gamma_0^\alpha|_{I_j}$ and $\gamma_0^\alpha|_{I_{j+1}}$ form a billiard path, i.e., obey the specular reflection law at $p(\gamma_0^\alpha(t_j))$;

- the contact is *tangential and isolated* (a grazing contact): then $s_j = t_j$ and $p(\gamma_0^\alpha|_{I_j \cup \{t_j\} \cup I_{j+1}})$ is a single straight segment, tangent to $\partial\mathcal{D}$ at $p(\gamma_0^\alpha(t_j))$;
- the geodesic *glides*: then $s_j > t_j$, $\gamma_0^\alpha|_{[t_j, s_j]}$ is a wall arc, and $p(\gamma_0^\alpha|_{\overline{I_j \cup [t_j, s_j] \cup I_{j+1}}})$ is a chord–arc–chord path of $\overline{\mathcal{D}}$ with tangential junctions, i.e., a geodesic of $d_{\mathcal{L}}$ (the two chords may lie in different sheets of K_0 , which is invisible on the table).

So we proceed as follows: we let the partition consist of the parameters t_j at which the contact is transversal. Between two consecutive such parameters, the projection of γ_0^α is a concatenation of straight segments and boundary arcs with tangential junctions, which is a geodesic of $d_{\mathcal{L}}$ by the above; at the parameters of the partition the projection obeys the specular reflection law. Hence $p \circ \gamma_0^\alpha$ is a billiard cycle, whose regular collision points are the transversal contacts of γ_0^α with the seam. This finishes the first part of the proof.

Conversely, we show that every billiard cycle of $\mathcal{D}$ is the image under p of a geodesic γ_0^α . Let σ be a billiard cycle, and $x_i = \sigma(t_i)$, $i = 0, 1, \dots, k-1$, its regular collision points, with $t_k = t_0$. Denote by $\mathcal{D}_{\text{up}}$ and $\mathcal{D}_{\text{down}}$ the two closed sheets of K_0 ; by Lemma 3.11, p restricts to isometries $(\mathcal{D}_{\text{up}}, d_0) \rightarrow (\overline{\mathcal{D}}, d_{\mathcal{L}})$ and $(\mathcal{D}_{\text{down}}, d_0) \rightarrow (\overline{\mathcal{D}}, d_{\mathcal{L}})$, and each $x_i \in \partial\mathcal{D}$ has a single preimage $\tilde{x}_i \in B(K)$. We lift σ to a curve σ_{up} in K_0 in the following way. We lift the first geodesic segment $\sigma|_{[t_0, t_1]}$ to the sheet $\mathcal{D}_{\text{up}}$, which can be done in a unique way, and obtain a d_0 -geodesic segment from $\tilde{x}_0$ to $\tilde{x}_1$. Then we lift the second geodesic segment $\sigma|_{[t_1, t_2]}$ to the sheet $\mathcal{D}_{\text{down}}$, and so on, alternating sheets at each regular collision point. By Lemma 3.48(iii), the concatenation at $\tilde{x}_1$ of two geodesic segments lying in different sheets is a local geodesic of K_0 exactly when the collision is specular; hence the lift of $\sigma|_{[t_0, t_2]}$ thus defined is a d_0 -geodesic, and so on.

If we proceed this way, we get a curve σ_{up} in K_0 from $\tilde{x}_0$ to $\tilde{x}_k = \tilde{x}_0$, which is locally minimising everywhere except maybe at its endpoints. If k is even, the last segment $\sigma|_{[t_{k-1}, t_k]}$ is lifted to $\mathcal{D}_{\text{down}}$, so that σ_{up} closes up with a sheet change at $\tilde{x}_0$, and the reflection being specular there, σ_{up} is a closed local geodesic of K_0 . If k is odd, the lifts of $\sigma|_{[t_0, t_1]}$ and $\sigma|_{[t_{k-1}, t_k]}$ lie in the same sheet $\mathcal{D}_{\text{up}}$, and σ_{up} is not locally minimising at $\tilde{x}_0$: the convexity of the scatterers means that for any two points $\sigma(t_0 - \varepsilon)$ and $\sigma(t_0 + \varepsilon)$ the straight segment

$$r(t) = \sigma(t_0 - \varepsilon) + t(\sigma(t_0 + \varepsilon) - \sigma(t_0 - \varepsilon)), \quad t \in [0, 1],$$

