7. Phase space

This will be a concept-heavy lecture. It involves changing how we picture billiards, in a way that's not immediately intuitive.

- We introduce phase space.
- We measure areas in phase space, and how the billiards dynamics affects those areas.
- As a consequence, we get an abstract but very general existence result about "almost periodic" trajectories.

(7a) Recurrence. Let's begin with the payoff, the Poincaré recurrence theorem:

THEOREM 7.1. *Inside any polygon, choose a point and a direction, in any way you want. Then, there is a billiards trajectory whose starting position and direction are arbitrarily close to the ones we picked, and which after some amount of bounces, returns to a position and direction arbitrarily close to the ones we picked.*

"Arbitrarily close" means: you get to specify a desired precision, and the theorem guarantees the existence of a trajectory that fits those specifications. Suppose that you declare the specification to be "positions at distance less than 0.001 from the one we fixed, and directions whose angle differs by less than 0.001° from the one we fixed". By the recurrence theorem, there is a trajectory that starts off like that, and after some unknown time, again satisfies that. If you now get more picky and refine your bounds to 0.000001, there is also a trajectory with those properties, but it will probably be much longer and complicated than the previous one. These trajectories won't usually be periodic: they are only "almost periodic", in the sense that they return to a state very close to their starting state.

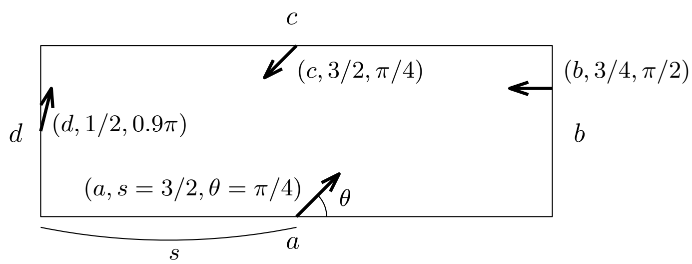
(7b) Phase space. Think about coding billiards on a computer. Clearly, it would be a waste of time to simulate the straight-line motion. We should just start at one bounce point, see what direction we take from there, and directly compute the next bounce point and direction. Mathematically, the idea is encoded into phase space. The formal definition of phase space (for billiards in a polygon P) is

$$\Omega = \bigsqcup_e (0, \mathrm{length}(e)) \times (0, \pi). \tag{7.1}$$

This is a disjoint (meaning non-overlapping) union of rectangles, one for each edge e of the polygon. A point in phase space is written as (e, s, θ), with the following meaning: e is the edge, let's say $e = \overline{v_{k-1}v_k}$ for some k, where the vertices have been ordered anticlockwise. Then, s specifies a point on the edge, through its distance from v_{k-1}; in other words, how far do we have to walk anticlockwise along the edge before reaching our desired point. Finally, θ specifies an inwards pointing direction at our point, obtained by rotating the vector $w = v_k - v_{k-1}$ anticlockwise by the angle θ.

EXAMPLE 7.2. *Let's take our polygon to be a* 3×1 *rectangle, with sides labeled abcd. Here are some examples of phase space coordinates:*

(7.2)



(7c) The billiards map. Take a point p in phase space, meaning a boundary point and inwards direction. Place the ball at the boundary point, and move it in the prescribed direction until we again hit the boundary of the polygon. Then, record the new position and the reflected direction (the direction in which the ball will continue after bouncing off). This defines another point in phase space, $T(p)$. We can think of this as a map from phase space to phase space, the billiards map

$$T : \Omega \dashrightarrow \Omega. \tag{7.3}$$

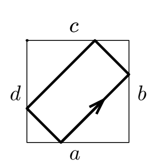
If we are interested in what happens to our trajectory as it keeps bouncing, we form $T(T(p))$, $T(T(T(p)))$, ... in that way, we have transformed continuous billiards motion into a problem of repeated application of the billiards map. There is also an inverse billiards map $T^{-1} : \Omega \dashrightarrow \Omega$, which is obtained by running the billiards motion in reverse. It satisfies

$$T(T^{-1}(p)) = T^{-1}(T(p)). \tag{7.4}$$

Why is the arrow in (7.3) dashed? Because we could run into a corner, in which case $T(p)$ is not defined. The same holds for T^{-1}, so the equalities in (7.4) hold only if no such catastrophe happens.

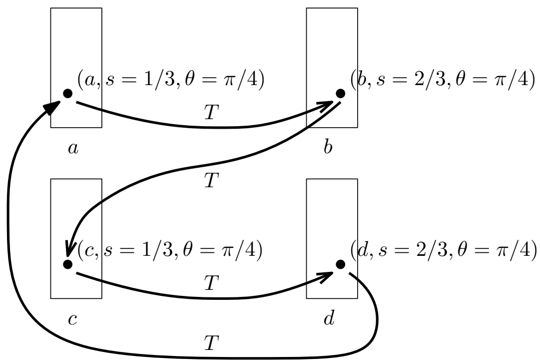
EXAMPLE 7.3. *We take a simple* 4*-bounce periodic billiards trajectory in a* 1×1 *square, where each bounce happens a third of the way off from a corner:*

(7.5)



