Università degli studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea Magistrale in Matematica

Birkhoff and symplectic billiards; an overview

Relatore:
Prof.ssa Olga Bernardi
Correlatore:
Prof. Alfonso Sorrentino

Laureanda:
Giorgia Testolina
Matricola 1207242

26 Febbraio 2021

Anno Accademico 2020-2021

Ringraziamenti

Vorrei ringraziare innanzitutto la mia relatrice, Professoressa Olga Bernardi, per avermi accompagnato con disponibilità e attenzione in questo lavoro di tesi. La sua passione per la materia e i preziosi consigli che mi ha fornito in questi mesi mi hanno motivato e guidato nella stesura di questo elaborato. Ringrazio, inoltre, il Professor Alfonso Sorrentino per la professionalità e la gentilezza con cui mi ha fornito il suo aiuto in molte occasioni.

Ringrazio la mia famiglia, fonte costante di coraggio e supporto. Mamma e papà perché, oltre ad essere il mio punto di riferimento, non mi hanno mai fatto mancare amore e sostegno, trasmettendomi gli insegnamenti fondamentali che mi hanno reso la persona che sono ora; i miei fratelli Giulia e Riccardo per esserci sempre nei momenti di bisogno e per rendere la mia vita molto più vivace. Ringrazio anche i miei nonni, per essere sempre presenti in ogni traguardo della mia vita.

Un grazie alle mie amiche del liceo, Elena, Elisa e Laura, per essere state al mio fianco in tutti questi anni, condividendo con me esperienze uniche e i momenti più importanti della mia vita. Sono orgogliosa di voi e delle splendide donne che siete diventate.

Ringrazio, anche, tutti gli amici che ho conosciuto in questo periodo universitario; avete reso questi anni indimenticabili e pieni di ricordi meravigliosi. In particolare, ringrazio Antonio, Chiara, Elena e Isabella per la loro impagabile amicizia e per essere stati fonte di ispirazione nei miei momenti di incertezza.

Infine, un ringraziamento speciale a Francesco, per ascoltarmi e comprendermi come nessun altro riesce a fare. Nei momenti più difficili, la tua forza e la tua bontà sono sempre riuscite a rimettermi in piedi. Grazie per essermi accanto e per condividere con me ogni momento della mia vita.

Contents

Introduction

A mathematical billiard is a dynamical system describing the motion of a mass point (the billiard ball) inside a planar region (the billiard table) with –in general– piecewise smooth boundary. The ball moves with costant –say, unit– speed and without friction, following a rectilinear path. Upon the impact with the boundary of the billiard table, two different situations can occur, according to the nature of the impact point. If the ball hits the boundary at a point of non-differentiability, then the motion stops; otherwise, the ball reflects elastically according to the standard reflection law angle of incidence is equal to the angle of reflection.
More generally, one could consider a Riemannian manifold with piecewise boundary as a billiard table. In that case, the ball moves along a geodesic of the manifold with constant speed and, at the regular points of the boundary, it reflects elastically in such a way that the tangential component of its velocity remains the same before and after the impact, while the normal one instantaneously changes its sign. In this way, mathematical billiards can be defined in higher dimensions and in geometries other than the Euclidean one. We refer to [31] and [33] for a more comprehensive introduction to the study of billiards.
The straightforwardness and versatility of this model has made mathematical billiards an object of interest in many different context. Indeed, they show a wide range of dynamical behaviour, such as integrability, regularity, caoticity, etc.; therefore they appear as a good mathematical "playground" to test techniques and solution methods, which can lately be adapted to other dynamical systems. Moreover, intriguing questions related to mathematical billiards are handled with several different techniques, from KAM theory to the theory of monotone twist maps.
We recall in the literature some classes of mathematical billiards, such as Birkhoff billiards, dual billiards, hyperbolic billiards and chaotic billiards.

In this thesis we focus on two classes of mathematical billiards: Birkhoff billiards and symplectic billiards.

Birkhoff billiards are a special class of mathematical billiards in which the billiard table is a stictly convex planar domain with smooth boundary. They owe

their name to the american mathematician G. D. Birkhoff, who, in his work *"On the periodic motions of dynamical systems"* of 1927 (see [6]), was the first one to consider these billiards as a paradigmatic example of Hamiltonian systems.
In Figure 1, it is represented the reflection at the boundary, according to the reflection law angle of incidence equals angle of reflection.

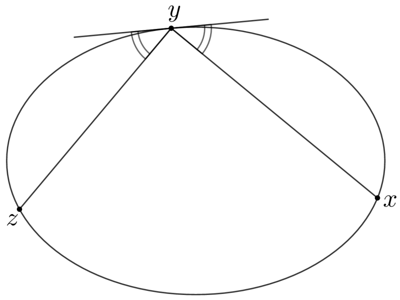

Figure 1: The Birkhoff billiard map

Birkhoff billiards have many interesting properties. In particular, the law describing the dynamics (the so-called billiard map) is an exact symplectic map with minus the Euclidean length as a generating function and, moreover, it is a monotone twist map.
One of the most discussed topics in the class of Birkhoff billiards is the one concerning the concept of caustic, i.e., a curve in the billiard table with the property that if a trajectory of the billiard map is tangent to it, then it remains tangent after every each reflection.
Closely related to the topic of caustics is the concept of integrability. Indeed, a Birkhoff billiard is said to be integrable if its billiard table is foliate by smooth convex caustics. Examples of integrable billiards are the billiards in the circle and in the ellipse.
The study of Birkhoff billiards integrability generated one of the most famous open problem in the field of mathematical billiards, the so-called Birkhoff conjecture, that states that the only integrable billiards are the circular and the elliptic ones.

Symplectic billiards are much more recent than Birkhoff billiards; they were introduced for the first time by P. Albers and S. Tabachnikov in their article *"Introducing symplectic billiards"* (see [1]) of 2017.
As in the Birkhoff case, the billiard table is a strictly convex planar domain with smooth boundary, but the dynamics does not follow the standard reflection law, but it is described as follows. Let x, y, z be three consecutive impact points on the boundary; the segment xy reflects to the segment yz if xz is parallel to the tangent line to the boundary at y, see Figure 2.

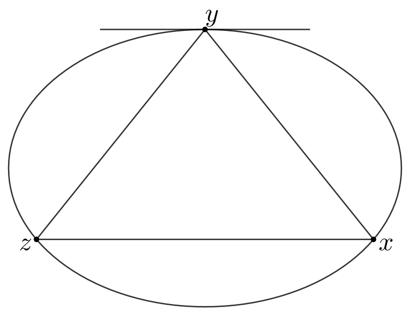

Figure 2: The symplectic billiard map

This new class of billiards has many properties in common with Birkhoff billiards. In particular, also the symplectic billiard map admits a generating function, which is essentialy the standard area form. The name "symplectic billiard" is actually due to this fact, since the standard area form is the symplectic 2-form on the plane. Moreover, the symplectic billiard map is a monotone twist map.
As in the Birkhoff case, the concepts of caustics and integrability can be introduced. In particular, both circular and elliptic symplectic billiards are integrable, although the dynamics of the elliptic billiard is different from the one of the corresponding Birkhoff billiard.
Finally, a remarkable property of symplectic billiards is that they can be easily generalized to linear symplectic spaces.

This thesis is organized as follows. Chapter 1 is devoted to the study of Birkhoff billiards. We define the billiard map, proving that it is an exact symplectic map with minus the Euclidean distance as a genarating function and that it is a monotone twist map. We introduce the concept of caustic for a Birkhoff billiard, discussing some results about their existence. We then describe Birkhoff billiards in a circle and an ellipse, introducing the topic of integrable billiards and Birkhoff conjecture.
In Chapter 2, we introduce symplectic billiards in the plane. We prove that the symplectic billiard map admits as a generating function the standard area form and that it is a monotone twist map. Finally, we focus on circular and elliptic billiards; in both cases, we prove that –as for Birkhoff billiards– the corresponding discrete dynamical systems are integrable.
Chapter 3 is entirely devoted to two results for caustics of symplectic billiards. The first one is a non-existence result, which applies when the boundary of the billiard table has points of zero curvature and whose proof is inspired by a result from Mather concerning Birkhoff billiards (see [22]). The second one is an existence result, applying when the boundary of the billiard table has everywhere positive curvature, and the proof follows the steps –inspired by KAM theory– carried out

by Lazutkin in [18].
In Chapter 4 we generalize the definition of symplectic billiards to symplectic vector spaces.
Finally, in Chapter 5 we study Birkhoff and symplectic billiards in the framework of Aubry-Mather theory. This theory was developed independently by S. Aubry and J. Mather in the Eighties, and it is concerned with the study of orbits of a monotone twist maps minimizing the action functional. In particular, we focus on the study of the average minimal action (the so-called Mather's β-function) and its properties for Birkhoff and symplectic billiards. This chapter constitutes the original part of the thesis, and it represents a good starting point for future development.

Introduzione

Un biliardo matematico è un sistema dinamico che descrive il moto di un punto materiale (la pallina del biliardo) all'interno di una regione del piano (il tavolo da biliardo) con bordo –in generale– regolare a tratti. La pallina si muove con velocità costante –che possiamo quindi considerare unitaria– e senza attrito, seguendo una linea retta. Al momento dell'urto con il bordo del tavolo si verificano due differenti situazioni, a seconda del punto di contatto. Se l'urto avviene in un punto di non differenziabilità del bordo, il moto si interrompe; altrimenti, la pallina rimbalza elasticamente in accordo con la legge di rifrazione, ovvero l'angolo d'incidenza è uguale all'angolo di riflessione.
Più genericamente, si può considerare come tavolo da biliardo una varietà Riemanniana con bordo regolare a tratti. In questo caso, la pallina si muove lungo una geodetica della varietà con velocità costante e, nei punti regolari del bordo, rimbalza elasticamente in maniera tale che la componente tangenziale della sua velocità rimanga la stessa prima e dopo l'urto, mentre la componente normale cambia istantaneamente verso. Questa generalizzazione permette di definire i biliardi matematici anche in più dimensioni e in geometrie diverse da quella euclidea. Per una estensiva trattazione sui biliardi matematici ci riferiamo a [31] e [33].
La semplicità e versatilità di questo modello ha reso i biliardi matematici un oggetto di interesse in diversi contesti. Infatti, mostrano un'ampia varietà di comportamenti dinamici, quali integrabilità, regolarità, caoticità, ecc.; si presentano quindi come un ottimo "terreno di gioco" per testare tecniche e metodi di soluzione, che vengono poi adattati ad altri sistemi dinamici. Inoltre, intriganti questioni relative ai biliardi matematici vengono affrontate usando diverse tecniche, dalla teoria KAM alla teoria delle mappe monotone twist.
Ricordiamo in letteratura alcune classi di biliardi matematici, come i biliardi di Birkhoff, i biliardi esterni, i biliardi iperbolici e quelli caotici.

In questa tesi ci concentreremo sullo studio di due classi di biliardi matematici: i biliardi di Birkhoff e i biliardi simplettici.

I biliardi di Birkhoff sono una speciale classe di biliardi matematici in cui il tavolo da biliardo è un dominio strettamente convesso del piano con bordo regolare. Devono il loro nome al matematico statunitense G. D. Birkhoff, che nel suo lavoro *"On the periodic motions of dynamical systems"* del 1927 ([6]), fu il primo a considerare questi biliardi come esempio paradigmatico di sistemi Hamiltoniani. In Figura 3, è rappresentata la riflessione nel bordo, in accordo con la legge di rifrazione angolo di incidenza uguale angolo di riflessione.

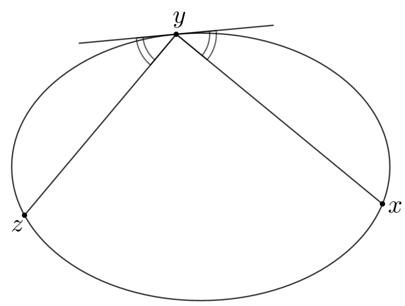

Figura 3: La mappa del biliardo di Birkhoff

I biliardi di Birkhoff presentano numerose interessanti proprietà. In particolare, la legge che descrive la dinamica (la cosiddetta mappa del biliardo) è una mappa esattamente simplettica con meno la distanza euclidea come funzione generatrice e, inoltre, è una mappa monotona twist.
Uno degli argomenti più discussi all'interno della classe dei biliardi di Birkhoff è quello riguardate il concetto di caustica, ovvero una curva del tavolo da biliardo con la proprietà che se una traiettoria della mappa del biliardo è tangente ad essa, allora vi rimane tangente dopo ogni riflessione.
Strettamente legato all'argomento delle caustiche è il concetto di integrabilità. Infatti, un biliardo di Birkhoff è detto integrabile se il suo tavolo da biliardo è foliato da caustiche regolari e convesse. Esempi di biliardi integrabili sono i biliardi nel cerchio e nell'ellisse.
Lo studio dell'integrabilità nei biliardi di Birkhoff ha generato uno dei più famosi problemi aperti nel campo dei biliardi, la cosiddetta congettura di Birkhoff, che afferma che gli unici esempi di biliardi integrabili siano quelli circolari ed ellittici.

Molto più recenti rispetto ai biliardi di Birkhoff sono i biliardi simplettici, introdotti per la prima volta nel 2017 da P. Albers e S. Tabachnikov nel loro articolo *"Introducing symplectic billiards"* ([1]).
Come nel caso Birkhoff, il tavolo da biliardo è un dominio del piano strettamente convesso, ma la dinamica non segue la regola di riflessione standard, ma è descritta nel seguente modo. Dati x, y, z tre punti sul bordo del dominio, xy riflette su yz se il segmento xz è parallelo alla retta tangente al bordo nel punto y, vedi Figura 4. Questa nuova classe di biliardi presenta numerose caratteristiche comuni a quelle dei biliardi di Birkhoff. In particolare, anche la mappa del biliardo simplettico

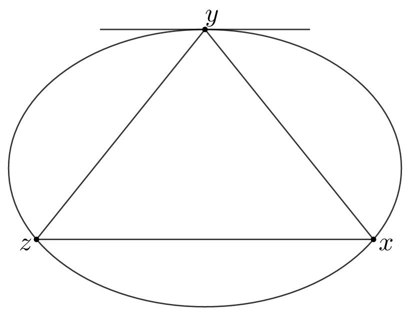

Figura 4: La mappa del biliardo simplettico

ammette una funzione generatrice, che è essenzialmente la forma d'area standard. Il nome di "biliardo simplettico" è proprio dovuto a questo fatto, dato che la forma d'area standard è la 2-forma simplettica nel piano. Inoltre, la mappa del biliardo simplettico è una mappa monotona twist.
Come per i biliardi di Birkhoff, possono essere introdotti i concetti di caustica e integrabilità. In particolare, sia il biliardo simplettico circolare sia quello ellittico sono integrabili, anche se nel caso dell'ellise la dinamica è diversa da quella del corrispettivo biliardo di Birkhoff.
Infine, una proprietà notevole dei biliardi simplettici è quella di poter essere facilmente generalizzati a spazi lineari simplettici.

Questa tesi è organizzata come segue. Il Capitolo 1 è dedicato allo studio dei biliardi di Birkhoff. Descriviamo la mappa del biliardo, dimostrando che è una mappa esattamente simplettica con meno la distanza euclidea come funzione generatrice e che è una mappa monotona twist. Introduciamo il concetto di caustica per i biliardi di Birkhoff, discutendo alcuni risultati riguardo la loro esistenza. Descriviamo poi i biliardi di Birkhoff nel cerchio e nell'ellisse, introducendo l'argomento dei biliardi integrabili e la congettura di Birkhoff.
Nel Capitolo 2 introduciamo i biliardi simplettici nel piano. Dimostriamo che la mappa del biliardo simplettico ammette come funzione generatrice la forma d'area standard e che è una mappa monotona twist. Infine, ci concentriamo sui biliardi circolari ed ellittici; in entrambi i casi, dimostriamo che –come per i biliardi Birkhoff– i corrispondenti sistemi dinamici discreti sono integrabili.
Il Capitolo 3 è interamente dedicato a due risultati per le caustiche dei biliardi simplettici. Il primo è un risultato di non esistenza, che si applica quando il bordo del tavolo da biliardo ha punti di curvatura nulla e la cui dimostrazione si ispira ad un risultato di Mather per i biliardi di Birkhoff (vedi [22]). Il secondo è un risultato di esistenza, che si applica quando il bordo del tavolo da biliardo ha ovunque curvatura positiva; la dimostrazione segue i passaggi –ispirati dalla teoria KAM–

svolti da Lazutkin in [18].
Nel Capitolo 4 generalizziamo la definizione di biliardo simplettico a spazi vettoriali simplettici.
Infine, nel Capitolo 5 studiamo i biliardi di Birkhoff e quelli simplettici all'interno della teoria Aubry-Mather. Questa teoria è stata sviluppata indipendentemente da S. Aubry e J. Mather negli anni Ottanta, e si occupa dello studio delle orbite di mappe monotone twist che minimizzano l'azione del funzionale. In particolare, ci concentriamo sullo studio dell'azione minima media (la cosiddetta Mather's β-function) e delle sue proprietà per i biliardi di Birkhoff e simplettici. Questo capitolo costituisce la parte originale della tesi, e rappresenta un buon punto di partenza per sviluppi futuri.

Chapter 1

Birkhoff billiards

Abstract. This chapter is devoted to the study of Birkhoff billiards on the plane. We give the definition and prove that the billiard map is a monotone twist map with the negative Euclidean distance as the generating function. We introduce the concept of caustics for Birkhoff billiards, discussing some results about their existence. We describe circular and elliptic billiards, introducing the topic of integrable billiards and one of the most famous open problem in the field of mathematical billiards, the so-called Birkhoff conjecture, which states that billiards in the circle and in the ellipse are the only integrable billiards.

1.1 Birkhoff billiard: definition

Let D be a strictly convex domain in $\mathbb{R}^2$ with $\mathcal{C}^r$ boundary ∂D, with $r \geqslant 3$. The phase space M of the Birkhoff billiard inside D is the space of all unit vectors (x, v) whose foot points are on ∂D and which have inward directions. The billiard ball map is

$$T : M \to M \qquad (x, v) \mapsto (x', v'),$$

where x' is the point in which the trajectory starting at x with velocity v hits the boundary ∂D again, and v' is the reflected velocity, according to the standard reflection law angle of incidence equals angle of reflection, see Figure 1.1.

Remark 1.1.1. More generally, one could consider, as a billiard table, a Riemannian manifold with smooth boundary $(M, \partial M, g)$. The billiard ball moves along a geodesic line in M with unit velocity until it hits the boundary and reflects in such a way that the tangential component of its velocity remains the same while the normal component instantaneously changes sign. Observe that in the Euclidean planar case, this gives exactly the standard reflection law that we have described above.

Remark 1.1.2. Observe that the boundary of the billiard table is required to be at least $\mathcal{C}^3$. Indeed, Halpern in [15] provides an example in which, if this condition (actually, it is enough $\mathcal{C}^2$ plus bounded third derivative) is not satisfied, the billiard ball trajectory will hit the boundary curve an infinite number of times in a finite period of time.

Assume that the length $l(\partial D)$ of the boundary curve is normalized to 1 and let $\gamma : \mathbb{S}^1 \to \mathbb{R}^2$ be the arc-length parametrization of the boundary ∂D, denoting by t the arc-length parameter. Without any loss of generalization, we orientate γ counterclockwise. Let φ be the angle between v and the tangent line to ∂D at x. Thus, the phase space M is identified with the annulus $\mathbb{S}^1 \times (0, \pi)$ and the billiard map becomes

$$T : \mathbb{S}^1 \times (0, \pi) \to \mathbb{S}^1 \times (0, \pi) \qquad (t, \varphi) \mapsto (t', \varphi').$$

We observe that the billiard map T can be continuously extended to the closure $\mathbb{S}^1 \times [0, \pi]$ by fixing $T(t, 0) = (t, 0)$ and $T(t, \pi) = (t, \pi)$ for all $t \in \mathbb{S}^1$.

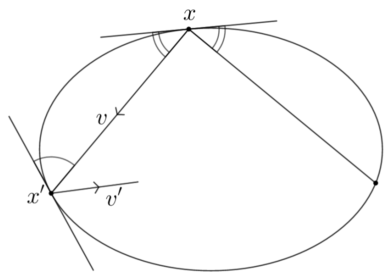

Figure 1.1: The billiard map

Let us denote by

$$h(t, t') = -\|\gamma(t) - \gamma(t')\| \tag{1.1}$$

the negative Euclidean distance between two points on ∂D.

Proposition 1.1.3. *Let φ, φ' be the angles that the vector from $\gamma(t)$ to $\gamma(t')$ makes with the tangent lines to γ respectively at $\gamma(t)$ and $\gamma(t')$. The next equalities hold*

$$\begin{cases} \frac{\partial h}{\partial t}(t, t') = \cos\varphi \\ \frac{\partial h}{\partial t'}(t, t') = -\cos\varphi'. \end{cases} \tag{1.2}$$

Proof. The proof easily follows by the computation of the partial derivative of h with respect to t, and the following definition of scalar product between two vectors v, w

$$v \cdot w = \|v\|\|w\| \cos\theta, \tag{1.3}$$

where θ is the angle between them.
Let $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ and $\gamma(t') = (\gamma_1(t'), \gamma_2(t'))$. Then,

$$h(t,t') = -\|\gamma(t) - \gamma(t')\| = -\sqrt{(\gamma_1(t) - \gamma_1(t'))^2 + (\gamma_2(t) - \gamma_2(t'))^2}$$

and

$$\begin{aligned}\frac{\partial h}{\partial t}(t,t') &= \frac{\partial}{\partial t}(-\|\gamma(t) - \gamma(t')\|) \\ &= -\frac{2(\gamma_1(t) - \gamma_1(t'))\frac{d\gamma_1}{dt} + 2(\gamma_2(t) - \gamma_2(t))\frac{d\gamma_2}{dt}}{2\sqrt{(\gamma_1(t) - \gamma_1(t'))^2 + (\gamma_2(t) - \gamma_2(t'))^2}} \\ &= -\frac{\gamma(t) - \gamma(t')}{\|\gamma(t) - \gamma(t')\|} \cdot \frac{d\gamma}{dt}\end{aligned} \tag{1.4}$$

Observe now that $u := \frac{\gamma(t)-\gamma(t'))}{\|\gamma(t)-\gamma(t')\|}$ is the unit vector from $\gamma(t)$ to $\gamma(t')$, $\frac{d\gamma}{dt}$ is the unit tangent vector at γ in $\gamma(t)$, and the angle between them is clearly the angle φ.
Therefore,

$$\begin{aligned}\frac{\partial h}{\partial t}(t,t') &= -u \cdot \frac{d\gamma}{dt} \\ &\overset{1.3}{=} \|u\| \left\| \frac{d\gamma}{dt} \right\| \cos\varphi \\ &= \cos\varphi.\end{aligned} \tag{1.5}$$

The other equality is given by a similar computation, differentiating with respect to t' instead. $\square$

A remarkable property of the billiard map is that it admits an invariant area form Ω.

