

Billiards inside a cusp

Jonathan L. King

University of Florida, Gainesville 32611-2082, squash@math.ufl.edu

INTRODUCTION

Here is an anecdote about how unexciting homework problems led a student studying calculus to a Good Question, and to the mathematics it engendered.

The standard calculus curriculum spends quite a bit of time on logarithms. Yet it is no hyperbole that the curve “ $y = 1/x$ ” –and the standard drill questions concerning its properties– leave many students comatose. One such was David Feldman –then an undergraduate at Berkeley– who, having complained to his dad (the mathematician Jacob Feldman) that all the homework was dull, dull, dull, was challenged in return to invent an *interesting* problem. David came up with this:

Can a mathematical cueball (a point), fired into the symmetric funnel between $y = +\frac{1}{x}$ and $y = -\frac{1}{x}$, escape in any direction other than flat out?

As shown in figure 1 below, the cueball ricochets off the two “cushions” so that the angle of incidence always equals the angle of reflection. “Escape” means that the x -coordinate of the cueball increases *monotonically* to $+\infty$. Evidently a cueball shot along the x -axis escapes. But can any cueball which actually hits the cushions avoid being eventually turned around?

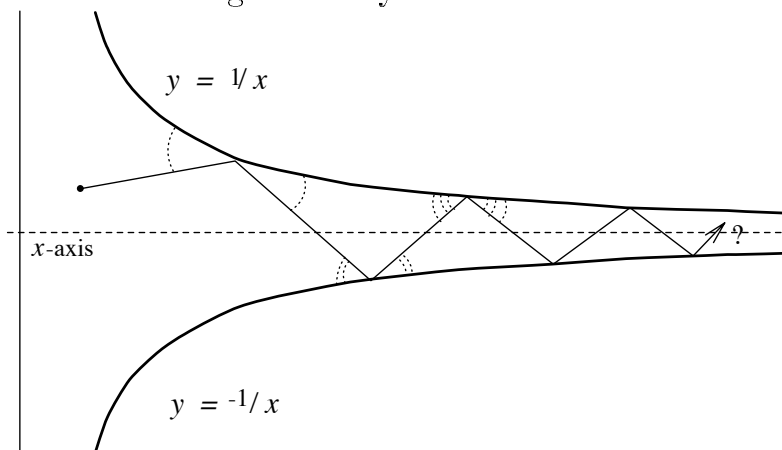


FIGURE 1 For what initial position and direction will a cueball escape to infinity?

Not long after it was posed, this problem was solved by Benjamin Weiss. Unaware of its origin, nor that it had been solved, nor that it would eventually connect with what became my field of study, I was intrigued by this problem when Paul Shields posed it to me during my graduate studies. Since the problem will lead to a harder question and to the tool which is the theme of this

Partially supported by NSF grant DMS-9112595.

article, I'm going to forthwith present the bare hands solution I found at the time—so if you want to think about the problem further, read no farther ...

A FIRST SHOT AT BILLIARDS IN A CUSP

Since the problem arose in a calculus class, to get the ball rolling let's see what information can be gleaned by methods taught in an introductory calculus course. The argument below has become a nice capstone to the section on "*Convergence tests for infinite series*" in my own classes.

Since the funnel is symmetric, we can without loss of generality reflect the trajectory over the x -axis and thus consider the cueball to be bouncing off "cushions" $y = f(x) = 1/x$ and the x -axis. So as to focus attention on the horizontal cusp at $x = +\infty$, let us do away with the vertical cusp at $x = 0$ by altering the upper-cushion f near the origin so that it is bounded. This does not affect whether a cueball can escape, since we keep that $f(x) = 1/x$ for all large x .

Suppose, for the sake of contradiction, that in figure 2 the cueball \mathbf{v} bounces so that its x -coordinate never decreases:

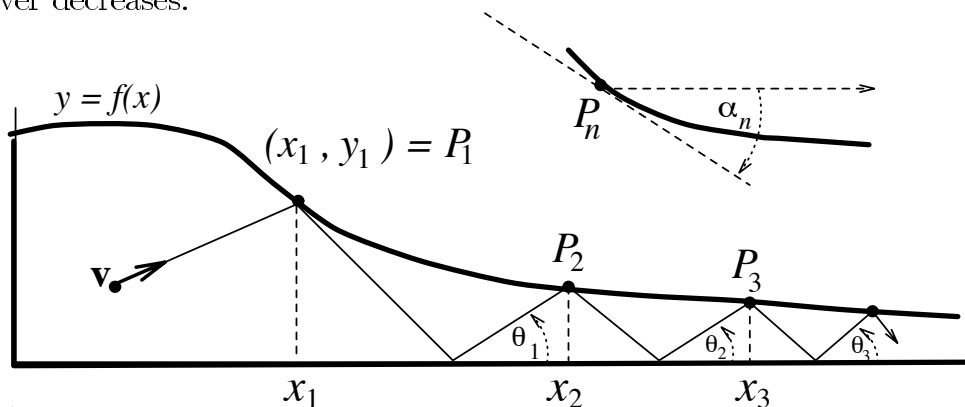


FIGURE 2 Let $P_n = (x_n, y_n)$ be the coordinates of the n th reflection of the orbit of cueball \mathbf{v} off the graph of f . Let α_n be the angle that the (tangent to the) curve at P_n makes with the horizontal ie. $\alpha_n = \arctan(|f'(x_n)|)$. After P_n the trajectory hits the "floor", from which the cueball rises at angle θ_n .

Thus $0 < \theta_n < \pi/2$ for all n , where θ_n is the angle after the n th reflection off the floor. Elementary geometry shows that $\theta_n = \theta_{n-1} + 2\alpha_n$, and so the situation in figure 2 implies this summation condition:

$$\sum_{n=1}^{\infty} \alpha_n < \infty. \quad (3a)$$

Consequently $\alpha_n \rightarrow 0$, which forces —since $f(\cdot)$ is convex— that $x_n \rightarrow \infty$. Thus $y_n \searrow 0$.

