SPECULAR AND NO-SLIP BILLIARDS WITH CUSPS

Clayton Boone, Christopher Cox, and Ed Smith
Tarleton State University
Department of Mathematics
Box T-0470, Stephenville, TX 76401
clcox@tarleton.edu

Abstract

Billiard dynamical systems, in which particles move freely inside a closed domain and reflect elastically upon colliding at boundaries, have proven widely applicable over the last century in providing physical models for a variety of systems and in providing a testing ground to better understand many questions in pure mathematics. We numerically investigate the behavior of a family of billiard tables with cusps, considering both the much studied specular collision model as well as the lesser known alternative no-slip model, in which angular and linear momentum may be exchanged when rotating particles collide with boundaries.

1 Introduction

Billiards systems are mathematical models in which particles move freely and without dissipation of energy in a two-dimensional domain until hitting a boundary, then reflect according to a collision law. The most often studied model uses specular collisions, sometimes called optical collisions, in which the angle of reflection at a collision is equal to the angle of incidence, relative to the tangent line at the point of collision. From this simple law a vast tapestry of behaviors emerges. For example, if two edges come together in a cusp, the particle may undergo a large number of collisions in the cusp before exiting. Investigating some of the consequences of this behavior is the focus of this paper.

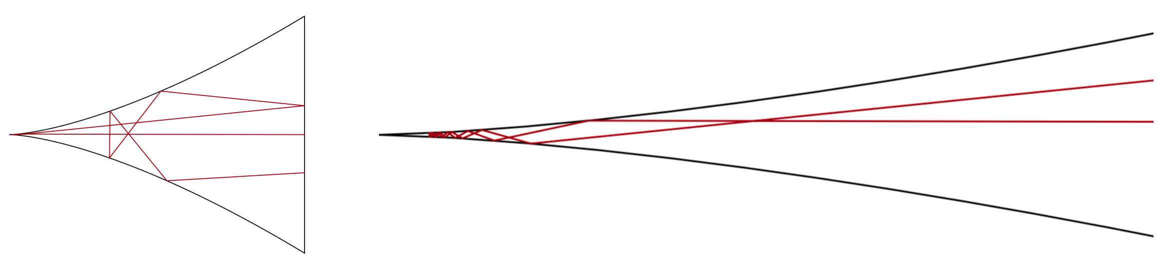

Figure 1: Left: A cusp billiard, starting near the center of the right vertical wall and ending on the lower half of the same wall. Note that at collisions, the entry angle relative to the tangent equals the exit angle. Right: A close-up of the cusp shows the particle actually collides many times before exiting.

As a tool for understanding dynamics, billiard systems are nearly as old as the field of dynamical systems itself, dating back at least to George D. Birkhoff (see for example his 1927 paper [3]). The study of mathematical billiards has consistently spawned whole new areas of inquiry over the last century, including polygonal billiards [15], convex billiards [1], chaotic billiards [7], and quantum billiards [2], as well as a host of generalizations. As an explicit model, billiards have found applicability in a wide range of fields, in recent years even extending to such unlikely realms as microbiology [18] and motion planning.

Meanwhile, in statistical mechanics, billiard dynamics has emerged as a tool both for fundamental inquiry and for understanding physical models. Foundational to this pursuit was understanding *ergodic* billiards, a class of billiards in which starting from almost any state the system may arrive at almost any other. (We will define ergodicity more formally in Section 2.) Perhaps counterintuitively, the chaotic behavior of ergodic billiards is precisely what facilitates a deeper understanding of certain aspects of their behaviour. Of central importance is the fact that while, on a microscopic or short-term scale, such systems are deterministic, on a macroscopic or long-term scale they are effectively random, allowing the use of many well-known statistical techniques. We will use this idea in Section 2 in our discussion of Birkhoff sums.

