Spectral gap for a class of random billiards

Renato Feres[*], Hong-Kun Zhang[†]

Abstract

A *random billiard* is a random dynamical system similar to an ordinary billiard system except that the standard specular reflection law is replaced with a more general stochastic operator specifying the post-collision distribution of velocities for any given pre-collision velocity. We consider such collision operators for certain random billiards that we call *billiards with microstructure*. Collisions modeled by these operators can still be thought of as elastic and time reversible. The operators are canonically determined by a second (deterministic) billiard system that models "microscopic roughness" on the billiard table boundary. Our main purpose here is to develop some general tools for the analysis of the collision operator of such random billiards. Among the main results, we give geometric conditions for these operators to be Hilbert-Schmidt and relate their spectrum and speed of convergence to stationary Markov chains with geometric features of the microscopic billiard structure. The relationship between spectral gap and the shape of the microstructure is illustrated with several simple examples.

1 Introduction

Billiards are widely studied dynamical systems and natural model systems in classical and statistical mechanics. In the standard set-up (see [7]) a point particle moves freely inside a domain in $\mathbb{R}^2$ with piecewise smooth boundary—*the billiard table*—reflecting elastically at the boundary according to the usual law of equal angles of incidence and reflection. More generally, one may consider billiard systems in which the reflection law is probabilistic. In this case, the particle is reflected at a random direction after each collision, according to some probability distribution that depends in general on the direction of incidence. We call systems of this kind *random billiards*. An ordinary deterministic billiard is an extreme special case for which the distribution of post-collision velocities is concentrated on the mirror image of the pre-collision velocity. Another example of random reflection often studied in classical kinetic theory of gases is the so-called *cosine law* ([13, 12]), according to which the distribution of post-collision velocities is given by the probability measure ν defined by $d\nu(v) = \rho(v)dS(v)$, where S is the area measure on the space of inward pointing unit vectors at the collision point and $\rho(v)$ is, up to a normalization constant, the cosine of the angle that v makes with the inward-pointing normal to the boundary at that point. The measure ν, which will play an important role in this paper, is often called the *(Knudsen) cosine law*. It turns out to be the (often unique) *equilibrium* (or *invariant*) scattering distribution for the types of random billiards that will be studied here.

This paper is concerned with a special class of random billiards, which we call *billiards with microstructure*. The random reflection law for these systems is defined on the basis of a geometric model of "microscopic surface texture" with which the billiard particle interacts at a collision point. This microstructure, in turn, is specified by an associated deterministic billiard system on a second

[*]Department of Mathematics, Washington University, Campus Box 1146, St. Louis, MO 63130
[†]Department of Mathematics & Statistics, University of Massachusetts, Amherst, MA 01003

billiard table Q, referred to as the *billiard cell*. Heuristically, the inner surface of the walls of the first ("macroscopic") billiard table is imagined as "microscopically rough" (see Figure 1), and the billiard cell Q is the motif of a periodic pattern defining this microscopic roughness. Throughout the paper, we suppose that $Q \subset \mathbb{R}^2$ has piecewise smooth boundary of finite total length, having one special boundary segment Γ_0, which is a flat segment. Without loss of generality, we identify Γ_0 with the interval $[0, 1]$. Further assumptions on Q will be stated later.

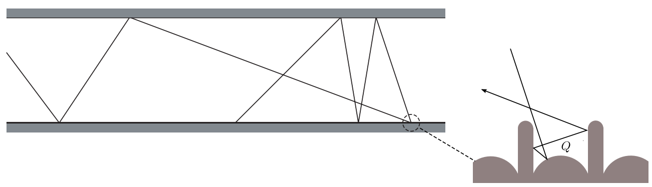

Figure 1: Heuristic description of a billiard with microstructure. The microstructure is periodic and the length scales of the "macro-" and "micro-" billiard tables (the channel and a cell of the periodic structure, respectively) are not comparable, so the precise position at which a particle enters a cell Q is not defined. This position is thus assumed to be a uniformly distributed random variable. The particular Q shown here and in Figure 2 is arbitrary.

The random reflection for a billiard with microstructure with cell Q is then defined as follows. Let $\theta \in V = [0, \pi]$ be the angle that the pre-collision velocity vector makes with the tangent to the "macroscopic" billiard table at a collision point. (This table will not play a role in the present paper, so we do not give it a special symbol, but we suggest that the reader keep in mind the channel of Figure 1. We refer to it in this introduction as the *macro-table*.) One imagines the flat side Γ_0 of Q to be aligned with the boundary of the macro-table in such a way that a billiard particle at a collision point enters Q through Γ_0, as in Figure 2. The particle then undergoes ordinary billiard motion inside Q until it returns to Γ_0 with angle $\Theta(r, \theta)$, where r is the point on Γ_0 at which the particle entered Q. The key assumption that makes Θ a random function of θ is that r is a uniformly distributed random variable, chosen independently at each collision. This is heuristically justified by the idea that micro- and macro-scales are not comparable so that we have no knowledge of the position on Γ_0 a particle enters Q.

Definition 1 (**Reflection map for billiard with microstructure**). *For a given direction of incidence of the billiard particle, represented by the angle θ as in Figure 2, the distribution of directions along which the billiard particle leaves the cell through Γ_0, represented by the angle Θ, is the probability measure ν_θ on V defined by*

$$\nu_\theta(A) = \int_0^1 \mathbb{1}_A(\Theta(r, \theta))\, dr$$

for any Borel subset $A \subset V$. From these scattering measures *ν_θ an operator P can be defined on (say, bounded) functions on V by extending the following definition on indicator functions:*

$$(P\mathbb{1}_A)(\theta) = \nu_\theta(A).$$

We are particularly concerned with P on $L^2(V, \nu)$, where ν is the Knudsen law measure defined above. It is easily shown that P is a bounded operator on $L^2(V, \nu)$. The dual operator on measures is indicated by

$$(\eta P)(f) = \eta(Pf).$$

Here η is a Borel measure on V and $\eta(f)$ is the integral of f with respect to η.

The random reflection law of a billiard with microstructure has a number of special features: (1) it depends canonically on a choice of billiard cell, leading naturally to the problem of relating the shape of the cell and properties of the associated Markov chain; (2) reflection off a billiard microstructure is still "elastic" in a certain sense. For example, the angle component of the invariant measure of deterministic billiards (i.e., the measure ν), is a stationary distribution of directions for the collision operator (or *Markov operator*) P, and it is typically unique; (3) the Markov chain comprising the sequence of random collision angles with initial distribution ν (say, in the channel shown in Figure 1) is time reversible and the process satisfies detailed balance. As shown later, this implies for Q bilaterally symmetric that P is a self-adjoint bounded (of norm 1) operator on $L^2(V, \nu)$.

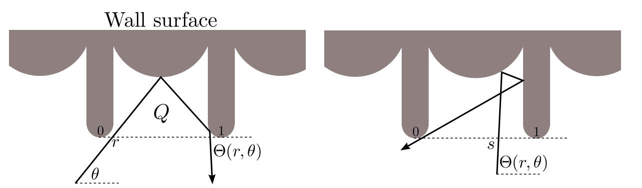

Figure 2: Angle convention for the pre- and post-collision vectors of two successive collision events. For our present purposes it is useful to think that the entire system consists of a single billiard cell Q, and that the segment Γ_0 on the boundary of Q, which is identified with $[0, 1]$, is also reflecting as in an ordinary billiard. But at each collision with Γ_0 the particle jumps instantaneously to another point of Γ_0 (indicated by s on the right-hand cell), chosen randomly and independently of the other variables, and continues its motion inside Q. From this point of view both Θ and θ measure angles from the horizontal segment counterclockwise.

The main focus of this paper is on properties of P, on the associated Markov chains, and on the relationship they have with the shape of the billiard cell. We develop tools that can help in the general analysis of these operators, and apply these tools to estimate spectral gap and rate of decay of correlations under certain geometric conditions. One geometric feature of Q that affects the spectral properties of P in an important way is the curvature of the part of the boundary of Q adjacent to the flat side Γ_0, hence the part that is *most exposed* to collisions with the billiard particle. By the *(normalized) curvature* we mean the reciprocal of the ratio of the radius of curvature of the arc by the length of Γ_0. Note that P is not affected by homotheties of Q so geometric quantities of interest are all dimension-free ratios.

It turns out that the methods for studying the effect of curvature on P are very different when the curvature is large and when it is small. For a first look at the case of small curvature see [17], where much sharper results are obtained for a special family of billiards using a perturbative method. Here we are mainly interested in large values of the curvature.

This paper is the basis for investigating other problems of interest, such as obtaining diffusion characteristics of the random billiard system in channels (as in Figure 1). The diffusion problem in channels is something that will be taken up in a subsequent paper. Thus the concern here is with the billiard cell Q and the operator P, and not with the random billiard itself. With this in mind, it is convenient to describe our random billiard in a slightly different, but equivalent, way that only refers to Q, as follows: The billiard particle moves inside Q as in an ordinary billiard system, until it collides with Γ_0, at which time the particle jumps instantaneously to a randomly chosen point on Γ_0 and continues inside Q after standard reflection on Γ_0.

Many of the more general facts proven in the subsequent sections (concerning the stationary distribution and basic operator properties of P) can be shown to hold for much more realistic models of surface roughness, including random media in dimension 3. In order to keep definitions to a minimum we only consider for now the periodic case in dimension 2 as described above.

1.1 Main results

We describe here the main results concerning convergence to the stationary measure and spectral gap. More results of general interest are given later in the paper.

The following are *standing assumptions*: $Q \subset \mathbb{R}^2$ is a domain with piecewise smooth boundary of finite length and a finite number of smooth boundary segments. The smooth segments of the boundary of Q are grouped into four sets: $\partial Q = \Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, where Γ_0 is a flat segment, Γ_1 and Γ_2 are dispersing segments adjacent to Γ_0, whose curvatures are bounded away from 0, and Γ_3 is the union of all the remaining segments. The segments Γ_1 and Γ_2 are allowed to intersect Γ_0 (at each of the endpoints of Γ_0) at angles in $[0, \pi/2)$. Without loss of generality, we often assume that Γ_0 has length 1 and identify it with the interval $I = [0, 1]$. It is convenient, but not essential, to also assume that Q is bilaterally symmetric, i.e., invariant under reflection on the line passing through the middle of Γ_0 and perpendicular to it.

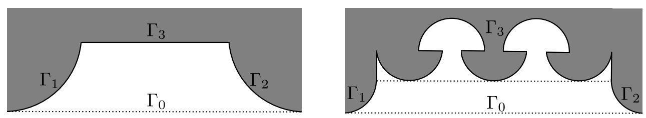

Figure 3: Examples of billiard cells for which most of the main theorems apply. The simpler cell on the left will be used for illustration purposes at a few places in the paper. For the cell on the right, Γ_3 consists of the entire contour above the (shorter) dashed line.

Let λ be the Lebesgue measure on the unit interval I and $M := I \times V$. The return billiard map is defined for a.e. point of M and it leaves invariant the probability measure $\mu := \lambda \otimes \nu$. Let $\Psi_\theta(r) := \Theta(r, \theta)$, where Θ is as shown in Figure 2. Each coordinate slice $I \times \{\theta\}$ is split by the singular set of the return map (see Section 2.1) into at most countably many intervals $W_{\theta,j}$ and on each such interval $\Psi_{\theta,j} := \Psi_\theta|_{W_{\theta,j}}$ is a diffeomorphism onto its image $V_{\theta,j} \subset V$. Now define the *expansion factor* $\Lambda_{\theta,j}$ on $V_{\theta,j}$ by

$$\Lambda_{\theta,j}(\varphi) := \frac{1}{2}\left|\Psi'_\theta(\Psi^{-1}_{\theta,j}(\varphi))\right| \sin\varphi.$$

The following simple fact is a special case of the main observation of Section 2.6.

Proposition 1. *If $\Psi'_\theta(r) \neq 0$ for μ-a.e. (r, θ) in M, then P is an integral operator on $L^2(V, \nu)$ with integral kernel*

$$\omega(\theta, \varphi) := \sum_j \mathbb{1}_{V_{\theta,j}}(\varphi)/\Lambda_{\theta,j}(\varphi).$$

When P is an integral operator, then P is a compact operator if its integral kernel ω lies in $L^2(V \times V, \nu \otimes \nu)$. A general criterion for compactness was given in [14].

Assumption 1. *In addition to the standing assumptions suppose that Γ_1 and Γ_2 are arcs of circle of equal radius, subtended angle $\pi/2$ (quarter circles), both tangent to Γ_0 at each of the endpoints of Γ_0, and the straight line connecting the endpoints of Γ_1 and Γ_2 not in Γ_0 separates Γ_3 from Γ_0. (See the billiard cells of Figure 3.)*

This assumption can likely be relaxed without significantly changing the conclusions of the below theorems, but the cost would be to add technical complications that we wish to avoid for the sake of clarity. For example, in some of the case studies discussed in Section 1.2, Q contains vertical line segments in positions that are not strictly permitted by Assumption 1, although it would take relatively small changes in the analysis so as to allow this feature.

The *spectral gap* of P and of other stochastic operators to appear shortly, is defined here as 1 minus the spectral radius of the restriction of P to the orthogonal complement of the constant functions in $L^2(V,\nu)$. The operator will be called *quasi-compact* if its essential spectral radius is strictly less than the spectral radius (which is equal to 1).

By *curvature*, in particular the constant curvature K of Γ_1 and Γ_2 under Assumption 1, it is always understood the ratio of the length of Γ_0 (typically taken without loss of generality to be 1) by a radius of curvature.

Theorem 1. *Under Assumption 1, the collision operator P is quasi-compact and its spectral gap γ satisfies*

$$\gamma \geq \frac{\sqrt{2}-\epsilon}{K}$$

for any $\epsilon > 0$ and all sufficiently large curvature K. There are examples (below) for which $\gamma \leq 2/K$.

Most of the analysis of P involves another operator, P_1, that captures the isolated effect of the scattering bumps Γ_1 and Γ_2. This is based on a general method that we refer to as *conditioning*, which can be used more generally to extricate the effect on P of various geometric features of Q. Conditioning is used later in several instances for a number of different purposes.

In particular, focusing attention on the bumps leads to a P_1 defined roughly as follows. Let M_1 be the subset of M consisting of initial conditions whose billiard trajectories only undergo one collision with the boundary of Q before returning to Γ_0, and this collision is with Γ_1 or Γ_2. Then define P_1 as P conditional on the event M_1. It turns out that P_1 is a bounded self-adjoint operator on $L^2(V,\nu_1)$ of norm 1, where ν_1 is the unique stationary measure for P_1, and is absolutely continuous with respect to ν. Similarly define P_2 from $M_2 = M \setminus M_1$. Define $\alpha_i(\theta) = \lambda(M_{i,\theta})$, the Lebesgue measure of the intersection $M_{i,\theta}$ of M_i and the slice $[0,1]\times\{\theta\}$, and let as before μ denote the probability measure on M invariant under the return billiard map. We refer to M_1, M_2 as the *special partition* of M. Other partitions will come up in the course of the paper. A common feature of many of them is that elements of the partition are cut out by singular curves of the return billiard map to M.