Since $\tan(\alpha)/\alpha \rightarrow 1$ as $\alpha \rightarrow 0$, summation (3a) can be restated as $\sum_{n=1}^{\infty} |f'(x_n)| < \infty$. Letting s_n be the absolute-value of the slope of the line joining point P_n with P_{n+1} , the convexity of f implies that $s_n \leq |f'(x_n)|$. Consequently,

$$\sum_{n=1}^{\infty} s_n < \infty. \quad (3b)$$

Computing the slope s_n . We have used the fact, when the cueball hits the upper cushion f , that the angles of incidence and reflection are equal. But somewhere our argument had better use that these angles are equal when the ball bounces up off the floor! Here it is:

$$\tan(\theta_n) = \frac{y_n + y_{n+1}}{x_{n+1} - x_n}.$$

The upshot is that

$$\frac{1}{x_{n+1} - x_n} = \frac{\tan(\theta_n)}{y_n + y_{n+1}} \geq \frac{\tan(\theta_1)}{y_n + y_{n+1}}.$$

As a consequence,

$$s_n \stackrel{\text{note}}{\geq} \frac{y_n - y_{n+1}}{x_{n+1} - x_n} \geq \tan(\theta_1) \cdot \frac{y_n - y_{n+1}}{y_n + y_{n+1}}.$$

So our summation condition mutates one last time, to become

$$\sum_{n=1}^{\infty} \frac{y_n - y_{n+1}}{y_n + y_{n+1}} < \infty, \quad \text{with } y_n \searrow 0. \quad (3c)$$

But this cannot be—such a sum as this last one must always be infinite. Its M th partial sum is

$$\begin{aligned} \mathcal{P}_M &= \sum_{n=M}^{\infty} \frac{y_n - y_{n+1}}{y_n + y_{n+1}} \geq \sum_{n=M}^{\infty} \frac{y_n - y_{n+1}}{y_M + y_M} \\ &\geq \frac{1}{y_M + y_M} \left(y_M - \lim_{n \rightarrow \infty} y_{n+1} \right) = \frac{1}{2}. \end{aligned}$$

Since the partial sums $\{\mathcal{P}_M\}_{M=1}^{\infty}$ do not go to zero, conditions (3c,b,a) were all impossible, as was figure 2. Any cueball shot out the cusp must turn around. ♦

Post mortem reflection. This proof can be readily shown to a second-semester calculus class and gives a non-traditional and curious use for a series-divergence test. All that is used about the upper cushion $y = f(x)$ of the table is that f is an eventually-convex differentiable function which is asymptotic to the x -axis.

However, the argument is unsatisfactory from the point of view of understanding “why” the cueball had to turn around. One test of the strength of a method of argument is whether it can be used on related questions. Suppose we remove the convexity condition and allow the upper-cushion to have wiggles.

$$\text{Can cueballs wander monotonically out the cusp for the table determined by, say, } f(x) = (3 + \sin(\sqrt{x})) / (x + 1)^2? \quad (4a)$$

While one could possibly use a series-divergence argument to exhibit a *specific* cueball which fails to escape, such an approach might require real delicacy to make a substantial general assertion.

Yet another natural question for which the series-divergence approach looks ill-adapted focuses on a stronger sense in which cueballs might fail to escape.

$$\text{Do cueballs return arbitrarily close (in both position and direction) to where they started?} \quad (4b)$$

By the way, a cueball which infinitely-often returns arbitrarily near to its initial state is called **recurrent**. Having developed more powerful tools, we will come back to recurrence later.

Philosophy. Answering questions such as (4a,b) for an individual cueball may be difficult. Yet nearby cueballs have nearby trajectories—for a while—and so it may be profitable to make assertions about *collections* of cueballs. This suggests finding a useful measure on the space of cueballs—a measure which is preserved under the action of “rolling” and “bouncing off the cushion”. It turns out—this is well-known to those who study dynamical systems but is not a commonplace among mathematicians in general—that the “billiard flow” on any billiard table has a natural invariant volume. The theme of this article is the tool of an *invariant measure* hidden inside a problem which, on the surface, has no mention of measures. Along the way we will encounter a few elementary but useful tools from dynamical systems.

Anatomy. Section 1 defines the billiard flow and gives a pictorial proof that billiard measure, which is a type of volume, is indeed invariant under the flow. Using this measure, §2 presents Weiss’s solution to Feldman’s problem when the cusp has finite area and gives an almost-everywhere solution to questions (4a,b).

In order to handle cusps with infinite area it is advantageous to view the billiard measure differently, and for that reason §3 introduces the notion of the cross-sectional measure “induced” by the billiard flow. The article culminates in §4 by using this *induced measure*, a type of area, to prove that on a “pinched” table, even a table of infinite area, almost-every cueball rolls recurrently. This result, which is illustrated in figure 17, appears to be new.

The APPENDIX contains brief connections to ergodic theory, and ends with an open problem.

History. Originally, Feldman’s Billiard Problem was part of a longer article with the same theme of hidden invariant measures. The other problems have been split off into a companion paper, *Three Problems in search of a Measure*, [1], which applies the tool of invariant measures to Poncelet’s Theorem, Tarski’s Plank Problem, and Gelfand’s Question. The APPENDIX of the current article describes a connection, in the case of an elliptical billiard table, between the induced measure of §3 and the “Poncelet measure” of [1].

Idiosyncrasy. Use “ $a := b$ ” to mean “ a is defined to be b ”. Symbol $\bigsqcup_{k=1}^{\infty} B_k$ indicates the sets $\{B_k\}_k$ in the union happen to be disjoint. For a measure of “area” or “volume”, a *nullset* will be a set which has zero area or volume. When a statement “holds almost-everywhere” (a.e.), this means that it holds except for a nullset of points.

Reflection problems such as David Feldman’s are called “billiard problems”; some curve or collection of curves form the boundary, the *cushions*, and the closed 2-dimensional region Γ that they bound is the billiard *table*. A mathematical *cueball* $\mathbf{v} = \langle v; \theta \rangle$ will be a point $v \in \Gamma$ on the table together with a direction θ . If v —sometimes called the “footpoint” of \mathbf{v} —is on the boundary of the table, then θ is restricted to the semi-circle of angles pointing into the table.

All our spaces are metric spaces. A measure-space (Ω, μ) means that μ is a Borel measure on space Ω ; all sets and functions are tacitly Borel measurable. A *transformation* $T: \Omega \rightarrow \Omega$ is a measurable map; we think of $T^n(\omega) := T(T(\dots T(\omega) \dots))$ as the location of ω at time n . A measure μ is *T-invariant*, or *T preserves* μ , if $\mu(T^{-1}S) = \mu(S)$ for each set S .