These directions were opened last century in part through the efforts of Sinai, who showed that dispersing billiards are ergodic [17] and of his student Bunimovich, who showed that certain convex billiards meeting defocusing criteria, most famously his eponymous stadium, are also ergodic [5].[1]

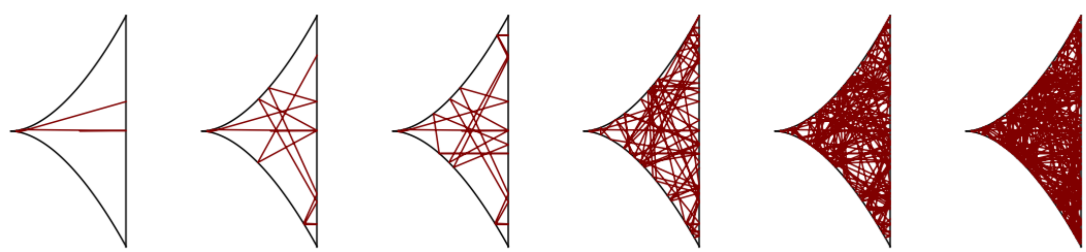

Figure 2: A cusp billiard with walls given by $y = \pm x^{1.8}$, iterated with identical initial collisions through (left to right) 10, 20, 30, 100, 200, and 300 collisions. As the progression suggests, the billiard satisfies Sinai dispersing criteria and is ergodic. However, the orbits may include long periods in the cusp, as shown in the leftmost figure in which eight of the ten collisions occur within the cusp, possibly leading to a slow mixing rate.

The cusp billiards of interest in this paper have one vertical boundary and two other sides defined by a portion of the curves $y = \pm\alpha x^{\beta}$, where $\alpha > 0$ and $1 < \beta \leq 2$. Such billiards (using

[1]Forty-four years later, Bunimovich showed that these are are in fact the only mechanisms for creating ergodicity in billiards with the standard specular collision model [6].

the specular collision model) are ergodic by the Sinai criteria: the two curved edges are concave, creating dispersion, while the third edge is flat, effectively neutral. (See Figure 2.) Nonetheless, ergodicity makes no guarantee about the rate of mixing, only about what happens in the limit. It can be shown that the particle will never become permanently stuck in the cusp (see, for example, Section 2.4 of [7]), however it may be slowed for long periods. These conflicting properties create a billiard worth studying.

In Section 2 we review the collision map, phase portraits, and ergodic billiards, and give results of numerical experiments. In particular, we look at Birkhoff sums for cusp billiards with β between 1 and 2, a case where the behavior is still not understood analytically. In Section 3, we consider cusp billiards with the no-slip collision model. After a brief review of the no-slip model, we introduce a projection for looking at phase portraits of no-slip billiards. We then give numerical evidence that no-slip cusp billiards are not ergodic.

2 Specular billiards with cusps

In this section we review the billiard map and phase portraits before formally defining ergodicity and showing how it can be informally determined by looking at phase portraits, including in the case of cusp billiards. We conclude the section by looking at one example of how deterministic billiards can mimic a random process, and give the results of numeric work on cusps investigating Birkhoff sums and random processes for cases still not understood analytically.

2.1 The billiard flow, the billiards map, and phase portraits

Looking at the phase space can be a useful tool in understanding the dynamics of a billiard system. Generally in mechanical systems, the phase space is comprised of the set of all possible positions and momenta, but here we make the assumption, common in billiard dynamics, of unit mass. Hence, for billiard tables with two dimensional domains, the phase space corresponds to the set of positions and velocities, and a priori is four dimensional. The dynamics of the billiard then might be encapsulated in the *billiard flow*, a continuous map (smooth except for collisions at which the velocity is usually discontinuous) with image in the complete phase space, a subset of $\mathbb{R}^4$. The flow is can be viewed as a parametric curve giving the position and velocity at any time t,

$$\Phi(t) = ((x(t), y(t)), (\dot{x}(t), \dot{y}(t))).$$

However, two observations allow us to reduce the dimension of the system. First, nothing particularly interesting happens between collisions in the usual model.[2] In fact, knowing the list of all collision points gives no less information than knowing all the trajectories.

[2]More interesting behavior between collisions might be observed if the billiards are placed on a more interesting manifold, with curvature, or if an external force is added. Indeed, such models are studied, but here we consider only billiards in the plane with no force.