Lemma 1.1.4. *The billiard map T preserves the area form*

$$\Omega := \sin\varphi \, d\varphi \wedge dt.$$

Proof. First we observe that $\sin\varphi > 0$ in M and therefore $\Omega = \sin\varphi \, d\varphi \wedge dt$ is an area form.
Let $T(t,\varphi) = (t', \varphi')$. To prove the invariance of Ω, we have to show that

$$\sin\varphi \, d\varphi \wedge dt = \sin\varphi' \, d\varphi' \wedge dt'.$$

From (1.2) it follows that

$$dh = \frac{\partial h}{\partial t}(t,t')\, dt + \frac{\partial h}{\partial t'}(t,t')\, dt' = \cos\varphi \, dt - \cos\varphi' \, dt'. \tag{1.6}$$

Taking differentials, we obtain

$$0 = d^2h = -\sin\varphi\, d\varphi \wedge dt + \sin\varphi'\, d\varphi' \wedge dt'$$

that gives us the desired result. $\square$

Observe that the area form Ω is an exact symplectic form on M, that is

$$\Omega = \sin\varphi\, d\varphi \wedge dt = d(-\cos\varphi dt) = d\alpha,$$

where $\alpha := -\cos\varphi dt$ is a 1-form on M.
Moreover, if T^* denotes the pull-back of T, it follows from formula (1.6) that

$$T^*\alpha - \alpha = -\cos\varphi'\, dt' + \cos\varphi\, dt = dh,$$

that is, $T^*\alpha - \alpha$ is an exact 1-form.
A map $T : M \to M$ with these properties (preserving an exact symplectic form $\Omega = d\alpha$ and such that $T^*\alpha - \alpha = dh$ is an exact 1-form) is said to be exact symplectic, and the function h is called a generating function of T. Therefore, the billiard map T is an exact symplectic map with the negative Euclidean distance h as a generating function.

Remark 1.1.5. Consider three consecutive collision points $(t_0, \varphi_0), (t_1, \varphi_1) = T(t_0, \varphi_0), (t_2, \varphi_2) = T(t_1, \varphi_1)$. It follows from (1.2) that

$$\frac{\partial h}{\partial t_1}(t_0, t_1) = -\cos\varphi_1 \qquad \frac{\partial h}{\partial t_1}(t_1, t_2) = \cos\varphi_1$$

and consequently

$$\frac{\partial h}{\partial t_1}(t_0, t_1) + \frac{\partial h}{\partial t_1}(t_1, t_2) = -\cos\varphi_1 + \cos\varphi_1 = 0.$$

This formula has the following variational interpretation. The points (t_0, φ_0), (t_1, φ_1) and (t_2, φ_2) are consecutive if and only if (t_1, φ_1) is a critical point of the functional

$$t_1 \mapsto h(t_0, t_1) + h(t_1, t_2).$$

Consider now the lift of the billiard map T to the universal cover $\mathbb{R} \times (0, \pi)$ of $\mathbb{S}^1 \times (0, \pi)$ and introduce new coordinates $(x, y) = (t, -\cos\varphi)$.

Proposition 1.1.6. *In the coordinates (x, y), the billiard map*

$$T : \mathbb{R} \times (-1, 1) \to \mathbb{R} \times (-1, 1)$$

is a monotone twist map with h as a generating function and it preserves the area form $dx \wedge dy$.

Proof. Let $(x', y') = T(x, y)$. We have to show that it holds the monotone twist condition

$$\frac{\partial x'}{\partial y} > 0. \tag{1.7}$$

According to the chain rule,

$$\frac{\partial x'}{\partial y} = \frac{\partial \varphi}{\partial y}\frac{\partial t'}{\partial \varphi} = \frac{1}{\sin\varphi}\frac{\partial t'}{\partial \varphi}.$$

Clearly, $\sin\varphi > 0$ in M.

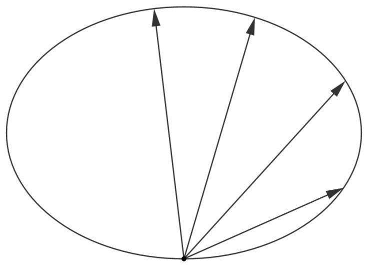

Figure 1.2: Twist condition for convex billiards

Fix a vertical line $t = \mathrm{const}$ in M. This line corresponds to a fixed position of the billiard ball on the billiard curve with all possible directions φ allowed. The convexity of the billiard curve implies that, if we increase the angle φ, the arc from the fixed point to the point at which the ball hits the boundary increases as well (see Figure 1.2) that is

$$\frac{\partial t'}{\partial \varphi} > 0,$$

Therefore, the twist condition 1.7 is satisfied. □

We conclude this section by determining the Ω-area of the phase space. In the next proposition we do not assume that the length $l(\partial D)$ of the boundary curve is normalized to 1.

Proposition 1.1.7. *Let l be the length of the billiard curve ∂D. The area of the phase space equals $2l$.*

Proof. The area of M is given by

$$\int_0^l \int_0^\pi \sin\varphi \, d\varphi \, dt = \int_0^l (-\cos\varphi)\big|_0^\pi \, dt = \int_0^l 2\, dt = 2l. \qquad \square$$

1.2 Birkhoff billiards in the space of rays of the plane

In this section we present a different approach –inspired by geometrical optics– to the study of Birkhoff billiards, based on the oriented lines (or rays) of the plane. We refer to [31, Section 1.3] and [33, pp. 34-36] for a more detailed explanation.
Let N be the set of all oriented lines in the plane. Any oriented line can be characterized by its direction ψ and its signed distance p from the origin of the plane. Thus, (ψ, p) are coordinates on the space of rays, and N can be identified with the infinte cylinder $\mathbb{S}^1 \times \mathbb{R}$.
The space of oriented rays in the plane admits a unique –up to a constant factor– area form invariant under the motions of the plane (see [33, Lemma 3.5]). In our notation, this area form is given by $\Lambda = d\psi \wedge dp$.
Consider now the set $U \subset N$ of oriented lines that intersect the billiard table D. U is given by the inequality $|p| \leqslant f(\psi)$, where $f(\psi)$ depends on the shape of M, and therefore it is diffeomorphic to an annulus.
We define the billiard map T' in U as follows. The ray that contains a segment of the trajectory of the billiard ball, oriented by the direction of its motion, is sent to the ray that contains the next segment of this trajectory after the reflection in the boundary.
The phase space M of the billiard, introduced in the previous section, is identified with U by the map

$$\phi : M \to U$$

that associates a point $(x, v) \in M$ with an oriented line through x with direction v. Thus, we can identify U with M, as well as the two billiard map T and T'. Moreover, it holds the following

Lemma 1.2.1. *The area forms Ω and Λ are equal. That is, $\phi^*(\Lambda) = \Omega$.*

Proof. Let (t, φ) be coordinates in M, and (ψ, p) the respective coordinates in N. In order to prove the lemma, we have to show that

$$d\psi \wedge dp = \sin\varphi \, d\varphi \wedge dt.$$

Denote by $\vartheta(t)$ the direction of the positive tangent line to the billiard curve γ at $\gamma(t)$, and let $\gamma_1(t)$ and $\gamma_2(t)$ be the two components of $\gamma(t)$. It follows that

$$\psi = \varphi + \vartheta(t) \tag{1.8}$$

and

$$p = \gamma_1(t) \sin\psi - \gamma_2(t) \cos\psi, \tag{1.9}$$

see Figure 1.3.
Differentiating equations (1.8) and (1.9), we obtain

$$d\psi = d\varphi + \vartheta'(t)\,dt$$

and

$$dp = [\gamma_1'(t)\sin\psi - \gamma_2'(t)\cos\psi]\,dt + [\gamma_1(t)\cos\psi + \gamma_2(t)\sin\psi]\,d\psi$$

and therefore

$$\begin{aligned} d\psi \wedge dp &= [\gamma_1'(t)\sin\psi - \gamma_2'(t)\cos\psi]\,d\varphi \wedge dt \\ &\quad + \vartheta'(t)[\gamma_1'(t)\sin\psi - \gamma_2'(t)\cos\psi]\,dt \wedge dt \\ &\quad + [\gamma_1(t)\cos\psi + \gamma_2(t)\sin\psi]\,d\psi \wedge d\psi \\ &= [\gamma_1'(t)\sin\psi - \gamma_2'(t)\cos\psi]\,d\varphi \wedge dt. \end{aligned}$$

Finally, observe that $(\gamma_1'(t), \gamma_2'(t)) = (\cos\vartheta(t), \sin\vartheta(t))$, which implies

$$[\gamma_1'(t)\sin\psi - \gamma_2'(t)\cos\psi] = \cos\vartheta\sin\psi - \sin\vartheta\cos\psi = \sin(\psi - \vartheta) = \sin\varphi.$$

Therefore,

$$d\psi \wedge dp = \sin\varphi\, d\varphi \wedge dt,$$

as claimed. □

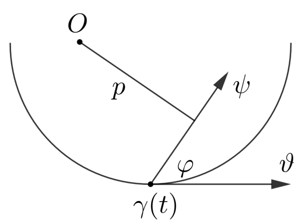

Figure 1.3: Relating the area forms Ω and Λ

As a consequence, the billiard map T' preserves the area form Λ of the space of rays in the plane.

1.3 Caustics for Birkhoff billiards

In this section we define caustics for Birkhoff billiards and discuss some results about their existence.

Let $D \in \mathbb{R}^2$ be a strictly convex domain with smooth boundary. As in Section 1.2, we identify the phase space M of the billiard map T with the set of oriented lines in the plane that intersects D.

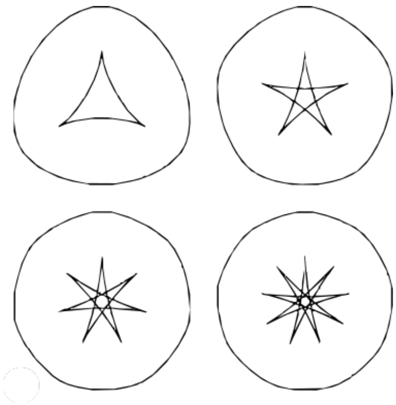

Figure 1.4: Non-convex caustics for a billiard of constant width

Definition 1.3.1. A curve $\delta \subset M$ is an invariant circle for the billiard map T if it is T-invariant and isotopic to a boundary component of M.

Observe that both boundary components of M are trivial invariant circles. Since the billiard map T is a monotone twist map, its invariant circles cannot take any form, but they are subject to the following Birkhoff's theorem (see [29, Chapter 1]).

Theorem 1.3.2. *Any invariant circle of a monotone twist map is the graph of a Lipschitz function.*

An invariant circle δ of M can be identified with a 1-parameter family of rays intersecting the billiard table (this identification is due to the projective duality between the plane and the space of oriented lines in this plane, see [33, pp. 86-87] for more details).
Denote with D_δ the intersection of the closed left half planes generated by the family of rays. By definition, D_δ is a convex closed set and it is contained in D. We say that the invariant circle δ is convex if every ray of the family is a supporting line of D_δ.

Definition 1.3.3. Given a convex invariant circle δ, a convex caustic is the boundary of the convex set D_δ

$$\Gamma = \partial D_\delta.$$

Clearly, Γ is a $\mathcal{C}^1$ simple closed convex curve in the interior of D. It follows from the definition that every time a trajectory is tangent to Γ, it remains tangent after every each reflection.

Remark 1.3.4. It is possible to consider a more general definition of caustics, which does not require them to be convex or even close. Given an invariant curve

δ of the billiard map, we identify it with a 1-parameter family of oriented lines intersecting the billiard table, and define the caustic corresponding to δ as the envelope of the rays of the family.
In [14, Section 3], it is provided an example of a billiard table with a non convex caustic. Consider a curve of constant width; there exists a chord, in any direction, that is perpendicular to the the curve at both ends. These chords are 2-periodic billiard trajectories and their envelope is a caustic, which corresponds to the involute of the billiard curve. In general, the involute of such a curve has cusps, see Figure 1.4.

Constructing a billiard with at least one convex caustic is always possible, thank to the so-called string construction. Let Γ be a curve and wrap a closed non-stretchable string around it, pull it tight at a point and move this point around Γ; the curve that one obtains correspond to a billiard domain that has Γ as a caustic.
Moreover, next theorem, proved by Lazutkin in [18], assures the existence of infinitely many convex caustics if the boundary of the billiard curve is sufficiently regular and has everywhere positive curvature.

Theorem 1.3.5. *Let D be a strictly convex planar domain and suppose that the boundary curve ∂D is $\mathcal{C}^6$ and has everywhere positive curvature. Then there exists a positive measure set of caustics accumulating to the boundary of the billiard table.*

The idea of the proof is to introduce the following change of coordinates

$$x = C^{-1}\int_0^t \rho^{-2/3}(\tau)\,d\tau \qquad y = 4\,C^{-1}\rho^{1/3}(t)\,\sin\frac{\varphi}{2}$$

where ρ denotes the radius of curvature of ∂D and $C = \int_{\partial D}\rho^{-2/3}(t)dt$, under which the billiard map becomes

$$T(x,y) = (x + y + \mathcal{O}(y^3), y + \mathcal{O}(y^4)).$$

Near the boundary $\{y = 0\}$ this map can be seen as a small perturbation of the integrable, area-preserving map $(x,y) \mapsto (x+y,y)$. Therefore, a version of KAM theory for twist maps can be applied and the existence of a positve measure Cantor set of invariant circles near the boundary is assured. This set translates into a positive measure set of caustics accumulating to the boundary of the billiard table.
Observe that, to ensure the existence of caustics for a Birkhoff billiard, it is essential that the billiard curve has everywhere positive curvature. Indeed, a theorem by Mather [22] shows the non-existence of caustics if the billiard curve has a point of null curvature.

Theorem 1.3.6. *If the curvature of a $\mathcal{C}^2$ smooth convex billiard curve vanishes at some point, then the billiard transformation has no invariant circles.*

In [14], Gutkin and Katok provided an alternative proof of this result. Moreover, they investigate how the shape of the domain determines the location of caustics, establishing that there exists an open region inside the billiard table "free of caustics" and estimating from below the area of this region.

The next examples describe Birkhoff billiards respectively in a disk and in an ellipse.

Example 1.3.7 (Circular billiards). Consider a disc of radius R. The billiard map T is completely determined by the angle of reflection, which remains constant at each reflection. Denoting by $t \in \mathbb{S}^1$ the arc-length parameter and by $\varphi \in (0, \pi/2]$ the angle of reflection, the billiard map can be written in the simple form (see Figure 1.5)

$$T(t, \varphi) = (t + 2R\,\varphi, \varphi).$$

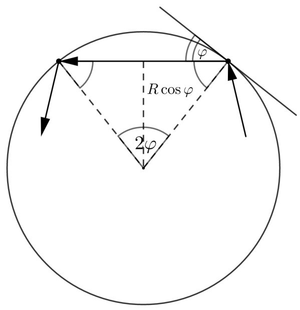

Figure 1.5: The billiard map in the disk

The angle φ remains constant along the orbit and therefore it is an integral of motion for T.
The properties of the orbits are determined by the value of the angle φ:

- (i) if $\varphi = \frac{p}{q}\pi$, with $\frac{p}{q} \in (0, 1/2]$ in lowest terms, the orbit is periodic with minimal period q and makes p turns about the circle;
- (ii) if φ is irrational the orbit is not periodic and it hits the boundary on a dense set of points.

In particular, all orbits determined by the same angle φ are tangent to the same concentric circle of radius $R\cos\varphi$. This concentric circle is a caustic for the circular billiard and it corresponds to an invariant circle for the billiard map on the phase space, which is topologically a cylinder. The phase space is completely foliated by these invariant circles and correspondingly the billiard table is completely foliated by caustics, see Figure 1.6.

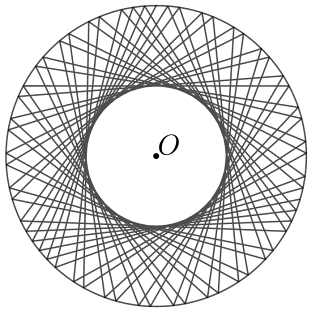

Figure 1.6: Birkhoff billiard in a circle

Example 1.3.8 (Elliptic billiards). Consider an ellipse with foci F_1 and F_2. In this case, writing the billiard map explicitly is more complicated than in the circular case, but using the optical properties of conics –already known in ancient Greece– it is possible to describe the dynamics of an elliptic billiard table.
Each trajectory that does not pass through a focal point is always tangent to a fixed confocal conics. More precisely, each trajectory either:

(i) never intersects the segment F_1F_2 between the two foci and is then always tangent to a confocal ellipse (see Figure 1.7a);

(ii) always intersects the segment F_1F_2 between the two foci and is then always tangent to a confocal hyperbola (see Figure 1.7b);

(iii) always passes through one of the two foci F_1, F_2 alternately and tends asymptotically to the major semiaxis (see Figure 1.7c).

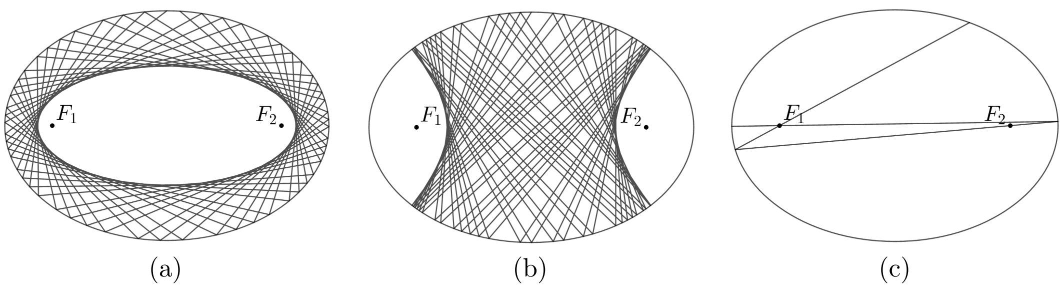

Figure 1.7: Birkhoff billiard in an ellipse

Observe, in addition, that the major and minor axes of the ellipse are 2-periodic orbits for the billiard.
Confocal ellipses are examples of caustics and they foliate the billiard table, except

for the segment between the two foci.
Observe that hyperbolae can also be considered caustics for the elliptic billiard. However, they do not correspond to invariant circles of the billiard map, but to contractible invariant curves.

1.4 Integrability in Birkhoff billiards

Integrability in Birkhoff billiards can be defined in different ways:

(i) through the existence of an integral of motion, globally or locally in the phase space;

(ii) through the existence of a smooth foliation of the phase space (globally or locally), consisting of invariant curves of the billiard map, which translates –under suitable conditions– into a foliation of the billiard table (globally or locally), consisting of smooth convex caustics.

Observe that both circular and elliptic billiard tables are foliate by convex caustics. In the first case, they are concentric circles and the foliation is global (see Example 1.3.7); in the second, confocal ellipses are the caustics and they foliate everything but the segment between the two foci (see Example 1.3.8). Thus, circular and elliptic billiards are integrable.
According to Birkhoff conjecture, these are the only examples of integrable billiards.

Conjecture (Birkhoff). *Circular and elliptic billiards are the only examples of integrable Birkhoff billiards.*

Several attempts had been made to prove this conjecture, but so far it remains open. An attempt in this direction is a theorem by Bialy in [5], asserting the uniqueness of circular billiards.

Theorem 1.4.1. *If the phase space of the billiard ball map is fully foliated by continuous invariant circles, then it is a circular billiard.*

Wojtkowski provided an alternative proof of this theorem by means of the mirror equation, see [34].
Another approach to this open problem is represented by the perturbative Birkhoff's conjecture, which focus only on a particular class of Birkhoff billiards, whose domain can be considered as perturbations of ellipses.

Conjecture (perturbative Birkhoff). *A smooth strictly convex domain that is sufficiently close (with respect to some topology) to an ellipse and whose corresponding billiard map is integrable, is necessarily an ellipse.*

Recently, this conjecture was proved by Avila, De Simoi and Kaloshin in [3] for domains that are sufficiently close to a circular billiard. The complete proof for domains sufficiently close to an ellipse of any eccentricity was proved by Kaloshin and Sorrentino in [17].

Chapter 2

Symplectic billiards

Abstract. In this chapter we introduce symplectic billiards on the plane. We give and discuss the definition. We establish that the symplectic billiard map admits a generating function, which is essentially the standard area form. We prove that the symplectic billiard map is a monotone twist map. Finally, we focus on circular and elliptic billiards; in both cases, we prove that –as for Birkhoff billiards– the corresponding discrete dynamical systems are integrable.

2.1 The standard area form

Consider two vectors $v = (v_1, v_2)$ and $w = (w_1, w_2)$ in $\mathbb{R}^2$ anchored at a point $x \in \mathbb{R}^2$. We define the standard (oriented) area form ω of v and w as the area –up to the sign– of the parallelogram spanned by v and w

$$\omega(x; v, w) = det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = v_1 w_2 - v_2 w_1,$$

see Figure 2.1.

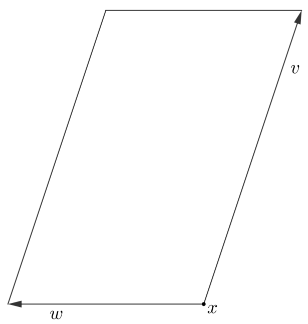

Figure 2.1: The standard area form

Here some basic properties of the standard area form:

(i) bilinearity

$$\omega(x; \alpha u + \beta v, w) = \alpha\, \omega(x; u, v) + \beta\, \omega(x; v, w),$$
$$\omega(x; u, \alpha v + \beta w) = \alpha\, \omega(x; u, v) + \beta\, \omega(x; v, w)$$

for any $x, u, v, w \in \mathbb{R}^2$ and $\alpha, \beta \in \mathbb{R}$;

(ii) antisymmetry

$$\omega(x; v, w) = -\omega(x; w, v)$$

for any $x, v, w \in \mathbb{R}^2$;

(iii) $\omega(x; v, w) = 0$ if and only if $w = \lambda\, v$, with $\lambda \in \mathbb{R}$.

In the sequel, if the point x is the origin of $\mathbb{R}^2$, we denote the oriented area form of two vectors v and w in $\mathbb{R}^2$ simply as $\omega(v, w)$.

Lemma 2.1.1. *The next equality holds*

$$\omega(\tau; v, w) = \omega(v, w) + \omega(v - w, \tau) \tag{2.1}$$

for every $\tau, v, w \in \mathbb{R}^2$.

