Chapter 4

Billiards

Billiards are very widely studied model systems. The study of billiards has a double parallel history. On the one hand, starting at least with G. Birkhoff, they are seen as simple examples of dynamical systems and a tool to understand issues of integrability (billiard in an ellipse, polygonal billiards) and tool to understand strongly irregular motion (Sinai and Bunimovich Billiards). We will concentrate on the second class of models here.

In general, billiards consist of a material point confined to some region of $\mathbb{R}^n$ or $\mathbb{T}^n$ with piecewise smooth boundaries;[1] in the simplest situation such a point moves with constant velocity until it reaches the boundary, and at the boundary it undergoes an elastic reflection. Such models include, e.g., a system of n hard spheres that interacts via elastic collisions (see section 4.4); the importance of such a systems as a basic model in statistical mechanics can be hardly overestimated.

These systems are conceptually extremely simple, yet they have an unpleasant feature: they lack smoothness. As we will see in the following there are three main type of non-smoothness: a) tangent collisions; b) collision with a corner; c) accumulation of infinitely many collisions in a finite time. Due to such pathologies these models, in spite of their simplicity, may present some incredibly annoying complications in their treatment.

Let $\mathbb{B}$ be the region in which the point is allowed to move and suppose that $\partial\mathbb{B}$ is a finite union of smooth manifolds with boundary. Clearly the motion can be seen as a flow ϕ_t on the unitary tangent bundle of $\mathbb{B}$ (in fact, given the initial position and the initial velocity the following motion is

[1] Although one can easily consider billiards in a region of a Riemannian manifold with piecewise smooth boundaries, in this case the motion in the interior is just the geodesic flow; see [12] for such a general setting.

uniquely determined, moreover the modulus of the velocity will be constant through the motion, so it can be assumed equal to one without loss of generality).[2]

It can be checked directly that the flow is symplectic (Hamiltonian) in $\widetilde{X} := \mathcal{B} \times \mathbb{R}^n$ (see problem Problem 4.6). So, calling m the measure induced by Lebesgue on $X \equiv \mathcal{B} \times \{v \in \mathbb{R}^n \mid \|v\| = 1\}$, (X, ϕ_t, m) is a smooth flow with collisions (crf. Examples 2.4.1).

4.0.1 Examples

Polygonal Billiards

The name is self-explanatory: the domain $\mathcal{B}$ is a polygon. The simplest case is probably a rectangle: $\mathcal{B} = [0, a] \times [0, b] \subset \mathbb{R}^2$. Although the notion is fairly trivial, to study it we will employ a neat trick that has many other applications. Consider a trajectory $x+vt$ that reaches the wall $\ell_1 := \{(a, y)\}$. The law of reflection states that, if $v = (v_1, v_2)$, the reflected velocity is $(-v_1, v_2)$. Now define the map $R_a(x, y) = (2a - x, y)$. This is a reflection (R_a^2 = identity) with respect to the wall $\{(a, y)\}$. Remark that $R_a\mathcal{B} = [a, 2a] \times [0, b]$, moreover $DR_a(-v_1, v_2) = v$. This means that, in the reflected box $R_a\mathcal{B}$, the reflected velocity is equal to the velocity before reflection.

The above algebraic discussion corresponds to a very intuitive geometrical fact: if the wall is a mirror, then the trajectory in the mirror is the continuation of the trajectory before collision.

After noticing this it is quite clear that one can understand better the trajectory in the "universal covering' of the box obtained be reflecting the box repeatedly with respect to its walls. In this covering the trajectory is simply a straight line and the trajectory in the original box is obtained by undoing the reflections (for the more mathematical inclines let us say that the plane is covered by equal boxes that are identified via reflections, see Problem 4.2). It is then obvious that, given the original velocity v only four velocities are possible: $(\pm v_1, \pm v_2)$. In fact, if we identify the opposite sides we obtain exactly a flat torus with sides twice as long as the ones of the original rectangle. In addition, the motion on such a torus corresponds precisely to the flow at unit speed in direction v. In other words, the motion is equivalent to rigid translations (geodesic flow) of the associated torus.

Accordingly, the motion is ergodic only if $\frac{v_1 b}{v_2 a}$ is irrational.

[2] A little thought will convince the reader that two motions with initial velocities that differ only in modulus will be exactly the same apart from the fact that they are run at different speeds.

Circular Billiards

In this case $\mathcal{B}$ is a disk of radius r. For convenience, let us center it at the origin of a Cartesian coordinate frame. Let us consider a point that has just collided with the boundary at the position $rn(\theta) := r(\cos\theta, \sin\theta)$, where θ is the angle with the x axis counted counterclockwise, and has velocity $v(\theta-\varphi) := (-\sin(\theta-\varphi), \cos(\theta-\varphi))$, which means that the velocity forms an angle φ with the tangent at the collision point. Accordingly, the trajectory will move along the cord of length $2r\sin\varphi$ and collide with the angle $\pi-\varphi$ which, after reflection, will be φ again.

This phenomena is nothing else than the conservation of the angular momentum (for the mechanical inclined) or of the Claroit integral (for the differential geometers).