Hence, we can parametrize the boundary and keep track of the location $r_n \in [0,S)$ where the n^{th} collision occurred. Secondly, for purposes of studying the dynamics, the speed is unimportant, so we may also make an assumption of unit velocity and only look at the angle giving direction after a collision. One might consider keeping track of the direction in a fixed frame (for example, the ambient plane $\mathbb{R}^2$), but it turns out to be convenient to let the direction be given by $\varphi_n \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, the angle relative to the inward normal at the n^{th} collision. With this we may understand the dynamics in terms of the *billiard map* (or *collision map*) $\mathcal{F}$, instead of the flow, a discrete map on the reduced collision space $\mathcal{M}$ which has dimension 2. Specifically,

$$\mathcal{M} = [0,S) \times [-\frac{\pi}{2}, \frac{\pi}{2}]$$

where S is the total length of all the edges, and

$$\mathcal{F} : \mathcal{M} \to \mathcal{M}$$

is the discrete map from one collision point to the next.

A *phase portrait* is then given by the set of all points (r_n, ϕ_n) corresponding to collisions in the orbit of some starting point. Particularly in the case of an ergodic billiard, one might wish to look at a single orbit, choosing a single starting point (r_1, ϕ_1) and running a large number of iterations of the billiards map, or alternatively (especially for non-ergodic billiards) run a large sampling of orbits, starting with many points in phase space and applying the billiard map on each for many iterations.

2.2 Ergodicity

Informally, it is generally unambiguous to describe a particular billiard as ergodic, but more accurately a measure preserving map is ergodic (or non-ergodic) with respect to a particular measure. We are interested in the *invariant* sets under the billiard map, that is, sets which are closed under the application of the map. We then define the following:

Definition 2.1. *A measure preserving map is ergodic if the only invariant sets are full measure or measure zero.*

Numerically, this implies that for a general starting point, if we continue to iterate the billiard map the orbit will almost completely fill up the phase space. There may be orbits, however, restricted to a measure zero subset. For example, there may still be periodic points, for which the orbits are limited to repeatedly passing through a finite number of points, repeating the same pattern.

For specular billiards, the canonical measure is $\cos\phi \, dr d\phi$, where r and ϕ are respectively the parametrized position and outgoing angle with respect to the inward normal for a point of collision, as previously defined.[3] A prerequisite for the consideration of ergodicity is the following lemma (see [7], Lemma 2.12).

[3] Note that one might also divide the canonical measure by $2S$ to normalize and get a probability measure, in which case invariant sets have measure zero or one in the presence of ergodicity.

Lemma 2.1. *The billiard map $\mathcal{F}$ preserves the canonical billiard measure.*

Analytically, proving that a billiard is ergodic is in general quite difficult. Indeed, as previously mentioned, there are only two known mechanisms for generating ergodic behavior, the Sinai type or dispersing billiard and Bunimovich's defocusing billiards. The former condition is easy to verify, as it merely requires that all boundaries are either concave (curving inward), resulting in dispersion, or at least linear, resulting in neither focusing nor dispersion. On the other had, Bunimovich's defocusing type encompasses a variety of possible billiards. One subset of this category, however, may be simply described: any billiard consisting entirely of arcs of circles and straight lines is defocusing (and therefore ergodic) if for any arc completing the circle would yield a circle entirely inside the billiard domain. This includes, for example, the stadium and n-petaled flower billiards. However, defocusing may also be manifested in more complicated forms, which in practice are harder to identify.