Theorem 2. *Let M_1, M_2 be the special partition and $\mathfrak{I}_i : L^2(V,\nu) \to L^2(V,\nu_i)$ the inclusion operator. This is a bounded isomorphism with bounded inverse. Then P decomposes as*

$$P = \mu(M_1)\mathfrak{I}_1^* P_1 \mathfrak{I}_1 + \mu(M_2)\mathfrak{I}_2^* P_2 \mathfrak{I}_2.$$

and the following hold:

1. *For any $\epsilon > 0$ and all sufficiently big K*

$$\|\mu(M_2)\mathfrak{I}_2^* P_2 \mathfrak{I}_2\|_2 \leq 1 - (\sqrt{2}-\epsilon)/K.$$

2. *P_1 is a Hilbert-Schmidt operator of norm 1 whose spectral gap satisfies*

$$\gamma_1 \geq 1 - O(K^{-1/N})$$

for some positive integer N, and all K sufficiently big. (More detailed information is obtained in the course of the proof.) Therefore, the spectral gap of P_1 approaches its maximum allowed value as the normalized curvature becomes very big.

3. *The stochastic operator P_1 is* geometrically ergodic*: If ν_1 denotes its invariant probability measure (described explicitly later) and η is any probability measure not having atoms at 0 and π then, for some positive constant C_η,*

$$\|\eta P_1^m - \nu_1\|_{TV} \leq C_\eta r^{m/3}$$

for all positive integers m, where $r = 1-(2-\sqrt{2})(1-O(1/K))$. P_1 restricted to $L^\infty(V,\nu_1)$ satisfies exponential decay of correlations with the same rate $r^{1/3}$.

Our results concerning P have parallels with certain results about chaotic billiard maps in the deterministic setting. Chaotic billiards have been studied by Sinai, Bunimovich, Chernov [19, 1, 3, 4] and many others, who have established ergodicity and strong statistical properties under very general assumptions on the boundary of the table [6, 18, 21, 22]. But under our assumptions on Q the deterministic billiard map need not have as strong statistical properties as we show for P. For example, we allow Γ_3 to contain mushroom-shaped curves (see the right-hand side of Figure 3), which is shown in [2] to produce invariant islands in the phase space of the deterministic billiard. In fact, under our assumptions on Q yielding positive spectral gap for P, the deterministic billiard may have arbitrarily slow decay of correlations and not even be ergodic. Compare with [9, 18, 10, 11]. Of course, our Markov operators are in many respects much simpler than similar objects considered in deterministic chaotic billiards (we are, in effect, looking at random 1-dimensional systems), yet they are canonically associated to such billiards. Thus we believe that our operator P is a natural object of study which, on one hand, captures some interesting aspects of the chaotic dynamics of billiard maps while being, on the other hand, technically much simpler and amenable to more explicit analysis compared to related operators of the deterministic theory, such as transfer operators. Another context in which our collision operators may be of interest is in the theory of the Boltzmann equation, where they naturally define boundary conditions that model gas-surface interaction. See, for example, [5].

Further facts of general interest concerning P are described in the subsequent sections, particularly Section 2. The already mentioned technique of conditioning, in particular, provides useful information about P, such as a Lebesgue decomposition (Proposition 7) giving conditions for P to be an integral operator. The technique can be used much more broadly for the purpose of investigating how different geometric features of the billiard cell Q affect the spectral theory of P. We finish this introduction with some examples that illustrate this point and other properties.

1.2 Examples

We now describe a few examples and compare some of the analytical conclusions with numerical observations. The numerical experiments were based on finite rank approximations of the collision operators based on simulation of the billiard dynamics.

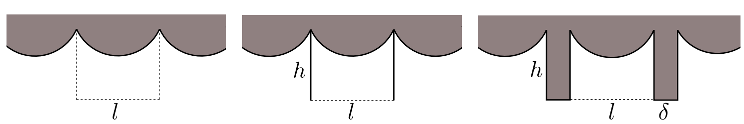

Figure 4: Three examples of microstructures to illustrate the idea of conditioning. Each shows a billiard cell and parts of adjacent cells. The first on the left consists of simple bumps. The dashed vertical lines are immaterial in this case. The second example adds vertical walls of height h to the first example, and in the third example the vertical walls have thickness δ. The explanations are in the text.

A useful technique employed throughout the paper is to consider in addition to P other operators obtained by making P conditional on the event that trajectories satisfy a given property, thus allowing one to focus attention on isolated geometric features. The main purpose of this subsection is to illustrate the use of this idea in estimating spectral gap. The examples all involve relatively simple shapes derived from circular bumps and straight walls.

Consider the billiard cells of Figure 4. (In this discussion, the partition M_1, M_2 and the operators P_1, P_2 are not those of Theorem 2. The general definition of these and related concepts are found in Section 2.) The reflection operator associated to the cell on the right will be denoted $P_{h,\delta}$. This cell consists of the region containing the circular bump and one entire wall. We denote the operator for the middle cell by $P_{h,0}$, and the operator for the cell on the left by P. A common geometric feature of all three cells, and the only feature of that on the left, is the circular bump. We wish to describe $P_{h,\delta}$ and $P_{h,0}$ in terms of P.

We first relate $P_{h,\delta}$ and $P_{h,0}$ for a positive δ. Let M_1 be the set of initial conditions in M corresponding to trajectories that enter the region of the bump, i.e., that do not hit the lower flat bottom of the wall. Now form the partition $M_1, M_2 = M \setminus M_1$ of M. Then α_1 (see the general definition in Subsection 2.1) is constant (independent of θ) and equal to $l/(l+\delta)$, and $\alpha_1 = \mu(M_1)$. Conditional on M_1, reflection is described by $P_{h,0}$, and conditional on M_2 reflection (in terms of our angles convention) is given by the identity operator. Therefore,

$$P_{h,\delta} = \frac{l}{l+\delta} P_{h,0} + \frac{\delta}{l+\delta} I.$$

Since $P_1 := P_{h,0}$ and $P_2 := I$ commute and the α_i are constants, any question concerning the spectrum of $P_{h,\delta}$ immediately reduces to understanding the same question for $P_{h,0}$.

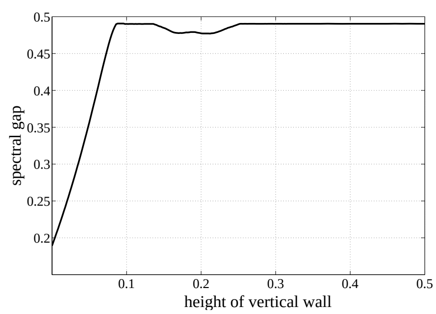

Figure 5: Numerically obtained spectral gap for the example of the middle microstructure of Figure 4 as a function of wall height. The circular bump used is such that the slope of its tangent line at a corner point is 0.4. The length l is set to 1 and the horizontal axis represents h. Observe that the gap grows very fast for small h and suddenly reaches a plateau when h is nearly the height of the circular bump itself.

For example, the spectral gaps of $P_{h,\delta}$ and $P_{h,0}$ are clearly related by

$$\gamma(P_{h,\delta}) = \frac{l}{l+\delta} \gamma(P_{h,0}),$$

and the factor multiplying $\gamma(P_{h,0})$ can be expressed in terms of the normalized curvature of the circular bumps approximately as a constant times $1/K$. Observe that K tends to infinity as l/δ tends to 0 as one expects from Theorem 1. (Assumption 1 does not, strictly, hold in this case, but it would be easy to extend the theorem so as to apply to this example.) We note in passing the similarity between the operator for this class of billiard cells and the well-known Maxwell-Smoluchowsky boundary operator for the Boltzmann equation describing a gas contained by solid walls. The latter operator has the form $(1-a)P_\nu + aI$, where a is the *slip probability* and P_ν is the rank-one operator that maps any probability measure to the stationary (Knudsen's cosine) distribution ν.

We now wish similarly to reduce $P_{h,0}$ to P and understand the effect of adding the vertical walls to the simpler cell on the left side of the figure. In this case, let M_1 consist of the initial conditions for billiard trajectories in the cell corresponding to $P_{h,0}$ for which the exit angle is exactly the same as it would be for trajectories entering the cell on the left with the same conditions. Thus conditional on M_1, $P_{h,0}$ can be replaced with P. And conditional on the complementary event $M_2 := M \setminus M_1$, the operator $P_{h,0}$ is equal to PJ, where J is the unitary involution on functions of the angle induced by the map $\theta \mapsto \pi - \theta$. It will be seen later that P and J commute. (See Subsection 2.4.) Note that for large values of h one expects α_1 and α_2 to be close to 1/2 for most angles, although the analytic expression of these probabilities may be difficulty to describe explicitly.

Thus we have

$$P_{h,0} = \alpha_1 P + (1 - \alpha_1)PJ. \tag{1.1}$$

We can derive from this the following observation about the spectral gap of $P_{h,0}$. We say that the billiard cell associated to P (which may be arbitrary in this discussion) satisfies the *second-odd* condition if there exists a simple eigenvalue λ of P with eigenfunction ψ such that (1) ψ is J*-odd*, i.e., $J\psi = -\psi$, and (2) the spectral radius of P restricted to the orthogonal complement of the span of $\{1, \psi\}$ in $L^2(V, \nu)$ (where ν is the probability measure invariant under P) is strictly less than $|\lambda|$. Although we do not offer a characterization of this property here, we note that [17] gives strong indication that the particular P illustrated in the above figure satisfies it, and our numerical calculations show that the property holds very generally. See Figure 5, which corresponds to the middle microstructure of Figure 4. Under this condition we have the following.

Proposition 2. *If P satisfies the second-odd condition, then by adding walls of height $h > 0$ to the cells defining P, the spectral gap of $P_{h,0}$ cannot be smaller than the gap of P.*

Proof. If ψ is an eigenfunction of P associated to a simple eigenvalue λ, then ψ is either J-even or J-odd. It ψ is J-even, then it is easily checked using 1.1 that ψ is an eigenfunction for $P_{h,0}$ with the same eigenvalue, for all values of h. On the other hand, if ψ is J-odd (the relevant case for the proposition), then $P_{h,0}\psi = \lambda(2\alpha_1 - 1)\psi$. Since the norm of the multiplication operator by $2\alpha_1 - 1$ is no greater than 1, we have the claim. Note that this norm can decrease to 0 as h grows and α_1 is expected to approach 1/2. □

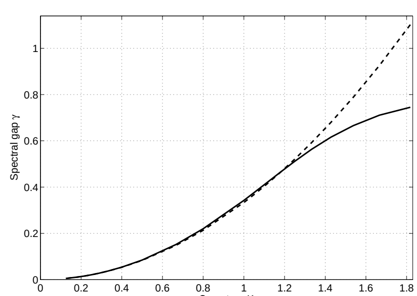

Figure 6: Spectral gap $\gamma(K)$ (solid line) and $\frac{1}{3}K^2$ (dashed line) for the circular bumps cells (left-hand side of Figure 4), for small values of K. By this approximation, in the experiment of Figure 5 for which K is approximately 0.7, we have $\gamma(K)$ slightly below 0.2

A study of the microstructure on the left-hand side of Figure 4 for small values of the normalized curvature was done in [17] using very different techniques and more detailed information was obtained about the spectrum of the Markov operator. Figure 6 contains one observation from that paper regarding the dependence of the spectral gap γ on the normalized curvature for that family of circular bumps. The asymptotic relation $\gamma(K) \sim K^2/3$ for small K was explained there.

Consider now the family of cells shown in Figure 7. The first two from left to right are not (bilaterally) symmetric. If P_+ and P_- denote the respective Markov operators, then

$$P_-^* = P_+ = JP_-J,$$

where J is the unitary involution introduced above. These two operators are associated to that of the fourth cell of Figure 7 conditional on the events that the trajectories enter into the left or the right side of the cell. Therefore,

$$P_0 = \frac{1}{2}\left(P_+ + JP_+J\right),$$

which is now self-adjoint (and, by our theorem, compact).

The relationship between the operators associated to the second and third cells in Figure 7 has already been discussed. It was noted above that the addition of the wall should not cause the spectral gap to decrease. In the present example there is actually little change in the gap. The obtained numerical values are: 0.694 for the spectral gap of P_0 and 0.697 for the spectral gap of P_1.

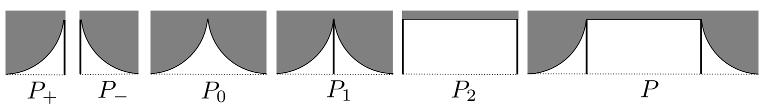

Figure 7: Some billiard cells discussed in the text with their respective collision operators.

Note that P_1 can be written in terms of P_0 as

$$P_1 = \alpha_e P_0 + (1 - \alpha_e)P_0 J$$

where $\alpha_e(\theta) = \lambda\left(M_e \cap (I \times \{\theta\})\right)$ and M_e is the event that the outgoing direction is the same for P_0 and P_1. So

$$M_e = M_0 \cup M_2 \cup M_4 \cup \cdots$$

where M_n is the set of initial conditions of trajectories that bounce off the vertical wall exactly n times before reemerging from the cell.

The left-most cell in Figure 7 can be analyzed in terms of P_1 and P_2, the latter being the rectangular cell second to last in that figure. We have

$$P_2 = \alpha I + (1 - \alpha)J$$

where $\theta \mapsto \alpha(\theta)$ is the probability that a billiard trajectory reemerges from the rectangular cell with the outgoing angle equal to the incoming angle. By our angle conventions, this amounts to specular reflection and is represented by the identity operator. It is not difficulty to show (by a standard "unfolding" argument) that α is given as follows: Let $l = 1 - 2/K$ be the length of the entry/exit segment, $h = 1/K$ the height of the rectangular cell, and write

$$\frac{2h}{l\tan\theta} = k(\theta) + s(\theta), \quad \alpha(\theta) = \begin{cases} s(\theta) & \text{if } k \text{ is odd} \\ 1 - s(\theta) & \text{if } k \text{ is even} \end{cases}$$

where $k(\theta) \in \mathbb{Z}$ is the integer part of $2h/(l\tan\theta)$ and $s(\theta) \in [0, 1)$ is its fractional part. It is clear that the function $\alpha_K = \alpha$ converges uniformly to 1 as K tends to ∞ on any compact subset of $(0, \pi)$. Therefore, P_2 approaches the identity in the strong operator topology of $L^2(V, \nu)$. Now, changing notations slightly to make explicit the dependence on the normalized curvature K, the operator $P_K = P$ can be expressed in terms of the previous ones:

$$P_K = \frac{2}{K} P_1 + \left(1 - \frac{2}{K}\right) P_{2,K}$$

where $\alpha_K := \alpha = k + s$ is the function that enters in the expression of $P_{2,K} = P_2$.

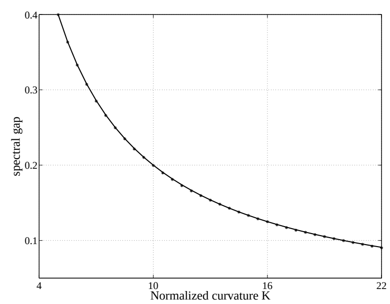

Figure 8: Spectral gap of the operator associated to the cell obtained by taking away from the right-most cell of Figure 7 the vertical walls. The solid line is the graph of $2/K$ and the values marked with an asterisk are numerically obtained.

We observe that essentially the same proof used for part (2) of Theorem 2 gives that P_1 is a Hilbert-Schmidt operator.

Proposition 3. *Let P_K be the Markov operator of the right-most billiard cell in Figure 7. Then the spectral gap of P_K is bounded above by $2/K$.*

Proof. Let $L^2(V, \nu) = H^+ \oplus H^-$ be the orthogonal decomposition of the Hilbert space into the eigenspaces of J. The compact operator P_1 commutes with J, so P_1 has a basis of eigenfunctions adapted to the orthogonal splitting. If ϕ_{even} is any eigenfunction of P_1 in H^+ with eigenvalue λ_{even}, then $P_K \phi_{\text{even}} = (1 - (2/K)(1 - \lambda_{\text{even}}))\phi_{\text{even}}$, so the spectral gap of P_K is no greater than $2/K$. □

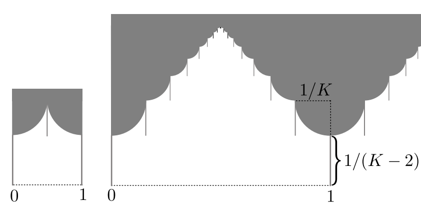

Figure 9: The Markov operator for the cell on right-hand side, for which the curvature parameter K of Γ_1 and Γ_2 can be made arbitrarily large, coincides with the operator for the cell on the left, for which the corresponding parameter is constant, equal to 2. It should be kept in mind when comparing billiard cells that the operators are invariant under homotheties of Q.