All the above implies that, if $\frac{\varphi}{\pi} \in \mathbb{Q}$, then the motion will be periodic, otherwise the collision point will perform an irrational rotation on the boundary. In fact, let us choose as coordinates the distance τ form the last collision point computed along the trajectory; the distance s, computed along the circumference, of the last collision point from a fixed point on the circumference; and the angle φ. Then the phase space is

$$X = \{(\tau, s, \varphi) \in [0, r] \times S^1 \times [0, \pi] \mid 0 \le \tau \le 2r\sin\varphi\}$$

and the flow is noting else than a suspension flow with ceiling function $2r\sin\varphi$ constructed on the map T defined by

$$T(s, \varphi) = (s + r(\pi - 2\varphi), \varphi).$$

At the same time the middle point of the cords between two consecutive collisions will describe an irrational rotation on the circle of radius $r\cos\varphi$. This last circle is called *caustic*; the name derives from optic because if the trajectory is run by a beam of light that is the place with the highest luminosity.[3] Note that this means that the trajectory under consideration (if $\varphi/\pi \notin \mathcal{Q}$) covers densely a two dimensional torus in the three dimensional space and it is ergodic restricted to it.

[3]In ancient Greek caustic ($\kappa\alpha\upsilon\sigma\tau\iota\kappa\acute{o}\zeta$) means "that burns". Of course, that would be an important concept if you want, e.g., burn a Roman ship (to be honest, we do not know if Archimedes really knew and used burning mirrors against the Romans. Nor if he had the knowledge to do so, since his work on optic, if ever existed, has been lost. Yet, his work on conics shows that he was not so far off [2, On the sphere and the cylinder and Quadrature of the parabola]).

The above examples correspond to very regular motions ("integrable motion") that is exactly the opposite of what we mean to investigate. Unfortunately, to progress in the direction we are interested in many more technical tools are needed. Yet, before going on with general facts and definitions, let us anticipate two concrete examples that will be particularly relevant.

4.1 Sinai Billiard

The simplest example of Sinai billiards (introduced in [58] and studied in [59]) are given when $\mathbb{B} \subset \mathbb{T}^2$. More precisely, given a disk D, centered at the origin and with diameter $r < \frac{1}{2}$, let $\mathbb{B} = \mathbb{T}^2 \backslash D$. Calling $(x, v) \in \mathbb{B} \times \mathbb{R}^2$ the position and the velocity, respectively, the motion is described by a free flow

$$\phi_t(x, y) = (x + vt, v), \tag{4.1.1}$$

provided $\|x + vt\| \geq r$, that is provided the motion does not exist $\mathcal{B}$. When $x \in \partial\mathcal{B} = \partial D$ a collision takes place. Of course, at the collision, it must be $\langle x, v\rangle \leq 0$, the velocity points toward D, otherwise the point would not have reached the obstacle D but rather would be flowing away from it. The collision law is, as already said, an elastic collision–namely, the total energy and the momentum tangential to the collision plane must be preserved. Thus, calling v_- the velocity before collision and v_+ the velocity after collision, we require

$$\|v_+\| = \|v_-\| \; ; \quad \langle Jx, v_-\rangle = \langle Jx, v_+\rangle,$$

where

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

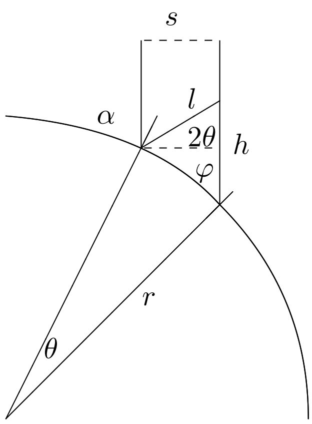
so that $\langle Jx, x\rangle = 0$, that is $r^{-1}Jx$ is a unit vector tangent to the disk and oriented counterclockwise. This implies:

$$v_+ = v_- - \frac{2}{r^2}\langle x, v_-\rangle x. \tag{4.1.2}$$

4.1.1 Flow

From the above discussion, it is clear that (X, ϕ^t, m) is a smooth flow with collisions, the only property that needs to be checked is (2.4.3).

Let us call $V(x, v) = (v, 0)$ the vector field generating ϕ^t. A useful fact is the following.



$\theta = \frac{s}{r \sin\varphi}$

$\alpha = \varphi + \theta$

Figure 4.1: Collision

Lemma 4.1.1 *If $w \in \mathcal{T}_\xi X$ and $\langle w, V(\xi)\rangle = 0$, then $\langle d\phi^t w, V(\phi^t(\xi))\rangle = 0$.*

PROOF. If no collision takes place, then the statement it is obvious by equation (4.5.9) and since for each $w = (w_1, w_2) \in \mathcal{T}X$ it must be $\langle w_1, v\rangle = 0$ (just differentiate $\|v\|^2 = 1$). Let us see what happens at collision.

Given the tangent vector $w = (w_1, w_2)$ at the point $\xi \in X$, we can consider the curve $\gamma(s) = \xi + ws$ that generates it ($\gamma'(0) = w$). Suppose that the next collision takes place with an angle φ. If we refer to the Figure 4.1 all we need to compute is the relation between h and l. A bit of geometry shows that

$$h = s \arctan\varphi + \mathcal{O}(s^2); \quad l = \frac{s}{\cos\theta}\arctan\varphi + \mathcal{O}(s^2) = s\arctan\varphi + \mathcal{O}(s^2).$$

Thus, if τ is the collision time of the trajectory starting at ξ and $\tilde{\gamma}(s) = \phi^{\tau_+}(\gamma(s))$, we have $\tilde{\gamma}'(0) = d_\xi\phi^{\tau_+}w := \tilde{w}$, and, calling v_+ the velocity after reflection, $\langle v_+, \tilde{w}\rangle = 0$, which proves the lemma. $\square$

This means that in this case there is a particularly simple way to quotient out the flow direction: consider only vectors perpendicular to the flow.