After a cueball \mathbf{v} has rolled for t seconds, let $\Phi^t(\mathbf{v})$ denote the resulting cueball. This mapping Φ is called a “flow” and satisfies that if one flows for s seconds followed by t seconds, the same result is obtained by flowing $(t + s)$ seconds. Specifically, a *flow*—which is a continuous-time analogue of a transformation—on a space Ω is a measurable map

$$\Phi: \mathbb{R} \times \Omega \rightarrow \Omega \quad \text{satisfying} \quad \Phi^t(\Phi^s(\omega)) = \Phi^{t+s}(\omega)$$

such that each Φ^t is a transformation of Ω , and Φ^0 is the identity. Saying the flow is measure-preserving means that each Φ^t is a μ -preserving transformation.

We use boldface lowercase letters for individual cueballs, e.g., \mathbf{v} , \mathbf{w} , \mathbf{e} , and boldface uppercase letters for *sets* of cueballs, e.g., \mathbf{S} , \mathbf{B} , $\mathbf{\Sigma}$, $\mathbf{\Gamma}$. We use slanted lowercase letters for (foot)points in the plane v , w , e , and slanted uppercase letters for sets of points, S , B , Σ , Γ .

§1 THE BILLIARD FLOW

Superficial Question: What is the simplest possible billiard table? *Superficial Answer:* One with no cushions. We first consider this primordial case of billiards.

Here the table Γ is the entire plane $\mathbb{R} \times \mathbb{R}$. Interpret $\mathbb{K} = [0, 2\pi)$ as the circle of directions (angles) θ , equipped with arclength measure “ $d\theta$ ”. The space of cueballs

$$\mathbf{\Gamma} := \Gamma \times \mathbb{K}$$

is thus 3-dimensional, and has a natural product-measure

$$\text{vol} := \text{area} \times \text{arclength},$$

which simply measures 3-dimensional volume. Given an arbitrary set \mathbf{S} of cueballs, its ***cross-section*** in direction θ is

$$[\mathbf{S}]_\theta := \{v \in \Gamma \mid \langle v; \theta \rangle \in \mathbf{S}\} \quad \text{and so} \quad \text{vol}(\mathbf{S}) = \int_{\mathbb{K}} \text{area}([\mathbf{S}]_\theta) d\theta,$$

by Fubini’s theorem. Cueball space $\mathbf{\Gamma}$ also has a natural topology. Letting $\theta_{\mathbf{v}}$ denote the direction of cueball \mathbf{v} , a metric on $\mathbf{\Gamma}$ is

$$\text{dist}(\mathbf{v}, \mathbf{w}) := \text{dist}(v, w) + \text{dist}(\theta_{\mathbf{v}}, \theta_{\mathbf{w}}),$$

where the righthand side uses the metrics on the plane and the “circle of directions”, respectively.

Billiard flow Φ . To write a formula for $\Phi^t(\mathbf{v})$, the location of cueball \mathbf{v} after it has “rolled for t seconds” at unit speed, interpret for a moment $\vec{\theta}$ as the unit-vector in direction θ . Then the billiard flow on the plane is the continuous map

$$\Phi^t(\mathbf{v}) := \langle v + t\vec{\theta}_{\mathbf{v}}; \theta_{\mathbf{v}} \rangle.$$

Since area is translation-invariant, the flow leaves volume invariant:

$$\begin{aligned} \text{vol}(\Phi^t(\mathbf{S})) &= \int_{\mathbb{K}} \text{area}([\Phi^t \mathbf{S}]_\theta) d\theta \\ &= \int_{\mathbb{K}} \text{area}([\mathbf{S}]_\theta + t\vec{\theta}) d\theta = \int_{\mathbb{K}} \text{area}([\mathbf{S}]_\theta) d\theta = \text{vol}(\mathbf{S}). \end{aligned} \tag{5}$$

Notice also that

$$\textit{The set of cueballs which ever flow through any particular point in the table is 2-dimensional; hence it has zero volume.} \tag{6}$$

Billiards tables with cushions. Moving to a more advanced case of billiards, we now glance at tables where reflection is possible.

Suppose our table $\Gamma \subset \mathbb{R} \times \mathbb{R}$ is the closure of an open set and whose boundary, $\partial\Gamma$, is a nice continuously differentiable curve. When a cueball hits this cushion, it keeps its tangential component of velocity but reverses its normal component. So cueball space $\mathbf{\Gamma}$ essentially consists of $\Gamma \times \mathbb{K}$ with an identification of cueball \mathbf{v}_1 with \mathbf{v}_2 if they have the same footpoint $v \in \partial\Gamma$, the same tangential component of velocity and opposite normal component.

In light of (6), one can freely permit the cushion to have finitely-many “corners” (e.g. the origin, in figure 2) simply by deleting from $\mathbf{\Gamma}$ the nullset of cueballs which ever roll into a corner. Thus the boundary $\partial\Gamma$ need only be *piecewise* continuously differentiable.

Because of the presence of corners, for a fixed t the “space map” $\mathbf{v} \mapsto \Phi^t \mathbf{v}$ can jump discontinuously as the trajectory of \mathbf{v} is moved across a corner. On the other hand, the “time map” $t \mapsto \Phi^t \mathbf{v}$ is always continuous.

The billiard flow leaves $\text{vol}(\cdot)$ invariant. In the case $\partial\Gamma$ consists of a single straight line, the argument of (5) still applies, since a reflection of the plane does not change area. Together with (6), this shows that when the table’s cushion is a polygon, billiard measure is Φ -invariant. The lemma we shall need is that for *any* cushion, volume is flow-invariant.