Yet another example is given by the cell on the right-hand side of Figure 9. On the left-hand side we have a variant of the bumps with walls. (We have added extra vertical walls to the forth cell of Figure 7, associated to P_1.) Recall that the walls increase the spectral gap of the operator P_1. The cell on the right-hand side is obtained from the first one by scaling (by $(K-2)/K$) and nesting. It is not difficulty to see, using the idea of conditioning, that the Markov operator P of the cell on the right and the operator P_1 of the cell on the left satisfy the recursive relation

$$P = \frac{2}{K} P_1 + \frac{K-2}{K} P. \tag{1.2}$$

Proposition 4. *The two billiard cells depicted in Figure 9 have the same associated Markov operator. In particular, the spectral gap of the cell on the right-hand side of the figure does not depend on K.*

Proof. This follows immediately from the recursive relation 1.2 involving P and P_1. □

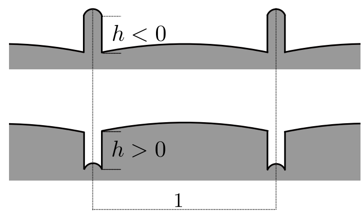

Figure 10: Example of a multi-parameter family of cells for which spectral gap can change very abruptly as a function of one of the parameters. See Figure 11.

There are many other geometric features one would like to explore and many interesting numerical observations for which we cannot as yet provide an analytical explanation. There are also many issues of a general nature to understand; for example, under what conditions is the spectral gap a continuous function of the geometric parameters?

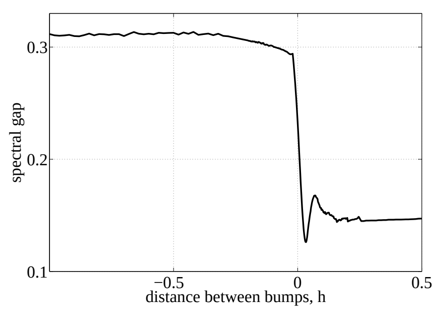

Figure 11: Spectral gap as a function of the relative height of the short and long bumps for the cell of Figure 10. The sharp transition roughly occurs at $h = 0$, when the two bumps are nearly at the same height.

The example of Figure 10 shows, numerically, that the spectral gap can change very quickly for certain values of a parameter. The figure depicts a billiard cell in which there are two competing curvature constants associated to two arc segments. They are placed at different positions specified by another (relative height) parameter h so that one or the other is more exposed to collision. Bigger curvature is expected to cause bigger spectral gap, so we should see a transition from relatively large to relatively small gap as the height parameter changes from large to small (as in figure Figure 10.) The graph of Figure 11 shows that the expected transition is very sharp. One would like to obtain estimates that better quantify and explain phenomena of this kind.

Another among many questions of interest to us concerns the maximum size of the spectral gap. Is it possible to find billiard cells for which $\gamma(P)$ is arbitrarily close to the upper bound 1? The largest value we have obtained so far in our numerical experiments is 0.91.

2 The random billiard Markov chain

This section describes in more detail the Markov process determined by P and how it relates to the return billiard map to the wall Γ_0.

2.1 The reduced billiard map

We review some basic facts concerning billiard maps. For details, see [7]. Let $\Gamma = \partial Q = \bigcup_i \widetilde{\Gamma}_i$ be the decomposition of the boundary of the billiard cell Q into its smooth component curves, or *walls*. The boundary is oriented so that Q lies to the left of each $\widetilde{\Gamma}_j$. Let $\mathbf{n}$ be the unit normal vector field on each $\widetilde{\Gamma}_j$ pointing into Q. Define the *collision space* $\mathcal{M} = \bigcup_i \mathcal{M}_i$, where $\mathcal{M}_i$ is the set of pairs $(q,v) \in \widetilde{\Gamma}_j$ such that $\langle v, \mathbf{n}(q)\rangle \geq 0$. With a little abuse of notation let $\partial\mathcal{M}$ denote the union of the boundaries of the $\mathcal{M}_i$ and $\mathcal{M}^\circ = \mathcal{M} \setminus \partial\mathcal{M}$. We denote by $\mathcal{F} : \mathcal{M} \to \mathcal{M}$ the *billiard map*, whose precise definition can be found in [7]. If $x = (q,v) \in \mathcal{M}^\circ$ and the first intersection q' of the ray $q + tv$, $t > 0$, with ∂Q is not a corner point of ∂Q (i.e., an endpoint of a wall) nor is v tangent to ∂Q at q' (*grazing collision*), then the billiard map is simply $\mathcal{F}(x) = (q', v')$, where v' is the orthogonal reflection of v on the tangent space to ∂Q at q'. Also define $\mathcal{S}_0 = \partial\mathcal{M}$, $\mathcal{S}_{\pm 1} = \mathcal{S}_0 \cup \{x \in \mathcal{M}^\circ : \mathcal{F}^{\pm 1}(x) \notin \mathcal{M}^\circ\}$ and, inductively,

$$\mathcal{S}_{\pm(m+1)} = \mathcal{S}_{\pm m} \cup \mathcal{F}^{\mp m}(\mathcal{S}_{\pm m}).$$

It can be shown that $\mathcal{M}\backslash\mathcal{S}_{\pm 1}$ are open sets and that $\mathcal{F} : \mathcal{M}\backslash\mathcal{S}_1 \to \mathcal{M}\backslash\mathcal{S}_{-1}$ is a smooth diffeomorphism. We refer to $\mathcal{S}_{\pm 1}$ as the *singularity set* of $\mathcal{F}^{\pm 1}$. The set $\widetilde{\mathcal{M}} := \mathcal{M} \setminus \cup_{j=-\infty}^{\infty}\mathcal{S}_j$ is a dense G_δ-subset of $\mathcal{M}$ of full Lebesgue measure on which $\mathcal{F}$ and all its positive and negative iterates are smooth. Furthermore, $\mathcal{F}$ leaves invariant a smooth probability measure on $\mathcal{M}$.

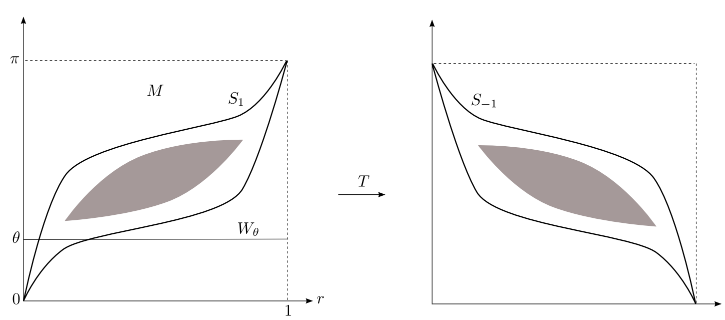

Figure 12: A qualitative description of the singular sets S_+ and S_- of the return map T to M. The shaded region indicates parts of the singular sets which will not enter into the analysis. If the incoming direction of the billiard particle is small enough (close to 0 or π), the shaded regions are not seen.

Let $M = \mathcal{M}_0$ be the subset of $\mathcal{M}$ consisting of (q,v) such that $q \in \Gamma_0$. We denote by μ the probability measure on M obtained by restricting to M (and normalizing) the invariant measure on $\mathcal{M}$. By Poincaré recurrence, there is a subset $E_0 \subset M \cap \widetilde{\mathcal{M}}$ of full μ-measure such that billiard orbits starting in E_0 return to M and are non-singular. Since these orbits return to M in a finite number of steps, there is an open neighborhood in M of each $x \in E_0$ all of whose points return to

M in the same number of steps as the orbit of x, and the return map on this open set is smooth. Therefore, E_0 is actually an open subset of M of full measure on which the first return map to M is well-defined and smooth. We can enlarge E_0 to a set E that also includes singular orbits that eventually still return to M; i.e., possibly grazing orbits that do not get trapped in Q and do not hit a corner point along the way. Note that, due to the standing assumptions on ∂Q (specifically, that Γ_1 and Γ_2 are tangent to Γ_0 and have positive curvature bounded away from 0), the first return map $T : E \to M$ is defined for all vectors that arrive at Γ_0 at angles close enough to 0 or π, and E_0 contains all vectors based at the two endpoints of Γ_0.

Denoting by S_1 the singular set of T and by S_{-1} the singular set of T^{-1} (these are compact subsets of M), then $T : M \setminus S_1 \to M \setminus S_{-1}$ is a smooth diffeomorphism. A more precise description of $S_{\pm 1}$, at least for vectors that are relatively close to tangent vectors on Γ_0, will be given later. For simplicity, we often refer to T as a map on M itself. We call T the *reduced billiard map* on M. Notice that the probability measure μ on M defined as the (normalized) restriction to M of the Liouville measure is T-invariant.

Coordinates on $\mathcal{M}$ and on M can be introduced by representing the base-point of a vector by the arc length parameter r (increasing in the sense defined by the orientation of Γ) and the angle θ that the vector makes with the positive tangent unit vector to the boundary. In particular, we parametrize M by identifying it with $I \times [0, \pi]$. This definition of θ is different than the common choice (for example, [7]) of measuring the angle from $\mathbf{n}$, but it is more convenient for our purposes. In this set of coordinates the measure μ has the form $\mu = \lambda \otimes \nu$, where λ is the Lebesgue measure on $[0, 1]$ and $d\nu(\theta) = \frac{1}{2}\sin(\theta)\, d\theta$.

2.2 Some operators associated to T

Summarizing basic notation: $V = [0, \pi]$, $I = [0, 1]$, $M = I \times V$, $\pi_2 : M \to V$ is the coordinate projection, and ν is the probability measure on V such that $d\nu = \frac{1}{2}\sin(\theta)d\theta$. Let P be the operator on $L^\infty(V, \nu)$ defined by

$$(Pf)(\theta) = \int_I f(\pi_2 \circ T(s, \theta))\, ds.$$

If A is a measurable subset of V, then

$$P(A|\theta) := (P\mathbb{1}_A)(\theta_0)$$

is interpreted as the conditional probability that the angle of reflection after a random collision with a "surface with micro-structure" defined by Q (the surface being parallel to Γ_0) lies in A. This P, which we alternatively refer to as the *Markov operator*, or the *collision operator* of the random billiard, is the transition probabilities operator for Markov chains to be considered shortly.

The same P also naturally acts on signed measures on V as already indicated: if η is a measure on V and $f \in L^\infty(V, \nu)$, then $(\eta P)(f) = \eta(Pf)$, where $\eta(g)$ is the integral of a function g with respect to η. If η is a probability measure representing the distribution of angles of an incoming billiard particle, then ηP is the distribution of angles after the random reflection. Equivalently, it is easily seen that $\eta P = (\pi_2 \circ T)_* \lambda \otimes \eta$, where the lower asterisk indicates the usual push-forward operation on measures.

Observe that $\nu P = (\pi_2)_* T_* \mu = (\pi_2)_* \mu = \nu$, due to T-invariance of μ. Thus ν is P-invariant. In the context of Markov chains with transition probabilities P, we say that ν is a *stationary* probability on V.

By a standard application of Jensen's inequality and the fact that ν is P-invariant, it can be shown that P also defines a bounded operator on $L^2(V, \nu)$ having norm $\|P\|_2 = 1$. Here $L^2(V, \nu)$ is the Hilbert space with inner product $\langle f, g\rangle := \int_V fg\, d\nu$. (We typically consider real-valued functions and so omit the complex conjugate bar; the notation $\overline{g}$ will be used below for the pull-back: $\overline{g} := g \circ \pi_2$.) Invariance of μ under T also gives the following expression for the L^2-dual, P^*, of the operator P:

$$(P^*g)(\theta) = \int_I g(\pi_2 \circ T^{-1}(s,\theta))\, ds.$$

The dual operator is naturally identified with the action of P on measure densities. In fact, let ν_g be, for any $g \in L^2(V,\nu)$, the measure on V such that $d\nu_g = g d\nu$. Then it is immediate from the definitions that for each $f \in L^2(V,\nu)$,

$$(\nu_g P)(f) = \langle Pf, g\rangle = \langle f, P^*g\rangle = \nu_{P^*g}(f).$$

2.3 Markov chains associated to P

We wish to regard P as the transition probabilities operator for Markov chains corresponding to the angle process of successive random collisions. Let $R_0, R_1, \dots$ be a sequence of independent uniformly distributed random variables taking values in $[0,1]$ and let Θ_0 be a random variable taking values in V with probability distribution η. Define a sequence $\Theta_1, \Theta_2, \dots$ of random angles by

$$\Theta_{n+1} = \pi_2(T(R_n, \Theta_n))$$

for $n = 0, 1, \dots$. This is the *angles Markov chain* with initial probability η. It is a simple check that the transition probabilities are given by

$$\text{Prob}(\Theta_{n+1} \in A | \Theta_n = \theta) = P(A|\theta) := (P\mathbb{1}_A)(\theta),$$

and that the probability distribution of Θ_n is given by

$$\text{Prob}(\Theta_n \in A) = (\eta P^n)(\mathbb{1}_A).$$

When $\eta = \nu$, the Markov chain is stationary.

Expressed somewhat differently, let $\Omega = [0,1]^{\mathbb{Z}}$ be equipped with a Borel σ-algebra induced from the product topology and the product measure $\lambda^{\mathbb{Z}}$. We can now think of the random variables Θ_n as measurable functions on Ω as follows. Let $\sigma : \Omega \to \Omega$ denote the shift map $\sigma(\mathbf{r})_i = r_{i+1}$ for each $\mathbf{r} = (\cdots, r_{-1}, r_0, r_1, \cdots)$. Writing $\Psi_{\mathbf{r}} : V \to V$ for the map $\Psi_{\mathbf{r}}(\theta) = \pi_2(T(r_0, \theta))$, then

$$\Theta_n(\mathbf{r}) = \Psi_{\sigma^{n-1}(\mathbf{r})} \circ \cdots \circ \Psi_{\sigma^0(\mathbf{r})}(\Theta_0)$$

is a random variable on Ω with law

$$(\Theta_n)_* \lambda^{\mathbb{Z}} = \eta P^n.$$

As already noted, under our standing assumptions on Γ (or more generally as in [14]) these random billiard Markov chains are irreducible and aperiodic, so ηP^n converges to ν as $n \to \infty$. Our main goal is to better understand this convergence quantitatively in relation to the geometry of Q.

2.4 Symmetries of P

One can relate P and its L^2-adjoint P^* via certain symmetries of the billiard system. Let J and $\mathcal{J}$ be the maps on M defined by $J(r,\theta) = (r, \pi - \theta)$ and $\mathcal{J}(r,\theta) = (1-r, \pi - \theta)$. Notice that $\mathcal{J}$ is the transformation that the map $(x,y) \mapsto (1-x, y)$ of $\mathbb{R}^2$ induces on M regarded as a subset of the tangent bundle $T\mathbb{R}^2$. We say that Q is *symmetric* if it is invariant under $\mathcal{J}$, i.e., if it is mapped to itself under the reflection across the line $x = 1/2$. Although the condition that Q be symmetric is part of our standing assumptions (made mostly for convenience), we temporarily allow Q to be general in order to clarify the meaning of this restriction.

