

Survival Probability for the Stadium Billiard

Carl P. Dettmann ^{*1} and Orestis Georgiou ^{†1}

¹*School of Mathematics, University of Bristol, United Kingdom*

We consider the open stadium billiard, consisting of two semicircles joined by parallel straight sides with one hole situated somewhere on one of the sides. Due to the hyperbolic nature of the stadium billiard, the initial decay of trajectories, due to loss through the hole, appears exponential. However, some trajectories (bouncing ball orbits) persist and survive for long times and therefore form the main contribution to the survival probability function at long times. Using both numerical and analytical methods, we concur with previous studies that the long-time survival probability for a reasonably small hole drops like $Constant \times (time)^{-1}$; here we obtain an explicit and exact expression for the *Constant*.

I. INTRODUCTION

Chaotic billiard theory was introduced by Yakov Sinai in 1970 [1]. In the past 38 years it has developed to become a well established theory for dynamical systems. A billiard is a dynamical system in which a particle alternates between motion in a straight line and specular reflections from the domain's boundary. The sequence of reflections is described by the billiard map which completely characterizes the motion of the particle, hence billiards have their boundaries as a natural Poincaré section. Billiard systems are convenient models for many physical phenomena, for example where one or more particles move inside a container and collide with its walls. An excellent and comprehensive mathematical introduction to chaotic billiard theory can be found in the book of Chernov and Markarian [2].

In the early 80's, mathematicians suggested investigating open systems, systems with holes or leakages, as a means of generating transient chaos [3], retrieving information from distributions [4], and deducing facts about the equivalent closed systems. The key distributions of interest classically are the escape probability density $p_e(t)$, which is given by the trajectories that leave the billiard at time t , where $t \geq 0$, and $t \in \mathbb{R}$, and also the survival probability of orbits $P(t)$ up to time t . These two are related by $P(t) = \int_t^\infty p_e(t') dt'$. Such investigations have naturally been extended to billiard systems as well. Links between billiards and geometrical acoustics [5], quantum chaos [6], controlling chaos [7], atom optics [8], hydrodynamical flows [9–11], astronomy [12] and cosmology [13], have been established in the context of open dynamical billiards. Furthermore, it has become apparent over the past few years, that the subject of open billiards and their distributions provide a pathway towards understanding chaos and may even open doors to old, but not so forgotten problems such as the Riemann hypothesis [14].

The stadium billiard (see Figure 1 below) is a seemingly simple dynamical system, introduced by Leonid Bunimovich in 1974 [15]. The billiard's boundary consists of two parallel straight lines and two semicircular arcs. It was later proven by him to be ergodic, to be mixing, to have the Kolmogorov property [16], and by Nikolai Chernov to have the Bernoulli property [17]. It has been described as a system with “fully developed chaos” [18]. Its entropy has been numerically estimated in [19] and theoretically in [20]. The stadium billiard is a special case of a chaotic billiard. Being constructed from two fully integrable billiard segments, the circle and the rectangle, it is remarkable that the system remains completely chaotic no matter how short its parallel segments are. It is a limiting case of the larger set of Hamiltonian systems, which Bunimovich refers to as “mushrooms” [21] with cleanly divided phase-space areas, regular and chaotic. The stadium is the fully chaotic limit of the natural mushroom billiard, while the circle is the fully regular limit. If the parallel segments are of length $2a$ say, where $a > 0$, and $a \in \mathbb{R}$, then the Lyapunov exponent $\lambda(a) \rightarrow 0$ in both limiting cases of $a \rightarrow 0$ and $a \rightarrow \infty$. Also, it is well known that the defocusing mechanism, which is one of the two sources of chaos in billiards [22] (the other being the dispersing mechanism), characteristic of all Bunimovich type billiards, requires $a > 0$ in order for any wave-front to defocus and therefore exhibit hyperbolicity.

The very existence of the parallel segments of the boundary is also the source of the intermittent behaviour found in the stadium billiard. They allow for the existence of a set of marginally stable periodic orbits of zero measure but indeed of great importance. They are the main reason why the stadium is not uniformly hyperbolic. Also, though it is classically and quantum mechanically ergodic, it does not have the property of unique ergodicity [23, 24]. This means that not *all* eigenfunctions are uniformly distributed and therefore this causes scarring [25]. This is due to the existence of the so called “bouncing ball” orbits, sometimes called “sticky orbits”. Semi-classically, they have caused much trouble in the treatment of the system [26]. It has also been shown by Bálint and Gouëzel [27] that the Birkhoff

^{*} Carl.Dettmann@bristol.ac.uk

[†] maxog@bristol.ac.uk

sums of a sufficiently smooth generic observable with zero mean of the stadium, satisfy a non-standard limit theorem where its convergence to a Gaussian distribution requires a $\sqrt{n \log n}$ normalization.

Even though the stadium billiard exhibits strong chaotic properties such as initial exponential decays of the escape times distribution and decays of correlations of initial conditions, it has been shown to experience a cross-over at longer times, towards an asymptotic power-law behavior [28, 29]. Hence, the stadium billiard is an example of a transient chaotic system which exhibits intermittency. Intermittency in open systems is being increasingly discussed and researched in the context of non-linear Hamiltonian systems. These investigations tend to focus on the asymptotic behaviour of distributions and often use the stadium as one of their main examples [30, 31]. It is an ideal model for studying the influence of almost regular dynamics near marginally stable boundaries both theoretically [26] and numerically [32]. Armstead *et al.* [33], give a detailed investigation into the stadium dynamics and show that $P(t) \sim \frac{Const}{t}$ for long times but do not calculate the *Constant*.

In this paper we explicitly calculate the measure of the set of orbits, causing the asymptotic power-law behavior and obtain an analytic expression for the *longer times* survival probability function of the stadium billiard. In contrast with Ref.[14], we do not assume or require that the hole is vanishingly small and in contrast to Ref.[33], we do not use a probabilistic description of the dynamics. The paper is organized as follows. In section II we state the reasoning and the main ideas of this paper and we also set up our problem and define all the variables and sets needed. In section III and IV we consider the two main sets of initial conditions which contribute to the survival probability at long times while in section V we introduce and explain the details of the approximation method used. Finally, in section VI we present our numerical results from computer simulations and compare with the analytical ones. Conclusions and discussions appear at the end, in section VII, where future work is also discussed.

II. THE MAIN IDEAS AND SET-UP

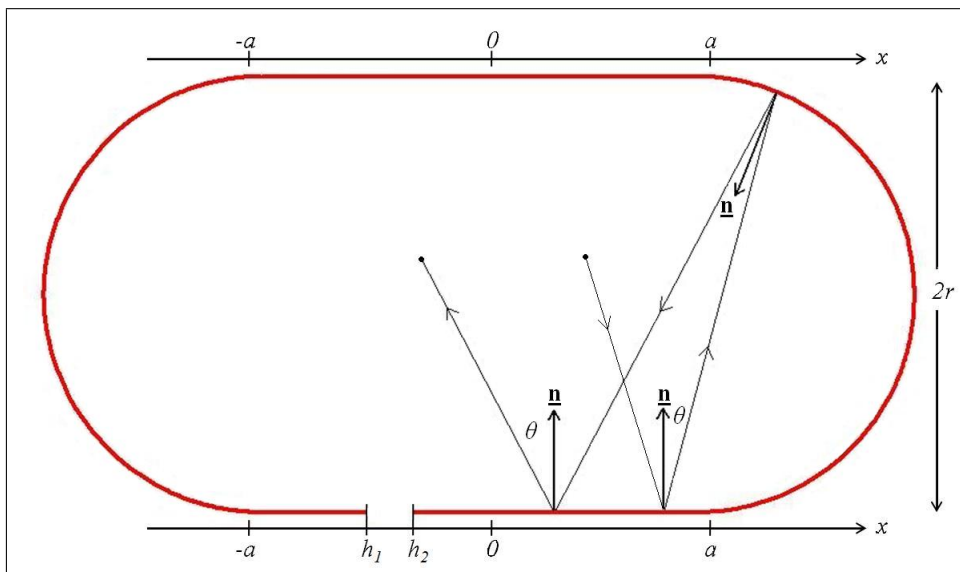


FIG. 1: The Set-Up of the stadium billiard.