4.1.2 Reeb flows

A more general way to understand and contextualize Lemma 4.1.1 is to realize that billiards are an example of Reeb flow.

Definition 13 *Given a $2d+1$ dimensional manifold M equipped with a one form ω such that $\omega \wedge (d\omega)^d \neq 0$ (that is, a* contact manifold*), we call a Reeb* *flow a flow generated by a vector field V such that*

$$\omega(V) = 1$$
$$\omega(V, v) = 0 \text{ for all tangent vectros } v$$

Let ϕ_t be the Reb flow, generated by the vector field V.

Lemma 4.1.2 *For each $t \in \mathbb{R}$ we have $\phi_t^*\omega = \omega$.*

PROOF. By Cartan formula, we have

$$\begin{aligned}\frac{d}{dt}(\phi_t^*\omega) &= L_V(\phi_t^*\omega) = d(i_V[\phi_t^*\omega]) + i_V(d[\phi_t^*\omega]) \\ &= d(i_V[\phi_t^*\omega]) + i_V([\phi_t^* d\omega]) = d([\phi_t^* i_V\omega]) + ([\phi_t^* i_V d\omega]) = 0\end{aligned}$$

Thus ω is invariant for $D\phi_t$. □

As we want to extend the idea of Reeb flows to piecewise smooth flows, it is natural to say that a piecewise smooth flow is Reeb is $\omega(V) = 1$, where it makes sense and $\omega(D\phi_t w) = \omega(w)$, again where it makes sense.

We can then prove that Billiards are Reeb flows on the constant energy surface. First of all, note that the energy is just the Kinetic energy, hence $M_E = \{(q,p) \in \mathbb{R}^2 : \|p\|^2 = 2E\}$ is an invariant surface for the flow. Note that $(\delta q, \delta p) \in TM_E$ iff $\langle \delta p, p\rangle = 0$. We then consider the one form $\omega(\delta q, \delta p) = \frac{1}{2E}\langle p, \delta q\rangle$. Note that the vector field $V = (p, 0)$ generates the flow away from collisions, and $\omega(V) = 1$. Note that

$$\langle V, (\delta q, \delta p)\rangle = \langle p, \delta q\rangle = \omega((\delta q, \delta p)),$$

thus being Reeb automatically implies the result in the previous section as a special case. It remains to check the invariance. If the flow does not experience collisions, then

$$D\phi_t(\delta q, \delta p) = (\delta q + t\delta p, \delta p).$$

Hence,

$$\omega(D\phi_t(\delta q, \delta p)) = \langle p, \delta q + t\delta p\rangle = \langle p, \delta q\rangle = \omega((\delta q, \delta p)).$$

It remains to see what happens at a collision. First of all, note that a curve with tangent vector $(\delta q, \delta p)$ in general consists of trajectories that collide at different times. We want then to flow each point along the flow direction the exact amount that makes the curve collide simultaneously. This means that the new curve will have the tangent vector $(\widetilde{\delta q}, \widetilde{\delta p}) = (\delta q, \delta p) + \tau V$, for some τ determined by the condtion $\langle \widetilde{\delta q}, q\rangle = 0$. Yet,

$$\omega((\widetilde{\delta q}, \widetilde{\delta p})) = \omega((\delta q, \delta p)) + \tau.$$

By (4.1.2) it follows that

$$\begin{aligned}
\widetilde{\delta q}_+ &= \widetilde{\delta q} \\
\widetilde{\delta p}_+ &= \widetilde{\delta p} - \frac{2}{r^2}\left[\langle \widetilde{\delta q}, p_-\rangle q + \langle q, \widetilde{\delta p}\rangle q + \langle q, p_-\rangle \widetilde{\delta q}\right].
\end{aligned}$$

Thus, using (4.1.2) again,

$$\omega((\widetilde{\delta q}_+, \widetilde{\delta p}_+)) = \langle p_+, \widetilde{\delta q}_+\rangle = \langle p_-, \delta q_+\rangle = \omega((\widetilde{\delta q}, \widetilde{\delta p})).$$

After the collision, we have to subtract the time shear that we introduced, and this yields

$$\omega((\delta q_+, \delta p_+)) = \omega((\delta q, \delta p)$$

which is the wanted time invariance.

4.1.3 Poincaré map

For many purposes it is useful to view the Sinai billiards as a symplectic map from a two dimensional domain to itself. Such a reduction is obtained via a general technique widely used in dynamical system: a Poincaré section (see 1.2). A Poincaré section consists in introducing some codimension one manifolds in the phase space X and then defining a map from such manifolds to themselves in such a way that to each point is associated its first return to the manifolds (if it exists). Let us be more concrete.

Historically the choice of the section to realize a Poincaré map as been based on $\partial\mathcal{B}$. In our case this consists of the boundary of the disk, that is a circle. Of course, it is also necessary to specify the velocity. Clearly there are two possibilities: one can consider velocities just before collision, which means $\langle x, v\rangle \leq 0$, (this is the Poincaré map from before collision to before collision) or one can consider the velocity just after collision, meaning $\langle x, v\rangle \geq 0$, (that is the Poincaré map from just after collision to just after the next collision). The two choices are equivalent, let us make the second.

If we define the velocity by the angle φ between v and the tangent (directed clockwise) to the disk at the collision point, then the phase space is $\mathcal{M} = S^1 \times [0, \pi]$.

We can then define a map $T : \mathcal{M} \to \mathcal{M}$ in the following way: for each $\xi \in \mathcal{M}$, let $T\xi$ be the point just after the next reflection (if such a reflection exists). Note that, if no reflections would occur, almost all the trajectories would fill $\mathbb{T}^2$ densely,[4] hence T is defined almost everywhere.