is entirely contained in $\mathcal{D}$, lifts to $\mathcal{D}_{\text{up}}$, and has Euclidean length smaller than $l_0(\sigma|_{[t_0 - \varepsilon, t_0]}) + l_0(\sigma|_{[t_0, t_0 + \varepsilon]})$. To deal with this, we consider the curve σ_{down} , the lift of σ constructed in the same way but starting in the sheet $\mathcal{D}_{\text{down}}$; it goes from $\tilde{x}_0$ to $\tilde{x}_0$, starting in $\mathcal{D}_{\text{down}}$ and ending in $\mathcal{D}_{\text{up}}$. The concatenation $\sigma_{\text{up}} * \sigma_{\text{down}}$ (which traverses σ twice) changes sheet at every passage through $\tilde{x}_0$ and at every other regular collision point, with the specular law, and is therefore a closed local geodesic of K_0 .

Finally, we let α be the free homotopy class of the closed local geodesic just constructed (σ_{up} or $\sigma_{\text{up}} * \sigma_{\text{down}}$ according to the parity of k). A closed local geodesic of the locally CAT(0) space K_0 lifts to an axis of the corresponding element of $\pi_1(K)$, hence is the unique closed geodesic in its class (Lemma 3.38(iii), Lemma 3.39); thus it equals γ_0^α . In the even-parity case $p(\gamma_0^\alpha) = \sigma$; in the odd-parity case $p \circ \gamma_0^\alpha$ traverses σ once in each sheet, so it coincides with σ as a subset of $\overline{\mathcal{D}}$ and $l_0(\gamma_0^\alpha) = 2l(\sigma)$. In either case p maps γ_0^α onto the billiard cycle σ , which shows the surjectivity. $\square$

B.1. The enriched length functional. In this section we think of $\mathcal{D}$ as a periodic billiard table on $\mathbb{R}^2$, and consequently of K as a periodic surface in $\mathbb{R}^3$. Our goal is to describe billiard cycles as minima of a length functional, in the spirit of Appendix A. To that end, we keep the same notation: a cycle σ visits a periodic sequence of scatterers $\Gamma_{\rho_0}, \dots, \Gamma_{\rho_{q-1}}$, $\rho_k \in \{1, \dots, m\}$, and we let $\rho = \rho(\sigma) := (\rho_k)_{k=0, \dots, q-1}$. Lifting σ to $\tilde{\mathcal{D}}$ starting in the cell $(0, 0)$, we denote by $\{(\tilde{i}_k, \tilde{j}_k)\}_{k=0, \dots, q}$ the labels of the cells of the first $(q+1)$ scatterers visited, with $(\tilde{i}_0, \tilde{j}_0) = (0, 0)$, and we set $I_k := (\tilde{i}_k, \tilde{j}_k) - (\tilde{i}_{k-1}, \tilde{j}_{k-1})$ for $k \in \{1, \dots, q\}$ and $I = I(\sigma) := (I_k)_{k=1, \dots, q}$. Finally, on each scatterer Γ_{ρ_k} the cycle follows a boundary arc (possibly reduced to a point) from an *entry point* to an *exit point*, in one of the two orientations of Γ_{ρ_k} ; we let $\epsilon_k \in \{+, -\}$ be that orientation and $\epsilon = \epsilon(\sigma) := (\epsilon_k)_{k=0, \dots, q-1}$ (when the arc is a point, either sign may be chosen).

For $i \in \{1, \dots, m\}$ and $\epsilon \in \{+, -\}$, let $D_i^\epsilon: \mathbb{T}_i \times \mathbb{T}_i \rightarrow [0, l_i)$, $(e, s) \mapsto D_i^\epsilon(e, s)$, be the length of the arc of Γ_i from $\gamma_i(e)$ to $\gamma_i(s)$ traversed in the direction ϵ , i.e., $D_i^+(e, s) = (s - e) \bmod l_i$ and $D_i^-(e, s) = (e - s) \bmod l_i$; it is smooth away from the diagonal $\{e = s\}$, where it jumps (the length of the shortest arc, $\min(D_i^+, D_i^-)$, would not suffice: a gliding arc of a closed geodesic of K_0 may well be longer than half the perimeter of the scatterer).