Consider an open stadium as shown in Figure 1. As discussed in the Introduction, due to the chaotic, ergodic nature of the stadium billiard, most trajectories are expected to escape exponentially fast as they soon come to occupy all regions of the billiard phase space. This is a result of the defocusing mechanism. However, ergodicity also guarantees that all orbits must spend arbitrarily long times in both integrable areas of phase space. Trajectories almost tangent to the circular arcs (“whispering gallery orbits, or rolling orbits”) are of bounded path lengths and therefore do not contribute to long time tails. Bouncing ball orbits on the contrary can, and are the ones that determine the asymptotic tail [33, 34] of $Const(\text{time})^{-1}$ behaviour of the survival probability. These trajectories are characterised by small angles θ (near vertical) that remain small for relatively long periods of time. Nikolai Chernov describes them in his book [2] as orbits with a large number of *nonessential collisions*. Their behaviour is controlled by a set of parabolic, non-isolated periodic orbits [35].

When a bouncing ball orbit, with an angle close to zero (near vertical) reaches one of the circular ends of the stadium, it will reflect. It is known that the reflected orbit will have a new angle which will again be very close to

zero [36]. Assuming that an orbit starts off with a small enough angle, then we expect a large number of reflections on either of the circular ends of the billiard before this angle starts to grow significantly. However, if the stadium has a hole situated somewhere on one of its flat segments, we only need to consider two simple scenarios. The first one is when a trajectory is moving towards the hole and the second one is when the trajectory is moving away from the hole, reflects at one of the circular ends, and then returns to escape through the hole. Hence, we calculate the measure of the set of orbits which experience up to two transitions between the integrable segments of the stadium billiard and use it to obtain an explicit expression for the survival probability function for long times (equation (33) below).

The set up of the problem is straight forward (see Figure 1). We have a classical particle of unit mass and unit speed which experiences elastic collisions on Γ , the boundary of the billiard table. The length of the parallel sides is $2a$, while the radius of the circular segment is r . A hole of size ϵ is punched onto one of the straight segments of Γ with x coordinates $x \in (h_1, h_2)$, $h_2 = h_1 + \epsilon$. x is the position coordinate which we only need to take values along the straight segments, $x \in [-a, a]$. We define the inward pointing normal vector $\hat{\mathbf{n}}$ which defines the angle θ made by the reflected particle and $\hat{\mathbf{n}}$. $\theta \in (-\pi/2, \pi/2)$, and is positive in the clockwise sense from $\hat{\mathbf{n}}$. The billiard flow conserves the phase volume and the corresponding invariant measure along the straight segments is $d\mu = C^{-1} \cos \theta d\theta dx$, where C is some normalization constant which we calculate in the next section. This is also the distribution of initial conditions. For the purpose of this paper, no parametrization of the position coordinate is needed along the circular segments of the billiard. Finally, if the particle hits the hole, it will escape; as noted in the Introduction, we are interested in the long time behaviour of the survival probability.

III. CASE I: MOVING TOWARDS THE HOLE

We start by considering trajectories initially on the right of the hole with $x \in (h_2, a)$ moving towards it. These trajectories will prove to be only a part of the survival probability function for long times. However, they are essential in order to construct a complete and accurate expression for the asymptotic limit of the full survival probability function. To ensure that such trajectories will escape when they reach the hole's vicinity, they must satisfy the following condition:

$$|\theta| < \arctan\left(\frac{\epsilon}{4r}\right), \quad (1)$$

hence they will definitely not jump over the hole. The set of initial conditions (x, θ) for trajectories which will escape in exactly time t satisfies:

$$a - h_2 \geq x - h_2 \geq t \sin |\theta| - \delta 4r \tan |\theta|, \quad (2)$$

where $0 < \delta < 1$. For long times however, $\delta 4r \tan |\theta|$ will shrink to zero ($\sim 1/t$) as the set of surviving trajectories is limited to near vertical angles. Hence we drop this nonsignificant term in what follows. Notice that we will be using physical time t for our calculations but the equations are set up as if considering a map between $n \in \mathbb{N}$ collisions, with $t = \frac{2rn}{\cos \theta}$. This way, we do not need to define equations for the billiard map, which in any case would just be described by the usual reflection map. From equations (1) and (2) we can deduce that the angles must satisfy:

$$|\theta| < \min\left\{\arctan\left(\frac{\epsilon}{4r}\right), \arcsin\left(\frac{a - h_2}{t}\right) + \mathcal{O}(1/t^2)\right\}. \quad (3)$$

The second term in (3) is the dominant one for long times. This leads to the following integral for the conserved measure of the billiard map:

$$I_r = 2 \int_0^{\arcsin\left(\frac{a - h_2}{t}\right) + \mathcal{O}(1/t^2)} \left(\int_{h_2 + t \sin \theta + \mathcal{O}(1/t)}^a \cos \theta dx \right) d\theta, \quad (4)$$

where the subscript r stands for *right*. We are integrating over the set of initial conditions, on the right of the hole, which will not escape until time t . Hence we are considering escape times greater than or equal to t . We have also dropped the modulus sign from θ and multiplied the whole expression by a factor of 2, due to the vertical symmetry of the problem. This simplifies to

$$I_r = \frac{(a - h_2)^2}{t} + \mathcal{O}(1/t^2). \quad (5)$$

This result is valid for trajectories satisfying:

$$\arcsin\left(\frac{a-h_2}{t}\right) + \mathcal{O}(1/t^2) < \arctan\left(\frac{\epsilon}{4r}\right),$$

that is

$$t \gtrsim \frac{8ar}{\epsilon}, \quad (6)$$

since the supremum of $a - h_2$ is $2a$. We continue with this calculation by adding the analogous contribution I_l from the small angle trajectories starting on the *left* of the hole with $x \in (-a, h_1)$ moving towards it. This operation can easily be calculated from equation (5) by simply sending $h_1 \mapsto -h_2$ and $h_2 \mapsto -h_1$.

$$I_l = \frac{(a + h_1)^2}{t} + \mathcal{O}(1/t^2).$$

Adding the two integrals gives the measure of all initial conditions moving towards the hole that survive until time t :

$$I_{r+l} = \frac{(a + h_1)^2}{t} + \frac{(a - h_2)^2}{t} + \mathcal{O}(1/t^2). \quad (7)$$

Finally, in order to turn this into a probability, we must divide equation (7) by the total measure of the billiard map given by:

$$C = \int_{-\pi/2}^{\pi/2} \cos \theta d\theta \int_{\Gamma} dr = 2|\Gamma| = 2(4a + 2\pi r). \quad (8)$$

Hence, part of the canonical Survival Probability function due to the nonessential orbits initially approaching the hole from either side, for long times satisfying condition (6) is:

$$P_1(t) = \frac{(a + h_1)^2 + (a - h_2)^2}{2(4a + 2\pi r)t} + \mathcal{O}(1/t^2). \quad (9)$$

This expression is essentially a sum of contributions from two families of bouncing ball orbits, each proportional to the square of the available length.