It is natural to use as coordinate on the boundary the distance s, computed counterclockwise along the circle, from a given point. If we want to compute the induced invariant measure on the Poincaré section $\mathcal{M}$, we can to introduce the change of coordinates $\Xi : \mathcal{M} \times [0, \delta] \to X$ defined by

$$\Xi(s, \varphi, t) = (rn(sr^{-1}) + v(sr^{-1} + \varphi)t, sr^{-1} + \varphi - \frac{\pi}{2}),$$

where $n(\theta) = (\cos\theta, \sin\theta)$, $v(\theta) = (\sin\theta, -\cos\theta)$. In this coordinates a point is determined by its collision data (s, φ) and the time t past from the last collision.

A direct computation shows that

$$\begin{aligned}
\det \Xi &= \begin{vmatrix} -v(sr^{-1}) + r^{-1}n(sr^{-1} + \varphi)t & n(sr^{-1} + \varphi) & v(sr^{-1} + \varphi) \\ r^{-1} & 1 & 0 \end{vmatrix} \\
&= \begin{vmatrix} -v(sr^{-1}) & n(sr^{-1} + \varphi) & v(sr^{-1} + \varphi) \\ 0 & 1 & 0 \end{vmatrix} \\
&= \begin{vmatrix} -v(sr^{-1}) & v(sr^{-1} + \varphi) \end{vmatrix} = \sin\varphi.
\end{aligned}$$

So, given a set $A \subset \mathcal{M}$, calling $A_\varepsilon = \cup_{t=0}^{\varepsilon} \phi^t A$,

$$\mu(A) := \frac{1}{\varepsilon} m(A_\varepsilon) = \frac{1}{\varepsilon} \int_{A_\varepsilon} |det(\Xi)| ds d\varphi dt = \int_A \sin\varphi ds d\varphi.$$

Thus $d\mu = \sin\varphi ds d\varphi$ and $(\mathcal{M}, T, \mu)$ is a Dynamical Systems.

It is interesting to notice that μ becomes degenerate for $\varphi \in \{0, \pi\}$, which correspond to tangent collisions. Another annoying feature of the above choice is that some trajectories never meet the boundary of the disk (for example consider the initial condition $x = (1, 0)$, $v = (0, 1)$) and other will travel an arbitrarily long time before the next collision.[5] These facts, although not catastrophic, may look unpleasant to someone. It is therefore relevant to notice that there are several other possible choices for the Poincaré section, each one with is own advantages and disadvantages. Let us see a couple of them.

Consider the fundamental domain $Q = [-\frac{1}{2}, \frac{1}{2}]^2$ of $\mathbb{T}^2$, choose ∂Q as the basis of the Poincaré section. Of course, ∂Q is not a smooth manifold (it consists of four lines). This problem is easily solved by extending the

[4] Since for almost all velocities v we would have an irrational translation on $\mathbb{T}^2$.

[5] This property is called *infinite horizon*. We will discuss it further in the sequel.

concept of Poincaré section to the case in which the section Σ is a finite (or even countable) union of smooth manifolds; the reader can see that this generalization is indeed immediate. This section has the advantage of a simple structure, that there is a maximal time from Σ to itself, yet it does not solve the problem of the degeneration of the measure. Here, we also have problems with the trajectories that meet the section at very small angles.

To overcome such a problem one can choose the section $\Sigma \times [\delta, \pi - \delta]$. It is easy to see that if δ is chosen small enough then the only effect is to skip at most one crossing of the boundary Σ.

We will keep using the relation between the two dynamical systems (X, m, ϕ_t) and $(\mathcal{M}, \mu, T)$. In particular it is convenient to define the cone family on all $\mathcal{T}X$ instead that only on $\mathcal{T}\mathcal{M}$. We will see that an invariant cone family on $\mathcal{T}X$ induces an invariant cone family on $\mathcal{T}\mathcal{M}$.

4.1.4 Singularity manifolds

In this subsection we will study more precisely the singularities of the system and we will verify that they belong to the general setting developed in 10. We will consider two different Poincaré section to give the reader a more complete idea of the situation.

Let us start with the classical section *just after collision*. As already mentioned, the phase space is $\mathcal{M} = S^1 \times [0, \pi]$. Clearly, the only singularities of the map correspond to coordinates where the next collision is a tangent one. To analyze such a pathology, it is more convenient to look at the billiard on the universal covering of the torus. In such a covering, the obstacles will form a lattice and the particles moves along a straight line between collisions.[6]

The particle with coordinates (s, φ), just after collision, will move in the direction $v(r^{-1}s + \varphi)$ with unit speed.[7] Hence, if $C \in \mathbb{R}^2$ is the coordinate of the center of the obstacle with which the next collision will take place, the condition for a tangent collision reads

$$rn(r^{-1}s) + tv(r^{-1}s + \varphi) = C \pm rn(r^{-1}s + \varphi). \tag{4.1.3}$$

Where $t = t(s, \varphi)$ is the collision time. From (4.1.3) we can immediately extract two interesting informations multipling it by $n(r^{-1}s + \varphi)$ and

[6] This trick is very similar to the one used at the beginning of the chapter to discuss rectangular billiards, only now we take advantage of the periodicity of the torus rather than the invariant properties of the domain with respect to the reflections.

[7] Remember the convention $n(\theta) := (\cos\theta, \sin\theta)$ and $v(\theta) := (\sin\theta, -\cos\theta)$.