Definition B.3 (Link sets and visibility domain). Given a pair of indices $(i, j) \in \{1, \dots, m\}^2$ and a cell offset $J \in \mathbb{Z}^2$, the *link set* $L_{ij}^J \subset \mathbb{T}_i \times \mathbb{T}_j$ is the set of pairs (s, e) such that the open segment joining $\gamma_i(s)$ to $\gamma_j(e) + J$ in $\mathbb{R}^2$ is contained in the interior of the lifted table $\tilde{\mathcal{D}}$. It is a closed subset of $\mathbb{T}_i \times \mathbb{T}_j$, possibly disconnected (a third scatterer may hide part of $\Gamma_j + J$ from Γ_i). Given data (ρ, I) as above, the *visibility domain* is

$$\mathbb{L}_\rho^I := \{S = (e_0, s_0, e_1, s_1, \dots, e_{q-1}, s_{q-1}) \in (\mathbb{T}_{\rho_0} \times \mathbb{T}_{\rho_0}) \times \dots \times (\mathbb{T}_{\rho_{q-1}} \times \mathbb{T}_{\rho_{q-1}}) : \\ (s_{k-1}, e_k) \in L_{\rho_{k-1}\rho_k}^{I_k} \text{ for } k = 1, \dots, q\},$$

with the convention that $(e_q, s_q) = (e_0, s_0)$ and I_q is the offset returning to the initial cell.

To a point $S \in \mathbb{L}_\rho^I$ and orientations ϵ we associate the closed curve c_S of $\tilde{\mathcal{D}}$ obtained by concatenating, for each k , the chord from $\gamma_{\rho_{k-1}}(s_{k-1})$ to $\gamma_{\rho_k}(e_k)$ (unobstructed, by definition of the link sets) with the arc of Γ_{ρ_k} from $\gamma_{\rho_k}(e_k)$ to $\gamma_{\rho_k}(s_k)$ traversed in the direction ϵ_k .

Definition B.4 (Enriched length functional). The enriched length functional associated to the data (ρ, I, ϵ) is

$$\text{EL}_\rho^{I, \epsilon}: \begin{cases} \mathbb{L}_\rho^I \rightarrow \mathbb{R} \\ S \mapsto \sum_{k=1}^q \tau_{I_k}(s_{k-1}, e_k) + D_{\rho_k}^{\epsilon_k}(e_k, s_k) \\ \quad = L_\rho^I(s_0, e_1, \dots, s_{q-1}, e_q) + \sum_{k=1}^q D_{\rho_k}^{\epsilon_k}(e_k, s_k), \end{cases}$$

where τ_{I_k} and L_ρ^I are the chord lengths and the length functional of Appendix A (there denoted L), evaluated on the $2q$ -tuple of chord endpoints. By construction, $\text{EL}_\rho^{I, \epsilon}(S)$ is the $d_{\mathcal{L}}$ -length of the closed curve c_S .

Remark that if $s = (s_0, \dots, s_{q-1})$ is a periodic orbit of the billiard, that is, a critical point of L_ρ^I (a non-degenerate local minimum, by Lemma A.1), then $L_\rho^I(s) = \text{EL}_\rho^{I, \epsilon}(S)$ for $S = (s_0, s_0, s_1, s_1, \dots, s_{q-1}, s_{q-1})$ and any ϵ . More generally, we have the following.

Proposition B.5. *Let σ be a billiard cycle with data (ρ, I, ϵ) , let $S_\sigma \in \mathbb{L}_\rho^I$ be the point formed by its entry and exit points, and let $\mathbb{L}_\sigma$ be the connected component of $\mathbb{L}_\rho^I$ containing S_σ . Then S_σ is a global minimum of $\text{EL}_\rho^{I, \epsilon}$ on $\mathbb{L}_\sigma$, and it is the only one: $\text{EL}_\rho^{I, \epsilon}(S) > \text{EL}_\rho^{I, \epsilon}(S_\sigma) = l(\sigma)$ for every $S \in \mathbb{L}_\sigma \setminus \{S_\sigma\}$. In particular, the lengths of the billiard cycles of $\mathcal{D}$, hence the enriched marked length spectrum, are determined by the functionals $\text{EL}_\rho^{I, \epsilon}$, which are defined in terms of the table only.*