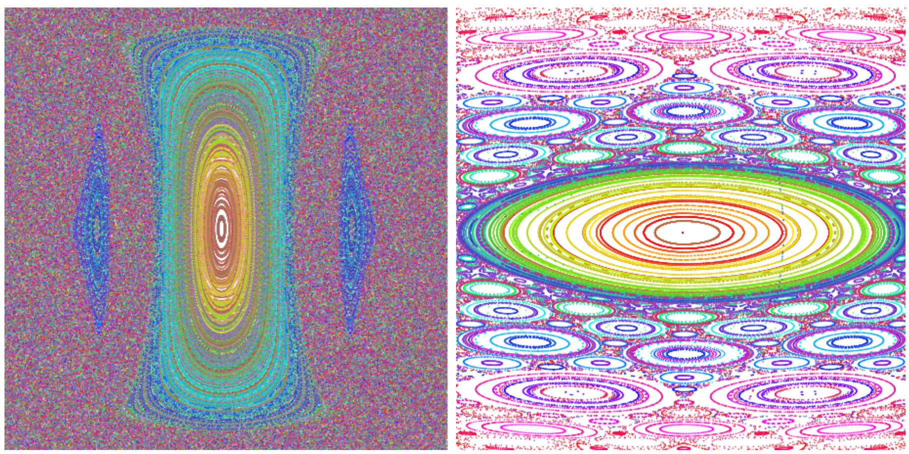

Figure 3: Colors indicate distinct orbits. Left: A partial portrait of a moon shaped billiard has elliptic islands and is not ergodic. Note however the "ergodic sea" surrounding the islands. This is a positive measure ergodic component, but since it is not full measure the billiard is not ergodic. Right: A partial phase portrait of a lemon shaped billiard comprised of two arcs is composed mostly of invariant loops around periodic points, and is clearly not ergodic.

Fortunately, looking at phase portraits often gives an accurate picture of ergodic properties. Recall that for cusp billiards, as in Figure 2, we noted that the trajectories completely fill in the billiard. While this is a necessary condition–if there were large open gaps after a large number of collisions, it is unlikely that the billiard is ergodic–it is not sufficient. It may be that starting at any point the orbit will eventually fill in the entire billiard domain, therefore hitting almost all values of r, but without hitting each with almost all possible angles. In fact, the trajectory graph might fill in while the collisions only include a very small number of angles. For example, if the billiard table is a rectangle with rational side lengths and the initial velocity has an irrational slope, almost all of the interior of the billiard will fill; however, only four angles will be used. For this reason, a better idea of ergodicity is given by investigating the phase portrait.

A strong numerical indication that a billiard is ergodic is if the entire phase space appears to fill when we start with a random point continue applying the billiard map over a large number of iterations. Of course, any actual plot will necessarily use a finite number of points of nonzero radius, rendering only a numeric approximation. If it happens that there is a very small but positive measure invariant set, its existence may not even be detectable numerically, resulting in a false positive in this test for ergodicity. (See for example [10].) Nonetheless, in most cases the phase portrait is instructive, and in particular the presence of elliptic periodic points, in which invariant loops of *quasiperiodic* points form around periodic points as in Figure 3, is a very strong indication that the billiard is not ergodic.

In the case of the cusp, there are no elliptic islands and the phase portrait suggests the billiard is ergodic, as indeed it must be by the dispersing criteria. However, this only guarantees the long term behavior will be completely mixing. In particular, the parametric value corresponding to the cusp may stand out, at least for relatively low iterations, as in Figure 4.

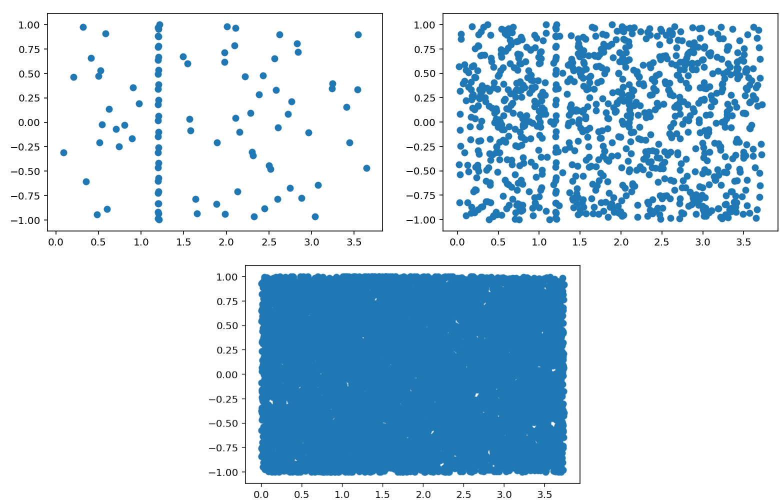

Figure 4: One orbit in the phase portrait of the cusp billiard bounded by $y = \pm\frac{2}{3}x^{\frac{3}{2}}$ along with a vertical line, plotting the cosine of the collision angle relative in the inward normal versus the parametrized point of collision, for 100, 1,000, and 10,000 collisions. Through 100 collisions, the points corresponding to the trip through the cusp dominate, but eventually the orbit mixes throughout phase space.