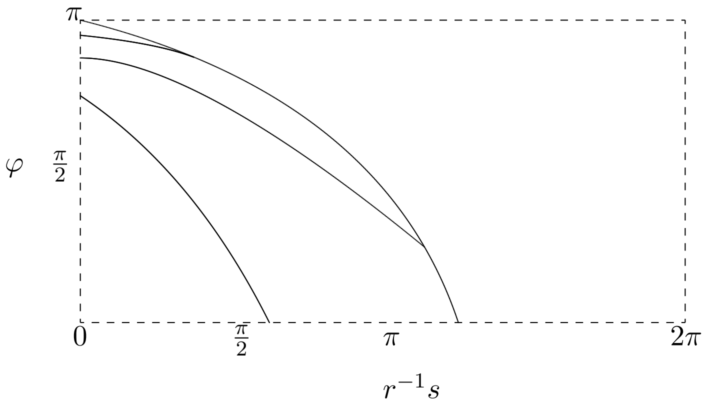


Figure 4.2: Few discontinuity lines in the Poicaré map

$v(r^{-1}s + \varphi)$ respectively

$$\langle C, v(r^{-1}s + \varphi)\rangle = t + r \sin\varphi > 0$$
$$F(s, \varphi) := \langle C, n(r^{-1}s + \varphi)\rangle - r\cos\varphi \pm r = 0$$

Taking the derivative of F with respect to φ we get

$$-r\sin\varphi + \langle C, v(r^{-1}s + \varphi)\rangle = t > 0,$$

thus we can apply the implicit function theorem and conclude that the manifold corresponding to this discontinuity can be represented as the graph of a function $\varphi(s)$. In addition,

$$\frac{d\varphi}{ds} = -(\frac{1}{r} + \frac{\sin\varphi}{t}) < 0. \tag{4.1.4}$$

Since there are infinitely many obstacles with which the next collision can take place, there must be countably many discontinuity line (some of them are schematically represented in figure 4.2)

To analize the preimages of the boundary of the section one can proced in analogy with what we have done before, equation (4.1.3) in this case beomes

$$rn(r^{-1}s) + tv(r^{-1}s + \varphi) = C \pm rn(r^{-1}s + \varphi \pm \delta). \tag{4.1.5}$$

From (4.1.5) we obtain

$$\frac{d\varphi}{ds} = -(\frac{1}{r} + \frac{\sin\varphi}{t + r\sin\delta}) < 0. \tag{4.1.6}$$

Clearly, the map is smooth up to, and including, this type of discontinuity, not so for the tangencies. In fact, it is easy to verify that the map is continuously crossing a tangency line (see Problem 4.8) but we will see immediately that it is not differentiable. By the discussion of section 5.1 (see in particular fromulae (5.1.1) and (5.1.2)) it follows that if the next collision takes place with an angle $\varphi \notin [\delta, \pi - \delta]$, then calling τ_1 the time up to the tangent collision and τ_2 the time from the tangent collision to the next, we have the formula

$$DT = \begin{pmatrix} 1 + \frac{2\tau_1}{r\sin\varphi} & \frac{2}{r\sin\varphi} \\ \tau_2(1 + \frac{2\tau_1}{r\sin\varphi}) & 1 + \frac{2\tau_2}{r\sin\varphi} \end{pmatrix}.$$

Clearly, the norm of DT is bounded by a constant time $\frac{1}{\sin\varphi}$ (if in doubt do Problem 4.9). Now, if a point has distance ε from the singularity line, it will land at a distance $\sqrt{\varepsilon}$ from the tangency, which means that there exists a constant $c_t > 0$ such that, calling $\mathcal{S}$ the singularity line and ξ the point

$$|\sin\varphi| \geq c_t\sqrt{d(\xi, \mathcal{S})}.$$

This means that the Derivative blows up only as a square root getting close to the singularity. By similar considerations, it is possible to verify also that the second derivative blows up polynomially (see Problem 4.10).

4.2 Bunimovich Stadium

These billiards have been introduced [10] and first studied [11] by Bunimovich. In this case the table of the billiard is a convex subset of $\mathbb{R}^2$. The simplest, and original, one consists in two half circles joined by two straight lines (see Figure 4.3).

The name "stadium" is due to the shape of the domain $\mathbb{B}$ in which the motion takes place. The only difference is that now the curvature of the boundary is either zero (collisions against the straight segments) or negative (collision against the half circles).

4.2.1 Flow

We have seen that the flow in a square or in a circle is well defined and rather regular. Clearly the only relevant discontinuity in the Bunimovich Billiard arise when a trajectory hit the joining between the circumference and the straight lines.

Figure 4.3: Bunimovich stadium

4.3 Different Tables, different games

Let us start with a bit of classification.

Definition 14 *Here are some standard classes of billiards:*

- dispersing *billiards are billiards with the boundary $\partial\mathbb{B}$ of the table is a finite union of stricly convex manifolds with boundary (this are often called Sinai billiards as well)*
- semi-dispersing *billiards are billiards with the boundary $\partial\mathbb{B}$ of the table is a finite union of stricly convex manifolds with boundary*
- convex *billiards are billiards where the tale $\mathbb{B}$ is a convex set.*

The remainder of the section is devoted to a more explicit description of several concrete examples of the above cases.

4.3.1 Dispersing

We have already seen the standard Sinai billiard in section 4.1. In general several convex obstacles may be present and they are not necessarily disjoint. One main issue in this class of billiards is the distinction between finite and infinite horizon. Finite Horizon means that there is a maximal time after which a collision must take place, infinite horizon means that there exists trajectories that never experience a collision.[8] The relevance of such a concept stems from the fact that orbit with no collision have zero Lyapunov exponents, hence the corresponding billiards cannot be uniformly hyperbolic.