IV. CASE II: MOVING AWAY FROM THE HOLE

Numerical simulations confirm that $P_1(t)$ in equation (9) is indeed not the full expression for the long time survival probability function of the open stadium billiard. The set of orbits accounted for in the previous section is only a fraction of all the marginally stable periodic orbits discussed in the introduction of this paper. In this section we will be considering orbits initially moving away from the hole, so that they experience a reflection process when they collide with the right semicircular end. If the initial angles are small, then the final angles (after being reflected at the wings of the stadium) will remain small and therefore survive for long times and account for the remaining set of orbits and built up the long time survival probability. In this section we will investigate and identify exactly the initial conditions which survive for long times t .

Throughout this and the following section we will be using (x_i, θ_i) as the coordinates of our initial conditions which lie on the right of the hole, $x_i \in (h_2, a)$, and move away from it. Due to the stadium's symmetry, we only need consider the case $\theta_i > 0$. We will use (x_1, θ_1) to indicate the position and angle of a trajectory right after its final collision on a flat segment, while still moving away from the hole. Therefore, the next collision of such trajectories will be on the right semi-circular segment of the billiard. This helps to distinguish between the initial conditions and their transformed final values. Notice that $\theta_i = |\theta_1|$.

We begin by formulating the time to escape

$$T(x_i, \theta_1) = \frac{2rn}{\cos \theta_1} + \frac{2rm}{\cos \theta_f} + D_f, \quad (10)$$

where n and m are the respective numbers of non essential collisions before and after the reflection process on the right semicircular end, and are defined as:

$$n = \left(\frac{a-x_1}{2r \tan |\theta_1|} - \delta_i \right)$$

$$m = \left(\frac{a-h_2}{2r \tan |\theta_f|} - \delta_f \right),$$

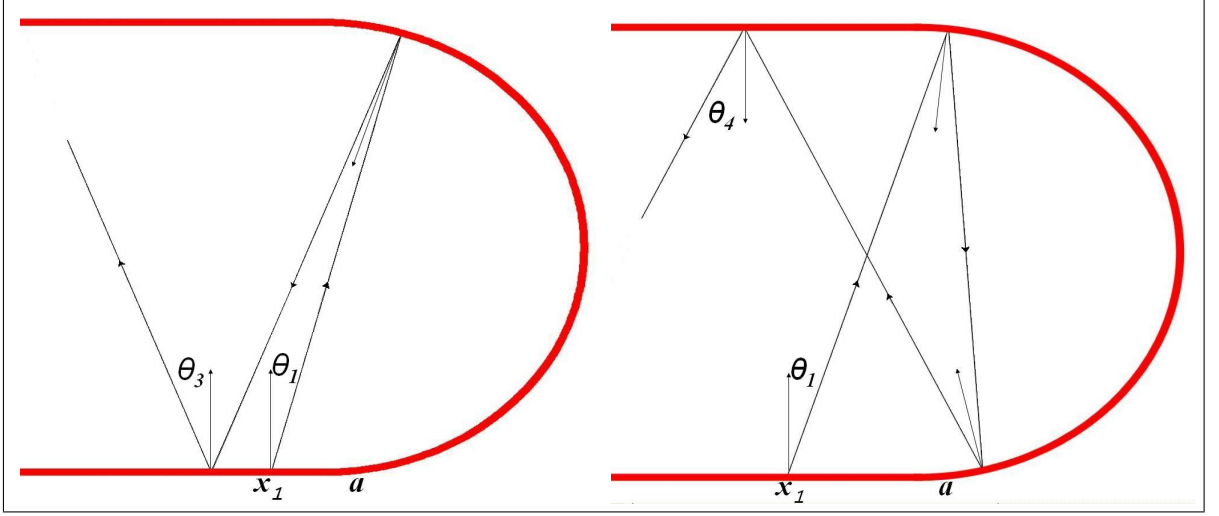


FIG. 2: The two possible scenarios of reflection from the semicircles, where $d_1 = a - x_1 = a - x_i - 2rn\theta_1 > 0$.

with $0 < \delta_{i,f} < 1$, and $f = 3, 4$. We chose these indices for f (3 and 4) to indicate the number of collisions comprising the reflection process as the angles change two or three times respectively. D_f is the time taken for the reflection process at the semicircular end, and is bounded by:

$$4r < D_3 < 2r \left(\frac{1}{\cos \theta_1} + \frac{1}{\cos \theta_3} \right)$$

$$6r < D_4 < 2r \left(1 + \frac{1}{\cos \theta_1} + \frac{1}{\cos \theta_4} \right).$$

The number of reflections on the curved boundary alternates between two scenarios, as shown in Figure 2, the case with one collision on the semicircle and the case with two collisions depending on the initial conditions of the trajectory x_i and θ_i . Specifically, θ_f can be found and defined by the use of small angle approximations as:

$$\theta_3 = \frac{2d_1}{r} - 3\theta_1 < 0, \quad (11)$$

$$\theta_4 = \frac{4d_1}{r} - 5\theta_1 > 0, \quad (12)$$

where $d_1 = a - x_1 = a - x_i - 2rn\theta_1 > 0$, is the horizontal distance between $x = x_1$ and $x = a$ [33, 36]. Taking only the leading order terms of θ_1 and θ_f from our expressions is justified by the fact that in long time scenarios, unstable periodic orbits have very small angles (before and after being reflected), tracing out near vertical trajectories. Using the above information, equation (10) can be expressed in the following way:

$$T(x_i, \theta_1) = \frac{a - x_1}{|\theta_1|} + \frac{a - h_2}{|\theta_f|} + \Delta_f, \quad (13)$$

where $\Delta_f = D_f - 2r \left(\frac{\delta_i}{\cos \theta_i} + \frac{\delta_f}{\cos \theta_f} \right)$.

Before we continue, it is essential to find the boundaries of validity for the functions of θ_f . These will define the geometry of the two scenarios. We ask the question: When do we see one and when two collisions at the semicircular ends? This is answered by considering a number of inequalities. The first and most obvious one is $\theta_1 \geq \frac{d_1}{2r}$ which requires the next collision to be on the circular segment. We also note that $\theta_1 = \frac{3d_1}{4r}$ is the transition line between θ_3 and θ_4 and is the case where the reflected particle will hit exactly the point $x = a$, where the circular segment meets the straight segment. Finally, if we are also to satisfy condition (1) we must form two more inequalities: equation (11) gives

$$\theta_1 < \frac{2d_1}{3r} + \frac{\arctan \left(\frac{\epsilon}{4r} \right)}{3}, \quad (14)$$

and equation (12) gives

$$\theta_1 < \frac{4d_1}{5r} - \frac{\arctan\left(\frac{\epsilon}{4r}\right)}{5}. \quad (15)$$

These inequalities enclose a small area in the $x_i\theta_1$ plane, the plane of initial conditions. Notice that d_1 is a function of x_i but also depends on n , the number of collisions before the reflection process, which if not equal to zero introduces an extra θ_1 term. This means that inequalities (14) and (15) have to be solved for θ_1 for every $n = 0, 1, 2, 3, \dots$. This is done and shown in Figure 3, up to and including $n = 2$, where z is taken to be equal to $\arctan(\epsilon/4r)$. One would expect, as in section IV, any dependence on the size of the hole ϵ , to vanish for long enough times. These boundaries of validity define the set of orbits which escape after being reflected at the right semicircular segment of the billiard. It is not immediately clear from Figure 3, but the peaks of these spikes are of the same height $\theta_1 = 3z$. However, we note that the set of orbits that survive up to time T , where T is large, is not identical with the former set.