Lemma 1. *The reduced billiard map T satisfies $J \circ T = T^{-1} \circ J$. If Q is symmetric, then $\mathfrak{J} \circ T = T \circ \mathfrak{J}$ also holds. Thus for a symmetric billiard cell $\pi_2 \circ T^{-1}(r, \theta) = \pi_2 \circ T(1 - r, \theta)$.*

Proof. The first relation expresses the fact that T is time reversible, and the second amounts to the geometric observation that billiard trajectories are mapped to billiard trajectories under $\mathfrak{J}$. The third relation is an immediate consequence of the first two. □

We also use J to denote the unitary involution on $L^2(V, \nu)$ defined by $(Jf)(\theta) = f(\pi - \theta)$, and the map on V defined by $J(\theta) = \pi - \theta$. In particular, it makes sense to write $\pi_2 \circ J = J \circ \pi_2$.

Proposition 5. *The L^2-adjoint of P is $P^* = JPJ$. If Q is symmetric, P is self-adjoint.*

Proof. Let $f, g \in L^2(V, \nu)$. Then

$$\begin{aligned}
\langle Pf, g\rangle &= \int_V \left(\int_I f(\pi_2 T(r, \theta))\, d\lambda(r) \right) g(\theta)\, d\nu(\theta) \\
&= \int_M f(\pi_2 T(r, \theta)) g(\pi_2(r, \theta))\, d\mu(r, \theta) \\
&= \int_M f(\theta) g(\pi_2 T^{-1}(r, \theta))\, d\mu(r, \theta) \\
&= \int_M f(\theta) g(J \pi_2 T J(r, \theta))\, d\mu(r, \theta)) \\
&= \int_V f(\theta) \left(\int_I g(J\pi_2 T(r, J(\theta)))\, d\lambda(r) \right) d\nu(\theta) \\
&= \int_V f(\theta)(JPJg)(\theta)\, d\nu(\theta) = \langle f, JPJg\rangle.
\end{aligned}$$

If Q is symmetric, the integral on the second line above becomes

$$\langle Pf, g\rangle = \int_M f(\theta) g(\pi_2 T(1 - r, \theta))\, d\mu(r, \theta) = \langle f, Pg\rangle$$

due to the identity $\pi_2 T^{-1}(r, \theta) = \pi_2 T(1 - r, \theta)$ and that $r \mapsto 1 - r$ preserves Lebesgue measure on the unit interval. □

If follows from the proposition that the operator PJ is self-adjoint and, as P has L^2-norm 1 and J is unitary, the spectrum of PJ in the general case, and of P in the symmetric case, is contained in $[-1, 1]$.

When P is an integral operator of the form $(Pf)(\theta) = \int_V \omega(\theta, \phi) f(\phi)\, d\nu(\phi)$, then the symmetries J and $\mathfrak{J}$ are reflected in the kernel ω as follows: $\omega(\theta, \phi) = \omega(\pi - \phi, \pi - \theta)$ due to time reversibility, and $\omega(\theta, \phi) = \omega(\phi, \theta)$ if Q is symmetric.

2.5 Conditional collision operators

Let $M_1, M_2, \ldots$ form a measurable partition of M. Let $M_j(\theta) := \{r \in I : (r, \theta) \in M_j\}$. Define $\alpha_j(\theta) := \lambda(M_j(\theta))$ for each j and $\theta \in [0, \pi]$. For each $f \in L^\infty(V, \nu)$, define

$$(2.1) \qquad (P_j f)(\theta) = \begin{cases} \frac{1}{\alpha_j(\theta)} \int_{M_j(\theta)} f(\pi_2 T(r, \theta))\, dr & \text{if } \alpha_j(\theta) \neq 0 \\ 0 & \text{if } \alpha_j(\theta) = 0. \end{cases}$$

We call P_j the *conditional operators* associated to the partition M_j. Notice that $P_j(\mathbb{1}_A)(\theta)$ is the conditional probability that the angle of reflection is in $A \subset V$ given that the pre-collision angle is

θ and that the event M_j holds. Let ν_j denote the measure on V such that $d\nu_j = \frac{\alpha_j}{\mu(M_j)} d\nu$. Then ν_j is a conditional measure:

$$\nu_j(A) = \nu(A|M_j) := \mu((I \times A) \cap M_j)/\mu(M_j).$$

It will generally be assumed that the partition is *reversible*: if $(r, \theta) \in M_j$, then $JT(r, \theta) \in M_j$. For example, M_j may represent the set of initial conditions of orbits that avoid collisions with specified walls of ∂Q. Then the orbit traced in reverse would have the same property. In addition, we assume that the partitions are *symmetric*, i.e., each M_j is $\mathcal{J}$-invariant, where $\mathcal{J}$ (defined earlier) is the mirror reflection about the middle line perpendicular to Γ_0. (Recall that $\mathcal{J}(r, \theta) = (1 - r, \pi - \theta)$.) This simplifies the discussion to some extent but is a less fundamental assumption. See Proposition 6. M_j being symmetric implies

$$\alpha_j(\theta) = \alpha_j(\pi - \theta),$$

and reversibility of M_j implies that

$$\lambda(W_{\pi-\theta} \cap M_j) = \lambda(W_\theta \cap TM_j),$$

where $W_\theta = I \times \{\theta\}$. J sends ν_j to $J_*\nu_j$ such that $d(J_*\nu_j) = \mu(M_j)^{-1}(\alpha_j \circ J)\, d\nu$. Let $\|\cdot\|_{\nu_j}$ be the norm of $L^2(V, \nu_j)$ and $\|\cdot\|_\nu$ the norm of $L^2(V, \nu)$. The induced map $J : L^2(V, \nu_j) \to L^2(V, J_*\nu_j)$ is an isometric isomorphism.

Proposition 6. *Let P_j be the conditional operators associated to a reversible partition $M_1, M_2, \ldots$ of M, and ν_j the conditional measures defined above. Then*

1. *$P_j : L^2(V, J_*\nu_j) \to L^2(V, \nu_j)$ has norm $\|P_j\| = 1$;*
2. *$\nu_j P_j = J_*\nu_j$;*
3. *The adjoint of P_j is $P_j^* = JP_jJ$.*

If, in addition, the M_j are symmetric, then each P_j is a self-adjoint operator on $L^2(V, \nu_j)$ of norm $\|P_j\|_{\nu_j} = 1$ commuting with the unitary involution J of $L^2(V, \nu_j)$, and ν_j is P_j-stationary, i.e., $\nu_j P_j = \nu_j$.

Proof. These are straightforward calculations similar to the proof of Proposition 5. We prove the first item to illustrate the main ideas. By Jensen's inequality, $(P_j f)^2 \leq P_j f^2$. Therefore,

$$\begin{aligned}
\|P_j f\|_{\nu_j}^2 \leq \int_V (P_j f^2)(\theta)\, d\nu_j(\theta) &= \frac{1}{\mu(M_j)} \int_V \int_{M_j(\theta)} f^2(\pi_2 T(r, \theta))\, d\lambda(r)\, d\nu(\theta) \\
&= \frac{1}{\mu(M_j)} \int_M \mathbb{1}_{M_j}(r, \theta) f^2(\pi_2 T(r, \theta))\, d\mu(r, \theta) \\
&= \frac{1}{\mu(M_j)} \int_M \mathbb{1}_{TM_j}(r, \theta) f^2(\theta)\, d\mu(r, \theta) \\
&= \frac{1}{\mu(M_j)} \int_V \lambda(W_\theta \cap TM_j) f^2(\theta)\, d\nu(\theta) \\
&= \frac{1}{\mu(M_j)} \int_V \alpha_j(\pi - \theta) f^2(\theta)\, d\nu(\theta) = \|f\|_{J_*\nu_j}^2.
\end{aligned}$$

Since $P_j 1 = 1$, the upper bound is realized. The other items are proved by similar manipulations. $\square$

The reason for introducing the operators P_j is that they provide a useful decomposition of P, as follows. First note that if $f \in L^2(V,\nu)$, it makes sense to write

$$Pf = \sum_j \alpha_j P_j f.$$

By then focusing attention on P_j we may be able to study the effect on P of different geometric features of Q. For example, one may expect, as turns out to be the case, that the exposed bumps formed by $\Gamma_1 \cup \Gamma_2$ are an important factor affecting the mixing properties of the Markov process. So we can isolate the effect they have by conditioning the process on the event that billiard orbits only hit those two walls.

It is necessary to describe this decomposition with a little more care. The inclusion operator of $L^2(V,\nu)$ into $L^2(V,\nu_j)$ will be denoted $\mathcal{I}_j$. This is a bounded operator whose adjoint is multiplication by $\alpha_j/\mu(M_j)$. Multiplication by α_j will be written $\mathcal{M}_{\alpha_j} : L^2(V,\nu_j) \to L^2(V,\nu)$. It is also bounded and $\mathcal{M}^*_{\alpha_j} = \mu(M_j)\mathcal{I}_j$. We can then write

$$P = \sum_j \mu(M_j)\mathcal{I}_j^* P_j\, \mathcal{I}_j, \tag{2.2}$$

which is a sum of bounded self-adjoint operators from $L^2(V,\nu)$ to itself. With a little abuse of notation we often omit the inclusions $\mathcal{I}_j$ and write more suggestively $P = \sum_j \alpha_j P_j$. Each term of the decomposition of P has norm no greater than $\|\alpha_j\|_\infty$.

Recall that Theorem 2 refers to the operator P_1 defined as above for the special partition M_1, M_2 of M such that M_1 is the subset consisting of initial conditions of billiard trajectories that only collide with Γ once, and the collision is at a point in $\Gamma_1 \cup \Gamma_2$. This is clearly a reversible partition.

2.6 A Lebesgue decomposition of P

We apply here the remarks of the previous section to the partition M_1, $M_2 = M \setminus M_1$, where M_1 is the open subset in M consisting of (r,θ) such that $\Psi'_\theta(r) \neq 0$, and $\Psi_\theta(r) := \Psi(r,\theta) := \pi_2 T(r,\theta)$. (The M_1 of the special partition assumed in Theorem 2 is a subset of the present M_1; the conclusions we derive here will apply to that set as well.) Recall that there is an open subset E_0 of full measure in M where T is smooth. For simplicity, we disregard the complement of E_0 and assume that T is also smooth on M_2. A simple exercise in implicit differentiation shows that M_1 is invariant under $J \circ T$, i.e., the partition is reversible.

Then $P = \alpha_1 P_1 + \alpha_2 P_2$ can be interpreted as a Lebesgue decomposition of P into an absolutely continuous part P_1 with respect to ν and a singular part P_2, as shown below. The absolutely continuous part can then be written as an integral operator on $L^2(V,\nu_1)$ of the form

$$(P_1 f)(\theta) = \int_V \omega_1(\theta,\phi) f(\phi)\, d\nu_1(\phi).$$

(See Proposition 7 below.) Much of the analysis later in the paper will be concerned with this absolutely continuous part or, in fact, a "part" of it associated to the reversible subset of M_1 representing trajectories that only hit Γ once and do not hit walls other than Γ_1 and Γ_2.

We need some more notation. Let $W_\theta = I \times \{\theta\}$. We identify W_θ with $[0,1]$ for convenience. Let $W^i_\theta := \{r \in [0,1] : (r,\theta) \in M_i\}$. Then $W^1_\theta \cup W^2_\theta$ is an open set of full Lebesgue measure in I and W^1_θ is a countable (or finite) union of open intervals $W_{\theta,j}$. Moreover, $\Psi'_\theta(r) = 0$ for all r in W^2_θ. The restriction $\Psi_{\theta,j} := \Psi_\theta|_{W_{\theta,j}}$ is a diffeomorphism from $W_{\theta,j}$ to its image $V_{\theta,j} \subset V$.

Under the standing assumptions on Γ, $W_\theta \cap M_1$ has measure $\alpha_1(\theta) = \lambda(W_\theta \cap M_1) > 0$ for all $\theta \in (0,\pi)$, and $\alpha_1(\theta) = 1$ for all θ close enough to 0 or π. The set $W_\theta \cap M_2$ can have positive measure if, for example, ∂Q contains *exposed* flat walls, i.e., if there is a billiard orbit that collides only with flat walls of Q. The cell on the left-hand side of Figure 3 is an example. In this case, the measure $P(\cdot|\theta)$ on V has atoms. (In the case of polygonal billiard cells, P becomes purely atomic.)

We now indicate how the decomposition of P associated to the reversible partition M_1, M_2 can be interpreted as a Lebesgue decomposition. Define $\Gamma_\theta(\varphi) = \sum_j \mathbb{1}_{V_{\theta,j}}(\varphi)\Lambda_{\theta,j}(\varphi)^{-1}$, where

$$\Lambda_{\theta,j}(\varphi) = \frac{1}{2}\left|\Psi'_\theta(\Psi^{-1}_{\theta,j}(\varphi))\right|\sin\varphi$$

and $\mathbb{1}_{V_{\theta,j}}$ denotes the indicator function of $V_{\theta,j}$.

Proposition 7. *Let M_1, M_2 be the reversible partition defined above in this section and set*

$$\omega_1(\theta,\varphi) := \mu(M_1)\Gamma_\theta(\varphi)/(\alpha_1(\theta)\alpha_1(\varphi)).$$

Then the operator P_1 on $L^2(V,\nu_1)$ is given by

$$(P_1 f)(\theta) = \int_V \omega_1(\theta,\varphi)f(\varphi)\,d\nu_1(\varphi)$$

and P_2 is singular with respect to ν in the sense that $P_2(\cdot|\theta)$ and ν are singular measures for each θ.

Proof. Let $A \subset V$ be measurable and $A_{\theta,j} = \{r \in W_{\theta,j} : \Psi_{\theta,j}(r) \in A\}$, so $\Psi_{\theta,j}(A_{\theta,j}) = A \cap V_{\theta,j}$. Then, as $\Psi_{\theta,j} : W_{\theta,j} \to V_{\theta,j}$ is a diffeomorphism,

$$\begin{aligned}
\int_A \Gamma_\theta(\varphi)\,d\nu(\varphi) &= \sum_j \int_{A\cap V_{\theta,j}} \Lambda_{\theta,j}(\varphi)^{-1}\,d\nu(\varphi) \\
&= \sum_j \int_{\Psi_{\theta,j}(A_{\theta,j})} \left|\Psi'_\theta\left(\Psi^{-1}_{\theta,j}(\varphi)\right)\right|^{-1} d\varphi \\
&= \sum_j \int_{A_{\theta,j}} d\lambda(r) \\
&= \lambda(\{r \in W^1_\theta : \Psi_\theta(r) \in A\}).
\end{aligned}$$

But $\lambda(\{r \in W^1_\theta : \Psi_\theta(r) \in A\}) = P(A|\theta) - \eta^\perp_\theta(A)$, where $\eta^\perp_\theta(A) := \lambda(\{r \in W^2_\theta : \Psi_\theta(r) \in A\})$. So

$$P(A|\theta) = \int_A \Gamma_\theta(\varphi)\,d\nu(\varphi) + \eta^\perp_\theta(A).$$

The measure $\eta^\perp_\theta$ is singular with respect to ν, due to Sard's theorem, so the above indeed corresponds to a Lebesgue decomposition of $P(\cdot|\theta)$, for each θ. The proposition now follows from the definitions. □

As already noted, we later work not with the full absolutely continuous part of P above, but with a part of it that is easier to identify dynamically, defined by the M_1 of the special partition of M. The kernel of the integral operator P_1 then involves sums of the functions $\Lambda_{\theta,j}(\varphi)^{-1}$ over a subset of indices j. In this special case, the kernel of P_1 can be given very explicitly, as shown next.