Proof. Let $\tilde{\sigma}$ be the closed local geodesic of K_0 constructed in the proof of Theorem B.2, and α its free homotopy class; recall that $\tilde{\sigma} = \gamma_0^\alpha$ and that $p \circ \tilde{\sigma}$ traverses σ once or twice according to the parity of the number of regular collision points of σ ; we treat the even case, the odd one being identical after doubling. For $S \in \mathbb{L}_\rho^I$, lift the curve c_S to a closed curve $\tilde{c}_S$ of K_0 by lifting its chords to the sheets prescribed by the sheet pattern of $\tilde{\sigma}$ (the k -th chord of c_S is lifted to the sheet containing the k -th chord of $\tilde{\sigma}$) and its arcs to the seam; the pattern closes up, so $\tilde{c}_S$ is a closed curve, which depends continuously on S , and $\tilde{c}_{S_\sigma} = \tilde{\sigma}$. Since the lifts of the chords are chords of flat sheets and the lifts of the arcs are wall arcs, the d_0 -length of $\tilde{c}_S$ is the $d_{\mathcal{L}}$ -length of c_S , namely $\text{EL}_\rho^{I, \epsilon}(S)$ (Theorem 3.6). By continuity, all the curves $\tilde{c}_S$, $S \in \mathbb{L}_\sigma$, are freely homotopic to $\tilde{\sigma}$, i.e., lie in the class α ; therefore

$$\text{EL}_\rho^{I, \epsilon}(S) = l_0(\tilde{c}_S) \geq \ell(\alpha) = l_0(\gamma_0^\alpha) = \text{EL}_\rho^{I, \epsilon}(S_\sigma),$$

with equality only if $\tilde{c}_S$ is a closed curve of minimal length in its class: its lift to $\widetilde{K}_0$ then joins a point x to gx , with $g \in \alpha$, and has length $\ell(g) = \min d_0(\cdot, g \cdot)$, so that it is an axis segment of g , and $\tilde{c}_S$ is the closed geodesic $\gamma_0^\alpha = \tilde{\sigma}$ (Lemma 3.38(iii)). As the entry and exit points of $\tilde{c}_S$ are those of S , this forces $S = S_\sigma$. $\square$

Acknowledgements. We are grateful to the LESET/MATH-AMSUD project, coordinated by Jérôme Buzzi, Pablo D. Carrasco, and Radu Saghin, for making this collaboration possible. We thank the Centre de Mathématiques Laurent Schwartz (CMLS) at École polytechnique for its hospitality during the development of this project. We also thank Andrey Gogolev for his valuable insights and discussions related to this work, and Mickaël Kourganoff for permission to reproduce the figures from [36]. We are very grateful to David Lautenschlager and Peter Albers for their careful reading of a previous version of this work, as well as for their insightful comments regarding inaccuracies in some statements. In the preparation of the revised version, the authors used Claude Opus (Anthropic) to improve the exposition and to draft the proofs of some technical lemmas in Section 4. The authors reviewed and take full responsibility for all content.

D.F. was supported by CAPES Grant No. 88887.898617/2023-00 as part of the LESET/MATH-AMSUD project, and would like to acknowledge support from the Max Planck Society as well, through the award of a Max Planck Partner Group for Geometry and Probability in Dynamical Systems. M.L. was supported by the ANR AAPG 2021 PRC CoSyDy (Conformally symplectic dynamics, beyond symplectic dynamics; Grant No. ANR-CE40-0014) and by the ANR JCJC PADAWAN (Parabolic dynamics, bifurcations, and wandering domains; Grant No. ANR-21-CE40-0012).

REFERENCES

- [1] Alberto Abbondandolo and Marco Mazzucchelli. Length spectrum rigidity and flexibility of spheres of revolution with one equator, 2026. arXiv:2601.16804v1.
- [2] Uwe Abresch and Wolfgang T. Meyer. Injectivity radius estimates and sphere theorems. In *Comparison Geometry*, volume 30, pages 1–47. MSRI Publications, 1997.
- [3] Karl Andersson and Richard Melrose. The propagation of singularities along gliding rays. *Inventiones Mathematicae*, 41:197–232, 1977.