Proof. Using the bilinearity, the antisimmetry of ω and the fact that $\omega(\tau, \tau) = 0$, we have

$$\begin{aligned}\omega(\tau; v, w) &= \omega(v + \tau, w + \tau)\\ &= \omega(v, w) + \omega(v, \tau) + \omega(\tau, w) + \omega(\tau, \tau)\\ &= \omega(v, w) + \omega(v, \tau) - \omega(w, \tau)\\ &= \omega(v, w) + \omega(v - w, \tau).\end{aligned}$$ □

2.2 Symplectic billiard: definition

From now on, $x \in \mathbb{R}^2$ is always the origin of the Cartesian plane.
Let γ be a smooth, strictly convex, closed and positively (that is counterclockwise) oriented curve. This curve γ is the boundary of the billiard table D, that is $+\partial D = \gamma$.
In the sequel, we denote by $T_x\gamma$ the tangent line at γ in the point x. A property of strictly convex curves is that for every point $x \in \gamma$ there exists a unique point $x^* \in \gamma$ such that $T_x\gamma = T_{x^*}\gamma$. These points are called opposite and clearly $(x^*)^* = x$. We refer to Figure 2.2.
We observe that, if ν_x denotes the outer normal in $x \in \gamma$, then $T_x\gamma = T_y\gamma$ if and only if $\omega(\nu_x, \nu_y) = 0$. Finally, since the curve γ is positively oriented, if we

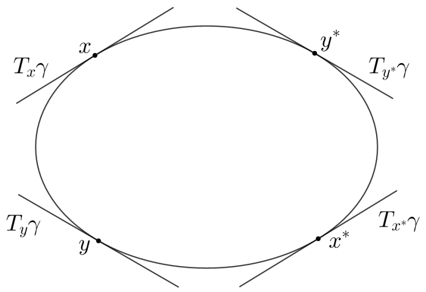

Figure 2.2: Opposite points

fix a point $x \in \gamma$ then for all the points $y \in \gamma$ such that $x < y < x^*$ it holds $\omega(\nu_x, \nu_y) > 0$.

We first introduce the phase space of the symplectic billiard map.

Definition 2.2.1. The phase space of the symplectic billiard map is

$$\mathcal{P} := \{(x, y) \in \gamma \times \gamma \mid x < y < x^*\} = \{(x, y) \in \gamma \times \gamma \mid \omega(\nu_x, \nu_y) > 0\}.$$

In order to describe the discrete dynamical system (the so-called "symplectic billiard") on $\mathcal{P}$, we need the next

Lemma 2.2.2. *Let $(x, y) \in \mathcal{P}$. Then there exists a unique point $z \in \gamma$ with $z - x \in T_y\gamma$. Moreover, $(y, z) \in \mathcal{P}$.*

Proof. We first prove existence and uniqueness of such a point $z \in \gamma$. For $(x, y) \in \mathcal{P}$, by definition, we have that the tangent spaces $T_x\gamma$ and $T_y\gamma$ are transversal. Since γ is convex, there exists a unique point $z \in \gamma$ such that $(x+T_y\gamma) \cap \gamma = \{x, z\}$. Moreover $z \neq x$; otherwise, it would be $T_y\gamma = T_z\gamma = T_x\gamma$, which contradicts the hypothesis $(x, y) \in \mathcal{P}$.

In order to conclude the proof, we have to show that $(y, z) \in \mathcal{P}$. We easily remark that if y is close to x then so is z and therefore $\omega(\nu_x, \nu_y) > 0$ implies $\omega(\nu_y, \nu_z) > 0$. For $(x, y) \in \mathcal{P}$ suppose –by contradiction– that $(y, z) \notin \mathcal{P}$, i.e., $\omega(\nu_y, \nu_z) \leqslant 0$. By continuity and moving y close to x, we can arrange $\omega(\nu_y, \nu_z) = 0$. This means that $T_y\gamma = T_z\gamma$ which is equivalent to $x = y$. Since we have contradicted the hypothesis $(x, y) \in \mathcal{P}$, we have concluded the proof. □

We refer to Figure 2.3.

Definition 2.2.3. The symplectic billiard map is

$$\phi : \mathcal{P} \to \mathcal{P} \qquad (x, y) \mapsto (y, z)$$

where $z \in \gamma$ is the unique point satisfying $z - x \in T_y\gamma$.

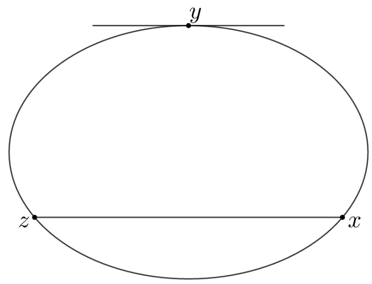

Figure 2.3: The unique point $z \in \gamma$ with $z - x \in T_y\gamma$

Remark 2.2.4. It is clearly possible to consider the (negative part of the) phase space, that is

$$\mathcal{P}^- := \{(x, y) \in \gamma \times \gamma \mid x^* < y < x\}.$$

The corresponding symplectic billiard map on $\mathcal{P}^-$ is the same as in Definition 2.2.3 once the orientation of γ is reversed.

Remark 2.2.5. We observe that the map $\phi : \mathcal{P} \to \mathcal{P}$ can be extended by continuity to the closure

$$\bar{\mathcal{P}} = \{(x, y) \in \gamma \times \gamma : \ x \leqslant y \leqslant x^*\}.$$

In fact,

$$\lim_{y \to x} \phi(x, y) = (x, x)$$

which means that the map extends to the identity and

$$\lim_{y \to x^*} \phi(x, y) = (x^*, x)$$

which follows from the fact that, due to the convexity of γ, the function $y \mapsto z(y)$ is monotone and $\lim_{y \to x^*} T_y\gamma = T_{x^*}\gamma$.

Lemma 2.2.6. *The continuous extension $\phi(x, x^*) = (x^*, x)$ is characterized by the 2-periodicity. This means that $\phi(x, y) = (y, x)$ is equivalent to $y \in \{x, x^*\}$.*

Proof. Let $(x, y) \in \mathcal{P}$ and $\phi(x, y) = (y, x)$. Then from Lemma 2.2.2 we have

$$(x + T_y\gamma) \cap \gamma = \{x, y\}$$

with $x \neq y$. This is a contraddiction because $T_y\gamma \cap \gamma = y$. So, the only possibilities are $(x, y) = (x, x)$ or $(x, y) = (x^*, x)$.
Suppose now that $y \in \{x, x^*\}$. It follows from the previuous remark that $\phi(x, x) = (x, x)$ and $\phi(x, x^*) = (x^*, x)$. $\square$

The envelope of the 1-parameter family of chords xx^* is a caustic of the symplectic billiard. This envelope is the so-called centre symmetry set of a curve γ which – in case of a strictly convexity– is defined as the envelope of lines joining opposite points of the curve, see [10, p.92]. Clearly, in the case of centrally symmetric curves (as the circle and the ellipse) the centre symmetry set is the centre of symmetry.

2.3 Symplectic billiard: generating function

With the next lemma we establish that the symplectic billiard map ϕ admits a generating function, which involves the standard area form $\omega(x, y)$.

Lemma 2.3.1. *The generating function for the symplectic billiard map* ϕ *is*

$$S : \mathcal{P} \to \mathbb{R} \qquad (x, y) \mapsto S(x, y) := \omega(x, y).$$

This means that

$$\phi(x, y) = (y, z) \;\Leftrightarrow\; \frac{d}{dy}[S(x, y) + S(y, z)] = 0. \tag{2.2}$$

Proof. Recall that, given $x = (x_1, x_2)$ and $y = (y_1, y_2) \in \mathbb{R}^2$, we have that

$$\omega(x, y) = x_1 y_2 - x_2 y_1$$

and so

$$\frac{d}{dy} S(x, y) = \frac{\partial}{\partial y_1}\omega(x, y)\, dy_1 + \frac{\partial}{\partial y_2}\omega(x, y)\, dy_2 = -x_2\, dy_1 + x_1\, dy_2$$

and

$$\frac{d}{dy} S(y, z) = \frac{\partial}{\partial y_1}\omega(y, z)\, dy_1 + \frac{\partial}{\partial y_2}\omega(y, z)\, dy_2 = z_2\, dy_1 - z_1\, dy_2.$$

It follows

$$\frac{d}{dy}[S(x, y) + S(y, z)] = (z_2 - x_2)\, dy_1 + (x_1 - z_1)\, dy_2.$$

Hence, given $x, y, z \in \gamma$ with $x \neq z$, we have that

$$\begin{aligned} \frac{d}{dy}[S(x, \cdot) + S(\cdot, z)] = 0 &\Leftrightarrow (z_2 - x_2, x_1 - z_1) \perp T_y\gamma \\ &\Leftrightarrow z - x \in T_y\gamma \end{aligned}$$

which corresponds exactly to Definition 2.2.3. □

Remark 2.3.2. The formula

$$\frac{d}{dy}[S(x,y)+S(y,z)]=0$$

has the the following variational interpretation. Given two points $x, z \in \gamma$, suppose that the billiard ball starts at the point x and –after a collision at $y \in \gamma$– it arrives at the point z. Then the point y is the critical point of the function $y \mapsto S(x,y)+S(y,z)$.

Let $A(x,y,z)$ be the area of the triangle of vertices $x, y, z \in \gamma$, see Figure 2.4. From the next lemma, we conclude that

$$S(x,y)+S(y,z)=2A(x,y,z)-\omega(z,x) \tag{2.3}$$

that is, the quantity involved in (2.2) differs from $2A(x,y,z)$ by the function $\omega(z,x)$, which has no effect on the partial derivative with respect to y.

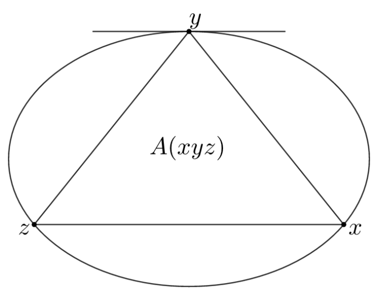

Figure 2.4: The triangle of vertices $x, y, z \in \gamma$

Lemma 2.3.3. *The area of the triangle xyz is*

$$\frac{1}{2}[\omega(x,y)+\omega(y,z)+\omega(z,x)]$$

Proof. If x is the origin of the Cartesian axis, we have

$$\omega(0,y)+\omega(y,z)+\omega(z,0)=\omega(y,z)=2A(x,y,z),$$

see Figure 2.5.
In the general case, it is sufficient to apply Lemma 2.1.1

$$\begin{aligned}
\omega(x,y) &= \omega(0+x,(y-x)+x)=\omega(x-y,x)\\
\omega(y,z) &= \omega((y-x)+x,(z-x)+x)=\omega(y-x,z-x)+\omega(y-z,x)\\
\omega(z,x) &= -\omega(x,z)=-\omega(x-z,x)
\end{aligned}$$

and therefore, by a trivial calculation, we obtain

$$\omega(x,y)+\omega(y,z)+\omega(z,x)=\omega(y-x,z-x)=2A(x,y,z). \qquad \square$$

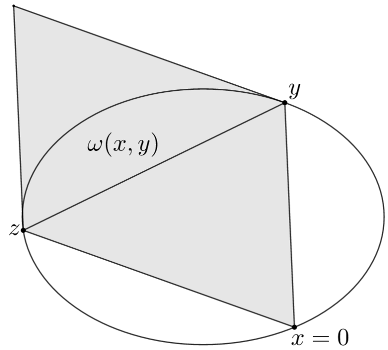

Figure 2.5: The area $\omega(y,z) = 2A(x,y,z)$

From previous lemma, equality 2.3 immediately follows.
The name “symplectic billiard” is due to the fact that the generating function of the map describing the dynamics involves the standard area form, which is the symplectic 2-form on the plane. Consequently, as explained in Chapter 4, the definition of symplectic billiard can be extended to the linear symplectic space $(\mathbb{R}^{2n}, \omega)$. Finally, we remind that –in the case of Birkhoff billiards– the generating function corresponds to the length of the trajectory segment joining two consecutive collision points.

2.4 Symplectic billiard as a monotone twist map

In this section we prove that the symplectic billiard map (i) admits an invariant area form Ω (Lemma 2.4.1) and (ii) is a monotone twist map (Proposition 2.4.2). Moreover –see Theorem 2.4.6– we calculate the Ω-area of the phase space $\mathcal{P}$.

Assume that the length of the symplectic billiard curve γ is normalized to 1 and parametrize γ by the arc length $t \in \mathbb{S}^1$. Denote by $t^* \in \mathbb{S}^1$ the unique parameter such that $\gamma(t)$ and $\gamma(t^*)$ are opposite points (that is, the tangents at $\gamma(t)$ and $\gamma(t^*)$ are parallel). Hence, a state of the system is now represented by the pair (t_1, t_2) where $t_1 < t_2 < t_1^*$ and the phase space $\mathcal{P}$ can be identified with $\mathbb{S}^1 \times \mathbb{S}^1$. According to Lemma 2.3.1, the generating function becomes

$$S(t_1, t_2) = \omega(\gamma(t_1), \gamma(t_2)) \tag{2.4}$$

In the sequel, we denote by $S_i := \frac{\partial S}{\partial t_i}, S_{ij} := \frac{\partial^2 S}{\partial t_i \partial t_j}$ for $i, j = 1, 2$.

Lemma 2.4.1. *The 2-form*

$$\Omega := S_{12}(t_1, t_2) dt_1 \wedge dt_2$$

is an area form on $\mathcal{P}$ which is ϕ-invariant, that is $\phi^*\Omega = \Omega$.

Proof. Let identify $\mathbb{R}^2$ with $\mathbb{C}$ and observe that multiplying the coordinates of a vector by $-i$ equals to clockwise rotating it by an angle of $\frac{\pi}{2}$.
We first prove that the 2-form Ω is an area form, that is $S_{12}(t_1,t_2) > 0$. Given a point $\gamma(t)$, let indicate by $\gamma'(t)$ the tangent vector of γ in $\gamma(t)$. Consequently, since γ is positively oriented, $-i\gamma'(t)$ represents the outward normal vector of γ in $\gamma(t)$.
Moreover, we recall that, for a given state of the system $(t_1,t_2) \in \mathbb{S}^1 \times \mathbb{S}^1$, it holds

$$\omega(\nu_{\gamma(t_1)}, \nu_{\gamma(t_2)}) = \omega(-i\gamma'(t_1), -i\gamma'(t_2)) > 0$$

see Definition 2.2.1.
Hence

$$S_{12}(t_1,t_2) = \omega(\gamma'(t_1), \gamma'(t_2)) = \omega(-i\gamma'(t_1), -i\gamma'(t_2)) > 0 \tag{2.5}$$

We finally prove that the 2-form Ω is ϕ-invariant. By equality (1.1),

$$\phi(t_1,t_2) = (t_2,t_3) \Leftrightarrow S_2(t_1,t_2) + S_1(t_2,t_3) = 0.$$

Consequently, we need to show that

$$S_{12}(t_1,t_2)\, dt_1 \wedge\, dt_2 = S_{12}(t_2,t_3)\, dt_2 \wedge\, dt_3.$$

We take the exterior derivative of $S_2(t_1,t_2) + S_1(t_2,t_3) = 0$

$$\begin{aligned} &d(S_2(t_1,t_2) + S_1(t_2,t_3)) \\ &\quad = S_{21}(t_1,t_2)\, dt_1 + S_{22}(t_1,t_2)\, dt_2 + S_{11}(t_2,t_3)\, dt_2 + S_{12}(t_2,t_3)\, dt_3 \\ &\quad = 0 \end{aligned}$$

and then we right-wedge-multiply by dt_2

$$\begin{aligned} &S_{12}(t_1,t_2)\, dt_1 \wedge dt_2 + S_{12}(t_2,t_3)\, dt_3 \wedge dt_2 \\ &\quad = S_{12}(t_1,t_2)\, dt_1 \wedge dt_2 - S_{12}(t_2,t_3)\, dt_2 \wedge dt_3 \\ &\quad = 0 \end{aligned}$$

We have obtained the desired equality. □

The next proposition is a direct consequence of inequality (2.5).

Proposition 2.4.2. *The symplectic billiard map is a (negative) monotone twist map.*

Proof. Let denote

$$s_1 := -S_1(t_1,t_2) \qquad s_2 := S_2(t_1,t_2)$$

First of all we observe that (t_1, s_1) are global coordinates in $\mathcal{P}$. Indeed, since

$$\frac{\partial s_1}{\partial t_2} = -S_{12}(t_1, t_2) < 0,$$

the Jacobian matrix of the map $(t_1, t_2) \mapsto (t_1, s_1)$

$$\begin{pmatrix} 1 & 0 \\ \frac{\partial s_1}{\partial t_1} & \frac{\partial s_1}{\partial t_2} \end{pmatrix}$$

is not degenerate. In these coordinates, the dynamics is given by $(t_1, s_1) \mapsto (t_2, s_2)$ and the twist condition

$$\frac{\partial t_2}{\partial s_1} < 0$$

immediately follows from $\frac{\partial s_1}{\partial t_2} = -S_{12}(t_1, t_2) < 0$. □

Remark 2.4.3. As a consequence of previous proposition, symplectic billiards can be handled also within the rich context of monotone twist maps. We refer to [4] and [12] for an exhaustive treatment of the matter.

In the sequel, we establish the Ω-area of the phase space $\mathcal{P}$. We premise some technical facts.

Let D be the billiard table bounded by the curve γ, that is $+\partial D = \gamma$. Choose an origin O in the interior of D and parametrize the curve γ by using the direction α of its tangent line. Let $p(\alpha)$ be the so-called support function of γ, i.e., the distance between O and the tangent line in $\gamma(\alpha)$, with direction $\alpha - \frac{\pi}{2}$, see Figure 2.6.

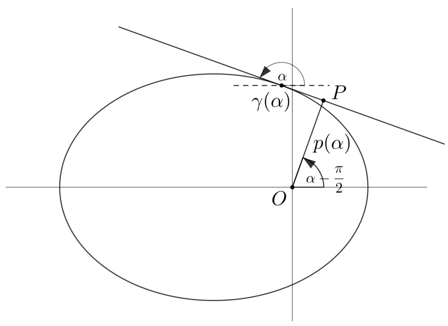

Figure 2.6: The support function $p(\alpha)$ of the curve γ

The curve γ can be seen as the envelope of the family of its tangent lines, whose equation is

$$x \cos\left(\alpha - \frac{\pi}{2}\right) + y \sin\left(\alpha - \frac{\pi}{2}\right) - p = 0,$$

which is

$$x \sin\alpha - y \cos\alpha - p(\alpha) = 0. \tag{2.6}$$

The envelope of the family is obtained by 2.6 and the derivative

$$x \cos\alpha + y \sin\alpha - p'(\alpha) = 0. \tag{2.7}$$

Thus the parametric representation of γ is

$$x(\alpha) = p(\alpha)\sin\alpha + p'(\alpha)\cos\alpha \qquad y(\alpha) = -p(\alpha)\cos\alpha + p'(\alpha)\sin\alpha \tag{2.8}$$

where $(x(\alpha), y(\alpha))$ are the coordinates of the contact point $\gamma(\alpha)$ between the tangent line and the curve γ.
By formulas (1.6), (1.7) in [27, p.3], the perimeter of γ and the area of D are given, respectively, by the integrals:

$$\text{perimeter}(\gamma) = \int_0^{2\pi} p(\alpha)\, d\alpha$$

and

$$\text{area}(D) = \frac{1}{2}\int_0^{2\pi} [p(\alpha) + p''(\alpha)]p(\alpha)\, d\alpha.$$

We observe that the tangent lines at γ with direction α and $\alpha + \pi$ are parallel that is, $\gamma(\alpha)$ and $\gamma(\alpha + \pi)$ are opposite points. Thus, the phase space $\mathcal{P}$ consists of pairs (α_1, α_2) with $\alpha_1 < \alpha_2 < \alpha_1 + \pi$, and the generating function becomes $S(\alpha_1, \alpha_2) = \omega(\gamma(\alpha_1), \gamma(\alpha_2))$.
In order to state the result about the Ω-area of the phase space, we need to recall the definitions of (i) Minkowski sum of two sets and (ii) centrally symmetric domain of a set.

Definition 2.4.4. The Minkowski sum of two convex sets X, Y is defined as

$$X +_{\mathcal{M}} Y = \{\, x + y \mid x \in X, y \in Y \,\}.$$

Definition 2.4.5. The centrally symmetric domain X_c of a convex set X is the reflection of X in the origin O.
Observe that, if $p(\alpha)$ is the support function of X, then $p(\alpha + \pi)$ is the support function of X_c.

Moreover, we recall the following two properties of the support function:

(i) if p_X is the support function of a convex set X, then the support function of λX is $p_{\lambda X} = \lambda p_X$, for any $\lambda \in \mathbb{R}$, $\lambda \geqslant 0$;

(ii) if p_X and p_Y are the support functions respectively of X and Y, then the support function of their Minkowski sum $X +_{\mathcal{M}} Y$ is $p_{XY} = p_X + p_Y$.