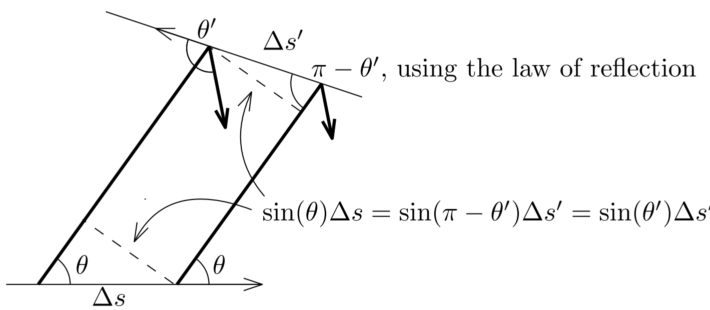
Each segment of the trajectory (or rather, the starting point and direction of that segment) corresponds to a point in phase space. The map T *takes the point corresponding to a segment to that* *for the subsequent segment. The outcome is that we get four phase space points permuted by T:*

(7.6)



What can one say about the billiards map in general? Let's write $T(e, s, \theta) = (e', s', \theta')$. If we vary s a little, and keep θ fixed, the trajectory gets displaced to a parallel one:

(7.7)



As the diagram shows, we have

$$T(e, s + \Delta s, \theta) = (e', s' + \Delta s', \theta') = \big(e', s' - \tfrac{\sin(\theta)}{\sin(\theta')}\Delta s, \theta'\big). \tag{7.8}$$

If we keep s fixed and vary θ, then θ' decreases by the same amount; and s' also changes, but in a more complicated way, which we won't try to write down.

$$T(e, s, \theta + \Delta\theta) = (e', s' + \textit{something}, \theta' - \Delta\theta). \tag{7.9}$$

In calculus language,

$$\begin{aligned} &\frac{\partial s'}{\partial s} = -\frac{\sin(\theta)}{\sin(\theta')}, && \frac{\partial s'}{\partial \theta} = (\textit{something}), \\ &\frac{\partial \theta'}{\partial s} = 0, && \frac{\partial \theta'}{\partial \theta} = -1. \end{aligned} \tag{7.10}$$

(7d) Conservation of areas and the recurrence theorem. At this point, we change coordinates on the phase space a little, keeping s, but replacing $\theta \in (0, \pi)$ by $t = -\cos(\theta) \in (-1, 1)$. In these coordinates, phase space looks like this:

$$(e, s, t) \in \Omega = \bigsqcup_e (0, \mathrm{length}(e)) \times (-1, 1). \tag{7.11}$$

While apparently arbitrary, this change leads to a crucial insight into the nature of the billiards map. Writing $T(s,t) = (s',t')$ in our new coordinates, the chain rule says that

$$\frac{\partial t'}{\partial t} = \frac{\partial t'}{\partial \theta'}\frac{\partial \theta'}{\partial \theta}\Big(\frac{\partial t}{\partial \theta}\Big)^{-1} = \sin(\theta')(-1)\sin(\theta)^{-1} = -\frac{\sin(\theta')}{\sin(\theta)}. \tag{7.12}$$

Let's summarize the situation in a matrix of derivatives

$$\begin{pmatrix} \frac{\partial s'}{\partial s} & \frac{\partial s'}{\partial t} \\ \frac{\partial t'}{\partial s} & \frac{\partial t'}{\partial t} \end{pmatrix} = \begin{pmatrix} -\frac{\sin(\theta)}{\sin(\theta')} & (\mathit{something}) \\ 0 & -\frac{\sin(\theta')}{\sin(\theta)} \end{pmatrix}. \tag{7.13}$$

Even though we haven't computed one of the entries, we can see that the matrix has determinant 1. By general change-of-coordinate formulas, this implies what's known as Liouville's theorem:

THEOREM 7.4. *In (s,t) coordinates on the phase space, the billiards map is area-preserving.*

We now return to the Poincaré recurrence theorem, in a formulation that uses phase space, and which is more mathematically precise than before. This is the same statement, even though it may take some time to recognize that!

THEOREM 7.5. *Let $U \subset \Omega$ be any subset with positive area. Then, there is some $p \in U$ and an $M > 0$ such that*

$$T^M(p) = \overbrace{T(T(\cdots}^{M\ \mathit{times}}(p))\cdots) \in U. \tag{7.14}$$

The proof is easy: take a natural number N and look at the sets

$$U,\, T(U),\, T^2(U), \dots, T^{N-1}(U). \tag{7.15}$$

Working in (s,t) coordinates, all those have the same area by Liouville's theorem. If they didn't overlap, the area of their union would be N times the area of U. On the other hand, the area of the entire phase space is finite: it's twice the perimeter of the polygon. If N is really large, that's a contradiction, so our sets must overlap after all. To say it more precisely, there must be $1 \leq N_1 < N_2 \leq N$ such that $T^{N_1}(U) \cap T^{N_2}(U)$ has positive area. Now we apply the inverse billiards map, which also preserves areas:

$$\mathrm{area}(T^{N_1}(U) \cap T^{N_2}(U)) = \mathrm{area}(T^{N_1-1}(U) \cap T^{N_2-1}(U) = \cdots = \mathrm{area}(U \cap T^{N_2-N_1}(U)). \tag{7.16}$$

Since the last intersection has positive area, there must be a point in it. This was our desired goal, and we got $M = N_2 - N_1$. There's a possible objection: we've argued as if T was a one-to-one (invertible) map from Ω to itself, whereas really it's not everywhere defined, and the same for T^{-1}. However, the badly-behaved points form one-dimensional subsets of phase space, which have zero area, so they don't affect our argument after all.

18.900 Geometry and Topology in the Plane
Spring 2023