Billiard Lemma. *Suppose the cushion of a billiard table is piecewise continuously-differentiable. Then volume-measure is invariant under the billiard flow.*

Sketch of proof of Billiard Lemma. Since we can partition both the cushion and $\mathbf{\Gamma}$ into small pieces, it suffices to check that $\text{vol}(\Phi^{-t}\mathbf{C}) = \text{vol}(\mathbf{C})$ when the cushion is the graph of a function $f: [0, 1] \rightarrow \mathbb{R}$ which is continuously differentiable, t is some fixed time, and \mathbf{C} is a set of cueballs each of which hit f exactly once as time goes from 0 to $-t$. Actually, we need but verify this inequality:

$$\text{vol}(\Phi^{-t}\mathbf{C}) \leq \text{vol}(\mathbf{C}). \quad (7)$$

For then analogous reasoning gives the same inequality with “ $-t$ ” replaced by t and “ \mathbf{C} ” replaced by $\Phi^{-t}\mathbf{C}$. This replacement gives $\text{vol}(\Phi^t\Phi^{-t}\mathbf{C}) \leq \text{vol}(\Phi^{-t}\mathbf{C})$, forcing equality in (7).

Furthermore, we may assume that \mathbf{C} is a cube, since the cubes generate the Borel-field on $\mathbf{\Gamma}$. A “cube” is of the form $\mathbf{C} = C \times I$, where $I \subset \mathbb{K}$ is an interval of directions and C is a square in the plane. We must therefore prove that

$$\text{vol}(\mathbf{S}) \leq \text{vol}(\mathbf{C}), \quad \text{whenever } \mathbf{C} = \Phi^t(\mathbf{S}) \text{ is a cube} \quad (7')$$

as in figure 8 below.

Suppose $g: [0, 1] \rightarrow \mathbb{R}$ is a piecewise linear approximation of f , and let Φ_g^t denote the transformation of flowing for time t but bouncing off the graph of g rather than f .

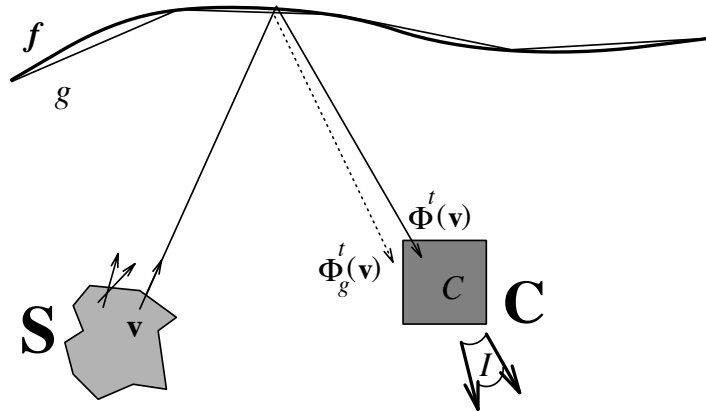


FIGURE 8 A small set \mathbf{S} of cueballs flows for time t . It bounces off the cushion, f , exactly once to form a small cube $\mathbf{C} = C \times I$, whose interval I of directions points to the southeast. Polygonal approximation g is sufficiently close to f that every cueball of \mathbf{S} hits g exactly once. The solid line shows the true trajectory of \mathbf{v} . The dotted segment shows the altered trajectory of \mathbf{v} when it bounces off of g rather than f .

Since f is *continuously* differentiable, given ε we may take g uniformly close to f in both position and slope so as to arrange, for any $\mathbf{v} \in \mathbf{S}$, that $\text{dist}(\Phi_g^t(\mathbf{v}), \Phi^t(\mathbf{v})) \leq \varepsilon$. (The nullset of cueballs \mathbf{v} which hit a vertex of g is immaterial.) This implies that $\Phi_g^t(\mathbf{S})$ is a subset of the set $\text{Ball}_\varepsilon := \text{Ball}_\varepsilon(\mathbf{C})$ of cueballs which are within distance ε of some cueball in \mathbf{C} . But billiard measure is preserved when bouncing off the polygonal cushion g and so

$$\text{vol}(\mathbf{S}) = \text{vol}(\Phi_g^t(\mathbf{S})) \leq \text{vol}(\text{Ball}_\varepsilon).$$

And the volume of Ball_ε tends to the volume of the cube \mathbf{C} , as $\varepsilon \searrow 0$. ◆

§2 A SECOND SHOT AT BILLIARDS

The series-divergence proof of the INTRODUCTION showed that if the (piecewise smooth) upper-cushion $f: [0, \infty) \rightarrow \mathbb{R}_+$ of figure 2 is eventually-convex, then the set of cueballs which escape is empty. For a more general f this **escape-set** $\mathbf{E} = \mathbf{E}(f)$, the set of cueballs \mathbf{e} such that

$$\liminf_{t \rightarrow \infty} x\text{-coord}(\Phi^t \mathbf{e}) = +\infty,$$

might not be empty but may nonetheless be small in another sense.

A finite-area cusp has a null escape-set: The Squeeze Play. Replacing the convexity of f with a finite area requirement, $\int_1^\infty f(x) dx < \infty$, allows the weaker conclusion that \mathbf{E} is a nullset.

“A gallon of water won’t fit inside a pint-sized cusp” is the proof: Pick x_0 sufficiently large that $\text{area}(S)$ is pint-sized, where S consists of those points $(x, y) \in \Gamma$ with $x \geq x_0$; indeed, $\text{area}(S)$ is to be taken so small that

$$\text{area}(S) \cdot 2\pi < \text{vol}(\mathbf{E}).$$

But $\text{vol}(S \times \mathbb{K}) = \text{area}(S) \cdot 2\pi$. Hence $S \times \mathbb{K}$ has strictly less volume than \mathbf{E} . But this contradicts that

$$\text{vol}(\Phi^t(\mathbf{E}) \cap (S \times \mathbb{K})) \rightarrow \text{vol}(\mathbf{E}), \quad \text{as } t \rightarrow \infty,$$

which follows from the definition of the escape-set. So no such x_0 exists and thus $\text{vol}(\mathbf{E}) = 0$. ◆

A second proof that $\text{vol}(\mathbf{E})$ is zero: *Recurrence*. For a continuous flow Φ on a metric space Ω , a point ω is (topologically) **recurrent** if $\Phi^{t_i}(\omega) \rightarrow \omega$ along some sequence of times $t_i \rightarrow \infty$.

Showing that almost-every cueball is recurrent would emphatically prove that the escape-set is null. We will not, however, be able to prove that \mathbf{E} is *empty*, since a table of finite area –even a bounded table– need not have all its cueballs recurrent[†].