[8] Note that the other possibility (all the trajectories experience a collision in finite time, but there does not exist an upper bound for such a time) cannot take place (see Problem 4.12).

Infinite Horizon

As already mentioned we have already seen the prototypical example in this class, yet it may be instructive to analize its properties a bit further. Consider the Sinai billiard described in section 4.1 and let r_1 the radius of the obstacle. Clearly it is necessary $r_1 < 1/2$ to have no self intersections of the obstacle. It is also obvious that if $r_1 < 1/2$ then there are trajecotirs that never collide. Let us study such trajectories a bit more in detail. First of all, since the sistem has a square simmetry, it is enough to consider trajectories with velocity in the first half of the first quadrant, i.e. velocities parallel to the directions $(1, \omega)$ with $\omega \in [0, 1]$.[9]

Let us consider the motion with no obstacle (a traslation on the torus) and see if there are trajectories that never enter in the region $\|(x, y)\| \leq r_1$. Clearly such trajetories are trajectories for the billiard systems as well and precisely the trajecotries that never experince a collision. For such trajectories it is particularly convenient to consider the poincarè section determined by the line $S : \{x = -1/2\}$. If we look at the motion only when the particle intersects such a line we have that the corresponding map is given by $Ty = y + \omega \mod 1$, that is a rotation by ω of the circle $(-1/2, 1/2]$. If $\omega \notin \mathcal{Q}$ then the map T is ergodic and this means that the trajectory will eventually collide. On the contrary, if $\omega = p/q$, $p, q \in \mathbb{N}$ and with no common divisors, then all the orbits will be periodic of period q and it may be possible that some of them never collide.

Notice that a point in S with velocity parallel to $(1, \omega)$ will experience before a collision before meeting S again only if $y \in [-\omega/2 - r_1\sqrt{1 + \omega^2}, -\omega/2 + r_1\sqrt{1 + \omega^2}]$. On the other hand, since the orbit of the point y has lenght q and because it is restricted to points of the type $y + n/q \mod 1$, which are exactly q, it follows that the orbit consists exactly of all such points. Accordingly, the orbit can avoid only intervals of size less than $1/q$. We can then conclude that there are orbits of the type p/q that never collide if and only if

$$2r_1\sqrt{1 + \frac{p^2}{q^2}} < \frac{1}{q}. \tag{4.3.7}$$

If $q = 1$ then for $p = 0$, we have the already know result $r_1 < 1/2$; for $p = 1$ there can be no collisions only if $r_1 < \frac{1}{2\sqrt{2}}$. For $q \leq 2$ there are always collisions if $r_1 > [2\sqrt{5}]^{-1}$.

[9] If this it is not clear, read again the discussion of polygonal billiards in section 4.0.1.

Finite Horizon

The simplest case of Sinai billiard with a finite horizon is obtained by employing two circular obstacles. We have thus a torus of size one together with a circular obstacle at the point $(0, 0)$ with radius r_1 and a circular obstacle at the point $(1/2, 1/2)$ with radius r_2.[10] Clearly

$$r_1 + r_2 < \frac{1}{\sqrt{2}}$$

in order for the obstacels not to intersect each other. By the discussion of the infinite horizon case it follows that we can choose $r_1 > 1/(2\sqrt{2})$ and $0 < r_2 < 1/\sqrt{2} - r_1$ to have a Sinai billiard with disjoint obstacles and finite horizon. For example one could choose $r_1 = 3/7$ and $r_2 = 1/4$.

In the following we will need a more in depth undertanding of the above model. Let us consider a regularized Poincarè section of the type introduced in section 4.1.4 and discuss the structure of the singularity lines for such a section.

The first step is to understand multiple consecutive tangencies. Let us start with a double tangency, the first of which is with the central obstacle. By simmetry, one can limit the analysis to the case in which the second takes place with the upper right copy of the obstacle. The position of the particle at time t is given by $r_1 n(\theta) + v(\theta)t$, where $n(\theta) = (\cos\theta, \sin\theta)$ and $v(\theta) = (\sin\theta, -\cos\theta)$, so we have the next two equations

$$\begin{aligned} &\|r_1 n(\theta) + v(\theta)t - p\| = r_2 \\ &\langle r_1 n(\theta) + v(\theta)t - p, v(\theta)\rangle = 0, \end{aligned}$$

where $p = (1/2, 1/2)$ are the coordiantes of the center of teh second obstacle (of course we are working in the universal covering). The first equation determine the value of t for which the second collision takes place while the second impose that the collision is tangent. Solving the above equations yields

$$\frac{1}{\sqrt{2}} \cos\left(\frac{\pi}{4} - \theta\right) = \langle n(\theta), p\rangle = r_1 \pm r_2.$$

Accordingly we have four solutions: $\frac{\pi}{4} - \theta = \pm(\cos^{-1}\sqrt{2}(r_1 \pm r_2))$. In fact, only two are really relevant since the other two are obtained by simmetry around the line joining the two centers. It remains to check that the above tajectories do not intersect any other obstacle between the two