2.3 Birkhoff sums for cusp billiards

Looking at Figure 4, the 10,000 iteration portrait suggests that there is some sense in which the cusp billiard does indeed mimic a random a random process, while the 100 iteration portrait suggests that the imitation is far from perfect. One way to more precisely investigate this duality, an approach by no means limited to billiards systems, is through looking at *Birkhoff sums*. It is then possible to analytically investigate the convergence of a Birkhoff sum to a random process.

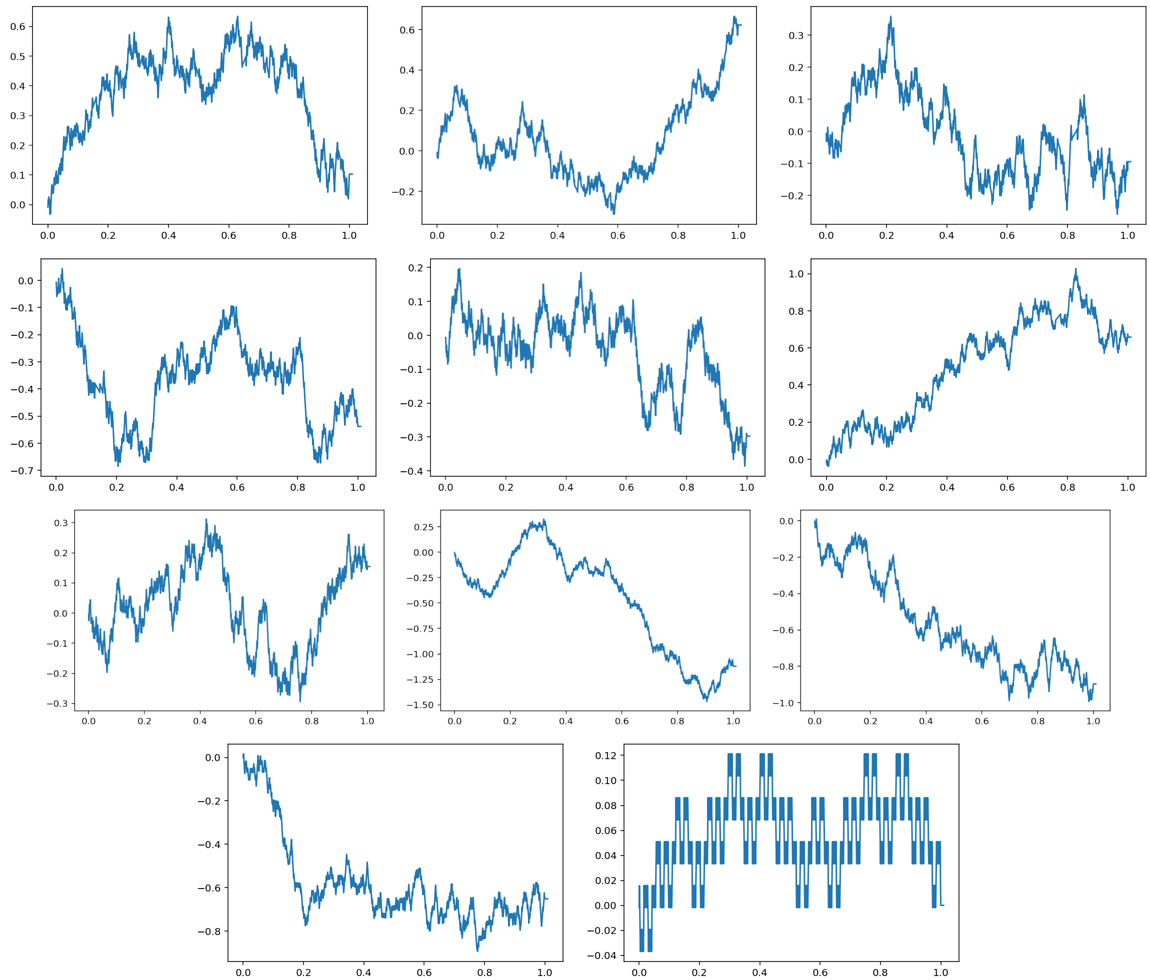

Figure 5: Birkhoff sums with $n = 2000$, and β varying from 2 in the upper left down to 1 (hence triangular) in the bottom row, varying by increments of 0.1. There is no apparently change in qualitative behavior until β reaches 1.