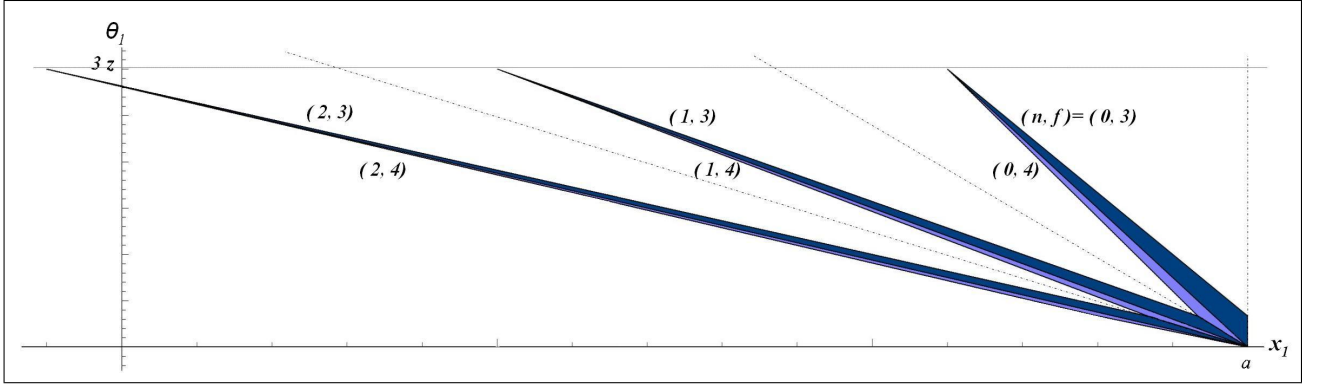


FIG. 3: (Colour online) The set of initial conditions initially on the right of the hole, which collide and reflect on the right semicircular segment of the stadium and do not cross over the hole are defined by the boundaries of validity. These are shown here for $n = 0, 1, 2$. The dotted diagonal lines are given by $\theta_1 = (a - x_i)/2r(n + 1)$ and separate the plane into the relative areas of n . The top triangle of each spike (dark blue) is for $f = 3$ (one collision on semicircle) while the bottom (light blue) is for $f = 4$ (two collisions on semicircle). They are separated by the straight lines given by $\theta_1 = 3(a - x_i)/2r(3n + 2)$. The remaining two sets of lines which define the spikes are given by the solutions of equations (14) and (15). Notice that all spikes have the same maximum height of $3z = 3\arctan(\epsilon/4r)$.

To motivate what is to follow we have a look momentarily to Figure 4 below, which on its left panel shows a numerical simulation which identifies the set of initial conditions (x_i, θ_1) which survive until time $t = 50$. Notice that the peaks of these spikes increase in height as we move from left to right away from the end of the flat segment at $x = a$, and therefore in a sense increasing the number of pre-reflection collisions n .

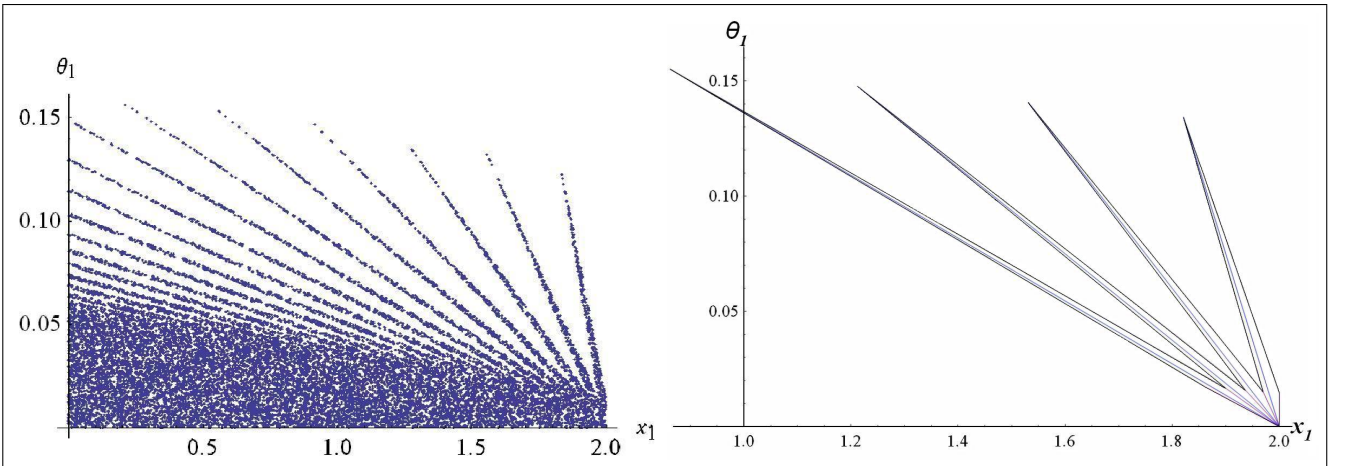


FIG. 4: *Left*: Numerical simulation identifying the set of initial conditions that survive until time $t = 50$. *Right*: Area enclosed by the hyperbolas for times $t = 50$, for $n = 0, 1, 2$ and 3 . The parameters used in these simulations were: $a = 2, r = 1, \epsilon = 0.2, h_1 = -\epsilon, h_2 = 0$.

To find the survival probability function, $P_2(t)$, of these trajectories we must now consider $T(x_i, \theta_1)$ as the parameter t . Hence we can rearrange equation (13) to form two expressions which depend not only on x_i , θ_1 and n , but also on the time t .

$$f_3(x_1, \theta_1, t) \equiv (a - x_i) \left(\frac{2d_1}{r} - 3\theta_1 \right) - (a - h_2)\theta_1 + (\Delta_3 - t) \left(\frac{2d_1}{r} - 3\theta_1 \right) \theta_1 = 0 \quad (16)$$

$$f_4(x_1, \theta_1, t) \equiv (a - x_i) \left(\frac{4d_1}{r} - 5\theta_1 \right) + (a - h_2)\theta_1 + (\Delta_4 - t) \left(\frac{4d_1}{r} - 5\theta_1 \right) \theta_1 = 0 \quad (17)$$

These expressions are conic sections as they are quadratic in both x_i and θ_1 and describe hyperbolas in the plane of initial conditions. The two hyperbolas, (16) and (17), approach each other as an effect of increasing the time t and tilt and shift discontinuously when increasing n . We notice that if we impose these hyperbolas onto the boundaries of validity we found earlier (see Figure 3), we are essentially imposing a time constraint on the set of initial conditions which will survive up to time t . Their effect will be for smaller n to erode the area enclosed by the inequalities, therefore sharpening them, and for larger n to thicken them from either side, effectively shaping them into a series of spikes, allowing for larger and larger values of θ_1 as we increase n from left to right.

This effect can be seen on the right of Figure 4, where the boundaries of validity for $n = 0, 1, 2$ and 3 are eroded by the hyperbolas which are time dependent, causing each spike to grow taller as we move away from the edge of the straight segment of the billiard. It can be easily seen that the pictures in Figure 4 are almost identical to a very small error, confirming that we are measuring the correct set of initial conditions and therefore the orbits they describe. The thickening effect is a consequence of the survival probability function's set up. Trajectories which fall just outside of the area enclosed by the boundaries of validity but for large enough n , are trapped between the two hyperbolas, will not eventually escape through the hole, i.e. they will jump over it, but they will still survive until the given time t .