Let now $\bar{D}$ be the so-called symmetrization of the domain D, which is the Minkowski sum of D with its centrally symmetric domain D_c, scaled by a factor 1/2

$$\bar{D} := \frac{1}{2}(D +_{\mathcal{M}} D_c).$$

Then, by the properties (i) and (ii) of the support function, it follows that the support function of $\bar{D}$ is

$$\bar{p}(\alpha) = \frac{1}{2}[p(\alpha) + p(\alpha + \pi)]. \tag{2.9}$$

We are now ready to prove the next

Theorem 2.4.6. *The Ω-area of the phase space $\mathcal{P}$ equals four times the area of the symmetrization $\bar{D}$.*

Proof. We start by recalling that, by Lemma 2.4.1, the area form is

$$\begin{aligned}\Omega &= S_{12}(\alpha_1, \alpha_2)\, d\alpha_1 \wedge d\alpha_2 \\ &= \omega(\gamma'(\alpha_1), \gamma'(\alpha_2))\, d\alpha_1 \wedge d\alpha_2 \\ &= det(\gamma'(\alpha_1), \gamma'(\alpha_2))\, d\alpha_1 \wedge d\alpha_2.\end{aligned}$$

Consequently, the Ω-area of the phase space $\mathcal{P}$ is given by

$$\int_0^{2\pi} \int_{\alpha_1}^{\alpha_1+\pi} S_{12}(\alpha_1, \alpha_2)\, d\alpha_2\, d\alpha_1.$$

From formula (2.8) we have

$$\gamma'(\alpha) = [p'(\alpha) + p''(\alpha)](\cos\alpha, \sin\alpha)$$

so that

$$\begin{aligned}S_{12}(\alpha_1, \alpha_2) &= [p(\alpha_1) + p''(\alpha_1)][p(\alpha_2) + p''(\alpha_2)](\sin\alpha_2 \cos\alpha_1 - \sin\alpha_1 \cos\alpha_2) \\ &= [p(\alpha_1) + p''(\alpha_1)][p(\alpha_2) + p''(\alpha_2)] \sin(\alpha_2 - \alpha_1).\end{aligned}$$

Therefore the Ω-area of the phase space is

$$\int_0^{2\pi} \int_{\alpha_1}^{\alpha_1+\pi} [p(\alpha_1) + p''(\alpha_1)][p(\alpha_2) + p''(\alpha_2)] \sin(\alpha_2 - \alpha_1)\, d\alpha_2\, d\alpha_1. \tag{2.10}$$

We first consider the inner integral

$$\int_{\alpha_1}^{\alpha_1+\pi} p(\alpha_2) \sin(\alpha_2 - \alpha_1)\, d\alpha_2 + \int_{\alpha_1}^{\alpha_1+\pi} p''(\alpha_2) \sin(\alpha_2 - \alpha_1)\, d\alpha_2 \tag{2.11}$$

and we use integration by parts twice on the second summand

$$\begin{aligned}
&\int_{\alpha_1}^{\alpha_1+\pi} p''(\alpha_2)\sin(\alpha_2-\alpha_1)\,d\alpha_2 \\
&\qquad = p'(\alpha_2)\sin(\alpha_2-\alpha_1)\big|_{\alpha_1}^{\alpha_1+\pi} - \int_{\alpha_1}^{\alpha_1+\pi} p'(\alpha_2)\cos(\alpha_2-\alpha_1)\,d\alpha_2 \\
&\qquad = -\int_{\alpha_1}^{\alpha_1+\pi} p'(\alpha_2)\cos(\alpha_2-\alpha_1)\,d\alpha_2 \\
&\qquad = -p(\alpha_2)\cos(\alpha_2-\alpha_1)\big|_{\alpha_1}^{\alpha_1+\pi} - \int_{\alpha_1}^{\alpha_1+\pi} p(\alpha_2)\sin(\alpha_2-\alpha_1)\,d\alpha_2 \\
&\qquad = p(\alpha_1)+p(\alpha_1+\pi) - \int_{\alpha_1}^{\alpha_1+\pi} p(\alpha_2)\sin(\alpha_2-\alpha_1)\,d\alpha_2.
\end{aligned}$$

Consequently, the inner integral (2.11) equals to $p(\alpha_1)+p(\alpha_1+\pi)$, and the Ω-area of the phase space (2.10) corresponds to

$$\int_0^{2\pi}[p(\alpha)+p''(\alpha)][p(\alpha)+p(\alpha+\pi)]\,d\alpha \overset{(2.9)}{=} 2\int_0^{2\pi}[p(\alpha)+p''(\alpha)]\bar{p}(\alpha)\,d\alpha.$$

We finally recall that – see [27, p.3] – $p(\alpha_1)+p''(\alpha_1)$ is the radius of curvature of γ in $\gamma(\alpha_1)$ and that

$$p(\alpha_1)+p''(\alpha_1) = p(\alpha_1+\pi)+p''(\alpha_1+\pi).$$

Thus

$$\begin{aligned}
&2\int_0^{2\pi}[p(\alpha)+p''(\alpha)]\bar{p}(\alpha)\,d\alpha \\
&\qquad = 2\int_0^{2\pi}[\frac{1}{2}(p(\alpha)+p''(\alpha))+\frac{1}{2}(p(\alpha+\pi)+p''(\alpha+\pi))]\bar{p}(\alpha)\,d\alpha \\
&\qquad \overset{(2.9)}{=} 2\int_0^{2\pi}[\bar{p}(\alpha)+\bar{p}''(\alpha)]p(\alpha)\,d\alpha = 4\,\text{area}(\bar{D}).
\end{aligned}$$

This is the desired equality. □

We remind that –in the case of Birkhoff billiards– the area of the phase space equals twice the perimeter of the boundary curve (see Chapter 1, Proposition 1.1.7).

2.5 Circular and elliptic billiards

This section is devoted to symplectic billiards in a disc and in an ellipse. We prove that they are both integrable.

We first consider the case of a disc of radius $R > 0$. Referring to Figure 2.7, consider two pairs $(x, y), (y, z) \in \mathcal{P}$ such that $(y, z) = \phi(x, y)$. Then, by definition, the tangent to the circle in y is parallel to the chord xz. The radius Oy is perpendicular to the tangent and hence to the chord xz. We call s the intersection point between Oy and xz.
The triangle xOz is isosceles and the segment Os is the height relative to the basis xz. In particular, Os is the bisector of the angle xOz, and thus we have that the angles $\widehat{xOy}, \widehat{yOz}$ are equivalent, that is, $\widehat{xOy} = \widehat{yOz} = \alpha$.
This shows that the corresponding dynamics is the rotation of the angle $\alpha \in [0, \pi)$. As a consequence, the dynamics is the same as the one of Birkhoff billiards in a disc, see Section 1.3. Therefore the billiard table is completely foliated by caustics, that is circular symplectic billiards are integrable.

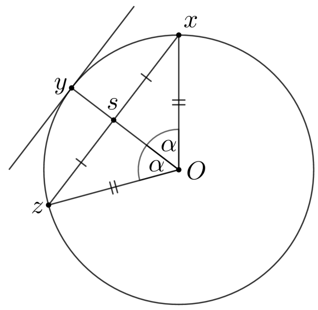

Figure 2.7: The symplectic billiard map in the disk

Moreover, we remark that the standard area form $\omega(x, y)$ is constant along orbits (i.e., it is a global integral of motion) because the triangles xOy and yOz are congruent and $\omega(x, y) = 2A(xOy)$.

The integrability of elliptic symplectic billiards easily follows from the integrability of circular symplectic billiards. The key element to prove the integrability of elliptic symplectic billiard is the notion of affine transformation of the plane.

Definition 2.5.1. An affine trasformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ of the plane is a function of type

$$f(x) := Lx + b$$

where $L \in GL(2, \mathbb{R})$ is an invertible matrix and $b \in \mathbb{R}^2$ is a translation vector.

Among the several properties of affine transformations, we recall that they map lines to lines and they preserve the property of parallelism between lines and the ratio of length of two parallel segments. Consequently, since the definition of the symplectic billiard map involves parallel lines, it follows that the symplectic billiard map commutes with affine transformations of the plane.

Moreover, from the next proposition we deduce that any circle can be mapped into an ellipse by an affine transformation.

Proposition 2.5.2. *Given any ellipse, there exists an affine transformation mapping the ellipse to the unit circle* $x_1^2 + x_2^2 = 1$.

Proof. Consider the ellipse centered at (h, k) and with major and minor axes of lenght $2a$ and $2b$. We can apply a rotation to align the major and minor axis of the ellipse with the axis of the Cartesian plane, and a translation of vector $b = \begin{pmatrix} -h \\ -k \end{pmatrix}$ to translate the center of the ellipse to the origin, obtaining an ellipse of equation:

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

Applying the affine transformation

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1/a \\ x_2/b \end{pmatrix}$$

the ellipse is mapped to the unit circle

$$x_1^2 + x_2^2 = 1. \qquad \square$$

By the previous properties, it becomes clear that –from the point of view of symplectic billiard maps– there is no difference between a circle or an ellipse and elliptic symplectic billiards are integrable as well.
Indeed, consider the affine transformation

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a/R & 0 \\ 0 & b/R \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{R}\begin{pmatrix} ax_1 \\ bx_2 \end{pmatrix}$$

that maps the circle of equation $x_1^2+x_2^2 = R^2$ to the ellipse of equation $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$.

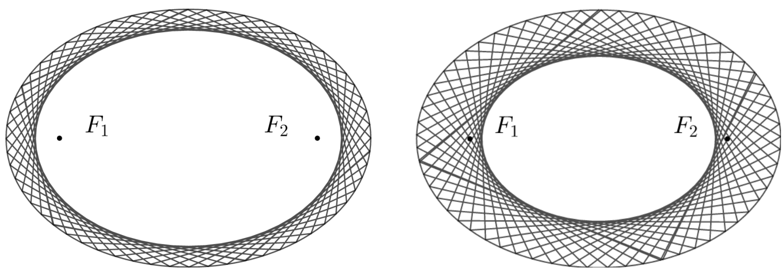

Figure 2.8: Symplectic billiard in an ellipse

The caustics of the circular symplectic billiard are all the concentric circles of equation

$$x_1^2 + x_2^2 = r^2$$

with $0 < r < R$. The affine transformation f maps these circles to ellipses of equation

$$\frac{x_1^2}{\tilde{a}^2} + \frac{x_2^2}{\tilde{b}^2} = 1$$

where

$$\tilde{a} = \frac{r}{R}a \qquad \tilde{b} = \frac{r}{R}b.$$

Since the symplectic billiard map commutes with affine transformations of the plane, these ellipses are caustics for the elliptic symplectic billiard.
In particular, from $0 < r/R < 1$, it follows that $0 < \tilde{a} < a$ and $0 < \tilde{b} < b$, and therefore the elliptic symplectic billiard table is completely foliate by caustics, see Figure 2.8.

Chapter 3

Caustics for symplectic billiards

Abstract. This chapter is devoted to two results for caustics of symplectic billiard. The first one is a non-existence result, which applies when the boundary of the billiard table has points of zero curvature, and is the generalization of Mather's theorem for Birkhoff billiards. The second one is an existence result, applying when the boundary of the billiard table has everywhere positive curvature, and the proof follows the steps –inspired by KAM theory– carried out by Lazutkin.

3.1 Caustics in a symplectic billiard

Let δ be an invariant curve of the symplectic billiard map. This curve can be thought of as a 1-parameter family of chords of the billiard curve γ (that is, each chord of this family corresponds to a point of the invariant curve). We define the caustic corresponding to δ as the envelope of the lines containing the chords of the family.
Therefore, a caustic for the symplectic billiard is a curve with the property that each trajectory that is tangent to it stays tangent after each reflection. Moreover, caustics lie inside the billiard table, as we show in the next

Lemma 3.1.1. *Let Γ be the caustic corresponding to an invariant curve δ of the symplectic billiard map. Then Γ lies inside the billiard table.*

Proof. Let δ be an invariant curve of the symplectic billiard map. According to Birkhoff's Theorem 1.3.2, invariant curves of monotone twist maps are graphs of Lipschitz functions. Therefore, since the symplectic billiard map is monotone twist, δ is a Lipschitz graph.
Consider the 1-parameter family of chords of the symplectic billiard corresponding to δ, and let x_1x_2 be one of these chord. Consider a nearby chord $\bar{x}_1\bar{x}_2$ of the same family and suppose that $\bar{x}_1$ has moved along the billiard curve γ in the positive direction from x_1. Because of the graph property, also $\bar{x}_2$ has moved in the positive direction from x_2 and, therefore, the chords x_1x_2 and $\bar{x}_1\bar{x}_2$ intersect inside the billiard table.
Since the caustic Γ is the envelope of these chords, it follows that Γ lies inside the billiard table. $\square$

In Chapter 2 we have already encountered caustics for symplectic billiards. For example, for strictly convex curves of constant width, the envelope of the 1-parameter family of 2-periodic orbits is a caustic and it corresponds to the centre symmetry set of the curve. Moreover, in Section 2.5, we studied in detail circular and elliptic billiards, showing that their tables are completely foliated by caustics, concentric circles in the first case, and confocal ellipses in the second.

3.2 Points of zero curvature

The result on non-existence of caustics proved in the next section involves points of zero curvature for the curve γ, boundary of the billiard table.

Clearly, if γ is a strictly convex curve –as in Section 2.2– than γ can have only isolated points of zero curvature. We present here below an example of such a curve with four isolated points of zero curvature.

Example 3.2.1. Consider the curve γ implicitly define by the equation

$$x^4 + y^4 = 1 \tag{3.1}$$

see Figure 3.1.

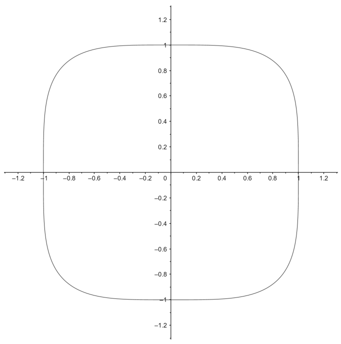

Figure 3.1: The curve $x^4 + y^4 = 1$ is strictly convex with points of zero curvature

We recall –see [11, pp.636-637]– that the curvature of an implicit curve $f(x, y) = 0$, $f \in C^2(\mathbb{R}^2; \mathbb{R})$, is given by the formula

$$k = \frac{|f_{xx}(f_y)^2 - 2f_{xy}f_xf_y + (f_x)^2 f_{yy}|}{|(f_x)^2 + (f_y)^2|^{3/2}}$$

Thus, the curvature of γ defined in (3.1) is

$$k = \frac{3x^2y^6 + 3x^6y^2}{(x^6 + y^6)^{\frac{3}{2}}}$$

which is zero at the points of intersection of γ with the axes $x = 0$ and $y = 0$.

Otherwise, if γ is a convex curve (not strictly convex), then γ can have nonzero measure subsets with zero curvature. We recall that the symplectic billiard map can be generalized to convex curves. In this case, Definition 2.2.3 is no longer applicable; in fact, if we consider two consecutive collision points x, y on a nonzero measure subset with zero curvature, the next collision point z is not uniquely determined because the tangent lines at x and y are the same. This problem can be solved by simply defining the symplectic billiard map via its generating function as in formula (2.2).

3.3 Non-existence of caustics

This section is devoted to prove Theorem 3.3.2, which is the generalization to symplectic billiards of Mather's theorem for Birkhoff billiards, see Theorem 1.3.6. We will give two proofs of Theorem 3.3.2. The first is by contradiction, assuming that a caustic exists and showing that, if the billiard curve has a point of zero curvature, the caustic does not lie inside the billiard table, contradicting Lemma 3.1.1. The second is based on Mather's analytic necessary condition for the existence of invariant curves of monotone twist maps, see [22, pp. 401-402]. We recall here below this result for reader's convenience.

Lemma 3.3.1. *Let*

$$f : \mathbb{S}^1 \times [a, b] \to \mathbb{S}^1 \times [a, b],$$

where $[a, b]$ is an interval of $\mathbb{R}$, be a C^1 diffeomorphism. Let (t, s) be coordinates on $\mathbb{S}^1 \times [a, b]$, with dynamics

$$(t_k, s_k) := f^k(t, s), \qquad k \in \mathbb{Z}$$

Suppose that f is area preserving and monotone twist, and denote with $S : \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{R}$ its generating function. Then invariant curves for f exist only if the next condition holds:

$$S_{22}(t_{-1}, t) + S_{11}(t, t_1) < 0, \qquad \forall t \in \mathbb{S}^1 \tag{3.2}$$

Since caustics correspond to invariant curves of the symplectic billiard map (see Section 3.1), showing that Mather's necessary condition (3.3.1) is violated when the boundary of the billiard table has a point of zero curvature implies the non-existence of caustics.

Theorem 3.3.2. *Let γ be a smooth closed convex curve with a point of zero curvature. Then the symplectic billiard in γ has no caustics.*

Proof 1. Assume, by contradiction, that a caustic Γ for the symplectic billiard exists. Let x_1x_2 and x_2x_3 be two consecutive trajectory segments tangent to the caustic Γ and suppose that the curvature at x_2 vanishes.
Consider a chord $\bar{x}_1\bar{x}_2$ infinitesimally close to the chord x_1x_2 and tangent to the same caustic Γ. Since the curvature at x_2 vanishes, in the linear approximation, the tangent line at $\bar{x}_2$ is the same as the one at x_2. Therefore, in the same linear approximation, the line $\bar{x}_1\bar{x}_3$ is parallel to x_1x_3.

The caustic Γ is defined as the envelope of the chords tangent to it and must lie inside the billiard table, see Lemma 3.1.1. Clearly the chords x_1x_3 and $\bar{x}_1\bar{x}_3$ do not intersect inside the billiard table, and therefore the caustic does not lie inside the billiard table. We have a contradiction and the result follows.

Proof 2. In the proof, to avoid confusion between vectors and scalars, we will denote the vectors with an underline bar.

Parametrize the curve γ by the arc length t. The curvature $k(t)$ at a point $\gamma(t)$ is defined as the norm of the normal vector at $\gamma(t)$

$$k(t) = \|\underline{\gamma}''(t)\|$$

Thus, denoting by $\underline{n}(t) = \frac{\underline{\gamma}''(t)}{\|\underline{\gamma}''(t)\|}$ the inner normal unit vector, we have

$$\underline{\gamma}''(t) = k(t)\,\underline{n}(t) \tag{3.3}$$

We now identify $\mathbb{R}^2$ with $\mathbb{C}$ and remind that multiplying the coordinates of a vector by i equals to counterclockwise rotating it by an angle of $\pi/2$. Consequently, since $\|\underline{\gamma}'(t)\| = 1$, the inner normal unit vector is

$$\underline{n}(t) = \frac{i\underline{\gamma}'(t)}{\|\underline{\gamma}'(t)\|} = i\underline{\gamma}'(t)$$

and identity (3.3) becomes

$$\underline{\gamma}''(t) = i\,k(t)\,\underline{\gamma}'(t)$$

We recall now that the symplectic billiard map is defined via its generating function S by formula (2.2). Hence, we need to show that the corresponding Mather's analytic necessary condition (3.3.1)

$$S_{22}(t_1, t_2) + S_{11}(t_2, t_3) < 0$$

does not hold everywhere on $\mathbb{S}^1$.

Recalling that $S(t_1, t_2) = \omega(\gamma(t_1), \gamma(t_2))$, see formula (2.4), we have

$$S_{22}(t_1, t_2) = \omega(\underline{\gamma}(t_1), \underline{\gamma}''(t_2)) = i\,k(t_2)\,\omega(\underline{\gamma}(t_1), \underline{\gamma}'(t_2))$$

and

$$S_{11}(t_2, t_3) = \omega(\underline{\gamma}''(t_2), \underline{\gamma}(t_3)) = i\,k(t_2)\,\omega(\underline{\gamma}'(t_2), \underline{\gamma}(t_3))$$

Clearly, if the curvature in t_2 is zero, i.e., $k(t_2) = 0$, we get

$$S_{22}(t_1, t_2) + S_{11}(t_2, t_3) = 0$$

violating Mather's criterion. $\square$

3.4 Existence of caustics

This section is devoted to an existence result of caustics for the symplectic billiard map. The proof follows the steps –inspired by KAM theory– carried out by Lazutkin in [18] in order to prove the existence of caustics for Birkhoff billiards (see Theorem 1.3.5).
Lazutkin's work is based on Moser's twist theorem, a version of KAM theorem for area-preserving twist maps of the annulus (see [25]). Initially, Moser's twist theorem was proven for $\mathcal{C}^k$ maps, with $k \leqslant 333$, but later, this number was reduced to 5 by H. Rüssman, see [26]. For simplicity, we will present here a version of Moser's twist theorem for $\mathcal{C}^\infty$ maps, following [7, pp. 20-21].

Let f_0 be a $\mathcal{C}^\infty$ integrable, area-preserving, twist map of the annulus $\mathbb{S}^1 \times A$, $A \subset \mathbb{R}$, of the form

$$f_0(x,y) = (x + \mu(y), y)$$

where $\mu : A \to \mathbb{R}$ is a diffeomorphism. Observe that the curves $\{y = \text{const}\}$ are invariant for the map f_0.
Consider now a small perturbation of f_0

$$f_\epsilon(x,y) = (x + \mu(y) + \epsilon g(x,y), y + \epsilon h(x,y))$$

and observe that this map is still an area-preserving twist map.
Moser's theorem shows that, for small ϵ and under a suitable condition, if f_ϵ is $\mathcal{C}^\infty$-near f_0, a large proportion of the invariant curves for f_0 survive the perturbation and gets slightly perturbed, becoming invariant curves of the map f_ϵ. The required condition is the following
Diophantine condition: for given constants $\tau > 1$ and $\lambda > 0$

$$|\mu(y) - \frac{p}{q}| \geqslant \frac{\lambda}{q^{\tau+1}}, \qquad \forall \frac{p}{q} \text{ rationals with } q > 0 \tag{3.4}$$

Let us denote the set of all such $\mu(y)$ with $\mathbb{R}_{\tau,\lambda} \subset \mathbb{R}$, noting that this set is closed.
Let $\Lambda := \mu(A) \subset \mathbb{R}$ and define the closed intervals (and therefore compact)

$$\Lambda_\lambda := \{\mu(y) \in \Lambda \mid \operatorname{dist}(\mu(y), \partial\Lambda) \geqslant \lambda\}$$

and

$$\Lambda_{\tau,\lambda} := \Lambda_\lambda \cap \mathbb{R}_{\tau,\lambda}.$$

Cantor-Bendixson theorem states that any closed subset of $\mathbb{R}$ can be written as the disjoint union of a perfect set and a countable set. In particular, $\Lambda_{\tau,\lambda}$ is the union of a Cantor set and a countable set. Moreover, $\Lambda_{\tau,\lambda}$ is nowhere dense, and hence 'topologically' small. However, the Lebesgue measure of $\Lambda_{\tau,\lambda}$ is large for $\lambda \downarrow 0$,

$$\operatorname{meas}(\Lambda \setminus \Lambda_{\tau,\lambda}) \leqslant \text{const}\, \lambda \sum_{q \geqslant 1} q^{-\tau} = \mathcal{O}(\gamma) \quad \text{as } \lambda \downarrow 0.$$

This estimate implies that the union $\bigcup_{\tau,\lambda} \Lambda_{\tau,\lambda}$ is of full measure in Λ.
Finally, notice that the map β pulls back the set $\Lambda_{\tau,\lambda}$ to a subset $A_{\tau,\lambda} \subset A$, which is a Cantor set.