The key to showing that a.e. cueball is recurrent is to define a measure-theoretic notion of recurrence. Consider a measure-preserving flow Φ on measure-space (Ω, μ) . A point $\omega \in S$ “recurs to S ” if $\Phi^t \omega \in S$ for arbitrarily large times t . A set S is **Poincaré-recurrent** if a.e. $\omega \in S$ recurs to S . Flow Φ is **conservative** if every set $S \subset \Omega$ is Poincaré-recurrent. Define conservativity for a transformation analogously.

Motivated by his study of the 3-body problem, Henri Poincaré made this simple, but tremendously useful, observation.

Poincaré-recurrence theorem. *If Φ is a measure-preserving flow on a finite[‡] measure space, then Φ is conservative. A measure-preserving transformation T on a finite measure space is likewise conservative.*

PROOF. Fix time $\tau > 0$ and let $B \subset S$ consist of those points which never recur to S after time τ . Thus $\Phi^t(B)$ is disjoint from B for all $t \geq \tau$. Consequently these sets

$$B, \Phi^\tau(B), \Phi^{2\tau}(B), \Phi^{3\tau}(B), \dots$$

are mutually disjoint. Since they all have the same mass, B must have been a nullset. ◆

On finite-area table, that almost-every cueball is recurrent is a consequence of the following elementary fact, which is left as an exercise.

Lemma 9. *Suppose Ω is a separable metric space and μ is a (finite or infinite) measure. Then under any conservative measure-preserving flow/transformation on (Ω, μ) , almost-every point is topologically recurrent.*

As a consequence, since table $f(x) = (3 + \sin(\sqrt{x})) / (x+1)^2$ of question (4a) has finite area, its escape-set is null. I do not know whether it is empty. One can certainly manufacture a finite-area non-convex upper-cushion f which coaxes one particular cueball \mathbf{v} monotonically out to infinity; simply draw the desired orbit of \mathbf{v} first, then draw f to match the desired slope at the reflection points of the orbit. With a bit of extra effort, one can even arrange for f to have negative slope everywhere.

Weiss’s proof of empty escape-set. Sometimes an “everywhere” rabbit can be pulled out of an “almost-everywhere” hat. A case in point is the neat proof by my friend Benjamin Weiss that under an eventually-convex f of finite area the escape-set is indeed *empty*.

[†]Consider a table bounded by a non-circular ellipse. A cueball hit along the major axis has a periodic orbit. Conversely, a cueball \mathbf{v} pointed at a focus –but with footpoint not on the major axis– has an orbit which converges to the periodic orbit along the major axis. So \mathbf{v} is not a recurrent point. Notice, though, that this example exhibits only a *nullset* of non-recurrent points, since the set of cueballs pointing at a focus is but 2-dimensional.

[‡]A flow on an *infinite* measure space need not be conservative; witness $\Phi^t(x) := x + t$ on the real line equipped with Lebesgue measure.

The strategy is to show that if even *one* cueball, \mathbf{v} , escapes, then an entire *open set* of cueballs escape. We may assume that the orbit of this $\mathbf{v} = \langle v; \theta_{\mathbf{v}} \rangle$ is already in the convex part of the cusp where the slope of f is always negative (as shown in figure 2), and that $0 < \theta_{\mathbf{v}} < \pi/2$.

Consider any cueball $\mathbf{e} = \langle e; \theta \rangle$ with $\theta > 0$ and having a shallower angle than \mathbf{v} ; by “shallower” we mean that $0 < |\theta| < |\theta_{\mathbf{v}}|$. Moreover, we ask that \mathbf{e} lie “further right” than \mathbf{v} in the sense that its footpoint, e , lie on the southeast side of the line through v in direction $\theta_{\mathbf{v}}$. Compare the orbits of \mathbf{e} and \mathbf{v} : When they first hit the upper-cushion, \mathbf{e} hits to the right of \mathbf{v} and consequently bounces off with a shallower angle than does \mathbf{v} . Thus, after they bounce off the floor, cueball \mathbf{e} is still further right, and is shallower than, \mathbf{v} . Iterating shows that \mathbf{e} escapes.

Were such a \mathbf{v} to exist, this reasoning would hold for the above open set of cueballs \mathbf{e} , which perforce has positive volume. The inescapable conclusion is that \mathbf{E} is empty. \blacklozenge

§3 POINCARÉ SECTION OF A FLOW

Imagine a large tube submerged horizontally in a river, through which water flows in some complicated way. Place a wire-mesh “surface” across the upstream end of the tube—in, say, the form of a hemisphere. Through each subregion of the mesh flows some number of gallons-per-minute, the “flux through the surface”, which therefore induces a *measure* on this surface. If we place a mesh also across the downstream end, we get a map from the upstream surface to the downstream surface simply by watching molecules of water flow from the one to the other. Since the flow preserves volume (water being incompressible), this “induced map” is flux-preserving.

The above description is meant to motivate the definition below, where cross-section $\Sigma \subset \mathbf{\Gamma}$ is a “surface”, $\text{vol}(\Sigma) = 0$, which is “transverse” to the flow in an appropriate sense. Let $\Phi^{(0,t]}(\Sigma)$ be the set of cueballs swept out as Σ flows for t seconds. (More generally, for any subset $W \subset \mathbb{R}$ of time, let $\Phi^W(\Sigma)$ denote the union of $\Phi^t \Sigma$, over all $t \in W$.) The *flux* of Σ is the limiting rate that the volume of $\Phi^{(0,t]}(\Sigma)$ grows. By this we mean

$$\text{flux}(\Sigma) := \lim_{t \searrow 0} \frac{1}{t} \text{vol}(\Phi^{(0,t]}(\Sigma)). \quad (10)$$

In order to show that this limit exists in $[0, \infty]$, we employ a type of argument which is often useful in dynamics: subadditivity. The function $V[t] := \text{vol}(\Phi^{(0,t]} \Sigma)$ is *subadditive* because

$$\begin{aligned} V[t] + V[s] &= \text{vol}(\Phi^{(0,t]} \Sigma) + \text{vol}(\Phi^{(t,t+s]} \Sigma) \\ &\geq \text{vol}(\Phi^{(0,t+s]} \Sigma) \\ &\geq V[t+s]. \end{aligned}$$