We notice that for finite t , n is also finite. The key of the relation between t and n lies in the conic section equations (16)-(17). Solving them for n we discover that for $n = \frac{-3r+2t-2\Delta_3}{4r}$ and $n = \frac{-5r+4t-4\Delta_4}{8r}$ respectively, the conic sections are no longer hyperbolas but turn into negative parabolas. Hence, we can define the maximum number of pre-reflection collisions for finite time as:

$$N_{max}(t) = \left\lfloor \min \left\{ \frac{-3r+2t-2\Delta_3}{4r}, \frac{-5r+4t-4\Delta_4}{8r} \right\} \right\rfloor = \left\lfloor \frac{-5r+4t-4\Delta_4}{8r} \right\rfloor \quad (18)$$

where the lower square brackets $\lfloor h \rfloor$, are defined as the integer part of h (also known as the floor function). Actually, as we shall find out in the next section, this term (N_{max}) is never reached in practice (see N_3 in (21) below).

It seems that we have the area of interest well defined and bounded. We thus need to integrate over the area of each spike and then sum them all up to N_{max} for any given t . Multiplying the result by a factor of 2 (vertical symmetry), would eventually give the measure of the orbits initially on the right of the hole moving away from it.

Integrating hyperbolas and then summing their enclosed areas is a lengthy and unpleasant process. This calculation has been done numerically and an accurate result has been obtained successfully and is presented in section VII. However, this calculation can only be carried out numerically, as an analytical result is in our opinion impossible to obtain. Therefore we shall present a simpler approximation method for $P_2(t)$, which is analytically tractable, but still accurate to leading order in $1/t$.

V. APPROXIMATING HYPERBOLAS

In this section we will take equations (16) and (17) and argue that for large enough times t , the sections of the hyperbolas that are of interest can be simply and accurately described by straight lines. We can visually confirm this from Figure 4; however, further investigations have shown that the distance between the foci of each hyperbola converges to zero faster than the lengths of the integration limits (on $\theta_1 = (a - x_i)/2rn$ and $\theta_1 = (a - x_i)/2r(n + 1)$) as $t \rightarrow \infty$. In fact, for any n we find that the corresponding rates are $\sim t^{-1.5}$ and $\sim t^{-1}$. This in turn shows that the error made by approximating hyperbolas by straight lines, after integrating and summing over n , is still negligible with respect to the leading term (t^{-2} rather than t^{-1}). This method will later be verified by further numerical simulations where we calculate the error between the approximate solution and the exact numerical result.

In this section we will use $\tau = t - \Delta$ without any subscript to avoid unnecessary confusion given that for long times, Δ will completely vanish from our results. We consider the same initial conditions as in the previous section. For the first step of the approximation method, we must find the coordinates where equations (16) and (17) meet with $\theta_1 = \frac{3(a-x_i)}{2(1+n)r}$. We also need to find the coordinates where equation (17) meets with $\theta_1 = \frac{a-x_i}{2(1+n)r}$ and finally equation

(16) with $\theta_1 = \frac{a-x_i}{2nr}$. These three points along with $(a, 0)$ define the four corners of a quadrilateral in phase space shown in Figure 5. Here are their coordinates:

$$\begin{aligned} A &= \left(a + \frac{6(a-h_2)(2+3n)r}{(4+6n)r-3\tau}, \frac{9(a-h_2)}{3\tau-(4+6n)r} \right), \\ B &= \left(a + \frac{2(a-h_2)nr}{3(2nr-\tau)}, \frac{(a-h_2)}{3(\tau-2nr)} \right), \\ C &= \left(a + \frac{2(a-h_2)(1+n)r}{3(2(1+n)r-\tau)}, \frac{(a-h_2)}{3(\tau-2(1+n)r)} \right). \end{aligned}$$

The next step is to form four equations, one for each side of the quadrilateral by using these coordinates for $n \geq 0$. Afterwards, elementary integration methods in x_1 are used to produce explicit functions for the area of each spike with t and n as the only parameters:

$$Area_1 = \frac{(a-h_2)^2 r}{(2nr-\tau)(2(1+n)r-\tau)}. \quad (19)$$

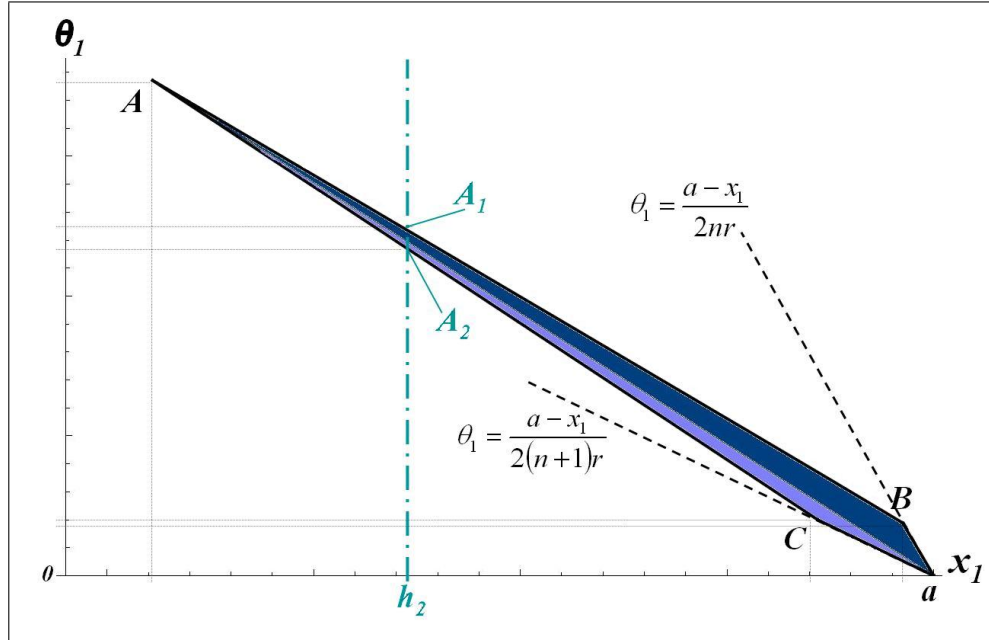


FIG. 5: (Colour online) The corners which make up the polygons which approximate the hyperbolas, for any n , are defined by the coordinates of points A, B, C and the corresponding coordinate of the endpoint of the straight segment (in this case $(a, 0)$). The x_1 and θ_1 coordinates of these corners are clearly marked by the dotted lines. The colour coding is similar to that of Figure 3. The green/dashed vertical line indicates the hypothetic position of the closest edge of the hole (in this case $x_1 = h_2$) which acts and changes the shape of the spike by defining two new corners A_1 and A_2 instead of A .