Theorem 3.4.1. *In the notations given above, assume that $\tau > 1$ and $\lambda > 0$ is sufficiently small.*
If f_ϵ is sufficiently close to f_0 in the $\mathcal{C}^\infty$-topology, there exists a $\mathcal{C}^\infty$-diffeomorphism $\psi_\epsilon : \mathbb{S}^1 \times A \to \mathbb{S}^1 \times A$ such that:

(i) ψ_ϵ is $\mathcal{C}^\infty$-near the identity map and depends $\mathcal{C}^\infty$-ly on ϵ;

(ii) the image of the union of f_0-invariant circles $\mathbb{S}^1 \times A_{\tau,\lambda}$ is f_ϵ-invariant under ψ_ϵ, and the restricted map $\widehat{\psi_\epsilon} := \psi_\epsilon|_{\mathbb{S}^1 \times A_{\tau,\lambda}}$ conjugates f_0 to f_ϵ, that is

$$f_\epsilon \circ \widehat{\psi_\epsilon} = \widehat{\psi_\epsilon} \circ f_0.$$

In the proof of Theorem 3.4.2 we will determine an appropriate change of variables $(s, r) \in \mathbb{S}^1 \times \mathbb{R}^+$ under which, near the boundary of the phase space $\{r = 0\}$, the symplectic billiard map can be seen as a small perturbation of the integrable area-preserving twist map

$$(s, r) \mapsto (s + r, r).$$

Thus, Moser's twist theorem 3.4.1 assures the existence of a positive measure set of invariant circles for the symplectic billiard map, which accumulate on $\{r = 0\}$ and on which the motion is conjugated to a rigid rotation with Diophantine rotation number.
As Lazutkin shows in his work, this set of invariant circles translates into the existence of a positive measure set of smooth caustics accumulating to the boundary of the billiard table.
In order to determine the change of variables $(s, r) \in \mathbb{S}^1 \times \mathbb{R}^+$, we introduce the concept of affine parametrization (see [28]).
Let $\gamma : [0, 1] \to \mathbb{R}^2$ be an embedded closed curve, parametrized by $p \in [0, 1]$. Choose a reparametrization of $\gamma(p)$ to a new parametr s such that

$$det\,(\gamma_s, \gamma_{ss}) = 1. \tag{3.5}$$

This relation is invariant under special affine transformations, i.e., affine transformations that preserves area. The parameter s is called affine arc-length, and, denoting

$$g(p) = [(\gamma_p, \gamma_{pp})]^{1/3},$$

the s is explicitly given by

$$s(p) := \int_0^p g(\xi) d\xi.$$

We assume that g is non-vanishing, which is satisfied for strictly convex curves.
Differentiating (3.5), we obtain

$$det\,(\gamma_s, \gamma_{sss}) = 0,$$

that is, the vectors γ_s and γ_{sss} are linearly independent:

$$\gamma_{sss} + \kappa(s)\gamma_s = 0. \tag{3.6}$$

The term $\kappa(s)$ is called the affine curvature of γ and is equal to

$$\kappa(s) = det(\gamma_{ss}, \gamma_{sss}). \tag{3.7}$$

If the initial parameter p is the Euclidean arc-length parameter t, then the relation between the affine arc-length parameter s and t is given by

$$ds = k^{1/3} dt \tag{3.8}$$

where k is the Euclidean curvature.

Theorem 3.4.2. *Suppose that the billiard curve γ is infinitely smooth and has everywhere positive curvature. Then, there exists a positive measure Cantor set of invariant curves for the symplectic billiard map accumulating to the boundary of the phase space.*

Proof. Let $\gamma(s)$ be an affine parametrization of the curve γ. A chord $\gamma(s_1)\gamma(s_2)$ is characterized by s_1 and $r_1 = s_2 - s_1$. Since by the Definition 2.2.1 of the phase space, $s_1 \leqslant s_2 \leqslant s_1^*$, r_1 is positive and bounded above by some function of s. We notice that the Jacobian matrix of the map $(s_1, s_2) \mapsto (s_1, r_1 = s_2 - s_1)$

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

is non degenerate and therefore (s_1, r_1) are global coordinates on $\mathbb{S}^1 \times \mathbb{R}^+$.
The symplectic billiard map in these coordinates becomes

$$(s, r) \mapsto (s + r, \delta(s, r)) \tag{3.9}$$

where $\delta(s, r)$ is a function on the phase space. In the sequel, we prove that

$$\delta(s, r) := r + r^3 f(s, r)$$

with f smooth.
Let $\delta(s, r) = a_0 + a_1 r + a_2 r^2 + a_3 r^3 + \mathcal{O}(r^4)$, where a_i $(i = 1, \ldots, 4)$ are smooth functions of s.
The case $r = 0$ corresponds to $s_1 = s_2$. Thus, $\delta(s, 0) = a_0$ must be zero.
In the sequel we will denote the determinant of two vectors by the brackets $[\cdot, \cdot]$.
Consider three consecutive collision points $\gamma(s - r), \gamma(s), \gamma(s + \delta)$, see Figure 3.2. By definition of symplectic billiard map, the segment $\gamma(s + r) - \gamma(s - r)$ is parallel to the tangent line at γ in $\gamma(s)$, and this condition can be expressed as

$$[\gamma(s + \delta) - \gamma(s - r), \gamma'(s)] = 0. \tag{3.10}$$

Expand $\gamma(s + \delta)$ and $\gamma(s - r)$ in Taylor series up to the 4th derivative and substitute in condition (3.10)

$$\left[(\delta + r)\, \gamma'(s) + \left(\frac{\delta^2 - r^2}{2} \right) \gamma''(s) + \left(\frac{\delta^3 + r^3}{6} \right) \gamma'''(s) + \left(\frac{\delta^4 - r^4}{24} \right) \gamma''''(s), \gamma'(s) \right] = 0. \tag{3.11}$$

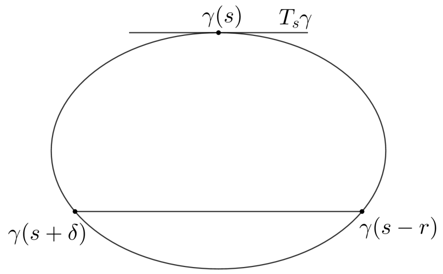

Figure 3.2: The symplectic billiard map in the new variables (s, r)

The first term $(\delta + r)[\gamma'(s), \gamma'(s)]$ clearly satisfies (3.11) since $[\gamma'(s), \gamma'(s)] = 0$.
Consider now the second term

$$\left(\frac{\delta^2 - r^2}{2}\right)[\gamma''(s), \gamma'(s)].$$

Recalling that for an affine parametrization $[\gamma'(s), \gamma''(s)] = 1$, the quadratic term $\delta^2 - r^2$ must be zero. Since $\delta^2 = a_1^2 r^2 + \mathcal{O}(r^3)$, we have that $\delta^2 - r^2 = (a_1^2 - 1)\, r^2 + \mathcal{O}(r^3) = 0$ if and only if $a_1 = 1$.
Also the third term

$$\left(\frac{\delta^3 + \epsilon^3}{6}\right)[\gamma'''(s), \gamma'(s)].$$

satisfies condition (3.11). In fact, since $\gamma'''(s) = -\kappa(s)\gamma'(s)$ (see formula (3.6)), it follows

$$[\gamma'''(s), \gamma'(s)] = -\kappa(s)[\gamma'(s), \gamma'(s)] = 0.$$

Finally, we consider the fourth term $\left(\frac{\delta^4 - r^4}{24}\right)[\gamma''''(s), \gamma'(s)]$. We have

$$\gamma''''(s) = \frac{d}{ds}\gamma'''(s) = \frac{d}{ds}(-\kappa(s)\gamma'(s)) = -\kappa'(s)\gamma'(s) - \kappa(s)\gamma''(s)$$

and therefore:

$$[\gamma''''(s), \gamma'(s)] = -\kappa'(s)[\gamma'(s), \gamma'(s)] - \kappa(s)[\gamma''(s), \gamma'(s)] = \kappa \neq 0.$$

Then, $\delta^4 - r^4$ is necessarily zero. From $\delta^4 = r^4 + 4a_2\, r^5 + \mathcal{O}(r^6)$, it follows that $\delta^4 - r^4 = 4\, a_2\, r^5 + \mathcal{O}(r^6) = 0$ if and only if $a_2 = 0$.
In conclusion, $a_0 = 0 = a_2$ and $a_1 = 1$, so that $\delta(s, r) = r + r^3 f(s, r)$.
As a consequence, the symplectic billiard map (3.9) is given by

$$(s, r) \mapsto (s + r, r + r^3 f(s, r))$$

which is a small perturbation of the integrable area-preserving twist map $(s, r) \mapsto (s + r, r)$. $\square$

Chapter 4

Symplectic billiards in a symplectic space

Abstract. This chapter is devoted to the generalization of symplectic billiards to the linear symplectic vector space $(\mathbb{R}^{2n}, \omega)$. We give the definition and prove that the symplectic billiard map admits a generating function, which –as in the planar case– involves the symplectic form ω. We prove the complete integrability of symplectic billiards in an ellipsoid and describe the symplectic billiard dynamics in the special case of the unit sphere.

4.1 Linear symplectic spaces

Let V be a real finite dimensional vector space. A symplectic form on V is a map $\omega : V \times V \to \mathbb{R}$ with the following properties:

(i) bilinearity

$$\omega(\alpha u + \beta v, w) = \alpha\, \omega(u, v) + \beta\, \omega(x; v, w),$$

$$\omega(u, \alpha v + \beta w) = \alpha\, \omega(u, v) + \beta\, \omega(v, w)$$

for any $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$;

(ii) skew-symmetry

$$\omega(v, w) = -\omega(w, v)$$

for any $v, w \in V$;

(iii) non-degenearcy

$$\omega(v, w) = 0 \quad \forall\, w \in V \;\Rightarrow\; v = 0.$$

Then, we call (V, ω) a symplectic vector space.
The following are immediate properties of the symplectic structure ω (see [8, Section 1]):

(i) the vector space V is even dimensional, $dimV = 2n$;

(ii) there always exists a basis $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ of V such that

$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0, \quad \omega(e_i, f_j) = \delta_{ij} \qquad i, j = 1, \dots, n$$

Such a basis is called symplectic basis and the symplectic form is represented by the block matrix

$$\omega = \begin{pmatrix} 0 & id_n \\ -id_n & 0 \end{pmatrix}.$$

The prototype of a symplectic vector spaces is $(\mathbb{R}^{2n}, \omega_0)$ with ω_0 such that the canonical basis of $\mathbb{R}^{2n}$

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, \underbrace{1}_{n}, 0, \dots, 0)$$
$$f_1 = (0, \dots, 0, \underbrace{1}_{n+1}, 0, \dots, 0), \dots, f_n = (0, \dots, 0, 1)$$

is a symplectic basis.

Any symplectic vector space can be endowed with a complex structure compatible with the symplectic form ω.

Definition 4.1.1. Let V be a real vector space. A complex structure on V is an automorphism

$$J : V \to V$$

such that $J^2 = -id$.

By identifying the map J with multiplication by i, we endow V with the structure of a complex vector space. Complex scalar multiplication can be defined by

$$(a + ib)v = av + bJv,$$

where $a, b \in \mathbb{R}$ and $v \in V$.

Definition 4.1.2. Let (V, ω) be a symplectic vector space. A complex structure J on V is said to be ω-compatible if

$$g(v, w) := \omega(v, Jw) \tag{4.1}$$

defines an inner product on V for any $v, w \in V$. That is, J is ω-compatible if and only if

$$\omega(Jv, Jw) = \omega(v, w)$$
$$\omega(v, Jv) > 0 \quad \forall\, v \neq 0.$$

Consider now the standard symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$ with the canonical symplectic basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$. Let J_0 be the complex structure on $\mathbb{R}^{2n}$ defined by $J_0 e_i = f_i$ and $J_0 f_i = -e_i$. The block matrix representing J_0 is

$$J_0 = \begin{pmatrix} 0 & -id_n \\ id_n & 0 \end{pmatrix}.$$

This complex structure is compatible with the standard inner product $\langle \cdot, \cdot \rangle$ of $\mathbb{R}^{2n}$

$$\langle v, w \rangle = v^t w$$

where v^t denotes the transpose of v.
Indeed, an easy computation shows that

$$\begin{aligned} \omega_0(v, J_0\, w) &= v^t \begin{pmatrix} 0 & id_n \\ -id_n & 0 \end{pmatrix} J_0\, w \\ &= v^t \begin{pmatrix} 0 & id_n \\ -id_n & 0 \end{pmatrix} \begin{pmatrix} 0 & -id_n \\ id_n & 0 \end{pmatrix} w \\ &= v^t w = \langle v, w \rangle. \end{aligned}$$

In particular, for any $v, w \in \mathbb{R}^{2n}$, it holds the following relation

$$\omega_0(v, w) = \langle J_0\, v, w \rangle. \tag{4.2}$$

4.2 Symplectic billiard in a symplectic space: definition

Let $(\mathbb{R}^{2n}, \omega)$ be a linear symplectic space and consider a smooth closed hypersurface $M \subset \mathbb{R}^{2n}$ bounding a strictly convex domain.
The definition of a symplectic billiard map –similar to the one of the planar case– in the higher dimensional setting presents an obviuos difficulty, that is choosing a tangent direction at every point of M. This is canonically provided by the symplectic structure. Indeed, since M is odd-dimensional, the restriction of the symplectic form ω to M is not non degenerate anymore, but it has a one-dimesional kernel on the tangent hyperplane T_xM for every $x \in M$. This kernel is called the characteristic direction of M at x and we denote it by

$$R(x) := ker\omega|_{T_xM \times T_xM}.$$

We choose this characteristic direction as the tangent line at $x \in M$.
Observe that the strict convexity of M gives rise to an involution $M \ni x \mapsto x^* \in M$ such that $R(x) = R(x^*)$.
As in the planar case, let ν_x denote the outer normal in $x \in M$. Orientate the hypersurface M with respect to the outer normal.

Lemma 4.2.1. *Given two points $x, y \in M$, the relation $R(y) \subset T_xM$ is equivalent to $\omega(\nu_x, \nu_y) = 0$. Moreover, this relation is symmetric in x and y, that is*

$$R(y) \subset T_xM \Leftrightarrow R(x) \subset T_yM.$$

Proof. Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ through the standard complex structure J. Observe that applying J on a vector equals to rotating it by an angle of $\pi/2$. Thus, the characteristic direction at $y \in M$ is obtained by applying J to the outer normal ν_y, that is $R(y) = \mathbb{R} \cdot (J\nu_y)$.
Therefore,

$$R(y) \subset T_xM$$

is equivalent to say that the vectors $J\nu_y$ and ν_x are orthogonal and thus, their inner product is null

$$\langle J\nu_y, \nu_x \rangle = 0.$$

But by relation (4.2), it follows that $\omega(\nu_y, \nu_x) = \langle J\nu_y, \nu_x \rangle = 0$, which is the desired result. The symmetry of the relation follows from the skew-symmetry of ω. □

We introduce now the phase space of the symplectic billiard map.

Definition 4.2.2. The phase space of the symplectic billiard map is

$$\mathcal{P} := \{(x_1, x_2) \in M \times M \mid \omega(\nu_{x_1}, \nu_{x_2}) > 0\}.$$

Lemma 4.2.3. *The phase space $\mathcal{P}$ is connected.*

Proof. Consider the Gauss map

$$\mathcal{G} : M \to \mathbb{S}^{2n-1} \qquad x \mapsto \nu_x.$$

that maps the hypersurface $M \subset \mathbb{R}^{2n}$ to the unit sphere $\mathbb{S}^{2n-1} \subset \mathbb{R}^{2n}$, associating any point of M to its outer normal vector at M. Observe that, because of the strict convexity of M, the Gauss map is an homeomorphism. Therefore, the phase space $\mathcal{P}$ is mapped to

$$\mathcal{P}_0 := \{(a, b) \in S^{2n-1} \times S^{2n-1} \mid \omega(a, b) > 0\}.$$

Clearly, $b \neq \pm a$, or otherwise $\omega(a, \pm a) = 0$. Let

$$a_t := \frac{a + t(b-a)}{\|a + t(b-a)\|} \in S^{2n-1}, \qquad t \in (0, 1].$$

Using the properties of the symplectic form ω, it follows that

$$\begin{aligned} \omega(a, a_t) &= \omega\left(a, \frac{a + t(b-a)}{\|a + t(b-a)\|}\right) \\ &= \frac{\omega(a, a) + t\omega(a, b) - t\omega(a, a)}{\|a + t(b-a)\|} \\ &= \frac{t\omega(a, b)}{\|a + t(b-a)\|}, \end{aligned}$$

showing that we can move the second factor close to a. That is, we can move any point in $\mathcal{P}_0$ into a tubular neighborhood of the diagonal, which is therefore connected.
Since the Gauss map is an homeomorphism, its inverse maps connected sets into connected sets. Thus, $\mathcal{P} = \mathcal{G}^{-1}(\mathcal{P}_0)$ is connected. $\square$

Remark 4.2.4. It is possible to define the negative part of the phase space

$$\mathcal{P}^- := \{(x, y) \in M \times M \mid \omega(\nu_x, \nu_y) < 0\}.$$

Arguing as in Lemma 4.2.3, one can show that $\mathcal{P}^-$ is also connected.

In order to define the symplectic billiard map, we need first the following lemma

Lemma 4.2.5. *Let $(x, y) \in \mathcal{P}$. Then there exists a unique point $z \in M$ such that $(x + R(y)) \cap M = \{x, z\}$. Moreover, $(y, z) \in \mathcal{P}$ and $z - x = tJ\nu_y$, with $t > 0$.*

Proof. The first statement is a direct consequence of the convexity of M and the definition of the phase space $\mathcal{P}$. Indeed, the convexity of M implies that $(x + R(y)) \cap M = \{x, z\}$ with possibly $x = z$. By definition of the phase space, if $(x, y) \in \mathcal{P}$, $R(y) \not\subset T_xM$, and therefore $x \neq z$. Observe that equation $(x + R(y)) \cap M = \{x, z\}$ is equivalent to $z - x \in R(y)$.

Next, we have to show that $(y, z) \in \mathcal{P}$. We easily remark that if y is close to x then so is z and therefore $\omega(\nu_x, \nu_y) > 0$ implies $\omega(\nu_y, \nu_z) > 0$. For $(x, y) \in \mathcal{P}$ suppose –by contradiction– that $(y, z) \notin \mathcal{P}$, i.e., $\omega(\nu_y, \nu_z) \leqslant 0$. Using the connectedness of $\{(x, y) \in M \times M \mid \omega(\nu_x, \nu_y) < 0\}$, and moving y close to x, we can arrange $\omega(\nu_y, \nu_z) = 0$. This means that $R(y) \subset T_zM$ which contradicts $x \neq z$.

Finally we have to show that $z - x = tJ\nu_y$ with $t > 0$.
Consider the sets $M_x^+ := \{y \in M \mid \omega(\nu_x, \nu_y) > 0\}$ and $M_x^- := \{y \in M \mid \omega(\nu_x, \nu_y) < 0\}$. The Gauss map maps these two sets to hemispheres, and thus they are both connected. The definition of the phase space can be rephrase as

$$(x, y) \in \mathcal{P} \Leftrightarrow x \in M_y^- \Leftrightarrow y \in M_x^+,$$

and the request that $(y, z) \in \mathcal{P}$ is equivalent to $z \in M_y^+$.
By the observation that $(x + R(y)) \cap M = \{x, z\}$ is equivalent to $z - x \in R(y)$, and recallig from Lemma 4.2.3 that the characteristic direction at y is $R(y) = \mathbb{R} \cdot J\nu_y$, it follows that

$$z - x = t(y)J\nu_y,$$

where we think of x as fixed and y as a variable.
First, we observe that the sign of $t(y)$ does not depend on y as long as $(x, y) \in \mathcal{P}$, or equivalently, $y \in M_x^+$. Indeed, suppose that $t(y)$ changes sign or vanishes. Since M_x^+ is connected, we can always find a point y such that $t(y) = 0$. But this is equivalent to $z = x$, which contradicts the first assertion of the proof.
Hence, the sign of $t(y)$ is fixed, and we compute it at an appropriate point. Choose y such that $\nu_y = aJ\nu_x$, with $a > 0$. The existence and the uniqueness of this point is due

to the Gauss map, which is a diffeomorphism. Recalling the relation (4.2) between the inner product and the symplectic form and that $\|\nu_x\| = 1$ since $\nu_x \in \mathbb{S}^{2n-1}$, it follows that

$$\omega(\nu_x, \nu_y) = \omega(\nu_x, aJ\nu_x) = \langle J\nu_x, aJ\nu_x\rangle = \langle \nu_x, a\nu_x\rangle = a\|\nu_x\|^2 = a > 0.$$

Finally, we observe that

$$z - x = t(y)J\nu_y = at(y)J^2\nu_x = -at(y)\nu_x.$$

Therefore, since ν_x is the outer normal of M at x, the convexity of M implies $t(y) > 0$. This concludes the proof. $\square$

Definition 4.2.6. The symplectic billiard map is

$$\phi : \mathcal{P} \to \mathcal{P}, \qquad (x, y) \mapsto (y, z)$$

where $z \in M$ is the unique point such that $z - x \in R(y)$.

Remark 4.2.7. The symplectic billiard map $\phi : \mathcal{P} \to \mathcal{P}$ can be extend continuously to the closure

$$\bar{\mathcal{P}} := \{(x, y) \in M \times M \mid \omega(\nu_x, \nu_y) \geq 0\}$$

by

$$\phi(x, x) := (x, x)$$

and

$$\omega(\nu_x, \nu_y) = 0 \Rightarrow \phi(x, y) := (y, x).$$

As in the planar case, the symplectic billiard map ϕ admits a generating function, that involves the symplectic form ω.