- [4] Viviane Baladi, Mark F Demers, and Carlangelo Liverani. Exponential decay of correlations for finite horizon sinai billiard flows. *Inventiones mathematicae*, 211(1):39–177, 2018.
- [5] Péter Bálint, Jacopo De Simoi, Vadim Kaloshin, and Martin Leguil. Marked length spectrum, homoclinic orbits and the geometry of open dispersing billiards. *Communications in Mathematical Physics*, 374(3):1531–1575, 2020.
- [6] Werner Ballmann. *Lectures on Spaces of Nonpositive Curvature*. Oberwolfach Seminars. Birkhäuser Basel, 1995.
- [7] Anja Bankovic and Christopher Leininger. Marked-length-spectral rigidity for flat metrics. *Transactions of the American Mathematical Society*, 370(3):1867–1884, October 2017.
- [8] Francis Bonahon. Bouts des Variétés Hyperboliques de Dimension 3. *Annals of Mathematics*, 124(1):71–158, 1986.
- [9] Marc Bourdon. Structure conforme au bord et flot géodésique d’un CAT(−1)-espace. *Enseign. Math. (2)*, 41(1-2):63–102, 1995.
- [10] Marc Bourdon. Sur le birapport au bord des CAT(−1)-espaces. *Publications Mathématiques de l’Institut des Hautes Études Scientifiques*, 83(1):95–104, December 1996.
- [11] Martin R. Bridson and André Häfliger. *Metric Spaces of Non-Positive Curvature*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2011.
- [12] Leonid A. Bunimovich. Variational principle for periodic trajectories of hyperbolic billiards. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(2):349–355, June 1995.
- [13] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2022.
- [14] Aaron Calderon, Solly Coles, Diana Davis, Justin Lanier, and Andre Oliveira. How to hear the shape of a billiard table. *arXiv preprint arXiv:1806.09644*, 2018.
- [15] Nikolai Chernov and Roberto Markarian. *Chaotic Billiards*. American Mathematical Soc., 2006.
- [16] David Constantine. Marked length spectrum rigidity in non-positive curvature with singularities. *Indiana Univ. Math. J.*, 67(6):2337–2361, 2018.
- [17] M. Coornaert, T. Delzant, and A. Papadopoulos. *Géométrie et théorie des groupes: les groupes hyperboliques de Gromov*, volume 1441 of *Lecture Notes in Mathematics*. Springer, 1990.
- [18] Christopher B Croke. Rigidity for surfaces of non-positive curvature. *Commentarii Mathematici Helvetici*, 65(1):150–169, 1990.
- [19] Françoise Dal’bo and Inkang Kim. Marked length rigidity for symmetric spaces. *Comment. Math. Helv.*, 77(2):399–407, 2002.
- [20] Jacopo De Simoi, Vadim Kaloshin, and Martin Leguil. Marked length spectral determination of analytic chaotic billiards with axial symmetries. *Inventiones mathematicae*, 233(2):829–901, 2023.
- [21] Jacopo De Simoi, Vadim Kaloshin, Qiaoling Wei, and Hamed Hezari. Dynamical spectral rigidity among $\mathbb{Z}_2$ -symmetric strictly convex domains close to a circle (appendix b coauthored with h. hezari). *Annals of Mathematics*, 186(1):277–314, 2017.
- [22] Moon Duchin, Viveka Erlandsson, Christopher J Leininger, and Chandrika Sadanand. You can hear the shape of a billiard table: Symbolic dynamics and rigidity for flat surfaces. *Commentarii Mathematici Helvetici*, 96(3):421–463, 2021.
- [23] Moon Duchin, Christopher J. Leininger, and Kasra Rafi. Length spectra and degeneration of flat metrics. *Invent. Math.*, 182(2):231–277, 2010.
- [24] Todd Fisher and Boris Hasselblatt. *Hyperbolic Flows*. Zurich lectures in advanced mathematics. European Mathematical Society, 2019.
- [25] Alex Furman. Coarse-Geometric Perspective on Negatively Curved Manifolds and Groups. In Marc Burger and Alessandra Iozzi, editors, *Rigidity in Dynamics and Geometry: Contributions from the Programme Ergodic Theory, Geometric Rigidity and Number Theory, Isaac Newton Institute for the Mathematical Sciences Cambridge, United Kingdom, 5 January – 7 July 2000*, pages 149–166. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.
- [26] Etienne Ghys. Flots d’Anosov sur les 3-variétés fibrées en cercles. *Ergodic Theory and Dynamical Systems*, 4(1):67–80, March 1984.
- [27] Colin Guillarmou, Gerhard Knieper, and Thibault Lefeuvre. Geodesic stretch, pressure metric and marked length spectrum rigidity. *Ergodic Theory and Dynamical Systems*, 42(3):974–1022, March 2022.
- [28] Colin Guillarmou and Thibault Lefeuvre. The marked length spectrum of Anosov manifolds. *Annals of Mathematics*, 190(1):321–344, July 2019. Publisher: Department of Mathematics of Princeton University.