In a series of papers, Zhang [20, 21] along with Jung [16] have looked at questions related to convergence to random processes for billiards with flat points, including cusps of the form $y = \pm\alpha x^{\beta}$ with $\beta \geq 2$. Among their results, they prove that properly normalized Birkhoff sums converge to a particular random process, specifically an α-stable Levy process. It is not currently known if the result extends to the cases when $1 < \beta < 2$, and in particular what happens in the limiting case when β approaches 1, at which point the cusp becomes a corner and the billiard is triangular. We investigate this family of cusps numerically.

In general, a Birkhoff sum can be expressed as a real-valued function $S_n(t)$ with domain $[0,1]$ given by

$$S_n(t) = \frac{1}{c_n} \sum_{i=0}^{\lfloor nt \rfloor} f(F^i(\varphi)),$$

where φ is an observable, iterated by F, and input into a function f of a certain analytic class. For our purposes, however, the answer will not fundamentally change if we let f be the identity, $\varphi = \phi$, the angle relative to the normal, and F the billiard map. We choose $c_n = \sqrt{n}$, thus giving

$$S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=0}^{\lfloor nt \rfloor} \phi_i.$$

One qualitative characteristic of a Levy process is periodic jumps. Figure 5 gives graphs of Birkhoff sums for ten cusp billiards with β varying from 2 (which is proven analytically to converge to a stable Levy process) down to 1 at increments of one tenth. With the exception of $\beta = 1.3$, there does not appear to be a noticeable qualitative change.

3 No-slip cusp billiards

We now consider the cusp billiard with *no-slip* collisions.

In recent decades, as questions of billiard dynamics that once seemed intractable have yielded to new approaches, there has been growing interest in billiard systems with alternative collision models to the well-known specular model, in which the angle of reflection equals the angle of incidence. For example, random billiards in which collisions are determined by a probability distribution have proven useful in statistical mechanics [8, 9]. And, of interest here, no-slip billiards, using a model in which angular and linear momentum may be exchanged upon collisions at boundaries. We begin with a brief history of the model, then consider ergodicity.

3.1 The no-slip model

The no-slip model has appeared in physical models since at least as early as Garwin's 1969 paper on Super-Balls[4] [14], which includes the results of a simulation of a no-slip square using the APL programming language, the first no-slip simulation to the best of our knowledge. However, no systematic attempt to understand the dynamics of no-slip billiards was made until three decades later, when Broomhead and Gutkin [4] showed that the uniform mass no-slip strip was bounded and Wojtkowski [19] gave conditions for stable

[4]registered by Wham-O Corporation, 835 E El Monte St., San Gabriel, CA 91776

2-periodic points on a Sinai type disperser.

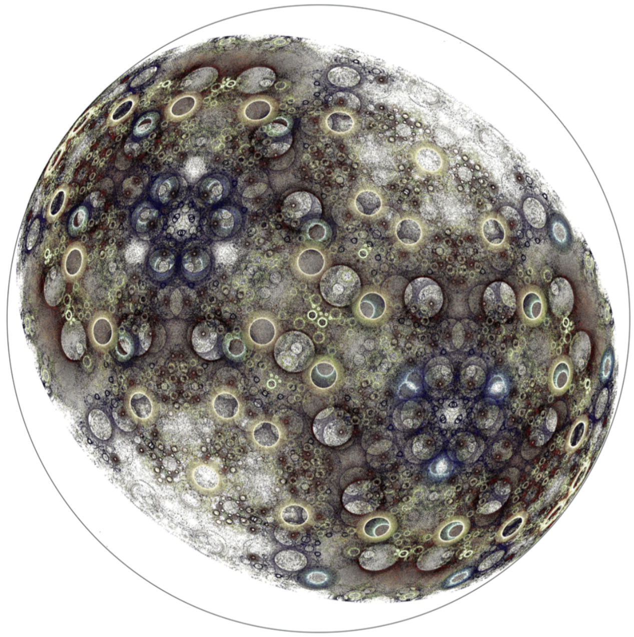

Figure 6: Elliptic islands in no-slip billiards, which preclude ergodicity, are very common. Shown here is the velocity phase portrait of a uniform mass no-slip pentagon.