Now fix a positive t . Given any smaller positive time s , let N be the integer such that $Ns \geq t > (N-1)s$. By subadditivity, $V[s] \geq \frac{1}{N} V[Ns]$. Thus

$$\begin{aligned} \frac{1}{s} V[s] &\geq \frac{1}{Ns} V[Ns] \geq \frac{1}{Ns} V[t] \\ &\geq \frac{N-1}{N} \cdot \frac{1}{t} V[t]. \end{aligned}$$

Sending $s \searrow 0$ along any sequence sends the associated N to infinity, and so $\liminf_{s \searrow 0} \frac{1}{s} V[s]$ dominates $\frac{1}{t} V[t]$. Taking a supremum over positive t shows that the limit in (10) always exists in $[0, \infty]$, and shows that

$$\text{flux}(\Sigma) = \sup_{t > 0} \frac{1}{t} \text{vol}(\Phi^{(0,t)}(\Sigma)). \quad (11a)$$

Figure 17 illustrates how this induced measure, $\text{flux}(\cdot)$, will be used to prove the conservativity result of this article, theorem 14.

How to visualize the induced measure. An important special case is when we have a curve Σ which is a subarc of the table's boundary $\partial\Gamma$, and Σ is the set of inward-pointing cueballs with footpoint on Σ . We can get an explicit integral for $\text{vol}(\Sigma)$ by describing a cueball \mathbf{v} on the boundary in terms of “relative angle”. Write $\mathbf{v} = (v; \rho)$, where $\rho \in (-\frac{\pi}{2}, \frac{\pi}{2})$ denotes the angle that \mathbf{v} makes relative to the inward normal of $\partial\Gamma$ at v . Figure 12 shows an example in which Σ is a line-segment.

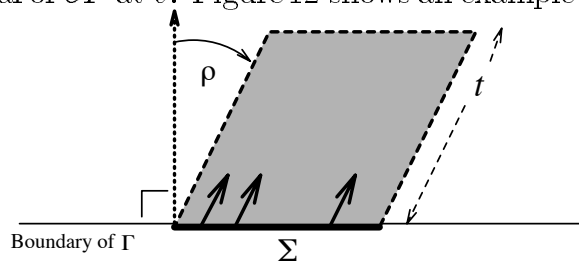


FIGURE 12 In a fixed relative direction ρ , assume that the points of Σ can be moved a distance t without encountering the cushion. Then the shaded parallelogram shows the location of the footpoints of those \mathbf{v} in Σ of relative angle ρ , after they have flowed for at most t seconds.

The area of the above parallelogram is $t \cdot \cos(\rho)$ times the length of Σ . Multiplying by $1/t$ and then integrating over ρ gives

$$\text{flux}(\Sigma) = \iint_{(v;\rho) \in \Sigma} \cos(\rho) \, d\rho \, dv, \quad (11b)$$

where “ dv ” denotes arclength measure along $\partial\Gamma$ and “ $d\rho$ ” is arclength on $(-\frac{\pi}{2}, \frac{\pi}{2})$.

It is routine to check that this formula remains valid for a general arc Σ by first approximating the arc by line segments and then sending $t \searrow 0$. This last step uses that Σ has a **first-return function** $\mathcal{R}_\Sigma: \Sigma \rightarrow [0, \infty]$ which is everywhere positive, where

$$\mathcal{R}_\Sigma(\mathbf{v}) := \sup\{t > 0 \mid \Phi^{(0,t)}(\mathbf{v}) \text{ is disjoint from } \Sigma\}.$$

It turns out that for an arbitrary set \mathbf{L} of cueballs, the condition $\mathcal{R}_\mathbf{L} > 0$ is a reasonable *definition* that cross-section \mathbf{L} is “transverse” to the flow. The proposition below is certainly plausible on physical grounds; in any case it follows from standard approximation arguments applied to (11a), and so we omit its proof.

Flux Proposition, 13. *Suppose that $\mathbf{L} \subset \Gamma$ has an everywhere positive first-return function $\mathcal{R}_\mathbf{L}$. Then*

- (a) $\text{flux}(\cdot)$ is a measure on the subsets of \mathbf{L} .

- (b) Volume-measure locally near \mathbf{L} is the product-measure of flux cross Lebesgue-measure on “time”. Specifically, suppose $\mathbf{U} \subset \mathbf{L}$ is a subset whose first-return is uniformly positive, that is, the number $\tau := \liminf_{\mathbf{v} \in \mathbf{U}} \mathcal{R}_{\mathbf{U}}(\mathbf{v})$ is positive. Then

$$\text{For any } \mathbf{S} \subset \Phi^{(-\tau, 0]}(\mathbf{U}): \quad \text{vol}(\mathbf{S}) = \int_0^\tau \text{flux}(\mathbf{U} \cap \Phi^t \mathbf{S}) dt.$$

Inducing a transformation. For an arbitrary cueball set Σ , its first-return function tells us *when* a cueball returns to Σ . The *induced map*, T_Σ , tells us where. It is defined on the subset of those $\mathbf{v} \in \Sigma$ which actually return to Σ at time $\mathcal{R}_\Sigma(\mathbf{v})$,

$$T_\Sigma(\mathbf{v}) := \Phi^{\mathcal{R}_\Sigma(\mathbf{v})}(\mathbf{v}).$$

(Of course if Σ is a closed subset of \mathbf{F} then the domain of T_Σ is simply where \mathcal{R}_Σ is finite.) As an illustration, were Σ the union of the two meshes at the ends of our submerged tube, then T_Σ would be defined just on the upstream surface and map it in a 1-to-1 fashion to the downstream surface.

As suggested by the physical situation of water flowing through a tube, the induced map preserves flux.

- (c) For an arbitrary set Σ of cueballs, the induced map T_Σ is measure-preserving wherever it is defined: If \mathbf{B} is included in the range of T_Σ then

$$\text{flux}(T_\Sigma^{-1}(\mathbf{B})) = \text{flux}(\mathbf{B}).$$

This too is an approximation argument, achieved by splitting Σ into countably many pieces whose first-return functions are nearly constant and then using that Φ flows at constant speed.

§4 CONSERVATIVITY ON AN INFINITE CUSP

A particular case where the induced map T_Σ is everywhere defined is when Σ consists of all cueballs on the boundary $\partial\Gamma$. A symbiosis exists between this induced transformation and the flow:

Transformation $T_{\partial\Gamma}$ is conservative iff Φ is conservative.