3 Estimating transition probabilities and densities

In this section we derive geometric estimates (in terms of the curvature K) for various transition probabilities and probability densities associated to P_1 and its powers.

3.1 A more detailed look at the function Ψ_θ

Recall that $\Theta = \Psi_\theta(r)$ is the angle of the outgoing direction (through Γ_0) of a billiard particle that enters Q through Γ_0 with initial state (r, θ). We need to understand this function in more detail. We are mainly interested in $\Psi_\theta(r)$ for the values of r such that (r, θ) lies in the set M_1 of the special partition of M.

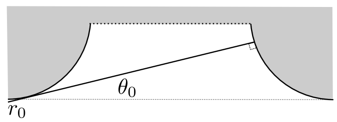

Figure 13: Definition of (r_0, θ_0).

Due to the symmetry of M_1 the angle function satisfies $\Psi_\theta(r) = \pi - \Psi_{\pi-\theta}(1-r)$. Therefore, it is sufficient to study Ψ_θ for $0 \leq \theta \leq \pi/2$. The form of Ψ_θ is especially simple for *small values* of θ as will be seen in a moment. More precisely, define θ_0 as the angle in $[0, \pi/2]$ such that

$$\sin\theta_0 = \frac{1}{K}.$$

This angle is characterized by the property that a billiard particle with initial condition (r_0, θ_0), where

$$r_0 = 1 - \sqrt{1 - K^{-2}},$$

will graze Γ_1, bounce off Γ_2, and return to Γ_0 exactly at r_0 with angle $\pi - \theta_0$. (See Figure 13.) The study of Ψ_θ will be divided into the cases $\theta \leq \theta_0$ and $\theta > \theta_0$. The expression *small* θ will refer to the former case.

It is not difficult to determine the overall qualitative features of the graph of Ψ_θ, which is sketched in Figure 14. The left-hand side of the figure shows the small θ case and the right-hand side suggests where further complications may lie when θ is not small. The figure also introduces some notation that is used later. These graphs may be compared with Figure 12, where the shaded region also indicates the region containing the rest of the singular set.

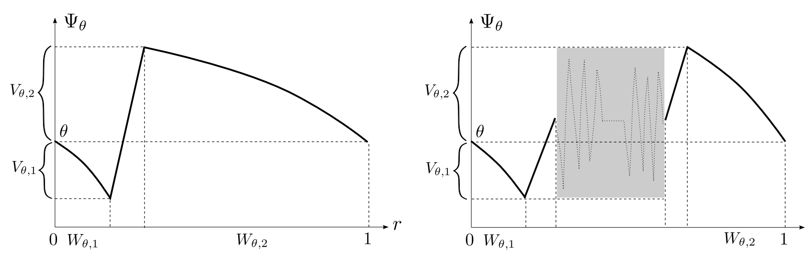

Figure 14: On the left is a sketch of the graph of Ψ_θ for relatively small values of θ. For a general θ something more complicated may happen somewhere in the middle of the interval $[0, 1]$, as suggested by the shaded box. In this section we obtain more detailed quantitative information about these graphs in terms of the curvature of the sides Γ_1 and Γ_2.

It is necessary to estimate the lengths of the intervals $V_{\theta,i}$, $W_{\theta,i}$ shown in Figure 14 in terms of the normalized curvature parameter K of the sides Γ_1 and Γ_2. The endpoints of these intervals will be denoted as follows (see also Figure 15):

$$W_{\theta,1} = [0, r_{\theta,1}],\ W_{\theta,2} = [r_{\theta,2}, 1],\ V_{\theta,1} = [\Theta_{\theta,1}, \theta],\ V_{\theta,2} = [\theta, \Theta_{\theta,2}].$$

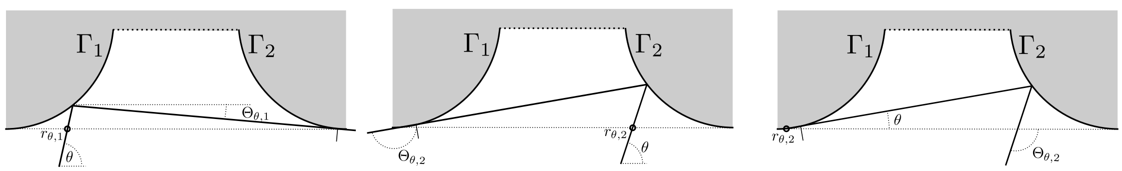

Figure 15: The numbers $r_{\theta,i}$ and $\Theta_{\theta,i}$ are endpoints of the intervals $W_{\theta,i}$ and $V_{\theta,i}$ as indicated in the text. They are obtained by regarding the trajectories that have a grazing second or first collision with $\Gamma_1 \cup \Gamma_2$.

The needed length estimates will be derived from the following proposition.

Proposition 8. *Suppose that Q satisfies Assumption 1. Let $(r, \theta) \in M$, $\theta \leq \pi/2$, be the initial state of a billiard trajectory that collides with ∂Q at most twice (not counting as collisions the initial and final states on Γ_0), and all collisions happen on $\Gamma_1 \cup \Gamma_2$. Suppose that one collision is regular and a second, if it occurs, is grazing. Set $\Theta = \Psi_\theta(r)$ and let (s, Θ) be the state at which the particle returns to Γ_0. Then the following hold:*

I *First (regular) collision with Γ_1 and possibly a second (grazing) collision with Γ_2. In this case,*

$$K r \sin\theta = \cos\left(\frac{\theta + \Theta}{2}\right) - \cos\theta. \tag{3.1}$$

If a second collision does occur and is grazing (see the graph on the left of Figure 15), then

$$K \sin\Theta = 1 - \cos\left(\frac{\theta + \Theta}{2}\right). \tag{3.2}$$

II *First (regular) collision with Γ_2, a possibly second (grazing) collision with Γ_1. In this case,*

$$K(1 - r)\sin\theta = \cos\theta - \cos\left(\frac{\theta + \Theta}{2}\right). \tag{3.3}$$

If a second collision does occur and is grazing (see the graph on the middle of Figure 15), then

$$K \sin(\Theta) = 1 + \cos\left(\frac{\theta + \Theta}{2}\right). \tag{3.4}$$

III *If the first collision (with Γ_1) is grazing and the second (with Γ_2) is regular (see the graph on the right of Figure 15; this may happen for $\theta \leq \theta_0$), then*

$$K r \sin\theta = 1 - \cos\theta \text{ and } \cos\left(\frac{\theta + \Theta}{2}\right) = 1 - K\sin\theta. \tag{3.5}$$

Proof. This is proved by elementary geometry and trigonometry. We omit the details. □

Corollary 1. *Under the assumptions of Proposition 8,*

$$\Psi_\theta'(r) = -\frac{2K\sin\theta}{\sin\left(\frac{\theta+\Theta}{2}\right)}. \tag{3.6}$$

Proof. This follows from implicit partial differentiation with respect to r of Equations 3.1 and 3.3 of Proposition 8. □

We now refine our description of Ψ_θ with further estimates of $\Theta_{\theta,i}$ in terms of K and θ. Keep in mind throughout that we are here concerned with large K. Unless otherwise stated, we assume that K is at least 2. We also typically assume in these estimates that $\theta \leq \pi/2$ and deal with other values by symmetry when needed.

For $\theta \leq \theta_0$ defined the map $\sigma(\theta) := \theta^2/8K$. The inverse of σ is $\sigma^{-1}(\theta) = \sqrt{8K\theta}$, which makes sense if θ is sufficiently small, say, less than $1/8K^3$. The next proposition states that, for θ small, the range of values of Ψ_θ closely corresponds to the interval $[\sigma(\theta), \sigma^{-1}(\theta)]$.

Proposition 9. *Suppose that $K \geq 2$. If $\theta \leq \theta_0$, then*

$$\left(1 - \frac{\theta^2}{48}\right)\sigma(\theta) \leq \Theta_{\theta,1} \leq (1+\theta)\,\sigma(\theta). \tag{3.7}$$

If $\theta \leq \sigma(\theta_0)$, then

$$\left(1 - \frac{\theta^2}{12}\right)\sigma^{-1}(\theta) - \theta \leq \Theta_{\theta,2} \leq \left(1 + \frac{K\theta}{9}\right)\sigma^{-1}(\theta) - \theta. \tag{3.8}$$

Let now $\theta_0 < \theta \leq \pi/2$. Then

$$\Theta_{\theta,1} + \pi - \Theta_{\theta,2} \leq 0.73K^{-1}. \tag{3.9}$$

Proof. From Proposition 8 we obtain (recall that $\Theta_{\theta,1} < \theta$)

$$K\Theta_{\theta,1} \geq K\sin\Theta_{\theta,1} = 1 - \cos\left(\frac{\theta + \Theta_{\theta,1}}{2}\right) \geq 1 - \cos(\theta/2)$$

so that

$$\Theta_{\theta,1} \geq \frac{1-\cos(\theta/2)}{K} = \frac{8(1-\cos(\theta/2))}{\theta^2}\sigma(\theta) \geq \left(1 - \frac{\theta^2}{48}\right)\sigma(\theta).$$

A similar kind of approximation gives the upper bound for $\Theta_{\theta,1}$ and the upper and lower bounds for $\Theta_{\theta,2}$, where the numerical coefficients depend on the assumption $K \geq 2$. Note that for the bounds for $\Theta_{\theta,2}$ one starts from the identity $\Theta_{\theta,2} = 2\arccos(1 - K\sin\theta) - \theta$.

Now suppose that $\theta_0 \leq \theta \leq \pi/2$ and let Θ stand for either $\Theta_{\theta,1}$ or $\pi - \Theta_{\theta,2}$. Observe that $\Theta \leq \overline{\Theta} := \Theta_{\pi/2,1}$. Then, by Proposition 8,

$$\Theta \leq \frac{\overline{\Theta}}{K\sin\Theta}\left[1 - \cos\left(\frac{\pi}{4} + \frac{\overline{\Theta}}{2}\right)\right] \tag{3.10}$$

Iterating the inequality $\overline{\Theta} \leq \arcsin((1-\cos(\pi/4+\overline{\Theta}/2))/K)$, with the first step, say, $\overline{\Theta} \leq \pi/2$, yields for $K \geq 2$ that $\overline{\Theta} < 0.18$ and $\overline{\Theta}/\sin\overline{\Theta} < 1.006$. Then 3.9 follows from $\Theta_{\theta,1} + \pi - \Theta_{\theta,2} \leq 2\overline{\Theta}$ and the inequality 3.10. □

3.2 The kernel of P_1 for the special partition

Let $V_\theta = V_{\theta,1} \cup V_{\theta,2}$. From Proposition 7 and the general discussion surrounding it, and Corollary 1, the integral kernel of the operator P_1 can be written as follows (here we use the special partition of M and, accordingly, the α_1 of Proposition 10):

$$\omega_1(\theta, \Theta) = \frac{\mu(M_1)\sin\left(\frac{\theta+\Theta}{2}\right)\mathbb{1}_{V_\theta}(\Theta)}{K\alpha_1(\theta)\alpha_1(\Theta)\sin\theta\sin\Theta} \tag{3.11}$$

and we have, for $f \in L^2(V,\nu_1)$, that

$$(P_1 f)(\theta) = \int_V \omega_1(\theta,\Theta) f(\Theta)\, d\nu_1(\Theta) = \frac{1}{2K\alpha_1(\theta)\sin\theta}\int_{V_\theta} \sin\left(\frac{\theta+\Theta}{2}\right) f(\Theta)\, d\Theta. \tag{3.12}$$

It is not difficulty to see that $\omega_1(\theta,\Theta) = \omega_1(\Theta,\theta)$.

From the calculations of the previous section we obtain the following estimate of $\alpha_1(\theta)$. Recall that θ_0 is defined by the equation $\sin\theta_0 = 1/K$.

Proposition 10. *Suppose $K \geq 2$ and let $\alpha_1(\theta)$ be associated to the special partition of M.*

1. *If $\theta \leq \theta_0$, then*
$$1 - \left(1 + 1.15K^{-1}\right)\frac{\theta}{8K} \leq \alpha_1(\theta) \leq 1.$$

2. *If $\theta_0 \leq \theta \leq \pi/2$, then*
$$\left(1 - 0.37K^{-1}\right)\frac{\sqrt{1+\sin\theta}}{K\sin\theta} \leq \alpha_1(\theta) \leq \frac{\sqrt{1+\sin\theta}}{K\sin\theta}.$$

Proof. First suppose $\theta \leq \theta_0$. Proposition 8 implies

$$\alpha_1(\theta) = r_{\theta,1} + 1 - r_{\theta,2} = 1 - \frac{1-\cos\left(\frac{\theta+\Theta_{\theta,1}}{2}\right)}{K\sin\theta} = 1 - \frac{\sin\Theta_{\theta,1}}{\sin\theta} \geq 1 - \frac{\Theta_{\theta,1}}{\sin\theta}.$$

Now by Proposition 9,

$$\alpha_1(\theta) \geq 1 - \frac{\theta(1+\theta)}{\sin\theta}\frac{\theta}{8K} \geq 1 - (1 + 1.15/K)\frac{\theta}{8K}.$$

Now let $\theta_0 \leq \theta \leq \pi/2$. Proposition 8 implies

$$\alpha_1(\theta) = r_{\theta,1} + 1 - r_{\theta,2} = \frac{\cos\left(\frac{\theta+\Theta_{\theta,1}}{2}\right) - \cos\left(\frac{\theta+\Theta_{\theta,2}}{2}\right)}{K\sin\theta} = \frac{\cos\left(\frac{\theta+\Theta_{\theta,1}}{2}\right) + \sin\left(\frac{\theta-(\pi-\Theta_{\theta,2})}{2}\right)}{K\sin\theta}.$$

As both $\Theta_{\theta,1}$ and $\pi - \Theta_{\theta,2}$ are less than $\pi/2$ we conclude

$$\alpha_1(\theta) \leq \frac{\cos(\theta/2) + \sin(\theta/2)}{K\sin\theta} = \frac{\sqrt{1+\sin\theta}}{K\sin\theta}.$$

A similar argument gives the lower bound

$$\alpha_1(\theta) \geq \left(1 - \frac{\Theta_{\theta,1} + \pi - \Theta_{\theta,2}}{2}\right)\frac{\sqrt{1+\sin\theta}}{K\sin\theta}.$$

The claim in part 2 then follows from 3.9 of Proposition 9. □

Let $\rho_\theta(\varphi) := d(\delta_\theta P_1)/d\varphi$ be the density with respect to the Lebesgue measure $d\varphi$ of the probability $\delta_\theta P_1$. Then,

$$\rho_\theta(\varphi) = \sin\left(\frac{\theta+\varphi}{2}\right)\left[2K\alpha_1(\theta)\sin\theta\right]^{-1}\mathbb{1}_{V_\theta}(\varphi). \tag{3.13}$$

This follows from the general expression of the integral kernel $\omega_1(\theta,\varphi)$. It is convenient to introduce the function $\tilde\rho_\theta(\varphi) := \sin\left((\theta+\varphi)/2\right)/(2K\sin\theta)$ so that $\rho_\theta(\varphi) = \tilde\rho_\theta(\varphi)\frac{\mathbb{1}_\theta(\varphi)}{\alpha_1(\varphi)}$.

Proposition 11. *We suppose that $K \geq 2$, $\theta \leq \pi/2$.*

1. *Let $\theta \leq \theta_0$ and $\varphi \in V_\theta$. Then*

$$(4K)^{-1} \leq \tilde{\rho}_\theta(\varphi) \leq \rho_\theta(\varphi) \leq (1 + 0.21\theta/K)\,\tilde{\rho}_\theta(\varphi).$$

   If in addition $\varphi \in V_{\theta,1}$, then $\tilde{\rho}_\theta(\varphi) \leq (2K)^{-1}$ and $P_1(V_{\theta,1}|\theta) \leq (1 + 0.43K^{-2})3\theta(8K)^{-1}$.