In [11], the authors establish that two dimensional no-slip collisions arise naturally as a special case in a general theory of rigid body collisions, enumerating the possible reflection laws under the assumption of *strict* collision maps, requiring linearity, time reversibility, conservation of energy, and conservation of linear and angular momentum while allowing a possible exchange at collisions. Under these assumptions, there are only two possible collision laws in dimension two, standard specular collisions and no-slip collisions. Note that since the angular momentum depends on the mass distribution, the reflection law does as well. The theory in [11] allows for an arbitrary mass distribution, but commonly disks of uniform mass distribution are used, including in [4, 19, 12]. Here, we use the somewhat broader assumption of disks with radially symmetric mass distribution, parametrized by γ to include a spectrum from point mass at $\gamma = 0$ (giving specular collisions) to $\gamma = \frac{1}{\sqrt{2}}$ for uniform, to $\gamma = 1$ for particles with the mass concentrated at the outward edge.

Elliptic periodic points in no-slip billiards are ubiquitous (see [12] and Figure 6 above), and there are currently no known examples of ergodic no-slip billiards. In particular, the standard dispersing and defocusing criteria do not ensure ergodicity under the no-slip collision model. It was shown in [13] that no-slip polygons are never ergodic, always having small invariant regions. For this reason,it seems unlikely that the no-slip cusp is ergodic. In the next section we give numerical evidence in support of this hypothesis.

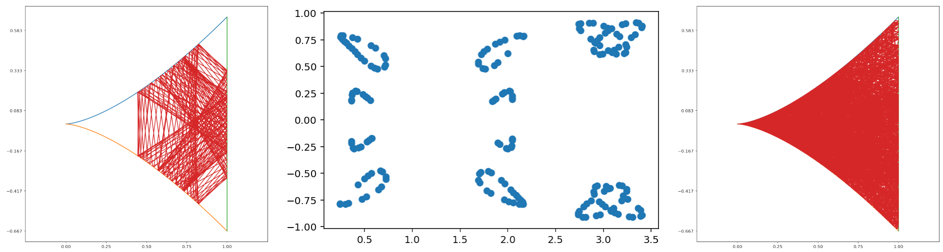

Figure 7: The trajectories (left) and phase portrait (center) for a no-slip cusp with $\beta = 1.5$ and $\gamma = 0.5$. The trajectories suggest an invariant region precluding ergodicity, while the loops in the phase portrait suggest elliptic periodic points, similarly precluding ergodicity. Note that the phase portrait is actually a projection, so that unlike in the specular case orbits may overlap. On the other hand, from the trajectories on the right it is not clear that the billiard is not ergodic.

3.2 Ergodicity in no-slip billiards

As Figure 7 shows, examples suggesting invariant sets in no-slip cusps are not difficult to find, but in other cases it is less clear that the cusp billiard is not ergodic. For that reason we focus on phase portraits. With the addition of the rotational dimension, the no-slip flow will be a curve in $\mathbb{R}^6$. However, as before we may reduce by two dimensions, by parametrizing the boundary and normalizing the velocity. Additionally, the angular position will have no bearing on the dynamics, and may be ignored. This leaves a three dimensional reduced phase space. We eliminate one more dimension–possibly losing information–by ignoring the rotational velocity, essentially projecting onto the specular phase space $\mathcal{M}$. (Or, as in Figure 6, we might project in the plane of the rotational and linear velocity.) Note that if $\mathcal{M}$ is completely filled, there may be empty components in the three dimensional phase space that will mean the billiard is not ergodic. Conversely, though, if the projection does not completely fill $\mathcal{M}$, this provides strong numeric evidence that the billiard is not ergodic. It turns out that through a wide range of values for γ and β, such gaps appear, though they do decrease in size as γ approaches zero. Apparently, no-slip cusps are never ergodic. (See Figure 8.)