Lemma 4.2.8. *The generating function for the symplectic billiard map ϕ is*

$$S : \mathcal{P} \to \mathbb{R} \qquad (x, y) \mapsto S(x, y) := \omega(x, y).$$

This means that

$$\phi(x, y) = (y, z) \Leftrightarrow \frac{\partial}{\partial y}[S(x, y) + S(y, z)] = 0. \tag{4.3}$$

Proof. Let $x = (x_\alpha, x_\beta) \in \mathbb{R}^{2n}$, where $x_\alpha, x_\beta \in \mathbb{R}^n$.
Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and observe that

$$Jx = \begin{pmatrix} 0 & -id_n \\ id_n & 0 \end{pmatrix} \begin{pmatrix} x_\alpha \\ x_\beta \end{pmatrix} = \begin{pmatrix} -x_\beta \\ x_\alpha \end{pmatrix},$$

and therefore

$$\omega(x, y) = \langle Jx, y\rangle = (Jx)^t y = (-x_\beta, x_\alpha)^t (y_\alpha, y_\beta) = x_\alpha y_\beta - x_\beta y_\alpha.$$

Hence, we have that

$$S(x, y) = \omega(x, y) = x_\alpha y_\beta - x_\beta y_\alpha$$

and

$$S(y, z) = \omega(y, z) = y_\alpha z_\beta - y_\beta z_\alpha.$$

Consequently,

$$\frac{\partial}{\partial y} S(x, y) = \frac{\partial}{\partial y}(x_\alpha y_\beta - x_\beta y_\alpha) = x_\alpha dy_\beta - x_\beta dy_\alpha$$
$$\frac{\partial}{\partial y} S(y, z) = \frac{\partial}{\partial y}(y_\alpha z_\beta - y_\beta z_\alpha) = -z_\alpha dy_\beta + z_\beta dy_\alpha$$

and

$$\begin{aligned}\frac{\partial}{\partial y}[S(x, y) + S(y, z)] &= (x_\alpha dy_\beta - x_\beta dy_\alpha) + (-z_\alpha dy_\beta + z_\beta dy_\alpha)\\ &= (z_\beta - x_\beta)dy_\alpha + (x_\alpha - z_\alpha)dy_\beta.\end{aligned}$$

Hence, if $x \neq z$, we have that

$$\begin{aligned}\frac{\partial}{\partial y}[S(x, \cdot) + S(\cdot, z)] = 0 &\Leftrightarrow (z_\beta - x_\beta, x_\alpha - z_\alpha) \perp T_y M\\ &\Leftrightarrow z - x \in R(y)\end{aligned}$$

as needed. □

We conclude this section by showing that the symplectic billiard map ϕ admits an invariant closed 2-form

$$\Omega_{(x,y)} := \frac{\partial^2 S}{\partial x \partial y}(x, y) dx \wedge dy \in \Omega^2(M \times M).$$

Lemma 4.2.9. *The symplectic billiard map preserves the closed* 2*-form* Ω*.*

Proof. By equality 4.3,

$$\phi(x, y) = (y, z) \;\Leftrightarrow\; \frac{\partial}{\partial y}[S(x, y) + S(y, z)] = 0.$$

Consequently, to prove the ϕ-invariance of Ω, we need to show that

$$\frac{\partial^2 S}{\partial x \partial y}(x, y) dx \wedge dy = \frac{\partial^2 S}{\partial y \partial z}(y, z) dy \wedge dz.$$

We take the exterior derivative of $\frac{\partial}{\partial y}[S(x, y) + S(y, z)] = 0$

$$\begin{aligned}&d\left[\frac{\partial}{\partial y}(S(x, y) + S(y, z))\right]\\ &= \frac{\partial^2}{\partial x \partial y} S(x, y) dx + \frac{\partial^2}{\partial y^2} S(x, y) dy + \frac{\partial^2}{\partial y^2} S(y, z) dy + \frac{\partial^2}{\partial z \partial y} S(y, z) dz\\ &= 0\end{aligned}$$

and then, recalling that the wedge product is antisymmetric and $dy \wedge dy = 0$, we right-wedge-multiply by dy

$$\begin{aligned}\frac{\partial^2}{\partial x \partial y} S(x,y) dx \wedge dy &+ \frac{\partial^2}{\partial z \partial y} S(y,z) dz \wedge dy \\ &= \frac{\partial^2}{\partial x \partial y} S(x,y) dx \wedge dy - \frac{\partial^2}{\partial z \partial y} S(y,z) dy \wedge dz \\ &= 0. \end{aligned}$$

□

4.3 Symplectic billiard in an ellipsoid

This section is devoted to prove the complete integrability of symplectic billiards in an ellipsoid, using the integrability of Birkhoff billiards in an ellipsoid. The idea of the proof is, given a trajectory of the symplectic billiard in an ellipsoid, to construct a polygonal line by connecting every second consecutive impact points (i.e., always skip one impact point), and then apply a linear transformation –that depends only on the original ellipsoid– that transforms the polygonal line into a Birkhoff billiard trajectory on another ellipsoid, that again depends only on the original ellipsoid.

Remark 4.3.1. The complete integrability of the Birkhoff billiard map in an ellipsoid is a consequence of the complete integrability of the geodesic flow in an ellipsoid.
An ellipsoid $E \subset \mathbb{R}^{n+1}$, given by the equation

$$\sum_{i=1}^{n+1} \frac{x_i^2}{a_i^2} = 1$$

can be considered as a degenerate ellipsoid, the limit of the ellipsoids

$$\sum_{i=1}^{n+2} \frac{x_i^2}{a_i^2} = 1$$

as $a_{n+2} \to 0$. Thus the billiard map is obtained from the geodesic flow, and the complete integrability follows (see [31, Section 2.3] and [32] for a detailed explanation).

Consider $\mathbb{R}^{2n}$ with linear coordinates $\{x_{\alpha,1}, \dots, x_{\alpha,n}, x_{\beta,1}, \dots, x_{\beta,n}\}$ and the symplectic structure $\omega_0 = \sum dx_{\alpha,j} \wedge dx_{\beta,j}$.
Let $M \in \mathbb{R}^{2n}$ be an ellipsoid. Applying a linear symplectic transformation and homothety, we may assume that the ellipsoid is given by the equation

$$\frac{x_{\alpha,1}^2 + x_{\beta,1}^2}{a_1} + \frac{x_{\alpha,2}^2 + x_{\beta,2}^2}{a_2} + \dots + \frac{x_{\alpha,n}^2 + x_{\beta,n}^2}{a_n} = 1 \qquad a_1, \dots, a_n > 0. \tag{4.4}$$

By applying the diagonal linear transformation

$$x_{\alpha,j} \mapsto \sqrt{a_j} x_{\alpha,j}, \quad x_{\beta,j} \mapsto \sqrt{a_j} x_{\beta,j}, \qquad j = 1, \dots, n,$$

the ellipsoid M is transformed into the unit sphere $\mathbb{S}^{2n-1}$ of equation

$$x_{\alpha,1}^2 + x_{\beta,1}^2 + x_{\alpha,2}^2 + x_{\beta,2}^2 + \cdots + x_{\alpha,n}^2 + x_{\beta,n}^2 = 1,$$

while the symplectic form ω_0 is mapped into the symplectic form

$$\omega = \sum a_j dx_{\alpha,j} \wedge dx_{\beta,j}.$$

Thus, we consider the symplectic billiard inside $\mathbb{S}^{2n-1}$ defined by the symplectic form ω. Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and let $z_j = x_{\alpha,j} + ix_{\beta,j}$, $j = 1, \ldots, n$, be linear coordinates on $\mathbb{C}^n$. Let $z = (z_1, \ldots, z_n) = (x_{\alpha,1} + ix_{\beta,1}, \ldots, x_{\alpha,n} + ix_{\beta,n})$ be a point of $\mathbb{S}^{2n-1}$, and let $R : \mathbb{C}^n \to \mathbb{C}^n$ be a complex linear operator, represented by the diagonal matrix

$$R = \begin{pmatrix} i/a_1 & & & \\ & i/a_2 & & \\ & & \ddots & \\ & & & i/a_n \end{pmatrix},$$

such that the characteristic direction at $\mathbb{S}^{2n-1}$ in z is $\mathbb{R} \cdot R(z)$.
Consider now the complex linear operator $R^{-1} : \mathbb{C}^n \to \mathbb{C}^n$, represented by the diagonal matrix

$$R = \begin{pmatrix} -ia_1 & & & \\ & -ia_2 & & \\ & & \ddots & \\ & & & -ia_n \end{pmatrix},$$

and observe that

$$w := R^{-1}(z) = (-ia_1 z_1, \ldots, -ia_n z_n) = (a_1(x_{\beta,1} - ix_{\alpha,1}), \ldots, a_n(x_{\beta,n} - ix_{\alpha,n})).$$

Let $w_j = a_j(x_{\beta,j} - ix_{\alpha,j})$, $j = 1, \ldots, n$, be the coordinates of $w \in \mathbb{C}^n$. It follows that

$$\|w_j\|^2 = \|a_j(x_{\beta,j} - ix_{\alpha,j})\|^2 = a_j^2(x_{\alpha,j}^2 + x_{\beta,j}^2).$$

and thus

$$x_{\alpha,j}^2 + x_{\beta,j}^2 = \frac{\|w_j\|^2}{a_j^2}.$$

Consequently, the linear map R^{-1} transforms the unit sphere $\mathbb{S}^{2n-1}$ into the ellipsoid E given by the equation

$$\frac{\|w_1\|^2}{a_1^2} + \frac{\|w_2\|^2}{a_2^2} + \cdots + \frac{\|w_n\|^2}{a_n^2} = 1. \tag{4.5}$$

Theorem 4.3.2. *Let $(\ldots, Z_0, Z_1, Z_2, \ldots)$ be a trajectory of the symplectic billiard map in the unit sphere $\mathbb{S}^{2n-1}$ with respect to the symplectic form $\omega = \sum a_j dx_{\alpha,j} \wedge dx_{\beta,j}$. Then the sequence $(\ldots, R^{-1}(Z_0), R^{-1}(Z_2), R^{-1}(Z_4), \ldots)$ is a usual billiard trajectory in E. Conversely, a usual billiard trajectory $(\ldots, W_0, W_1, W_2, \ldots)$ in E corresponds to a unique symplectic billiard trajectory $(\ldots, Z_0, Z_1, Z_2, \ldots)$ in $\mathbb{S}^{2n-1}$, with $Z_j = R(W_j)$.*

Proof. Consider the points $Z_0, Z_2, Z_4 \in \mathbb{S}^{2n-1}$ of a symplectic billiard trajectory. Recalling from Lemma 4.2.5 that $Z_2 - Z_0 = tR(Z_1)$, with $t > 0$, it follows that Z_1 is uniquely determined

$$R(Z_1) = t_1(Z_2 - Z_0), \quad t_1 > 0. \tag{4.6}$$

Appling the linear operator R^{-1} to both members of equation (4.6), we get

$$Z_1 = t_1 R^{-1}(Z_2 - Z_0).$$

Normalizing the equation (observe that $\|Z\| = 1$ if $Z \in \mathbb{S}^1$), we can uniquely determined $t_1 > 0$

$$1 = \|Z_1\|^2 = t_1^2 \|R^{-1}(Z_2 - Z_0)\|^2 \Rightarrow t_1 = \frac{1}{\|R^{-1}(Z_2 - Z_0)\|},$$

Hence

$$Z_1 = \frac{R^{-1}(Z_2 - Z_0)}{\|R^{-1}(Z_2 - Z_0)\|}. \tag{4.7}$$

In a similar way, we obtain

$$Z_3 = \frac{R^{-1}(Z_4 - Z_2)}{\|R^{-1}(Z_4 - Z_2)\|}. \tag{4.8}$$

The symplectic billiard reflection law implies that

$$R(Z_2) = t(Z_3 - Z_1) = t\left(\frac{R^{-1}(Z_4 - Z_2)}{\|R^{-1}(Z_4 - Z_2)\|} + \frac{R^{-1}(Z_0 - Z_2)}{\|R^{-1}(Z_0 - Z_2)\|}\right).$$

Let $W_j := R^{-1}(Z_j)$, $j = 0, 2, 4$, and rewrite the last equation as

$$R^2(W_2) = t\left(\frac{W_4 - W_2}{\| W_4 - W_2 \|} + \frac{W_0 - W_2}{\| W_0 - W_2 \|}\right). \tag{4.9}$$

The vector $R^2(W_2)$ is normal to the ellipsoid E, and therefore equation (4.9) describes the billiard reflection in E at point W_2 that takes W_0W_2 to W_2W_4.

Conversely, given a segment of the billiard trajectory W_0, W_2, W_4 in E, we define $Z_j = R(W_j)$, $j = 0, 2, 4$. By equations (4.7) and (4.8), it follows that

$$Z_1 = \frac{R^{-1}(Z_2 - Z_0)}{\|R^{-1}(Z_2 - Z_0)\|} = \frac{R^{-1}(Z_2) - R^{-1}(Z_0)}{\|R^{-1}(Z_2) - R^{-1}(Z_0)\|} = \frac{W_2 - W_0}{\|W_2 - W_0\|}$$

and similarly

$$Z_3 = \frac{W_4 - W_2}{\|W_4 - W_2\|}.$$

Finally, equation (4.9) becomes

$$R(Z_2) = t(Z_3 - Z_1),$$

which describes the symplectic billiard map in $\mathbb{S}^{2n-1}$. Hence, $Z_0, \ldots, Z_4$ is a segment of a symplectic billiard trajectory. □

4.4 Symplectic billiard in a sphere

In this section we describe in details the symplectic billiard dynamics in the unit sphere.

Let M be the unit sphere $\mathbb{S}^{2n-1}$ of center O. The symplectic billiard map ϕ can be explicitely determined.

Proposition 4.4.1. *Let $(z_0, z_1) \in \mathcal{P}$. The symplectic billiard map ϕ of the sphere is given by the formula*

$$z_2 = z_0 + 2\omega(z_0, z_1)Jz_1, \tag{4.10}$$

where J is multiplication by i.

Proof. The normal direction at the sphere in $z \in \mathbb{S}^{2n-1}$ is the radial direcion, that is, the direction of the line joining the center of the sphere O with z. Recall that applying J on z equals to rotating the vector z of an angle of $\pi/2$. It follows then, that the characteristic direction at $\mathbb{S}^{2n-1}$ in z is the direction of Jz.
Therefore, from Lemma 4.2.5, it holds

$$z_2 = z_0 + tJz_1 \tag{4.11}$$

with $t > 0$. We have to show that $t = 2\omega(z_0, z_1)$.
Normalize equation (4.11), recalling that for any point $z \in \mathbb{S}^{2n-1}$ it holds $\langle z, z\rangle = 1$.

$$\begin{aligned}
1 &= \langle z_2, z_2\rangle \\
&= \langle z_0 + tJz_1, z_0 + tJz_1\rangle \\
&= \langle z_0, z_0\rangle + t\langle z_0, Jz_1\rangle + t\langle Jz_1, z_0\rangle + t^2\langle Jz_1, Jz_1\rangle \\
&= 1 + t\langle z_0, Jz_1\rangle + t\langle z_0, Jz_1\rangle + t^2\langle z_1, z_1\rangle \\
&= 1 + 2t\langle z_0, Jz_1\rangle + t^2.
\end{aligned}$$

Therefore, we have a second order equation

$$t^2 + 2t\langle z_0, Jz_1\rangle = 0$$

with solutions $t = 0$ –which is not ammissible since we are looking for $t > 0$– and

$$t = -2\langle z_0, Jz_1\rangle = 2\langle Jz_0, z_1\rangle = 2\omega(z_0, z_1),$$

which is the desired result. □

The quantity $\omega(z_0, z_1)$ is an integral for the symplectic billiard map, that is $\omega(z_0, z_1) = \omega(z_1, z_2)$. Indeed, if we substitute the value of z_2 given by formula (4.10) in $\omega(z_0, z_1)$, we obtain

$$\begin{aligned}
\omega(z_1, z_2) &= \omega(z_1, z_0 + 2\omega(z_0, z_1)Jz_1) \\
&= \omega(z_1, z_0) + 2\omega(z_0, z_1)\omega(z_1, Jz_1) \\
&= -\omega(z_0, z_1) + 2\omega(z_0, z_1)\langle Jz_1, Jz_1\rangle \\
&= -\omega(z_0, z_1) + 2\omega(z_0, z_1) \\
&= \omega(z_0, z_1).
\end{aligned}$$

We will denote this integral symply by ω. Clearly, because of the definition of the phase space $\mathcal{P}$, $\omega \geqslant 0$. Moreover, by the Cauchy-Schwartz inequality

$$\omega(z_0, z_1) = \langle Jz_0, z_1\rangle \leqslant \|Jz_0\|\|z_1\| = 1.$$

Thus, $0 \leqslant \omega \leqslant 1$ and we set $\omega := \sin\alpha$ with $\alpha \in [0, \pi/2]$.
The dynamics of the symplectic billiard in the sphere in the cases $\omega = 0$ and $\omega = 1$ is special.
The first case is characterized by the 2-periodicity, since for $\omega = 0$ we have that $\phi(z_0, z_1) = (z_1, z_0)$.
In the second case, the ϕ-orbit is 4-periodic. Indeed, if $\omega(z_0, z_1) = \langle Jz_0, z_1\rangle = 1$, it follows that $z_1 = Jz_0$. Hence, we obtain the sequence of points

$$z_0 \mapsto z_1 = Jz_0 \mapsto z_2 = -z_0 \mapsto z_3 = -Jz_0 \mapsto z_4 = z_0 \mapsto \dots$$

In the genaral case, the dynamics of the symplectic billiard in the sphere is described in Proposition 4.4.3. The main idea of the proof is to regard equation 4.10 as a second order linear recurrence with constant coefficients. Then the sequence of impact points $z_0, z_1, z_2, \dots$ can be determined by the next

Theorem 4.4.2. *Suppose a sequence $z_0, z_1, z_2, \dots$ satisfies a second order linear recurrence relation*

$$z_k = Az_{k-1} + Bz_{k-2}$$

for all $k \in \mathbb{Z}$, $k > 2$ and with $B \neq 0$. If its associated characteristic equation

$$\lambda^2 - A\lambda - B = 0$$

has distint roots λ_1, λ_2, then the sequence $z_0, z_1, z_2, \dots$ is given by the explicit formula

$$z_n = C\lambda_1^n + D\lambda_2^n \tag{4.12}$$

where C and D are determined by the initial values z_0, z_1 by

$$\begin{cases} z_0 = C + D \\ z_1 = C\lambda_1 + D\lambda_2. \end{cases}$$

Proposition 4.4.3. *Let $\omega = \sin\alpha < 1$, $\alpha \in [0, \pi/2)$, and define $\lambda_1 := e^{i\alpha}$, $\lambda_2 := -e^{-i\alpha}$. One has*

$$z_n = \frac{\lambda_1^{n-1} - \lambda_2^{n-1}}{\lambda_1 - \lambda_2} z_0 + \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} z_1. \tag{4.13}$$

The ϕ-orbit of a point lies on the union of two circles. The orbit is periodic if α is π-rational and dense on the two circles otherwise. If $\alpha = 2\pi(p/q)$, where p/q is in the lowest terms, then the period equals q for even q, and $2q$ for odd q.

Proof. Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and rewrite equation (4.10) as

$$z_2 = z_0 + 2i\omega z_1. \tag{4.14}$$

This equation is a second order linear recurrence with constant coefficients $A = 2i\omega$ and $B = 1$, and it generates the sequence $z_0, z_1, z_2, \dots$.
The characterist equation associated to (4.14) is

$$\lambda^2 - 2i\omega\lambda - 1 = 0, \tag{4.15}$$

and its roots are

$$\lambda_{1,2} := i\omega \pm \sqrt{1-\omega^2}.$$

Since $\omega = \sin\alpha$, we have that

$$\lambda_1 = i\omega + \sqrt{1-\omega^2} = i\sin\alpha + \sqrt{1-\sin^2\alpha} = i\sin\alpha + \cos\alpha = e^{i\alpha}$$

and

$$\lambda_2 = i\omega - \sqrt{1-\omega^2} = i\sin\alpha\sqrt{1-\sin^2\alpha} = i\sin\alpha - \cos\alpha = -e^{-i\alpha}.$$

Finally, we determine constants C, D by solving the system

$$\begin{cases} z_0 = C + D \\ z_1 = C\lambda_1 + D\lambda_2. \end{cases}$$

By the first equation, we have $C = z_0 - D$, and substituting into the second we obtain

$$z_1 = (z_0 - D)\lambda_1 + D\lambda_2 = z_0\lambda_1 + D(\lambda_2 - \lambda_1).$$

Hence,

$$D = \frac{\lambda_1 z_0 - z_1}{\lambda_1 - \lambda_2} \qquad C = \frac{z_1 - \lambda_2 z_0}{\lambda_1 - \lambda_2}.$$

Consequently, the sequence $z_0, z_1, z_2, \dots$ is given by

$$\begin{aligned} z_n &= C\lambda_1^n + D\lambda_2^n \\ &= \frac{z_1 - \lambda_2 z_0}{\lambda_1 - \lambda_2}\lambda_1^n + \frac{\lambda_1 z_0 - z_1}{\lambda_1 - \lambda_2}\lambda_2^n \\ &= \frac{\lambda_1^{n-1} - \lambda_2^{n-1}}{\lambda_1 - \lambda_2} z_0 + \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} z_1, \end{aligned}$$

which is the desired result.

The symplectic billiard map $\phi : (z_0, z_1) \mapsto (z_1, z_2)$ is a complex self-linear map of $\mathbb{C}^{2n}$, and it is represented by the block matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 2i\omega 1 \end{pmatrix},$$

where each entry is an $n \times n$ block. The eigenvalues of ϕ are the roots λ_1, λ_2 of the characteristic equation (4.15), each with multiplicity n.
Let (u, v) be a vector of $\mathbb{C}^{2n}$, with $u, v \in \mathbb{C}^n$. The vector (u, v) is an eignevector of ϕ associated to the eigenvalue λ_i, $i = 1, 2$, if it holds

$$\phi \begin{pmatrix} u \\ v \end{pmatrix} = \lambda_i \begin{pmatrix} u \\ v \end{pmatrix} \Leftrightarrow \begin{pmatrix} v \\ u + 2i\omega v \end{pmatrix} = \begin{pmatrix} \lambda_i u \\ \lambda_i v \end{pmatrix}.$$

Thus, the eigenvectors associated to λ_i are given by the system

$$\begin{cases} v = \lambda_i u \\ u + 2i\omega v = \lambda_i v \end{cases} \Rightarrow \begin{cases} v = \lambda\, u \\ (\lambda_i^2 - 2i\omega\lambda_i - 1)u = 0. \end{cases}$$

Clearly, $\lambda_i^2 - 2i\omega\lambda_i - 1 = 0$, and therefore the second equation of the system equals 0 for any $u \in \mathbb{C}^n$.
Consequently, the eigenvectors associated to the eigenvalue λ_i are all the vectors

$$\begin{pmatrix} \xi \\ \lambda_i \xi \end{pmatrix}, \qquad \xi \in \mathbb{C}^n.$$

Finally, the eigenspace associated to the eigenvalue λ_i is the vector space

$$V_{\lambda_i} := \left\{ \begin{pmatrix} \xi \\ \lambda_i \xi \end{pmatrix} \mid \xi \in \mathbb{C}^n \right\} \simeq \mathbb{C}^n, \qquad i = 1, 2.$$

Let (a, b) be a vector of $\mathbb{C}^{2n}$, with $a, b \in \mathbb{C}^n$, and let $(u_1, v_1) \in V_{\lambda_1}$, $(u_2, v_2) \in V_{\lambda_2}$. Then (a, b) can be decompose into the sum of eigenvectors

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}.$$

Then,

$$\begin{cases} a = u_1 + u_2 \\ b = v_1 + v_2 \\ v_1 = \lambda_1 u_1 \\ v_2 = \lambda_2 u_2 \end{cases} \Leftrightarrow \begin{cases} a = u_1 + u_2 \\ b = \lambda_1 u_1 + \lambda_2 u_2 \\ v_1 = \lambda_1 u_1 \\ v_2 = \lambda_2 u_2. \end{cases}$$

Observing that $\lambda_1 \cdot \lambda_2 = -1$, we obtain

$$\begin{cases} u_1 = \frac{b - \lambda_2 a}{\lambda_1 - \lambda_2} \\ v_1 = \frac{\lambda_1 b + a}{\lambda_1 - \lambda_2} \end{cases} \qquad \begin{cases} u_2 = \frac{b - \lambda_1 a}{\lambda_2 - \lambda_1} \\ v_2 = \frac{\lambda_2 b + a}{\lambda_1 - \lambda_2}. \end{cases}$$

Thus,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} b - \lambda_2 a \\ \lambda_1 b + a \end{pmatrix} + \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} b - \lambda_1 a \\ \lambda_2 b + a \end{pmatrix}.$$

Now, writing a vector accordingly as (u, v) with $u \in V_{\lambda_1}$ and $v \in V_{\lambda_2}$, and recalling that $\lambda_1 = e^{i\alpha}$ and $\lambda_2 = -e^{-i\alpha}$, one has

$$\phi \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} e^{i\alpha} u \\ -e^{-i\alpha} v \end{pmatrix}. \tag{4.16}$$

The orbit of (u, v) lies on the union of two circles (e^{itu}, e^{-itv}) and $(e^{itu}, -e^{-itv})$, where $t \in \mathbb{R}$. The orbit is finite if α is π-rational, and dense on the two circles otherwise. Let

$$\alpha = 2\pi \frac{p}{q}$$

where p and q are coprime, and assume that $v \neq 0$. If q is even, the orbit closes up after q iterations, but if q is odd, one needs twice as many, due to the alternating sign of the second component in 4.16. $\square$

Chapter 5

The minimal action in Birkhoff and symplectic billiards

Abstract. In this chapter we study Birkhoff and symplectic billiards in the framework of Aubry-Mather theory. This theory is concerned with the study of orbits of a monotone twist map minimizing the action functional. In particular, we focus on the study of the average minimal action (the so-called Mather's β-function) and its properties for Birkhoff and symplectic billiards.