Even though Poincaré’s recurrence theorem does not apply to this transformation –the measure it preserves being infinite because $\partial\Gamma$ has infinite length– nonetheless, on a finite-area table, $T_{\partial\Gamma}$ inherits conservativity from the associated billiard flow Φ .

We conclude this article by turning the implication around, in that we will use an induced transformation to prove conservativity of the flow.

Pinched-cusp Theorem, 14. *The billiard flow under a pinched cusp, even one of infinite area, is conservative.*

A flow is *pinched* if, arbitrarily far out the cusp, there are cross-sections \mathbf{L} of arbitrarily small flux. So our billiard flow is pinched exactly when

$$\liminf_{x \rightarrow \infty} f(x) = 0,$$

since the value $f(x)$ is proportional to the flux of the set of cueballs with footpoint on the vertical line-segment going from $(x, 0)$ up to $(x, f(x))$.

As an example of a pinched cusp of infinite area, consider

$$f(x) := x[1 - \sin(x)] + \frac{1}{x+1}.$$

Even though for this cushion the supremum of $f(x)$ is infinite, nonetheless the theorem asserts that a cueball placed at a random location and then hit in a random direction will pass arbitrarily near to its starting position and direction. In contrast, it would seem difficult to show by means of the calculus technique of the INTRODUCTION that for this cushion there is even a single (non-periodic) recurrent trajectory.

Squeeze Play on an Infinite Cusp. Intuitively, conservativity on a finite-area table came from being unable to squeeze a gallon into a pint-sized bottle. This time, our bottle has infinite volume but, being vague for a moment, it still has in some sense a pint-sized neck. Our gallon of water will not be able to squeeze through this bottleneck because –if the gallon flows non-recurrently– it has an *intrinsic* positive cross-sectional flux that can never diminish.

PROOF OF THE PINCHED-CUSP LEMMA. Supposing Φ not conservative, there is some cueball set \mathbf{S} of positive volume and a positive time τ so that

$$\Phi^{[\tau, \infty)}(\mathbf{S}) \text{ is disjoint from } \mathbf{S}. \quad (15)$$

Moreover, \mathbf{S} can be taken to lie to the left of some vertical line, say, left of $x = 1$. After deleting a nullset we can assume that for all $\mathbf{v} \in \mathbf{S}$,

$$\limsup_{t \rightarrow \infty} x\text{-coord}(\Phi^t \mathbf{v}) = +\infty. \quad (16)$$

If not, then (by dropping to a positive-mass subset) all cueballs in \mathbf{S} forever stay left of some line, $x = 100$ say, and $\Phi^{(-\infty, \infty)}(\mathbf{S})$ would be a Φ -invariant set of finite volume—to which Poincaré Recurrence would apply, contradicting (15).

Consequently the situation is as figure 17 illustrates. For each positive number x , let \mathbf{L}_x denote the set of cueballs with footpoint on the vertical line-segment from $(x, 0)$ up to $(x, f(x))$ and which point to the right, ie. their directions are between $-\pi/2$ and $\pi/2$. This “line” is a 2-dimensional subset of cueball space. Because of (16), eventually $\Phi^t \mathbf{S}$ will have positive volume lying to the right of \mathbf{L}_1 . Thus there exists a time t_0 such that $\text{flux}(\Sigma)$ is positive, where

$$\Sigma := \mathbf{L}_1 \cap \Phi^{t_0}(\mathbf{S}).$$

(This follows from breaking \mathbf{L}_1 into countably many pieces \mathbf{U} and applying (13b) to each.) In addition, since assertion (15) is flow invariant, we may conclude that $\Sigma \cap \Phi^{[\tau, \infty)}(\Sigma)$ is empty.

Strategy. We have progressed from a non-recurrent set \mathbf{S} of positive volume to a non-recurrent cross-section Σ of positive flux. We will obtain a contradiction by examining the *bad* set

$$\mathbf{B} := \{\mathbf{v} \in \Sigma \mid \mathcal{R}_\Sigma(\mathbf{v}) = \infty\}$$

of cueballs which never come back to Σ .

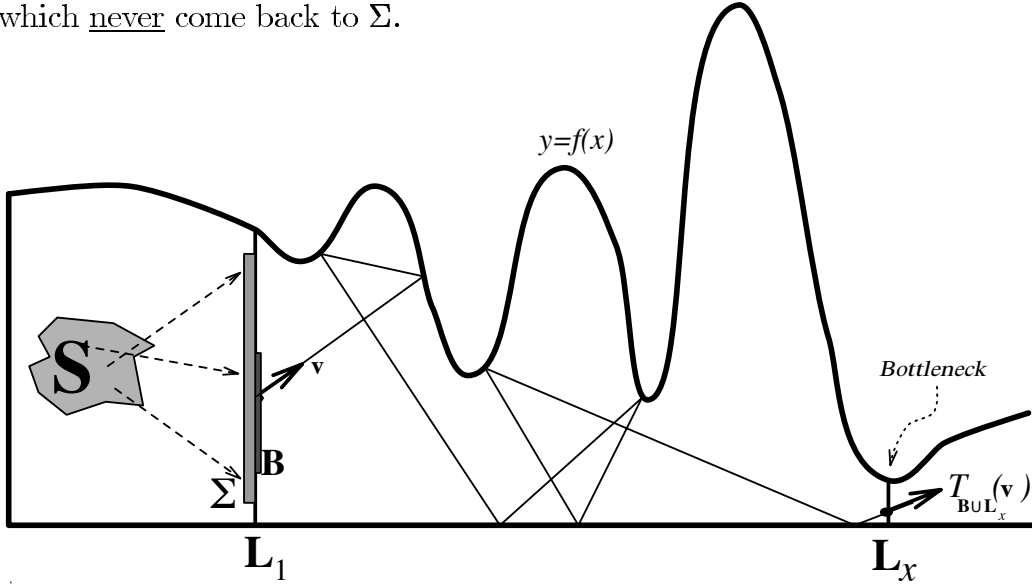


FIGURE 17 A pinched-cusp. To make them visible, sets Σ and \mathbf{B} are shown thickened—they are actually subsets of “line” \mathbf{L}_1 . The forward trajectory of a cueball $\mathbf{v} \in \mathbf{B}$ will never again touch Σ (although it might conceivably hit \mathbf{L}_1 elsewhere) and will sooner or later cross any given \mathbf{L}_x . But if \mathbf{L}_x is chosen to be a sufficiently small bottleneck, not all of \mathbf{B} will be able to squeeze through.