2. *Now suppose that $\theta \geq \theta_0$. Then*

$$\rho_\theta(\varphi) \geq c_\varphi := \frac{1}{2}\min\left\{\sin\frac{\varphi}{2}, \cos\frac{\varphi}{2}\right\}.$$

Proof. These are all straightforwardly derived from Proposition 10. For part 2, observe:

$$\rho_\theta(\varphi) = \frac{\tilde{\rho}_\theta(\varphi)}{\alpha_1(\theta)} \geq \frac{\sin\left(\frac{\theta+\varphi}{2}\right)}{2(1+\sin\theta)^{1/2}} = \frac{1}{2}\frac{\sin\frac{\theta}{2}\cos\frac{\varphi}{2} + \cos\frac{\theta}{2}\sin\frac{\varphi}{2}}{\sin\frac{\theta}{2} + \cos\frac{\theta}{2}} \geq c_\varphi$$

□

3.3 Upper bound for the action of P_1 on probability densities

We need further notation before continuing. Recall the map $\sigma(\varphi) = \varphi^2/8K$. Define

$$a_n = \sigma^n(1/K) = 8K\left(8K^2\right)^{-2^n}$$

for $n = 0, 1, \ldots$ and $A_n = [a_n, \pi - a_n]$. Recall that $\sin\theta_0 = 1/K$, so a_0 is approximately θ_0.

We are interested in the action of powers of P_1 on probability density functions g with respect to the Lebesgue measure m on $[0,\pi]$. Let m_g be the measure on $[0,\pi]$ having density function g. Let P_1^* be the Markov operator on densities so, by definition, $P_1^* g$ is the density of the measure $m_g P_1$. It follows from the definition of $\Theta_{\theta,i}$ that $\varphi \in V_\theta = [\Theta_{\theta,1}, \Theta_{\theta,2}]$ if and only if $\theta \in V_\varphi$. With this in mind one obtains

$$(P_1^* g)(\theta) = \frac{1}{2K}\int_{V_\theta} \sin\left(\frac{\varphi+\theta}{2}\right)[\alpha_1(\varphi)\sin\varphi]^{-1} g(\varphi)\,d\varphi.$$

Proposition 12. *Let $K \geq 2$, $\theta \leq \pi/2$, and g a probability density on $[0,\pi]$ with respect to m. If $\theta \leq \Theta_{a_0,1}$,*

$$(P_1^* g)(\theta) \leq 0.93\frac{\|g|_{V_\theta \cap A_0^c}\|_\infty}{K}\sigma^{-1}(\theta)$$

and if $\theta > \Theta_{a_0,1}$,

$$(P_1^* g)(\theta) \leq 1.45\frac{\|g|_{V_\theta \cap A_0^c}\|_\infty}{K} + 0.62$$

where we set $\|g|_{V_\theta \cap A_0^c}\|_\infty = 0$ if $V_\theta \cap A_0^c$ is empty. In particular,

$$\|P_1^* g\|_\infty \leq 1.45\frac{\|g\|_\infty}{K} + 0.62.$$

The value $0.99a_1$ can be substituted for $\Theta_{a_0,1}$ in the above conditions on θ.

Proof. We write $(P_1^* g)(\theta) = \mathcal{I}_1 + \mathcal{I}_2$, where

$$\mathcal{I}_1 = \frac{1}{2K}\int_{V_\theta\cap A_0} \frac{\sin\left(\frac{\varphi+\theta}{2}\right)}{\alpha_1(\varphi)\sin\varphi} g(\varphi)\,d\varphi, \quad \mathcal{I}_2 = \frac{1}{2K}\int_{V_\theta\cap A_0^c} \frac{\sin\left(\frac{\varphi+\theta}{2}\right)}{\alpha_1(\varphi)\sin\varphi} g(\varphi)\,d\varphi.$$

Using Proposition 10, $\sin((\theta+\varphi)/2)/\sqrt{1+\sin\varphi} \leq 1$, and $K \geq 2$, we obtain

$$\mathcal{I}_1 \leq 0.62 m_g(V_\theta \cap A_0), \quad \mathcal{I}_2 \leq 0.53 \|g|_{V_\theta\cap A_0^c}\|_\infty K^{-1}\mathcal{I}_\theta,$$

where $\mathcal{I}_\theta = \int_{V_\theta\cap A_0^c} \sin((\varphi+\theta)/2)/\sin\varphi\, d\varphi$. We now estimate $\mathcal{I}_\theta$ by writing $\mathcal{I}_\theta = \mathcal{J}_1 + \mathcal{J}_2$, where $\mathcal{J}_1$ is the integral of the argument of $\mathcal{I}_\theta$ over $V_{\theta,1}\cap A_0^c$ and $\mathcal{J}_2$ is the integral of that same function over $V_{\theta,2}\cap A_0^c$. As $\varphi+\theta \leq 2\theta$ on $V_{\theta,1}$ and $\varphi + \theta \leq 2\varphi$ on $V_{\theta,2}$, then

$$\mathcal{J}_1 \leq \sin\theta \int_{V_{\theta,1}\cap A_0^c} \frac{d\varphi}{\sin\varphi}, \quad \mathcal{J}_2 \leq m(V_{\theta,2}\cap A_0^c).$$

A straightforward computation, which uses the approximation $0 \leq -x\ln x < 0.74\sqrt{x}$, $0 < x < 1$, shows that $\mathcal{J}_1 \leq c_1 \sin\theta[(\theta\wedge a_0)/(\Theta_{\theta,1}\wedge a_0)]^{1/2}$, where the wedge symbol represents the minimum of the two numbers and c_1 can be taken to be 0.78 if $\theta \leq a_0$ and 0.74 if $\theta \leq a_1$. (These numbers arise from multiplying 0.74 by an upper-bound for $\varphi/\sin\varphi$ over $\varphi \leq a_0$ and $\varphi \leq a_1$, respectively.)

If $\theta \leq a_0$ one shows that $\sin\theta[(\theta\wedge a_0)/(\Theta_{\theta,1}\wedge a_0)]^{1/2} \leq \theta^{3/2}/\Theta_{\theta,1}^{1/2} \leq c_2\sigma^{-1}(\theta)$, where c_2 can be taken, based on Proposition 9, to equal 1.003 for $\theta \leq a_0$ and 1.00001 for $\theta \leq a_1$. Thus $\mathcal{J}_1 \leq c\sigma^{-1}(\theta)$, where $c = c_1c_2$, if $\theta \leq a_0$. Also observe that $m(V_{\theta,2}\cap A_0) \leq \Theta_{\theta,2}\wedge a_0 \leq 1.004\sigma^{-1}(\theta)\wedge a_0$, where we have used Proposition 9 and $K \geq 2$. The first inequality of the proposition now follows, in which the number 0.93 is seen to come from $0.53(1.004 + 0.74)$. The cases $\Theta_{\theta,1} \leq a_0 \leq \theta$ and $\Theta_{\theta,1} \geq a_0$ are similarly treated, leading to the general expression

$$(P_1^* g)(\theta) \leq 0.62 m_g(V_\theta\cap A_0) + s_\theta \frac{\|g|_{V_\theta\cap A_0^c}\|_\infty}{K}$$

where $s_\theta = 1.45$ for $\theta \geq a_1$ and $s_\theta = 0.93\sigma^{-1}(\theta)$ for $\theta \leq a_1$. The proposition results immediately from this inequality. □

3.4 Lower bounds for transitions to middle intervals

Let $P_1(A|\theta) := (\delta_\theta P_1)(A)$, representing the probability that the scattered angle lies in A given that the angle of incidence is θ and that the event M_1 holds. Similarly, define $P_1^n(A|\theta) := (\delta_\theta P_1^n)(A)$ for all $n \geq 1$. The following notation is used: $Z_\varphi := [\varphi, \pi - \varphi]$, for some $\varphi \in [0, \pi/2]$, and $Z_0 := Z_{\theta_0}$. Z_φ^c denotes the complement in $[0,\pi]$.

Lemma 2. *We assume that $K \geq 2$, θ, φ both lie in $[0, \pi/2]$, and that $\varphi \leq \theta_0$.*

1. *If $\theta \in Z_0$, then*

$$P_1(Z_\varphi^c|\theta) \leq \left(1 + 0.38K^{-1}\right)\frac{\varphi\left(1+\varphi\right)^{1/2}}{2} \leq 0.75\varphi.$$

2. *If $\theta \in Z_0^c$, then*

$$P_1(Z_\varphi^c|\theta) \leq \left(1 + 0.43K^{-2}\right)\frac{\varphi\left(1+\frac{\varphi}{\theta}\right)}{4K}.$$

Proof. Keeping in mind the expression 3.13, note that

$$P_1(Z_\varphi^c|\theta) = \int_0^\varphi [\rho_\theta(\psi) + \rho_\theta(\pi-\psi)]\, d\psi \leq \frac{1}{2K\alpha_1(\theta)\sin\theta}\int_0^\varphi \left[\sin\left(\frac{\theta+\psi}{2}\right) + \cos\left(\frac{\theta-\psi}{2}\right)\right] d\psi.$$

It follows from the identity

$$\sin\left(\frac{\theta+\psi}{2}\right) + \cos\left(\frac{\theta-\psi}{2}\right) = (1+\sin\theta)^{1/2}(1+\sin\psi)^{1/2}$$

and Proposition 10 that

$$P_1(Z_\varphi^c|\theta) \leq \frac{\varphi\,(1+\sin\varphi)^{1/2}(1+\sin\theta)^{1/2}}{2K\alpha_1(\theta)\sin\theta} \leq (1+0.38/K)\,\frac{\varphi(1+\sin\varphi)^{1/2}}{2}$$

as claimed. For part 2, observe that $\Theta_{\varphi,2} \leq \Theta_{\theta_0,2} = \pi - \theta_0 \leq \pi - \varphi$. Therefore,

$$P_1(Z_\varphi^c|\theta) = \int_0^\varphi \rho_\theta(\psi)\, d\psi \leq \frac{\varphi\sin\left(\frac{\theta+\phi}{2}\right)}{2K\sin\theta\alpha_1(\theta)}.$$

Using the estimates of $\alpha_1(\theta)$ from Proposition 10 (recall that $\alpha_1(\theta) := \lambda(M_1(\theta))$) and assuming $K \geq 2$, which is used when simplifying expressions such as $\theta/\sin\theta$ and taking the reciprocal of $\alpha_1(\theta)$), we can obtain 2. ◻

Let $\Theta_0, \Theta_1, \ldots$ be the Markov chain for P_1 with initial state $\Theta_0 = \theta$. The next proposition quantifies the tendency of the Markov chain to move points in $(0,\pi)$ towards the middle range of angles.

Proposition 13. *We suppose that $K \geq 2$. Let $\varphi \leq \theta_0$ and $\theta \in Z_{\sigma(\varphi)}$ be fixed. Then*

$$P_1^{n+2}(Z_\varphi^c|\theta) \leq \frac{1.77 + 0.10n\varphi^2}{K}.$$

for all $n \geq 1$.

Proof. We prove the proposition by induction. Let $\Theta_0, \Theta_1, \ldots$ be the Markov chain for P_1 with initial state $\Theta_0 = \theta \in Z_{\sigma(\varphi)}$. For the induction step we need to estimate

$$\mathcal{I}_\theta := P_1(\Theta_2 \in Z_\varphi^c|\Theta_0 = \theta) = \int_0^\pi \rho_\theta(\theta')P_1(Z_\varphi^c|\theta')\, d\theta'.$$

Assume first that $\sigma(\varphi) \leq \theta \leq \varphi$ and write $\mathcal{I}_\theta = \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3$, where

$$\mathcal{I}_1 = \int_0^{\varphi^2/2} \rho_\theta(\theta')P_1(Z_\varphi^c|\theta')\, d\theta',\ \mathcal{I}_2 = \int_{\varphi^2/2}^{\varphi} \rho_\theta(\theta')P_1(Z_\varphi^c|\theta')\, d\theta',\ \mathcal{I}_3 = \int_\varphi^\pi \rho_\theta(\theta')P_1(Z_\varphi^c|\theta')\, d\theta'.$$

Transition probabilities involved in these and other integrals below can be derived from Lemma 2. The following, in particular, will be needed:

(TP1) If $\theta \in Z_\varphi$, then $P_1(Z_\varphi^c|\theta) \leq 0.75\varphi$ and $P_1(Z_{\varphi^2/2}^c|\theta) \leq 0.38\varphi^2$

(TP2) If $\theta \in Z_{\varphi^2/2} \setminus Z_\varphi$, then $P_1(Z_\varphi^c|\theta) \leq 0.70K^{-1}$

(TP3) If $\theta \in Z_{\sigma(\varphi)} \setminus Z_\varphi$, then $P_1(Z_{\varphi^2/2}^c|\theta) \leq 0.63\varphi^2$

(TP4) If $\theta \in Z_{\sigma(\varphi)}$, then $P_1(Z^c_{\sigma(\theta)}|\theta) \leq 0.10\varphi^2/K$.

Then using TP3 above,

$$\mathfrak{I}_1 \leq \int_0^{\varphi/2} \rho_\theta(\theta')\, d\varphi \leq P_1([0, \varphi^2/2]|\theta) \leq P_1(Z^c_{\varphi^2/2}|\theta) \leq 0.63\varphi^2.$$

Using TP2 we have $\mathfrak{I}_2 \leq (0.70/K)P_1([\varphi^2/2, \varphi]|\theta) \leq 0.70/K$. Note that $\mathfrak{I}_3$ only needs to be integrated over Z_φ since $\Theta_{\theta,2} \leq \pi - \varphi$. Using TP1, we have $\mathfrak{I}_3 \leq 0.75\varphi P_1(Z_\varphi|\theta) \leq 0.75\varphi$. Approximating θ_0 by $1/K$, we obtain the sum $\mathfrak{I}_\theta \leq \frac{1.77}{K}$. Now let $\theta \in Z_\varphi$ and write $\mathfrak{I}_\theta = \mathfrak{I}_1 + \mathfrak{I}_2$, where

$$\mathfrak{I}_1 = \int_{Z_{\varphi^2/2}} \rho_\theta(\theta')P_1(Z^c_\varphi|\theta')\, d\theta' \leq \frac{0.75}{K} P_1(Z_{\varphi^2/2}|\theta) \leq \frac{0.75}{K}$$

where we have used TP1 and TP2, and

$$\mathfrak{I}_2 = \int_{Z^c_{\varphi^2/2}} \rho_\theta(\theta')P_1(Z^c_\varphi|\theta')\, d\theta' \leq \int_{Z^c_{\varphi^2/2}} \rho_\theta(\theta')\, d\theta' = P_1(Z^c_{\varphi^2/2}|\theta) \leq 0.38\varphi^2$$

where we have used TP1. Thus $\mathfrak{I}_\theta \leq \frac{0.94}{K}$. Therefore, if $\theta \in Z_{\sigma(\varphi)}$, $\varphi \leq \theta_0$, and $K \geq 2$, then

$$P_1(\Theta_2 \in Z^c_\varphi|\Theta_0 = \theta) \leq \frac{1.77}{K}.$$

Now set $C_2 = 1.77$ and, for each positive integer k suppose that $P_1^{k+2}(Z^c_\varphi|\theta)$ is bounded above by C_{k+2}/K, for $\varphi \leq \theta_0$, $\theta \in Z_{\sigma(\varphi)}$, and $k = 0, 1, \ldots, n$. Then $P_1^{k+3}(Z^c_\varphi|\theta)$ can be estimated as follows:

$$\begin{aligned}
\int_0^\pi \rho_\theta(\theta')P_1^{k+2}(Z^c_\varphi|\theta')\, d\theta' &= \int_{Z_{\sigma(\varphi)}} \rho_\theta(\theta')P_1^{k+2}(Z^c_\varphi|\theta')\, d\theta' + \int_{Z^c_{\sigma(\varphi)}} \rho_\theta(\theta')P_1^{k+2}(Z^c_\varphi|\theta')\, d\theta' \\
&\leq \frac{C_{k+2}}{K} \int_{Z_{\sigma(\varphi)}} \rho_\theta(\theta')\, d\theta' + \int_{Z^c_{\sigma(\varphi)}} \rho_\theta(\theta')\, d\theta' \\
&= \frac{C_{k+2}}{K} P_1(Z_{\sigma(\varphi)}|\theta) + P_1(Z^c_{\sigma(\varphi)}|\theta) \leq \frac{C_{k+2} + 0.10\varphi^2}{K}
\end{aligned}$$

where at the last step we have used TP4. Therefore, $C_{k+3} \leq C_{k+2} + 0.10\varphi^2$, concluding the proof. □

In the proof of the next corollary and later in the paper we use $\rho_\theta^{(n)}$ to denote the density of the probability measure $\delta_\theta P_1^n$ relative to the Lebesgue measure on $[0, \pi]$.