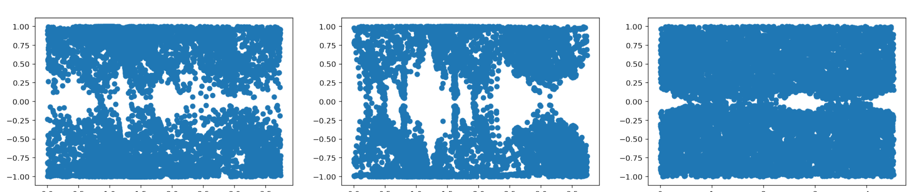

Figure 8: The projection of the no-slip cusp appears to contain gaps for any mass distribution and curvature. Notice that the gaps correspond to angles near the normal and positions far into the corners or cusp. In the right example, γ is very near to zero.

4 Acknowledgements

The paper reports results from a student/faculty project in a math modeling course on billiard dynamics at Tarleton State University, discussed in the second author's talk "Opportunities for Students in Numerical Simulation of Billiards" presented at the 31^{st} International Conference on Technology in Collegiate Math in March 2019.

References

[1] A. Avila, J. De Simoi, V. Kaloshin, *An integrable deformation of an ellipse of small eccentricity is an ellipse*. Annals of Mathematics, Second Series, 184 (2016) No. 2, 527–558.

[2] M. V. Berry, *Quantizing a classically ergodic system: Sinai's billiard and the KKR method*. Ann. Physics 131 (1981) 163–216.

[3] G. D. Birkhoff, *On the periodic motions of dynamical systems*. Acta Math, 50 (1) (1927) 359–379.

[4] D.S. Broomhead, E. Gutkin, *The dynamics of billiards with no-slip collisions*. Physica D 67 (1993) 188–197.

[5] L.A. Bunimovich, *The ergodic properties of certain billiards*. Funkt. Anal. Prilozh. 8 (1974) 73–74.

[6] L.A. Bunimovich, *Mechanisms of chaos in billiards: dispersing, defocusing and nothing else*. Nonlinearity 31 (2018) no.2, 78–92.

[7] N. Chernov, R. Markarian, *Chaotic billiards*. Mathematical Surveys and Monographs, V. 127, American Mathematical Society, 2006.

[8] T. Chumley, S. Cook, and R. Feres, *From billiards to thermodynamics*. Comput. Math. Apple 65(10) 1596-1613 (2013).

[9] S. A. Cook, R. Feres, *Random Billiards with wall temperatures and associated Markov chains*. Nonlinearity 25 (2012) No. 9, 2503.

[10] M. F. Correia, H. K. Zhang, *Stability and ergodicity of moon billiards*. Chaos, 25 083110 (2015).

[11] C. Cox, R. Feres, *Differential geometry of rigid bodies collisions and non-standard billiards*. Discrete and Continuous Dynamical Systems-A, 33 (2016) no. 11, 6065–6099.

[12] C. Cox, R. Feres *No-slip billiards in dimension two*, Dynamical Systems, Ergodic Theory, and Probability: in Memory of Kolya Chernov, Contemporary Mathematics, vol. 698, Amer. Math. Soc., Providence, RI, 2017, 91-110.

[13] C. Cox, R. Feres, H.-K. Zhang, *Stability of periodic orbits of no-slip billiards*. Nonlinearity 33 (2018) no10, 4443–4471.

[14] R.L. Garwin, *Kinematics of an Ultraelastic Rough Ball*. American Journal of Physics, 37, no. 1 January 1969, pages 88–92.

[15] E. Gutkin, *Billiards in polygons: Survey of recent results*. Journal of Statistical Physics, 83 (1996) no. 1-2, pp 7–26.

[16] P. Jung, H.-K. Zhang, *Stable laws for chaotic billiards with cusps at flat points*. Annales Henri Poincaré, 19 (2018) 12, pp 3815–3853.

[17] Y. G. Sinai, *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*. Uspehi Mat. Nauk 25 (1970), no. 2 (152), 141–192.

[18] S. E. Spagnolie, C. Wahl, J. Lukasik, J. L. Thiffeault, *Microorganism billiards*. Physica D 341 (2017) 33–44.

[19] M. Wojtkowski, *The system of two spinning disks in the torus*. Physica D 71 (1994) 430–439.

[20] H.-K. Zhang, *Decay of correlations for billiards with flat points part I: channel effect*, to appear in Contemporary Mathematics.

[21] H.-K. Zhang, *Decay of correlations for billiards with flat part II: cusp effect* to appear in Contemporary Mathematics.