Before we insert equation (19) into a sum, we must figure out the upper limit of n for which this expression is valid. We do this by finding the smallest integer value of n for which the x_i coordinate of point A is smaller than h_2 and call it N_1 . This is because, for increasing values of n , all the corners (A, B, C) shift to the left causing the spike to tilt and stretch but when $n \geq N_1$ point A is no longer a valid coordinate. A closer look at the situation reveals that point A will split in to two points, A_1 and A_2 say, both situated on the line $x_i = h_2$. For $n > N_1$, the quadrilateral is replaced by a pentagon as the the spike's peak (point A) overshoot the vicinity of the hole's location. This process is best described diagrammatically in Figure 5 where the four corners A, B, C and $(a, 0)$ are replaced by A_1, A_2, B, C and $(a, 0)$ which are then fed back to the integration method to produce a new expression describing the area of the truncated spike. The same will happen to all corners, and different combinations of them will be necessary to produce the corresponding area functions. In other words, as the x_i coordinates of corners A, C and B , in this order, overshoots the hole's location at $x_i = h_2$, as we increase n , a new area function ($Area_j$, $j = 1 - 4$) via integration,

with a new expiration number (summation limit) via equation solving in n , will be required. This process produces three more area functions, and therefore four summations, each with different limits:

$$\begin{aligned}
Area_2 = & - \left((a - h_2)^2 \left[1152n^4r^4 + (-224r^3 + 313r^2\tau - 114r\tau^2 + 9\tau^3)\tau + 192n^3r^3(14r - 9\tau) \right. \right. \\
& \left. \left. + 4n^2r^2(512r^2 + 225\tau^2 - 762r\tau) + 4nr(128r^3 + 270r\tau^2 - 45\tau^3 - 400r^2\tau) \right] \right) \\
& \div \left(r \left(16(1+n)(2+3n)r - 3(5+8n)\tau \right) (2nr - \tau) \left(2(1+n)r - \tau \right) \left[-9\tau + 4n((4+6n)r - 3\tau) \right] \right), \quad (20)
\end{aligned}$$

$$Area_3 = \frac{(a - h_2)^2 \left(32nr^2 + 16n^2r^2 - (14r + 3\tau)\tau \right)}{4r(1+n)(\tau - 2nr) \left[-9\tau + 4n((4+6n)r - 3\tau) \right]}, \quad (21)$$

$$Area_4 = \frac{(a - h_2)^2}{4(n + n^2)r}. \quad (22)$$

Having all the *puzzle* pieces at hand, we form an expression for the invariant measure of all the initial conditions moving away from the hole from the right. The sum over the areas of quadrilaterals ($Area_1$) added to the sum of pentagons ($Area_2$) and another sum of quadrilaterals ($Area_3$) and finally the infinite sum of triangles, $Area_4$, gives:

$$Area_{Right} = \sum_{n=0}^{N_1} Area_1 + \sum_{n=N_1}^{N_2} Area_2 + \sum_{n=N_2}^{N_3} Area_3 + \sum_{n=N_3}^{\infty} Area_4, \quad (23)$$

where

$$N_1 = \left\lfloor \frac{3t - 3\Delta - 16r}{24r} \right\rfloor, \quad N_2 = \left\lfloor \frac{3t - 3\Delta - 8r}{8r} \right\rfloor, \quad \text{and} \quad N_3 = \left\lfloor \frac{3t - 3\Delta}{8r} \right\rfloor.$$

All these sums, except the second one, where simplified as follows by *Mathematica v.6*, by allowing the summation limits to acquire their non-integer values. This is allowed since the upper limit $N_l \sim t$ ($l = 1, 2, 3$), and therefore losing or gaining a term from the end of each summation effectively makes no difference whatsoever for long times.

$$\sum_{n=0}^{N_1} Area_1 = \frac{(a - h_2)^2(8r + 3\tau)}{2\tau(9\tau - 8r)}, \quad (24)$$

$$\sum_{n=N_2}^{N_3} Area_3 = \frac{32(a - h_2)^2r(192r^2 + 56r\tau + 3\tau^2)}{(8r + 3\tau)(-8r - \tau)(64r^2 - 9\tau^2 - 72r\tau)}, \quad (25)$$

$$\sum_{n=N_3}^{\infty} Area_4 = \frac{(a - h_2)^2}{4rN_3}. \quad (26)$$

The simplification of the second sum (that of $Area_2$) requires a more lengthy and tricky process, as it can not be simplified explicitly by any conventional means. This is so, not only because $Area_2$ has the most complicated of the four expressions, but also because its sum covers the largest range over n ($N_2 - N_1 \sim t/4$). Therefore, for large t , n is never small. By using a substitution of the form $u = 1/t$, assuming u to be small for large t and then substituting $n = s/u$, where s is of $\mathcal{O}(1)$, before expanding $Area_2$ into a power series effectively incorporates the effect of large n into the leading order term of the series. We get:

$$Area_2 = \sum_{k=0}^{\infty} \alpha_k u^k = - \frac{\left((a - h_2)^2(1 - 16sr + 32s^2r^2) \right) u^2}{32(s^2r(-1 + 2sr)^2)} + \mathcal{O}(u^3), \quad (27)$$

We reverse the substitution, and simplify the sum to obtain:

$$\sum_{n=N_1}^{N_2} Area_2 = (a - h_2)^2 \left(\frac{12ru(\Psi^{(0)}(z_1) - \Psi^{(0)}(z_2) + \Psi^{(0)}(z_3) - \Psi^{(0)}(z_4))}{32r} + \frac{-\Psi^{(1)}(z_1) - \Psi^{(1)}(z_2) + \Psi^{(1)}(z_3) + \Psi^{(1)}(z_4)}{32r} \right). \quad (28)$$

where

$$\begin{aligned} z_1 &= \frac{8r - 9/u - 3\Delta}{24r}, \\ z_2 &= \frac{8r + 3/u - 3\Delta}{24r}, \\ z_3 &= \frac{3(1/u - \Delta)}{8r}, \\ z_4 &= -\frac{1/u + 3\Delta}{8r}, \end{aligned}$$

and $\Psi^{(k)}$'s are polygamma functions. The polygamma function of order k is defined as the $(k+1)$ th derivative of the logarithm of the gamma function:

$$\Psi^{(k)}(z) = \frac{d^{(k+1)}}{dz^{(k+1)}} \ln \Gamma(z).$$

Fortunately the polygamma functions are of the form $z = \frac{a}{bu} + c$, where a, b and c are constants, and can be expanded as a Taylor series to leading order as follows:

$$\begin{aligned} \Psi^{(0)}\left(\frac{a}{bu} + c\right) &= \ln(a/b) - \ln u + \mathcal{O}(u), \\ \Psi^{(k \geq 2)}\left(\frac{a}{bu} + c\right) &= (-1)^{(k-1)}(k-1)! \left(\frac{bu}{a}\right)^k + \mathcal{O}(u^{k+1}). \end{aligned}$$

Substituting these expressions into equation (28) will simplify the expression dramatically, finally leaving us with the desired result. We substitute $t = 1/u$ back in to get:

$$\sum_{n=N_1}^{N_2} Area_2 = \frac{(a - h_2)^2(9 \ln 3 - 4)}{12t} + \mathcal{O}(1/t^2). \quad (29)$$

In light of equations (24),(25),(26) and (29), we can now simplify (23) to first order to get:

$$Area_{Right} = \frac{(a - h_2)^2(3 \ln 3 + 2)}{4t} + \mathcal{O}(1/t^2). \quad (30)$$

To find an expression for $Area_{Left}$ we must use the same approach used in section III to calculate equation (7). This gives:

$$Area_{Total} = 2(Area_{Right} + Area_{Left}),$$

$$Area_{Total} = \frac{(3 \ln 3 + 2)((a + h_1)^2 + (a - h_2)^2)}{2t} + \mathcal{O}(1/t^2). \quad (31)$$

Dividing by (8), the total measure of the billiard map, we can obtain an approximate result for the long time survival probability of all initial conditions initially moving away from the hole:

$$P_2(t) = \frac{(3 \ln 3 + 2)((a + h_1)^2 + (a - h_2)^2)}{4(4a + 2\pi r)t} + \mathcal{O}(1/t^2). \quad (32)$$

VI. RESULT AND NUMERICAL SIMULATION

It remains to add the probability measure of the two types of trajectories to obtain the asymptotic limit of the survival probability function:

$$P_s(t) = P_1(t) + P_2(t),$$

where the subscript s stands for the *straight* lines we have approximated the hyperbolas with. This gives:

$$P_s(t) = \frac{(3 \ln 3 + 4) \left((a + h_1)^2 + (a - h_2)^2 \right)}{4(4a + 2\pi r)t} + O(1/t^2), \quad (33)$$

which is valid only for trajectories satisfying (6), i.e. sufficiently large t .