- [29] Colin Guillarmou, Thibault Lefeuvre, and Gabriel P Paternain. Marked length spectrum rigidity for anosov surfaces. *Duke Mathematical Journal*, 174(1):131–157, 2025.
- [30] Sa’ar Hersensky and Frédéric Paulin. On the rigidity of discrete isometry groups of negatively curved spaces. *Comment. Math. Helv.*, 72(3):349–388, 1997.
- [31] Luc Hillairet. Contribution of periodic diffractive geodesics on surfaces with conical singularities. *Journal of Functional Analysis*, 226:48–89, 2005.
- [32] Morris W. Hirsch. *Differential Topology*. Graduate Texts in Mathematics. Springer New York, 2012.
- [33] In Kang Kim. Marked length rigidity of rank one symmetric spaces and their product. *Topology*, 40(6):1295–1323, 2001.
- [34] In Kang Kim. Cross-ratio in higher rank symmetric spaces. *Forum Math.*, 22(1):179–186, 2010.
- [35] Wilhelm Klingenberg. Riemannian Manifolds With Geodesic Flow of Anosov Type. *Annals of Mathematics*, 99(1):1–13, 1974. Publisher: [Annals of Mathematics, Trustees of Princeton University on Behalf of the Annals of Mathematics, Mathematics Department, Princeton University].
- [36] Mickaël Kourganoff. Anosov Geodesic Flows, Billiards and Linkages. *Communications in Mathematical Physics*, 344(3):831–856, June 2016.
- [37] John Lee. *Introduction to Riemannian Manifolds*. Graduate Texts in Mathematics. Springer International Publishing, 2019.
- [38] Takehiko Morita. The symbolic representation of billiards without boundary condition. *Transactions of the American Mathematical Society*, 325(2):819–828, 1991.
- [39] Jean-Pierre Otal. Le Spectre Marque des Longueurs des Surfaces a Courbure Negative. *Annals of Mathematics*, 131(1):151–162, 1990.
- [40] Jean-Pierre Otal. Sur la géométrie symplectique de l’espace des géodésiques d’une variété à courbure négative. *Revista Matemática Iberoamericana*, 8(3):441–456, December 1992.
- [41] Vesselin M. Petkov and Luchezar Stoyanov. *Geometry of the Generalized Geodesic Flow and Inverse Spectral Problems*. John Wiley & Sons, 2 edition, 2017.
- [42] Yu Reshetnyak. Two-dimensional manifolds of bounded curvature. In *Geometry IV*, volume 70 of *Encyclopaedia of Mathematical Sciences*, pages 3–163. Springer, 1993.
- [43] Luis Santaló. *Integral Geometry and Geometric Probability*. Cambridge University Press, 2 edition, 2004.
- [44] Marie-France Vignéras. Variétés riemanniennes isospectrales et non isométriques. *Annals of Mathematics*, 112(1):21–32, 1980.
- [45] W. Walter. *Ordinary Differential Equations*. Graduate Texts in Mathematics. Springer New York, 2013.
- [46] Amie Wilkinson. Lectures on marked length spectrum rigidity. Lecture notes from the IAS/PCMI Summer School, 2013, 2013.
- [47] Maciej Wojtkowski. Two applications of Jacobi fields to the billiard ball problem. *Journal of Differential Geometry*, 40(1):155–164, January 1994. Publisher: Lehigh University.

CENTRO DE MATEMÁTICA, COMPUTAÇÃO E COGNIÇÃO - UFABC, ALAMEDA DA UNIVERSIDADE
 | 09606-405 SÃO BERNARDO DO CAMPO, SP, BRAZIL.
Email address: `douglas.finamore@pm.me`

CENTRE DE MATHÉMATIQUES LAURENT SCHWARTZ - ÉCOLE POLYTECHNIQUE, ROUTE DE SACLAY
 | 91128 PALAISEAU CEDEX, FRANCE.
Email address: `martin.leguil@polytechnique.edu`