Every such \mathbf{v} , as (16) reminds, eventually hits any particular line \mathbf{L}_x to the right of \mathbf{L}_1 . Thus the induced map $T_{\mathbf{BUL}_x}$ is defined on all of \mathbf{B} and maps it into \mathbf{L}_x . The flow-invariance of flux now gives the key inequality that

$$\text{flux}(\mathbf{L}_x) \geq \text{flux}(\mathbf{B}), \quad \text{for all } x > 1.$$

But $\text{flux}(\mathbf{L}_x)$ is proportional to $f(x)$, by (11b). Thus the above inequality will flatly contradict that $f()$ is pinched against the x -axis, *if* we can rule out $\text{flux}(\mathbf{B})$ being zero. In order to do this, we now look at those cueballs whose behavior is antithetical to the bad set.

The infinitely-often set. Consider the set \mathbf{I} of cueballs which, under T_Σ , return to Σ infinitely often,

$$\mathbf{I} := \Sigma - (\mathbf{B} \cup T_\Sigma^{-1}(\mathbf{B}) \cup T_\Sigma^{-2}(\mathbf{B}) \cup T_\Sigma^{-3}(\mathbf{B}) \cup \dots).$$

Of necessity, T_Σ maps \mathbf{I} into \mathbf{I} and so $\mathcal{R}_\mathbf{I}$ is finite on all of \mathbf{I} . But $\mathbf{I} \subset \Sigma$ and, since $\Sigma \cap \Phi^{[\tau, \infty)}(\Sigma)$ is empty, we see that $\mathcal{R}_\mathbf{I}$ must in fact be everywhere less than the constant τ . Consequently, no cueball $\mathbf{v} \in \mathbf{I}$ could satisfy (16). Thus \mathbf{I} is empty and, in consequence,

$$\Sigma = \mathbf{B} \cup T_\Sigma^{-1}(\mathbf{B}) \cup T_\Sigma^{-2}(\mathbf{B}) \cup T_\Sigma^{-3}(\mathbf{B}) \cup \dots$$

But $\text{flux}(\Sigma) > 0$, and so some $T_\Sigma^{-n}(\mathbf{B})$ has positive flux. The third part of the Flux Proposition, (13c), implies that $\text{flux}(\mathbf{B}) \geq \text{flux}(T_\Sigma^{-n}(\mathbf{B}))$. Thus $\text{flux}(\mathbf{B})$ is positive, completing the proof of the Pinched-cusp Theorem. \blacklozenge

APPENDIX

The observation that a Poincaré-section provides a fast proof of conservativity of the billiard-flow in a convex ∞ -measure cusp, arose in a discussion with my colleague Albert Fathi. That argument lead naturally to the

generalization for flows on an arbitrary pinched cusp. Although it was illustrated here with a billiard flow, the conservativity result holds *mutatis mutandis* for any measure-preserving continuous flow whose induced measure is pinched. The illustrations in this article were drawn with the excellent computer facilities at the MATHEMATICAL SCIENCES RESEARCH INSTITUTE, whom I thank for its hospitality.

(A1) In the setting of LEMMA 9, under a continuous flow/transformation the set of recurrent points—in addition to being a full-measure set—must be *residual* (must include a dense G_δ set), once one adds the natural assumption that μ gives positive measure to every non-empty open set.

In contrast, if there is no such conservative invariant measure μ , then the set of recurrent points need not be residual. Nonetheless, Birkhoff established that *there is at least one recurrent point* under any continuous flow or transformation on a compact space. The transformation $x \mapsto x + 1$ on the topological circle $\mathbb{R} \cup \{\infty\}$ shows that no more can be guaranteed.

(A2) An unexplained coincidence occurs for flux measure in the special case where our billiard table Γ is bounded by an elliptical cushion $C := \partial\Gamma$. It turns out that the set \mathbf{C} of inward pointing cueballs breaks up into $T_{\mathbf{C}}$ -invariant subsets; one for each ellipse E which is inside of, and has the same foci as, C . The invariant set consists of those cueballs $\mathbf{v} \in \mathbf{C}$ whose flow trajectory will pass tangent to E before it again hits C .

This invariant decomposition of \mathbf{C} implies that $\text{flux}(\cdot)$, on \mathbf{C} , breaks up into measures parameterized by confocal ellipses E . When suitably normalized, each of these measures turns out to be the “Poncelet CE -measure” of [1], which arises from what appears to be an entirely unrelated construction.

(A3) Billiard flows are a kind of geodesic flow on surfaces of only zero and infinite curvature. A stronger result (see, for example, DONNAY 1988) is known for the geodesic flow on the surface-of-revolution around the x -axis generated by a differentiable $f: [0, \infty) \rightarrow \mathbb{R}_+$. If the surface is “pinched”, $\liminf_{x \rightarrow \infty} f(x) = 0$, then every geodesic orbit is bounded—except for the obvious ones which flow directly out the cusp.

(A4) An open and probably difficult research question is suggested by SULLIVAN’S 1982 result on the geodesic flow Φ on a cusp of constant negative curvature. Letting $\text{dist}(\mathbf{v})$ denote the distance of the footpoint of \mathbf{v} to some chosen point on the surface, Sullivan gives an explicit speed function $D(t)$ such that

$$\limsup_{t \rightarrow \infty} \frac{\text{dist}(\Phi^t \mathbf{v})}{D(t)} = 1, \quad \text{for a.e. } \mathbf{v}.$$

Paul Shields raised the tantalizing question of characterizing the finite-area cuspidal billiard tables which have such a speed function.

REFERENCES

- V.J. Donnay, *Geodesic flow on the two-sphere, Part I: Positive measure entropy*, Ergodic Theory and Dynamical Systems **8** (1988), 531–553.
 J.L. King, *Three Problems in search of a Measure*, Amer. Math. Monthly **101** #7 (1994), 609–628.
 D. Sullivan, *Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics*, Acta Math. **149** (1982), 215–237.