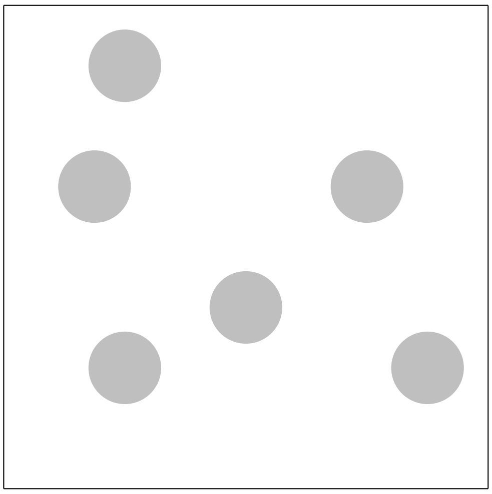
[10] Remember that the coordinates are in the universal covering of the torus and that the points $(1/2, -1/2)$, $(1/2, 1/2)$, $(-1/2, 1/2)$, $(-1/2, -1/2)$ are identified.

tangencies. In fact, it turns out that the trajectories of the type $\frac{\pi}{4} - \theta = \pm(\cos^{-1}\sqrt{2}(r_1 - r_2))$ have a tangent collision with the central scatterer before colliding with the corner one. It is then easy to see that there can be at most four consecutive tangencies after that the next collision will take place with an angle of more than 70 degree.

4.4 Hard spheres

We have already mentioned that the motions of several discs or balls that collide elastically among themselves are an example of billiards. The study of such models goes back at least to Boltzmann, who proposed studying the properties of a gas, imagining that it consists of balls colliding elastically.

We start by looking at the simplest possible case.



A two dimensional gas of particles in a box

4.5 A gas with two particles

The (seemingly ridiculous) simplest case is a gas of two particles in two dimensions. For simplicity, let us consider two particles of radius $r < \frac{1}{2}$ in a torus of size one. Let $x_1, x_2 \in \mathbb{T}^2$ be the coordinate of the center of the disks, the velocity changes at collision according to the law

$$\begin{cases} v_1^+ = v_1^- - \langle n, v_2^- - v_1^- \rangle n \\ v_2^+ = v_2^- + \langle n, v_2^- - v_1^- \rangle n \end{cases} \tag{4.5.8}$$

where n is a unit vector in the direction $x_2 - x_1$.[11]

Here, there are three integrals of motion: the energy $E = \frac{1}{2}(\|v_1\|^2 + \|v_2\|^2)$ and the total momentum $P = v_1 + v_2$. Thus, if we want to obtain an ergodic system, we have to reduce the system. We will then consider that phase spaces

$$X_{E,P} = \left\{ (x_1, x_2, v_1, v_2) \in \mathbb{T}^4 \times \mathbb{R}^4 \;\middle|\; \frac{1}{2}(\|v_1\|^2 + \|v_2\|^2) = E;\ v_1 + v_2 = P \right\}.$$

Since, in the velocity space, the previous conditions correspond to the intersection between the surface of a four-dimensional sphere (S^3) and a two-dimensional linear space, the velocity vectors $(v_1 + v_2)$ are contained in a one-dimensional circle. Thus, topologically, $X_{E,P} = \mathbb{T}^4 \times S^1$.[12] It is then natural to choose an angle θ as coordinate on S^1, moreover, since

$$2E = \|v_1\|^2 + \|v_2\|^2 = \frac{1}{2}\|v_1 - v_2\|^2 + \frac{1}{2}\|P\|^2,$$

it is hard to resist setting $v_2 - v_1 = v(\theta)$.[13] Hence,

$$\begin{cases} v_1 = \frac{1}{2}(P - v(\theta)) \\ v_2 = \frac{1}{2}(P + v(\theta)). \end{cases}$$

The free motion is then given by

$$\begin{cases} x_1(t) = x_1(0) + \frac{1}{2}(P - v(\theta))t \\ x_2(t) = x_2(0) + \frac{1}{2}(P + v(\theta))t. \end{cases}$$

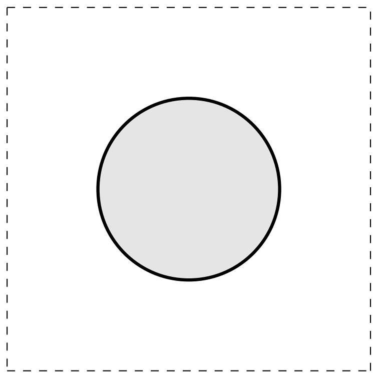
Accordingly,

$$\begin{cases} x_1(t) + x_2(t) = x_1(0) + x_2(0) + Pt \mod 1 \\ x_2(t) - x_1(t) = x_2(0) - x_1(0) + v(\theta)t \mod 1. \end{cases}$$

[11] To be precise $x_2 - x_1$ has no meaning since $\mathbb{T}^2$ it is not a linear space. Yet, at collision, the distance between the two disks is $2r$, so the global structure of $\mathbb{T}^2$ is irrelevant, and we can safely confuse it with a piece of $\mathbb{R}^2$.

[12] Of course, we are considering only the cases $E \neq 0$.

[13] As usual $v(\theta) = (\sin\theta, \cos\theta)$.



The Siani billiard $\mathcal{B}$

Figure 4.4: Sinai Billiard with infinite horizon

It is then clear the need to introduce the two new variables $Q = x_1 + x_2$ and $\xi = x_2 - x_1$. The variable Q performs a translation on the torus, such a motions are completely understood, and we can then disregard it. The only relevant motion is the one in the variables (ξ, θ). The reduced phase space is then $\mathcal{B} \times S^1$ where $\mathcal{B} = \mathbb{T}^2 \backslash \{\|\xi\| \le 2r\}$, that is, the torus minus a disk of radius $2r$. The domain $\mathcal{B}$ is represented in the next Figure and, apart from the different choices of the fundamental domain, it corresponds exactly to the simplest Sinai billiard. The free motion corresponds to the free motion of a point as well, while at collision, from (4.5.8), we have

$$v(\theta^+) = v(\theta^-) - 2\left\langle \frac{\xi}{2r}, v(\theta^-) \right\rangle v(\theta^-)$$

that is exactly the elastic reflection from the disk!