In the left panel of Figure 6 below, we compare P_s (equation (33)) with P_d which is obtained by a *direct* numerical simulation using *Mathematica v6.*, consisting of 1.5 million initial conditions, uniformly distributed along the boundaries of the stadium billiard. We see that P_s is a good representation of the numerical survival probability for long times P_d . We have tested this result with other values of the parameters: a, r, h_1, h_2 . What is even more important however is that equation (33) is found to be an exact formula. This is shown in the right panel of Figure 6, where we have plotted the $(P_s - P_h)$ at regular intervals of time, and fitted it to an inverse time curve D/t^2 , where D is some constant. Here, P_h is the exact result obtained by numerically summing over the areas of each spike (see Figure 4), found by the integrated difference of $\theta_3(x_1, t, n)$ and $\theta_4(x_1, t, n)$ which define the *hyperbolas* in the $x_1\theta_1$ plane. We find that the $(P_s - P_h)$ fits perfectly into D/t^2 , where D needs to be calculated numerically. D is simply the coefficient of the second order term in:

$$P(t) = \frac{(3 \ln 3 + 4) \left((a + h_1)^2 + (a - h_2)^2 \right)}{4(4a + 2\pi r)t} + \frac{D}{t^2} + O(1/t^3). \quad (34)$$

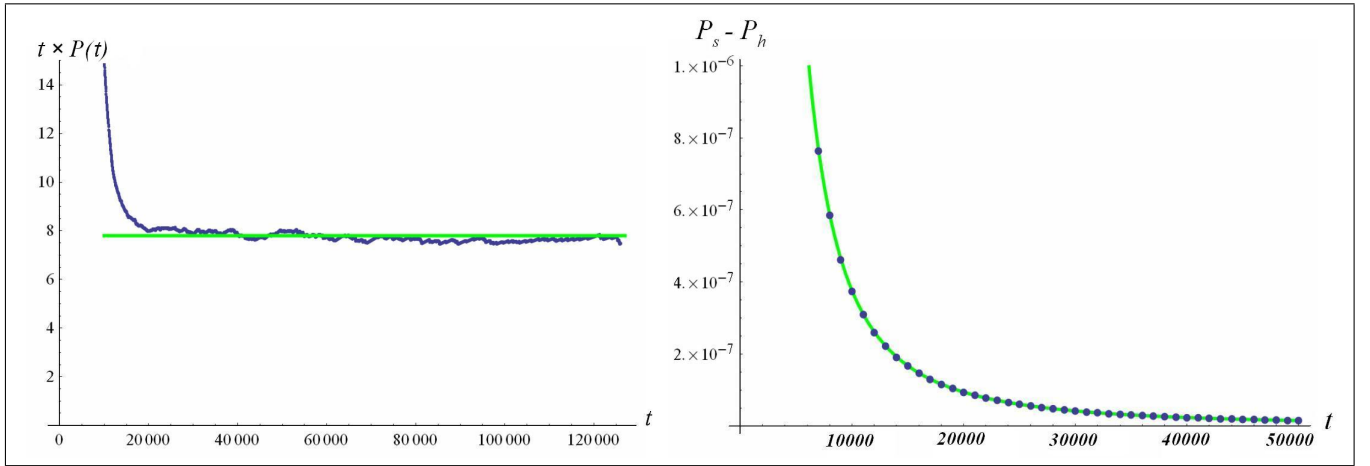


FIG. 6: (Colour online) *Left*: Plot comparing the survival probability P_s (light green/horizontal line) found by equation (33), with the numerical survival probability P_d (dark blue/curve), found by direct numerical simulation, both multiplied by the time t . *Right*: The difference between equation P_s and P_h which is found by numerically integrating hyperbolas and summing the relative areas under the spikes created, decays as D/t^2 . D is the coefficient of the second order term in equation (34).

VII. CONCLUSION AND DISCUSSION

In this paper, we have investigated the open stadium billiard and managed to derive an expression consisting of the two main contributions to the long time asymptotic tail of the survival probability function of the stadium billiard. Both expressions are exact. This has been confirmed numerically, since the error between numerical simulations and our result decays as $\sim t^{-2}$ which indicates that to leading order (33) is an exact result. In total, we confirm

that the survival probability of the stadium for long times goes as $\frac{Constant}{t}$, and we find that the *Constant* depends quadratically on the lengths of the parallel segments of the billiard and on the position of the hole but not on its size (see equation (33)).

In the context of stadia, there is a variety of possible shapes for which one can observe similar properties. In this paper we only considered the standard stadium that is a construction of two parallel straight lines and two complete semicircular arcs. It is also possible to construct different stadiums by using circular arcs of lengths less than πr or by using elliptical arcs. In both cases we expect ergodicity and strong chaotic decays followed by an asymptotic power law decay for longer times, provided that the parallel straight sides are still present, and that the dynamics remain defocusing [37]. Any similar version of the standard stadium billiard would hence result in a more or less similar asymptotic survival probability function as that of equation (33).

This is evident by the $\ln 3$ term, which first appeared in equation (29). Similar terms were found in both Gouëzel's work [27] and in Armstead's *et al.* work [33] as well as several other papers which relate to the stadium. It appears, that the $\ln 3$ term is a direct consequence of the geometry of the stadium billiard. More specifically, the curvature of the boundary near the straight segments, leads to a reflected final angle $|\theta_f| \in (|\theta_i|/3, 3|\theta_i|)$, if θ_i is small enough; this follows from equations (11) and (12) above. Hence any change to that specific section of the billiard's boundary would change the dynamics quantitatively but not qualitatively.

Further work on this subject may include researching the open stadium with holes on the circular segments. Such an example is expected to behave very similarly to the case described in the present paper as is numerically shown in [32]. This is because the trajectories which dominate and survive for long times, again will be characterised by small (near vertical) angles. Their collisions will mainly be with the straight segments of the billiard, but also on very short segments of the semicircular arcs. What is obviously different in such a situation is that the number of collisions on the semicircular arcs is not restricted to only one, as was the case here. This fact will complicate the dynamics substantially. Therefore one might prefer to choose a probabilistic approach to such a problem rather than an analytic one. One might also be interested in addressing this from a different perspective such as a semiclassical approach or even a quantum mechanical one and hopefully obtain some sort of quantitative correspondence between results.

Acknowledgements

We would like to thank Uzy Smilansky for helpful discussions, and OG's EPSRC Doctoral Training Account number SB1715.