Corollary 2. *Let c_θ be as defined in Proposition 11 and $C_{n+2} = 1.77 + n\theta_0^2$ the constant that appears in Proposition 13 (setting $\varphi = \theta_0$). Let η be a probability measure on $(0, \pi)$ whose support is contained in $Z_{\sigma(\theta_0)}$. Then for all $n \geq 3$ the density function of ηP_1^n, relative to the Lebesgue measure, satisfies*

$$\frac{d(\eta P_1^n)}{d\theta} \geq c_\theta \left(1 - \frac{C_{n-1}}{K}\right).$$

for all $\theta \in Z_{\theta_0}$.

Proof. Observe that

$$\eta P_1^n = \int \delta_{\theta'} P_1^n\, d\eta(\theta'),$$

so it suffices to prove the lemma for measures of the form $\eta = \delta_{\theta'}$. In this case, if $\theta' \in Z_{\sigma(\theta_0)}$, then

$$\rho^{(n)}_{\theta'}(\theta) \geq \int_{Z_0} \rho^{(n-1)}_{\theta'}(\psi)\rho_\psi(\theta)\, d\psi \geq c_\theta P_1^{n-1}(Z_0|\theta').$$

The desired conclusion now follows from Proposition 13. □

We have already introduced the numbers $a_n = \sigma^n(1/K) = 8K(1/8K^2)^{2^n}$ and the intervals $A_n = [a_n, \pi - a_n]$. Define in addition $b_n = a_n/(2K)^{n+2}$, for $n = 0, 1, \ldots$ and $B_n = [b_n, \pi - b_n]$. Then $b_n > a_{n+1} > b_{n+1}$, so that

$$B_n \subset A_{n+1} \subset B_{n+1}.$$

Lemma 3. *We define $\delta_n := 3/(2K)^n$. Let $n \geq m \geq 1$ and $\theta \in B_n$. Then*

$$P_1^m(B_{n-m}|\theta) \geq \prod_{i=0}^{m-1}(1 - \delta_{n-i}).$$

Proof. The proof is by induction in m. The case $m = 1$ amounts to the following consequence of Proposition 2: for $\theta \in B_n$ and $K \geq 2$

$$P_1(B^c_{n-1}|\theta) \leq \left(1 + \frac{0.43}{K^2}\right) \frac{b_{n-1}\left(1 + \frac{b_{n-1}}{b_n}\right)}{4K} \leq \delta_n.$$

So suppose that the claim holds for $m - 1$, that is, $P_1^{m-1}(B_{n-m+1}|\theta) \geq (1 - \delta_{n-m+2}) \ldots (1 - \delta_n)$. Then, making use of Proposition 2 again and the induction hypothesis, we obtain

$$\begin{aligned} P_1^m(B_{n-m}|\theta) &\geq \int_{B_{n-m+1}} \rho^{(m-1)}_\theta(\varphi) P_1(B_{n-m}|\varphi)\, d\varphi \\ &\geq (1 - \delta_{n-m+1}) \int_{B_{n-m+1}} \rho^{(m-1)}_\theta(\varphi)\, d\varphi \\ &= (1 - \delta_{n-m+1}) P_1^{m-1}(B_{n-m+1}|\theta) \\ &\geq \prod_{i=0}^{m-1}(1 - \delta_{n-i}), \end{aligned}$$

as claimed. □

Proposition 14. *Assume that $K \geq 2$. Let $\theta \in A_n$, $n \geq 0$, and suppose that $m \geq n + 2$. Then*

$$P_1^m(A_0|\theta) \geq \left(1 - \frac{2.06 + 0.10(m - n - 2)\theta_0^2}{K}\right) \prod_{i=2}^{n}(1 - \delta_i).$$

Proof. It can be derived from Proposition 2 that $P_1(A_1^c|b_1) \leq 0.29/K$. If $\varphi \in B_1$ and $m \geq n + 2$, it then follows from Proposition 13 and this inequality (in the third and last inequalities below, respectively),

$$
\begin{aligned}
P_1^{m-n+1}(A_0|\varphi) &\geq P_1^{m-n+1}(A_0|b_1) \\
&\geq \int_{A_1} \rho_{b_1}(\psi) P_1^{m-n}(A_0|\psi)\, d\psi \\
&\geq \left(1 - \frac{1.77 + 0.10(m-n-2)\theta_0^2}{K}\right) \int_{A_1} \rho_{b_1}(\psi)\, d\psi \\
&= \left(1 - \frac{1.77 + 0.10(m-n-2)\theta_0^2}{K}\right) P_1(A_1|b_1) \\
&\geq 1 - \frac{2.06 + 0.10(m-n-2)\theta_0^2}{K}.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
P_1^m(A_0|\theta) &\geq \int_{B_1} \rho_\theta^{(n-1)}(\varphi) P_1^{m-n+1}(A_0|\varphi)\, d\varphi \\
&\geq \left(1 - \frac{2.06 + 0.10(m-n-2)\theta_0^2}{K}\right) P_1^{n-1}(B_1|\theta).
\end{aligned}
$$

But by Lemma 3, $P_1^{n-1}(B_1|\theta) \geq (1-\delta_2)\dots(1-\delta_n)$, proving the claim. □

Corollary 3. *Assume $K \geq 2$. If $\theta \in A_n$, $\varphi \in A_0$, and $m \geq n+2$, then*

$$
\rho_\theta^{(m+1)}(\varphi) \geq c_\varphi \left(1 - \frac{2.06 + 3(K-1)^{-1} + 0.10(m-n-2)K^{-2}}{K}\right).
$$

Proof. If $\varphi \in A_0$ and $\psi \in A_0$, then by Proposition 11 (2) $\rho_\psi(\varphi) \geq c_\varphi$, where c_φ was defined in that proposition. Now let $\theta \in A_n$, $\varphi \in A_0$, and $m \geq n+2$. Then

$$
\begin{aligned}
\rho_\theta^{(m+1)}(\varphi) \geq \int_{A_0} \rho_\theta^{(m)}(\psi)\rho_\psi(\varphi)\, d\psi \geq c_\varphi \int_{A_0} \rho_\theta^{(m)}(\varphi)\, d\varphi = c_\varphi P_1^m(A_0|\theta) \\
\geq c_\varphi \left(1 - \frac{2.06 + 0.10(m-n-2)\theta_0^2}{K}\right) \prod_{i=2}^{n} (1-\delta_i).
\end{aligned}
$$

In the last step above we have used Proposition 14. The product on the right-hand side of the above inequality can be estimated by $\prod_{i=2}^n (1-\delta_i) \geq 1 - 3/K(K-1)$ for all n. Replacing $1/K$ for θ_0 we obtain the claim. □

We summarize the main point of the above observations in the following lemma.

Lemma 4. *Let $n \geq 0$ and $m \geq n+2$. Let η be any probability measure on $(0, \pi)$. Then there exists a positive constant C_{m-n} depending on $m-n$ but independent of η, such that, for $\theta \in A_0$,*

$$
\frac{d(\eta P_1^m)}{d\theta} \geq c_\theta \left(1 - \frac{C_{m-n}}{K}\right) \eta(A_n).
$$

If $K \geq 2$, we can choose $C_l = 5.01 + 0.03l$.

Proof. This is due to the above corollary and that $(d(\eta P_1^m)/d\theta)(\theta) \geq \int_{A_n} \rho_\varphi^{(m)}(\theta)\, d\eta(\varphi)$. □

4 Mixing speed of the billiard Markov processes

Recall from Lemma 12 that $\|P_1^* g\|_\infty \leq 1.45(\|g\|_\infty/K) + 0.62$, where $P_1^* g$ is the density of $m_g P_1$ with respect to Lebesgue measure m on $[0, \pi]$. Thus, if $K \geq 2$ and $\|g\|_\infty \leq 2.3$, then $\|P_1^* g\|_\infty \leq 2.3$. Set $\gamma = 1.45/K$ and $\delta = 0.62/(1-\gamma)$. Observe that the map $F(x) = \gamma(x-\delta) + \delta$ is a contraction with fixed point δ, and $F^n(x) - \delta = \gamma^n(x - \delta)$. (As always, we assume $K \geq 2$.) Thus for all $\alpha > 0$ there exists an n_0 such that $\|P_1^{*n} g\|_\infty \leq \alpha(\|g\|_\infty - \delta) + \delta$ for all $n \geq n_0$. In particular, for n_0 big enough, we can assume that $\|P_1^{*n} g\|_\infty \leq 2.3$, for all $n \geq n_0$.

From Lemma 4,

$$(P_1^{*3} g)(\theta) \geq c_\theta \left(1 - \frac{C_2}{K}\right)(m_g P_1)(A_0)$$

and, assuming $\|g\|_\infty \leq 2.3$, then $m_g P_1(A_0^c) \leq 4.6/K$, from which we obtain

$$(P_1^{*3} g)(\theta) \geq c_\theta \left(1 - \frac{C}{K}\right),$$

where $C = C_2 + 4.6 < 9.7$. Define

$$p := (2 - \sqrt{2})(1 - C/K), \quad h(\theta) := c_\theta \left(1 - \frac{C}{K}\right) \mathbb{1}_{A_0}(\theta).$$

Notice that $m_h(A_0) = (2\cos a_0 - \sqrt{2})(1 - C/K))$, so by making C slightly bigger, and K sufficiently large, the higher order terms of $\cos a_0 = 1 + O(K^{-2})$ can be absorbed into C and we have $p = m_h(A_0)$.

Definition 2. *A probability density g will be called* proper *if $g|_{A_0} \geq h$ and $\|g\|_\infty \leq 2.3$.*

Lemma 5. *If g is a proper probability density then the probability density g_1 defined by*

$$m_{g_1} := \frac{m_g - m_h}{1-p} P_1^3$$

is also proper, assuming that K is big enough (say, $K > 10$).

Proof. This is an immediate consequence of the above observations. Note that $(m_g - m_h)/(1-p)$ is indeed a probability measure since $m_g([0,\pi]) - m_h([0,\pi]) = 1 - p$, and it has density $(g-h)/(1-p)$, which is positive as g is proper. □

We can then iterate the operation $g \mapsto g_1$ defined in Lemma 5 to obtain the following for any proper density g. Let g_n be the resulting density after n iterations. Then g_n is also proper and

$$m_g P_1^{3n} = (1-p)^n m_{g_n} + \sum_{j=0}^{n-1} (1-p)^j m_h P_1^{3(n-j)}. \tag{4.1}$$

Proposition 15. *Suppose K is big enough (say, $K > 10$) and let $\eta = m_g$ for any probability density g on $[0,\pi]$. Then there exists a positive integer r such that*

$$\|\eta P_1^{r+3n} - \nu_1\|_{TV} \leq 2(1-p)^n$$

for all positive integers n, where $\|\cdot\|_{TV}$ is the total variation norm and ν_1 is the invariant probability associated to P_1. (See Section 2.5.)

Proof. After applying P_1^r to m_g for r big enough we obtain a probability measure whose density is proper. The invariant measure ν_1 is itself proper. Then using the expansion 4.1 for both ηP_1^r and ν_1 we obtain $\eta P_1^{r+3n} - \nu_1 = (1-p)^n(\eta' - \nu')$, where η' and ν' are probability measures. Now recall that $\|\eta' - \nu'\|_{TV} \leq 2$. □

If f, g are essentially bounded functions and $g \geq 0$ such that $m_g([0,\pi]) = 1$, it follows from Proposition 15 that there exists a constant $C = C(g)$ such that

$$|(m_g P_1^n)(f) - \nu_1(f)| \leq C\|f\|_\infty (1-p)^{n/3} \tag{4.2}$$

for all positive n.

Corollary 4. *Assume that K is sufficiently big (say, $K > 10$). Let $f \in L^\infty([0,\pi],\nu_1) \cap L^2([0,\pi],\nu_1)$, where ν_1 is the P_1-stationary probability measure, and suppose that $\nu_1(f) = 0$. Then there exists a constant $C_f > 0$ such that*

$$|\langle P_1^n f, f\rangle_{\nu_1}| \leq C_f (1-p)^{n/3}$$

for every integer n.

Proof. Recall that $(d\nu_1/dm)(\theta) = \alpha_1(\theta)\sin\theta/\mu_1(M_1)$, where the notations are defined in Section 2.1 and m is the Lebesgue measure. We decompose $f = f_+ - f_-$ into its positive and negative parts and define $g_\pm(\theta) := f_\pm(\theta)(d\nu_1/dm)(\theta)$, which are essentially bounded functions with respect to m. Then

$$|\langle P_1^n f, f\rangle_{\nu_1}| \leq |(m_{g_+} P_1^n)(f)| + |(m_{g_-} P_1^n)(f)|.$$

If we now set $C_f = C(g_+) + C(g_-)$, then the claim follows from 4.2. □

5 Compactness and spectral gap

We show here that P_1 is a Hilbert-Schmidt operator, hence compact, and P is in general quasi-compact, and provide an estimate of the spectral gap.