5.1 Aubry-Mather theory

Aubry-Mather theory, developed independently by Serge Aubry and John Mather in the Eighties, is concerned with the study of orbits of a monotone twist map minimizing the action functional. We refer to [29, Chapter 1] for a detailed explanation on this topic.

Definition 5.1.1. Let $a, b \in \mathbb{R} \cup \{\pm\infty\}$, $a < b$. A $\mathcal{C}^1$ diffeomorphism $f : \mathbb{S}^1 \times (a,b) \to \mathbb{S}^1 \times (a,b)$ is a monotone twist map if its lift to the universal cover

$$\begin{aligned} \tilde{f} : \mathbb{R} \times (a,b) &\to \mathbb{R} \times (a,b) \\ (x_0, y_0) &\mapsto (x_1, y_1) \end{aligned}$$

satisfies the following properties:

(i) $\tilde{f}(x_0 + 1, y_0) = \tilde{f}(x_0, y_0) + (1, 0)$;

(ii) $\tilde{f}$ preserves orientation and the boundaries of $\mathbb{R} \times (a,b)$, in the sense that $y_1(x_0, y_0) \to a, b$ as $y_0 \to a, b$;

(iii) if a or b are finite, then $\tilde{f}$ extends continuously to the boundary $\mathbb{R} \times [a,b]$ by a rotation: $\tilde{f}(x,a) = (x + \rho_-, a)$ and $\tilde{f}(x,b) = (x + \rho_+, b)$;

(iv) $\tilde{f}$ satisfies a monotone twist condition

$$\frac{\partial x_1}{\partial y_0} > 0; \tag{5.1}$$

(v) $\tilde{f}$ is exact symplectic; there is a $\mathcal{C}^2$ 1-periodic function h, called a generating function for $\tilde{f}$, such that

$$y_1 dx_1 - y_0 dx_0 = dh(x_0, x_1). \tag{5.2}$$

The interval (ρ_-, ρ_+) is called the twist interval of f.

In particular, from (5.2) we have that

$$\begin{cases} y_1 = \frac{\partial h}{\partial x_1}(x_0, x_1) \\ y_0 = -\frac{\partial h}{\partial x_0}(x_0, x_1). \end{cases} \tag{5.3}$$

Denoting by ∂_i the partial derivative of a function with respect to the i–th variable, it follows that an orbit $((x_i, y_i))_{i\in\mathbb{Z}}$ of a monotone twist map is completely determined by the sequence $(x_i)_{i\in\mathbb{Z}}$ via the relations

$$y_i = \partial_i h(x_{i-1}, x_i) = -\partial_i h(x_i, x_i + 1).$$

Similarly, an arbitrary sequence $(x_i)_{i\in\mathbb{Z}}$ corresponds to an orbit of a monotone twist map $\tilde{f}$ if and only if

$$\partial_i h(x_{i-1}, x_i) + \partial_i h(x_i, x_{i+1}) = 0.$$

Thus, orbits of a monotone twist map correspond to 'critical points' of the discrete action functional

$$(x_i)_{i\in\mathbb{Z}} \mapsto \sum_{i\in\mathbb{Z}} h(x_i, x_{i+1}) \tag{5.4}$$

on $\mathbb{R}^{\mathbb{Z}}$.
Aubry-Mather theory is concerned with the study of orbits that minimize this action functional.

Definition 5.1.2. A sequence $(x_i)_{i\in\mathbb{Z}}$ is called minimal if every finite segment minimizes the action with fixed end points, i.e, if

$$\sum_{i=k}^{l-1} h(x_i, x_{i+1}) \leqslant \sum_{i=k}^{l-1} h(\xi_i, \xi_{i+1})$$

for all finite segments $(\xi_k, \ldots, \xi_l) \in \mathbb{R}^{l-k+1}$ with $\xi_k = x_k$ and $\xi_l = x_l$.
In particular, each minimal sequence $(x_i)_{i\in\mathbb{Z}}$ corresponds to an orbit $((x_i, y_i))_{i\in\mathbb{Z}}$ of f; these are called minimal orbits of f.

We can associate to an orbit a rotation number defined as follows

Definition 5.1.3. The rotation number of an orbit $((x_i, y_i))_{i\in\mathbb{Z}}$ is given by

$$\pi\rho = \lim_{i\to\pm\infty} \frac{x_i}{i}$$

if this limit exists.

The rotation number always exists for periodic orbits, i.e., orbits $((x_i, y_i))_{i\in\mathbb{Z}}$ with

$$x_{i+q} = x_i + p$$

for all $i \in \mathbb{Z}$, where p, q are integers with $q > 0$. In order to have q as the minimal period one assumes that p and q are relatively prime. Then the rotation number is given by

$$\rho = \frac{p}{q}$$

Moreover, Birkoff proved that monotone twist maps possess periodic orbits for each rational rotation number in their twist interval.

Theorem 5.1.4. *Let f be a monotone twist map with twist interval (ρ_-, ρ_+). For any rational number $p/q \in (\rho_-, \rho_+)$ in lowest terms, f possesses at least two geometrically distinct periodic orbits with rotation number p/q.*

The main result of Aubry-Mather theory is the generalization of this result to orbits of any given rotation number in the twist interval.

Theorem 5.1.5. *A monotone twist maps possesses minimal orbits for every rotation number in its twist interval (ρ_-, ρ_+). For rational numbers there are always at least two periodic minimal orbits. Moreover, every minimal orbit lies on a Lipschitz graph over the x-axis.*

We now introduce the minimal average action (or Mather's β function), which will play a central role in our discussion.

Definition 5.1.6. The minimal action of a monotone twist map f with generating function h is defined as the map $\beta : (\rho_-, \rho_+) \to \mathbb{R}$ such that

$$\beta(\rho) := \lim_{N\to\infty} \frac{1}{2N} \sum_{i=-N}^{N-1} h(x_i, x_{i+1}).$$

The minimimal average action contains information about the dynamics of the monotone twist maps.
Here we list some properties:

(i) β is strictly convex and, hence, continuous (see[9]);

(ii) β is differentiable at all irrationals numbers (see [21]);

(iii) β is differentiable at a rational p/q if and only if there exists an invariant circle consisting of periodical minimal orbits of rotation number p/q (see [21]).

It follows from the convexity of β that the minimal average action possesses a convex conjugate

$$\alpha(c) := \beta^*(c) = \sup_{\rho \in (\rho_-, \rho_+)} (\rho c - \beta(\rho)).$$

Since β is strictly convex, the supremum is actually a maximum, and the function α – generally called Mather's α-function– is a convex, real-valued function with $\alpha'(\beta'(\rho)) = \rho$, whenever $\beta'(\rho)$ exists.

5.2 The minimal action in Birkhoff billiards

This section is concerned with the study of Mather's β-function for Birkhoff billiards. This function corresponds to the minimal average action of orbits with a given rotation number and, at least for rational rotation number, it can be related to the maximal length of periodic orbits with a given rotation number (the so-called marked length spectrum). We study its main properties and provide explicit expressions of the coefficients of its (formal) Taylor expansion at zero, only in terms of the curvature of the boundary.

Let D be a strictly convex planar domain with smooth boundary ∂D. An orbit $\{x_i\}_{i\in\mathbb{Z}}$ of the Birkhoff billiard map T corresponds to a critical configuration of the length functional

$$\{x_i\}_{i\in\mathbb{Z}} \mapsto \sum_{i\in\mathbb{Z}} h(x_i, x_{i+1})$$

where the generating function h is minus the Euclidean distance (see 1.1.5). Therefore, the action of the orbit corresponds –up to a sign– to the length of the trajectory traced by the ball on the billiard table D. Consequently, action minimization can be rephrased in terms of length maximization.
An n-periodic orbit of the Birkhoff billiard map T is an n-gon $P = (x_1, \dots, x_n)$ of extremal length inscribed in the billiard curve.
Any periodic orbit admits a rotation number defined as follows

Definition 5.2.1. The rotation number of a billiard periodic orbit is the rational number

$$\frac{m}{n} = \frac{\text{winding number}}{\text{number of reflections}} \in (0, \frac{1}{2}]$$

where the winding number m is defined as follows. Fix the positive orientation of ∂D and pick any reflection point of the periodic orbit on ∂D; then follow the trajectory and measure how many times it goes around ∂D in the positive direction until it comes back to the starting point.

Remark 5.2.2. Notice that, inverting the direction of motion, any periodic billiard trajectory with rotation number $m/n \in (0, 1/2]$ can be seen as one with rotation number $(n-m)/n \in [1/2, 1)$.

Applied to Birkhoff billiards, Theorem 5.1.4 shows that for every m/n in lowest terms there are at least two periodic orbits of rotation number m/n; one of them is an inscribed n–gon with winding number m maximizing the total length amongst all such n–gons, the other is obtained by min-max methods (see [6]).

Example 5.2.3. The ellipse is an example of Birkhoff billiard with exactly two 2-periodic orbits (that is, with rotation numeber $1/2$), which correspond to the two semi-axes of the ellipse (see Example 1.3.8).
There are also cases of billiards with more than two periodic orbits for any given rotation number. For example, in a circular billiard, due to the existence of a 1-dimensional group of symmetries (rotations), there exists a 1-parameter family of periodic orbits for any given rotation number m/n; for example, all diameters are 2-periodic orbits (see Example 1.3.7).

Given these examples, a natural question that one could ask is whether the domain of a Birkhoff billiard can be recover from the knowledge of periodic orbits and their length. To get closer to the comprehension of this problem, we first define the length spectrum and the marked length spectrum associated to a domain D.

Definition 5.2.4. The length spectrum of D is the set of multiples of the lengths of all periodic orbits and multiples of the perimeter $l(\partial D)$ of the domain D

$$\mathcal{L}_D := \mathbb{N}\{\text{lengths of periodic orbits in } D\} \cup \mathbb{N}\, l(\partial D). \tag{5.5}$$

Remark 5.2.5. The length spectrum $\mathcal{L}_D$ of a strictly convex domain D is closely related to the spectrum of the Laplace operator with Dirichlet boundary conditions

$$\begin{cases} \Delta u = \lambda^2 u & \text{in } D \\ u|_{\partial D} = 0. \end{cases} \tag{5.6}$$

From the physical point of view, the eigenvalues λ are the eigenfrequencies of the membrane D with a fixed boundary.
Let $\mathrm{Spec}_\Delta(D) := \{0 < \lambda_1 \leqslant \lambda_2 \leqslant \dots\}$ be the Laplace spectrum of eigenvalues solving this problem. The following theorem by K. Andersson and R. Melrose (see [2]) provides the relation between length and Laplace spectra and it implies that, at least for generic domains, one can recover the length spectrum from the Laplace one.

Theorem 5.2.6. *Let $D \in \mathbb{R}^2$ be a strictly convex compact domain with smooth boundary and let $\mathcal{L}_D$ denote its length spectrum. Then, the wave trace*

$$w(t) := Re\Bigg(\sum_{\lambda_n \in Spec_\Delta(D)} e^{i\lambda_n t} \Bigg)$$

is well-defined as a distribution and smooth away from the length spectrum, that is

$$sing.supp.(w(t)) \subseteq \pm\mathcal{L}_D \cup \{0\}.$$

So if $l > 0$ belongs to the singular support of this distribution, then there exists either a closed billiard trajectory of length l, or a closed geodesic of length l in the boundary of the billiard table.
The relation between periodic orbits and spectral properties of the domain immediately recalls the famous question of Kak "*Can you hear the shape of a drum?*" (see [16]). More precisely, is it possible to recover information about the shape of a drumhead (i.e., a domain) from the sound it makes (i.e., the list of basic harmonics/ eigenvalues of the Laplace operator with Dirichlet or Neumann boundary conditions)? In general, the answer to this question is negative, but it is still an open and interesting problem if one restricts the shape of the domain.

Definition 5.2.7. The marked length spectrum of D is the map

$$\mathcal{ML}_D : \mathbb{Q} \cap (0, \frac{1}{2}] \to \mathbb{R}_+ \tag{5.7}$$

that associates to any rotation number m/n the maximal lenghts of periodic orbits with rotation number m/n.

Due to the relation between action minimization and length maximization that we described at the beginning of the section, periodic orbits of maximal length correspond to minimal orbits of the length functional, and for any $m/n \in (0, 1/2]$, the marked length spectrum is essentialy Mather's β-function

$$\beta(\frac{m}{n}) = -\frac{1}{n}\mathcal{ML}_D(\frac{m}{n}).$$

In [29, Chapter 3], the properties of Mather's β-function and its conjugate Mather's α-function for Birkhoff billiards have been studied in depth. Here we recall some of the most important, pointing out how some geometric properties of the domain D can be express in terms of the β and α functions.

(i) β is strictly convex in $[0, 1]$ and symmetric with respect to the point $1/2$;

(ii) β is differentiable at m/n if and only if there exists a caustic of rotation number m/n;

(iii) if Γ_ρ is a caustic with rotation number $\rho \in (0, 1/2]$, then β is differentiable at ρ and $\beta'(\rho) = -\text{length}(\Gamma_\rho) =: -|\Gamma_\rho|$. In particular, β is always differentiable at 0 and $\beta'(0) = -l(\partial D)$;

(iv) given a caustic Γ_ρ of rotation number $\rho \in (0, 1/2]$, the Lazutkin parameter of Γ_ρ is defined as

$$Q(\Gamma_\rho) := |A - P| + |P - B| - |AB|$$

where P is any point on ∂D, $A, B \in \Gamma_\rho$ are the points of tangency of Γ_ρ seen from P and $|AB|$ denotes the length of the arc on the caustic joining A to B.
This quantity is connected to the value of the α-function

$$Q(\Gamma_\rho) = \alpha(\beta'(\rho)) = \alpha(-|\Gamma_\rho|).$$

Next theorem provides explcit expressions for the Taylor expansions of Mather's β and α functions, respectively at $\rho = 0$ and $c = -l_0$, where l_0 denotes the length of the boundary ∂D (see [30, Theorem 1.3]). In particular, the coefficients of these expressions are only in terms of the curvature of the boundary.

Theorem 5.2.8. *Let D be a strictly convex domain in $\mathbb{R}^2$ with smooth boundary. Denote by $k(t)$ the curvature of ∂D with arc-length parametrization t. Let $l_0 := |\partial D|$ be the length of the boundary and denote*

$$\mathcal{I}_1 := \int_0^{l_0} dt = l_0$$

$$\mathcal{I}_3 := \int_0^{l_0} k^{2/3} dt$$

$$\mathcal{I}_5 := \int_0^{l_0} \left(9\,k^{4/3} + \frac{8\,\dot{k}^2}{k^{8/3}}\right) dt$$

$$\mathcal{I}_7 := \int_0^{l_0} \left(9\,k^2 + \frac{24\,\dot{k}^2}{k^2} + \frac{24\,\ddot{k}^2}{k^4} - \frac{144\,\dot{k}^2\,\ddot{k}}{k^5} + \frac{176\,\dot{k}^4}{k^6}\right) dt$$

$$\mathcal{I}_9 := \int_0^{l_0} \left(\frac{281}{44800} k^{8/3} + \frac{281\,\dot{k}^2}{8400\,k^{4/3}} + \frac{167\,\ddot{k}^2}{4200\,k^{10/3}} - \frac{167\,\dot{k}^2\,\ddot{k}}{700\,k^{13/3}} + \frac{\dddot{k}^2}{42\,k^{16/3}} + \frac{559\,\dot{k}^4}{2100\,k^{16/3}} \right.$$
$$\left. - \frac{473\,\ddot{k}^3}{4725\,k^{19/3}} - \frac{10\,\dddot{k}\,\ddot{k}\,\dot{k}}{21\,k^{19/3}} + \frac{5\,\dddot{k}\,\dot{k}^3}{7\,k^{22/3}} + \frac{13142\,\dot{k}^2\,\ddot{k}^2}{4725\,k^{22/3}} - \frac{10777\,\dot{k}^4\,\ddot{k}}{1575\,k^{25/3}} + \frac{521897\,\dot{k}^6}{127575\,k^{28/3}}\right) dt$$

Then

- *the formal Taylor expansion of β at $\omega = 0$, $\beta(\omega) \sim \sum_{k=0}^{\infty} \beta_k \frac{\omega^k}{k!}$, has coefficients:*

$$\beta_{2k} = 0 \quad \textit{for all}\, k$$

$$\beta_1 = -\mathcal{I}_1$$

$$\beta_3 = \frac{1}{4}\mathcal{I}_3^3$$

$$\beta_5 = -\frac{1}{144}\mathcal{I}_3^4\,\mathcal{I}_5$$

$$\beta_7 = \frac{1}{320}\mathcal{I}_3^5\left(\frac{14}{81}\mathcal{I}_5^2 - \mathcal{I}_3\,\mathcal{I}_7\right)$$

$$\beta_9 = -7\,\mathcal{I}_3^6\left(\mathcal{I}_3^2\,\mathcal{I}_9 - \frac{1}{5600}\mathcal{I}_3\,\mathcal{I}_5\,\mathcal{I}_7 + \frac{7}{583200}\mathcal{I}_5^3\right);$$

- *the (formal) Taylor expansion of $(c+l_0)^{-3/2}\alpha(c)$ at $c = -l_0$ (note that α has in fact a square-root type singularity at the boundary), $(c+l_0)^{-3/2}\alpha(c) \sim \sum_{k=0}^{\infty} \alpha_k \frac{(c+l_0)^k}{k!}$,* *has coefficients:*

$$
\begin{aligned}
\alpha_0 &= \frac{4\sqrt{2}}{3}\,\mathcal{I}_3^{-3/2}\\
\alpha_1 &= \frac{\sqrt{2}}{135}\,\mathcal{I}_3^{-7/2}\,\mathcal{I}_5\\
\alpha_2 &= \frac{1}{56700\sqrt{2}}\left(\frac{72\,\mathcal{I}_3\,\mathcal{I}_7 + 7\,\mathcal{I}_5^2}{\mathcal{I}_3^{11/2}}\right)\\
\alpha_3 &= \frac{1}{826686000\sqrt{2}}\left(\frac{261273600\,\mathcal{I}_3^2\,\mathcal{I}_9 + 21384\,\mathcal{I}_3\,\mathcal{I}_5\,\mathcal{I}_7 + 1001\,\mathcal{I}_5^3}{\mathcal{I}_3^{15/2}}\right).
\end{aligned}
$$

A direct consequence of this theorem is the following

Corollary 5.2.9. *Let D be a strictly convex domain in $\mathbb{R}^2$ with smooth boundary. Then*

$$\beta_3 + \pi^2\beta_1 \leqslant 0$$

and equality holds if and only if D is a disc.

Proof. The proof follows from the expressions of β_1 and β_3 found in Theorem 5.2.8

$$\beta_1 = -\mathcal{I}_1 \qquad \beta_3 = \frac{1}{4}\mathcal{I}_3^3.$$

Indeed, it follows that

$$\beta_3 + \pi^2\beta_1 \leqslant 0 \;\Rightarrow\; \mathcal{I}_3^3 - 4\pi^2\mathcal{I}_1 \leqslant 0.$$

Using Hölder inequality (with $p = \frac{3}{2}$ and $q = 3$)

$$
\begin{aligned}
\mathcal{I}_3 &= \int_0^{l_0} k^{2/3ds}\\
&\leqslant \left(\int_0^{l_0} (k^{2/3})^{3/2} ds\right)^{2/3} \left(\int_0^{l_0} 1^3 ds\right)^{1/3}\\
&= (2\pi)^{2/3} l_0^{1/3}\\
&= (4\pi^2\mathcal{I}_1)^{1/3}.
\end{aligned}
$$

Moreover, equality holds if and only if it holds in Hölder inequality. This means that k must be constant (and strictly positive) and therefore, the curve must be a circle. □

It is clear from the corollary above, that the Mather's β-function univocally determines discs amongst all possible Birkhoff billiards.
A similar result for elliptic billiards has not been discovered yet. However, Mather's β-function determines univocally a given ellipse in the family of all ellipses (see [30, Proposition 1]).

Proposition 5.2.10. *If $\mathcal{E}_1$ and $\mathcal{E}_2$ are two ellipses such that $\beta_{\mathcal{E}_1} \equiv \beta_{\mathcal{E}_2}$, then $\mathcal{E}_1$ and $\mathcal{E}_2$ are the same ellipse. More generally, if the Taylor coefficients $\beta_{\mathcal{E}_1,1}, \equiv \beta_{\mathcal{E}_2,1}$ and $\beta_{\mathcal{E}_1,3}, \equiv \beta_{\mathcal{E}_2,3}$, then the same conclusion remains true.*

In the next example we determine Mather's β-function for circular billiards.