-
- [1] Ya. G. Sinai, Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards *Math. Surv.* **25**, 137-189 (1970).
 - [2] N. Chernov and R. Markarian, *Chaotic Billiards*, AMS, (2006).
 - [3] G. Pianigiani, and J. A. Yorke, Expanding maps on sets which are almost invariant: decay and chaos, *Trans. American. Math. Soc.*, **252**, 351 (1979).
 - [4] J. Schneider, T. Tél, and Z. Neufeld, Dynamics of leaking Hamiltonian systems, *Phys. Rev. E*, **66**, 066218 (2002).
 - [5] W. Bauer and G.F. Bertsch, Decay of ordered and chaotic systems, *Phys. Rev. Lett.* **65**, 2213 (1990);
O. Legrand and D. Sornette, Fractal Set of Recurrent Orbits in Billiards, *Europhys. Lett.* **11**, 583 (1990);
O. Legrand and D. Sornette, Coarse-grained properties of the chaotic trajectories in the stadium, *Physica D* **44**, 229 (1990);
O. Legrand and D. Sornette, First return, transient chaos, and decay in chaotic systems, *Phys. Rev. Lett.* **66**, 2172 (1991);
F. Mortessagne, O. Legrand and D. Sornette, Transient chaos in room acoustics, *Chaos* **3**, 529 (1993).
 - [6] E. Doron and U. Smilansky, Chaotic spectroscopy, *Phys. Rev. Lett.* **68**, 1255 (1992);
E. Doron and U. Smilansky, Chaotic spectroscopy, *Chaos* **2**, 117 (1992).
 - [7] V. Paar and N. Pavin, Bursts in average lifetime of transients for chaotic logistic map with a hole, *Phys. Rev. E* **55**, 4112 (1997);
V. Paar and H. Buljan, Bursts in the chaotic trajectory lifetimes preceding controlled periodic motion, *Phys. Rev. E* **62**, 4869 (2000);
H. Buljan and V. Paar, Many-hole interactions and the average lifetimes of chaotic transients that precede controlled periodic motion, *Phys. Rev. E* **63**, 066205 (2001).
 - [8] V. Milner, J. L. Hanssen, W. C. Campbell, and M. G. Raizen, Optical Billiards for Atoms, *Phys. Rev. Lett.* **86**, 1514 - 1517 (2001).
 - [9] R. T. Pierrehumbert, On tracer microstructure in the large-eddy dominated regime, *Chaos Sol. Fractals* **4**, 1091 (1994).

- [10] Z. Neufeld, P. Haynes, and G. Picard, The effect of forcing on the spatial structure and spectra of chaotically advected passive scalars, *Phys. Fluids* **12**, 2506 (2000).
- [11] J. Schneider, T. Tél, and Z. Neufeld, Dynamics of “leaking” Hamiltonian Systems, *Phys. Rev. E* **66**, 066218 (2002);
J. Schneider and T. Tél, Extracting Flow Structures out of Tracer Data, *Ocean Dynamics* **53** 64-72 (2003);
I. Tuval, J. Schneider, O. Piro, and T. Tél, Opening up fractal structures of three-dimensional flows via leaking, *Europhys. Lett.* **65**, 633 (2004);
J. Schneider, V. Fernandez, E. Hernandez-Garcia, Leaking method approach to surface transport in the Mediterranean Sea from a numerical ocean model, *J. Marine Syst.* **57**, 111 (2005);
J. Schneider, J. Schmalzl, and T. Tél, Lagrangian avenues of transport in the Earth’s mantle, *Chaos* **17**, 033115 (2007).
- [12] J. Nagler, Crash test for the Copenhagen problem, *Phys. Rev. E* **69**, 066218 (2004);
J. Nagler, Crash test for the restricted three-body problem, *Phys. Rev. E* **69**, 026227 (2005).
- [13] A. E. Motter and P. S. Letelier, Mixmaster chaos, *Phys. Lett. A* **285** 127 (2001).
- [14] L. Bunimovich and C. Dettmann, Open circular billiards and the Riemann hypothesis, *Phys. Rev. Lett.* **94**, 100201 (2005);
L. Bunimovich and C. Dettmann, Peeping at chaos: Nondestructive monitoring of chaotic systems by measuring long-time escape rates, *Europhys. Letts.* **80** 40001 (2007).
- [15] L. A. Bunimovich, The ergodic properties of certain billiards. *Funk. Anal. Prilozh.* **8**, 73-73 (1974).
- [16] L. A. Bunimovich, On ergodic properties of nowhere dispersing billiards, *Comm. Math. Phys.* **65**, 295-312 (1979).
- [17] N. Chernov and C. Haskell, Nonuniformly hyperbolic K-systems are Bernoulli, *Ergodic Theory and Dynamical Systems*, **16**, 19-44 (1996).
- [18] S. Bleher, E. Ott, and C. Grebogi, Routes to chaotic scattering, *Phys. Rev. Lett.* **63**, 919 - 922 (1989).
- [19] G. Benettin and J. M. Strelcyn, Numerical experiments on the free motion of a point mass moving in a plane convex region: Stochastic transition and entropy, *Phys. Rev. A*, **17**, No 2, 773-785 (1978).
- [20] P. V. Elyutin, Kolmogorov entropy of billiard “stadiums”, *Dokl. Akad. Nauk SSSR*, **291**, No 3, 595-598 (1986).
- [21] L. A. Bunimovich, Mushrooms and other billiards with divided phase space, *Chaos*, vol. **11**, 802-808, (2001).
- [22] L. A. Bunimovich, Conditions of Stochasticity of Two Dimensional Billiards, *Chaos*, vol. **1**, 187-193, (1991).
- [23] A. Bäcker, Quantum chaos and quantum ergodicity, *Ergodic Theory, Anal. and Eff. Sim. of Dyn. Sys.*, B. Fiedler (ed.), 717-752, Springer-Verlag Berlin/Heidelberg (2001).
- [24] A. Hassell, L. Hillairet, Ergodic billiards that are not quantum unique ergodic , pre-print, arXiv:0807.0666v3 [math.AP] (2008).
- [25] G. Carlo, E. Vergini and P. Lustemberg, Scar functions in the Bunimovich stadium billiard, *J. Phys. A: Math. Gen.* **35** 7965-7982 (2002).
- [26] G. Tanner, How chaotic is the stadium billiard? A semi-classical analysis, *J. Phys. A: Math. Gen.* **30**, 2863-2888, (1997).
- [27] P. Bálint and S. Gouëzel, Limit Theorems in the Stadium Billiard, *Comm. in Math. Phys.*, **263**, Vol 2, 461-512 (2006).
- [28] B. Chirikov, Poincaré Recurrences in Microtron and the Global Critical Structure, pre-print, arxiv:nlin/0006013v1 [nlin.CD] (2000).
- [29] G. Zaslavsky, Fractional Kinetics, and Anomalous Transport, *Physics Reports* **371**, 461 (2002).
- [30] A. S. Pikovsky, Escape exponent for transient chaos and chaotic scattering in non-hyperbolic Hamiltonian systems, *Phys. A Math. Gen.* **25**, 477-481, (1992). **30**, 2863-2888 (1997).
- [31] E. G. Altmann and Tamás Tél, Poincaré recurrences and transient chaos in leaked systems, pre-print, arXiv:0808.3785 (2008).
- [32] R. S. Dumont and P. Brumer, Decay of a chaotic dynamical system, *Chemical physics letters*, vol. **188**, 565-571 (1992).
- [33] D. N. Armstead, B. R. Hunt and E. Ott, Power-law decay and self-similar distributions in stadium-type billiards , *Physica D* **193**, 96-127 (2004).
- [34] F. Vivaldi, G. Casati and I. Guarneri, Origin of Long-Time Tails in Strongly Chaotic Systems, *Phys. Rev. Lett.* **51**, 727 - 730 (1983).
- [35] M. V. Berry, ‘Regular and Irregular Motion’ in *Topics in Nonlinear Mechanics*, ed. S. Jorna, Am.Inst.Ph.Conf.Proc No.46, 16-120, (1978).
- [36] K.C. Lee, Long-Time Tails in a Chaotic System, *Phys. Rev. Lett.* **60**, 1991 - 1994 (1988).
- [37] L. A. Bunimovich and G. Del Magno, Semi-focusing billiards: ergodicity, Cambridge University Press, *Ergodic Theory and Dynamical Systems*, **28**, 1377-1417 (2008).