Proposition 16. *The operator P_1 on $L^2([0,\pi],\nu_1)$ is Hilbert-Schmidt .*

Proof. Set $V = [0,\pi]$ and recall that $(P_1 f)(\theta) = \int_V \omega_1(\theta,\varphi) f(\varphi)\, d\nu_1(\varphi)$ for $f \in L^2(V,\nu_1)$, where

$$\omega_1(\theta,\varphi) = \frac{\mu_1(M_1)\sin\left(\frac{\theta+\varphi}{2}\right)\mathbb{1}_{V_\theta}(\varphi)}{K\alpha_1(\theta)\alpha_1(\varphi)\sin\theta\sin\varphi}$$

and $V_\theta = V_{\theta,1} \cup V_{\theta,2}$. (See the beginning of Section 3 for the notations.) We need to verify that the integral kernel ω_1 is a square integrable function on $V \times V$ with the product measure $\nu_1 \times \nu_1$. Due to the symmetry $\omega_1(\theta,\varphi) = \omega_1(\pi-\theta,\pi-\varphi)$, the L^2-norm of ω_1 is

$$2\int_0^{\pi/2}\int_V \omega_1^2(\theta,\varphi)\, d\nu_1(\varphi)\, d\nu_1(\theta) = \frac{1}{2K^2}\int_0^{\pi/2}\int_{V_\theta} \frac{\sin^2\left(\frac{\theta+\varphi}{2}\right)}{\alpha_1(\theta)\alpha_1(\varphi)\sin\theta\sin\varphi}\, d\varphi\, d\theta$$

is finite. (We have used here the expression for ν_1 given at the beginning of Subsection 2.1.) Observe that if $\theta_0 \leq \theta \leq \pi/2$ then both θ and $\varphi \in V_\theta$ are at a bounded distance away from 0 and π. The quantities $\alpha_1(\theta)$ and $\varphi_1(\varphi)$ are also bounded away from 0 for all angles. (In fact, they approach 1 for angles close to 0 or π.) So it suffices to check that $\mathcal{I} := \mathcal{I}_1 + \mathcal{I}_2$ is finite, where $\mathcal{I}_i$ is defined by

$$\mathcal{I}_i := \int_0^{\theta_0}\int_{V_{\theta,i}} \frac{\sin^2\left(\frac{\theta+\varphi}{2}\right)}{\sin\theta\sin\varphi}\, d\varphi\, d\theta.$$

Note that $\sin\left(\frac{\theta+\varphi}{2}\right) \leq \sin\theta$ on $V_{\theta,1} = [\Theta_{\theta,1},\theta]$, and $\sin\left(\frac{\theta+\varphi}{2}\right) \leq \sin\varphi$ on $V_{\theta,2} = [\theta,\Theta_{\theta,2}]$. So

$$\mathcal{I}_1 \leq \int_0^{\theta_0}\int_{\Theta_{\theta,1}}^{\theta} (\sin\theta/\sin\varphi)\, d\varphi\, d\theta \leq C\int_0^{\theta_0} \sin\theta \ln(\theta/\Theta_{\theta,1})\, d\theta.$$

Keeping in mind the bounds on $\Theta_{\theta,i}$ obtained in Proposition 9, we have $\Theta_{\theta,1} \geq c\theta^2$, for some positive constant c, from which it follows that the integral is finite. Now

$$\mathfrak{I}_2 \leq \int_0^{\theta_0} \int_{\Theta_{\theta,1}}^{\theta} (\sin\varphi/\sin\theta)\, d\varphi\, d\theta \leq \int_0^{\theta_0} \frac{1}{\sin\theta} \left[\frac{\Theta_{\theta,2}^2 - \theta^2}{2}\right] d\theta$$

and we know that $\Theta_{\theta,2}$ is bounded above by a constant times $\sqrt{\theta}$. Thus $\mathfrak{I}_2$ is also finite. □

We now turn to the estimate of the spectral gap for P_1.

Lemma 6. *Let T be a bounded self-adjoint operator of norm 1 in the Hilbert space $H := L^2(V, \mu)$, where μ is a probability measure on the Lebesgue space V. Let H_0 be the orthogonal complement of the constant functions and define the spectral gap $\gamma(T) = 1 - \rho_0(T)$, where $\rho_0(T)$ is the spectral radius of $T|_{H_0}$. Also suppose that $\langle T^n f, f\rangle \leq C_f \lambda^n$ for a positive $\lambda < 1$ and all $f \in H_0 \cap L^\infty(V, \mu)$ and positive integer n. Then*

1. *$\|T^n f\|_2 \leq \|f\|_2 \lambda^n$ for all $f \in H_0$ and positive integer n;*
2. *$\gamma(T) \geq 1 - \lambda$.*

Proof. We may assume that $\|f\|_2 = 1$. Since $L^\infty(V, \mu)$ is dense in H, we may also assume that the functions are essentially bounded. Let $T = \int_\sigma \lambda\, dE(\lambda)$ be the spectral decomposition of T, where $\sigma \subset \mathbb{R}$ denotes the spectrum of T and E is the corresponding projection-valued measure. Write $d\mu_f(\lambda) := \langle dE(\lambda) f, f\rangle$. Then for $p \geq 1$

$$\|T^n f\|_2^2 = \int_\sigma \lambda^{2n} d\mu_f(\lambda) = \left(\left[\int_\sigma (\lambda^{2n/p})^p\, d\mu_f(\lambda)\right]^{1/p}\right)^p \geq \left(\int_\sigma \lambda^{2n/p}\, d\mu_f(\lambda)\right)^p = \|T^{n/p} f\|_2^{2p}.$$

We have used that $\|g\|_1 \leq \|g\|_p$ for probability measures. Now let $p = n/m$, where $n \geq m$. Then

$$\|T^m f\|_2 \leq \|T^n f\|_2^{m/n} = \langle T^{2n} f, f\rangle^{m/2n} \leq C_f^{m/2n} \lambda^m.$$

By letting n go to ∞ we conclude (1).

Now recall that $\rho_0(T) = \lim_{m\to\infty} \|(T|_{H_0})^n\|_2^{1/n}$. For any ϵ_n such that $\lim \epsilon_n = 0$ we can choose $f_n \in H_0$ of norm 1 for which

$$\|(T|_{H_0})^n\|_2^{1/n} \leq \|T^n f_n\|_2^{1/n} + \epsilon_n \leq \lambda + \epsilon_n,$$

where we have used part (1). By passing to the limit we obtain (2). □

Thus we obtain for the spectral gap of P_1 the estimate $\gamma \geq 1 - (1-p)^{1/3}$ where, we recall, $p = (2 - \sqrt{2})(1 - C/K)$ whenever K is sufficiently big. Thus for large values of K we get

$$\gamma \geq 0.25 - 0.35C/K.$$

Proposition 17. *For the same $C > 0$ as above (see paragraph just prior to Definition 2) and some $N \geq 3$, the spectral gap of P_1 satisfies $\gamma \geq 1 - (C/K)^{1/N}$ for all $K > 0$. In particular, the gap must approach the maximum possible value 1 for very large values of K.*

Proof. We know from Proposition 15 that for any probability measure η absolutely continuous with respect to the Lebesgue measure ηP_1^n converges to ν_1 exponentially fast in the total variation norm. In fact, any measure η that does not have atoms at 0 and π satisfies this property since one application of P_1 makes the measure absolutely continuous. With this in mind, we can now go back to the argument at the beginning of Section 4 and improve on the estimate used in the specification of p. More precisely, it is now possible to find a positive integer N such that, for any given probability density g on V with respect to the Lebesgue measure m

$$P_1^{*N} g \geq \left(1 - \frac{C}{K}\right) \frac{d\nu_1}{dm}.$$

That is, we have replaced the function c_θ with the density of ν_1. Accordingly, we redefine $h(\theta)$ and p, so that $p = 1 - C/K$. (This has been achieved at the cost of increasing the number of iterations of P_1 needed for each step in the coupling argument.) The spectral gap γ can now be estimated, just as done above (using Corollary 4 and Lemma 6), by

$$\gamma \geq 1 - (1-p)^{1/N} = 1 - (C/K)^{1/N}$$

as claimed. □

With this remark, all the claims involved in Theorem 2 have been proved. Theorem 1 follows from Theorem 2 and the next general observation. (See Theorem 9.9 in [20].)

Proposition 18. *Let K and T be bounded self-adjoint operators on a Hilbert space and suppose that K is compact. Then the essential spectrum of $T + K$ is contained in the essential spectrum of T. In particular, if $\|T + K\|_2 = 1$ and $\|T\|_2 < 1$, then the spectral gap of $T + K$ satisfies*

$$\gamma(T + K) \geq \min\{1 - \|T\|_2, \gamma(K)\}.$$

The theorem follows from letting $K = \alpha_1 P_1$ and $T = \alpha_2 P_2$. Observe that, due to Proposition 10 we have $1 - \|T\|_2 \geq \inf \alpha_1 \geq (1 - 0.37/K)\sqrt{2}/K$.

6 Glossary of symbols

We summarize below the notation used across more than one section of the text. Symbols not listed here are likely to be used only locally, and their first occurrence should be near the places where they are employed.

1. Measures
   - λ Lebesgue measure on the unit interval $I = [0, 1]$; occasionally denoted ds, dx, etc.
   - m (non-normalized) Lebesgue measure on $V = [0, \pi]$; often denoted $d\theta$, $d\phi$, etc.
   - m_g finite measure on V with density $g \in L^1(V, m)$; i.e., $dm_g/dm = g$
   - $\nu = m_g$, where $g(\theta) = \frac{1}{2}\sin\theta$
   - ν_g finite measure on V with density $g \in L^1(V, \nu)$; i.e., $d\nu_g/d\nu = g$
   - $\mu = \lambda \otimes \nu$ probability measure on $M = I \times V$ invariant under the return billiard map T
   - ν_j conditional measures associated to a measurable partition $\{M_j\}$ of M (Subsection 2.5)
   - δ_θ Dirac delta measure on V concentrated at θ
   - η symbol often used for general purpose measure
2. Sets
   - $Q \subset \mathbb{R}^2$ the (micro) billiard cell
   - Γ_i parts of the boundary of Q (Subsections 1.1 and 2.1)

- $I = [0, 1]$, $V = [0, \pi]$, $M = I \times V$
- $\mathcal{M}$ phase space of the billiard system on Q (Subsection 2.1)
- $\mathcal{M}_i$ parts of $\mathcal{M}$ associated to the Γ_i (Subsection 2.1)
- $\mathcal{S}_m$ singular sets associated to the billiard map $\mathcal{F}$ on $\mathcal{M}$ (Subsection 2.1)
- $W_\theta = I \times \{\theta\}$, W_θ^i, $W_{\theta,j}$, $V_{\theta,j}$ (Subsection 2.6 for a general partition)
- $W_{\theta,j}$, $V_{\theta,j}$, $V_\theta = [\Theta_{\theta,1}, \Theta_{\theta,2}]$ (Subsection 3.1 for the special partition)
- $Z_\varphi = [\varphi, \pi - \varphi]$, $Z_0 = Z_{\theta_0}$
- $A_n = Z_{a_n}$, where $a_n = \sigma^n(1/K)$, $B_n = Z_{b_n}$, where $b_n = a_n/(2K)^{n+2}$ (Subsection 3.3)
- M_i elements of a measurable partition of M (Subsection 2.5). Different partitions are employed at different places in the text, and are defined where they arise; for the main theorems and throughout most of Section 3 we use the *special partition*, which is defined just prior to the statement of Theorem 2.

3. Functions, maps, and operators

- $\mathcal{F}$ billiard map on $\mathcal{M}$
- T return billiard map on M
- $\Psi_r(\theta) = \Theta(r, \theta)$ angle component of $T(r, \theta)$
- P Markov (reflection) operator of random billiard (Definition 1)
- P_i conditional operator associated to a partition M_i of M (Subsection 2.5)
- α_i probability function associated to a partition M_i of M (Subsection 2.5)
- $\pi_2 : M \to V$ the natural component projection
- $\mathcal{J}$ used for the map $(r, \theta) \to (1 - r, \pi - \theta)$ on M
- J used for the maps $(r, \theta) \to (r, \pi - \theta)$ on M, $\theta \to \pi - \theta$ on V, and the induced unitary operator on $L^2(V, \nu)$
- $\mathcal{I}_i : L^2(V, \nu) \to L^2(V, \nu_i)$, inclusion operator
- $\mathcal{M}_i : L^2(V, \nu_i) \to L^2(V, \nu)$, multiplication by α_i
- $\omega_1(\theta, \varphi)$, integral kernel of P_1 (Subsection 2.6 general; Subsection 3.2 for the special partition)
- $\Theta_1, \Theta_2, \dots$ Markov chain for the angle process with Markov operator P

4. Miscellaneous

- γ, $\gamma(K)$ spectral gap of an operator K; $\gamma_1 = \gamma(P_1)$
- K scale-invariant curvature of Γ_i (defined just prior to Theorem 1)
- θ_0 is defined by $\sin \theta_0 = 1/K$
- $r_{\theta,i}$, $\Theta_{\theta,i}$ (Subection 3.1)
- $\mathbb{1}_A$ the indicator function of A

Acknowledgment. Hong-Kun Zhang is partially supported by NSF DMS-0901448.

Email addresses: feres@math.wustl.edu, hongkun@math.umass.edu

References

[1] L.A, Bunimovich, *On the ergodic properties of nowhere dispersing billiards*, Comm. Math. Phys. **65** (1979), 295–312.

[2] L.A, Bunimovich, *Mushrooms and other billiards with divided phase space*, Chaos 11 (2001), 802808.

[3] L.A, Bunimovich, Ya.G. Sinai, & N.I. Chernov, *Markov partitions for two-dimensional hyperbolic billiards,* Russian Math. Surveys **45** (1990) 105–152.

[4] L.A, Bunimovich, Ya.G. Sinai, & N.I. Chernov, *Statistical properties of two-dimensional hyperbolic billiards,* Russian Math. Surveys **46** (1991) 47–106.

[5] C. Cercignani and D.H. Sattinger, *Scaling Limits and Models in Physical Process*, Birkhäuser, DMV Seminar, Band 28, 1998.

[6] N.I. Chernov, *Decay of correlations in dispersing billiards*, J. Stat. Phys. **94** (1999), 513–556.

[7] N.I. Chernov and R. Markarian, *Chaotic Billiards*, Mathematical Surveys and Monographs, **127**, AMS, Providence, RI, 2006.

[8] N.I. Chernov and L. S. Young, *Decay of correlations for Lorentz gases and hard balls.* Hard ball systems and the Lorentz gas. Encyclopaedia Math. Sci. **101**, Springer, Berlin, (2000), 89–120.

[9] N.I. Chernov and H.-K. Zhang, *Billiards with polynomial mixing rates*, Nonlinearity **18** (2005), 1527–1553.

[10] N.I. Chernov and H.-K. Zhang, *A family of chaotic billiards with variable mixing rates*, Stochast. Dynam. **5** (2005), 535–553.

[11] N. Chernov and H.-K. Zhang, *Improved estimate for correlations in billiards,*, Communications in Mathematical Physics, **277** (2008), 305–321.

[12] F. Comets, S. Popov, G. Schutz and M. Vachkovskaia, *Knudsen gas in a finite random tube: transport diffusion and first passage properties* J. Stat. Phys. (2010) 140: 948-984.

[13] F. Comets, S. Popov, G.M. Schutz, M. Vachkovskaia *Billiards in a general domain with random reflections*, Arch. Ration. Mech. Anal. **191** (2008), 497-537

[14] R. Feres, *Random walks derived from billiards*, Dynamics, Ergodic Theory, and Geometry, Ed. B. Hasselblatt. Mathematical Sciences Research Institute Publications **54** (2007), 179–222.

[15] R. Feres and G. Yablonsky, *Knudsen's cosine law and random billiards*, Chemical Engineering Science, **59** (2004), 1541–1556.

[16] R. Feres and G. Yablonsky, *Probing surface structure via time-of-escape analysis of gas in Knudsen regime*, Chemical Engineering Science, **61** (2006), 7864–7883.

[17] R. Feres and H.-K. Zhang, *The spectrum of the billiard Laplacian of a family of random billiards*, J. Stat. Phys., Volume 141, Number 6, (2010) 1039-1054

[18] R. Markarian, *Billiards with polynomial decay of correlations*, Ergodic Theory Dynam. Syst. **24** (2004), 177–197.

[19] Ya.G. Sinai, *Dynamical systems with elastic reflections. Ergodic properties of diepersing billiards,* Russian Math. Surveys **25** (1970) 137–189.

[20] J. Weidmann, *Linear operators in Hilbert spaces,* (1980), Springer-Verlag, New York Inc.

[21] L.-S.Young, *Statistical properties of systems with some hyperbolicity including certain billiards*, Ann. Math., **147** (1998), 585–650.

[22] L.-S.Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.