Example 5.2.11. Let D be a disc of radius R. As we have seen in Example 1.3.7, the angle of reflection $\varphi \in (0, \pi/2]$ remains constant along an orbit. In particular, if the orbit has rotation number $\rho \in (0, \frac{1}{2}]$, then $\varphi = \pi\rho$.
Therefore, the segment joining two consecutive collision points has always length $2R\sin(\pi\rho)$, and from the Deefinition 5.1.6 of β, it follows that

$$\beta(\rho) = -\lim_{N\to\infty} \frac{1}{2N} \sum_{i=-N}^{N-1} 2R\sin(\pi\rho) = -2R\sin(\pi\rho).$$

Its Taylor expansion is

$$\begin{aligned}\beta(\rho) &= -2R\sin(\pi\rho)\\ &= -2R\pi\rho + \frac{1}{3!}(2R\pi^3)\rho^3 - \frac{1}{5!}(2R\pi^5)\rho^5 + \frac{1}{7!}(2R\pi^7)\rho^7 - \frac{1}{9!}(2R\pi^9)\rho^9 + \mathcal{O}(\rho^{11})\end{aligned}$$

Recalling that the curvature is $k \equiv \frac{1}{R}$ and the length of the billiard curve is $l_0 = 2\pi R$, the invariants of Theorem 5.2.8 are

$$\begin{aligned}\mathcal{I}_1 &= 2\pi R\\ \mathcal{I}_3 &= 2\pi R^{1/3}\\ \mathcal{I}_5 &= 18\pi R^{-1/3}\\ \mathcal{I}_7 &= 18\pi R^{-1}\\ \mathcal{I}_9 &= \frac{281}{22400}\pi R^{-5/3}.\end{aligned}$$

Substituting these values in the expressions of Taylor expansion's coefficients of Mather's β-function, one can check that they match with (5.2.11).

Mather's β-function plays a crucial role in the comprehension of different rigidity phenomena that appear in the study of convex billiards. Moreover, some open questions and conjectures can be rephrased in terms of this function.
For example, because of the relation between the differentiability properties of Mather's β-function at rational rotation numbers and the existence of invariant circles consisting of periodic points, Birkhoff conjecture 1.4 can be rephrased as

Conjecture (Birkhoff revisited). *Let D be a strictly convex planar domain with smooth boundary and assume that β_D is differentiable in $[0, 1/2)$. Is it true that D is an ellipse?*
More generally, if β_D is differentiable in $[0, \epsilon)$ for some small $0 < \epsilon < 1/2$, is it true that D is an ellipse?

Observe that if β_D is $\mathcal{C}^\infty([0,\epsilon))$, then the billiard map is locally integrable near the boundary. Indeed, β_D will be differentiable at all rationals in $(0,\epsilon)$ and therefore there will be caustics corresponding to these rotation number. By semi-continuity arguments, one obtains caustics corresponding to irrational rotation number and hence a family of caustics that foliate a neighbourhood of the boundary.
Another open question that one can reformulate in terms of Mather's β-function is the one concerning the spectral rigidity of the domain (see [13] for more details)

Question (**Marked length spectrum rigidity**). *Let D_1 and D_2 be two strictly convex planar domains with smooth boundaries and assume that they are isospectral, i.e., $\mathcal{ML}_{D_1} \equiv \mathcal{ML}_{D_2}$. Is it true that D_1 and D_2 are isometric?*
More generally, if $\mathcal{ML}_{D_1}(\rho) \equiv \mathcal{ML}_{D_2}(\rho)$ for all $m/n \in \mathcal{Q} \cap [0,\epsilon)$ for some $0 < \epsilon \leqslant 1/2$, is it true that D_1 and D_2 are isometric?

In terms of Mather's β-function, this question can be rewritten as

Question (**Marked length spectrum rigidity revisted**). *Let D_1 and D_2 be two strictly convex planar domains with smooth boundaries and assume that $\beta_{D_1} \equiv \beta_{D_2}$. Is it true that D_1 and D_2 are isometric?*
More generally, if $\beta_{D_1}(\rho) \equiv \beta_{D_2}(\rho)$ for all $\rho \in (0,\epsilon)$ for some small $\epsilon > 0$, is it true that D_1 and D_2 are isometric?

5.3 The minimal action in symplectic billiards

In this section we address the topic of minimal action in symplectic billiards. Similarly to the Birkhoff case, Mather's β-function, at least for rational rotation number, can be related to the maximal area of periodic orbits with a given rotation number (the so-called marked area spectrum). In particular, we give explicit expressions of the Taylor expansion at 0 of Mather's β-function for symplectic billiards in the circle and the ellipse.

Let γ be a smooth, strictly convex, closed curve in $\mathbb{R}^2$. The generating function S of the symplectic billiard map ϕ is the standard area form ω. With such a generating function, the symplectic billiard map ϕ turns out to be a negative monotone twist map (see Proposition 2.4.2).
If we consider as a generating function minus the standard area form ω, that is

$$S(x,y) = -\omega(x,y)$$

for any $(x,y) \in \mathcal{P}$, the results from Chapter 2 remain essentialy the same, but the symplectic billiard map becomes a positive monotone twist map.
In the sequel, to stay in the assumptions of Aubry-Mather theory (so, with a positive monotone twist map, see Definition 5.1.1), we will work with minus the standard area form as the generating function.

A periodic orbit $\{x_i\}_{i=1}^n$ of the symplectic billiard map ϕ corresponds to a critical configuration of the area functional

$$\{x_i\}_{i=1}^n \mapsto -\sum_{i=1}^{n-1} \omega(x_i, x_{i+1}).$$

Observe that $\omega(x_i, x_{i+1}) \geqslant 0$ in the phase space $\mathcal{P}$. Thus, the action of the orbit coincides –up to a factor -2– with the area enclosed by the trajectory traced by the ball on the billiard table.
Consequently, an n-periodic orbit $\{x_i\}_{i=1}^n$ corresponds to an n-gon $P = (x_1, \ldots, x_n)$ of extremal area inscribed in the billiard curve.
In particular, given this relation between action and area, minimizing the action functional is equivalent to maximizing the area enclosed by the orbit.
Note that the role played by the Euclidean length h in Birkhoff billiards is here played by the standard area form ω. This induces us to define the area spectrum and marked area spectrum associated to a symplectic billiard in a similar way to how we defined length spectrum and marked length spectrum for Birkhoff billiards (see Definitions 5.2.4 and 5.2.7).
The rotation number associated to a symplectic periodic orbit is defined in the same way as the rotation number for Birkhoff periodic orbit (see Definition 5.2.1), i.e., as the rational number

$$\frac{m}{n} = \frac{\text{winding number}}{\text{number of reflections}} \in (0, \frac{1}{2}],$$

where, once fixed the positive orientation of the symplectic billiard curve γ, the winding number m is the number of times the trajectory goes around γ in the positive direction until it comes back to the starting point.

Remark 5.3.1. If x is the starting point of a ϕ-periodic orbit, due to the definition of the phase space $\mathcal{P}$, the next reflection point y must be between x and its opposite point x^*. Then, the maximal number of rotation is $1/2$, which corresponds to the case $y = x^*$. Unlike Birkhoff billiards, where, inverting the direction of motion, for any periodic orbit of rotation number $m/n \in (0, 1/2]$ we obtain an equivalent orbit of rotation number $(n-m)/n \in [1/2, 1)$, in the symplectic case inverting the direction of motion we obtain an orbit for the negative phase space $\mathcal{P}^-$.

We define the area spectrum and the marked area spectrum associated to the symplectic billiard in γ. Let D be the billiard table enclosed by γ.

Definition 5.3.2. The area spectrum of D is the set of multiples of the areas of all periodic orbits and multiples of the area $A(\partial D)$ of the domain D

$$\mathcal{A}_D := \mathbb{N}\{\text{area of periodic orbits in } D\} \cup \mathbb{N}\, A(\partial D). \tag{5.8}$$

Definition 5.3.3. The marked area spectrum of D is the map

$$\mathcal{MA}_D : \mathbb{Q} \cap (0, \frac{1}{2}] \to \mathbb{R}^+ \tag{5.9}$$

that associates to any rotation number m/n the maximal area of the periodic orbits with rotation number m/n.

As we have pointed out at the beginning of the section, periodic orbits of maximal area correspond to minimal orbits of the action functional. Therefore, for any $m/n \in (0, 1/2]$, the marked area spectrum corresponds to the Mather's β-function for symplectic billiards

$$\beta(\frac{m}{n}) = -\frac{2}{n}\mathcal{MA}_D(\frac{m}{n}). \tag{5.10}$$

Let now $A_n := \mathcal{MA}_D(m/n)$ be the maximal area of the periodic orbits with rotation number m/n. The following theorem provides an asymptotic expansion for A_n as $n \to \infty$.

Theorem 5.3.4. *Let γ be a smooth, strictly convex, closed curve in $\mathbb{R}^2$, parametrized by affine arc-length s. Denote by $L = \int_\gamma ds$ the total affine length and by $\kappa(s)$ the affine curvature.*
Let A_n be the maximal area of the periodic orbits of the symplectic billiard map ϕ with rotation number m/n.
The asymptotic expansion of A_n for $n \to \infty$ is given by

$$A_n \sim a_0 + \frac{a_1}{n^2} + \frac{a_2}{n^4} + \frac{a_3}{n^6} + \dots \tag{5.11}$$

where

(i) a_0 *is the area of the billiard table enclosed by* γ,

(ii) $a_1 = \frac{L^3}{12}$,

(iii) $a_2 = -\frac{L^4}{240}\int_0^L \kappa(s)ds$

The asymptotic expansion of A_n is due to the theory of interpolating Hamiltonians. Applied to symplectic billiards, this theory implies that the symplectic billiard map ϕ equals an integrable symplectic map, the time-one map of a Hamiltonian vector field, composed with a smooth symplectic map that fixes the boundary of the phase space point-wise to all orders, see [[20], [24]]. The coefficients a_1, a_2 were respectively found in [[23], [19]].

Remark 5.3.5. The length spectrum of the billiard problem is closely related to the spectrum of the Laplace operator with Dirichlet or Neumann boundary conditions on the billiard curve (see Remark 5.2.5).
An interesting open question is whether the area spectrum od symplectic billiards is related to the spectrum of some differential operator. Notice that such an operator, if it exists, must be invariant under area-preserving affine transformations

Next corollary is an immediate consequence of the affine isoperimetric inequality (see [28, Section 4]), that we recall here below for reader's convenience.

Lemma 5.3.6. *Given a strictly convex close curve parametrized by affine arc-length s, let L be its total affine length and A the area enclosed by the curve. The next inequality holds*

$$L^3 \leqslant 8\pi^2 A \tag{5.12}$$

with equivalence only for ellipses.

Corollary 5.3.7. *Let D be a strictly convex planar domain with smooth boundary. Then*

$$3a_1 \leqslant 2\pi^2 a_0$$

and equality holds if and only if D is an ellipse.

Proof. The expressions of the coefficients a_0, a_1 imply that

$$a_0 = A \qquad L^3 = 12a_1.$$

Substituting into the affine isoperimetric inequality 5.12, we obtain the desired result. □

The next examples provide explicit expressions of the Taylor expansion at $\rho = 0$ of Mather's β-function for symplectic billiards in the circle and the ellipse.

Example 5.3.8. Let D be a disc of radius R. As we have seen in Section 2.5, the dynamics of the symplectic billiard in D is the same as the one of the Birkhoff billiard in the disc, that is, the rotation of angle $\alpha \in (0, \pi]$.
Therefore, for an orbit of rotation number $\rho \in (0, 1/2]$, it follows that $\alpha = 2\pi\rho$. Moreover, the standard area form is constant along the orbit and, if x, y are two consecutive collision points, it equals twice the area of the triangle xOy

$$\omega(x, y) = 2A(xOy).$$

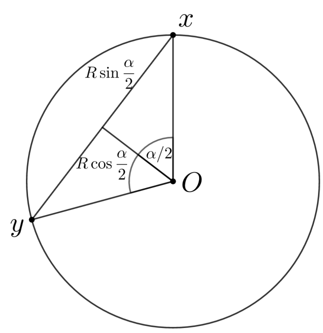

Figure 5.1: The area $A(xOy)$

The area of the triangle xOy is given by (see Figure 5.1)

$$\begin{aligned} A(xOy) &= \frac{1}{2}(2R\sin\frac{\alpha}{2})(R\cos\frac{\alpha}{2}) \\ &= \frac{R^2}{2}(2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}) \\ &= \frac{R^2}{2}\sin\alpha \\ &= \frac{R^2}{2}\sin(2\pi\rho). \end{aligned}$$

Thus,

$$\omega(x,y) = 2A(xOy) = R^2\sin(2\pi\rho).$$

From the Definition 5.1.6 of Mather's β-function, it follows

$$\beta(\rho) = -\lim_{N\to\infty}\frac{1}{2N}\sum_{i=-N}^{N-1} R^2\sin(2\pi\rho) = -R^2\sin(2\pi\rho).$$

and therefore, its Taylor expansion at $\rho = 0$ is

$$\begin{aligned} \beta(\rho) &= -R^2\sin(2\pi\rho) \\ &= -2R^2\pi\rho + \frac{1}{3!}(8R^2\pi^3)\rho^3 - \frac{1}{5!}(32R^2\pi^5)\rho^5 + \mathcal{O}(\rho^7). \end{aligned} \tag{5.13}$$

The coefficients a_0, a_1, a_2 of the asymptotic expansion (5.11) of A_n in the case of the circle correspond, up to a factor, to the first three coefficients $\beta_0, \beta_1, \beta_2$ of the Taylor expansion of Mather's β-function.
Indeed, by the affine isoperimetric inequality (5.12), it follows that

$$L^3 = 8\pi^2 A = 8\pi^2(\pi R^2) = 8\pi^3 R^2 \Rightarrow L = 2\pi R^{2/3}.$$

Moreover, the affine curvature of the circle is $\kappa = R^{-4/3}$ (see [28, p. 89]), and therefore

$$\int_0^L \kappa(s)ds = \int_0^{2\pi R^{2/3}} R^{-4/3}ds = R^{-4/3}\int_0^{2\pi R^{2/3}} ds = 2\pi R^{-2/3}.$$

Therefore, the coefficients a_0, a_1, a_2 are given by

$$\begin{aligned} a_0 &= \pi R^2 = -\frac{1}{2}\beta_0 \\ a_1 &= \frac{L^3}{12} = \frac{8}{12}\pi^3 R^2 = \frac{2}{3}\pi^3 R^2 = \frac{1}{2}\beta_1 \\ a_2 &= -\frac{L^4}{240}\int_0^L \kappa(s)ds = -\frac{1}{240}(2\pi R^{2/3})^4(2\pi R^{-2/3}) = -\frac{2}{15}\pi^5 R^2 = \frac{1}{2}\beta_2. \end{aligned}$$

Example 5.3.9. Mather's β-function for symplectic billiard in the ellipse follows easily from the one for the symplectic billiard in the circle.
Consider the affine transformation

$$\begin{pmatrix}\tilde{x}_1\\ \tilde{x}_2\end{pmatrix} = f\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}a/R & 0\\ 0 & b/R\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \frac{1}{R}\begin{pmatrix}ax_1\\ bx_2\end{pmatrix}$$

that maps the circle of equation $x_1^2 + x_2^2 = R^2$ to the ellipse of equation $\frac{\tilde{x}_1^2}{a^2} + \frac{\tilde{x}_2^2}{b^2} = 1$. Let L be the matrix representing such affine transformation.
The standard area form between two collision points in the circle, $\omega(x, y) = R^2 \sin(2\pi\rho)$, is mapped by this affine transformation to

$$\begin{aligned}\tilde{\omega}(\tilde{x}, \tilde{y}) &= \tilde{\omega}(Lx, Ly)\\ &= \det(L)\omega(x, y)\\ &= \frac{ab}{R^2}R^2\sin(2\pi\rho)\\ &= ab\sin(2\pi\rho).\end{aligned}$$

Therefore, the Mather's β-function for the ellipse is given by

$$\beta_{\text{ellipse}}(\rho) = \det(L)\beta_{\text{disk}}(\rho) = -ab\sin(2\pi\rho).$$

Its Taylor expansion at $\rho = 0$ is

$$\begin{aligned}\beta(\rho) &= -ab\sin(2\pi\rho)\\ &= -2ab\pi\rho + \frac{1}{3!}(8ab\pi^3)\rho^3 - \frac{1}{5!}(32ab\pi^5)\rho^5 + \mathcal{O}(\rho^7).\end{aligned} \tag{5.14}$$

The coefficients a_0, a_1, a_2 of the asymptotic expansion of A_n for the ellipse can be compute in a similar way to the ones of the circle. The only differences are that the area A of the ellipse is πab and the affine curvature is $\kappa = (ab)^{-2/3}$. Therefore,

$$\begin{aligned}a_0 &= \pi ab = -\frac{1}{2}\beta_0\\ a_1 &= \frac{L^3}{12} = \frac{2}{3}\pi^3 ab = \frac{1}{2}\beta_1\\ a_2 &= -\frac{L^4}{240}\int_0^L \kappa(s)ds = \frac{2}{15}\pi^5 ab = \frac{1}{2}\beta_2\end{aligned}$$

and they correspond, up to a factor, to the first three coefficients $\beta_0, \beta_1, \beta_2$ of the Taylor expansion of Mather's β-function.

Remark 5.3.10. The fact that the first coefficients of the Taylor expansions of Mather's β-function for the circular and elliptic symplectic billiards are essentialy the same as the coefficients of the asymptotic expansion of A_n is clearly due to the relation (5.10) between the marked area spectrum and the β-function. Thus, this correspondence hold

for a generic β-function.
In particular, we observe that these coefficients depend only on the area of the billiard table and the affine total length and affine curvature of the symplectic billiard curve, which are invariant for affine special transformation. This reminds us of the affine plane curve evolution, see [28]. According to this evolution, a closed convex curve evolves converging to an ellipse. Moreover, the evolution equation of area, total affine length and affine curvature were found in [28, Section 6].
Since the only integrable symplectic billiards known are the circle and the ellipse, an idea could be to study integrability in relation to the affine evolution of the Taylor expansion's coefficients of Mather's β-function.

Bibliography

[1] P. Albers and S. Tabchnikov. “Introducing symplectic billiards”. In: *Adv. Math.* 333 (2018), pp. 822–867.

[2] K. G. Andersson and R. Melrose. “The propagation of singularities along gliding rays”. In: *Invent. Math.* 41.3 (1977), pp. 197–232.

[3] A. Avila, J. De Simoi, and V. Kaloshin. “An integrable deformation of an ellipse of small eccentricity is an ellipse”. In: *Ann. Math.* 184.2 (2016), pp. 527–558.

[4] V. Bangert. “Mather sets for twist maps and geodesic on tori”. In: *Dynamics Reported* 1 (1988). Ed. by U. Kirchgraber and H. O. Walther, pp. 1–56.

[5] M. Bialy. “Convex billiards and a theorem by E. Hopf”. In: *Math. Z.* 214.1 (1993), pp. 147–154.

[6] G. D. Birkhoff. “On the periodic motions of dynamical systems”. In: *Acta Math.* 50 (1927), pp. 359–379.

[7] H. W. Broer and M. B. Sevryuk. *KAM Theory: quasi-periodicity in dynamical systems.* 2007.

[8] A. Cannas da Silva. *Lectures on symplectic geometry.* Springer-Verlag, 2006.

[9] G. Forni and J. N. Mather. “Action minimizing orbits in Hamiltonian systems”. In: *Transition to Chaos in Classical and Quantum Mechanics.* Ed. by S. Graffi. Springer, 1992.

[10] P. Giblin and P. Holtom. “The centre symmetry set”. In: *Banach Center Publ.* 50 (1999), pp. 91–105.

[11] R. Goldman. “Curvature formulas for implicit curves and surfaces”. In: *Comput. Aided Geom. Des.* 22 (2005), pp. 632–658.

[12] C. Golé. *Symplectic twist maps. Global variational techniques.* World Scientific Publ. Co., 2001.

[13] V. Guillemin and Melrose R. “A cohomological invariant of discrete dynamical systems”. In: *E. B. Christoffel.* Ed. by Butzer P.L. and Fehér F. Birkhäuser, Basel, 1981.

[14] E. Gutkin and A. Katok. “Caustics for inner and outer billiards”. In: *Commun. Math. Phys.* 173 (1995), pp. 101–133.

[15] B. Halpern. “Strange billiard tables”. In: *Trans. Amer. Math. Soc.* 232 (1977), pp. 297–305.

[16] M. Kak. “Can one hear the shape of a drum?” In: *Amer. Math. Monthly* 73.4 (1966), pp. 1–23.

[17] V. Kaloshin and Sorrentino A. “Kaloshin, V. and Sorrentino A.” In: *Ann. Math.* 188 (2018), pp. 315–380.

[18] V. F. Lazutkin. “The existence of caustics for a billiard problem in a convex domain”. In: *Izv. Akad. Nauk SSSR Ser. Mat.* 7 (1973), pp. 185–214.

[19] M. Ludwig. “Asymptotic approximation of convex curves”. In: *Arch. Math.* 63 (1994), pp. 377–384.

[20] S. Marvizi and R. Melrose. “Spectral invariants of convex planar regions”. In: *J. Differ. Geom.* 17 (1982), pp. 475–502.

[21] J. N. Mather. “Differentiability of the minimal average action as a function of the rotation number”. In: *Bol. Soc. Bras. Mat.* 21 (1990), pp. 59–70.

[22] J. N. Mather. “Glancing billiards”. In: *Ergod. Theory Dyn. Syst.* 2 (1982), pp. 297–403.

[23] D. McClure and R. Vitale. “Polygonal approximation of plane convex bodies”. In: *J. Math. Anal. Appl.* 51 (1975), pp. 326–358.

[24] R. Melrose. “Equivalence of glancing hypersurfaces”. In: *Invent. Math.* 37 (1976), pp. 165–192.

[25] J. K. Moser. “On invariant curves of area-preserving mappings of an annulus”. In: *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl.* (1962), pp. 1–20.

[26] H. Rüssman. “Kleine Nenner I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes”. In: *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl.* 5 (1970), pp. 67–65.

[27] L. A. Santaló. *Integral geometry and geometric probability.* Addison-Wesley Pub. Co., Advanced Book Program, 1976.

[28] G. Sapiro and A. Tannenbaum. “On affine plane curve evolution”. In: *J. Funct. Anal.* 119 (1994), pp. 79–120.

[29] K. F. Siburg. *The principle of least action in geometry and dynamics.* In: *Lecture Notes in Mathematics.* Vol. 1844. Springer, 2004, pp. xiii+128.

[30] A. Sorrentino. “Computing Mather’s β-function for Birkhoff billiards”. In: *Discrete Contin. Dyn. Syst.* 35.10 (Oct. 2015).

[31] S. Tabachnikov. “Billiards”. In: *Panor. Synth.* (1995), pp. vi+ 142.

[32] S. Tabachnikov. “Ellipsoids, complete integrability and hyperbolic geometry”. In: *Mosc. Math. J.* 2.1 (2002), pp. 185–198.

[33] S. Tabachnikov. *Geometry and billiards.* Ed. by American Mathematical Society. Vol. 30. Student mathematical library. 2005, pp. xii+ 176.

[34] M. P. Wojtkowski. “Two applications of Jacobi fields to the billiard ball problem”. In: *J. Differ. Geom.* 40.1 (1994), pp. 155–164.