It is then natural to consider the general problem of a particle moving in a region with reflecting boundary conditions. Let $\mathcal{B} \subset \mathbb{R}^d$ (or $\mathcal{B} \subset \mathbb{T}^d$) be the region and suppose that the boundary $\partial\mathcal{B}$ is made of finitely many smooth manifolds. Calling $(x, v) \in \mathcal{B} \times \mathbb{R}^d$ the position and the velocity, respectively, the motion inside $\mathcal{B}$ is described by a free flow

$$\phi_t(x, v) = (x + vt, v), \tag{4.5.9}$$

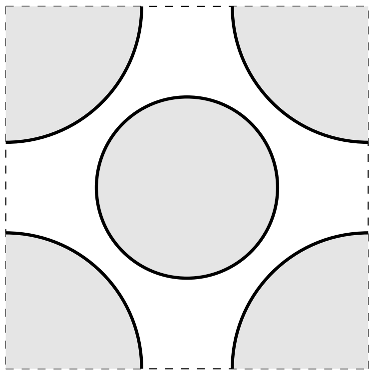
When $x \in \partial\mathcal{B}$, a collision takes place. If $n \in \mathbb{R}^d$, $\|n\| = 1$, is the normal to $\partial\mathcal{B}$ at x, then, calling v_- and v_+ the velocities before and after collision, respectively, the elastic collision is described by

$$v_+ = v_- - 2\langle v_-, n \rangle n.$$

Remark 4.5.1 *Here, I will provide a few ideas on billiards and hyperbolicity. This should allow the reader to be able to easily takle a more complete account of the theory (in particular [16]).*

4.6 Some Billiard tables

In the two-dimensional case, there are many possible billiards tables that have been studied. The two most famous are the Sinai Billiard and the Bunimovich Stadium.



Sinai Billiard with finite horizon

Bunimovich stadium

Further interesting billiard tables can be found in [64, 14, 16] and references therein.

Problems

4.1 Check that the maps ϕ^t generated by a billiard flow are symplectic. (Hint: It is obvious for the free flow, but it remains to be checked for the reflections. This can be done by using formulae like (5.1.3)).

4.2 Given a rectangular box $\mathbb{B}$, with its sides labeled by $\{1, 2, 3, 4\}$ and let $R_i(\mathbb{B})$ be the reflection with respect to the side i of the box $\mathbb{B}$.[14] Let R_0 be the identity. Consider $G = \cup_{n=0}^{\infty}\{0, 1, 2, 3, 4\}^n$, if $g \in G$ then we define $R_g(\mathbb{B}) = R_{g_1}(\cdots R_{g-n}(\mathbb{B})\cdots)$ and, for each $g^i \in \{0, 1, 2, 3, 4\}$, $i \in \{1, 2\}$, $g = g^2 \circ g^1 \in \{0, 1, 2, 3, 4\}^{n_1+n_2}$, is defined by $g_k = g_k^1$ for $k \leq n_1$ and $g_k = g_{n-k}^2$, for $k > n_1$. Verify that $R_{g^2}(R_{g^1}(\mathbb{B})) = R_g(\mathbb{B})$. Introduce the equivalence relation $g_1 \sim g_2$ iff $R_{g^1}(\mathbb{B}) = R_{g^2}(\mathbb{B})$. Let $\tilde{G}$ be the collection of the equivalence classes. Verify the $\tilde{G}$ is a commutative group with respect to the operation $\circ$. (hint: Note that the geometrical meaning is simply that the final position of the box after a certain number of reflections does not depend on the order of the reflections. Clearly, it suffices to check such a property for two reflections.)

4.3 Study the motion in a triangular billiard when the angles defining the triangle are all rational multiples of π. (Hint: use reflections again)

4.4 Study the motion in an elliptical billiard. (Hint: Verify that there exists an integral of motion.)

4.5 Verify that the caustics correspond to a two dimensional torus.

4.6 Check that the maps ϕ^t generated by a billiard flow are symplectic. (Hint: It is obvious for the free flow, but it remains to be checked for the reflections. This can be done by using formulae (5.1.3))

4.7 Find a change of variable that transforms the symplectic form in a regularized boundary section in the standard symplectic form.

4.8 Verify that, in a regularized boundary section, the map is continuous across a singularity line corresponding to a tangency.

---

[14] The labels attached to the sides of the reflected boxes are the ones obtained naturally from the old ones.

4.9 Prove that, given an $n \times n$ matrix A the norm $\|A\| := \sup_{v\in\mathbb{R}^n} \frac{\|Av\|}{\|v\|}$ where $\|v\| := \sqrt{\sum_n v_n^2}$ satisfies

$$\|A\| \leq \text{constant} \max_{ij} |A_i j|$$

and compute explicitely the optimal constant.

4.10 Determine the rate at which the second derivative explode as one gets near a tangency singularity in the Sinai billiard with one circular obstacle in the torus.

4.11 Compute the number of collisions in a convex angle.

4.12 Show that in a Sinai billiard on $\mathbb{T}^2$ for which there exist trajectories that spend an arbitrarily long time without colliding, there must exist trajectories that never collide. (Hint: some continuity...)

4.13 Study two disks with different masses.

4.14 Prove that the Poincarè map for the Sinai billiard is piecewise Hölder of Hölder exponent $\frac{1}{2}$.